

Talentierte symbiontische Naturstoffproduzenten als Quelle neuer Wirkstoffe

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Ursula Steffens

aus Wittlich

Bonn, 2019

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Gutachter:

1. Prof. Dr. Jörn Piel
2. Prof. Dr. Hans-Georg Sahl
3. Prof. Dr. Gabriele Bierbaum
4. Prof. Dr. Ulrich Kubitschek

Tag der Promotion: 29.11.2019

Erscheinungsjahr: 2020

Inhaltsverzeichnis

Inhaltsverzeichnis.....	3
Abbildungsverzeichnis.....	6
Tabellenverzeichnis.....	10
1 Zusammenfassung.....	12
2 Abstract.....	14
3 Einleitung.....	16
3.1 Die Biosynthese mikrobieller Sekundärmetabolite.....	17
3.1.1 Die Polyketid-Biosynthese.....	17
3.1.2 Nichtribosomale Peptidsynthetasen.....	24
3.1.3 NRPS/PKS Hybride.....	29
3.1.4 Genome mining.....	30
3.2 Mikrobielle Symbiosen.....	30
3.3 Bakterielle Symbionten - potente „Naturstofffabriken“.....	32
3.3.1 Mikrobielle Schwammsymbionten des Schwammes <i>Theonella swinhoei</i>	33
3.3.2 Mikrobielle Symbionten der <i>Paederus/Paederidus</i> -Käfer.....	41
3.4 Systematik von Bakterien.....	45
3.4.1 Phylogenie.....	45
3.4.2 Evolutionsanalyse von Bakterien.....	47
4 Zielsetzung.....	53
4.1 ' <i>Candidatus Entotheonella factor</i> '.....	53
4.2 <i>Symbiont 2</i> des Schwammes <i>Theonella swinhoei</i>	53
4.3 ' <i>Candidatus Pseudomonas paederi</i> '.....	53
5 Ergebnisse und Diskussion.....	54
5.1 ' <i>Candidatus Entotheonella factor</i> '.....	54
5.1.1 Assemblierung des ' <i>Entotheonella factor</i> ' Genoms.....	54
5.1.2 Umweltanalysen.....	60
5.1.3 Phylogenetische Analysen.....	69

5.1.4	Isolierung von 'Candidatus Entotheonella' aus Bodenproben	85
5.2	Symbiont 2 des Schwammes <i>Theonella swinhoei</i>	102
5.3	'Candidatus Pseudomonas paederi'	108
5.3.1	Schließen der Lücken des 'Pseudomonas paederi' Genoms	108
5.3.2	Analyse eines NRPS/PKS Hybriden	110
6	Zusammenfassung und Ausblick	118
6.1	'Candidatus Entotheonella'	118
6.2	Symbiont 2 des Schwammes <i>Theonella swinhoei</i>	120
6.3	'Candidatus Pseudomonas paederi'	121
7	Material und Methoden	123
7.1	Organismen und Vektoren	123
7.1.1	Organismen	123
7.1.2	Vektoren	123
7.2	Medien und Zusätze	123
7.3	Mikrobiologische Techniken	124
7.3.1	Kultivierung von <i>E. coli</i>	124
7.3.2	Erstellen von Gefrierkulturen	124
7.3.3	Konzentrationsbestimmung	124
7.3.4	Transformation in <i>E. coli</i>	125
7.4	Molekularbiologische Techniken	126
7.4.1	Polymerase-Kettenreaktion (PCR)	126
7.4.2	Agarose-Gelelektrophorese	135
7.4.3	Klonierung	137
7.4.4	Restriktion	139
7.4.5	Isolierung von Plasmid-DNA aus <i>E. coli</i>	140
7.4.6	Sequenzierung	140
7.5	Proteinexpression	141
7.5.1	Sequenzalignment zur Festlegung der Domänensequenz	141
7.5.2	Expression der A-Domänen	142
7.5.3	Zellaufschluss	142

Inhaltsverzeichnis

7.5.4	Aufreinigung von His-Tag-Fusionsproteinen.....	142
7.5.5	Entsalzen und Aufkonzentrieren von Proteinen	144
7.5.6	SDS-Polyacrylamidgelelektrophorese (SDS-PAGE).....	144
7.5.7	Proteinkonzentrationsbestimmung	145
7.5.8	Bestimmung der Spezifität von A-Domänen: [γ - $^{18}\text{O}_4$]-ATP-Assay	145
7.5.9	Massenspektrometrische Untersuchung von Proteinen.....	146
7.5.10	Expression des gesamten NRPS/PKS Genclusters	146
7.6	Untersuchung von Bodenproben.....	146
7.6.1	Aufreinigung metagenomischen DNA aus Umweltproben	146
7.6.2	Nachweis von ' <i>Candidatus Entotheonella</i> ' mittels PCR	147
7.7	Isolierung von Bakterien aus Umweltproben.....	147
7.7.1	Isolierung aus Bodenproben.....	147
7.8	Mikroskopie.....	148
7.8.1	DAPI-Färbung.....	148
7.9	Fluoreszenz in situ Hybridisierung (FISH)	148
7.9.1	Catalyzed reporter deposition (CARD)-FISH.....	149
7.10	Verwendete Chemikalien.....	152
7.11	Verwendete Geräte und Materialien	155
8	Abkürzungsverzeichnis.....	157
9	Literaturverzeichnis	159
10	Anhang.....	171

Abbildungsverzeichnis

Abbildung 3-1: Darstellung des sogenannten „discovery void“ in der Antibiotikaentwicklung.	17
Abbildung 3-2: Strukturelle Vielfalt der Polyketide.	18
Abbildung 3-3: Substrate von FAS (A) und PKS (B).	19
Abbildung 3-5: Aktivierung der ACP-Domäne.	20
Abbildung 3-6: Übertragung der Starter- bzw. Elongationseinheit auf das <i>holo</i> -ACP einer modularen PKS.	20
Abbildung 3-7: Elongation des wachsenden Polyketids durch die PKS.	21
Abbildung 3-8: Mögliche Reduktions-, bzw. Dehydratisierungsreaktionen während der Polyketidsynthese.	21
Abbildung 3-9: Klassifizierung der Polyketidsynthasen.	22
Abbildung 3-10: <i>cis</i> -AT PKS und <i>trans</i> -AT PKS.	23
Abbildung 3-11: Pederin.	23
Abbildung 3-12: Naringenin-chalcon.	24
Abbildung 3-13: Katalysierte Reaktionen der Grunddomänen einer NRPS.	25
Abbildung 3-14: Pharmakologisch relevante nichtribosomale Peptide (NRP).	26
Abbildung 3-15: Bleomycin.	29
Abbildung 3-16: Der Stammbaum des Lebens nach Hug, L. A. <i>et al.</i> ⁸⁹	31
Abbildung 3-17: Schematischer Aufbau und Übersicht verschiedener Vertreter der Klasse <i>Demospongiae</i>	35
Abbildung 3-18: Chemotypen des Schwammes <i>Theonella swinhoei</i>	36
Abbildung 3-19: Bioaktive Naturstoffe aus dem Schwamm <i>Theonella swinhoei</i>	37
Abbildung 3-20: Phylogenetische Einordnung von ' <i>Candidatus Entotheonella factor</i> ' anhand verschiedenener Markergene.	40
Abbildung 3-21: Mikroskopische Aufnahme filamentöser Symbionten des Schwammes <i>Theonella swinhoei</i>	41
Abbildung 3-22: NRPS-PKS aus ' <i>Pseudomonas paederii</i> '.	44
Abbildung 3-23: Sekundärstruktur der 16S RNA.	46
Abbildung 3-24: Arbeitsschritte einer phylogenetischen Analyse.	48
Abbildung 3-25: Beschreibung monophyletischer, paraphyletischer, polyphyletischer Gruppen.	50
Abbildung 3-26: Übersicht über Methoden zur Rekonstruktion von phylogenetischen Stammbäumen.	51
Abbildung 3-27: Prinzip des bootstrap	52

Abbildung 5-1: Schließen der Lücken innerhalb von Scaffolds.	55
Abbildung 5-2: Orbiculamid A.	55
Abbildung 5-3: Schließung der Lücken im Orbiculamid-Gencluster.	56
Abbildung 5-4: Kombinatorische PCR zur Schließung von Sequenzlücken.	57
Abbildung 5-5: Biosynthesecluster aus 'Entotheonella factor' und 'Entotheonella gemina'. ¹³¹	59
Abbildung 5-6: Agarosegel zur Überprüfung der Nachweismethodik.	60
Abbildung 5-7: Screenen erster Umweltproben.	62
Abbildung 5-8: Alignment der 16S rRNA-Gensequenz aus Probe 5 und der Sequenz von 'E. factor'.	63
Abbildung 5-9 Schematische Darstellung der <i>nested</i> -PCR Strategie.	64
Abbildung 5-10: <i>Nested</i> -PCR zur Identifizierung von 'Candidatus Entotheonella' in Bodenproben.	64
Abbildung 5-11: Untersuchung von Wasserproben auf 'Candidatus Entotheonella'.	65
Abbildung 5-12: Ausschnitt aus einem 16S rRNA-Gensequenz Alignment zur Erstellung eines phylogenetischen Stammbaums am Beispiel der Sequenzen aus Bodenprobe 8.	69
Abbildung 5-13: Erste phylogenetische Einordnung von 'E. factor' (Dr. Susanne Schmitt)....	70
Abbildung 5-14: Erste phylogenetische Einordnung der 'Tectomicrobia'-Sequenzen aus Umweltproben.	71
Abbildung 5-15: Detailansicht der Klade „Schwamm-assoziierte Bakterien“.	72
Abbildung 5-16: Detailansicht der Klade 'Entotheonella'.	73
Abbildung 5-17: Verbreitung von 'Tectomicrobia' anhand einer Datenbankanalyse. ²¹³	74
Abbildung 5-18: Phylogenetische Analyse des Kandidatenphylums 'Tectomicrobia' mittels Neighbour Joining.	76
Abbildung 5-19: Phylogenetische Analyse des Kandidatenphylums 'Tectomicrobia' mittels Maximum Likelihood.	77
Abbildung 5-20: Vergleich der „Klade 1“ (NJ) mit den „Kladen A und B“ (ML).	78
Abbildung 5-21: Vergleich der „Klade 4“ (NJ) mit den „Kladen E und F“ (ML).	80
Abbildung 5-22: Vergleich der „Klade H“ (ML) mit den „Kladen 6 und 7“ (NJ).	82
Abbildung 5-23: Phylogenetische Analyse des Kandidatenphylums 'Tectomicrobia' von Eike Peters.	83
Abbildung 5-24: Untersuchungen von Wurzel und Pflanzenmaterial auf 'Candidatus Entotheonella'.	86
Abbildung 5-25: Isolierung von 'Candidatus Entotheonella' durch differentielle Zentrifugation.	87
Abbildung 5-26: Bakterienfraktionen nach differentieller Zentrifugation.	87
Abbildung 5-27: Prinzip der Dichtegradientenzentrifugation mittels Nycodenz.	88

Abbildung 5-28: Screening der Bakterienfraktionen nach differentieller Zentrifugation und Dichtegradientenzentrifugation.	89
Abbildung 5-29: Lichtmikroskopieaufnahmen einiger Bakterienfraktionen.	90
Abbildung 5-30: CARD-FISH der Bodenproben 5 und 26.	93
Abbildung 5-31: Zusammensetzung des Mikrobioms der Probe 5 vor und nach der Fraktionierung.	95
Abbildung 5-32: Zusammensetzung des Mikrobioms der Probe 26 vor und nach der Fraktionierung.	96
Abbildung 5-33: Zusammensetzung des Mikrobioms der Probe I vor und nach der Fraktionierung.	96
Abbildung 5-34: Prozentualer Anteil von ' <i>Candidatus Entotheonella</i> ' am Mikrobiom.	99
Abbildung 5-35: Sequenzalignment der bei der Amplikon-Sequenzierung gewonnenen ' <i>Candidatus Entotheonella</i> '-16S-Sequenzen vor der Aufreinigung.	100
Abbildung 5-37: PKS und NRPS-Cluster Contigs des <i>Symbiont 2</i> Genoms.	105
Abbildung 5-38: Weitere annotierte NRPS-Domänen einiger <i>Symbiont 2</i> -Contigs.	106
Abbildung 5-39: Assemblierte Cluster-Contigs.	107
Abbildung 5-40: Assemblierung der Contigs 42 und 73 ' <i>P. paederii</i> ' Genoms.	109
Abbildung 5-41: Assemblierung der Contigs 62 und 59 ' <i>P. paederii</i> ' Genoms.	109
Abbildung 5-42: Assemblierung der Contigs 95 und 26 ' <i>P. paederii</i> ' Genoms.	109
Abbildung 5-43: Assemblierung der Contigs 69, 1 und 51 des ' <i>P. paederii</i> ' Genoms.	110
Abbildung 5-44: Assemblierung der Contigs 72 und 22 des ' <i>P. paederii</i> ' Genoms.	110
Abbildung 5-45: Testexpression der PppE-A-Domäne.	111
Abbildung 5-46: Expression der PppE-A-Domäne.	112
Abbildung 5-47: Prinzip des [γ - $^{18}\text{O}_4$]-ATP-Assays nach Phelan et al.	113
Abbildung 5-48: Beispiel der Massenspektrometrische Darstellung des [γ - $^{18}\text{O}_4$]-ATP-Assays bei Akzeptanz des Substrats aus der Literatur.	114
Abbildung 7-1: GeneRuler 1 kb Plus DNA Ladder (Thermo Fisher Scientific).	136
Abbildung 7-2: pBluescript II SK (+/-).	138
Abbildung 7-3: pGEM-T Easy Vektor.	139
Abbildung 7-4: Dichtegradientenmedium Nycodenz.	148
Abbildung 7-5: Fluoreszenz-in-situ-Hybridisierung (FISH).	149
Abbildung 7-6: CARD-FISH.	149
Abbildung 10-1: Aufspaltung der „Klade 1“ im NJ-Stammbaum.	189
Abbildung 10-2: Aufspaltung der „Klade 2“ im NJ-Stammbaum.	190
Abbildung 10-3: Aufspaltung der „Klade 3“ im NJ-Stammbaum.	191
Abbildung 10-4: Aufspaltung der „Klade 4“ im NJ-Stammbaum.	192
Abbildung 10-5: Aufspaltung der „Klade 5“ im NJ-Stammbaum.	193

Abbildung 10-6: Aufspaltung der „Klade 6“ im NJ-Stammbaum.	194
Abbildung 10-7: Aufspaltung der „Klade 7“ im NJ-Stammbaum.	195
Abbildung 10-8: Aufspaltung des Kandidatengenus 'Ca. Entotheonella' („Klade 8“ im NJ-Stammbaum).....	196
Abbildung 10-9: Sequenzalignment A-Domäne PppE.	197
Abbildung 10-10: Sequenzalignment A-Domäne PppB.....	198
Abbildung 10-11: pHIS8-Vektor. ²³²	199
Abbildung 10-12: pWEB-Vektor.....	199
Abbildung 10-13: Anteile der Phyla am Mikrobiom vor und nach der Fraktionierung.	208
Abbildung 10-14: Prozentualer Anteil der Klasse/Ordnung/Familie Deltaproteobacteria am Mikrobiom der Proben vor und nach der Fraktionierung.	210
Abbildung 10-15: Beschaffenheit einiger untersuchter Bodenproben.....	212
Abbildung 10-16: MALDI-TOF- Analyse nach Expression des gesamten Genclusters.....	212
Abbildung 10-17: Wiederholung der MALDI-TOF Analyse.....	213
Abbildung 10-18: Wiederholung der MALDI-TOF Analyse nach Expression des gesamten Genclusters.	213
Abbildung 10-19: UV-Chromatogramm (254 nm) der Ethylacetat-Extraktionen der Zellen..	214
Abbildung 10-20: UV-Chromatogramm (220 nm) der Ethylacetat-Extraktionen der Zellen..	215
Abbildung 10-21: UV-Chromatogramm (254 nm) der Ethylacetat-Extraktion des Mediums.	216
Abbildung 10-22: UV-Chromatogramm (220 nm) der Ethylacetat-Extraktion des Mediums.	217
Abbildung 10-23: Global Natural Products Social Molecular networking (GNPS).....	218
Abbildung 10-24: MALDI-TOF Spektren verschiedener $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assays.	219

Tabellenverzeichnis

Tabelle 3-1: Open reading frames (ORFs) eines PKS-NRPS Genclusters aus 'P. paederi'. .45	45
Tabelle 5-1: Untersuchte Umweltproben.....61	61
Tabelle 5-2: ' <i>Candidatus Entotheonella</i> ' 16S rRNA-Gensequenzen verschiedener Bodenproben.....66	66
Tabelle 5-3: Phylogenetische Klassifizierung.....95	95
Tabelle 5-4: Anzahl der ' <i>Candidatus Entotheonella</i> ' -spezifischen Sequenzen vor und nach der Fraktionierung.98	98
Tabelle 5-5: AntiSMASH-Analyse des <i>Symbiont 2</i> Genoms.102	102
Tabelle 5-6: Getestete Substrate und deren Aktivierung durch die exprimierte A-Domäne. 114	114
Tabelle 7-1: Verwendete Organismen.123	123
Tabelle 7-2: Verwendete Vektoren.123	123
Tabelle 7-3: Verwendete Medien.....123	123
Tabelle 7-4: Verwendete Antibiotika.124	124
Tabelle 7-5: Verwendete Medienzusätze.....124	124
Tabelle 7-6: Puffer zur Herstellung chemisch kompetenter Zellen.125	125
Tabelle 7-7: PCR-Reaktionsansätze.....127	127
Tabelle 7-8: PCR-Programm Taq-Polymerase.127	127
Tabelle 7-9: PCR-Programm Phusion-Polymerase.....127	127
Tabelle 7-10: Primer zum Lückenschluß im 'E. factor'-Genom durch <i>Primer-Walking</i>128	128
Tabelle 7-11: Primer zum Lückenschluß im 'E. factor'-Genom.....129	129
Tabelle 7-12: Primer zum Lückenschluß im 'E. factor'-Genom durch Screenen einer Metagenom-Bibliothek.....130	130
Tabelle 7-13: Primer zum Lückenschluß innerhalb des 'E. factor' Plasmids.....131	131
Tabelle 7-14: Primer zum Nachweis von ' <i>Candidatus Entotheonella</i> '.....132	132
Tabelle 7-15: Primer zum Lückenschluß des 'P. paederi'-Genoms.....132	132
Tabelle 7-16: Primer zum Lückenschluß des <i>Symbiont 2</i> Genoms.....134	134
Tabelle 7-17: Ligationsansatz.....137	137
Tabelle 7-18: Ligationsansatz pGEM-T Easy.....139	139
Tabelle 7-19: Dephosphorylierung von Vektor-DNA.140	140
Tabelle 7-20: Restriktionsansatz.140	140
Tabelle 7-21: PCR-Programm zur Amplikon-Sequenzierung.....141	141
Tabelle 7-22: Puffer zur His-Tag Aufreinigung von Proteinen143	143
Tabelle 7-23: Assaypuffer des [γ - $^{18}\text{O}_4$]-ATP-Assay.....145	145
Tabelle 7-24: Verwendeten Sonde zur Fluoreszenz-in-situ-Hybridisierung.....150	150

Tabelle 7-25: Puffer CARD-FISH.....	151
Tabelle 10-1: Beschreibung der 16S rRNA-Sequenzen der NCBI-Datenbankrecherche. ...	171
Tabelle 10-2: Ergebnis der semiquantitativen Untersuchung der aufgereinigten Bodenproben.....	200
Tabelle 10-3: Mikrobiom der Bodenproben: Phyla.	201
Tabelle 10-4: Anzahl der Sequenzen und prozentualer Anteil der zehn häufigsten Spezies in den vier untersuchten Bodenproben.....	211
Tabelle 10-5: Ergebnis der Korngrößenbestimmung einiger untersuchter Bodenproben. ...	211

1 Zusammenfassung

Zunehmende Resistenzen und gleichzeitig sinkende Zahlen neu zugelassener Antibiotika lassen den Bedarf an neuen antibiotischen Wirkstoffen steigen. Aber auch in anderen medizinischen Bereichen wie z. B. der Onkologie ist der Bedarf an neuen pharmakologisch wirksamen Leitstrukturen hoch. Ein Ansatz, an neue bioaktive Molekülstrukturen zu gelangen, ist die Naturstoffforschung. Zwei Klassen von Naturstoffen, Polyketide und nichtribosomalen Peptide sind dabei von besonderem Interesse, da viele dieser Stoffe pharmakologische Wirkungen aufweisen. Diese werden meist als Sekundärmetabolite von symbiontisch lebenden Bakterien gebildet und nicht wie meist zunächst angenommen von den Wirtsorganismen wie z. B. Schwämmen oder anderen Invertebraten. So bieten diese meist noch unbekanntes Bakterienspezies ein enormes, oft noch ungenutztes Potenzial in der Wirkstoffforschung. In dieser Arbeit wurden einige dieser potenten Naturstoffproduzenten untersucht. Aus dem Schwamm *Theonella swinhoei*, aus dem bereits viele Naturstoffe bekannt waren, wurde in Vorarbeiten ein filamentöses Bakterium untersucht. Erste Experimente gaben Hinweise darauf, dass viele dieser Naturstoffe nicht vom Schwamm selbst sondern tatsächlich von diesem Bakterium produziert werden. Im Rahmen dieser Arbeit wurden in Zusammenarbeit mit der Arbeitsgruppe von Prof. Dr. Jörn Kalinowski (Universität Bielefeld, Verbundprojekt GenBioCom) und Dr. Agus Uria, Dr. Micheal Wilson und Nina Heycke das Genom dieses Symbionten sequenziert und verbleibende Sequenzlücken mithilfe kombinatorischer PCR geschlossen. So konnten viele der Naturstoffgencluster assembliert und '*Candidatus Entotheonella factor*' (lat. *factor*, der Produzent) als Produzent nahezu aller Polyketide und nichtribosomalen Peptide aus dem Schwamm *Theonella swinhoei* bestätigt werden. Erste phylogenetische Analysen dieses Naturstoffproduzenten gaben Anhaltspunkte dafür, dass weitere Mitglieder der Gattung '*Entotheonella*' auch in anderen Habitaten zu finden sind (z. B. NCBI GenBank: JN178754, AY493913, JN825509). Um die Verbreitung und das metabolische Potential dieser Spezies zu analysieren, wurden in dieser Arbeit verschiedene Umweltproben untersucht. Dabei zeigte sich, dass mehr als die Hälfte (62,9%) der untersuchten Bodenproben nah verwandte Spezies zu '*E. factor*' (Mitglieder desselben Genus, derselben Familie bzw. Ordnung) enthalten. Zur weiteren Untersuchung dieser bisher noch unbekanntes Spezies der Bodenproben wurde versucht, diese aus den Umweltproben zu isolieren. Mithilfe einer spezifischen Sonde für '*E. factor*' konnten in einem CARD-FISH-Experiment Bakterien in einigen Bodenproben markiert und visualisiert werden. Dieses Ergebnis konnte jedoch nicht reproduziert werden. Amplikon-Sequenzierungen zeigten, dass der Anteil an '*Candidatus Entotheonella*' - bzw. '*Candidatus Tectomicrobia*'-Spezies am Mikrobiom der Proben gering

ist (0,0008-0,002%). Versuche, diese Spezies zu separieren und aufzukonzentrieren, waren nur teilweise erfolgreich (max. Anteil dieser Bakterien am Mikrobiom auf 0,15%). Daher kann bisher keine Aussage über das Potential dieser Bakterien als Naturstoffproduzenten getroffen werden. Phylogenetische Analysen bestätigten, dass sowohl 'E. factor', als auch die verwandten Spezies aus den Bodenproben, nicht wie bisher vermutet dem Phylum der Proteobakterien angehören, sondern ein eigenes, bisher unbeschriebenes Kandidatenphylum bilden ('Tectomicrobia', lat.: *tegere*, sich verbergen, etwas schützen). Mithilfe einer phylogenetischen Analyse der aus den Bodenproben gewonnenen 78 16S rRNA-Gensequenzen und weiteren 562 Sequenzen einer Datenbankrecherche konnte das neue Kandidatenphylum 'Tectomicrobia' erstmals charakterisiert werden.

In Vorarbeiten wurde, ebenfalls aus dem Schwamm *Theonella swinhoei*, ein weiteres filamentöses Bakterium (*Symbiont 2*) isoliert. Erste Genom-Sequenzierungen und anschließende Analysen zeigten, dass auch im Genom dieses Bakteriums eine enorme Anzahl an Naturstoffgenclustern enthält. In dieser Arbeit konnten insgesamt 49 Contigs mit Fragmenten von Naturstoffclustern identifiziert werden, davon 36 PKS und NRPS Gencluster. Insgesamt konnten 39 A-Domänen, 11 KS-Domänen, 35 C-Domänen annotiert werden. Einige dieser Contigs wurden mittels kombinatorischer PCR verbunden.

Das Genom des als Pederin-Produzent bekannte *Paederus*-Symbiont '*Candidatus Pseudomonas paederii*' wurde in Vorarbeiten sequenziert. In dieser Arbeit konnten einige der verbliebenen Sequenzlücken durch kombinatorische PCR geschlossen werden. Das Genom beherbergt neben dem Biosynthesecluster für Pederin einen weiteren PKS/NRPS-Gencluster. Zur Vorhersage des Naturstoffes ist es notwendig, die Spezifität der A-Domäne und damit der in der Synthese eingebauten Bausteine zu kennen. Dazu sollte die bioinformatische Vorhersage der Spezifität *in vitro* experimentell überprüft werden. In dieser Arbeit konnte eine der beiden enthaltenen A-Domänen heterolog exprimiert werden. Aufgrund fehlender Aktivität konnte jedoch keine Aussage zur Spezifität getroffen werden.

2 Abstract

Increasing resistance and a decreasing number of newly approved antibiotics raise the need for new antibiotic drugs. Also in other medical fields, e.g. oncology, the necessity of new pharmacologically effective leading structures is high. One approach to find such molecules is the analysis of natural products. Two classes of natural products, polyketides and nonribosomal peptides, are of particular interest since many of these substances have pharmacological properties. It has been shown that these molecules are produced as secondary metabolites by symbiotically living bacteria and not, as often assumed, by the host organisms, e.g. sponges or other invertebrates. Thus, these mostly unknown bacteria offer an enormous, unused potential in drug discovery. In this thesis, some of these potent natural product producers have been studied. In previous work a filamentous bacterium was isolated from the sponge *Theonella swinhoei*. First experiments showed that many of the natural products known from this sponge are actually produced by this bacterium. In this work, in collaboration with the group of Prof. Dr. Jörn Kalinowski (University Bielefeld, "Verbundprojekt" GenBioCom) and Dr. Agus Uria, Dr. Micheal Wilson and Nina Heycke, the genome of this bacterium was sequenced and many of the remaining sequence gaps were closed by combinatorial PCR. By this procedure, many of the natural product gene clusters could be assembled and their bacterial origin of these polyketides and nonribosomal peptides could be confirmed. Because of its ability to produce an enormous number of natural products this bacterium was named '*Candidatus Entotheonella factor*' (lat. *factor*, the producer). First phylogenetic analyses showed evidence that these '*Candidatus Entotheonella*' -bacteria can also be found in other habitats. To examine the dissemination and the metabolic potential of these species, various environmental samples were analysed. It was found that more than half of the soil samples (62.9%) contained related species to '*E. factor*' (belonging to the same genus, order or family). In order to characterize these species, which have not been described yet, isolation was attempted. With the help of a specific probe in a CARD-FISH experiment, bacteria could be labeled and visualized in some samples. But these results could not be reproduced. Amplicon sequencing could show that the proportion of '*Candidatus Entotheonella*' species in the microbiome of the analysed soil samples is low (0,0008-0,002%). Trials to isolate and concentrate these bacteria were only partially successful (maximum increase in the proportion to 0.15%). For this reason, it is impossible to make any statement about the potential of these bacteria as natural product producers. Phylogenetic analyses confirmed that '*E. factor*' as well as its related species from soil do not belong to the phylum of the deltaproteobacteria, as hitherto presumed, but form their own, so far undescribed candidate phylum ('Tectomicrobia'; *latin*: tegere, conceal, protect something).

By means of a phylogenetic analysis of the seventy 16S rRNA gene sequences isolated from the soil samples and further 570 sequences from a database investigation, the new candidate phylum 'Tectomicrobia' could be characterized for the first time.

In previous experiments, another filamentous bacterium (*Symbiont 2*) was isolated from the sponge *Theonella swinhoei*. Genome sequencing and subsequent analysis showed that this genome also harbours an enormous number of natural product gene clusters. A total of 49 contigs containing fragments of natural product clusters, including 36 PKS and NRPS gene clusters, could be identified in this work. A total of 39 A domains, 11 KS domains and 35 C domains have been annotated so far. Some of the contigs have been assembled by using combinatorial PCR.

Previously, the genome of the *Paederus*-symbiont 'Candidatus *Pseudomonas paederii*', known as the producer of Pederin, could be sequenced. In this work some of the remaining sequence gaps could be closed by combinatorial PCR. In addition to the pederin biosynthesis gene cluster, the genome contains another PKS/NRPS gene cluster. For the prediction of the natural product, it is necessary to know the specificity of the A domain and thus the used building blocks in the synthesis. For this purpose, the bioinformatic prediction of the specificity should be tested experimentally *in vitro*. In this work one of the two contained A-domains was heterologously expressed. However, due to the lack of activity, no specificity could be detected.

3 Einleitung

Sonar, Echolot und Propeller, aber auch solch einfache Alltagsgegenstände wie Netze oder der Klettverschluss - viele technische Dinge haben wir Menschen uns im Laufe der Zeit von der Natur abgeschaut. Mittlerweile hat sich daraus sogar eine eigene Wissenschaft entwickelt, die Bionik. Aber auch in anderen Bereichen ist es nicht neu, sich bereits bestehender Prinzipien aus der Umwelt zu bedienen. So nutzt man auch in der Medizin das vorhandene Potential der Natur, um an neue Wirkstoffe zu gelangen. Bereits vierzig Jahrhunderte vor Entdeckung des Penicillins durch Alexander Fleming im Jahr 1928 behandelten die alten Ägypter oberflächliche Entzündungen erfolgreich mit verschimmeltem Brot.^{1,2} Wie wir heute wissen, sind es die von einigen Schimmelpilzen produzierten antibiotischen Substanzen, die bakterielle Infektionen solcher Hautläsionen verhindern. Aufgrund zunehmender Resistenzen bakterieller Erreger gegen viele der heute verwendeten Antibiotika ist der Bedarf an neuen Wirkstoffen ungebrochen.³ Allein in Europa sterben nach Zahlen der WHO von 2014 jährlich ca. 25000 Menschen aufgrund resistenter Erreger.³ Auch wirtschaftlich ist die Belastung aufgrund von Resistenzen enorm. Nach Schätzungen liegt diese allein in der EU bei ca. 1,5 Milliarden Euro.³ Während die Zahl der Antibiotika-resistenten Bakterien zunimmt, sinkt gleichzeitig die Zahl der Zulassungen neuer Antibiotika drastisch.⁴ Gründe dafür sind zum einen die für die Pharmaindustrie ökonomisch wenig attraktive Entwicklung neuer Antibiotika. Zum anderen machen es Resistenzbildungen z. B. durch Veränderung der Targetstruktur notwendig, dass neue Wirkstoffe auf unterschiedliche Strukturen der Bakterienzellen wirken sollten. Dies stellt hohe Ansprüche an die Molekülstruktur des Wirkstoffes. Zudem änderte die Pharmaindustrie in den 1990er Jahren ihre bisherige Strategie der Naturstoffforschung zur Akquirierung neuer Antibiotika, hin zum Target-abhängigen Screenen bereits vorhandener Molekül-Bibliotheken.⁵ Jedoch konnten seit dem Ende der 80iger Jahre keine neuen Antibiotika zugelassen werden (Abbildung 3-1). Doch einige Pharmakonzerne reagieren bereits mit unterschiedlichen Ansätzen. Unter anderem erlebt die Naturstoffforschung eine Renaissance. Nicht nur in Forschungseinrichtungen, sondern auch in der Industrie wurde das bisher ungenutzte Potential erkannt, das besonders bisher nicht kultivierbare und damit wenig erforschte Bakterien bieten.⁵ Einige dieser Bakterien produzieren antibiotisch wirksame Substanzen aber auch andere pharmakologisch wirksame Naturstoffe, die, wie die meisten Naturstoffe eine größere strukturelle Diversität aufweisen als es durch klassische oder kombinatorische Synthesen erreicht werden kann.⁶ Trotzdem sind diese wichtigen Naturstoffproduzenten kaum erforscht. Diese Arbeit soll einen Beitrag dazu leisten, dies zu ändern und einen

Einblick in diese sehr interessanten und immer wichtiger werdenden Mikroorganismen als natürliche Quelle neuer Wirkstoffe zu geben.

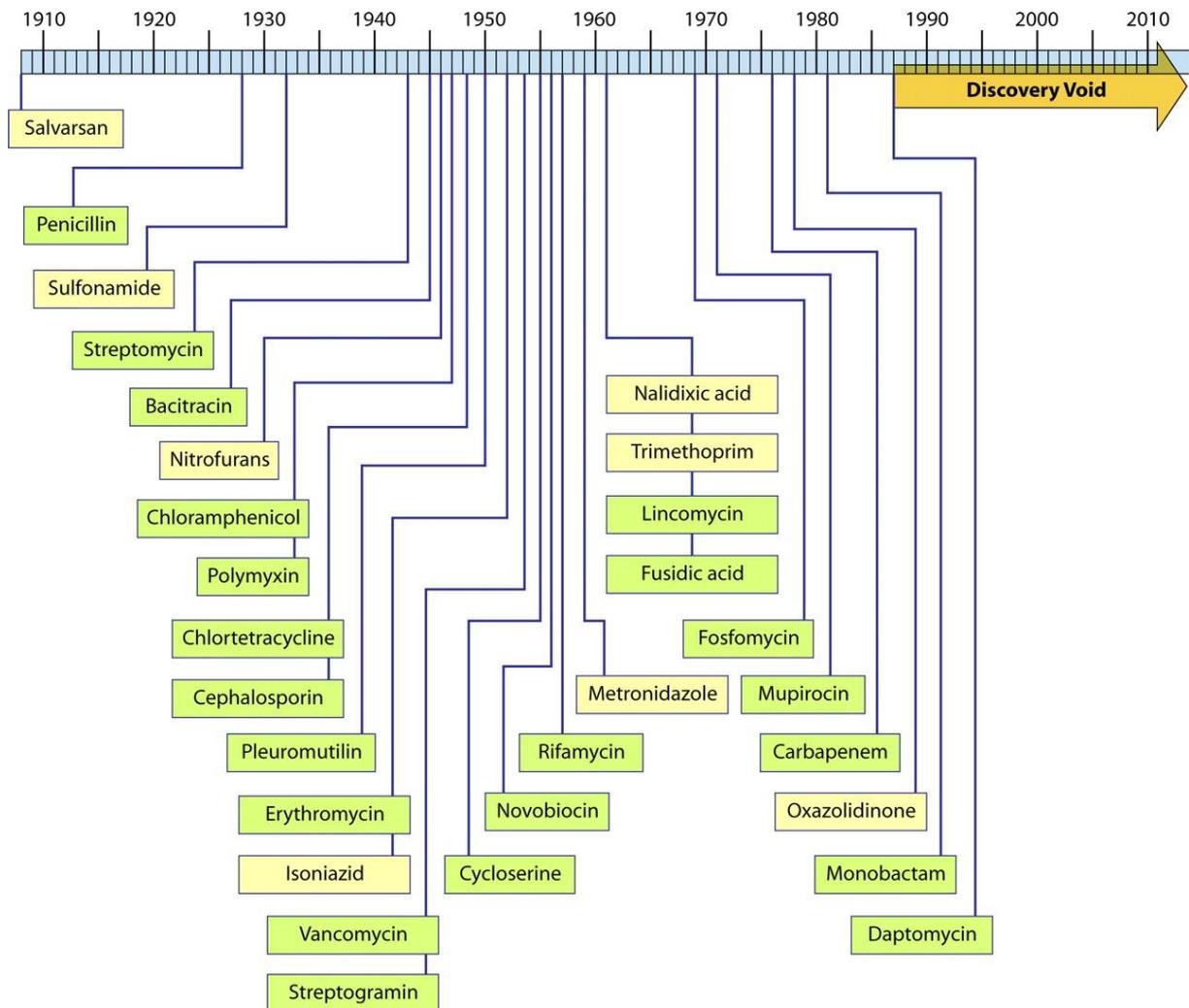


Abbildung 3-1: Darstellung des sogenannten „discovery void“ in der Antibiotikaentwicklung.⁴

3.1 Die Biosynthese mikrobieller Sekundärmetabolite

Viele pharmakologisch wirksame, mikrobielle Sekundärmetabolite gehören zu den Naturstoffklassen der Polyketide und nichtribosomalen Peptide bzw. sind Hybride dieser beiden Klassen. Das folgende Kapitel soll einen Einblick in die Biosynthese dieser oft sehr komplexen Moleküle geben.

3.1.1 Die Polyketid-Biosynthese

Namensgeber und Entdecker der Polyketide war der Chemiker James Collie im Jahr 1893.⁷ Er postulierte, dass die Synthese dieser neuen Naturstoffklasse über die Zyklisierung von Poly- β -keto-Derivaten aus „Polyketene“.⁸ Den endgültigen Beweis dieser Synthese lieferte Birch im Jahr 1967 und prägte den Namen Polyketide.⁹ Polyketide zeichnen sich durch ihre

große strukturelle Vielfalt aus. Die daraus resultierende Spanne an pharmakologischen Eigenschaften, die diese Naturstoffe aufweisen, reicht von antibiotisch (wie z. B. Erythromycin¹⁰ **1**) und antimykotisch (wie z. B. Amphotericin B¹¹ **2**), über antiparasitär (wie z. B. Avermectin⁶ **3**) und cholesterinsenkend (wie z. B. Lovastatin¹² **4**), bis hin zu zytostatisch (wie z. B. Doxorubicin¹³ **5**) und immunsuppressiv (wie z. B. Rapamycin¹⁴ **6**, Cyclosporin A¹⁵). Diese Eigenschaften macht diese Klasse v.a. für die medizinisch-pharmakologische Forschung interessant. Beispielhaft sind einige Vertreter in Abbildung 3-2 dargestellt. Polyketide werden als Sekundärmetabolite von Pilzen, Pflanzen und Bakterien produziert. Sie dienen u.a. der Verteidigung und Kommunikation, meist ist ihre ökologische Funktion jedoch unbekannt.^{16,17}

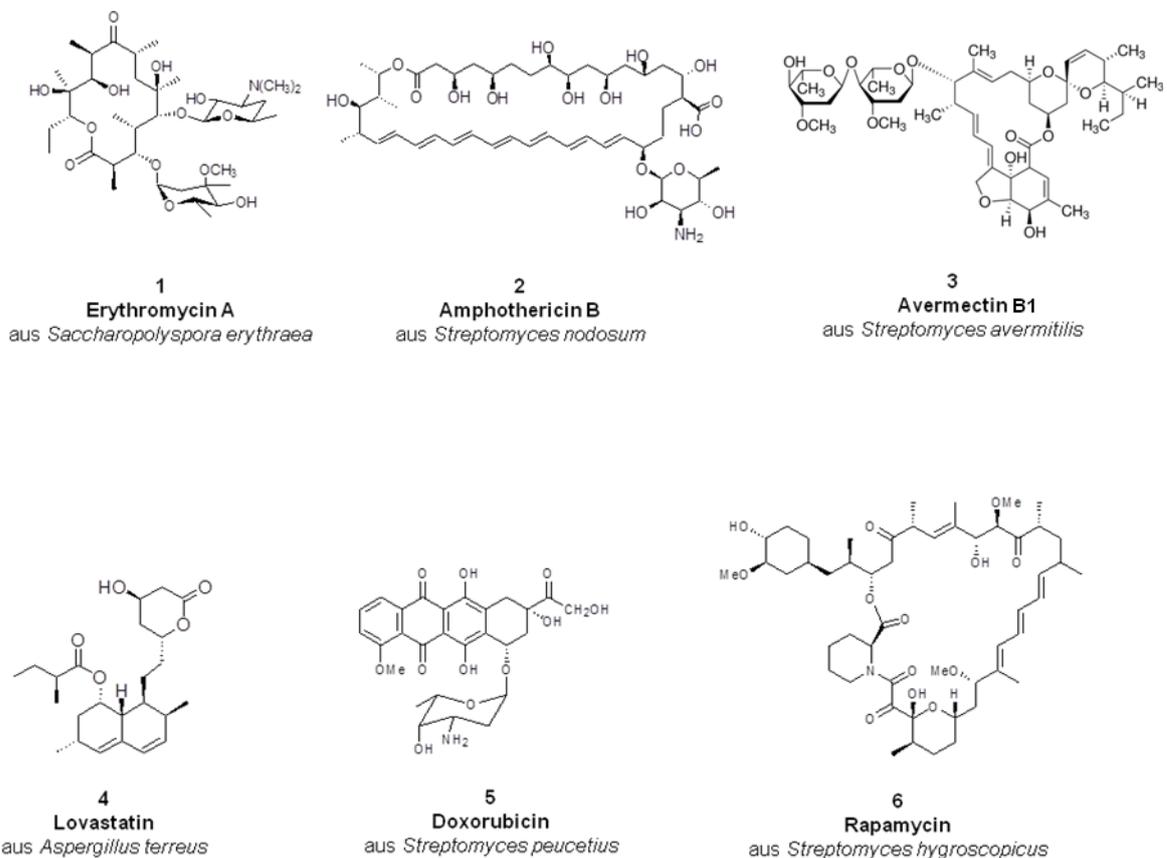


Abbildung 3-2: Strukturelle Vielfalt der Polyketide.^{6,10-14,18}

Bemerkenswert ist, dass die strukturelle Diversität nur durch die Auswahl, Verknüpfung und Modifikation einfacher Carbonsäure-Bausteine durch Multienzymkomplexe, den sogenannten Polyketidsynthasen (PKS), entsteht.^{6,15} Die Funktionsweise der PKS ähnelt dabei sehr der von Fettsäuresynthasen (FAS). Diese katalysieren eine Claisen-artige Verknüpfung von Acetyl-CoA und Malonyl-CoA-Monomeren zu langen Fettsäureketten.^{6,19}

PKS folgen dem gleichen Syntheschema, sind in ihrer Substratspezifität aber wesentlich flexibler und können neben den Substraten der FAS (**7** und **8**) weitere aktivierte Substrate prozessieren (Abbildung 3-3).²⁰

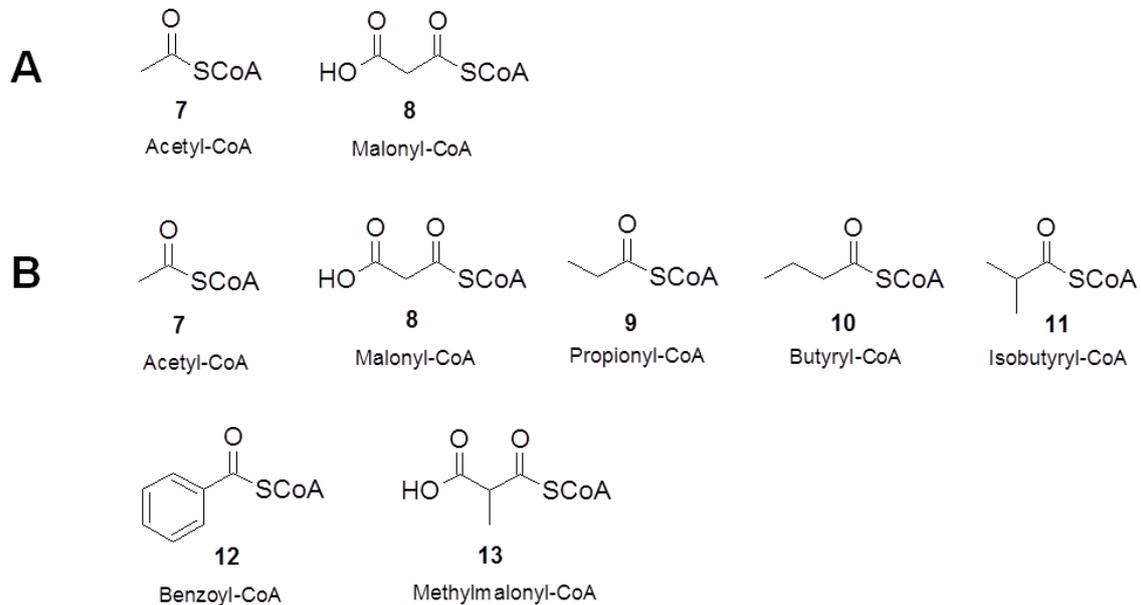


Abbildung 3-3: Substrate von FAS (A) und PKS (B).²⁰

Im Folgenden soll die Biosynthese von Polyketiden anhand der modularen Typ I Polyketidsynthese (Typ I PKS) beschrieben werden. Auf die Unterschiede im Biosynthesemechanismus, nach denen die Polyketidsynthasen klassifiziert werden, wird in Kapitel 3.1.1.1 näher eingegangen.

Während der Biosynthese sind die aktivierten Bausteine kovalent über eine Thioesterbindung eines 4'-Phosphopantethein-Linkers (PPant) an ein Acylträgerprotein (engl.: Acylcarrierprotein, ACP) gebunden. Dazu wird der PPant-Rest in einer Phosphopantetheintransferase (PPTase)-katalysierten Reaktion von CoA auf einen konservierten Serinrest des ACP übertragen. Das ACP wird so von der inaktiven *apo*-Form in das aktive *holo*-ACP überführt (Abbildung 3-4).^{22,23}

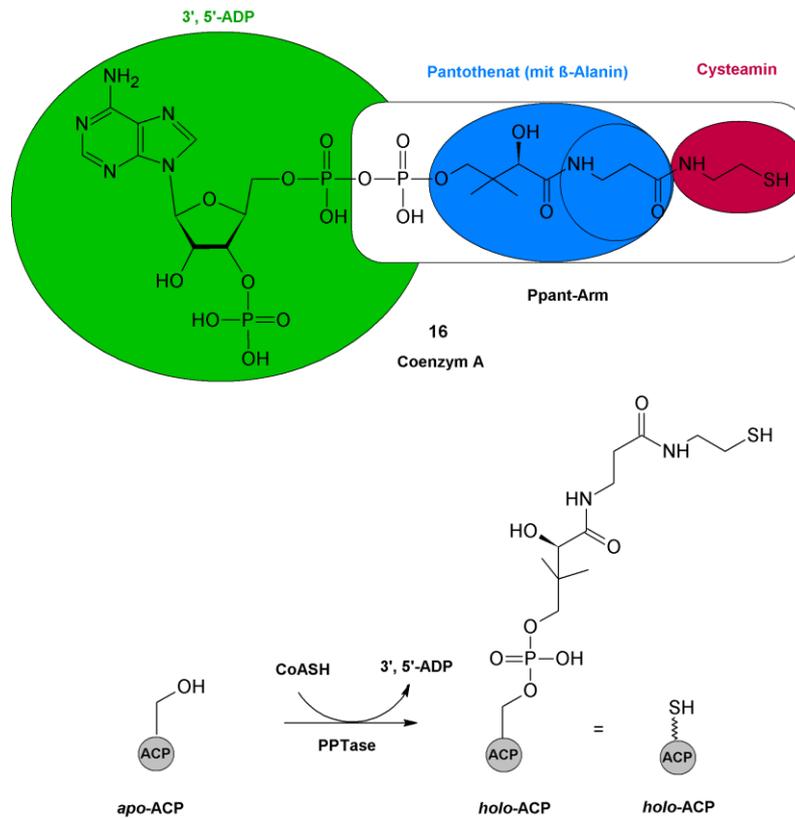


Abbildung 3-4: Aktivierung der ACP-Domäne. Schematisch dargestellt ist die PPTase-katalysierte Aktivierung des apo-ACP zum holo-ACP. Der Phosphopantetheinarm ist als gezackte Linie mit abschließender Thiolgruppe dargestellt. Ebenfalls abgebildet ist die Coenzym A (CoA oder CoASH)-Struktur die sich aus mehreren Teilstrukturen (farbig hinterlegt) zusammensetzt.^{23,24}

Die Übertragung der Starter- bzw. Elongationseinheiten auf das holo-ACP wird durch eine Acyltransferase (AT) katalysiert (Abbildung 3-5).

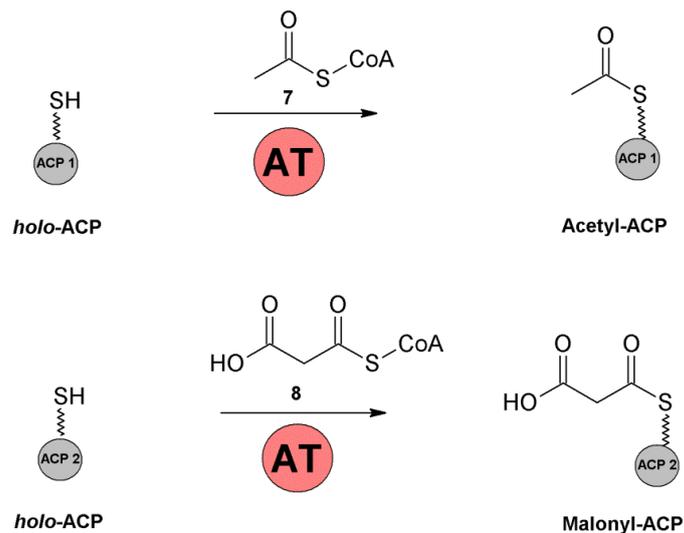


Abbildung 3-5: Übertragung der Starter- bzw. Elongationseinheit auf das holo-ACP einer modularen PKS. ACP 1 stellt stark vereinfacht das ACP des Startermoduls dar. ACP 2 stellt das ACP des darauffolgenden Moduls (Elongationsmodul) dar.²⁴

Vom ACP wird die Startereinheit nun auf den konservierten Cysteinrest einer Ketosynthase (KS) übertragen.²³ Die KS katalysiert dann in einer Claisen-Kondensation die irreversible Verknüpfung der Startereinheit mit einer Elongationseinheit unter Freisetzung von Kohlendioxid.²¹ Die Elongationseinheit wird von einem weiteren ACP zur Verfügung gestellt (Abbildung 3-6).

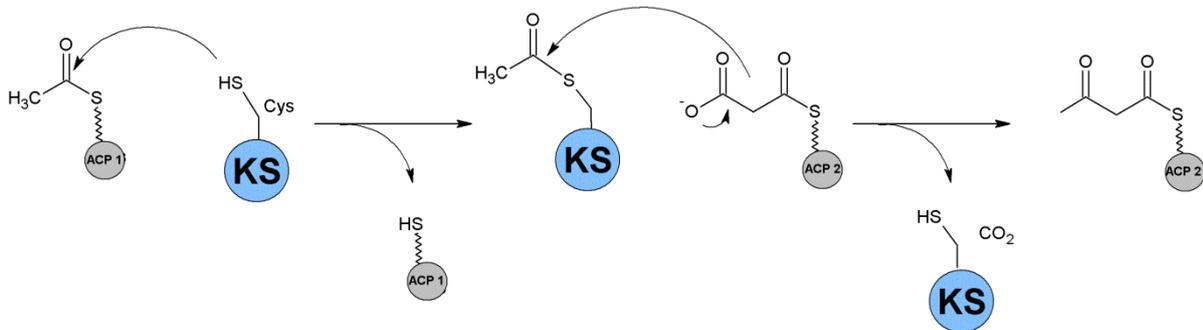


Abbildung 3-6: Elongation des wachsenden Polyketids durch die PKS. Abgebildet ist die Übertragung der Startereinheit auf die KS und die Verknüpfung dieses Starterbausteins mit einem Elongationsbaustein. Dabei stellen ACP 1 und ACP 2 Acylcarrierproteine zweier aufeinander folgender Module dar. Bei der Elongation der wachsenden Kette ist statt Acetyl-CoA das entsprechende Intermediat an ACP 1 gebunden.²¹

Neben dieser „Grundausstattung“ eines PKS-Moduls (aus AT, KS und ACP) können optional bei Typ I PKS weitere Domänen vorhanden sein. So können neben Methyltransferasen (MT), die Methylgruppen in das Polyketid einführen, auch weitere Domänen vorhanden sein, die die β -Ketogruppen teilweise oder vollständig reduzieren bzw. dehydratisieren können. Die Ketogruppe kann durch eine Ketoreduktase-Domäne (KR) zu einer Hydroxylgruppe reduziert werden. Eine Dehydratase-Domäne (DH) kann dann unter Wasserabspaltung die Bildung einer Doppelbindung katalysieren. Schließlich kann in einem weiteren Reaktionsschritt die Doppelbindung zu einer Einfachbindung reduziert werden. Diese Reaktion wird durch eine Enoylreduktase (ER) katalysiert.²¹

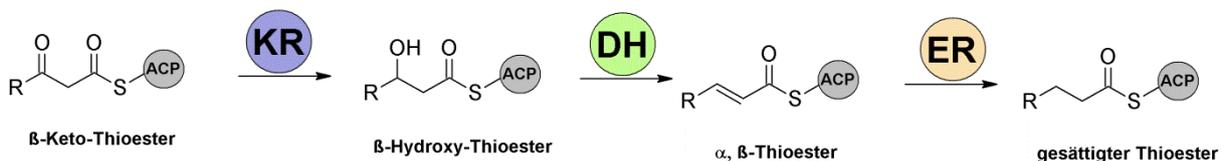


Abbildung 3-7: Mögliche Reduktions-, bzw. Dehydratisierungsreaktionen während der Polyketidsynthese. Ketoreduktasen (KR), Dehydratasen (DH) und Enoylreduktasen (ER) können die Ketogruppe sowohl vollständig als auch teilweise reduzieren.²¹

Während bei der Biosynthese von Fettsäuren die Ketogruppen nach jedem Elongationsschritt in der Regel vollständig reduziert werden, sind die Reduktionsschritte bei der Polyketidbiosynthese durch eine Typ I PKS optional.^{6,7,24,25} Dies trägt ebenfalls zur strukturellen Diversität der Polyketide bei.

Hat die wachsende Kette den gesamten Multienzymkomplex durchlaufen, wird es schließlich durch eine Thioesterase-Domäne (TE) abgespalten.²⁶ Nach der Biosynthese des Polyketids durch die PKS kann das Molekül noch weiter durch sogenannte Tailoring-Enzyme modifiziert werden. Diese können Zucker-, Methyl-, und Hydroxylgruppen in das Polyketid einführen.²⁷

3.1.1.1 Klassifizierung der Polyketidsynthesen

Polyketidsynthesen können anhand ihres Funktionsmechanismus in drei unterschiedliche Typen (Typ I-III) eingeteilt werden. Außerdem unterscheidet man iterative und modulare Typen anhand ihrer Verwendung der einzelnen Domänen. Abbildung 3-8 gibt eine Übersicht über die verschiedenen PKS-Typen.

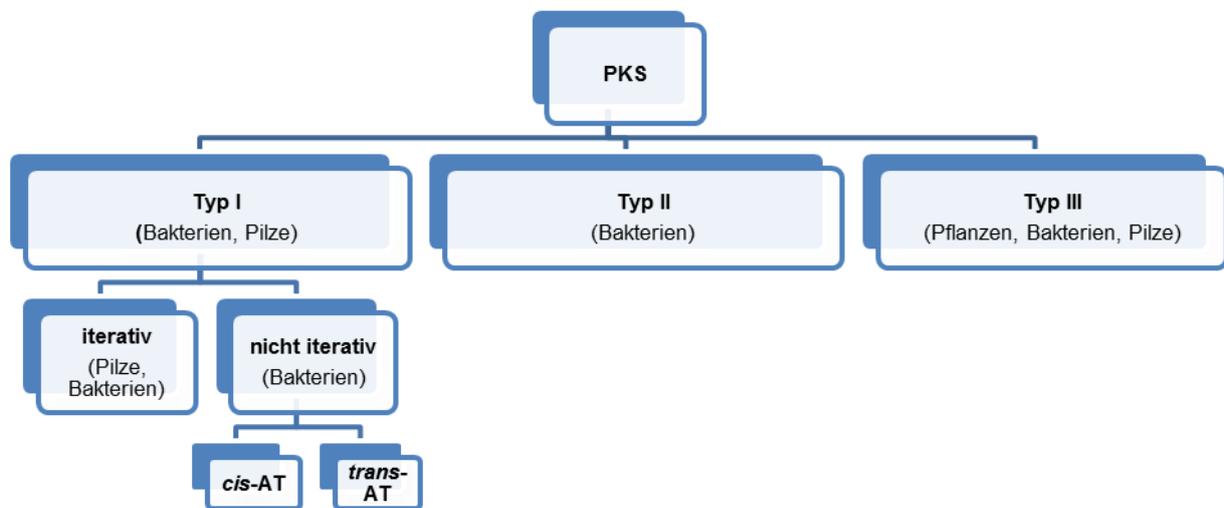


Abbildung 3-8: Klassifizierung der Polyketidsynthesen.

Typ I-PKS

Modulare Typ I PKS

Modulare Typ I-PKS bestehen aus festen Enzymsätzen, den sogenannten Modulen, die, neben weiteren optionalen Domänen mindestens aus einer KS-, einer AT- und einer ACP-Domäne bestehen. Jedes Modul verlängert die wachsende Kette normalerweise um je einen neuen Baustein. Dadurch bestimmen der modulare Aufbau und die Domänenzusammensetzung der Module unmittelbar die Struktur des Polyketids. Dies ermöglicht eine Vorhersage der Molekülstruktur des Polyketids anhand des PKS-Aufbaus und umgekehrt (Kolinearitätsregel).²⁸ Die Kolinearitätsregel lässt sich allerdings nur für eine Subklasse der modularen Typ I PKS anwenden, den sogenannten *cis*-AT PKS. Bei *cis*-AT PKS ist, wie oben beschrieben, neben einer KS- und einer ACP-Domäne in jedem Modul eine AT-Domäne vorhanden. Anders verhält es sich bei den *trans*-AT PKS, bei denen die AT-Domänen innerhalb der Module fehlen. Hier übernehmen freistehende Enzyme mit AT-Aktivität das Beladen der Module.^{29,30}

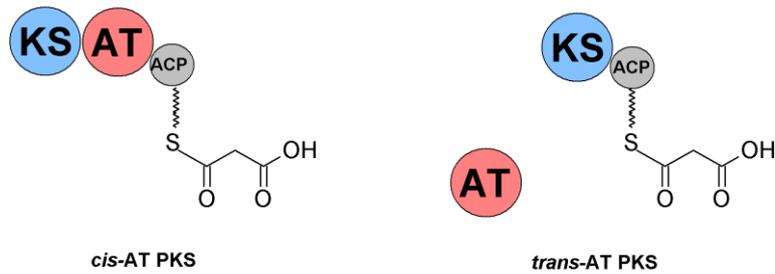


Abbildung 3-9: *cis*-AT PKS und *trans*-AT PKS.^{29,30}

Auch das PKS/NRPS-Hybrid des Polyketids Pederin enthält solche *trans*-AT Module. Pederin (Abbildung 3-10) wird von einem Symbionten des Käfers *Paederus fuscipes* produziert.³¹ Auf diesen Symbionten wird in Kapitel 3.3.2 näher eingegangen.

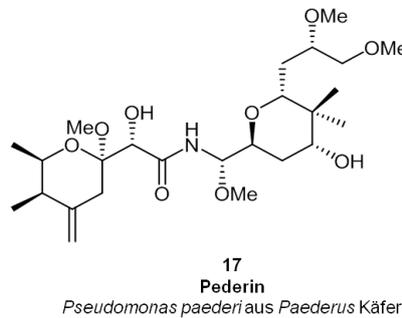


Abbildung 3-10: Pederin.³¹

Iterative Typ I-PKS

Iterative Typ I-PKS bestehen, wie auch die zuvor beschriebenen modularen PKS, aus einer Abfolge von Enzymdomänen, die allerdings nur ein einzelnes Modul bilden. Dieses Modul wird, wie der Name dieser PKS schon erahnen lässt, iterativ, also für jeden Elongationszyklus wiederholt verwendet.³² Anders als in den ähnlich funktionierenden FAS vom Typ I müssen nicht alle optionalen Domänen wie KR, DH, ER oder MT in jedem Zyklus aktiv sein. So kann auch hier die Ketogruppe sowohl unvollständig als auch vollständig reduziert sein. Daher werden PKS dieses Typs in nicht reduzierende (engl.: *non-reducing*, NR), teilweise reduzierende (engl.: *partially reducing*, PR) und stark reduzierende (engl.: *highly reducing*, HR) PKS unterteilt.^{33,34} Anders als die modularen Typ I-PKS kommen iterative PKS meist in Pilzen vor.³⁵ Ein von diesem PKS-Typ synthetisiertes Polyketid ist Lovastatin (Abbildung 3-2, 4) aus dem Pilz *Aspergillus terreus*.³⁶

Typ II-PKS

Typ II-PKS ähneln in ihrem Aufbau den Typ II-FAS in Bakterien und Pflanzen. Sie setzen sich aus wenigen monofunktionalen Einzelproteinen zusammen, die, anders als bei anderen PKS, auf separaten Genen codiert werden.³⁷ Eine Typ II-PKS besteht dabei mindestens aus zwei KS-ähnlichen Einheiten ($KS\alpha$ und $KS\beta$, auch englisch *chain-length-factor*, CLF genannt) und einem ACP. Optional können weitere enzymatische Einheiten wie KR, Zyklasten, Aromatasen, Oxygenasen, Glycosyltransferasen und Methyltransferasen vorhanden sein.³⁸ Ein bekanntes, von diesem PKS Typ synthetisiertes Polyketid ist das in der Tumorthherapie eingesetzte Doxorubicin (Abbildung 3-2, **5**).^{39,40}

Typ III-PKS

Anders als bei den bisher beschriebenen PKS-Typen agieren die Typ III-PKS (Chalcon-Synthase-ähnliche PKS) ACP-unabhängig, d.h. sie nutzen die freien CoA-Ester direkt als Substrat. Ihre funktionellen Einheiten bestehen aus Homodimeren, die iterativ genutzt werden.⁴¹ Diese katalysieren sowohl Decarboxylierung und Kondensation als auch Zyklisierung und Aromatisierung. Durch ihre hohe Substrattoleranz und durch Variation der Elongationszyklen und des Zyklisierungsmechanismus können diese Enzymkomplexe eine Vielzahl strukturell unterschiedlicher Moleküle synthetisieren.⁴² Weitere Modifikation nach der Synthese durch die PKS tragen zur Diversität dieser Polyketide bei. Ein Beispiel sekundärer Pflanzenstoffe ist das Flavonoid Naringenin-chalcon (Abbildung 3-11, **18**).⁴¹

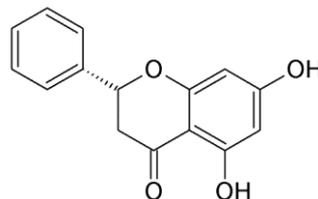


Abbildung 3-11: Naringenin-chalcon.⁴¹

3.1.2 Nichtribosomale Peptidsynthetasen

Neben den bisher beschriebenen PKS gibt es noch eine weitere Gruppe multimodularer Enzymkomplexe, der eine bedeutende Rolle in der Arzneimittelforschung zukommt: die nichtribosomalen Peptidsynthetasen (NRPS).⁴³ Diese Multienzymkomplexe sind in der Lage, Peptide mit einer wesentlich höheren Strukturvielfalt als ribosomale Peptide zu synthetisieren. Diese Strukturdiversität wird durch den zusätzlichen Einbau von nicht-proteinogenen Aminosäuren (AS), heterozyklischen Elementen, Fettsäure- und Glycosylresten und durch Methylierungen erzeugt.⁴⁴ Bei der ribosomalen Peptidsynthese werden hingegen ausschließlich proteinogene Aminosäuren in das Peptid eingebaut.⁴⁵ Eine

Erhöhung der Diversität kann hier durch Posttranslationale Modifikationen (PTMs) erzeugt werden. Auch die NRPS bestehen wie die modularen Typ I-PKS aus einzelnen Enzymdomänen, die sich zu einem Modul zusammensetzen und die wachsende Peptidkette um je einen Elongationsbaustein verlängern. Die Minimalbesetzung eines solchen Moduls besteht aus einer Adenylierungsdomäne (A-Domäne), einem Peptidylcarrierprotein (PCP oder T) und einer Kondensationsdomäne (C-Domäne).⁴⁶ Analog zur AT-Domäne der PKS ist die A-Domäne für die Auswahl, die Aktivierung und die Beladung der Aminosäure auf die NRPS verantwortlich. Die PCP-Domäne überträgt, wie die ACP-Domäne, die wachsende Kette von Modul zu Modul und verankert sie an der Synthase. Die C-Domäne verknüpft die Elongationsbausteine mit der wachsenden Kette.⁴⁷ Anders als die KS-Domäne katalysiert die C-Domäne allerdings die Bildung von Peptidbindungen, da es sich bei den Bausteinen um Aminosäuren und nicht wie bei der Bildung von Polyketiden um Carbonsäuren handelt.⁴⁸ Das letzte Modul einer NRPS bildet, wie auch bei einer PKS, eine Thioesterasedomäne (TE-Domäne), die das fertige Peptid vom Enzymkomplex ablöst.²⁶ Optional können Epimerase-(E), Methyltransferase (MT) und Zyklisierungsdomänen (Cy) innerhalb einer NRPS vorhanden sein.

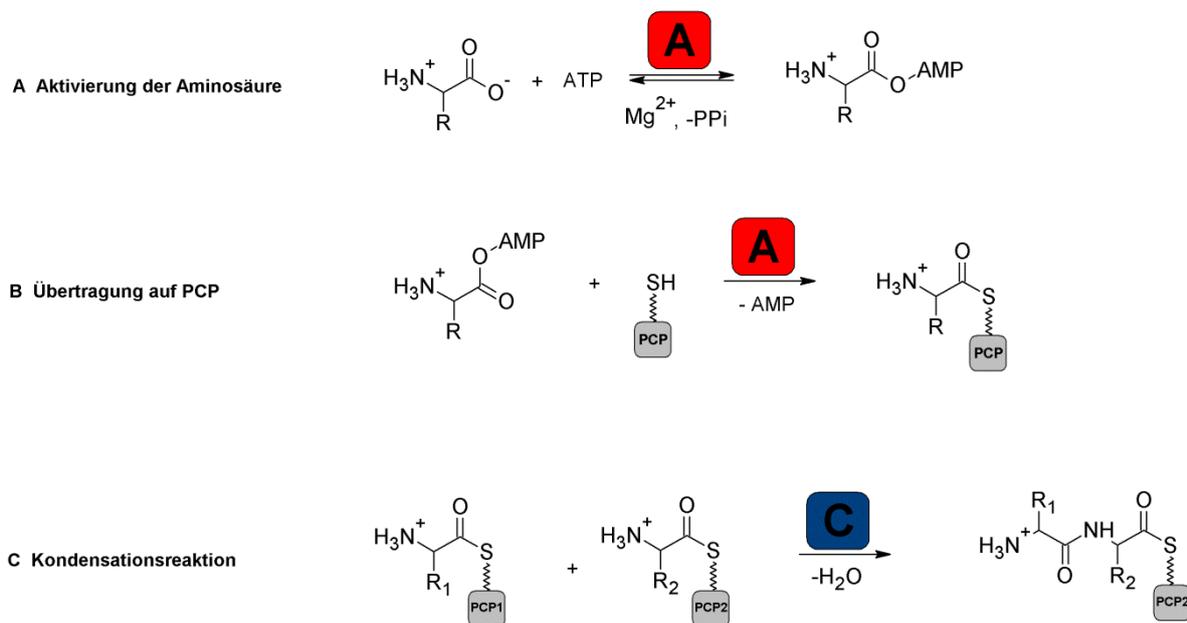


Abbildung 3-12: Katalysierte Reaktionen der Grunddomänen einer NRPS. Nach Aktivierung der Aminosäure durch die A-Domäne (A) wird diese auf eine PCP-Domäne übertragen (B). In einer Kondensationsreaktion verknüpft eine C-Domäne den neuen Aminosäurebaustein mit der wachsenden Kette (C).⁴⁶⁻⁴⁸

Wie bei den modularen PKS kann auch bei den meisten NRPS aufgrund ihrer Modularchitektur die Kolinearitätsregel angewendet werden. Wie bereits in Kapitel 3.1.1.1 für PKS beschrieben, kann dabei von der Modulabfolge auf die Molekülstruktur geschlossen werden und umgekehrt.^{28,49} Um eine solche Vorhersage treffen zu können, ist es wichtig,

Informationen über die Substratspezifität der einzelnen Module zu haben. Diese Substratspezifität wird maßgeblich vom aktiven Zentrum der Adenylierungsdomäne (A-Domäne) definiert. Auf die Rolle der A-Domänen und ihrer Substratspezifität wird in Kapitel 3.1.2.1 näher eingegangen. Der Entschlüsselung dieses sogenannten „nichtribosomalen Codes“ ist es zu verdanken, dass heute oft mithilfe von bioinformatischen Methoden die verwendeten Aminosäurebausteine und somit die Molekülstruktur anhand der Spezifität der A-Domänen vorhergesagt werden kann.^{50–52} So können Gencluster bereits bekannter Peptide identifiziert und durch sogenanntes „genome mining“ neue Naturstoffe anhand ihrer Gencluster prognostiziert und ggf. isoliert werden.^{53–55} Allerdings gibt es auch bei NRPS Ausnahmen, bei denen die Kolinearitätsregel nicht angewendet werden kann. Beispielsweise werden bei der Synthese von Di- oder Trimeren alle Module zweimal durchlaufen und die Monomere schließlich von der Thioesterase verbunden.⁵⁶ Wie die Polyketide weisen auch viele nichtribosomale Peptide (NRP) häufig pharmakologische Eigenschaften auf, was sie für die pharmakologische und medizinische Forschung interessant macht.⁴³ Beispiele solcher pharmakologisch wirksamer NRP sind das Glycopeptid Vancomycin,^{20,57} das als Antibiotikum eingesetzt wird und Cyclosporin (**21**), das als Immunsuppressivum Verwendung findet⁵⁸ (Abbildung 3-13).

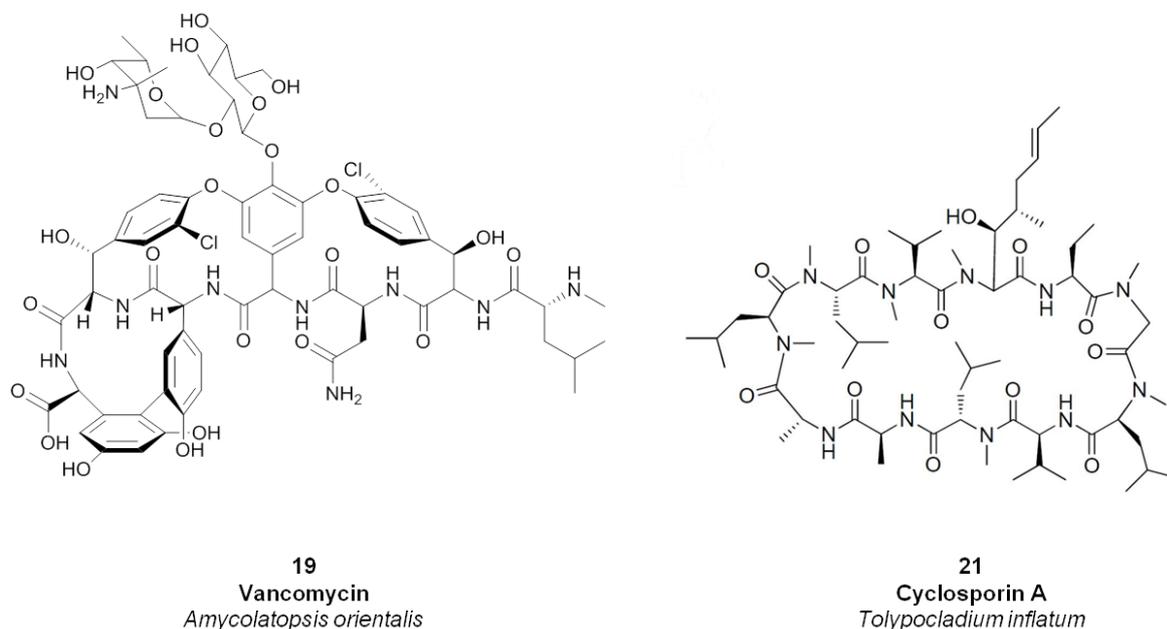


Abbildung 3-13: Pharmakologisch relevante nichtribosomale Peptide (NRP).^{20,57,58}

3.1.2.1 Die Adenylierungsdomänen (A-Domänen)

Wie bereits erwähnt, ist die A-Domäne für die Substratspezifität des Multienzymkomplexes verantwortlich. Dabei aktiviert sie die einzubauende Aminosäure in einer Adenylierungsreaktion durch Bildung eines Aminoacyl-Intermediates aus Aminosäure und Adenosyltriphosphat (ATP). In einem zweiten Schritt wird die so aktivierte Aminosäure durch

Abspaltung von Adenosinmonophosphat (AMP) auf die terminale Sulfhydrylgruppe der PCP-Domäne übertragen (Abbildung 3-12). Erste Kristallstrukturanalysen einer A-Domäne (aus der Gramicidinsynthese GrsA) zeigten, dass die Domäne aus zwei Subdomänen besteht: Einer kleineren ca. 100 Aminosäuren großen C-terminalen Subdomäne und einer aus 400 Aminosäuren bestehenden N-terminalen Subdomäne.^{59,60} Trotz ihrer geringen Sequenzhomologie weisen einige Adenylierungsenzyme eine hohe strukturelle Verwandtschaft zueinander auf. Diese Enzyme werden zu einer Enzymsuperfamilie, der sogenannten ANL-Superfamilie, zusammengefasst. Der Name ANL leitet sich von den drei wichtigsten Subfamilien ab: Acyl-CoA-Synthasen, NRPS Adenylierungsdomänen und Luciferasen.⁶⁰ Ihnen gemein ist der Aufbau aus einer kleinen und einer großen Subdomäne, die durch eine hochkonservierte Region miteinander verbunden sind.^{52,59} In dieser hochkonservierten Region liegt auch das aktive Zentrum der Adenylierungsenzyme.⁵⁹ Bei der Analyse der GrsA A-Domäne konnten in der Bindetasche 10 Aminosäurereste identifiziert werden, die mit der aktivierten Aminosäure interagieren.^{52,59} Bioinformatische Analysen zeigten, dass sich diese Aminosäurereste je nach Substratspezifität der A-Domäne unterscheiden. Diese Tatsache ermöglicht es häufig, allein anhand dieser Aminosäuresequenz auf die Substratspezifität der A-Domäne zu schließen (nichtribosomaler Code).^{50,51} Da sich später zeigte, dass zwei der Positionen bei nahezu allen sequenzierten A-Domänen identisch waren, wurde der Code auf acht Aminosäuren reduziert. Dies erlaubt gleichzeitig die Substratspezifität einer A-Domäne gentechnisch so zu ändern, dass eine andere Aminosäure in das NRP eingebaut wird, was sich allerdings praktisch oft schwierig gestaltet.^{51,61,62} Allerdings ist die Substratspezifität von Aminosäuren kein starres Dogma. So können in einem gewissen Rahmen auch andere Aminosäuren als Substrate erkannt werden. Dies betrifft besonders hydrophobe Aminosäuren, was eine natürliche Strukturdiversität von nichtribosomalen Peptiden zur Folge hat.⁵⁰

Wie dargestellt, ist es zur Vorhersage von NRPS-Produkten erforderlich, die Substratspezifität und die damit verbundene Bausteinauswahl zu kennen. Aufgrund der gewissen Substratflexibilität ist es notwendig, die bioinformatische Substratvorhersage (anhand des nichtribosomalen Codes) experimentell zu überprüfen. Hierfür stehen verschiedene Ansätze zur Verfügung. Ursprünglich wurde ein *in vitro* (d.h. außerhalb des Organismus) durchgeführter ATP/Pyrophosphat-Austauschassay genutzt. Dabei wird radioaktives ³²P-Pyrophosphat im Überschuss zugegeben und dessen Austausch mit dem natürlichem Pyrophosphat gemessen.⁶³ Anhand der Rate dieses Austauschs kann der Grad der Aminosäureaktivierung bestimmt werden. Phelan *et al.* entwickelten einen weiteren Assay, bei dem ATP vierfach durch ein schweres Sauerstoffisotop (¹⁸O) markiert ist.⁶⁴ Dieser Assay wurde auch in dieser Arbeit verwendet und in Kapitel 7.5.8 detailliert beschrieben. In einem dritten, von Dorrestein *et al.* entwickelten Assay wird der A-Domäne eine Vielzahl von

Substraten angeboten. Bei der anschließenden Detektion kann dann die tatsächlich gebundene Aminosäure identifiziert werden.⁶⁵

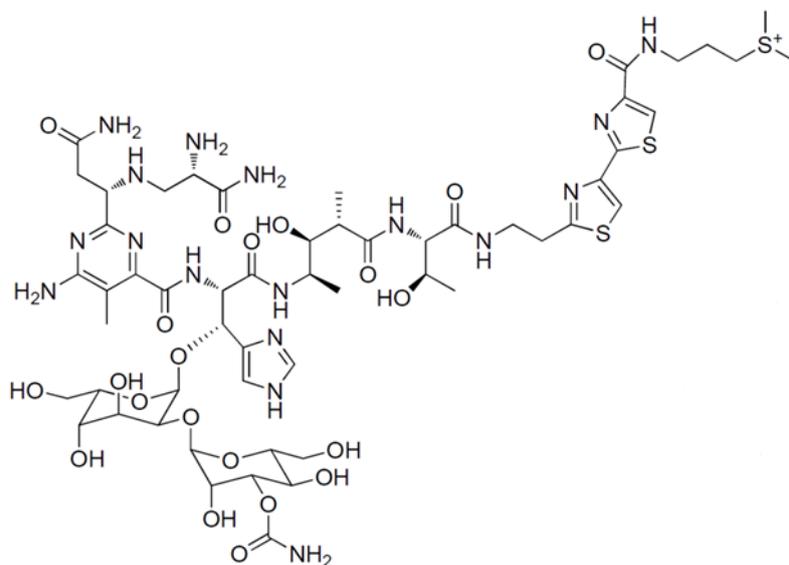
3.1.2.2 *MbtH-artige Proteine*

Weitere in vielen NRPS-Clustern codierte Proteine sind die sogenannten MbtH-artigen Proteine (engl.: *MbtH-like proteins*, MLP). Ihr Name geht auf ihre Erstentdeckung im Biosynthesecluster von Mycobactin des *Mycobacterium tuberculosis* zurück.⁶⁶ Durch Analyse aller bis dato sequenzierten MbtH-Homologen stellte sich heraus, dass diese besonders häufig in Actinobakterien (hier vor allem in Streptomyzeten und Mycobakterien) auftreten, sie aber auch in Bacillales und Proteobakterien zu finden sind.⁶⁷ In Actinomyceten werden sie einzeln in NRPS-Clustern codiert, eine bekannte Ausnahme bildet das Teicoplanin-Gencluster. Hier sind zwei MbtH-artige Proteine codiert.⁶⁸ Die ca. 70-80 AS großen MbtH-artigen Proteine⁶⁹ sind in ihrer Gensequenz z. T. hoch konserviert.⁶⁷ Ein Sequenzvergleich zeigte u. a. drei in allen Proteinen konservierte Tryptophane.⁶⁹ Mehrere *in vivo* Inaktivierungs- und Komplementierungsstudien haben gezeigt, dass MbtH-artige Proteine in vielen NRP-Biosynthesen wichtig, in einigen sogar essentiell sind.^{70,71} Sie zeigten aber auch, dass MbtH-artige Proteine eher unspezifisch wirken, da der Effekt einer Inaktivierung durch heterologe Expression nicht-nativer MbtH-artiger Proteine aufgehoben werden konnte.⁷¹ Die genaue Funktion von MbtH-artigen-Proteinen ist bisher unbekannt, allerdings geben Versuche von Felngale *et al.* wichtige Hinweise hierzu.⁷² Sie konnten zeigen, dass heterolog exprimierte A-PCP-Domänen *in vitro* nur dann aktiv sind, wenn das entsprechende MbtH-Protein coexprimiert wurde. Außerdem konnte beobachtet werden, dass bei der Proteinaufreinigung über eine Nickel-NTA Säule das MbtH-Protein und die A-PCP-Domäne im Verhältnis 1:1 gemeinsam eluiert werden. Auch separat exprimiertes und im Assay zugegebenes MbtH-Protein kann die A-Domäne aktivieren, allerdings nicht in dem Maße wie das coexprimierte. Der Austausch eines der drei hoch konservierten Tryptophane am C-Terminus (gegen Alanin) verhinderte eine gemeinsame Elution. Dies lässt den Schluss zu, dass dieser Aminosäurerest essentiell für die Interaktion ist. Des Weiteren zeigte die Studie aber auch, dass MbtH-artige Proteine für andere A-Domänen hingegen nicht essentiell sind. Diese waren auch ohne die Zugabe von MbtH-artigen Proteinen aktiv. Andere Studien bestätigten diese Beobachtungen.^{73,74} Außerdem konnten sie belegen, dass den drei konservierten Tryptophanen eine Schlüsselrolle in der Interaktion mit anderen Enzymdomänen zukommt und MbtH-artige Proteine eine stabilisierende (Chaperon-artige) Funktion im NRPS Cluster übernehmen. Neuere Arbeiten konnten weiter Hinweise zur Funktionsweise von MbtH-artigen Proteinen geben: Ein Sequenzvergleich der aktiven Zentren einer MbtH-abhängigen A-Domäne und einer MbtH-unabhängigen A-Domäne zeigte, dass sich diese lediglich in einer Aminosäure unterschieden.⁷⁵ Tauschte man diese

Aminosäure in der MbtH-abhängigen A-Domäne aus, kam es tatsächlich zu einer MbtH-Unabhängigkeit. In einer weiterführenden Kristallstrukturanalyse der Interaktion einer A-Domäne mit einem MbtH-artigen Protein konnte diese Bindung näher charakterisiert werden.⁷⁶ Dazu wurde das A/MbtH-Fusionsprotein des Streptolydigin-Biosyntheseclusters (PKS/NRPS Hybrid) exprimiert und durch Kristallisationsstrukturanalyse charakterisiert. Diese Analyse konnte die wichtige Rolle der konservierten Tryptophanreste besonders von Trp-25 und Try-35 in der Interaktion mit der A-Domäne bestätigen. Zu einer direkten Interaktion zwischen der MbtH-Domäne und dem Substrat kommt es nicht. Die genaue Rolle von MbtH-artigen Proteinen ist aber weiterhin unklar. Eine Hypothese ist, dass es durch die Bindung der MbtH-artigen Proteine zu einer Konformationsänderung der A-Domäne kommt, was deren Aktivität erhöht (allosterischer Regulator).^{75,76}

3.1.3 NRPS/PKS Hybride

Häufig treten Multienzymkomplexe auf, die sowohl Module der PKS als auch NRPS-Domänen aufweisen. Diese werden als NRPS/PKS Hybride bezeichnet. Beispiele von Naturstoffen, die von solchen Komplexen synthetisiert werden, sind die antitumoral wirkenden Substanzen Bleomycin (**20**)⁷⁷ und Pederin (Abbildung 3-10).³¹



20
Bleomycin
Streptomyces verticillus

Abbildung 3-14: Bleomycin.⁷⁸

Auch in dieser Arbeit wurde der Gencluster eines solchen NRPS/PKS-Hybriden untersucht (Kapitel 3.3.2, Abbildung 3-21).

3.1.4 Genome mining

PKS, NRPS und deren Hybride synthetisieren eine Vielzahl bioaktiver Substanzen, denen aufgrund ihrer pharmakologischen Eigenschaften wichtige Bedeutung für die Entwicklung neuer Wirkstoffe zukommt. Wie bereits zu Beginn der Einleitung angeklungen ist, bemühte sich auch die Pharmaindustrie seit Mitte des 20. Jahrhunderts um die Erschließung dieser Substanzen. Meist beschränkte man sich auf die Untersuchung von Extrakten (aus Pflanzen, Pilzen und Mikroorganismen) nach Bioaktivität. Diese aufwändige und kostenintensive Herangehensweise führte jedoch immer häufiger zu Wiederentdeckungen bereits bekannter Naturstoffe statt der Isolierung neuer Substanzen.⁷⁹

Die Zunahme von Sequenzierungen ganzer Genome und Metagenome, besonders im Bereich der Mikrobiologie, eröffnet jedoch ganz neue Möglichkeiten zur Detektion von Naturstoffgenclustern. So können Genom- und Metagenomdaten gezielt anhand konservierter Sequenzmuster nach PKS und NRPS-Genclustern durchsucht werden. Dieses sogenannte *Genome mining* ermöglicht die automatisierte Analyse großer Datensätze mithilfe bioinformatischer Programme. Anhand der prognostizierten Domänenstruktur kann dann oft, wie in Kapitel 3.1.1 und 3.1.2 beschrieben, der resultierende Naturstoff vorhergesagt werden.⁸⁰ So können sowohl prognostizierte Metabolite einfacher isoliert als auch Gencluster bereits bekannter Naturstoffe identifiziert werden.^{54,81,82} Ein Beispiel für ein solches bioinformatischen Programm ist das auch in dieser Arbeit verwendete antiSMASH (antibiotics & Secondary Metabolite Analysis SHell).^{83–85}

3.2 Mikrobielle Symbiosen

Viele Bakterien leben dauerhaft in enger Assoziation (Symbiose) mit anderen Bakterien oder anderen Organismen.⁸⁶ Symbiosen werden anhand unterschiedlicher Kriterien eingeteilt. Ist das Zusammenleben für das Bakterium lebensnotwendig, spricht man von einer obligaten Symbiose im Gegensatz zur fakultativen, bei der ein Zusammenleben nicht lebensnotwendig ist. Aufgrund der Auswirkung der Symbionten auf ihren Wirt werden Symbiosen in *mutualistisch* (Eine Beziehung zwischen zwei Arten, aus denen beide Partner einen Benefit ziehen, „+/+ Interaktion“⁸⁷) und *parasitär* (der Mikroorganismus zieht alleinigen Nutzen aus dem Zusammenleben) eingeteilt. In manchen Symbiosen hat der Symbiont keinen erkennbaren Effekt auf seinen Wirt. Man spricht von *Kommensalismus*. Des Weiteren wird zwischen Ekto- und Endosymbiosen unterschieden, je nachdem, ob die Symbionten außer- oder innerhalb der Wirtszellen vorkommen.^{86–88} Wie genau es zur Aufnahme solcher Bakterien und der Bildung einer Symbiose kommt, ist häufig ungeklärt. Neben der Weitergabe der Symbionten von einem Vorfahr an die nachfolgende Generation (vertikale Transmission) können die Symbionten auch über horizontale Transmission wiederholt aus der Umgebung aufgenommen werden.⁸⁶

Eine neue Studie zum Stammbaum des Lebens zeigt die Bedeutung bakterieller Diversität. Zusätzlich zu Sequenzen bereits bekannter Bakterienspezies wurden zur Berechnung dieses Stammbaums auch Sequenzen bisher nicht kultivierter Bakterien herangezogen. Hierbei handelt es sich häufig um bakterielle Symbionten (Abbildung 3-15).⁸⁹

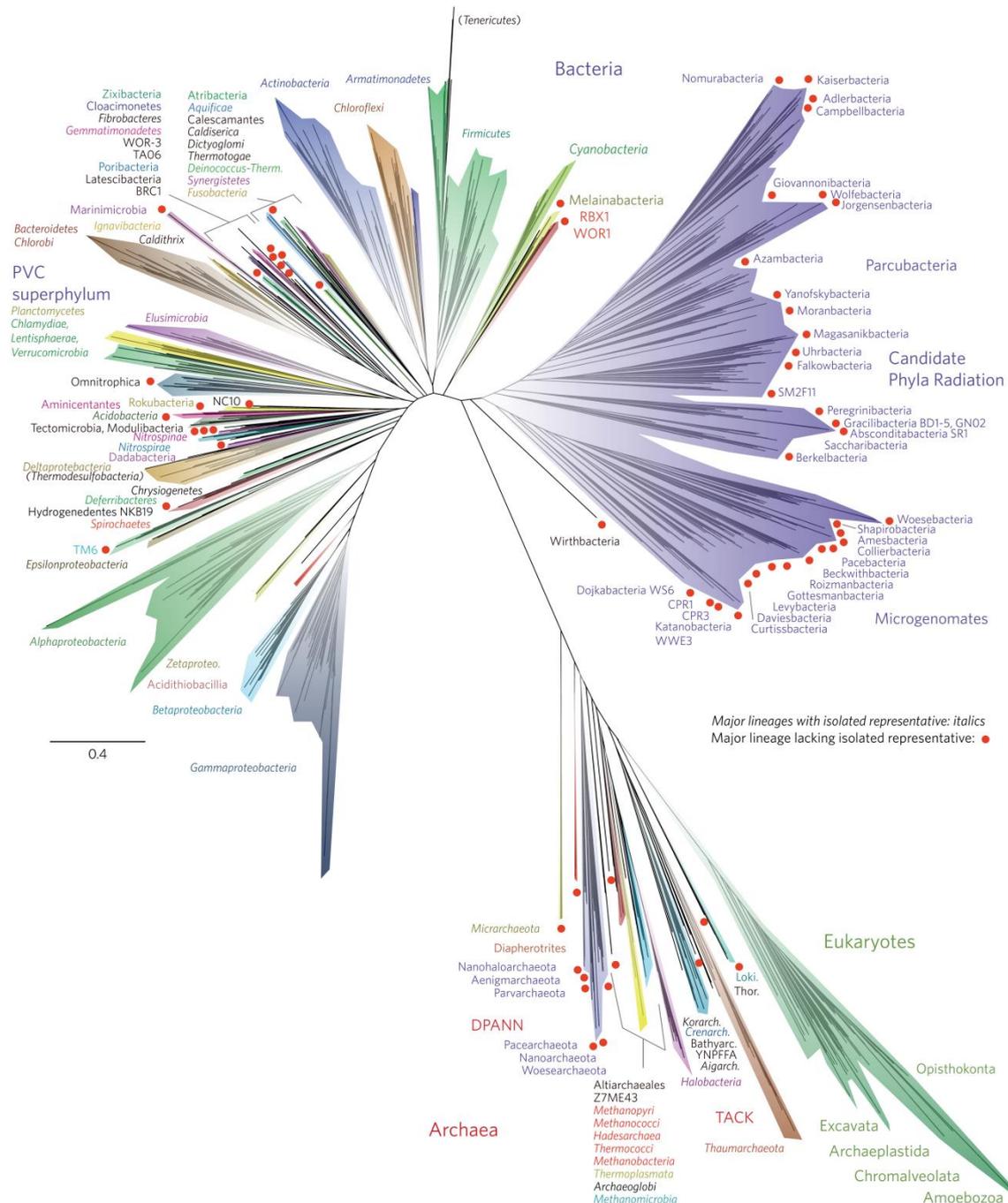


Abbildung 3-15: Der Stammbaum des Lebens nach Hug, L. A. et al.⁸⁹

Mikrobielle Symbiosen sind in nahezu allen Arten von Lebewesen zu finden: Ein Beispiel einer Symbiose zwischen Mikroorganismen sind Flechten. Diese mutualistischen Lebensgemeinschaften bestehen aus einem Pilz (Mycobiont), der mit einer Photosynthese

betreibenden Grünalge oder einem Cyanobakterium (Photobiont) assoziiert ist.⁹⁰ Auch im Pflanzenreich sind sehr häufig Assoziationen mit Bakterien zu finden. Diese ebenfalls oft mutualistischen Symbiosen verschaffen der Pflanze besseren Zugang zu Nährstoffen oder steigern ihre Abwehr gegenüber Pathogenen.⁸⁶ Aber auch der Mensch selbst beherbergt (wie andere Säugetieren auch) ein großes Mikrobiom.⁸⁶ Der wohl wichtigste und gleichzeitig diverseste Teil unseres Mikrobioms ist im Intestinum zu finden. Neben der Verwertung von Nahrungsbestandteilen^{91,92} hat das Mikrobiom des Intestinaltrakts jedoch noch viele weitere Auswirkungen auf den menschlichen Organismus, die z. T. bis heute noch nicht vollständig aufgeklärt werden konnten. Dazu gehört die Abwehr enteropathogener Bakterien und die Modulation des Immunsystems⁹³ um hier nur einige zu nennen. Neuere Studien an Mäusen zeigen, dass die Gabe von Antibiotika mit einer Verringerung der hippocampalen Neurogenese korreliert. Neben einem direkten negativen Effekt durch die Antibiotikagabe zeigt die Studie, dass das zerstörte Mikrobiom ebenfalls einen Einfluss auf die Neurogenese haben könnte: Die Gabe ausgewählter Bakterienstämmen konnte den negativen Effekt der Antibiotikagabe z.T. aufheben.⁹⁴ Auch mikrobielle Symbiosen mit Insekten oder im marinen Habitat in Assoziation mit Wirbellosen sind möglich. Drei Beispiele hierfür stehen im Fokus dieser Arbeit und werden in Kapitel 3.3 näher beschrieben.

3.3 Bakterielle Symbionten - potente „Naturstofffabriken“

Viele Naturstoffe, wie die in Kapitel 3.1 beschriebenen Polyketide und nichtribosomalen Peptide, werden häufig aus Umweltproben und höheren Organismen wie Pilzen und Invertebraten isoliert. Immer häufiger stellt sich heraus, dass es sich bei den eigentlichen Produzenten dieser oft pharmakologisch wirksamen Naturstoffe um bakterielle Symbionten handelt.⁹⁵ Ein Beispiel hierfür sind die antitumoral wirkenden Substanzen der Pederin-Familie. So wird z. B. das bereits erwähnte, für diese Substanzfamilie namensgebende Pederin (**17**) von dem Bakterium *Pseudomonas paederi* synthetisiert, das in Symbiose mit Käfern der Gattung *Paederus* und *Paederidus* lebt. Hierauf wird in Kapitel 3.3.2 genauer eingegangen.¹⁷ Ein weiterer Vertreter dieser Familie ist das Nosperin. Dieses wird von einem *Nostoc*-Photobionten produziert, der in enger Symbiose mit der Flechte *Peltigera membranacea* assoziiert ist.⁹⁶ Andere Vertreter dieser Substanzfamilie wie z. B. Psymberin und Onnamid A konnten aus Schwämmen isoliert werden.³⁰ Auf den Produzenten von Onnamiden wird in Kapitel 3.3.1 näher eingegangen. Auch viele andere pharmakologisch aktive Substanzen werden von bakteriellen Symbionten synthetisiert. so z. B. Mutanobactin, das von dem in der menschlichen Mundhöhle beheimateten *Streptococcus mutans* synthetisiert wird. Dieses, von einer Hybrid-PKS-NRPS produziertes Molekül inhibiert die Biofilmbildung durch *Candida albicans*.^{97,98} Auch viele der im Boden zu findenden Bakterien leben in symbiotischer Beziehung zueinander. So vermutlich auch das Bakterium *Eleftheria*

terrae, das nur mithilfe des sogenannten iChips in einem speziellen Kultivierungsverfahrens im ursprünglichen Habitat kultiviert werden konnte.^{99,100} Das von ihm produzierte antibiotisch wirksame Teixobactin sorgte für große Beachtung, da bisher keine Resistenzen gegenüber dieser Substanz beobachtet werden konnten.¹⁰⁰ Bei den meisten dieser Naturstoffe beantwortet sich die Frage, warum diese so häufig von diesen Symbionten produziert werden, anhand ihrer Bioaktivität. Häufig dienen diese Polyketide und nichtribosomalen Peptide der Kommunikation der Bakterien oder zu deren Verteidigung (Kapitel 3.1).^{16,17} So scheint es, zumindest bei Letzterem offensichtlich, dass sie so auch für die Symbiosepartner von Vorteil sein könnten. Aufgrund des pharmakologischen Interesses an diesen Wirkstoffen rücken diese Symbionten immer mehr in den Fokus der Forschung. Wie bereits beschrieben, bieten diese Bakterien eine große z. T. noch unentdeckte Quelle für neue Leitstrukturen in der Arzneimittelforschung. Die aus der symbiontischen Lebensweise resultierenden komplexen Wachstumsbedingungen machen es allerdings sehr schwierig diese Mikroorganismen unter Laborbedingungen zu kultivieren. Dies erschwert die Arbeit an und mit diesen Naturstoffproduzenten. Auch diese Arbeit beschäftigt sich mit diesen Polyketide und nichtribosomalen Peptide produzierenden Symbionten. Die folgenden Abschnitte sollen einen kurzen Überblick über die in dieser Arbeit untersuchten bakteriellen Symbionten geben.

3.3.1 Mikrobielle Schwammsymbionten des Schwammes *Theonella swinhoei*

Viele marine Invertebraten wie z. B. Schwämme gelten als wichtige Quelle für neue bioaktive Substanzen.¹⁰¹ Diese Substanzen spielen eine große Rolle als chemisches Verteidigungsmittel dieser immobilen marinen Lebewesen. Einige dieser Naturstoffe sind biomedizinisch relevant bzw. haben hohes pharmakologisches Potential, was sie für die Forschung interessant macht.¹⁰² Bereits Anfang der 1950er Jahre konnten aus dem Schwamm *Cryptotethia crypta* zwei zytotoxisch aktive Nukleoside (Spongothymidin und Spongouridin) isoliert werden, aus denen später das Virostatikum Vidarabin und Zidovudin sowie das Zytostatikum Cytarabin entwickelt wurde.^{103–106} Ein erst vor wenigen Jahren entwickeltes Medikament marinen Ursprungs ist Eribulin (strukturell vereinfachtes synthetisches Analogon des in Meeresschwämmen vorkommenden Halichondrin B), das als Chemotherapeutikum bei der Behandlung von Brustkrebs angewendet wird. Diese Substanz wurde in den 1980er Jahren aus dem japanischen Schwamm *Halicondria okadai* isoliert. Bis März 2018 wurden sieben Substanzen marinen Ursprungs, drei davon aus Schwämmen (die bereits erwähnten Eribulin, Vidurabin und Cytarabin), von der FDA (Food and Drug Administration) zugelassen. Insgesamt befinden sich 23 marine Naturstoffe in den klinischen Phasen. Eine Polyketid aus dem Schwamm *Lithoplocamia lithistoide* ist in der klinischen

Phase II.^{107–109} Diese Zahlen machen das pharmakologische Potential dieser Naturstoffe deutlich.

Häufig zeigen einige Schwamm-Naturstoffe auffallende strukturelle Ähnlichkeit zu Substanzen bakterieller Herkunft.^{30,110–112} So auch das bereits erwähnte *Halichondrin* und das ebenfalls antitumoral wirkende Discodermolid aus dem Schwamm *Discodermia dissoluta*.^{112,113} Es ist also naheliegend, dass auch hier, wie bereits beschrieben, nicht die Schwämme selbst, sondern symbiontische Bakterien oder andere Mikroorganismen die eigentlichen Produzenten dieser interessanten Naturstoffe sind.

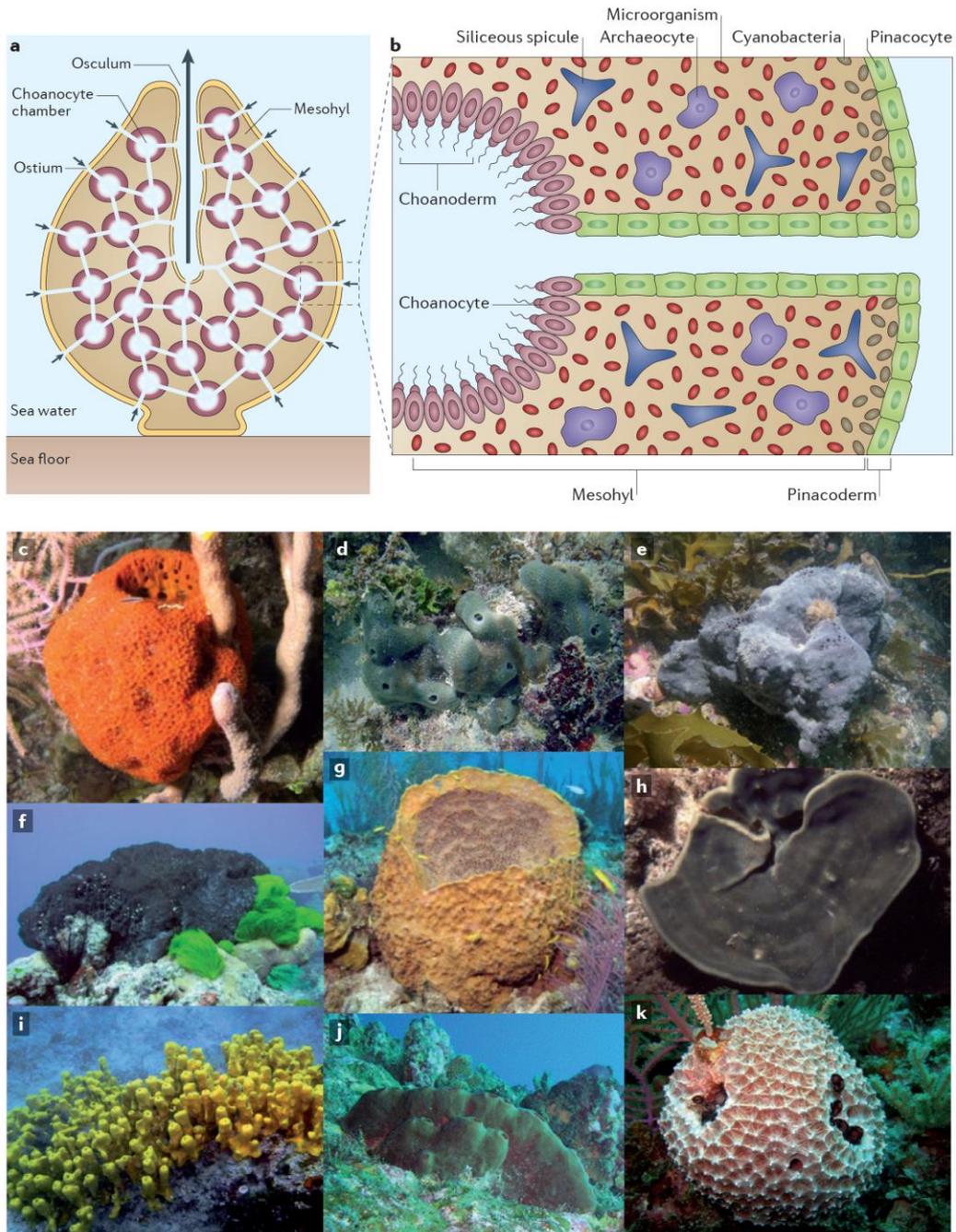


Abbildung 3-16: Schematischer Aufbau und Übersicht verschiedener Vertreter der Klasse

Demospongiae.¹¹⁴ a: Schematische Übersicht über den Aufbau eines Hornkieselschwammes (*Demospongiae*). b: Vergrößerte Ansicht der inneren Struktur eines Demospongiae-Gewebes. Das Seewasser wird durch kleine Poren (*Ostia*) im Pinacoderm aufgenommen. Nach dem Herausfiltrieren der Mikroorganismen wird das nahezu sterile Wasser durch das *Osculum* wieder herausgeführt. Chanozyten dienen der Aufnahme von Nahrungspartikeln wie Bakterien, Algen und auch Viren. Diese werden dann in die innere Mesohylschicht transportiert und dort von Archaeocyten aufgenommen. In vielen Demospongiae ist die Mesohylschicht von vielen symbiotischen Mikroorganismen besiedelt, wobei Photosynthese-betreibende Cyanobakterien aufgrund der Lichtverhältnisse meist nur in der äußeren Region zu finden sind. Kieselnadeln (lat.: *Siliceous spicules*) bilden das sogenannte *Skerlit* und bestimmen somit die Form des Schwammes. c-k: Unterwasseraufnahmen einiger Vertreter der Klasse Demospongiae. *Mycale laxissima* (c), *Amphimedon queenslandica* (d), *Ancorina alata* (e), *Rhopaloeides odorabile* (f), *Xestospongia muta* (g), *Cymbastela concentrica* (h), *Aplysina aerophoba* (i), *Theonella swinhoei* (j), *Ircinia felix* (k).

Schwämme besitzen einen einfachen, multizellulären Aufbau, der perfekt an das Leben eines sogenannten Filtrierers angepasst ist: Ihre Oberfläche besteht aus vielen kleinen

Poren, die das Eindringen von Wasser erlauben (*Ostia*). Dieses zirkuliert dann durch ein Kanalsystem, in dem Mikroorganismen und kleine organische Partikel zur Nahrung herausgefiltert werden. Mikroorganismen, die sowohl diesem Verdauungsprozess als auch dem Immunsystem des Schwammes entgehen, können diese Schwämme besiedeln.¹¹⁵ 16S rRNA-Gen-Analysen konnten bisher 14 verschiedene Bakterienphyla in Schwämmen identifizieren. Die am häufigsten vorkommenden Phyla sind Acidobacteria, Actinobacteria und Chloroflexi.¹⁰² Schwammsymbionten kommen sowohl intra- als auch extrazellulär vor: Extrazelluläre Symbionten, die weitaus größere Gruppe, sind in der Außenschicht, als sogenannte Exosymbionten oder in der extrazellulären Matrix (Mesohyl) als sogenannte Endosymbionten zu finden. Intrazelluläre bzw. intranukleäre Symbionten hingegen kommen, wie der Name schon erahnen lässt, ausschließlich innerhalb der Wirtszellen bzw. dessen Kern vor.¹⁰² Wie genau es zu einer Symbiose zwischen Mikroorganismen und Schwämmen kommt, konnte bisher nicht endgültig geklärt werden. Verschiedene Szenarien sind hier vorstellbar: 1. Selektive Aufnahme eines bestimmten Bakteriums; 2. der Symbiont vermehrt sich schneller als andere Mikroorganismen; 3. der Symbiont wird durch vertikalen Transfer von den adulten Schwämmen an die Larven weiter gegeben.^{102,116} Obwohl Schwämme diese symbiontischen Bakterien über denselben Mechanismus aufnehmen wie die Mikroorganismen, die ihnen als Nahrung dienen, werden sie kaum verdaut.^{102,117,118} Auch aus Schwämmen der Ordnung *Lithistida* (Klasse: *Demospongiae*) wurden bereits über 300 Naturprodukte beschrieben, darunter Sterole, Makrolide, Polyketide, Alkaloide und Peptide.^{119–121} Ein wichtiger und gut untersuchter Vertreter dieser Ordnung ist der Steinschwamm *Theonella swinhoei*.



Photo by J. Tanaka

***Theonella swinhoei* WA**



Photo by Y. Nakao

***Theonella swinhoei* Y**

Abbildung 3-17: Chemotypen des Schwammes *Theonella swinhoei*. In der Abbildung links abgebildet ist der weiße Chemotyp (W, engl.: *white*), rechts der gelbe Chemotyp (Y, engl.: *yellow*) des Schwammes *Theonella swinhoei*.

Wie viele andere Schwämme beheimatet auch er eine Vielzahl bislang unkultivierbarer Bakterien verschiedenster Phylotypen.^{102,114,122} Außerdem ist *Theonella swinhoei* als umfangreiche Quelle diverser Naturstoffe bekannt. Schwämme dieser Art, aber

unterschiedlicher Herkunft zeigen verschiedene Stoffwechselprofile und bilden somit unterschiedliche Chemotypen. Ein besonders interessanter Vertreter ist eine Variante, die vor der japanische Insel Hachijo-jima (etwa 300 km südlich von Tokio) beheimatet ist. Aufgrund der gelben Färbung seines Inneren wird dieser Chemotyp auch als *Theonella swinhoei* Y (engl.: *yellow*) bezeichnet. Aus diesem Chemotyp konnten insgesamt über 40 bioaktive Polyketide und Peptide isoliert werden, darunter die Serin Protease-Hemmer Cyclotheonamide (**22**) und Nazumamid A (**23**) sowie die zytotoxischen Moleküle Onnamid (**24**) und das antifungale Aurantosid (**25**)^{123–127}, außerdem das Peptid Konbamid (**26**), ein Calmodulin-Antagonist¹²⁸ und die außergewöhnlichen Peptide Polytheonamid A und B (**27**), die, trotz ihres ribosomalen Ursprungs D-Aminosäuren enthalten.^{128–130}

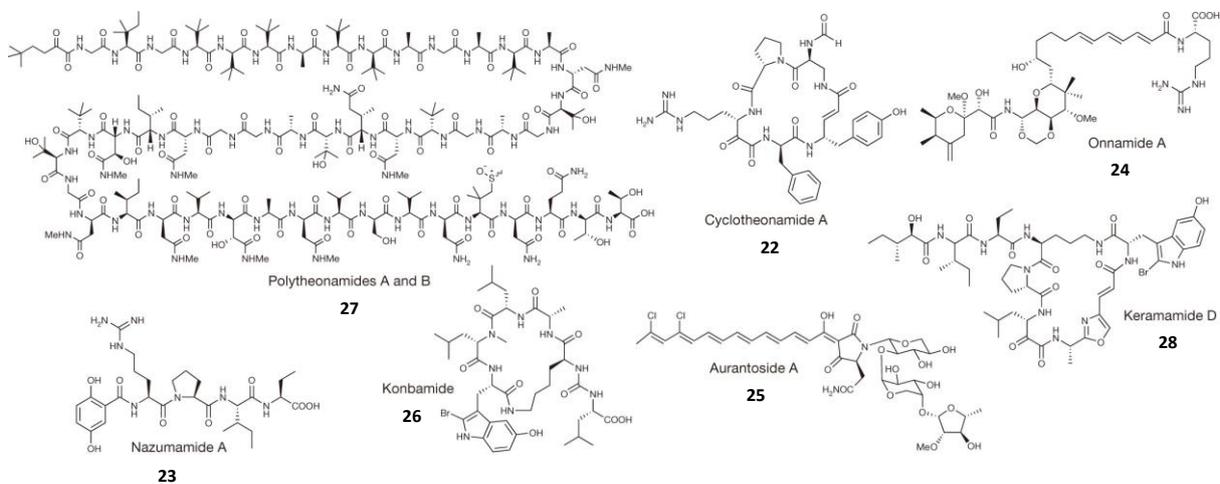


Abbildung 3-18: Bioaktive Naturstoffe aus dem Schwamm *Theonella swinhoei*.¹³¹

Bei einigen dieser Naturstoffe war schon zu Beginn dieser Arbeit klar, dass diese nicht vom Schwamm selbst produziert werden, sondern bakteriellen Ursprungs sind. In vorangegangenen Arbeiten konnten bereits anhand metagenomischer Sequenzanalysen die Gencluster von Onnamid und Polytheonamid identifiziert und analysiert werden, eine Identifizierung und taxonomische Einordnung des bzw. der potentiellen Produzenten waren allerdings nicht möglich.^{129,132} Um den oder die Produzenten bestimmen zu können, sollte zunächst eine sogenannte Einzelzellanalyse durchgeführt werden.¹³¹ Diese Strategie ermöglicht es, Mikroorganismen anhand funktionaler Gene (in diesem Fall anhand der Naturstoffcluster) phylogenetischen einzuordnen.^{133–135} Dazu wurden alle Bakterien aus dem Mesohyl von *Theonella swinhoei* Y durch differentielle Zentrifugation anhand ihrer spezifischen Zelldichte in verschiedenen Fraktionen angereichert.^{131,136} Die anschließende mikroskopische Untersuchung der Fraktionen zeigte eine Anreicherung von gram-negativen filamentösen Bakterien (2-3 µm große miteinander verbundene Zellen) in der Fraktion hoher Dichte. Diese Filamente wiesen unter UV-Licht eine Eigenfluoreszenz auf. Morphologisch

ähnelten sie stark einem Symbionten aus einem *Theonella swinhoei* Chemotyp vor dem Inselstaat Palau. Dieser Symbiont, '*Candidatus Entotheonella palauensis*', ist vermutlich Produzent einiger antifungaler Peptide ist.^{136,137} Durch eine „Multiple Displacement Amplification“ (MDA)-PCR konnte die Einzelgenom-DNA der filamentösen Bakterien für anschließende PCR-Ansätze amplifiziert werden.¹³¹ Um schließlich die Onnamid- und Polytheonamid- produzierenden Bakterien identifizieren zu können, wurde eine PCR mit spezifischen Primern der für die Synthese dieser Naturstoffe verantwortlichen Synthasen (Kapitel 3.1) durchgeführt. Dabei konnten einige der filamentösen Bakterien als Produzenten identifiziert werden. Phylogenetische Untersuchungen mithilfe spezifischer '*Candidatus Entotheonella*' -16S ribosomalen RNA (rRNA) Primern bestätigten, dass es sich dabei um Bakterien der Gattung '*Candidatus Entotheonella*' handelt. Neben diesen für alle drei Primer (Onnamid-, Polytheonamid- und '*Candidatus Entotheonella*' -Primer) positiven Kandidaten konnten allerdings auch Bakterien beobachtet werden, die nur für einzelne Cluster positiv waren.¹³¹ Um dieses Phänomen zu klären und auszuschließen, dass einzelne Bakterien aufgrund von unterrepräsentierten Genomabschnitten (infolge der MDA-Methodik) falsch negativ waren, sollten die Ergebnisse durch Metagenom-Sequenzierung weiter verifiziert werden. Hierzu wurde die Fraktion der filamentösen Bakterien in mehreren Illumina, 454, PacBio und Sanger-Sequenzierungen analysiert. 78,3% der sequenzierten Fragmente (engl.: *reads*) konnten zu 18098 Sätzen überlappender Fragmente (engl.: *contigs*) von min. 500 bp Länge assembliert werden. Die übrigen *Reads* konnten nicht assembliert werden, was darauf schließen ließ, dass die Phylotypen dieser Fragmente nur in geringer Konzentration in der Bakterienfraktion vorkamen. Diese Hypothese wurde auch durch die hohe Varianz in der mittleren Abdeckung (engl: *coverage*) belegt. Binning-Analysen der *Contig* und *Scaffold* Sequenzen (aus *Contigs* assemblierte, größere Sequenzstücke) und deren anschließender Einordnung deuteten darauf hin, dass es sich um zwei große Bakterienpopulationen mit einem GC-Gehalt von 55% handelte. Viele der untersuchten Gene wiesen jeweils zwei sehr ähnliche Homologe mit Sequenzübereinstimmungen von 85-91% auf. Zusammengefasst entspricht dieser Sequenzbereich einer Synthenie über ca. 4,5 Mb. In diesem Bereich konnten jeweils exakt zwei Orthologe von 35 Einzelkopie-Genen identifiziert werden, wobei viele dieser Gene als phylogenetische Marker dienen. Diese Ergebnisse deuten darauf hin, dass die sequenzierte Probe die Genome zweier nah verwandter '*Candidatus Entotheonella*'-Spezies enthielt. Die zunächst als TSY1 (später '*Entotheonella factor*') und TSY2 (später '*Entotheonella gemina*') bezeichneten Populationen wiesen eine Sequenzübereinstimmung der 16S rRNA von 97,6% auf. Die Übereinstimmung zu '*Entotheonella palauensis*' lag bei 97%. Die *Coverage* der chromosomalen DNA lies darauf schließen, dass es sich bei TSY1 um die dominierende Spezies handelte. Einige *Contigs* bzw. *Scaffolds* zeigten im Verhältnis zu den restlichen Sequenzen eine stark erhöhte *Coverage* auf, was darauf hin deutet, dass

die beiden Bakterienspezies neben der genomischen DNA noch mindestens ein Plasmid enthalten.¹³¹ Erste Analysen nach der vierten Sequenzierungsrunde ließen bereits das hohe Potential dieser Spezies als Naturstoffproduzent erahnen: 59 der 3700 Contigs (>1 kb) enthielten PKS bzw. NRPS Module (mit ca. 30 KS-, 37 C- und 77 A-Domänen) (Dr. Micheal C. Wilson). Bereits in anderen Untersuchungen von '*Candidatus Entotheonella*'-Spezies enthaltenden Schwämmen konnten interessante Naturstoffe isoliert werden. Beispielsweise das Hybrid Polyketid-Peptid Theopalauamid in der angereicherten '*Entotheonella palauensis*'-Fraktion des *Theonella swinhoei* Palau-Chemotyps und das antitumor Polyketid Discodermolid aus dem ebenfalls '*Candidatus Entotheonella*'-enthaltenden Schwamm *Discodermia dissoluta*.^{136,137} Um herauszufinden ob '*Candidatus Entotheonella*' -Spezies auch in weiteren Schwammspezies zu finden sind und ggf. eine generelle Rolle in der Naturstoffsynthese übernehmen wurden 37 taxonomisch unterschiedliche Schwammspezies aus 20 unterschiedlichen Regionen mittels PCR untersucht. 28 der 37 untersuchten Schwämme zeigten Amplikons mit Sequenzübereinstimmungen zu TSY1 ('E. factor') von 95,5-99,9%. Um das Naturstoff-Potential dieser Bakterien besser einschätzen zu können, wurde ein weiterer *Theonella swinhoei* Chemotyp (Typ WA aufgrund der weißen Farbe des Mesoyls) auf die Existenz von PKS-Clustern untersucht. Aus diesem Schwamm konnte bereits in anderen Arbeiten das antitumoral wirkende Makrolid Misakinolid A (ein Dimer) isoliert werden, das im Chemotyp Y nicht zu finden ist.^{138,139} Eine PCR-Analyse der metagenomischen Schwamm-DNA konnte lediglich Fragmente einer in Schwämmen ubiquitär verbreiteten putativen FAS (*sup*-Gene; engl.: *sponge symbiont ubiquitous pks*) detektiert werden.^{131,140} Im Gegensatz dazu konnten aus einer angereicherten '*Candidatus Entotheonella*'-Fraktion dieses Schwammes sechs Genfragmente von PKS amplifiziert werden. In TSY1 und TSY2 konnten keine Homologe dieser PKS gefunden werden. Das lässt auf eine komplexe und diverse Chemie der unterschiedlichen '*Candidatus Entotheonella*'-Phylotypen schließen.¹³¹

Um die verschiedenen '*Candidatus Entotheonella*'-Spezies taxonomisch einordnen zu können, wurde eine phylogenetische Analyse der 16S rRNA-Sequenzen und verschiedener Markergene durchgeführt. Dabei wurden insgesamt 243 Sequenzen sowohl aus den untersuchten Schwämmen dieser Studie als auch aus Datenbank-Analysen miteinander verglichen. Dabei zeigte sich, dass diese Sequenzen eine eigene separate Klade bilden, mit einer Übereinstimmung von nur 82% zu Vertretern bekannter Phyla. Das lässt darauf schließen, dass diese Spezies ein neues, eigenes Kandidatenphylum bilden. Aufgrund der bisherigen unkultivierbarkeit seiner Mitglieder und deren Fähigkeit, Abwehrstoffe zu produzieren, wurde dieses Kandidatenphylum 'Tectomicrobia' (lat.: *tegere*, sich verbergen, etwas schützen) getauft.¹³¹ Dieses Kandidatenphylum spaltet sich in drei Kladen auf: Die größte Klade beinhaltet '*Candidatus Entotheonella*'-Sequenzen hauptsächlich aus

Schwämmen und Seewasser, die zweite beinhaltet verwandte Sequenzen aus Schwämmen, die nicht der Gattung '*Candidatus Entotheonella*' angehören und die dritte besteht aus Sequenzen aus Bodenproben. Im Stammbaum des Lebens nimmt dieses Phylum die Schwesterposition zu einem ebenfalls neuen Phylum '*Nitrospirae*' ein.⁸⁹

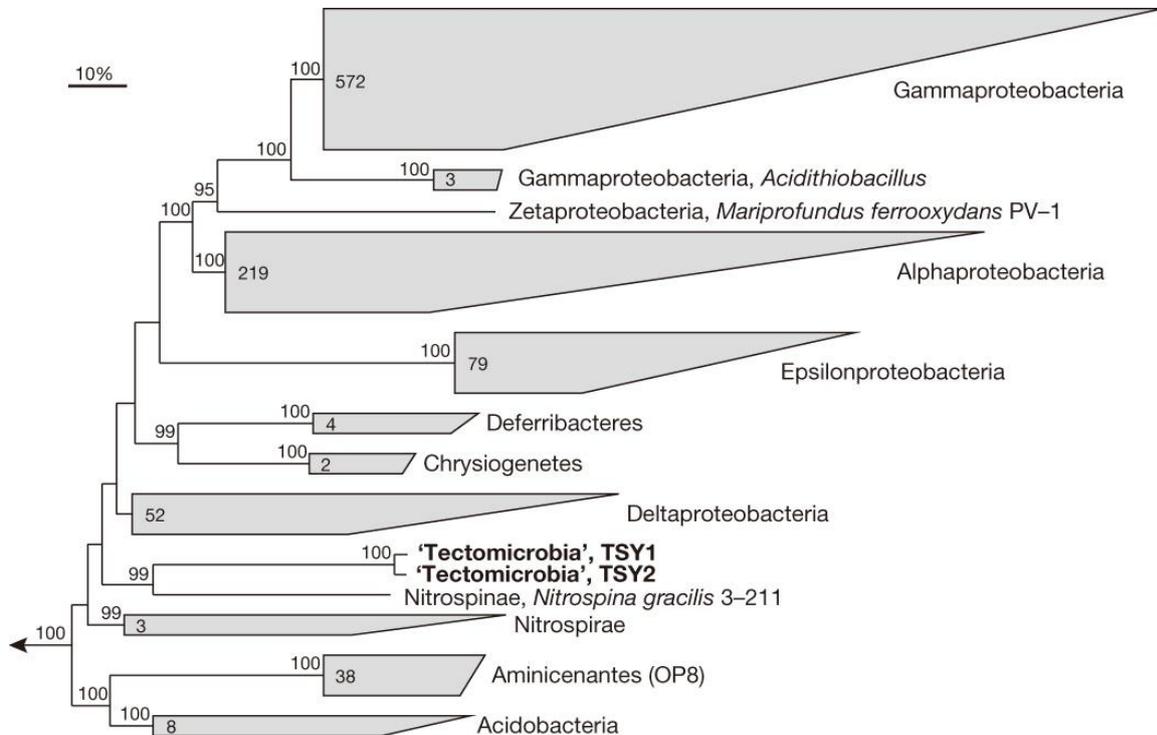


Abbildung 3-19: Phylogenetische Einordnung von '*Candidatus Entotheonella factor*' anhand verschiedenerer Markergene.¹³¹

Neben diesem '*E. factor*' konnten Dr. Tetsushi Mori und Prof. Dr. Haruko Takeyama einen weiteren filamentösen Symbionten im Schwamm *Theonella swinhoei* über Raman-Mikroskopie identifizieren (im Folgenden *Symbiont 2* genannt). Die Raman-Mikroskopie kombiniert die optische Mikroskopie mit Raman-Spektroskopie. Sie dient u. a. der nicht invasiven Analyse von Bakterien. Dabei werden durch Laserlicht Moleküle in charakteristische Schwingungen versetzt, was zu einer Strahlung im roten Wellenbereich führt.¹⁴¹ Die Raman-Spektren können Auskunft über die chemische Zusammensetzung, Bindungssituation, Symmetrien, Strukturen und physikalische Parameter wie beispielsweise die Bindungslänge geben.¹⁴²

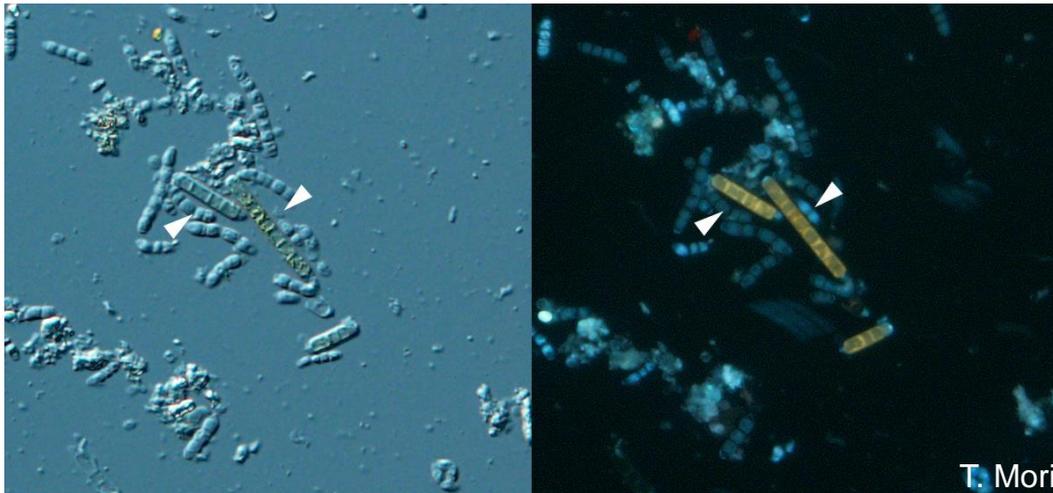


Abbildung 3-20: Mikroskopische Aufnahme filamentöser Symbionten des Schwammes *Theonella swinhoei*. Mit Pfeilen markiert ein weiteres filamentöses Bakterium. Bei den nicht markierten Filamenten handelt es sich um 'E. factor' (Aufnahme: Prof. Dr. Tetsushi Mori, Waseda University, Tokyo).

Im Unterschied zu 'E. factor' zeigen diese Filamente im Lichtmikroskop eine gelbliche Färbung (Abbildung 3-20, links). In der Raman-Mikroskopie zeigen sie das typische Spektrum von Aurantosiden, die auch die gelbe Färbung verursachen. Einzelfilament-Sequenzierungen und erste phylogenetische Einschätzungen legen nahe, dass es sich bei diesen filamentösen Bakterien wahrscheinlich um ein neues Kandidatengenus in der Nachbarschaft zum Genus *Caldilinea* handelt. Dieses ist phylogenetisch zur Familie *Caldilineaceae*, Ordnung *Caldilineales*, Klasse *Caldilineae* und Phylum *Chloroflexi* einzuordnen.^{143,144} Das große Phylum *Chloroflexi* besteht aus sechs Klassen: *Chloroflexi*, *Anaerolineae*, *Caldilineae*, *Ktedonobacteria*, *Dehalococcoidia* und *Thermomicrobia*. Eine typische Eigenschaft der Bakterien der Klassen *Chloroflexi*, *Anaerolineae*, *Caldilineae*, und *Ktedonobacteria* ist die filamentöse Morphologie. Phototroph sind nur die Mitglieder der Klasse *Chloroflexi*.¹⁴⁵ In kultivierungsunabhängigen Untersuchungen zur mikrobiellen Diversität konnten Bakterien dieses Phylums in nahezu allen natürlichen Habitaten identifiziert werden.¹⁴⁶

3.3.2 Mikrobielle Symbionten der *Paederus/Paederidus*-Käfer

Die Diversität von Insekten-assoziierten Bakterien ist enorm. 10% aller 800.000 Insektenpezies beherbergen Symbionten, die sich vermutlich auch evolutionär zusammen mit ihrem Wirt weiterentwickeln. Dies kann wiederum neue individuelle Spezies hervorbringen.^{88,147} So gehört auch '*Candidatus Pseudomonas paederii*', ein weiteres prominentes Beispiel eines Naturstoffproduzenten, zu den mit Insekten in Symbiose lebenden Bakterien. '*P. paederii*' ist in Käfern der Gattung *Paederus* und *Paederidus* zu finden. Diese Käfer bevorzugen warme und feuchte Klimata und sind auf der ganzen Welt beheimatet. Schon vor der Entdeckung des Symbionten war bekannt, dass Käfer dieser Gattung starke, entzündliche Hautreaktionen, die sogenannte Paederus-Dermatitis,

hervorrufen können.¹⁴⁸ Ursächlich für die beschriebene Dermatitis ist das Polyketid Pederin (17, Abbildung 3-10), das aus den Käfern extrahiert werden konnte.¹⁴⁹ Pharmakologisch wirkt Pederin wie alle Vertreter der Pederinfamilie antitumoral durch Inhibition der Proteinbiosynthese.^{150–152} Ein weiteres Mitglied dieser Familie ist das bereits in Kapitel 3.3.1 erwähnte Onnamid aus dem Schwamm *Theonella swinhoei*.¹⁷ Zunächst ging man davon aus, dass Pederin von den Käfern selbst produziert wird und ihnen als chemische Abwehr gegen Fressfeinde dient.¹⁵³ Erst spätere Studien von Kellner *et al.* gaben erste Hinweise darauf, dass nicht das Insekt selbst, sondern ein Bakterium der eigentliche Produzent von Pederin sein könnte.^{154,155} Die Arbeit von Prof. Dr. Jörn Piel aus dem Jahr 2002 brachte schließlich den endgültigen Beweis durch die Klonierung und genaue Beschreibung des Synthese-Genclusters von Pederin.³¹ Phylogenetische Analysen der 16S rRNA zeigten, dass es sich bei diesem bakteriellen Endosymbionten um eine *Pseudomonas*-Spezies handelt, deren nächster Verwandter *Pseudomonas aeruginosa* ist.^{31,155–157} Trotz aller Anstrengungen ist auch '*Candidatus Pseudomonas paederii*' bisher nicht kultivierbar. Ein Grund dafür könnte das aufgrund der Symbiose im Vergleich zum nächsten Verwandten reduzierte Genom sein. Dieses Phänomen tritt bei vielen symbiotisch lebenden Bakterien auf.^{158,159}

'*Pseudomonas paederii*' kann wie *Pseudomonas aeruginosa* phylogenetisch in die Gattung *Pseudomonas* (Familie *Pseudomonadaceae*, Ordnung *Pseudomonadales*) eingeordnet werden. Dies ist eine Gattung gerader oder leicht gekrümmter, stäbchenförmiger Gram-negativer Bakterien. Polare Geißeln ermöglichen ihnen eine aktive Fortbewegung. Pseudomonaden sind fast ausschließlich aerob, auch wenn einige Spezies durch Nitrat, Fumarat oder andere Elektronenakzeptoren auch unter anaeroben Bedingungen wachsen können.^{86,160} Eine weitere charakteristische Eigenschaft dieser Bakterien ist ihre Fähigkeit, viele verschiedene organische Stoffe als Kohlenstoff und Energiequelle nutzen zu können (einige Spezies nutzen bis zu 100 verschiedene Stoffe). Allerdings fehlen ihnen hydrolytische Enzyme, um Polymere in ihre Monomere aufspalten zu können. Um die Vielzahl von Substraten verarbeiten zu können, beinhalten ihre Genome eine Reihe von Enzymgenen mit induzierbaren Operons. Ökologisch nehmen Pseudomonaden eine wichtige Rolle in Boden und Wasserökosystemen ein. Dort sind sie vermutlich für den Abbau von abgestorbenen Tier- und Pflanzenresten verantwortlich.⁸⁶

Unter den Pseudomonaden sind auch einige pathogene Vertreter. *Pseudomonas aeruginosa* ist wohl der bekannteste und am besten untersuchte unter ihnen. Er verursacht häufig Infektionen des Urogenital- und des respiratorischen Traktes. Häufig treten auch Infektionen bei Brandopfern bzw. Patienten mit traumatisch bedingten Hautläsionen (hier kann es auch zu systemische Infektionen kommen) und bei Patienten mit Zystischer Fibrose (auch Mucoviszidose genannt) auf.^{161,162} *P. aeruginosa* ist fakultativ pathogen, d.h. er bedingt nur Infektionen bei bereits immunsupprimierten Menschen. Diese Infektionen sind aufgrund

seiner natürlichen Resistenzen gegenüber vielen Antibiotika schwer zu behandeln. Die Resistenzen werden durch ein Resistenzplasmid vermittelt, das für Proteine codiert, die in der Lage sind, β -Lactam-Antibiotika zu hydrolysieren oder die Antibiotika aus der Zelle heraus pumpen.^{86,163}

Wie bereits erwähnt, gehört '*Pseudomonas paederii*' ebenfalls zur oben beschriebenen Gattung *Pseudomonas*. Wie einige andere Vertreter ist '*P. paederii*' allerdings nicht freilebend, sondern lebt in einer mutualistischen Symbiose mit Käfern der Gattung *Paederus*. Hier liegt der Vorteil dieser Symbiose für '*P. paederii*' in der geschützten Umgebung der weiblichen Käfer, in der diese Bakterien leben. Für die Käfer besteht der Nutzen, wie bereits erwähnt, aus der bakteriellen Produktion von Pederin, das sie als „chemische Waffe“ gegen Predatoren schützt.^{153,155,164} In einer Studie von Keller und Dettner zeigte sich allerdings, dass nicht alle der untersuchten weiblichen Käfer Pederin enthalten: Nur in 90% der Weibchen konnte Pederin nachgewiesen werden. Nur diese sogenannten (+)-Weibchen waren auch in der Lage, diese Eigenschaft an ihre Nachkommen weiter zu geben (vertikale Transmission). Dies gab einen ersten Hinweis auf eine bakterielle Produktion von Pederin. Mithilfe von Fütterungsversuchen konnte die These einer bakteriellen Produktion von Pederin untermauert werden. Dabei wurden die Larven von (-)-Weibchen mit Eiern der (+)-Weibchen gefüttert. Die weiblichen Käfer dieser Larven zeigten im adulten Zustand nun auch ein Pederin-Vorkommen. Behandelte man die Eier der (+)-Weibchen vor der Fütterung jedoch mit bestimmten Antibiotika oder setzte sie Hitze bzw. Kälte aus, blieb eine Pederin-Produktion aus.¹⁵² Das Ergebnis dieses Versuches verdeutlicht auch, dass es sich bei '*P. paederii*' um einen sogenannten sekundären Symbionten handelt, d.h. im Gegensatz zu primären Symbionten ist ihre Anwesenheit für eine Reproduktion ihre Wirtes nicht zwingend erforderlich.⁸⁶

Wie bereits beschrieben, ergab die Analyse der ribosomalen DNA und des Genoms, dass es sich bei *P. aeruginosa* um den nächsten Verwandten von '*P. paederii*' handelt.^{31,155–157} Auf dieser Tatsache beruht auch folgende Hypothese für die Entstehung der Käfer-*Pseudomonas*-Symbiose: Eine frühe Infektion der Käfer durch *P. aeruginosa* könnte zur Entwicklung des Endosymbionten '*P. paederii*' geführt haben. Diese These wird gestützt von den opportunistisch-pathogenen Eigenschaften von *P. aeruginosa*. Es ist denkbar, dass der horizontale Transfer des Pederin-Genclusters in *P. aeruginosa* einen evolutionären Vorteil der infizierten Käfer gegenüber nicht infizierten Spezies darstellte. Durch die toxische Wirkung von Pederin und der damit möglichen Abwehr gegenüber Fressfeinden konnten sich diese Individuen deutlich schneller und besser fortpflanzen.¹⁵⁶

Die Sequenzierung des '*P. paederii*'-Genoms durch Piel und Lipmann (Institut FLI, Jena; unveröffentlicht) ergab eine vorläufige Genomgröße von 5,45 Mbp (schätzungsweise 95-99% der endgültigen Gesamtgröße). Die einzelnen Reads konnten zu 100 Contigs assembliert

werden. Zur Sequenzierung wurde eine Mischung unterschiedlicher Symbiontenstämme gewählt, die aus unterschiedlichen *Paederus*-Käfern isoliert worden waren. Dieses diverse Templat führte zu störenden Sequenzüberlagerungen mit negativen Auswirkungen auf das Sequenzierergebnis und die Assemblierung.¹⁵⁷ Im Vergleich zum *P. aeruginosa*-Genom ist das Genom ca. 20% kleiner und hat mit 5993 vorhergesagten codierenden DNA-Regionen (coding DNA sequence-CDS) ca. 6% weniger CDS als bei *P. aeruginosa*.¹⁵⁷ Wie bereits erwähnt, sind solche reduzierten Genome bei Symbionten aufgrund der symbiontischen Lebensweise häufig zu beobachten.^{158,159} Neben dem Pederin-Gencluster sind noch weitere Naturstoffcluster im 'P. paederi'-Genom codiert, unter anderem ein kleiner NRPS-PKS Gencluster (Abbildung 3-21, Prof. Dr. Jörn Piel).

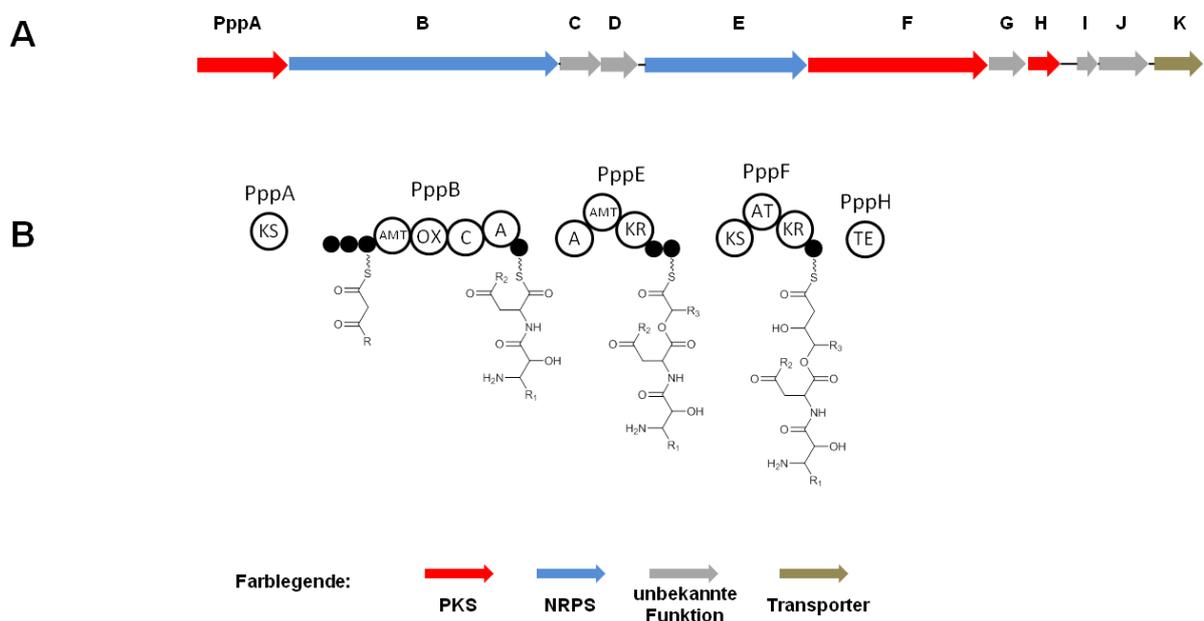


Abbildung 3-21: NRPS-PKS aus 'Pseudomonas paederi'.

Unter **A** dargestellt ist das NRPS-PKS Gencluster (PppA-K). Die Farben kennzeichnen die putative Funktion der ORFs (*open reading frame*). Unter **B** dargestellt ist die jeweilige Domänenzusammensetzung der Synthesemodule. **KS**=Ketosynthase; **AMT**=Aminotransferase; **OX**=Oxidase; **C**=Kondensation, **A**=Adenylierung; **KR**=Ketoreduktase; **TE**=Thioesterase.

Dieser besteht aus 11 open reading frames (ORF), *pppA-K*, und hat eine Größe von 24695 bp. Das eigentliche PKS-NRPS-Hybrid wird von fünf der 11 ORFs (*pppA*, *pppB*, *pppE*, *pppF*, *pppH*) codiert, die je ein Modul umfassen.

Tabelle 3-1: Open reading frames (ORFs) eines PKS-NRPS Genclusters aus 'P. paederi'.

ORF	vorhergesagte Funktion	Sequenzähnlichkeit (Protein, Organismus)	Query cover/ Identity	Zugangsnummer
pppA	PKS (KS) Hybrid NRPS/Typl PKS	Modulare PKS , <i>Paederus fuscipes</i>	29%/99%	AAQ91796
pppB	(ACP, ACP, AMT, OX, C, A, PCP)	Hybrid NRPS/Typl PKS , <i>Stigmatella aurantiaca</i>	95%/33%	WP_002613306
pppC	Phytanoyl-CoA-Dioxygenase	Phytanoyl-CoA-Dioxygenase, <i>Mycobacterium kansasii</i>	96%/48%	ORB82811
pppD	unbekannt Hybrid NRPS/Typl PKS			
pppE	(A, AMT, KR, PCP, PCP)	NRPS , <i>Mastigocoleus testarum</i>	89%/32%	WP_058183473
pppF	PKS (KS, AT, KR, ACP)	AT-Domäne, <i>Stigmatella erecta</i>	98%/37%	SEU18803
pppG	alpha-ketoglutarate-dependent halogenase	chlorinating enzyme, <i>Mycobacterium kansasii</i>	89%/61%	WP_036406945
pppH	Thioesterase	Thioesterase, <i>Bacillus thuringiensis</i>	81%/27%	WP_003301888
pppI	GCN5-related N-acetyltransferase	GCN5-related N-acetyltransferase, <i>Herpetosiphon aurantiacus</i>	92%/38%	ABX05906
pppJ	C-Domäne	NRPS, <i>Paenibacillus alvei</i>	91%/24%	WP_005550894
pppK	MFS Transporter	MFS Transporter, <i>Photorhabdus luminescens</i>	96%/59%	WP_011145292

Anhand der Domänenstruktur konnte mithilfe des nichtribosomalen Codes (Kapitel 3.1.1 und 3.1.2) bereits bioinformatisch eine erste Prognose zur Substratspezifität gegeben werden. Die A-Domäne des zweiten Moduls könnte Glutamin oder Glutaminsäure aktivieren, die potenziellen Substrate der zweiten A-Domäne (PppE) sind 4-Methyl-2-oxopentansäure, 3-Methyl-2-oxobutansäure und 2-Oxo-propionsäure (Pyruvat).

3.4 Systematik von Bakterien

3.4.1 Phylogenie

Das Verständnis evolutionärer Beziehungen (Phylogenie) ist Voraussetzung für die Charakterisierung neuer Organismen aber auch unbekannter Gene oder Proteine.^{86,165} Anders als bei Tieren und Pflanzen ist die phylogenetische Analyse von Bakterien anhand ihrer Morphologie nicht möglich. Trotzdem erfolgen erste systematische Einordnungen von Bakterien ausschließlich über physiologische und biochemische Untersuchung von Reinkulturen. Führt man sich allerdings vor Augen, dass nur ein Bruchteil aller Bakterien sich mit heutigen Methoden kultivieren lassen ist (z. B. aus Salzwasser: 0,001-0,1%¹⁶⁶⁻¹⁶⁸ Süßwasser: 0,25%, Sediment: 0,25%,¹⁶⁹ Boden: 0,1-0,5%¹⁷⁰), so wird klar, dass mit dieser Methodik eine riesige Lücke in der zu beschreibenden Bakteriendiversität entstand.⁸⁸ Außerdem korreliert eine solche Klassifizierung bei Mikroorganismen nicht immer mit dem tatsächlichen evolutionären Verwandtschaftsverhältnissen, wie man heute weiß.^{171,172} Daher finden phylogenetische Bestimmungen heutzutage indirekt über die Analyse von Sequenzierungsdaten statt.¹⁶⁵ Grundlage dafür ist die Annahme, dass alle Organismen über gemeinsame Vorfahren miteinander verwandt sind und die Nukleotidsequenz diese Phylogenie (durch evolutionär bedingte Sequenzvariationen) widerspiegelt.⁸⁶ Zur phylogenetischen Analyse können verschiedene Gene genutzt werden. Das für Prokaryoten meist genutzte und auch in dieser Arbeit verwendete Gen ist das Gen der 16S ribosomalen RNA (rRNA). Woese *et al.* zeigten 1980, dass die kleine Untereinheit des

bakteriellen Ribosoms aufgrund folgender Eigenschaften sehr gut zur phylogenetischen Analyse geeignet ist: Die 16S rRNA ist universell (kommt in allen Bakterien vor), sie hat eine feste Funktion, sie ist ausreichend konserviert und ihre Sequenz ist ausreichend lang, um eine substantielle Aussage über Verwandtschaftsgrade geben zu können.¹⁷³

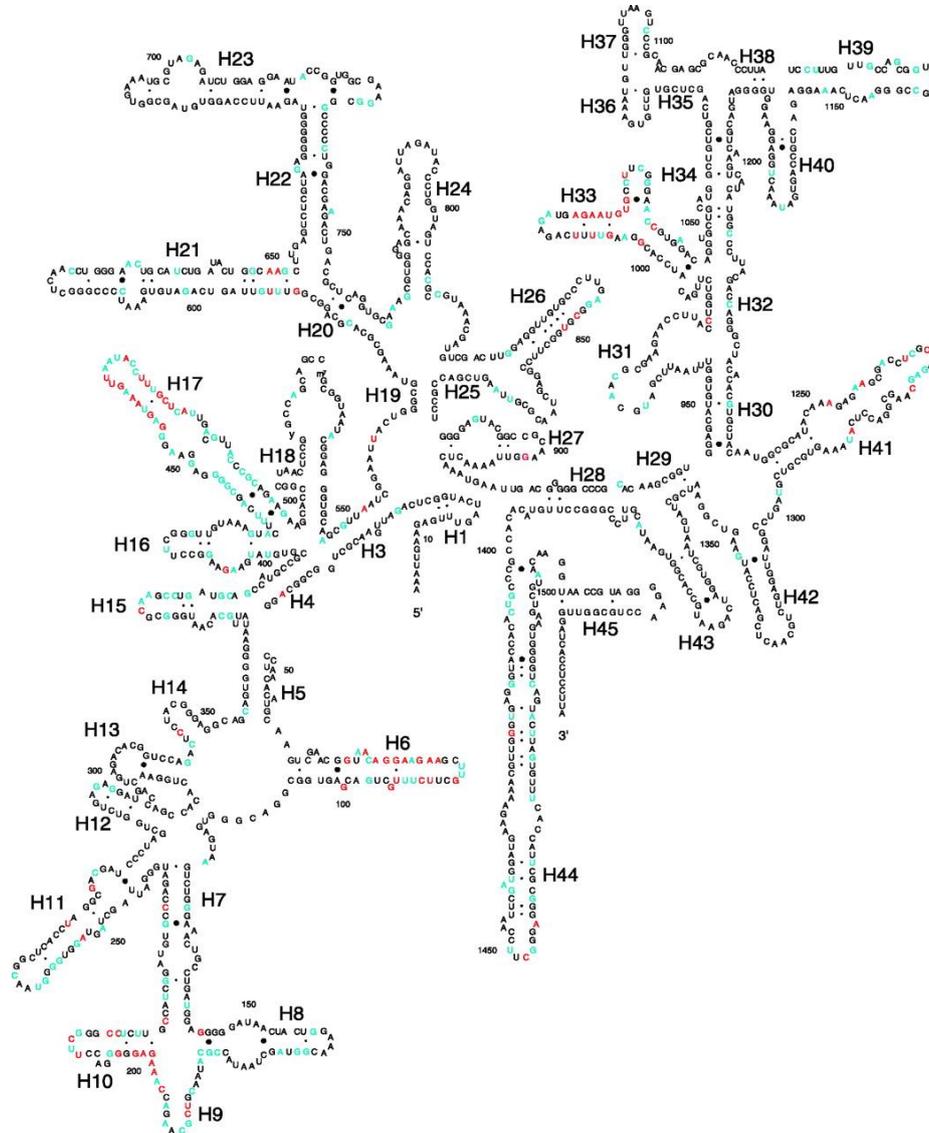


Abbildung 3-22: Sekundärstruktur der 16S RNA.¹⁷⁴ Farblich markiert sind heterogene Positionen, d.h. Nukleotidsubstitutionen, die innerhalb eines Organismus (blau) bzw. zwischen Organismen auftreten (rot). Helices (H) sind entsprechend der Nomenklatur von Cannone *et al.*¹⁷⁵ nummeriert.

Trotz dieser guten Eignung kann die Varianz des 16S rRNA Gens unzureichend sein, um zwei nah verwandte Spezies voneinander unterscheiden zu können. Daher können auch weitere Genloci in sogenannten *Multilocus sequence typing* (MLST)-Analysen untersucht oder ganze Genomanalysen zur phylogenetischen Einordnung durchgeführt werden.⁸⁶

Die Molekulare Uhr

Grundlage einiger Methoden zur Rekonstruktion phylogenetischer Stammbäume ist die von Zuckerkandl und Pauling Mitte der 60er Jahre entwickelte Theorie der „Molekularen Uhr“. Sie beobachteten, dass die Aminosäuresequenz des Hämoglobins immer stärker differierte, je länger sich zwei Arten bereits evolutionär unabhängig voneinander entwickelten. Aus dieser Beobachtung schlossen sie, dass sich die Sequenz bestimmter Gene (z. B. der 16S rRNA) evolutionär mit zeitlich konstanter Mutationsrate verändert.^{176–178} Nach dieser Hypothese wäre es dann möglich, anhand der DNA bzw. AS-Sequenz den Zeitpunkt der Aufspaltung zweier Arten von ihrem gemeinsamen Vorfahren zu bestimmen. Denn je mehr Mutationen entstanden sind, desto länger ist die Evolutionszeit seit der Aufspaltung („Mutationsrate~Zeit“).⁸⁶ Diese Annäherung wurde beispielsweise genutzt, um den Zeitpunkt der Aufspaltung entfernt verwandter Organismen zu bestimmen (Aufspaltung der Domänen Archaea und Eukaryota vor ca. 2,8 Milliarden Jahren).⁸⁶ Nachteil dieser Hypothese ist, dass Mutationraten eben nicht konstant sind und somit eine direkte Korrelation mit der Zeit schwierig ist. Die Mutationsrate und damit die Molekulare Uhr werden von verschiedensten Faktoren beeinflusst, die berücksichtigt werden müssen: 1. Generationszeit (kurze Generationszeiten beschleunigen die Molekulare Uhr, da sich Mutationen schneller manifestieren können.); 2. Populationsgröße (eine große Population verlangsamt die Geschwindigkeit der molekularen Uhr, da es länger dauert, bis sich Mutationen durchsetzen.); 3. Spezies (unterschiedliche Polymerasen und andere Spezies-spezifische biologische Eigenschaften können die Molekulare Uhr verlangsamen oder beschleunigen.); 4. Evolutionär bedingte Veränderung der Proteinfunktion; 5. Selektionsdruck (der durch die Umwelt einwirkende Selektionsdruck ist nicht berechenbar und wirkt unterschiedlich stark auf verschiedene Spezies.).¹⁷⁹ Die meisten phylogenetischen Analysen beschäftigen sich jedoch mit relativen Verwandtschaftsverhältnissen zwischen verschiedenen Organismen. Daher spielt die Genauigkeit der Molekularen Uhr eine eher untergeordnete Rolle.⁸⁶

3.4.2 Evolutionsanalyse von Bakterien

Zur phylogenetischen Analyse muss zunächst das zu untersuchende Gen, z. B. das der 16S rRNA, mithilfe von Standardprimern in einer Polymerasenkettenreaktion (PCR) amplifiziert und anschließend sequenziert werden. Typischerweise wird hierzu genomische DNA verwendet, die entweder aus spezifischen Bakterienstämmen oder aus Umweltproben isoliert wurde (Abbildung 3-23).⁸⁶ In einem zweiten Schritt wird die zu analysierende Gensequenz mit homologen Sequenzen anderer Organismen verglichen (engl.: *Alignment*). Dabei werden evolutionär bedingte Nukleotid-Varianten, Insertionen und Deletionen (InDel) genau bestimmt.⁸⁶ Häufig hat dabei das Sequenz-Alignment einen größeren Einfluss auf das Ergebnis der phylogenetischen Analyse als die eigentliche Methode zur

Stammbaumkonstruktion.^{180,181} Für das Alignment von Sequenzen stehen Methoden mit unterschiedlichen Algorithmen zur Verfügung (z. B. ClustalW, MUSCLE).^{182–184} Anschließend wird über verschiedene Algorithmen anhand des Sequenzalignments ein sogenannter phylogenetischer Stammbaum (graphische Darstellung der verwandtschaftlichen Beziehungen) erstellt (Abbildung 3-23).⁸⁶

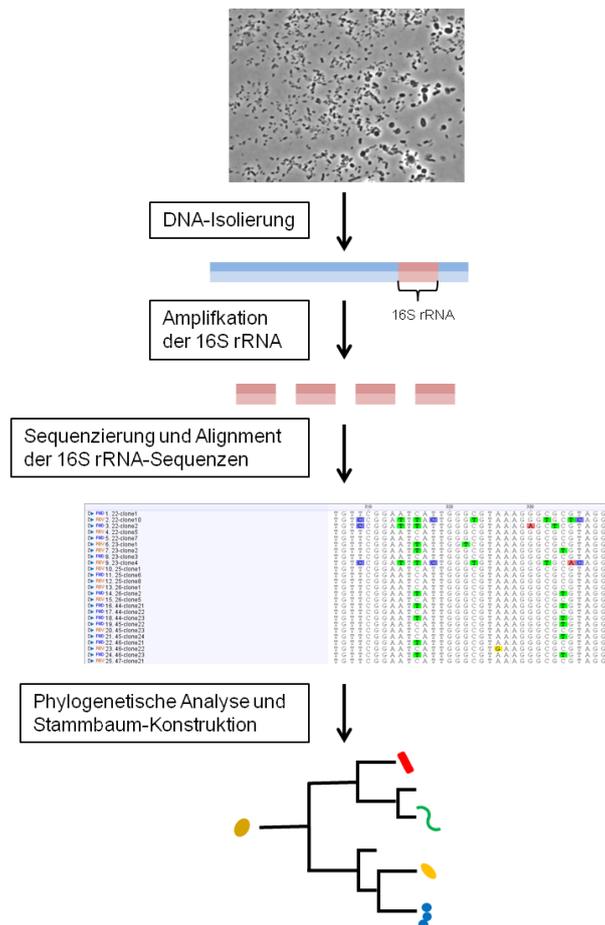


Abbildung 3-23: Arbeitsschritte einer phylogenetischen Analyse.⁸⁶ Nach der Isolierung genomischer DNA aus einer Umweltprobe oder einer Kultur wird mithilfe universeller 16S rRNA-Primer die 16S rRNA in einer PCR amplifiziert. Die Sequenzen der PCR-Produkte werden dann in einem sogenannten Sequenzalignment miteinander verglichen. Schließlich kann anhand dieses Alignments nach Wahl der entsprechenden Methode ein Stammbaum konstruiert werden.

Phylogenetische Stammbäume bestehen aus sogenannten Knoten (Verzweigungspunkte) und Ästen, an deren Ende die einzelnen Spezies stehen (engl.: *operational taxonomic units*, OTUs). Dabei ist die Astlänge proportional zur Anzahl der evolutionären Ereignisse.¹⁸⁵ Die äußeren Enden werden auch analog zur „Baum“-Darstellung als Blätter bezeichnet. Die Knoten repräsentieren den Punkt in der Evolution, an dem sich ein gemeinsamer Vorfahr in zwei neue Spezies (Schwestertaxa) aufspaltete, die sich separat entwickelten.¹⁸⁶ Als Regel gilt: Je mehr gemeinsame Vorfahren (Knoten) zwei Spezies haben, desto näher sind sie miteinander verwandt. In phylogenetischen Stammbäumen können entweder relative

Verwandtschaftsbeziehungen zwischen den untersuchten Organismen (engl.: *unrooted*) oder auch der evolutionäre Verlauf von einem gemeinsamen Vorfahren bis zum OTU (engl.: *rooted*) dargestellt werden. Eine solche „Wurzel“ kann durch Einführung einer sogenannten Außengruppe (engl.: *outgroup*) erzeugt werden. Diese Außengruppe wird von einem Organismus gebildet, der mit den Organismen der Analyse weiter entfernt verwandt ist als diese untereinander. Die Spezies der Außengruppe muss allerdings trotz der entfernten Verwandtschaft ein Homolog des untersuchten Gens besitzen.^{86,180}

Methoden zur Stammbaumrekonstruktion

Zur Rekonstruktion phylogentischer Stammbäume stehen verschiedene Methoden mit unterschiedlichen Ansätzen zur Verfügung. Aufgrund der Komplexität dieses Themas werden im Folgenden nur einige wichtige Methoden kurz vorgestellt. Generell unterscheidet man zwischen quantitativen *Distanz*-basierten Methoden (bestimmen die Ähnlichkeit zwischen zwei OTUs) und den qualitativen, kladistischen Methoden, der sogenannten *Maximum Parsimony* und den *Likelihood*-Methoden (charakterbasiert).^{165,187}

Distanzbasierte Methoden verwenden die „genetische Distanz“, die sich aus dem dem Vergleich der Sequenzen (dem oben beschriebenen Alignment) ergibt, um daraus eine sogenannte Distanz-Matrix zu erstellen. Dabei repräsentiert die Distanz die Unterschiedlichkeit zweier Taxa. Daraus wird dann ein Baummodell erstellt, was möglichst genau die genetische Distanz dieser Matrix widerspiegelt.¹⁸⁰ Distanz-basierte Algorithmen zur Berechnung eines phylogenetischen Stammbaums sind die sogenannte *Unweighted Pair Group Method with Arithmetic mean (UPGMA)*, das *Neighbour Joining*, und das Prinzip der *Minimum Evolution*.^{188–190}

In kladistischen Methoden werden phylogenetische Beziehungen anhand von Nukleotidänderungen an bestimmten, phylogenetisch informativen Sequenzpositionen definiert. Dadurch können die Spezies in sogenannte monophyletische Gruppen (Gruppen die von einem Vorfahren abstammen) eingeteilt werden (Abbildung 3-24).⁸⁶

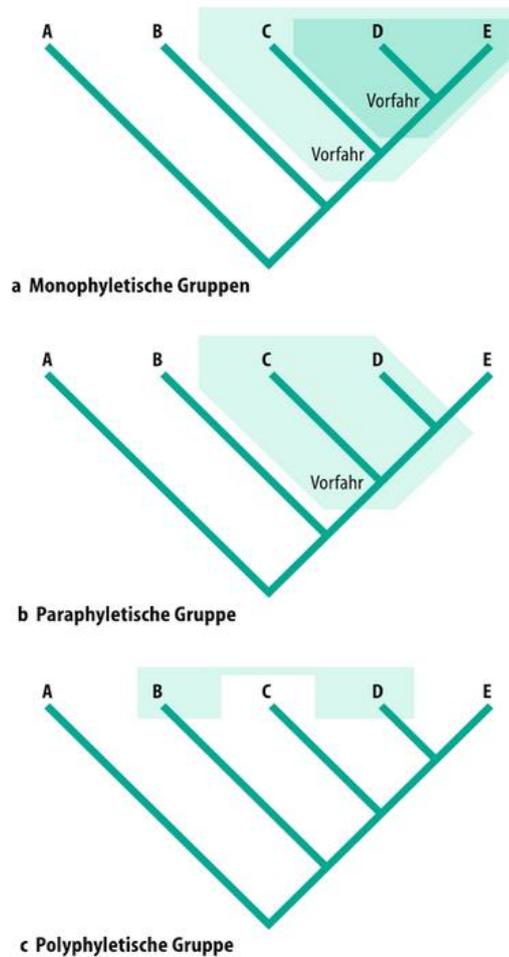


Abbildung 3-24: Beschreibung monophyletischer, paraphyletischer, polyphyletischer Gruppen.¹⁸⁵

a Monophyletische Gruppe (Klade): Alle Nachkommen eines gemeinsamen Vorfahrs, z. B. DE oder CDE.

b. Paraphyletische Gruppe: Einige, aber nicht alle Nachfahren eines gemeinsamen Vorfahrs. Ein Mitglied dieser Gruppe (D) ist näher mit einer Gruppe außerhalb dieses Taxons (E) als mit den anderen Mitgliedern innerhalb seiner Gruppe (C) verwandt.

c. Polyphyletische Gruppe: Mitglieder dieser Gruppe haben keinen gemeinsamen Vorfahren.

Die *Maximum Parsimony* beruht, wie der Name schon sagt, auf dem Prinzip der „einfachsten“/„sparsamsten“ Erklärung (Parsimonie).¹⁹¹ Bei dieser Methode werden aus den Datensätzen (z. B. dem Sequenzalignment) verschiedene mögliche Stammbäume konstruiert. Diese spiegeln die verschiedenen möglichen verwandtschaftlichen Beziehungen zwischen den OTUs wider. Die unterschiedlichen Stammbäume werden dann anhand der Anzahl der evolutionären Veränderungen (z. B. Nukleotidaustausche) bewertet, um das im Baum dargestellte Evolutionsszenario zu generieren. Nach dem Parsimonie-Prinzip gibt der Baum mit dem geringsten Wert (der einfachste/sparsamste Baum) die tatsächlichen evolutionären Beziehungen der OTUs am besten wieder.^{191–193}

Das Prinzip des *Maximum Likelihood* ist die differenzierteste der drei vorgestellten Methoden. Anders als in Distanz- und Parsimonie-basierten Methoden ermittelt die *Maximum Likelihood* Methode die Wahrscheinlichkeit, mit der ein möglicher Stammbaum die korrekte Topologie und Astlänge besitzt. Dabei werden auch die unterschiedlichen

Veränderungsraten der einzelnen Charaktere (auch Merkmale genannt; z. B. der Nukleotide) berücksichtigt, beispielsweise treten Transitionen (Purinbase zu Purinbase oder Pyrimidinbase zu Pyrimidinbase) häufiger auf als Transversionen (Purinbase zu Pyrimidinbase) und umgekehrt. Diese Parameter werden für jede einzelne Alignmentposition errechnet. Das macht diese Methode zum genauesten, aber gleichzeitig auch aufwendigsten Prinzip zur Stammbaumrekonstruktion.^{191,194–196}

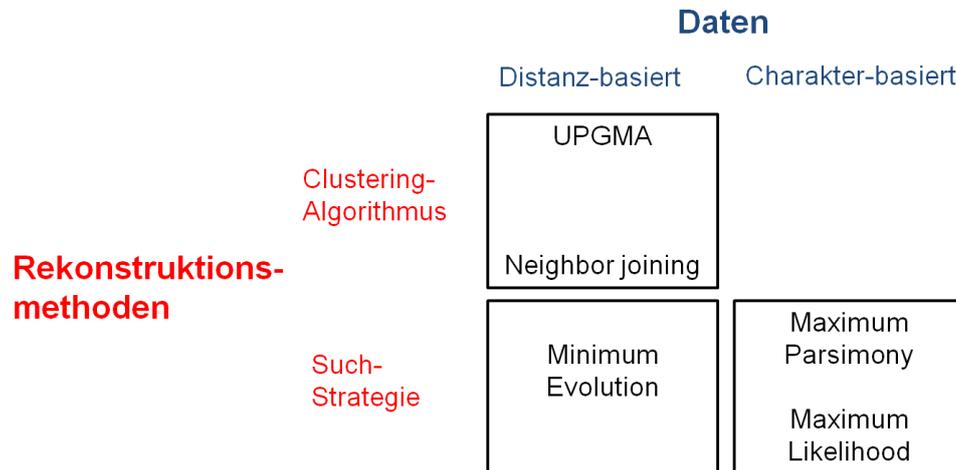
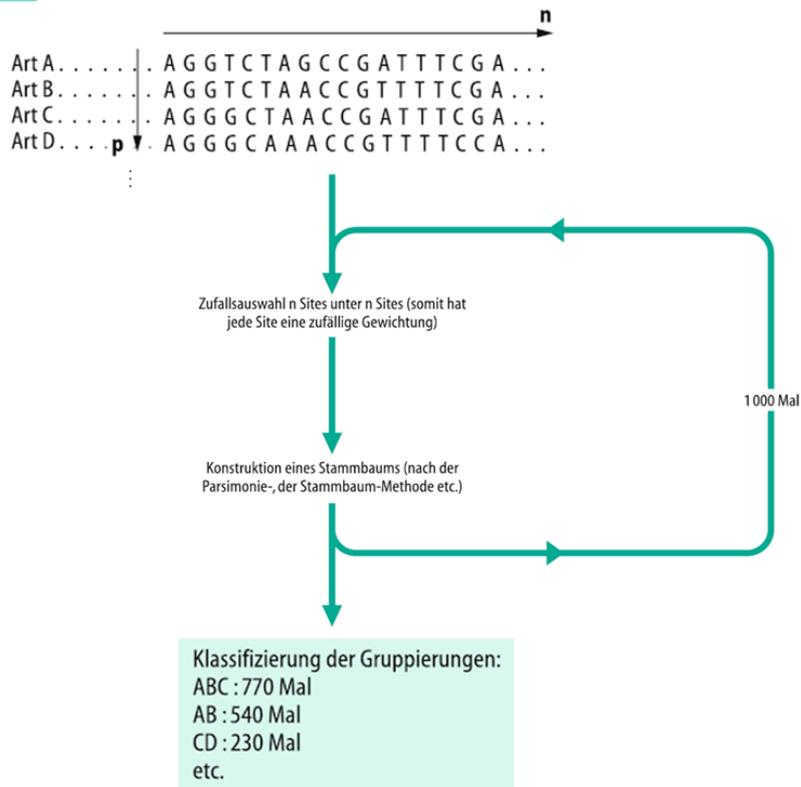


Abbildung 3-25: Übersicht über Methoden zur Rekonstruktion von phylogenetischen Stammbäumen.

Ziel der vorgestellten Methoden ist es, die Evolutionsgeschichte zu rekonstruieren und in Form einer Baumtopologie wiederzugeben. Dabei wird jedoch keine qualitative Bewertung des errechneten Stammbaums gegeben. Mithilfe verschiedener Verfahren kann die Qualität eines Stammbaums bewertet werden. Im Folgenden soll die in dieser Arbeit verwendete Methode des *bootstrap* kurz erläutert werden.^{197,198}

Bootstrap-Werte geben Auskunft darüber, wie robust die im Stammbaum wiedergegebene Evolutionshypothese ist. Das Prinzip des *bootstrap* besteht aus der zufälligen Auswahl und dem ebenso zufälligen Ausschluss von Charakteren bzw. Merkmalen bei einem erneuten Sequenzvergleich. Praktisch heißt dies, dass nach der Auswahl eines bestimmten Merkmals dieses mit so vielen weiteren Merkmalen verglichen wird, bis die Größe des Datensatz der Größe des Originaldatensatzes entspricht. Dabei können, wie bereits erwähnt, einzelne Merkmale mehrfach, andere dafür nicht berücksichtigt werden. Aus dieser Stichprobe wird nun ein neuer Stammbaum konstruiert. Dieses Vorgehen wird mehrfach wiederholt (einige hundert bis tausend Mal). Aus der Vielzahl der daraus resultierenden Stammbäume wird nun ein sogenannter Konsensusstammbaum berechnet. Darin wird für jeden Knoten der individuelle *bootstrap*-Wert angegeben. Dieser gibt den Prozentsatz der berechneten Stammbäume an, in dem dieses Evolutionsmodell, also diese Verzweigungspunkt, so berechnet wurde.¹⁸⁵

Sequenzvergleich



Hauptkonsenz

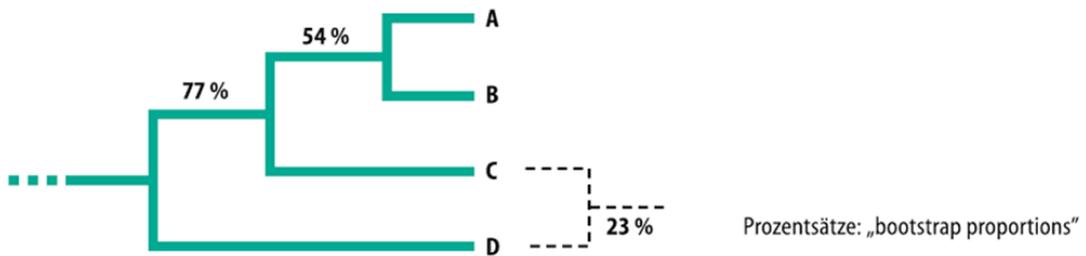


Abbildung 3-26: Prinzip des bootstrap¹⁸⁵ n=Merkmale, p= Arten

Abschließend ist zu sagen, dass jeder Stammbaum nur ein Modell der darzustellenden Evolutionsgeschichte ist. Es ist unmöglich, den einen Stammbaum zu konstruieren, der in der Lage ist, den tatsächlichen Verlauf der Evolution darzustellen.

4 Zielsetzung

Wie bereits in der Einleitung angeklungen, ist die Entwicklung neuer Arzneistoffe insbesondere von Antiinfektiva und Chemotherapeutika von enormer gesellschaftlicher, medizinischer und ökonomischer Bedeutung.^{3,4} Eine wichtige Quelle für pharmakologisch wirksame Substanzen bilden mikrobielle Symbionten.⁹⁵ In dieser Arbeit sollen drei dieser talentierten Naturstoffproduzenten genauer untersucht werden.

4.1 '*Candidatus Entotheonella factor*'

Aus dem Schwamm *Theonella swinhoei* konnte in Vorarbeiten ein filamentöses Bakterium, '*Candidatus Entotheonella factor*' isoliert werden. Erste Versuche legten nahe, dass dieses Bakterium der eigentliche Produzent vieler, aus dem Schwamm isolierter Naturstoffe ist (Kapitel 3.3.1). Im Rahmen dieser Arbeit sollten anhand der Sequenzinformationen aus Genomsequenzierungen durch die Arbeitsgruppe von Prof. Dr. Jörn Kalinowski in Zusammenarbeit mit Dr. Micheal Wilson und Nina Heycke die Gencluster dieser Naturstoffe assembliert und identifiziert werden.

Erste BLAST-Analysen zeigten, dass nahe Verwandte dieses Symbionten auch in anderen Habitaten zu finden sind (z. B. NCBI GenBank: JN178754, AY493913, JN825509) (Kapitel 3.3.1). Daher sollte, als weiteres Ziel, untersucht werden, in welchen Lebensräumen diese Bakterien zu finden sind. Dazu sollten verschiedene Umweltproben untersucht werden. Hierbei isolierte, nah verwandte Spezies könnten ebenfalls eine Rolle als vielversprechende Naturstoffproduzenten spielen.

4.2 *Symbiont 2 des Schwammes Theonella swinhoei*

Neben 'E. factor' und 'E. gemina' konnte ein weiteres filamentöses Bakterium aus dem Schwamm *Theonella swinhoei* isoliert werden. Auch mit diesem Symbionten wurden in vorangegangenen Arbeiten Genomsequenzierungen durchgeführt (Kapitel 3.3.1). Im Rahmen dieser Arbeit sollten die Genomdaten auf Naturstoffcluster hin untersucht und die nach der Assemblierung verbliebenen Sequenzlücken geschlossen werden.

4.3 '*Candidatus Pseudomonas paederii*'

Der Käfersymbiont '*Candidatus Pseudomonas paederii*' ist bekannt für die Produktion des Polyketids Pederin.^{149,155,156} Genomsequenzierungen zeigten, dass sein Genom neben diesem einen weiteren PKS-NRPS-Gencluster enthält (Kapitel 3.3.2). In dieser Arbeit sollte die Substratspezifität der beiden A-Domänen dieses Clusters bestimmt werden. Als weiteres Ziel sollten die noch vorhandenen Sequenzlücken innerhalb des Genoms durch kombinatorische PCR geschlossen werden.

5 Ergebnisse und Diskussion

5.1 '*Candidatus Entotheonella factor*'

Erste Arbeiten zeigten, dass viele Naturstoffe des Schwammes *Theonella swinhoei* vermutlich bakteriellen Ursprungs sind (Kapitel 3.3.1).^{131,199} Kapitel 5.1.1 beschreibt die Identifizierung und Assemblierung einiger dieser Naturstoffgencluster sowie das Schließen weiterer Lücken innerhalb des '*Entotheonella factor*'-Genoms. BLAST-Analysen gaben Hinweise darauf, dass nah verwandte Spezies auch in anderen Habitaten wie z. B. Salzseen und Bodenproben zu finden sind (z. B. NCBI GenBank: JN178754, AY493913, JN825509). Daher sollte die Verbreitung dieser Spezies, auch im Hinblick auf ihr mutmaßliches Potential als Naturstoffproduzenten, genauer untersucht werden. Zur Charakterisierung der Bakterien sollten diese dann ggf. isoliert werden. Kapitel 5.1.2 und 5.1.4 zeigen die Ergebnisse der Umweltanalysen und der Isolierungsversuche dieser '*Candidatus Entotheonella*'

5.1.1 Assemblierung des '*Entotheonella factor*' Genoms

Wie in Kapitel 3.3.1 beschrieben, konnte in Vorarbeiten '*Entotheonella factor*' aufgrund seiner charakteristischen Sedimentationseigenschaften durch differentielle Zentrifugation separiert und angereichert werden. Mit der so gewonnenen Zellfraktion aus *Theonella swinhoei* (gelber Chemotyp) wurden, in Kollaboration mit der Arbeitsgruppe von Prof. Dr. Jörn Kalinowski (Universität Bielefeld, Verbundprojekt GenBioCom), mehrere Illumina-, 454-, PacBio- und Sanger-Sequenzierungsläufe durchgeführt.^{131,199–204} Dabei zeigte sich, dass die metagenomische DNA, neben verschiedenen Eukaryoten, zwei sehr nah verwandte Genome der Gattung '*Candidatus Entotheonella*' (mit einer Übereinstimmung der 16S rRNA von 97,6%) beinhaltet.¹³¹ Die Sequenzhomologien dieser beiden sehr eng verwandten Genome und die Vielzahl an repetitiven Elementen machten eine endgültige, rein bioinformatische Assemblierung nur anhand der Sequenzierungsergebnisse unmöglich. Daher wurde in Zusammenarbeit mit Dr. Micheal Wilson und Nina Heycke versucht, die noch vorhandenen Lücken innerhalb der Contigs bzw. Scaffolds durch PCR zu schließen (Kapitel 7.4.1). Aufgrund der spezifischen Primer war das Schließen der Lücken innerhalb der Contigs/Scaffolds unproblematisch. Beispielhaft sind einige der PCR-Produkte nach deren Aufreinigung in Abbildung 5-1 dargestellt.

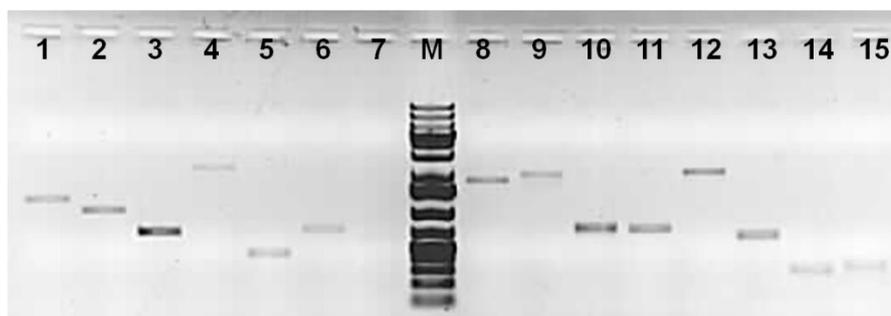
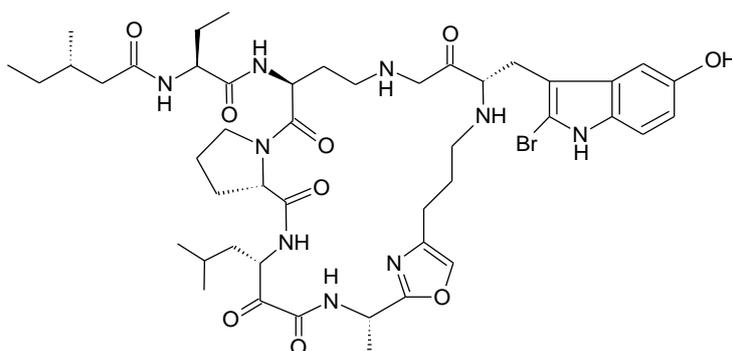


Abbildung 5-1: Schließen der Lücken innerhalb von Scaffolds. Das Agarosegel zeigt aufgereinigte PCR-Produkte zur Schließung von Sequenzlücken innerhalb einiger Scaffolds. 1-10 bezeichnen die verschiedenen PCR-Produkte; **M**: Fermentas 1 kb plus Ladder. (1: Scaffold 1; 2: Scaffold 3; 3: Scaffold 60; 4: Scaffold 84; 5: Scaffold 86; 6: Scaffold 98; 7: Scaffold 98 (zweite Lücke); 8: Scaffold 98 (dritte Lücke); 9: Scaffold 111; 10: Scaffold 111 (zweite Lücke); 11: Scaffold 114; 12: Scaffold 114 (zweite Lücke); 13: Scaffold 114 (dritte Lücke); 14: Scaffold 118; 15: Scaffold 333; 16: Scaffold 374). Mithilfe der Sequenzinformationen dieser PCR-Produkte konnten mit Ausnahme von 7 und 8 die entsprechenden Lücken durch Assemblierung geschlossen werden.

Nach dieser Methodik wurden in Zusammenarbeit mit Dr. Augustinus Uria auch die nach einer ersten Sequenzierung verbliebenen Sequenzlücken des Orbiculamid-Biosyntheseclusters geschlossen.



29

Abbildung 5-2: Orbiculamid A. ²⁰⁵

Orbiculamid (**29**) ist nah verwandt mit anderen bereits gezeigten, aus dem Schwammes *Theonella swinhoei* isolierten, zyklischen Peptiden (Keramamid D, **28**, Abbildung 3-18).

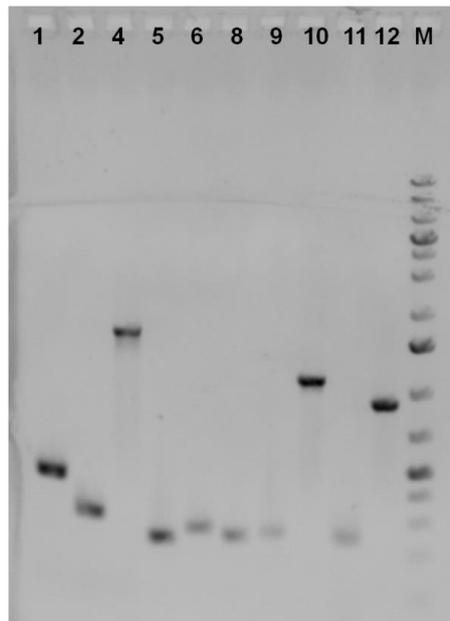


Abbildung 5-3: Schließung der Lücken im Orbiculamid-Gencluster. Das Agarosegel zeigt die aufgereinigten PCR-Produkte (1-12: entsprechen den unterschiedlichen Primerpaaren; M: Fermentas 1 kb plus Ladder).

Contigs bzw. Scaffolds wurde mithilfe von kombinatorischer PCR verbunden. Bei diesem Verfahren werden Primer erstellt, die jeweils mindestens 100 bp vom Contigende entfernt binden. Die Primer werden deshalb in einiger Entfernung vom Contigenden gelegt, da die Sequenzqualität der ersten Basenpaaren bei einer Sanger-Sequenzierung meist ungenügend ist und eine eindeutige Bestimmung der Sequenz in diesem Bereich unmöglich macht. Durch Vergleich der Sequenzdaten mit dem Genom des nächsten Verwandten und durch Analyse der Gene an den Enden der jeweiligen Contigs/Scaffolds können Vorhersagen zu deren Reihenfolge getroffen werden. Bei der Analyse der Contigenden liegt der Fokus auf unvollständigen Gensequenzen: Findet sich ein Teil der fehlenden Gensequenz am Ende eines anderen Contigs wieder, ist dies ein Hinweis darauf, dass es sich bei diesen beiden Contigs um direkte Nachbarn handeln könnte. Anhand dieser Vorhersagen werden dann alle möglichen Primerkombinationen in separaten PCR getestet (Kapitel 7.4.1). Anschließend werden die daraus resultierenden Banden nach der Gelelektrophorese aus den Agarosegelen ausgeschnitten und die DNA mittels Gelextraktion isoliert (Kapitel 7.4.2.1). Die DNA-Fragmente werden dann mittels Sanger-Sequenzierung sequenziert (Kapitel 7.4.6.1). Handelt es sich um die richtige Primerkombination, kann die Lücke zwischen den entsprechenden Contigs durch Assemblierung mit der so gewonnenen Sequenzinformation geschlossen werden (Abbildung 5-4).

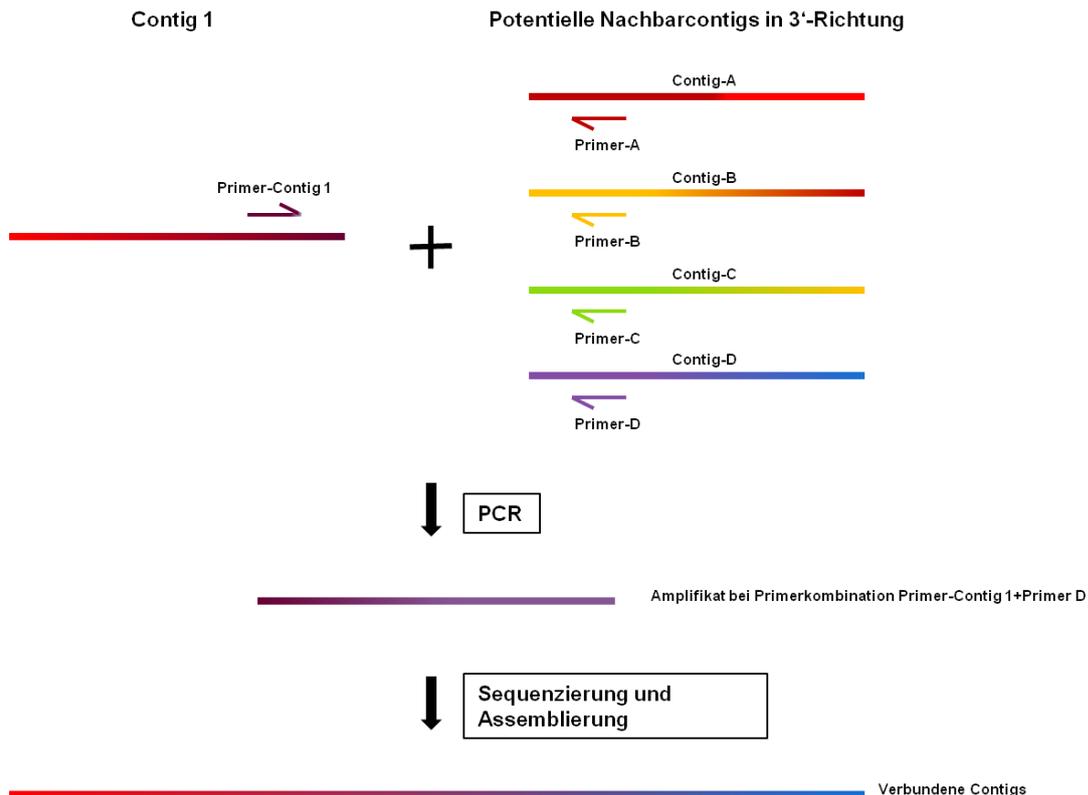


Abbildung 5-4: Kombinatorische PCR zur Schließung von Sequenzlücken.

Schematisch dargestellt ist die Verbindung zweier Contigs durch kombinatorische PCR. Dazu werden zunächst Primer erstellt, die ca. 100 bp von den jeweiligen Contig-Ende binden. In mehreren PCR werden dann verschiedene Primer miteinander kombiniert. Im Beispiel wird der Primer von Contig 1 mit den verschiedenen Primern der möglichen Nachbarcontigs (Primer A-D) kombiniert. Bei korrekter Kombination der Primer (im Beispiel Primer-Contig 1+ Primer D) erhält man ein spezifisches PCR-Produkt. Durch Sequenzierung des PCR-Produkts und anschließende Assemblierung des erhaltenen Sequenzfragments mit den beiden Contigs kann die Sequenzlücke geschlossen und so die beiden Nachbarcontigs miteinander verbunden werden.

Anhand der *Coverage* und des GC-Gehalts konnten Scaffolds des bereits in der Einleitung erwähnten Plasmids von Genom-Scaffolds unterschieden werden. Dies erleichterte den Lückenschluss zwischen den Scaffolds des Plasmids, da so weniger Kombinationen möglich waren. Allerdings konnten mit dieser kombinatorischen PCR-Methode viele Lücken, insbesondere innerhalb des Genoms, nicht geschlossen werden, daher wurden verschiedene weitere Ansätze getestet. Neben einer *primer walking*-Methodik,²⁰⁶ die nicht den gewünschten Erfolg brachte, wurde eine, in einer vorangegangenen Arbeiten erstellte, metagenomische Cosmid-Bibliothek (Prof. Dr. Jörn Piel, Dr. Cristian Gurgui) des Schwammes *Theonella swinhoei* mit entsprechenden Primern gescreent, um so weitere Lücken schließen zu können. Außerdem wurden verschieden DNA-Polymerasen mit unterschiedlichen Konditionen getestet. So konnten einige weitere Lücken geschlossen werden.

Zusammenfassend wurden durch Kombination der beschriebenen, unterschiedlichen und sehr zeitintensiven PCR-Ansätze viele der für Naturstoffgencluster codierende Contigs bzw.

Scaffolds assembliert (Abbildung 5-5). Einige Lücken konnten jedoch nur durch wiederholte Genom-Sequenzierungen geschlossen werden. Mit dem Nachweis der Gencluster innerhalb des Bakteriengenoms konnte die bakterielle Herkunft einiger der bekannter Naturstoffe (Kapitel 3.3.1, Abbildung 3-18, mit Ausnahme von Aurantosid, Abbildung 5-5) aus dem Schwamm *Theonella swinhoei* *in silico* gezeigt werden. Aufgrund dieser außergewöhnlich großen Menge an Naturstoffclustern innerhalb des Genoms wurde diese Spezies 'Candidatus Entotheonella factor' (lat. *factor*, der Produzent) genannt.

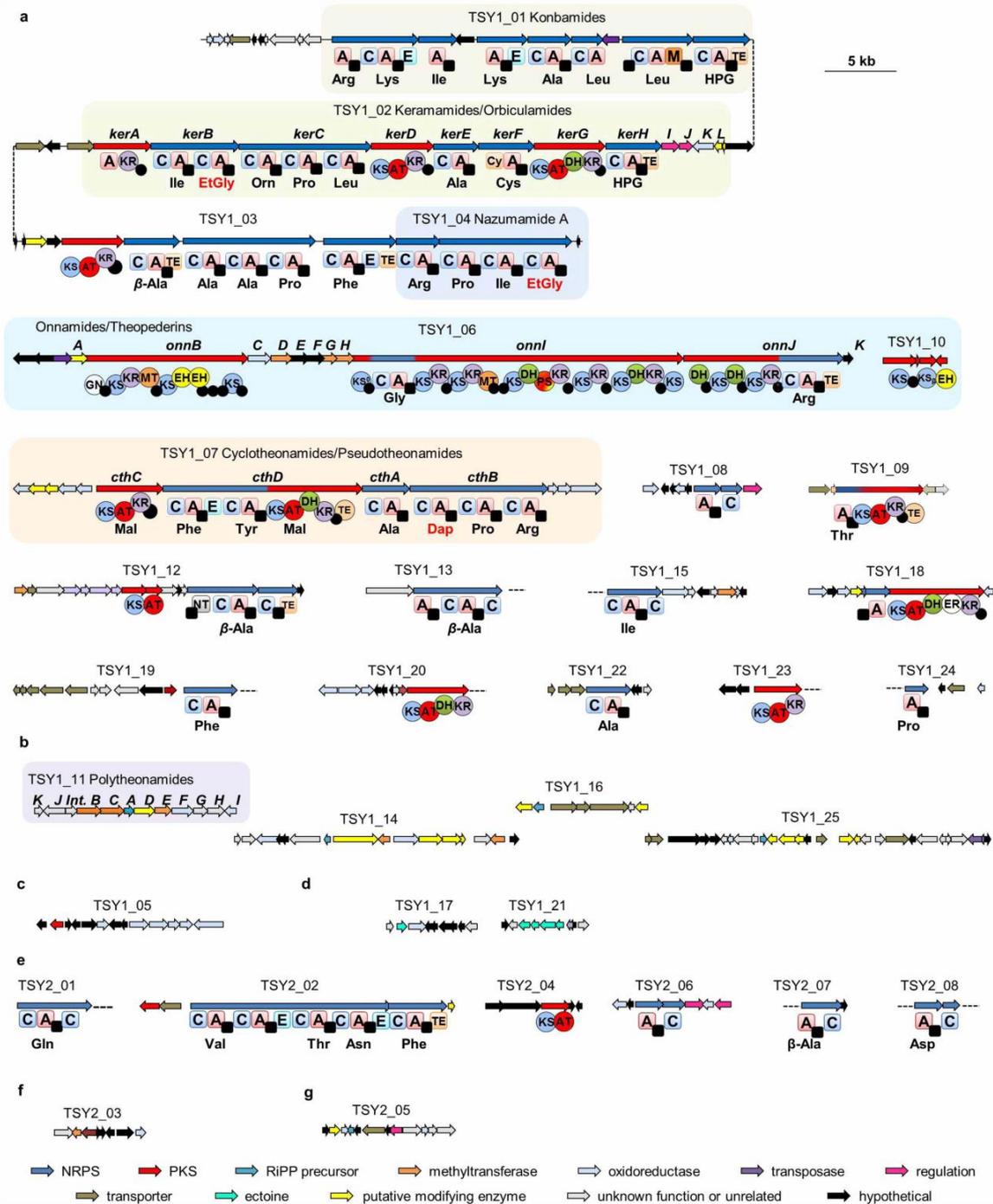


Abbildung 5-5: Biosynthesecluster aus 'Entotheonella factor' und 'Entotheonella gemina'.¹³¹ Dargestellt sind die z. T. assemblierten Naturstoffcluster aus 'Entotheonella factor' (TSY1) bzw. 'Entotheonella gemina' (TSY2). In Zusammenarbeit mit Dr. Agustinus Uriä, Dr. Micheal Wilson und Nina Heycke konnten durch kombinatorischer PCR und durch Screenen einer Cosmid-Bibliothek Lücken geschlossen und einzelne Contigs verbunden werden. Die Proteusin-Gencluster wurden von Dr. Christian Gurgui, Dr. Maximilian Helf und Dr. Michael Freeman assembliert.¹²⁹ Die *open reading frames* (ORF) wurden anhand der vorhergesagten Funktion des Gens farblich markiert. PKS-Domänen (Kugeln) sind: **AT**=Acyltransferase; **DH**=Dehydratase; **ER**=Enoylreduktase; **EH**=Enoyl-CoA-Hydratase; **MT**=Methyltransferase; **KR**=Ketoreduktase; **KS**=Ketosynthase; **TE**=Thioesterase; **GN**=GCN5-like N-Acetyltransferase; **PS**=Pyransynthase. Die Acylcarrierproteine sind als kleine schwarze Kugeln dargestellt. NRPS-Domänen (Vierecke) sind: **A**=Adenylierung; **C**=Kondensation; **E**= Epimerase; **T**=Thiolierungsdomäne; **TE**=Thioesterase. Die Peptidcarrierproteine sind als kleine schwarze Vierecke dargestellt. Die Spezifität der A-Domänen ist unter der jeweiligen Domäne im Dreibuchstabencode angegeben. Die Vorhersage erfolgte entweder bioinformatisch oder anhand der Strukturvorhersagen der Substanzen. Die Biosynthese-Cluster wurden eingeteilt in (a) NRPS-, modulare PKS- und NRPS/PKS Hybrid-Genloci aus 'Entotheonella factor'; (b) Proteusin-Genloci aus 'E. factor'; (c) Typ III-PKS aus 'E. factor'; (d) Ectoin Gencluster

aus 'E. factor'; (e) NRPS- und PKS-Loci aus 'E. gemina'; (f) Typ III-PKS aus 'E. gemina' (g) putative Proteosin-Loci aus 'E. gemina'.

5.1.2 Umweltanalysen

Durch phylogenetische Analysen war bekannt, dass '*Candidatus Entotheonella*' auch in nicht marinen Habitaten vorkommen (Kapitel 3.3.1, z. B. NCBI GenBank: JN178754, AY493913, JN825509).¹³¹ Aus dieser Tatsache heraus entwickelte sich die Fragestellung, ob und in welchem Umfang '*Candidatus Entotheonella*' in weiteren Habitaten zu finden sind und ob bzw. inwieweit sich diese Varianten bzw. Phylotypen morphologisch und metabolisch von 'E. factor' besonders im Hinblick auf die Synthese von Naturstoffen unterscheiden. Dazu sollten verschiedene Umweltproben untersucht werden.

Zur Identifizierung dieser Bakterien musste eine entsprechende Nachweismethode etabliert werden. Um '*Candidatus Entotheonella*' schnell und effizient nachweisen zu können, wurden die Proben mittels PCR (Kapitel 7.4.1) anhand spezifischer 'E. factor' 16S rRNA-Primern analysiert. Hierzu wurde zunächst die Gesamt-DNA der Proben aufgereinigt. Da es sich bei den ersten untersuchten Umweltproben um Bodenproben handelte, musste bei der Aufreinigung der DNA darauf geachtet werden, eventuell vorhandene Huminsäure zu entfernen, da sich diese hemmend auf eine PCR auswirkt.²⁰⁷ Das verwendete PowerSoil® DNA Isolation Kit der Firma MoBio wurde speziell für die Isolierung von DNA aus Böden entwickelt (Kapitel 7.6.1) und entfernt eventuell vorhandene Huminsäure während des Aufreinigungsprozesses. Zur Überprüfung der Methodik wurden zwei unterschiedliche, spezifische Primerpaare (basierend auf der 16S rRNA von 'Entotheonella factor', erstellt von Dr. Agustinus R. Uria; Kapitel 7.4.1.1) verwendet. In der Positivkontrolle wurde als Templat 1 µl des 'E. factor'-Zellpellets verwendet. Zur Kontrolle der Methodik wurde eine Bodenprobe mit 'E. factor' versetzt (Abbildung 5-6).

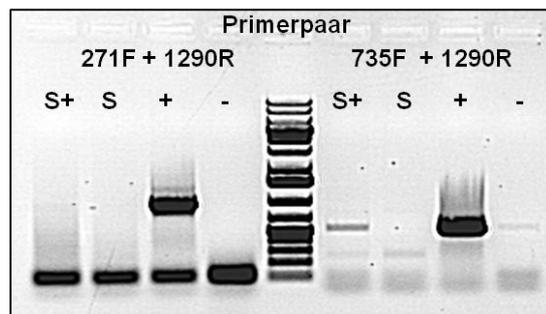


Abbildung 5-6: Agarosegel zur Überprüfung der Nachweismethodik. S kennzeichnet die Bodenprobe, S+ entspricht der Bodenprobe, der 'E. factor' zugegeben wurde, + kennzeichnet die Positivkontrolle ('E. factor'), - kennzeichnet die Negativkontrolle (Wasser).

Wie in (Abbildung 5-6) zu sehen, war diese PCR-Methode zumindest für das Primerpaar 735F und 1290R erfolgreich. Die schwach zu erkennende Bande in der Negativkontrolle

Ergebnisse und Diskussion

resultiert vermutlich aus einer Kontamination beim Beladen des Gels mit der Positivkontrolle. Bei Primerpaar 271F+1290R ist in der Kontrolle (S+) eine Bande nur zu erahnen. Anschließend wurden in einem ersten Screening einige Umweltproben analysiert (Tabelle 5-1, Abbildung 5-17)

Tabelle 5-1: Untersuchte Umweltproben.

Probe	alternativer Probenname	Datum	Art	Ort	Breitengrad	Längengrad	Bemerkung/Beschreibung
1		03.06.2012	Wasser	DE, Hamburg, Hafen	53.602342	9.992991	Elbe-Brackwasser
2		03.06.2012	Wasser	DE, Kiel, Hafen	54.311206	10.132981	
3		03.06.2012	Wasser	DE, Laboe	54.408404	10.221222	Ostsee-Strand
4		03.06.2012	Wasser	DE, Berlin	unbekannt	unbekannt	
5		26.05.2012	Boden	DE, Bausendorf	50.016165	6.993830	Garten
5.1		23.07.2012	Boden	DE, Bausendorf	50.016166	6.993831	Garten
5.2		23.07.2012	Boden	DE, Bausendorf	50.016167	6.993832	Garten
6		03.06.2012	Sediment	DE, Kalenborn	50.557967	7.000637	Bach Swist, Ufer
7		03.06.2012	Sediment	DE, Laboe, Ostsee	54.408404	10.221222	Strand
8		05.06.2012	Boden	DE, Berlin	unbekannt	unbekannt	
9		03.07.2012	Boden	NO, Lianvatnet, Trondheim	63.401791	10.315083	Ufer
10		03.07.2012	Wasser+ Sediment	NO, Lianvatnet, Trondheim	63.401750	10.315035	See
11		03.07.2012	Wasser	NO, Lianvatnet	63.402760	10.318428	Bach
12		03.07.2012	Boden	NO, Stor Leirsjøen	63.383249	10.284710	Ufer
13		03.07.2012	Wasser	NO, Stor Leirsjøen	63.383281	10.284548	See
14		06.07.2012	Boden	NO, Parkgata 14, Ålesund,	62.473128	6.162424	Park
15		08.07.2012	Boden	NO, Skomakerdiket, Bergen	63.383165	10.284703	Ufer
16		08.07.2012	Wasser	NO, Skomakerdiket, Bergen	63.383238	10.284634	See
17		09.07.2012	Wasser+ Sediment	NO, Blåmansvatnet, Bergen	60.399644	5.358201	See
18		11.07.2012	Boden	NO, Geilo	60.529827	8.192067	Fluss/Fjord/Ufer
19		11.07.2012	Wasser	NO, Geilo	60.529827	8.192068	Fluss/Fjord
20		12.07.2012	Wasser	NO, Nationalpark Geilo	60.458690	7.749951	Gletscherwasser
21		12.07.2012	Boden	NO, Nationalpark Geilo, Rallarvegen Buskernd	60.458690	7.749951	
22		13.07.2012	Boden	NO, Geilo	60.546011	8.241340	
23		14.07.2012	Boden	NO, Drammen			
24		14.07.2012	Boden	NO, Flughafen SandeFjord	59.177366	10.251133	Nadelwaldboden
25		24.07.2012	Boden	DE, Meckenheim	50.629311	7.048771	Garten
26		24.07.2012	Boden	DE, Meckenheim	50.629152	7.048804	Vorgarten
27	RD	26.07.2012	Boden	HR, Strand M. Duboka	43.183932	16.588840	
28		27.07.2012	Wasser	HR, Strand M. Duboka	43.188469	16.584163	
29	A	08.09.2012	Boden	DE, Neuwied (Industriegebiet)	50.438631	7.486529	Gras, Sträucher
30	B	08.09.2012	Boden	DE, Neuwied (Industriegebiet)	50.43848	7.486581	schwarze Erde mit Baumrinde
31	C	08.09.2012	Boden	DE, Neuwied (Zentrum)	50.427134	7.473694	Hecke
32	D	08.09.2012	Boden	DE, Neuwied (Zentrum)	50.427196	7.473873	Baum
33	E	08.09.2012	Boden	DE, Neuwied, Friedhof	50.424134	7.475721	
34	F	08.09.2012	Boden	DE, Neuwied, Friedhof	50.424405	7.475612	Kastanienbaum
35	G	08.09.2012	Boden	DE, Neuwied	50.42181	7.46568	Zimmerpflanze (Mango)
36	H	09.09.2012	Boden	DE, Neuwied (Deich am Rhein)	50.425845	7.457446	
37	I	09.09.2012	Boden	DE, Neuwied (Deich am Rhein)	50.426084	7.455733	Sand
38	J	09.09.2012	Boden	DE, Neuwied	50.425897	7.461494	Innenhof
39	K	09.09.2012	Boden	DE, Neuwied	50.42583	7.46404	Innenhof, Hecke, Nadelgehölz
40	L	09.09.2012	Boden	DE, Neuwied	50.425824	7.461559	Innenhof, Gras, sandiger Boden

Ergebnisse und Diskussion

41	11.09.2013	Boden	DE, Bonn, Kekulé Institut	50.733843	7.072509	
42	11.09.2013	Boden	DE, Bonn, Kekulé Institut	50.733983	7.071806	
43	19.09.2013	Boden	CH, Zürich, ETH Hönggerberg	47.407121	8.508973	
44	03.07.2015	Boden	DE, Bonn, UKB, IMMIP	50.698618	7.106535	
45	05.06..2015	Boden	ZA, Pretoria	-25.732617	28.22281	dunkler Sand
46	30.08.2013	Boden	DE, Oberstdorf (OT Reichenbach)	47.449034	10.293104	Weg
47	31.08.2013	Boden	DE, Oberstdorf (OT Reichenbach)	47.448989	10.293135	Wegrand

Zusätzlich zu Bodenproben wurden dabei auch Wasserproben untersucht. Dazu wurden die Wasserproben bei 14000 x g zentrifugiert, um alle festen Bestandteile, wie auch alle Bakterien, vom Wasser abzutrennen. Die DNA wurden dann aus dem Pellet, wie zuvor bei den Bodenproben, mittels PowerSoil® DNA Isolation Kit aus dem Pellet aufgereinigt.

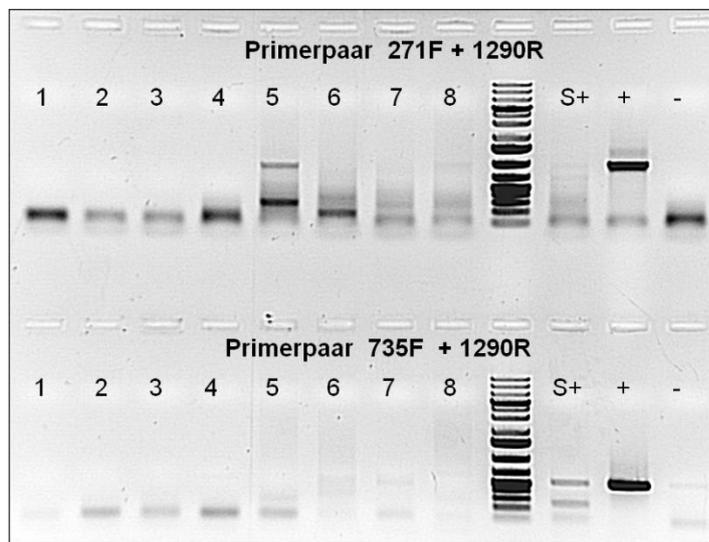


Abbildung 5-7: Screenen erster Umweltproben. 1-8 bezeichnen die unterschiedlichen Umweltproben, M kennzeichnet den Marker Fermentas 1 kb plus Ladder.

Bei diesem Screening konnte eine erste positive Bodenprobe detektiert werden (Abbildung 5-7, Probe 5). Bemerkenswert ist, dass diese positive Bande mit dem Primerpaar amplifiziert wurde, mit dem sowohl in der ersten Test-PCR als auch in dieser PCR keine Bande in der Kontrollprobe (S+) amplifiziert werden konnte. Zur systematischen Verifizierung und Einordnung des so amplifizierten 16S rRNA-Gens bzw. der Spezies wurde das PCR-Produkt isoliert und sequenziert (Kapitel 7.4.2.1; Kapitel 7.4.6). Die Sequenzanalyse ergab eine Übereinstimmung zur DNA des 16S rRNA-Gens von 'Entotheonella factor' von 97,4 % (Abbildung 5-8). Somit ist davon auszugehen, dass es sich bei der detektierten Spezies um ein Bakterium desselben Genus handelt.²⁰⁸

Ergebnisse und Diskussion

Abbildung 5-8: Alignment der 16S rRNA-Gensequenz aus Probe 5 und der Sequenz von 'E. factor'. Dargestellt ist ein Sequenzalignment der aus Probe 5 mithilfe spezifischer Primer amplifizierten DNA des 16S rRNA-Gens und der 16S rRNA-Gensequenz von 'E. factor' (Referenzsequenz). In der Probensequenz sind die Basen farblich hervorgehoben, in denen sich die Sequenz von der Referenzsequenz ('E. factor') unterscheidet. Verwendete Software: Geneious 7.1.

Wie sich in weiteren Probenscreenings herausstellte, konnten die PCR-Ergebnisse nicht konstant reproduziert werden: Zeigte eine Probe in einer ersten PCR eine positive Bande, war diese bei Wiederholung der PCR nicht mehr zu erkennen. Ein Grund hierfür könnte die geringe Konzentration des spezifischen Templates im Verhältnis zur restlichen, aus der Probe isolierten metagenomischen DNA sein. Daher wurde versucht, die Sensitivität der Nachweismethode weiter zu erhöhen. Dies wurde von der Bachelorstudentin Nilani Vickneswaran im Rahmen ihrer Bachelorarbeit durchgeführt.²⁰⁹ In Anlehnung an die Nachweismethode von '*Candidatus Entotheonella*' in Schwämmen (Kapitel 3.3.1), wurde eine *nested*-PCR-Strategie entwickelt, die das Problem der Reproduzierbarkeit und Sensitivität des PCR-Verfahrens behob. Bei diesem Verfahren wird in einem ersten Schritt alle der in der Probe enthaltenen 16S rRNA-Gene mittels universellen 16S rRNA-Genprimern amplifiziert (U-27-rRNA-F und U-1442-rRNA-R²¹⁰). Diese „Gesamt-DNA“ der 16S rRNA-Gene dient dann in einer zweiten PCR mit spezifischen 'E. factor'-Primern als Template (Abbildung 5-9).

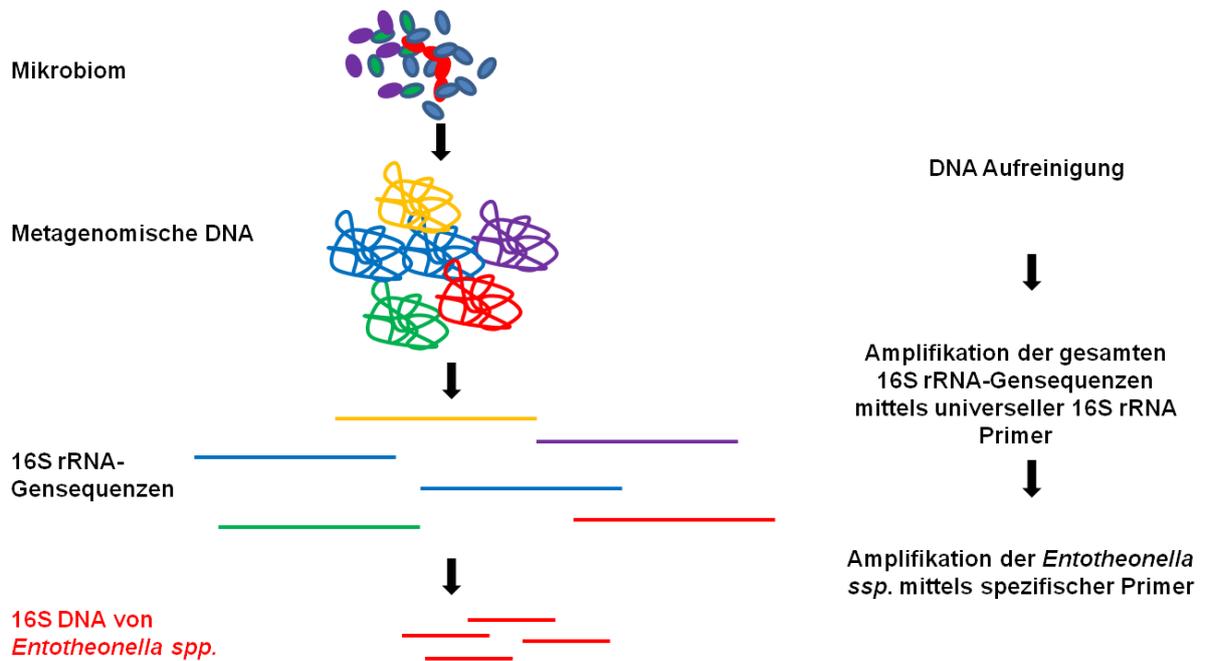


Abbildung 5-9 Schematische Darstellung der *nested*-PCR Strategie. Nach Aufreinigung der DNA aus den Bodenproben wird zunächst mittels universeller Primer die „Gesamt-DNA“ der 16S rRNA-Gene der Probe amplifiziert. In einer zweiten PCR wird diese nun als Templat verwendet, um mithilfe spezifischer 'E. factor'-Primer die spezifische 16S rRNA amplifizieren zu können. In rot dargestellt ein Bakterium der Gattung 'Candidatus Entotheonella' bzw. dessen DNA. Mit blau, grün, gelb und violett sind andere Spezies und deren DNA gekennzeichnet.

Zusätzlich testete sie weitere 'E. factor'-spezifische Primer^{137,172,199,210}, um die beste Primerkombination zu bestimmen.²⁰⁹ Dabei zeigte sich, dass mit der Kombination der Primer 238F bzw. 271F mit 1290R in der oben beschriebenen *nested*-PCR die besten Ergebnisse erzielt werden konnten.²⁰⁹ Daher wurden alle weiteren Proben mit diesen Primern und nach dieser Methode untersucht. Alle Proben sind in Tabelle 5-1 aufgelistet.

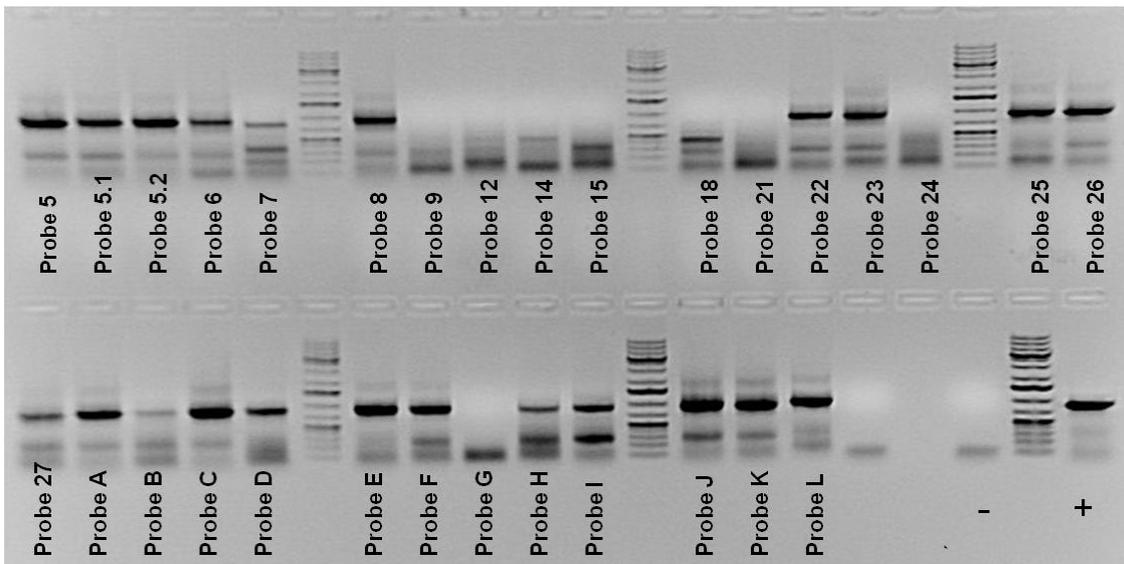


Abbildung 5-10: *Nested*-PCR zur Identifizierung von 'Candidatus Entotheonella' in Bodenproben.²⁰⁹

Wie in Abbildung 5-10 beispielhaft zu sehen, konnte in vielen Bodenproben mithilfe 'Candidatus Entotheonella' -spezifischer Primer ein spezifisches PCR-Produkt erzeugt werden.

Neben der Untersuchung von Bodenproben wurde auch versucht, 'Candidatus Entotheonella' in Wasserproben nachzuweisen. Die Wasserproben wurden dazu über eine Polyethersulfon-Membran (Sartolab BT Einmal-Vakuumfiltrationsgerät, 150 ml, Firma Sartorius) mit einer Porengröße von 0,22 µm gefiltert, um so alle Mikroorganismen und Partikel >0,22 µm auf dem Filter zurück zu halten. Die filtrierte Wassermenge war dabei von der jeweiligen Probe abhängig, die unterschiedliche viele Schwebeteilchen enthielten und den Filter zusetzten. Dieser wurde dann in kleine Stücke zerteilt und in ein mit Glaskügelchen gefülltes Reaktionsgefäß des „Power soil DNA Extraction Kits“ der Firma MoBio gegeben. Anschließend wurde die DNA der darauf befindlichen Mikroorganismen laut Protokoll aufgereinigt. Wie auch bei den Bodenproben wurde dann eine *nested*-PCR zur Identifizierung von 'Candidatus Entotheonella' -Spezies durchgeführt (Kapitel 7.4.1).

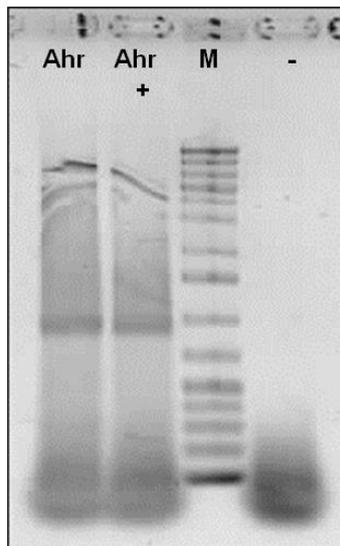


Abbildung 5-11: Untersuchung von Wasserproben auf 'Candidatus Entotheonella'.Untersucht wurde eine Wasserprobe der Ahr. Als Positivkontrolle (**Ahr+**) wurde ein Teil der Wasserprobe mit 'E. factor' versetzt und parallel zur unbehandelten Probe (**Ahr**) aufgereinigt. **M**: Fermentas 1 kb plus Ladder; **-**: Negativkontrolle.

In einer Wasserprobe der Ahr konnte mithilfe der Primer 238F und 1290R eine - allerdings sehr unscharfe - Bande amplifiziert werden (Abbildung 5-11). Nach Aufreinigung der DNA konnte in einer anschließenden Sequenzierung nur eine sehr schlechte Sequenzqualität von wenigen Basenpaaren erzielt werden. Vergleiche einzelner Sequenzabschnitte deuteten jedoch darauf hin, dass es sich hierbei auch um die 16 S rDNA eines Bakteriums der Gattung 'Candidatus Entotheonella' handeln könnte. Die Sequenzqualität konnte auch bei wiederholter Amplifikation und Aufreinigung nicht verbessert werden. Weitere Versuche mit unterschiedlichen Wasserproben konnten keine positiven Ergebnisse erzielen. Das Problem liegt hier vermutlich in der geringen Menge, mit der diese Bakterien im Verhältnis zum

gesamten Mikrobiom auftreten. Zur Gewinnung von ausreichend Biomasse, gerade bei Fließgewässern, müssten sehr große Volumina filtriert werden. Daher wurde der Fokus auf die Untersuchung von Bodenproben gelegt. Versuche mit Pflanzenmaterialien, wie beispielsweise Blättern, erzielten keine positiven Ergebnisse. Zusammenfassend konnte so in 22 (bzw. 23, zählt man das nicht eindeutige Ergebnis der Wasserprobe der Ahr hinzu) der 49 Umweltproben bzw. 22 der 35 Bodenproben (Tabelle 5-1) mittels spezifischer 'E. factor'-Primer PCR Produkte amplifiziert werden (Abbildung 5-10). Das entspricht mehr als jeder zweiten untersuchten Bodenprobe (62,9% der Bodenproben).

Die 16S rRNA-Genamplifikate der Bodenproben wurden dann, wie oben bereits beschrieben, isoliert und sequenziert (Kapitel 7.4.1.3, Kapitel 7.4.2.1, Kapitel 7.4.6.1). Aufgrund der potentiellen Diversität der Amplifikate, dass es sich also nicht um eine Einzelsequenz sondern eine Mischung von 16S rRNA-Sequenzen handeln könnte (respektive mehrere unterschiedliche Spezies bzw. Phylotypen in der Probe enthalten sind), wurden die PCR-Produkte in den Klonierungsvektor pGEMTeasy (Promega) kloniert (Kapitel 7.4.3.3) und nach Möglichkeit jeweils mindestens drei Klone sequenziert. Die so erhaltenen Sequenzen wurden mithilfe der Software Geneious (Version 7.1.9) mit der Referenzsequenz, der 16S rRNA-Gensequenz von 'Entotheonella factor', verglichen. Aus den Übereinstimmungen der einzelnen Nukleotide ergibt sich dann die sogenannte *pairwise identity* (PI). Die Ergebnisse sind in Tabelle 5-2 aufgeführt.

Tabelle 5-2: 'Candidatus Entotheonella' 16S rRNA-Gensequenzen verschiedener Bodenproben.

Bodenprobe	(alias)	Primerpaar	Klon	„pairwise identity“ zu 'Entotheonella factor' (%)
5		238F+1290R	5-clone1	88,6
		238F+1290R	5-clone4	88,3
		271F+1290R	MG1	97,4
6		238F+1290R	6-clone11	88,6
		238F+1290R	6-clone15	88,5
		238F+1290R	6-clone18	85,5
8		238F+1290R	8-clone4	88,1
		238F+1290R	8-clone7	88,7
		238F+1290R	8-clone23	89,1
		238F+1290R	8-clone27	88,5
		238F+1290R	8-clone30	88,6
		271F+1290R	clone801	96,3
		271F+1290R	clone808	96,1

Ergebnisse und Diskussion

		271F+1290R	clone811	96,2
		271F+1290R	clone814	95,9
		271F+1290R	clone822	96,5
		271F+1290R	clone825	95,6
22		238F+1290R	22-clone1	88,8
		238F+1290R	22-clone2	83,6
		238F+1290R	22-clone5	88,5
		238F+1290R	22-clone7	88,6
		238F+1290R	22-clone10	78,2
23		238F+1290R	23-clone1	88,1
		238F+1290R	23-clone2	91
		238F+1290R	23-clone3	88,5
		238F+1290R	23-clone4	79,2
25		238F+1290R	25-clone1	89
		238F+1290R	25-clone8	88,5
		238F+1290R	25-clone6	89,1
26		238F+1290R	26-clone1	88,8
		238F+1290R	26-clone2	97
		238F+1290R	26-clone5	89,1
29	A	238F+1290R	A-clone2	89
		238F+1290R	A-clone11	88,7
		238F+1290R	A-clone 15	88,9
30	B	238F+1290R	B-clone7	87,7
		238F+1290R	B-clone8	87,5
		238F+1290R	B-clone9	87,8
32	C	238F+1290R	C-clone6	88,8
		238F+1290R	C-clone7	88,6
		238F+1290R	C-clone8	88,9
32	D	238F+1290R	D-clone4	88,6
		238F+1290R	D-clone6	88,5
		238F+1290R	D-clone7	87,9
33	E	238F+1290R	E-clone1	88,3
		238F+1290R	E-clone20	89,2
		238F+1290R	E-clone23	88,9
34	F	238F+1290R	F-clone1	88,5
		238F+1290R	F-clone4	88,3

Ergebnisse und Diskussion

		238F+1290R	F-clone9	89
36	H	238F+1290R	H-clone1	89
37	I	238F+1290R	I-clone1	99,1
		238F+1290R	I-clone4	99,3
		238F+1290R	I-clone7	99,1
38	J	238F+1290R	J-clone1	90,7
		238F+1290R	J-clone3	93,1
		238F+1290R	J-clone7	88,6
39	K	238F+1290R	K-clone1	90,4
		238F+1290R	K-clone3	96
		238F+1290R	K-clone5	89,2
		238F+1290R	K-clone6	89,6
40	L	238F+1290R	L-clone2	88,7
		238F+1290R	L-clone4	89,9
		238F+1290R	L-clone7	88,7
		238F+1290R	L-clone10	89,2
		238F+1290R	L-clone11	89,3
44		238F+1290R	44-clone21	95,9
		238F+1290R	44-clone22	88,8
		238F+1290R	44-clone23	96,2
45		238F+1290R	45-clone 22	89
		238F+1290R	45-clone 23	88,7
		238F+1290R	45-clone 24	89
46		238F+1290R	46-clone 21	89,3
		238F+1290R	46-clone 22	89,1
		238F+1290R	46-clone 23	97,3
47		238F+1290R	47-clone 21	88,9
		238F+1290R	47-clone 22	88,2
		238F+1290R	47-clone 23	90,2

Wie aus den *pairwise identity*-Werten der Tabelle 5-2 hervorgeht, sind in vier der 22 positiven Bodenproben 'Entotheonella factor'-Phylotypen zu finden (*pairwise identity* \geq 97%) (Tabelle 5-1, Proben 5, 26, I, 46). Besonders bemerkenswert sind die Sequenzen aus Probe I, die mit knapp über 99% *pairwise identity* nahezu identisch zur 16S rRNA-Gensequenz von 'E. factor' sind. Definiert man die untere Grenze eines Genus bei 95%, so konnten in drei weiteren Proben 16S rRNA-Gensequenzen aus '*Candidatus Entotheonella*' -Spezies

identifiziert werden (Proben 8, K, 44). Alle weiteren 16S rRNA-Gensequenzen gehören zu Spezies derselben Familie (*pairwise identity* 90-95%), derselben Ordnung (*pairwise identity* 85-90%), derselben Klasse (*pairwise identity* 80-85%), bzw. desselben Phylums (*pairwise identity* >77%) wie 'E. factor'. Bemerkenswert ist, dass mit den 'E. factor'-spezifischen Primern nicht nur die 16S rRNA-Gensequenzen von Spezies desselben Genus sondern auch die weiter entfernt verwandte Spezies desselben Phylums detektiert werden konnten. Abbildung 5-12 zeigt beispielhaft das Alignment der aus Probe 8 gewonnenen Sequenzen. Hier wird deutlich, dass in der Probe nicht nur eine 'Candidatus Entotheonella'-Spezies enthalten ist, sondern unterschiedlich nah verwandte Mitglieder desselben Genus, derselben Familie bzw. Ordnung. Des verdeutlicht das Alignment, dass sich die Sequenzen anhand einer Insertion in zwei Gruppen einteilen lassen, worauf im folgenden Abschnitt näher eingegangen wird.



Abbildung 5-12: Ausschnitt aus einem 16S rRNA-Gensequenz Alignment zur Erstellung eines phylogenetischen Stammbaums am Beispiel der Sequenzen aus Bodenprobe 8. Die Sequenzen lassen sich anhand einer Insertion in zwei Gruppen aufteilen (roter Kreis)

5.1.3 Phylogenetische Analysen

Wie in Kapitel 3.3.1 beschrieben, zeigten erste phylogenetische Untersuchungen, dass 'Entotheonella factor' einem neuen, bisher unbeschriebenen Phylum zugehörig zu sein scheint. Dieses wiederum spaltet sich nach einer ersten phylogenetischen Analyse in drei Kladen auf (Abbildung 5-13, Dr. Susanne Schmitt, Ludwig-Maximilians-Universität München). Das neue Phylum wurde aufgrund der Eigenschaften seiner bisher bekannten Mitglieder (zum einen, dass diese bislang nicht kultiviert werden konnten und zum anderen ihrer Fähigkeit zur Produktion von Naturstoffen, die vermutlich der Abwehr dienen) 'Tectomicrobia' (*lat. tegere*: verbergen; schützen) genannt.¹³¹

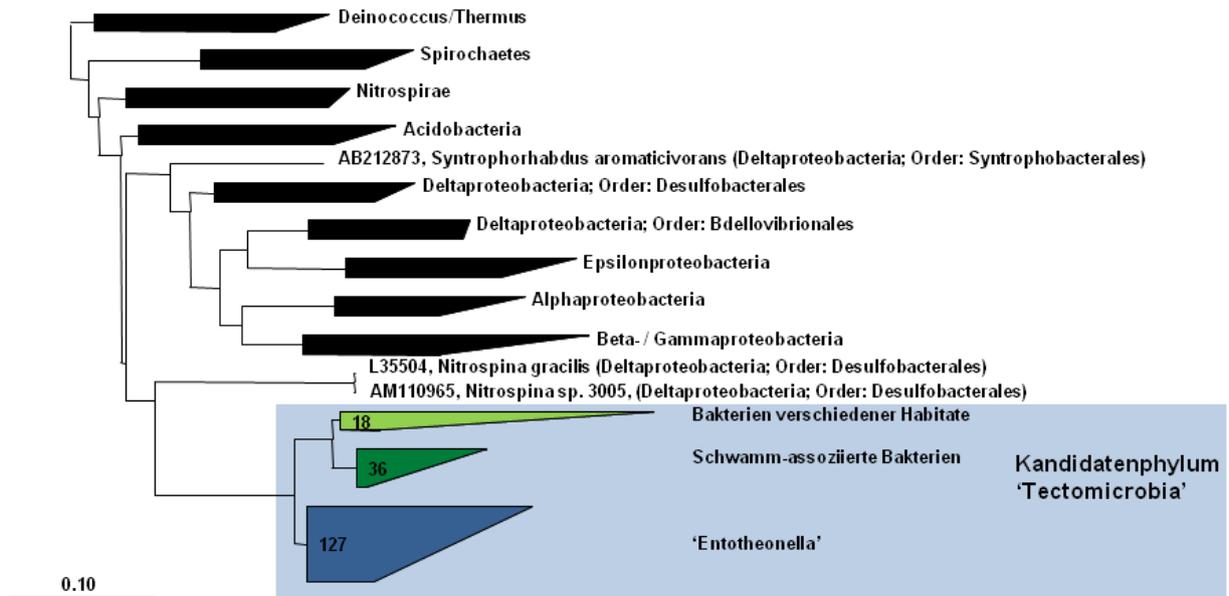


Abbildung 5-13: Erste phylogenetische Einordnung von 'E. factor' (Dr. Susanne Schmitt).

Um die aus den Bodenproben amplifizierten 16S rRNA-Gensequenzen in diese Kladen einordnen zu können, wurde zunächst ein Alignment erstellt (Software: Geneious 7.1.9). Dazu wurden die schon vorhandenen Sequenzen von Dr. Susanne Schmitt und die Sequenzen aus den Bodenproben mit der DNA des 16S rRNA-Gens von 'E. factor' verglichen. Auffällig bei diesem Alignment war, dass sich die Sequenzen anhand einer 29-31 bp großen Insertion in zwei Gruppen aufteilen ließen (Abbildung 5-12, roter Kreis). Sequenzen ohne Insertion weisen einen hohen Verwandtschaftsgrad zu E. factor (selbes Genus) auf, dessen 16S rRNA-Gensequenz ebenfalls keine Insertion aufweist. Anhand des Alignments wurde dann ein Stammbaum erstellt (Abbildung 5-14, Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, ohne *bootstrapping*).

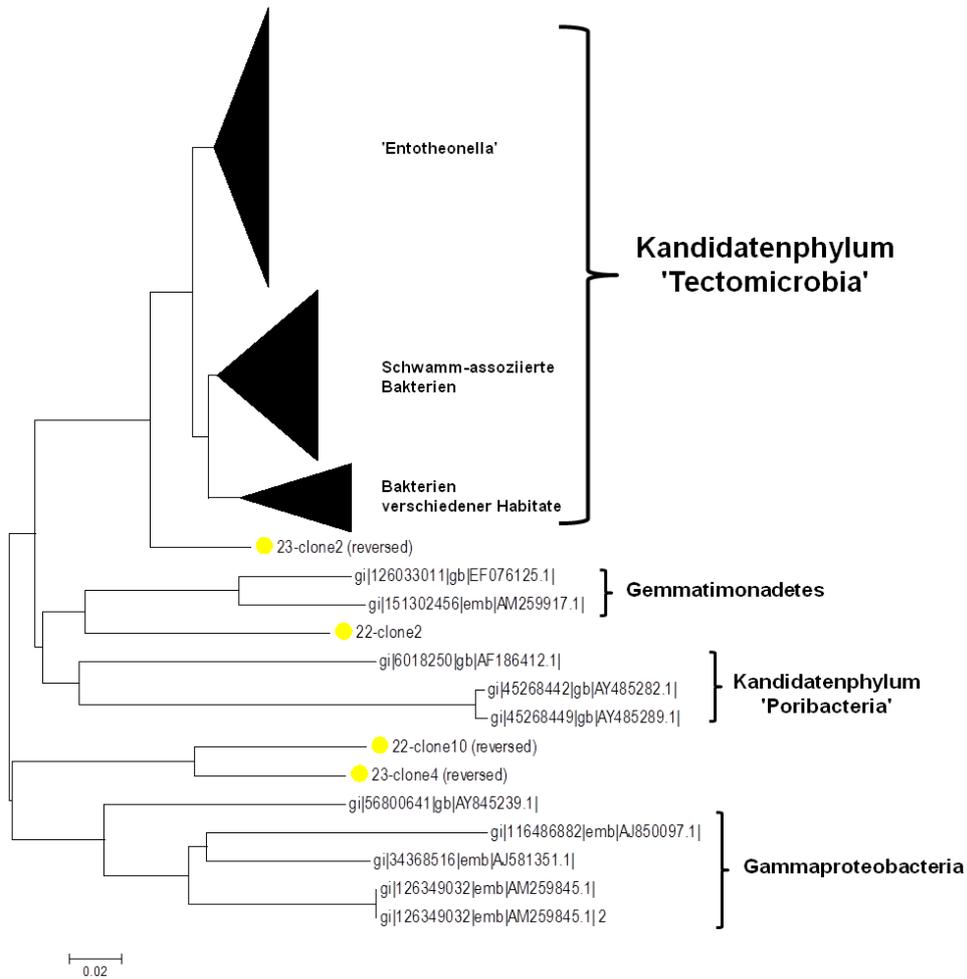


Abbildung 5-14: Erste phylogenetische Einordnung der 'Tectomicrobia'-Sequenzen aus Umweltproben. Phylogenetischer Stammbaum (Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood ^{190,211,212} ohne *bootstrapping*), der die 16S rRNA-Gensequenzen der Bodenproben in den Stammbaum von Frau Dr. Susanne Schmitt einordnet.

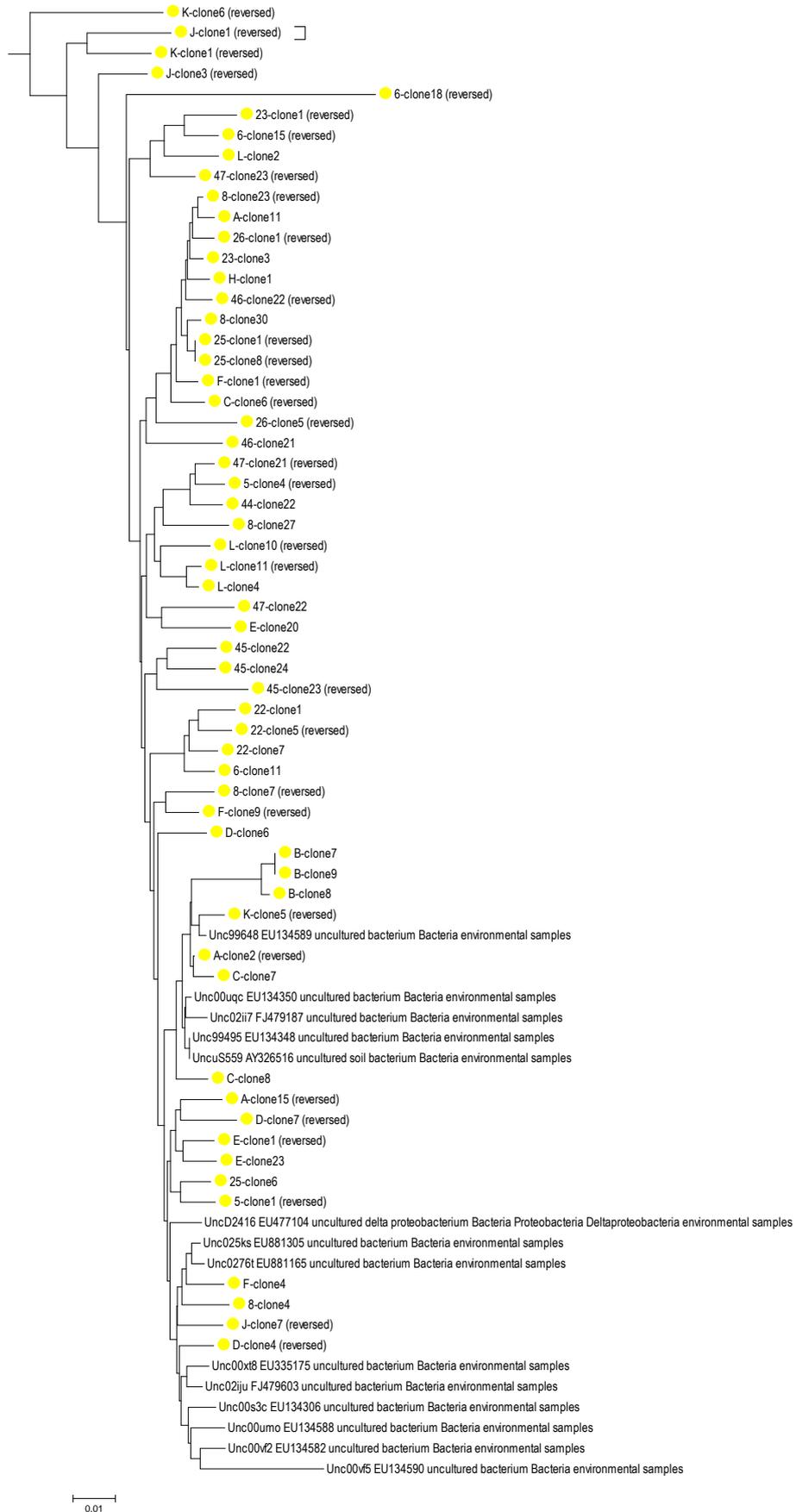


Abbildung 5-15: Detailansicht der Klade „Schwamm-assoziierte Bakterien“. Mit einem gelben Punkt markiert sind die im Rahmen dieser Arbeit gewonnen Sequenzen.

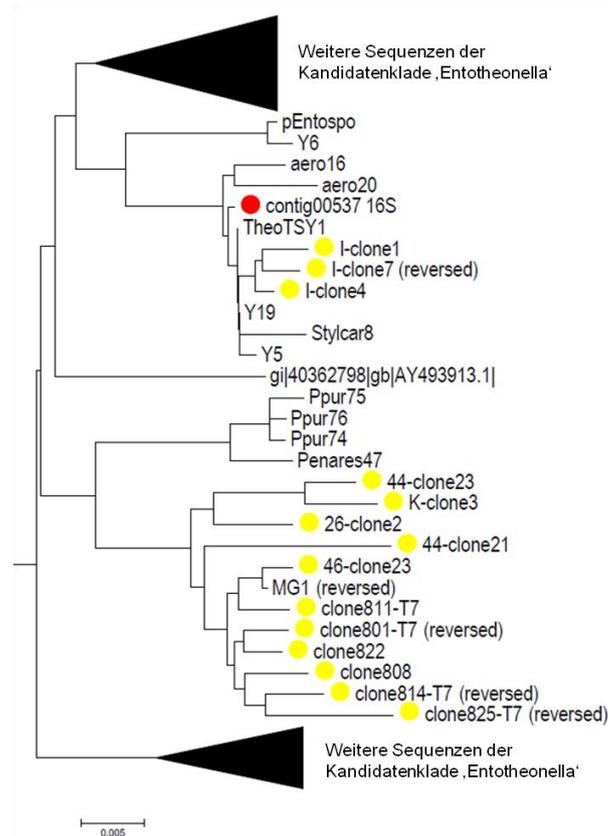


Abbildung 5-16: Detailansicht der Klade 'Entotheonella'. Mit einem gelben Punkt markiert sind die im Rahmen dieser Arbeit gewonnenen Sequenzen. In die 'Entotheonella'-Klade fällt auch 'E. factor' (rot markiert).

Wie aus der Analyse des Stammbaums ersichtlich wird, sind in den untersuchten Bodenproben 16S rRNA-Gensequenzen zweier unterschiedlicher Kladen („Schwamm-assoziierte Bakterien“ und „Bakterien verschiedener Habitate“) zu finden (Abbildung 5-14, Abbildung 5-15, Abbildung 5-16). Die oben beschriebene Unterscheidung zweier Gruppen anhand einer Insertion (Abbildung 5-12) spiegelt sich im Stammbaum wider: So fallen die Sequenzen mit Insertion in die eine („Schwamm-assoziierte Bakterien“), Sequenzen ohne Insertion (zusammen mit 'E. factor') in die andere Klade ('Ca. Entotheonella'). Versuche Kladen-spezifische Primer zu erstellen, scheiterten an der Diversität der Insertionssequenz.

Um einen Eindruck über die Vertreter dieses neuen Kandidatenphylums gewinnen zu können, wurde die 16S rRNA-Gensequenz von 'Entotheonella factor' mit der NCBI-Nukleotid-Datenbank (Bakterien) verglichen. Dabei fällt zunächst auf, dass die 1000 nächsten verwandten 16S rRNA-Gensequenzen (pairwise identity >77% zu 'Entotheonella factor') in ganz unterschiedlichen, z. T. extremen Habitaten, und in nahezu allen Teilen der Welt zu finden sind (Abbildung 5-17) Unter anderem sind Sequenzen in Salzwasserseen z. B. in Mexiko und Kalifornien (NCBI Accession number FJ152750, FJ152768, FJ152806, FJ152847), in Erdproben aus Uranabfällen (NCBI Accession number AJ519395, AJ519397, AJ519400, AJ519401, AJ532724, AJ532726, AJ532728), im Grundwasser von Goldminen in

Südafrika (NCBI Accession number DQ088768, DQ088769), auf Reis- aber auch auf Ölfeldern in China (NCBI Accession number EU589249, HQ190412) und in der Antarktis (NCBI Accession number GU235246, GU235266, GU235310) sowie in der ozeanischen Todeszone (u. a. NCBI Accession number GQ348755, GQ350305) isoliert worden, um nur einige Beispiele neben den hier bereits benannten Schwämmen und Böden zu nennen.

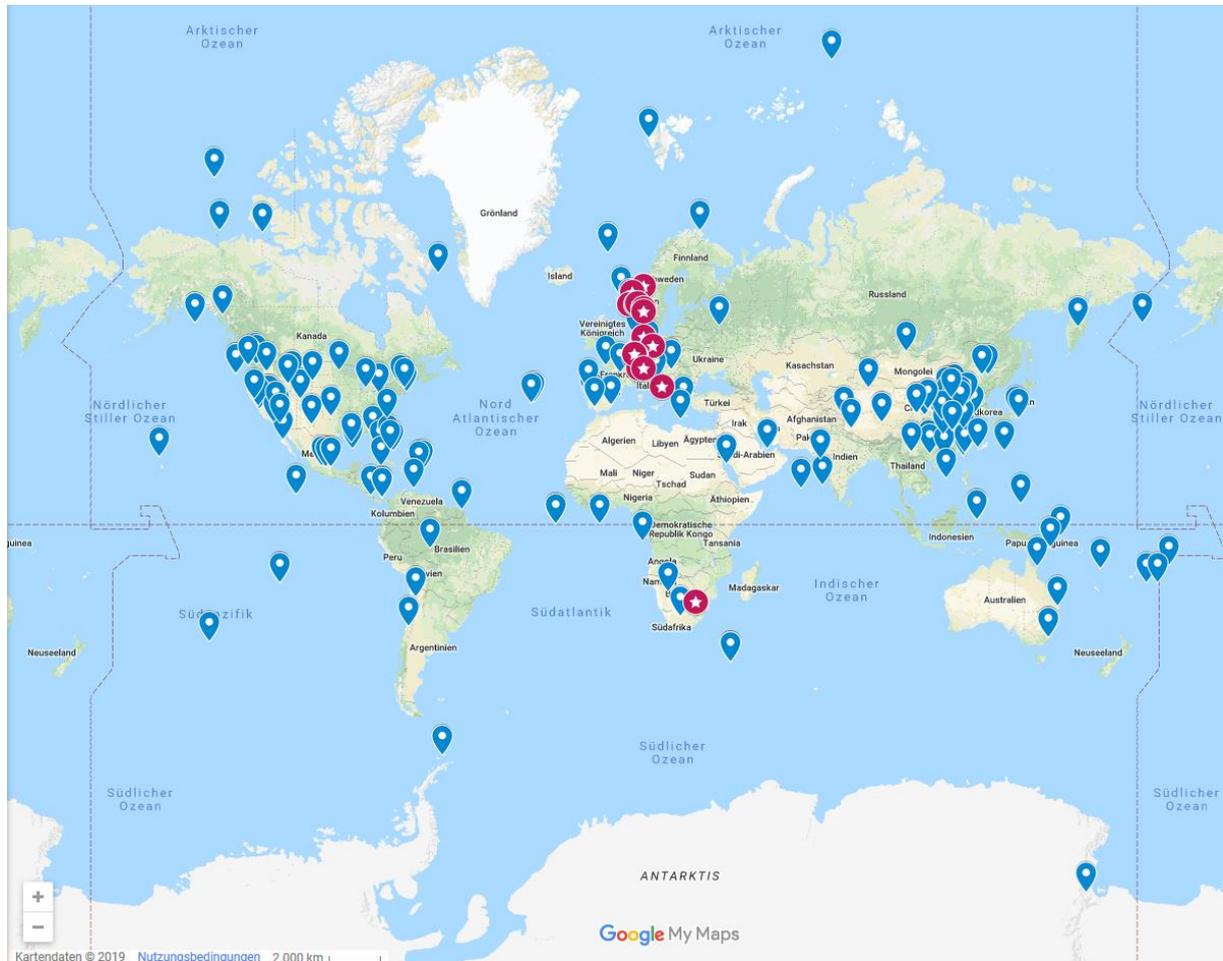


Abbildung 5-17: Verbreitung von 'Tectomicrobia' anhand einer Datenbankanalyse.²¹³

Dargestellt sind Vorkommen von 'Tectomicrobia' anhand einer Datenbankanalyse (**blaue Stecknadeln**). Diese wurden in von *Google Maps* bereitgestelltem Kartenmaterial (© 2019) eingepflegt. Eingetragen konnten nur die Spezies, deren Vorkommen im Rahmen der Datenbankanalyse ermittelt werden konnten. Die Qualität der Angaben war sehr inhomogen (Angaben schwankten von der Angabe des Landes bis hin zu exakten geographischen Koordinaten). Mit **pinken Sternen** markiert sind im Rahmen dieser Arbeit detektierte Vorkommen in Bodenproben.

Neuere Arbeiten von Lackner et al. zeigen, dass der Metabolismus von 'E. factor' auch für solche z. T. sehr lebensfeindliche Bedingungen ausgelegt ist.²¹⁴ Genomische, (meta)proteomische und chemische Analysen der beiden sequenzierten '*Candidatus Entotheonella*' Varianten des Schwammes *Theonella swinhoei* ('*Candidatus Entotheonella factor*' und '*Candidatus Entotheonella gemina*') zeigen das Stoffwechselmodell eines fakultativ anaeroben, organotrophen Bakteriums, das auch in der Lage ist, Methanol als Energiequelle zu nutzen. Es können sowohl organische Säuren, Alkohole, Polyalkohole und Polysaccharide als auch komplexe aromatische Verbindungen als Kohlenstoffquelle genutzt

werden. Zwar scheinen diese Bakterien für einige Vitamine auxotroph zu sein, die meisten Aminosäuren sowie einige Coenzyme wie z. B. das Coenzym F₄₂₀, das für die charakteristische Autofluoreszenz verantwortlich ist, können aber selbst synthetisiert werden. Auffällig ist außerdem das große Repertoire an Transporter und Regulatoren, die ebenfalls ein Indiz dafür sein können, dass diese Bakterien nicht auf eine bestimmte ökologische Nische beschränkt sind.²¹⁴ Diese untersuchten Vertreter des Kandidatenphylums Tectomicobia können sich also aufgrund ihres genetischen Repertoires vermutlich sehr gut an verschiedenste Umweltbedingungen anpassen. Das könnte eine Erklärung dafür sein, dass andere Vertreter dieses Kandidatenphylums in verschiedensten, z. T. extremen Habitaten zu finden sind. Es ist zu vermuten, dass diese ähnliche genetische bzw. metabolische Eigenschaften besitzen.

In phylogenetischen Analysen wurden einige 16S rRNA-Gensequenzen der Datenbankrecherche mit den isolierten Sequenzen aus den Bodenproben verglichen (Abbildung 5-18 und Abbildung 5-19). Von den rund 1000 Sequenzen der Datenbankrecherche wurden dazu 562 anhand ihrer Sequenzlänge ausgewählt, um ein Sequenzalignment maximaler Länge und ohne jegliche Sequenzlücken zu ermöglichen. Als Außengruppe wurden die 16S rRNA-Gensequenzen zweier *Nitrospinae*-Spezies verwendet. Das Phylum *Nitrospinae* ist das bislang nächst verwandte, bekannte Phylum zu 'Tectomicobia'. Das Sequenzalignment und die anschließende phylogenetische Analyse der 562 Sequenzen der Datenbank und der 78 neuen, in dieser Arbeit isolierten Sequenzen der Bodenproben wurden mithilfe der Software MEGA durchgeführt.²¹²

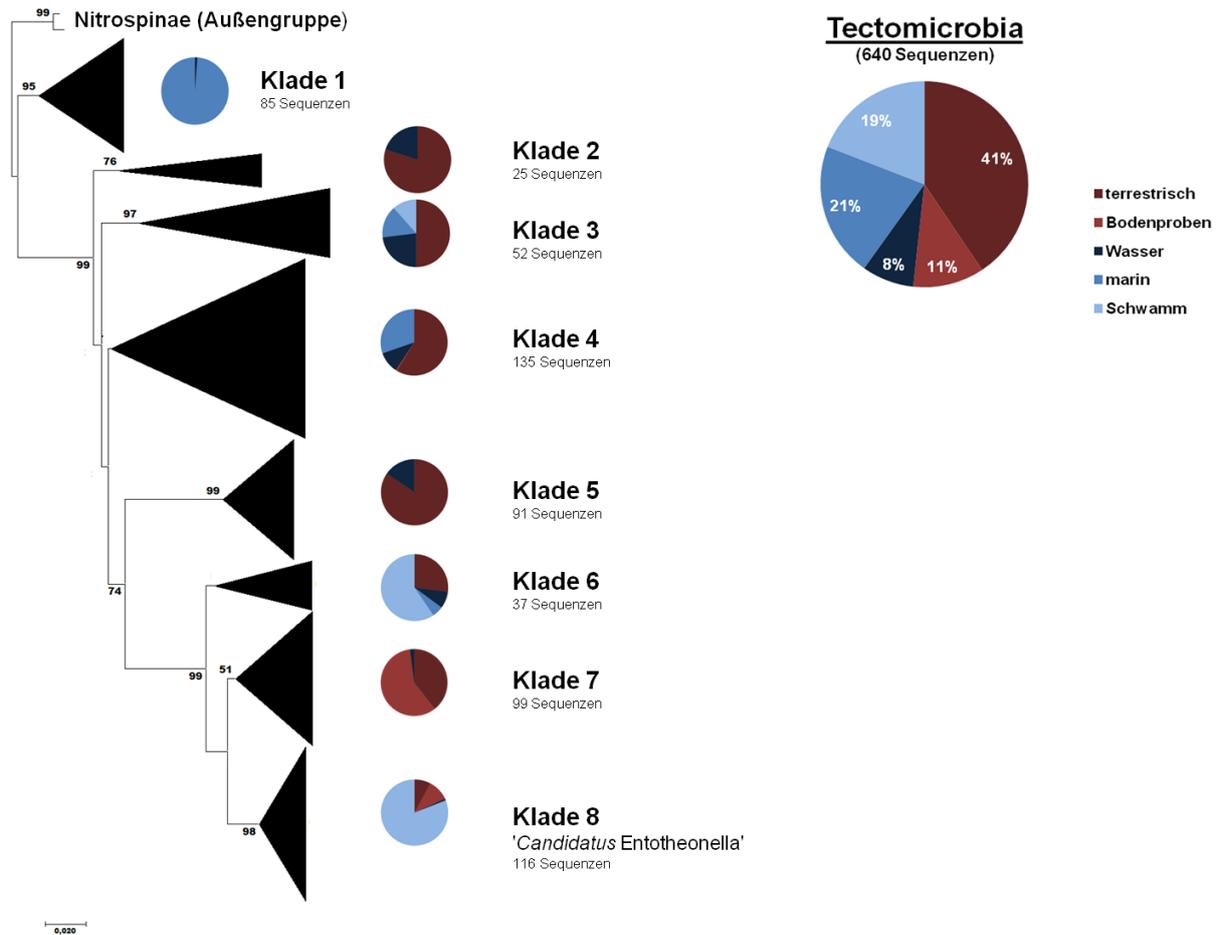


Abbildung 5-18: Phylogenetische Analyse des Kandidatenphylums 'Tectomicrobia' mittels Neighbour Joining.

Phylogenetischer Stammbaum (Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, bootstrapping: 1000x)^{190,198,211,212} der 16S rRNA-Gensequenzen aus den in dieser Arbeit untersuchten Bodenproben (Base 225-1130 (*E. coli*); enthält die variablen Regionen V3-V6) und weiterer verwandter 16S-Sequenzen von '*E. factor*' aus der NCBI Nukleotid-Datenbank. In einem Kuchendiagramm ist die jeweilige Verteilung der Habitate der in den Kladen enthaltenen Sequenzen dargestellt. Das Kuchendiagramm rechts zeigt die gesamte Verteilung der Habitate des Kandidatenphylums 'Tectomicrobia'.

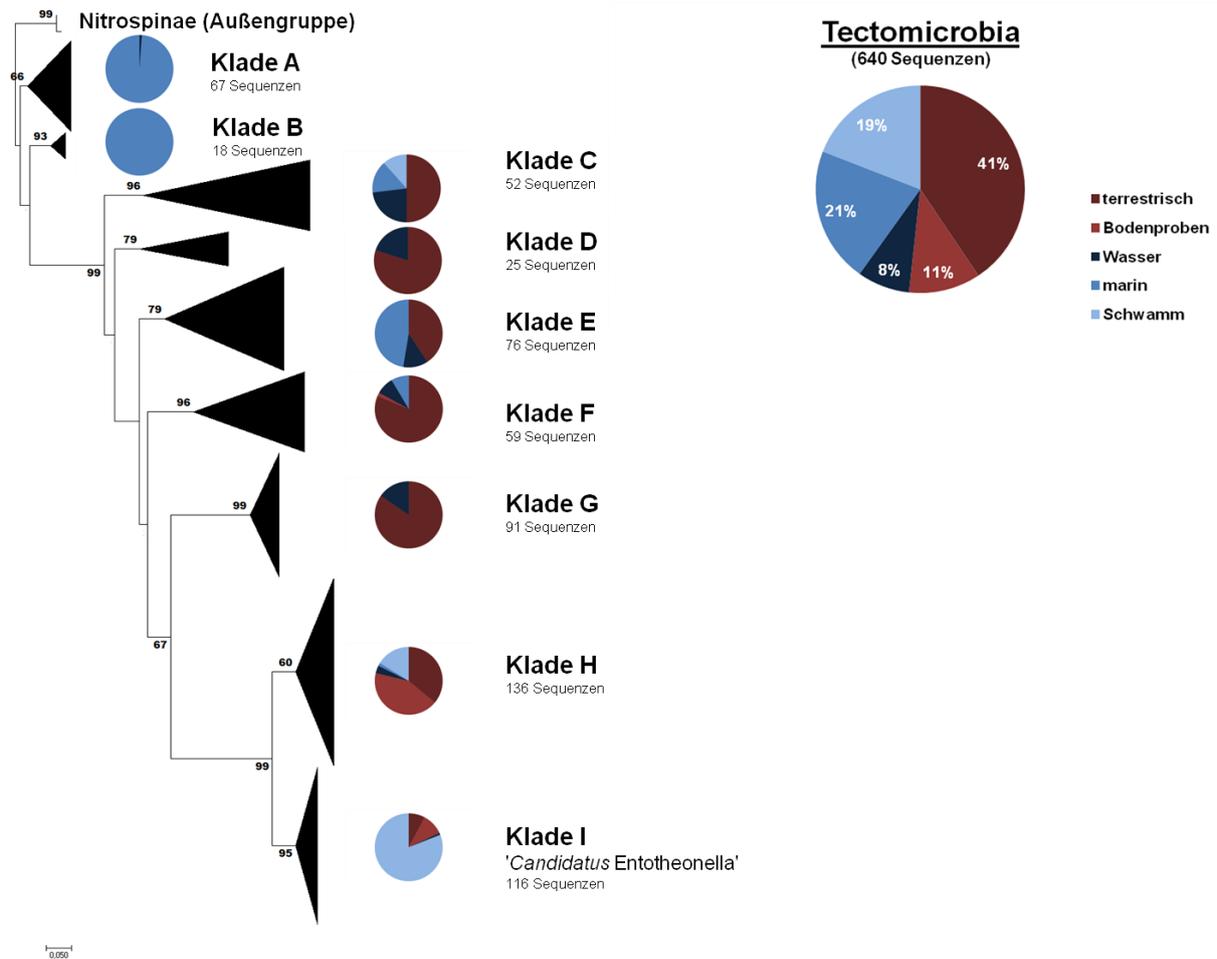


Abbildung 5-19: Phylogenetische Analyse des Kandidatenphylums 'Tectomicrobia' mittels Maximum Likelihood.

Phylogenetischer Stammbaum (Software: MEGA 7, Methode: Maximum Likelihood, Model: Maximum Composite Likelihood, Gamma Distribution, bootstrapping: 1000x) der 16S rRNA-Gensequenzen der aus den in dieser Arbeit untersuchten Bodenproben (Base 225-1130 (*E. coli*); enthält die variablen Regionen V3-V6) und weiterer verwandter 16S-Sequenzen von 'E. factor' aus der NCBI Nukleotid-Datenbank. In einem Kuchendiagramm ist die jeweilige Verteilung der Habitate der in den Kladen enthaltenen Sequenzen dargestellt. Das Kuchendiagramm rechts zeigt die gesamte Verteilung der Habitate des Kandidatenphylums 'Tectomicrobia'.

Die Analysen nach zwei unterschiedlichen Algorithmen (*Neighbour-Joining* und *Maximum-Likelihood*, Kapitel 3.4.2) zeigten eine ähnliche Aufspaltung des Stammbaums, es konnten jedoch auch Unterschiede beobachtet werden. Im Folgenden soll der Stammbaum nach *Neighbour-Joining* (NJ) näher betrachtet und die Unterschiede zum Stammbaum nach *Maximum-Likelihood* (ML) aufgezeigt werden.

Der NJ-Stammbaum zeigt eine Aufspaltung in zwei Hauptcluster (Abbildung 5-18, „Klade 1“ und „Klade 2-8“). Hohe *bootstrap*-Werte stützen diese Hypothese. Schaut man sich die Aufspaltung des ML-Stammbaums an, so zeigt sich, dass sich hier die Sequenzen aus Klade 1 (NJ) auf zwei Kladen aufteilen (Abbildung 5-19, Klade A und B), die anders als im NJ-Stammbaum keinem direkten gemeinsamen Ursprung haben (Abbildung 5-20).

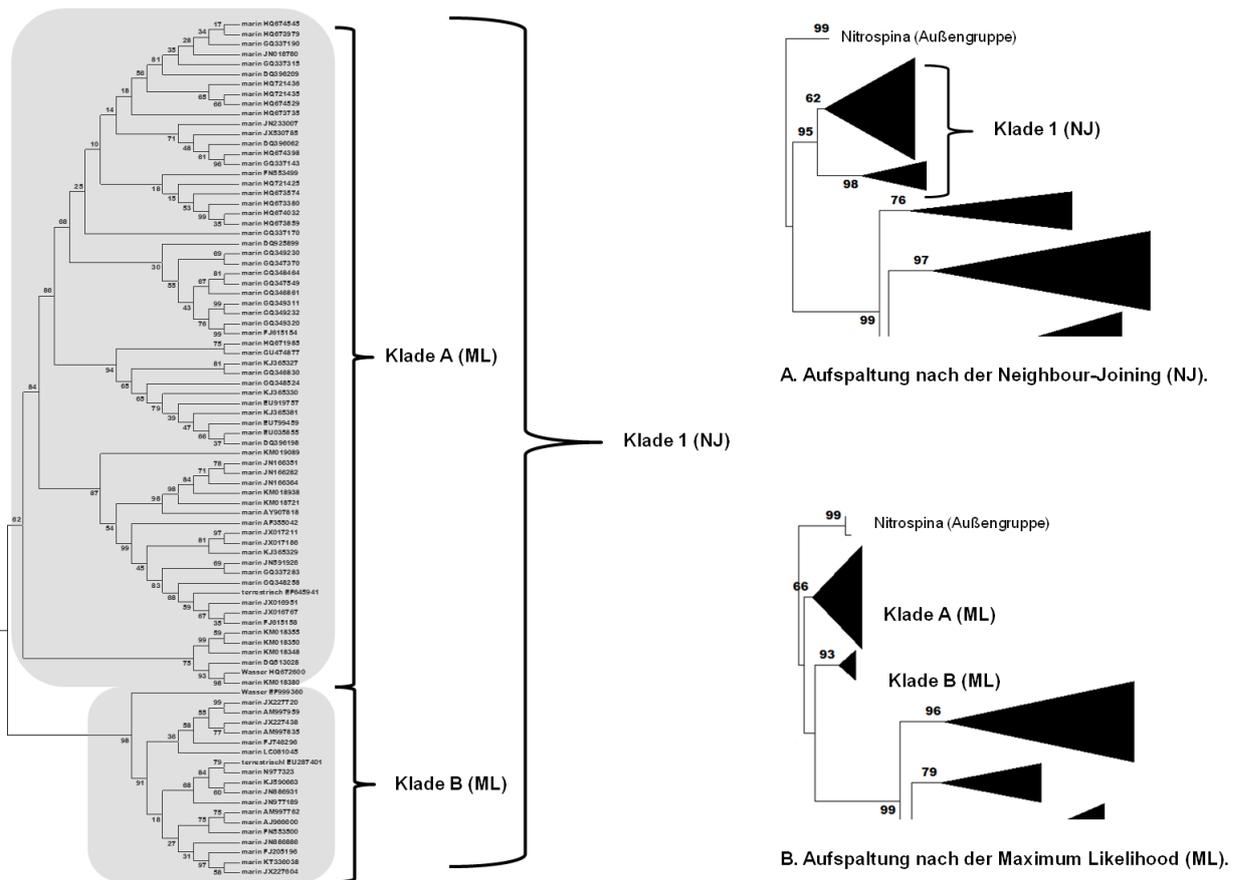


Abbildung 5-20: Vergleich der „Klade 1“ (NJ) mit den „Kladen A und B“ (ML). Links (und rechts oben) dargestellt ist die Topologie der „Klade 1“ (NJ), darin markiert sind die beiden Cluster, die im ML-Stammbaum keinen direkten gemeinsamen Ursprung haben (Abbildung links unten, „Klade A und B“).

„Klade 1“ enthält 85 Sequenzen mit einer *pairwise identity* (PI) von 93,7%, was der taxonomischen Klassifizierung einer Familie entspricht (Tabelle 5-3, Abbildung 5-18). Die PI der Kladen A und B (ML) beträgt 95,1% bzw. 95,8% und ist damit etwas höher als die der Klade 1 des NJ-Stammbaums (entsprechen der taxonomischen Klassifizierung von Genera). Betrachtet man jedoch die *bootstrap*-Werte der beiden Aufspaltungen, so scheint die Hypothese nach der NJ-Methode wahrscheinlicher (Abbildung 5-18, Abbildung 5-19, Abbildung 5-20). Alle Sequenzen dieser Klade(n) sind, mit einer Ausnahme (aus einer Süßwasserprobe), marinen Ursprungs und stammen aus Proben aller Weltmeere (detaillierte Aufspaltung im Anhang, Abbildung 10-1; Tabelle 10-1).

Das zweite Hauptcluster verästelt sich weiter in mehrere Untergruppen (NJ: „Klade 2 -8“; ML: „Klade C-I“; Abbildung 5-18, Abbildung 5-19). Diese sollen im Folgenden näher beschrieben werden.

Die erste Untergruppe (NJ: Klade 2, ML: Klade D) besteht aus 25 Sequenzen, die überwiegend aus terrestrischen Proben (80%) stammen. Fünf Sequenzen stammen aus Süßwasserproben. Die PI liegt bei 88,8% und entspricht damit der taxonomischen Klassifizierung einer Ordnung. Auffällig viele der Sequenzen stammen aus Proben

thermischen Ursprungs (z. B. Sequenz AB477994 aus geothermaletem Grundwasser; EF559004, EF559052, EF559064 aus einem Faulbehälter mit 55°C; FJ638546 aus dem Sediment einer heißen Quelle; vgl. Anhang Tabelle 10-1). Eine detaillierte Aufspaltung dieser Klade findet sich im Anhang (Abbildung 10-2; Tabelle 10-1).

Die 52 Sequenzen der zweiten Untergruppe (Abbildung 5-18 „Klade 3“, Abbildung 5-19 „Klade C“) stammen zu 50% aus terrestrischen Proben die andere Hälfte setzen sich aus Sequenzen aus Süßwasser-Proben (23%), Schwämmen (15%) und aus Proben mariner Habitats (12%) zusammen. Die Sequenzen der Schwammproben stammen aus Spezies der Gattung *Xestospongia* und *Geodia*. Die PI dieser Sequenzen liegt bei 84,1%. Dies entspricht ebenfalls der taxonomischen Klassifizierung einer Ordnung. Eine detaillierte Aufspaltung der Klade ist im Anhang zu finden (Abbildung 10-3, Tabelle 3-1).

„Klade 4“ (NJ; 135 Sequenzen) besteht überwiegend aus Sequenzen terrestrischen Ursprungs (59%), gefolgt von Sequenzen aus Meeresproben (30%) und Süßwasserproben (10%) (Abbildung 5-18). Eine Sequenz stammt aus einer Bodenprobe, die im Rahmen dieser Arbeit untersucht wurde (Tabelle 5-1, Probe 22, NO-Geilo). Auch die Sequenzen dieses Clusters weisen eine niedrige PI auf (83,4%). In der phylogenetischen Analyse nach dem Maximum Likelihood Algorithmus spalten sie sich in zwei Kladen („E und F“) auf, die anders als im NJ-Stammbaum nicht einem direkten gemeinsamen Knoten entspringen (Abbildung 5-21). Die *bootstrap*-Werte liegen bei beiden Hypothesen unter 50%. Daher kann hier keine Aussage dazu getroffen werden, welche der beiden Aufspaltungen wahrscheinlicher ist. Die PI der beiden „Klades E und F“ des Maximum Likelihood-Stammbaums liegen bei 86,6 und 88,8%. Auffallend viele Sequenzen dieses Clusters stammen aus salinischen Habitats (z. B. AB533815 aus einem Salzbecken des Death Valley, DQ206407 aus einem Natronsee in den USA, EU531787 von einem Schlammvulkan in China u. a.; Tabelle 3-1). Eine detaillierte Aufspaltung dieser Klade ist in Abbildung 10-4 bzw. Tabelle 10-1 im Anhang zu finden.

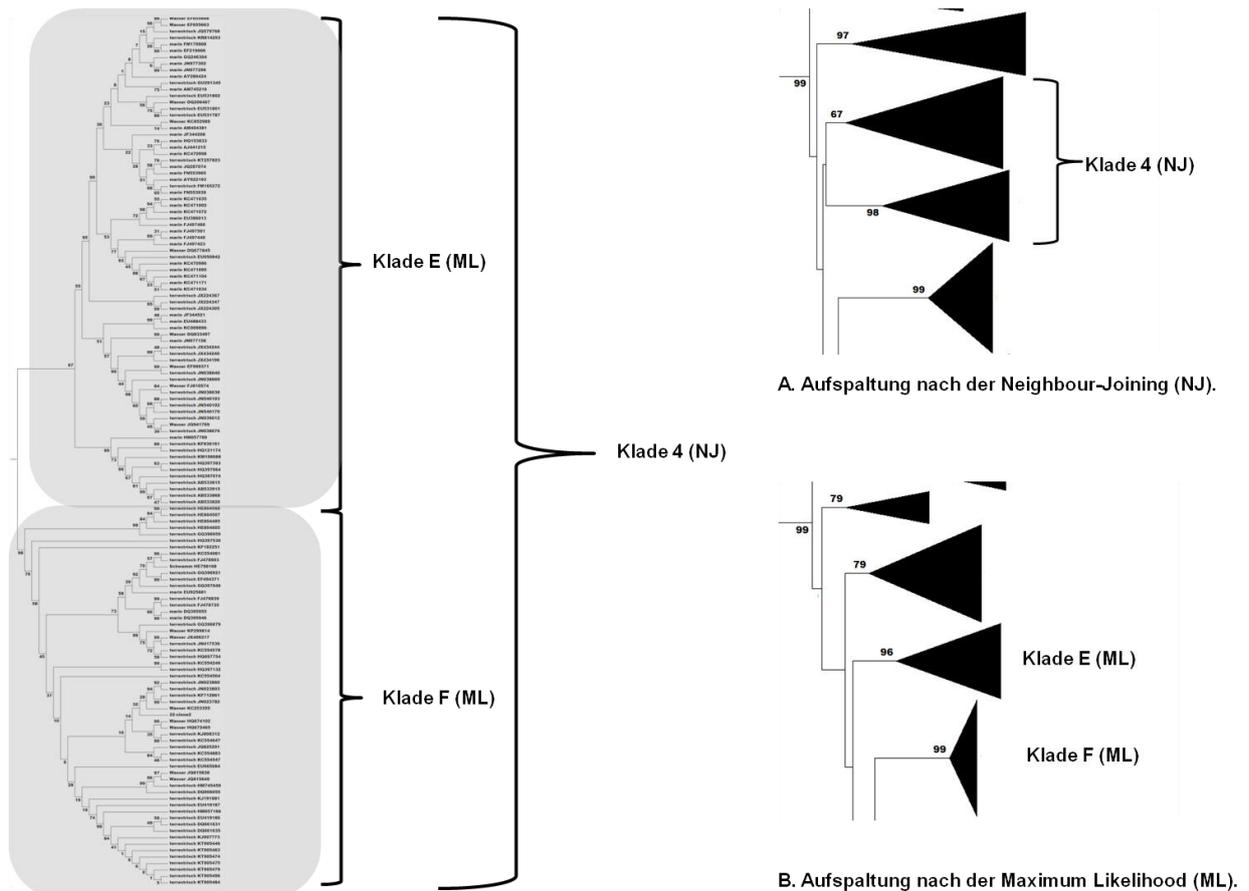


Abbildung 5-21: Vergleich der „Klade 4“ (NJ) mit den „Klades E und F“ (ML). Links (und rechts oben) dargestellt ist die Topologie der „Klade 4“ (NJ), darin markiert sind die beiden Cluster, die im ML-Stammbaum keinen direkten gemeinsamen Ursprung haben (Abbildung links unten, „Klade E und F“).

Die nächste Untergruppe (NJ: „Klade 5“; ML: „Klade G“) besteht ausschließlich aus 91 Sequenzen aus terrestrischen- (84%) und Süßwasserproben (16%) (Abbildung 5-18). Die PI liegt bei 94,6 % (Taxon: Familie). Eine interessante Gruppe von Sequenzen dieses Clusters stammt aus Müllhaufen von Blattschneideameisen aus Panama (z. B. LN566343). Weitere Sequenzen dieser Probe finden sich auch in „Klade 6, 7 und 8“ (NJ) bzw. in „Klade H und I“ (ML) (Abbildung 5-18, Abbildung 5-19). Sequenzen aus einer weiteren eher ungewöhnlichen Probe wurden aus dem Kot von Pinguinen isoliert (Tabelle 10-1, AY188328, AY218680). Auffallend viele Sequenzen dieses Clusters stammen aus Kulturböden (z. B. EU193059, EU276549, EU440654, FJ444679; detaillierte Aufspaltung: Abbildung 10-5; Tabelle 10-1).

Die Untergruppe „Klade 6“ des NJ-Stammbaums bestehend aus 37 Sequenzen mit einer PI von 93,6% (Taxon: Familie) und setzt sich aus Sequenzen aus Schwämmen (60%), terrestrischen Proben (27%), Wasserproben (8%) und marinen Proben (5%) zusammen (Abbildung 5-18). Die Sequenzen aus Schwämmen stammen aus folgenden Arten: *Xestospongia testudinaria*, *Xestospongia muta*, *Astrosclera willeyana*, *Geodia barretti*, und *Ircinia strobilina*. Viele der Sequenzen terrestrischen Ursprungs stammen von einer

biologischen Forschungsstation in der Prärie von Oklahoma (Kessler Farm Field Laboratory research station; detaillierte Aufspaltung Abbildung 10-6; Tabelle 10-1).

Die Untergruppe „Klade 7“ besteht aus 99 Sequenzen mit einer pairwise identity von 94,2% (Taxon: Familie) (Abbildung 5-18). Sie beinhaltet 58 Sequenzen aus 21 Bodenproben dieser Arbeit (Tabelle 5-1, Proben 5, 6, 8, 22, 23, 25, 26, 44, 45, 46, 47, A, B, C, D, E, F, H, J, K, L). Dies entspricht einem Anteil von 59% aller Sequenzen dieses Clusters. Die verbleibenden 41% setzen sich aus Sequenzen zusammen, die mit einer Ausnahme (eine Sequenz aus einer Süßwasserprobe) ebenfalls terrestrischen Ursprungs (39%) sind. Auch in dieser Klade finden sich insgesamt acht Sequenzen aus Müllhaufen von Blattschneideameisen (siehe auch „Klade 5, 7 und 8“, NJ bzw. „Klade H und I“, ML; Tabelle 10-1; weitere Details zur Aufspaltung auch in Abbildung 10-7 im Anhang).

Die beiden Untergruppen „Kladen 6 und 7“ fallen in der Stammbaumanalyse nach dem Maximum Likelihood Algorithmus in ein gemeinsames Cluster (Klade H, Abbildung 5-19, Abbildung 5-22). Die Aufspaltung unterscheidet sich von der im NJ-Stammbaum. Die hohen *bootstrap*-Werte der NJ-Stammbaumhypothese indizieren jedoch, dass diese Aufspaltung wahrscheinlicher ist als die der ML-Methode.

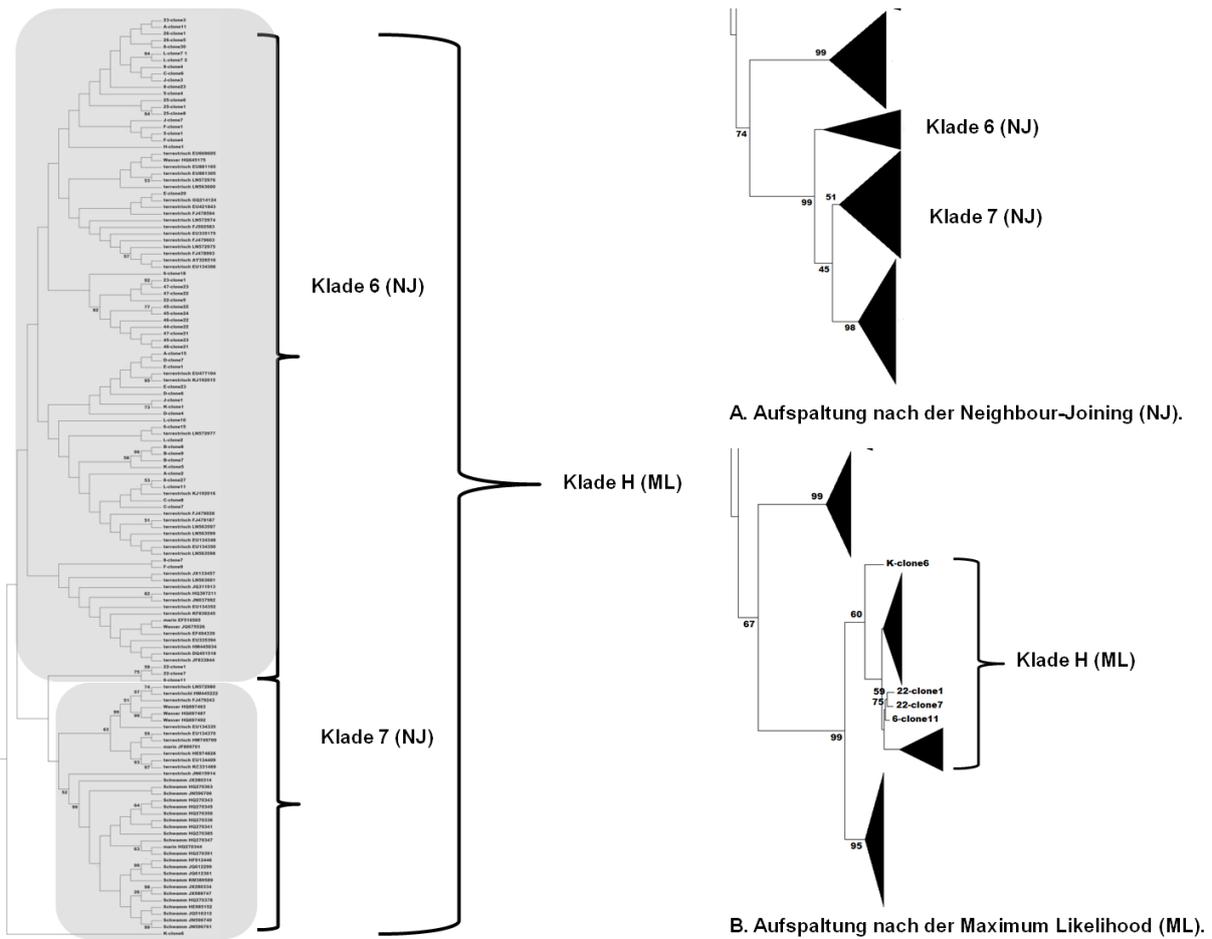


Abbildung 5-22: Vergleich der „Klade H“ (ML) mit den „Kladen 6 und 7“ (NJ). Links (und rechts unten) dargestellt ist die Topologie der „Klade H“ (ML), darin markiert sind die beiden Cluster, die im NJ-Stammbaum keinen direkten gemeinsamen Ursprung haben (Abbildung links, „Klade 6 und 7“).

Die letzte Untergruppe („Klade 8“ bzw. „Klade I“; 'Ca. Entotheonella') besteht aus 116 Sequenzen mit einer pairwise identity von 95,3% (Taxon: Genus) besteht zu einem überwiegenden Teil (81%) aus Sequenzen aus Schwämmen. Darunter befinden sich auch die u. a. in dieser Arbeit untersuchten 'E. factor'-Sequenz aus dem Schwamm *Theonella swinhoei*, sowie weitere bereits identifizierte 'Entotheonella'-Spezies (z. B. 'Candidatus Entotheonella palauensis', Abbildung 5-18, Abbildung 5-19, Tabelle 10-1) Weitere Sequenzen stammen aus 28 weiteren Schwämmen, die bereits im Rahmen der Arbeit von Wilson et al. isoliert und untersucht wurden (Tabelle 10-1).¹³¹ 10% der Sequenzen stammen aus sechs der im Rahmen dieser Arbeit gesammelten Bodenproben (Probe 5, 8, 26, 44, 46, I). Weitere 8% der Sequenzen sind ebenfalls terrestrischen Ursprungs. Davon erneut sechs Sequenzen aus Müllhaufen von Blattschneideameisen (siehe auch „Klade 5, 6 und 7“, NJ bzw. „Klade H und I“, ML; Abbildung 5-18, Abbildung 5-19, Tabelle 10-1). Die einzige nicht marine Wasserprobe stammt aus einem Salzwassersee in Mexiko (detaillierte Aufspaltung: Abbildung 10-8; Tabelle 10-1). Auffallend ist auch hier, dass sich die 'Ca. Entotheonella'-Sequenzen, wie auch schon bei der ersten phylogenetische Einordnung der 'Candidatus Entotheonella'-Sequenzen (Abbildung 5-12, Abbildung 5-14) beobachtet, anhand einer 29-

31 bp großen, variablen Insertionssequenz in der V3-Region von den anderen Kladen unterscheidet.

Parallel wurde von Eike Peters im Rahmen seiner Doktorarbeit ebenfalls eine phylogenetische Analyse mit zusätzlichen Schwammdaten durchgeführt.²¹⁵

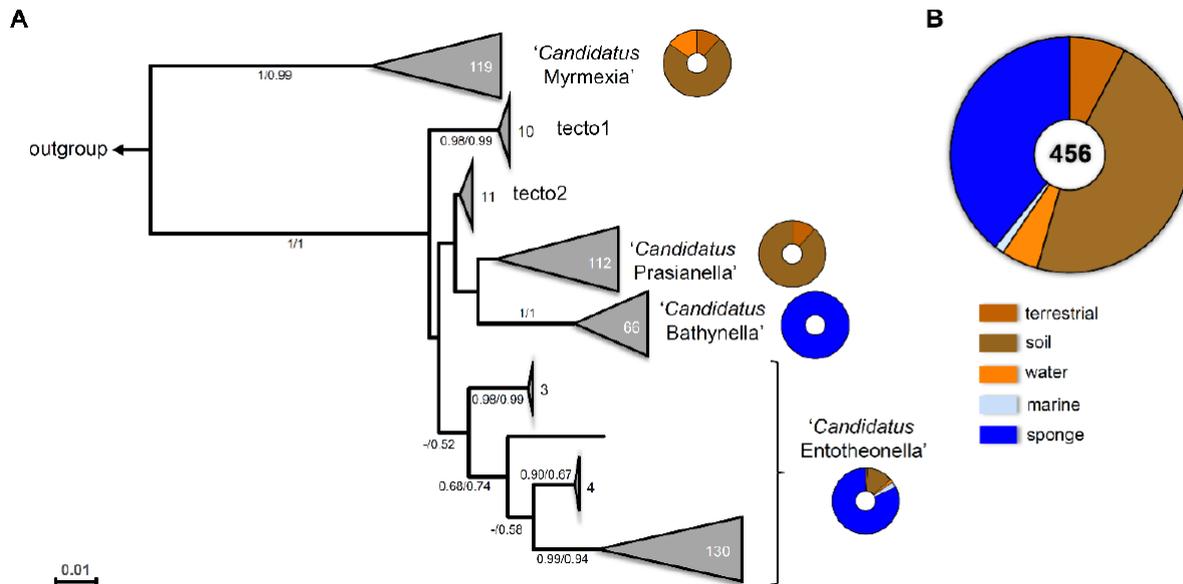


Abbildung 5-23: Phylogenetische Analyse des Kandidatenphylums 'Tectomicrobia' von Eike Peters. Neighbour Joining-Stammbaum anhand von 456 'Tectomicrobia' Sequenzen. Angegeben sind die bootstrap-Werte über 50 % (links nach der Neighbour Joining Methode, rechts nach Maximum Likelihood). In bzw. neben den jeweiligen Kladen ist die Anzahl der enthaltenen Sequenzen angegeben. Als Außengruppe wurden *Nitrospinae* verwendet. Die jeweilige Verteilung der Sequenzen anhand der Habitate ist in Kuchendiagrammen dargestellt.

Dabei wurden 456 Sequenzen in die Analyse einbezogen, u.a. auch einige neue Sequenzen aus verschiedenen Schwämmen der Gattung *Geodia*, die in den Stammbaumanalysen dieser Arbeit noch nicht berücksichtigt werden konnten. Im Folgenden soll kurz auf die Gemeinsamkeiten und Unterschiede des daraus Stammbaums mit dem in dieser Arbeit entwickelten Stammbaum eingegangen werden. Dabei fällt zunächst auf, dass in beiden Stammbäumen die Verteilung der Sequenzen aus terrestrischen (Bodenproben/terrestrisch) und aquatischen (Wasser/marin/Schwamm) Proben nahezu bei 50% liegt. Einige Kladen unterscheiden sich kaum. So fallen die Sequenzen der Klade 5 (NJ) bzw. G (ML) dieser Arbeit in die Klade 'Candidatus Myrmexia' (Abbildung 5-23). Die PI der als Familie klassifizierten Klade 5 bzw. G lag mit 94,6% nur knapp unter der taxonomischen Grenze eines Genus (95%). Das Kandidatengenus 'Myrmexia' enthält zusätzlich zu diesen Sequenzen 28 weitere und weist eine PI von 95,4% auf.

Ähnlich verhält es sich bei den Sequenzen der Klade 7 aus dieser Arbeit. Die PI liegt auch hier mit 94,2% nur knapp unter der Genusgrenze. Die Sequenzen dieser Klade fallen im

Stammbaum von Eike Peters in das Kandidatengenus 'Prasianella'. Dieses enthält zusätzlich noch weitere 13 Sequenzen und erreicht eine PI von 96,3%.

Eike Peters beschreibt in seiner Stammbaum-Analyse auch das Kandidatengenus 'Bathynella' (PI: 96,6%). Dieses enthält ausschließlich Sequenzen aus marinen Schwämmen; rund die Hälfte der Sequenzen wurden im Rahmen seiner Arbeit von Dr. Phil Crews und Dr. Paco Cardenas zur Verfügung gestellt und stammen von verschiedenen Schwämmen der Gattung *Geodia*. In die Stammbaumanalyse dieser Arbeit sind diese Sequenzen nicht mit eingeflossen. Eike Peters konnte bei der Aufspaltung des Kandidatengenus 'Bathynella' eine Korrelation mit der Wassertiefe der Proben beobachten. Im Stammbaum der hier vorliegenden Arbeit sind ausschließlich Sequenzen der in der Stammbaumanalyse von Eike Peters als „mesopelagische *Xestospongia*“ bezeichneten Gruppe zu finden. Diese fallen gemeinsam in Klade 6 (NJ). Eine Korrelation zwischen Phylogenie und Wassertiefe konnte daher nicht beobachtet werden. Anders als 'Ca. Bathynella', stammen die Sequenzen der Klade 6 nicht ausschließlich aus Schwämmen, wobei eine Clusterung der Schwammsequenzen innerhalb dieser Klade beobachtet werden kann (Abbildung 10-6, PI 97,7%). Die restlichen Sequenzen der Klade 6 finden sich im Stammbaum von Eike Peters in den Kladen tecto1 und tecto2 wieder.

Die phylogenetische Analyse des neuen Kandidatenphylums 'Tectomicrobia' zeigt, dass seine Mitglieder nicht nur in Schwämmen unterschiedlichster geographischer Regionen zu finden sind¹³¹, sondern weltweit in vielen unterschiedlichen Habitaten (Abbildung 5-17; Abbildung 5-18 bzw. Abbildung 5-19). So sind 52% aller untersuchten 16S rRNA-Gensequenzen terrestrischen Ursprungs, 8% aus Süßwasserproben, 21% aus Meerwasserproben und lediglich 19% stammen aus Schwämmen. Auffällig ist, dass in einer Probe meist verschiedene 'Tectomicrobia'-Phylotypen zu finden sind. Die Analysen nach Neighbour Joining und Maximum Likelihood zeigten an einigen Verzweigungspunkten unterschiedliche phylogenetische Hypothesen (Klade 1 (NJ) bzw. Klade A und B (ML); Klade 4 (NJ) bzw. Klade E und F (ML); Klade 6 und 7 (NJ) bzw. Klade H (ML)). Hier sind weitere Analysen zusätzlicher Sequenzen notwendig, um ein genaueres Bild der verwandtschaftlichen Beziehungen der einzelnen Kladen zueinander zu erreichen. Dies wird auch beim Vergleich mit der phylogenetischen Analyse von Eike Peters deutlich.

Bei einigen Kladen ist eine Korrelation zwischen natürlichem Vorkommen und Phylogenie zu erkennen. So kommen Mitglieder der Klade 1 nahezu ausschließlich in marinen Habitaten vor. Dies impliziert eine starke Anpassung an hohe Salzgehalte. Gleiches gilt auch für Mitglieder der Kladen 4 (bzw. Klade E und F) und 8 ('*Candidatus Entotheonella*'). Deren nächste Verwandte, Mitglieder der Klade 7, kommen hingegen ausschließlich in Böden ohne hohen Salzgehalt vor. Innerhalb des Kandidatengenus 'Entotheonella' ist eine Clusterung der

wenigen terrestrischen Sequenzen erkennbar. Eine Korrelation zwischen Salinität des Habitats und der Phylogenie ist auch in anderen Phyla zu finden. Ein Beispiel dafür ist das Genus *Salinispora*, des Phylums *Actinobacteria*, für dessen Wachstum Meerwasser obligat ist.²¹⁶ Auch diese Spezies stehen im Fokus der Wissenschaft, da sie potente Naturstoffproduzenten sind.²¹⁷ Mitglieder der Klade 2 (bzw. Klade D) und 5 (bzw. Klade G) sind zu einem überwiegenden Teil in terrestrischen Habitaten (80 bzw. 84%) beheimatet, einige jedoch auch in Süßwasserhabitaten. Hier ist innerhalb der Klade keine Clusterung nach Habitat zu erkennen (Tabelle 3-1, Abbildung 10-2, Abbildung 10-5). Die Mitglieder der Klade 2 (bzw. Klade D) sind, wie bereits beschrieben, bevorzugt in thermalen Habitaten zu finden. Die unterschiedlichen Habitate sprechen, wie bereits ausgeführt, ebenfalls für die große Anpassungsfähigkeit dieser Spezies.

Alle bisherigen Erkenntnisse zu Mitgliedern des Kandidatengenus 'Entotheonella' zeigen, dass dieses Genus eine vielversprechende Quelle für bioaktive Naturstoffe ist. Die bisher untersuchten Naturstoff-produzierenden 'Entotheonella'-Spezies stammen jedoch ausschließlich aus Schwämmen. Daher ist es besonders interessant, dass einige der im Rahmen dieser Arbeit isolierten Sequenzen ebenfalls in diese Klade fallen. Wie bereits erwähnt, clustern einige der Sequenzen terrestrischen Ursprungs innerhalb des Genus zusammen (Abbildung 10-8). Die Sequenzen der Probe I nehmen eine besondere Stellung ein. Mit einer PI von 99,1 bzw. 99,3% fallen diese Sequenzen gemeinsam in ein Cluster mit 'E. factor' und zwei weiteren Sequenzen aus den Schwämmen *Aplysina aerophoba* und *Stylissa carteri*. Auch in diesen Schwämmen konnten bereits Naturstoffe nachgewiesen werden.^{218,219} Die Sequenzen der anderen Bodenproben innerhalb der 'Entotheonella' Klade fallen gemeinsam mit Sequenzen aus einem Müllhaufen von Blattschneideameisen. Die enge Phylogenese von potenten Naturstoffproduzenten aus Schwämmen wie 'E. factor' und 16S rRNA-Gensequenzen aus Spezies terrestrischen Ursprungs, lässt mutmaßen, dass diese ebenfalls eine Quelle bioaktiver Naturstoffe sind.

5.1.4 Isolierung von 'Candidatus Entotheonella' aus Bodenproben

Nach dem Nachweis von 'Candidatus Entotheonella'-16S rRNA-Gensequenzen (Kapitel 5.1.2) in verschiedenen Bodenproben sollte in einem nächsten Schritt versucht werden, diese Bakterien aus den Proben zu isolieren, um eine morphologische und metabolische Charakterisierung der Bakterien zu ermöglichen. Dazu wurde zunächst Probe 5 genauer untersucht. Diese Bodenprobe wurde in einem Nutzgarten direkt unter einer Rhabarber-Pflanze (*Rheum Rhabarbarum*) entnommen. Um nähere Erkenntnisse über die genaue Lokalisation der 'Candidatus Entotheonella' zu gewinnen, wurden verschiedene Tiefen (Probe 5: Bodenoberfläche; Probe 5.1: Probe aus ca. 15 cm Tiefe, Probe 5.2: Probe aus ca. 30 cm Tiefe) auf 'Candidatus Entotheonella' untersucht. Wie in Abbildung 5-10 zu erkennen

ist, konnte in allen drei Proben ein spezifisches PCR-Produkt erzeugt werden. Um herauszufinden, ob diese '*Candidatus Entotheonella*'-Phylotypen ggf. Wurzel-assoziiert sind, wurde zusätzlich eine Bodenprobe samt Wurzel der Rharbarber-Pflanze genommen. Anschließend wurde sowohl die anhaftende Erde, die unbehandelte Wurzel, als auch die Wurzel nach Entfernung der Erde (durch Abspülen) auf '*Candidatus Entotheonella*'-DNA untersucht. Außerdem wurde auch das Pflanzenmaterial (Stengel) selbst untersucht. Dabei zeigte sich, dass die in dieser Probe nachgewiesenen '*Candidatus Entotheonella*' nicht wurzel-assoziiert sind (Abbildung 5-24).

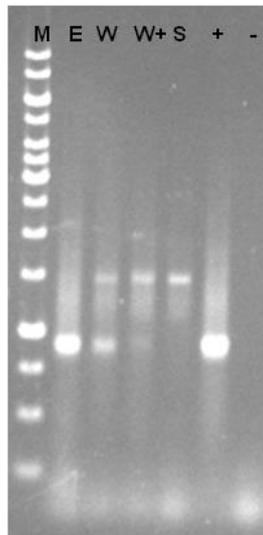


Abbildung 5-24: Untersuchungen von Wurzel und Pflanzenmaterial auf '*Candidatus Entotheonella*'.

Untersucht wurden die Wurzel einer Rharbarber-Pflanze vor (**W**) und nach Reinigung (**W+**), die anhaftende Erde (**E**) und der Stengel der Pflanze (**S**). Als Positivkontrolle wurde die aufgereinigte DNA der Bodenprobe 5 verwendet (**+**) (**-** kennzeichnet die Negativkontrolle). Gezeigt ist die PCR mit den '*Candidatus Entotheonella*' spezifischen Primern 238F und 1290R.

Um weitere Hinweise über die Bodenhabitats zu erlangen wurde versucht die Bodenarten anhand der Korngröße zu bestimmen (Kapitel 7.6). Bei vielen war dies jedoch aufgrund der Methodik der Probennahme (es handelt sich um so genannte Mischproben; Kapitel 7.6) nicht möglich. Die Ergebnisse sind im Anhang in Tabelle 10-5 aufgeführt.

Zur Isolierung der '*Candidatus Entotheonella*'-Phylotypen musste zunächst ein neues Protokoll etabliert werden. Dazu wurde die jeweilige Bodenprobe in einer isotonischen NaCl-Lösung resuspendiert. In Anlehnung an die Isolierung von 'Entotheonella factor' aus dem Schwamm *Theonella swinhoei* (Kapitel 3.3.1) wurde die Lösung dann durch ein Netzgewebe (32 µm Porengröße) gefiltert, um grobe Erdpartikel zu entfernen.¹³¹ Anschließend sollte '*Candidatus Entotheonella*' durch differentielle Zentrifugation von anderen Bakterienspezies separiert werden (Abbildung 5-25). Durch die Isolierung von 'E. factor' war bekannt, dass diese u. a. für die meisten Bakterien eher ungewöhnliche Sedimentationseigenschaften aufweisen: Sie sedimentieren bereits bei einer Zentrifugationsgeschwindigkeit von 100 x g.

Die meisten anderen unizellulären Bakterien sedimentieren bei 4500 x g.¹³¹ Die z. T. hohe Übereinstimmung der aus den Bodenproben gewonnenen 16S rRNA-Gensequenzen zur 'E. factor' (z. B. Sequenz aus Probe I: 99,3% (Tabelle 5-2)) lassen einen hohen Verwandtschaftsgrad (selbe Spezies) vermuten. Daher wurden, wie bereits bei der Isolierung von 'E. factor' aus dem Schwamm *Theonella swinhoei* drei verschiedene Zentrifugationsgeschwindigkeiten von 100 x g, 1000 x g und 4500 x g gewählt.

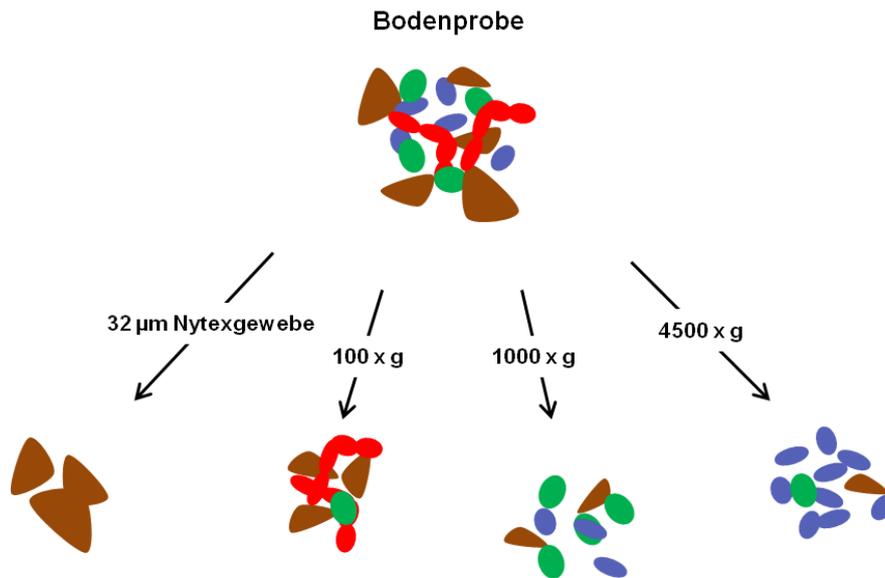


Abbildung 5-25: Isolierung von 'Candidatus Entotheonella' durch differenzielle Zentrifugation. In braun dargestellt sind die Bodenpartikel, in blau und grün unterschiedliche Bakterienspezies und in rot die vermuteten filamentösen 'Candidatus Entotheonella'.

Wie auch beim Screenen der Bodenproben wurden nach der Zentrifugation alle Fraktionen mittels PCR auf 'Candidatus Entotheonella' untersucht (Abbildung 5-26). Dazu wurde auch hier zunächst die DNA mit dem *Power Soil DNA Extraction Kit* (MoBio) isoliert, um eventuell noch vorhandene Huminsäure zu entfernen.

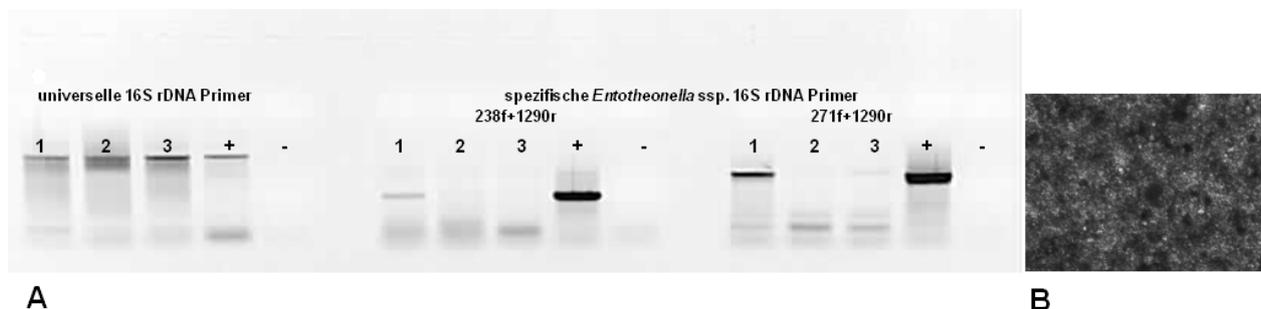


Abbildung 5-26: Bakterienfraktionen nach differentieller Zentrifugation. A: *Nested*-PCR-Ansatz zum Nachweis von 'Candidatus Entotheonella' nach Durchführung des Isolierungsprotokolls. 1: 100 x g-Fraktion, 2: 1000 x g-Fraktion, 3: 4500 x g-Fraktion, +: Positivkontrolle, -: Negativkontrolle. B: Mikroskopische Untersuchung der 100 x g-Fraktion.

Wie beispielhaft an Bodenprobe 5 in Abbildung 5-26 (A) gezeigt, konnte die 16S rRNA-Gensequenzen der gesuchten Bakterien in den Fraktionen 100 x g nachgewiesen werden. Allerdings ist ebenfalls eine leichte Bande in der Fraktion 4500 x g (Abbildung 5-26,A, spezifische Primer 271F+1290R; Spur 3) zu erkennen. Erste Versuche zeigten, dass die Verunreinigungen der Bakterienfraktionen mit kleinsten Erdpartikeln eine morphologische Beurteilung der separierten Bakterien unter dem Mikroskop unmöglich machten (Abbildung 5-26, B). Daher wurde anschließend eine Dichtegradientenzentrifugation durchgeführt. Bei diesem physikalischen Verfahren werden die Partikel aufgrund ihrer Sedimentationseigenschaften in einem Dichtegradientenmedium, in diesem Fall Nycodenz (60% w/v; Serva Elektrophoresis; Kapitel 7.7.1, Abbildung 7-4), voneinander getrennt.²²⁰

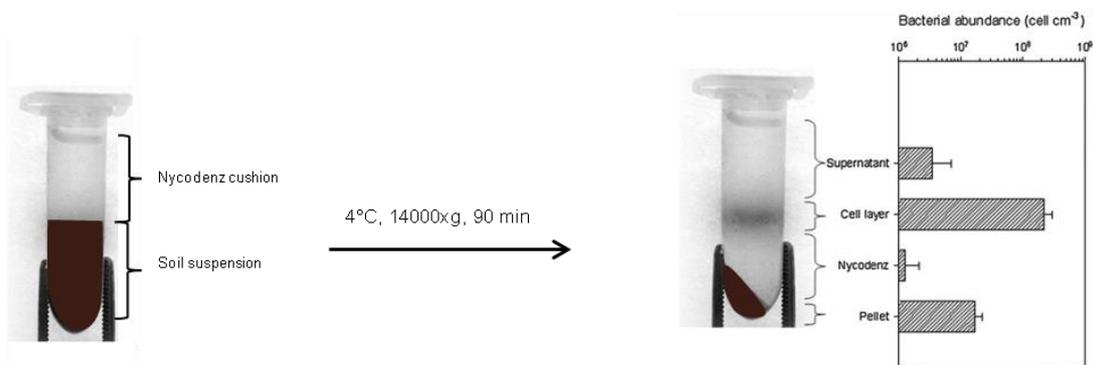


Abbildung 5-27: Prinzip der Dichtegradientenzentrifugation mittels Nycodenz.²²⁰ Im Reaktionsgefäß links in der Abbildung ist die in NaCl gelöste Bodenprobe, die bereits unter das Dichtegradientenmedium Nycodenz pipettiert wurde, dargestellt. Durch einen Zentrifugationsschritt bei 4 °C und 14000 x g für 90 min kann ein großer Teil der Bakterienzellen (*cell layer*) von den restlichen Probenbestandteilen abgetrennt werden (*Pellet*). Die Zellzahl pro Kubikzentimeter innerhalb der jeweiligen Schichten ist im rechten Säulendiagramm dargestellt (**Supernatant**: Überstand; **Cell layer**: Bakterienzellschicht; **Pellet**: hauptsächlich feste Bestandteile wie z. B. Erdpartikel der Bodenprobe).

Anschließend wurde wiederum eine PCR durchgeführt (Abbildung 5-28), um den Erfolg der Aufreinigung bzw. die Probe auf die Anwesenheit der '*Candidatus Entotheonella*' -16S rRNA-Gensequenz zu überprüfen (Abbildung 5-28). Bei der Isolierung der Bakterien wurde der Fokus auf die Proben gelegt, die beim Screening und der anschließenden Sequenzierung Sequenzen mit einer hohen Übereinstimmung zur 16S rRNA-Gensequenz von '*Entotheonella factor*' und damit einen hohen Verwandtschaftsgrad (selbe Spezies) zu '*E. factor*' aufwiesen (Tabelle 5-2). Dabei handelt es sich um die Proben 5 (96,5% pairwise identity), 26 (97% pairwise identity) und I (99,3% pairwise identity) (Tabelle 5-2). Die Bodenprobe 5 stammt wie bereits beschrieben aus einem Nutzgarten, in direkter Nähe einer Rhabarberpflanze. Bodenprobe 26 wurde in einem Ziergarten entnommen. Bei Probe I handelt es sich um eine Sandprobe aus den Fugen eines Kopfsteinpflasters.

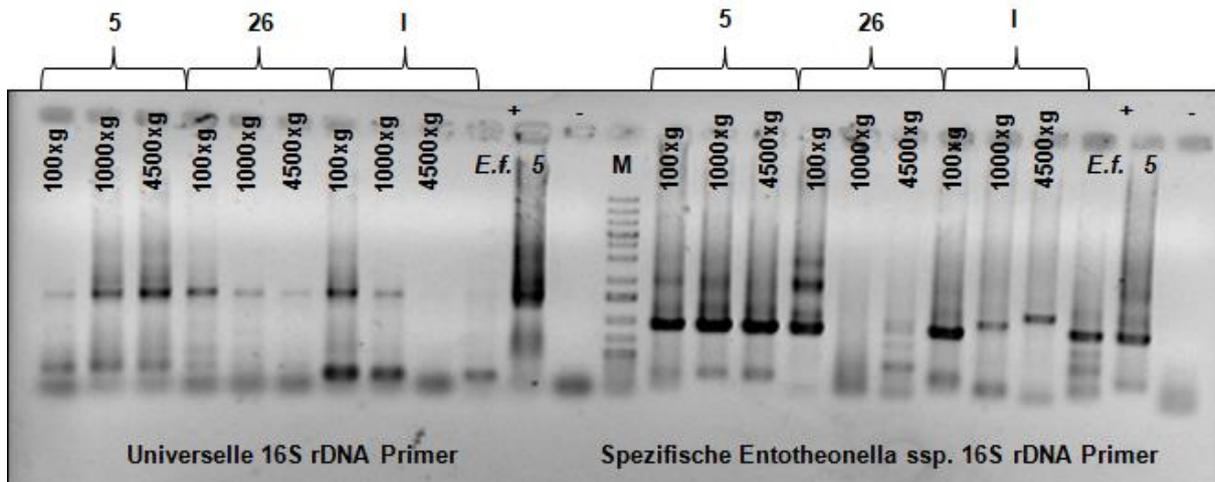


Abbildung 5-28: Screening der Bakterienfraktionen nach differentieller Zentrifugation und Dichtegradientenzentrifugation. Nach Aufreinigung der Bodenproben durch differentielle Zentrifugation und Dichtegradientenzentrifugation, wurden '*Candidatus Entotheonella*' durch Amplifikation des 16S rRNA-Gens in einer *nested*-PCR nachgewiesen. 5, 26, und I bezeichnen die jeweiligen Bodenproben; 100 x g, 1000 x g, 4500 x g bezeichnen die jeweilige Zentrifugationsfraktion; E.f.: 'Entotheonella factor' als Positivkontrolle; 5: aufgereinigte metagenomische DNA aus Bodenprobe 5 als Positivkontrolle; -: Negativkontrolle.

Wie in Abbildung 5-28 zu sehen, konnten auch hier nach der Aufreinigung die spezifischen 16S rRNA-Gensequenzen mittels PCR amplifiziert werden. Dieses Ergebnis wurde in einer anschließenden Sequenzierung der Amplifikate verifiziert. Die dabei gewonnen Sequenzen stimmten mit den ursprünglichen Sequenzierergebnissen aus dem Screening der Bodenproben überein (Tabelle 5-2).

Visuelle Untersuchungen der Bakterienfraktionen zeigten, dass die Diversität der Proben immer noch hoch war. Da sowohl die Konzentration der gesuchten Spezies als auch deren genaue Morphologie unbekannt waren, sollten diese zur genauen Identifizierung mittels fluoreszenzmarkierter Sonden visualisiert werden.

5.1.4.1 CARD-FISH

Aufgrund des scheinbar hohen Verwandtschaftsgrads der zu isolierenden Bakterien zu 'E. factor' (hohe Sequenzübereinstimmung, engl.: *pairwise identity*, des 16S rRNA-Gens; Tabelle 5-2) wurde zunächst der Fokus auf filamentöse Bakterien gelegt. Bei der Untersuchung der so isolierten Bakterienfraktionen mittels Durchlichtmikroskopie konnten erste Kandidaten identifiziert werden.

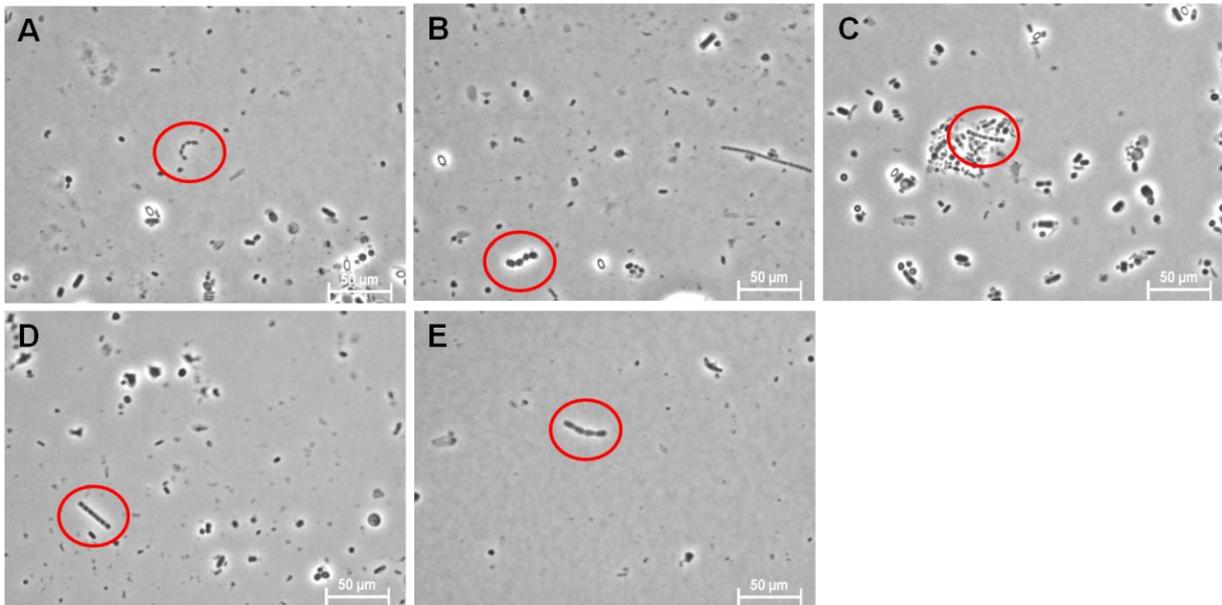
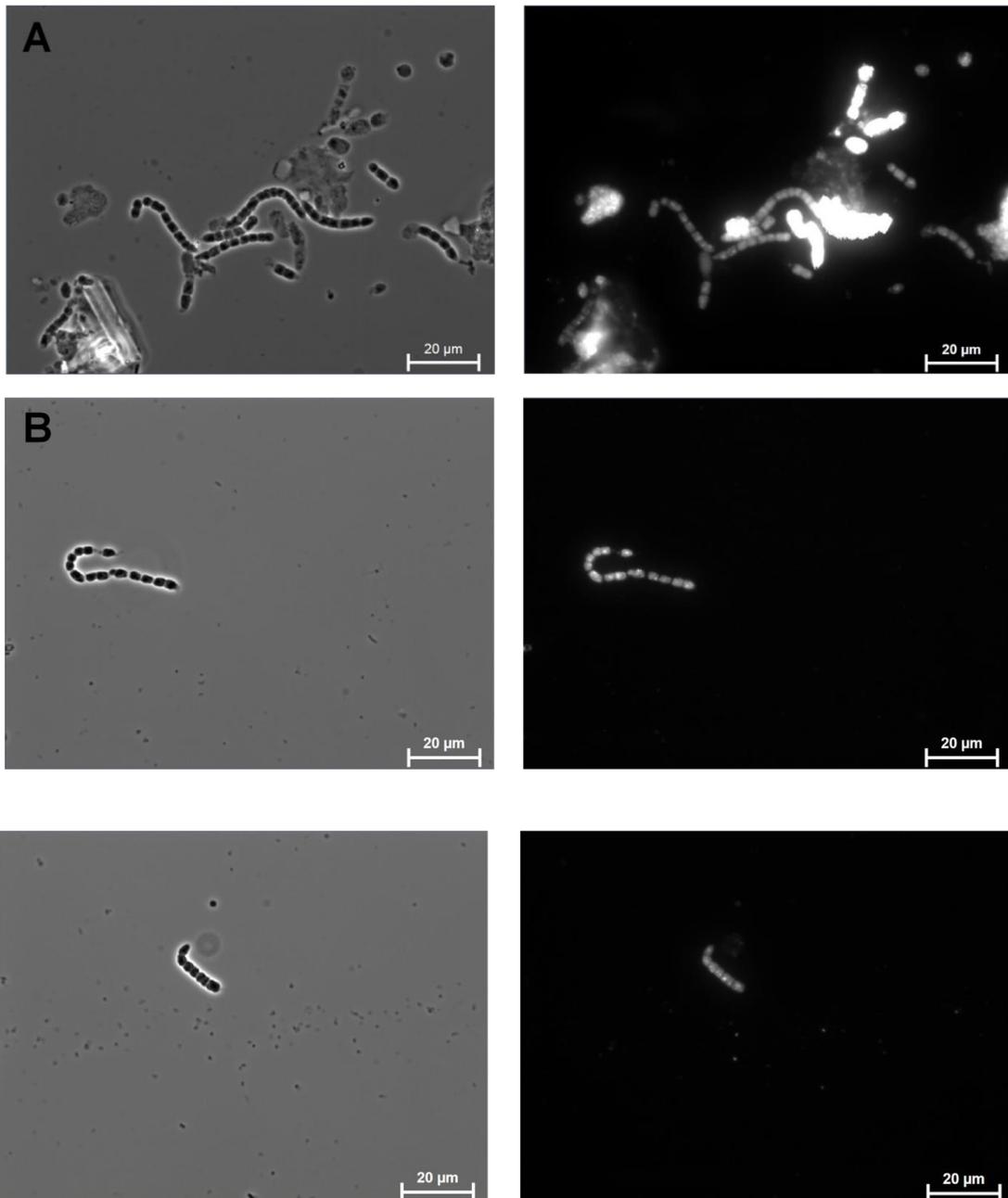
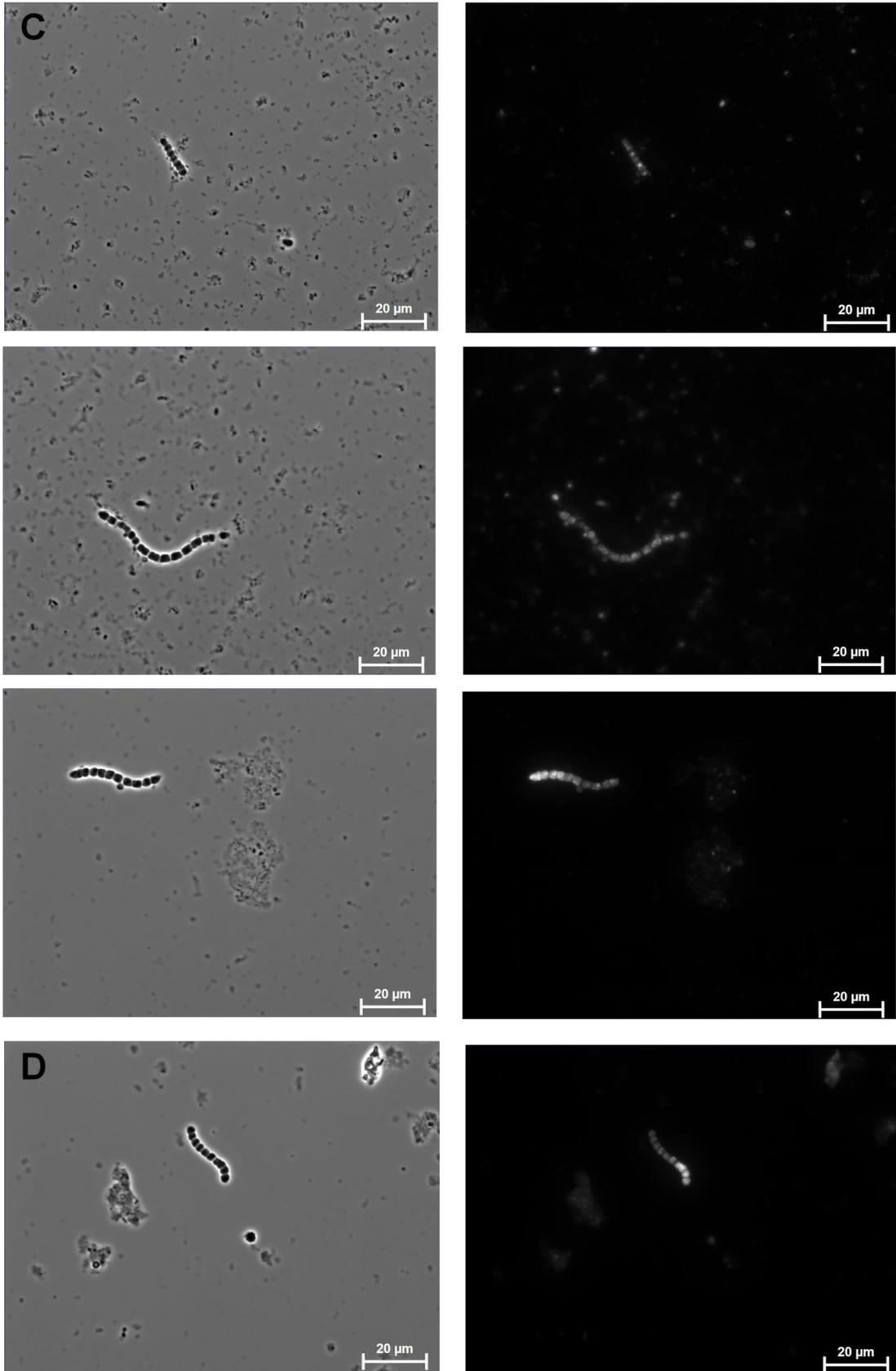


Abbildung 5-29: Lichtmikroskopieaufnahmen einiger Bakterienfraktionen. A und B: Probe 26 (100 x g); C: Probe 26 (1000 x g); D: Probe 26 (4500 x g), E: Probe I (100 x g). Rot eingekreist sind einige Bakterien, die morphologische Übereinstimmungen zu 'E. factor' aufweisen.

Die zu untersuchenden Spezies könnten zum einen allerdings auch einen anderen morphologischen Aufbau als ihrer Verwandten aus dem Schwamm haben, zum anderen bilden auch andere Bakterienpezies Filamente. Außerdem zeigten die Bakterienfraktionen auch nach der Aufreinigung eine hohe Diversität. Daher musste eine genauere Methodik zur Identifizierung der neuen '*Candidatus Entotheonella*' verwendet werden. Eike Edzard Peters (AG Piel, ETH Zürich) entwickelte bereits im Vorfeld anhand der 16S rRNA-Gensequenz eine spezifische Sonde für 'E. factor' (Tabelle 7-24). Mit dieser konnte 'E. factor' erfolgreich markiert und in einer sogenannten CARD (CAlyzed Reporter Deposition)-FISH (Fluorescence In Situ Hybridization) visualisiert bzw. identifiziert werden (Kapitel 7.9.1). Bei einer CARD-FISH werden spezifische Oligonukleotide erstellt, die, anders als bei einer herkömmlichen FISH, nicht direkt mit einem Fluoreszenzfarbstoff, sondern mit einer Peroxidase (der sogenannten Meerrettichperoxidase, engl.: *horseradish*, HRP) verknüpft sind. In einer Hybridisierungsreaktion binden die Sonden dann, wie bei einer „normalen“ FISH, an die Zielsequenz (in diesem Fall an die 16S rRNA). In einem zweiten Schritt wird nun Fluoreszenz-markiertes Thymidin zugegeben. Durch eine Radikalisierungsreaktion der Thymidinmoleküle durch die gebundene Peroxidase kommt es zur Anlagerung und somit zur Akkumulation des Fluoreszenzsignals an der Ziel-DNA. Das so erzeugte Signal kann so bis zu 1000-fach verstärkt werden (verglichen mit der Signalstärke einer FISH mit Fluoreszenz-markierten Sonden).¹³³ Ein Vergleich der Sondensequenz (Tabelle 7-24) mit den gewonnenen Sequenzen aus den Bodenproben zeigte, dass diese auch in einer CARD-FISH zur Identifizierung der bis dato unbekanntes '*Candidatus Entotheonella*' verwendet werden kann. Dieses CARD-FISH Experiment wurde nach einem von Eike Edzard Peters für

'Entotheonella factor' etablierten Protokoll durchgeführt (Kapitel 7.9). Als Positivkontrolle wurde zusätzlich zu den isolierten Bakterienfraktionen eine Probe von 'Entotheonella factor' auf ein Reaktionsfeld eines mit mehreren Reaktionsfeldern versehenen Objektträgers aufgetragen (Abbildung 5-30, A). Nach der Permeabilisierung der Bakterienmembranen, der Inaktivierung endogener Peroxidasen und der Hybridisierung der Sonde an die Ziel-DNA wurde die eigentliche Detektion der gebundenen Sonde durch Zugabe des markierten Tyramids durchgeführt. Die Detektion der so erzeugten Signale erfolgte an einem Zeiss Axioskop 2 Epifluoreszenz Mikroskop mithilfe eines entsprechenden Filtersets für den Fluoreszenzfarbstoff Alexa 647.





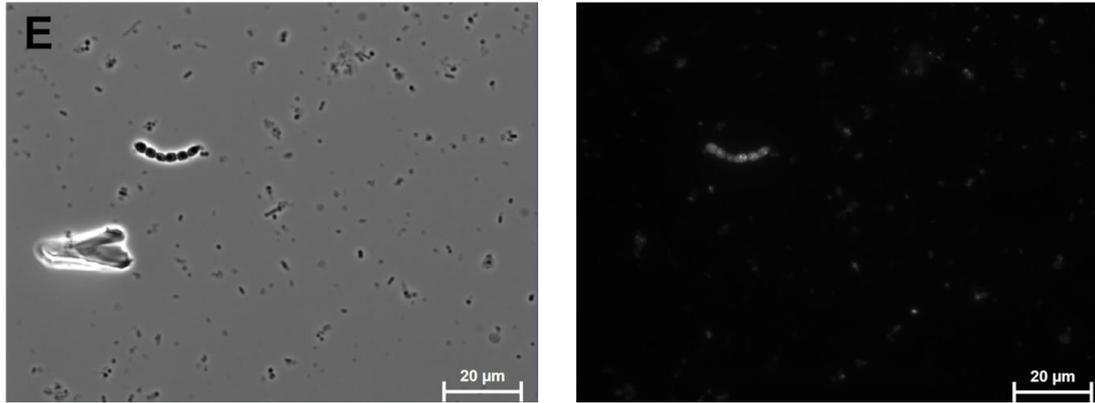


Abbildung 5-30: CARD-FISH der Bodenproben 5 und 26.

A: 'Entotheonella factor' als Positivkontrolle; **B:** Probe 5 (100 x g); **C:** Probe 5 (1000 x g); **D:** Probe 26 (100 x g); **E:** Probe 26 (1000 x g;).

In zwei der drei mit CARD-FISH untersuchten Bodenproben konnten so in einem ersten Versuch potentielle '*Candidatus Entotheonella*' anhand der spezifischen Sonden visualisiert werden (Abbildung 5-30, B-F).

Das Ergebnis dieser ersten CARD-FISH ließ sich bei Wiederholung des Versuchs jedoch nicht reproduzieren. Es konnten weder Filamente lichtmikroskopisch dargestellt noch durch Hybridisierung markiert werden. Eine Erklärung hierfür könnte eine Kreuzkontamination durch die Positivkontrolle ('*Entotheonella factor*') sein, die auf denselben Objektträger aufgetragen wurde. Gegen diese Theorie spricht allerdings, dass in der Vorbereitung zum ersten CARD-FISH-Versuch diese Filamente auch unter dem Lichtmikroskop zu sehen waren. Außerdem wurden die Proben nach dem Auftragen auf den Objektträger und vor der Hybridisierung durch Trocknung fixiert, was eine Kreuzkontamination in Folge unwahrscheinlich macht. Des Weiteren waren die markierten Filamente nicht in allen, sondern nur in vier der neun aufgetragenen Fraktionen (aus zwei der drei Bodenproben: Probe 5, 26 jeweils 100 x g, 1000 x g) zu finden (Abbildung 5-30). Eine weitere Möglichkeit, die dieses Phänomen erklären könnte, wäre eine Instabilität der Filamente selbst. Das heißt, die Filamente könnten durch das wiederholte Auftauen und Einfrieren degradiert bzw. lysiert sein. Bei Lagerung der DNA von '*E. factor*' unter -20 °C konnte beobachtet werden, dass diese z. T. so degradiert, dass ein Nachweis nicht mehr möglich ist. Eine Degradation der Filamente selbst konnte bisher jedoch nicht beobachtet werden. Dieses Beispiel zeigt jedoch, dass aufgrund der noch fehlenden Charakterisierung dieses Kandidatenphylums viele Phänomene noch nicht abschließend erklärt werden können.

Nachdem auch nach mehreren Wiederholungen keine spezifischen Signale an Bakterien detektiert werden konnten, musste aus zeitlichen Gründen auf weitere Versuche verzichtet werden.

5.1.4.2 Quantifizierung der '*Candidatus Entotheonella*'

Bis zur Durchführung der CARD-FISH-Experimente wurde die Existenz von '*Candidatus Entotheonella*' in den Bodenproben ausschließlich über eine *nested*-PCR nachgewiesen. Eine quantitative Spezifizierung dieser Bakterien wurde jedoch nicht durchgeführt. Eine Aussage über die Qualität der verwendeten Aufreinigungsmethode und deren Optimierung ist allerdings nur durch eine quantitative Bestimmung von '*Candidatus Entotheonella*' möglich. Zunächst wurde hierfür eine Quantifizierung über eine TaqMan-qPCR in Betracht gezogen. Stark vereinfacht wird bei dieser Methode die Zunahme des Fluoreszenzsignals einer spezifischen Sonde gemessen, das proportional zur Amplifikation des PCR-Produktes zunimmt.²²¹ Dies ermöglicht einen Rückschluss auf die Menge der eingesetzten Template-DNA. Da im Fall von '*Candidatus Entotheonella*' eine reproduzierbare Amplifikation allerdings nur über eine *nested*-PCR möglich ist, konnte ein solcher quantitativer Nachweis nicht durchgeführt werden. Um trotzdem einen raschen, allerdings nur semiquantitativen Eindruck davon zu bekommen, ob die Isolierung der '*Candidatus Entotheonella*' erfolgreich war, bzw. in welcher Konzentration '*Candidatus Entotheonella*' in den aufgereinigten Proben enthalten ist, wurden die gesamten 16S rRNA-Gensequenzen der aufgereinigten Bakterienfraktionen amplifiziert, kloniert und anschließend 20 der so erhaltenen Klone sequenziert (Kapitel 7.4.1, 7.4.2, 7.4.3, 7.4.3.3, 7.4.6.1). Schließlich wurden die so erhaltenen Sequenzen mit der 16S rRNA-Gensequenz von '*E. factor*' und der NCBI Nukleotid-Datenbank verglichen (Tabelle 10-2). Bei keiner der amplifizierten Sequenzen handelte es sich um die 16S rRNA-Gensequenz eines Bakteriums des Kandidatenphylums '*Tectomicrobia*' (Tabelle 10-2). Daher ist davon auszugehen, dass die Menge der '*Candidatus Entotheonella*' im Verhältnis zum restlichen, in der Probe vorhandenen Mikrobiom eher gering ist. Schließlich wurde zur genaueren Quantifizierung von '*Candidatus Entotheonella*' eine sogenannte Amplicon-Sequenzierung durchgeführt (Firma MR DNA, Illumina MiSeq; detaillierte Beschreibung in Kapitel 7.4.6.3). Diese Methode wird häufig zur Analyse verschiedenster Mikrobiome verwendet. Dabei macht man sich den Sequenzaufbau der 16S rRNA, bestehend aus neun variablen Sequenzabschnitten, flankiert von hoch konservierten Bereichen, zunutze. In einer PCR mit universellen 16S rRNA Primern innerhalb der konservierten Bereiche werden zwei 300 bp große Sequenzstücke (variable Regionen V1-V3 und V4) amplifiziert, sequenziert (20000 reads pro Probe) und schließlich anhand von Datenbank-Vergleichen den entsprechenden Phyla, Klassen, Ordnungen, Familien, Genera und Spezies zugeordnet (Kapitel 7.4.6.3).²²² Die Einordnung erfolgte nach folgenden phylogenetischen Grenzen:

Tabelle 5-3: Phylogenetische Klassifizierung.²²³

Übereinstimmung zur Referenzsequenz	Einordnung
>97%	Spezies
97%-95%	Genus
95%-90%	Familie
90%-85%	Ordnung
85%-80%	Klasse
80%-77%	Phylum
<77%	Unbekannt

In Tabelle 10.3 (siehe Anhang) sind die Anteile der einzelnen Phyla am Mikrobiom der untersuchten Proben vor und nach dem Aufreinigungsprotokoll aufgeführt und in den Abbildung 5-31, Abbildung 5-32 und Abbildung 5-33 visualisiert. Dabei ist zu beachten, dass bei der bioinformatischen Auswertung der Sequenzdaten anhand von bestehenden Datenbanken das Kandidatenphylum 'Tectomicrobia' noch nicht beschrieben war. Sequenzen, die '*Candidatus Entotheonella*' zugeordnet werden konnten, wurden dem Phylum *Proteobacteria* zugeschrieben.

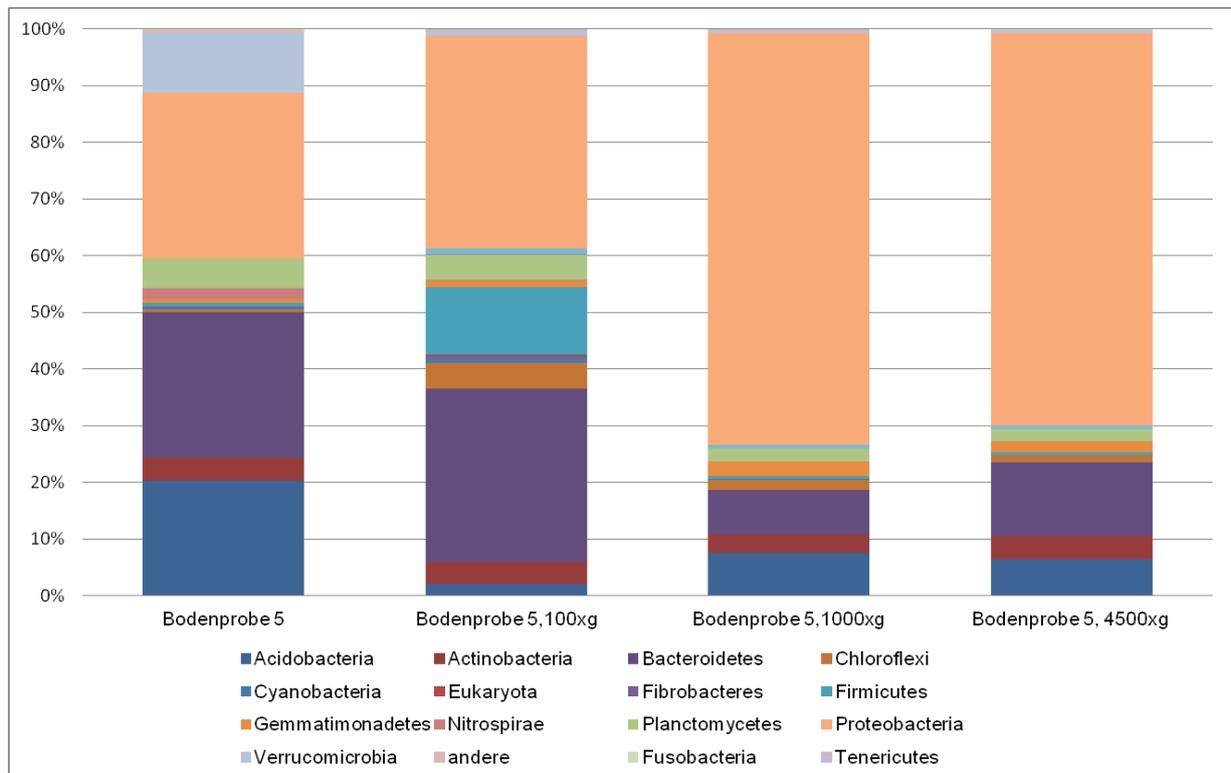


Abbildung 5-31: Zusammensetzung des Mikrobioms der Probe 5 vor und nach der Fraktionierung.

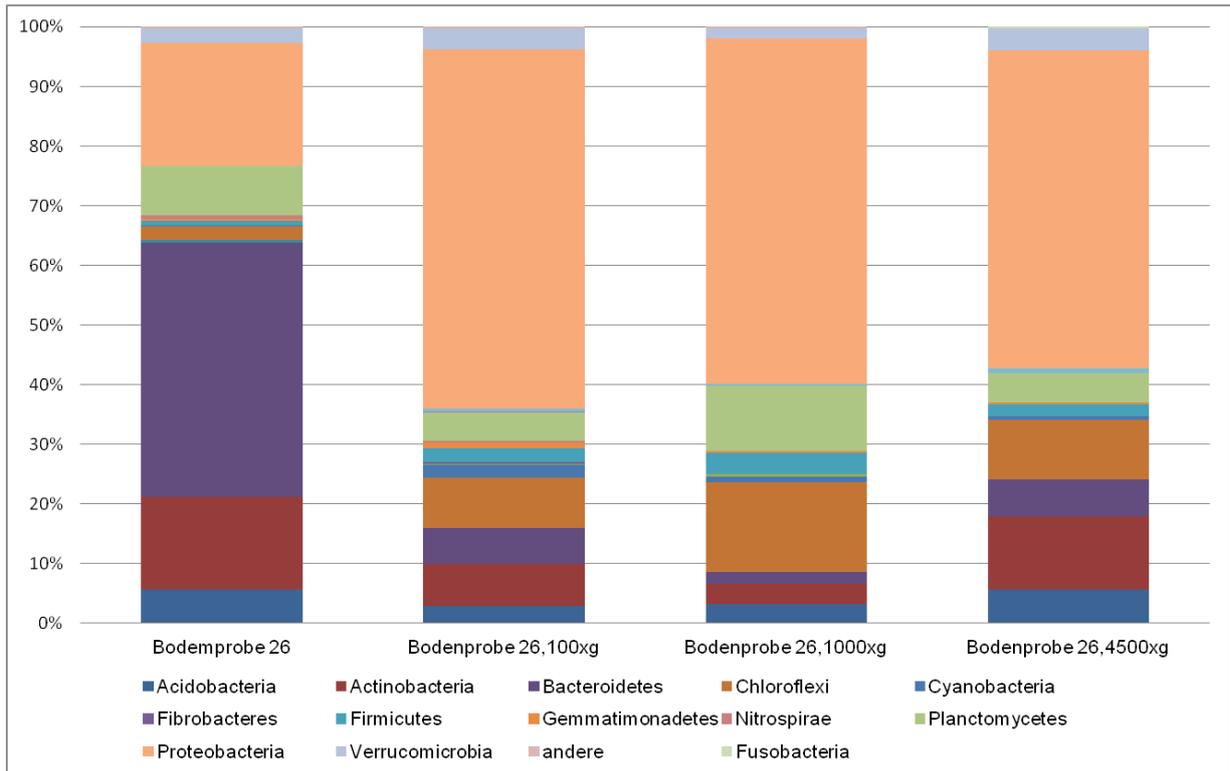


Abbildung 5-32: Zusammensetzung des Mikrobioms der Probe 26 vor und nach der Fraktionierung.

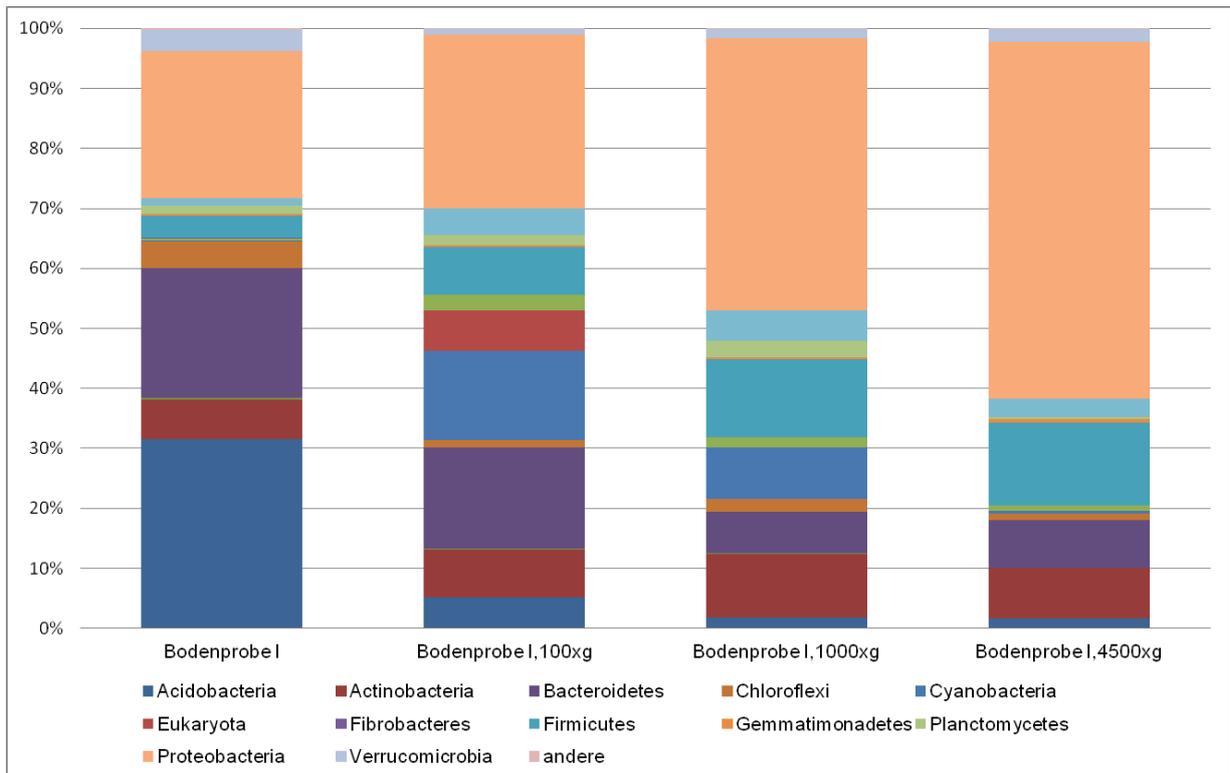


Abbildung 5-33: Zusammensetzung des Mikrobioms der Probe I vor und nach der Fraktionierung.

Allgemeine Studien zur mikrobiellen Zusammensetzung von Böden anhand gepoolter 16S rRNA-Sequenzdaten zeigten, dass Proteobacteria das dominierende Phylum ist, gefolgt von

Acidobacteria und Bacteroidetes. Das Phylum Proteobacteria umfasst eine große Gruppe Gram-negativer Bakterien. Vertreter dieses Phylums konnten aus Boden- und Wasserproben aber auch aus Tieren isoliert werden.⁸⁶ Acidobacteria kommen, anders als der Name impliziert, nicht nur in sauren, sondern in nahezu allen Böden vor. Viele Acidobacteria sind nicht kultivierbar und nur durch ihre 16S rRNA-Sequenzen bekannt.⁸⁶ Bacteroidetes bevölkern nahezu alle Habitate. Sie kommen, wie Proteobacteria, sowohl in Böden, Süß- und Salzwasser, aber auch als Teil der Darmflora in Tieren vor.⁹² Die Säulendiagramme (Abbildung 5-31, Abbildung 5-32, Abbildung 5-33) verdeutlichen die unterschiedlich großen Anteile der einzelnen Bakterien-Phyla am Mikrobiom der Bodenproben. Wie zu erwarten, sind die bereits erwähnten Phyla Proteobacteria, Actinobacteria und Bacteroidetes die vorherrschenden Gruppen. Das dominierende Phylum in Probe I sind Acidobacteria. In Probe 26 hingegen bilden Bacteroidetes die am stärksten vertretene Gruppe von Bakterien. Die Proben unterscheiden sich zwar in den Anteilen der verschiedenen Phyla, generell ist die Zusammensetzung jedoch sehr ähnlich. Auffällig ist, dass sich die Verhältnisse nach der Aufreinigung (Kapitel 7.7) bei allen untersuchten Proben zugunsten der Proteobacteria verschieben, meist zulasten der Bacteroidetes. Der Anteil der Klasse Deltaproteobacteria (Phylum Proteobacteria) am Mikrobiom, zu der, aufgrund der bis zum Zeitpunkt der Analyse fehlenden Beschreibung des Kandidatenphylums 'Tectomicrobia' auch '*Candidatus* Entotheonella' gezählt wurden, liegt vor der Aufreinigung zwischen 1,4% (Probe I) und 7,4% (Probe 5) und konnte durch das Aufreinigungsprozedere bei Probe 26 auf 6% und in Probe I auf knapp 8% gesteigert werden. Betrachtet man die Ebene der Klasse bzw. der Familie Deltaproteobacteria, so ist deren Anteil am Mikrobiom jedoch sehr gering. (zwischen 0% in Probe 5 und 0,05% in Probe I). Auch durch die Aufreinigung konnten die Werte hier nur auf maximal 0,42% in Probe 26 (4500 x g) und 0,93% in Probe I (100 x g) gesteigert werden. Dies lässt bereits erahnen, dass der Erfolg der Aufreinigung, auch aufgrund der vermutlich geringen Ausgangszahl, eher gering ist. Tabelle 5-4 und Abbildung 5-34 geben einen Überblick über die absolute Anzahl und den prozentualen Anteil der *Entotheonella*-spezifischen Sequenzen vor und nach der Isolierung (Kapitel 5.1.4).

Tabelle 5-4: Anzahl der '*Candidatus Entotheonella*' -spezifischen Sequenzen vor und nach der Fraktionierung.

Tabellarisch dargestellt ist die Auswertung der Amplikon-Sequenzierung. Von den vier untersuchten Bodenproben sind jeweils die *absolute* Anzahl der identifizierten '*Candidatus Entotheonella*' -Sequenzen und der *prozentuale* Anteil an der Gesamtzahl der Sequenzen (in Klammern hinter der absoluten Anzahl der '*Candidatus Entotheonella*' -spezifischen Sequenzen) angegeben.

<u>A. Bodenproben vor der Fraktionierung</u>		
Bodenprobe	absolut	prozentual
5	1 (von 93256 Sequenzen)	0.00107
26	2 (von 87165 Sequenzen)	0.00229
I	1 (von 127608 Sequenzen)	0.00078
8	1 (von 74411 Sequenzen)	0.00134
<u>B. Bodenproben nach der Fraktionierung</u>		
Bodenprobe	absolut	prozentual
5 (100xg)	1 (von 51373 Sequenzen)	0.00195
5 (1000xg)	0 (von 19982 Sequenzen)	0
5 (4500xg)	0 (von 50748 Sequenzen)	0
Bodenprobe	absolut	prozentual
26 (100xg)	92 (von 61014 Sequenzen)	0.15079
26 (1000xg)	0 (von 63195 Sequenzen)	0
26 (4500xg)	0 (von 54731 Sequenzen)	0
Bodenprobe	absolut	prozentual
I (100xg)	0 (von 38932 Sequenzen)	0
I (1000xg)	0 (von 62107 Sequenzen)	0
I (4500xg)	0 (von 60113 Sequenzen)	0

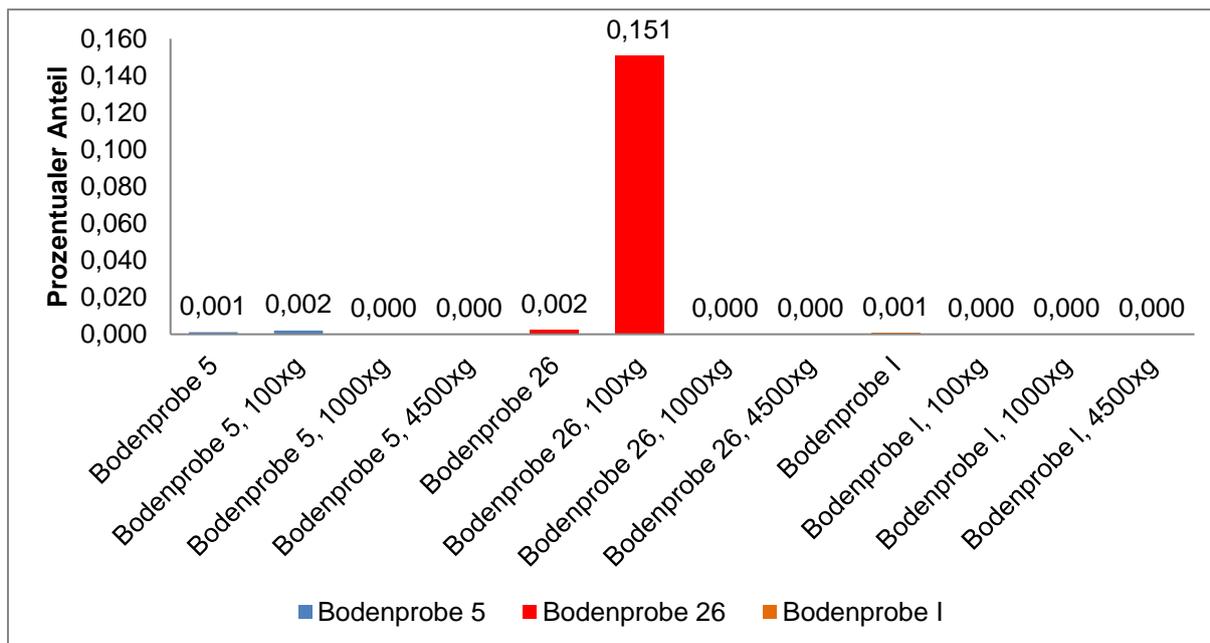


Abbildung 5-34: Prozentualer Anteil von '*Candidatus Entotheonella*' am Mikrobiom.

Die Amplikon-Sequenzierung konnte zeigen, dass der Anteil von '*Candidatus Entotheonella*' am Mikrobiom der untersuchten Bodenproben sehr gering ist. Allerdings ist bei allen gezeigten Ergebnissen der Amplikon-Sequenzierung zu bedenken, dass es hier zu einer systematischen Verzerrung (Bias) der Ergebnisse kommen kann. Dies beruht darauf, dass unterschiedliche Bakterienspezies, aber auch unterschiedliche Stämme einer Spezies eine unterschiedliche Anzahl an Kopien des 16S rRNA-Gens besitzen. Diese schwankt zwischen einer und fünfzehn Kopien.^{222,224,225} Auch die Auswahl der Primer hat Einfluss auf das Ergebnis. Bei der Amplikon-Sequenzierung mittels Illumina ist die Länge der Amplifikate beschränkt. Es kann also nicht die gesamte 16S rRNA, sondern nur ein Abschnitt amplifiziert und analysiert werden. Die verschiedenen variablen Regionen der 16S rRNA eignen sich jedoch unterschiedlich gut zur Abgrenzung verschiedener Spezies voneinander. So kann es je nach verwendeter Primerkombination zu einer Unterrepräsentation oder einer Selektion bestimmter Spezies kommen.^{222,226} Durch das in Kapitel 5.1.4 beschriebene Aufreinigerungsverfahren konnte der prozentuale Anteil von '*Candidatus Entotheonella*' am Mikrobiom in den Proben 5 und I nicht erhöht werden. Nur in Probe 26 konnte eine nennenswerte Erhöhung von '*Candidatus Entotheonella*' (pairwise identity zu '*E. factor*' >97%, Tabelle 5-2) verzeichnet werden. So konnte in Probe 26, 100 x g die Anzahl der '*Candidatus Entotheonella*' 16S rRNA statistisch auf mehr als eine pro 1000 Sequenzen erhöht werden. Daher könnte bei dieser Probe über eine Isolierung der Bakterien über Fluoreszenzmarkierung und anschließender FACS-Sortierung nachgedacht werden. Über den Grund, warum eine Erhöhung der Sequenzfrequenz, zumindest in geringem Maße, bei dieser Probe möglich war, bei allen anderen Proben aber nicht, kann mit dem heutigen

Wissensstand nur spekuliert werden. Zum einen kann dies Folge der oben beschriebenen systematischen Verzerrung sein, zum anderen enthält diese Probe eventuell '*Candidatus Entotheonella*', die sich in ihren Sedimentationseigenschaften so von den Spezies der anderen Proben unterscheiden, dass mit der hier angewendeten Methode eine bessere Aufreinigung möglich ist. Aufschluss über den tatsächlichen Grund würde erst ein Vergleich der Spezies nach erfolgreicher Isolierung und einer anschließenden Untersuchung geben. Ein Vergleich der bei der Amplikon-Sequenzierung gewonnenen 16S rRNA-Gensequenzen von '*Candidatus Entotheonella*' vor der Aufreinigung ist in Abbildung 5-35 gezeigt. Auffällig dabei ist, dass keine dieser Sequenzen die bei einigen Spezies vorkommende Insertion (Kapitel 5.1.2, Abbildung 5-12) aufweist. Eine Analyse der aus den Proben nach der Aufreinigung gewonnenen '*Candidatus Entotheonella*' Sequenzen ergab neben je einer Sequenz in Probe 5-100 x g und Probe I-4500 x g auch 153 Sequenzen in Probe 26-100 x g, von denen 126 Einzelsequenzen sind. Hier enthalten nur vier der Sequenzen die Insertion. In der Mehrzahl der Sequenzen, wie auch in der 16S rRNA-Gensequenz von '*E. factor*' fehlt dieses Sequenzstück.

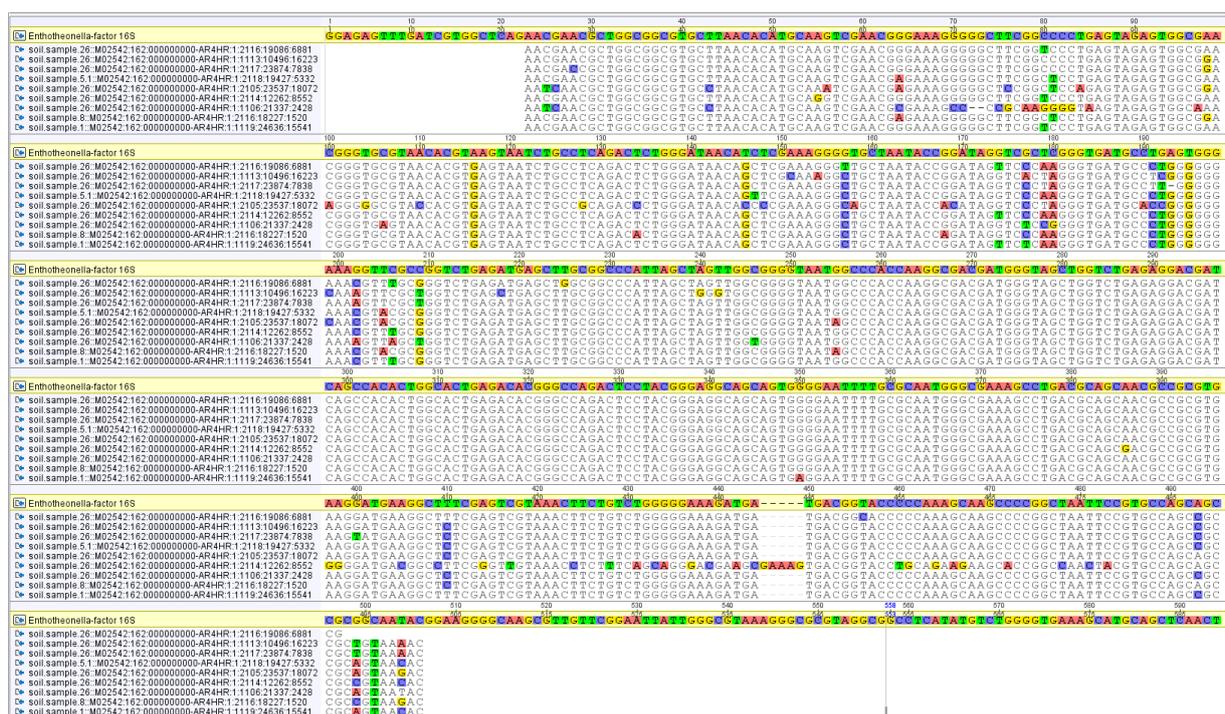


Abbildung 5-35: Sequenzalignment der bei der Amplikon-Sequenzierung gewonnenen '*Candidatus Entotheonella*'-16S-Sequenzen vor der Aufreinigung.

Die Isolierung von '*Entotheonella factor*' aus dem Schwamm *Theonella swinhoei* in vorangegangenen Arbeiten, die anschließende Sequenzierung sowie die, u. a. im Rahmen dieser Arbeit durchgeführte Assemblierung des Genoms konnten *in silico* beweisen, dass

viele der bisher dem Schwamm *Theonella swinhoei* zugeschriebenen Naturstoffe, tatsächlich von diesem außergewöhnlichen Bakterium produziert werden (Kapitel 5.1.1).

Phylogenetische Untersuchungen zeigten, dass 'E. factor' zu einem neuen, bisher unbekanntem Phylum '*Candidatus* Tectomicrobia' gehört (Kapitel 5.1.1).¹³¹ Die Analyse von insgesamt 640 16S rRNA-Gensequenzen (78 Sequenzen aus Bodenanalysen dieser Arbeit und 562 Sequenzen aus einer Datenbankrecherche) demonstriert, dass sich das Kandidatenphylum 'Tectomicrobia' aus verschiedensten Phylotypen, die weltweit in unterschiedlichsten Habitaten verbreitet sind, zusammensetzt (Kapitel 5.1.2 und 5.1.3; Abbildung 5-17 und Abbildung 5-18). Das Phylum 'Tectomicrobia' kann in verschiedene Kladen unterteilt werden. Bei einigen Kladen ist eine Korrelation zwischen Umgebungsfaktoren der Habitate und der Phylogenie zu beobachten (z. B. Assoziation mit Schwämmen, salinen oder thermalen Habitaten). Der hohe verwandtschaftliche Grad einiger aus Bodenproben isolierter 16S rRNA Gensequenzen zu der 16S rRNA-Gensequenz von 'Entotheonella factor' gibt Grund zu der Annahme, dass auch diese Spezies pharmakologisch relevante Naturstoffe produzieren könnten.

Quantitative Analysen zeigten, dass '*Candidatus* Entotheonella' nur einen geringen Anteil am gesamten Mikrobiom der untersuchten Bodenproben einnehmen (maximal 0,15% in Probe 26 nach Fraktionierung; Kapitel 5.1.4). Dies macht eine Isolierung dieser Spezies schwierig. Eine zu prüfende Möglichkeit könnte die Markierung dieser Spezies durch fluoreszenzmarkierte Sonden und die anschließende Isolierung mittels *Fluorescence activated cell sorting* (FACS) sein. Ein erster Versuch, einzelne Kandidaten in aufgereinigten Zellfraktionen mittels '*Candidatus* Entotheonella' spezifischer Sonde im Rahmen einer CARD-FISH zu markieren schien erfolgversprechend, konnte aber bislang nicht reproduziert werden. Eine Markierung und anschließende Isolierung mittels FACS könnte eventuell bessere und reproduzierbare Ergebnisse erzeugen, da hier eine große Probenmenge untersucht und '*Candidatus* Entotheonella' so trotz ihres geringen Anteils am Mikrobiom isoliert werden könnten. Eine Möglichkeit zur anschließenden Kultivierung und Charakterisierung dieser bislang unbekanntem Phylotypen, könnte der Einsatz des sogenannten iChips sein. Dieser Chip besteht aus vielen kleinen Bohrungen. Taucht man ihn in eine Bakteriensuspension (z. B. aus einer Bodenprobe) in flüssigem Agar, befindet sich statistisch gesehen in jeder dieser Bohrungen jeweils ein Bakterium.⁹⁹ „Verschlossen“ werden diese Bohrungen mithilfe einer semipermeablen Membran, die den Austausch von Nährstoffen und Wachstumsfaktoren ermöglicht, eine Kontamination aber verhindert. Der Chip kann dann in das ursprüngliche Habitat der Probe verbracht werden. Unter anderen Umständen unkultivierbare Symbionten, für die der Austausch von Metaboliten mit ihren Interaktionspartnern lebensnotwendig ist, können so u. U. doch in Reinkultur kultiviert werden. Der Anteil kultivierbarer Bakterien aus Bodenproben kann so von 1% (bei herkömmlicher

Kultivierung) auf 50% gesteigert werden.⁹⁹ Erste Erfolge konnten bereits bei der Isolierung des Teixobactin-Produzenten *Eleftheria terrae*, einem β -Proteobakterium, erzielt werden.¹⁰⁰

5.2 Symbiont 2 des Schwammes *Theonella swinhoei*

Während der Analysen des Mikrobioms des Schwammes *Theonella swinhoei* konnte ein weiteres filamentöses Bakterium identifiziert werden, das vermutlich dem Phylum Chloroflexi angehört (Dr. Tetsushi Mori, Prof. Dr. Haruko Takeyama; Kapitel 3.3.1). Nach Einzelzellsequenzierung (Dr. Tetsushi Mori, Prof. Dr. Haruko Takeyama) konnten die Genomdaten auf Naturstoffgencluster hin analysiert werden. Diese Analyse wurde mithilfe der Onlinesoftware antiSMASH (antibiotic & Secondary Metabolite Analysis SHell) durchgeführt.⁸⁴ Dazu wurden die bereits assemblierten Contigs (4829 Contigs aus ca. 43553 Einzelsequenzen) hochgeladen und analysiert. Insgesamt konnten so 40 Contigs mit Fragmenten von Naturstoffclustern identifiziert werden, davon 32 PKS und NRPS Gencluster, auf denen in dieser Arbeit der Fokus liegen soll. Detaillierte Ergebnisse der AntiSMASH-Analyse sind in Tabelle 5-5 aufgeführt.

Tabelle 5-5: AntiSMASH-Analyse des Symbiont 2 Genoms.

AT=Acyltransferase; **DH**=Dehydratase; **ER**=Enoylreduktase; **EH**=Enoyl-CoA-Hydratase; **MT**=Methyltransferase; **KR**=Ketoreduktase; **KS**=Ketosynthase; **TE**=Thioesterase. Die jeweiligen Aminosäuren sind im Dreibuchstabencode angegeben. **NRP**=Nichtribosomales Peptid (hier konnte aufgrund der unterschiedlichen Vorhersagen zur A-Domänen-Spezifität kein Konsensus gefunden werden).

Naturstoffcluster	Domänen	ORF	Contig	NRSPredictor2	Stachelhaus-code	Minowa	Consensus
				Rausch, C., <i>et al.</i> , 2005, Nucleic Acids Res. 33: 5799-5808 Röttig, M., <i>et al.</i> , 2011, Nucleic Acids Res. 39: W362-367.	Stachelhaus, T., Mootz, H. D. & Marahiel, M. A., Chem.Biol. 6, 493–505 (1999).	Minowa, Y., <i>et al.</i> , 2007, J. Mol. Biol. 368:1500-1517.	
NRPS-Typ1 PKS	A, nMT KR, ACP ER AT, DH KS A, PCP	59_4 59_5 59_6 59_7 59_12 59_13	59	Ile	Ile	Ile	Ile
				hydrophob-aliphatisch		Arg	NRP
NRPS	PCP C, A, PCP	104_5 104_6	104	hydrophob-aliphatisch	Ser	Ser	Ser
Trans-AT-PKS-NRPS	KS, ACP, trans-AT docking ACP ECH PCP, E, A	148_1 148_2 148_6 148_8 148_9	148			Orn	Orn
							Orn
NRPS	A, PCP, TE C C, A, PCP A nMT	251_1 251_2 251_3 251_4 251_6	251	Orn, Lys, Arg Asp, Asn, Glu, Gln, Aad Ala	Orn Lys Cys	Ala Arg Ala	NRP NRP Ala

Ergebnisse und Diskussion

NRPS-Typl PKS	C	303_3	303	Asp, Asn, Glu, Gln, Aad	Ser	Dab	NRP
	PCP	303_4					
	C, A	303_5					
	C	303_6					
	KS, AT, ACP, C, A, PCP	303_7		Asp, Asn, Glu, Gln, Aad	Ser	Ser	Ser
NRPS	C, A	368_3	368	Ser	Ser	Ser	Ser
	PCP, E	368_4					
	E	368_5					
		368_6					
NRPS-Typl PKS	PCP, C, A	435_1	435	hydrophil Asp	Phe Asn	Arg Asp	NRP Asp
	C, A	435_2					
	C	435_3		Asp	Asn	Asp	Asp
	C, A, PCP, E	435_4					
	C	435_8					
	PCP	435_9					
Typ I PKS	ACP KS, AT FkbH, PCP		490				
Typl PKS-NRPS	KS	538_9	538	Val, Leu, Ile, Abu, Iva	Gua	Guanidin Essig- säure	NRP
	ER	538_10					
	A, PCP, KS	538_11					
	AT, DH	538_12					
	KR, ACP	538_13					
	KS	538_14					
NRPS	FkbH	606_1	606	hydrophob-aliphatisch	Orn	Glu	NRP
	A, ACP, C	606_3					
NRPS	C, PCP, TE	637_2	637	Val, Leu, Ile, Abu, Iva	His	His	His
	A, PCP	637_3					
NRPS	PCP	786_3	786	Asp, Asn, Glu, Gln, Aad	Ser	Ser/Dab	Ser/NRP
	A, PCP, C, A	786_4					
	PCP, C	786_5					
	A	786_6					
NRPS	C, A, PCP, E, TE		1016	Ala	Ala	Ala	Ala
NRPS	A, PCP	1121_1	1121	Asn	Asn	Asn	Asn
	C	1121_2					
	A	1121_4					
			hydrophob-aliphatisch	Leu	Lys	NRP	
NRPS	C, A, PCP, C	1266_1	1266	N/A Ser/Asp, Asn, Glu,	Phe Ser/Ser	Ser Ser/Ser	NRP Ser/Ser
	C, A, PCP, C, A	1266_2					
	PCP	1266_3		Gln, Add, Cys	Cys	Cys	Cys
	A, PCP						
Typ I PKS-NRPS	ER KR, ACP KS, AT cMT, KR, ACP C, A	1267_1 1267_2 1267_3 1267_4	1267	Cys	Cys	Cys	Cys
NRPS	KR, DH	1452_1	1452	Val	Ile	Ile	Ile
	PCP, E	1452_2					
		1452_3					
	C, A, PCP, C	1452_4					
NRPS	C, A, PCP		1706	Val	Val	Val	Val

Ergebnisse und Diskussion

NRPS	A, C		2168	hydrophob-aliphatisch	Leu	Lys	NRP
NRPS	KS C, A, PCP		2190	Asp, Asn, Glu, Gln, Aad	Glu	Leu	NRP
NRPS	C, PCP A	2207_1 -rev	2207	N/A	N/A	Ser	NRP
	A	2207_2 -rev 2207_3 -rev 2207_5		N/A	Ser	Dab	NRP
NRPS	A PCP, C	2293-1 2293-2	2293	Asp, Asn, Glu, Gln, Aad	Ser	Ser	Ser
NRPS	A, PCP, C	2558_1	2558	Gly	Gly	Gly	Gly
NRPS	C, A	2878_1	2879	Gln	Gln	Gln	Gln
NRPS	A PCP C	2898_2 2898_3 2898_4	2899	Asp, Asn, Glu, Gln, Aad	Phe	Arg	NRP
NRPS	C, A	2899_2	2900	Gly	Gly	Gly	Gly
NRPS	PCP, C	3943_1	3543				
NRPS	C	3981_1	3981				
Trans-AT PKS	DH, KR, ACP, KS, AT-docking DH, ACP	1471_1 1471_2	1471				
Trans-AT PKS	KR KS_At-docking ACP	1496_1 1496_2 1496_3	1496				
Trans-AT-PKS	KS trans-AT docking? KR ACP KS	2468_1 2468_2 2468_5	2468				
Typ I PKS-NRPS			781				
Terpene			197				
Terpene			394				
Terpene			3265				
Lantipeptide			654				
Cyanobactin			663				
Cyanobactin			1023				

Ergebnisse und Diskussion

Cyanobactin	3547
Bacteriocin	1660
andere	1909
andere	1492
andere	3574

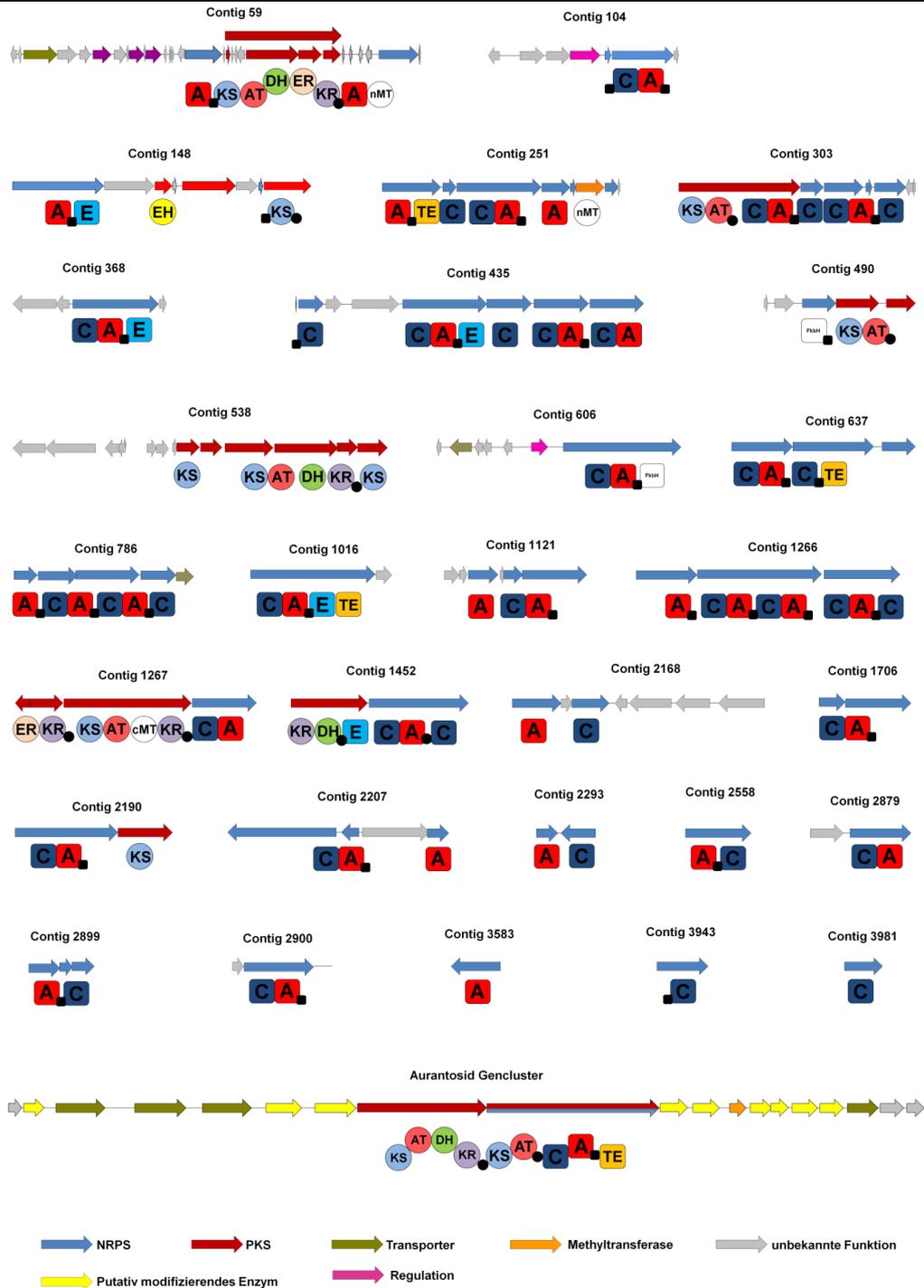


Abbildung 5-36: PKS und NRPS-Cluster Contigs des *Symbiont 2* Genoms. Dargestellt sind die Contigs des Bakteriums *Symbiont 2* aus dem Schwamm *Theonella swinhoei*, die Biosynthesegencluster von Polyketiden und nichtribosomalen Peptiden enthalten. Die *open reading frames* (ORF) wurden anhand der vorhergesagten

Funktion farblich markiert. PKS-Domänen (Kreise) sind: **AT**=Acyltransferase; **DH**=Dehydratase; **ER**=Enoylreduktase; **EH**=Enoyl-CoA-Hydratase; **MT**=Methyltransferase; **KR**=Ketoreduktase; **KS**=Ketosynthase; **TE**=Thioesterase. Die Acylcarrierproteine sind als kleine schwarze Kugelndargestellt. NRPS-Domänen (Vierecke) sind: **A**=Adenylierung; **C**=Kondensation; **E**= Epimerase; **TE**=Thioesterase. Die Peptidcarrierproteine sind als kleine schwarze Vierecke dargestellt. Die bioinformatisch prognostizierte Spezifität der A-Domänen ist unter der jeweiligen Domäne im im Dreibuchstabencode angegeben. Konnte keine eindeutige Vorhersage getroffen werden, so ist diese Domäne mit einem „?“ versehen (vgl. Tabelle 5-5). Der Aurantosid-Gencluster konnte von der AG Piel an der ETH Zürich assembliert werden. An dem in der AntiSMASH-Analyse (Tabelle 5-5) ebenfalls genannte *trans*-AT PKS-Gencluster wurde bzw. wird ebenfalls in der AG Piel an der ETH Zürich gearbeitet. Daher wird dieses in der obigen Abbildung nicht gezeigt.

Bei näherer Analyse der als „andere“ Synthesecuster in einer ersten antiSMASH-Analyse annotierten Contigs, zeigte sich, dass auch diese z. T. NRPS-Domänen enthalten.

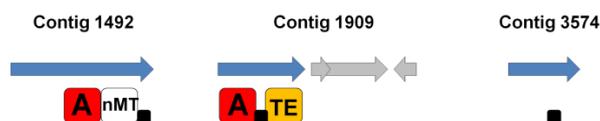


Abbildung 5-37: Weitere annotierte NRPS-Domänen einiger *Symbiont 2*-Contigs.

Die Domänenstruktur der Genclusterfragmente lässt darauf schließen, dass es sich neben dem Aurantosid und einem *trans*-AT-PKS-Gencluster um mindestens vier weitere NRPS/PKS-Gencluster handeln muss. Vier Fragmente enthalten eine Thioesterase-Domäne, die, wie in Kapitel 3.1.1 beschrieben, das Molekül vom Multienzymkomplex abspaltet und so ein terminales Modul bilden. Ein Initiationsmodul konnte lediglich bei einem Fragment eindeutig identifiziert werden. Hier sind im Randbereich weitere Gene zu finden, die eindeutig nicht dem Gencluster zugeschrieben werden können.

Die Read-Assemblierung der Gencluster-Contigs mussten zunächst noch einmal händig korrigiert werden, da die Sequenzen in den Überlappungsbereichen häufig ungenau waren. Die NRPS- und PKS-Cluster sollten dann durch kombinatorische PCR und anschließende Sequenzierung der PCR-Produkte miteinander verbunden werden. Anhand der assemblierten Cluster kann dann auf die von ihnen produzierten Naturstoffe geschlossen werden (Kolinearitätsregel, Kapitel 3.1.1.1). Dazu wurden, wie auch beim Lückenschluss des 'Entotheonella factor'-Genoms, Primer an die jeweiligen Enden der Contigs gelegt, um über kombinatorische PCR die Sequenzlücken zwischen zwei Contigs zu überbrücken (Tabelle 7-16, Kapitel 7.4.1.1, Abbildung 5-4). Bei 'E. factor' konnten zur Vorhersage potentiell zusammengehöriger Contigs die Molekülstruktur bereits aus dem Schwamm *Theonella swinhoei* isolierter Naturstoffe herangezogen werden. Aufgrund der Kolinearitätsregel konnte davon auf die Clusterarchitektur zurück geschlossen werden. Diese Möglichkeit bestand beim *Symbiont 2* nicht, da, mit Ausnahme von Aurantosid keine zusätzlichen Naturstoffe aus dem Schwamm isoliert und identifiziert werden konnten. Zunächst wurde deshalb untersucht, ob an den Enden der Contigs unvollständige Domänensequenzen liegen, anhand derer eine Vorhersage getroffen werden könnte. Da dies nicht der Fall war, konnten nur wenige

Contigkombinationen ausgeschlossen werden und es mussten viele Primerkombinationen ausgetestet werden. Nach Erstellung der Primer, wurde die PCR, wie in Kapitel 7.4.1 bereits für 'E. factor' beschrieben, mit dem KAPA2G Fast PCR Kit der Firma *KAPABiosystems* durchgeführt. Als Templat diente 1 µl einer 1:10 Verdünnung des aus dem Schwamm *Theonella swinhoei* aufgereinigten Zellpellets. Leider gestaltete sich das Verbinden der Contigs durch kombinatorische PCR, wie auch schon bei 'E. factor', schwierig. Oft ließen sich einzelne, spezifische PCR-Produkte amplifizieren, beim anschließenden Assemblieren konnte das Produkt jedoch nicht oder nur mit einem der beiden Contigs verbunden werden. Dies könnte u. a. mit dem genomisch komplexen Templat in Zusammenhang stehen. Dieses enthält neben dem untersuchten Genom noch mindestens zwei weitere Genome ('*Candidatus Entotheonella*'), die ebenfalls eine Vielzahl an Naturstoffgenclustern enthalten und so durch konservierte Sequenzen einzelner Domänen ebenfalls als Templat dienen könnten. Außerdem könnte es bei der Assemblierung der Reads bzw. der Rohdaten zu Fehlpaarungen gekommen sein. Dies ist häufig dann der Fall, wenn diese Stelle nur durch die Sequenz eines einzigen Reads abgedeckt wird bzw. es keinen oder nur einen kleinen Überlappungsbereich zum Nachbarread gibt. In einem solchen Fall ist das Ergebnis der Assemblierung in diesem Bereich nicht zuverlässig. Auch die bereits erwähnte fehlende Information zur Molekülstruktur der Produkte (Polyketide/nichtribosomale Peptide) erschwert eine Vorhersage der Contigkombinationen und damit die Verknüpfung der einzelnen Contigs mittels kombinatorischer PCR. Daher konnte mit dieser Methode aus den genannten Gründen und aufgrund zeitlicher Limitierung nur zwei Contigs nachweislich verbunden werden (Abbildung 5-38).

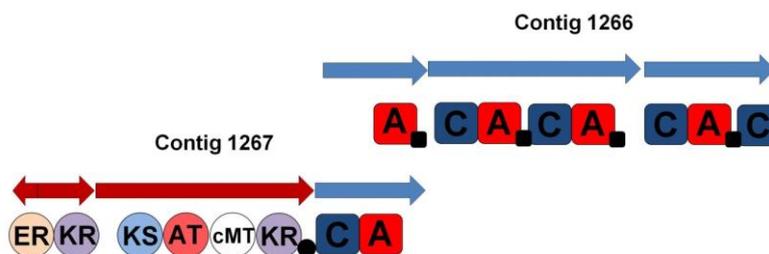


Abbildung 5-38: Assemblierte Cluster-Contigs.

Neben PKS und NRPS enthält das Genom auch Gencluster-Fragmente von Terpenen und Cyanobactinen, sowie je ein Gencluster-Fragment eines Lantibiotikums und eines weiteren Bacteriocins. Terpene sind eine sehr heterogene Gruppe sekundärer Naturstoffe, die sich formal vom Kohlenstoffgerüst des Isoprens ableiten. Neben wichtigen biologischen Funktionen wie z. B. als Membranbaustein (Sterole) oder als Steroidhormon weisen sie z. T. auch pharmakologische, wie z. B. antibakterielle Eigenschaften auf.²²⁷ Unter dem Oberbegriff Bacteriocine werden verschiedene, von Bakterien produzierte Proteine zusammengefasst. Mithilfe von Bacteriocinen können Bakterien nah verwandte Spezies inhibiert oder abgetötet.

Im Vergleich zu Antibiotika ist das Aktivitätsspektrum jedoch gering.⁸⁶ Zu dieser Gruppe proteinöser Toxine gehören auch Cyanobactine. Diese kleinen, zyklischen Peptide werden von verschiedenen Cyanobakterien produziert und haben unterschiedliche pharmakologische Wirkungen (z. B. antitumoral, anti-Malaria). Daher dienen sie häufig als Leitstrukturen neuer Medikamente.²²⁸ Lantibiotika, die ebenfalls zu den Bacteriocinen zählen, enthält die namensgebende Aminosäuren Lanthionin bzw. Methyllanthionin und werden neben der Behandlung von Infektionen auch als Konservierungsmittel eingesetzt.²²⁹

Im Rahmen dieser Arbeit wurden die Genomdaten eines weiteren filamentösen *Theonella swinhoei*-Symbionten auf Naturstoffcluster hin untersucht. Dabei wurden 40 Contigs identifiziert, die Fragmente von Naturstoffclustern enthalten (Tabelle 5-5). 32 dieser Contigs beinhalten Fragmente von PKS und NRPS Gencluster (Abbildung 5-36). Eine Analyse der Genclusterfragmente lässt darauf schließen, dass es sich neben dem Aurantosid und einem *trans*-AT-PKS-Gencluster um mindestens vier weitere NRPS/PKS-Gencluster handeln muss. Die Assemblierung der Contigs gestaltete sich schwieriger als bei den '*Candidatus Entotheonella*' Symbionten, da mit Ausnahme von Aurantosid und der bereits 'E. factor' zugeordneten Naturstoffe keine weiteren Naturstoffstrukturen aus *Theonella swinhoei* bekannt waren. So konnten trotz diverser Primerkombinationen in einer Vielzahl von PCR-Ansätzen lediglich zwei der Contigs miteinander verbunden werden. Um die verbleibenden Lücken schließen zu können, sollte in zukünftigen PCR-Ansätzen als Templat aufgereinigte DNA des Symbionten 2 verwendet werden. Metagenomische DNA mehrerer Symbionten sollte vermieden werden. Dieser Ansatz konnte im Rahmen dieser Arbeit aus zeitlichen Gründen nicht weiterverfolgt werden. Neben PKS und NRPS enthält das Genom auch Genfragmente von Terpenen und Cyanobactinen, sowie je ein Genfragment eines Lantibiotikums und eines weiteren Bacteriocins.

Zusammenfassend zeigen diese Ergebnisse, dass auch dieser *Theonella swinhoei*-Symbiont enormes Potential im Hinblick auf die Identifizierung und Charakterisierung neuer bioaktive Naturstoffe besitzt.

5.3 '*Candidatus Pseudomonas paederii*'

5.3.1 Schließen der Lücken des '*Pseudomonas paederii*' Genoms

Wie in Kapitel 3.3.2 beschrieben, konnte das Genom von '*P. paederii*', einem Symbionten aus Käfern der Gattung *Paederus*, bereits in vorangegangenen Arbeiten sequenziert werden.^{156,157} Die Sequenzierdaten konnten zu 100 Contigs assembliert werden.¹⁵⁷ Diese Contigs sollten nun durch Schließen der Sequenzlücken zusammengefügt werden. Dazu wurden zunächst möglichst viele Contigs eines zusätzlich vorhandenen Plasmids

herausgefiltert werden, da diese aufgrund der Sequenzvariabilität des Plasmids nur schwer zu assemblieren sind. Dies gestaltete sich im Vergleich zum 'Entotheonella factor'-Genom schwierig. Bei diesem konnte Dr. Micheal Wilson auf die Rohdaten zurückgreifen und anhand der Coverage (engl.: mittlere Abdeckung) der Contigs auf einen genomischen oder einen Plasmid-Ursprung schließen (Kapitel 5.1.1). Dies war hier nicht möglich, da keine Rohdaten zur Verfügung standen. Daher wurde versucht, einige Contigs anhand Plasmid-spezifischer Charakteristika auszuschließen. Zunächst wurde der GC-Gehalt aller Contigs untersucht, da dieser bei Plasmiden oft niedriger ist als im restlichen Genom.²³⁰ Allerdings konnten hier keine Contigs identifiziert werden, deren GC-Gehalt signifikant geringer war (mindestens 2 %).²³⁰ Des Weiteren wurden die Contigs auf Plasmid-spezifischen Genen wie z. B. *tra*-Gene eines F-Plasmids, Resistenz-/Virulenzgene oder Colicin-Gene durchsucht. Aber auch dadurch konnten nur wenige Contigs ausgeschlossen werden. So konnte die Menge an möglichen Contig-Kombinationen kaum reduziert werden. Zur Vorhersage benachbarter Contigs kamen verschiedene bioinformatische Strategien zur Anwendung: Zunächst wurde ein simples Assemblieren aller Contigs durchgeführt (Geneious 7.1.9). Dadurch konnten bereits einige Contigs verbunden werden. Experimentell wurden diese Ergebnisse mittels PCR (Kapitel 7.4.1), Gelextraktion der PCR-Produkte (Kapitel 7.4.2.1) und anschließender Sequenzierung (Kapitel 7.4.6.1) bestätigt. Die so verbundenen Contigs sind in Abbildung 5-39, Abbildung 5-40, Abbildung 5-41 dargestellt.

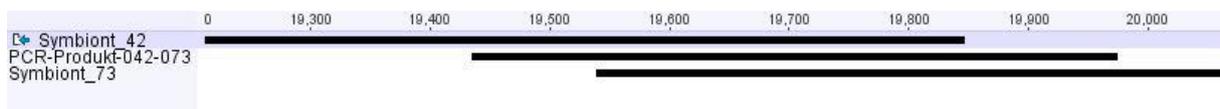


Abbildung 5-39: Assemblierung der Contigs 42 und 73 'P. paederii' Genoms.

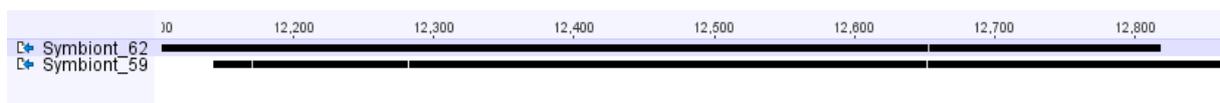


Abbildung 5-40: Assemblierung der Contigs 62 und 59 'P. paederii' Genoms.

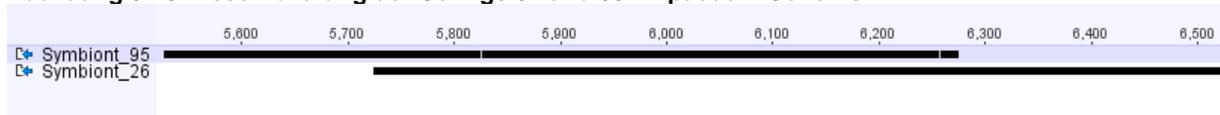


Abbildung 5-41: Assemblierung der Contigs 95 und 26 'P. paederii' Genoms.

In einem weiteren Ansatz wurde, analog zur Vorgehensweise bei der Genomassemblierung von 'E. factor', versucht Gene zu identifizieren, deren Sequenzfragmente an den Enden zweier Contigs zu finden sind (vgl. Kapitel 5.1.1 und 5.2). Diese können einen Hinweis darauf geben, dass es sich um zwei benachbarte Contigs handeln könnte. Dazu wurden zunächst die ersten und letzten 500 bp der Contigsequenzen extrahiert und mit diesen Sequenzfragmenten eine Nukleotid BlastX-Analyse gegen den nächsten Verwandten *P.*

aeruginosa durchgeführt. Die daraus resultierenden Top-Hits wurden dann analysiert. Deckt das jeweilige Genfragment nur einen Teil der Nukleotidsequenz des BlastX-Hits ab, wurde eine tBlastN-Analyse gegen die restlichen Contigs durchgeführt. Die so verbundenen Contigs sind in Abbildung 5-42 und Abbildung 5-43 dargestellt.

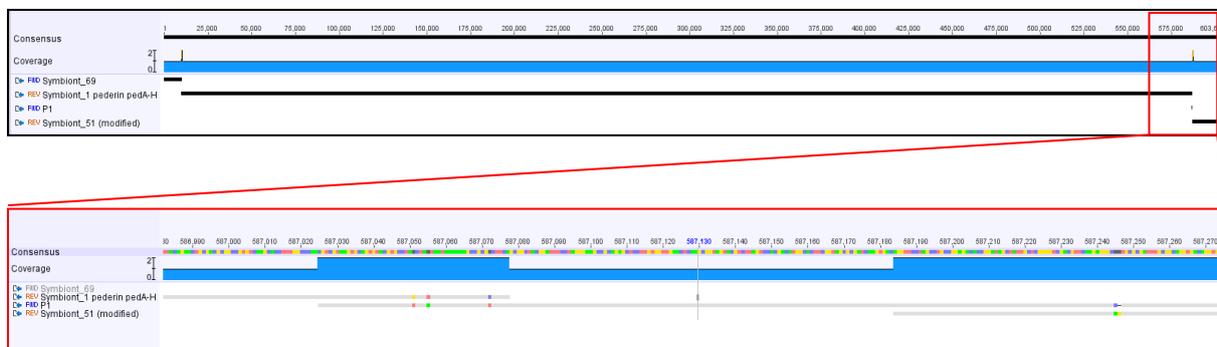


Abbildung 5-42: Assemblierung der Contigs 69, 1 und 51 des 'P. paederii' Genoms. Dargestellt ist die Assemblierung von drei Contigs (Symbiont_69, Symbiont_1 und Symbiont_51). Die Lücke zwischen dem mittleren Contig (Symbiont_1) und rechten Contig (Symbiont_51) wurden durch kombinatorische PCR (Abbildung 5-4) und Sequenzierung des resultierenden PCR Produkts (P1) geschlossen (vergrößerter Bereich).

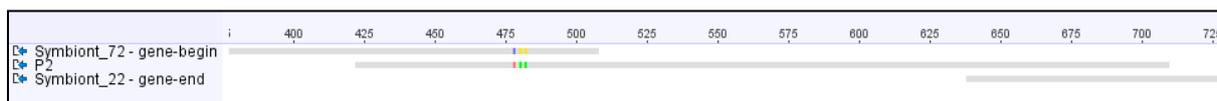


Abbildung 5-43: Assemblierung der Contigs 72 und 22 des 'P. paederii' Genoms.

5.3.2 Analyse eines NRPS/PKS Hybriden

Wie in Kapitel 3.3.2 beschrieben, beinhaltet das 'P. paederii' Genom neben dem bekannten Gencluster von Pederin ein weiteres PKS-NRPS-Gencluster, dessen Produkt bisher unbekannt ist (Abbildung 3-21). Zur Vorhersage der Struktur des gesuchten Metaboliten ist es notwendig, die Spezifität der A-Domänen zu kennen, die für die Auswahl der Molekülbausteine verantwortlich ist (Kapitel 3.1.2.1). Im Rahmen dieser Arbeit sollte daher die zuvor getroffene bioinformatische Vorhersage der Spezifität der beiden A-Domänen experimentell mithilfe eines $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assays (Kapitel 7.5.8) überprüft werden.⁶⁴

Für die Aktivität einiger A-Domänen ist das Vorhandensein von MbtH-artigen Proteinen unabdingbar (Kapitel 3.1.2.2). Diese kleinen Proteine werden oft im oder in direkter Nachbarschaft zum NRPS-Cluster codiert.^{67,69–72,75,76} Um festzustellen, ob ein solches Protein auch bei der Expression der A-Domänen dieses NRPS-Clusters eine Rolle spielen könnte, wurden die Sequenzen des MbtH-artigen Proteins HrmR aus der Hormaomycin-Biosynthese gegen das 'P. paederii'-Genom geblastet.²³¹ Dabei konnte kein MbtH-artiges Protein innerhalb des Genoms detektiert werden. Eine Coexpression mit einem solchen Protein schien daher zunächst nicht notwendig.

Zur Expression musste zunächst der codierende Bereich der beiden A-Domänen innerhalb des Genclusters über ein Sequenzalignment bestimmt werden (Anhang, Abbildung 10-9). Während der Amplifikation der Gensequenzen (Kapitel 7.4.1) wurden mittels modifizierter, spezifischer Primer zusätzliche Schnittstellen (*EcoRI* und *HindIII*) an den Enden der PCR-Produkte angefügt. Nach Überprüfung der PCR-Produkte durch Sequenzierung (Kapitel 7.4.6.1) wurden diese dann über die Restriktionsschnittstellen *EcoRI* und *HindIII* in den Expressionsvektor pHIS8²³² (Kapitel 7.1.2, Abbildung 10-11) kloniert (Kapitel 7.4.3). Die so entstandenen Konstrukte (pPppB-A und pPppE-A) sollten dann in den Expressionsstamm *E. coli* BL21 (DE3) transformiert werden. Die Ligation in den Expressionsvektor gestaltete sich - auch nach vorhergehenden Subklonierungen- schwierig. So konnten nur das PCR-Produkt der zweiten A-Domäne (ORF pppE) erfolgreich in den pHIS8-Vektor ligiert und *E. coli* BL21 (DE3) mit dem so entstandenen Plasmid transformiert werden. Die klonierte Sequenz wurde erneut mittels Sanger-Sequenzierung überprüft (Kapitel 7.4.6.1).

Das Protein wurde dann in *E. coli* BL21 (DE3) in einer ersten Testexpression überexprimiert (Kapitel 7.5.2) und über den N-terminalen Octahistidin-Tag mithilfe von Ni-NTA-Affinitätschromatographie aufgereinigt (Kapitel 7.5.4). Die Überprüfung der Expression erfolgte über eine SDS-Gelelektrophorese (Kapitel 7.5.6; Abbildung 5-44).

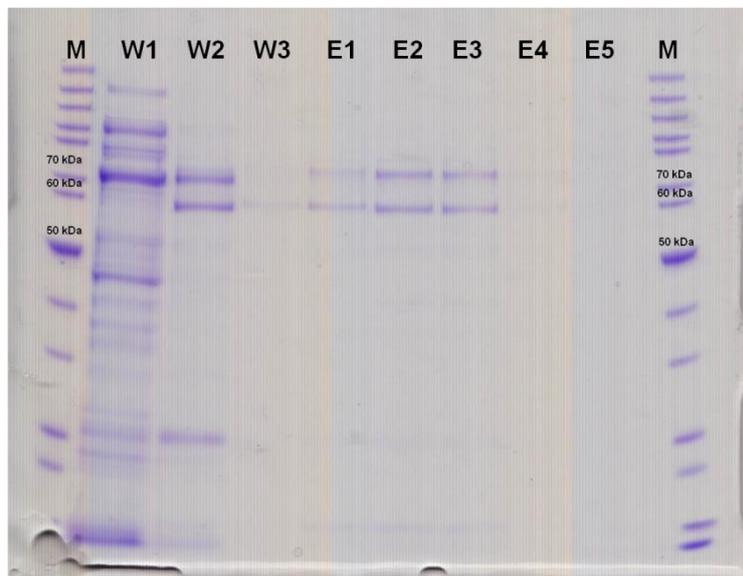


Abbildung 5-44: Testexpression der PppE-A-Domäne. M: PageRuler unstained Protein Ladder (*Thermo Fisher Scientific*); W1-W3: Waschfraktionen 1-3; E1-E5: Elutionsfraktionen 1-5.

Die kalkulierte Größe des exprimierten Proteins liegt bei 66 kDa. Wie in Abbildung 5-45 zu erkennen, werden offensichtlich zwei Proteine unterschiedlicher Größe aufgereinigt. Die Größen der Banden liegen bei etwa 60 kDa und etwas über 70 kDa. Zur Durchführung des Spezifitätsassays wurde ein größerer präparativer Expressionsansatz (2x 500 ml in TB-Medium (Kapitel 7.2)) durchgeführt, um genügend Protein aufzureinigen zu können (Kapitel

7.5.2ff.). Die Expression wurde durch Zugabe von 0,5 mM IPTG induziert und die Kultur über Nacht bei 16 °C inkubiert. Die Volumina des Lyse- und der Waschpuffer wurden entsprechend angepasst (20 ml Puffer).

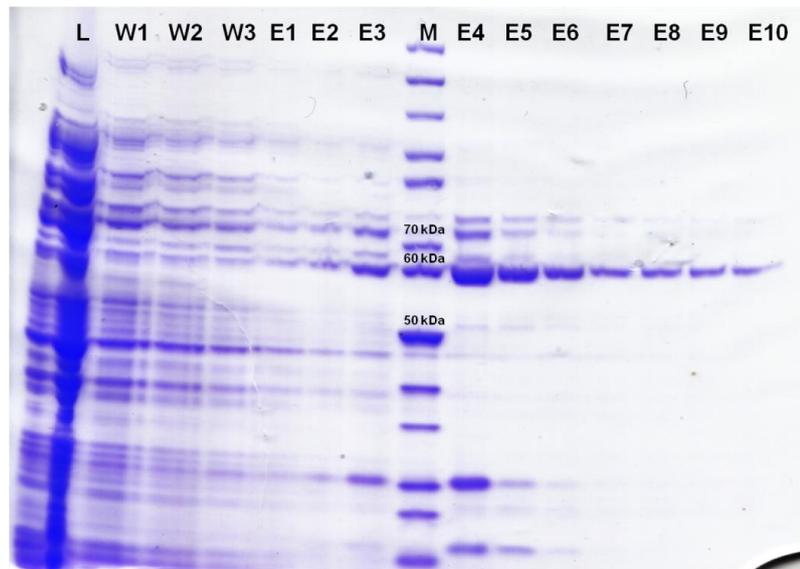


Abbildung 5-45: Expression der PppE-A-Domäne.M: PageRuler unstained Protein Ladder (*Thermo Fisher Scientific*); **W1-W3**: Waschfraktionen 1-3; **E1-E10**: Elutionsfraktionen 1-5.

Anders als in der Testexpression ist die Bande bei 60 kDa deutlich stärker als die in der Testexpression beobachtete Bande bei ca. 70 kDa. Dies spricht eher dafür, dass es sich bei der Bande >70 kDa um eine Coaufreinigung wirtseigener Proteine handelt. Häufig binden einige wirtseigene Proteine aufgrund ihrer Histidin-reichen Struktur ebenfalls an die Ni-NTA-Säule und werden zusammen mit den überexprimierten, His-getaggten Proteinen eluiert.^{233,234} Eines der häufigsten Histidin-reichen Proteine, die aus *E. coli* coeluiert werden, ist ArnA, das mit 74 kDa der auf dem Gel sichtbaren Bande entsprechen würde.²³⁵ Neben dieser sind in den Elutionsfraktionen auch niedrigere Banden zu erkennen. Ein weiteres Histidin-reiches Protein aus *E. coli* ist das 20 kDa große SlyD, das eventuell für eine dieser kleineren Banden verantwortlich ist. Da auch in einem zweiten Ansatz eine deutliche Bande bei 60 kDa zu erkennen war, wurde spekuliert, dass es sich hierbei um die überexprimierte A-Domäne handeln könnte. In einigen Fällen können Proteine auch niedriger laufen, als die Vergleichsbande des Markers. Dies kann dann der Fall sein, wenn das Protein aufgrund seiner Aminosäurezusammensetzung eine hohe Hydrophobizität aufweist und so mehr SDS bindet.²³⁶ Mit einem *GRAVY* (*Grand average of hydropathicity*) Wert von 0.063 ist die A-Domäne jedoch weder besonders hydrophob noch besonders hydrophil. Eine hier indizierte weitere Charakterisierung mittels MALDI oder Proteinsequenzierung konnte aus Zeitgründen nicht durchgeführt werden. Die Analyse zur Bestimmung der Spezifität wurde trotz der unklaren Proteinidentität durchgeführt. Dazu wurden die Elutionsfraktionen mit einem möglichst hohen Proteingehalt und möglichst schwachen Kontaminationsbanden über PD10-

Säulen in den Assaypuffer eluiert. Anschließend wurde das Protein mithilfe von Vivaspin-Säulen aufkonzentriert (Kapitel 7.5.5). In einer Massenspektrometrie-basierten Analyse sollte dann die Aktivität und Spezifität der exprimierten A-Domäne bestimmt werden (Kapitel 7.5.8). Wie in Kapitel 3.1.2.1 beschrieben, katalysieren A-Domänen in einer ATP-abhängigen Reaktion die Aktivierung einer spezifischen Aminosäure unter Bildung von Aminoacyl-AMP (**31**) und Pyrophosphat. Mithilfe von Isotopen markiertem $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ (**29**; $m/z=514$) kann die Rückreaktion nachverfolgt und zur Messung der Enzymaktivität verwendet werden. Das Prinzip dieses Assays beruht auf der Bildung und der Detektion von $[\gamma\text{-}^{16}\text{O}_4]\text{-markiertem ATP}$ (Abbildung 5-46, **32**) durch die von der A-Domäne katalysierte Rückreaktion mit im Überschuss zugegebenen Pyrophosphat (**30**), bei gleichzeitigem Verbrauch des $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$. Im Massenspektrum lässt sich dies durch eine Abnahme des $m/z=514$ ($[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$) mit proportionaler Zunahme des $m/z=506$ durch das entstehende ATP verfolgen (Abbildung 5-47).⁶⁴

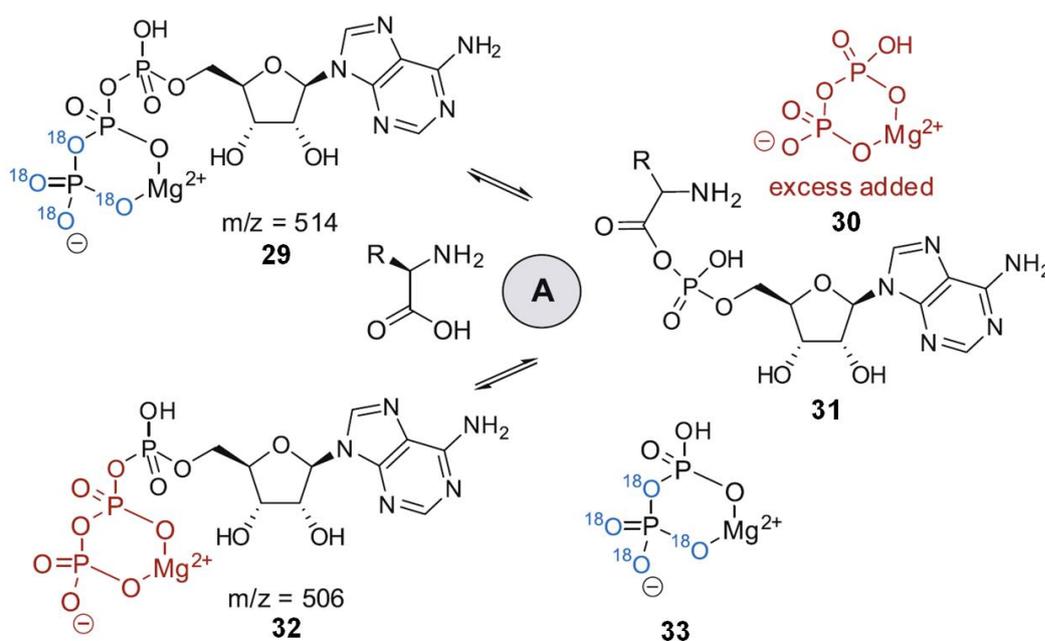


Abbildung 5-46: Prinzip des $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assays nach Phelan et al.⁶⁴ Die A-Domäne katalysiert in einer ATP-abhängigen Reaktion die Aktivierung der Aminosäure unter Bildung eines Aminoacyl-AMP (**31**) und Pyrophosphat. Das Prinzip des $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assays beruht auf die Detektion der Rückreaktion. Dabei wird die Bildung von $[\gamma\text{-}^{16}\text{O}_4]\text{-ATP}$; **32**, ($m/z=506$) durch im Überschuss zugegebenen Pyrophosphat (**30**) bei gleichzeitiger Abnahme des eingesetzten $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ (**29**; $m/z=514$) detektiert

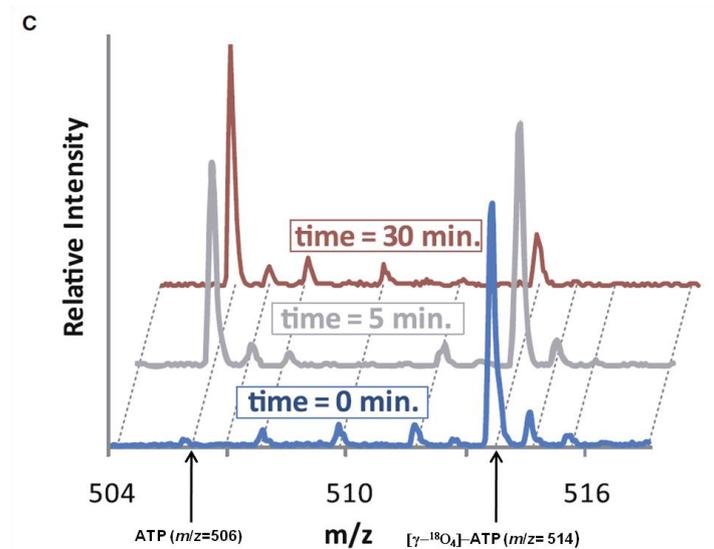


Abbildung 5-47: Beispiel der Massenspektrometrische Darstellung des $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assays bei Akzeptanz des Substrats aus der Literatur.⁶⁴ Dargestellt ist die „Verschiebung“ des Peaks bei $m/z=514$ ($[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$) zu $m/z=506$ ($[\gamma\text{-}^{16}\text{O}_4]\text{-ATP}$). Gezeigt werden die Spektren zu drei verschiedenen Zeitpunkten (nach 0, 5, und 30 min).

Die durch die AG Piel (ETH Zürich) bioinformatisch vorhergesagten, potentiellen Substrate sind α -Ketoisocapronsäure, α -Ketoisovaleriansäure und/oder α -Ketopropionsäure. Neben diesen Substraten wurden auch weitere Aminosäuren getestet (Tabelle 5-6).

Tabelle 5-6: Getestete Substrate und deren Aktivierung durch die exprimierte A-Domäne.

Substrat	Aktivierung %
4-Methyl-2-oxopentansäure	0,29
3-Methyl-2-oxobutansäure	0,18
Pyruvat	0,09
Prolin	0,08
Tryptophan	0,05
Tyrosin	0,02
Valin	0,07
Isoleucin	0,13
Leucin	0,19
Phenylalanin	0,12
Alanin	0,39
Arginin	0,15
Hisidin	0,15
Serin	0,09
Threonin	0,11
Lysin	0,19
Methionin	0,10
Asparagin	0,06
Cystein	0,04
Glutamin	0,16
Glutaminsäure	0,10

Glycin	0,09
--------	------

Keines der untersuchten Substrate zeigte eine Verschiebung des Peaks von $m/z= 514$ zu $m/z= 506$ (Abbildung 10-24). Die Expression und der Assay wurden mehrfach wiederholt. Es konnte jedoch in keinem der Ansätze eine Aktivierung einer Aminosäure bzw. eine daraus resultierende Verschiebung der Peaks im MALDI-Spektrum beobachtet werden. Hier sind verschiedene Gründe denkbar:

Da aus zeitlichen Gründen nach der heterologen Expression und der anschließenden Isolierung keine weitere Charakterisierung des Proteins durch MALDI oder Proteinsequenzierung durchgeführt wurde, kann nicht ausgeschlossen werden, dass es sich bei dem isolierten Protein u. U. nicht um die gewünschte A-Domäne handelt. Dafür spricht auch, dass die auf dem Gel der SDS-PAGE beobachtete Proteinbande nicht auf der erwarteten Höhe von 66 kDa sondern auf Höhe von etwa 60 kDa lag. Dagegen spricht die Intensität der Proteinbande bzw. die Menge an isoliertem Protein. Die Elution eines *E. coli*-eigenen, Histidin-reichen Proteins, das nicht überexprimiert wurde, würde eine deutlich schwächere Bande zeigen. Die zumindest auf dem Gel der SDS-PAGE zu beobachtende geringere Größe des Proteins könnte auch ein Hinweis auf eine Degradation der A-Domäne hindeuten. Für diese Theorie würden die ebenfalls zu erkennenden kleineren Proteinbanden sprechen. Allerdings sind diese kleinen Proteinbanden in Ihrer Intensität deutlich schwächer als die 60 kDa-Bande und würden zusammen mit dieser weit über der für die A-Domäne erwarteten Größe von 66 kDa liegen. Um diese Möglichkeiten auszuschließen, ist eine nähere Charakterisierung des isolierten Proteins zwingend erforderlich.

Eine weitere Hypothese ist eine Fehlfaltung bzw. inaktive Konformation der aufgereinigten A-Domäne *in vitro*. Grund hierfür könnte das Fehlen eines MbtH-ähnlichen Proteins sein, das für die Aktivität einiger A-Domänen essentiell ist (Kapitel 3.1.2.2).⁷² Die genaue Funktion dieser kleinen Proteine ist noch unbekannt. Es gibt aber Hinweise darauf, dass sie durch Bindung zu einer Konformationsänderung und somit zur Steigerung der Aktivität der A-Domäne führen (allosterischer Regulator).^{75,76} Zwar konnte in einer ersten Genomanalyse des 'Pseudomonas paederi'-Genoms kein MbtH-ähnliches Protein identifiziert werden, allerdings ist das Vorhandensein eines solchen Proteins aufgrund ihrer kurzen Gensequenz (ca. 200 bp) und der noch bestehenden Sequenzlücken im Genom nicht auszuschließen. Daher sollte in einem weiteren Ansatz die A-Domäne gemeinsam mit einem MbtH-ähnlichen Protein coexprimiert werden und der Assay erneut durchgeführt werden, um dies auszuschließen. Dieser Ansatz konnte bisher aus zeitlichen Gründen nicht weiter verfolgt werden.

Parallel zur Bestimmung der A-Domänen-Spezifität wurde versucht, den gesamten Gencluster heterolog zu exprimieren (Kapitel 7.5.10). In vorangegangenen Arbeiten konnte aus einer Genombibliothek ein Cosmid isoliert werden, das den gesamten Gencluster enthält (pPD7E4). Dieses wurde aus dem *E. coli* EPI300 isoliert (Kapitel 7.4.5) und in den Expressionsstamm *E. coli* BL21 transformiert (Kapitel 7.3.4). Als Negativkontrolle wurde das Insert durch einen Restriktionsverdau aus dem Vektor (pWEB) entfernt und dieser dann ebenfalls in *E. coli* BL21 kloniert. Eine Überexpression mithilfe des enthaltenen T7 Promotors konnte nicht durchgeführt werden, da eine Sequenzierung zeigte, dass dieser invers zum Gencluster orientiert ist. Beide Stämme wurden in 500 ml TB-Medium für 48 h bei 37 °C inkubiert. Überstand und Zellen wurden durch Zentrifugation getrennt und separat behandelt. Die Extraktion erfolgte mit 100 ml Ethylacetat (v/v). Nach dem Einrotieren der organischen Phase und Gefriertrocknen der entstandenen Interphase wurden beide in Methanol gelöst und mittels MALDI-TOF untersucht. Die erwartete Größe lag zwischen 303 und 348 Da. Im MALDI-Spektrum konnte unter Anwendung einer HCCA-Matrix (α -Cyano-4-hydroxymizinsäure) 6-Aza-2-thiothymin (ATT)-Matrix keine Massen in der erwarteten Größe beobachtet werden. In Zusammenarbeit mit Dr. Max Crüsemann wurde zusätzlich eine LC/MS-Messung durchgeführt (Abbildung 10-19, Abbildung 10-20, Abbildung 10-21, Abbildung 10-22). Die so gewonnenen Daten wurden dann mithilfe der Plattform GNPS (*Global Natural Products Social Molecular Networking*) analysiert (Abbildung 10-23).²³⁷ Dabei werden die verschiedenen Spektren miteinander verglichen und die Massen farblich gekennzeichnet, die nicht in einer der Negativkontrollen vorkommen. Jedoch konnten auch hier keine Kandidaten in der erwarteten Größe identifiziert werden.

Am wahrscheinlichsten ist, dass das nichtribosomale Peptid nicht synthetisiert wurde. Ursächlich dafür könnte ein inaktiver/defekter Gencluster, fehlende Substrate oder die Konditionen der Kultivierung bzw. die Wahl des Expressionssystems sein. Zur Verbesserung der Transkription könnte ein Promoter (z. B. ein IPTG-induzierbarer T5 Promotor) vor den Gencluster kloniert und/oder ein für die Expression besser geeignetes Plasmid gewählt werden. Auch die Wahl des Expressionssystems könnte variiert und optimiert werden (z. B. ein *Pseudomonas* Expressionssystem anstelle des hier gewählten *E. coli*-Expressionssystems). Neue Arbeiten zeigten auch einen positive Effekt durch eine Coexpression von Chaperonen auf die heterologen Expression von Symbionten DNA.²³⁸ Um prinzipiell festzustellen, ob der Naturstoff synthetisiert wird, wäre eine MALDI-Analyse der Bakterienzellen selbst zielführend. Ein weiterer Grund für die fehlende Detektierbarkeit des nichtribosomalen Peptids könnte eine unzureichende Aufreinigung des Moleküls sein. Aufgrund der fehlenden Information der A-Domänen-Spezifität kann keine genaue Aussage über die zu detektierende Molekülstruktur getroffen werden. Eine optimale Anpassung der

Konditionen zur Aufreinigung und Analyse ist daher kaum möglich. Allerdings könnte hier eventuell eine Variation des Lösungsmittels Abhilfe schaffen.

Im Hinblick auf die eher gering einzuschätzende Wahrscheinlichkeit einer erfolgreichen heterologen Expression, sollte der Fokus zur Charakterisierung zunächst auf die Bestimmung der A-Domänen-Spezifität gelegt werden.

Im Rahmen dieser Arbeit konnten acht der nach der Sequenzierung und Assemblierung verbliebenen Contigs des 'P. paederii' Genoms mittels kombinatorischer PCR verbunden werden (Kapitel 5.3.1). Aus zeitlichen Gründen konnte dies jedoch nicht für alle Contigs erfolgreich erfolgen. Der Ansatz einer kombinatorischen PCR sollte weiter verfolgt und so die verbleibenden Sequenzlücken geschlossen werden.

Des Weiteren wurde ein im 'P. paederii'-Genom enthaltenes NRPS-/PKS Gencluster genauer untersucht. Dazu sollte die Spezifität der beiden enthaltenen A-Domänen in einem *in vitro*-Ansatz bestimmen werden. Eine der A-Domänen konnte erfolgreich subkloniert und heterolog exprimiert werden. Eine Aktivität konnte in einem [γ - $^{18}\text{O}_4$]-ATP-Assays jedoch nicht gezeigt werden. Um Fehlfaltungen zu vermeiden, sollte in zukünftigen Versuchen die A-Domäne zusammen mit einem MbtH-ähnlichen Protein coexprimiert werden. Anschließend sollte vor Durchführung des Assays das aufgereinigte Protein analysiert und als die gewünschte A-Domäne identifiziert werden. Parallel zur Bestimmung der A-Domänen-Spezifität wurde in einem zweiten Ansatz versucht, den gesamten Gencluster heterolog zu exprimieren um das Produkt anschließend mithilfe von Ethylacetat aus der Kultur zu extrahieren. Dabei konnte unter den gewählten Bedingungen jedoch kein Molekül in der erwarteten Größe, weder mittels MALDI-TOF noch in einer ESI-LC/MS- Messung detektiert werden. Hier sollte versucht werden die Expressionsbedingungen zu optimieren und durch Variation des Lösungsmittels das Molekül erfolgreich aufzureinigen. Da der Erfolg einer heterologen Expression des Naturstoffclusters geringer einzuschätzen ist, sollte der Fokus zunächst auf die Bestimmung der A-Domänen-Spezifität gelegt werden, um so das Molekül näher charakterisieren zu können. Anhand dessen könnte dann auch eine Optimierung des Aufreinigungsprotokolls erfolgen. Diese Ansätze konnten im Rahmen dieser Arbeit aus zeitlichen Gründen nicht weiter verfolgt werden.

6 Zusammenfassung und Ausblick

Die Ergebnisse dieser Arbeit verdeutlichen das Potential bakterieller Symbionten als Quelle neuartiger bioaktiver Naturstoffe.

Im Folgenden sollen die Ergebnisse noch einmal zusammengefasst und ein kurzer Ausblick gegeben werden.

6.1 '*Candidatus Entotheonella*'

Die Genomsequenzierung und eine anschließenden Assemblierung der Naturstoffgenclustern des bakteriellen Symbionten '*Entotheonella factor*' konnten *in silico* zeigen, dass diese Spezies (Abbildung 5-5) alle bisher isolierten Naturstoffe des Schwammes *Theonella swinhoei* (Abbildung 3-18), mit Ausnahme von Aurantosid, synthetisiert (Kapitel 5.1.1).¹³¹ Dies macht diesen filamentösen Symbionten zu einem außergewöhnlich potenten Naturstoffproduzenten.

Aus ersten phylogenetischen Analysen war bekannt, dass '*Candidatus Entotheonella*' auch in nicht marinen Habitaten vorkommen. Anhand einer in dieser Arbeit zusammen mit der Bachelorstudentin Nilani Vickneswaran entwickelten, PCR-basierten Methode konnten in 22 von 35 Bodenproben (bzw. 49 Umweltproben) '*Candidatus Entotheonella*' nachgewiesen werden. Die anschließende taxonomische Einordnung der 16S rRNA-Gensequenzen anhand ihrer *pairwise identity* zu der von '*E. factor*' zeigte, dass vier der positiven Bodenproben '*E. factor*'-Phylotype und drei weiteren Bodenproben Spezies desselben Genus enthalten. Alle weiteren Bodenproben enthalten Spezies derselben Familie, derselben Ordnung, derselben Klasse bzw. desselben Phylums wie '*E. factor*' (Kapitel 5.1.2, Tabelle 5-1). Erste phylogenetische Stammbaumanalysen zeigten, dass '*E. factor*' dem neuen, bis dato unbeschriebenen Kandidatenphylum '*Tectomicrobia*' angehört.¹³¹ Um dieses neue Kandidatenphylum näher charakterisieren zu können, wurde eine phylogenetische Analyse mit insgesamt 640 16S rRNA-Gensequenzen (78 Sequenzen aus Bodenanalysen dieser Arbeit und 562 Sequenzen aus einer Datenbankrecherche) durchgeführt. Diese demonstriert, dass sich das Kandidatenphylum '*Tectomicrobia*' aus verschiedenen Kladen zusammensetzt, deren Mitglieder weltweit in unterschiedlichsten Habitaten zu finden sind. (Kapitel 5.1.2, Abbildung 5-17). Bei einigen Kladen ist eine Korrelation zwischen verschiedenen Umgebungsfaktoren und der Phylogenie zu beobachten (z. B. eine Assoziation mit Schwämmen, Salinität oder Wärme; Kapitel 5.1.3). Parallel wurde von Eike Peters im Rahmen seiner Doktorarbeit ebenfalls eine phylogenetische Analyse mit zusätzlichen Schwammdaten durchgeführt.²¹⁵ Sowohl das Ergebnis der Stammbaumanalysen dieser Arbeit anhand unterschiedlicher Algorithmen, als auch der Vergleich dieser Stammbäume mit der Stammbaumanalyse von Eike Peters zeigen, dass

weitere Analysen der einzelnen Kladen unter Hinzuziehen weiterer Sequenzen notwendig ist, um ein genaueres Bild der Phylogenie von 'Tectomicrobia' zu erhalten. Die Analyse des Kandidatengenus 'Entotheonella' zeigt, dass fast alle Sequenzen terrestrischen Ursprungs innerhalb dieses Genus zusammen clustern (ebenso wie die übrigen aus Schwämmen stammenden Sequenzen). Die Sequenzen aus einer der Bodenprobe (Probe I) nehmen eine besondere Stellung ein: Mit einer PI von 99,1% bzw. 99,3% fallen diese Sequenzen gemeinsam in ein Cluster mit 'E. factor'. Dies könnte ein Indiz dafür sein, dass auch diese Entotheonella-Phylotypen eine mögliche Quelle weiterer bioaktiver Naturstoffe sind. Häufig stellen diese Naturstoffe einen evolutionären Vorteil dar (z. B. wie im Fall von 'P. paederi' als Abwehrstoff in einer Symbiose oder durch Erschließung einer ökologischen Nische durch antibiotische Wirkung).^{214,239} Dies könnte, neben dem für 'E. factor' und zwei weiteren Vertreter des Kandidatengenus 'Entotheonella' bereits gezeigten, komplexen Stoffwechselrepertoire, ein weiterer Grund für die ubiquitäre Verbreitung dieser Bakterien sein. Für die beiden erwähnten weitere Mitglieder des Genus 'Entotheonella' ('E.serta' und 'E. gemina') konnte bereits in anderen Arbeiten gezeigt werden, dass auch ihr Genom ein enormes Repertoire an Naturstoffgenclustern enthält.^{214,240} Besonders erstaunlich ist das Genom von 'E.serta', das mit 308 Naturstoffdomänen die Anzahl von 'E. factor' mit 275 Domänen noch übersteigt. Dies zeigt das enorme Potential dieser Symbionten.²⁴⁰ Dabei ist nicht auszuschließen, dass auch andere Genera innerhalb des Phylums ebenfalls potente Naturstoffproduzenten beinhaltet. Allerdings fehlen hier derzeit noch entsprechende Genomdaten.

Versuche '*Candidatus Entotheonella*' aus anderen Habitaten wie z. B. Böden zu isolieren, gestalteten sich schwierig. Schon die Tatsache, dass '*Candidatus Entotheonella*' nur unter Anwendung einer two-step PCR konstant und reproduzierbar in Umweltproben nachweisbar war, deutet darauf hin, dass diese Spezies nur in sehr geringer Zahl im Verhältnis zum restlichen Mikrobiom vorkommen (Kapitel 5.1.2). Bestätigt wurde diese Vermutung durch eine Analyse des Mikrobioms einiger Bodenproben durch Amplicon-Sequenzierung (maximal 0,15% Anteil am gesamten Mikrobiom, nach Aufkonzentrierung; Tabelle 5-4, Kapitel 5.1.4.2). Dies erschwert eine Isolierung immens. Zwar konnten mittels FISH-Sonden einige Kandidaten markiert und visualisiert werden, diese Ergebnisse waren jedoch nicht reproduzierbar (Kapitel 5.1.4.1). Zur Charakterisierung dieser neuen Spezies sind weitere Versuche zur Isolierung notwendig. Dabei könnten auch Ansätze zur selektiven Kultivierung wie z. B. das sogenannte 'iChip'-Verfahren hilfreich sein. Dieses bietet durch Anwendung einer semiquantitativen Membran die Möglichkeit, bislang unter Laborbedingungen nicht kultivierbare Bakterien innerhalb ihres natürlichen Habitats in Reinkultur zu kultivieren.

6.2 Symbiont 2 des Schwammes *Theonella swinhoei*

Neben dem außergewöhnlichen Naturstoffproduzenten 'Entotheonella factor' konnten Dr. Tetsushi Mori und Prof. Dr. Haruko Takeyama einen weiteren filamentösen Symbionten im Schwamm *Theonella swinhoei* identifizieren. Neben dem bereits bekannten Aurantosid-Gencluster, das von der AG Piel an der ETH Zürich assembliert werden konnte, wurden in dieser Arbeit 32 weitere Contigs mit Fragmenten von PKS und NRPS-Clustern identifiziert (Abbildung 5-36). Neben NRPS- und PKS-Gencluster wurden Genfragmente weiterer Naturstoffe (Terpene: drei Contigs; einem Lanthipeptid: ein Contig; Cyanobactine: drei Contigs und mindestens einem weiteren Bacteriocin: ein Contig) identifiziert (Tabelle 5-5). Das zeigt, dass auch dieses Bakterium ein extrem potenter Naturstoffproduzent zu sein scheint. Da keine weiteren Naturstoffe aus *Theonella swinhoei* bekannt sind, wurde versucht, die Contigs nach dem Plausibilitätsprinzip mittels PCR miteinander zu verbinden. Dies gestaltete sich jedoch schwierig. Ein Grund hierfür war vermutlich die DNA-Probe, die für die PCR-Ansätze zur Verfügung stand. Dieses enthielt neben dem Genom dieses Symbionten noch DNA von mindestens zwei weitere ('Entotheonella factor' und 'Entotheonella gemina') Symbionten, die ebenfalls viele Naturstoffgencluster enthalten.²¹⁴ So kam es vermehrt zu unspezifischen Primerbindungen. Daher konnten lediglich zwei der identifizierten Contigs miteinander verbunden werden (Abbildung 5-38). Weitere Sequenzierungsläufe, wie im Fall von 'E. factor', könnten dabei helfen, weitere, noch vorhandene Lücken zu schließen. Alternativ dazu könnten eine inverse PCR zur Verknüpfung von Contigs genutzt werden. Bei dieser Methode wird die Templat-DNA mithilfe eines Restriktionsenzym verdaut und die dabei entstandenen Sequenzfragmente religiert. Die DNA-Konzentration wird dazu zuvor so verdünnt, dass die Fragmente vorzugsweise mit sich selbst ligieren. Es entstehen ringförmige Strukturen. Diese dienen in einer anschließenden PCR als Templat, wobei die jeweiligen Primer nach außen, also in entgegengesetzter Richtung, orientiert sind. Man erhält so ein PCR-Produkt, das Sequenzinformationen jenseits des untersuchten Contigendes enthält.²⁴¹ Eine Vorhersage einer Contigreihenfolge ist bei dieser Methode nicht nötig. Allerdings ist eine zu hohe Komplexität des Templats auch bei dieser Methode hinderlich.²⁴¹ Daher wäre eine aufgereinigte Probe, die ausschließlich DNA des *Symbiont 2* enthält, als Templat für weitere PCR-Ansätze zur Schließung der Lücken von Vorteil. Eine Analyse der NRPS- und PKS-Contigs zeigt, dass das Genom neben dem Aurantosid-Gencluster und einem *trans*-AT-PKS-Gencluster mindestens vier weitere NRPS/PKS-Gencluster enthält: Vier Contigs enthalten eine terminale Thioesterase-Domän (Kapitel 5.2). Ein Initiationsmodul konnte lediglich in einem Contig eindeutig identifiziert werden. Bei diesem sind im Randbereich weitere Gene zu finden, die eindeutig nicht dem Biosynthesecluster zugeschrieben werden können.

6.3 'Candidatus Pseudomonas paederii'

Nach der Sequenzierung des Genoms des Paederus Symbionten 'Pseudomonas paederii', einem weiteren Naturstoffproduzenten, konnte dieses in einer vorangegangenen Arbeit zu 100 Contigs assembliert werden.^{156,157} In dieser Arbeit konnten acht der verbliebenen Contigs durch kombinatorische PCR miteinander verbunden werden (Kapitel 5.3.1). Aus zeitlichen Gründen konnten nicht alle Lücken geschlossen werden. Um eine vollständige Assemblierung des Genoms zu erreichen sollte Ansatz der kombinatorischen PCR weiter verfolgt werden.

Außerdem wurde versucht, das neben dem Pederin- enthaltene NRPS-PKS-Gencluster näher zu charakterisieren. Dazu sollte die Spezifität der im Cluster enthaltenen A-Domänen anhand eines Aktivitätsassays bestimmt werden. Im Rahmen dieser Arbeit konnte eine der beiden A-Domänen exprimiert und mehrere Aktivitätsassays durchgeführt werden. Aufgrund fehlender Aktivität konnte allerdings keine Aussage zur Spezifität getroffen werden. In zukünftigen Versuchen die A-Domäne zusammen mit einem MbtH-ähnlichen Protein coexprimiert werden. Anschließend sollte vor Durchführung des Assays das aufgereinigte Protein analysiert und als die gewünschte A-Domäne identifiziert werden.

Dies kann verschiedene Gründe haben: Zum Einen kann nicht ausgeschlossen werden, sich bei dem isolierten Protein u. U. nicht um die gewünschte A-Domäne handelt, da aus zeitlichen Gründen keine weitere Charakterisierung des Proteins (z. B. durch MALDI oder Proteinsequenzierung) durchgeführt werden konnte. Dafür spricht, dass die auf dem Gel der SDS-PAGE beobachtete Proteinbande nicht auf der erwarteten Höhe von 66 kDa lag. Dagegen spricht die Intensität der Proteinbande bzw. die Menge an isoliertem Protein. Ein weiterer Grund könnte eine Degradation der A-Domäne sein, was auch die auf dem Gel der SDS-PAGE zu beobachtende geringere Größe des Proteins erklären würde. Für diese Hypothese würden auch die ebenfalls zu erkennenden kleineren Proteinbanden sprechen. Um diese Möglichkeiten auszuschließen ist eine nähere Charakterisierung des isolierten Proteins zwingend erforderlich.

Eine weitere Hypothese ist eine Fehlfaltung bzw. inaktive Konformation der aufgereinigten A-Domäne *in vitro*. Grund hierfür könnte das Fehlen eines MbtH-ähnlichen Proteins sein, das für die Aktivität einiger A-Domänen essentiell ist (Kapitel 3.1.2.2).⁷² In einer Genomanalyse des 'Pseudomonas paederii'-Genoms konnte kein MbtH-ähnliches Protein identifiziert werden, allerdings ist das Vorhandensein eines solchen Proteins aufgrund ihrer kurzen Gensequenz und der noch bestehenden Sequenzlücken im Genom nicht auszuschließen. Daher sollte in einem weiteren Ansatz ein MbtH-ähnliches Protein heterolog coexprimiert werden.

Parallel zur Bestimmung der A-Domänen-Spezifität wurde versucht, den gesamten Gencluster heterolog zu exprimieren und so eine direkte Charakterisierung des Moleküls zu

ermöglichen. Im direkten Vergleich der MALDI- bzw. LC/MS-Spektren konnte jedoch keine Masse im erwarteten Bereich beobachtet werden. Der wahrscheinlichste Grund hierfür ist, dass das nichtribosomale Peptid nicht synthetisiert wurde. Ursächlich dafür könnte ein inaktiver oder defekter Gencluster, fehlende Substrate bzw. die Konditionen der Kultivierung bzw. die Wahl des Expressionssystems sein. Hier könnte zur Verbesserung der Transkription ein Promoter vor den Gencluster kloniert und/oder ein für die Expression besser geeignetes Plasmid gewählt werden. Auch eine Variation und Optimierung des Expressionssystems oder eine Coexpression von Chaperonen könnte Abhilfe schaffen.²³⁸ Um prinzipiell festzustellen, ob der Naturstoff synthetisiert wird sollte eine MALDI-Analyse der Bakterienzellen selbst durchgeführt werden. Ein weiterer Grund für die fehlende Detektierbarkeit des nichtribosomalen Peptids könnte eine unzureichende Aufreinigung des Moleküls sein. Da aufgrund der fehlenden Information der A-Domänen-Spezifität keine genaue Aussage über die zu detektierende Molekülstruktur getroffen werden kann, ist eine Optimierung der Konditionen zur Aufreinigung und Analyse kaum möglich. Allerdings könnte eine Variation des Lösungsmittels zielführend sein. Im Hinblick auf die eher gering einzuschätzende Wahrscheinlichkeit einer erfolgreichen heterologen Expression, sollte der Fokus zur Charakterisierung zunächst auf die Bestimmung der A-Domänen-Spezifität gelegt werden.

7 Material und Methoden

7.1 Organismen und Vektoren

7.1.1 Organismen

Tabelle 7-1: Verwendete Organismen.

Stamm	Genotyp	Referenz/Herkunft
E. coli XL1-Blue	recA1 end A1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F'proAB lacIq ZΔM15 Tn10 (Tetr)]	Fa. Stratagene
E. coli DH5α	fhuA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	
E. coli BL21 (DE3)	(TetR)]F-, ompT, hsdSB(rB-mB-), gal, dcm, DE3	Fa. Invitrogen
Enthothionella factor		Theonella swinhoei
'P. paederi' aus Paederus Käfern		Paederus Käfer
<i>Symbiont 2</i>		Theonella swinhoei
verschiedenste Bakterienspezies		Bodenproben

7.1.2 Vektoren

Tabelle 7-2: Verwendete Vektoren.

Plasmid	Herkunft	Verwendung	Resistenz
pBluescript II SK (-)	Fa. Stratagene	Subklonierung	Ampicillin (100 µg/ml)
pGEM-T easy	Fa. Promega	Subklonierung	Ampicillin (100 µg/ml)
pHIS8	Prof. Bradley Moore	Proteinexpression	Kanamycin (50 µg/ml)

7.2 Medien und Zusätze

Tabelle 7-3: Verwendete Medien.

Medium	Substanz	Menge	pH
LB-Medium (engl.: <i>lysogeny broth</i>) ²⁴²	Natriumchlorid	10 g	7,5
	Trypton	10 g	
	Hefe-Extrakt	5 g	
	dH ₂ O	ad 1 L	
TB-Medium (Terrific Broth) ²⁴³	Trypton	12 g	
	Hefeextrakt	24 g	
	Glycerin	4 ml	

dH ₂ O	ad 900 ml
Nach dem Autoklavieren wurden 100 ml einer sterilen KH ₂ PO ₄ (0,17 M)/ K ₂ HPO ₄ (0,72 M) Lösung zugegeben.	

Tabelle 7-4: Verwendete Antibiotika.

Antibiotikum	Stockkonzentration	Endkonzentration
Ampicillin	100 mg/ml	100 µg/ml
Chloramphenicol	25 mg/ml	12,5 µg/ml
Kanamycin	50 mg/ml	50 µg/ml

Tabelle 7-5: Verwendete Medienzusätze.

Zusatz	Stockkonzentration	Endkonzentration
Isopropyl-β-D-thiogalactopyranosid (IPTG)	20 mg/ml	40 µg/ml
5-Brom-4-chlor-3-indoxyl-β-D-galactopyranosid (X-Gal)	20 mg/ml	40 µg/ml

7.3 Mikrobiologische Techniken

7.3.1 Kultivierung von *E. coli*

E. coli-Bakterien wurden sowohl auf LB-Agarplatten als auch in Flüssigkultur (5 ml) kultiviert. Diese wurden jeweils über Nacht im Brutschrank bzw. im Schüttelinkubator bei 37 °C (200 rpm) inkubiert. Die Medien wurden mit entsprechendem Antibiotikum versetzt.

7.3.2 Erstellen von Gefrierkulturen

Zur Lagerung von Bakterienstämmen über einen längeren Zeitraum wurden Gefrierkulturen (Kryokulturen) angelegt. Hierzu wurde 500 µl einer frischen Kultur des entsprechenden Bakterienstammes mit 500ml 30%igen Glycerins versetzt. Diese Glycerinkultur wurde dann in einem Kryoröhrchen bei -80 °C gelagert.

7.3.3 Konzentrationsbestimmung

Die Konzentrationsbestimmung einer Bakterienkultur erfolgte photometrisch. Dazu wurde die optische Dichte von 600 nm (OD₆₀₀) bestimmt. Als Referenz wurde das jeweilige sterile Medium verwendet.

7.3.4 Transformation in *E. coli*

7.3.4.1 Herstellung chemisch kompetenter Zellen

Zur Herstellung chemisch kompetenter Zellen wurde eine 5 ml Kultur mit *E. coli* XL-1 Blue wie unter 7.3.1 beschrieben kultiviert. Anschließend wurden 2,5 ml dieser Übernachtskultur in eine 100 ml LB-Hauptkultur inokuliert und bei 37 °C und 200 rpm für 1,5 Stunden inkubiert. Bei einer OD₆₀₀ von 0,4-0,5 wurden die Zellen in zwei 50 ml Reaktionsgefäße überführt und für 6 min bei 5000 rpm und 4 °C abzentrifugiert. Nach Verwerfen des Überstands wurden die Zellpellets in je 12,5 ml TFB I auf Eis gewaschen. Schließlich wurde der Überstand verworfen und die Zellpellets in 2 ml TFB II gelöst. Die Zellsuspension wurde in gekühlte 1,5 ml Reaktionsgefäßen zu je 90 µl aliquotiert, schockgefroren und bei -80 °C gelagert.

Tabelle 7-6: Puffer zur Herstellung chemisch kompetenter Zellen.

	Substanz	Stockkonzentration	Endkonzentration
TFB I	Kaliumacetat	5 M	30 mM
	Calciumchlorid	1 M	10 mM
	Kaliumchlorid	1 M	100 mM
	Manganchlorid	1 M	50 mM
	Glycerin	50%	15%
	dH ₂ O		ad Endvolumen
TFB II	Morpholinopropansulfonsäure (MOPS)	1 M	10 mM
	Calciumchlorid	1 M	75 mM
	Kaliumchlorid	1 M	10 mM
	Glycerin	50%	15%
	dH ₂ O		ad Endvolumen

7.3.4.2 Chemische Transformation

Die chemische Transformation dient dem Einbringen von Fremdplasmiden in chemisch kompetente Bakterien. Das Prinzip dieser Methode basiert auf Erhöhung der Permeabilität der Zellmembran durch einen Hitzeschock. Dazu wurde der gesamte Ligationsansatz (10 µl) zu 90 µl chemisch kompetenten *E. coli* zugegeben und vorsichtig gemischt. Der Ansatz wurde dann für 30 min auf Eis inkubiert, dann folgte der Hitzeschock bei 42 °C für 90 s. Anschließend wurden die Zellen für 2 min auf Eis abgekühlt. Zur Regeneration wurden dann 900 µl Antibiotika-freies LB-Medium zugegeben und die Kultur für 1 h bei 37 °C und 250 rpm im Schüttelinkubator inkubiert. Schließlich wurden unterschiedliche Volumina des Ansatzes auf Selektionsplatten ausplattiert und über Nacht bei 37 °C inkubiert.

7.3.4.3 Herstellen elektrokompetenter Zellen

Zur Herstellung elektrokompetenter Zellen wurde zunächst eine 5 ml Kultur mit *E. coli* XL-1 Blue bzw. *E. coli* BL21 bei 37 °C und 200 rpm im Schüttelinkubator über Nacht inkubiert. Am

Folgetag wurden diese in 400 ml LB-Medium überimpft und bei 37 °C und 170 rpm bis zu einer OD₆₀₀ von 0,4 inkubiert. Anschließend wurden die Zellen in vier 50 ml Reaktionsgefäßen bei 5000 rpm und 4 °C für 5 min abzentrifugiert. Nach Abnehmen des Überstands wurden die Zellpellets auf Eis in je 40 ml 10%igen Glycerin resuspendiert und wieder bei 5000 rpm und 4 °C für 5 min abzentrifugiert. Dieser Waschschrift wurde mit 20 ml wiederholt. Die so entstandenen Zellpellets wurden in 10 ml resuspendiert und je zwei zu 20 ml vereinigt. Nach wiederholtem Abzentrifugieren bei 5000 rpm und 4 °C für 5 min wurden die beiden Zellpellets in 1 ml 10%igen Glycerins vorsichtig resuspendiert. Die Zellsuspension wurde dann in 70 µl Aliquots auf vorgekühlte 1,5 ml Reaktionsgefäße aufgeteilt und in flüssigem Stickstoff schockgefroren. Die elektrokompenten Zellen wurden bei -80 °C gelagert.

7.3.4.4 Elektroporation

Die Elektroporation dient dem Einbringen von Plasmiden in elektrokompente *E. coli*. Bei dieser Methode wird die Permeabilität der Zellmembran gegenüber Fremd-DNA durch Anlegen eines starken elektrischen Feldes erhöht. Für die Transformation wurden 3-5 µl des Ligationsansatzes zu 90 µl auf Eis aufgetauten, elektrokompenten *E. coli* zugegeben und vorsichtig gemischt. Der Transformationsansatz wurde dann für ca. 5 min auf Eis inkubiert. Anschließend wurde der Ansatz in eine gekühlte Küvette überführt. Diese wurde dann in einem Elektroporator für kurze Zeit einer Spannung von 2,5 kV ausgesetzt. Unmittelbar danach wurden die Zellen in 500 µl LB-Medium aufgenommen und in 1,5 ml-Gefäß überführt. Zur Regeneration wurde die Kultur für 1 h bei 37 °C und 250 rpm im Schüttelinkubator inkubiert. Schließlich wurden zwei verschiedene Volumina des Ansatzes auf entsprechende Selektionsplatten ausplattiert und über Nacht bei 37 °C inkubiert.

7.3.4.4.1 Blau-Weiß-Screening

Durch das im pBluescript II SK (-) enthaltene lacZ-Gen war es möglich, ein sogenanntes Blau-Weiß-Screening durchzuführen. Dazu wurde der Selektionskulturplatte zusätzlich 100 µl IPTG (20 mg/ml) und 40 µl X-Gal (20 mg/ml) zugegeben (Kapitel 7.2). Nach Inkubation der Platten bei 37°C über Nacht konnten Kolonien, die das gewünschte Konstrukt aufgenommen hatten, anhand ihrer weißen Farbe von den anderen Kolonien (blau) unterschieden werden.

7.4 Molekularbiologische Techniken

7.4.1 Polymerase-Kettenreaktion (PCR)

Die Polymerase-Kettenreaktion dient der Vervielfältigung (Amplifikation) einer Ziel-DNA z. B. zur Klonierung oder Sequenzierung.²⁴⁴

Zur Amplifikation von DNA-Fragmenten mittels PCR wurden je nach Ansatz verschiedene Polymerasen entsprechend der Herstellerangaben verwendet. Bei den meisten PCR wurden

entweder eine *Taq*-Polymerase oder das Phusion High Fidelity System benutzt. Vorteile der Phusion High Fidelity Polymerase gegenüber einer *Taq*-Polymerase ist die geringere Fehlerrate (10^{-6} , *Taq*-Polymerase: 10^{-5}) und die höhere Synthesegeschwindigkeit.

Tabelle 7-7: PCR-Reaktionsansätze.

	taq-Polymerase	Phusion-Polymerase
DNA	0,5 µl	2 µl
Puffer (10x bzw. 5x)	2,5 µl	10 µl
BSA (10x)	2,5 µl	/
Vorwärtsprimer (50 µl)	0,5 µl	0,5 µl
Rückwärtsprimer (50 µl)	0,5 µl	0,5 µl
Enzym	0,2 µl	0,5 µl
dNTP (10 mM)	0,5 µl	1 µl
dH ₂ O	ad 25 µl	ad 50 µl

Die PCR wurden in einem Thermocycler mit den entsprechenden Temperaturprogrammen (primer-, polymerase-, templat- und produktabhängig) durchgeführt.

Tabelle 7-8: PCR-Programm Taq-Polymerase.

Taq-Polymerase		
	Temperatur	Dauer
Heizdeckel	106 °C bzw. 99°C	
Denaturierung	95 °C	2 – 5 min
Denaturierung	95 °C	30 sec
Annealing	X °C	1 min
Elongation	72 °C	1 – 3 min
Elongation	72 °C	10 min
Abkühlen	4 °C	~

} 35 Zyklen

Tabelle 7-9: PCR-Programm Phusion-Polymerase.

Phusion-Polymerase		
	Temperatur	Dauer
Heizdeckel	106 °C bzw. 99°C	
Denaturierung	98°C	30 s
Denaturierung	98°C	10 s
Annealing	X °C	30 s
Elongation	72 °C	15 s
Elongation	72 °C	5 min
Abkühlen	4 °C	~

} 35 Zyklen

Das Ergebnis der PCR wurde mittels Agarose-Gelelektrophorese überprüft.

7.4.1.1 Verwendete Primer

Tabelle 7-10: Primer zum Lückenschluß im 'E. factor'-Genom durch *Primer-Walking*.

Scaffold	Primer	Primersequenz	Primertyp
	T2	TAGTCATGCCATGTCAGTTCTAGCATCTGNNNNCGTCTTC	CLP
	NRL1	TAGTCATGCCATGTCAGTTCTAGCATCTGNNNNCTGCTGA	CLP
	NRL3	TAGTCATGCCATGTCAGTTCTAGCATCTGNNNNCAGCAGT	CLP
	NRL5	TAGTCATGCCATGTCAGTTCTAGCATCTGNNNNGCTCTCA	CLP
	NRL7	TAGTCATGCCATGTCAGTTCTAGCATCTGNNNNCAGCTGA	CLP
	RL	TAGTCATGCCATGTCAGTTCTAGCATCTG	LFP
Scaffold00001	01_SP_5	CCAGCGATGCAGAGTTCCGCC	SP
Scaffold00001	01_SP_3	TGTTGACAGAGGCCGGAACGC	SP
Scaffold00001	01_LSP_5	GAGTTCGCCGGGAATACCGGGGGGTAAGG	LSP
Scaffold00001	01_LSP_3	GCGGAACGCCGGCAGCTTCTGGCCGAGTG	LSP
Scaffold00019	19_SP_3	GCGGTTGCGGTCGTTTTTGG	SP
Scaffold00019	19_LSP_3	CGTTTTTGCCAGACGCAGTGGCGCTATC	LSP
Scaffold00098	98_SP_3	CGCCGCACACGATTCACATC	SP
Scaffold00098	98_LSP_3	ATTCACATCTTTTCCGACCTTTCATTTCA	LSP
Scaffold00111	111_SP_3	TCAACCCAGAAGCAGTGGGA	SP
Scaffold00137	137_SP_3	TGCCGCTTTGGCTCGTGAC	SP
Scaffold00195	195_SP_3	GGTGTGCCTCAATCCCTGC	SP
Scaffold00273	273_SP_3	GCGACCGTGACATACTGCC	SP
Scaffold00296	296_SP_3	CCAGTGTCTATCCC GCGCCT	SP
Scaffold00296	296_SP_5	GTGCAAGTCCGGCAAGGGAG	SP
Scaffold00425	425_SP_5	TCGGGAAGCGGTAGTCGGTC	SP
Scaffold00514	514_SP_5	CCTGAGGTGATTGTCGCCGT	SP
Scaffold00514	514_SP_3	TCCAGCTCCACAAAGGCCGA	SP
Scaffold00588	588_SP_3	CGGTGGATCGCCACTTGCC	SP
Scaffold00588	588_SP_5	CGATGCGCTGCTGGTGAAGA	SP

Tabelle 7-11: Primer zum Lückenschluß im 'E. factor'-Genom.

A. Primer zum Lückenschluß zwischen den Genom-Scaffolds			
Scaffold	Primer	Primersequenz	Herkunft
Scaffold00001	001_Ext_3	TCAGGCAATCCCAGTTTTGCTGAGGTGC	Micheal Wilson
	001_Ext_5	TCCACTACTTCAGTCATGTGCTGC	Micheal Wilson
Scaffold00019	019_Ext_3	CTGACAAGGAAGTCGAGCCATGAG	Micheal Wilson
Scaffold00111	111_Ext_3	CAGCGGAATACAGGCATTAGGTG	Micheal Wilson
Scaffold00160	160_Ext5	TCTGGAGACAGGCCACACCA	Micheal Wilson
Scaffold00195	195_Ext_3	TCTACCCAAATCTCTTGCCTGTGG	Micheal Wilson
Scaffold00203	203_Ext3	TACGCACTGCCGAACATCGC	Micheal Wilson
	203_Ext5	CAGAAATTGCGAGGCCGGTGC	Micheal Wilson
Scaffold00224	224_Ext3	CGCCGATGATTTTCGCGGTGT	Micheal Wilson
	224_Ext5	CGCTTGTGGAAGAACTCCTCGC	Micheal Wilson
Scaffold00273	273_Ext_3	TTCCGGTCGTTTGGGTTAAGGATG	Micheal Wilson
Scaffold00288	288_Ext3	GCTGCCTATACCGTCAAGTGTG	Micheal Wilson
	288_Ext5	GCGGTAGAGTGGCCTTGCTG	Micheal Wilson
Scaffold00296	296_Ext_3	CCACTCTGTTATCGTCATCGCTCC	Micheal Wilson
	296_Ext_5	GGTGTCAGCAAAGGAAGCTCAGAC	Micheal Wilson
Scaffold00321	321_Ext3	GGCATCACGGGCAGACCAC	Micheal Wilson
	321_Ext5	CGCCGTTTGTTCGCGCCTAT	Micheal Wilson
scaffold00376	376_Ext_3	GCCTGACACCCTACTTATTGCCTG	Micheal Wilson
Scaffold00394	394_Ext3	GAAGGGGCACGTTGTCTCA	Micheal Wilson
	394_Ext5	GGCCAGACCGAACGCTTCAC	Micheal Wilson
Scaffold00451	451_Ext3	GTGACCTGTAATGCATGGATATCGT	Micheal Wilson
	451_Ext5	AGCCTGGATGCCGATGCCTT	Micheal Wilson
Scaffold00514	514_Ext_3	TGTCCTCAAAGCCCACAACCTC	Micheal Wilson
	514_Ext_5	GGTGCAGGAGGTGCTTCAACA	Micheal Wilson
Scaffold00588	588_Ext_5	ATAAATCCACACTTCCCGCCCATC	Micheal Wilson
scaffold00762	762_Ext_5	GGAGTTGTCCTGAAATGGCTTGATTG	Micheal Wilson
B. Primer zum Lückenschluß innerhalb der Genom-Scaffolds			
Scaffold	Primer	Primersequenz	Herkunft
Scaffold00001	001_Kon1F	TCCACTCTGATACCGAGCAGT	Micheal Wilson
	001_Kon1R	TCGACGGAATCCGGCAGGGG	Micheal Wilson
Scaffold00003	003_01F	CCCGACGATGGGCGTTCTTG	Micheal Wilson
	003_01R	GGTGAGCGGCACCTGGGATA	Micheal Wilson
Scaffold00060	060_01F	GTTCAGGGGCAAGGGTTCGC	Micheal Wilson
	060_01R	CGTCGAGCATCGGCGCATT	Micheal Wilson
Scaffold00076	076_01_01F	TTCGTGCCCCAATAACCCG	Micheal Wilson
	076_01_01F	TTCGTGCCCCAATAACCCG	Micheal Wilson
	076_01_01R	TTCAGACCTGCGGATCGTGC	Micheal Wilson
	076_01_01R	TTCAGACCTGCGGATCGTGC	Micheal Wilson
Scaffold00084	084_01F	CACACCTGCCGCCACTTTGT	Micheal Wilson
	084_01R	TTGCGATCTTGCGACTGCC	Micheal Wilson
Scaffold00086	086_01F	ATCAGGGCACAATCGCGCTG	Micheal Wilson
	086_01R	GGGCAACGGCACCTCTTCTG	Micheal Wilson
Scaffold00098	098_01_01F	CGCGGACTTGCTGTTCCCTC	Micheal Wilson
	098_01_01R	CGGAATTTGCGGTGGCCGTA	Micheal Wilson
	098_01_01R	CGGAATTTGCGGTGGCCGTA	Micheal Wilson
	098_02_01F	CAGGCCAAAAGGGGTGCAGG	Micheal Wilson
	098_02_01R	TGGCTCGGCCACGAGATAGG	Micheal Wilson
	098_02_01R	TGGCTCGGCCACGAGATAGG	Micheal Wilson
	098_03_01F	CCGGATCTCCTGCCGTTGA	Micheal Wilson
	098_03_01R	GCCGGATAGGCAGACAGGGT	Micheal Wilson
	098_03_01R	GCCGGATAGGCAGACAGGGT	Micheal Wilson
Scaffold00111	111_01_01F	GCCGTCGAGCAAGGCATGAA	Micheal Wilson

	111_01_01R	GGGCAATACCACCGGCGAAG	Micheal Wilson
	111_02_01F	CAGCCCGCGACACCTTTCTC	Micheal Wilson
	111_02_01R	AGATCGACGGCGAGCGCTAT	Micheal Wilson
Scaffold00114	114_01_01F	TCTCACGGCGGATGATGCCT	Micheal Wilson
	114_01_01R	CCAGCGAATGTCCTGGTGCC	Micheal Wilson
	114_02_01F	GATTCTGCTGCGGGACGGTG	Micheal Wilson
	114_02_01R	GCTGCTCCAAGCATTGCGGA	Micheal Wilson
	114_03_01F	TCCGTTGACGACCTGCCTGA	Micheal Wilson
	114_03_01R	TGATTGCCCAACCGGGGTGT	Micheal Wilson
Scaffold00118	118_01F	GGGTGGTCACCGAGGCGTTG	Micheal Wilson
	118_01R	GGTTCGATCACCGGGCTGGC	Micheal Wilson
Scaffold00333	333_01F	GTCGCAGCAGCACCACCATC	Micheal Wilson
	333_01R	AGCGGCCAGCCAGAGCTATT	Micheal Wilson
Scaffold00345	345_01_01F	CTGCTGCCAATCACCACCA	Micheal Wilson
	345_01_01F	CTGCTGCCAATCACCACCA	Micheal Wilson
	345_01_01R	TATCGCGTCTATGCCACCGC	Micheal Wilson
	345_01_01R	TATCGCGTCTATGCCACCGC	Micheal Wilson
Scaffold00376	376_01F	TCCCATCACCCGCCAACCTT	Micheal Wilson
	376_01R	CACGCTAAACGCAGCCCAT	Micheal Wilson
Scaffold00425	425GapF	CTGTTCAAGCTCTGTCATCGCGTC	Micheal Wilson
	425GapR	TTTCCGGGCAGGGTTCACAATATC	Micheal Wilson

Tabelle 7-12: Primer zum Lückenschluß im 'E. factor'-Genom durch Screenen einer Metagenom-Bibliothek.

Scaffold	Primer	Primersequenz	Herkunft
Scaffold00001	001_Screening_01F	GCTACCGCTCCGACGGC	Micheal Wilson
	001_Screening_01R	CGTGACGTGAGCCAAATCGTCC	Micheal Wilson
	001_Screening_02F	TGTTGCGCCTTGAATGGGTGCG	Micheal Wilson
	001_Screening_02R	GATTTTCGTTTCAGAGCGTGCTGC	Micheal Wilson
Scaffold00002	002_Screening_01R	AGAGTCCCGGTAGTTCACGTCG	Micheal Wilson
	002_Screening_01F	CACCTTTGCGGTCAATGTGGC	Micheal Wilson
	002_Screening_02F	CTCCAGCGAAGAGACCCATACG	Micheal Wilson
	002_Screening_02R	AAAGTTCCTGTGGTAAGGTCGGG	Micheal Wilson
Scaffold00019	019_Screening_01F	TGAGGACGGTCTTACCTTTGCC	Micheal Wilson
	019_Screening_01R	CCCCAACTCATTGACAAGGGG	Micheal Wilson
	019_Screening_02F	AAC TTGCTGGTGGCGTACTTGG	Micheal Wilson
	019_Screening_02R	CGAGCAACCGGAAGGCATGG	Micheal Wilson
Scaffold00098	098_Screening_01F	TGCGAGCCCTGGTTGCC	Micheal Wilson
	098_Screening_01R	ACAATATCAAAGCCGTGCCGC	Micheal Wilson
	098_Screening_02F	TGGCAGGTCAACGACGATGC	Micheal Wilson
	098_Screening_02R	TCTCCTGAATCCACAGTGCTGC	Micheal Wilson
Scaffold00111	111_Screening_01F	GAGTTGGGTGAAATCGAAACGG	Micheal Wilson
	111_Screening_01R	TGAAGAGGGTTGCATTGGTGG	Micheal Wilson
Scaffold00137	137_Screening_01F	TGAGGTTTGC GGATGGTATGGC	Micheal Wilson
	137_Screening_01R	GCAAGCAATGTCTTCGAGCACC	Micheal Wilson
Scaffold00173	173_Screening_01F	GACTGGCCGTAGCCGTGG	Micheal Wilson
	173_Screening_01R	GGCTCTCAAAGATCATCGTCAAGG	Micheal Wilson
Scaffold00195	195_Screening_01F	GCAACCCGACGTTTGACGC	Micheal Wilson
	195_Screening_01R	TGGCTGCCCAACTGC	Micheal Wilson
Scaffold00273	273_Screening_01F	GTGGTTGACATTTGAGTGGCC	Micheal Wilson
	273_Screening_01R	TGCAATACCTGGAACACCTTGGG	Micheal Wilson
	273_Screening_02F	GTTTGTGCCGGATGCCTTTGG	Micheal Wilson
	273_Screening_02R	GGCCAATAGAGAATGTCCGCC	Micheal Wilson
Scaffold00296	296_Screening_01F	GATGATGGTCTCAGTCGAGGCG	Micheal Wilson
	296_Screening_01R	CCCAACTCAATACGCAGACCCC	Micheal Wilson
	296_Screening_02F	CACGACCCTGTTTGATTTGTCCG	Micheal Wilson
	296_Screening_02R	TGGTCTTTCAATGCCGTTTGCG	Micheal Wilson
Scaffold00514	514_Screening_01F	GACGTGTAAGAGACGAGCGGG	Micheal Wilson

	514_Screening_01R	GACCTTCATGCTGGCTGACACC	Micheal Wilson
Scaffold00588	588_Screening_01F	AGTGGCGTATGTGGTCTCAGG	Micheal Wilson
	588_Screening_01R	ATCTCGACTTGGAGTTGACCG	Micheal Wilson

Tabelle 7-13: Primer zum Lückenschluß innerhalb des 'E. factor' Plasmids.

A. Primer zum Lückenschluß zwischen den Plasmid-Scaffolds

Scaffold	Primer	Primersequenz	Herkunft
Scaffold0002	002_Ext3	ACTAGAGGCGCGCTTGAAA	Micheal Wilson
	002_Ext5	GACCTGCCTGTCTCTGTTTTCATTTT	Micheal Wilson
Scaffold0009	009_Ext3	ACAGCCTGCTTTCTGGAGCG	Micheal Wilson
	009_Ext5	GCGCCGTCAATGATCTGCTTGT	Micheal Wilson
Scaffold00073	073_Ext3	CGTCGGCAATATCCTCAAACGC	Micheal Wilson
	073_Ext5	AGCCCGATCACCAGTCTGG	Micheal Wilson
Scaffold00160	160_Ext3	ACCCCTTTCCTGGTCTGTGTGG	Micheal Wilson
	160_Ext5	TCTGGAGACAGGCCACACCA	Micheal Wilson
Scaffold00203	203_Ext3	TACGCACTGCCGAACATCGC	Micheal Wilson
	203_Ext5	CAGAAATTGCGAGGCCGGTCCG	Micheal Wilson
Scaffold00224	224_Ext3	CGCCGATGATTTCCGCGGTGT	Micheal Wilson
	224_Ext5	CGCTTGTTGGAAGAACTCCTCGC	Micheal Wilson
Scaffold00288	288_Ext3	GCTGCCTATACCGTCAAGTGTG	Micheal Wilson
	288_Ext5	GCGGTAGAGTGGCCTTGCTG	Micheal Wilson
Scaffold00321	321_Ext3	GGCATCACGGGCAGACCAC	Micheal Wilson
	321_Ext5	CGCCGTTTGTTCGCGCCTAT	Micheal Wilson
Scaffold00394	394_Ext3	GAAGGGGCACGTTGTCTCTCA	Micheal Wilson
	394_Ext5	GGCCAGACCGAACGCTTCAC	Micheal Wilson
Scaffold00451	451_Ext3	GTGACCTGTAATGCATGGATATCGT	Micheal Wilson
	451_Ext5	AGCCTGGATGCCGATGCCTT	Micheal Wilson

B. Primer zum Lückenschluß innerhalb der Plasmid-Scaffolds

Scaffold	Primer	Primersequenz	Herkunft
Scaffold00073	073_01F	AGCCGACAAGCGGGATCTGCAA	Ursula Steffens
	073_01F_in	CACGATCTACCGGCGTCCA C	Ursula Steffens
	073_01R	TGGCCCATGGCGCATACTCACT	Ursula Steffens
Scaffold00160	160_01F	ACAGCGCGGATTCCTACTGCGT	Ursula Steffens
	160_01F_in	CTG CGC TTC AAT GAG ATT C	Ursula Steffens
	160_02F_in	GGA TTC TAC TGT CTT AGA G	Ursula Steffens
	160_01R	TGAGGTGTCACCCCGGCAATGT	Ursula Steffens
	160_02F	ACGCGGACACCACAAGTCGAGA	Ursula Steffens
	160_02R	AGCCAGCGGCAGGTTGTTTCA	Ursula Steffens
Scaffold00203	203_01F	AAATCGTCGCTTAGCCCGCGT	Ursula Steffens
	203_01F_in	CTA TGC GCC AGG AAC TCG TG	Ursula Steffens
	203_01R	TCCGCAACGCACAAGGCGAA	Ursula Steffens
Scaffold00224	224_01F	GGTTGTAAGATCGCCCGCTAGGC	Ursula Steffens
	224_01R	TCCTCCACACTCGTGCCGTTT	Ursula Steffens
	224_02R	CGTCCATCCATTGTAACGGCGCT	Ursula Steffens
	224_02F	TTACCTGCGTTCTGATGGCTCGC	Ursula Steffens
Scaffold00321	321_01F	ACGCTGCTGTTGTCGCCATCT	Ursula Steffens
	321_01R	GCGACCGAGTAGCAGAACTCTTCG	Ursula Steffens
	321_02F	TCG GCT GCA GCG AGC AGT G	Ursula Steffens
	321_02R	GCTTATCCATGTGGCGCAGTCGG	Ursula Steffens
Scaffold00394	394_01F	ACGAGCGGGCGGATTTGAACA	Ursula Steffens
	394_01F_in	GTC TCT ATG CCG GCT TGG TG	Ursula Steffens
	394_01R	ACGAAAAAGGCCAGACGCACACA	Ursula Steffens

Tabelle 7-14: Primer zum Nachweis von 'Candidatus Entotheonella'

Primer	Primersequenz	Herkunft
U-rRNA-27-F	AGAGTTTGATCMTGGCTCAG	Lane et al. (1985)
U-rRNA-1492-R	GGYTACCTTGTTACGACTT	Lane et al. (1985)
GenEnto271-f	GGGAAAGGTTCCGCCGGTCTG	Agustinus Uria
GenEnto735-f	GCATTAAGCCTTGAAACTGT	Agustinus Uria
GenEnto1290-r	GCCCCGCATCACCCGGTA	Agustinus Uria
238F	CCGGTCTGAGATGAGCTTGC	Agustinus Uria
1442R	TCACCCCAATCACCCGC	Agustinus Uria
Envir-01-spez-f	GTAGCTGGTCTGAGAGGA	Ursula Steffens
Envir-01-spez-r	CTGACAGGAGTTTACAACC	Ursula Steffens
Envir-02-r	GCACCGACGGGATCAAG	Ursula Steffens
Envir-deg-03-f	GGYYGCGRGGGAAAACG	Ursula Steffens
Envir-deg-03-r	GATTRGCAACCYCTYRCGAG	Ursula Steffens
sponge-01-f	GGGAAAGCAGGCCTCT	Ursula Steffens
sponge-01-r	CGGCGTACTGCGTCAGGCTT	Ursula Steffens
sponge-02-f	TAAAGTGCGAACGGGTG	Ursula Steffens
sponge-02-r	CCAGCCATCCGAACTGAG	Ursula Steffens
sponge-deg-03-f	BARGYCGRGGGGARAGCA	Ursula Steffens
sponge-deg-03-r	GYACAGRCTCCGARGAGAAGACA	Ursula Steffens

Tabelle 7-15: Primer zum Lückenschluß des 'P. paederi'-Genoms.

Contig	Primer	Sequenz
Symbiont_001	001_Ext5	TCGACGATTGAGCTGAGATC
	001_Ext3	CGTAGTAAGCGTTCAGTTGC
Symbiont_002	002-Ext5	TCATGGATGCGGTGAATACT
	002_Ext3-new	GCGATATCCCCATACTCAG
Symbiont_003	003_Ext5-new	GTTGCTGGGTGAAAACATGT
Symbiont_004	004_Ext5	CAGCTCACTTTTCTGTTGG
Symbiont_005	005_Ext5-new	GACTGAGGAATGTTTGTGCC
	005_Ext3	GTGCGTCAATACAGGAATGG
Symbiont_006	006_Ext3-new	GGAGATCGTCGGCTATGAAT
Symbiont_007	007_Ext3	AGATCGTCCGTAATTTGGCT
	007_Ext5	ACGCACTTCATCATCAGCTA
Symbiont_008	008_Ext3-new	TCTCTAAACCTCCTCCCACA
	008_Ext5	GCCTCGTGAATATGCTTGTC
Symbiont_011	011_Ext5-new	TCACTTGAGAGCAGGATGAC
Symbiont_012	012_Ext3-new	TGATCACTTACCACGTCCTG
Symbiont_013	013_Ext3-new	ATTGGCTTTACTCCGCTACA
	013_Ext5-new	CGACTACTGTAACCGTGACA
Symbiont_014	014_Ext3-new	TGTGGACCTTGATCTCGATG
	014_Ext5	CAATGTACTGTAGCCCGGAT
Symbiont_015	015_Ext5-new	GAAGTTCGGGTCACAACATG
Symbiont_017	017_Ext5	AAGTCGAGCAGATCCAAACC
	017_Ext3	GAACTGTTGATCGATGGCTC
Symbiont_018	018_Ext3-new	TGATAGGTGCTGGATTGCTC
Symbiont_020	020_Ext3-new	CGGACTTGAACATGTAGGGA
	020_Ext5-new	TCAAATTAACACGGTCGCC
Symbiont_021	021_Ext5	ATCAGGGGGAATAGGAGAGG
Symbiont_022	022_Ext3-new2	TCCAGATGATTAAGGGTGCC
	022_Ext3-new	TTCCCTTGTGACGTGGTAAA
Symbiont_023	023_Ext5-new	CGGATTCGGAACCTCAGGTAA

Material und Methoden

Symbiont_024	024_Ext3-new	GCATTCAACAATCCAGCACT
Symbiont_025	025_Ext3-new	TCAACTCGTCCTCAATCTCG
	025_Ext5	CCCAGCAATCACCTGTATGA
Symbiont_026	026_Ext5-new	ATCCAGACCCGTGAATACAC
Symbiont_031	031_Ext5-new	CCTAGTTTCGGGGTGACTAC
Symbiont_035	035_Ext5-new	GCATATCGACGTCAAACAGG
	035_Ext3-new	CTTTTCGGACGATTATGGCC
Symbiont_038	038_Ext5-new	ATGAGACCATTACCACCGTC
Symbiont_040	040_Ext5	GAGCCTGTGAACAATACAGC
Symbiont_042	042_Ext3-new	TCGAAAACAAACAGCTCGAC
	042-Ext5	ACGCGATAAGAAATTGGTCG
Symbiont_044	044_Ext3	CAGTTAGGGCTTCAAAGTCG
Symbiont_045	045_Ext5-new	GCCAGCTATTAGAATGCAGC
Symbiont_047	047_Ext5-new	GGAGGTAAGGAGGGTGAAG
Symbiont_048	048_Ext3	ACGCATTTTTCATACTCGCC
Symbiont_049	049_Ext5-new	GTTTTCAGGGCTTCATCGAC
Symbiont_050	050_Ext5	ACGACCGTATTAACCTCCCC
Symbiont_051	051_Ext3	AGGGTTCATGCCTATCTACG
Symbiont_053	053_Ext5	CGATGGATGATTCTCGAGGT
Symbiont_054	054_Ext3	GTACAAAGTCGCCACACTTG
Symbiont_056	056_Ext5-new	GCCCGAGTAAACAATTGCTT
Symbiont_058	058_Ext5-new	CGATTTTTCCGCAGTGACTT
Symbiont_059	059_Ext5	GATGTTGGACTCATAGGCGA
Symbiont_060	060_Ext3	CCAAAGAGAATATGCACCGC
	060_Ext5-new	CTGCCGGCTAGAATTCAAAG
Symbiont_061	061_Ext5	CAGATTCCCGGGTAGAACG
	061_Ext3	CGCGAGCTACTACGGATTTA
Symbiont_062	062_Ext3-new	TTGGGAAGTCATCGGGAAAA
	062_Ext5-new	CCAGATGATCTGCCATTCT
Symbiont_063	063_Ext3	ACCATCACGCCGGATATTG
Symbiont_064	064_Ext3-new	TGTATGACACTCCTGCTGTC
Symbiont_067	067_Ext5-new	GCTTCGCTACGTGTATGAAC
Symbiont_069	069_Ext3-new	TGATGATGTCACCCTGGATG
Symbiont_070	070_Ext5	CTCCTGGTCTCTTGCGAAA
Symbiont_071	071_Ext5	TAAATACGCTCAGGCTCAGG
Symbiont_072	072_Ext5-new2	TAGGGGAATTGATGATGGCC
Symbiont_073	073_Ext5-new	CGGGTCATAAAGGTCTTCGA
	073_Ext3-new	TTGATCCATTCCCATTTCGC
Symbiont_077	077_Ext5	ATGGCGAAAACCTGAGGAAGA
Symbiont_078	078_Ext3-new	CCACTTCCGGCATTTCATT
Symbiont_080	080_Ext5-new	GGTAATTGCAAGCCGAAGTT
Symbiont_081	081_Ext3-new	TGGTAATGCAGCCCGAATAT
Symbiont_082	082_Ext3	TCAGACATAGATTGCTCGCC
Symbiont_083	083_Ext5	TCTGTGTCACAACCTGGGAAA
Symbiont_084	084_Ext3	CTACCGTGGTCTGGAAAAGA
	084_Ext5	TTTCAGTTACCTGTTCCGCA
Symbiont_085	085_Ext5	GATAGCGACTTTGTTGGACG
Symbiont_086	086_Ext5-new	CATTCAATGCCAGTCGTTT
Symbiont_088	088_Ext3	TTAAGGAGGGCAAACAGGTT
	088_Ext5	CGGTCAGCTTGAGATCTTGA
Symbiont_091	091_Ext3	TTGCAACTCTCGTACTGTGA
Symbiont_094	094_Ext5-new	TCGAGCGTTCTCTACTCCTA
Symbiont_095	095_Ext3-new	GTGATACGAGTGTTGGGTCT
Symbiont_096	096_Ext3-new	GGCTCAGTGACTTCGACTAA
Symbiont_099	099_Ext3	ACTAACGAATGGGGACTGC
Symbiont_100	100_Ext3	AGCTGAACCACTATCACTAGC

Tabelle 7-16: Primer zum Lückenschluß des *Symbiont 2* Genoms.

Contig	Primer	Sequenz
Contig_0059	Chloro_0059_Ext5	CTCAACCGAACACTTGTGTC
Contig_0104	Chloro_0104_Ext3	GTAGTGCTTCGAGAACAACG
Contig_0148	Chloro_0148_Ext5	AGCGTATATCAAGGTGGTGG
Contig_0251	Chloro_0251_Ext3	ATCAAAATCCCCAGAACATTC
	Chloro_0251_Ext5	CACTGTCTAGCCAATCTTAATC
Contig_0303	Chloro_0303_Ext3	AACGGGCTATATACTGCTGG
	Chloro_0303_Ext5	GTGCCATAGTTGAACAGCG
Contig_0368	Chloro_0368_Ext3	CTACCAAGATGTCGCTATCTCTT
	Chloro_0368_Ext5	CAATCTGGAGAAGTTGAACGAG
Contig_0435	Chloro_0435_Ext3	CCAGATACCAGCGATGACTT
	Chloro_0435_Ext5	ATCTTGCCGATGGACAGATT
Contig_0490	Chloro_0490_Ext3	ACAAAGATTCGGTGCCTTA
	Chloro_0490_Ext5	TGATTCTGTCCCTGAAGTGG
Contig_0538	Chloro_0538_Ext3	TGACCGAACAGCTCATCTTT
Contig_0606	Chloro_0606_Ext3	AACAAATGTGGGTCAACTGC
	Chloro_0606_Ext5	CCTCAAAGGAAGGAAGTGTG
Contig_0637	Chloro_0637_Ext3	TACAGGGCGAGATGTTCTTC
	Chloro_0637_Ext5	AGTGGTGAAGAAAGCATGA
Contig_0786	Chloro_0786_Ext3	TCTCAAATGTGCGACCATG
	Chloro_0786_Ext5	TTCTTTGTGCTTTTGCAGC
Contig_1016	Chloro_1016_Ext3	CTTGCCTTAAGTTTTGCCGT
	Chloro_1016_Ext5	TCATGCTCGAATGGTTACGA
Contig_1121	Chloro_1121_Ext3	TTTGCATAGCCATTGAGCTG
	Chloro_1121_Ext5	GATGACAAGGAAACCACACG
Contig_1266	Chloro_1266_Ext3	CCGATTGCTCAAAAATGGGT
	Chloro_1266_Ext5	GGAATCGCAACTCACCTACT
Contig_1267	Chloro_1267_Ext3	ATTTAACCACAACGCAGCTC
	Chloro_1267_Ext5	GGACCGCATCTTAATTCACG
Contig_1452	Chloro_1452_Ext3	TTTGGCCCTGTTTGTCAATG
	Chloro_1452_Ext5	ACTTGCGAATTTTGAGGCTG
Contig_1706	Chloro_1706_Ext3	AATGAGAGGGGTGGAAACTC
	Chloro_1706_Ext5	ACCGGATCTCAAGGGATAGA
Contig_2168	Chloro_2168_Ext3	CCATAGTTGTTTGTGGTGCC
	Chloro_2168_Ext5	TAATCAGCAGACCACAACCC
Contig_2190	Chloro_2190_Ext3	CAAAGATTGCCTCGACCATC
	Chloro_2190_Ext5	CTTACAAGTTGGACGGCATG
Contig_2207	Chloro_2207_Ext3	CGCATGATACTGGAGGAAGT
	Chloro_2207_Ext5	CCATCAAGACATGCCCTTTG
Contig_2293	Chloro_2293_Ext3	GTGTCCACCAATTCGAAGA
	Chloro_2293_Ext5	GACAGAGATTGGGCGGTAG
Contig_2558	Chloro_2558_Ext3	GACTACGCCTATTTTCAGCG
	Chloro_2558_Ext5	AGCTTGCGTATCCTCAATCA
Contig_2879	Chloro_2879_Ext3	GTGATGTCCTTTGCCTGAAC
	Chloro_2879_Ext5	TGAATTGGGTGAGATCGAGG
Contig_2899	Chloro_2899_Ext3	AGAGTTGGAAGGGCTGATTG
	Chloro_2899_Ext5	TCGGTGCTGTACTIONAAATGA
Contig_2900	Chloro_2900_Ext3	ATATATTGAACGCCCCGCAA
	Chloro_2900_Ext5	CCAAAGCGAGCTGAATGATT
Contig_3583	Chloro_3583_Ext3	TTGTTCTCTCTCCAGCTTG
	Chloro_3583_Ext5	CGGAAAGTCAACAGCTTACG
Contig_3943	Chloro_3943_Ext3	GTGCCAACTGTGAGTAGTCA
	Chloro_3943_Ext5	AGCGGCTCTCATTTATGCT

7.4.1.2 Kolonie PCR

Um Ligationen bzw. Transformationen überprüfen zu können, wurde eine sogenannte Kolonie-PCR durchgeführt. Dabei wird statt aufgereinigter DNA eine aus den Kolonien entnommenen Zellprobe direkt zum PCR-Ansatz zugegeben. Bei *E. coli* ist es ausreichend, diese Zellen in einem ersten PCR-Schritt bei 95 °C für 5 min aufzubrechen, sodass die DNA frei vorliegt. Die weiteren Schritte entsprechen einem Standard-PCR Programm (Kapitel 7.4.1).

7.4.1.3 Aufreinigung von PCR-Produkten

Alternativ zur Aufreinigung von PCR-Produkten über die Auftrennung in einer Agarose-Gelelektrophorese und anschließender Isolierung aus dem Agarosegel (Kapitel 7.4.2) wurden PCR-Aufreinigungskits der Firma *Fermentas* oder der Firma *PeqLab* verwendet.

7.4.2 Agarose-Gelelektrophorese

DNA und DNA-Fragmente wurden in 1%igen (w/v) Agarosegelen aufgetrennt. Dazu wurde Agarose (1% (w/v)) in 1x TAE-Puffer (0,12% (v/v) Eisessig, 0,48% (w/v) Tris, 1 mM EDTA, pH 8) durch Aufkochen gelöst. Die Auftrennung erfolgte in einer mit 1x TAE-Puffer gefüllte Elektrophoresekammer bei einer Spannung von 100-130 V. Zum Beladen des Gels wurden die Proben mit einem 10x Ladepuffer versetzt. Als Größenstandard wurden 3-6 µl eines 1 kb-DNA Marker (1 kb plus ladder, *Fermentas*) verwendet (Abb. 5-1).

Zur Anfärbung der DNA wurde das Gel vor oder nach der elektrophoretischen Auftrennung mittels Ethidiumbromid behandelt. Dazu wurde entweder nach Abkühlen der Agaroselösung auf ca. 60 °C 0,1 µg/ml Ethidiumbromid zugegeben oder das Agarosegel nach der Auftrennung in einem Ethidiumbromid-Färbebad für 15-30 min inkubiert. Die DNA-Banden konnten schließlich unter UV-Licht detektiert werden.

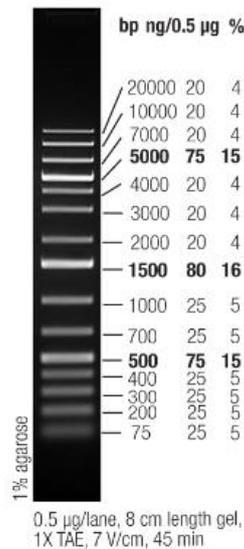


Abbildung 7-1: GeneRuler 1 kb Plus DNA Ladder (Thermo Fisher Scientific).²⁴⁵

7.4.2.1 Isolierung von DNA-Fragmenten aus Agarosegelen

Zur Isolierung von DNA-Fragmenten wurden diese zunächst auf ein Agarosegel aufgetragen und, wie unter 7.4.2 beschrieben, aufgetrennt. Anschließend wurde ein kleiner Teil der Probenasche abgeschnitten und dieses Stück in einem Ethidiumbromid-Bad (1x TAE-Puffer, 0,1 µg/ml Ethidiumbromid) angefärbt. Die DNA-Banden dieser Gelstücke wurden unter UV-Licht detektiert und entsprechend markiert. Schließlich wurden die Gelstücke wieder zusammengefügt und der entsprechende Gelbereich mittels der zuvor angefügten Markierung ausgeschnitten. Die ausgeschnittenen Gelbänder wurden dann in ein 1,5 ml Reaktionsgefäß überführt und die DNA mit einem Gel-Extraktionskit aufgereinigt (Firma *PeqLab*, *Thermo Fisher Scientific*, *BioBudget*).

Alternativ zu dieser Methode wurde zum Anfärben der DNA der Farbstoff GelGreen verwendet. Dieser Farbstoff bindet ebenfalls an die DNA und kann unter Licht im blauen Wellenbereich detektiert werden. Der Vorteil gegenüber Ethidiumbromid besteht darin, dass dieser laut Hersteller nicht in die DNA interkaliert. Zudem kann die Bande unmittelbar während des Detektierens unter Blau-Licht ausgeschnitten werden, da es anders als bei UV-Licht nicht zu einer DNA-Schädigung kommt.²⁴⁶

Als weitere Alternative zur Detektion mittels Ethidiumbromid und UV-Licht wurde das Blau/Grün-LED-Transilluminator-System der Firma *Nippon Genetics* verwendet. Die DNA wurde dabei mit dem Farbstoff „Midori Green Direct“ angefärbt. Dabei wird der Farbstoff, anders als bei den zuvor genannten Farbstoffen, direkt mit der DNA-Probe vermischt. Ein etwaiger Hintergrund durch Anfärben des gesamten Gels entfällt somit. Durch die Anregung der Probe durch ein Spektrum an Wellenlängen (Blau-Grüner Bereich) wird die Sensitivität zusätzlich erhöht.²⁴⁷

7.4.3 Klonierung

7.4.3.1 Ligation

Die enzymatische Ligation von DNA-Fragmenten wurde durch eine T4 DNA-Ligase (*Jena Bioscience* oder *NEB*) katalysiert. Der Reaktionsansatz hatte folgende Zusammensetzung (das Volumen von Vektor und Insert wurde individuell, abhängig von der Größe des Vektors und des Inserts, kalkuliert):

Tabelle 7-17: Ligationsansatz

Ligationsansatz	
T4-DNA Ligase	1 µl
T4-DNA Ligase Puffer	1 µl
Vektor	x µl
Insert	x µl
dH ₂ O	ad 10 µl
Verhältnis Vektor:Insert = 1:7	

Der Ligationsansatz wurde dann über Nacht bei 16 °C bzw. 4-7 °C inkubiert. Am Folgetag wurde die Ligase durch Inkubation für 15 min bei 65 °C inaktiviert. Die Selektion erfolgte auf LB-Agarplatten mit entsprechender Antibiotikazugabe.

7.4.3.2 TA-Klonierung

Zur besseren Sequenzierbarkeit wurden einige DNA-Fragmente in selbst hergestellte TA-Vektoren oder kommerziell erhältliche Klonierungsvektoren subkloniert. Dabei wird sich die Eigenschaft der *Taq*-Polymerase zu Nutze gemacht, während der PCR Fragmente mit 3'-dATP zu generieren. Dies ermöglicht es, die PCR-Fragmente in Vektoren mit 3'-dTTP Überhängen zu ligieren.

Zur Konstruktion eines TA-Klonierungsvektors diente das Phagenplasmid (Phagemid) pBluescript II SK (-), das durch Linearisierung und anschließendem Anhängen von 3'-dTTP modifiziert wurde.²⁴⁸ Nach der Präparation des Vektors (Kapitel 7.4.5) wurde dieser, wie unter 7.4.4 beschrieben, mithilfe des Restriktionsenzym *EcoRV* linearisiert. Anschließend wurde dieses durch Inkubation für 20 min bei 80 °C inaktiviert. Durch Zugabe von *Taq*-Polymerase (1u/µl Plasmid-DNA), dTTPs (2 mM) und anschließender Inkubation für zwei Stunden bei 70 °C wurden 3'-dTTP-Überhänge angefügt. Nach Aufreinigung der Plasmid-DNA mithilfe eines PCR-Aufreinigungskits (Firma *Thermo Fisher Scientific, PeqLab*) wurde die DNA in dH₂O eluiert und bei -20 °C gelagert.

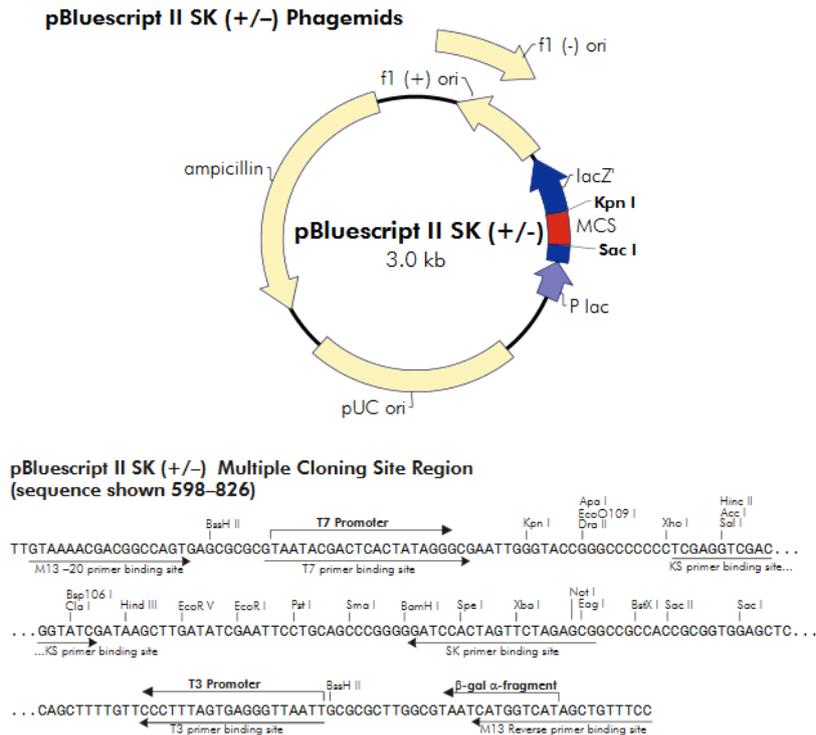


Abbildung 7-2: pBluescript II SK (+/-).²⁴⁹

7.4.3.2.1 Ligation in den TA-Vektor

Die enzymatische Ligation der PCR-Produkte in den TA-Vektor wurde durch eine T4 DNA-Ligase (*Jena Bioscience* oder *NEB*) katalysiert. Der Reaktionsansatz hatte die in Tabelle 7-17 aufgeführte Zusammensetzung.

Die Selektion erfolgte auf LB-Agarplatten mit 100 µg/ml Ampicillin (Amp100). Zum Blau-Weiß-Screening wurde IPTG (40 µg/ml) und X-Gal (40 µg/ml) zugegeben. Wie in Kapitel 7.3.4.4.1 beschrieben, enthalten weiße Kolonien meist das gewünschte Klonierungskonstrukt. Mit einigen dieser Kolonien wurde eine Kolonie-PCR durchgeführt (Kapitel 7.4.1.2) und/oder Übernachtskulturen angesetzt. Anschließend wurde die Plasmid-DNA, wie in Kapitel 7.4.5 beschrieben, isoliert und über einen Restriktionsverdau (Kapitel 7.4.4) und eine Sequenzierung (Kapitel 7.4.6) überprüft.

7.4.3.3 Ligation in pGEM-T Easy Vektor

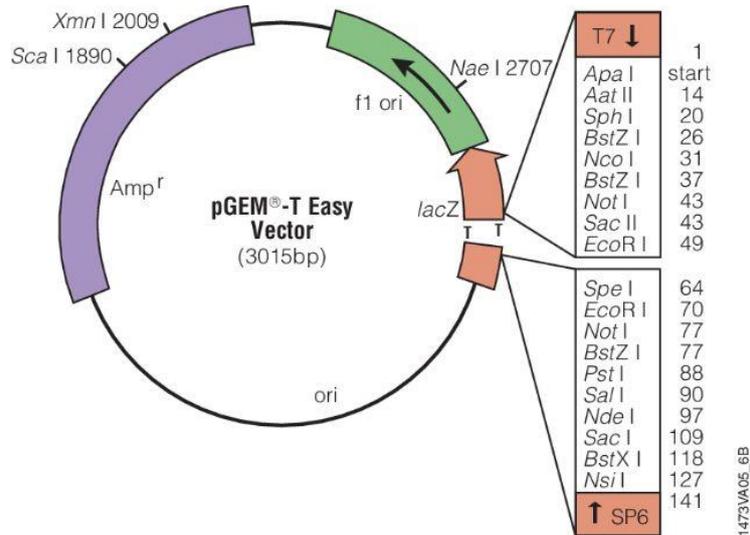


Abbildung 7-3: pGEM-T Easy Vektor.²⁵⁰

Alternativ zum pBluescript II SK (-) TA-Klonierungsvektor (Abbildung 7-2) wurde der kommerziell hergestellte pGEM[®]-T Easy Vektor der Firma Promega zur Subklonierung entsprechend folgendem Ansatz verwendet:

Tabelle 7-18: Ligationsansatz pGEM-T Easy.

Ligationsansatz	
T4-DNA Ligase	1 µl
2x Rapid Ligation Puffer	5 µl
pGEM-T easy	0,5 µl
Insert	x µl
dH ₂ O	ad 10 µl

7.4.4 Restriktion

Zum Restriktionsverdau von DNA wurden je nach Zielsequenz unterschiedliche Endonukleasen der Firma NEB entsprechend ihrer Reaktionsbedingungen verwendet. Die Inkubation erfolgte ebenfalls nach Herstellerangaben, meist bei 37 °C, die Inkubationszeit variierte. Zur Kontrolle von Plasmiden wurde ein Reaktionsansatz von 10 µl, zur Isolierung von Fragmenten ein 50 µl Ansatz verwendet.

7.4.4.1.1 Dephosphorylierung

Um falsch positive Kolonien aufgrund einer Religation des Vektors zu verhindern, wurde der geschnittene Vektor (Kapitel 7.4.4) dephosphoryliert. Dazu wurde Antarctic Phosphatase in folgendem 10 µl Ansatz verwendet und für 45-60 min bei 37 °C inkubiert.

Tabelle 7-19: Dephosphorylierung von Vektor-DNA.

Substanz	Volumen
DNA	8 µl
Antarctic Phosphatase (5 u/µl)	1 µl
10x Puffer Antarctic Phosphatase	1 µl

Nach der Inkubation wurde noch einmal 1 µl Antarctic Phosphatase hinzu gegeben und weitere 45- 60 min bei 37 °C inkubiert. Zur Inaktivierung des Enzyms wurde der Ansatz bei 60 °C für 5 min inkubiert. Anschließend konnte die DNA für Ligationen (Kapitel 5.15) eingesetzt werden.

Tabelle 7-20: Restriktionsansatz.

Substanz	Volumen
DNA	2 µl
Puffer (10x)	1 µl
ggf. BSA (10x)	1 µl
Enzym	0,2-0,5 µl
dH ₂ O	ad 10 µl

7.4.5 Isolierung von Plasmid-DNA aus *E. coli*

Zur Isolierung von Plasmid-DNA aus *E. coli* wurden 5 ml LB-Medium mit dem entsprechenden Stamm (aus Kryokultur oder von einer Agarplatte) unter Antibiotika-Selektion angeimpft. Diese Kultur wurde wie unter 7.3.1 beschrieben inkubiert. Anschließend wurden 1-2 ml der Kultur bei 14000 rpm für 30 s abzentrifugiert und der Überstand verworfen. Die Plasmid-DNA wurde dann mithilfe des „peqGOLD Plasmid Miniprep Kit I (C-Line)“ der Firma PEQLAB bzw. des „GeneJET Plasmid Miniprep Kit“ der Firma Thermo Fisher aufgereinigt.

7.4.6 Sequenzierung

7.4.6.1 Sequenzierung von kurzen DNA-Abschnitten

Die Sequenzierungen von kleinen DNA-Fragmenten bzw. -Abschnitten (ca. 1000 bp) wurde von der Firma GATC mittels Sanger-Sequenzierung durchgeführt. Dazu wurden entweder eigene Primer innerhalb der PCR-Produkte verwendet oder Standardprimer der Firma GATC.

7.4.6.2 Genomsequenzierung

Die Genomsequenzierung von 'Entotheonella factor' wurde am „Center for Biotechnology“ der Universität Bielefeld von Dr. Christian Rückert unter der Leitung von Prof. Dr. Kalinowski durchgeführt.

Zur Sequenzierung wurden folgende Methoden angewandt:

1. 454-Pyrosequenzierung
2. MiSeq-Sequenzierung (Illumina)
3. PacBio

7.4.6.3 Amplikon Sequenzierung (durch die Firma MR DNA)^{223,251–256}

Zur Amplikon-Sequenzierung wurden je 20 µl der aufgereinigten metagenomischen DNA der Bodenproben 5, 26, und I an die Firma MR DNA (Shallowater, TX, USA) gesendet. Das im Folgenden geschilderte Prozedere wurde von der Firma MR DNA durchgeführt. Zunächst wurde eine 30 Zyklen PCR mit den universellen 16S rRNA Primer durchgeführt. Hierzu wurde das HotStarTaq Plus Master Mix Kit (Quiagen, USA) unter folgenden Bedingungen verwendet:

Tabelle 7-21: PCR-Programm zur Amplikon-Sequenzierung.

HotStarTaq			
	Temperatur	Dauer	
Denaturierung	94 °C	3 min	} 28 Zyklen
Denaturierung	94 °C	30 s	
Annealing	53 °C	40 s	
Elongation	72 °C	1 min	
Elongation	72 °C	1 min	
Abkühlen	4 °C	~	

Nach Überprüfung der PCR auf einem 2%igen Agarosegel wurden gleiche Mengen mehrere Proben zusammengefasst (basierend auf dem Molekulargewicht und der Konzentration der DNA). Diese Mischproben wurden dann mithilfe von „Ampure XP beads“ aufgereinigt. Mit den so auf gereinigten PCR Produkte wurde schließlich nach dem „Illumina TrueSeq DNA library preparation“-Protokoll eine DNA Bibliothek erstellt. Die Sequenzierung erfolgte an einem MiSeq-Sequenziersystem der Firma Illumina. Nach Bereinigung der Sequenzen wurden die Sequenzen in die verschiedenen OTUs eingeteilt (97% Sequenzübereinstimmung). Anschließend erfolgte die taxonomische Einordnung über BLASTn gegen entsprechende Datenbanken von GreenGenes, RDP II und NCBI.^{257–259}

7.5 Proteinexpression

7.5.1 Sequenzalignment zur Festlegung der Domänensequenz

Wie alle Enzymdomänen der Multienzymkomplexe NRPS und PKS sind auch die Sequenzen der Adenylierungsdomänen nicht in einzelnen Genen kodiert. Da die Sequenz also nicht durch Start- und Stopcodon flankiert wird, wurde zur Festlegung der Sequenz ein Alignment mit anderen NRPS-A-Domänen durchgeführt werden. Dazu wurde die Software der Firma Geneious benutzt.

7.5.2 Expression der A-Domänen

Zunächst wurde eine 5 ml Übernachtskultur entweder von einer Agarplatte oder aus einer Kryokultur angeimpft. Mit dieser Übernachtskultur wurde am nächsten Morgen 500 ml (bzw. 200 ml bei einer Testexpression) TB-Medium inokuliert. Die dabei verwendeten Schikanekolben hatten mindestens doppeltes Kulturvolumen zur besseren Sauerstoffzufuhr. Die Kultur wurde dann bei 170-200 rpm und 37 °C bis zu einer OD₆₀₀ von 1,5-1,7 kultiviert. Dann wurde die Kultur auf 16 °C herunter gekühlt und die Expression durch Zugabe von 1 mM IPTG (Endkonzentration) induziert. Anschließend wurde Kultur für 15-20 Stunden bei 16°C und 170-200 rpm inkubiert. Schließlich wurde die Zellen bei 4 °C und 5000 x g für 5 min pelletiert in Lysepuffer resuspendiert (ca. 2,5 ml pro Gramm Zellpellet) und bis zur weiteren Verwendung bei -20°C gelagert.

7.5.3 Zellaufschluss

Zum Aufschluss der Zellen wurden diese bei 4°C mit einem Ultraschallhomogenisator behandelt. Die 5-10 Ultraschallintervalle von je 10 sec wurden durch 30-sekündige Phasen auf Eis unterbrochen, um eine Erwärmung des Lysats und die eventuell damit verbundene Degradation von Proteinen zu verhindern. Um eventuell störende Nukleinsäuren aus dem Lysat zu entfernen, wurde dieses mit einer Benzonase-Nuklease (25 units pro Milliliter Lysat) behandelt und für 30 min auf Eis inkubiert. Anschließend wurden die unlöslichen Zellfragmente von gelösten Zellbestandteilen, wie den Proteinen, durch zwei Zentrifugationsschritte bei 15000 x g für je 15 min getrennt.

7.5.4 Aufreinigung von His-Tag-Fusionsproteinen

Die Fusion von Gensequenzen mit Histidin-Resten erlaubt eine einfache Aufreinigung des darauf kodierten Proteins nach dessen Expression. Dazu wird die Gensequenz in einen Expressionsvektor vor oder hinter einen darin enthaltenen Histidin(His)-Tag kloniert. Die hohe Bindungsaffinität von Histidin an Cu²⁺, Ni²⁺, Co²⁺ oder Zn²⁺ erlaubt nach der Expression der der Fusionsproteine eine schnelle und einfache Aufreinigung mittels Metallchelat-Chromatographie.^{260,261}

In dieser Arbeit wurden die Proteine mithilfe von Ni-NTA-Agarose selektiv aufgereinigt. Dazu wurde der Überstand der lysierten Expressionskulturen (Abschnitt 7.5.3) mit 200 µl (Testexpression) bzw. 1000 µl versetzt und für 1 h und 100 min⁻¹ auf Eis inkubiert. Während dieser Inkubationszeit kann der His-Linker an die ihrerseits am Säulenmaterial gebundenen Ni²⁺-Ionen binden. Nach Äquilibrieren der Polyprep-Säule mit Lysepuffer wurde die Suspension aus Lysat und Ni-NTA-Agarose auf die Säule gegeben. Dann wurde die Säule mit je 2 ml (Testexpression) bzw. 20 ml Lysepuffer, 2 bzw. 20 ml Waschpuffer I und 2 bzw. 20 ml Waschpuffer II (Tabelle 7-22) gewaschen. Schließlich wurde mit je 300 µl Elutionspuffer zehnmal eluiert.

Tabelle 7-22: Puffer zur His-Tag Aufreinigung von Proteinen

Lysepuffer	
NaH ₂ PO ₄	25 mM
NaCl	500 mM
Imidazol	10 mM
Glycerin	10% (v/v)
pH = 8,0 (HCl)	
Waschpuffer I	
NaH ₂ PO ₄	25 mM
NaCl	500 mM
Imidazol	20 mM
Glycerin	10% (v/v)
pH = 8,0 (HCl)	
Waschpuffer II	
NaH ₂ PO ₄	25 mM
NaCl	500 mM
Imidazol	40 mM
Glycerin	10% (v/v)
pH = 8,0 (HCl)	
Elutionspuffer	
NaH ₂ PO ₄	25 mM
NaCl	500 mM
Imidazol	300 mM
Glycerin	10% (v/v)
pH = 8,0 (HCl)	

7.5.5 Entsalzen und Aufkonzentrieren von Proteinen

Zur Entsalzung und Umpufferung der eluierten Proteinfractionen (Abschnitt 7.5.4) wurden die sogenannten PD-10 Desalting Columns der Firma GE Healthcare verwendet. Dabei wurde das im Handbuch beschriebene „gravity protocol“ verwendet.

Zur Aufkonzentrierung der Proteinfractionen wurden die Zentrifugationssäulen „Vivaspin 500“ der Firma Satorius verwendet. Diese besitzen eine vertikale PES-Membran mit unterschiedlichen Porengröße (MWCO). Die Porengröße wurde so gewählt, dass diese maximal die halbe zu erwartende Proteingröße erreicht.

7.5.6 SDS-Polyacrylamidgelelektrophorese (SDS-PAGE)

Zur Analyse der Proteinaufreinigung (Abschnitt 7.5.4) wurde die sogenannte SDS-Polyacrylamidgelelektrophorese (SDS-PAGE) verwendet. Bei dieser Methode werden die Proteine aufgrund ihrer Molekularmasse aufgetrennt²⁶². SDS (sodium dodecyl sulfat), ein anionisches Detergenz, hebt die hydrophoben Wechselwirkungen innerhalb eines Proteins auf, sodass Quartär-, Tertiär- und Sekundärstrukturen verloren gehen und das Protein linearisiert. Disulfidbrücken werden bereits zuvor durch die im Probenpuffer befindlichen niedermolekularen Thiole wie z. B. β -Mercaptoethanol (β -ME) reduziert.

Dazu wurden 10-20 μ l sowohl der Elutionsfraktionen als auch der Waschfraktionen (Abschnitt 7.5.4) 1:1 mit „Laemmli sample buffer“ der Firma Biorad (nach Zugabe von β -ME laut Protokoll) versetzt, gemischt und für 5 min bei 95 °C inkubiert. Die Proben wurden dann auf ein NuPAGE™ 10% Bis-Tris Protein Gel (1 mm, 10-15 wells) der Firma *Thermo Fisher Scientific* aufgetragen. Als Größenstandard wurde der Marker „PageRuler unstained“ (*Thermo Fisher Scientific*) mit aufgetragen. Das SDS-Gel wurden dann vertikal in die entsprechende Gelkammer gegeben und diese mit NuPAGE™-SDS-Laufpuffer gefüllt. Zur Trennung der Proteine wurde eine Spannung von 180 V angelegt. Anschließend wurde das Gel in einer Comassie-Färbelösung (10% (v/v) Eisessig, 45% (v/v) Methanol, 0,25% (v/v) Brilliant-Blau) für 1 h inkubiert. Dann wurde das Gel in eine Entfärbelösung (10% (v/v) Eisessig, 45% (v/v) Methanol) überführt, für zwei Stunden entfärbt und schließlich über Nacht in destilliertem Wasser inkubiert. Zur Dokumentation wurde das Gel in Klarsichtfolie eingeschweißt.

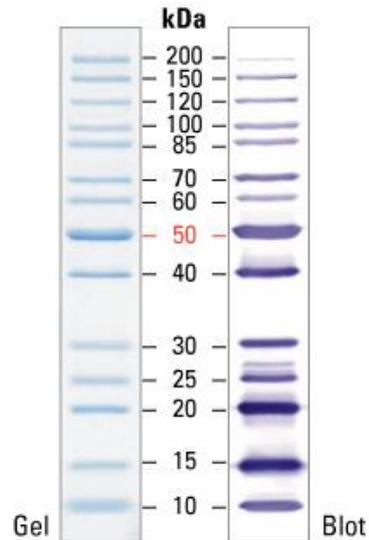


Abb. 7-1 "PageRuler unstained" Proteinmarker (*Thermo Fisher Scientific*)

7.5.7 Proteinkonzentrationsbestimmung

Die Proteinkonzentrationen wurden photometrisch an einem NanoDrop bzw an einem Photometer bestimmt

7.5.8 Bestimmung der Spezifität von A-Domänen: $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assay

Um die Spezifität der jeweiligen A-Domäne bestimmen zu können wurde diese mit je einem potentiellen Substrat getestet. Die Ansätze wurden nach folgendem Pipettierschema in 200 μl Reaktionsgefäße pipettiert:

- 2 μl $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Lösung
(3 mM $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$, 15 mM MgCl_2 , 20 mM Tris-HCl (pH 7,5))
- 2 μl Aminosäure-Lösung
(3mM des Testsubstrats, 15 mM Na-Pyrophosphat (PPi), 20 mM Tris HCl (pH 7,5))
- 2 μl der aufgereinigten und aufkonzentrierten A-Domäne in Assaypuffer

Tabelle 7-23: Assaypuffer des $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assay

Assaypuffer	
NaH_2PO_4	25 mM
NaCl	500 mM
Glycerin	10% (v/v)
pH = 8,0 (HCl)	

Der Ansatz wurde dann zwei Stunden bei Raumtemperatur inkubiert. Zum Abstoppen der Reaktion wurden 6 µl 9-Aminoacridin (in Aceton, 10 mg/ml), ein Matrixmaterial zur MALDI-TOF MS Analyse zugegeben.

7.5.9 Massenspektrometrische Untersuchung von Proteinen

Sowohl die Analyse der Expression des Genclusters (Kapitel 7.5.10), als auch die Auswertung des Spezifitätsassays (Kapitel 7.5.8) wurden in einem matrix-assisted laser desorption/ionization – time of flight-(MALDI-TOF) Massenspektrometer (Biflex 2) der Firma Bruker durchgeführt. Als Matrix wurde 9-Aminoacridin bzw. α -Cyano-4-hydroxycimtsäure (HCCA) verwendet.

7.5.10 Expression des gesamten NRPS/PKS Genclusters

Zur Expression des gesamten NRPS/PKS Genclusters wurde das Cosmid pPD7E4 einer 'P. paederii' Cosmid-Bibliothek in den *E. coli* Expressionsstamm BL21 kloniert. Anschließend wurde mit diesem Klon eine 500 ml Kultur angeimpft und für 48 h bei 37 °C und 180 rpm inkubiert. Nach dem Abzentrifugieren der Zellen bei 6800 x g für 5 min wurde sowohl der Überstand als auch das Zellpellet 100 ml Ethylacetat (v/v) extrahiert. Anschließend wurde die organische Phase einrotiert. Die entstandene Interphase wurde gefriergetrocknet. Nach Lösen in Methanol wurden die Proben im MALDI-TOF analysiert.

7.6 Untersuchung von Bodenproben

Bei den in dieser Arbeit verwendeten Bodenproben handelt es sich nicht um sogenannte ungestörte Proben wie sie beispielsweise mithilfe eines Stechzylinders gewonnen werden, sondern um sogenannte gestörte Beutelproben (Abbildung 10-15). Dadurch kann es bei der Bodenbestimmung anhand der Korngröße durch Siebung zu Verfälschungen kommen. Zeit und Ort der Probenentnahme sind in Tabelle 5-1 aufgeführt. Durchgeführt wurde die Korngrößenbestimmung von Camilla Kurt (Steinmann Institut Geologie, Universität Bonn). Eingeordnet und bestimmt wurden die Bodenarten anhand der Bodenkundlichen Kartieranleitung.²⁶³

7.6.1 Aufreinigung metagenomischer DNA aus Umweltproben

Die metagenomische DNA der Proben wurde mittels „PowerSoil®DNA Isolation Kit“ der Firma MO BIO Laboratories Inc. nach Herstellerangaben aufgereinigt. Dabei wurde für den mechanischen Zellaufschluss ein Flachbett-Vortexer verwendet. Die „PowerBeadTubes“ wurden mit Klebeband fixiert. Die aufgereinigte DNA wurde mit 100 µl eluiert bei -20 °C gelagert. Der Erfolg der Aufreinigung wurde mittels Gelelektrophorese überprüft.

7.6.2 Nachweis von '*Candidatus Entotheonella*' mittels PCR

Zur Überprüfung, ob die vorhandenen Bodenproben '*Candidatus Entotheonella*' enthält, wurde die isolierte metagenomische DNA mit spezifischen Primern untersucht. Dazu wurde zunächst die gesamte metagenomische 16 S DNA mit den universellen Primern U-rRNA-27-F und U-rRNA-1492-R^{210,264} amplifiziert und in einer nachfolgenden PCR mit *Entotheonella* spp. spezifischen Primern analysiert. Dabei wurde 1 µl der ersten PCR als Templat eingesetzt.

7.7 Isolierung von Bakterien aus Umweltproben

Um die Bakterien des neuen Kandidatenphylums '*Tectomicrobia*' näher charakterisieren zu können, sollten diese aus den gesammelten Proben isoliert werden.

7.7.1 Isolierung aus Bodenproben

Zur Isolierung der Bakterien aus den Bodenproben, musste das Material zunächst in einer isotonischen Lösung (0,9% NaCl-Lösung) gelöst werden. Dazu wurde die Suspension für ca. 30 min auf einem Magnetrührer gerührt, 10 min sedimentieren lassen und der Überstand schließlich über ein Nytex Gewebe (Porengröße 0,32 µm) dekantiert. Anschließend wurde der Rückstand nochmals für 10 min in 0,9%iger NaCl-Lösung suspendiert und 10 min sedimentieren gelassen. Nach dem Dekantieren über das Nytex Gewebe, wurde der Überstand in 50 ml Reaktionsgefäß abgefüllt und für 10 min bei 4 °C und 100 x g zentrifugiert. Der Überstand wurde in ein neues 50 ml Reaktionsgefäß überführt und für weitere 10 min bei 4 °C und 1000 x g zentrifugiert. Auch dieser Überstand wurde wiederum in ein neues 50 ml Reaktionsgefäß überführt und schließlich für 10 min bei 4 °C und 4500 x g zentrifugiert. Alle aus den Zentrifugationsschritten resultierenden Pellets wurden in 5 ml 0,9%iger NaCl-Lösung resuspendiert.

Um weitere Verunreinigungen durch kleinste Erdpartikel zu eliminieren, wurde anschließend eine Dichtegradientenzentrifugation durchgeführt. Wie in Kapitel 7.7.1 beschrieben, wird in diesem physikalischen Verfahren die Partikel aufgrund ihrer unterschiedlichen Sedimentationseigenschaften in einem Dichtegradientenmedium voneinander getrennt. Dazu wurde in dieser Arbeit eine Nycodenzlösung (60 % (w/v) der Firma SERVA verwendet. Unter 1 ml der Zellsuspension wurde mithilfe einer Injektionsnadel 1 ml der Nycodenzlösung (1,3 g/ml) gegeben. Anschließend wurde die Probe bei 4°C und 14000 x g für 30-90 min zentrifugiert. Die dabei entstandene Interphase konnte anschließend abgenommen und die darin enthaltenen Zellen isoliert werden.^{220,265} Bei größeren Ansätzen wurden die Volumina entsprechend angepasst.

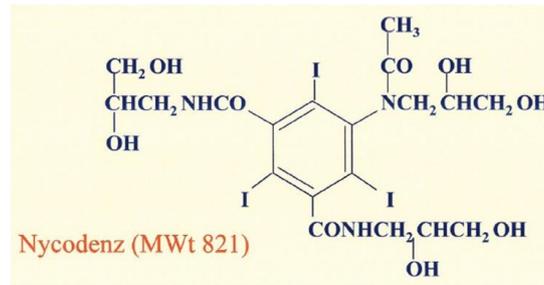


Abbildung 7-4: Dichtegradientenmedium Nycodenz.²⁶⁶

7.8 Mikroskopie

Zur visuellen Charakterisierung der Bakterien wurde das Zeiss Mikroskop Axio Vert. A1 benutzt.

7.8.1 DAPI-Färbung

Um die Bakterien besser von anderem organischen Material unterscheiden zu können, wurden diese mit dem Farbstoff 4',6-Diamidin-2'-phenylindol (DAPI) angefärbt.²⁶⁷ Zunächst wurden dazu die Zellen eines Zellpellets in 3 ml Wasser resuspendiert und 900 µl dieser Zellsuspension mit 100 µl Formaldehyd (37%) gemischt und die Zellen für zwei Stunden bei 4°C fixiert. Zum Färben der Zellen wurden schließlich 7 µl einer DAPI-Lösung (100 µg/ml) zugegeben und diese für 30 min unter Lichtabschluss inkubiert.

7.9 Fluoreszenz in situ Hybridisierung (FISH)

Die Fluoreszenz in situ Hybridisierung dient als molekularbiologischer Nachweis von spezifischen Nukleinsäuren. Dabei binden fluoreszenzmarkierte Oligonukleotide, sogenannte Sonden, an deren komplementäre Zielsequenz. Die so markierten Zielstrukturen (z. B. Chromosome, Zellen) können dann durch Fluoreszenzmikroskopie detektiert werden (Abbildung 7-5). So kann FISH zur Identifizierung, Quantifizierung und zur Charakterisierung von Mikrobiomen beitragen.^{133,268}

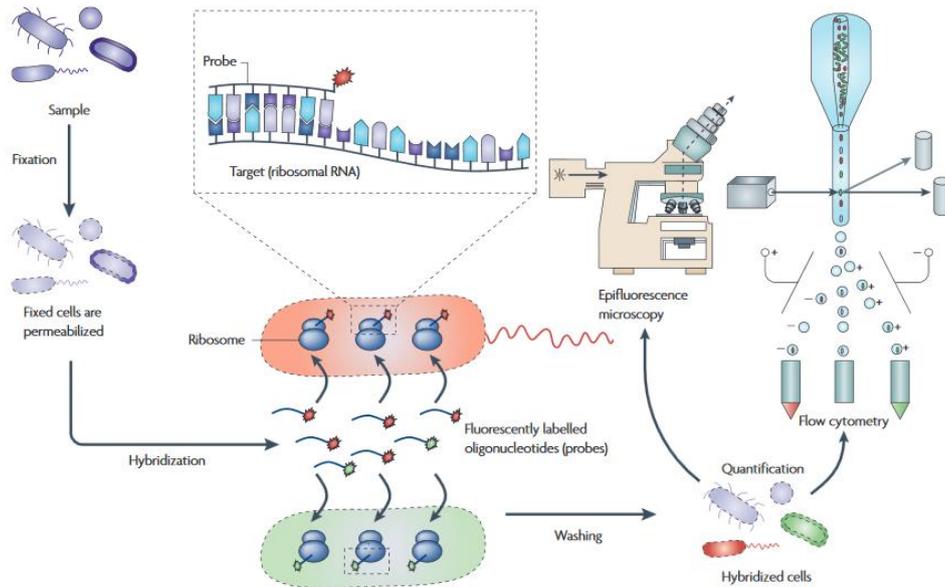


Abbildung 7-5: Fluoreszenz-in-situ-Hybridisierung (FISH).¹³³

7.9.1 Catalyzed reporter deposition (CARD)-FISH

Um eine höhere Sensitivität der FISH zu erreichen kann die sogenannte CARD-FISH Methode angewendet werden. Hierbei werden Oligonukleotidsonden verwendet, an die eine Meerrettichperoxidase (engl.: *horseradish* (HRP)) gebunden ist. Die Signalverstärkung basiert auf einer Radikalisierung zugegebener fluoreszenzmarkierter Tyramid-Moleküle (Abbildung 7-6).¹³³

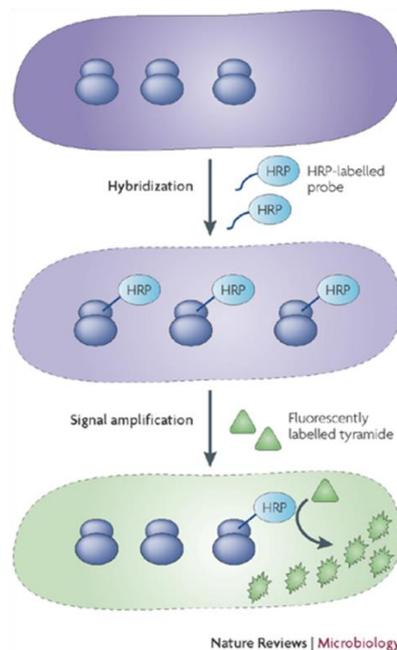


Abbildung 7-6: CARD-FISH.¹³³

Zur Durchführung einer CARD-FISH mussten die aufgereinigten Bakterienzellen zunächst fixiert werden. Dazu wurden zwei verschiedene Methoden angewendet: Fixation mit Paraformaldehyd (PFA) und Fixation mit Ethanol.

Zur PFA-Fixation wurden die Zellen bei 5000 x g für 3 min abzentrifugiert. Nach Abnahme des Überstands wurden 100 µl PFA (4%) zugegeben und für 1 h bzw. über Nacht fixiert. Nach zweimaligem Waschen mit PBS wurden das Zellpellet schließlich in einem PBS/Ethanol-Gemisch (50% v/v) resuspendiert und bei -20°C eingefroren.

Zur Ethanol-Fixation wurden die Zellen bei 5000 x g für 3 min abzentrifugiert. Nach Abnahme des Überstands wurde das Zellpellet in 40 µl dH₂O gelöst. Anschließend wurden 960 µl Ethanol (absolut) zugegeben und der Ansatz über Nacht bei 4°C inkubiert. Nach Waschen mit PBS wurden das Zellpellet schließlich in einem PBS/Ethanol-Gemisch (50% v/v) resuspendiert und bei -20°C eingefroren.

Vor Durchführung der CARD-FISH wurde eine Durchlichtmikroskopie durchgeführt, um die Zelldichte abschätzen und die Zellsuspensionen ggf. verdünnen zu können. Dann wurden je 1 µl der Proben in je ein Reaktionsfeld eines Objektträgers pipettiert und für 20 min trocknen gelassen. Zur Permeabilisierung der Zellmembranen wurde 0,5 mg/ml Lysozym (in Lysozypuffer) hinzugegeben und für 10 bzw. 30 min bei RT inkubiert. Durch Zugabe der Inaktivierungslösung (Tabelle 7-25) und Inkubation für 10 min bei RT wurden endogene Peroxidasen inaktiviert. Anschließend wurde der Objektträger dreimal gewaschen (2 min - 1x PBS, 5 min - 1x PBS, 1 min - MilliQ H₂O). Nach nochmaligem Waschen mit 96% igem Ethanol wurde der Objektträger für 20 min bei RT getrocknet. Zur Hybridisierung wurde die DNA-Sonde in Hybridisierungspuffer (Tabelle 7-25) (≤ 5 ng/ μ l) zugegeben und für 2 h bei 35 °C im Dunkeln inkubiert.

Tabelle 7-24: Verwendeten Sonde zur Fluoreszenz-in-situ-Hybridisierung.

FISH-Sonde	CCGCAAGCYCATCTCAGACC
------------	----------------------

Es folgten wiederum zwei Waschschritte (1x Waschpuffer (Tabelle 7-25) für 30 min bei 37°C, 1x PBS-T (0,01% Triton X-100) für 30-45 min bei 37°C). Zur Amplifikation des Sonden-Signals wurde die Amplifikationslösung (1 ml Amplifikationspuffer (Tabelle 7-25), 10 µl H₂O₂ [200 µl 1x PBS, 1 µl 30% H₂O₂], 2 µl markiertes Tyramid) hinzugegeben und für 1 h bei 37 °C inkubiert. Anschließend wurde mit je 1x PBS für 1 min, 2 min, 5 min, und 2x 10 min bei 46°C gewaschen. Schließlich wurde ein weiteres Mal für 1 min mit MilliQ H₂O gewaschen. Der Objektträger wurde dann mit 96%igem Ethanol gespült und bei RT getrocknet.

Die anschließende Auswertung erfolgte an einem Zeiss Axioskop 2 Epifluoreszenz Mikroskop mithilfe eines entsprechenden Filtersets für den Fluoreszenzfarbstoff Alexa 633.

Tabelle 7-25: Puffer CARD-FISH.

Inaktivierungslösung	
HCl (37 %)	250 µl
MilliQ H ₂ O	ad 50 ml
Lysozympuffer	
Tris-HCl (1 M, pH 8)	1 ml
EDTA (0,5 M, pH 8)	1 ml
MilliQ H ₂ O	ad 10 ml
Hybridisierungspuffer	
Formaldehyd	55%
TRIS (1 M, pH 8)	400 µl
Dextransulfat	2 g
SDS (10%)	50 µl
NaCl (5 M)	3,6 ml
Blocking reagent for nucleic acids (10%; Roche)	2 ml
MilliQ H ₂ O	3 ml
<ul style="list-style-type: none"> - Mischen aller Substanzen außer SDS, „blocking reagent“ und Fomaldehyd in einem 50 ml Reaktionsgefäß. - Inkubation im Schüttelinkubator für 1 h bei 48°C. - Zugabe der restlichen Substanzen und Lagerung bei -20 °C 	
Amplifikationslösung	
PBS (10 x)	2 ml
NaCl (5 M)	8 ml
Blocking reagent for nucleic acids (10%; Roche)	200 µl
Dextransulfat	4 g
MilliQ H ₂ O	ad 20 ml
<ul style="list-style-type: none"> - Mischen aller Substanzen außer SDS, „blocking reagent“ und Fomaldehyd in einem 50 ml Reaktionsgefäß. - Inkubation im Schüttelinkubator für 1 h bei 48°C. - Zugabe des „blocking reagent“ und Filtrieren durch einen 0,22 µm Filter. 	

Waschpuffer	
TRIS (1 M, pH 8)	1 ml
EDTA (0,5 M, pH 8)	0,5 ml
SDS (10%)	50 µl
NaCl (5 M)	30 µl
MilliQ H ₂ O	ad 50 ml

7.10 Verwendete Chemikalien

Agar	Roth, Karlsruhe
Agarose NEEO Qualität	Roth, Karlsruhe
Ampicillin-Natriumsalz	Roth, Karlsruhe
L-Alanin	Fluka / Riedel-de-Haën, Seelze
L-Arginin	AppliChem, Darmstadt
L-Asparagin,	AppliChem, Darmstadt
L-Asparaginsäure	Acros Organics, Geel (B
Antarctic Phosphatase + 10x Puffer	NEB, Frankfurt/Main
Benzonase Nuclease	Merck, Darmstadt
Bovin Serum Albumin (BSA-Lösung 10 mg x mL ⁻¹)	NEB, Frankfurt/Main
Brilliant-Blau R250	Roth, Karlsruhe
5-Brom-4-chlor-3-indoxyl-β-D-galactopyranosid (X-Gal)	AppliChem, Darmstadt
Bromphenolblau-Natriumsalz	AppliChem, Darmstadt
Calciumchlorid	KMF, Lohmar
Chloramphenicol	Roth, Karlsruhe
Cycle-Pure Kits, peqGOLD	Peqlab bzw. VWR , Darmstadt
L-4- L-Cystein	Roth, Karlsruhe
Desoxyadenosin-5'-triphosphat (dATP)	Invitrogen, Karlsruhe
Desoxythymidin-5'-triphosphat (dTTP)	Invitrogen, Karlsruhe
Desoxynukleotide (dNTPs)	Invitrogen, Karlsruhe
Dimethylsulfoxid (DMSO)	Roth, Karlsruhe
Dinatriumhydrogenphosphat-Dihydrat	Fluka / Riedel-de-Haën, Seelze
1kb plus DNA ladder	Fermentas, St. Leon-Rot
Eisessig	Merck, Darmstadt
Ethidiumbromid	Roth, Karlsruhe
Ethylendiamintetraessigsäure (EDTA)	KMF, Lohmar
Ethanol absolut	VWR, Darmstadt

Ethylacetat	Acros Organics, Geel (B)
Expand High FidelityPLUS PCR System	Roche, Grenzach-Wyhlen
Gel Green	Biotium, Hayward (USA)
GeneJET PCR Purification Kit	ThermoFisher Scientific, Schwerte/Darmstadt
GeneJET Plasmid Miniprep Kit	ThermoFisher Scientific, Schwerte/Darmstadt
D(+)-Glucose Monohydrat	Fluka / Riedel-de-Haën, Seelze
4',6-Diamidin-2'-phenylindol (DAPI)	Roth, Karlsruhe
Laemmli Sample Buffer	Bio-Rad, München
L-Glutamin	AppliChem, Darmstadt
L-Glutaminsäure-Mononatriumsalz	AppliChem, Darmstadt
Glycerin	Merck, Darmstadt
Glycin	Fisher-Scientific, Schwerte
Hefeextrakt	Becton Dickinson, Heidelberg
L-Histidin	AppliChem, Darmstadt
Imidazol	Roth, Karlsruhe
L-Isoleucin	Roth, Karlsruhe
Isopropylthiogalactosid (IPTG)	Roth, Karlsruhe
Kaliumacetat	ABCR; Karlsruhe
Kaliumdihydrogenphosphat	Roth, Karlsruhe
Dikaliumhydrogenphosphat	Roth, Karlsruhe
Kaliumchlorid	Acros Organics, Geel (B)
Kanamycin-Sulfat	Roth, Karlsruhe
α -Ketoisocaprionsäure	Sigma-Aldrich, Seelze
α -Ketoisovaleriansäure	Sigma-Aldrich, Seelze
α -Ketopropionsäure	Sigma-Aldrich, Seelze
L-Leucin	Sigma-Aldrich, Seelze
L-Lysin	Sigma-Aldrich, Seelze
Magnesiumchlorid-Hexahydrat	Fluka / Riedel-de-Haën, Seelze
Manganchlorid-Dihydrat	Merck, Darmstadt
L-Phenylalanin	Sigma-Aldrich, Seelze
Methanol	VWR, Darmstadt
Midori Green Direct	Nippon Genetics, Düren
Morpholinpropansulfonsäure (MOPS)	Sigma-Aldrich, Seelze

Natriumacetat	Acros Organics, Geel (B)
Natriumchlorid	Roth, Karlsruhe
Natriumdodecylsulfat (SDS)	Roth, Karlsruhe
Natriumhydroxid	KMF, Lohmar
Ni-NTA Agarose	Macherey/Nagel, Duisburg
NuPAGE® MOPS SDS Running Buffer (20X)	ThermoFisher Scientific, Schwerte/Darmstadt
Nycodenz®, 60 % (w/v) solution in water	Serva Electrophoresis GmbH, Heidelberg
Orange G	Sigma-Aldrich, Seelze
pGEM-Teasy Kit	Promega, Mannheim
PageRuler™ Unstained Protein Ladder	ThermoFisher Scientific, Schwerte/Darmstadt
L-Phenylalanin	AppliChem, Darmstadt
Paraformaldehyd	Roth, Karlsruhe
peqGOLD Plasmid Miniprep Kit I	Peqlab bzw. VWR, Darmstadt
Primer (DNA-Oligonucleotide)	Invitrogen, Karlsruhe
L-Prolin	Acros Organics, Geel (B)
Prolin	Alfa Aesar, Karlsruhe
Phusion® High-Fidelity DNA Polymerase	NEB, Frankfurt/Main
PowerSoil®DNA Isolation Kit	MoBio bzw. Quiagen, Hilden
Restriktionsenzyme + 10x Puffer + 100x BSA	NEB, Frankfurt/Main
Salzsäure	Fluka / Riedel-de-Haën, Seelze
L-Serin	Roth, Karlsruhe
T4 DNA Ligase + 10x Puffer	Jena Bioscience, Jena/ NEB, Frankfurt/Main
<i>Taq</i> DNA Polymerase + 10x Puffer	NEB, Frankfurt/Main
L-Threonin	Roth, Karlsruhe
Tris-(hydroxymethyl)-aminomethan (Tris)	Roth, Karlsruhe
Trypton	Becton Dickinson, Heidelberg
L-Tryptophan	Acros Organics, Geel (B)
L-Tyrosin	Roth, Karlsruhe
L-Valin	AppliChem, Darmstadt

7.11 Verwendete Geräte und Materialien

Agarosegelelektrophorese-Kammer +	
Spannungsquelle Standard Power Pack P25	Biometra, Göttingen
Analysenwaage CP225D	Sartorius, Göttingen
Autoklav V65	Systemec, Wettenberg
Elektroporationsküvette 2mm	Bio-Rad, München
Elektroporator MicroPulser	Bio-Rad, München
Geldokumentation Genius	Syngene, Cambridge (UK)
Geldokumentation FAS Digi	Nippon Genetics, Düren
Inkubationsschüttler Certomat BS-1	Sartorius, Göttingen
Inkubator B12	Thermo, Langenselbold
Mikroküvette UltraVette (70-850 µl)	Roth, Karlsruhe
Mikrowelle	Lifetec Medion, Essen
Mikrozentrifuge ungekühlt Mikro200	Hettich, Tuttlingen
Mikrozentrifuge, gekühlt 5417R	Eppendorf, Hamburg
Mikrozentrifuge, gekühlt Mikro200R	Hettich, Tuttlingen
NuPAGE™ 10% Bis-Tris Protein Gel	ThermoFisher Scientific, Schwerte/Darmstadt
Standard PowerPack P25	
Photometer (Biofotometer)	Eppendorf, Hamburg
Photometer BioMate 3	Thermo Electron, Cambridge(UK)
Pipetten Pipetman P2 - P10 mL	Gilson, Middleton (USA)
Poly-Prep Säule (leer)	Bio-Rad, München
Speedvac / Concentrator 5301	Eppendorf, Hamburg
Sterilbank Biowizard	Kojair, Vilppula (FIN)
Sterilfilter (0,2 µm, Celluloseacetat, FP 30/0,2),	Whatman, Dassel
Thermocycler T-Gradient	Biometra, Göttingen
Thermomixer comfort	Eppendorf, Hamburg
Tischzentrifuge Rotina	35R Hettich, Tuttlingen
Tischzentrifuge Z513K	Hermle, Wehingen
Ultraschallhomogenisator Sonopuls HD2070	Bandelin, Berlin
UV-Crosslinker CL1000 UVP	Cambridge (UK)
Vortex-Mixer VTX-3000L	LMS, Tokio (J)
Waage 440-47N	Kern, Balingen-Frommern
Waage BP110	Sartorius, Göttingen
Wasserbad	GFL, Burgwedel

XCell SureLock™ Mini-Cell Electrophoresis System

ThermoFisher Scientific,
Schwerte/Darmstadt

Zentrifugaleinheiten Vivaspin 500 (versch. MWCO)

Sartorius, Göttingen

8 Abkürzungsverzeichnis

A	Adenylierungs-Domäne
ACP	Acyl-Carrier-Protein
ADP	Adenosindiphosphat
Ala	Alanin
AMP	Adenosinmonophosphat
AMP	Ampicillin
Arg	Argenin
AS	Aminosäur
AT	Acyltransferase
Asn	Asparagin
Asp	Asparaginsäure
ATP	Adenosintriphosphat
BLAST	basic local alignment search tool
bp	Basenpaar(e)
BSA	Bovine Serum Albumin
bzw.	Beziehungsweise
C	Kondesations-Domäne
C	Cytosin
Cys	Cystein
Cam	Chloramphenicol in der Endkonzentration 12,5 µg/mL
CoA	Coenzym A
dATP	Desoxyadenosintriphoshat
dNTP	Desoxynukleotidtriphosphat
dTTP	Desoxythymidintriphosphat
DH	Dehydratase
DNA	Desoxyribonukleinsäure (desoxyribonucleic acid)
dNTP	Desoxynukleotid
dTTP	Desoxythymintriphoshat
EDTA	Ethylendiamin-tetraessigsäure (ethylene diamine tetracetic acid)
engl.	englisch
ER	Enoylreduktase
Fa.	Firma
FAS	Fettsäuresynthase (fatty acid synthase)
Gly	Glycin
G	Guanosin
Gln	Glutamin
Glu	Glutaminsäure
griech.	griechisch
h	Stunden
dH ₂ O	destilliertes (<i>hier</i> : deionisiertes) Wasser
Ile	Isoleucin
IPTG	Isopropylthiogalactosid
kb	Kilobasen
KS	Ketosynthase

Kn	Kanamycin
KR	Ketoreduktase
lat.	Lateinisch
LB Medium	Luria-Bertani Medium
Leu	Leucin
Lys	Lysin
MALDI	matrix-assisted laser desorption/ionization
Mb	Megabasen
Met	Methionin
min	Minuten
MLP	MbtH-like protein
MT	Methyltransferase
MWCO	molecular weight cut-off
NRPS	Nichtribosomale Peptidsynthetase
OD	Optische Dichte
PCP	Peptidyl-Carrier-Protein
PCR	Polymerase-Kettenreaktion
Phe	Phenylalanin
PI	Pairwise identity
PKS	Polyketidsynthase
PPTase	Phosphopantethein-Transferase
RNA	Ribonukleinsäure (Ribonucleic acid)
RNase A	Ribonuklease A
rpm	Drehzahl (revolutions per minute)
RT	Raumtemperatur
RT	Reverse Transkriptase
SDS	Natriumdodecylsulfat (sodiumdodecylsulphate)
SDS-PAGE	sodium dodecylsulfate polyacrylamide gelelectrophoresis
s	Sekunden
Ser	Serin
T	Thymin
Thr	Threonin
TE	Thioesterase
TOF	time of flight
Tris-HCl	Tris-(hydroxymethyl)-aminomethan, pH-Wert eingestellt mit Salzsäure (HCl)
Trp	Tryptophan
Tyr	Tyrosin
u. a.	unter anderem
u. U.	unter Umständen
UV	Ultraviolett
Val	Valin
vgl.	Vergleiche
X-Gal	5-Brom-4-chlor-3-indoxyl- β -D-galactopyranosid

9 Literaturverzeichnis

1. Fleming, A. On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their Use in the Isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **10**, 226 (1929).
2. Wainwright, M. Moulds in ancient and more recent medicine. *Mycologist* **3**, 21–23 (1989).
3. World Health Organization. *Antimicrobial Resistance-Global Report on Surveillance*. WHO Press (2014).
4. Silver, L. L. Challenges of Antibacterial Discovery. *Clin. Microbiol. Rev.* **24**, 71–109 (2011).
5. Livermore, D. M. *et al.* Discovery research: the scientific challenge of finding new antibiotics. *J. Antimicrob. Chemother.* **66**, 1941–4 (2011).
6. Hertweck, C. Die biosynthetische Grundlage der Polyketid-Vielfalt. *Angew. Chemie* **121**, 4782–4811 (2009).
7. Staunton, J. & Weissman, K. J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **18**, 380–416 (2001).
8. Cane, D. E. Introduction: Polyketide and Nonribosomal Polypeptide Biosynthesis. From *Collie* to *Coli*. *Chem. Rev.* **97**, 2463–2464 (1997).
9. Birch, A. J. Biosynthesis of polyketides and related compounds. *Science* **156**, 202–6 (1967).
10. Katz, L. & Ashley, G. Translation and Protein Synthesis: Macrolides. *Chem. Rev.* **105**, 499–528 (2005).
11. Lemke, A., Kiderlen, A. F. & Kayser, O. Amphotericin B. *Appl Microbiol Biotechnol* **68**, 151–162 (2005).
12. Folkers, K. *et al.* Lovastatin decreases coenzyme Q levels in humans. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 8931–4 (1990).
13. Grimm, A., Madduri, K., Ali, A. & Hutchinson, C. R. Characterization of the *Streptomyces peucetius* ATCC 29050 genes encoding doxorubicin polyketide synthase. *Gene* **151**, 1–10 (1994).
14. Schwecke, T. *et al.* The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7839–43 (1995).
15. Cane, D. E. & Walsh, C. T. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem. Biol.* **6**, R319–R325 (1999).
16. Shen, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr. Opin. Chem. Biol.* **7**, 285–295 (2003).
17. Piel, J. *et al.* Exploring the chemistry of uncultivated bacterial symbionts: Antitumor polyketides of the Pederin family. in *Journal of Natural Products* **68**, 472–479 (2005).
18. Steffens, U. Genetische Studien an einer ungewöhnlichen trans -AT Polyketidsynthase eines Cyanobionten. (Rheinische Friedrich-Wilhelms-Universität Bonn, 2011).
19. Heath, R. J. & Rock, C. O. The Claisen condensation in biology. *Nat. Prod. Rep.* **19**, 581–596 (2002).
20. McDaniel, R., Welch, M. & Hutchinson, C. R. Genetic approaches to polyketide antibiotics. 1. *Chem. Rev.* **105**, 543–558 (2005).
21. Hopwood, D. A. Cracking the Polyketide Code. *PLoS Biol.* **2**, e35 (2004).
22. Sattely, E. S., Fischbach, M. a & Walsh, C. T. Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. *Nat. Prod. Rep.* **25**, 757–793 (2008).
23. Lai, J. R., Koglin, A. & Walsh, C. T. Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. *Biochemistry* **45**, 14869–14879 (2006).

24. Rawlings, B. J. Biosynthesis of fatty acids and related metabolites. *Nat. Prod. Rep.* **15**, 275–308 (1998).
25. Rawlings, B. J. Type I polyketide biosynthesis in bacteria (Part B) (1995 to mid-2000). *Nat. Prod. Rep.* **18**, 231–281 (2001).
26. Du, L. & Lou, L. PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **27**, 255–78 (2010).
27. Rix, U., Fischer, C., Remsing, L. L. & Rohr, J. J. Modification of post-PKS tailoring steps through combinatorial biosynthesis. *Nat. Prod. Rep.* **19**, 542–580 (2002).
28. Katz, L. & Donadio, S. Polyketide Synthesis : Prospects for Hybrid Antibiotics. *Annu. Rev. Microbiol* **47**, 875–912 (1993).
29. Cheng, Y.-Q., Tang, G.-L. & Shen, B. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3149–3154 (2003).
30. Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat. Prod. Rep.* **27**, 996 (2010).
31. Piel, J. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14002–14007 (2002).
32. Crawford, J. M., Dancy, B. C. R., Hill, E. A., Udvary, D. W. & Townsend, C. A. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc. Natl. Acad. Sci.* **103**, 16728–16733 (2006).
33. Evans, B. S., Robinson, S. J. & Kelleher, N. L. Surveys of non-ribosomal peptide and polyketide assembly lines in fungi and prospects for their analysis in vitro and in vivo. *Fungal Genet. Biol.* **48**, 49–61 (2011).
34. Cox, R. J. & Simpson, T. J. *Methods in Enzymology.* **459**, (2009).
35. Campbell, C. D. & Vederas, J. C. Mini Review: Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers* **93**, 755–763 (2010).
36. Hendrickson, L. *et al.* Lovastatin biosynthesis in *Aspergillus terreus*: Characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chem. Biol.* **6**, 429–439 (1999).
37. Hopwood, D. A. Genetic Contributions to Understanding Polyketide Synthases. *Chem. Rev.* **97**, 2465–2498 (1997).
38. Zhang, W. & Tang, Y. *In vitro analysis of type II polyketide synthase. Methods in enzymology* **459**, (2009).
39. Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chemie Int. Ed.* **48**, 4688–4716 (2009).
40. Rawlings, B. J. Biosynthesis of polyketides (other than actinomycete macrolides). *Nat. Prod. Rep.* **16**, 425–484 (1999).
41. Austin, M. B. & Noel, J. P. The chalcone synthase superfamily of type III polyketide synthases. *Nat Prod Rep* **20**, 79–110 (2003).
42. Austin, M. B., O'Maille, P. E. & Noel, J. P. Evolving biosynthetic tangos negotiate mechanistic landscapes. *Nat. Chem. Biol.* **4**, 217 (2008).
43. Felnagle, E. A. *et al.* Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Molecular Pharmaceutics* **5**, 191–211 (2008).
44. Schwarzer, D., Finking, R. & Marahiel, M. a. Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* **20**, 275–287 (2003).
45. McQuillen, K., Roberts, R. B. & Britten, R. J. Synthesis of nascent protein by ribosomes in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **45**, 1437–47 (1959).
46. Walsh, C. T. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **303**, 1805–1810 (2004).
47. Finking, R. & Marahiel, M. A. Biosynthesis of nonribosomal peptides. *Annu. Rev. Microbiol.* **58**, 453–488 (2004).
48. Meier, J. L. & Burkart, M. D. The chemical biology of modular biosynthetic enzymes. *Chem. Soc. Rev.* **38**, 2012–45 (2009).
49. Marahiel, M. a, Stachelhaus, T. & Mootz, H. D. Modular Peptide Synthetases Involved

- in Nonribosomal Peptide Synthesis. *Chem. Rev.* **97**, 2651–2674 (1997).
50. Challis, G. L., Ravel, J. & Townsend, C. A. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* **7**, 211–224 (2000).
 51. Stachelhaus, T., Mootz, H. D. & Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **6**, 493–505 (1999).
 52. Von Döhren, H., Dieckmann, R. & Pavela-Vrancic, M. The nonribosomal code. *Chemistry and Biology* **6**, (1999).
 53. Awakawa, T., Yang, X. L., Wakimoto, T. & Abe, I. Pyranonigrin E: A PKS-NRPS hybrid metabolite from *Aspergillus niger* identified by genome mining. *ChemBioChem* **14**, 2095–2099 (2013).
 54. Bergmann, S. *et al.* Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol.* **3**, 213–217 (2007).
 55. Tokuoka, M. *et al.* Identification of a novel polyketide synthase-nonribosomal peptide synthetase (PKS-NRPS) gene required for the biosynthesis of cyclopiazonic acid in *Aspergillus oryzae*. *Fungal Genet. Biol.* **45**, 1608–1615 (2008).
 56. Watanabe, K. *et al.* Total biosynthesis of antitumor nonribosomal peptides in *Escherichia coli*. *Nat. Chem. Biol.* **2**, 423–428 (2006).
 57. Donadio, S. & Sosio, M. Biosynthesis of glycopeptides: prospects for improved antibacterials. *Curr. Top. Med. Chem.* **8**, 654–66 (2008).
 58. Survase, S. A., Kagliwal, L. D., Annapure, U. S. & Singhal, R. S. Cyclosporin A - A review on fermentative production, downstream processing and pharmacological applications. *Biotechnology Advances* **29**, 418–435 (2011).
 59. Conti, E., Stachelhaus, T., Marahiel, M. A. & Brick, P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* **16**, 4174–4183 (1997).
 60. Gulick, A. M. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chemical Biology* **4**, 811–827 (2009).
 61. Eppelmann, K., Stachelhaus, T., Marahiel, M. A. & Philipps-uni, V. Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry* **41**, 9718–9726 (2002).
 62. Chen, C.-Y., Georgiev, I., Anderson, A. C. & Donald, B. R. Computational structure-based redesign of enzyme activity. *Proc. Natl. Acad. Sci.* **106**, 3764–3769 (2009).
 63. Otten, L. G., Schaffer, M. L., Villiers, B. R. M., Stachelhaus, T. & Hollfelder, F. An optimized ATP/PPI-exchange assay in 96-well format for screening of adenylation domains for applications in combinatorial biosynthesis. *Biotechnol. J.* **2**, 232–240 (2007).
 64. Phelan, V. V., Du, Y., McLean, J. A. & Bachmann, B. O. *Adenylation Enzyme Characterization Using γ -1804-ATP Pyrophosphate Exchange*. *Chemistry & Biology* **16**, (2009).
 65. Dorrestein, P. C. *et al.* Activity screening of carrier domains within nonribosomal peptide synthetases using complex substrate mixtures and large molecule mass spectrometry. *Biochemistry* **45**, 1537–1546 (2006).
 66. Quadri, L. E., Sello, J., Keating, T. A., Weinreb, P. H. & Walsh, C. T. Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chem. Biol.* **5**, 631–645 (1998).
 67. Baltz, R. H. Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. *Journal of Industrial Microbiology and Biotechnology* **38**, 1747–1760 (2011).
 68. Sosio, M. *et al.* Organization of the teicoplanin gene cluster in *Actinoplanes teichomyceticus*. *Microbiology* **150**, 95–102 (2004).
 69. Drake, E. J. *et al.* The 1.8 Å crystal structure of PA2412, an MbtH-like protein from the pyoverdine cluster of *Pseudomonas aeruginosa*. *J. Biol. Chem.* **282**, 20425–20434

- (2007).
70. Lautru, S., Oves-Costales, D., Pernodet, J. L. & Challis, G. L. MbtH-like protein-mediated cross-talk between non-ribosomal peptide antibiotic and siderophore biosynthetic pathways in *Streptomyces coelicolor* M145. *Microbiology* **153**, 1405–1412 (2007).
 71. Wolpert, M., Gust, B., Kammerer, B. & Heide, L. Effects of deletions of mbtH-like genes on clorobiocin biosynthesis in *Streptomyces coelicolor*. *Microbiology* **153**, 1413–1423 (2007).
 72. Felnagle, E. A. *et al.* MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. *Biochemistry* **49**, 8815–8817 (2010).
 73. Zhang, W., Heemstra, J. R., Walsh, C. T. & Imker, H. J. Activation of the pacidamycin *pacI* adenylation domain by MbtH-like proteins. *Biochemistry* **49**, 9946–9947 (2010).
 74. Imker, H. J., Krahn, D., Clerc, J., Kaiser, M. & Walsh, C. T. N-Acylation during glidobactin biosynthesis by the tridomain nonribosomal peptide synthetase module GlbF. *Chem. Biol.* **17**, 1077–1083 (2010).
 75. Boll, B., Taubitz, T. & Heide, L. Role of MbtH-like proteins in the adenylation of tyrosine during aminocoumarin and vancomycin biosynthesis. *J. Biol. Chem.* **286**, 36281–36290 (2011).
 76. Herbst, D. A., Boll, B., Zocher, G., Stehle, T. & Heide, L. Structural Basis of the Interaction of MbtH-like Proteins, Putative Regulators of Nonribosomal Peptide Biosynthesis, with Adenylating Enzymes. *J. Biol. Chem.* **288**, 1991–2003 (2013).
 77. Du, L., Sánchez, C., Chen, M., Edwards, D. J. & Shen, B. The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. *Chem. Biol.* **7**, 623–642 (2000).
 78. Galm, U. *et al.* Antitumor antibiotics: Bleomycin, enediynes, and mitomycin. *Chemical Reviews* **105**, 739–758 (2005).
 79. Zerikly, M. & Challis, G. L. Strategies for the Discovery of New Natural Products by Genome Mining. *ChemBioChem* **10**, 625–633 (2009).
 80. Van Lanen, S. G. & Shen, B. Microbial genomics for the improvement of natural product discovery. *Current Opinion in Microbiology* **9**, 252–260 (2006).
 81. Bode, H. B. & Müller, R. The impact of bacterial genomics on natural product research. *Angewandte Chemie - International Edition* **44**, 6828–6846 (2005).
 82. Gross, H. Strategies to unravel the function of orphan biosynthesis pathways: Recent examples and future prospects. *Applied Microbiology and Biotechnology* **75**, 267–277 (2007).
 83. Weber, T. *et al.* AntiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **43**, W237–W243 (2015).
 84. Blin, K. *et al.* antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic acids* **41**, W204–W212 (2013).
 85. Medema, M. H. M. *et al.* antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic acids* **39**, W339–46 (2011).
 86. Madigan, M. T., Martinko, J. M., Stahl, D. A. & Clark, D. P. *Brock Biology of Microorganisms. International Microbiology* (2012). doi:10.1038/hr.2014.17
 87. Boucher, D. H., James, S. & Keeler, K. H. The Ecology of Mutualism. *Annu. Rev. Ecol. Syst.* **13**, 315–347 (1982).
 88. Amann, R. I., Ludwig, W. & Schleifer, K. H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**, 143–69 (1995).
 89. Hug, L. A. *et al.* A new view of the tree of life. *Nat. Microbiol.* **1**, 16048 (2016).
 90. Rikkinen, J., Oksanen, I. & Lohtander, K. Lichen guilds share related cyanobacterial symbionts. *Science* **297**, 357 (2002).
 91. Sonnenburg, J. L. Glycan Foraging in Vivo by an Intestine-Adapted Bacterial Symbiont. *Science (80-.)*. **307**, 1955–1959 (2005).
 92. Thomas, F., Hehemann, J.-H., Rebuffet, E., Czjzek, M. & Michel, G. Environmental

- and gut bacteroidetes: the food connection. *Front. Microbiol.* **2**, 93 (2011).
93. Olszak, T. *et al.* Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**, 489–493 (2012).
 94. Möhle, L. *et al.* Ly6Chi Monocytes Provide a Link between Antibiotic-Induced Changes in Gut Microbiota and Adult Hippocampal Neurogenesis. *Cell Rep.* **15**, 1945–1956 (2016).
 95. Crawford, J. M. *et al.* Bacterial symbionts and natural products. *Chem. Commun. (Camb)*. **47**, 7559–7566 (2011).
 96. Kampa, A. *et al.* Metagenomic natural product discovery in lichen provides evidence for a family of biosynthetic pathways in diverse symbioses. *Proc. Natl. Acad. Sci.* **110**, E3129–E3137 (2013).
 97. Joyner, P. M. *et al.* Mutanobactin A from the human oral pathogen *Streptococcus mutans* is a cross-kingdom regulator of the yeast-mycelium transition. *Org. Biomol. Chem.* **8**, 5486 (2010).
 98. Wang, X., Du, L., You, J., King, J. B. & Cichewicz, R. H. Fungal biofilm inhibitors from a human oral microbiome-derived bacterium. *Org. Biomol. Chem.* **10**, 2044 (2012).
 99. Nichols, D. *et al.* Use of ichip for high-throughput in situ cultivation of ‘uncultivable’ microbial species. *Appl. Environ. Microbiol.* **76**, 2445–2450 (2010).
 100. Ling, L. L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 455–459 (2015).
 101. Sabdono, A. & Radjasa, O. K. Microbial symbionts in marine sponges: Marine natural product factory. *J. Coast. Dev.* **11**, 57–61 (2008).
 102. Taylor, M. W., Radax, R., Steger, D. & Wagner, M. Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiol. Mol. Biol. Rev.* **71**, 295–347 (2007).
 103. Bergmann, W. & Feeney, R. J. The isolation of a new thymine pentoside from sponges. *Journal of the American Chemical Society* **72**, 2809–2810 (1950).
 104. Bergmann, W. & Feeney, R. J. Contributions to the study of marine products. XXXII. the nucleosides of sponges. I. *J. Org. Chem.* **16**, 981–987 (1951).
 105. Proksch, P., Edrada, R. A. & Ebel, R. Drugs from the seas - Current status and microbiological implications. *Applied Microbiology and Biotechnology* **59**, 125–134 (2002).
 106. Bowling, J. J., Kochanowska, A. J., Kasanah, N. & Hamann, M. T. Nature’s bounty – drug discovery from the sea. *Expert Opin. Drug Discov.* **2**, 1505–1522 (2007).
 107. Mayer, A. M., Nguyen, M., Newman, D. J. & Glaser, K. B. The Marine Pharmacology and Pharmaceuticals Pipeline in 2015. *FASEB J* **30**, 932.7- (2016).
 108. Marine Pharmacology: Clinical Development. (2017). at <<http://marinepharmacology.midwestern.edu/clinPipeline.htm>>
 109. Cortes, J. *et al.* Eribulin monotherapy versus treatment of physician’s choice in patients with metastatic breast cancer (EMBRACE): a phase 3 open-label randomised study. *Lancet* **377**, 914–923 (2011).
 110. Schmidt, E. W., Bewley, C. a & Faulkner, D. J. Theopalauamide, a Bicyclic Glycopeptide from Filamentous Bacterial Symbionts of the Lithistid Sponge *Theonella swinhoei* from Palau and Mozambique. *J. Org. Chem.* **63**, 1254–1258 (1998).
 111. Piel, J. Metabolites from symbiotic bacteria. *Nat. Prod. Rep.* **26**, 338–362 (2009).
 112. Piel, J. Bacterial symbionts: prospects for the sustainable production of invertebrate-derived pharmaceuticals. *Curr. Med. Chem.* **13**, 39–50 (2006).
 113. Brück, W. M., Sennett, S. H., Pomponi, S. a, Willenz, P. & McCarthy, P. J. Identification of the bacterial symbiont *Entotheonella* sp. in the mesohyl of the marine sponge *Discodermia* sp. *ISME J.* **2**, 335–9 (2008).
 114. Hentschel, U., Piel, J., Degan, S. M. & Taylor, M. W. Genomic insights into the marine sponge microbiome. *Nat. Rev. Microbiol.* **10**, 641–654 (2012).
 115. Lee, Y. K., Lee, J. & Lee, H. K. Microbial symbiosis in marine sponges. *J. Microbiol.* **39**, 254–264 (2001).
 116. Webster, N. S. & Hill, R. T. R. T. The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by an α -Proteobacterium.

- Mar. Biol.* **138**, 843–851 (2001).
117. Douglas, A. E. *Symbiotic Interactions*. University Press, Oxford (Oxford University Press, 1995). doi:10.1016/0169-5347(94)90326-3
 118. Wilkinson, C. R. Microbial associations in sponges. III. Ultrastructure of the in situ associations in coral reef sponges. *Mar. Biol.* **49**, 177–185 (1978).
 119. Fusetani, N. & Matsunaga, S. Bioactive sponge peptides. *Chem. Rev.* **93**, 1793–1806 (1993).
 120. Bewley, C. A. & Faulkner, D. J. Lithistid Sponges: Star Performers or Hosts to the Stars. *Angew. Chemie Int. Ed.* **37**, 2162–2178 (1998).
 121. Prakasa Rao, T. S. *et al.* New polyhydroxy sterols from the marine sponge *Callyspongia fibrosa* (Ridley & Dendly). *Tetrahedron Letters* **51**, (2010).
 122. Hentschel, U. *et al.* Molecular Evidence for a Uniform Microbial Community in Sponges from Different Oceans Molecular Evidence for a Uniform Microbial Community in Sponges from Different Oceans. *Appl. Environ. Microbiol.* **68**, 4431–4440 (2002).
 123. Fusetani, N., Matsunaga, S., Matsumoto, H. & Takebayashi, Y. Cyclotheonamides, Potent Thrombin Inhibitors, from a Marine Sponge *Theonella* sp. *J. Am. Chem. Soc.* **112**, 7053–7054 (1990).
 124. Nakao, Y., Matsunaga, S. & Fusetani, N. Three more cyclotheonamides, C, D, and E, potent thrombin inhibitors from the marine sponge *Theonella swinhoei*. *Bioorg. Med. Chem.* **3**, 1115–1122 (1995).
 125. Fusetani, N., Nakao, Y. & Matsunaga, S. Nazumamide A, a thrombin-inhibitory Tetrapeptide, from a marine sponge, *theonella* sp. *Tetrahedron Lett.* **32**, 7073–7074 (1991).
 126. Matsunaga, S., Fusetani, N. & Nakao, Y. Eight new cytotoxic metabolites closely related to onnamide A from two marine sponges of the genus *Theonella*. *Tetrahedron* **48**, 8369–8376 (1992).
 127. Matsunaga, S., Fusetani, N., Kato, Y. & Hirota, H. Aurantosides A and B: cytotoxic tetramic acid glycosides from the marine sponge *Theonella* sp. *J. Am. Chem. Soc.* **113**, 9690–9692 (1991).
 128. Kobayashi, J. *et al.* *Konbamide, a novel peptide with calmodulin antagonistic activity from the Okinawan marine sponge Theonella sp.* *J. Chem. Soc., Chem. Commun.* (Royal Society of Chemistry, 1991). doi:10.1039/C39910001050
 129. Freeman, M. F. *et al.* Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. *Science* **338**, 387–90 (2012).
 130. Hamada, T., Matsunaga, S. & Yano, GenFusetani, N. Polytheonamides A and B, highly cytotoxic, linear polypeptides with unprecedented structural features, from the marine sponge, *Theonella swinhoei*. *J. Am. Chem. Soc.* **127**, 110–118 (2005).
 131. Wilson, M. C. *et al.* An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* **506**, 58–62 (2014).
 132. Piel, J. *et al.* Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 16222–7 (2004).
 133. Amann, R. & Fuchs, B. M. Single-cell identification in microbial communities by improved fluorescence in situ hybridization techniques. *Nat. Rev. Microbiol.* **6**, 339–48 (2008).
 134. Siegl, A. *et al.* Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges. *ISME J.* **5**, 61–70 (2011).
 135. Stepanauskas, R. Single cell genomics: an individual look at microbes. *Curr. Opin. Microbiol.* **15**, 613–620 (2012).
 136. Bewley, C. a, Holland, N. D. & Faulkner, D. J. Two classes of metabolites from *Theonella swinhoei* are localized in distinct populations of bacterial symbionts. *Experientia* **52**, 716–22 (1996).
 137. Schmidt, E. W., Obratsova, A. Y., Davidson, S. K., Faulkner, D. J. & Haygood, M. G. Identification of the antifungal peptide-containing symbiont of the marine sponge *Theonella swinhoei* as a novel δ -proteobacterium, 'Candidatus Enttheonella

- palauensis'. *Mar. Biol.* **136**, 969–977 (2000).
138. Sakai, R., Higa, T. & Kashman, Y. Misakinolide-A, an antitumor macrolide from the marine sponge *Theonella* sp. *Chem. Lett.* **15**, 1499–1502 (1986).
139. Kato, Y. *et al.* Antitumor macrodiolides isolated from a marine sponge *theonella* sp.: Structure revision of misakinolide A. *Tetrahedron Lett.* **28**, 6225–6228 (1987).
140. Hochmuth, T. *et al.* Linking chemical and microbial diversity in marine sponges: Possible role for poribacteria as producers of methyl-branched fatty acids. *ChemBioChem* **11**, 2572–2578 (2010).
141. Beltramo, G. Forschungszentrum Jülich - Raman-Mikroskopie. at <http://www.fz-juelich.de/ics/ics-7/DE/Forschung/MethodenUndVerfahren/RamanMikroskopie/ramanMikroskopie_node.html>
142. Schuster, K. C., Urlaub, E. & Gapes, J. R. Single-cell analysis of bacteria by Raman microscopy: spectral information on the chemical composition of cells and on the heterogeneity in a culture. *J. Microbiol. Methods* **42**, 29–38 (2000).
143. Sekiguchi, Y. *et al.* *Anaerolinea thermophila* gen. nov., sp. nov. and *Caldilinea aerophila* gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain bacteria at the subphylum level. *Int. J. Syst. Evol. Microbiol.* **53**, 1843–1851 (2003).
144. Yamada, T. *et al.* *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the . *Int. J. Syst. Evol. Microbiol.* **56**, 1331–1340 (2006).
145. Hanada, S. in *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*. (eds. Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E. & Thompson, F.) 515–532 (Springer Berlin Heidelberg, 2014). doi:10.1007/978-3-642-38954-2_165
146. Rappé, M. S. *et al.* The Uncultured Microbial Majority. *Annu. Rev. Microbiol.* **57**, 369–394 (2003).
147. Dasch, G. A., Weiss, E. & Chang, K. P. in *Bergey's Manual of Systematic Bacteriology* **1**, 881–883 (1984).
148. Borroni, G., Brazzelli, V., Rosso, R. & Pavan, M. *Paederus fuscipes* dermatitis: A histopathological study. *Am. J. Dermatopathol.* **13**, 467–474 (1991).
149. Pavan, M. & Bo, G. Pederin, toxic principle obtained in the crystalline state from the beetle *Paederus fuscipes* Curt. *Physiol. Comp. Oecologia* **3**, 307–312 (1953).
150. Brega, A., Falaschi, A., Carli, L. De & Pavan, M. Studies on the mechanism of action of pederine. *J. Cell Biol.* **36**, 485–496 (1968).
151. Carrasco, L., Fernandez-Puentes, C. & Vazquez, D. Antibiotics and compounds affecting translation by eukaryotic ribosomes. Specific enhancement of aminoacyl-tRNA binding by methylxanthines. *Mol. Cell. Biochem.* **10**, 97–122 (1976).
152. Kellner, R. L. L. & Dettner, K. Allocation of pederin during lifetime of *Paederus* rove beetles (Coleoptera: Staphylinidae): Evidence for polymorphism of hemolymph toxin. *J. Chem. Ecol.* **21**, 1719–1733 (1995).
153. Kellner, R. L. L. & Dettner, K. Differential efficacy of toxic pederin in deterring potential arthropod predators of *Paederus* (Coleoptera: Staphylinidae) offspring. *Oecologia* **107**, 293–300 (1996).
154. Kellner, R. L. L. Suppression of pederin biosynthesis through antibiotic elimination of endosymbionts in *Paederus sabaeus*. *J. Insect Physiol.* **47**, 475–483 (2001).
155. Kellner, R. L. L. Molecular identification of an endosymbiotic bacterium associated with pederin biosynthesis in *Paederus sabaeus* (Coleoptera: Staphylinidae). *Insect Biochem. Mol. Biol.* **32**, 389–395 (2002).
156. Piel, J., Hofer, I. & Hui, D. Evidence for a Symbiosis Island Involved in Horizontal Acquisition of Pederin Biosynthetic Capabilities by the Bacterial Symbiont of *Paederus fuscipes* Beetles. *J. Bacteriol.* **186**, 1280–1286 (2004).
157. Lagkouvardos, I. Genomic Analysis of the Primary Metabolism of the uncultivated symbiont “*Pseudomonas paederii*”. (Rheinische Friedrich-Wilhelms-Universität Bonn, 2006).

158. McCutcheon, J. P. & Moran, N. A. Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10**, 13 (2011).
159. Sakharkar, K. R. & Chow, V. T. K. Strategies for genome reduction in microbial genomes. *Genome Inform.* **16**, 69–75 (2005).
160. Arai, H. Regulation and Function of Versatile Aerobic and Anaerobic Respiratory Metabolism in *Pseudomonas aeruginosa*. *Front. Microbiol.* **2**, 103 (2011).
161. Lyczak, J. B., Cannon, C. L. & Pier, G. B. Lung infections associated with cystic fibrosis. *Clin. Microbiol. Rev.* **15**, 194–222 (2002).
162. Bodey, G. P., Bolivar, R., Fainstein, V. & Jadeja, L. Infections caused by *Pseudomonas aeruginosa*. *Rev. Infect. Dis.* **5**, 279–313
163. Hancock, R. E. W. & Speert, D. P. Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and impact on treatment. *Drug Resist. Updat.* **3**, 247–255 (2000).
164. Haine, E. R. Symbiont-mediated protection. *Proc. R. Soc. B Biol. Sci.* **275**, 353 (2008).
165. Delsuc, F., Brinkmann, H. & Philippe, H. Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* **6**, 361–375 (2005).
166. Ferguson, R. L., Buckley, E. N. & Palumbo, A. V. Response of marine bacterioplankton to differential filtration and confinement. *Appl. Environ. Microbiol.* **47**, 49–55 (1984).
167. Kogure, K., Simidu, U. & Taga, N. A tentative direct microscopic method for counting living marine bacteria. *Can. J. Microbiol.* **25**, 415–20. (1979).
168. Kogure, K., Simidu, U. & Taga, N. Distribution of viable marine bacteria in neritic seawater around Japan. *Canadian journal of microbiology* **26**, 318–23 (1980).
169. Jones, J. G. The effect of environmental factors on estimated viable and total populations of planktonic bacteria in lakes and experimental enclosures. *Freshw. Biol.* **7**, 67–91 (1977).
170. Torsvik, V., Goksoyr, J. & Daae, F. L. High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* **56**, 782–787 (1990).
171. Society for General Microbiology. Symposium (32nd : 1981 : University of Edinburgh), M. J. (Michael J., Carlile, M. J. (Michael J., Collins, J. F. (John F. & Moseley, B. E. B. *Molecular and cellular aspects of microbial evolution : Thirty-second Symposium of the Society for General Microbiology held at the University of Edinburgh, September 1981.* (Cambridge University Press, 1981).
172. Lane, D. J. *et al.* Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 6955–9 (1985).
173. Stackebrandt, E. & Woese, C. R. The phylogeny of prokaryotes. *Microbiol. Sci.* **1**, 117–122 (1984).
174. Case, R. J. *et al.* Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.* **73**, 278–88 (2007).
175. Cannone, J. J. *et al.* The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3**, 2 (2002).
176. Zuckerkandl, E. & Pauling, L. Evolutionary divergence and convergence in proteins. *Evol. genes proteins* **97**, 97–166 (1965).
177. Zuckerkandl, E. & Pauling, L. Molecular Disease, Evolution, and Genic Heterogeneity. *Horizons in Biochemistry* 189–222 (1962). doi:10.1016/S0140-6736(54)91105-8
178. Zuckerkandl, E. On the molecular evolutionary clock. *J. Mol. Evol.* **26**, 34–46 (1987).
179. Ayala, F. J. Molecular clock mirages. *Bioessays* **21**, 71–75 (1999).
180. Weiß, M. & Göker, M. in *The Yeasts* 159–174 (Elsevier, 2011). doi:10.1016/B978-0-444-52149-1.00012-4
181. Morrison, D. A. & Ellis, J. T. Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of apicomplexa. *Mol. Biol. Evol.* **14**, 428–41 (1997).
182. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–82 (1997).
183. Thompson, J. D. *et al.* *Multiple Sequence Alignment Using ClustalW and ClustalX.*

- Current Protocols in Bioinformatics* **00**, 2.3.1-2.3.22 (John Wiley & Sons, Inc.).
184. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–7 (2004).
 185. Lecointre, G. & Guyader, H. *Biosystematik*. (Springer-Verlag Berlin Heidelberg, 2006).
 186. Campbell, N. A. & Reece, J. B. *Biologie: gymnasiale Oberstufe*. (Pearson, 2010).
 187. Mount, D. W. D. Choosing a method for phylogenetic prediction. *Cold Spring Harb. Protoc.* **3**, (2008).
 188. Sokal, R. R. & Michener, C. D. A statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.* **38**, 1409–1438 (1958).
 189. Kidd, K. K. & Sgaramella-Zonta, L. A. Phylogenetic analysis: concepts and methods. *Am. J. Hum. Genet.* **23**, 235–52 (1971).
 190. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
 191. Barton, N. H., Briggs, D. E. G., Eisen, J. A., Goldstein, D. B. & Patel, N. H. *Evolution*. (Cold Spring Harbor Laboratory Press, 2007).
 192. Camin, J. H. & Sokal, R. R. A Method for Deducing Branching Sequences in Phylogeny. *Evolution (N. Y.)* **19**, 311 (1965).
 193. Fitch, W. M. Toward Defining the Course of Evolution: Minimum Change for a Specific Tree Topology. *Syst. Biol.* **20**, 406–416 (1971).
 194. Ludwig, W. *et al.* Bacterial phylogeny based on comparative sequence analysis (review). *Electrophoresis* **19**, 554–568 (1998).
 195. Sforza, C. & Edwards, A. W. F. Analysis of human evolution. *Genet. Today* **3**, 923–933 (1964).
 196. Felsenstein, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376 (1981).
 197. Efron, B. *The Jackknife, the Bootstrap and Other Resampling Plans*. (Society for Industrial and Applied Mathematics, 1982). doi:10.1137/1.9781611970319
 198. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution (N. Y.)* **39**, 783–791 (1985).
 199. Uria, A. R. Investigating Natural Product Biosynthesis in Uncultivated Symbiotic Bacteria of the Marine Sponge Theonella Swinhoei. *Dissertation* (2012).
 200. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–80 (2005).
 201. Droege, M. & Hill, B. The Genome Sequencer FLX™ System-Longer reads, more applications, straight forward bioinformatics and more complete data sets. *Journal of Biotechnology* **136**, 3–10 (2008).
 202. Rothberg, J. M. & Leamon, J. H. The development and impact of 454 sequencing. *Nat. Biotechnol.* **26**, 1117–1124 (2008).
 203. DiGuistini, S. *et al.* De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biol.* **10**, R94 (2009).
 204. Uria, A. & Piel, J. Cultivation-independent approaches to investigate the chemistry of marine symbiotic bacteria. *Phytochem. Rev.* **8**, 401–414 (2009).
 205. Fusetani, N., Sugawara, T., Matsunaga, S. & Hirota, H. Orbiculamide A: a novel cytotoxic cyclic peptide from a marine sponge Theonella sp. *J. Am. Chem. Soc.* **113**, 7811–7812 (1991).
 206. Luo, P., Su, T., Hu, C. & Ren, C. A Novel and Simple PCR Walking Method for Rapid Acquisition of Long DNA Sequence Flanking a Known Site in Microbial Genome. *Mol. Biotechnol.* **47**, 220–228 (2011).
 207. Wilson, I. G. Inhibition and facilitation of nucleic acid amplification. *Appl. Environ. Microbiol.* **63**, 3741–51 (1997).
 208. Stackebrandt, E. *et al.* Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. *Int. J. Syst. Evol. Microbiol.* **44**, 846–849 (1994).
 209. Vickneswaran, N. *Studien zur Verbreitung von - einer neuen Gruppe von Naturstoffproduzenten*. (2013).
 210. Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. 16S ribosomal DNA

- amplification for phylogenetic study. *J. Bacteriol.* **173**, 697–703 (1991).
211. Tamura, K., Nei, M. & Kumar, S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci.* **101**, 11030–11035 (2004).
212. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
213. *Goggle Maps, Kartendaten* © 2019.
214. Lackner, G., Peters, E. E., Helfrich, E. J. N. & Piel, J. Insights into the lifestyle of uncultured bacterial natural product factories associated with marine sponges. *Proc. Natl. Acad. Sci.* **114**, E347–E356 (2017).
215. Peters, E. E. Genome-Mining of Uncultivated Bacteria as Source for New Bioactive Natural Products - ETH Bibliotheken. (2018).
216. Maldonado, L. A. *et al.* *Salinispora arenicola* gen. nov., sp. nov. and *Salinispora tropica* sp. nov., obligate marine actinomycetes belonging to the family Micromonosporaceae. *Int. J. Syst. Evol. Microbiol.* **55**, 1759–1766 (2005).
217. Udway, D. W. *et al.* Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10376–81 (2007).
218. Sacristán-Soriano, O., Banaigs, B., Casamayor, E. O. & Becerro, M. A. Exploring the links between natural products and bacterial assemblages in the sponge *Aplysina aerophoba*. *Appl. Environ. Microbiol.* **77**, 862–70 (2011).
219. Inbaneson, S. J. & Ravikumar, S. In vitro antiplasmodial activity of marine sponge *Stylissa carteri* associated bacteria against *Plasmodium falciparum*. *Asian Pacific J. Trop. Dis.* **2**, 370–374 (2012).
220. Amalfitano, S. & Fazi, S. Recovery and quantification of bacterial cells associated with streambed sediments. *J. Microbiol. Methods* **75**, 237–243 (2008).
221. Mülhardt, C. *Der Experimentator: Molekularbiologie, Genomics*. (Springer Spektrum, 2013).
222. Stallmach, A. & Vehreschild, M. J. G. T. *Mikrobiom: Wissensstand und Perspektiven*. (De Gruyter, 2016).
223. *Methods and overviews*. (Molecular Research LP, Shallowater, USA).
224. Klappenbach, J. A., Saxman, P. R., Cole, J. R. & Schmidt, T. M. rrndb: the Ribosomal RNA Operon Copy Number Database. *Nucleic Acids Res.* **29**, 181–4 (2001).
225. Větrovský, T. & Baldrian, P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One* **8**, e57923 (2013).
226. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1–e1 (2013).
227. Knöss, W. Terpenbiosynthese in Bakterien und Pflanzen - neue Perspektive. *Pharm. Unserer Zeit* **28**, 247–254 (2007).
228. Sivonen, K., Leikoski, N., Fewer, D. P. & Jokela, J. Cyanobactins-ribosomal cyclic peptides produced by cyanobacteria. *Appl. Microbiol. Biotechnol.* **86**, 1213–25 (2010).
229. McAuliffe, O., Ross, R. P. & Hill, C. Lantibiotics: structure, biosynthesis and mode of action. *FEMS Microbiol. Rev.* **25**, 285–308 (2001).
230. Rocha, E. P. C. & Danchin, A. Base composition bias might result from competition for metabolic resources. *Trends in Genetics* (2002). doi:10.1016/S0168-9525(02)02690-2
231. Crüsemann, M., Kohlhaas, C. & Piel, J. Evolution-guided engineering of nonribosomal peptide synthetase adenylation domains. *Chem. Sci.* **4**, 1041–1045 (2013).
232. Jez, J. M., Ferrer, J. L., Bowman, M. E., Dixon, R. A. & Noel, J. P. Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry* **39**, 890–902 (2000).
233. Bolanos-Garcia, V. M. & Davies, O. R. Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *Biochim. Biophys. Acta - Gen. Subj.* **1760**, 1304–1313 (2006).
234. Wülfing, C., Lombardero, J. & Plückthun, A. An *Escherichia coli* protein consisting of a domain homologous to FK506-binding proteins (FKBP) and a new metal binding motif. *J. Biol. Chem.* **269**, 2895–901 (1994).

235. Andersen, K. R., Leksa, N. C. & Schwartz, T. U. Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins* **81**, 1857–61 (2013).
236. *Protein Gel Electrophoresis Tips and Troubleshooting Guide. Electrophoresis* (2011).
237. Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. (2016). doi:10.1038/nbt.3597
238. Crüsemann, M. *et al.* Heterologous Expression, Biosynthetic Studies, and Ecological Function of the Selective Gq-Signaling Inhibitor FR900359. *Angew. Chemie Int. Ed.* **57**, 836–840 (2018).
239. Thomas-Poulsen, M. *et al.* New Genomic Insights into ‘Entotheonella’ Symbionts in *Theonella swinhoei*: Mixotrophy, Anaerobic Adaptation, Resilience, and Interaction. *Z Front. Microbiol* **7**, (2016).
240. Mori, T. *et al.* Single-bacterial genomics validates rich and varied specialized metabolism of uncultivated *Entotheonella* sponge symbionts. *Proc. Natl. Acad. Sci.* **115**, 1718–1723 (2018).
241. Mülhardt, C. *Der Experimentator Molekularbiologie/Genomics. Genomics* **7**, (2013).
242. Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293–300 (1951).
243. Sambrook, J., Fritsch, E. & Maniatis, T. *Molecular cloning: a laboratory manual/J. Sambrook, EF Fritsch, T. Maniatis.* (1989). at <<http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=CICY.xis&method=post&formato=2&cantidad=1&expresion=mfn=000178>>
244. Green, M. R. (Michael R., Sambrook, J. & Sambrook, J. *Molecular cloning: a laboratory manual.* (Cold Spring Harbor Laboratory Press, 2012).
245. Dhuffman. *Manual: pBluescript II Phagemid Vectors.* 1–27 (2008). at <<http://www.chem-agilent.com/pdf/strata/212205.pdf>>
246. <https://biotium.com/technology/gelred-gelgreen-nucleic-acid-gel-stains/>.
247. <http://www.nippongenetics.de/produkte/imaging-systeme/transilluminatoren/blaugruen-led-transilluminator/>.
248. Marchuk, D., Drumm, M., Saulino, A. & Collins, F. S. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19**, 1154 (1991).
249. pBluescript II Phagemid Vectors INSTRUCTION MANUAL LIMITED PRODUCT WARRANTY. 894–1304 (2122).
250. Promega. pGEM®- T and pGEM®- T Easy Vector Systems. *Tech. Man.* 1–28 (2010).
251. Capone, K. A., Dowd, S. E., Stamatias, G. N. & Nikolovski, J. Diversity of the human skin microbiome early in life. *J. Invest. Dermatol.* **131**, 2026–32 (2011).
252. Dowd, S. E. *et al.* Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiol.* **8**, 125 (2008).
253. Dowd, S. E., Sun, Y., Wolcott, R. D., Domingo, A. & Carroll, J. A. Bacterial Tag-Encoded FLX Amplicon Pyrosequencing (bTEFAP) for Microbiome Studies: Bacterial Diversity in the Ileum of Newly Weaned Salmonella-Infected Pigs. *Foodborne Pathog. Dis.* **5**, 459–472 (2008).
254. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
255. Eren, A. M. *et al.* Exploring the diversity of *Gardnerella vaginalis* in the genitourinary tract microbiota of monogamous couples through subtle nucleotide variation. *PLoS One* **6**, (2011).
256. Swanson, K. S. *et al.* Phylogenetic and gene-centric metagenomics of the canine intestinal microbiome reveals similarities with humans and mice. *ISME J.* **5**, 639–649 (2011).
257. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–72 (2006).
258. Ribosomal Database Project. RDP release 11 -sequence analysis tools. *Michigan State University* (2014). at

- <<https://rdp.cme.msu.edu/index.jsp%0Ahttp://rdp.cme.msu.edu/>>
259. National Center for Biotechnology Information. at <<https://www.ncbi.nlm.nih.gov/>>
260. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. & Stüber, D. Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. *Nat. Biotechnol.* **6**, 1321–1325 (1988).
261. Hengen, P. N. Purification of His-Tag fusion proteins from *Escherichia coli*. *Trends Biochem. Sci.* **20**, 285–286 (1995).
262. Luttmann, W., Bratke, K., Küpper, M. & Myrtek, D. *Der Experimentator: Immunologie*. (Springer Spektrum, 2014). doi:10.1007/978-3-642-41899-0
263. Eckelmann, W., Sponagel, H., Grottenthaler, W. & Hartmann, K.-J. u. a. Bodenkundliche Kartieranleitung. - 5. verbesserte und erweiterte - Auflage. (2005). at <<http://www.schweizerbart.de/publications/detail/artno/184050681>>
264. Stackebrandt, E., Liesack, W. & Goebel, B. M. Bacterial diversity in a soil sample from a subtropical Australian environment as determined by 16S rDNA analysis. *FASEB J.* **7**, 232–6 (1993).
265. Lindahl, V. & Bakken, L. R. Evaluation of methods for extraction of bacteria from soil. *FEMS Microbiol. Ecol.* **16**, 135–142 (1995).
266. <http://www.cosmobioussa.com/axis-shield-density-gradient-nycodenz.html>.
267. Russell, W. C., Newman, C. & Williamson, D. H. A simple cytochemical technique for demonstration of DNA in cells infected with mycoplasmas and viruses. *Nature* **253**, 461–462 (1975).
268. Wagner, M., Horn, M. & Daims, H. Fluorescence in situ hybridisation for the identification and characterisation of prokaryotes. *Curr. Opin. Microbiol.* **6**, 302–309 (2003).

10 Anhang

Tabelle 10-1: Beschreibung der 16S rRNA-Sequenzen der NCBI-Datenbankrecherche.

Klade1 (NJ)	Klade A (ML)	Klade B (ML)	title	Sample specification	localization
AF355042	AF355042		Phylogenetic composition of bacterioplankton assemblages from the Arctic Ocean	obtained from a depth of 55m from Arctic Ocean (SCICEX 95 cruise)	Arctic Ocean
AJ966600		AJ966600	Construction and screening of environmental metagenomic library derived from 5274m depth sediment of Pacific nodule province	deep sea sediment	China:nodule province of Pacific Ocean
AM997762		AM997762	Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean		Atlantic Ocean:South-Atlantic Ocean,Guinea Basin
AM997835		AM997835	Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean		Atlantic Ocean:South-Atlantic Ocean,Guinea Basin
AM997959		AM997959	Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean		Atlantic Ocean:South-Atlantic Ocean,Guinea Basin
AY907818	AY907818		Molecular identification of picoplankton populations in contrasting waters of the Arabian Sea	water	Arabian Sea
DQ396062	DQ396062		unclassified sequences; environmental samples.	deep-sea octacoral	Gulf of Alaska seamounts
DQ396198	DQ396198		unclassified sequences; environmental samples.	deep-sea octacoral	Gulf of Alaska seamounts
DQ396209	DQ396209		unclassified sequences; environmental samples.	deep-sea octacoral	Gulf of Alaska seamounts
DQ513028	DQ513028		unclassified sequences; environmental samples.	ridge flank crustal fluid	Baby Bare Seamount, a 3.5 Ma old outcrop located in the north-east Pacific Ocean on the eastern flank of the Juan de Fuca Ridge.
DQ925899	DQ925899		Microbes and Thermal Gradients in High Temperature Hydrothermal Chimneys	hydrothermal vent chimneys	Guaymas Basin
EF645941	EF645941		Potential interactions of particle-associated anammox bacteria with bacterial and archaeal partners in the Namibian upwelling system	Namibian upwelling system	Namibia
EF999360		EF999360	Vertical distribution and diversity of sulfate-reducing prokaryotes in the Pearl River estuarine sediments, Southern China	Pearl River Estuary sediments at 22cm depth	Guandong Province
EU035855	EU035855		Insight into the microbial community structure of a Norwegian deep-water coral reef environment	water from deep-sea coral reef	Norway
EU287401		EU287401	Bacterial diversity in surface sediments from the Pacific Arctic Ocean	arctic surface sediment	Pacific Arctic Ocean
EU799459	EU799459		It's all relative: ranking the diversity of aquatic bacterial communities	Newport Harbour, RI	USA
EU919757	EU919757		Community composition of the marine bacterioplankton in Kongsfjorden (Spitsbergen) as revealed by 16S rRNA gene analysis	ocean water	Arctic Ocean: Kongsfjorden, western Spitsbergen, Svalbard
FJ205196		FJ205196	Prokaryotic Diversity in the deep-sea hydrothermal region of the East Lau Spreading Centre	inactive hydrothermal field sediments, depth:2350m	Pacific Ocean: Lau Basin
FJ615154	FJ615154		Microbial community dynamics in a seasonally anoxic fjord: Saanich Inlet, British Columbia	Saanich Inlet, depth of 100m in anoxic marine fjord	British Columbia
FJ615158	FJ615158		Microbial community dynamics in a seasonally anoxic fjord: Saanich Inlet, British Columbia	Saanich Inlet, depth of 100m in anoxic marine fjord	British Columbia
FJ746296		FJ746296	Microbial diversity and stratification of South Pacific abyssal marine sediments	ocean sediment, 5306 m water depth, during Cruise Knox02rr	South Pacific
FN553499	FN553499		Bacterial sulfur cycling shapes microbial communities in surface sediments of an ultramafic hydrothermal vent field	Logatchev hydrothermal vent field,oceanic sediment, watercolumn depth = 3034 m, sediment depth = 0-1 cm	Atlantic Ocean:Mid-Atlantic Ridge, Logatchev hydrothermal vent field
FN553500		FN553500	Bacterial sulfur cycling shapes microbial communities in surface sediments of an ultramafic hydrothermal vent field	Logatchev hydrothermal vent field,oceanic sediment, watercolumn depth = 3034 m, sediment depth = 0-1 cm	Atlantic Ocean:Mid-Atlantic Ridge, Logatchev hydrothermal vent field
GQ337143	GQ337143		Hydrography shapes bacterial biogeography of the deep Arctic Ocean	water at depth of 1000m	Arctic Ocean: Canada Basin
GQ337170	GQ337170		Hydrography shapes bacterial biogeography of the deep Arctic Ocean	water at depth of 410m	Arctic Ocean: Canada Basin
GQ337190	GQ337190		Hydrography shapes bacterial biogeography of the deep Arctic Ocean	water at depth of 1000m	Arctic Ocean: Canada Basin
GQ337315	GQ337315		Hydrography shapes bacterial biogeography of the deep Arctic Ocean	water at depth of 252m	Arctic Ocean: Nansen Basin
GQ346830	GQ346830		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 100 m depth	48.5883 N 123.5037 W
GQ346861	GQ346861		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 100 m depth	48.5883 N 123.5037 W
GQ347370	GQ347370		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 215 m depth	48.5883 N 123.5037 W
GQ347549	GQ347549		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 215 m depth	48.5883 N 123.5037 W
GQ348258	GQ348258		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 120 m depth	48.5883 N 123.5037 W

Anhang

GQ348524	GQ348524		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 10 m depth	48.5883 N 123.5037 W
GQ349230	GQ349230		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 120 m depth	
GQ349232	GQ349232		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 120 m depth	48.5883 N 123.5037 W
GQ349311	GQ349311		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 120 m depth	48.5883 N 123.5037 W
GQ349320	GQ349320		Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones	Saanich Inlet, 120 m depth	48.5883 N 123.5037 W
GU474877	GU474877		Time-series analyses of Monterey Bay coastal microbial picoplankton using a 'genome proxy' microarray	Hawaii Oceanographic Time-series study site ALOHA	USA
HQ671985	HQ671985		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 10m depth	48°58.2'N, 130°40.0'W
HQ672600	HQ672600		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 500m depth	48°58.2'N, 130°40.0'W
HQ673380	HQ673380		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P4, 1300m depth	48°39.0'N, 126°4.0'W
HQ673574	HQ673574		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 1000m depth	48°58.2'N, 130°40.0'W
HQ673735	HQ673735		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 1000m depth	48°58.2'N, 130°40.0'W
HQ673859	HQ673859		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 2000m depth	48°58.2'N, 130°40.0'W
HQ673979	HQ673979		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 2000m depth	48°58.2'N, 130°40.0'W
HQ674032	HQ674032		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 1000m depth	48°58.2'N, 130°40.0'W
HQ674398	HQ674398		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 2000m depth	48°58.2'N, 130°40.0'W
HQ674529	HQ674529		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 2000m depth	48°58.2'N, 130°40.0'W
HQ674545	HQ674545		Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P12, 2000m depth	48°58.2'N, 130°40.0'W
HQ721425	HQ721425		Microbial diversity and stratification of oligotrophic abyssal marine sediments at the southern edge of the South Pacific Gyre	abyssal seawater	South Pacific Gyre
HQ721435	HQ721435		Microbial diversity and stratification of oligotrophic abyssal marine sediments at the southern edge of the South Pacific Gyre	abyssal seawater	South Pacific Gyre
HQ721436	HQ721436		Microbial diversity and stratification of oligotrophic abyssal marine sediments at the southern edge of the South Pacific Gyre	abyssal seawater	South Pacific Gyre
JN018780	JN018780		Natural gas and temperature structured a microbial community response to the Deepwater Horizon oil spill		Gulf of Mexico
JN166282	JN166282		Non-random assembly of bacterioplankton communities in the subtropical north pacific ocean The mean state and temporal variability in bacterioplankton communities spanning the surface to bathypelagic zones of the oligotrophic North Pacific subtropical gyre	depth=350m; Hawaii Ocean Time series cruise 157, station 2, cast 3, bottle 7	22.75 N 178.00 W
JN166351	JN166351		Non-random assembly of bacterioplankton communities in the subtropical north pacific ocean The mean state and temporal variability in bacterioplankton communities spanning the surface to bathypelagic zones of the oligotrophic North Pacific subtropical gyre	depth=350m; Hawaii Ocean Time series cruise 157, station 2, cast 3, bottle 8	22.75 N 178.00 W
JN166364	JN166364		Non-random assembly of bacterioplankton communities in the subtropical north pacific ocean The mean state and temporal variability in bacterioplankton communities spanning the surface to bathypelagic zones of the oligotrophic North Pacific subtropical gyre	depth=350m; Hawaii Ocean Time series cruise 157, station 2, cast 3, bottle 9	22.75 N 178.00 W
JN233007	JN233007		Composition and enzymatic function of particle-associated and free-living bacteria: a coastal/offshore comparison Extracellular enzymatic function and 16S rDNA and 16S rRNA clone library composition of marine particle-associated and free-living bacterial communities	ocean water collected from the continental slope off Cape Lookout, NC from 505m depth	USA; 34.18 N 75.82 W
JN591926	JN591926		Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota	surface seawater, Puget Sound	USA: WA
JN886886		JN886886	Microbial communities in the semi-consolidated carbonate sediment at the South West Indian Ridge	carbonate sediments	South West Indian Ridge
JN886931		JN886931	Microbial communities in the semi-consolidated carbonate sediment at the South West Indian Ridge	carbonate sediments	South West Indian Ridge
JN977189		JN977189	Diversity of bacterial community in Jiaozhou Bay sediments investigated by 16S rRNA gene libraries	sediment	Jiaozhao Bay
JX016767	JX016767		Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom	marine bulk water	Nordsee; 54.19167 N 7.9 E
JX016951	JX016951		Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom	marine bulk water	Nordsee; 54.19167 N 7.9 E
JX017186	JX017186		Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom	marine bulk water	Nordsee; 54.19167 N 7.9 E
JX017211	JX017211		Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom	marine bulk water	Nordsee; 54.19167 N 7.9 E
JX227438		JX227438	A comparison of microbial communities in deep-sea polymetallic nodules and the surrounding sediments in the Pacific Ocean	sediment collected from station WS0904 in the Clarion-Clipperton Fracture Zone	Pacific Ocean
JX227604		JX227604	A comparison of microbial communities in deep-sea polymetallic nodules and the surrounding sediments in the Pacific Ocean	sediment collected from station WS0902 in the Clarion-Clipperton Fracture Zone	Pacific Ocean
JX227720		JX227720	A comparison of microbial communities in deep-sea polymetallic nodules and the surrounding sediments in the Pacific Ocean	sediment collected from station WS0902 in the Clarion-Clipperton Fracture Zone	Pacific Ocean

Anhang

JX530785	JX530785		Response of bacterioplankton to iron fertilization of the Southern Ocean, Antarctica Effect of iron fertilization on bacterial communities in Southern Ocean	300m depth water samples filtered on 0.22 micron filter paper; derived from the Southern ocean iron fertilization experiment (LOHAFEX)	47.5015 S 15.4450 W
KJ365327	KJ365327		Winter bloom of a rare betaproteobacterium in the Arctic Ocean	sea water	Canada: Amundsen Gulf, 71.300 N 124.483 W
KJ365329	KJ365329		Winter bloom of a rare betaproteobacterium in the Arctic Ocean	sea water	Canada: Amundsen Gulf, 71.300 N 124.483 W
KJ365330	KJ365330		Winter bloom of a rare betaproteobacterium in the Arctic Ocean	sea water	Canada: Amundsen Gulf, 71.300 N 124.483 W
KJ365381	KJ365381		Winter bloom of a rare betaproteobacterium in the Arctic Ocean	sea water	Canada: Amundsen Gulf, 71.300 N 124.483 W
KJ590663		KJ590663	Effects of Chlorine Dosage on Marine Biofilm Formation	bio-ball filter	
KM018348	KM018348		Diversity of Methanogens and Sulfate-Reducing Prokaryotes in Five Interfaces of Deep-Sea Brines and Seawater in the Red Sea	Brine-seawater interface of Atlantis II Deep brine pool, 1998 m	Red Sea; 21.20 N 38.40 E
KM018350	KM018350		Diversity of Methanogens and Sulfate-Reducing Prokaryotes in Five Interfaces of Deep-Sea Brines and Seawater in the Red Sea	Brine-seawater interface of Atlantis II Deep brine pool, 1998 m	Red Sea; 21.20 N 38.40 E
KM018355	KM018355		Diversity of Methanogens and Sulfate-Reducing Prokaryotes in Five Interfaces of Deep-Sea Brines and Seawater in the Red Sea	Brine-seawater interface of Atlantis II Deep brine pool, 1998 m	Red Sea; 21.20 N 38.40 E
KM018380	KM018380		Diversity of Methanogens and Sulfate-Reducing Prokaryotes in Five Interfaces of Deep-Sea Brines and Seawater in the Red Sea	Brine-seawater interface of Atlantis II Deep brine pool, 1998 m	Red Sea; 21.20 N 38.40 E
KM018721	KM018721		Diversity of Methanogens and Sulfate-Reducing Prokaryotes in Five Interfaces of Deep-Sea Brines and Seawater in the Red Sea	Brine-seawater interface of Discovery Deep brine poo, 2081 ml	Red Sea; 21.16 N 38.30 E
KM018938	KM018938		Diversity of Methanogens and Sulfate-Reducing Prokaryotes in Five Interfaces of Deep-Sea Brines and Seawater in the Red Sea	Brine-seawater interface of Erba Deep brine pool, 2183 m	Red Sea; 20.44 N 38.10 E
KM019089	KM019089		Diversity of Methanogens and Sulfate-Reducing Prokaryotes in Five Interfaces of Deep-Sea Brines and Seawater in the Red Sea	Brine-seawater interface of Kebrit Deep brine pool, 1467 m	Red Sea; 24.43 N 36.16 E
KT336038		KT336038	Manganese-Cycling Microbial Communities Inside Deep-Sea Manganese Nodules	manganese nodule	
LC081045		LC081045	Dense microbial community on a ferromanganese nodule from the ultra-oligotrophic South Pacific Gyre: Implications for biogeochemical cycles	Inside part of ferromanganese nodule	South Pacific Gyre
Klade 2 (NJV)					
Klade C (ML)			title	Sample specification	localization
AB477994			Microbial methane production in deep aquifer associated with the accretionary prism in Japan	geothermal groundwater	Japan: Shizuoka, Shimada;
AF507887			Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California	Mono Lake at a depth of 35m from station 6 in July 2000	USA: California
AM777948			Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization	alkaline groundwater	Portugal; 39.14 N 7.58 W
AM777965			Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization	alkaline groundwater	Portugal; 39.14 N 7.58 W
AM777966			Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization	alkaline groundwater	Portugal; 39.14 N 7.58 W
AM777972			Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization	alkaline groundwater	Portugal; 39.14 N 7.58 W
DQ088768			Planktonic microbial communities associated with fracture-derived groundwater in a deep gold mine of South Africa	0.7 to 1.4 kmbls section of the continental crust	
EF559004			Microbial functional groups in a thermophilic anaerobic solid waste digester revealed by stable isotope probing	thermophilic anaerobic digester at 55 degrees celsius	
EF559052			Microbial functional groups in a thermophilic anaerobic solid waste digester revealed by stable isotope probing	thermophilic anaerobic digester at 55 degrees celsius	
EF559064			Microbial functional groups in a thermophilic anaerobic solid waste digester revealed by stable isotope probing	thermophilic anaerobic digester at 55 degrees celsius	
EF585980			Methanol-assimilating bacteria in anaerobic solid waste digester	solid waste digester fed with methanol	
FJ638546			Temperature dependent succession of sulfate reducing communities associated with a muddy hot spring in southwestern Taiwan	hot spring sediment	Taiwan: Tainan
GU455241			Understanding short-chain fatty acids accumulation enhanced in waste activated sludge alkaline fermentation: kinetics and microbiology	anaerobic fermentation reactor with waste activated sludge at thermophilic and pH 8.0	
GU455242			Understanding short-chain fatty acids accumulation enhanced in waste activated sludge alkaline fermentation: kinetics and microbiology	anaerobic fermentation reactor with waste activated sludge at thermophilic and pH 8.1	
GU455245			Understanding short-chain fatty acids accumulation enhanced in waste activated sludge alkaline fermentation: kinetics and microbiology	anaerobic fermentation reactor with waste activated sludge at thermophilic and pH 8.2	
GU455246			Understanding short-chain fatty acids accumulation enhanced in waste activated sludge alkaline fermentation: kinetics and microbiology	anaerobic fermentation reactor with waste activated sludge at thermophilic and pH 8.3	
GU455247			Understanding short-chain fatty acids accumulation enhanced in waste activated sludge alkaline fermentation: kinetics and microbiology	anaerobic fermentation reactor with waste activated sludge at thermophilic and pH 8.4	
JF417913			Characteristic microbial community of a dry thermophilic methanogenic digester: its long-term stability and change with feeding (Microbial community in a thermophilic dry anaerobic cellulose digester and the change of community with feeding of garbage stillage discharged from an ethanol production proces)	dry anaerobic digester	
JF935174			Microbial diversity in alkaline hot springs of Ambitle Island, Papua New Guinea	hot spring sediment	Papua New Guinea: Ambitle Island, Kapkai

Anhang

JQ087001			Microbial communities along biogeochemical gradients in a hydrocarbon-contaminated aquifer	push core sediment sample from the vadose zone of a hydrocarbon contaminated aquifer	Germany: Leuna
JX240597			The abundance of functional genes, cbbL, nifH, amoA and apsA, and bacterial community structure of intertidal soil from Arabian Sea	coastal soil	India; 22.1187 N 72.2088 E
KF912980			Microbial Diversity of Chromium-Contaminated Soils and Characterization of Six Chromium-Removing Bacteria	soil contaminated with chromium	
KP150691			Molecular biological method evaluation of thermophilic fermentation samples	thermophilic chicken dung - cow slurry fermentation	
KP150706			Molecular biological method evaluation of thermophilic fermentation samples	thermophilic chicken dung - cow slurry fermentation	
KP150709			Molecular biological method evaluation of thermophilic fermentation samples	thermophilic chicken dung - cow slurry fermentation	
Klade 3 (Klade D)			Titel	Probenspezifikation	Lokalisation
AB696548			Phylogenetic diversity of 16S rRNA from uncultured bacteria	soil	China: Taklamakan desert
AM176872			Recovery of novel bacterial diversity from mangrove sediment	mangrove sediment	China: Shenzhen Futian
AM902640			Diversity of Acidobacteria and Planctomycetes in a subsurface radioactive thermal mineral spring in the Central Alps, Austria	subsurface thermal spring	Austria: Bad Gastein
DQ404605			Environmental whole-genome amplification to access microbial populations in contaminated sediments	contaminated sediment	
DQ404735			Environmental whole-genome amplification to access microbial populations in contaminated sediments	contaminated sediment	
DQ404768			Environmental whole-genome amplification to access microbial populations in contaminated sediments	contaminated sediment	
DQ404785			Environmental whole-genome amplification to access microbial populations in contaminated sediments	contaminated sediment	
DQ404794			Environmental whole-genome amplification to access microbial populations in contaminated sediments	contaminated sediment	
EU246285			Miniprimer PCR, a new lens for viewing the microbial world	hypersaline microbial mat	Puerto Rico: Cabo Rojo, Candelaria lagoon
EU652503			Seasonal variation of microbial diversity in the Yellow Sea sediment	sediment	Yellow Sea
EU881356			The bacterial community of cropland (maize-sweet potato cropping field) in Karst region	maize-sweet potato cropland soil	China: Guangxi Autonomous Region
FJ269262			Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	Xestospongia testudinaria (sponge)	Indonesia: Manado
FJ269291			Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	Xestospongia testudinaria (sponge)	Indonesia: Manado
FJ269336			Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	Xestospongia muta (sponge)	USA: Key Largo, Florida
FJ269336			Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	Xestospongia muta (sponge)	USA: Key Largo, Florida
FJ444679			Analysis of bacterial communities in rhizosphere soil of healthy and diseased cotton (<i>Gossypium</i> sp.) at different plant growth stages	cotton rhizosphere	China
FJ748784			Vertical distribution and diversity of sulfate-reducing prokaryotes in the Pearl River estuarine sediments, Southern China	Pearl River Estuary sediments	China: Guangdong Province
FJ936824			16S rDNA analysis of environmental samples from volcano mud taken at Avachinsky (Kamchatka)	volcano mud	Avachinsky, islands Kamchatka
GQ472366			Bacterial communities in sediments of the shallow Lake Dongping in China	lake sediment	Lake Dongping, China
GU208397			Bacterial communities in sediments of the shallow Lake Dongping in China	lake sediment	Lake Dongping, China
HF912745			Investigation of sponge-associated bacteria from marine cold-water sponges	<i>Geodia barretti</i> (marine sponge)	
HF922363			Growth of anaerobic methane-oxidizing archaea and sulfate-reducing bacteria in a high-pressure membrane capsule bioreactor	high pressure reactor	Germany: Eckernförde Bay, Baltic Sea
HM187168			Vertical stratification of subsurface microbial community composition across geological formations at the Hanford Site	saturated zone of the Hanford Site 300 Area subsurface	USA
HM187297			Vertical stratification of subsurface microbial community composition across geological formations at the Hanford Site	saturated zone of the Hanford Site 300 Area subsurface	USA
HM243779			Vertical distribution of prokaryotes and responses to their environment in Honghu Lake sediments	upper sediment from Honghu Lake	China
HM243870			Vertical distribution of prokaryotes and responses to their environment in Honghu Lake sediments	middle sediment from Honghu Lake	China
HM243983			Comparison of Bacterial Diversity in Azorean and Hawaiian Lava Cave Microbial Mats (Investigation of Novel Microbial Diversity in Azorean and Hawaiian Lava Tubes)	middle sediment from Honghu Lake	China
HM445421			Comparison of Bacterial Diversity in Azorean and Hawaiian Lava Cave Microbial Mats (Investigation of Novel Microbial Diversity in Azorean and Hawaiian Lava Tubes)	yellow microbial mat from lava tube walls	Azores, Terceira, Gruta dos Principiantes Lava Tube
HM445462			Comparison of Bacterial Diversity in Azorean and Hawaiian Lava Cave Microbial Mats (Investigation of Novel Microbial Diversity in Azorean and Hawaiian Lava Tubes)	yellow microbial mat from lava tube walls	Azores, Terceira, Gruta dos Principiantes Lava Tube
HQ183989			Prokaryotic diversity, composition structure, and phylogenetic analysis of microbial communities in leachate sediment ecosystems	leachate sediment	Dongyang, China
HQ190396			Bacterial diversity of Zhongyuan oil field, China	Zhongyuan oil field	China

Anhang

HQ330534			Role of sediment bacteria and archaea in biogeochemical cycling in Lake Wivenhoe, Australia	sediment	Lake Wivenhoe, Australia
HQ445759			Bacterial Associations with Weathered Minerals in Deep Tropical Saprolite Formed from Quartz Diorite	deep saprolite and saprock from LG-1 sampling location	Puerto Rico: eastern, Luquillo Mountains
JF833887			Bacteria; environmental samples.	potassium mine soil	China: Yuexi, Anhui
JN038798			Alnus trabeculosa alters rhizosphere bacterial community of Phragmites australis in Chongxi tidal wetland, the Yangtze Estuary	Chongxi wetland soil	Yangtze Estuary, China
JN039006			Alnus trabeculosa alters rhizosphere bacterial community of Phragmites australis in Chongxi tidal wetland, the Yangtze Estuary	Chongxi wetland soil	Yangtze Estuary, China
JN496642			Phylogenetic stratigraphy in the Guerrero Negro hypersaline microbial mat	microbial mat, hypersaline evaporation pond	Mexico: Guerrero Negro
JN886848			Microbial communities in the semi-consolidated carbonate sediment at the South West Indian Ridge	arborate sediments at the South West Indian Ridge	South West Indian Ridge
JN977251			Diversity of bacterial community in Jiaozhou Bay sediments investigated by 16S rRNA gene libraries	Jiaozhou Bay sediment	Jiaozhou Bay, China
JN977257			Diversity of bacterial community in Jiaozhou Bay sediments investigated by 16S rRNA gene libraries	Jiaozhou Bay sediment	Jiaozhou Bay, China
JN977318			Diversity of bacterial community in Jiaozhou Bay sediments investigated by 16S rRNA gene libraries	Jiaozhou Bay sediment	Jiaozhou Bay, China
JQ612212			Metatranscriptomics of the marine sponge Geodia barretti: tackling phylogeny and function of its microbial community	Geodia barretti (marine sponge)	Norway (59.58.7N 05.22.89E)
JQ612380			Metatranscriptomics of the marine sponge Geodia barretti: tackling phylogeny and function of its microbial community	Geodia barretti (marine sponge)	Norway (59.58.7N 05.22.89E)
JQ668585			Diversity and distribution of bacteria community in different typeoil reservoirs in China	oil reservoir	China
JQ861367			PGPR from peanut	peanut rhizosphere soil	
JX120432			Biostimulation induces syntrophic interactions that impact C, S and N cycling in a sediment microbial community	subsurface aquifer sediment	USA: Colorado, Rifle
JX120436			Biostimulation induces syntrophic interactions that impact C, S and N cycling in a sediment microbial community	subsurface aquifer sediment	USA: Colorado, Rifle
JX120437			Biostimulation induces syntrophic interactions that impact C, S and N cycling in a sediment microbial community	subsurface aquifer sediment	USA: Colorado, Rifle
JX120438			Biostimulation induces syntrophic interactions that impact C, S and N cycling in a sediment microbial community	subsurface aquifer sediment	USA: Colorado, Rifle
KF906488			Seasonal dynamics of bacterial community composition in PAH-contaminated riverbank sediments	PAH-contaminated sediments	USA: Mahoning River Lowellville, OH
KF906489			Seasonal dynamics of bacterial community composition in PAH-contaminated riverbank sediments	PAH-contaminated sediments	USA: Mahoning River Lowellville, OH
KJ081618			Initial copper stress strengthens the resistance of soil microorganisms to a subsequent copper stress	copper contaminated soil	China
KM205507			Impact of beta-Carotene Transgenic Rice with Four Synthetic Genes on Rhizosphere Enzyme Activities and Bacterial Communities at Different Growth Stages	rhizosphere of a-carotene transgenic rice AH33 and its conventionally bred comparator Zhonghua 11 at different growth stages	
KP174433			Biodiversity of moderately halophilic bacteria in Yuncheng Salt Lake of Shanxi Province, China	Yuncheng Salt Lake	China
Klade 4 (NJ)	Klade E (ML)	Klade F (ML)	title	Sample specification	localization
22-clone2		22-clone2	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
AB533815	AB533815		Diversity of Bacteria and Archaea in hypersaline sediment from Death Valley National Park, California	salt pan	USA:California, Death Valley
AB533820	AB533820		Diversity of Bacteria and Archaea in hypersaline sediment from Death Valley National Park, California	salt pan	USA:California, Death Valley
AB533868	AB533868		Diversity of Bacteria and Archaea in hypersaline sediment from Death Valley National Park, California	salt pan	USA:California, Death Valley
AB533915	AB533915		Diversity of Bacteria and Archaea in hypersaline sediment from Death Valley National Park, California	salt pan	USA:California, Death Valley
AJ441215	AJ441215		Phylogenetic characterization of the bacterial assemblage associated with mucous secretions of the hydrothermal vent polychaete Paralvinella palmiformis	Paralvinella palmiformis mucus secretions	USA: Axial Volcano, Juan de Fuca
AM404381	AM404381		Hydrocarbon-fueled sulfate reduction can sustain long-lived tubeworm aggregations at cold seeps in the Gulf of Mexico	marine sediment	USA:Gulf of Mexico, Green Canyon
AM745216	AM745216		Impact of natural oil and higher hydrocarbons on microbial diversity, distribution, and activity in Gulf of Mexico cold-seep sediments	marine sediment	USA:Gulf of Mexico, Green Canyon, site GC232, station 4463
AY280424	AY280424		Microbial Diversity Associated with a Paralvinella sulficola Tube and the Adjacent Substratum on an Active Deep-Sea Vent Chimney	exterior of a deep-sea hydrothermal vent sulfide chimney grey whale bone, Pacific Ocean, SantaCruz Basin (N33.30 W119.22) depth 1674 meters	Juan de Fuca Ridge, Pacific Ocean
AY922193	AY922193		Comparative metagenomics of microbial communities		Pacific Ocean
DQ206407	DQ206407		Sulfide oxidation coupled to arsenate reduction by a diverse microbial community in a soda lake	soda lake water	USA: California, Mono Lake
DQ395946		DQ395946	unclassified sequences; environmental samples.	deep-sea octacoral	
DQ395955		DQ395955	unclassified sequences; environmental samples.	deep-sea octacoral	

Anhang

DQ661631		DQ661631	Community Analysis of Moderately Thermophilic Microflora Enriched by Different Energy Sources	bioreactor	China
DQ661635		DQ661635	Community Analysis of Moderately Thermophilic Microflora Enriched by Different Energy Sources	bioreactor	China
DQ677845	DQ677845		Bacterial diversity associated with Blood Falls, a subglacial outflow from the Taylor Glacier, Antarctica	subglacial outflow	Antarctica: McMurdo Dry Valleys, Taylor Glacier, Blood Falls
DQ833497	DQ833497		Bacterial diversity in sediments of the eutrophic Guanting Reservoir, China, estimated by analyses of 16S rDNA sequence	sediment of Guanting Reservoir	
DQ906055		DQ906055	Novel nickel resistance genes from the rhizosphere metagenome of plants adapted to acid mine drainage	rhizosphere	Tinto River, Huelva, Spain
EF219006	EF219006		Bacterial and Eukaryotic Diversity at a Ridge Flank Abyssal Hill Hydrothermal Site on the East Pacific Rise	Ridge Flank Abyssal Hill hydrothermal site on the east Pacific rise	
EF494371		EF494371	Interplay between landscape position and microbial abundance impacts geochemical characteristics of soil and saprolite	Nunnock River granitic landscape	Australia
EF655663	EF655663		Complete inhibition of sulfate-reducing bacteria with incomplete inhibition of mercury methylation by molybdate in salt marsh sediments	Walker Marsh	USA: CA, Tomales Bay
EF655668	EF655668		Complete inhibition of sulfate-reducing bacteria with incomplete inhibition of mercury methylation by molybdate in salt marsh sediments	Walker Marsh	USA: CA, Tomales Bay
EF999371	EF999371		Vertical distribution and diversity of sulfate-reducing prokaryotes in the Pearl River estuarine sediments, Southern China	Pearl River Estuary sediments at 22cm depth	Guandong Province
EU050842	EU050842		Bacterial, archaeal and eukaryotic diversity in Arctic sediment as revealed by 16S rRNA and 18S rRNA gene clone libraries analysis	sediment	Kings Bay, Svalbard, Arctic
EU386013	EU386013		Biogeographical distribution and diversity of bacterial communities in surface sediments of the South china sea	subseafloor sediment	South China Sea
EU419186		EU419186	Microbial diversity and ecology of selected andesitic hydrothermal environments	acidic, hydrothermally modified volcanic soil	Yellowstone National Park
EU419187		EU419187	Microbial diversity and ecology of selected andesitic hydrothermal environments	acidic, hydrothermally modified volcanic soil	Yellowstone National Park
EU488433	EU488433		Characterization of the lucinid bivalve-bacteria symbiotic system: the significance of the geochemical habitat on bacterial symbiont diversity and phylogeny	siliciclastic sediment from Thalassia sea grass bed	USA: lucinid bivalve collection site, Dog Island, Cedar Key, Florida
EU531787	EU531787		Diversity of unculturable bacteria isolated from mud volcanos in Xinjiang	mud volcano (salty environment)	China: Xinjiang Province, northwest
EU531801	EU531801		Diversity of unculturable bacteria isolated from mud volcanos in Xinjiang	mud volcano (salty environment)	China: Xinjiang Province, northwest
EU531802	EU531802		Diversity of unculturable bacteria isolated from mud volcanos in Xinjiang	mud volcano (salty environment)	China: Xinjiang Province, northwest
EU665084		EU665084	Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities	agricultural soil	Mexico, near lake Texcoco
EU925881		EU925881	Phylogenetic diversity of sediment bacteria in the northern Bering Sea	sediment from station DBS1	northern Bering Sea
FJ478735		FJ478735	Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ478839		FJ478839	Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ478903		FJ478903	Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ497423	FJ497423		Microbial Ecology of Fe-rich Mats and Basaltic Rock from Vailulu'u Seamount, American Samoa	Vailulu'u Seamount	American Samoa: South Pacific
FJ497440	FJ497440		Microbial Ecology of Fe-rich Mats and Basaltic Rock from Vailulu'u Seamount, American Samoa	Vailulu'u Seamount	American Samoa: South Pacific
FJ497460	FJ497460		Microbial Ecology of Fe-rich Mats and Basaltic Rock from Vailulu'u Seamount, American Samoa	Vailulu'u Seamount	American Samoa: South Pacific
FJ497501	FJ497501		Microbial Ecology of Fe-rich Mats and Basaltic Rock from Vailulu'u Seamount, American Samoa	Vailulu'u Seamount	American Samoa: South Pacific
FJ810574	FJ810574		Subsurface ecosystem resilience: long-term attenuation of subsurface contaminants supports a dynamic microbial community	coal tar waste-contaminated groundwater	rural area in South Glens Falls, New York,
FM165272	FM165272		Bacterial diversity associated with the tube of a cold seep vestimentiferan	chitinous tube of the vestimentiferan tubeworm Lamelibrachia sp. from cold seeps	S5.798, E9.711, 3152 depth
FM179908	FM179908		Biogeochemical processes and microbial diversity of the Gullfaks and Tommeliten methane seeps (Northern North Sea)	marine sediments	Norway:North Sea, Gullfaks oil and gas fields
FN553939	FN553939		Bacterial sulfur cycling shapes microbial communities in surface sediments of an ultramafic hydrothermal vent field	Logatchev hydrothermal vent field, site F, watercolumn depth = 3000 m	Atlantic Ocean:Mid-Atlantic Ridge, Logatchev hydrothermal vent field
FN553965	FN553965		Bacterial sulfur cycling shapes microbial communities in surface sediments of an ultramafic hydrothermal vent field	Logatchev hydrothermal vent field, site F, watercolumn depth = 3000 m	Atlantic Ocean:Mid-Atlantic Ridge, Logatchev hydrothermal vent field
GQ246304	GQ246304		Temporal and spatial distribution characterization of bacterial community in North Yellow Sea sediments, China	North Yellow Sea sediments	
GQ396879		GQ396879	Functional shifts in unvegetated, perhumid, recently-deglaciated soils do not correlate with shifts in soil bacterial community composition	soil	Mendenhall Glacie
GQ396921		GQ396921	Functional shifts in unvegetated, perhumid, recently-deglaciated soils do not correlate with shifts in soil bacterial community composition	soil	Mendenhall Glacie
GQ396959		GQ396959	Functional shifts in unvegetated, perhumid, recently-deglaciated soils do not correlate with shifts in soil bacterial community composition	soil	Mendenhall Glacie
GQ397046		GQ397046	Functional shifts in unvegetated, perhumid, recently-deglaciated soils do not correlate with shifts in soil bacterial community composition	soil	Mendenhall Glacie

Anhang

GU291345	GU291345		Anaerobic sulfide-oxidizing bacteria couple acidification and nitrogen cycling in heavy metal contaminated wetlands	acid mine drainage contaminated salt marsh sediments	USA: San Francisco Bay
HE798168		HE798168	Bacterial diversity in Maize Rhizosphere	Zea mays rhizosphere	
HE804485		HE804485	Coupled geochemical and microbiological assessment on bacteria-mediated bioreduction of As(V)-doped lepidocrocite in wetland soil	arsenic-doped Fe-oxyhydroxide (lepidocrocite) in wetland soil, incubated in anaerobic columns	France:Kervidy-Naizin catchment, Brittany
HE804507		HE804507	Coupled geochemical and microbiological assessment on bacteria-mediated bioreduction of As(V)-doped lepidocrocite in wetland soil	arsenic-doped Fe-oxyhydroxide (lepidocrocite) in wetland soil, incubated in anaerobic columns	France:Kervidy-Naizin catchment, Brittany
HE804560		HE804560	Coupled geochemical and microbiological assessment on bacteria-mediated bioreduction of As(V)-doped lepidocrocite in wetland soil	arsenic-doped Fe-oxyhydroxide (lepidocrocite) in wetland soil, incubated in anaerobic columns	France:Kervidy-Naizin catchment, Brittany
HE804605		HE804605	Coupled geochemical and microbiological assessment on bacteria-mediated bioreduction of As(V)-doped lepidocrocite in wetland soil	arsenic-doped Fe-oxyhydroxide (lepidocrocite) in wetland soil, incubated in anaerobic columns	France:Kervidy-Naizin catchment, Brittany
HM057166		HM057166	Autotrophic, sulfur-oxidizing actinobacteria in acidic environments	geothermal soil	Greece: Milos
HM057789	HM057789		Seasonal variation of microbial diversity in Yellow Sea water	ocean water from	Yellow Sea
HM745459		HM745459	Geomicrobiology of La Zarza-Perrunal Acid mine effluent (Iberian Pyritic Belt, Spain)	acid mine effluent	Spain: Iberian Pyritic Belt, La Zarza-Perrunal
HQ121174	HQ121174		Optimizing the indirect extraction of prokaryotic DNA from soils	loamy sand collected from a field planted with tomatoes at the Be-Wise Ranch, Escondido, CA	USA
HQ153833	HQ153833		Composition and diversity of microbial mats at shallow hydrothermal vents on Volcano 1, South Tonga Arc	white microbial film on pebble substrate in the aphotic zone of a shallow hydrothermal vent, Volcano 1, 189.1 m	Tonga trench
HQ397064	HQ397064		Bacterial diversity in haloalkaline soil	haloalkaline soil	21.99932 N 72.07703 E
HQ397074	HQ397074		Bacterial diversity in haloalkaline soil	haloalkaline soil	21.99932 N 72.07703 E
HQ397132		HQ397132	Bacterial diversity in normal soil	soil	21.91613 N 71.30910 E
HQ397393	HQ397393		Comparative molecular analysis of chemolithoautotrophic bacterial diversity and community structure from coastal saline soils, Gujarat, India	saline soil from Gujarat coast	India
HQ397530		HQ397530	Comparative molecular analysis of chemolithoautotrophic bacterial diversity and community structure from coastal saline soils, Gujarat, India	agricultural soil	India
HQ672465		HQ672465	Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P4, 500m depth	48°39.0'N, 126°4.0'W
HQ674102		HQ674102	Diversity and population structure of Marine Group A bacteria in the Northeast subarctic Pacific Ocean	Northeast subarctic Pacific Ocean, Station P4, 500m depth	48°39.0'N, 126°4.0'W
HQ697754		HQ697754	Bioremediation of Petroleum-Contaminated Saline-Alkali Soils Using Maize Straw and the Effect on Microbial Community Structure	hydrocarbon contaminated saline-alkali soil	
JF344208	JF344208		Characterization of the anaerobic microbial community in oil-polluted subtidal sediments: aromatic biodegradation potential after the Prestige oil spill	petroleum spot over marine sediments from figueiras beach	Spain: Cies Islands-Galicia
JF344521	JF344521		Characterization of the anaerobic microbial community in oil-polluted subtidal sediments: aromatic biodegradation potential after the Prestige oil spill	hydrocarbon polluted marine sediments from figueiras beach	Spain: Cies Islands-Galicia
JN023782		JN023782	Bacterial communities in soil under lichen and moss crusts	temperate highland grassland	Pico Del Águila
JN023803		JN023803	Bacterial communities in soil under lichen and moss crusts	temperate highland grassland	Pico Del Águila
JN023860		JN023860	Bacterial communities in soil under lichen and moss crusts	temperate highland grassland	Pico Del Águila
JN038638	JN038638		Alnus trabeculosa alters rhizosphere bacterial community of Phragmites australis in Chongxi tidal wetland, the Yangtze Estuary	Chongxi wetland soil	Yangtze Estuary, China
JN038640	JN038640		Alnus trabeculosa alters rhizosphere bacterial community of Phragmites australis in Chongxi tidal wetland, the Yangtze Estuary	Chongxi wetland soil	Yangtze Estuary, China
JN038669	JN038669		Alnus trabeculosa alters rhizosphere bacterial community of Phragmites australis in Chongxi tidal wetland, the Yangtze Estuary	Chongxi wetland soil	Yangtze Estuary, China
JN038676	JN038676		Alnus trabeculosa alters rhizosphere bacterial community of Phragmites australis in Chongxi tidal wetland, the Yangtze Estuary	Chongxi wetland soil	Yangtze Estuary, China
JN039012	JN039012		Alnus trabeculosa alters rhizosphere bacterial community of Phragmites australis in Chongxi tidal wetland, the Yangtze Estuary	Chongxi wetland soil	Yangtze Estuary, China
JN417536		JN417536	Effect of Paclotbutrazol on soil bacterial communities	soil	23° 59'N, 120° 33'E
JN540102	JN540102		Relationships between Soil Organic Matter, Nutrients, Bacterial Community Structure, And the Performance of Microbial Fuel Cells	soil from an agricultural site	USA: MA, Petersham, Harvard Forest
JN540103	JN540103		Relationships between Soil Organic Matter, Nutrients, Bacterial Community Structure, And the Performance of Microbial Fuel Cells	soil from an agricultural site	USA: MA, Petersham, Harvard Forest
JN540170	JN540170		Relationships between Soil Organic Matter, Nutrients, Bacterial Community Structure, And the Performance of Microbial Fuel Cells	soil from an agricultural site	USA: MA, Petersham, Harvard Forest
JN977158	JN977158		Diversity of bacterial community in Jiaozhou Bay sediments investigated by 16S rRNA gene libraries	sediment	Jiaozhao Bay, China
JN977296	JN977296		Diversity of bacterial community in Jiaozhou Bay sediments investigated by 16S rRNA gene libraries	sediment	Jiaozhao Bay, China
JN977302	JN977302		Diversity of bacterial community in Jiaozhou Bay sediments investigated by 16S rRNA gene libraries	sediment	Jiaozhao Bay, China
JQ287074	JQ287074		Life and death of deep-sea vents: bacterial diversity and ecosystem succession on inactive hydrothermal sulfides		East Pacific Rise
JQ579768	JQ579768		Characterization of the anaerobic microbial community in oil-polluted subtidal sediments: aromatic biodegradation potential after the Prestige oil spill	sediments from Figueiras Beach	Spain: Cies islands-Galicia
JQ815649		JQ815649	Microbial diversity of anaerobic zones of Tinto River combining culture dependent and non-dependent methods	Tinto River sediment (extreme acid environment)	Spain

Anhang

JQ815838		JQ815838	Microbial diversity of anaerobic zones of Tinto River combining culture dependent and non-dependent methods	Tinto River sediment (extreme acid environment)	Spain
JQ825201		JQ825201	Qinghai-Tibet Plateau, alpine-cold swamp meadow, alpine sandy grassland, freeze-thaw cycle, clone library	Qinghai-Tibet Plateau	
JQ941769	JQ941769		Typical microbial community compositions of a freshwater lake in China	diseased leaf	China: Taihu Lake
JX224305	JX224305		Subsurface microbial community response to acetate amendment	subsurface aquifer sediment	USA: Rifle, Colorado
JX224347	JX224347		Subsurface microbial community response to acetate amendment	subsurface aquifer sediment	USA: Rifle, Colorado
JX224367	JX224367		Subsurface microbial community response to acetate amendment	subsurface aquifer sediment	USA: Rifle, Colorado
JX406217		JX406217	Temporal variation of planktonic bacteria community composition in Lake Taihu (N2)	lake water	Lake Taihu, China
JX434196	JX434196		Fracture water associated microbial community from Sanford Lab/Homestake Mine, Black Hills, USA	gold mine borehole	Black Hills, USA
JX434240	JX434240		Fracture water associated microbial community from Sanford Lab/Homestake Mine, Black Hills, USA	gold mine borehole	Black Hills, USA
JX434244	JX434244		Fracture water associated microbial community from Sanford Lab/Homestake Mine, Black Hills, USA	gold mine borehole	Black Hills, USA
KC009896	KC009896		Stable Isotope Probing of Carbon Dioxide Fixing Bacterial Communities off the French Guiana coast	shallow fluidized muds	French Guiana coast
KC253355		KC253355	Effects of a cyanobacterial bloom on bacterial community composition of freshwater of pond	freshwater ponds	
KC470986	KC470986		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC470998	KC470998		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC471002	KC471002		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC471034	KC471034		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC471035	KC471035		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC471072	KC471072		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC471095	KC471095		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC471104	KC471104		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC471171	KC471171		A comparative study of microbial diversity and community structure in marine sediments using poly(A) tailing and reverse transcription-PCR	marine sediment	Yonaguni Knoll, Okinawa Trough, China
KC554081		KC554081	Bacteria; environmental samples.	soil of Yanshan Mountain	China: Hebei Province
KC554246		KC554246	Bacteria; environmental samples.	soil of Yanshan Mountain	China: Hebei Province
KC554504		KC554504	Bacteria; environmental samples.	soil of Yanshan Mountain	China: Hebei Province
KC554547		KC554547	Bacteria; environmental samples.	soil of Yanshan Mountain	China: Hebei Province
KC554578		KC554578	Bacteria; environmental samples.	soil of Yanshan Mountain	China: Hebei Province
KC554647		KC554647	Bacteria; environmental samples.	soil of Yanshan Mountain	China: Hebei Province
KC554883		KC554883	Bacteria; environmental samples.	soil of Yanshan Mountain	China: Hebei Province
KC852989	KC852989		Microbial arsenic transformation associated with a soda Lake in Khovsgol, Mongolia	lake sediment	Mongolia: Khovsgol, Tuh Lake
KF182251		KF182251	Biodiversity of soil around the coal gangue dump	coal gangue dump soil	
KF712861		KF712861	Analysis of bacterial community in rhizosphere of <i>Paeonia ostii</i> by 16S ribotyping	<i>Paeonia ostii</i> rhizosphere soil	
KF836161	KF836161		Analysis of bacterial community in rhizosphere of <i>Paeonia ostii</i> by 16S ribotyping	prings and wells fed by a deep, fractured rock aquifer in the Mojave Desert	
KJ191881		KJ191881	Effects of pinewood nematode (<i>Bursaphelenchus xylophilus</i>) infected- <i>Pinus massoniana</i> on soil bacterial communities	<i>Pinus massoniana</i> soil	
KJ808312		KJ808312	Impacts of sludge bulking on functional bacteria in a full-scale municipal wastewater treatment plant	activated sludge	
KT905446		KT905446	Analysis of different geological periods of soil microbial diversity in Nihewan Basin	soil	Nihewan Basin, China
KT905463		KT905463	Analysis of different geological periods of soil microbial diversity in Nihewan Basin	soil	Nihewan Basin, China
KT905474		KT905474	Analysis of different geological periods of soil microbial diversity in Nihewan Basin	soil	Nihewan Basin, China
KT905475		KT905475	Analysis of different geological periods of soil microbial diversity in Nihewan Basin	soil	Nihewan Basin, China
KT905479		KT905479	Analysis of different geological periods of soil microbial diversity in Nihewan Basin	soil	Nihewan Basin, China
KT905484		KT905484	Analysis of different geological periods of soil microbial diversity in Nihewan Basin	soil	Nihewan Basin, China

Anhang

KT905496		KT905496	Analysis of different geological periods of soil microbial diversity in Nihewan Basin	soil	Nihewan Basin, China
KJ907773		KJ907773	Metagenomic diversity survey of photosynthetic microbial communities from thermal and acidophilic biofilms	thermal and acidic green biofilms from a fumarole	Mexico: Michoacan, Los Azufres
KM108688	KM108688		Analysis of Microbial Diversity of the Tomato Soil in Yunnan E'shan	soil	Yunnan E'shan, China
KP299814		KP299814	A metagenomic approach to study the bacterial communities of bioaugmented and non-bioaugmented sediments	shrimp pond sediments	
KR814253	KR814253		Activity and interactions of methane seep microorganisms assessed by parallel transcription and FISH-NanoSIMS analyses	mud volcano sediment (Mound 12) anoxic bottle incubation	Costa Rica
KT257823	KT257823		Bacterial diversity and successional patterns during biofilm formation on freshly exposed basalt surfaces at diffuse-flow deep-sea vents	diffuse flow vent surfaces	East Pacific Rise
Klade 5 (N.J) / Klade G (ML)			title	Sample specification	Localization
AM934748			Bacterial community changes during bioremediation of aliphatic hydrocarbon-contaminated soil	pilot-scale bioremediation process of a hydrocarbon-contaminated soil	France
AY188328			16S rDNA Diversity of Bacteria in Penguin Droppings Sediments from Ardley Island, Antarctica	penguin droppings sediment	Antarctica: Ardley Island
AY218680			16S rDNA Diversity of Bacteria in Penguin Droppings Sediments from Ardley Island, Antarctica	penguin droppings sediment	Antarctica: Ardley Island
AY921749			Comparative metagenomics of microbial communities	farm soil adjacent to a silage storage bunker	USA: Minnesota
EF590039			Change of bacterial communities in sediments along Songhua River in Northeastern China after a nitrobenzene pollution event	nitrobenzene-contaminated river sediment	
EF632885			Unique microbial communities in contrasting aquatic environments of the high altitude Andean Altiplano (northern Chile)	freshwater sediment	Chile: Altiplano, Bofedal de Parinacota
EU193059			Molecular characterization of soil bacterial communities in contrasting zero tillage systems	agricultural soil	Mexico
EU202842			PCR-DGGE and Phylogenetic Approach to Analyze the Effect of Agricultural Practices on Soil Bacterial Communities	agricultural soil	Mexico
EU234231			Bacterial communities in penicillin G production wastewater and the receiving river by using culture and unculture techniques	downstream of Wang Yang River, receiving penicillin G production wastewater	China: Shijiazhuang City, Hebei Province
EU276549			Molecular characterization of soil bacterial communities in contrasting zero tillage systems	agricultural soil	Mexico: Texcoco
EU276550			Molecular characterization of soil bacterial communities in contrasting zero tillage systems	agricultural soil	Mexico: Texcoco
EU440654			Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities	agricultural soil	Mexico
EU440657			Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities	agricultural soil	Mexico
EU440658			Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities	agricultural soil	Mexico
EU449585			Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities	agricultural soil	Mexico
EU449591			Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities	agricultural soil	Mexico
EU449741			Molecular characterization of soil bacterial communities in contrasting zero tillage systems PCR-DGGE and Phylogenetic Approaches to Analyze the Effect of Agricultural Practices on Soil Bacterial Communities	agricultural soil	Mexico
EU665102			Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities	agricultural soil	Mexico
EU665104			Phylogenetic and multivariate analyses to determine the effects of different tillage and residue management practices on soil bacterial communities	agricultural soil	Mexico
EU979110			Effects of intercropping and Rhizobium inoculation on microbial biomass and bacterial diversity in rhizosphere of faba bean (<i>Vicia faba</i> L.)	soil	
FJ152768			Changes in the bacterial populations of the highly alkaline saline soil of the former lake Texcoco (Mexico) following flooding	alkaline saline soils of the former lake Texcoco	Former lake Texcoco, Mexico
FJ152847			Changes in the bacterial populations of the highly alkaline saline soil of the former lake Texcoco (Mexico) following flooding	alkaline saline soils of the former lake Texcoco	Former lake Texcoco, Mexico
FJ444679			Analysis of bacterial communities in rhizosphere soil of healthy and diseased cotton (<i>Gossypium</i> sp.) at different plant growth stages	cotton rhizosphere	JiningCity, China
FJ444686			Analysis of bacterial communities in rhizosphere soil of healthy and diseased cotton (<i>Gossypium</i> sp.) at different plant growth stages	cotton rhizosphere	JiningCity, China
FJ444693			Analysis of bacterial communities in rhizosphere soil of healthy and diseased cotton (<i>Gossypium</i> sp.) at different plant growth stages	cotton rhizosphere	JiningCity, China
FJ479166			Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ479410			Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ478799			Vertical distribution and diversity of sulfate-reducing prokaryotes in the Pearl River estuarine sediments, Southern China	Pearl River Estuary sediments	China: Guangdong Province
FJ889267			The effect of tillage and residues management on soil bacterial communities: A phylogenetic analysis	agricultural soil	Mexico

Anhang

FJ889268		The effect of tillage and residues management on soil bacterial communities: A phylogenetic analysis	agricultural soil	Mexico
FJ889269		The effect of tillage and residues management on soil bacterial communities: A phylogenetic analysis	agricultural soil	Mexico
FJ889270		The effect of tillage and residues management on soil bacterial communities: A phylogenetic analysis	agricultural soil	Mexico
GQ264485		Application of molecular techniques to elucidate the influence of cellulosic waste on the bacterial community structure at a simulated low-level-radioactive-waste site	simulated low level waste site	USA: Idaho National Labs
GQ472385		Bacterial communities in sediments of the shallow Lake Dongping in China	lake sediment	China, Dongping Lake
GU208418		Bacterial communities in sediments of the shallow Lake Dongping in China	lake sediment	China, Dongping Lake
HE650095		Characterization and comparison of bacterial communities operating in conventional activated sludge and membrane bioreactor pilot plants fed with municipal wastewater: a focus on Nitrospira and Planctomycetes bacterial phyla	MBR wastewater treatment pilot plant	
HE974816		Mobility of Arsenic, Chromium, Molybdenum and Vanadium in Ajka red mud contaminated soils during progressive anoxia	red mud contaminated soil	Hungary
HM438457		Modifications of bacterial populations in anthracene contaminated soil	soil contaminated with anthracene	
HM438458		Modifications of bacterial populations in anthracene contaminated soil	soil contaminated with anthracene	
HM438460		Modifications of bacterial populations in anthracene contaminated soil	soil contaminated with anthracene	
HM445053		Comparison of Bacterial Diversity in Azorean and Hawaiian Lava Cave Microbial Mats Investigation of Novel Microbial Diversity in Azorean and Hawaiian Lava Tubes	white microbial mat from lava tube walls	Portugal: Azores, Terceira, Gruta da Achada
HM445452		Comparison of Bacterial Diversity in Azorean and Hawaiian Lava Cave Microbial Mats Investigation of Novel Microbial Diversity in Azorean and Hawaiian Lava Tubes	yellow microbial mat from lava tube walls	Azores, Terceira, Gruta dos Principiantes
HM447781		Population analysis of soil from permanent or conventional tilled beds	agricultural soil	
HQ014654		Bacteria; environmental samples.China, wastewater treatment plant	wastewater treatment plant	
HQ114020		Bacterial community composition and function in a full-scalevermifilter system treated with continuous rural sewage biofilms in a full-scale vermifilter	biofilms in a full-scale vermifilter	
HQ114035		Bacterial community composition and function in a full-scalevermifilter system treated with continuous rural sewage biofilms in a full-scale vermifilter	biofilms in a full-scale vermifilter	
HQ114125		Bacterial community composition and function in a full-scalevermifilter system treated with continuous rural sewage biofilms in a full-scale vermifilter	biofilms in a full-scale vermifilter	
HQ132470		Bacterial community of heavy metal contaminated estuarine sediment from Xiangjiang River estuary in China heavy metal contaminated estuarine sediment	heavy metal contaminated estuarine sediment	
HQ190412		Bacterial diversity of Zhongyuan oil field, China Zhongyuan oil field	Zhongyuan oil field	China
HQ640626		The microbial community composition and qualifying analysis in the laboratory-scale stable partial nitrifying-ANAMMOX municipal wastewater reactor partial nitrifying-ANAMMOX municipal wastewater reactor	partial nitrifying-ANAMMOX municipal wastewater reactor	China: Beijing
HQ904185		Bacterial community composition during anaerobic degradation of Microcystis blooms and surface sediment mixture under differenttemperatures	bacterial community composition of day 60 sample during anaerobic degradation of Microcystis blooms and surface sediment mixture under 25 degrees centigrade	China
JF265839		Bacteria; environmental samples.	white microbial mat lava tube wall	Portugal: Gruta dos Montanheiros, Pico, Azores
JF776952		Bacterial community diversity in rhizosphere soil of eggplant in Dalian	rhizosphere soil	China
JF776954		Bacterial community diversity in rhizosphere soil of eggplant in Dalian	rhizosphere soil	China
JF833707			potassium mine soil	China: Yuexi, Anhui
JN408875			potassium mine soil	China: Yuexi, Anhui
JN409092		Bacterial community diversity in rhizosphere soil of cucumber in Dalian		China
JN559177		Soil microbial communities after conversion of abandoned lands in the Loess Plateau, China		China: Loess Plateau, Shanxi
JN825451		Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline lake Alchichica (Mexico)	microbialites from Alchichica alkaline lake maintained in aquarium	Mexico
JQ426554		Bacterial diversity in an alkaline saline soil spiked with anthracene	soil	
JQ427474		Bacterial diversity in an alkaline saline soil spiked with anthracene	soil	
JQ427578		Bacterial diversity in an alkaline saline soil spiked with anthracene	soil	
JQ428008		Bacterial diversity in an alkaline saline soil spiked with anthracene	soil	
JQ800855		Bacterial diversity of soil samples from the Yellow River Delta	soil	China: Yellow River Delta

Anhang

JQ978890			The impact of inorganic nitrogen on soil atrazine bioremediation	atrazine-contaminated soil added with inorganic nitrogen and DAT1	China
JX133285			Uncultured bacteria in soil from banana plantation infected by banana fusarium wilt	diseased soil	
JX489855			p-Coumaric acid influenced cucumber rhizosphere soil microbial communities and the growth of <i>Fusarium oxysporum</i> f.sp. <i>Cucumerinum</i> Owen	soil	China: Heilongjiang, Harbin
KC138688			environmental samples	saffron bulk soil	
KC172195			Microbial Communities Diversity of Ginger Continuous Cropping Soil in Anqiu, Shandong	ginger continuous cropping soil	China: Anqiu, Shandong
KC172204			Microbial Communities Diversity of Ginger Continuous Cropping Soil in Anqiu, Shandong	ginger continuous cropping soil	China: Anqiu, Shandong
KC331530			Mulching practices altered soil bacterial community structure and improved orchard productivity and apple quality after five growing seasons	apple orchard	China: Loess Plateau
KC331541			Mulching practices altered soil bacterial community structure and improved orchard productivity and apple quality after five growing seasons	apple orchard	China: Loess Plateau
KC541080			Bacterial community structure in river sediments associated with mining	river sediment	
KF494473			Vertical changes of the structure of bacterial communities through a permafrost core profile from Qinghai-Tibet Plateau	permafrost soil	China: Qinghai-Tibet Plateau
KF712824			Analysis of bacterial community in rhizosphere of <i>Paeonia ostii</i> by 16S riboprinting	<i>Paeonia ostii</i> rhizosphere soil	
KJ192061			Effects of pinewood nematode (<i>Bursaphelenchus xylophilus</i>) infected- <i>Pinus massoniana</i> on soil bacterial communities	<i>Pinus massoniana</i> soil	
KJ578055			<i>Typha</i> rhizosphere bacteria	Yong Ding river	China: Beijing
KJ782244			Diversity of 16S rDNA genes from bacteria of ramie rhizosphere soil	ramie rhizosphere soil	China
KJ783023			Bacterial 16S rRNA sequences obtained from the petrochemical wastewater treatment plant activated sludge by clone library sequencing technology	petrochemical wastewater treatment plant sludge	
KM200436			Bacteria from tobacco rhizospheric soil applied with fungicide Metalaxyl-mancozeb	tobacco rhizospheric soil applied with Metalaxyl-mancozeb	
KM200760			Shifts in a bacterial community composition of a mesotrophic peatland after wildfire	mesotrophic peatland Galitski Mokh, unburned site	Russia
KN566343			Draft genome of the hookworm <i>Oesophagostomum dentatum</i>	<i>Oesophagostomum dentatum</i> , hookworm	Germany: Lower Saxony
LN567854			Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 12 layer 4	Panama: Gamboa, Pipeline Road harboring diverse microbial assemblages; 9.13 N 79.72 W
LN567855			Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 12 layer 5	Panama: Gamboa, Pipeline Road harboring diverse microbial assemblages; 9.13 N 79.72 W
LN567857			Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 12 layer 6	Panama: Gamboa, Pipeline Road harboring diverse microbial assemblages; 9.13 N 79.72 W
LN567860			Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 12 layer 7	Panama: Gamboa, Pipeline Road harboring diverse microbial assemblages; 9.13 N 79.72 W
LN571069			Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 2	Panama: Gamboa, Pipeline Road harboring diverse microbial assemblages; 9.15 N 79.73 W
LN572017			Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 3	Panama: Gamboa, Pipeline Road harboring diverse microbial assemblages; 9.15 N 79.73 W
LN572018			Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 4	Panama: Gamboa, Pipeline Road harboring diverse microbial assemblages; 9.15 N 79.73 W
Klade H (ML)	Klade 6 (NJ)	Klade 7 (NJ)	title	Sample specification	localization
22-clone1		22-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
22-clone5		22-clone5	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
22-clone7		22-clone7	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
23-clone1		23-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
23-clone3		23-clone3	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
25-clone1		25-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
25-clone6		25-clone6	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
25-clone8		25-clone8	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
26-clone1		26-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1

Anhang

26-clone5		26-clone5	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
44-clone22		44-clone22	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
45-clone22		45-clone22	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
45-clone23		45-clone23	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
45-clone24		45-clone24	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
46-clone21		46-clone21	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
46-clone22		46-clone22	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
47-clone21		47-clone21	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
47-clone22		47-clone22	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
47-clone23		47-clone23	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
5-clone1		5-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
5-clone4		5-clone4	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
6-clone11		6-clone11	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
6-clone15		6-clone15	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
6-clone18		6-clone18	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
8-clone23		8-clone23	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
8-clone27		8-clone27	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
8-clone30		8-clone30	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
8-clone4		8-clone4	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
8-clone7		8-clone7	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
A-clone11		A-clone11	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
A-clone2		A-clone2	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
A-clone15		A-clone15	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
AY326516		AY326516	Bacterial diversity of terra preta and pristine forest soil from the Western Amazon	soil	Jamari National Forest, of Rondonia, Brazil 8°45'0 S, 63°27'0W
B-clone7		B-clone7	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
B-clone8		B-clone8	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
B-clone9		B-clone9	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
C-clone6		C-clone6	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
C-clone7		C-clone7	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
C-clone8		C-clone8	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1

Anhang

D-clone4		D-clone4	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
D-clone6		D-clone6	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
D-clone7		D-clone7	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
DQ451518		DQ451518	Bacterial Community Composition in Fushan Forest Soils of Taiwan	forest soil	
E-clone1		E-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
E-clone20		E-clone20	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
E-clone23		E-clone23	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
EF494339		EF494339	Interplay between landscape position and microbial abundance impacts geochemical characteristics of soil and saprolite	Nunnock River granitic landscape	Australia
EF516565		EF516565	Despite strong seasonal responses, soil microbial consortia are more resilient to long-term changes in rainfall than overlying grassland	grassland soil- fälschlicherweise als marin gekennzeichnet	USA: northern California, Angelo Coast Range Reserve
EU134306		EU134306	Novelty and uniqueness patterns of rare members of the soil biosphere Phylogenetic diversity and evolutionary relationships between rare and abundant members of the bacterial community in tallgrass prairie soil	soil from an undisturbed mixed grass prairie preserve	USA: Oklahoma, Kessler farm
EU134335	EU134335		Novelty and uniqueness patterns of rare members of the soil biosphere	soil from an undisturbed mixed grass prairie preserve	USA: Oklahoma, Kessler farm
EU134348		EU134348	Novelty and uniqueness patterns of rare members of the soil biosphere Phylogenetic diversity and evolutionary relationships between rare and abundant members of the bacterial community in tallgrass prairie soil	soil from an undisturbed mixed grass prairie preserve	USA: Oklahoma, Kessler farm
EU134350		EU134350	Novelty and uniqueness patterns of rare members of the soil biosphere Phylogenetic diversity and evolutionary relationships between rare and abundant members of the bacterial community in tallgrass prairie soil	soil from an undisturbed mixed grass prairie preserve	USA: Oklahoma, Kessler farm
EU134352		EU134352	Novelty and uniqueness patterns of rare members of the soil biosphere Phylogenetic diversity and evolutionary relationships between rare and abundant members of the bacterial community in tallgrass prairie soil	soil from an undisturbed mixed grass prairie preserve	USA: Oklahoma, Kessler farm
EU134370	EU134370		Novelty and uniqueness patterns of rare members of the soil biosphere	Novelty and uniqueness patterns of rare members of the soil biosphere	USA: Oklahoma, Kessler farm
EU134409	EU134409		Novelty and uniqueness patterns of rare members of the soil biosphere	Novelty and uniqueness patterns of rare members of the soil biosphere	USA: Oklahoma, Kessler farm
EU335175		EU335175	Changes in bacterial and archaeal community structure and functional diversity along a geochemically variable soil profile	saturated C horizon soil aggregate	Melton Branch Watershed, Oak Ridg
EU335394		EU335394	Changes in bacterial and archaeal community structure and functional diversity along a geochemically variable soil profile	saturated C horizon soil aggregate	Melton Branch Watershed, Oak Ridg
EU421843		EU421843	Un-culturable diversity of soil sample collected from Lahaul-Spiti Valley of The Indian Himalayas	soil sample collected under a glacier from Lahaul-Spiti Valley	India: Himachal Pradesh, Himalayas cold desert
EU477104		EU477104	Molecular characterization of soil bacterial communities in contrasting zero tillage systems PCR-DGGE and Phylogenetic Approaches to Analyze the Effect of Agricultural Practices on Soil Bacterial Communities	agricultural soil	Mexico
EU669605		EU669605	Bacteria; environmental samples	soil	China: Guangxi, Huanjiang, typical karst peak-cluster depression
EU881165		EU881165	Soil bacterial community composition and diversity respond to cultivation in Karst ecosystems	primary forest soil	China: Guangxi Autonomous Region
EU881305		EU881305	The bacterial community of cropland (maize-sweet potato cropping field) in Karst region	maize-sweet potato cropland soil	China: Guangxi Autonomous Region
F-clone1		F-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
F-clone4		F-clone4	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
F-clone9		F-clone9	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
FJ478594		FJ478594	Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ478993		FJ478993	Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ479028		FJ479028	Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ479187		FJ479187	Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station
FJ479243	FJ479243		Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler farm
FJ479603		FJ479603	Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys	undisturbed tall grass prairie, top 5 cm	USA: Oklahoma, Kessler Farm Biological Station

Anhang

FJ592583		FJ592583	Fumarole-supported islands of biodiversity within a hyperarid, high-elevation landscape on Socompa Volcano, Puna de Atacama, Andes	cold-fumarole soil (elev. 5824 m) collected at Socompa Volcano, Andes	Socompa Volcano, Andes
GQ214124		GQ214124	Bacterial diversity of loess in China	Loess	China
H-clone1		H-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
HE974828	HE974828		Mobility of Arsenic, Chromium, Molybdenum and Vanadium in Ajka red mud contaminated soils during progressive anoxia	red mud contaminated soil	Hungary
HE985152	HE985152		Phylogenetic diversity and community structure of the symbionts associated with the coralline sponge <i>Astrosclera willeyana</i> of the Great Barrier Reef	sponge tissue	Australia:Great Barrier Reef, Yonge Reef
HF012446	HF012446		BAC End Sequencing of Barley cv. Morex	<i>Hordeum vulgare</i> subsp. <i>Vulgare</i>	
HM445034		HM445034	Comparison of Bacterial Diversity in Azorean and Hawaiian Lava Cave Microbial Mats Investigation of Novel Microbial Diversity in Azorean and Hawaiian Lava Tubes	white microbial mat from lava tube walls	Portugal: Azores, Terceira, Gruta da
HM445222	HM445222		Comparison of Bacterial Diversity in Azorean and Hawaiian Lava Cave Microbial Mats Investigation of Novel Microbial Diversity in Azorean and Hawaiian Lava Tubes	yellow microbial mat from lava tube walls	Portugal: Azores, Terceira, Gruta Branca Opala
HM749700	HM749700		Microbial diversity of the lava tubes in Pico Island, Azores	yellow microbial mat from lava tube walls	Portugal: Azores, Terceira, Gruta Branca Opala
HQ270336	HQ270336		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado, 01°32 N, 124°55 E
HQ270341	HQ270341		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado, 01°32 N, 124°55 E
HQ270343	HQ270343		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado, 01°32 N, 124°55 E
HQ270344	HQ270344		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado, 01°32 N, 124°55 E
HQ270345	HQ270345		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado, 01°32 N, 124°55 E
HQ270347	HQ270347		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado, 01°32 N, 124°55 E
HQ270350	HQ270350		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado, 01°32 N, 124°55 E
HQ270363	HQ270363		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado, 01°32 N, 124°55 E
HQ270378	HQ270378		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia muta</i> (marine sponge)	USA: Key Largo, Florida, 24°56.82 N, 80°27.40 W)
HQ270385	HQ270385		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia muta</i> (marine sponge)	USA: Key Largo, Florida, 24°56.82 N, 80°27.40 W)
HQ270391	HQ270391		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia muta</i> (marine sponge)	USA: Key Largo, Florida, 24°56.82 N, 80°27.40 W)
HQ397211		HQ397211	Bacterial diversity in normal soil	Soil	21.91613 N 71.30910 E
HQ645175		HQ645175	Shifting prokaryotic community structure along different land management in meadow in the Qinghai-Tibet Plateau, China	soil samples from meadow	China: Tibet Plateau
HQ697463	HQ697463		Bromate reduction and bacterial communities in biologically activated carbon	biologically activated carbon for drinking water treatment	
HQ697487	HQ697487		Bromate reduction and bacterial communities in biologically activated carbon	biologically activated carbon for drinking water treatment	
HQ697492	HQ697492		Bromate reduction and bacterial communities in biologically activated carbon	biologically activated carbon for drinking water treatment	
J-clone1		J-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
J-clone3		J-clone3	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
J-clone7		J-clone7	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
JF809701	JF809701		Low Bacterial Diversity and High Labile Organic Matter Concentrations in the Sediments of the Medee Deep-sea Hypersaline Anoxic Basin	Medea hypersaline basin, Mediterranean Sea	Mediterranean Ridge
JF833844		JF833844		potassium mine soil	China: Yuexi, Anhui
JN037992		JN037992		potassium mine soil	China: Yuexi, Anhui
JN596706	JN596706		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado
JN596740	JN596740		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado
JN596761	JN596761		Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts	<i>Xestospongia testudinaria</i> (marine sponge)	Indonesia: Manado
JN615914	JN615914		Lava cave bacterial diversity	tan microbial mat from lava cave wall	Portugal: Gruta da Madre de Deus, Terceira, Azores
JQ311913		JQ311913	Identification of cellulose-responsive bacterial and fungal communities in geographically and edaphically different soils by using stable isotope probing	soil microcosm	USA, HickoryCorners, MI, Wilmington, NC
JQ516312	JQ516312		The <i>Montastraea faveolata</i> microbiome: ecological and temporal influences on a Caribbean reef-building coral in decline	diseased tissue	17°56.093 N and 67°02.931 W
JQ612299	JQ612299		Metatranscriptomics of the marine sponge <i>Geodia barretti</i> : tackling phylogeny and function of its microbial community	<i>Geodia barretti</i> (marine sponge), 200–300 m depth	Norway, 59.58.7N 05.22.89E

Anhang

JQ612361	JQ612361		Metatranscriptomics of the marine sponge <i>Geodia barretti</i> : tackling phylogeny and function of its microbial community	<i>Geodia barretti</i> (marine sponge), 200–300 m depth	Norway, 59.58.7N 05.22.89E
JQ675526		JQ675526	Bacterial Diversity of Isolated Carbonate Cave Pools in New Mexico	water sample	USA: Lechuguilla Cave, Sugarlands Pool, Carlsbad, New Mexico
JX133457		JX133457	Uncultured bacteria in soil from healthy banana plantation	healthy soil	
JX280314	JX280314		Biogeography and Host Fidelity of Bacterial Communities in <i>Ircinia</i> spp. from the Bahamas	<i>Ircinia strobilina</i> (marine sponge)	Bahamas: Sweeting's Cay, 26°33.578 N, 077°53.036 W
JX280334	JX280334		Biogeography and Host Fidelity of Bacterial Communities in <i>Ircinia</i> spp. from the Bahamas	<i>Ircinia strobilina</i> (marine sponge)	Bahamas: Sweeting's Cay, 26°33.578 N, 077°53.036 W
JX988747	JX988747		Characterizing the bacterial associates of three Caribbean sponges along a gradient from shallow to mesophotic depths	marine sponge	19°42.03'N, 80°03.25'W
KC331469	KC331469		Mulching practices altered soil bacterial community structure and improved orchard productivity and apple quality after five growing seasons	apple orchard	China: Loess Plateau
K-clone1		K-clone1	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
K-clone5		K-clone5	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
K-clone6		K-clone6	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
KF836245		KF836245	Planktonic Microbial Communities of High Volume Springs: Ash Meadows Nevada, US Great Basin	springs and wells fed by a deep, fractured rock aquifer in the Mojave Desert	USA
KJ192015		KJ192015	Effects of pinewood nematode (<i>Bursaphelenchus xylophilus</i>) infected-Pinus massoniana on soil bacterial communities	Pinus massoniana soil	
KJ192016		KJ192016	Effects of pinewood nematode (<i>Bursaphelenchus xylophilus</i>) infected-Pinus massoniana on soil bacterial communities	Pinus massoniana soil	
KM389589	KM389589		Polyketide Synthases in the Microbiome of the Marine Sponge <i>Plakortis halichondrioides</i> : A Metagenomic Update	<i>Plakortis halichondrioides</i> (marine sponge)	24°35.167' N; 75°58.476' W
L-clone10		L-clone10	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
L-clone11		L-clone11	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
L-clone2		L-clone2	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
L-clone7		L-clone7	Siehe Tabelle 5.1	Siehe Tabelle 5.1	Siehe Tabelle 5.1
LN563587		LN563587	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 11 layer 4, <i>Atta colombica</i>	Panama: Gamboa, Pipeline Road, 9.13 N 79.72 W
LN563597		LN563597	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 11 layer 4, <i>Atta colombica</i>	Panama: Gamboa, Pipeline Road, 9.13 N 79.72 W
LN563598		LN563598	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 11 layer 4, <i>Atta colombica</i>	Panama: Gamboa, Pipeline Road, 9.13 N 79.72 W
LN563599		LN563599	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 11 layer 4, <i>Atta colombica</i>	Panama: Gamboa, Pipeline Road, 9.13 N 79.72 W
LN563600		LN563600	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 11 layer 4, <i>Atta colombica</i>	Panama: Gamboa, Pipeline Road, 9.13 N 79.72 W
LN563601		LN563601	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 11 layer 4, <i>Atta colombica</i>	Panama: Gamboa, Pipeline Road, 9.13 N 79.72 W
LN572974		LN572974	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 4	Panama: Gamboa, Pipeline Road, 9.15 N 79.73 W
LN572975		LN572975	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 5	Panama: Gamboa, Pipeline Road, 9.15 N 79.73 W
LN572976		LN572976	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 6	Panama: Gamboa, Pipeline Road, 9.15 N 79.73 W
LN572977		LN572977	Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 7	Panama: Gamboa, Pipeline Road, 9.15 N 79.73 W
LN572980	LN572980		Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 4, <i>Atta colombica</i>	Panama: Gamboa, Pipeline Road, 9.15 N 79.73 W
Klade 8 (NJ)			title	Sample specification	localization
AB683979			Calyxamides A and B, cytotoxic cyclic peptides from the marine sponge <i>Discodermia calyx</i>	<i>Discodermia calyx</i> (sponge)	Japan: Tokyo, Shikine-jima
AF142626			Identification of the antifungal peptide-containing symbiont of the marine sponge <i>Theonella swinhoei</i> as a novel delta-proteobacterium, 'Candidatus Entotheonella palauensis'	<i>Theonella swinhoei</i>	Philippines, Western Caroline Islands
AY493913			Microarray analysis of bacterial diversity and distribution in aggregates from a desert agricultural soil	Soil	Imperial Valley, California
AY897120			Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge <i>Discodermia dissoluta</i>	<i>Discodermia dissoluta</i> (sponge)	Curacao, Netherlands Antilles
AY897121			Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge <i>Discodermia dissoluta</i>	<i>Discodermia dissoluta</i> (sponge)	Curacao, Netherlands Antilles
AY897122			Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge <i>Discodermia dissoluta</i>	<i>Discodermia dissoluta</i> (sponge)	Curacao, Netherlands Antilles
AY897123			Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge <i>Discodermia dissoluta</i>	<i>Discodermia dissoluta</i> (sponge)	Curacao, Netherlands Antilles
AY897125			Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge <i>Discodermia dissoluta</i>	<i>Discodermia dissoluta</i> (sponge)	Curacao, Netherlands Antilles

Anhang

GU444089		Effects of cotton straw amendments on soil fertility and microbial community structure in cucumber, <i>Cucumis sativus</i> L., production systems	cotton straw treated soil	China
JN825509		Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline lake Alchichica (Mexico)	microbialites from Alchichica alkaline lake, 14m depth	Mexico, Laguna Alchichica
KF926701		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Aplysia aerophoba</i> (marine sponge)	Croatia, 45.13 N 13.66 E
KF926705		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Aaptos ciliata</i> (marine sponge)	Japan, 28.52 N 129.33 E
KF926706		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Aaptos ciliata</i> (marine sponge)	Japan, 28.52 N 129.33 E
KF926707		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Aaptos ciliata</i> (marine sponge)	Japan, 28.52 N 129.33 E
KF926708		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Agelas dilatata</i>	Bahamas, 24.57 N 78.97 W
KF926710		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Agelas nakamura</i>	Japan, 30.48 N 130.19 E
KF926711		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Agelas nakamura</i>	Japan, 30.48 N 130.19 E
KF926712		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Agelas nakamura</i>	Japan, 30.48 N 130.19 E
KF926713		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Cliona raromicrosclera</i> (marine sponge)	Japan, 33.47 N 132.26 E
KF926714		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Cliona raromicrosclera</i> (marine sponge)	Japan, 33.47 N 132.26 E
KF926715		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Asteropus simplex</i>	Japan, 34.32 N 139.22 E
KF926717		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Asteropus simplex</i>	Japan, 34.32 N 139.22 E
KF926718		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Amphimedon</i> sp	Japan, 30.48 N 130.19 E
KF926720		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Cacospongia mycofijiensis</i>	Vanuatu, 17.72 S 168.23 E
KF926721		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Cacospongia mycofijiensis</i>	Vanuatu, 17.72 S 168.23 E
KF926722		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Cacospongia mycofijiensis</i>	Vanuatu, 17.72 S 168.23 E
KF926723		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Ceratopsion</i> sp.	Japan, 29.79 N 130.32 E
KF926727		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Callyspongia vaginalis</i>	Bahamas, 24.57 N 78.97 W
KF926728		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Callyspongia vaginalis</i>	Bahamas, 24.57 N 78.97 W
KF926729		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dysidea avara</i>	Croatia, 45.13 N 13.66 E
KF926730		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dysidea avara</i>	Croatia, 45.13 N 13.66 E
KF926732		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Discodermia calyx</i>	Japan, 34.61 N 138.82 E
KF926733		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Discodermia calyx</i>	Japan, 34.61 N 138.82 E
KF926734		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Discodermia calyx</i>	Japan, 34.61 N 138.82 E
KF926735		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dysidea etheria</i>	Bahamas, 24.57 N 78.97 W
KF926736		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dysidea etheria</i>	Bahamas, 24.57 N 78.97 W
KF926737		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dysidea etheria</i>	Bahamas, 24.57 N 78.97 W
KF926738		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dysidea etheria</i>	Bahamas, 24.57 N 78.97 W
KF926739		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Discodermia kiensis</i>	Japan, 34.61 N 138.82 E
KF926740		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dercitus simplex</i>	Japan, 28.52 N 129.33 E
KF926741		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dercitus simplex</i>	Japan, 28.52 N 129.33 E
KF926742		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Dercitus simplex</i>	Japan, 28.52 N 129.33 E
KF926743		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Erylus placenta</i>	Japan, 33.07 N 139.78 E
KF926744		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Erylus placenta</i>	Japan, 33.07 N 139.78 E
KF926745		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Epipolasis</i> sp.	Japan, 26.15 N 127.31 E
KF926746		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Epipolasis</i> sp.	Japan, 26.15 N 127.31 E
KF926747		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Epipolasis</i> sp.	Japan, 26.15 N 127.31 E
KF926748		An environmental bacterial taxon with a large and distinct metabolic repertoire	<i>Epipolasis</i> sp.	Japan, 26.15 N 127.31 E

Anhang

KF926754		An environmental bacterial taxon with a large and distinct metabolic repertoire	Hexadella sp.	Japan, 30.48 N 130.19 E
KF926755		An environmental bacterial taxon with a large and distinct metabolic repertoire	Hexadella sp.	Japan, 30.48 N 130.19 E
KF926756		An environmental bacterial taxon with a large and distinct metabolic repertoire	Hexadella sp.	Japan, 30.48 N 130.19 E
KF926757		An environmental bacterial taxon with a large and distinct metabolic repertoire	Ircinia felix	Bahamas, 24.57 N 78.97 W
KF926758		An environmental bacterial taxon with a large and distinct metabolic repertoire	Ircinia felix	Bahamas, 24.57 N 78.97 W
KF926769		An environmental bacterial taxon with a large and distinct metabolic repertoire	Mycale magellanica	Japan, 34.64 N 138.92 E
KF926770		An environmental bacterial taxon with a large and distinct metabolic repertoire	Mycale magellanica	Japan, 34.64 N 138.92 E
KF926772		An environmental bacterial taxon with a large and distinct metabolic repertoire	Niphates digitalis	Bahamas, 24.57 N 78.97 W
KF926773		An environmental bacterial taxon with a large and distinct metabolic repertoire	Niphates digitalis	Bahamas, 24.57 N 78.97 W
KF926775		An environmental bacterial taxon with a large and distinct metabolic repertoire	Penares aff. incrustans	Japan, 33.13 N 139.80 E
KF926776		An environmental bacterial taxon with a large and distinct metabolic repertoire	Penares aff. incrustans	Japan, 33.13 N 139.80 E
KF926777		An environmental bacterial taxon with a large and distinct metabolic repertoire	Penares aff. incrustans	Japan, 33.13 N 139.80 E
KF926778		An environmental bacterial taxon with a large and distinct metabolic repertoire	Penares sp.	Japan, 28.05 N 129.22 E
KF926779		An environmental bacterial taxon with a large and distinct metabolic repertoire	Penares sp.	Japan, 28.05 N 129.22 E
KF926780		An environmental bacterial taxon with a large and distinct metabolic repertoire	Penares sp.	Japan, 28.05 N 129.22 E
KF926781		An environmental bacterial taxon with a large and distinct metabolic repertoire	Pseudoceratina purpurea	Japan, 29.83 N 129.85 E
KF926782		An environmental bacterial taxon with a large and distinct metabolic repertoire	Pseudoceratina purpurea	Japan, 29.83 N 129.85 E
KF926783		An environmental bacterial taxon with a large and distinct metabolic repertoire	Pseudoceratina purpurea	Japan, 29.83 N 129.85 E
KF926784		An environmental bacterial taxon with a large and distinct metabolic repertoire	Pseudoceratina purpurea	Japan, 29.83 N 129.85 E
KF926785		An environmental bacterial taxon with a large and distinct metabolic repertoire	Pseudoceratina purpurea	Japan, 29.83 N 129.85 E
KF926786		An environmental bacterial taxon with a large and distinct metabolic repertoire	Pseudoceratina purpurea	Japan, 29.83 N 129.85 E
KF926787		An environmental bacterial taxon with a large and distinct metabolic repertoire	Psammocinia sp.	Papua New Guinea, 9.53 S 150.28 E
KF926788		An environmental bacterial taxon with a large and distinct metabolic repertoire	Psammocinia sp.	Papua New Guinea, 9.53 S 150.28 E
KF926789		An environmental bacterial taxon with a large and distinct metabolic repertoire	Psammocinia sp.	Papua New Guinea, 9.53 S 150.28 E
KF926790		An environmental bacterial taxon with a large and distinct metabolic repertoire	Psammocinia sp.	Papua New Guinea, 9.53 S 150.28 E
KF926792		An environmental bacterial taxon with a large and distinct metabolic repertoire	Psammocinia sp.	Papua New Guinea, 9.53 S 150.28 E
KF926793		An environmental bacterial taxon with a large and distinct metabolic repertoire	Ptilocaulis sp.	Bahamas, 24.57 N 78.97 W
KF926794		An environmental bacterial taxon with a large and distinct metabolic repertoire	Ptilocaulis sp.	Bahamas, 24.57 N 78.97 W
KF926795		An environmental bacterial taxon with a large and distinct metabolic repertoire	Ptilocaulis sp.	Bahamas, 24.57 N 78.97 W
KF926802		An environmental bacterial taxon with a large and distinct metabolic repertoire	Stylissa carteri	Saudi Arabia, 22.38 N 39.05 E
KF926803		An environmental bacterial taxon with a large and distinct metabolic repertoire	Stylissa carteri	Japan, 30.81 N 130.32 E
KF926804		An environmental bacterial taxon with a large and distinct metabolic repertoire	Stylissa carteri	Japan, 30.81 N 130.32 E
KF926805		An environmental bacterial taxon with a large and distinct metabolic repertoire	Stylissa carteri	Japan, 30.81 N 130.32 E
KF926806		An environmental bacterial taxon with a large and distinct metabolic repertoire	Topsentia sp.	Japan, 31.54 N 131.42 E
KF926807		An environmental bacterial taxon with a large and distinct metabolic repertoire	Topsentia sp.	Japan, 31.54 N 131.42 E
KF926808		An environmental bacterial taxon with a large and distinct metabolic repertoire	Topsentia sp.	Japan, 31.54 N 131.42 E
KF926809		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoei chemotype W	Japan, 33.14 N 139.73 E
KF926810		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoei chemotype W	Japan, 33.14 N 139.73 E
KF926812		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoei chemotype W	Japan, 33.14 N 139.73 E
KF926813		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoei chemotype W	Japan, 33.14 N 139.73 E

Anhang

KF926814		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoi chemotype W	Japan, 33.14 N 139.73 E
KF926815		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoi chemotype W	Japan, 33.14 N 139.73 E
KF926816		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoi chemotype W	Japan, 33.14 N 139.73 E
KF926818		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoi chemotype Y	Japan, 34.61 N 138.82 E
KF926819		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoi chemotype Y	Japan, 34.61 N 138.82 E
KF926820		An environmental bacterial taxon with a large and distinct metabolic repertoire	Theonella swinhoi chemotype Y	Japan, 34.61 N 138.82 E
KF926821		An environmental bacterial taxon with a large and distinct metabolic repertoire	Xestospongia muta	Bahamas, 24.57 N 78.97 W
KF926822		An environmental bacterial taxon with a large and distinct metabolic repertoire	Xestospongia muta	Bahamas, 24.57 N 78.97 W
KJ192017		Effects of pinewood nematode (<i>Bursaphelenchus xylophilus</i>) infected-Pinus massoniana on soil bacterial communities	Pinus massoniana soil	
LC026816		Abundance and community structure of bacteria	aerosol	China:Beijing
LN572981		Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 4, Atta colombica	Panama:Gamboa, Pipeline Road, 9.15 N 79.73 W
LN572982		Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 4, Atta colombica	Panama:Gamboa, Pipeline Road, 9.15 N 79.73 W
LN572983		Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 4, Atta colombica	Panama:Gamboa, Pipeline Road, 9.15 N 79.73 W
LN572984		Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 4, Atta colombica	Panama:Gamboa, Pipeline Road, 9.15 N 79.73 W
LN572985		Leaf-cutter ant refuse dumps are nutrient reservoirs harboring diverse microbial assemblages	refuse dump Nest 9 layer 4, Atta colombica	Panama:Gamboa, Pipeline Road, 9.15 N 79.73 W

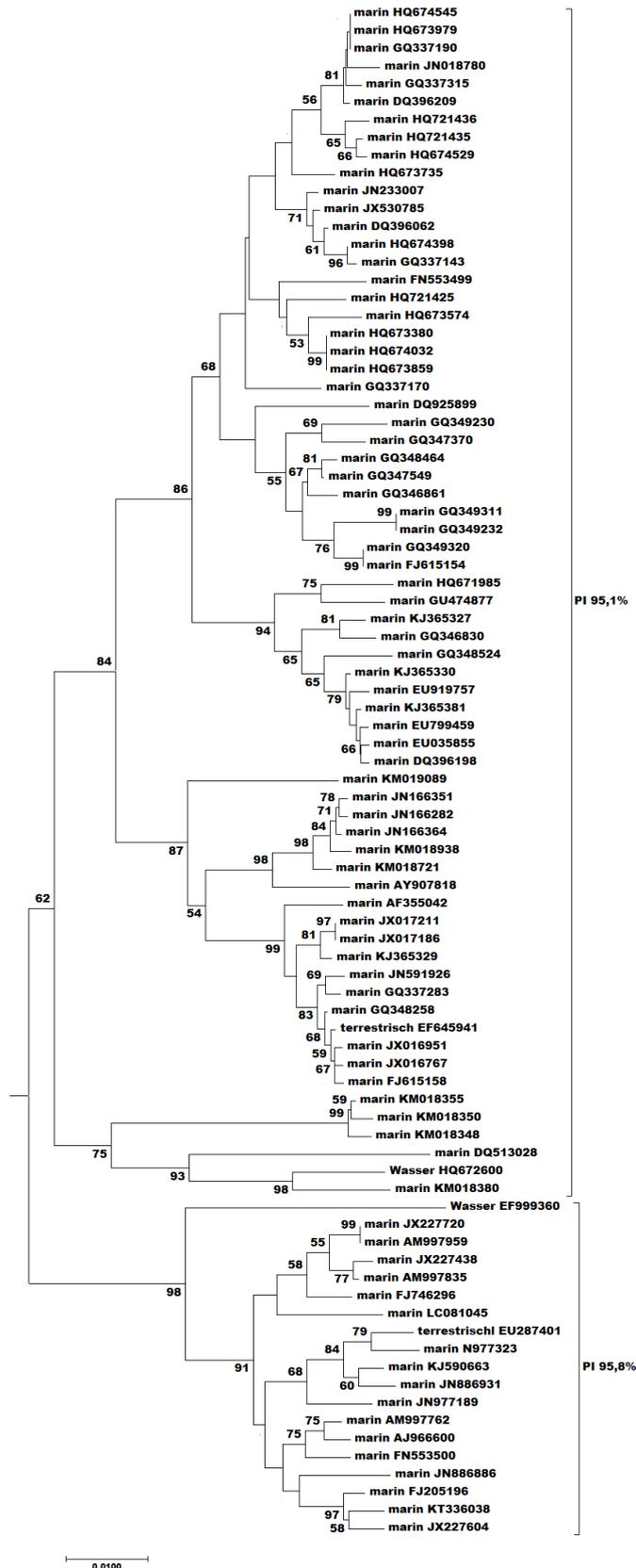


Abbildung 10-1: Aufspaltung der „Klade 1“ im NJ-Stammbaum.

Ausschnitt aus dem Phylogenetischer Stammbaum des Kandidatenphylums 'Tectomicrobia' (Abbildung 5-18, Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, bootstrapping: 1000x^{190,198,211,212}; Base 225-1130, *E. coli*; enthält die variablen Regionen V3-V6). Angegeben sind bootstrap-Werte >50%. Rechts angegeben sind die *pairwise identity* (PI)-Werte einiger Cluster innerhalb der Kandidatenfamilie.

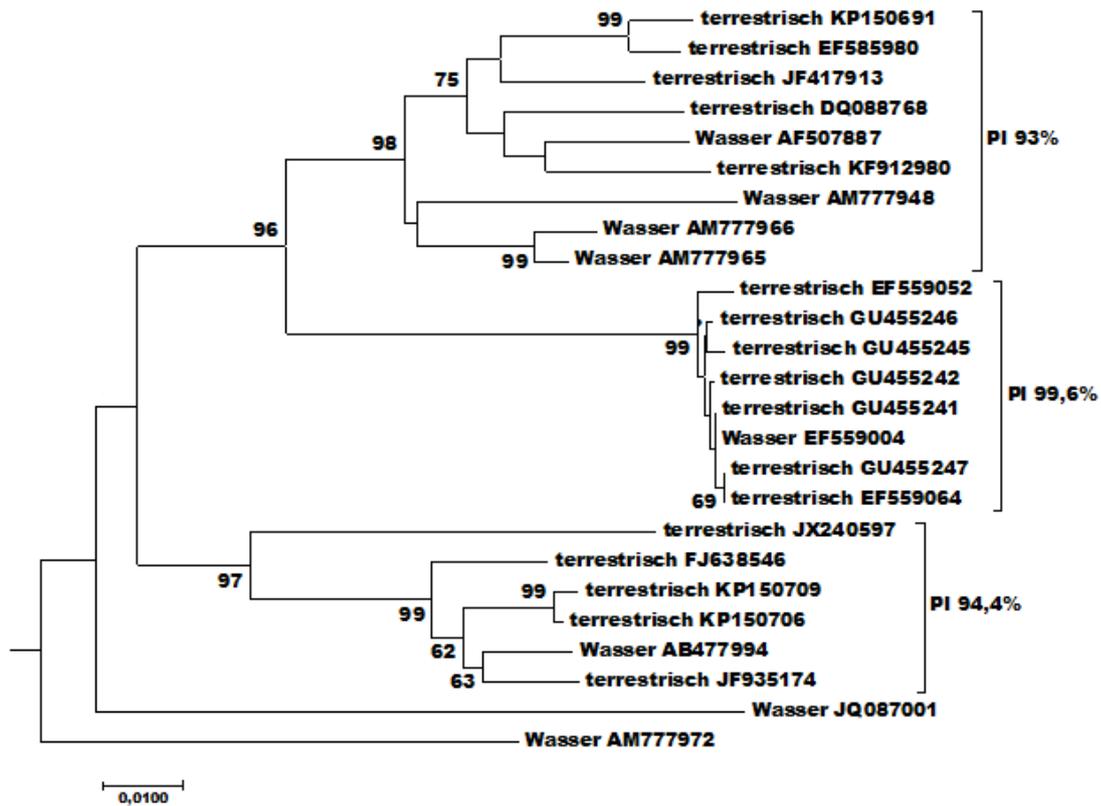


Abbildung 10-2: Aufspaltung der „Klade 2“ im NJ-Stammbaum.

Ausschnitt aus dem Phylogenetischen Stammbaum des Kandidatenphylums 'Tectomicrobia' (Abbildung 5-18, Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, bootstrapping: 1000x^{190,198,211,212}, Base 225-1130, *E. coli*; enthält die variablen Regionen V3-V6). Angegeben sind bootstrapping-Werte >50%. Rechts angegeben sind die *pairwise identity* (PI)-Werte einiger Cluster innerhalb der Kandidatenordnung.

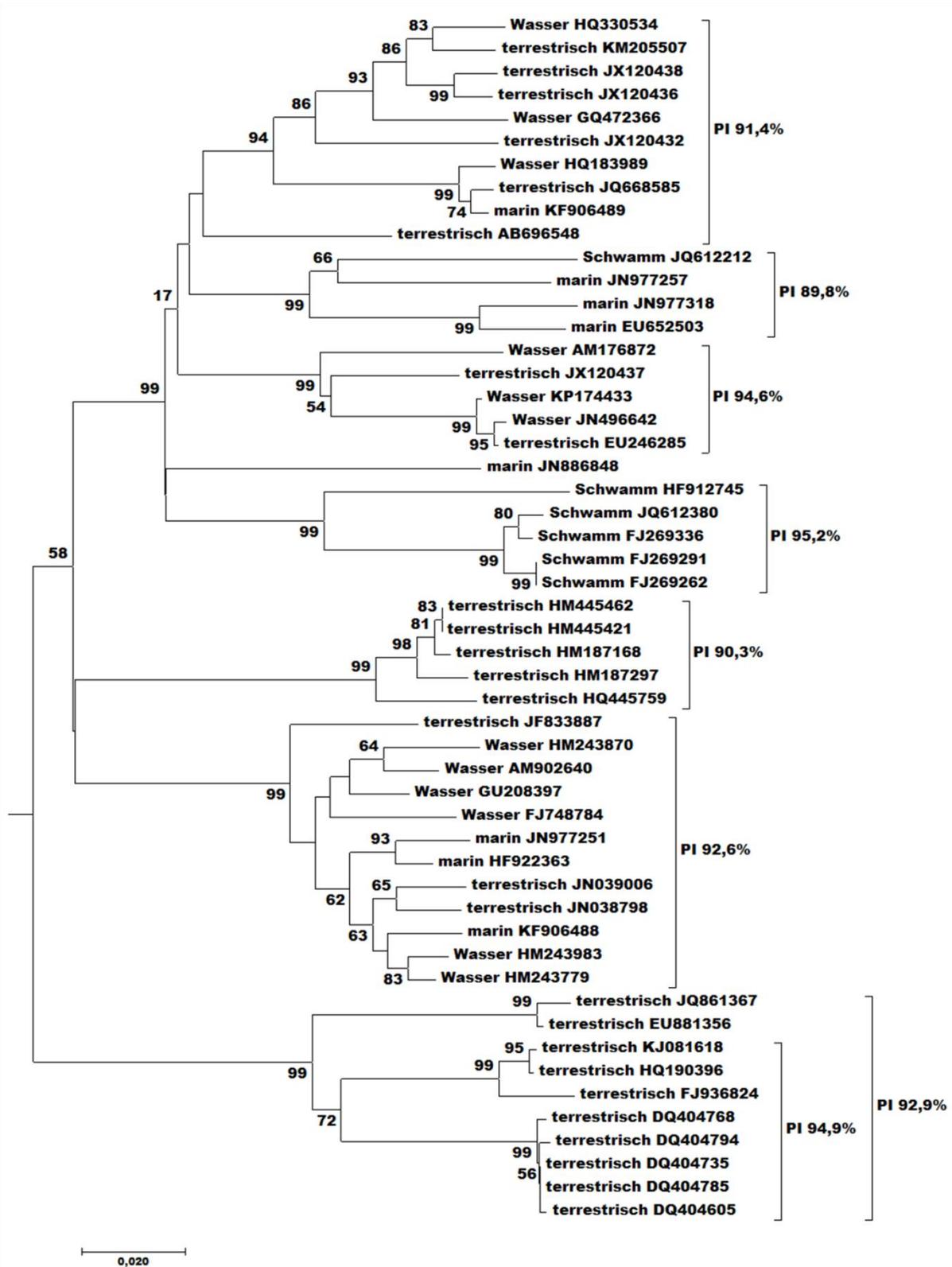


Abbildung 10-3: Aufspaltung der „Klade 3“ im NJ-Stammbaum. Ausschnitt aus dem Phylogenetischer Stammbaum des Kandidatenphylums 'Tectomicrobia' (Abbildung 5-18, Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, bootstrapping: 1000x^{190,198,211,212}; Base 225-1130, *E. coli*; enthält die variablen Regionen V3-V6). Angegeben sind bootstrap-Werte >50%. Rechts angegeben sind die *pairwise identity* (PI)-Werte einiger Cluster innerhalb der Kandidatenordnung.

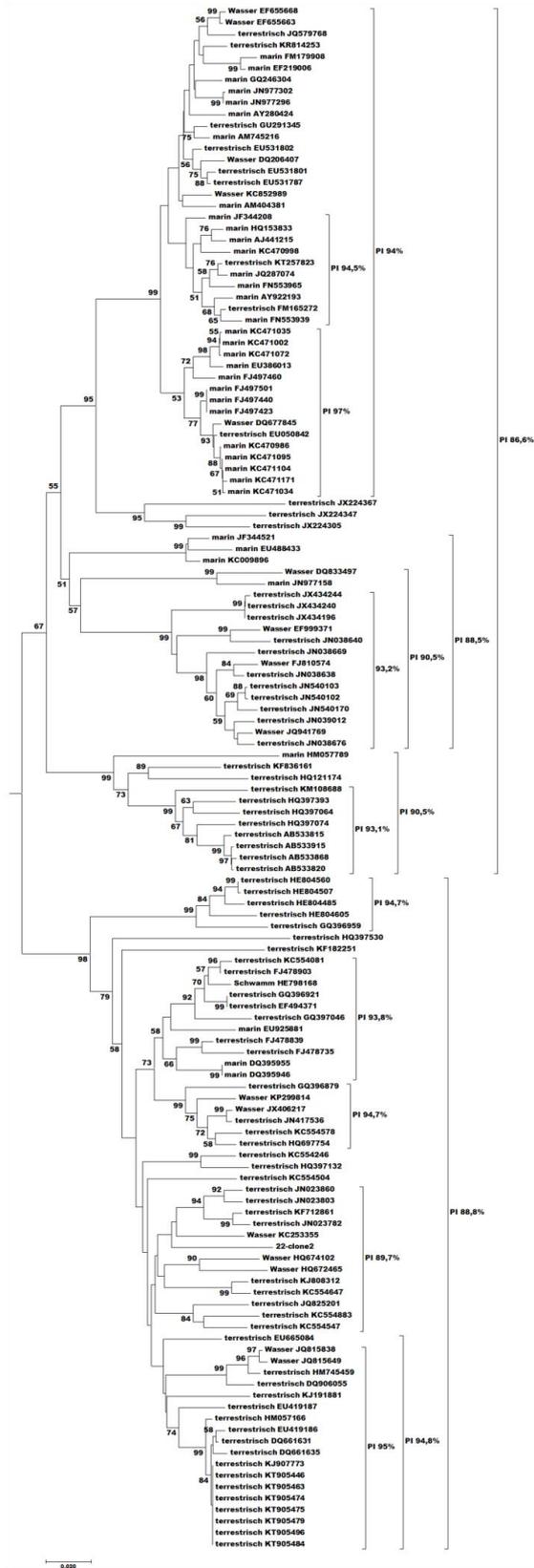


Abbildung 10-4: Aufspaltung der „Klade 4“ im NJ-Stammbaum. Ausschnitt aus dem Phylogenetischer Stammbaum des Kandidatenphylums 'Tectomicrobia' (Abbildung 5-18, Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, bootstrapping: 1000x^{190,198,211,212}, Base 225-1130, *E. coli*; enthält die variablen Regionen V3-V6). Angegeben sind bootstrap-Werte >50%. Rechts angegeben sind die *pairwise identity* (PI)-Werte einiger Cluster innerhalb der Kandidatenklasse.

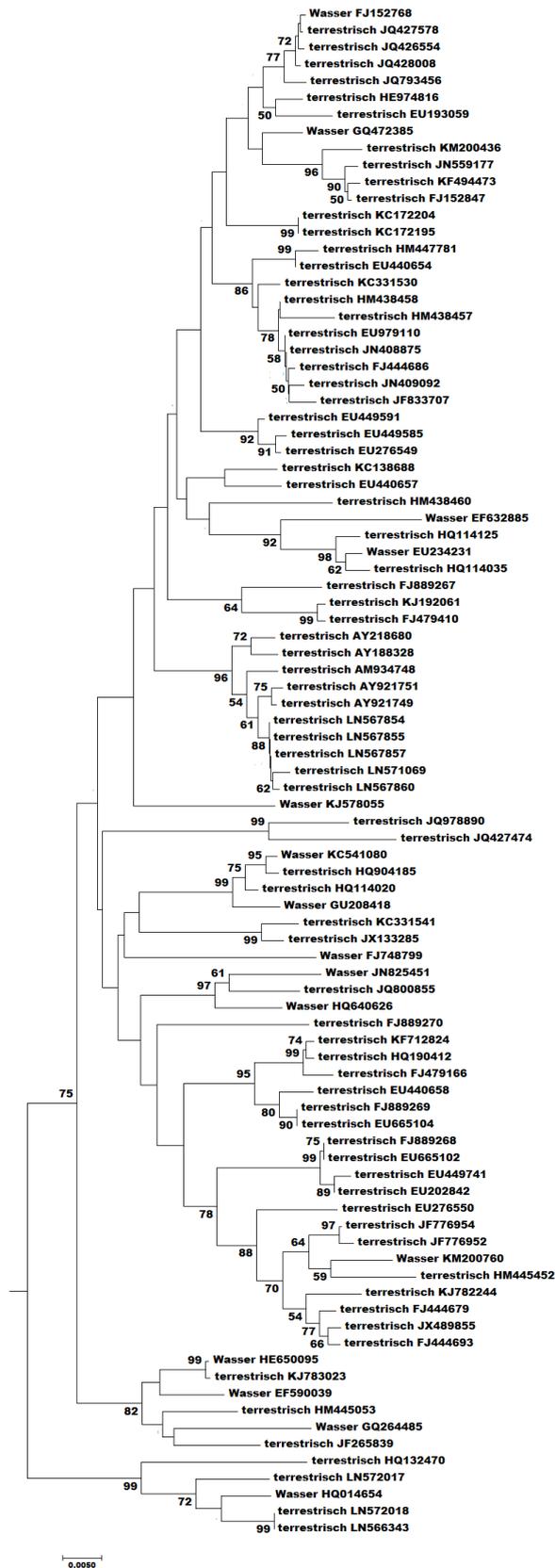


Abbildung 10-5: Aufspaltung der „Klade 5“ im NJ-Stammbaum. Ausschnitt aus dem Phylogenetischer Stammbaum des Kandidatenphylums 'Tectomicrobia' (Abbildung 5-18, Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, bootstrapping: 1000x^{190,198,211,212}; Base 225-1130, *E. coli*; enthält die variablen Regionen V3-V6). Angegeben sind bootstrap-Werte >50%. Rechts angegeben sind die *pairwise identity* (PI)-Werte einiger Cluster innerhalb der Kandidatenklasse.

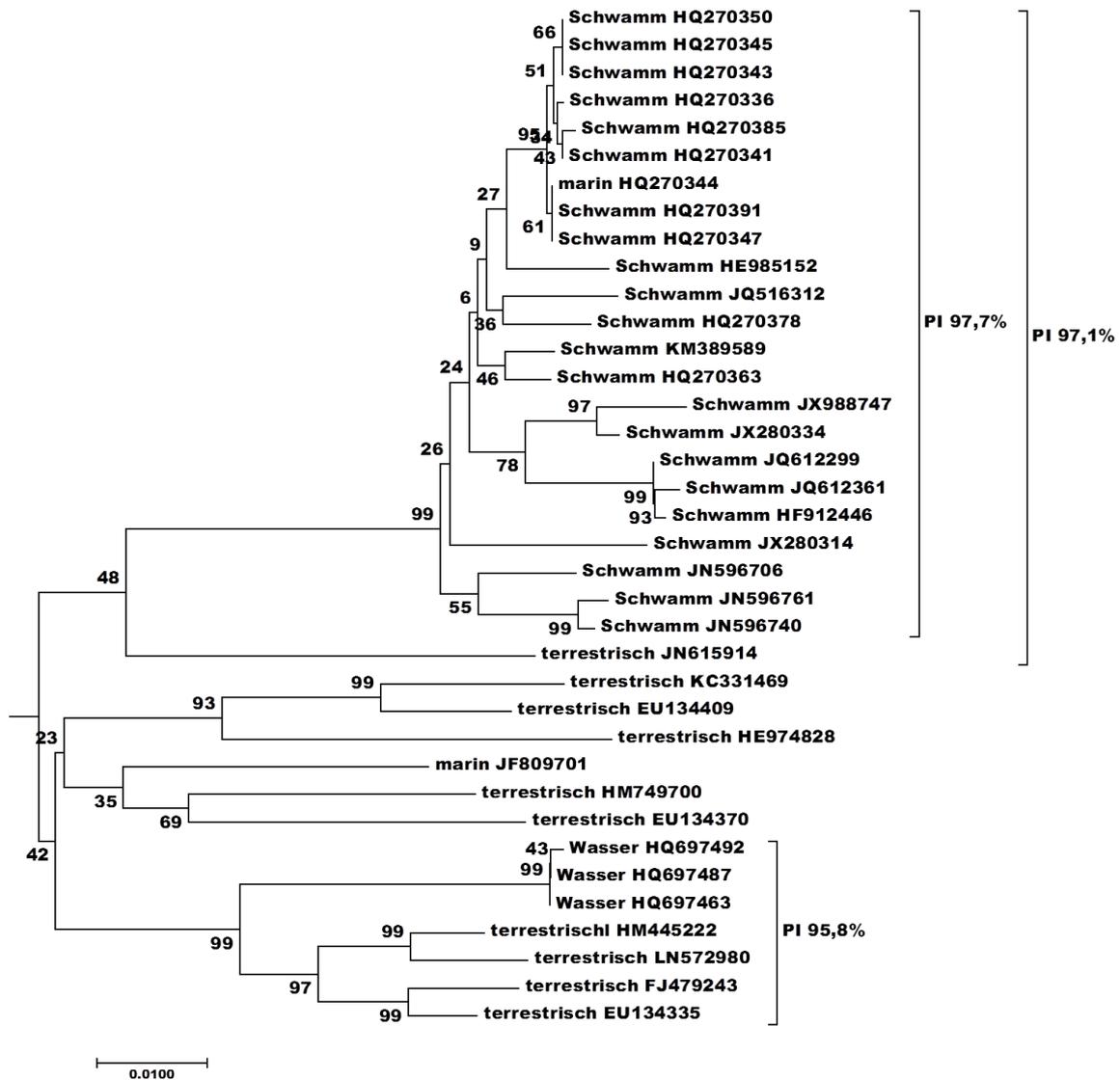


Abbildung 10-6: Aufspaltung der „Klade 6“ im NJ-Stammbaum. Ausschnitt aus dem Phylogenetischer Stammbaum des Kandidatenphylums 'Tectomicrobia' (Abbildung 5-18, Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, bootstrapping: 1000x^{190,198,211,212}, Base 225-1130, *E. coli*; enthält die variablen Regionen V3-V6). Angegeben sind bootstrap-Werte >50%. Rechts angegeben sind die *pairwise identity* (PI)-Werte einiger Cluster innerhalb der Kandidatenfamilie.

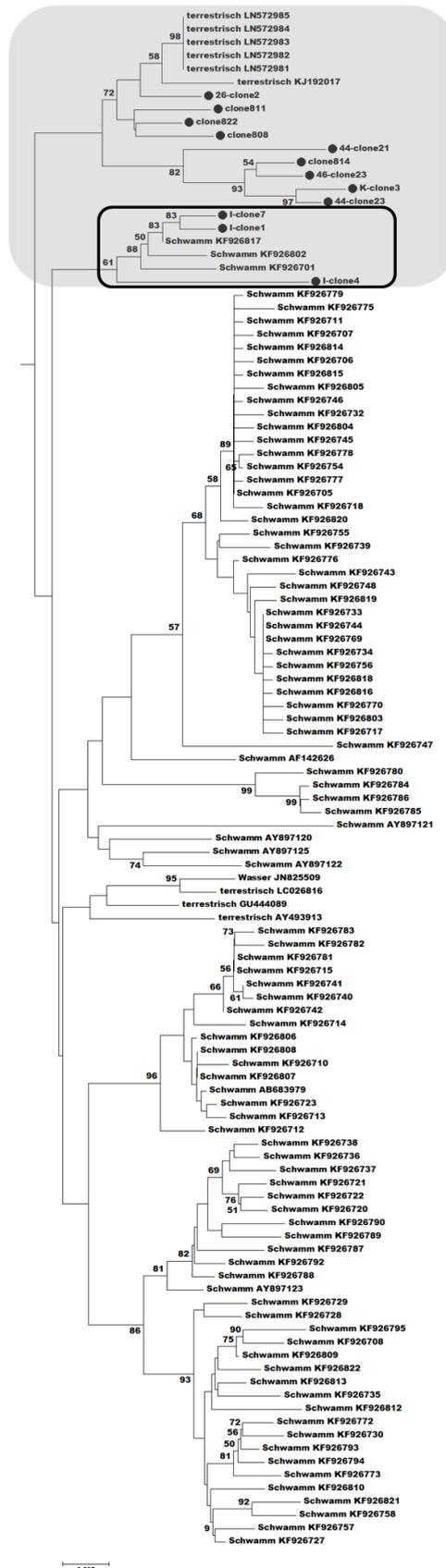


Abbildung 10-8: Aufspaltung des Kandidatengenus 'Ca. Entotheonella' („Klade 8“ im NJ-Stammbaum). Ausschnitt aus dem Phylogenetischer Stammbaum des Kandidatenphylums 'Tectomicrobia' (Abbildung 5-18, Software: MEGA 7, Methode: Neighbour-Joining, Model: Maximum Composite Likelihood, bootstrapping: 1000x^{190,198,211,212}, Base 225-1130, *E. coli*; enthält die variablen Regionen V3-V6). Angegeben sind bootstrap-Werte >50%. Rechts angegeben sind die *pairwise identity* (PI)-Werte einiger Cluster innerhalb des Kandidatengenus. Grau hinterlegt die Gruppe vonSequezen terrestrischen Ursprungs. Schwarz umrandet die

Anhang

Sequenzen der Probe I, die gemeinsam in ein Cluster mit 'E. factor' und zwei weiteren Sequenzen aus den Schwämmen *Aplysina aerophoba* und *Stylissa carteri* fallen. Auch in diesen Schwämmen konnten bereits Naturstoffe nachgewiesen werden.^{218,219} Die enge Phylogenese von potenten Naturstoffproduzenten aus Schwämmen wie 'E. factor' und 16S rRNA-Gensequenzen aus Spezies terrestrischen Ursprungs, lässt mutmaßen, dass diese ebenfalls eine Quelle bioaktiver Naturstoffe sind.



Abbildung 10-9: Sequenzalignment A-Domäne PppE.

Anhang



Abbildung 10-10: Sequenzalignment A-Domäne PppB.

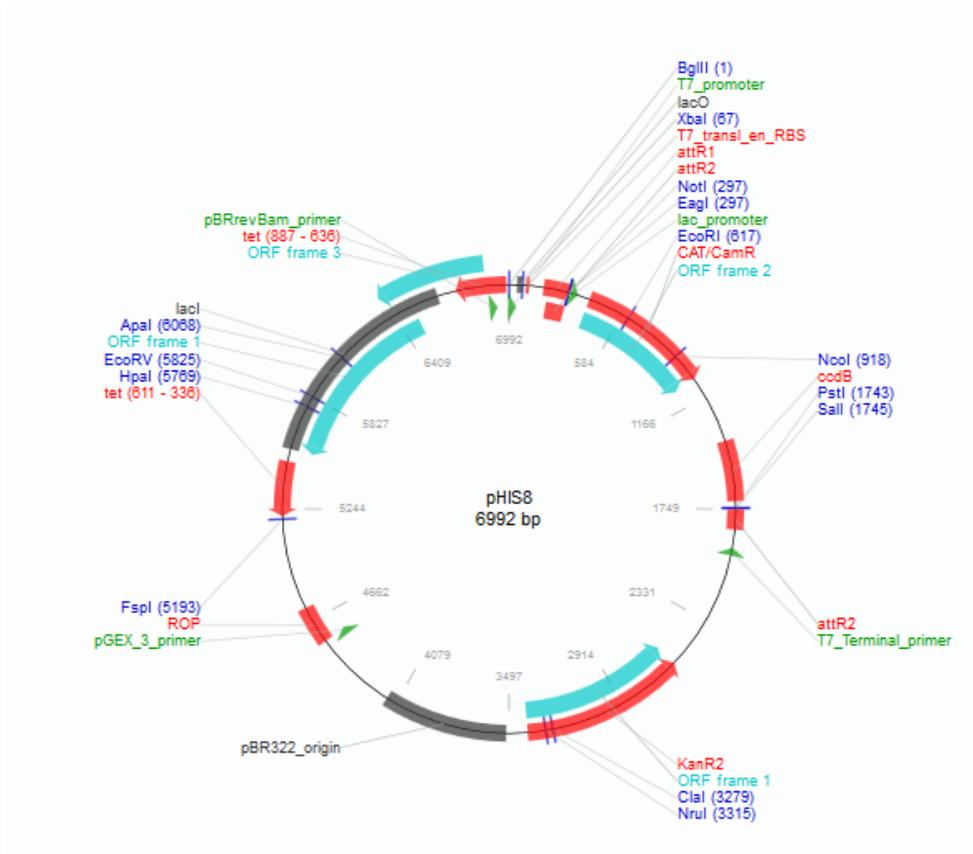


Abbildung 10-11: pHIS8-Vektor.²³²

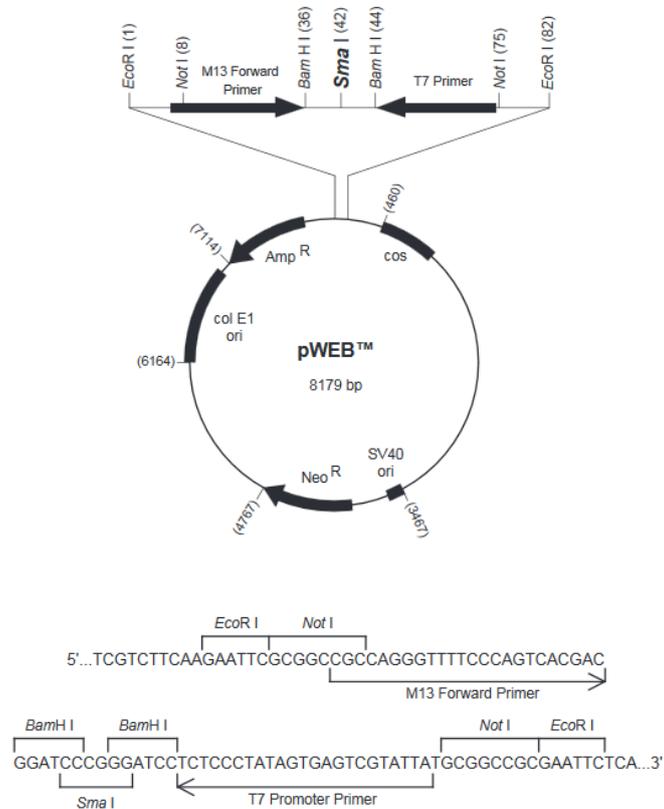


Abbildung 10-12: pWEB-Vektor.

Tabelle 10-2: Ergebnis der semiquantitativen Untersuchung der aufgereinigten Bodenproben.

Probe	Klon	Pairwise Identity zu 'E. factor'	Ergebnis BLASTN , (% pairwise identity)
5	1	75,1	uncultured Gemmatimonadetes bacterium
	2	83,9	uncultured Acidobacteria bacterium
	3	82,2	uncultured bacterium
	4		beta proteobacterium, Hydrogenopha sp.
	5	78,9	beta proteobacterium, uncultured bacterium
	6	81,1	Gammaproteobacteria, Xanthomonas
	7	80	Alphaproteobacteria, Mesorhizobium
	8	78,7	Alphaproteobacteria
	10	79,4	Actinobacteria, Nacardioides sp.
	11	79,9	Alphaproteobacteria, Bradyrhizobium
	12	83,5	uncultured bacterium
	13	81,1	Gammaproteobacteria, Pseudomonas brassicacearum (99,8%)
	14	78,3	Actinobacteria, Rhodococcus opacus (99,6%)
	15	81,7	Actinobacteria, Cellulomonas terrae strain (99,7%)
	16	79	uncultured bacterium (99,7%)
	17	78,6	uncultured bacterium (99,6%)
	18	76,3	beta proteobacterium, uncultured bacterium (97,4%)
	19	80,3	uncultured bacterium (95,6%), Acidobacteria
	20	78,3	uncultured bacterium (98,3%), Acidobacteria
	21	78,3	uncultured bacterium
	26	1	77,3
2		80	uncultured bacterium/ marine bacterioplankton (94,2%/92,5%)
3		78	uncultured bacterium (95,1%)
4		74,4	uncultured bacterium (97,9%)
5		76,6	uncultured bacterium (97,5%)
6		81,9	Alpha proteobacteria, Rhizobium sp. (99,3%)
8		77,6	uncultured bacterium (98,4%)
9		77,2	uncultured bacterium (97%)
10		77,5	Actinobacteria, Mycobacterium sp (97,3%)
11		77,3	uncultured bacterium (97, 9%)
12			uncultured bacterium (95,1%)
14		81,5	uncultured bacterium (98,7%)
15		79	Beta Proteobacteria, Variovorax sp. (99,6%)
16		78,8	Gammaproteobacteria, Xanthomonas (96,5%)
17		79,1	uncultured bacterium (98%)
18		75,1	uncultured bacterium (97,7%)
19		83,4	Alpha proteobacteriauncultured bacterium (98,8%)
20		74,8	uncultured baterium (99,4%)
21		81,8	Delta Proteobacteria (98,8%)
22		79,6	Alpha Proteobacteria, uncultured bacterium (99,2%)

Anhang

I	1	75,5	beta Proteobacteria, uncultured bacterium (99%)
	4	81,4	uncultured bacterium (99,1%)
	5	76,9	Gamma proteobacteria, uncultured bacterium (93,3%)
	6	80,6	Alpha Proteobacteria, uncultured bacterium 899,4%)
	7	81,7	Actinobacteria, Rhodococcus sp. (99,8%)
	8	76,5	uncultured bacterium (96,8%)
	11	78,7	
	12	79,1	
	13	67,1	uncultured bacterium, Sphingobacteria, Chitinophagaceae
	14	75,5	uncultured bacterium, actinobacterium
	15	72,8	uncultured Gemmatimonadetes sp.
	16	72,5	uncultured bacterium, Paracoccus sp.
	17	75,5	Kineosporia rhizophila strain
	18	71	uncultured bacterium, Acidovorax sp.
	19	70,3	Pseudomonas putida strain (99,3%)
	31	70,3	uncultured bacterium, Carnobacterium sp.
	32	69,8	uncultured cyanobacterium
	33	72,3	uncultured delta proteobacterium
	34	70,1	Polaromonas sp. (99,6%)
	36	64,7	uncultured bacterium, Shingobacteria
	37	68,2	Exiguobacterium sp. (99,9%)
	38	67,5	uncultured bacterium, beta proteobacterium
	40	72,1	uncultured bacterium, Sphingomonas

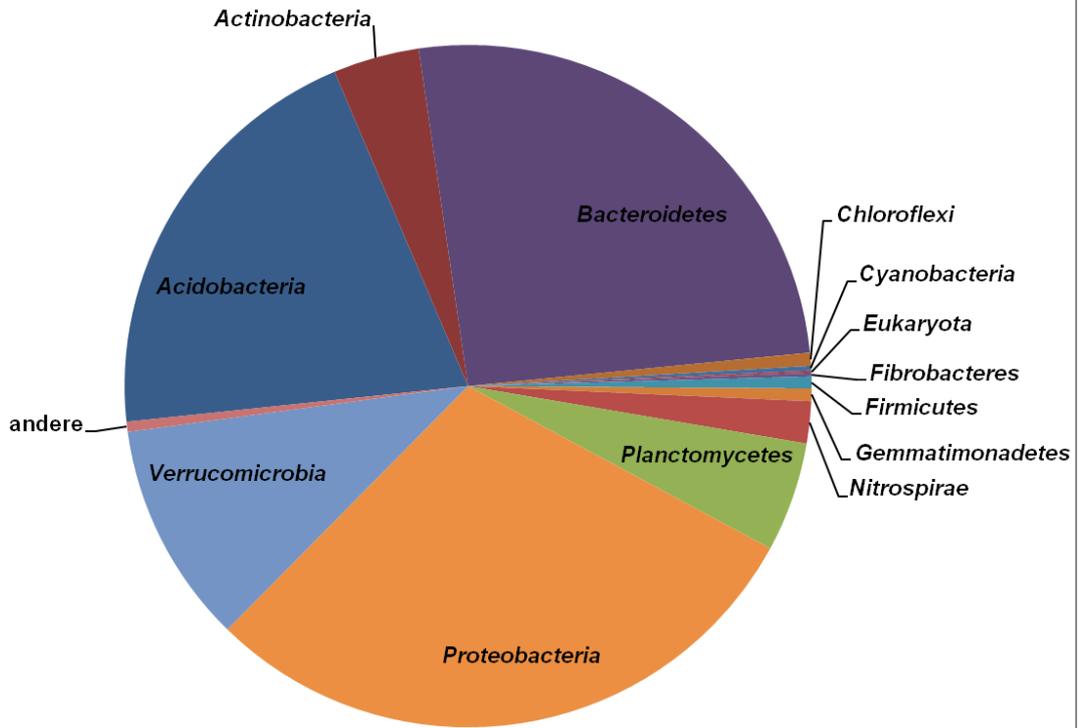
Tabelle 10-3: Mikrobiom der Bodenproben: Phyla.

	Bodenprobe 5		Bodenprobe 8	
	absolut	prozentual	absolut	Prozentual
<i>Acidobacteria</i>	18990	20,34	25615	34
<i>Actinobacteria</i>	3787	4,06	3707	4,92
<i>Aquificae</i>	11	0,01	0	0
<i>Armatimonadetes</i>	26	0,03	90	0,12
<i>Arthropoda</i>	5	0,01	9	0,01
<i>Bacillariophyta</i>	6	0,01	31	0,04
<i>Bacteroidetes</i>	24068	25,78	11850	15,73
' <i>Candidatus Saccharibacteria</i> '	21	0,02	32	0,04
<i>Chlamydiae</i>	44	2,03	8	0,01
<i>Chloroflexi</i>	578	0,62	1141	1,51
<i>Chlorophyta</i>	7	0,01	12	0,01
<i>Cyanobacteria</i>	205	0,22	271	0,36
<i>Deinococcus thermus</i>	27	0,03	17	0,02
<i>Elusimicrobia</i>	31	0,03	8	0,01
<i>Eukaryota</i>	87	0,10	868	1,15
<i>Fibrobacteres</i>	168	0,18	31	0,04
<i>Firmicutes</i>	524	0,56	676	0,9
<i>Fusobacteria</i>	0	0,00	0	0
<i>Gemmatimonadetes</i>	515	0,60	375	0,5

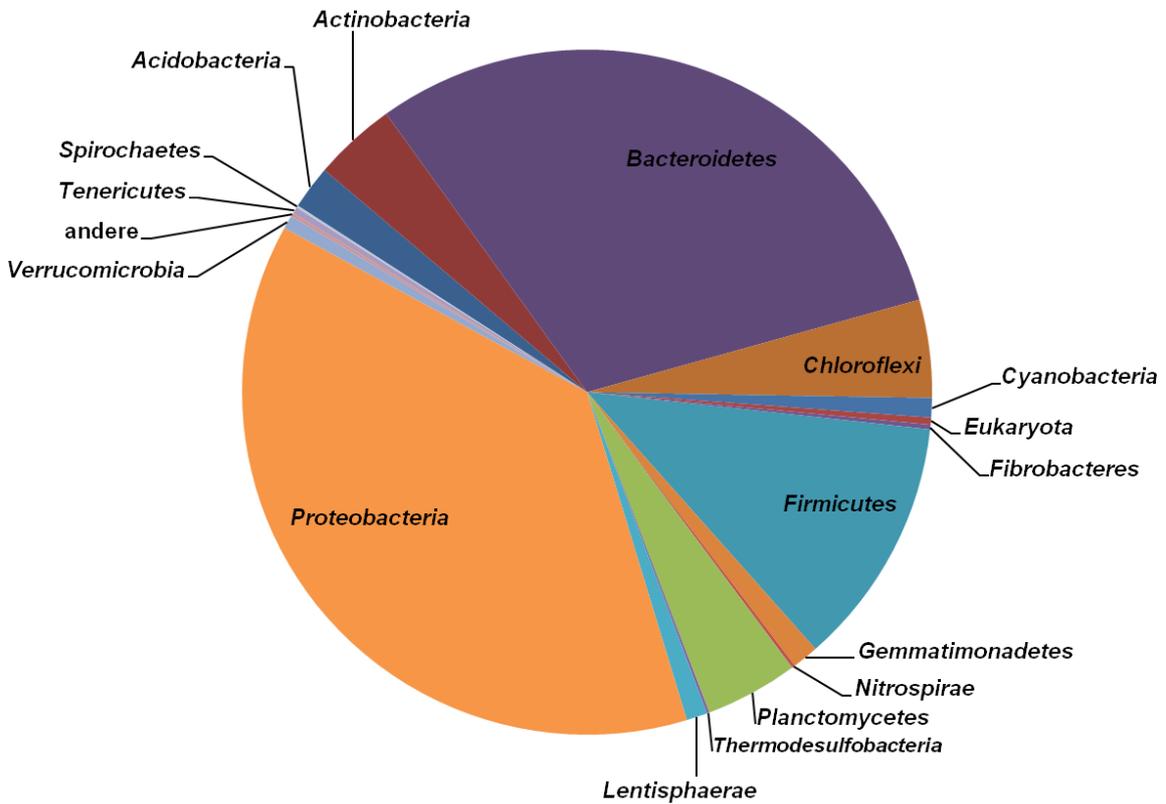
Anhang

<i>Ignavibacteriae</i>	17	0,02	10	0,01
<i>Lentisphaerae</i>	79	0,08	39	0,05
<i>Nitrospirae</i>	1862	1,99	2051	2,72
<i>Planctomycetes</i>	4861	5,20	5257	6,98
<i>Proteobacteria</i>	27559	29,52	15747	20,90
<i>Spirochaetes</i>	24	0,03	5	0,01
<i>Streptophyta</i>	0	0,00	5	0,01
<i>Tenericutes</i>	6	0,01	4	0,01
<i>Thaumarchaeota</i>	2	0,00	0	0
<i>Thermodesulfobacteria</i>	47	0,06	84	0,11
<i>Verrucomicrobia</i>	9815	10,51	7393	9,81
	Bodenprobe 26		Bodenprobe I	
	absolut	prozentual	absolut	prozentual
<i>Acidobacteria</i>	4907	5,62	40252	31,5
<i>Actinobacteria</i>	13583	15,57	8495	6,65
<i>Aquificae</i>	0	0	13	0,01
<i>Armatimonadetes</i>	4	0	230	0,18
<i>Arthropoda</i>	21	0,02	2	0
<i>Bacillariophyta</i>	2	0	7	0,01
<i>Bacteroidetes</i>	37184	42,61	27656	21,65
' <i>Candidatus Saccharibacteria</i> '	369	0,42	105	0,08
<i>Chlamydiae</i>	0	0	0	0
<i>Chloroflexi</i>	1986	2,28	5715	4,47
<i>Chlorophyta</i>	25	0,03	85	0,07
<i>Cyanobacteria</i>	133	0,15	180	0,14
<i>Deinococcus thermus</i>	33	0,04	332	0,26
<i>Elusimicrobia</i>	0	0	7	0,01
<i>Eukaryota</i>	36	0,04	47	0,04
<i>Fibrobacteres</i>	5	0,01	167	0,13
<i>Firmicutes</i>	769	0,88	4879	3,82
<i>Fusobacteria</i>	0	0	3	0
<i>Gemmatimonadetes</i>	240	0,28	315	0,25
<i>Ignavibacteriae</i>	0	0	1	0
<i>Lentisphaerae</i>	38	0,04	1497	1,17
<i>Nitrospirae</i>	464	0,53	89	0,07
<i>Planctomycetes</i>	7232	8,3	1834	1,44
<i>Proteobacteria</i>	18011	20,64	31390	24,57
<i>Spirochaetes</i>	0	0	0	0
<i>Streptophyta</i>	2	0	13	0,01
<i>Tenericutes</i>	24	0,03	16	0,01
<i>Thaumarchaeota</i>	5	0	2	0
<i>Thermodesulfobacteria</i>	19	0,02	3	0
<i>Verrucomicrobia</i>	2164	2,48	4429	3,47

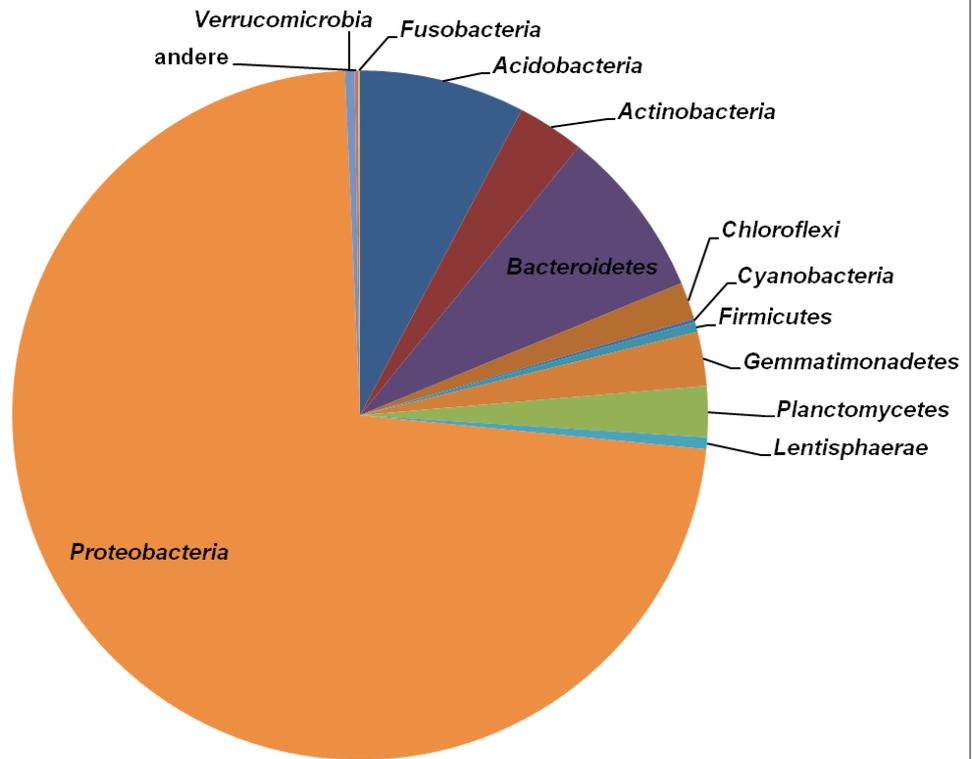
Bodenprobe 5



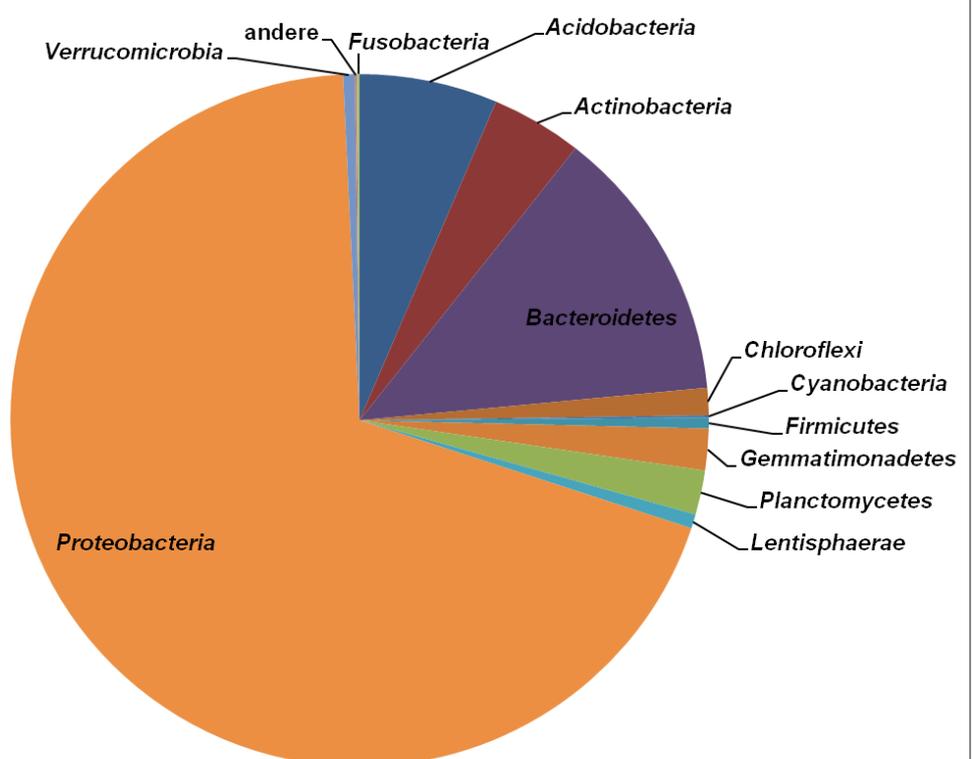
Bodenprobe 5, 100xg



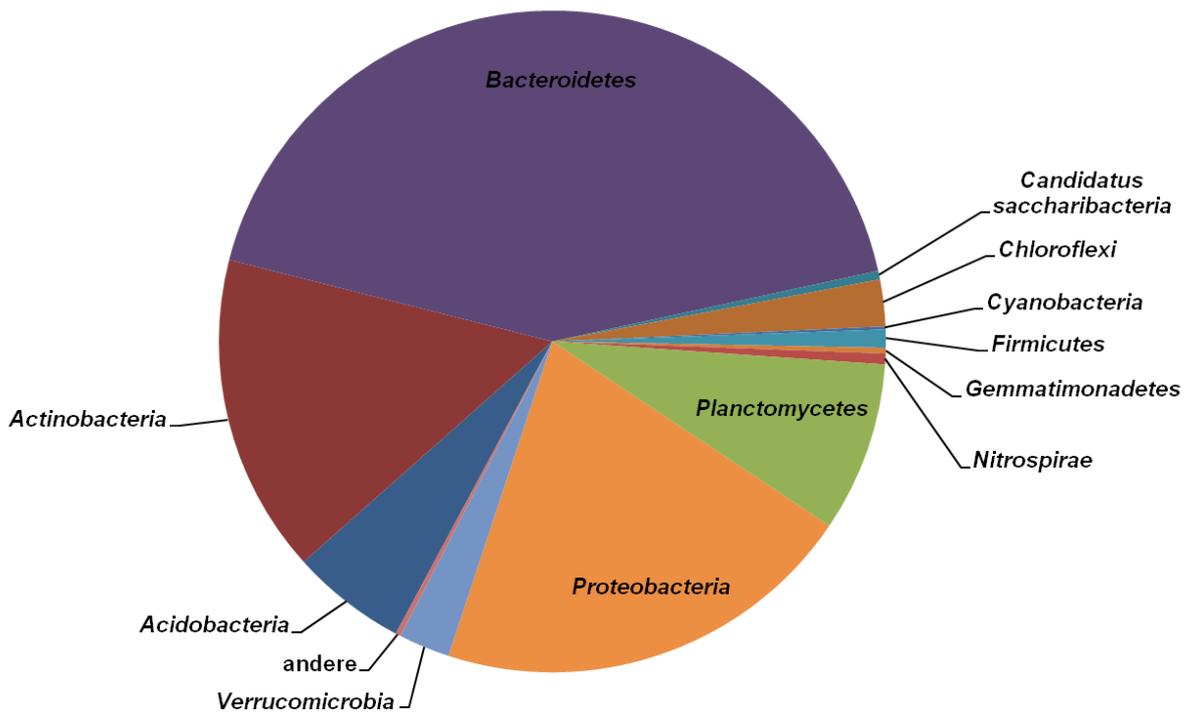
Bodenprobe 5, 1000xg



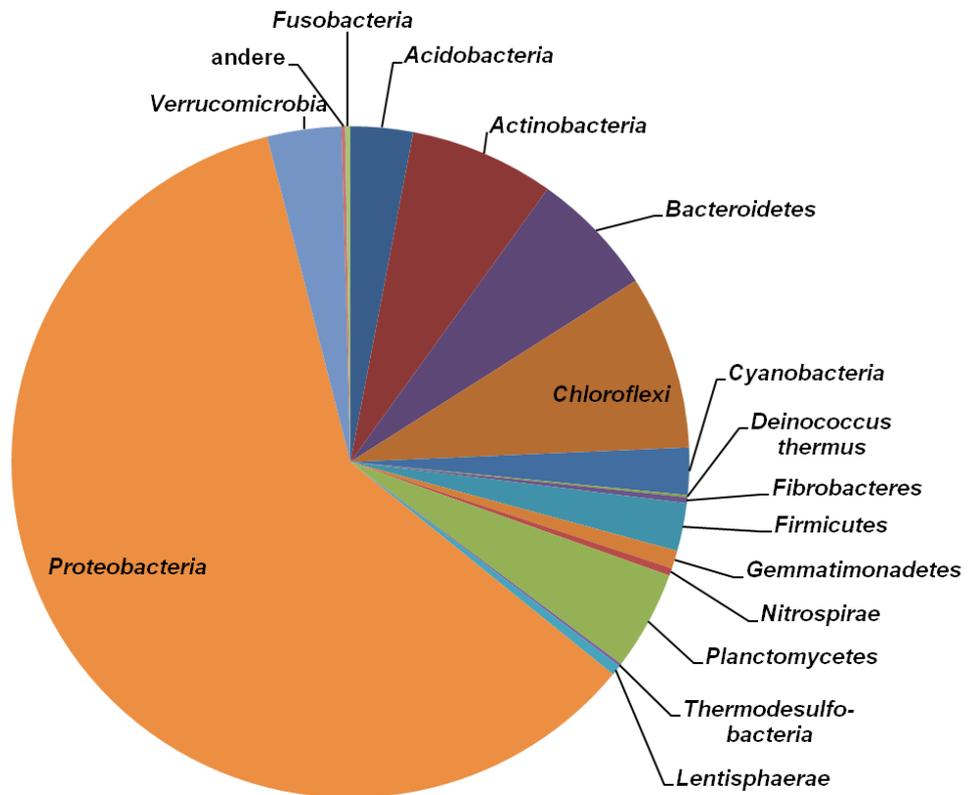
Bodenprobe 5, 4500xg



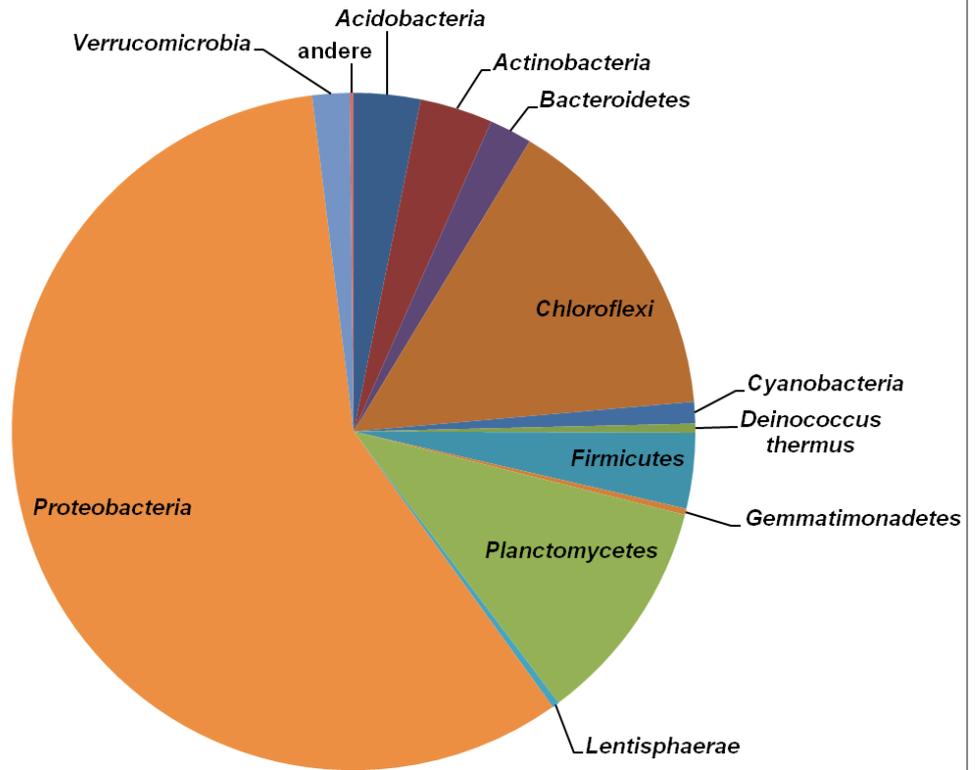
Bodenprobe 26



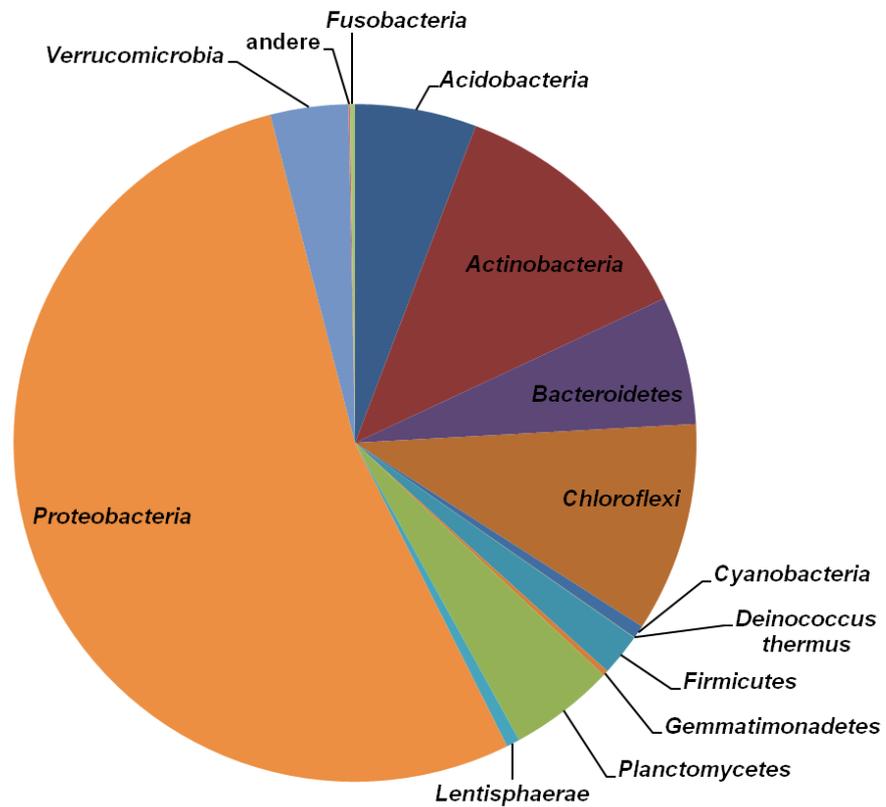
Bodenprobe 26, 100xg



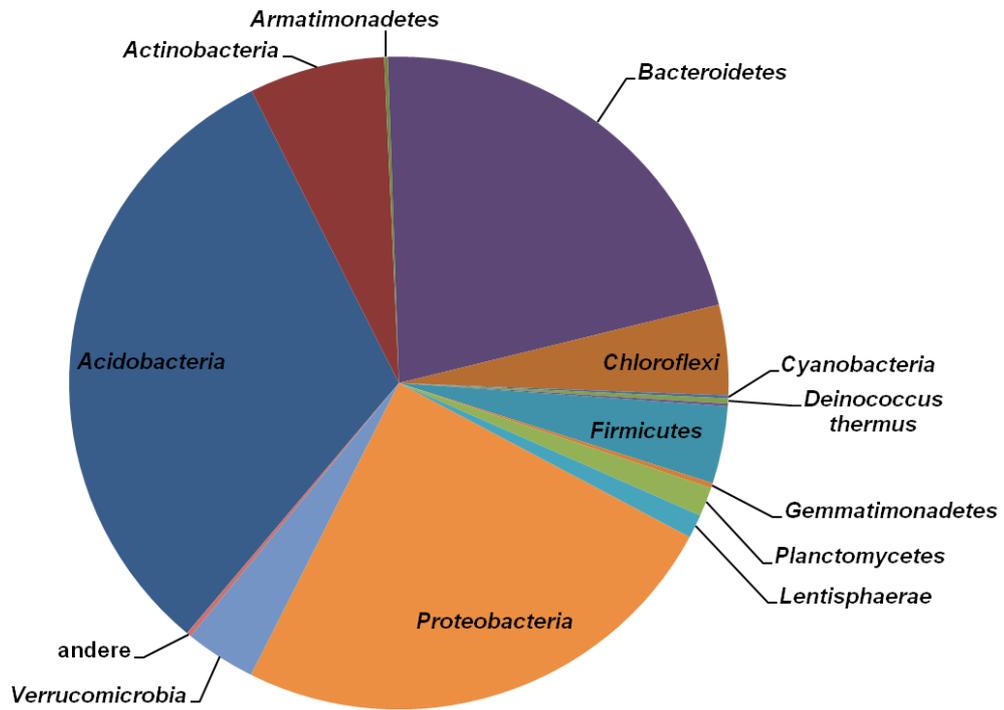
Bodenprobe 26, 1000xg



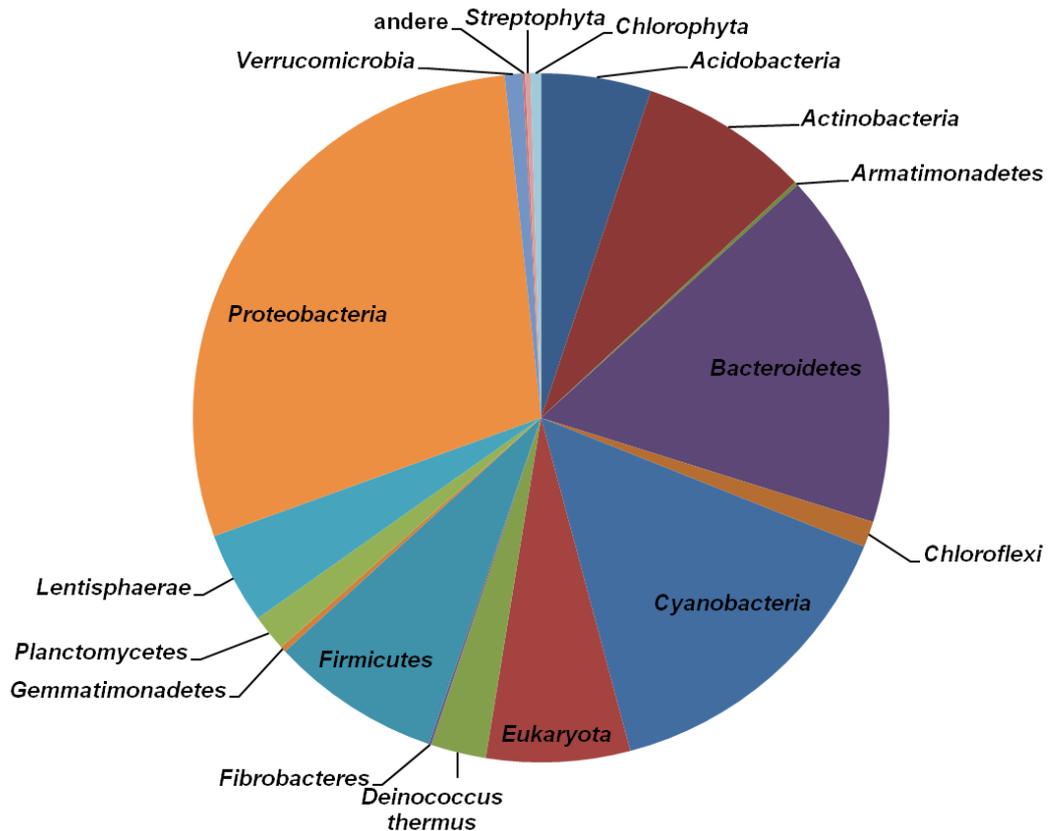
Bodenprobe 26, 4500xg



Bodenprobe I



Bodenprobe I, 100xg



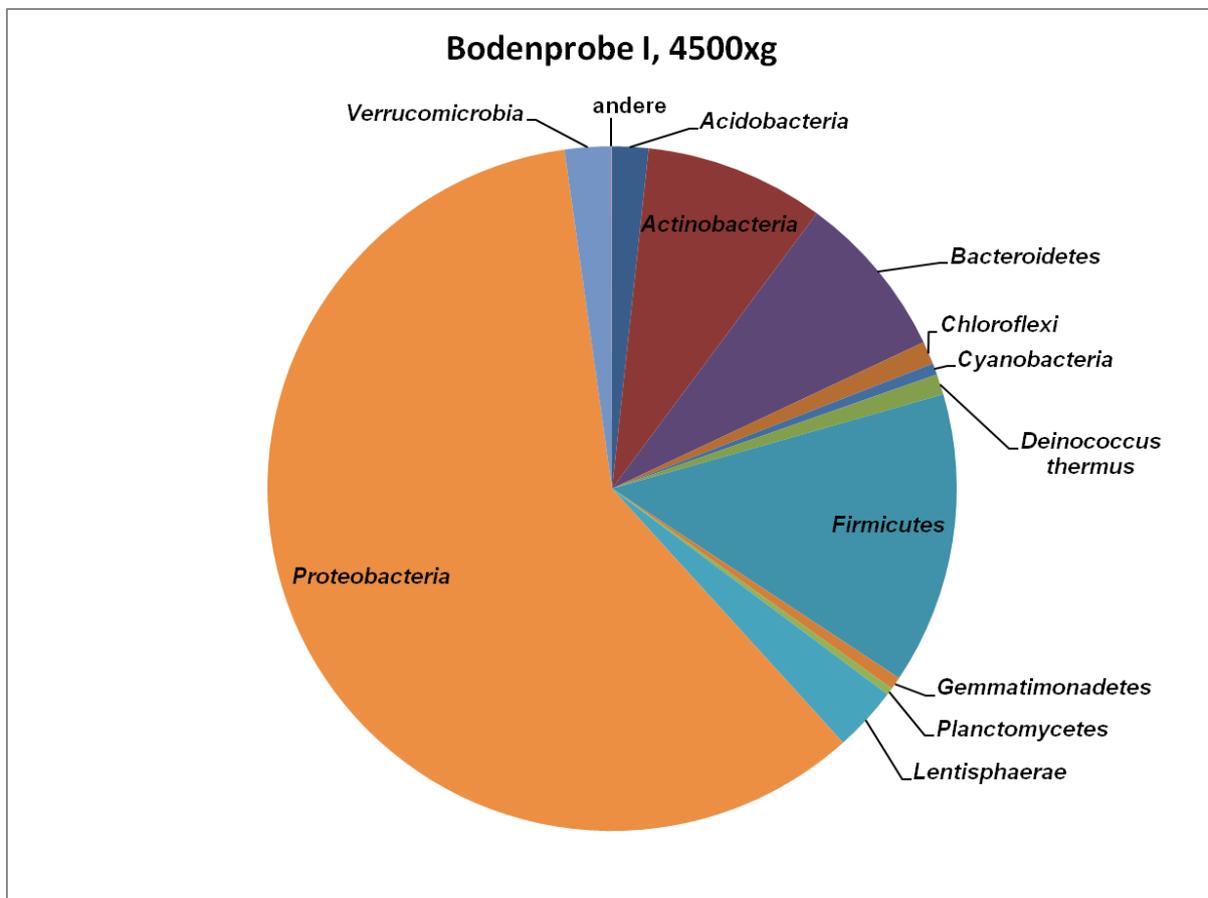
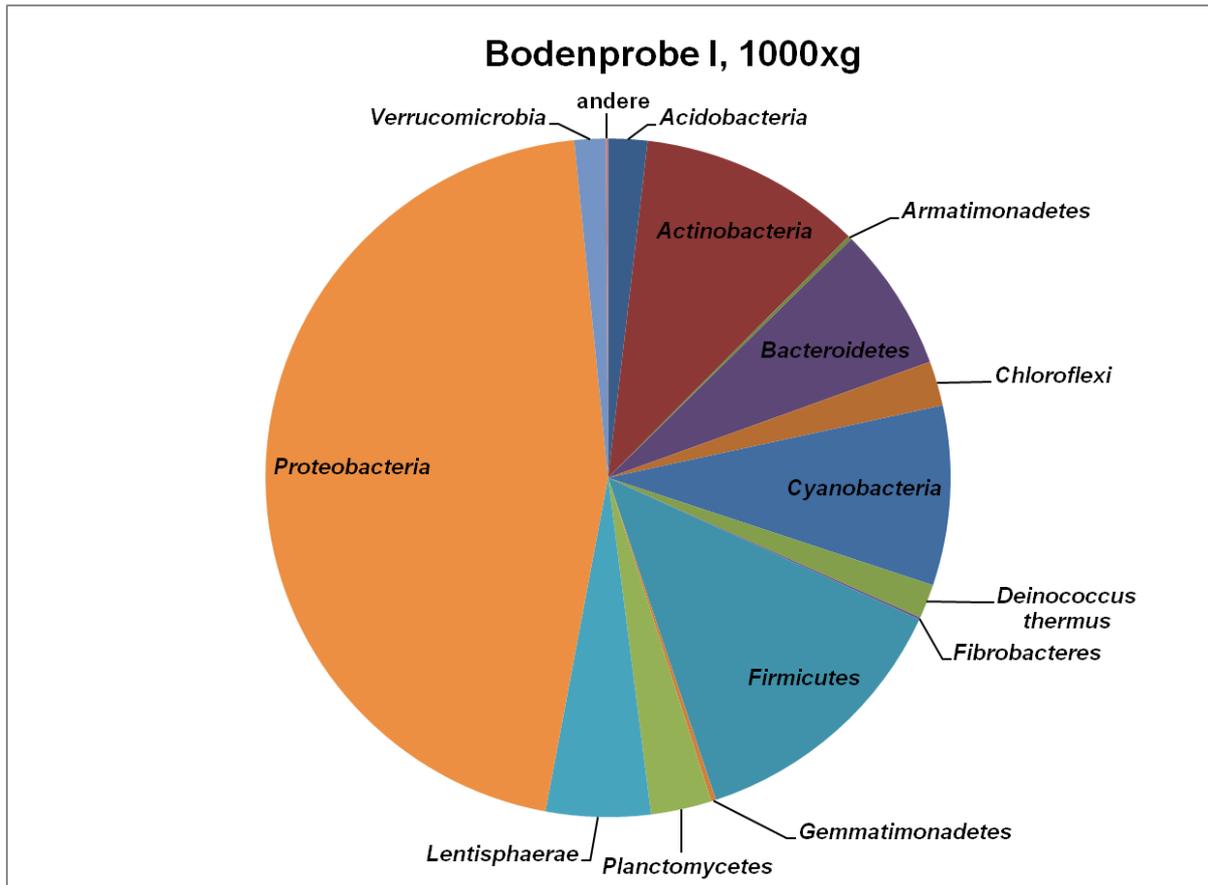
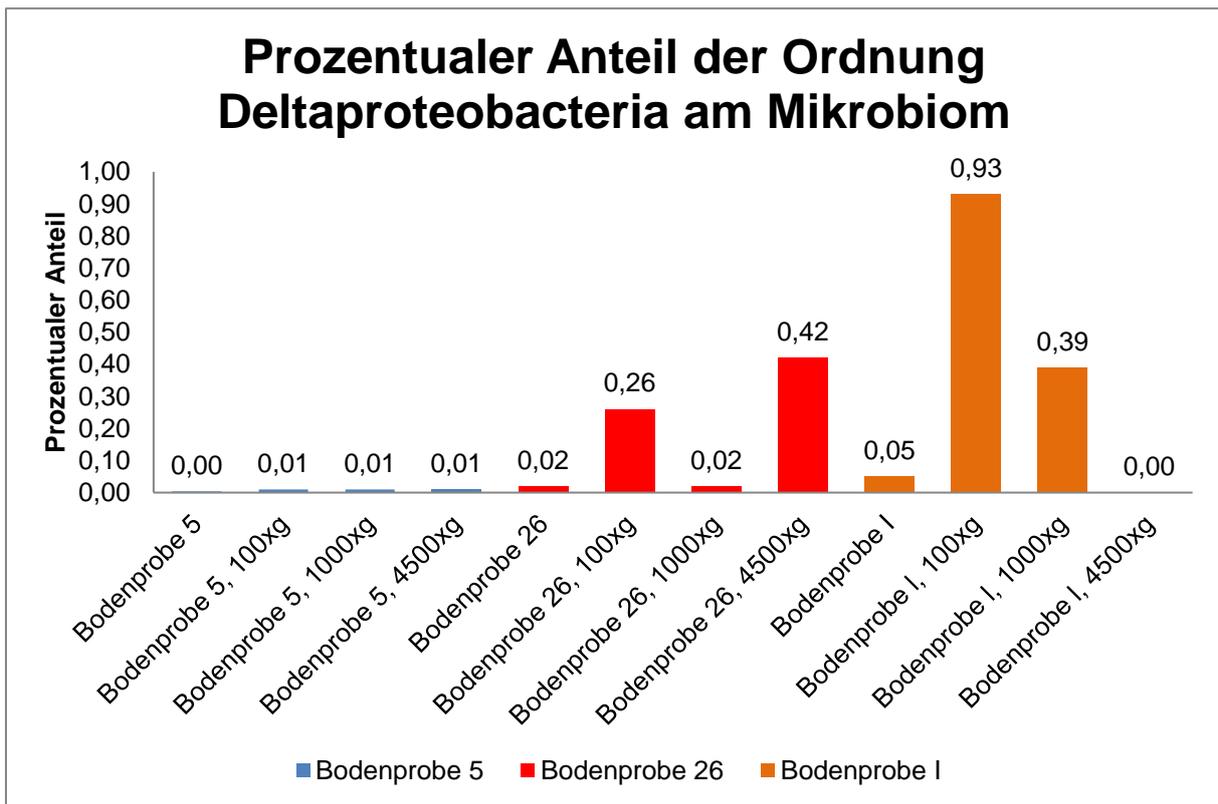
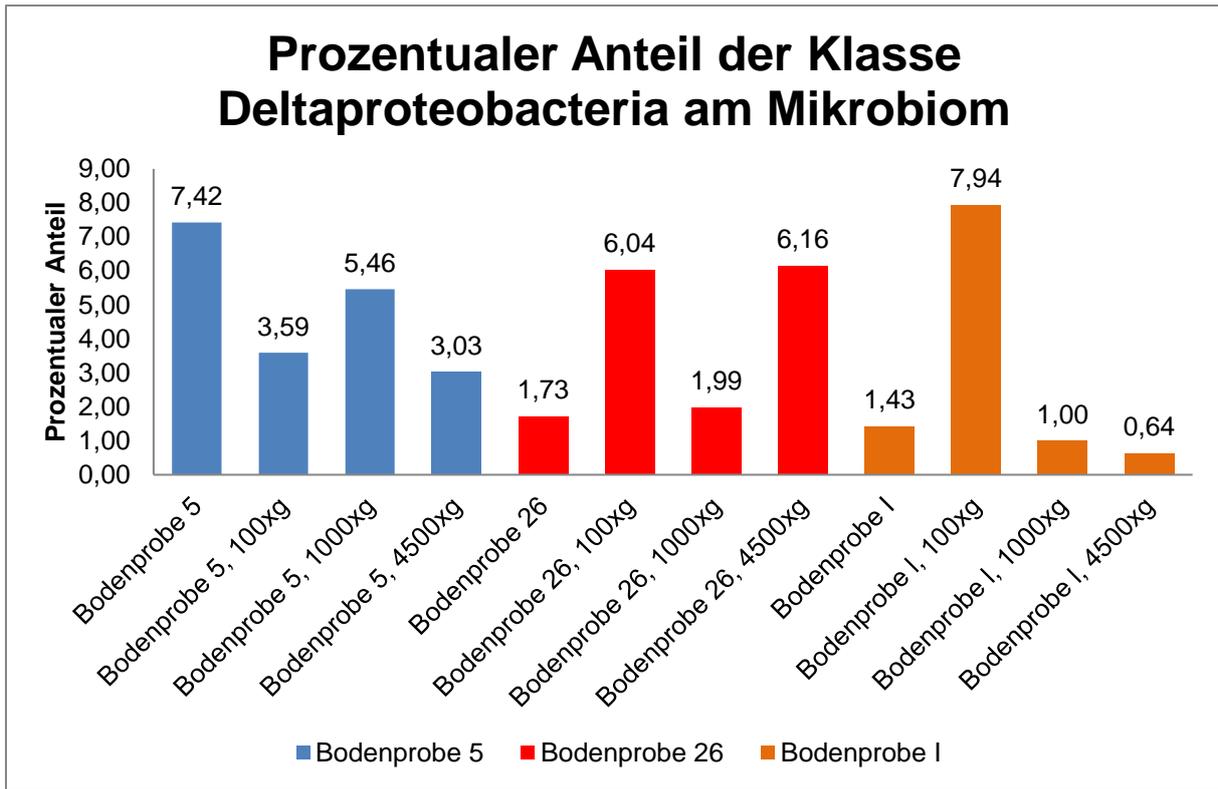


Abbildung 10-13: Anteile der Phyla am Mikrobiom vor und nach der Fraktionierung.



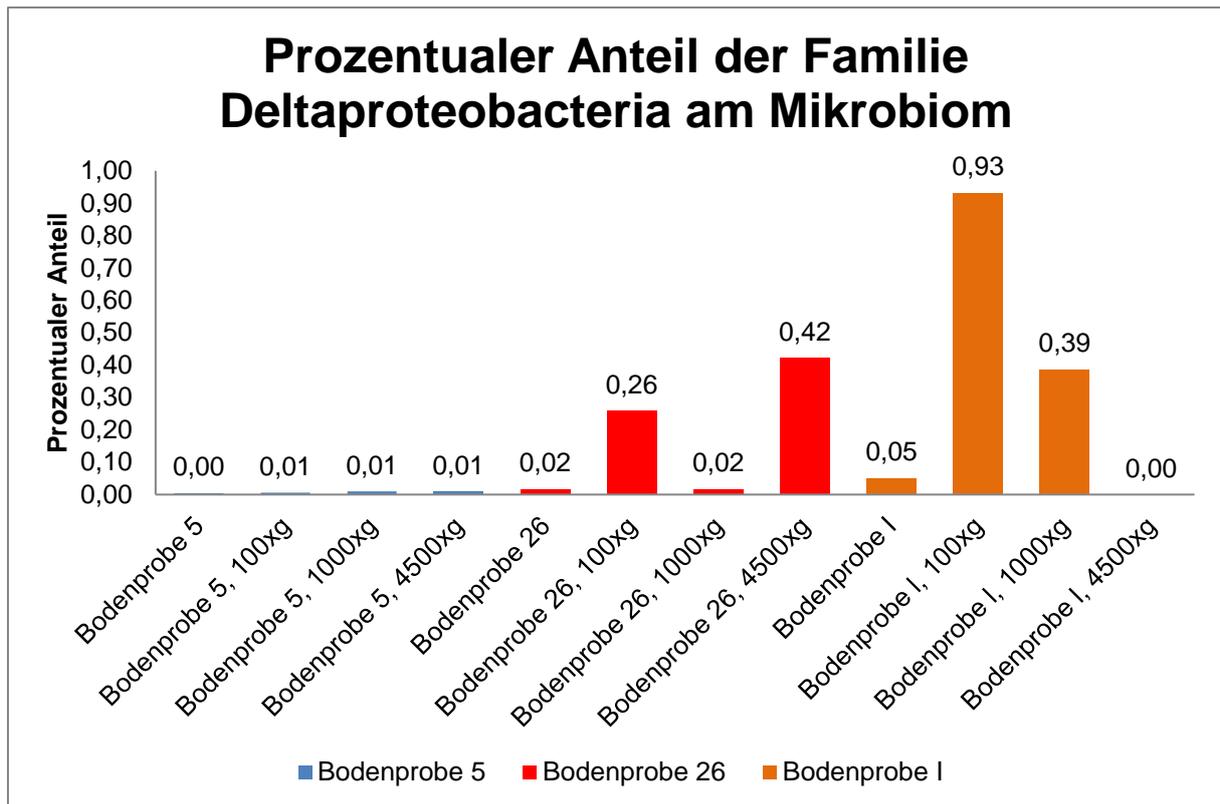


Abbildung 10-14: Prozentualer Anteil der Klasse/Ordnung/Familie Deltaproteobacteria am Mikrobiom der Proben vor und nach der Fraktionierung.

Tabelle 10-4: Anzahl der Sequenzen und prozentualer Anteil der zehn häufigsten Spezies in den vier untersuchten Bodenproben.

Bodenprobe 5			Bodenprobe 8		
Spezies	absolut	prozentual	Spezies	absolut	prozentual
<i>Acidobacterium</i> spp.	16236	20,52	<i>Acidobacterium</i> spp.	23876	21,36
<i>Flexibacter</i> spp.	4649	5,88	<i>Verrucomicrobium</i> spp.	4470	4,00
<i>Sphingobacterium</i> spp.	3889	4,91	<i>Sphingobacterium</i> spp.	2704	2,42
<i>Opitutus</i> spp.	3750	4,74	<i>Pirellula</i> spp.	2106	1,88
<i>Ohtaekwangia</i> spp.	3142	3,97	<i>Nitrospira</i> spp.	2049	1,83
<i>Steroidobacter</i> spp.	2494	3,15	<i>Flexibacter</i> spp.	1634	1,46
<i>Rhizobium</i> spp.	2226	2,81	<i>Rhodoplanes</i> spp.	1076	0,96
<i>Pirellula</i> spp.	1983	2,51	<i>Ohtaekwangia</i> spp.	1058	0,95
<i>Nitrospira</i> spp.	1842	2,33	<i>Chryseolinea</i> sphingobacteria bacterium	930	0,83
<i>Niastella</i> spp.	1823	2,30	<i>Anaeromyxobacter</i> spp.	913	0,82

Bodenprobe 26			Bodenprobe I		
Spezies	absolut	prozentual	Spezies	absolut	prozentual
<i>Salegentibacter</i> spp.	10332	27,93	<i>Blastocatella</i> fastidiosa	21342	30,11
<i>Acidobacterium</i> spp.	4527	12,24	<i>Acidobacterium</i> spp.	14495	20,45
<i>Flavobacterium</i> spp.	4196	11,34	<i>Flavisolibacter</i> spp.	7591	10,71
<i>Pirellula</i> spp.	3873	10,47	<i>Flavisolibacter</i> sp.	6877	9,70
<i>Salegentibacter</i> sp.	3068	8,29	<i>Aridibacter</i> acidobacteria bacterium	4359	6,15
<i>Aquiflexum</i> spp.	2422	6,55	<i>Kouleoethrix</i> spp.	4299	6,07
<i>Flavobacterium</i> sp.	2357	6,37	<i>Microvirga</i> spp.	3553	5,01
<i>Pseudomonas</i> stutzeri	2194	5,93	<i>Rubellimicrobium</i> spp.	3253	4,59
<i>Conexibacter</i> spp.	2144	5,80	<i>Carnobacterium</i> spp.	2639	3,72
<i>Ohtaekwangia</i> spp.	1881	5,08	<i>Chthoniobacter</i> flavus	2473	3,49

Tabelle 10-5: Ergebnis der Korngrößenbestimmung einiger untersuchter Bodenproben.

Bodenprobe	Bodenart	Kartieranleitung Bodenkunde
5	mittel schluffiger Sand	Su3
6	sandig-lehmiger Schluff	Us
7	schluffig-lehmiger Sand	Su3
9	sandiger Schluff	Us
18	stark schluffiger Sand	Su4
22	mittel schluffiger Sand	Su3-Su4
25	sandiger Schluff	Us
A	stark schluffiger Sand	Su4 (Su3)
B	stark schluffiger Sand	Su4
D	schwach schluffiger Sand	Su2 (Su3)
F	sandig-lehmiger Schluff	Us
G	stark schluffiger Sand	Su4 (Su3)
H	schwach schluffiger Sand	Su2 (Su3)
I	reiner Sand	Ss
K	Schluffsande	Su3 (Su4)



Abbildung 10-15: Beschaffenheit einiger untersuchter Bodenproben.

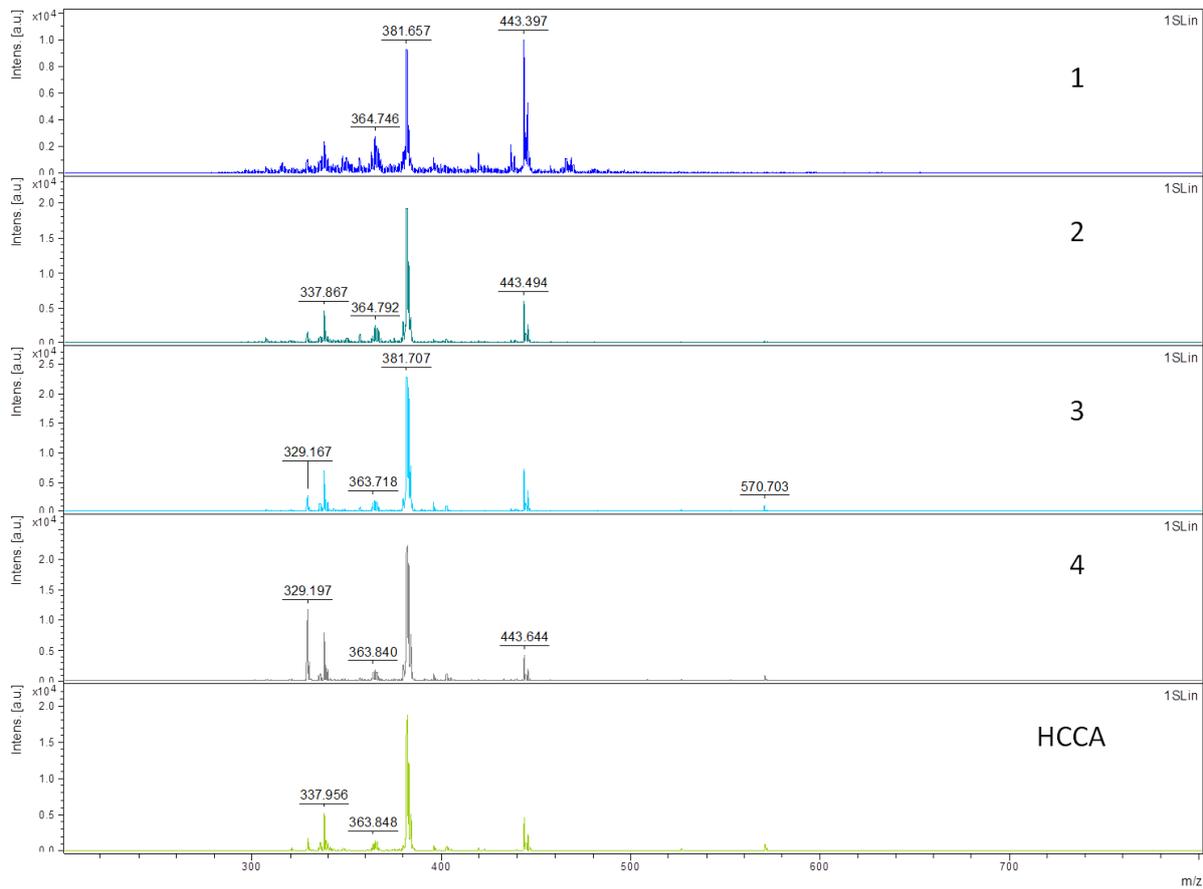


Abbildung 10-16: MALDI-TOF- Analyse nach Expression des gesamten Genclusters.

Nach heterologer Expression und Extraktion mit Ethylacetat wurden die in Methanol gelösten Extrakte zur Analyse in einem MALDI-TOF auf eine HCCA-Matrix aufgetragen. **1:** Überstand der Kultur mit Leervektor (pWEB), **2:** Überstand der Kultur mit Konstrukt aus Gencluster im pWEB-Cosmid, **3:** Zellpellet der Kultur mit Leervektor (pWEB), **4:** Zellpellet der Kultur mit Konstrukt aus Gencluster im pWEB-Cosmid.

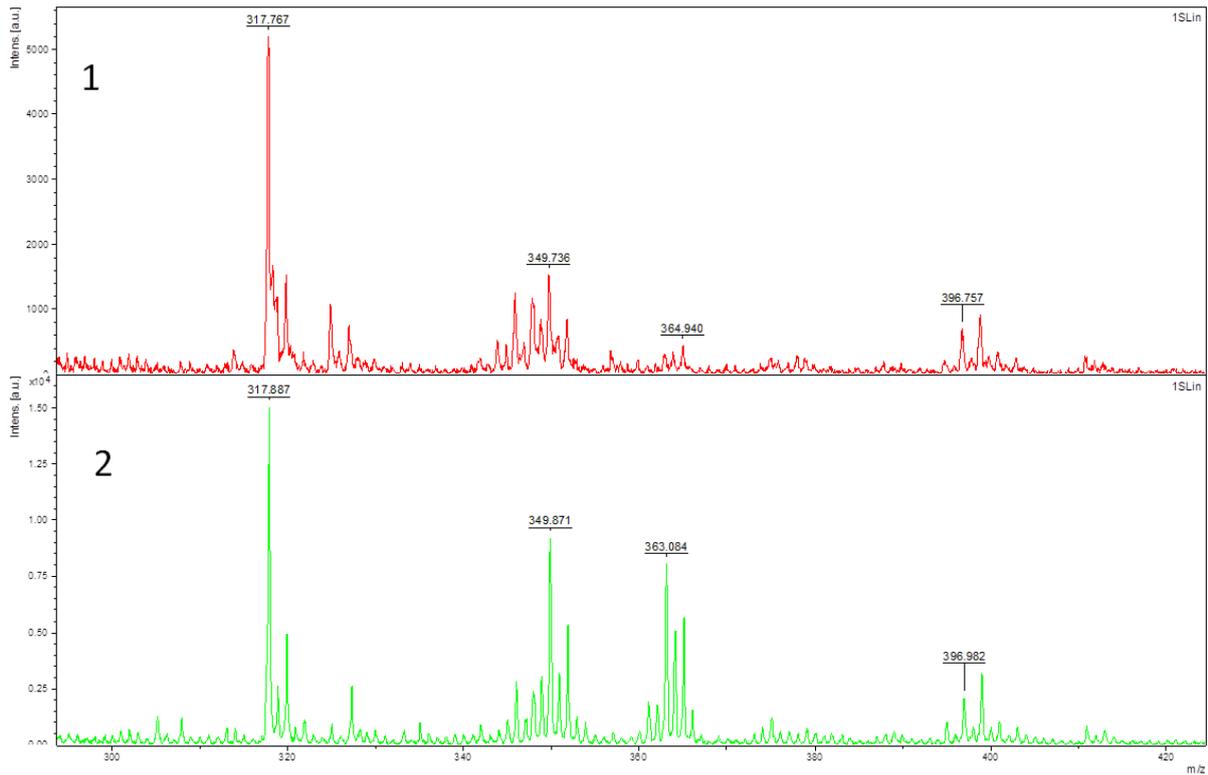


Abbildung 10-17: Wiederholung der MALDI-TOF Analyse.

Wiederholung der MALDI-TOF Analyse mit ATT (6-Aza-2-thiothymidin). Proben 1: Überstand der Kultur mit Leervektor (pWEB) und 2: Überstand der Kultur mit Konstrukt aus Gencluster im pWEB-Cosmid.

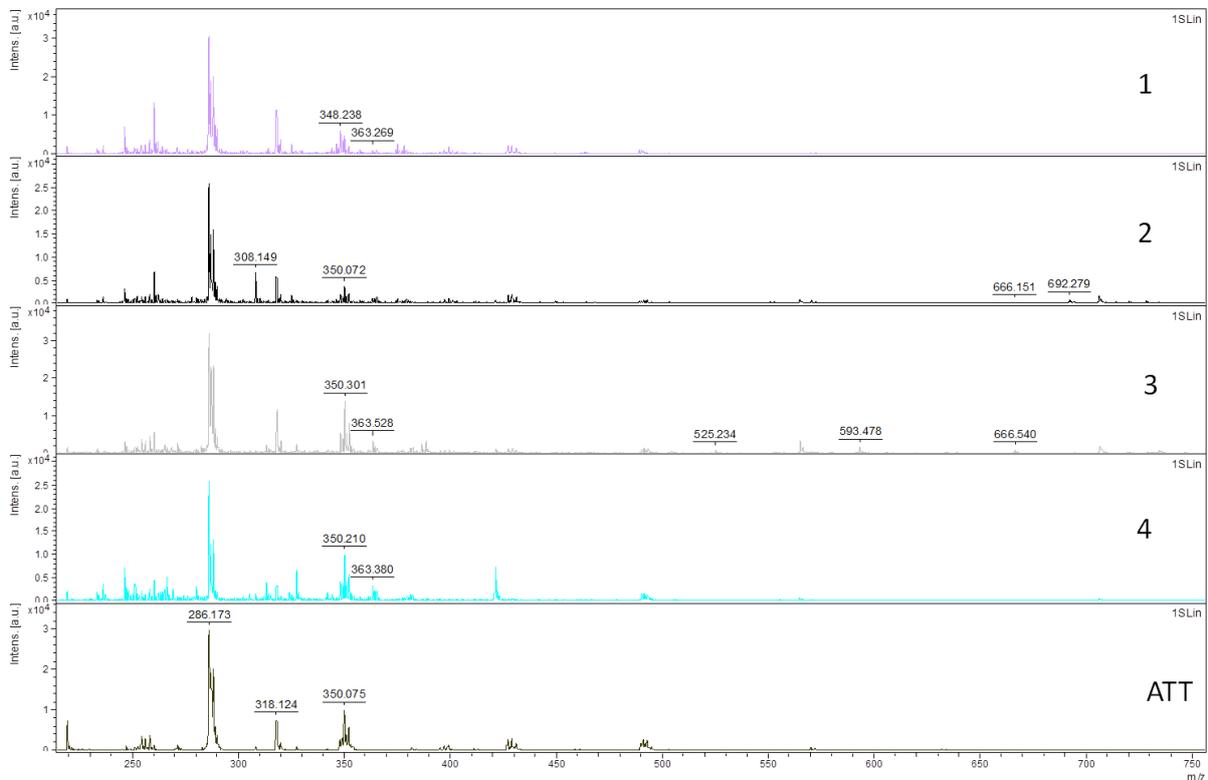


Abbildung 10-18: Wiederholung der MALDI-TOF Analyse nach Expression des gesamten Genclusters.

Wiederholung der MALDI-TOF Analyse mit ATT (6-Aza-2-thiothymidin) Proben 1: Überstand der Kultur mit Leervektor (pWEB) und 2: Überstand der Kultur mit Konstrukt aus Gencluster im pWEB-Cosmid; 3: Zellpellet der Kultur mit Leervektor (pWEB); 4: Zellpellet der Kultur mit Konstrukt aus Gencluster im pWEB-Cosmid.

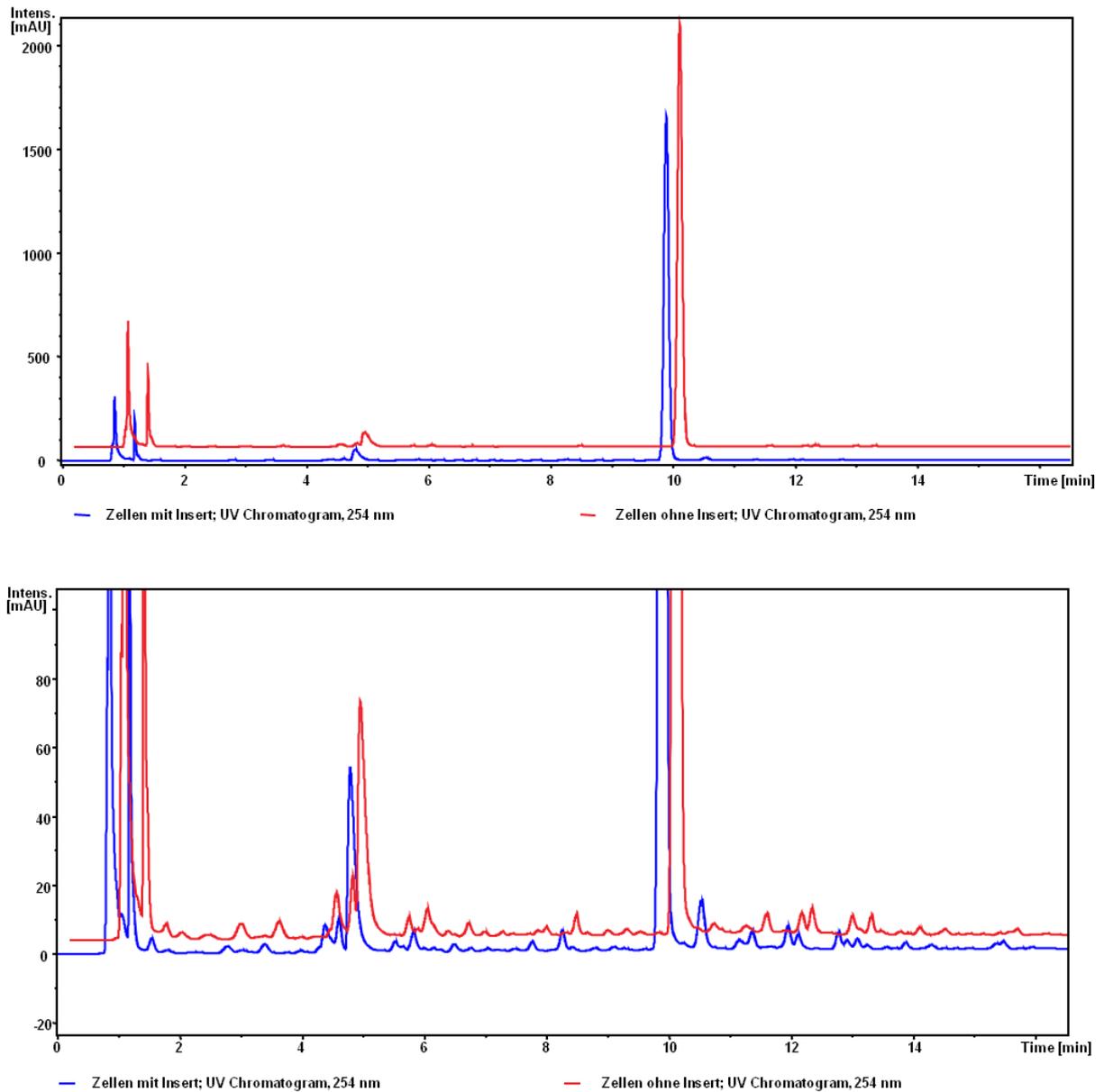


Abbildung 10-19: UV-Chromatogramm (254 nm) der Ethylacetat-Extraktionen der Zellen. LC/MS-Messung der Ethylacetat-Extraktionen der Zellen nach Expression des Plasmids mit ohne ohne PKS/NRPS-Gencluster. In **rot** als Negativkontrolle das Extrakt der Zellen ohne Insert (PKS/NRPS-Cluster) in **blau** das Extrakt der Zellen mit Insert (PKS/NRPS-Cluster). Die untere Abbildung zeigt das Spektrum in höherer Auflösung.

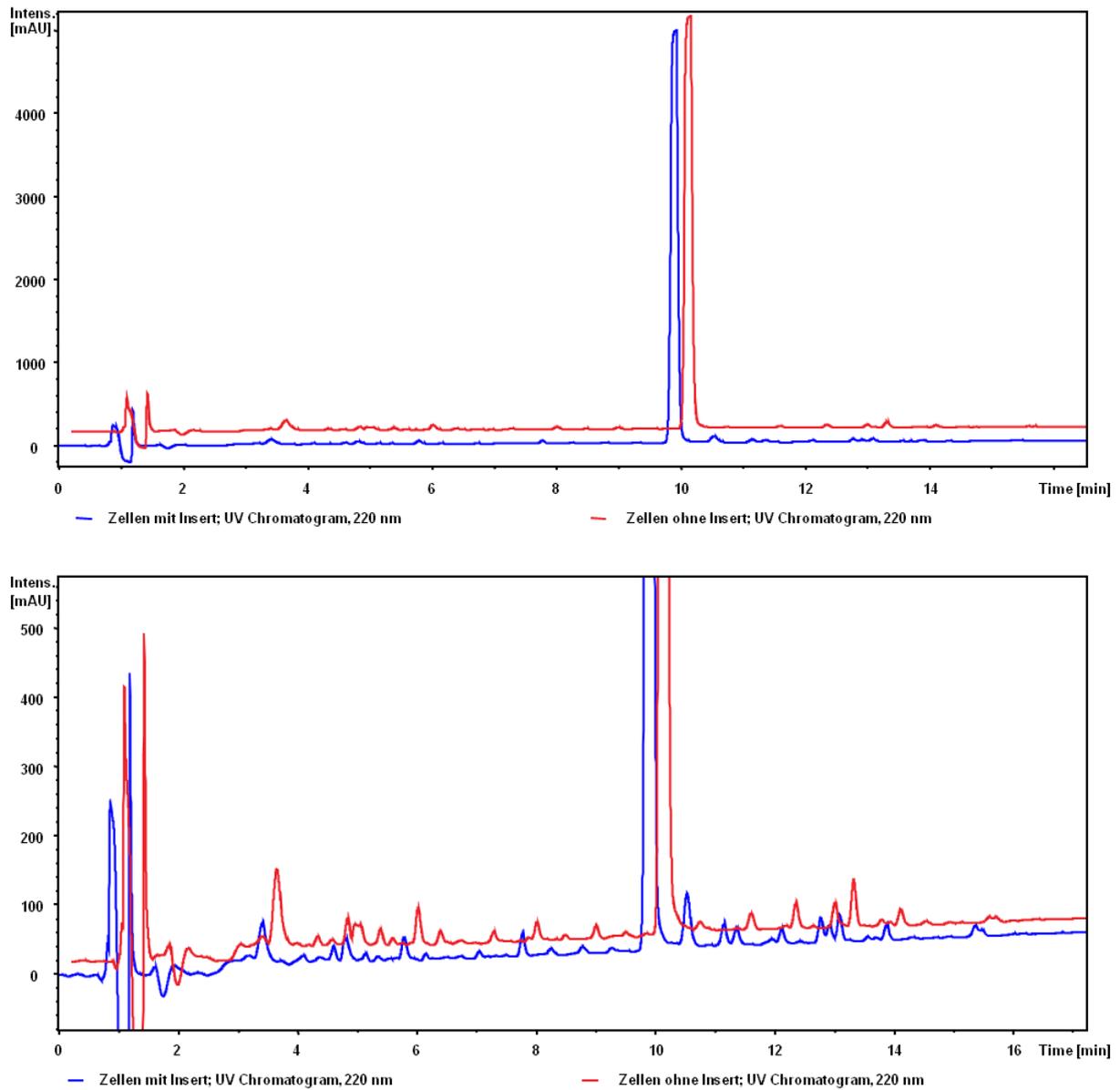


Abbildung 10-20: UV-Chromatogramm (220 nm) der Ethylacetat-Extraktionen der Zellen. LC/MS-Messung der Ethylacetat-Extraktionen der Zellen nach Expression des Plasmids mit ohne ohne PKS/NRPS-Gencluster. In **rot** als Negativkontrolle das Extrakt der Zellen ohne Insert (PKS/NRPS-Cluster) in **blau** das Extrakt der Zellen mit Insert (PKS/NRPS-Cluster). Die untere Abbildung zeigt das Spektrum in höherer Auflösung.

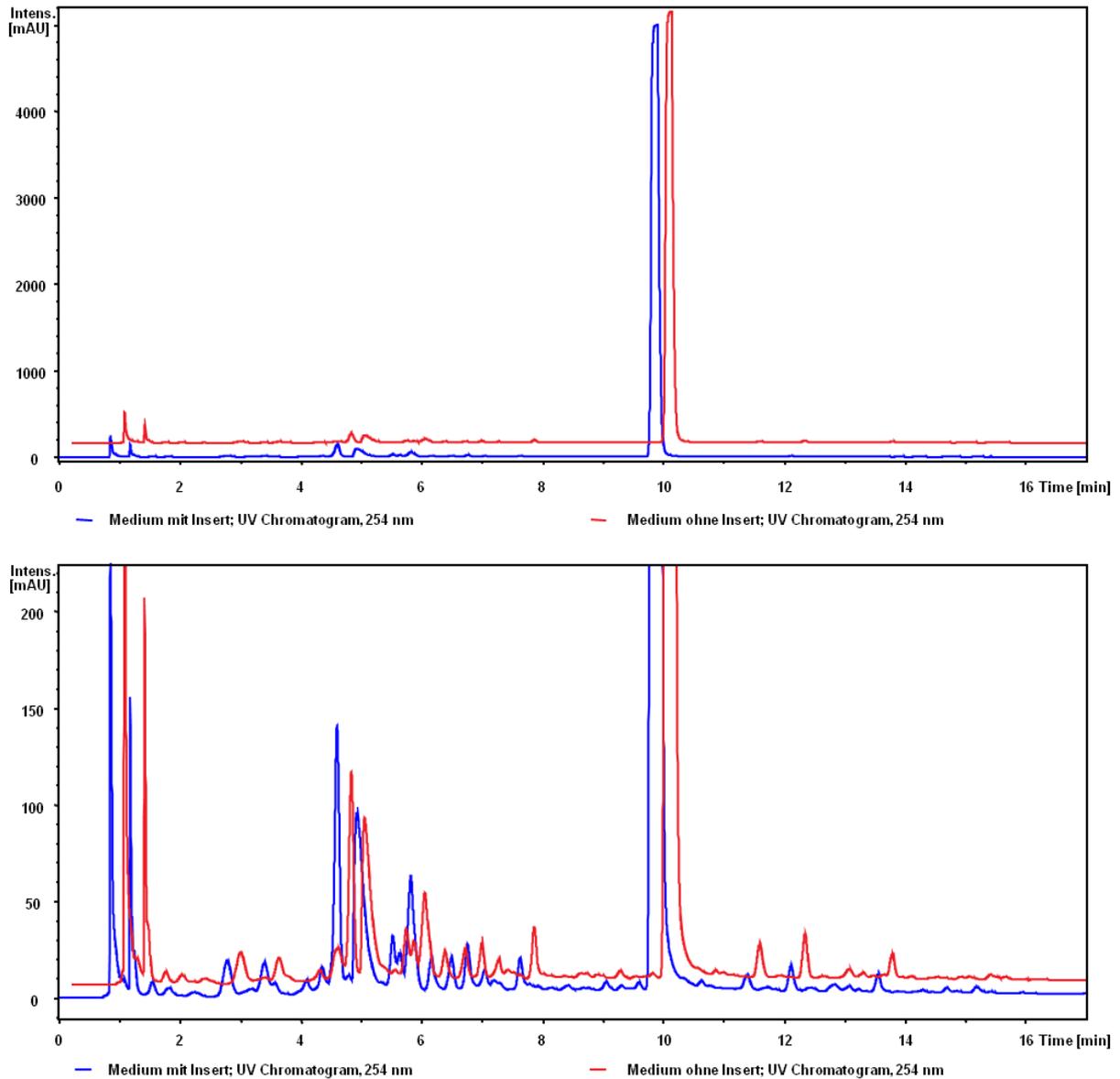


Abbildung 10-21: UV-Chromatogramm (254 nm) der Ethylacetat-Extraktion des Mediums. LC/MS-Messung der Ethylacetat-Extraktion der Zellen nach Expression des Plasmids mit ohne ohne PKS/NRPS-Gencluster. In **rot** als Negativkontrolle das Extrakt der Zellen ohne Insert (PKS/NRPS-Cluster) in **blau** das Extrakt der Zellen mit Insert (PKS/NRPS-Cluster). Die untere Abbildung zeigt das Spektrum in höherer Auflösung.

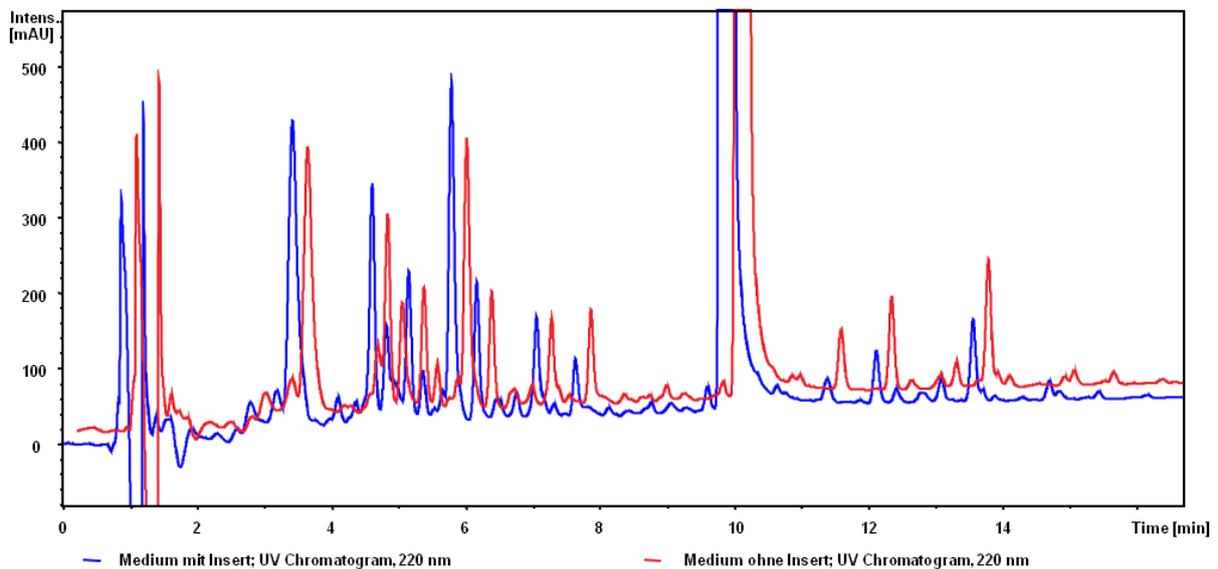
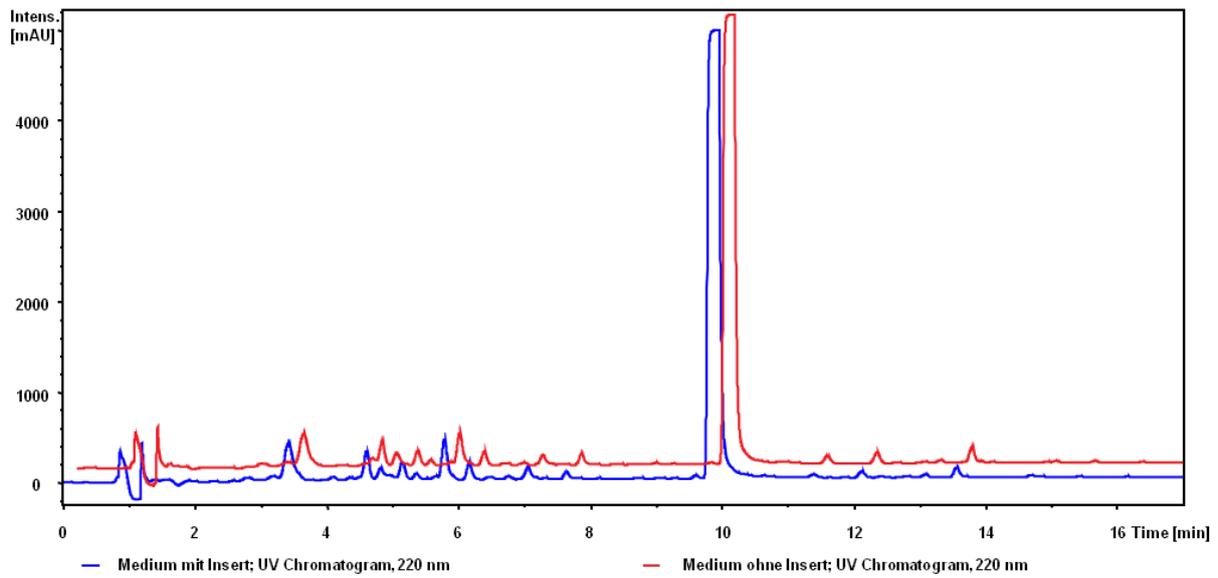


Abbildung 10-22: UV-Chromatogramm (220 nm) der Ethylacetat-Extraktion des Mediums. LC/MS-Messung der Ethylacetat-Extraktion der Zellen nach Expression des Plasmids mit ohne ohne PKS/NRPS-Gencluster. In **rot** als Negativkontrolle das Extrakt der Zellen ohne Insert (PKS/NRPS-Cluster) in **blau** das Extrakt der Zellen mit Insert (PKS/NRPS-Cluster). Die untere Abbildung zeigt das Spektrum in höherer Auflösung.

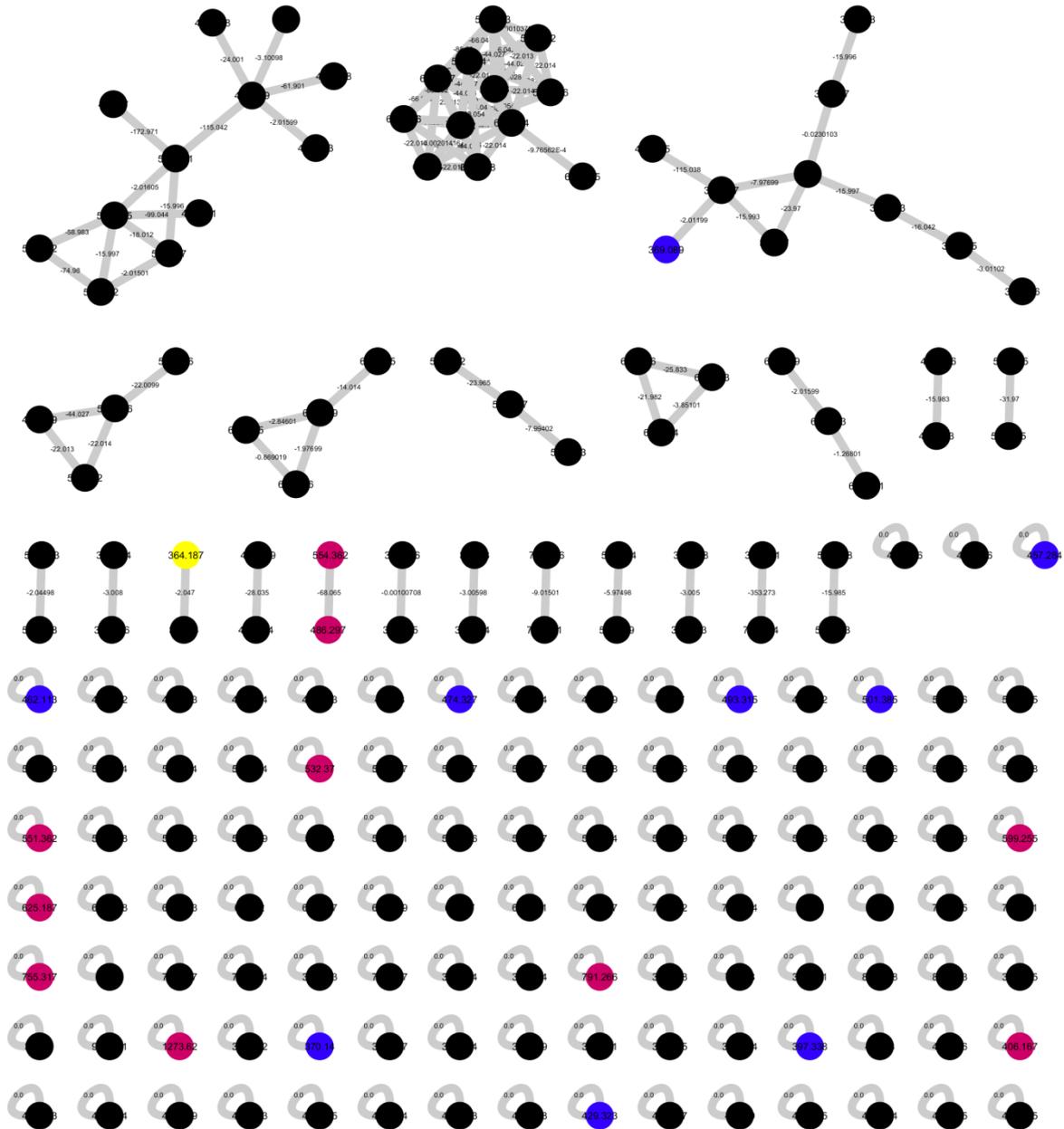


Abbildung 10-23: Global Natural Products Social Molecular networking (GNPS). Verglichen wurden die in Abbildung 10-19, Abbildung 10-20, Abbildung 10-21 und Abbildung 10-22 gezeigten Spektren. In **schwarz** sind die Massen dargestellt, die mindestens in einer der beiden Negativkontrollen vorkommen, **blau** sind die Massen die nur im Medium der Expression des Genclusters vorhanden sind und die in **rot** gezeigten Massen sind ausschließlich in den Zellen der Expression des Genclusters zu finden.

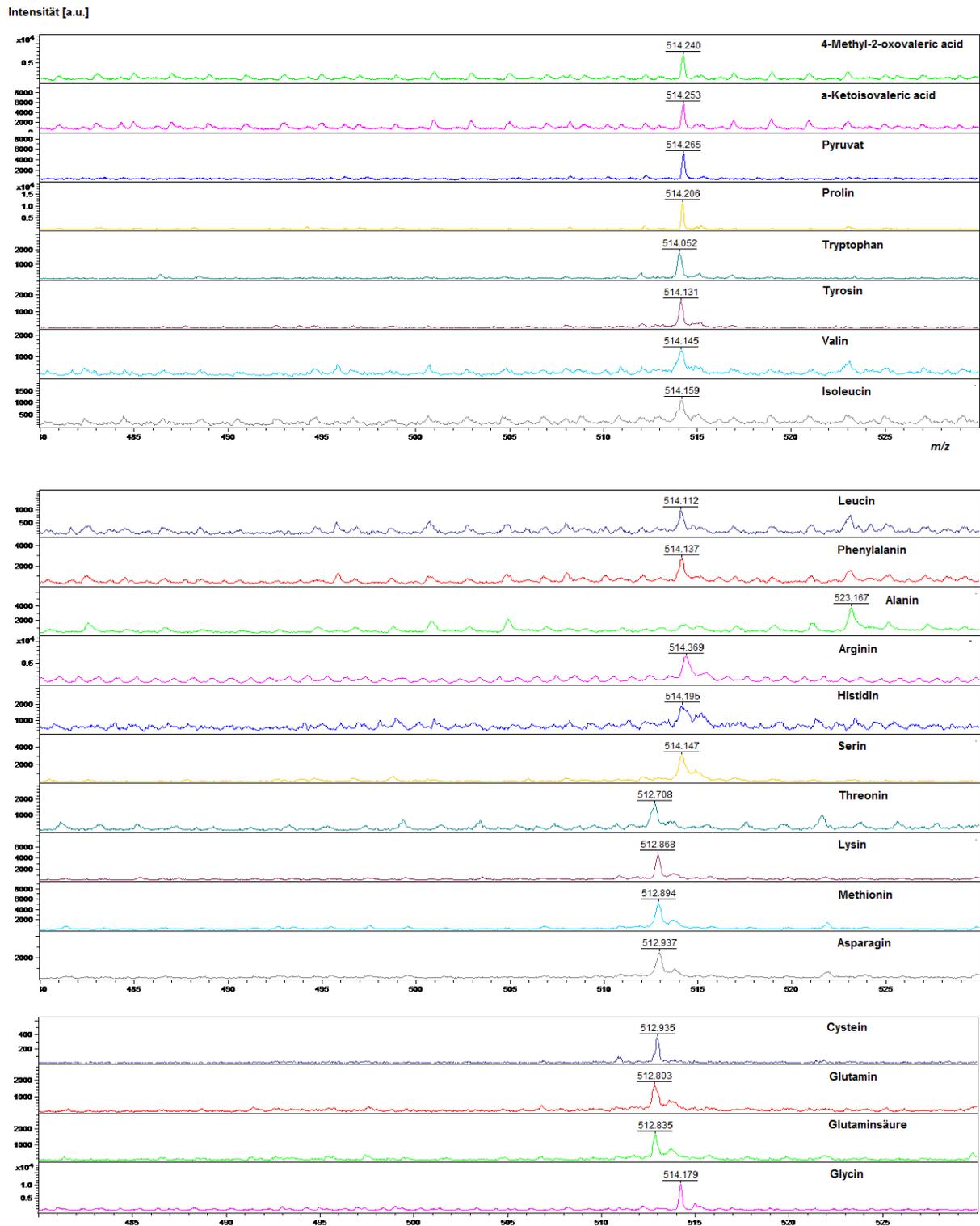


Abbildung 10-24: MALDI-TOF Spektren verschiedener $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assays. Dargestellt sind beispielhaft die MALDI-Spektren mehrere $[\gamma\text{-}^{18}\text{O}_4]\text{-ATP}$ -Assays zur Bestimmung der Substratspezifität der pppE A-Domäne. Es wurden verschiedene Substrate getestet.