

# **Molekularbiologische Charakterisierung von nicht-kultivierten Crenarchaeota und Acidobacteria in Metagenombanken aus Bodenproben**

vom Fachbereich Biologie der  
Technischen Universität Darmstadt

zur

Erlangung des akademischen Grades eines

*Doctor rerum naturalium*

Genehmigte

**DISSERTATION**

vorgelegt von

Achim Quaiser

1. Berichterstatterin: Prof. Dr. Felicitas Pfeifer

Mikrobiologie, TU Darmstadt

2. Berichterstatter: Prof. Dr. H. Ulrich Göringer

Mikrobiologie, TU Darmstadt

Eingereicht am: 19.12.2002

Tag der mündlichen Prüfung: 11.02.2003

Darmstadt 2003

## **Erklärungen**

Die vorliegende Arbeit wurde unter der Leitung von Dr. Christa Schleper in der Arbeitsgruppe von Frau Professor Dr. Felicitas Pfeifer an der Technischen Universität Darmstadt am Institut für Mikrobiologie und Genetik in der Zeit von Oktober 1999 bis Dezember 2002 angefertigt.

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertation selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe. Weiterhin erkläre ich, dass ich bisher keinen Versuch unternommen habe, an einer deutschen oder ausländischen Universität zu promovieren.

Darmstadt, den 18.12.2002

(Achim Quaiser)

## Danksagung

**Christa Schleper** möchte ich besonders für die Möglichkeit ein so interessantes Thema zu bearbeiten danken. Durch ihre einmalige Art der Betreuung schafft sie es immer zu motivieren, und hilft einem so über Durststrecken hinweg. "Deine wissenschaftliche Begeisterungsfähigkeit ist wirklich einmalig.... "

**Frau Prof. F. Pfeifer** möchte ich für Ihre konstruktive Kritik und neue Denkanstöße danken. Sie ist für mich immer ein wichtiger beruhigender und verlässlicher Pol gewesen, wenn es manchmal heiss herging.

**Arnulf Kletzin** danke ich für seine wissenschaftliche Einbringung in meine Arbeit. Er hat immer ein offenes Ohr und hat mir so einige Dokumente gerettet.

**Herrn Prof. Dr. Jobst Gmeiner** möchte ich danken für seine Nachmittagsmüdigkeits-vertreibungsrundgänge und für seine Hilfe im Chaos und seine Diskussionsbereitschaft.

Torsten danke ich für die wahnsinnigen Ideen und Diskussionen, die wir während unserer wissenschaftlichen Nächte entwickelten.

Eri, Melli Alex und Riti danke ich nicht nur für die lustigen Momente.

Gabi Liebig danke ich für Geduld und ihr Wesen.

Holger Zinke und Jürgen Eck bin ich sehr dankbar für die klaren und ehrlichen Worte, ohne die ich noch nicht so weit wäre.

Diese Promotionsarbeit wurde durch ein dreijähriges Stipendium der Firma **BRAIN AG** unterstützt.

Der **BRAIN AG** danke ich ausserdem für den Wissenschaftserwerb während dieser Promotionszeit und vor allem, das gilt für alle, für die Diskussionen nicht nur während der Seminare, sondern auch bei diversen Grillabenden. Bernd und Guido danke ich für die fruchttragende Zusammenarbeit während meiner "BRAIN-Zeit".

Silke danke ich für Lustigkeiten. Christian danke ich auch für die vielen Sequenzen.

Hans-Peter Klenk danke ich für die FixA Phylogenie und Arne Zibat für die Subklonierung und Sequenzierung des 2914-Inserts.

Stefan Schuster, Günter Raddatz und Christa Lanz danke ich für die Sequenzen der Klone mit acidobakteriellem Insert.

Ute Hentschel und Lars Fieseler danke ich für die schöne Zeit in Banyuls/sur mer und für die dieses so interessante Projekt der Schwamm-assoziierten-Prokaryonten.

Ein besonderer Dank gilt meinen Eltern **Christa** und **Jakob Quaiser** für ihre kontinuierliche nicht nur finanzielle Unterstützung.

**Florence**, ma chère adorée de mon coeur et **Perle** ma Trognine, d'aujourd'hui et toujours demain...

## Veröffentlichungen

- Quaiser A.#, Ochsenreiter T.#, Lanz C., Schuster SC., Treusch AH., Eck J., and Schleper C. (2003) Acidobacteria form a Coherent but Highly Diverse Group within the Bacterial Domain: Evidence from Environmental Genomics; *Molecular Microbiology* (in press)  
# These authors contributed equally to this work
- Ochsenreiter T. Selesi D. Bonch-Ozmolovskaya L., Quaiser A., Schleper C. (2003) Diversity and Abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and Real Time PCR; *Environmental Microbiology* 5 (9), 787–797
- Quaiser A., Ochsenreiter T., Klenk HP., Kletzin A., Treusch AH., Meurer G., Eck J., Sensen CW. and Schleper C. (2002) First insight into the genome of an uncultivated crenarchaeote from soil. *Environmental Microbiology* 4(10):603-611.
- Quaiser A., Ochsenreiter T., Treusch AH., Schleper C., Lorenz P., Eck J (2002)., Isolation and cloning of DNA from uncultivated organisms, *Priority-establishing EP-patent application No. 02 01 8210.1: Brain AG, Zwingenberg*

## Inhaltsverzeichnis

<b>A. Zusammenfassung .....</b>	<b>4</b>
<b>B. Einleitung.....</b>	<b>6</b>
<b>C. Ergebnisse .....</b>	<b>16</b>
<b>1. Entwicklung einer Methode zur Isolierung von hochmolekularer, klonierbarer DNA aus Bodenproben .....</b>	<b>16</b>
1.1. Herstellung einer mikrobiellen Fraktion .....	16
1.2. Einbetten der Biomasse in Agarose-Nudeln.....	17
1.3. Lyse von Zellen in Agarose-Nudeln.....	18
1.4. Reinigung von hochmolekularer DNA aus Bodenproben: das 2-Phasengel .....	19
1.4.1. Isolierung von hochmolekularen DNA Fragmenten aus Agarosegelen .....	22
<b>2. Erstellung und Charakterisierung von Metagenombanken mit hochmolekularen DNA-Inserts .....</b>	<b>23</b>
2.1. Restriktionsanalyse von Fosmid-Klonen der Genbank Rud001.AQ .....	26
2.2. Anwendung der DNA-Präparationstechnik auf weitere Bodenproben .....	27
2.3. Bestimmung der Diversität, der in den Genbanken klonierten DNA.....	28
2.4. Isolierung und Klonierung hochmolekularer DNA aus Schwamm-assoziierten-Prokaryonten.....	29
2.4.1. Trennung der prokaryotischen Zellen von Schwammzellen .....	30
2.4.2. Einbetten der Biomasse in Agarose-Nudeln .....	30
2.4.3. Reinigung und Isolierung von hochmolekularer DNA von Schwamm-assoziierten-Prokaryonten .....	31
2.4.4. Klonierung der Schwamm-assoziierten-Prokaryonten DNA .....	33
<b>3. Identifizierung und Analyse von archaealen und acidobakteriellen Genomfragmenten aus Metagenombanken .....</b>	<b>35</b>
3.1. Identifizierung von archaealen Fosmid-Klonen .....	35
3.2. Analyse des archaealen Fosmid-Klons 2914 .....	37
3.3. Identifizierung von Genomfragmenten aus Acidobacteria .....	44

3.3.1. Identifizierung von Klonpools mit acidobakteriellen 16S rRNA-Genen.....	44
3.3.2. Klonierung und Sequenzierung der 16S rDNA Fragmente.....	46
3.3.3. Identifizierung von Einzelklonen mit acidobakteriellen 16S rRNA-Genen.....	48
3.4. Analyse des acidobakteriellen Klons 32D1 .....	49
3.4.1. Vergleichende Genomanalyse des kolinearen Bereichs auf Klon 32D1 .....	52
<b>D. Diskussion.....</b>	<b>58</b>
Isolierung von hochmolekularer DNA und Abtrennung von Humin- und Fulvinsäuren .....	59
Herstellung der Umwelt-Genbanken .....	61
Suche nach archaealen Genomfragmenten in den Genbanken .....	63
Suche nach acidobakteriellen Genomfragmenten.....	63
Charakterisierung eines Genomfragmentes von nicht-kultivierten Crenarchaeoten .	65
Charakterisierung eines Genomfragmentes von nicht-kultivierten Acidobacteria.....	67
Horizontaler Gentransfer zwischen Rhizobiaceae und Acidobacteria ? .....	70
Ausblick .....	71
<b>E. Material und Methoden.....</b>	<b>73</b>
<b>1. Material.. .....</b>	<b>73</b>
1.1. Chemikalien und Hilfsmittel.....	73
1.2. Synthetische Oligonukleotide.....	74
1.3. Enzyme .....	74
1.4. DNA-Größenstandards.....	74
1.5. Medien .....	75
1.6. Puffer und Lösungen .....	75
<b>2. Methoden.....</b>	<b>76</b>
2.1. Beschreibung der Bodenprobe .....	76
2.2. Isolierung von hochmolekularer, klonierbarer DNA aus Bodenproben und marinen Invertebraten .....	76
2.2.1. Gewinnung der mikrobiellen Fraktion .....	76
2.2.2. Herstellung von Agarose-Nudeln .....	77
2.2.3. Zelllyse .....	77

## Inhaltsverzeichnis

---

2.2.4. Gelelektrophoretische Reinigung der hochmolekularen DNA aus Boden: 2-Phasengel .....	77
2.2.5. Restriktion von hochmolekularer DNA in Agarose-Nudeln .....	78
2.2.6. Elektroelution von hochmolekularer DNA .....	78
2.2.7. Konzentrierung der DNA .....	78
2.3. Klonierung hochmolekularer DNA in Fosmide .....	79
2.3.1. Manipulation der DNA Enden .....	80
2.3.2. Ligation .....	81
2.3.3. Verpackung des Ligationsansatzes.....	81
2.3.4. Infektion und Ausplattierung .....	81
2.4. Herstellung von Glycerinkonserven der Genbanken.....	82
2.5. Herstellen von Matrizen-DNA für Multiplex-Polymerase-Kettenreaktion .....	82
2.6. Isolierung von rekombinanten Fosmiden .....	82
2.7. Restriktionsanalyse von isolierten Fosmid-DNAs .....	83
2.8. Polymerase-Kettenreaktion zur Amplifizierung von 16S rRNA-Genen .....	84
2.8.1. Isolierung von PCR-Produkten aus Agarosegelen .....	85
2.8.2. Klonierung von PCR-Produkten .....	85
2.9. DNA-Sequenzbestimmung .....	85
2.10. Herstellung DIG-markierter DNA-Sonden .....	87
2.11. Herstellung von Koloniefiltern.....	87
2.12. Filterhybridisierung .....	88
2.12.1. Hybridisierung .....	88
2.12.2. Waschen und Exposition: .....	88
2.13. Subklonierung und Sequenzierung .....	89
2.14. Phylogenetische Analyse.....	89
2.15. Sequenzannotation .....	90
2.16. Schwammmaterial.....	90
2.17. Trennung der prokaryotischen Zellen von Schwammzellen.....	90
<b>F. Literatur.....</b>	<b>91</b>

## A. Zusammenfassung

Molekular-phylogenetische Studien, basierend auf der Charakterisierung von 16S rRNA-Genen, haben gezeigt, dass Bodenhabitate besonders reich an mikrobieller Diversität sind. So wurde z.B. eine bestimmte Gruppe der Crenarchaeota (Archaea), die bislang als Extremophile galten, durch 16S rDNA-Studien häufig in Bodenhabitaten entdeckt. Weiterhin konnte das Phylum der Acidobacteria, das derzeit von ca. 200 16S rDNA Umwelt-Sequenzen und von nur 3 kultivierten Vertretern repräsentiert wird, definiert werden. Es ist aufgrund seiner phylogenetischen Breite mit den Proteobacteria vergleichbar. Acidobacteria stellen häufig eine dominierende Fraktion der Mikroorganismen in Bodenhabitaten dar, sind jedoch bis heute ebenso wie die mesophilen Crenarchaeota weitgehend uncharakterisiert. Aus Mangel an geeigneten Techniken bleiben bislang insgesamt mehr als 99% der Mikroorganismen in den meisten Habitaten uncharakterisiert, weil sie nicht im Labor kultiviert werden können.

Inspiziert von den schnellen Fortschritten der mikrobiellen Genomik wurde in dieser Arbeit die Strategie der Umweltgenomik auf Bodenhabitate angewendet, um erstmals mesophile, unkultivierte Crenarchaeota und Acidobacteria aus Böden zu charakterisieren. Voraussetzung dafür war die direkte Isolierung von hochmolekularer DNA aus Bodenproben, die aufgrund des Vorliegens von grossen Mengen an polyphenolischen Komponenten (Humin- und Fulvinsäuren) eine besondere Herausforderung darstellte. Es wurde eine Methodik entwickelt, die auf einer 2-phasigen Gelelektrophoresetechnik basierte, und die eine schnelle und reproduzierbare Isolierung von klonierbarer, hochmolekularer DNA aus unterschiedlichen Habitaten erlaubte. Von diesen Nukleinsäuren wurden komplexe Genbibliotheken mit großen Inserts konstruiert. Die Charakterisierung der Genbibliotheken zeigte, dass die in den Habitaten durch PCR-Studien dargestellte große mikrobielle Diversität tatsächlich auf genomischer Ebene abgebildet werden konnte.

Mit Archaea-spezifischen und Acidobacteria-spezifischen Sonden wurden je ein Genomfragment der mesophilen nicht-kultivierten Crenarchaeota und der Acidobacteria aus einer 960 Mbp großen Fosmid-Genbank einer Bodenprobe isoliert und charakterisiert.

Der archaeale Klon enthielt ein 34 kb Fragment, auf dem ein komplettes rRNA-Operon und 17 Protein-kodierende Gene lokalisiert waren. Die phylogenetischen Analysen der rRNA-Gene und eines Proteingens bestätigten die Zugehörigkeit des Genomfragments zur Gruppe der nicht-thermophilen Crenarchaeota. Die Genomstruktur und die auf dem Genomfragment lokalisierten Gene zeigten deutliche Unterschiede zu den bereits untersuchten, nahe-verwandten marinen Vertretern.



Die detaillierte Analyse eines 35 kb Fragments der Acidobacteria bestätigte die Zugehörigkeit des Genomfragmentes zu dieser noch unbeschriebenen, ökologisch aber relevanten Gruppe. Durch vergleichende Genomanalysen und phylogenetische Rekonstruktionen wurden Hinweise auf einen horizontalen Gentransfer zwischen *Rhizobiaceae* ( $\alpha$ -Proteobacteria) und Acidobacteria erhalten. Ein Bereich, der 10 kb mit 8 potentiellen Protein-kodierenden Genen überspannte, war colinear und zeigte große Sequenzähnlichkeiten mit einem Genombereich von *Rhodopseudomonas palustris*.

Die Charakterisierung der Protein-kodierenden Gene ermöglichte erste Einblicke in das physiologische Potential der mesophilen Crenarchaeota und der Acidobacteria und stellt eine Basis für weiterführende genomische Untersuchungen dar.

## B. Einleitung

Prokaryotische Mikroorganismen, Bacteria und Archaea, spielen auf unserem Planeten eine zentrale Rolle. Einerseits waren sie vor ca. 3,8 Milliarden Jahren die ersten Bewohner der Erde und aus ihnen sind höchstwahrscheinlich die Eukaryonten hervorgegangen. Andererseits ist auch heute die Existenz höherer Lebewesen von der Aktivität einer Vielzahl prokaryotischer Mikroorganismen abhängig. So ist der gesunde Mensch mit 10-100 mal mehr Bakterienzellen besiedelt als er eigene Zellen enthält. Ausserdem werden von prokaryotischen Mikroorganismen essentielle Schritte der globalen Stoffkreisläufe (Methanbildung, Stickstofffixierung, Nitrifikation, Denitrifikation) katalysiert und wichtige Bestandteile der Atmosphäre produziert. Prokaryonten dominieren auf unserem Planeten nicht nur bezüglich ihrer Zellzahl sondern nach neuesten Schätzungen machen sie auch mehr als die Hälfte der auf der Erde vorkommenden Biomasse aus (Whitman *et al.*, 1998). Etwa die Hälfte des Kohlenstoffs ist in Prokaryonten gebunden, bezogen auf Stickstoff und Phosphor sind es fast 90%.

Trotz ihrer im Vergleich zu allen anderen Lebewesen wesentlich längeren Entwicklungsgeschichte und ihrer enormen physiologischen Vielfalt wurde lange Zeit angenommen, dass der Hauptteil der Diversität des Lebens innerhalb der Eukaryonten zu finden ist. Erst phylogenetische Studien auf Basis der vergleichenden Sequenzanalyse der rRNA konnten aufzeigen, dass die Prokaryonten in Form der Archaea und Bacteria zwei von drei Hauptentwicklungslinien des Lebens bilden, deren phylogenetische Tiefe insgesamt sehr viel größer ist als die der Eukaryonten (Woese *et al.*, 1990). In krassem Widerspruch hierzu steht allerdings die Zahl der bisher beschriebenen Arten innerhalb der Prokaryonten. Während bereits über 1,33 Millionen eukaryotische Spezies beschrieben sind (davon alleine ca. 950000 Insektenarten), beschränkt sich die Zahl beschriebener Bakterien- und Archaea-Arten auf ca. 4500 (Torsvik *et al.*, 2002; Heywood, 1995). Diese Diskrepanz ist hauptsächlich darauf zurückzuführen, dass Prokaryonten bisher nach ihrer Isolierung mit klassischen Kultivierungsmethoden beschrieben wurden. Hiermit lassen sich jedoch nur ein Bruchteil (je nach Umweltprobe zwischen 0,001% und 15%) der tatsächlich in der Natur vorkommenden Prokaryonten in Reinkultur erhalten und einer taxonomischen Beschreibung zuführen (Wagner *et al.*, 1993). Schätzungen bezüglich der Gesamtzahl an existierenden Prokaryontenspezies gehen von 1,5 – 14 Millionen aus (Palleroni, 1997). Anhand des Unvermögens die meisten Prokaryonten anzuzüchten, wird die unzureichende Kenntnis über deren Wachstumsbedürfnisse deutlich.

Aufgrund der noch weitgehend unentdeckten Welt der Prokaryonten konzentriert sich das Interesse einiger Forschungsgruppen auf die Identifizierung von bisher nicht-kultivierten und nicht-kultivierbaren Archaea und Bacteria.

Der Einsatz molekulargenetischer Methoden, die unabhängig von der Kultivierbarkeit der Mikroorganismen sind, hat in den letzten Jahren bereits zum Nachweis einer großen Anzahl von Archaea und Bacteria in verschiedenen Habitaten geführt. Diese neuen Methoden beruhen auf der direkten Isolierung von Gesamtnukleinsäuren aus einer Umweltprobe und ihrer Verwendung als Matrize in einer Polymerase-Kettenreaktion (PCR), um 16S rRNA-Gene zu amplifizieren (Olsen *et al.*, 1986). Bei diesem Verfahren werden zu konservierten Bereichen der rRNA der gesuchten Organismengruppe (z.B. Prokaryonten, Archaea, Bacteria) komplementäre Oligonukleotide als Startermoleküle für die PCR benutzt. Die erhaltenen Produkte werden kloniert, sequenziert und die Ergebnisse zur Rekonstruktion phylogenetischer Stammbäume verwendet. Das rRNA-Molekül ist als phylogenetischer Marker besonders geeignet, da es ubiquitär verbreitet ist und ausserdem Spezies-spezifisch hochkonservierte wie auch variable Bereiche enthält und somit eine eindeutige Identifizierung einzelner Spezies bzw. 16S rRNA-Typen (Phylotypen) erlaubt.

Auch die heute gültige Einteilung der Organismen in die drei Domänen Bacteria, Archaea und Eukaryonten wurde durch Sequenzvergleiche der 16S bzw. 18S rRNA-Gene und dem daraus rekonstruierten phylogenetischen Stammbaum bestätigt (Woese *et al.*, 1990). Archaea bilden demnach neben Bacteria und Eukaryonten eine eigene Domäne. Sie besitzen Merkmale der Bakterien, wie den fehlenden Nukleus, das meist zirkuläre Chromosom und verschiedene, speziell prokaryotische Stoffwechselwege, aber auch Merkmale der Eukaryonten, z.B. in ihrem Transkriptionsapparat (Olsen und Woese, 1997).

Eine der beeindruckendsten Entdeckungen auf dem Gebiet der molekularen Ökologie war allerdings das ubiquitäre Vorkommen von archaealen rDNA-Sequenzen in mesophilen, marinen Habitaten, was die völlig unzureichende Charakterisierung der Crenarchaeota als "Extremophile" mittels klassischer, auf Kultivierung basierender Methoden verdeutlichte (DeLong, 1992; Fuhrman *et al.*, 1992; DeLong *et al.*, 1994; Fuhrman und Davis, 1997; Massana *et al.*, 1997).

Bislang wurden die Vertreter der beiden Königreiche der Archaea, die Crenarchaeota und Euryarchaeota (Woese *et al.*, 1990) als ausschließlich in extremen oder anaeroben Habitaten vorkommenden Mikroorganismen beschrieben. Die Vertreter der Euryarchaeota, die in Halophile, Thermophile und Methanogene eingeteilt werden, bilden eine physiologisch sehr heterogene Gruppe, im Gegensatz zu den Crenarchaeota deren kultivierte Vertreter nur in (hyper-)thermophilen Standorten zu finden sind.

Um so überraschender war das ubiquitäre Auftreten von rDNA-Sequenzen von "nicht-thermophilen" Crenarchaeota auch in mesophilen Habitaten. Die meisten dieser durch 16S rRNA-Analysen identifizierten und bislang nicht kultivierten Archaea bilden eine monophyletische Gruppe, die mit den extrem thermophilen Crenarchaeota assoziiert ist (Abbildung 1).

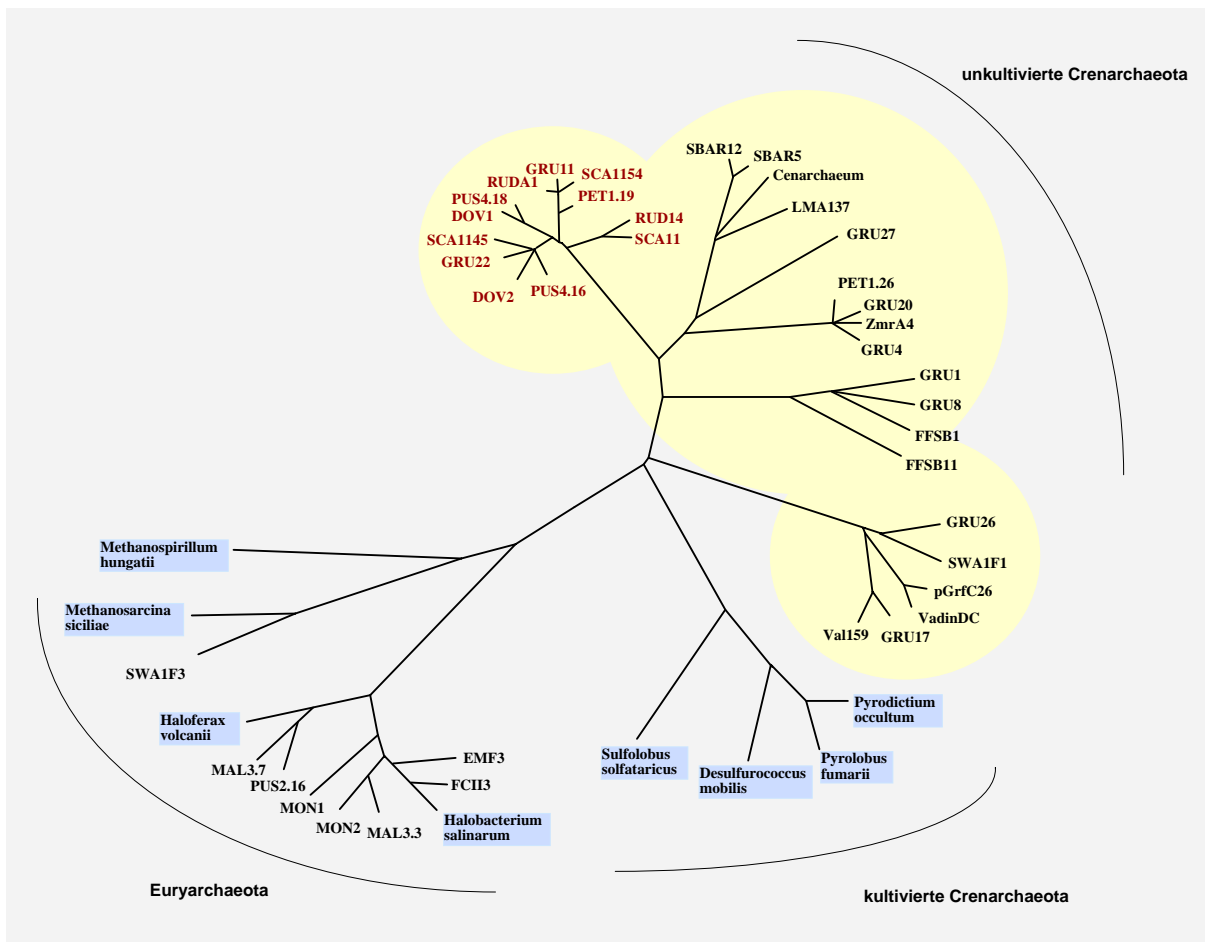


Abbildung 1: Phylogenetische Analyse von ausgewählten 16S rDNA-Sequenzen kultivierter und nicht-kultivierter Archaea. Die meisten dargestellten Sequenzen sind aus einer molekular-ökologischen Studie unseres Labors hervorgegangen (Selesi, 2000; Ochsenreiter *et al.*, 2002; Ochsenreiter *et al.*, eingereicht). Sequenzen, die direkt aus Umweltproben amplifiziert wurden, tragen Kürzel (Ausnahme: Cenarchaeum: *Cenarchaeum symbiosum*, der Symbiont eines marinen Schwammes): GRU: Wasserprobe der Prinz von Hessen Grube bei Darmstadt; RUD: aerobe Bodenschicht des Sandökosystems "Am Rotböhl" bei Darmstadt; PET: aerobe Bodenschicht bei Petropavlovsk, Kamtschatka/Russland; PUS: aerobe Bodenschicht bei Puschino, Kamtschatka; DOV Kalkboden in Dover, England; SWA Sumpf bei Moskau; MAL: Mikroorganismenmatte bei Malki, Kamtschatka; MON: alkalischer Boden, Mongolei; FCI: Salzgewinnungsanlage bei Fuencaliente, Spanien; EMF: Alkalischer Salzsee "Eight Miles Flat" in Nevada, USA; Die übrigen Abkürzungen entstammen publizierten Studien anderer Laboratorien (Übersicht in Jürgens *et al.*, 1999); rot dargestellt: eine spezifische Gruppe der nicht-thermophilen Crenarchaeota in Böden (Ochsenreiter, persönliche Mitteilung).

16S rRNA Sequenzen von Archaea wurden zuerst in marinem Pikoplankton identifiziert (DeLong, 1992; Fuhrman *et al.*, 1992). Auch in Seesedimenten, Waldböden und landwirtschaftlich genutzten Böden wurden archaeale Phyla entdeckt (Ueda *et al.*, 1995). Weiterhin konnten 16S rDNA-Sequenzen von "nicht-thermophilen" Crenarchaeota auch in Süßwassersedimenten (Hershberger *et al.*, 1996; MacGregor *et al.*, 1997; Ovreas *et al.*, 1997; Schleper *et al.*, 1997a; Vetriani *et al.*, 1999), im Darm einer Tiefseegurke (McInerney *et al.*, 1995) und im Gewebe eines marinen Schwammes (Preston *et al.*, 1996; Schleper *et al.*, 1998) nachgewiesen werden. Quantitative Analysen hatten gezeigt, dass Archaea im späten Winter bis zu 25% der prokaryotischen Biomasse des Oberflächenwassers des antarktischen Meeres ausmachen (DeLong, 1992; DeLong *et al.*, 1994).

Die große ökologische Verbreitung und Abundanz der mesophilen Crenarchaeota legt nahe, dass diese Gruppe der Archaea eine wesentliche Komponente der prokaryotischen Biomasse in moderaten Habitaten darstellen. Die Physiologie und ökologische Rolle dieser Organismen ist jedoch bislang unbekannt.

Auch unter den Bacteria konnten durch 16S rRNA Analysen viele neuen Gruppen identifiziert werden, so z.B. wurde das neue Phylum der Acidobacteria definiert (Hugenholtz *et al.*, 1998; Barns *et al.*, 1999). Acidobacteria konnten in den unterschiedlichsten Habitaten wie Böden, Süßwasser- und marinen Sedimenten identifiziert werden, wo sie häufig die dominierende Fraktion der Mikroorganismen ausmachten (Kuske *et al.*, 1997; Wise *et al.*, 1997): In einem trockenen Boden in Arizona fielen 51% der 16S rDNA-Sequenzen in dieses Phylum (Dunbar *et al.*, 1999), in einem Pinienwaldboden 55% (Kuske *et al.*, 1997), in einem landwirtschaftlich genutzten Boden machten sie 31% aus (Smit *et al.*, 2001) und in Weidelandböden 11% (McCaig *et al.*, 1999). Da Acidobacteria bisher in jeder 16S rDNA Analyse von Bodenhabitaten identifiziert wurden (Hugenholtz *et al.*, 1998), scheinen sie eine besonders wichtige Rolle in diesen Habitaten zu spielen.

Aus dem Phylum der Acidobacteria sind bisher nur drei kultivierte Vertreter bekannt und auch die wenigen physiologischen Kenntnisse, die man von diesen bereits hat, geben keine Hinweise auf Gemeinsamkeiten. Inzwischen sind ca. 200 16S rRNA Sequenzen von nicht-kultivierten Acidobacteria in den Datenbanken vorhanden. Im Gegensatz dazu steht eine einzige Proteinsequenz und die zugehörige Gensequenz eines Acidobacteriums, die in den Datenbanken zu finden sind. Es handelt sich um eine endo-1,4- $\beta$ -Xylanase von *Acidobacterium capsulatum* (Inagaki *et al.*, 1998). Die unzulängliche Kenntnis über diese große Gruppe von Mikroorganismen spiegelt sich in diesem Mißverhältnis wider.

Die 16S rDNA-Sequenzen der Acidobacteria lassen sich in mindestens 6 monophyletischen Gruppen einteilen. In seiner phylogenetischen Breite wird das Phylum mit Proteobacteria verglichen (Hugenholtz *et al.*, 1998). Von Proteobacteria sind im Vergleich

bereits ca. 1300 validierte Spezies beschrieben und von bereits mehr als 100 Spezies wird oder wurde bereits das Genom entschlüsselt. Anhand der phylogenetischen Breite und der Abundanz der Acidobacteria drängt sich eine enorme ökologische Relevanz vor allem in Bodenhabitaten auf. Aus diesem Grund sind sie nicht nur in unserem Labor zu interessanten Studienobjekten geworden.

Die Entwicklung der Techniken zur Analyse von 16S rRNA-Gensequenzen war ein wichtiger Schritt zur Identifizierung von noch unbekanntem Organismen und zur Erfassung der prokaryotischen Biodiversität. Allerdings können in den seltensten Fällen anhand von 16S rDNA-Sequenzen Aussagen über die Physiologie der zugehörigen Mikroorganismen gemacht werden, wodurch auch ihre ökologische Rolle unbekannt bleibt. Durch die Entwicklung neuerer Techniken, wie z.B. der Fluoreszenz-*in situ*-Hybridisierung (FISH) wird es zunehmend möglich, kultivierungsunabhängig quantitative Analysen über die mikrobielle Populationsstruktur zu machen und die Organismen in der Umweltprobe zu visualisieren (Simon *et al.*, 2000). Eine Kombination aus FISH und Mikroautoradiographie ermöglicht Einblicke in Aktivität und physiologische Eigenschaften von Bakterien ohne vorherige Kultivierung (Lee *et al.*, 1999). Hierbei wird die Umweltprobe mit radioaktiv-markiertem Substrat inkubiert, dann fixiert und anschließend mit fluoreszenzmarkierten Sonden hybridisiert. Zellen, die das radioaktiv-markierte Substrat umsetzen, werden detektiert und ihre Herkunft anschließend über FISH ermittelt. Damit wird einerseits die Lebensfähigkeit nachgewiesen und die phylogenetische Zuordnung der Organismen ermöglicht, und andererseits werden Hinweise über deren Physiologie und mögliche Nährsubstrate gewonnen.

Eine weitere bereits bekannte Methode, die *Differential Display* Technik, die erfolgreich bei der Untersuchung von eukaryotischer Genexpression angewendet wird, wurde zur Untersuchung von mikrobiellen Populationen unter Verwendung von 16S rRNA eingesetzt. Dadurch war es möglich, auf schnelle Art und Weise die Diversität von metabolisch aktiven Mikroorganismen in einem Habitat zu charakterisieren (Yakimov *et al.*, 2001). Auch die Expression von bestimmten Genen konnte in unterschiedlichen Bodenhabitaten über Extraktion von mRNA nachgewiesen werden (Chandler *et al.*, 1997; Rogers *et al.*, 2002).

Weiterhin wurden Techniken für quantitative Untersuchungen von unterschiedlichen Mikroorganismengruppen in Bodenhabitaten entwickelt, beispielsweise über *Slot Blot* Hybridisierungsanalysen (DeLong, 1992; Massana *et al.*, 1997) oder neuerdings mit Hilfe der *Real-Time* PCR (Hermansson und Lindgren, 2001; Ochsenreiter *et al.*, eingereicht).

Nach wie vor ist allerdings die Isolierung und Kultivierung von Mikroorganismen im Labor die einzige Möglichkeit, sie physiologisch und biochemisch genau zu charakterisieren. Die Entwicklung verbesserter Kultivierungsmethoden hat auch in den letzten Jahren

zur Isolierung neuer Mikroorganismen geführt (Janssen *et al.*, 2002; Kaeberlein *et al.*, 2002; Zengler *et al.*, 2002). Allerdings bleibt der Anteil kultivierbarer Spezies weiterhin gering, sicherlich nicht zuletzt, weil viele Mikroorganismen aufgrund ihrer speziellen Wechselwirkungen mit der Umwelt und mit anderen Spezies nicht in Reinkultur züchtbar sind.

Bei Betrachtung der erwähnten Techniken zur Charakterisierung von Mikroorganismen fällt auf, dass die gewonnenen Kenntnisse oft weitgehend voneinander unabhängige, kleine Mosaiksteinchen zum Verständnis der ökologischen Zusammenhänge in mikrobiellen Populationen darstellen. Aufgrund dessen ist es notwendig, weitere neue Methoden in der Mikrobiologie zu entwickeln, die eine nähere Charakterisierung von nicht-kultivierten oder nicht kultivierbaren Mikroorganismen erlauben und ganze Populationen in einem Habitat darstellen können, um Wechselwirkungen zu identifizieren.

Eine Möglichkeit der Annäherung an dieses Ziel kann durch die Umweltgenomik erreicht werden (Abbildung 2).

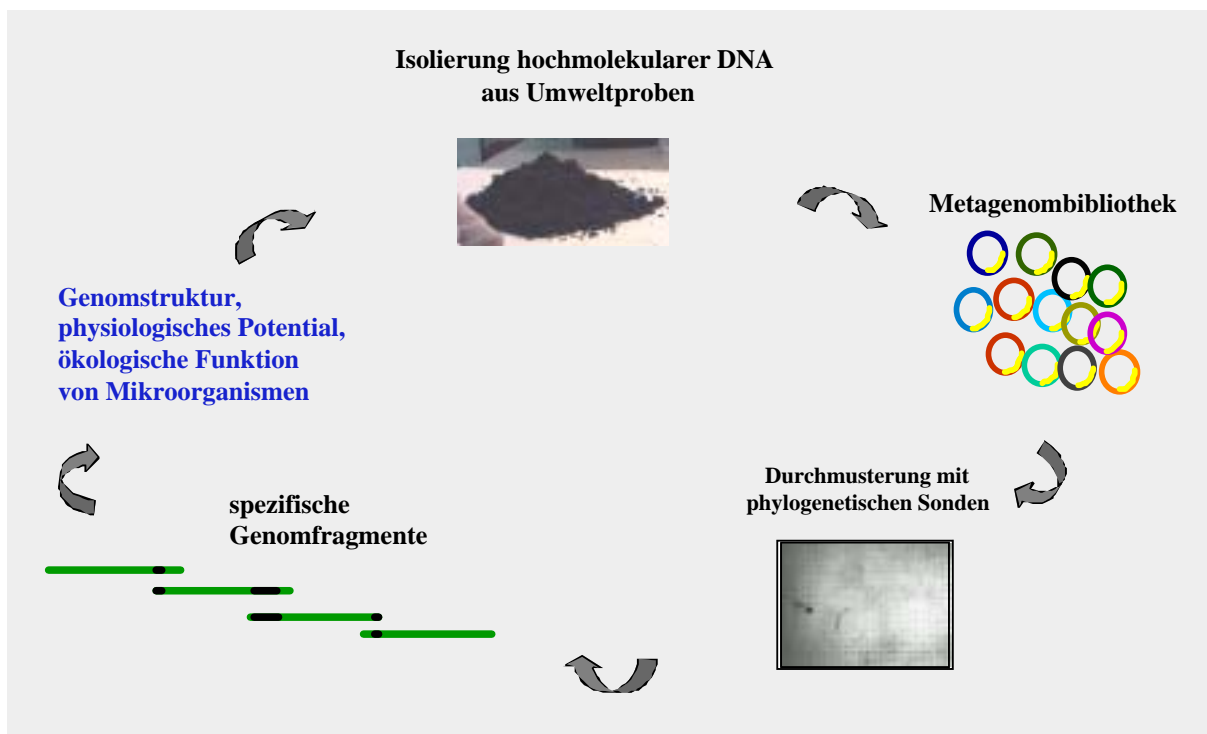


Abbildung 2 : Vereinfachte Darstellung der Strategie der Umweltgenomik

Dabei wird, in Anlehnung an die Sequenzierung und Analyse von Genomen von kultivierten Organismen, hochmolekulare chromosomale DNA von den in einer Umweltprobe vorhandenen Organismen kloniert und analysiert. Es wird angestrebt, die Gesamtheit der Genome in einer Population oder Umweltprobe – das sogenannte Metagenom

(Handelsman *et al.*, 1998; Schleper und Eck, 2000) - zu erfassen und in komplexen Genbanken abzulegen. Mit sogenannten phylogenetischen "Anker"- Genen, z.B. das 16S rRNA-Gen, werden dann spezifische Genomfragmente detektiert und analysiert. Die Möglichkeit zur Klonierung von hochmolekularen Inserts ist für die Umweltgenomik eine notwendige Voraussetzung. Besondere Bedeutung für die Etablierung der Umweltgenomik hatte daher die Entwicklung von Vektoren, die hochmolekulare DNA aufnehmen und stabil replizieren können. Vektoren, wie BAC- Fosmid- und Cosmid-Vektoren, werden insbesondere zur Analyse von komplexen eukaryotischen Genomen eingesetzt. Genbibliotheken, in denen große genomische DNAs konserviert sind, bilden eine wichtige Grundlage für Genomsequenzierungen, wo sie bei der Genom-Assemblierung und auch bei der Genomkartierung verwendet werden (Choi und Wing, 2000).

Die Methode der Umweltgenomik wurde bereits erfolgreich für die Charakterisierung von marinen Bacteria und Archaea eingesetzt. So konnte beispielsweise ein 40 kb großes Genomfragment eines planktonischen marinen Archaeons identifiziert werden (Stein *et al.*, 1996). Hierzu war DNA, die aus der Biomasse von Pikoplankton des Pazifik isoliert worden war, in einer Fosmid-Genbank abgelegt worden. Mittels PCR wurde hieraus ein Klon identifiziert, der ein archaeales 16S rRNA-Gen trug. Auf dieselbe Art und Weise wurden Genomfragmente eines symbiotischen, marinen Archaeons, *Cenarchaeum symbiosum*, aus einer Fosmid-Genbank isoliert, die aus Gesamt-DNA eines marinen Schwammes und der in ihm lebenden Organismen konstruiert worden war (Preston *et al.*, 1996; Schleper *et al.*, 1998). Durch die anschließende Sequenzierung und Analyse konnten die ersten Informationen über Proteingene und die Genomstruktur von *Cenarchaeum symbiosum* erhalten werden (Schleper *et al.*, 1998). Hierbei wurde eine DNA-Polymerase entdeckt, die in *E. coli* exprimiert wurde, wodurch es möglich war, das erste Protein eines nicht-kultivierten Archaeons zu charakterisieren (Schleper *et al.*, 1997b).

Weiterhin wurden umweltgenomische Studien auch für die Beschreibung von Genklustern mariner Proteobacteria eingesetzt. Dabei wurde ein neuer Rhodopsin-Typ, das Proteorhodopsin, das als lichtgetriebene Protonenpumpe fungiert, auf einem Genomfragment eines nicht-kultivierten marinen  $\gamma$ -Proteobakteriums gefunden (Beja *et al.*, 2000b; Beja *et al.*, 2001; Beja *et al.*, 2002a).

In dieser Arbeit sollte die Technik der Umweltgenomik erstmals auf Bodenbiotope angewendet werden. Bisher ist erstaunlich wenig über die im Boden stattfindenden Stofftransformationsprozesse und die mikrobiellen Populationen bekannt. In Bodenhabitaten kommen zwischen 3700-8800 unterschiedliche Spezies vor (Torsvik *et al.*, 2002). Man geht davon aus, dass der Hauptanteil der Stoffumsetzungen von Mikroorganismen



katalysiert wird, die in Böden, je nach Bodentyp, eine Konzentration von ca.  $10^9$  Zellen/g Boden ausmachen.

Die Isolierung von DNA aus Bodenproben ist kein einfaches Unterfangen. Verschiedene Humus-Bestandteile, die als Humine, Fulvin- und Huminsäuren bezeichnet werden, können mit DNA Molekülen komplexieren und behindern dadurch enzymatische Reaktionen während und nach der DNA Präparation. Huminsäuren sind hochmolekulare, heterozyklische Polyphenole (Abbildung 3), die starke Affinität zu Nukleinsäuren zeigen und nur schwer von DNA ablösbar sind.

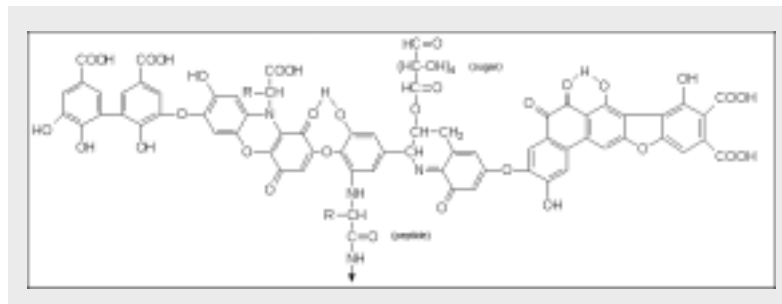


Abbildung 3: Hypothetische Modellstruktur von Huminsäuren nach Stevenson (Stevenson, 1982)

Um DNA in ausreichender Reinheit aus Boden zu erhalten, müssen besondere Reinigungsschritte durchgeführt werden. Es ist notwendig, die vorhandenen Inhibitoren, vor allem die Humin- und Fulvinsäuren, von den Nukleinsäuren zu trennen.

Außerdem unterscheiden sich unterschiedliche Böden im relativen Anteil der Humuskomponenten, was die Entwicklung von reproduzierbaren Methoden zur Isolierung von DNA aus Boden zusätzlich erschwert. Eine Reihe von Studien haben sich mit der Isolierung von DNA aus Bodenhabitaten beschäftigt (Zhou *et al.*, 1996; Krsek und Wellington, 1999; Miller *et al.*, 1999) und unterschiedliche Methoden wurden bereits zur Abtrennung der DNA von Humin- und Fulvinsäuren entwickelt. Die meisten Methoden basieren auf einer DNA-Extraktion aus wässrigen Lösungen (Zhou *et al.*, 1996). Inzwischen ist auch ein Kit (Fast-Präp Kit, Bio101) erhältlich, mit dem DNA mit guter Ausbeute, aber nur beschränkter Reinheit aus Bodenhabitaten isoliert werden kann. In einer Technik wurde Polyvinylpyrrolidon (PVP) in einem Agarosegel eingesetzt, um die phenolischen Komponenten während der Gelelektrophorese zu binden und dadurch die Humin- und Fulvinsäuren von der DNA zu trennen (Young *et al.*, 1993). Durch diese Reinigungsmethode war es möglich, hochgereinigte DNA zu isolieren.

Den erwähnten Methoden ist gemeinsam, dass die DNA aus Bodenproben zwar weitgehend, je nach Bodenbeschaffenheit, von Inhibitoren getrennt und somit in PCR Reaktionen eingesetzt werden konnte. Allerdings war die DNA bei der Präparation star-

ken Scherkräften ausgesetzt, wodurch nur niedermolekulare DNA isoliert werden konnte. Bisher gab es keine Methode, mit der verlässlich und reproduzierbar hochmolekulare, klonierbare DNA aus Bodenproben isoliert werden konnte.

Ziel dieser Studie war die Entwicklung einer Methode, die eine effiziente, reproduzierbare Reinigung und Isolierung von hochmolekularen, klonierbaren DNA-Fragmenten durch die Eliminierung der inhibitorischen polyphenolischen Humusbestandteile aus Bodenproben erlaubte. Durch Klonierung dieser DNA in einen Fosmid-Vektor sollten komplexe Genbibliotheken in *E. coli* abgelegt werden. Aus diesen Metagenombanken sollten Genomfragmente von Archaea und Acidobacteria mit Hilfe von 16S rDNA-Gensonden identifiziert und anschließend charakterisiert werden, um Informationen über die Genomstruktur und Genomorganisation zu erhalten.

Parallel zu dieser Arbeit wurde eine Studie veröffentlicht, in der ebenfalls hochmolekulare DNA kloniert wurde. Jedoch wurde in dieser Metagenombank nur nach biotechnologisch relevanten Proteingenen gesucht und kein phylogenetischer Zusammenhang hergestellt (Rondon *et al.*, 2000).

Die in der vorliegenden Arbeit verwendete Bodenprobe stammte aus Rotböhl, einem Standort bei Darmstadt, der zum Naturraum Hessische Rheinebene gehört (Abbildung 4). Dieser Standort ist typisch für ein Sandökosystem; er hat aufgrund eines Mangels an Kolloiden nur ein schwaches Nährstoffbindungsvermögen. Die Bodenprobe, die im Verlauf als RUD bezeichnet wird, entstammt einem ruderalisierten Bereich des Standorts (Storm *et al.*, 1998).



Abbildung 4: Naturraum Hessische Rheinebene am Rotböhl bei Darmstadt.

In einer vorangegangenen Studie war die mikrobielle Diversität von RUD anhand von amplifizierten 16S rRNA-Genen in unserem Labor untersucht worden (Selesi, 2000; Och-

senreiter *et al.*, eingereicht). Dabei wurden in einer 16S rRNA-Genbank bakterielle Sequenzen aus 10 unterschiedlichen bakteriellen Phyla identifiziert. 14% der Sequenzen gehörten zu den Acidobacteria. Aus einer archaealen 16S rRNA-Genbank von RUD wurden ausschließlich Sequenzen gefunden, die in die Gruppe der terrestrischen Crenarchaeota fielen (siehe Abbildung 1, rot dargestellt).

Ein Teil der im folgenden dargestellten Studie, die Entwicklung der Klonierungstechnik für DNA aus Bodenproben und die Charakterisierung eines archaealen Genomfragments, wurde bereits veröffentlicht (Quaiser *et al.*, 2002a; Quaiser *et al.*, 2002b) .

## C. Ergebnisse

### 1. Entwicklung einer Methode zur Isolierung von hochmolekularer, klonierbarer DNA aus Bodenproben

Das erste Ziel dieser Arbeit war die Entwicklung einer Technik, die es erlaubte, hochmolekulare, klonierbare DNA aus einer Bodenprobe zu isolieren. Die Mikroorganismen mussten zunächst von Bodenbestandteilen getrennt werden, um den Anteil an Humusbestandteilen zu reduzieren und um Biomasse anzureichern. Im nächsten Schritt wurde eine Methodik entwickelt, die es ermöglichte, die polyphenolischen Komponenten der Humusbestandteile durch eine elektrophoretische Auftrennung zu eliminieren und gleichzeitig hochmolekulare, klonierbare DNA zu isolieren.

#### 1.1. Herstellung einer mikrobiellen Fraktion

Um die prokaryotischen Zellen von anderen Bodenbestandteilen zu trennen, wurden 50 g der Bodenprobe des Habitats RUD (siehe E.2.1) in 300 ml eines Puffers, der  $\epsilon$ -Aminocapronsäure (10 mM) enthielt, über Nacht aufgeschlämmt (Abbildung 5).  $\epsilon$ -Aminocapronsäure ist ein Zwitterion mit einem hydrophoben Bereich, das ionische Bindungen neutralisieren kann. Es sollte die Dissoziation der Zellen von Bodenpartikeln unterstützen. Grobe Bodenbestandteile wurden durch Filtration des aufgeschlämmten Bodens entfernt (Abbildung 5). Das Filtrat wurde anschließend durch Zentrifugation sedimentiert.

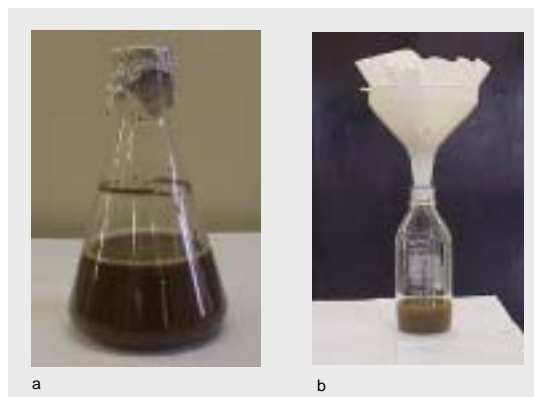


Abbildung 5: Bodenprobe RUD nach Aufschlammung (a) und nach Filtrierung (b) zur Herstellung einer mikrobiellen Fraktion; die braune Farbe der Suspension zeigt den hohen Gehalt an Humusbestandteilen.

Eine Bestimmung der Zellkonzentration konnte aufgrund der noch vorhandenen Bodenpartikel, die die mikroskopische Betrachtung verfälschten, nicht durchgeführt werden.

Um die Konzentration der Biomasse zu maximieren, wurde das sedimentierte Gemisch aus Biomasse und restlichen Bodenbestandteilen (ca. 1 g) in möglichst geringem Volumen Bodenpuffer (ca. 5 ml), an der Grenze der Löslichkeit, aufgenommen (siehe E.2.2.1).

### 1.2. Einbetten der Biomasse in Agarose-Nudeln

Um hochmolekulare DNA isolieren zu können, ist es notwendig, Scherkräfte, die durch noch vorhandene Bodenpartikel verstärkt auftreten, während der Extraktion zu minimieren. Um dies zu ermöglichen, wurde eine Technik optimiert, die häufig bei der Isolierung von hochmolekularer DNA aus Zellkulturen verwendet wird (Choi und Wing, 2000). Die Zellen werden hierbei in Agarose eingebettet und die Zellyse wird in sogenannten Agarose-Nudeln mit anschließender Dialyse durchgeführt. Im Gegensatz zur Zellyse in wässrigen Lösungen werden dadurch starke Scherkräfte, die durch das Pipetieren und Zentrifugieren bedingt sind, umgangen, wodurch die Isolierung von DNA im Größenbereich von mehreren Megabasenpaaren möglich wird. Bei DNA-Extraktionstechniken, die auf das Einbetten der Zellen in Agarose verzichten, sind Nukleinsäuren starken Scherkräften ausgesetzt (Lottspeich und Zorbas, 1998) und können nur bis maximal 150 kb isoliert werden (eigene Beobachtung).

Die vorliegende Suspension, bestehend aus der mikrobiellen Fraktion und restlichen Bodenbestandteilen, wurde in Agarose eingebettet, indem die Suspension abweichend von Standardprotokollen mit 4%iger Agarose im Verhältnis 7:1 vermischt wurde (Abbildung 6).



Abbildung 6: Agarose-Nudeln der mikrobiellen Fraktion aus der Bodenprobe RUD, die aus der Spritze in Puffer ausgedrückt wurde (Vordergrund). Bodenprobe RUD (Hintergrund) und DNA-Lösung (Reaktionsgefäße).

Durch den relativ geringen Volumenanteil an Agarose konnte die Konzentration der Biomasse maximiert werden. Das Gemisch, das eine Endkonzentration von 0,5% Agarose

se enthielt, wurde in sterilen 1 ml Spritzen aufgezogen und nach Polymerisierung auf Eis in den Zelllysepuffer ausgedrückt. An der braunen Farbe der Agarose-Nudeln war die noch starke Kontamination von Humusbestandteilen erkennbar.

### 1.3. Lyse von Zellen in Agarose-Nudeln

Die in den Agarose-Nudeln vorliegenden Zellen wurden nacheinander mit Lauroylsarcosin, Lysozym und Proteinase K aufgeschlossen. Im Vergleich zu Standardprotokollen, bei denen zur Lyse von Zellen aus Zellkulturen in Agarose-Nudeln bereits 0,1 mg/ml Proteinase K ausreichen (Choi und Wing, 2000), mussten bei der verwendeten Bodenprobe hohe Enzymkonzentrationen eingesetzt werden (1 mg/ml Lysozym, 2 mg/ml Proteinase K), um eine effiziente Lyse zu erhalten. Dies könnte darauf zurückzuführen sein, dass die noch vorhandenen Humin- und Fulvinsäuren die Zugänglichkeit der Enzyme zu den Zellen behinderten und zusätzlich hemmend auf die Enzyme wirkten.

Die Größe der DNA und die Effizienz der Zellyse wurden in einer analytischen Pulsfeldgelelektrophorese untersucht (Abbildung 7).

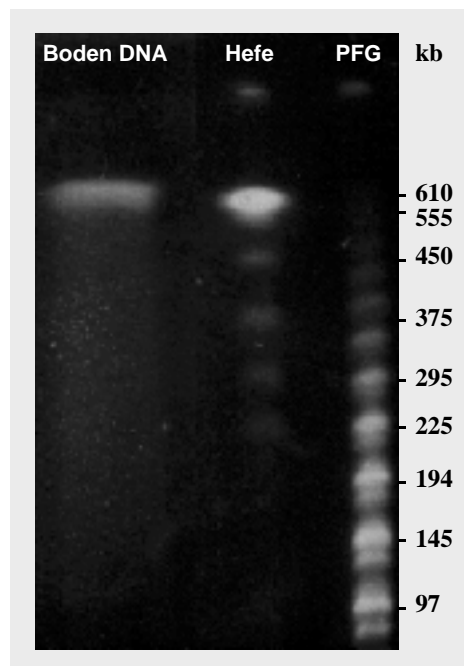


Abbildung 7: Pulsfeldgelelektrophorese von hochmolekularer DNA aus der Bodenprobe RUD: Boden-DNA aus Agarose-Nudeln; Hefe: Hefe-Chromosomen-Marker; PFG: *Mid Range PFG* Marker I; PFGE-Bedingungen: 200 V, 20 h, 10°C, Pulsfrequenz 5-50 s.

#### 1.4. Reinigung von hochmolekularer DNA aus Bodenproben: das 2-Phasengel

Die Kombination aus Aufschlämmen der Bodenprobe in Bodenpuffer mit  $\epsilon$ -Aminocaprinsäure, Einbetten der Zellen in Agarose-Nudeln und Verwendung von Lysozym sowie Proteinase K in hohen Konzentrationen bei der Zelllyse hatte es ermöglicht, DNA aus Bodenproben der Größe von über 600 kb zu erhalten (Abbildung 7).

Es wurden unterschiedliche Versuche durchgeführt, diese DNA mit Restriktionsenzymen zu hydrolysieren. In einem Protokoll wurde versucht, die DNA direkt in den Agarose-Nudeln zu hydrolysieren, und damit die Manipulierbarkeit zu überprüfen. Diese Vorgehensweise wird bei der Klonierung von hochmolekularer DNA aus Zellkulturen erfolgreich angewendet (Choi und Wing, 2000). Auch bei Einsatz von sehr großen Mengen an Restriktionsenzym (100 U *Hind* III, 50 U *Sau* 3A) konnte keine Hydrolyse der DNA festgestellt werden. In einem weiteren Protokoll wurde versucht, die DNA, die über ein Agarosegel aufgereinigt wurde, direkt in ausgeschnittenen Gelstücken zu hydrolysieren. Dabei konnte erst bei Einsatz von 200 U *Hind* III (für ca. 1  $\mu$ g DNA) eine unvollständige Hydrolyse detektiert werden. Auch die Hydrolyse von DNA, die nach Reinigung über ein Agarosegel aus dem Gelstück extrahiert wurde, war nicht erfolgreich. Bei Betrachtung der Agarosegele vor und nach der Elektrophorese konnte an einem braunen Schmier die Komigration von Humin- und Fulvinsäuren beobachtet werden (Abbildung 8), woraus man schließen konnte, dass diese die Aktivität der Restriktionsenzyme beeinträchtigt hatten.

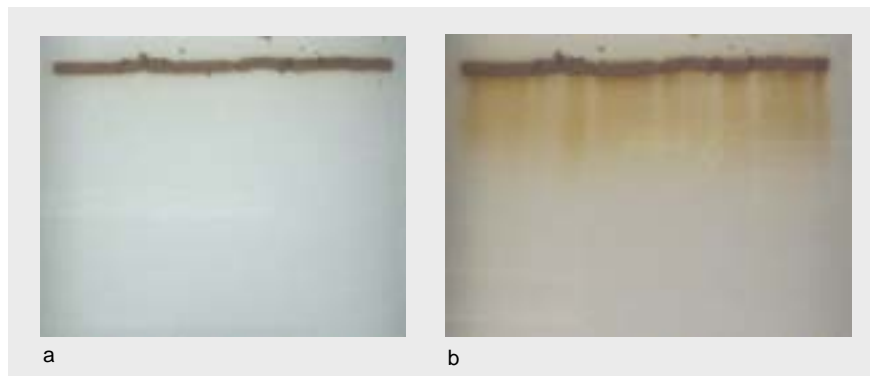


Abbildung 8: Agarosegel (ungefärbt) vor (a) und nach (b) der Gelelektrophorese mit in Agarose-Nudeln eingebetteter DNA aus Bodenproben. Nach der Gelelektrophorese war die Komigration der Humin- und Fulvinsäuren zu erkennen.

Polyvinylpyrrolidon wird häufig eingesetzt, um Huminsäuren oder andere Polyphenole abzutrennen. Young *et al.* (1993) konnte zeigen, dass durch den Einsatz von Polyvinylpyrrolidon (PVP) in einem Agarosegel Humin- und Fulvinsäuren soweit selektiv

abgetrennt werden, dass die erhaltene DNA durch PCR amplifizierbar ist (Young *et al.*, 1993).

Bei Anwendung dieser Methode auf unsere Bodenproben musste allerdings festgestellt werden, dass das PVP selbst auch inhibierende Wirkung auf Restriktionsenzyme zeigte (nicht gezeigt). Somit musste eine Möglichkeit gefunden werden, auch das PVP quantitativ zu eliminieren und dabei die DNA möglichst geringen Scherkräften auszusetzen.

Hierzu wurde ein 2-phasiges Gelelektrophoreseverfahren entwickelt (Quaiser *et al.*, 2002b). Das 2-Phasengel bestand aus einer oberen Phase, die neben 1% Agarose auch 2% Polyvinylpyrrolidon enthielt, und einer unteren Phase, die nur 1% Agarose enthielt (Abbildung 9).

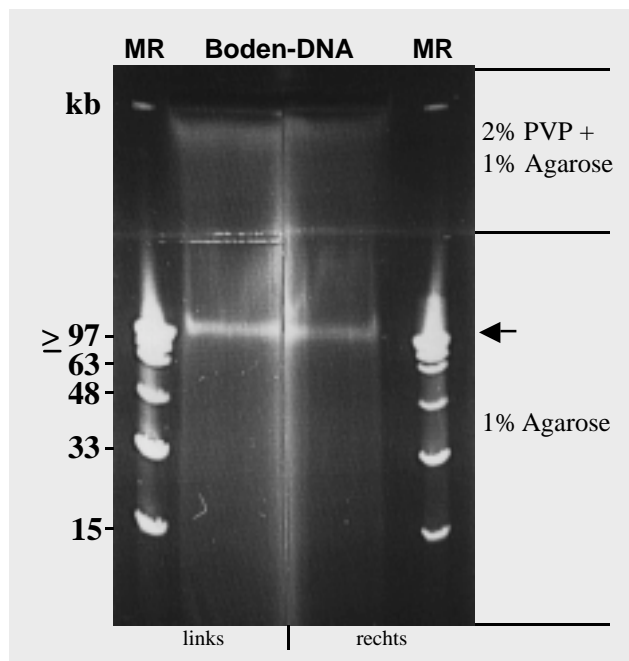


Abbildung 9: DNA der Bodenprobe RUD im 2-Phasengel nach Pulsfeldgelelektrophorese. DNA der beiden äußeren Seiten des Gels wurden aneinandergelegt und sichtbar gemacht (links und rechts). MR: *MidRange PFG* Marker I, Boden-DNA: DNA aus Bodenproben aus Agarose-Nudeln; Pfeil: Kompressionszone, DNA  $\geq 97$  kb; PFGE-Bedingungen: 4,5 V/cm, 1-4 s Puls; 36 h, 10°C. Das obere Drittel des Gels enthält 1% Agarose + 2% PVP, die unteren zwei Drittel bestehen nur aus 1% Agarosematrix.

Es wurden 3 Agarose-Nudeln nebeneinander in die obere Phase mit Agarose und PVP eingegossen. Die untere Phase wurde gegossen, nachdem die obere Phase polymerisiert war und die Kammerabtrennung entfernt werden konnte. Ein gleichmäßiges Laufverhalten der DNA während der Pulsfeldgelelektrophorese konnte bei nicht maximaler



Spannung (4,5 V/cm) und langer Laufdauer (36 h) erreicht werden. Das PVP verringerte selektiv das Ladungs-Massen-Verhältnis der Humin- und Fulvinsäuren, wodurch sie während der Gelelektrophorese von der DNA abgetrennt wurden. Die Bedingungen für die Pulsfeldgelelektrophorese wurden so gewählt, dass die DNA ca. 2 cm unter der Phasengrenze im Größenbereich von  $\geq 97$  kb der PVP-freien Phase komprimiert wurde (Abbildung 9).

Um die DNA nicht durch UV-Bestrahlung zu schädigen, wurden nur die beiden äußeren Seiten des Gels, auf denen sich auch die Marker befanden, mit Ethidiumbromid gefärbt und unter UV-Bestrahlung sichtbar gemacht (Abbildung 9). Während der Bestrahlung wurden die Bereiche der Kompressionszone, in der die DNA Fragmente von  $\geq 97$  kb komprimiert vorlagen, mit einem Skalpell markiert. Anschließend wurde das Gesamtgel wieder zusammengelegt und die Kompressionszone wurde anhand der Markierungen auf dem unbestrahlten Bereich des Gels ausgeschnitten.

In einer analytischen Pulsfeldgelelektrophorese konnte gezeigt werden, dass die im 2-Phasengel bei 97 kb komprimierte DNA, immer noch eine Größe von bis zu 290 kb aufwies (Abbildung 10).

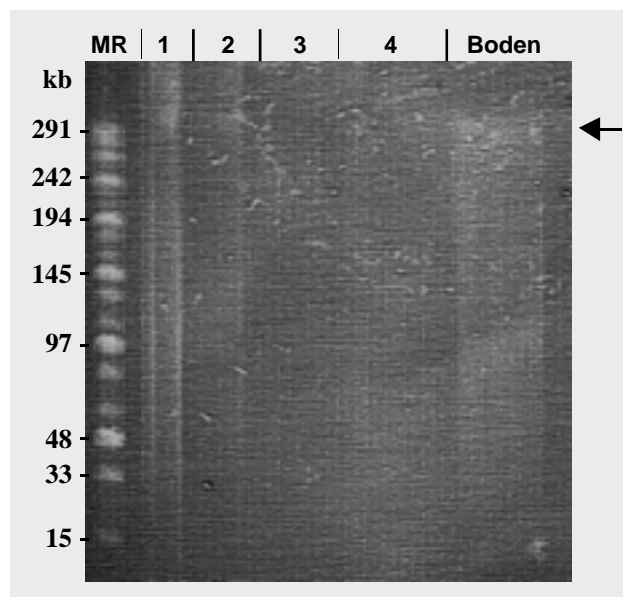


Abbildung 10: Analytisches Pulsfeldgel zur Größenbestimmung der in der Kompressionszone des 2-Phasengels ausgeschnittenen DNA. MR: *MidRange PFG* Marker I; 1-4: DNA aus anderen Untersuchungen; Boden: DNA aus Bodenproben aus der Kompressionszone des 2-Phasengels bis 290 kb groß (Pfeil); PFGE-Bedingungen: 6V/cm, 1-30 s Puls, 20 h, 10°C.

#### **1.4.1. Isolierung von hochmolekularen DNA Fragmenten aus Agarosegelen**

Um die hochmolekulare DNA aus Agarosegelen zu isolieren, wurden zwei unterschiedliche Verfahren getestet. Beim ersten Verfahren wurde das Enzym Gelase verwendet, wodurch die DNA nach Hydrolyse der Agarose freigelegt wurde. Ein Großteil der DNA verblieb dabei allerdings in nicht-hydrolysierbaren Gelresten, wodurch die Ausbeute wesentlich verringert wurde.

Bei der Etablierung der Elektroelution der hochmolekularen DNA aus Gelstücken, dem zweiten Verfahrens, wurde festgestellt, dass mit Huminsäuren kontaminierte DNA nur mit sehr geringen Ausbeuten auch unter PFGE-Bedingungen bei großer Feldstärke (6V/cm) elektroeluiert werden konnte. Auch DNA, die aus PVP-haltigen Gelstücken elektroeluiert wurde, konnte nur mit geringer Effizienz extrahiert werden. Da anfangs die Konzentrationen gereinigter hochmolekularer DNA in den Agarosegelen gering waren, wurden sehr große Gelvolumina (ca. 2x1x10 cm) verwendet, wodurch sich die Migrationsstrecke der DNA während der Elektroelution vergrößerte. Dabei wurde eine verstärkte Komprimierung der Nukleinsäuren, die nicht eluiert werden konnten, am Ende des Gelstücks beobachtet. Durch den Einsatz des 2-Phasengels konnte diese Methode allerdings verbessert werden. DNA, die aus dem 2-Phasengel stammte, war einerseits von PVP und Huminsäuren befreit und lag andererseits in ausreichender Konzentration vor, so dass die Gelvolumina und damit die Migrationsstrecke reduziert werden konnte. Trotz der noch vorhandenen Verluste an DNA, die in den Gelstücken zurückblieb, konnten abweichend von Standardprotokollen für niedermolekulare DNA unter geringer Feldstärke und langer Elektroelution (1,5 V/cm für 15 h bei Raumtemperatur) gute DNA-Ausbeuten erhalten werden. Die erhaltene DNA wurde nach Konzentrierung durch Zentrifugation gelelektrophoretisch aufgetrennt und die Ausbeute bestimmt (Abbildung 11). Insgesamt konnten ca. 3 µg hochmolekulare gereinigte DNA nach der Präparation aus 3 Agarose-Nudeln, die aus ca. 50 g Boden entstanden waren, isoliert werden.

## 2. Erstellung und Charakterisierung von Metagenombanken mit hochmolekularen DNA-Inserts

Bei der Klonierung von hochmolekularer DNA werden vor allem Fosmide und *Bacterial Artificial Chromosome* Vektoren (BAC) verwendet. Beide Vektoren enthalten den Replikationsursprung des F-Plasmids von *E. coli* und zeichnen sich durch eine hohe Stabilität in *E. coli* aus, die durch die niedrige Kopienzahl (1-2 pro Zelle) unterstützt wird. BAC-Vektoren haben den Vorteil, dass sie bis zu 300 kb Fremd-DNA aufnehmen können (Shizuya *et al.*, 1992). Fosmid-DNA hingegen wird in  $\lambda$ -Phagenköpfe verpackt, wodurch die Aufnahmekapazität durch die natürliche Größenselektion auf 32,5-43,5 kb beschränkt, aber andererseits auch gewährleistet wird. Das hier verwendete Fosmid besitzt im Gegensatz zu BAC-Vektoren in der Klonierungskassette eine glatte Restriktionsenzym-Schnittstelle, wodurch eine zusätzliche Hydrolyse, die bei Klonierungen über klebrige Enden durchgeführt werden muss, vermieden werden kann.

In Vorversuchen mit BAC-Vektoren wurden Genbanken mit genomischer DNA von *Streptomyces* spp. und *Sulfolobus solfataricus* erstellt. Dabei wurde festgestellt, dass die erhaltenen Klonzahlen mit BAC-Vektoren gering waren. Auch eine BAC-Genbank mit DNA aus Bodenproben wurde erstellt, allerdings war die durchschnittliche Insert-Größe mit 15 kb sehr klein. Hierbei hatte sich gezeigt, dass die Ausbeute von hochmolekularer DNA durch die hydrolytische Spaltung, stark reduziert wurde. In vergleichenden Untersuchungen konnte festgestellt werden, dass die erhaltenen Klonzahlen bei einem Fosmid mit einer glatten Schnittstelle wesentlich höher war, als bei den verwendeten BAC-Vektoren.

Aufgrund der höheren Klonierungseffizienz und der Gewährleistung der Insertgröße von 32,5-43,5 kb wurde die elektroeluierte und gereinigte DNA in ein Fosmid mit glatter Schnittstelle ligiert (siehe E.2.3, Abbildung 37, S.80). Die für die Klonierung notwendigen Modifizierungen der DNA, das Glätten und die Phosphorylierung der Enden wurden in einem Schritt mit Hilfe eines Enzymgemisches durchgeführt (siehe E.2.3.1). Es wurden 3  $\mu$ g DNA in die Reaktion eingesetzt, und nach Inkubation wurden die Enzymkomponenten durch eine Phenol-Chloroform-Isoamylalkohol Extraktion eliminiert. Die manipulierte DNA wurde einer 2. Größenselektion unterzogen (Abbildung 11), um kleine DNA-Fragmente zu eliminieren, die während der 1. Größenselektion zurückgehalten wurden (*trapping-effect*) oder durch Scherkräfte, während der Manipulation der Enden entstanden waren. Die Bedingungen der PFGE wurden so gewählt, dass sich DNA-Fragmente  $\geq$  50 kb in der Kompressionszone befanden. Die Kompressionszone wurde auf den beiden gefärbten Randbereichen des Gels markiert und nach Zusammensetzung des Gesamtgels die nicht bestrahlte DNA ausgeschnitten (Abbildung 11). Ein Großteil der DNA befand sich in der

Kompressionszone bei 50 kb, allerdings war auch fragmentierte DNA bis ca. 15 kb zu erkennen. Nach Elektroelution und Konzentrierung wurde die DNA zur Konzentrationsbestimmung und auch um eine eventuelle DNA-Degradation zu identifizieren, gelelektrophoretisch aufgetrennt (Abbildung 11).

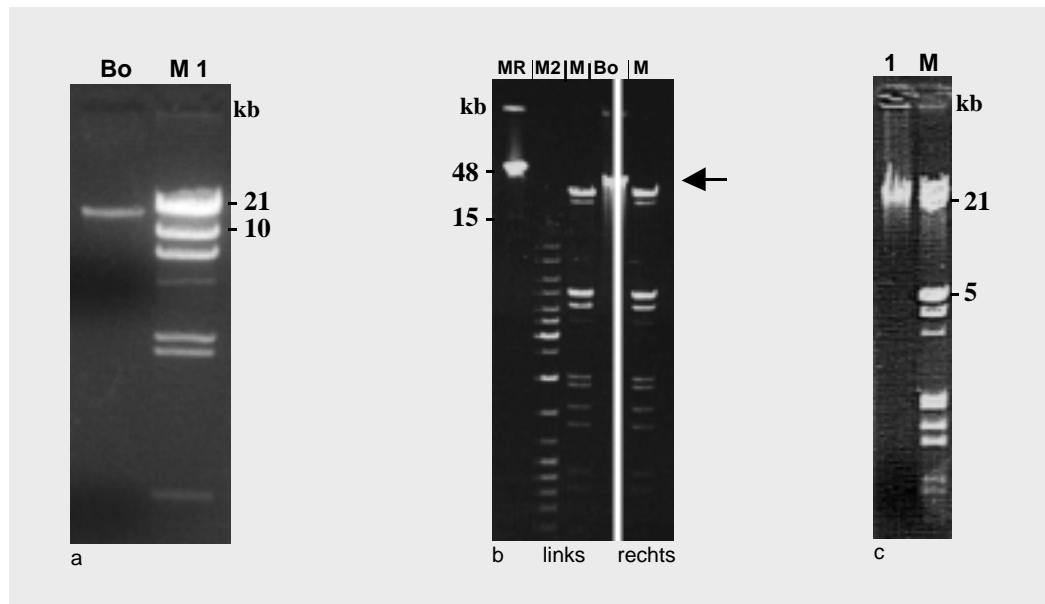


Abbildung 11: **(a)** Konzentrationsbestimmung der isolierten DNA der Bodenprobe RUD nach Elektroelution und Konzentrierung. **Bo**: 100 ng DNA der Bodenprobe; **(b)** Gefärbte Randbereiche des Pulsfeldgels der 2. Größenselektion der manipulierten DNA der Bodenprobe RUD. Der Pfeil markiert die Kompressionszone; links und rechts bezeichnen die UV-bestrahlten Randbereiche; **Bo**: DNA der Bodenprobe RUD auf beiden Randbereichen des Gels erkennbar (links, rechts). Der mittlere Teil blieb ungefärbt und DNA der Kompressionszone (Pfeil) wurde ausgeschnitten, PFGE-Bedingungen: 6 V/cm, 0,1-1s, 9 h, 10°C; **(c)** Konzentrationsbestimmung der manipulierten DNA der Bodenprobe RUD nach 2. Größenselektion. **1**: 200 ng DNA der Bodenprobe RUD, **M**: 500 ng Lambda DNA/*EcoRI*+*HindIII*. Verwendete DNA Größenstandards: **M1**: Marker, Lambda DNA/*EcoRI*+*EcoRV*; **MR**: *MidRange PFG* Marker I; **M2**: Marker, 10 kb DNA ladder (NEB).

Die erhaltene DNA zeigte keinen Abbau und die DNA-Konzentration wurde durch Vergleich zur  $\lambda$ -DNA auf ca. 200 ng/ $\mu$ l abgeschätzt, was eine Gesamtausbeute von ca. 2  $\mu$ g DNA aus 50 g Boden bedeutete.

In die folgende Ligation wurden 500 ng Insert DNA und 250 ng Vektor eingesetzt, was einem molaren Insert zu Vektor Verhältnis von ca. 1:3,5 entsprach. Anders als bei Klonierungen von niedermolekularer DNA hatte sich bei hochmolekularer DNA gezeigt, dass ein 5- bis 10-facher Vektorüberschuss zu besseren Klonierungseffizienzen führte (Choi und Wing, 2000). In diesem Fall wurde ein Mittelweg gewählt, bei dem ebenfalls gute Effizienzen erhalten wurden. Nach der Verpackung in  $\lambda$ -Phagenköpfe, Infektion von *E. coli* Zellen, Ausplattierung unter selektiven Bedingungen und Inkubation wurden aus diesem Ligationsansatz insgesamt ca. 17000 Klone erhalten. Das entsprach einer Klonie-

rungeffizienz von 68 Klonen/ng Vektor DNA. Daraufhin wurde die restliche DNA, die geglättet und phosphoryliert vorlag, unter gleichen Bedingungen kloniert.

Insgesamt wurden 25344 Klone der Genbank Rud001.AQ in 66 384er-Mikrotiterplatten geordnet abgelegt. Der Rest der Genbank wurde in Form von Ligationensätzen bei  $-70^{\circ}\text{C}$  konserviert (50000 potentielle Klone). Es hatte sich gezeigt, dass die Klonierungseffizienz durch die Lagerung des Ligationensatzes bei  $-70^{\circ}\text{C}$  nicht beeinträchtigt wurde. Dies bedeutete, dass man "Genbanken" auch in Form von Ligationensätzen konservieren konnte und dadurch eine Genbank erst niederlegen muss, wenn sie benötigt wird.

Um die Reproduzierbarkeit der entwickelten Technik zu bestätigen, wurden die Schritte der Reinigung von DNA aus der Bodenprobe des Rotböhl durch das 2-Phasengel bis zur Klonierung in das Fosmid wiederholt. Abgesehen von unterschiedlichen DNA-Ausbeuten bei der DNA Präparation konnten vergleichbare Klonzahlen und Klonierungseffizienzen erhalten werden (nicht gezeigt). In Abbildung 12 ist die entwickelte Methode in ihrer Gesamtheit dargestellt.

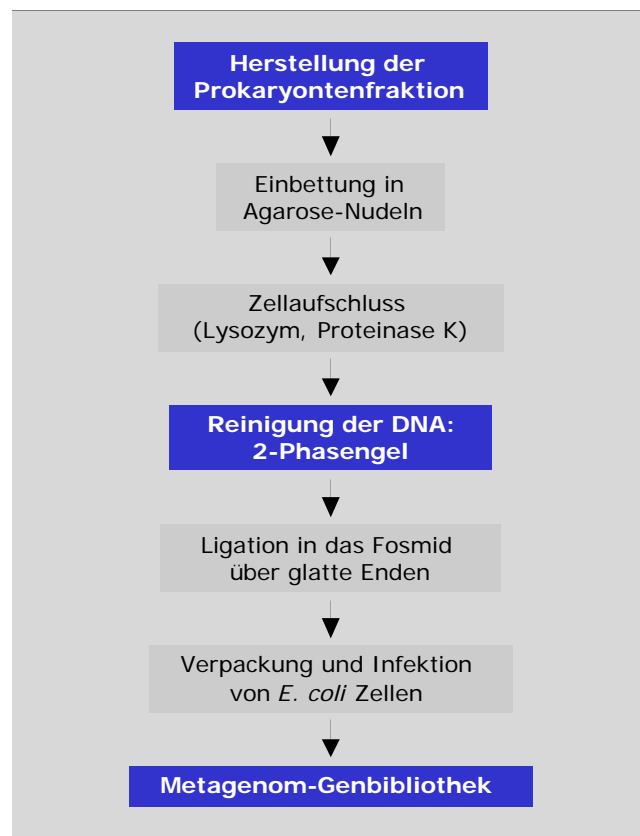


Abbildung 12: Flussdiagramm des entwickelten Verfahrens zur Isolierung und Klonierung hochmolekularer DNA aus Bodenproben.

## 2.1. Restriktionsanalyse von Fosmid-Klonen der Genbank Rud001.AQ

Um die Qualität der Genbank Rud001.AQ der Bodenprobe RUD abzuschätzen, wurde eine Restriktionsanalyse von 30 zufällig ausgewählten Einzelklonen durchgeführt. Gereinigte Plasmid-DNA wurde hierzu mit dem Restriktionsenzym *Not* I geschnitten und durch Pulsfeldgelelektrophorese aufgetrennt (Abbildung 13). *Not* I hat die Erkennungssequenz GC/GGCCGC. Der Vektor enthält zwei *Not* I Schnittstellen, an den Positionen 1 und 643, die die Klonierungskassette flankieren. Aufgrund dessen waren bei einer Restriktion mit *Not* I, neben Fragmenten des Inserts plus 643 bp des Vektors, eine Vektorbande der Größe von ca. 6900 bp zu erwarten.

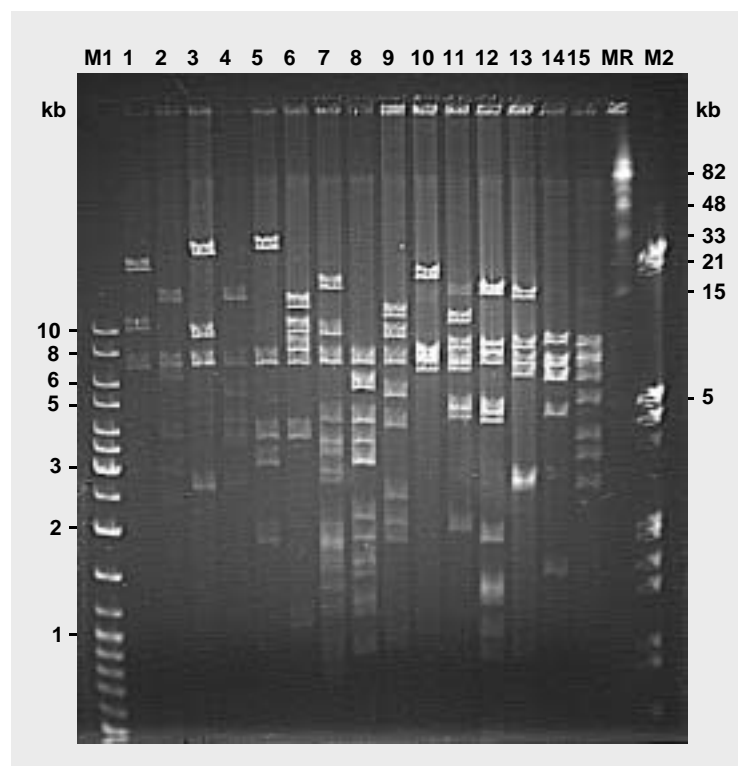


Abbildung 13: Restriktionsfragmente von 15 zufällig ausgewählten Fosmid-Klonen der Genbank Rud001.AQ nach Restriktion mit *Not* I. M1: 10 kb DNA Ladder (NEB); 1-15: Klone 1-15; MR: *PFG-MidRange* Marker I (NEB); M2:  $\lambda$ -DNA/*Eco*RI+*Hind*III. PFGE-Bedingungen: 6 V/cm, 0,1-2,5s, 6h, 10°C.

Die Analyse der Restriktionsfragmente zeigte, dass kein Restriktionsmuster der 30 untersuchten Klone identisch war. In allen Fällen war die Vektorbande bei ca. 7 kb zu erkennen. Die Größen der klonierten DNA-Fragmente wurden mittels einer Eichgeraden eines bekannten Größenstandards bestimmt. In 15 unterschiedlichen Klonen ergab sich eine durchschnittliche Insertgröße von 38,5 kb, wobei das kleinste Insert 33,5 kb und

das größte Insert 42,5 kb betrug. Die auffallend große Anzahl von *NotI*-Fragmenten bei vielen der Inserts ließ auf einen hohen GC-Gehalt der klonierten DNA schließen.

Aufgrund der durchschnittlichen Insertgröße von 38,5 kb und einer Gesamtzahl von 25344 konservierten Klonen wurde die konservierte "Metagenom"- Menge in der Genbank Rud001.AQ auf 960 Mio Basenpaare geschätzt. Mit der Annahme einer durchschnittlichen Größe der klonierten Prokaryontengenome von ca. 4 Mbp wurden demnach insgesamt ca. 200-250 prokaryotische Genome in der Genbank abgelegt. Die zusätzlich in Form von Ligationsansätzen konservierten Basenpaare wurden auf 1900 Mio geschätzt.

## 2.2. Anwendung der DNA-Präparationstechnik auf weitere Bodenproben

Um die Zuverlässigkeit der entwickelten DNA-Reinigungsmethode über das 2-Phasengel zu überprüfen, wurde DNA anderer Bodenhabitats isoliert und drei weitere Genbanken aus unterschiedlichen Bodenproben wurden erstellt (Tabelle 1). Die Genbank Rud003.AT wurde ebenso wie Rud001.AQ aus der Bodenprobe RUD erstellt. Die Bodenprobe zur Erstellung der Genbank S247 stammte von einem Graslandboden (Hähnlein, Deutschland), und die Bodenprobe für die Genbank S248 von einem Mischwaldboden (Gernsheim, Deutschland). Beide Proben wurden von Robert Pieper (BRAIN AG) zur Verfügung gestellt. Die drei Proben waren anhand ihrer unterschiedlichen Brauntöne, was auf eine unterschiedliche Zusammensetzung des Humin- und Fulvinsäuregehalts hindeutete, deutlich voneinander zu unterscheiden. Untersuchungen haben gezeigt, dass in einem Graslandboden der Anteil an Humin- zu Fulvinsäuren im Verhältnis von ca. 2:1 vorliegt, wobei in Waldböden ein Verhältnis von ca. 1:2 vorherrscht (Stevenson, 1982). Die Genbank Rud003.AT wurde von Alexander Treusch und die Genbanken S247 und S248 wurden in Zusammenarbeit mit Bernd Köhler (Brain AG) und Alexander Treusch erstellt. Die Details dieser Genbanken sind in Tabelle 1 zusammengefasst.

Tabelle 1: Übersicht der Genbanken aus Bodenproben, die unter Verwendung des 2-Phasengels entstanden sind.

Genbank	Konservierte Klone #	Konservierte DNA (in Mio bp)	Habitat
Rud001.AQ	25344	960	Sandökosystem (Weiterstadt); höherer Huminsäure- als Fulvinsäuregehalt
Rud003.AT	30336	1150	Sandökosystem (Weiterstadt)
S247	11520	440	Graslandboden (Hähnlein)
S248	19968	760	Mischwaldboden (Gernsheim) höherer Fulvinsäuregehalt als Huminsäuregehalt

# Anzahl, der in Mikrotiterplatten geordnet abgelegten Einzelklone

### 2.3. Bestimmung der Diversität, der in den Genbanken klonierten DNA

Um einen Einblick in die 16S rDNA Biodiversität der vier unterschiedlichen Genbanken aus Bodenproben zu erhalten, wurden in Zusammenarbeit mit Torsten Ochsenreiter und Alexander Treusch 16S rRNA-Genanalysen der unterschiedlichen DNAs, die in die Klonierung der Genbanken Rud001.AQ, Rud003.AT, S247 und S248 eingesetzt wurden, durchgeführt. Dabei wurden 16S rDNA-Fragmente der vier verschiedenen DNA-Präparationen mit *Bacteria*-spezifischen Oligonukleotiden amplifiziert, kloniert und sequenziert. Mit Hilfe der Sequenzdaten wurden phylogenetische Rekonstruktionen vorgenommen und diese auf der Ebene der Phyla ausgewertet (Abbildung 14).

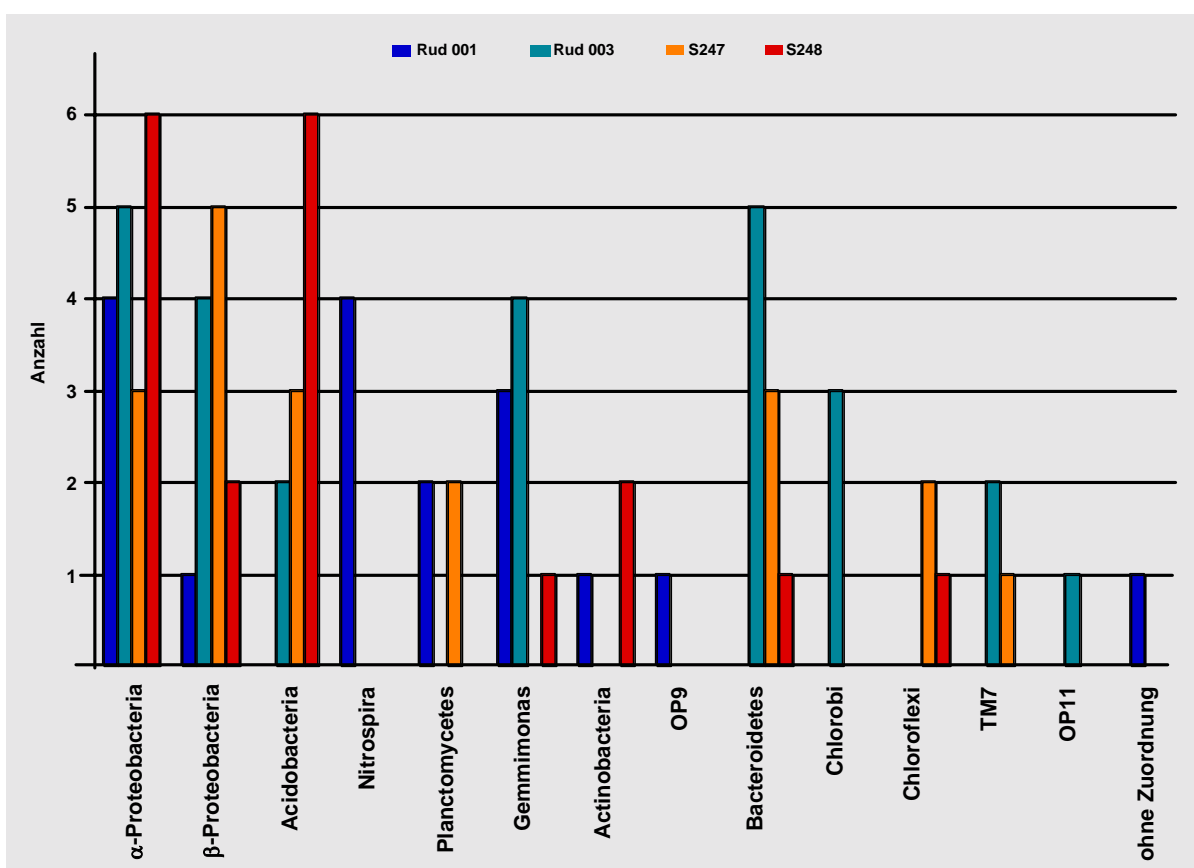


Abbildung 14: Balkendiagramm der Verteilung der 16S rDNA Diversität der Genbanken Rud001.AQ (blau), Rud003.AT (hellblau), S247 (orange) und S248 (rot) auf verschiedene bakterielle Phyla.

Insgesamt wurden aus den vier DNA-Matrizen 81 16S rDNA-Sequenzen generiert. Die phylogenetische Auswertung brachte hervor, dass 16S rDNA-Sequenzen aus jeweils 7 unterschiedlichen Phyla in den DNAs, die zur Erstellung der Genbanken Rud001.AQ und Rud003.AT eingesetzt wurden, vertreten waren (blau und hellblau). In den DNAs zur Generierung der Genbanken S247 und S248 waren jeweils 6 der insgesamt 34 unter-



schiedlichen bakteriellen Phyla vorhanden. Trotz der relativ kleinen Stichprobe von 81 Sequenzen aus den vier Präparationen konnte dadurch festgestellt werden, dass zur Herstellung der Genbanken DNA eingesetzt worden ist, die aus großer bakterieller Diversität stammte.

#### **2.4. Isolierung und Klonierung hochmolekularer DNA aus Schwamm-assozierten-Prokaryonten**

Dieses Projekt wurde in Kooperation mit der Arbeitsgruppe von Ute Hentschel (Universität Würzburg) durchgeführt. Lars Fieseler – aus der Arbeitsgruppe von Ute Hentschel - war dabei für die Bereitstellung der Schwamm-assozierten-Prokaryonten (SAPs) verantwortlich.

Die Anwendbarkeit der entwickelten Technik zur Isolierung von hochmolekularer DNA aus Bodenhabitaten wurde an einem völlig anderen Ökosystem und zwar an den marinen Schwämmen *Aplysina aerophoba* und *Axinella polypoides*, in denen ganz andere inhibitorischen Stoffe als im Boden vorherrschen, erprobt.

Die Biomasse von Schwämmen besteht aus bis zu 40% aus prokaryotischen Zellen (Hentschel *et al.*, 2001). Vor allem aber erfahren Schwämme aufgrund ihrer Fähigkeit, Naturstoffe zu produzieren besonderes Interesse. Diese Naturstoffe dienen ihnen häufig als Schutzmechanismen, in denen sie ihre Feinde zum Rückzug bewegen. So zum Beispiel wird von *Aplysina aerophoba* einem marinen Schwamm Aeroplysinin-1, ein Brom-Tyrosin Derivat, das cytotoxische Eigenschaften besitzt, gebildet (Koulman *et al.*, 1996). Es wird angenommen, dass an den Transformationsprozessen dieser Naturstoffe auch Prokaryonten beteiligt sein könnten. Die produzierten Abwehrstoffe, meist bromhaltige Alkanoide, aber auch Polysaccharide, stellen im Vergleich zu Humin- und Fulvinsäuren weitere inhibitorische Stoffe dar, die bei einer Präparation von hochmolekularer DNA eliminiert werden müssen.

Um Informationen über die assoziierten Prokaryonten zu erhalten, wurde hier der Ansatz der Umweltgenomik auf Schwamm-assozierte-Prokaryonten angewendet. Es sollten Metagenombanken von Schwamm-assozierten-Prokaryonten (SAPs) erstellt werden, um dann spezifische prokaryotische Gruppen genomisch charakterisieren zu können.

Da die Vorgehensweise bei beiden Schwämmen analog verlief, wird im Folgenden beispielhaft nur die Präparation der Prokaryontenfraction von *Aplysina aerophoba* dargestellt.

### 2.4.1. Trennung der prokaryotischen Zellen von Schwammzellen

Im ersten Schritt der DNA Aufbereitung wurden die prokaryotischen Zellen von Schwammzellen durch differentielle Zentrifugation getrennt. Dabei wurden ca. 5 mm<sup>3</sup> große Schwammstücke mit einem Skalpell in künstlichem Meerwasser (ASW) zugeschnitten. Nach Aufschlämmen der Schwammstücke und Zermalmen mit einem Mörser wurden durch zweimalige differentielle Zentrifugation die prokaryotischen Zellen von Schwammzellen aufgrund der unterschiedlichen Dichten getrennt (siehe E.2.17).

### 2.4.2. Einbetten der Biomasse in Agarose-Nudeln

Anschließend wurden die SAPs von *Aplysina aerophoba* in unterschiedlichen Konzentrationen (von 1 x 10<sup>9</sup> bis 0,05 x 10<sup>9</sup> Zellen/ml) in Agarose eingebettet und die Zellyse wie oben beschrieben unter Einsatz von hohen Enzymkonzentrationen und anschließender Dialyse durchgeführt (siehe C.1.3 und E.2.2.3). In einem analytischen Pulsfeldgel wurden die DNA-Konzentrationen und die Zellyse der unterschiedlichen Agarose-Nudeln durch elektrophoretische Auftrennung überprüft (Abbildung 15).

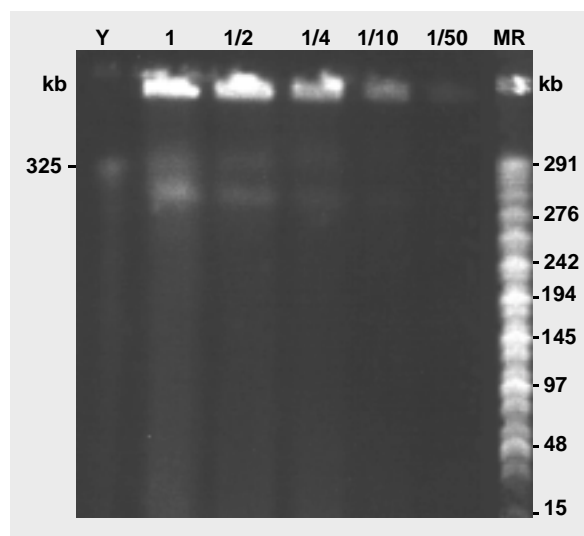


Abbildung 15: Pulsfeldgelelektrophorese der DNA aus Agarose-Nudeln mit unterschiedlichen Konzentrationen von Schwamm-assoziierten-Prokaryonten (SAPs) von *Aplysina aerophoba*. Y: Hefe-Chromosomen-Marker (NEB); 1 - 1/50: DNA aus Agarose-Nudeln der Zellkonzentrationen 1 - 0,05 x 10<sup>9</sup> Zellen/ml; MR: *Mid Range PFG Marker I*; PFGE-Bedingungen: 200 V, Pulsfrequenz 1-30 s, 24 h, 10°C.

Es wurde festgestellt, dass die Zellkonzentrationen von 1 x 10<sup>9</sup> Zellen/ml und 0,5 x 10<sup>9</sup> Zellen/ml eine gute DNA Ausbeute erlaubten. Andererseits war ein Großteil der DNA

in Agarose-Nudeln verblieben, was Hinweise auf noch vorhandene Hemmstoffe, die die Migration der DNA verhinderten oder auf einen Überladungseffekt gab. Auffallend war ausserdem, dass die Agarose-Nudeln der Zellkonzentration  $1 \times 10^9$  Zellen/ml nach der Zellyse noch trüb waren, aber die Agarose-Nudeln der Zellkonzentration  $0,5 \times 10^9$  Zellen/ml klar waren. Dies ließ den Schluss zu, dass im ersten Fall die Zellkonzentration für eine vollständige Lyse unter den verwendeten Bedingungen zu hoch war.

Um zu überprüfen, ob die DNA durch Restriktionsenzyme hydrolysierbar war, wurden ca. 20 µg der Agarose-Nudeln in einen Restriktionsansatz eingesetzt (siehe E.2.2.5) und anschließend die DNA elektrophoretisch aufgetrennt.

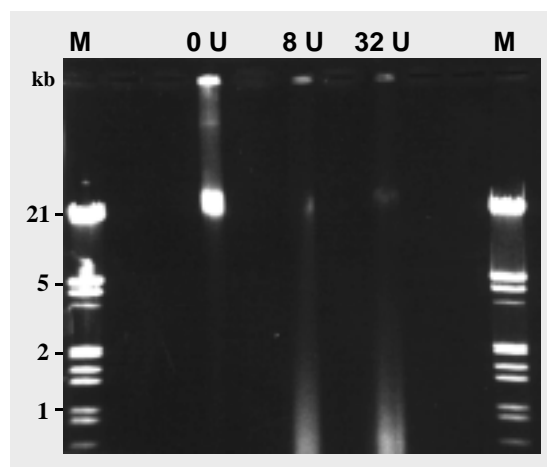


Abbildung 16: Elektrophoretische Analyse von Schwamm-assoziierten-Prokaryonten-DNA (SAPs) nach Hydrolyse in Agarosestückchen mit *Sau* 3a. M: Lambda DNA/*Eco*RI+*Eco*RV; 0U, 8U und 32U: DNA von jeweils 20 µg Agarose-Nudeln wurden mit 0U, 8U und 32U hydrolysiert.

Dabei konnte festgestellt werden, dass die DNA der noch trüben Agarose-Nudeln (mit  $1 \times 10^9$  Zellen/ml) nicht hydrolysierbar war (nicht gezeigt), die DNA der  $0,5 \times 10^9$  Zellen/ml Agarose-Nudeln hingegen war deutlich, aber nicht vollständig hydrolysierbar (Abbildung 16). Dies gab ebenfalls Hinweise auf noch vorhandene Hemmstoffe. Aufgrund dessen wurde ein weiterer Reinigungsschritt der DNA über das 2-Phasengel durchgeführt.

#### 2.4.3. Reinigung und Isolierung von hochmolekularer DNA von Schwamm-assoziierten-Prokaryonten

Die Präparation der hochmolekularen DNA von Schwamm-assoziierten-Prokaryonten (SAPs) erfolgte im Wesentlichen nach der oben beschriebenen Methode (siehe Abbildung 12, S.25). Zunächst wurden aus 50 g Schwamm (Frischgewicht) 30 Agarose-Nudeln der

SAPs von *Aplysina aerophoba* und 17 Agarose-Nudeln der SAPs von *Axinella polypoides* mit der Zellkonzentration von  $0,5 \times 10^9$  Zellen/ml hergestellt und die Zelllyse durchgeführt. Allerdings wurde nach Aufreinigung der DNA über das 2-Phasengel festgestellt (Abbildung 17), dass die isolierte DNA stark degradierte, wenn sie in Restriktionspuffern oder auch Modifikationspuffer inkubiert wurde. Daher wurden einige Abänderungen im Protokoll vorgenommen: Um vorhandene DNAsen zu eliminieren, wurde die DNA nach der Reinigung und Auftrennung im 2-Phasengel sofort für 2 h bei 6 V/cm und 4°C anstatt bei 1,5 V/cm, 15 h und bei Raumtemperatur elektroeluiert. Ausserdem wurde sofort danach eine Phenol-Chloroform-Isoamylalkohol (PCI) Extraktion durchgeführt.

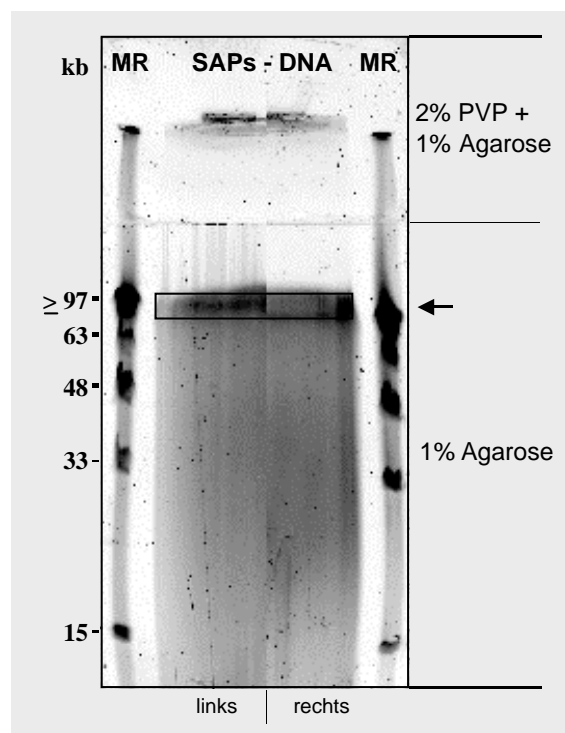


Abbildung 17: 2-Phasengel der DNA von Schwamm-assoziierten-Prokaryonten (SAPs) aus Agarose-Nudeln nach Pulsfeldgelelektrophorese: MR, *MidRange PFG* Marker I, SAPs-DNA: SAPs-DNA aus Agarose-Nudeln; Pfeil: Kompressionszone, DNA  $\geq 97$  kb; PFGE-Bedingungen: 4,5 V/cm, 1-4 s Puls; 36 h, 10°C; oberes Drittel enthält 1% Agarose + 2% PVP, die unteren zwei Drittel enthalten nur 1% Agarose. Das Gel wurde mit SyBr-Green gefärbt; DNA der beiden äußeren Seiten des Gels wurden aneinandergelegt und sichtbar gemacht (links und rechts).

Durch diese beiden abgeänderten Schritte konnte die DNA stabilisiert werden. Darauf folgte das Glätten und die Phosphorylierung der Enden der DNA Fragmente sowie die 2. Größenselektion der Nukleinsäuren (Abbildung 18). Dieser Schritt wurde analog zur oben beschriebenen Methode für Bodenproben durchgeführt (siehe C.2).

Am Schmier der DNA war deutlich zu erkennen, dass die DNA degradiert bzw. durch Scherkräfte teilweise zerkleinert wurde. Allerdings war dies vergleichbar mit isolierter DNA aus anderen Umweltproben. Der weitaus größte Anteil der DNA war in der Kompressionszone bei ca. 50 kb zu finden. Diese Kompressionszone wurde ausgeschnitten und die DNA für 2 h bei 4°C und 6 V/cm elektroeluiert. Daraufhin erfolgte die Konzentrierung der DNA und die Konzentrationsbestimmung durch gelelektrophoretische Auftrennung (Abbildung 18).

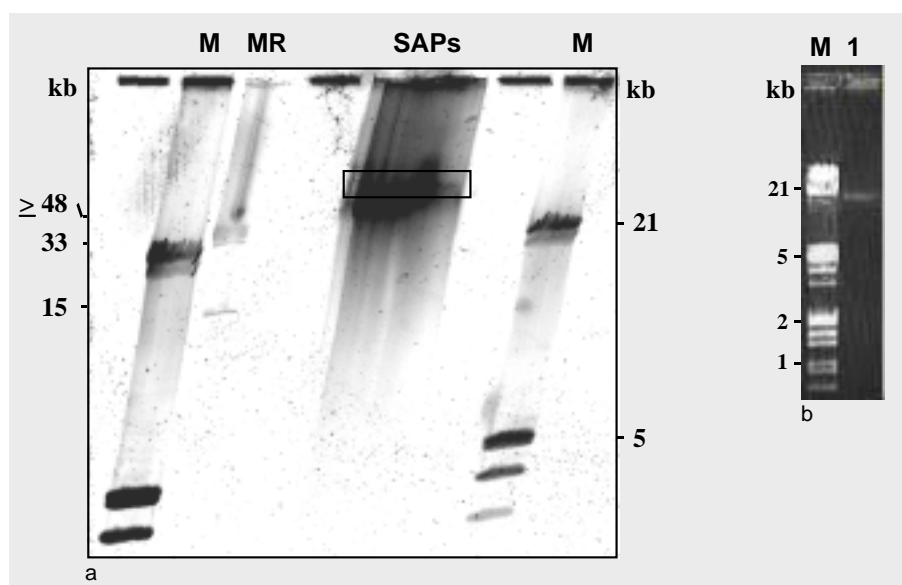


Abbildung 18: (a) Pulsfeldgelelektrophorese der manipulierten DNA aus Schwamm-assoziierten-Prokaryonten (SAPs) nach der 2. Größenselektion. Das Gel wurde mit SyBr-Green gefärbt. MR: *MidRange PFG* Marker I; SAPs: Schwamm-assoziierte-Prokaryonten DNA nach Manipulation der Enden; M: Lambda DNA/*EcoRI*+*HindIII*; PFGE-Bedingungen: 6 V/cm, 0,1-1s, 9 h, 10°C. Die inhomogene Migration der DNA war auf einen Defekt der PFGE-Apparatur zurückzuführen; (b) Konzentrationsbestimmung der SAPs-DNA nach Manipulation der Enden und 2. Größenselektion. M: 500 ng Lambda DNA/*EcoRI*+*HindIII*; 1: 30 ng SAPs-DNA

Die Konzentration wurde durch Vergleich mit dem Größenstandard auf 30 ng/µl abgeschätzt. Die Gesamtausbeute der DNA aus 3 Agarose-Nudeln, in denen ca. 5 g Schwamm (Frischgewicht) eingesetzt wurden, betrug 1,2 µg klonierbare DNA.

#### 2.4.4. Klonierung der Schwamm-assoziierten-Prokaryonten DNA

Zur Erstellung der Genbank wurde wiederum das Fosmid mit glatter Schnittstelle (siehe Abbildung 37, S.80) verwendet. In die Ligation mit DNA aus SAPs von *Aplysina aerophoba* wurden 300 ng Insert-DNA und 500 ng Vektor eingesetzt. Das entspricht einem molaren Verhältnis von 11:1. Die Klonierungseffizienz betrug 36 Klone/ng Vektor.

Insgesamt wurden 29568 Klone in 77 384er Mikrotiterplatten konserviert (siehe Tabelle 2).

Tabelle 2: Zusammenfassung der Genbanken aus Schwamm-assoziierten-Prokaryonten

Genbank	Ligation*	Konservierte Klone <sup>#</sup>	Konservierte DNA (in Mio bp)	Umweltprobe
ApAe002.AQ		29568	1120	<i>Aplysina aerophoba</i>
Axpo004.AQ	25440			<i>Axinella polypoides</i>

\* im Ligationsansatz konservierte, theoretisch vorhandene Anzahl der Klone;

<sup>#</sup> Anzahl, der in Mikrotiterplatten geordnet, abgelegten Klone

Somit wurden ca. 5 g Schwammmaterial (Frischgewicht) eingesetzt, um 1,2 µg klonierbare DNA zu erhalten, woraus insgesamt 29568 Klone hervorgegangen sind. Die durchschnittliche Insert-Grösse der Klone der Genbank Apae002.AQ wurde nach Restriktion mit *NotI* und Vergleich mit einem Standard auf 38 kb bestimmt. Somit wurden in der Genbibliothek der Prokaryontenfraktion von *Aplysina aerophoba* 1,12 Gbp des genomischen Gehalts konserviert.

Bei der Ligation der prokaryotischen DNA von *Axinella polypoides* wurden 350 ng Insert und 500 ng Vektor in die Ligation eingesetzt, was einem molaren Vektor zu Insert Verhältnis von 10:1 entspricht. Die Klonierungseffizienz betrug bei den SAPs von *Axinella polypoides* 51 Klone/ng Vektor. Die Genbank Axpo004.AQ wurde bisher nicht geordnet abgelegt und bestand theoretisch aus 25440 Klonen, die in Form des Ligationsansatzes bei -70°C konserviert wurde.

### **3. Identifizierung und Analyse von archaealen und acidobakteriellen Genomfragmenten aus Metagenombanken**

Um erste Einblicke in die Genome der mesophilen nicht-kultivierten Crenarchaeota und der Acidobacteria, die in Bodenhabitaten vorkommen zu erhalten, wurden Klone, die in der Insert-DNA ein spezifisches 16S rRNA-Gen tragen identifiziert. Dabei wurden jeweils 384 Klone einer Mikrotiterplatte der Genbanken nach Inkubation auf Selektionsmedium abgeschwemmt und die Fosmid-DNA isoliert (siehe E.2.5.). Jeweils 1 µl DNA der entstandenen Klonpools wurden anschließend als Matrizen-DNA in die Multiplex-PCR, die mit Archaea- oder Acidobacteria-spezifischen Oligonukleotiden durchgeführt wurde, eingesetzt.

#### **3.1. Identifizierung von archaealen Fosmid-Klonen**

Mit dem Archaea-spezifischen Oligonukleotid-Paar Arch20F-Arch958R (DeLong, 1992) wurden 66 Mikrotiterplatten der Genbank Rud001.AQ und 30 Mikrotiterplatten der Genbank S247 mittels Multiplex-PCR durchgemustert. Als Positiv-Kontrolle wurde chromosomale DNA von *Sulfolobus solfataricus* in verschiedenen Konzentrationen benutzt. Die Null-Kontrolle wurde mit bi-deionisiertem Wasser durchgeführt. Die aus der Genbank Rud001.AQ und S247 erhaltenen PCR-Produkte wurden auf Agarosegelen elektrophoretisch untersucht. Beispielhaft ist die elektrophoretische Auftrennung einiger der PCR-Produkte der Genbank Rud001.AQ gezeigt (Abbildung 19).

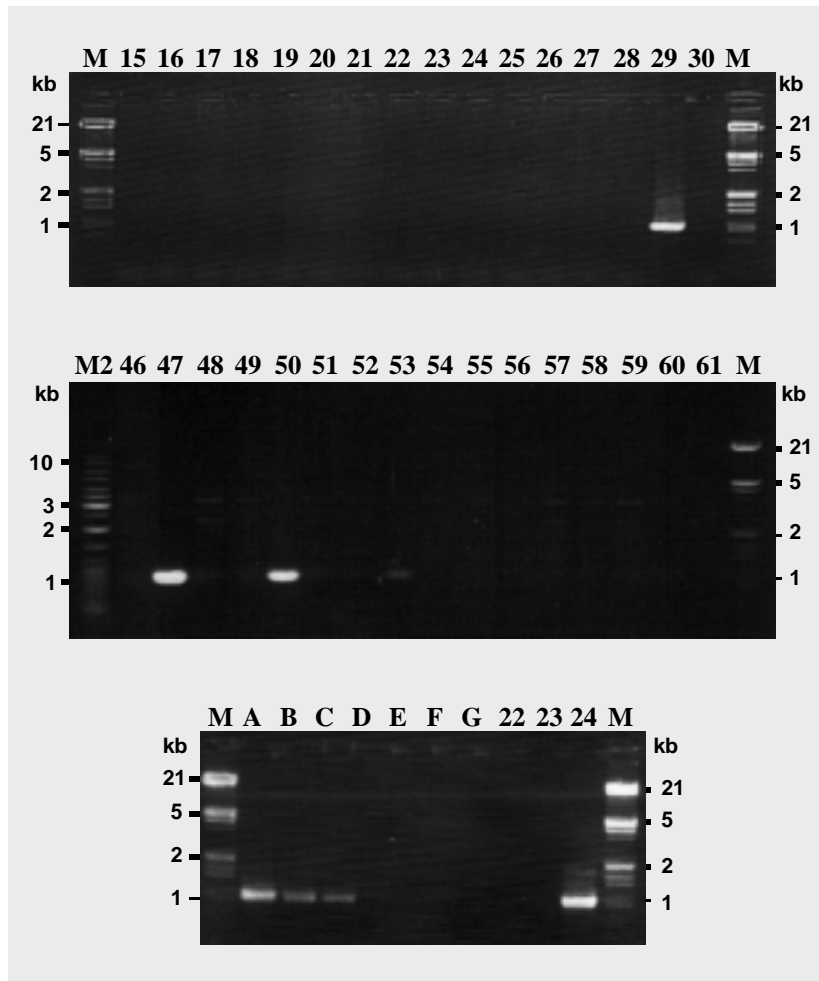


Abbildung 19: Gelelektrophoretische Auftrennung von PCR-Produkten aus der 16S rDNA-Durchmusterung der Genbank Rud001.AQ mit Archaea-spezifischen Oligonukleotiden. Die Pools 15-30 sowie 46-61 sind gezeigt. M: Lambda DNA/*EcoRI*+*HindIII*; M2: 10 kb DNA Ladder (New England Biolabs); Kontrollen: unterschiedliche Mengen chromosomaler DNA von *Sulfolobus solfataricus*, A = 100ng, B = 20ng, C = 2ng, D = 200pg, E = 50pg, F = 10pg; G = Nullkontrolle.

Amplifikate der erwarteten Größe von ca. 940 bp konnten in den Pools 29, 47, 50 und schwach auch in Pool 53 der Genbank Rud001.AQ erhalten werden. Die Positivkontrollen zeigten Amplifikate der richtigen Größe mit abnehmender Intensität bis zu einer eingesetzten Minimalmenge von 200 pg DNA.

In der Genbank S247 konnten in 6 Pools (7, 9, 11, 12, 27, 29) Amplifikate der erwarteten Größe von ca. 940 bp erhalten werden (nicht gezeigt). Diese Genbank wurde bisher nicht weiter analysiert.

Die erhaltenen Amplifikate der Pools 29, 47, 50 und 53 der Genbank Rud001.AQ wurden zur Herstellung von Klonpool-spezifischen Sonden verwendet und anschließend in die Filterhybridisierung eingesetzt. Beispielhaft ist die Filterhybridisierung des Filters



der Mikrotiterplatte 29 gezeigt (Abbildung 20). Die dabei eingesetzte Sonde wurde aus dem PCR-Amplifikat des Klonpools 29 generiert.

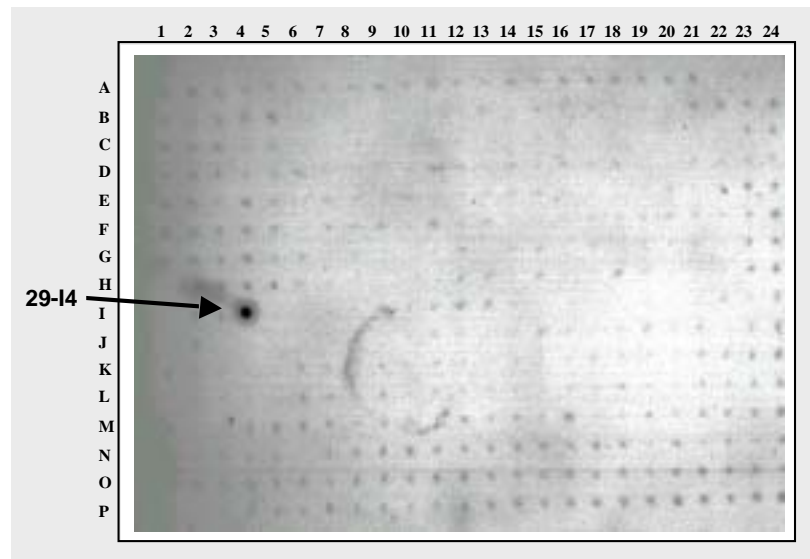


Abbildung 20: Filterhybridisierung der Mikrotiterplatte 29 der Genbank Rud001.AQ. Als Sonde diente das DIG-markierte, mit Archaea-spezifischen Oligonukleotiden erhaltene Amplifikat des Klonpools 29. Auf der Koordinate I4 ist ein positives Signal zu sehen.

Auf dem Röntgenfilm der Mikrotiterplatte 29 konnte ein positives Signal in der Kavität I4 identifiziert werden. Bei der Hybridisierung mit den Filtern 47, 50 und 53 waren jedoch keine eindeutigen Signale vorhanden (nicht gezeigt). Die Größe des Inserts des Klons 2914 wurde nach Spaltung mit unterschiedlichen Enzymen und elektrophoretischer Auftrennung anhand eines Größenstandards auf ca. 35 kb abgeschätzt.

### 3.2. Analyse des archaealen Fosmid-Klons 2914

Das Insert des Klons 2914 wurde subkloniert, vollständig sequenziert und ausgewertet. Die Subklonierung und Sequenzbestimmung wurde von Hans-Peter Klenk und dessen Mitarbeiter Arne Zibat (Epidaurus AG, Bernried) durchgeführt.

Die nachfolgend beschriebene Annotation und die Analyse wurde in unserem Labor durchgeführt. Das Insert des Klons 2914 enthielt 33925 bp und der Gesamt-GC-Gehalt betrug 40%. Auf dem Insert war ein vollständiges 16S und 23S ribosomales RNA-Operon, ein tRNA<sup>Arg</sup> Gen und 17 potentielle Protein-codierende Gene enthalten. Das 16S rRNA-Gen war zu 98,8% über 1402 Positionen identisch zu einer archaealen Sequenz (Sca1145; U62811) der Gruppe 1.1b der Crenarchaeota, die das terrestrische Cluster umfasst (nicht gezeigt, phylogenetische Analyse siehe Ochsenreiter, 2002), die aus einer PCR-Studie eines landwirtschaftlich genutzten Bodens in West Madison hervorging

(Bintrim *et al.*, 1997). Das ribosomale RNA-Operon von Klon 2914 bestand nur aus 16S und 23S rRNA-Genen ohne Verknüpfung mit 5S rRNA oder tRNA Genen, wie dies auch bei kultivierten thermophilen oder unkultivierten marinen Crenarchaeota zu finden ist. Die phylogenetische Analyse der kompletten 16S rRNA und 23S rRNA-Gene bestätigte die Zugehörigkeit des Klons 2914 zu den Crenarchaeota (phylogenetische 23S rRNA-Genanalyse siehe Abbildung 21 und 16S rRNA-Genanalyse siehe Ochsenreiter, 2002).

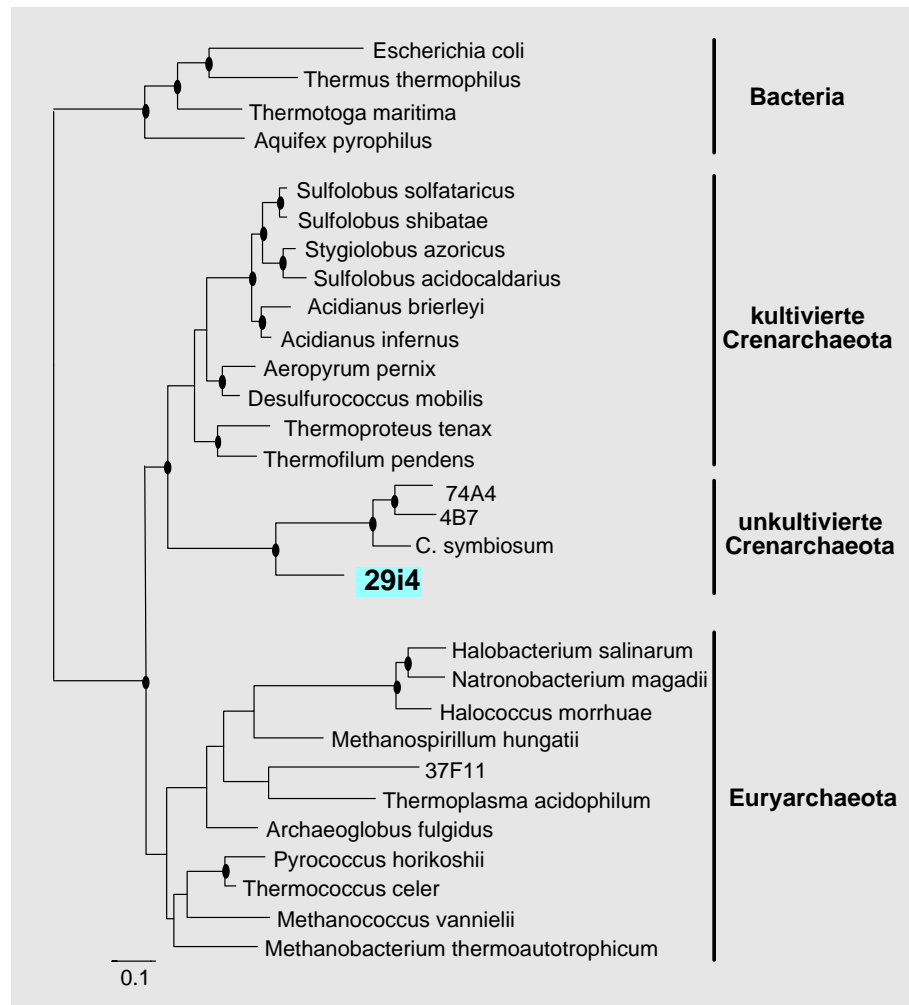


Abbildung 21: Phylogenetischer Baum basierend auf 23S rDNA Gesamtlängensequenzen von Bacteria, Archaea und Sequenzen von marinen Umweltklonen sowie 2914. Die gezeigte Baumtopologie basiert auf einer Maximum Parsimony Analyse (1048 konservierte Positionen wurden durch einen Variabilitätsfilter ausgewählt). Die ausgefüllten Ellipsen zeigen die Verzweigungspunkte an, die sich durch unterschiedliche phylogenetische Methoden und Filter bei der Rekonstruktion des 23S und dem zugehörigen 16S rDNA Baum bestätigt haben (siehe E.2.14). Die phylogenetische Rekonstruktion wurde von Torsten Ochsenreiter mit Hilfe des ARB-Programm-Pakets durchgeführt.

Die 23S rDNA-Sequenz von 2914 gruppierte mit Sequenzen von nicht-kultivierten marinen Crenarchaeota, den einzigen bisher bekannten Sequenzen von moderaten Crenarchaeota.

Zehn der 17 vorhergesagten Protein-codierenden Gene zeigten große Ähnlichkeiten zu Genen mit bekannter Funktion, zwei waren konservierte hypothetische Gene, fünf offene Leserahmen (ORFs) zeigten keine Ähnlichkeiten zu Sequenzen aus der Datenbank. Acht der abgeleiteten Aminosäuresequenzen zeigten größte Ähnlichkeiten mit archaealen Proteinen (Abbildung 22 und Tabelle 3).

## C. Ergebnisse

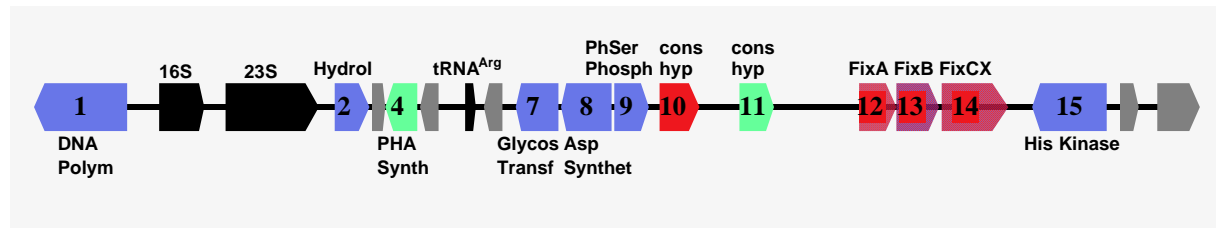


Abbildung 22: Schematische Darstellung des archaealen Fosmid-Klons 2914. Unterschiedliche Farben zeigen die phylogenetische Affinität der abgeleiteten Protein-codierenden Gene an: in Archaea (rot), Bacteria (grün), Bacteria und Archaea (rot/blau gestreift), Archaea, Bacteria und Eukaryonten (blau).

Tabelle 3: Potentielle RNA und Protein codierende Gene des archaealen Fosmids 2914

ORF	Nt Bereich	Protein-Größe	vorhergesagte Funktion	ähnlichstes Ortholog*	Phyl. aff. #	Bemerkung
01	1-2367	789 as C-termi. abgesch.	Familie B DNA-Polymerase pfam00136	AAC62689 Cenarchaeum symbiosum (0.0)	AEB	
RNA	3548-5018		16S RNA	98,8% Identität zu 16S RNA eines Umweltklons SCA1145		Paarweises Alignment über 1402 Positionen
RNA	5847-8818		23S RNA	77% Identität zu 23S RNA von Cenarchaeum symbiosum		
02	8968-9849	293 as	$\alpha/\beta$ -Hydrolase pfam00561	AAD02150 Pseudomonas stutzeri (e -17)	AEB	Kat. Triade Ser/Asp/His; nächste Homologe stammen von Bacteria
03	9888-10205	105 as	Hypothetisch	keine		
04	10219-11304	361 as	PHA-Synthase (TypIII) $\alpha/\beta$ Hydrolase, pfam00561	P45366 Thiocystis violacea (e -66)	B	Keine Homologen in Archaea, aber PHB Produkt. in Halobacteriaceae beschrieben; Typ III PHA-Synthase
05	11285-11866	193 as	Hypothetisch	keine		
RNA	12414-12502		tRNA <sup>Arg</sup>			
06	13017-13454	145 as	Hypothetisch	keine		
07	14324-15238	304 as	Glycosyltransferase Gruppe 1, pfam00534	AAK41834 Sulfolobus solfataricus (e -11)	AEB	Übertragung von ADP, UDP, GDP, CMP verknüpften Zucker
08	15716-17407	563 as	Asparagine Synthetase pfam 00310	AAB99117 Methanococcus jannaschii (e -67)	AEB	
09	17492-18157	221 as	Phosphoserin-Phosphatase pfam00702	AAB86099 Methanothermobacter thermoautotrophicus (e -14)	AEB	
10	18377-19588	403 as	Konserviert hypothetisch pfam 00535	CAB50138 Pyrococcus abyssi (e -91)	A	Homologe nur in P. abyssi, P. horikoschii, A. pernix gefunden Glycosyltransferase Gruppe 2
11	20630-21793	387 as	Transmembranprotein pfam 01173	BAB50489 Mesorhizobium loti (e -31)	B	
12	24580-25461	293 as	Fix A pfam01012	P53576 Azotobacter vinelandii (e -23)	AB	Paraloge von ETF $\beta$ : electron transfer flavoprotein $\beta$ subunit
13	25458-26741	427 as	Fix B pfam00766	P53578 Clostridium saccharobutylicum (e -45)	AB	Paraloge von ETF $\alpha$ : electron transfer flavoprotein $\alpha$ subunit
14	26738-28615	625 as	Fix CX	NP_454687 Salmonella enterica (e -36) FixX: Thermoplasma volcanium (e -08)	AB	Fusioniertes Protein aus FixC und FixX
15	29228-31465	745 as	Sensor-Transduktions-Histidinkinase, pfam00512	BAB73503 Nostoc sp. PCC 7120 (e -12)	AEB	sHPK der Familie 1a/b; cytoplasmatisch
16	32505-33023	172 as	Hypothetisches Protein	keine		
17	32918-33925	335 as (abges.)	Hypothetisch	keine		

\*Proteine sind durch ihre Proteinidentifizierungsnummer gefolgt vom Namen der Spezies bezeichnet. Die E-Werte von blastx sind in Klammern angegeben.

# Phyl. aff. = Phylogenetische Affinität, zeigt das Vorhandensein von Homologen in Archaea (A), Bacteria (B) oder Eucarya (E) an.

Eine abgeleitete Aminosäuresequenz (ORF 1) zeigte große Ähnlichkeit mit archaealen DNA-Polymerasen und teilte 46% identische Positionen mit dem nächst-verwandten Protein von *Cenarchaeum symbiosum* (Schleper *et al.*, 1997b). Die DNA-Polymerase wurde mit 11 DNA-Polymerasen von Crenarchaeota aligniert und die konservierten Bereiche wurden untersucht. Obwohl das C-terminale Ende von ungefähr 90 Aminosäuren nicht mehr auf dem Fosmid 2914 codiert war, waren auf der abgeleiteten Aminosäuresequenz alle konservierten Exonuklease- und Polymerase-Motive der Familie B DNA-Polymerasen vorhanden (Abbildung 23, Wong *et al.*, 1988; Blanco *et al.*, 1991).

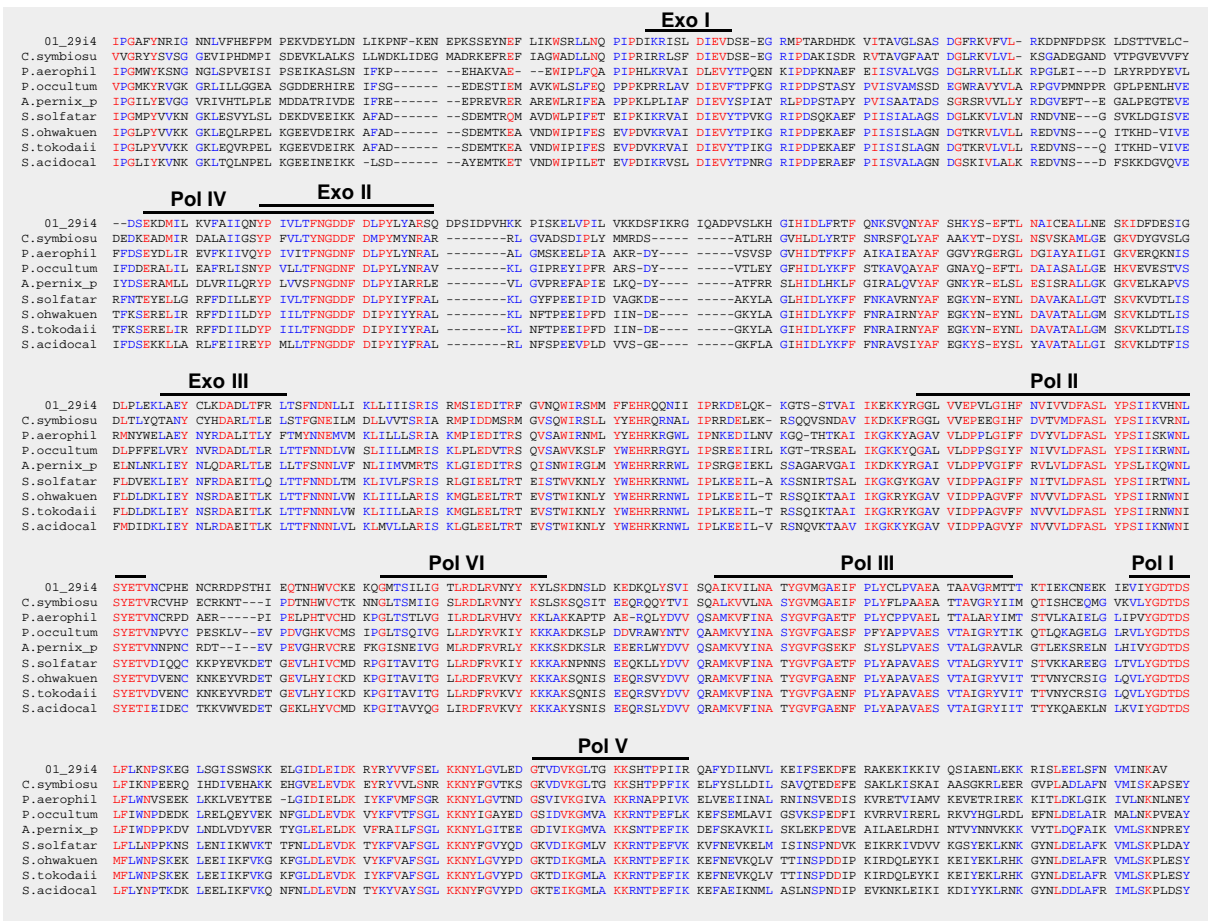


Abbildung 23: Proteinsequenzalignment von 9 archaealen Familie B DNA-Polymerasen. N-Terminale Regionen mit schwacher Konservierung wurden entfernt. Insgesamt sind 590 AS der 2914-DNA Polymerase vorhanden. N-Terminale wurden 155 Aminosäuren entfernt. Pol I-VI, konservierte Regionen der Familie B DNA-Polymerasen (Wong *et al.*, 1988; Braithwaite und Ito, 1993); Exo I-III, konservierte Motive von 3´-5´ Exonukleasen. Spezies Abkürzungen: C.symbiosu., *Cenarchaeum symbiosum*; S.solfatar., *Sulfolobus solfataricus* Pol B1; S.acidocal., *Sulfolobus acidocaldarius*; S. tokodaii, *Sulfolobus tokodaii*; P.occultum, *Pyrodictium occultum* Pol I; S.ohwakuen, *Sulfurisphaera ohwakuensis*, P. aerophil, *Pyrobaculum aerophilum*; A.pernix\_p, *Aeropyrum pernix* Pfu Polymerase.

Zwei weitere abgeleitete Proteine, die zu konservierten archaealen Proteinfamilien gehören, waren eine Asparagin-Synthetase (ORF 8) und eine Phosphoserin-Phosphatase (ORF 9), beides Enzyme, die am Aminosäurestoffwechsel beteiligt sind. Zwei vorausgesagte Glycosyltransferasen (ORF 7 und ORF 10) zeigten große Ähnlichkeiten zu Homologen des Crenarchaeoten *Sulfolobus solfataricus* bzw. des Euryarchaeoten *Pyrococcus* ssp. Hingegen zeigten eine potentielle Polyhydroxyalkanoat-Synthase (PHA-Synthase, ORF 4) und eine  $\alpha/\beta$  Hydrolase (ORF 2) große Ähnlichkeiten zu bakteriellen Proteinen. Bislang sind keine Homologen der PHA-Synthasen bei Archaea bekannt, allerdings wurde bereits PHA-Akkumulierung in Halobacteriaceae beobachtet (Kirk und Ginzburg, 1972; Fernandez-Castillo *et al.*, 1986).

Außerdem konnte ein Gencluster identifiziert werden, das große Ähnlichkeiten bezüglich seiner Anordnung und der Gensequenzen mit *fixABCX*-Operons zeigte, die in vielen symbiontischen Stickstoff-fixierenden Bodenbakterien vorkommen. Aufgrund der Ähnlichkeiten mit Komponenten der Flavoprotein-haltigen Elektronentransportkette, die an der  $\beta$ -Oxidation von Fettsäuren in Mitochondrien und einigen Bakterien beteiligt ist, wurde vorgeschlagen, dass das *fixABCX*-Operon eine Flavoprotein-haltigen Elektronentransportkette codiert (Weidenhaupt *et al.*, 1996). In symbiotischen Bakterien wird das Operon mit anderen Genen, die an der Stickstofffixierung beteiligt sind, co-reguliert (Gubler und Hennecke, 1988).

In einer vergleichenden Genomanalyse wurden *fixABCX*-Gene in Genomen von einigen Bacteria und einigen thermophilen und hyperthermophilen Archaea, wie z.B. in *Sulfolobus solfataricus*, *Thermoplasma acidophilum*, *Pyrobaculum aerophilum* und *Aeropyrum pernix*, die nicht Stickstoff fixieren, identifiziert. Die phylogenetische Analyse des potentiellen FixA Proteins von 2914 mit Homologen aus sequenzierten Genomen von Mikroorganismen zeigten eine enge Verwandtschaft von FixA des Fosmids 2914 zu den anderen archaealen Proteinen (Abbildung 24). Gemeinsam mit anderen FixA Proteinen von stickstofffixierenden Bakterien bildeten sie eine distinkte Untergruppe innerhalb der Etf $\beta$ /FixA Superfamilie.

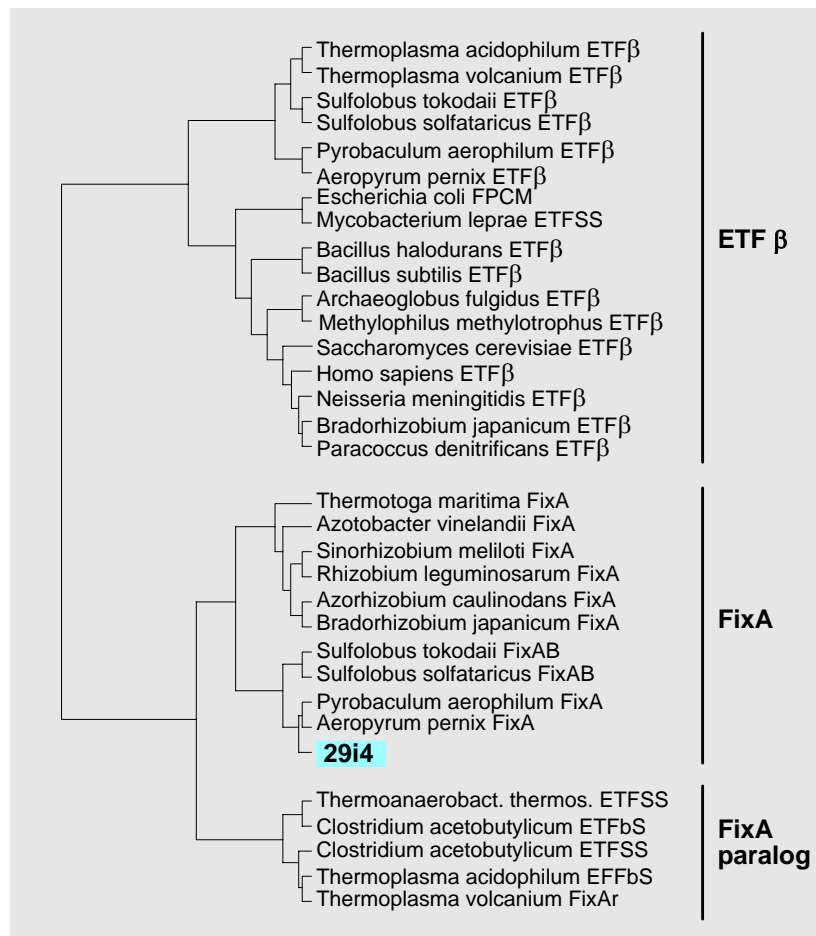


Abbildung 24: Phylogenetische Analyse mit ausgesuchten Sequenzen der Etf-ähnlichen Proteinfamilie. Homologe von FixA Proteinen von Archaea bilden eine monophyletische Gruppe mit den FixA Proteinen von Stickstoff-fixierenden Bacteria und *Thermotoga maritima*. Sie unterscheiden sich deutlich von "Haushalts"- Etf Proteinen (Etf β Homologe) der Bacteria, Archaea und Eukaryonten. Eine dritte Untergruppe (FixA Paraloge genannt) wird von einigen noch uncharakterisierten Sequenzen von Bacteria und Archaea gebildet. Die phylogenetische Rekonstruktion wurde von Hans-Peter Klenk durchgeführt (Details siehe E.2.14).

Eine Sensor-Histidinkinase konnte in der Nähe des fix-Operons auf 29i4 identifiziert werden. Während die C-terminale Hälfte des Proteins die typischen konservierten Motive von Sensor-Histidinkinasen der Superfamilie 1 a/b zeigte (Grebe und Stock, 1999), konnte in der 350 Aminosäuren langen N-Terminalen Hälfte keine Ähnlichkeiten zu bekannten Proteinen gefunden werden, die Hinweise auf eine Rolle bzgl. der Sensierung geben könnte. Aus der Tatsache, dass im Protein keine Transmembrandomänen zu fin-

den waren, konnte man schließen, dass es sich hierbei um eine cytoplasmatische Histidinkinase handelte.

Mit Oligonukleotiden, die für die Enden des Inserts von Fosmid 2914 und auch interne Gene spezifisch waren (DNA Polymerase-ORF 1, ORF 10, FixABC-ORF 12-14 und SHPK-ORF 15), konnten keine überlappenden genomischen Fragmente zu Klon 2914 in der Genbank Rud001.AQ identifiziert werden.

### **3.3. Identifizierung von Genomfragmenten aus Acidobacteria**

Um erste Einblicke in die Genome der Acidobacteria zu erhalten, wurden die Genbanken mit 16S rDNA-spezifischen Oligonukleotiden mittels Multiplex-PCR-Technik, analog wie in Kapitel C.3.1. für Archaea beschrieben, durchgemustert.

Bei einer Studie über die Verbreitung der Acidobacteria in unterschiedlichen Habitaten waren bereits eine Reihe von Acidobacteria-spezifischen Oligonukleotiden definiert worden (Barns *et al.*, 1999). Um Acidobacteria-spezifische 16S rDNA-Sequenzen zu amplifizieren wurde das Oligonukleotidpaar Ac31F-Univ1492R eingesetzt. Das Oligonukleotid Ac31F bindet spezifisch an acidobakterielle 16S rDNA, wobei das Oligonukleotid Univ1492R an 16S rDNA-Sequenzen der Prokaryonten bindet (Lane, 1991). In der vorliegenden Arbeit wurde festgestellt, dass der Einsatz dieses Oligonukleotidpaares oftmals zur Bildung von unspezifischen PCR-Produkten führte (nicht gezeigt). Andererseits konnten mit dem Oligonukleotid-Paar Ac31F-Univ533R unter Einsatz eines PCR-Zusatzes (5x GC-Solution, FastStart DNA Polymerase Zusatz, Roche) reproduzierbar Acidobacteria-spezifische PCR-Produkte erhalten werden. Das Oligonukleotid Univ533R (Reysenbach, 1995, modifiziert von Torsten Ochsenreiter) bindet wie Univ1492R spezifisch an 16S rDNA der Prokaryonten. Die optimale Anlagerungstemperatur der Oligonukleotide wurde in einer Gradienten-PCR bei 50°C identifiziert.

#### **3.3.1. Identifizierung von Klonpools mit acidobakteriellen 16S rRNA-Genen**

Die etablierten PCR-Bedingungen wurden auf Klonpools der Genbanken Rud001.AQ, Rud003.AT, S247, S248 die aus Bodenproben entstanden, sowie auf Klonpools der Genbank Apae002.AQ, die aus Schwamm-assoziierten-Prokaryonten von *Aplysina aerophoba* entstand, angewendet. Repräsentativ ist hier die Identifizierung von positiven Klonpools der Genbank Rud001.AQ gezeigt. In der Genbank Rud001.AQ wurden in 17 Pools Amplifikate der erwarteten Fragmentlänge von ca. 500 bp erhalten, wobei die Signale in Pool 17 und 63 im Vergleich zu den Anderen nur schwach waren (Abbildung 25). Außerdem wurde in Pool 24 ein unerwartet großes Produkt der Größe von ca. 800 bp erhalten. Als Kontrollen dienten DNAs, die direkt aus anderen Umweltproben extrahiert wurden (DAR,



## C.Ergebnisse

PUS und RUD; Ochsenreiter, 2002). Es hatte sich herausgestellt, dass in RUD und DAR, aber nicht in PUS Amplifikate der erwarteten Größe erhalten werden konnten.

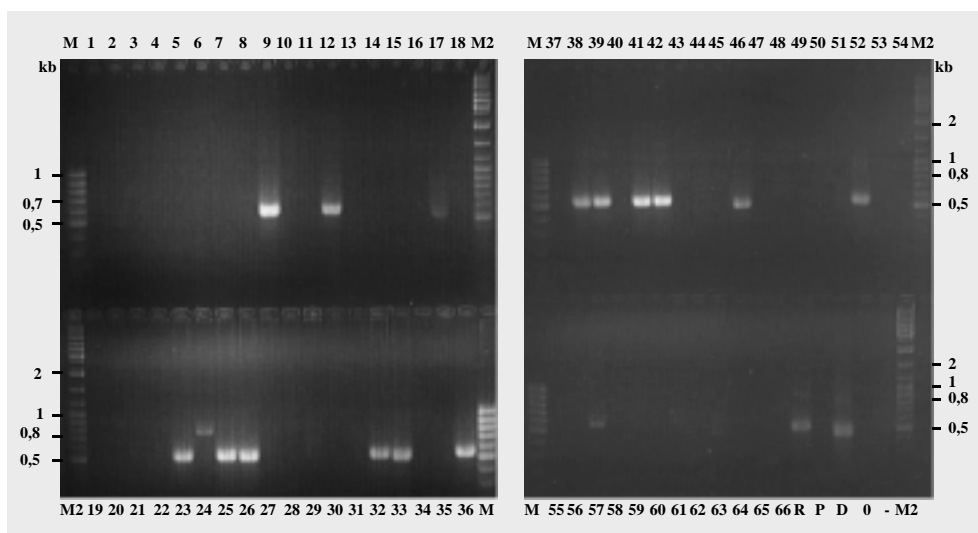


Abbildung 25: Gelelektrophoretische Auftrennung von PCR-Produkten aus der 16S rDNA Durchmusterung der Genbank Rud001.AQ mit Acidobacteria-spezifischen Oligonukleotiden. 1-66: Klonpools der Mikrotiterplatten 1-66, R = RUD und D = DAR positiv Kontrollen, P = PUS: Negativ Kontrolle; 0 = Nullkontrolle; M: 1 kb DNA Ladder (New England Biolabs); M2: 10 kb DNA Ladder (New England Biolabs).

Das Ergebnis der Durchmusterung der 5 Genbanken ist in Tabelle 4 zusammengefasst.

Tabelle 4: Zusammenfassung der Klonpools, in denen bei der Durchmusterung der Genbanken Rud001.AQ, Rud003.AT, S247 und S248, sowie Apae002.AQ mit Acidobacteria-spezifischen Oligonukleotiden ein Amplifikat der erwarteten Größe von ca. 500 bp erhalten wurde.

Genbank	Positive Pools	Verhältnis positive zu vorhandenen Pools
Rud001.AQ	9, 12, (17), 23, 25, 26, 32, 33, 36, 38, 39, 41, 42, 46, 52, 57, 63	17 von 66
Rud003.AT	9, 36, 46, 62, 72	5 von 79
S247	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 22, 29	15 von 30
S248	2, 6, 13, 16, (18), (21), 23, 25, 28, 32, 34, 39, (40), 41, 47, 49, 50	17 von 52
Apae002.AQ	1, 2, 5, 6, 7, 8, 15, 20, 21, 22, 24, 27, 28, 29, 34, 46, 52, 56, 62, 68, 69, 70, 71, 72, 75, 76, 77	27 von 77

Einzel aufgelistet sind Pools, bei denen ein Amplifikat erhalten wurde. In Klammern gesetzt sind Pools, bei denen das Amplifikat nur schwach war. Ausserdem ist das Verhältnis der positiven zur Gesamtzahl der in den Genbanken vorhandenen Pools angegeben. Auffallend war das unterschiedliche Verhältnis an positiven Pools zu Gesamt-

pools von 17/66 und 5/79 bei den Genbanken Rud001.AQ und Rud003.AT, die beide aus demselben Habitat stammten, allerdings auf unterschiedlichen DNA-Präparationsmethoden beruhten. Die DNA zur Erstellung der Genbank Rud003.AT wurde nach direkter Zellyse (Zhou *et al.*, 1996; Treusch, persönliche Mitteilung) erhalten, während die DNA zur Erstellung der Genbank Rud001.AQ nach Lyse der mikrobiellen Fraktion in Agarose-Nudeln extrahiert wurde. Dies deutete an, dass durch die beiden DNA-Präparationsmethoden auf unterschiedliche Prokaryontenphyla während des Zellaufschlusses selektiert wurde. Eine Bestätigung dieser Hypothese würde allerdings nur durch die genaue Bestimmung der vorhandenen Klone mit acidobakteriellem Insert in den Banken geben können.

### 3.3.2. Klonierung und Sequenzierung der 16S rDNA Fragmente

Um die phylogenetische Zugehörigkeit der amplifizierten 16S rRNA-Genfragmente zu den Acidobacteria zu bestätigen, wurden einige PCR-Produkte kloniert, sequenziert und phylogenetisch ausgewertet (Abbildung 26).

Die Acidobacteria werden in sechs unterschiedliche Gruppen unterteilt (Gruppe I-VI), wobei kultivierte Vertreter nur in Gruppe I (*Acidobacterium capsulatum*) und VI (*Holophaga foetida*, *Geothrix fermentans*) vorkommen (Barns *et al.*, 1999). Die Gruppe II wurde im Baum nicht dargestellt, da keine Sequenz in dieses Cluster fiel. Insgesamt wurden 16 Sequenzen in die phylogenetische Berechnung miteinbezogen. Sequenzen, die aus der Genbank Rud001.AQ generiert wurden, waren ausschließlich in der Gruppe V der Acidobacteria zu finden (Abbildung 26, rot dargestellt). Die zwei Sequenzen der Genbank Rud003.AT (P36K2 und P9K1) fielen in die Gruppen III und V (blau). Aus der Genbank S248 stammende Sequenzen verteilten sich auf die Gruppen I, IV und V (grün). Die Sequenz aus S247 (P11K2) war am nächsten verwandt zu Sequenzen der Gruppe V (lila). Auffallend war, dass 12 der 16 partiellen Sequenzen in die Gruppe V der Acidobacteria fielen. Aus Pool 25 der Genbank Rud001.AQ wurden 2 unterschiedliche Sequenzen generiert. Mit weiteren 9 Sequenzen konnten durch Datenbankvergleich (blastn) deutliche Hinweise erhalten werden, dass es sich hierbei ebenfalls um Sequenzen der Acidobacteria handelte (nicht gezeigt). Durch die phylogenetische Analyse konnte bestätigt werden, dass alle mit Ac31F-Univ533R amplifizierten 16S rRNA-Genfragmente zum Phylum der Acidobacteria gehörten und dass damit 16S rDNA Acidobacteria-spezifische Amplifikate erhalten werden konnten. Insgesamt konnten aus 24 unterschiedlichen Pools der 4 Genbanken aus Bodenproben 16S rDNA-Sequenzen der Acidobacteria generiert werden.

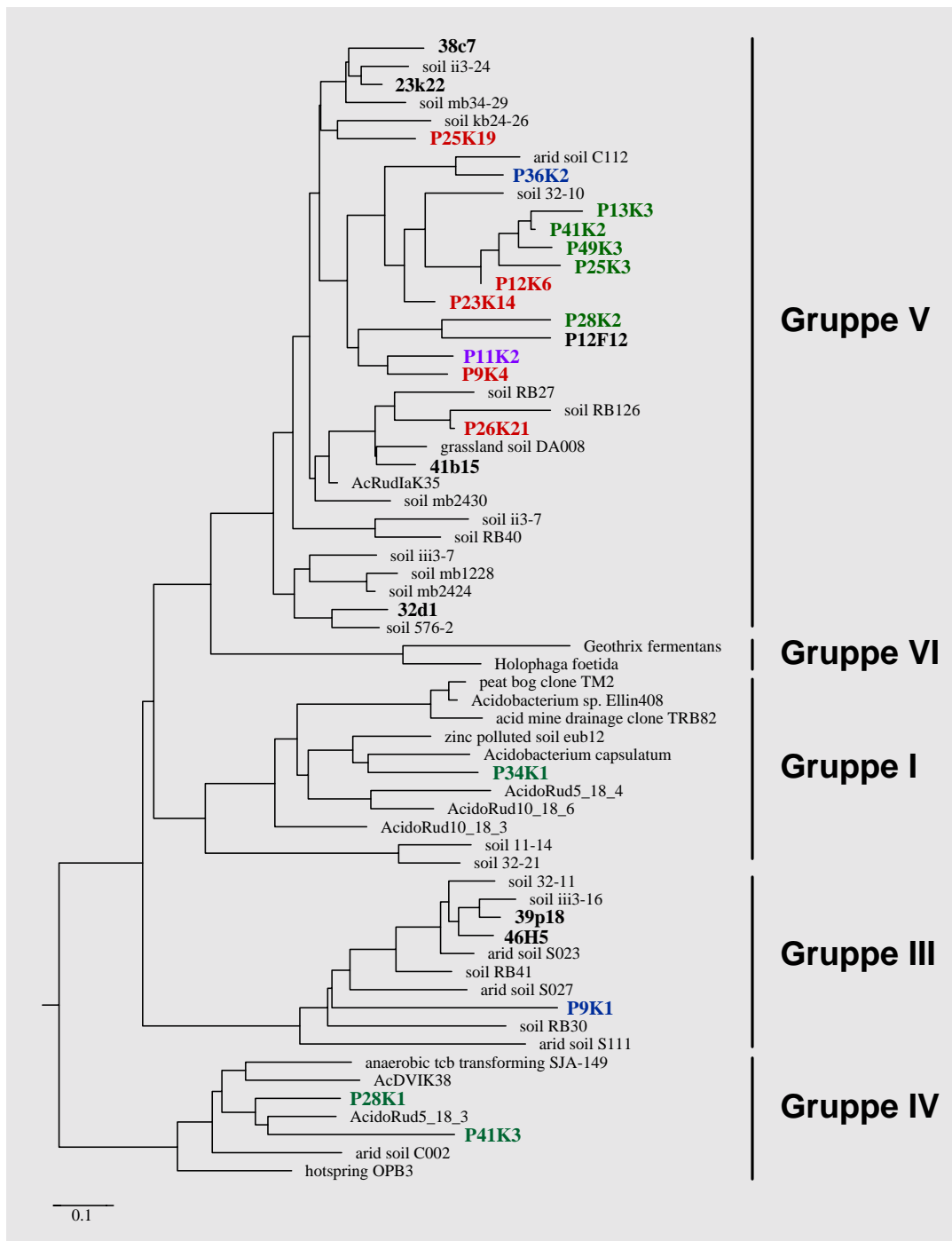


Abbildung 26: Phylogenetische Einordnung der in dieser Studie in den Genbanken Rud001.AQ (rot), Rud003.AT (blau), S247 (lila), S248 (grün) identifizierten acidobakteriellen Sequenzen. Farblich gekennzeichnete Sequenzen stammen von PCR-Pool-Amplifikaten. Fett dargestellte Sequenzen stammen von Einzelklonen mit acidobakteriellem Insert (siehe unten, S. 48). Der Baum wurde mit Sequenzen konstruiert, die mindestens eine Länge von 300 Nukleotiden aufwiesen (Position 37 bis 342 nach *E. coli* Nummerierung). Die Topologie stammt aus einer Neighbour Joining Analyse. Die phylogenetische Rekonstruktion wurde von Torsten Ochsenreiter ARB-Programm-Pakets durchgeführt.

### 3.3.3. Identifizierung von Klonen mit acidobakteriellen 16S rRNA-Genen

#### Identifizierung von Klonen durch Filterhybridisierung

Zunächst wurden spezifische Sonden aus PCR-Produkten von 14 verschiedenen Pools (siehe Tabelle 4) der Genbank Rud001.AQ (Pool 9, 12, 23, 25, 26, 32, 33, 36, 38, 39, 41, 42, 46, 52) und von 5 verschiedenen Pools der Genbank Rud003.AT (9, 36, 46, 62, 72) hergestellt. Aus den Pools 17, 57 und 63 der Genbank Rud001.AQ konnten keine Sonden hergestellt werden, da keine Amplifikate in dieser PCR erhalten wurden.

Mit den spezifischen Sonden wurde daraufhin eine Filterhybridisierung durchgeführt. Dabei musste im Vergleich zur Filterhybridisierung mit Archaea-spezifischen Sonden die Waschtemperatur von 60°C auf 68°C erhöht werden, um unspezifische Bindungen zu verringern (siehe E.2.12). Beispielhaft ist die Filterhybridisierung der Mikrotiterplatte 39 gezeigt, auf dem der Einzelklon 39P18 identifiziert werden konnte (Abbildung 27). Insgesamt wurden durch Filterhybridisierung 83 potentielle Einzelklone in den beiden Genbanken identifiziert. In einigen Fällen waren auf den Röntgenfilmen neben eindeutigen auch nur schwache Signale zu erkennen. Alle in Frage kommenden Einzelklone wurden untersucht.

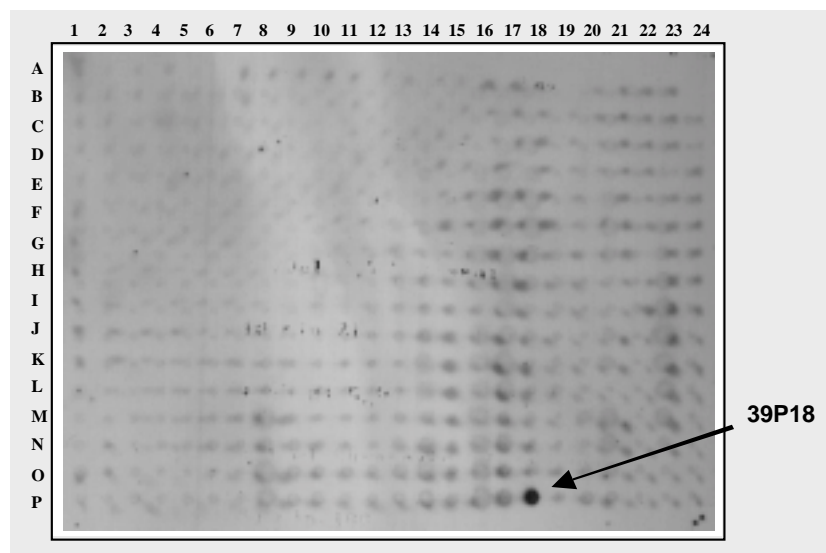


Abbildung 27: Filterhybridisierung der Mikrotiterplatte 39 der Genbank Rud001.AQ. Als Sonde diente das DIG-markierte, mit Acidobacteria-spezifischen Oligonukleotiden erhaltene Amplifikat des Klonpools 39. Auf der Koordinate P18 ist ein positives Signal zu sehen.

### **Analyse der Klone mit Genomfragmenten der Acidobacteria**

Plasmid-DNA von Einzelklonen, die in der Filterhybridisierung positiv waren, wurden extrahiert und durch PCR mit den Oligonukleotiden Ac31F-Univ533R auf das Vorhandensein eines ribosomalen RNA-Gens untersucht. Nach phylogenetischer Analyse oder durch Datenbankvergleich (blastn) konnten insgesamt 9 Fosmid-Klone identifiziert werden, die ein acidobakterielles 16S rRNA-Gen enthielten.

Dabei wurden 68 Einzelklone der Genbank Rud001.AQ und 15 Einzelklone der Genbank Rud003.AT analysiert und ausgewertet. Von 68 Klonen der Genbank Rud001.AQ zeigten 9 ein sehr starkes Amplifikat und 51 Einzelklone ein schwächeres Amplifikat. Nur bei 8 Einzelklonen war kein Amplifikat vorhanden. Von 15 Klonen der Genbank Rud003.AT zeigten 2 ein sehr starkes, 9 ein schwaches und 4 kein Amplifikat. Von 7 Einzelklonen der Genbank Rud001.AQ, sowie von 2 Einzelklonen der Genbank Rud003.AT, die in der PCR ein starkes Signal gezeigt hatten, wurde anschließend mit Hilfe des Oligonukleotids Ac31F die 16S rDNA-Sequenz bestimmt und ausgewertet. Die Einzelklone 32D1, 12F12 und 38C7 der Genbank Rud001.AQ konnten durch phylogenetische Analyse der Gruppe V der Acidobacteria zugeordnet werden (siehe Abbildung 26, fett dargestellt). Nur die Sequenzen des Klons 39P18 der Genbank Rud001.AQ und des Klons 46H5 der Genbank Rud003.AT waren zu Sequenzen der Gruppe III der Acidobacteria verwandt. Mit den weiteren Sequenzen der Klone 9O8, 25C19, 26L16 und 36O12 konnten durch Vergleich mit Sequenzen der Datenbanken (blastn) deutliche Hinweise auf die Zugehörigkeit zu den Acidobacteria erhalten werden. Insgesamt wurden 9 Fosmid-Klone identifiziert, die ein acidobakterielles 16S rRNA-Gen enthielten.

#### **3.4. Analyse des acidobakteriellen Klons 32D1**

Die Sequenzierung des Einzelklons 32D1 wurde am MPI in Tübingen (Stephan Schuster, Kooperationspartner) und die Annotation und Analyse in unserem Labor durchgeführt.

Der Klon 32D1 trug ein Insert von 35152 bp. Der GC-Gehalt des Inserts war 62,3%, in den Protein-codierenden Regionen sogar 63,7%. Es waren ein rRNA-Operon, zwei tRNAs und 22 vorhergesagte Protein-codierende Regionen auf dem Klon lokalisiert. Die Kodierungsdichte lag bei 86%. Das ribosomale RNA-Operon bestand aus einem 16S rRNA-Gen, tRNA<sup>Ala</sup> und tRNA<sup>Ile</sup> Genen, einem 23S rRNA-Gen und einem 5S rRNA-Gen. Das Vorkommen von tRNA<sup>Ala</sup> und tRNA<sup>Ile</sup> Genen zwischen 16S und 23S rRNA-Genen wird bei vielen bakteriellen rRNA-Operons gefunden (Chen *et al.*, 2000). Das 16S rRNA-Gen zeigte über eine Länge von 1331 bp 96% Sequenzidentität zur 16S rDNA des Klons "clone 576-2" (AF423279), der aus einer Studie eines landwirtschaftlich genutzten Bodens in

Kalifornien hervorging (Valinsky *et al.*, 2002). Die phylogenetische Analyse des kompletten 16S rRNA-Gens zeigte deutlich die Zugehörigkeit zur Gruppe V des Phylums der Acidobacteria (siehe Abbildung 26).

Vierzehn der 22 vorhergesagten Protein-codierenden Gene zeigten große Ähnlichkeiten mit Genen bekannter Funktion, fünf waren konservierte hypothetische Proteingene und drei zeigten nur schwache Ähnlichkeiten mit Sequenzen aus der Datenbank (Tabelle 5 und Abbildung 28). Siebzehn der 22 abgeleiteten Proteinsequenzen zeigten große Ähnlichkeiten mit Homologen der Proteobacteria, allein 12 davon zu  $\alpha$ -Proteobacteria und davon 10 zu *Rhodopseudomonas palustris*. Eine Region des Klons 32D1 zeigte über 10 kb mit 8 offenen Leserahmen Kolinearität, zu einem Genomfragment von *Rhodopseudomonas palustris*. Dieser kolineare Bereich wurde nur durch die Insertion eines Gens (ORF 3), das für ein MFS Transporterprotein (*Major Facilitator Subfamily* Transport Protein) codiert, unterbrochen. Die 8 hier gefundenen offenen Leserahmen codierten für eine Zink-Metallopeptidase, eine intrazelluläre Polyhydroxybutyrat Depolymerase, ein konserviert hypothetisches Protein, ein *Multidrug*-Transportsystem des ABC-Typs, ein Zwei-Komponenten System, bestehend aus Sensor-Histidinkinase und *Response*-Regulator und ein weiteres konserviert hypothetisches Protein (ORF 10). Die abgeleiteten Proteinsequenzen dieses kolinearen Bereichs zeigten in paarweisem Alignment mit Proteinsequenzen von *Rhodopseudomonas palustris* 63,9% identische Positionen über eine Länge von 2843 Aminosäurereste (siehe Kapitel D.3.4.1).

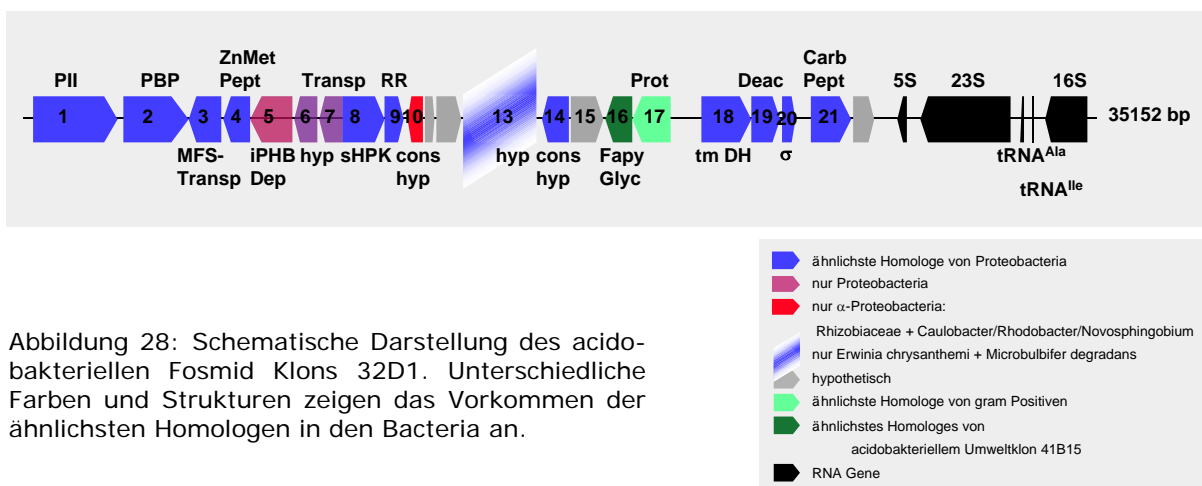


Abbildung 28: Schematische Darstellung des acidobakteriellen Fosmid Klons 32D1. Unterschiedliche Farben und Strukturen zeigen das Vorkommen der ähnlichsten Homologen in den Bacteria an.

## C. Ergebnisse

Tabelle 5: Potentielle RNA und Protein codierende Gene auf dem acidobakteriellen Fosmid 32D1.

ORF (O <sup>5</sup> )	Nt Bereich	Protein	Vorhergesagte Funktion	Ähnlichster Ortholog*	Phyl. aff.#	Bemerkung
01 (+)	330-3144	938 as	Protein PII Uridyltransferase pfam01842 pfam01909	ZP_00010189 Rhodopseudomonas palustris (0.0)	α-Proteobacteria	
02 (+)	3250-5356	702 as	Penicillin-binding protein pfam00912, pfam00905	ZP_00010175 Rhodopseudomonas palustris (0.0)	α-Proteobacteria	Transglycosylase Domäne Transpeptidase Domäne
03 (-)	6477-5358	373 as	major facilitator subfamily transporter pfam00083	ZP_00056223 Magnetospirillum magnetotacticum (e -48)	α-Proteobacteria	Zucker und "drug" Transporter Domäne GC_rich low_complexity repeat (1-21) bei 6154 -6174
04 (-)	7364-6563	361 as	Zink-Metallopeptidase pfam01863	ZP_00010174 Rhodopseudomonas palustris (e -59)	α-Proteobacteria	
05 (-)	8791-7390	467 as	Intrazelluläre PHB Depolymerase	ZP_00010173 Rhodopseudomonas palustris (e -178)	α-Proteobacteria	
06 (+)	9119-9518	133 as	Konserviert hypothetisch	ZP_00010172 Rhodopseudomonas palustris (e -75)	α-Proteobacteria	
07 (-)	10422-9666	252 as	ABC-Type Multidrug Transportsystem pfam00005	ZP_00010171 Rhodopseudomonas palustris (e -91)	α-Proteobacteria	smart00382 (AAA-ATPase) ATPase Komponente
08 (+)	10482-11796	438 as	Sensor Transduktion Histidine Kinase pfam02518 pfam00512	ZP_00010170 Rhodopseudomonas palustris (e -123)	α-Proteobacteria	HPK Subfamilie 3e;
09 (+)	11851-12424	191 as	Response-Regulator pfam00072	ZP_00010169 Rhodopseudomonas palustris (e -61)	α-Proteobacteria	smart00448, Helix-Turn-Helix Motiv, RegR bei B. japonicum
10 (-)	13079-12560	173 as	Konserviert hypothetisch	ZP_00010168 Rhodopseudomonas palustris (e -46)	α-Proteobacteria	
11 (+)	12941-13451	170 as	hypothetisch	ZP_00097019 Desulfitobacterium hafniense (e -10)	Gram +	pfam00190 pfam0204 pfam03079
12 (+)	13457-14354	299 as	Konserviert hypothetisch pfam03976	ZP_00081872 Geobacter metallireducens (e -95)	δ-Proteobacteria	DUF344
13 (-)	16780-14356	808 as	Hypothetisches Membranprotein pfam02687	CAD27338 Pectobacterium chrysanthemi (e -44)	γ-Proteobacteria	DUF214, Permease Domäne, repeat bei 14385-14405 - GC_rich Low_complexity (1-21)
14 (-)	17823-16947	292 as	Konserviert hypothetisch	NP_642719 Xanthomonas axonopodis str.306 (e -30)	γ-Proteobacteria	
15 (+)	17888-18953	355 as	hypothetisch pfam01207	At4g38890 Arabidopsis thaliana (e -47)	Eukaryonten	UPF0034
16 (-)	19867-18972	226 as	Formamidopyrimidine-DNA glycosylase pfam01149	acidobakterieller Umweltklon 41B15 65,4% Aminosäuresequenzidentität über 298 Positionen	Acidobacteria	Phylogenetische Analyse (Ochsenreiter, 2002)
17 (-)	21157-19876	427 as	Carboxy-terminal processing Protease	NP_244466 Bacillus halodurans (e -16)	Gram +	smart00245 - tail specific protease
18 (+)	22194-23784	530 as	Dehydrogenase pfam00732	NP_518892 Ralstonia solanacearum (e -53)	β-Proteobacteria	diese Familie bindet FAD als Cofaktor
19 (+)	23783-24740	319 as	Polysaccharide Deacetylase pfam01522	NP_457527 Salmonella enterica subsp. Enterica ser. Typhi (e -40)	γ--Proteobacteria	
20 (+)	24796-25285	163 as	RNA Polymerase sigma factor pfam00776	ZP_00009012 Rhodopseudomonas palustris (e -11)	α-Proteobacteria	Sigma70 ECF Subfamilie
21 (+)	25753-27049	432 as	Carboxypeptidase G2 Chain A, pfam01546	ZP_00096773 Novosphingobium aromaticivorans (e -63)	α-Proteobacteria	
22 (+)	27099-27822	241 as	Konserviert hypothetisch pfam01904	NP_294264 Deinococcus radiodurans (e -55)	Deinococcus-Thermus	DUF72
RNA	29040-28928		5S RNA			
RNA	32276-29309		23S RNA			
RNA	32668-32596		tRNA <sup>Ala</sup>			
RNA	32977-32903		tRNA <sup>Ile</sup>			
RNA	34779-33384		16S RNA	AF423279 unkultiviertes Bodenbakt. "clone 576-2"		96% identische Positionen auf 1300 bp

\*Proteine sind durch ihre Proteinidentifizierungsnummer gefolgt vom Name der Spezies bezeichnet. Die e-Werte von blast sind in Klammern angegeben.

# Phyl. aff. = Phylogenetische Affinität, zeigt das Vorhandensein von Homologen in Archaea (A), Bacteria (B) oder Eukaryonten (E) an.

§ = +/- Orientierung

pfam = Protein Familien Domänen Datenbank (Sonnhammer *et al.*, 1998); DUF = Domain of unknown function; smart = Simple Modular Architecture Research Tool (Schultz *et al.*, 1998)

Ein  $\sigma$ -Faktor (ORF 20) und eine Carboxypeptidase (ORF 21) zeigten neben dem kollearen Bereich ebenfalls hohe Ähnlichkeit mit  $\alpha$ -Proteobacteria. Hingegen zeigte ein abgeleitetes Protein, das für eine Protease (ORF 17) codiert, hohe Ähnlichkeit mit einem Proteingen von Gram-positiven Bakterien. Ein hypothetisches Membranprotein (ORF 13), in dem eine Sequenzwiederholung detektiert wurde, zeigte ausschließlich Sequenzähnlichkeiten mit Proteinen von *Pectobacterium chrysanthemi* (Synonym: *Erwinia chrysanthemi*) und *Microbulbifer degradans*. Ein weiteres Gen (ORF 19) codierte für eine potentielle Polysaccharid-Deacetylase. Die phylogenetische Analyse eines weiteren Proteins, das auf 32D1 codiert war, eine Formamidopyrimidine-DNA Glycosylase (ORF 16), zeigte die Zugehörigkeit zu einem homologen Protein auf einem weiteren acidobakteriellen Klon 41B15 (Ochsenreiter, 2002).

### 3.4.1. Vergleichende Genomanalyse des kollearen Bereichs auf Klon 32D1

Die vorausgesagte intrazelluläre Polyhydroxybutyrat-Depolymerase (ORF 5) zeigte Sequenzidentitäten zwischen 68-42% über 370 Nukleotide zu den insgesamt 23 in den Datenbanken abgelegten Homologen, die ausschließlich von  $\alpha$ -Proteobacteria abstammen. Die vergleichende Genomanalyse von Homologen der intrazellulären PHB-Depolymerase und der stromabwärts liegenden vorhergesagten Zink-Metallopeptidase (ORF 4) zeigte, dass beide Gene oft paarweise auftreten (Abbildung 29), was auf einen funktionellen Zusammenhang hindeuten könnte.

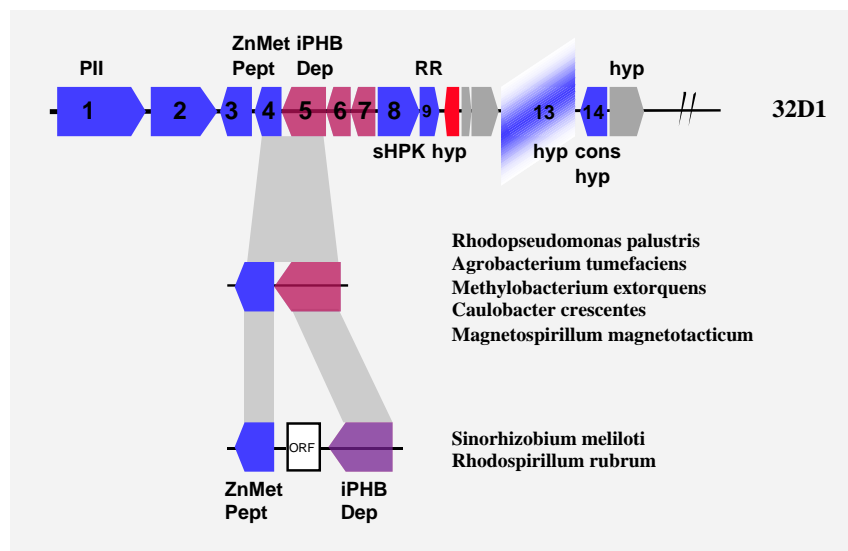


Abbildung 29: Schematische Darstellung des paarweisen Vorkommens der Zink-Metallopeptidase mit der intrazellulären Polyhydroxybutyrat-Depolymerase bei 32D1 und einigen  $\alpha$ -Proteobacteria.



Die vorhergesagte Histidinproteinkinase (ORF 8) zeigte die für diese Proteine typischen konservierten Sequenzmotive (Abbildung 30: H-, X-, N-, D- und F- und G-Box). Durch genauere Betrachtung der konservierten Motive konnten die für Sensor-Histidinproteinkinasen der Subfamilie 3e (sHPK3e) typischen Sequenzübereinstimmungen identifiziert werden (Grebe und Stock, 1999): Histidinproteinkinasen der Subfamilie HPK 3e besitzen in der H-Box 4 Alanin Reste vor dem konservierten Histidin (His-217) und außerdem einen für Histidinkinasen ungewöhnlichen Glycin Rest anstatt eines Arginins (R) oder Lysins (K) vor dem konservierten Prolin. Diese waren auch in der abgeleiteten Proteinsequenz des ORF 8 vorhanden. Durch das Alignment der konservierten Motive mit homologen Sequenzmotiven und durch den Sequenzvergleich mit RegS, das experimentell bereits gut charakterisierte Homolog von *Bradyrhizobium japonicum* (Emmerich *et al.*, 1999), konnte einerseits die potentielle Autophosphorylierungsstelle an einem Histidinrest (His-217) und außerdem die essentiellen Positionen der Kinase-Domäne identifiziert werden (Abbildung 30). Des weiteren wurden durch eine Transmembrananalyse für die vorausgesagte Sensor-Histidinkinase (ORF 8) 5 Transmembrandomänen identifiziert (nicht gezeigt), woraus man schließen konnte, dass es sich um eine membranständige Sensor-Histidinkinase handelte.

	H-Box	X-Box	N-Box	D-und F-Box	G-Box
HPK 3e	AAAAHELGTPL	X-box	hh .h.NLh .NAVdYA	h.h.h.DDG.G..P.hh.RhG-PFh	GhGLGL LLERsGA.h.F.N
prrb_rsh	LGVVAAAAHELGTPLATI	AELIREQADRCDILRSMGRA	RRPEVIHGLRNLIQNAVDFa	IAIRIVDDGEGYPPAIIGRI GDPPVREQ	EGMGLGLFIAKTLLESGAELSFANAADP
regb_rca	LGVVAAAAHELGTPLATI	AVLIREQAERCRIILRSMGRA	RYPELVHALRNLIQNAVDFa	IIVRVTDGCRGYSPVNLNRI GDPPFIST	EGMGLGLFIAKTLLERTGAKIRFANGSEP
acts_rme	LDGLAAAAHELGTPLATI	VHLLRSQSERCDILRRLTTL	RNAGILYGLGNLLENADVYA	VRVTIEDDGEFSPDILARIGEPYVTR	--LGLGLFIAKTLLESGARIRFENGGAK
regs_bja	LDGLAAAAHELGTPLATI	LKTLREQTQRCDILSKITQL	RNPAILYGVGNIVENADVFA	IELLISDDGPGIPDDILNRI GEPYLSR	RGLGLGVFIARTLLERTGAKVSPFNRIFFP
32d1_08	LDGLAAAAHELGTPMSTI	VKLLREQAARCREILKLTLEL	RNPAILYGLGNLLENADVFA	VSVTITDDGPGFAPEVIGRLGEPYV--	-GLGLGLFIAKTLLESGATLITFAN
	AAAA   G P		N	D_G_G	G_GL
	His-217				

Abbildung 30: Zuordnung der vorgeschlagenen Histidin-Proteinkinase (ORF 8) zur HPK Subfamilie 3e durch Alignment der C-terminal liegenden, charakteristischen, konservierten Motive von Histidin Protein Kinasen. Identifizierung der potentiellen Autophosphorylierungsstelle (His-217) sowie der potentiellen Kinase Domäne (N; D\_G\_G; G\_GL). Rot, identische Positionen; blau, ähnliche Positionen; HPK3e, Consensus Sequenz innerhalb der konservierten Motive charakteristisch für HPK der Subfamilie 3e; Abkürzungen: prrb\_rsh: *Rhodobacter sphaeroides*; regb\_rca: *Rhodobacter capsulatus*; acts\_rme: *Sinorhizobium meliloti*; regs\_bja: *Bradyrhizobium japonicum*; RegS\_32D1: acidobakterieller Umweltklon 32D1 (ORF 8).

Die abgeleitete Sequenz des ORFs 09 zeigte große Sequenzähnlichkeiten mit *Response*-Regulatoren von typischen bakteriellen Zwei-Komponentensystemen. Durch Datenbankvergleich (blastp) und Alignierungen konnten insgesamt 34 Homologe zum vorausgesagten *Response*-Regulator (ORF 9) identifiziert werden, die vergleichbare Aminosäuresequenzlängen und hohe Ähnlichkeiten aufwiesen. Alle 34 Sequenzen stammten von Proteobacteria. 18 Sequenzen, die von  $\alpha$ -Proteobacteria stammten, zeigten in paar-

C.Ergebnisse

weisem Alignment mit der abgeleiteten Sequenz des Klons 32D1 (ORF 9) Sequenzidentitäten von 80% bis 68% über 170 – 180 Positionen. Die weiteren 16 Homologen Sequenzen stammten von anderen Proteobacteria (nicht  $\alpha$ ) und zeigten geringere Sequenzidentitäten von 53,2% - 39,9% über 175 - 165 Positionen zu 32D1. Auf Basis des Alignments konnten alle Sequenzmotive eines *Response*-Regulators in ORF 9 identifiziert werden (Abbildung 31).

		D-70											
<b>32D1 09</b>	SG	SSARVALPAE	RSLLVIEDDA	SFLGRLLAKAL	EARGFAVTTA	GSVAEGLAQV	EASPPAFAVV	DMRLADGNGL	DVISAALKRRR	PDARGIILTG			
B. jap RegR	MN	AIAELNEQTD	RSLLVIEDDK	PFLERLSRAM	ETRGFAVTSC	DTVSDGLAQI	GKAAPAFAVV	DLRLGDGNGL	DVVSALKKKR	PDARAIIVLTG			
R. pal	MN	AIPELNDHTD	RSLLVIEDDK	PFLDLRLARAM	ETRGFAVTAC	DSVSSGLAQI	GKAAPAFAVV	DLRLGDGNGL	DVVSALKRER	PDCAIVLTG			
R. leg ActR	FA	ASDEDHICPD	ASLLIVDDDG	PFLRRLARAM	ETRGFQVETA	ESVAEGVAKS	KGRPPKYAVV	DLRLGDGNGL	DVIEAIRQRR	DDTRIVLTG			
A. tum	EP	VDDIDPIQPD	RSLLVIEDDG	PFLRRLARAM	ETRGFAVDTA	ETVSEGIARS	KAAPPKYAVV	DLRLADGNGL	EVIEAIRQNR	DDTQIVVLTG			
S. mel ActR	H	AADADLCPD	KSLLVIEDDT	AFLRRLARAM	EARGFAVEIA	ESVAEGIAKA	KTRPPKHAVI	DLRLSDGSL	DVIEAIRGRR	DDTRMIVLTG			
B. mel	LK	QEDTEAICND	PTLLLVDDDK	PFLQLRLARAM	ESRGFQVTTA	ESVEEGIAAV	KTSAPAYAVV	DMRLGDGNGL	NIEAIRSRR	ADTRAIIVLTG			
M. lot ActR	DE	NIGAMVECED	TSLLVIEDDK	PFLTRRLARAM	ETRGFVETA	ESVEEAVGKA	RANPPAYAVV	DMRLGDGNGL	DVVAIREKR	EDSRTIIVLTG			
R. sph RegA	MA	EDLVFELGAD	RSLLVIEDDE	PFLKRLAKAM	EKRGFVLETA	QSVAEKKAIA	QARPPAYAVV	DLRLEDGNGL	DVVEVLRER	PDCAIVVLTG			
R. cap RegA	MA	EEEFaelGSD	RSLLVIEDDN	AFLTRRLARAM	EKRGFQTEIA	ETVSAGKAIV	QNRAPAYAVI	DLRLEDGNGL	EVVEALRER	PEARIVVLTG			
R. sul RegA	MA	EQE-YEIGED	PSLLVIEDDE	PFLRRLARAM	EKRGFQPEMA	ETVAAGKAIA	SARPPAYAVV	DLRLEDGTGL	DVVEVLRER	PDCAIVVLTG			
R. den RegA	MP	DTNPielGED	KSLLLVIEDDE	PFLRRLAKAM	EKRGFQVETA	GSVAAGRAIA	TARPPAFAVV	DLRLEDGNGL	DVVEVLRER	PDTRVIVLTG			
unku Prot	M	ANEPLDIGED	KSLLVIEDDE	PFLKRLAKAM	EKRGFVLETA	QSVAEKKAIA	TARPPAFAVV	DLRLEDGNGL	DVVEVLRER	PDCAIVVLTG			
C. cre	AD	IGELVAALPD	KSLLVIEDDD	PLRTRLGRAL	EQRGFVETLA	ASVAEALTL	RTQAPAHAVL	DMRLEDGSL	KVVEAVRDAR	ADAKVIMLTG			
M. mag	LS	DTDPLSAFSD	RSLLVIEDDK	PFSTRRLARAM	EARGYEVHVA	ESVSDGVAV	EARAPAFAVI	DMRLGDGNGL	DVIARLKERR	PEARIVVLTG			

		Helix-Turn-Helix											
<b>32D1 09</b>	YGN	IATAVTA	VKLGAVDYL	KPADADDVVA	ALLALEGNKT	ELPENPMSA--DRVRW	EHIQRIYEMC	GRNVSETARR	LSMHRRTLQR	ILAKRAP			
B. jap RegR	YGN	IATAVTA	VKMGAIDYL	KPADADDVVA	ALLSTSAEKS	ELPTNPMSA--DRVRW	EHIQRIYEMC	NRNVSETARR	LNMHRRTLQR	ILAKRAPR			
R. pal	YGN	IATAVTA	VKMGAIDYL	KPADADDVVA	ALLSTSAEKS	EPPNPMSA--DRVRW	EHIQRIYEMC	NRNVSETARR	LNMHRRTLQR	ILAKRAPR			
R. leg ActR	YGN	IATAVTA	VKLGAVDYL	KPADADDVFS	ALTQRPGEKA	ELPENPMSA--DRVRW	EHIQRVYEMC	ERNVSETARR	LNMHRRTLQR	ILAKRAPK			
A. tum	YGN	IATAVTA	VKLGALDYLA	KPADADDVFN	ALTQRKGEKT	EVPENPMSA--DRVRW	EHIQRVYEMC	ERNVSETARR	LNMHRRTLQR	ILAKRAP			
S. mel ActR	YGN	IATAVTA	VKLGALDYLA	KPADADDILA	ALLQRPGEKV	EPPNPMSA--DRVRW	EHIQRVYEMC	ERNVSETARR	LNMHRRTLQR	ILAKRAPK			
B. mel	YGN	IATAVTA	VKLGALDYLA	KPADADDVFA	ALTRRPGEKV	APPENPMSA--DRVRW	EHIQRVYEMC	ERNVSETARR	LNMHRRTLQR	ILAKRAPR			
M. lot ActR	YGN	IATAVTA	VKLGAVDYL	KPADADDIFA	ALTRTSGERA	APPENPMSA--DRVRW	EHIQRVYEMC	ERNVSETARR	LNMHRRTLQR	ILAKRAPR			
R. sph RegA	YGA	IATAVAA	VKIATDYL	KPADANEVTH	ALLAKGESLP	PPPENPMSA--DRVRW	EHIQRIYEMC	DRNVSETARR	LNMHRRTLQR	ILAKRSPR			
R. cap RegA	YGA	IATAVAA	VKMATDYL	KPADANDITN	ALLAKGEALP	PPPENPMSA--DRVRW	EHIQRVYELC	DRNVSETARR	LNMHRRTLQR	ILAKRSPR			
R. sul RegA	YGA	IATAVAA	VKVGATDYL	KPADANDVTA	ALLSNGEALP	PPPENPMSA--DRVRW	EHIQRVYEQC	DRNVSETARR	LNMHRRTLQR	ILAKRSPR			
R. den RegA	YGA	IATAVAA	VKIATDYL	KPADATDIVN	ALLATGDDL	PPPENPMSA--DRVRW	EHIQRVYELC	DRNVSETARR	LNMHRRTLQR	ILAKRSPR			
unku Prot	YGA	IATAVAA	VKIATDYL	KPADAKDIVN	ALLATGDEL	PPPENPMSA--DRVRW	EHIQRVFELC	DRNVSETARR	LSMHRRTLQR	ILAKRSPR			
C. cres	YGN	IATAVAA	VKAGVDYL	KPADADDVAR	ALLAAKDAAP	APPENPMSA--DRVRW	EHIQRVYEMC	GHNVSETARR	LNMHRRTLQR	ILAKRAPR			
M. mag	YGN	IATAVTA	VKLGAVDYL	KPADADIEIHG	TLMATQPERA	DPENPMSA--DRVRW	EHIQRVYELC	GRNVSETARR	LNMHRRTLQR	ILAKRAPR			

Abbildung 31: Proteinsequenzalignment von 14 bakteriellen *Response*-Regulatoren der  $\alpha$ -Proteobacteria und von 32D1. Insgesamt sind 184 Reste des 32D1-*Response*-Regulators (ORF 9) gezeigt. N-Terminal wurden 7 Reste entfernt. Die potentielle Phosphorylierungsstelle liegt an Position D70. Das hochkonservierte Helix-Turn-Helix Motiv ist gekennzeichnet. Abkürzungen: 32D1 09, ORF 9 des acidobakteriellen Umweltklons 32D1; R.pal, *Rhodopseudomonas palustris*; B.jap RegR, *Bradyrhizobium japonicum*; M.lot ActR, *Mesorhizobium loti*; B.mel, *Bruceella melitensis*; A.tum, *Agrobacterium tumefaciens*; R.leg ActR, *Rhizobium leguminosarum* bv. viciae; S.mel, *Sinorhizobium meliloti*; R.den RegA, *Roseobacter denitrificans*; R.sph RegA, *Rhodobacter sphaeroides*; R.cap RegA, *Rhodobacter capsulatus*; R.sul RegA *Rhodovulvum sulfidophilum*; C.cre, *Caulobacter crescentens* CB15; unku Prot, nicht-kultiviertes Proteobacterium - mariner Umweltklon ; M.mag, *Magnetospirillum magnetotacticum*.

Einerseits konnte die potentielle Phosphorylierungsstelle bei D70 und andererseits auch das hochkonservierte Helix-Turn-Helix Motiv am C-Terminus (**NVSE-TARLSMHRRTLQRILA**), das typische Motive der AraC Familie zeigte (HTHScan HUSAR) bestimmt werden. Experimentell wurde bereits beim homologen *Response*-Regulator RegR von *Bradyrhizobium japonicum* die Phosphorylierungsstelle an dieser Position nachgewiesen (Emmerich *et al.*, 1999). Im Helix-Turn-Helix Motiv war auffallend, dass nur die Proteinsequenz des ORFs 09 und das Homologe von einem unkultivierten  $\alpha$ -Proteobacterium aus marinem Plankton (Beja *et al.*, 2002a) hier einen Serinrest (S) anstatt einen Argininrest (R) zeigten.

Durch phylogenetische Analyse des *Response*-Regulators von 32D1 mit 26 Homologen konnten zwei unterschiedliche Gruppen unter den nächsten Verwandten definiert werden (siehe Diskussion Abbildung 35). Die phylogenetische Analyse zeigte die Zugehörigkeit des *Response*-Regulators von 32D1 zu Homologen der Gruppe 1, die ausschließlich von *Rhizobiaceae* stammten. Die Gruppe 2 wurde von Homologen anderer  $\alpha$ -Proteobacteria gebildet. *Response*-Regulatoren von  $\beta$ - und  $\gamma$ -Proteobacteria, sowie Paraloge von  $\alpha$ -Proteobacteria zeigten nur eine schwache Verwandtschaft. Homologe zu diesem Zwei-Komponentensystem wurden teilweise bereits biochemisch sowie *in vivo* untersucht und charakterisiert (siehe Diskussion).

Der stromaufwärts liegende ORF 10, der hier als konserviert hypothetisches Protein annotiert wurde, zeigte große Sequenzähnlichkeiten mit einem Protein mit unbekannter Funktion, das ausschließlich bei  $\alpha$ -Proteobacteria vorkommt. Durch Sequenzvergleich konnten in den Datenbanken insgesamt 11 homologe Proteine identifiziert werden. Davon stammten 7 von  $\alpha$ -Proteobacteria der *Rhizobiaceae*, die eine monophyletische Gruppe unter den  $\alpha$ -Proteobacteria bilden (Ludwig und Klenk, 2001). Durch Alignierung konnten hochkonservierte Bereiche identifiziert werden (Abbildung 32).

## C. Ergebnisse

<b>32D1 10</b>	VTWSIRKPIE	APVDGRQSET	ALAIARGTSR	LLHAHGF	SELSLASGRR	ADLVAINSSC	EIWIVEIKSS	VADFRADQKW		
B. jap hyp	DSSARNIALV	PPDRRQSET	ALAIARGTAR	LRLSLGF	SELPLPSGRR	ADLVALNERG	EIWIVEIKSS	VEDLRADQKW		
R. pal	ESNA--VQMI	LPRDLRQSET	ALNIARGTQR	LRLSLGF	SELPLPSGRR	ADLVALNERG	EIWIVEIKSS	VEDLRADHKW		
A. tum	MTILSFTNNN	PLIDGRQSDK	AMLVRRGVQV	LLHEMRHSVL	PELPLASGRR	ADLISLSAKG	EIWIIEIKSS	IEDFRVDRKW		
R. leg	MTIFSVYNNN	PLIDGRQSDR	AMVRRGTQI	LLHEMRHVAVL	PELPLASGRR	ADLITLSEKQ	EVIWIEIKTS	IEDFRVDRKW		
S. mel	MTILSIHGDN	PLADGRQSER	AMKVRRGVQR	LIMELRHAAL	PELTLASGRR	ADLISLSPKG	EIWIIEIKTS	IEDFRVDRKW		
B. mel	MPIVIPNQTH	PLSDGRQSEN	ALLVRRGVQR	LFLAMGLATL	PELSLASGRR	ADLIAVSRKG	EIWIVEIKSS	IEDWKADHKW		
M. lot	MPIISPIAIN	PLIDGRQSER	AMLVRRGVQR	LLMAMGAHVL	PELSLATGRR	ADLVALTRQG	DIWIIEIKSS	IEDFRVDRKW		
C. cres	MDVI	IELAPSRPET	TLAVTRGAAR	LLVDLGYAPL	AEVTLPNRGR	ADLMALGPKG	DVLIVEVKSQ	LEDFRVDRKW		
R. sph	HDGAMQTLTV	PDSGLAPLLP	GQRLARGVCR	HLRSHDFVSV	VELVPAQGLR	VDVMALGPKG	EIVVVECKSS	RADYQSDRWK		
N. arom	MSESVSASAE	VTVEAAVERE	ARGVARGICR	LFARNDIWCL	PEMPLRSRGR	ADLMGIDAKG	QVVIVEIKVS	RADLLGDGKW		
M. magn			MR	LFAEMGCVTL	PEFSLANGRR	ADVIALCGVG	KLTIIVEIKSS	VADFRADRWK		
<b>32D1 10</b>	MDYRVHCDRL	FFATCQEVFC	-----E	IFP	PDTGLIVADA	FGASIVCEAP	EHRLPAPTRK	SMMLAFARVA	ALRLQALADP	H
B. jap hyp	HEYRAHCDRL	FFAFTQDLPC	-----E	IFP	QGTGLIADA	YGAHLHCEAP	EHKIAAATRK	QMTVRFAMAA	ALRINRLVDP	Q
R. pal	PDYRAHCDRL	FFAFTQDLPC	-----E	IFP	AETGLIADG	YGAHLHCEAP	EHKLPAATRK	VMMLRFALAA	AQRRLGRLSDP	Q
A. tum	PDYRLHCDRL	FFATHPEVPT	-----D	IFP	EECGLILSDG	YGAHMLRDAP	EHKLPPATRK	SVMLDFSRTA	ATRLMLAEWT	T
R. leg	PEYRLHCDRL	FFATHQDVPL	-----E	IFP	EECGLFLSDG	YGAHMIREAP	EHRMAPATRK	SVTLNFSRAA	AQRLLMMAEWA	N
S. mel	PEYRLHCDRL	FFATHAQVPL	-----D	IFP	EECGLMLSDG	YGAHMLREAS	EHRLPSTRK	SVTQAFARTA	AQRRLTAEWA	A
B. mel	PEYRAYCDRL	FFATHPAVPR	-----E	IFP	EECGFILSDG	YGAELIREAP	EHRLAGATRK	ALMLRFARAG	AARLMAAEWA	G
M. lot	PDYRLHSDRF	FFATHPGVPA	-----E	IFP	DECGFILSDG	YGAELIRDAP	EHRMAAATRK	ALMLRIRARAG	AARLLAAEWA	G
C. cres	GDYAPFCDAF	YFAVAPSFDP	-----G	ILP	EDPGLLVADG	FGGAVVREAP	LTPLAPARRK	ALTLAFGRLA	AWRAAGVNAE	R
R. sph	QGYLDWCDRF	FFAVDEAFV	-----D	LLP	EGTGLMLADG	YDAEILRIGP	EGR LAPARRK	AVIQGFARHA	ALRLQGLRDP	G
N. arom	TDYLDHCDRF	YWGLAPHLDR	AVLETE	DFLP	HACGVIVADG	YDAEIVRPAP	TVPLAAARRK	AEVERLARAS	LRRQLVGLDP	H
M. magn	PDYRDYCDRF	FFAHSRRLSR	-----	-----	-----	-----	-----	-----	-----	-----

Abbildung 32: Proteinsequenzalignment der Homologen des konserviert hypothetischen Proteins (ORF 10) auf 32D1. Bereiche schwacher Konservierung wurden entfernt. 173 Aminosäurereste des konserviert hypothetischen Proteins sind gezeigt, N-Terminal wurden 11 Aminosäuren und C-terminal 7 Aminosäuren entfernt; Abkürzungen: siehe Abbildung 31.

Aminosäuresequenzidentitäten wurden durch paarweise Alignierung zwischen ORF 10 auf Klon 32D1 und deren Homologen bestimmt: Homologe, die unter den *Rhizobiacae* zu finden waren zeigten 71%-57,8% Sequenzidentität über 145-135 Positionen (*Rhodopseudomonas palustris*, *Bradyrhizobium japonicum*, *Brucella melitensis*, *Mesorhizobium loti*, *Agrobacterium tumefaciens*, *Sinorhizobium meliloti*, *Rhizobium leguminosarum*). Die Identitäten zu 3 weiteren Homologen, die von anderen  $\alpha$ -Proteobacteria stammten, lagen hingegen nur bei 48,1-36,1% über 158-105 Positionen (*Caulobacter crescentens*, *Rhodobacter sphaeroides*, *Novosphingobium aromaticivorum*). Die Sequenz des hypothetischen Proteins von *Magnetospirillum magnetotacticum* war nur 104 Aminosäuren lang und zeigte nur eine schwache Konservierung.

In der vergleichenden Genomanalyse konnte eine auffällige Konservierung der Genomstruktur bezüglich des Zwei-Komponentensystems (ORF 8 und ORF 9) und dieses konservierten hypothetischen Proteins (ORF 10) identifiziert werden. Bei allen *Rhizobiaceae*, die dieses konservierte hypothetische Protein besitzen, und auch bei 32D1, lag das homologe Zwei-Komponentensystem direkt anliegend (bei *Bradyrhizobium japonicum*, *Rhizobium leguminosarum*, *Sinorhizobium meliloti*, *Agrobacterium tumefaciens*, und *Rhodopseudomonas palustris*) oder im Abstand von 4 bzw. 1 ORF (Abbildung 28; *Brucella melitensis* - Transposase und *Mesorhizobium loti* - H<sup>+</sup>/Na<sup>+</sup> -Transportergeren). Die vier weiteren Homologen, deren Gene in *Magnetospirillum magnetotacticum*, *Novosphingobium aromaticivorum*, *Caulobacter crescentens* und *Rhodobacter sphaeroides* gefunden wurden, lagen nicht in der Nähe der Gene des Zwei-Komponentensystems.

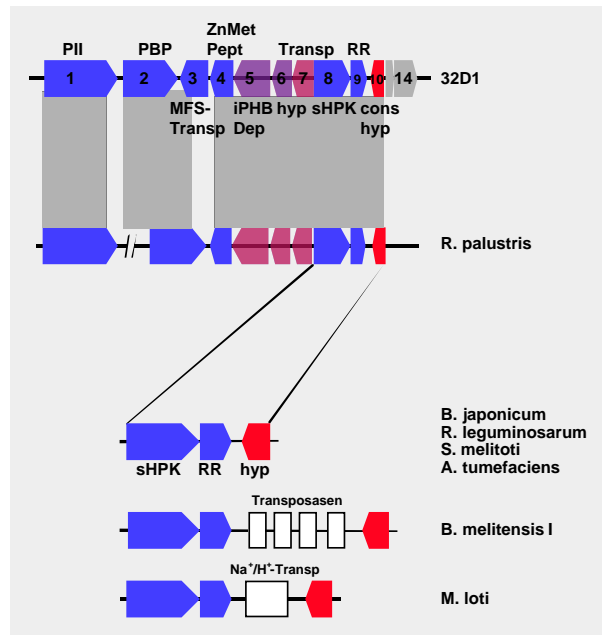


Abbildung 33: Schematische Darstellung des Fosmid Klon 32D1, sowie des kolinearen Bereichs bei *Rhodopseudomonas palustris* und bei *Rhizobiaceae*.

## D. Diskussion

Die Sichtweise der mikrobiellen Diversität in natürlichen Habitaten hat sich im letzten Jahrzehnt enorm gewandelt. Die Entwicklung von neuen Techniken in der molekularen Ökologie, vor allem die Möglichkeit zur Amplifizierung von 16S rRNA-Genen aus Umweltproben, hat zur Entdeckung von vielen neuen Mikroorganismen bzw. Phylotypen geführt (Pace, 1997). Inzwischen werden 13 der insgesamt 36 definierten Phyla ausschließlich durch 16S rDNA-Sequenzen ohne kultivierte Vertreter repräsentiert (Hugenholtz *et al.*, 1998). Neben der Amplifizierung von 16S rRNA-Genen werden auch zunehmend Protein-codierende Gene per PCR aus Umweltproben analysiert (Desaint *et al.*, 2000; Rondon *et al.*, 2000; Brady *et al.*, 2001; Fjellbirkelund *et al.*, 2001; Prieme *et al.*, 2002). Bedingt durch die limitierten technischen Möglichkeiten sind diese Studien aber weitgehend auf die Entdeckung von Variationen bereits bekannter Gene begrenzt und zeigen keine Verknüpfung zur Herkunft der entdeckten Enzyme. Somit bleiben der Zusammenhang zwischen Mikroorganismen und ihren spezifischen Stoffwechselwegen, die Relevanz von bestimmten Gruppen im Ökosystem und ihre speziellen Aufgaben noch vielfach verborgen.

Eine neue Strategie zur Charakterisierung von unkultivierten Mikroorganismen bildet die Umweltgenomik. Dabei wird angestrebt, einen Großteil der Genome der in einem Habitat vorkommenden Mikroorganismen in komplexen Genbanken zu konservieren und anschließend zu analysieren. Dadurch lassen sich auf Basis des genetischen Potentials metabolische und biochemisch relevante Eigenschaften ableiten. Vor allem lässt sich aber mit Hilfe geeigneter Markergene, wie z.B. der 16S rRNA die phylogenetische Verknüpfung zu bestimmten Mikroorganismen herstellen. Mit Hilfe erster umweltgenomischer Ansätze wurden in den letzten Jahren eine Reihe mariner Mikroorganismen analysiert. Im Zuge dieser Studien konnte beispielsweise das erste Protein eines mesophilen Crenarchaeoten, eine DNA Polymerase des Archaeons *Cenarchaeum symbiosum* biochemisch charakterisiert werden (Schleper *et al.*, 1997b). Ausserdem führte dieser Ansatz zur Identifizierung eines neuen Rhodopsin-Typs, des Proteorhodopsins (Beja *et al.*, 2001). Dieses wurde auf einem Genomfragment eines nicht-kultivierten marinen  $\gamma$ -Proteobacteriums der SAR86 Gruppe identifiziert, die eine der dominierenden Fraktionen in marinem Oberflächenwasser darstellt (Pernthaler *et al.*, 2002). Aufgrund dessen wird angenommen, dass dieses photoaktive Proteorhodopsin einen wesentlichen Einfluss auf den Kohlenstoff- und Energiestoffwechsel der Meere haben könnte (Beja *et al.*, 2002a). Damit bildet diese Entdeckung die Grundlage neuer Studien zur mikrobiellen Phototrophie im Meeresplankton.

In der vorliegenden Arbeit wurde der umweltgenomische Ansatz auf Bodenhabitats angewendet. Hierfür wurde ein Verfahren entwickelt, das die Isolierung von hochmolekularer, klonierbarer DNA aus verschiedenen Habitats erlaubte, deren Nucleinsäuren mit unterschiedlichen Hemmstoffen, wie Humin- und Fulvinsäuren, Polysacchariden oder bromhaltigen Alkaloiden kontaminiert waren. Die Anwendung dieses Verfahrens führte zur Erstellung von komplexen Genbibliotheken der Prokaryontenfraktionen aus unterschiedlichen Bodenhabitats sowie aus zwei marinen Schwämmen. Insgesamt wurden in dieser Studie mehr als 2 Gbp genomische DNA aus Umweltproben konserviert. Mit 16S rRNA-genspezifischen Sonden konnten aus der Genbank eines Sandökosystems ein Genomfragment eines Crenarchaeoten und 9 Genomfragmente der Acidobacteria identifiziert werden.

### **Isolierung von hochmolekularer DNA und Abtrennung von Humin- und Fulvinsäuren**

Das erste Ziel dieser Arbeit war die Entwicklung einer Methode, um hochmolekulare DNA von prokaryotischen Mikroorganismen effizient aus Bodenhabitats zu isolieren und zu reinigen. Bei der Isolierung von hochmolekularer DNA aus Bodenhabitats stellt die größte Herausforderung die Eliminierung von Humin- und Fulvinsäuren dar. Diese hochmolekularen, heterozyklischen Polyphenole komplexieren mit der isolierten DNA und inhibieren die folgenden, zur Klonierung notwendigen, enzymatischen Reaktionen.

Bei dem hier angewendeten Verfahren wurde zunächst eine Mikroorganismenfraktion des Sandökosystems RUD durch Aufschlammung in einem Puffer, der  $\epsilon$ -Aminocapronsäure enthielt, und durch anschließende Filtrierung gewonnen. Die  $\epsilon$ -Aminocapronsäure sollte die Dissoziation von an Bodenbestandteile gebundenen Mikroorganismen fördern. Um die DNA möglichst schonend zu behandeln, wurde die Zellyse mit in Agarose eingebetteter Biomasse durchgeführt, in Anlehnung an die Extraktion von hochmolekularer DNA aus Zellkulturen. Dabei wurde die höchstmögliche Menge an Biomasse eingesetzt. Die beiden Schritte zur Maximierung der Biomasse hatten sich als unerlässlich erwiesen, um genügend hohe DNA-Ausbeuten aus Bodenproben zu erhalten.

Die Zellen wurden mittels Lauroylsarkosin und hohen Konzentrationen von Lysozym und Proteinase K aufgeschlossen, wodurch eine vollständige Zellyse erhalten wurde. Im Vergleich zu Standardprotokollen wurde die 2-fache Konzentration an Lysozym und sogar die 20-fache Konzentration an Proteinase K eingesetzt. Zur Lyse der stark mit Bodensubstanzen kontaminierten Zellen in den braunen Agarose-Nudeln war anfangs der Einsatz dieser hohen Konzentrationen notwendig und wurde aufgrund des Gelingens der voll-

ständigen Lyse beibehalten. Die enormen Kosten, die dieser Schritt mit sich bringt, würden allerdings eine Optimierung auf die minimal notwendige Enzymkonzentration, vor allem der Proteinase K, sinnvoll machen.

Versuche, die in den Agarose-Nudeln vorliegende DNA direkt zu manipulieren, mißlingen allerdings. Auch nach einer zusätzlichen gelelektrophoretischen Aufreinigung war unter Einsatz großer Mengen an Restriktionsenzymen nur eine unvollständige Hydrolyse der DNA möglich. Erst der Einsatz von Polyvinylpyrrolidon in Agarosegelen konnte einen Weg zur Verbesserung der Reinigungsmethode aufzeigen. Unter vergleichender Betrachtung von aufgetrennter DNA aus Bodenproben in Agarosegelen mit und ohne PVP sowie vor und nach der Elektrophorese wurde die Wirkung des PVPs auf die Huminsäuren mit bloßem Auge sichtbar (siehe Abbildung 8). Unter Einsatz von PVP wurde beobachtet, dass der von Huminsäuren verursachte braune Schmier in Laufrichtung der DNA wesentlich verringert war, wobei die Migrationsgeschwindigkeit der DNA kaum verändert wurde. Dies bedeutete, dass das PVP die Komigration der Huminsäuren im elektrischen Feld selektiv verringerte (wie bereits von Young *et al.* 1993 beschrieben), wodurch die DNA von den Huminsäuren getrennt werden konnte. Es zeigte sich, dass nur hochmolekulares Polyvinylpyrrolidon, im Gegensatz zu niedermolekularem PVP, die Komigration von Huminsäuren effizient verhinderte. Aber auch nach diesem Reinigungsschritt war die DNA nur bedingt manipulierbar, da das mitgereinigte PVP inhibitorische Wirkungen auf nachfolgende enzymatische Reaktionen zeigte. Aufgrund dessen wurde ein neuartiger Gelaufbau, das 2-Phasengel, entwickelt. Es bestand aus einer oberen Phase, die neben Agarose auch PVP enthielt und einer unteren Phase, die nur Agarose enthielt. Durch die elektrophoretische Auftrennung in einem Pulsfeldgel wurde die mit Huminsäuren kontaminierte DNA selektiv in der ersten Phase gereinigt und durch die Wanderung in die zweite Phase zusätzlich das PVP von der DNA getrennt. Die Auftrennung mittels Pulsfeldgelelektrophorese war dabei unerlässlich, da der Reinigungseffekt in einer Standardgelelektrophorese wesentlich geringer war. Ein weiterer Vorteil der PFGE lag in der Möglichkeit, neben dem Reinigungseffekt auch gleichzeitig eine notwendige Größen-selektion zu erzielen.

In der Anwendung des 2-Phasengels wurden 4 Punkte, die für eine Klonierung von hochmolekularer DNA aus Bodenproben entscheidend waren, in einem einzigen gelelektrophoretischen Verfahren kombiniert: I) die Huminsäuren wurden von der DNA selektiv getrennt; II) eine Größenselektion wurde durchgeführt; III) das inhibierende PVP wurde eliminiert; IV) Scherkräfte und DNA-Verluste, die mit einer weiteren Gelelektrophorese einhergehen würden, wurden minimiert.

Durch die Anwendung des entwickelten Verfahrens auf Umweltproben anderer Habitate, in denen Huminsäuren in verschiedenen Konzentrationen und



Zusammensetzungen, oder andere Hemmstoffe wie Polysaccharide und bromierte Alkaloide vorlagen, konnte die vielfältige Einsatzmöglichkeit dieser Methode nachgewiesen werden. Es war möglich, reproduzierbar hochmolekulare, reine und klonierbare DNA aus Prokaryontenfraktionen unterschiedlicher Bodenhabitats und mariner Schwämme zu isolieren, während andere Reinigungstechniken auf viele unserer Umweltproben nicht angewendet werden konnten, wie beispielsweise die von Rondon *et al.* (2000) verwendete Methode. Ausserdem war die entwickelte Technik unabhängig von der Art der Zellyse. Reine und klonierbare DNA konnte nach direkter Zellyse (Zhou *et al.*, 1996), die zur Erstellung der Genbank Rud003.AT angewendet wurde, oder nach Lyse der mikrobiellen Fraktion in Agarose-Nudeln, wie hier diskutiert, erhalten werden.

Die Elution der hochmolekularen DNA aus Agarosegelen wurde mit zwei verschiedenen Verfahren getestet. Einerseits wurde die DNA durch hydrolytische Spaltung mit dem Enzym Gelase freigelegt. Bei Einsatz von großen Gelvolumina, was bei dieser Präparationsmethode unumgänglich war, verblieb ein Großteil der DNA in noch vorhandenen Gelresten. Andererseits wurde eine Elektroelution von DNA aus Gelstücken angewendet. Obwohl ebenfalls ein Teil der DNA in den Gelstücken zurückblieb, konnten mit dieser Methode hohe Ausbeuten hochmolekularer DNA erhalten werden.

Bei der Isolierung von hochmolekularer DNA aus Schwamm-assoziierten Prokaryonten (SAPs) hatten sich andere Schwierigkeiten ergeben. Die über das 2-Phasengel gereinigte DNA aus Agarose-Nudeln degradierte stark, wenn sie in Restriktionspuffern oder auch Modifikationspuffer inkubiert wurde. Diese Beobachtungen waren beschränkt auf DNA, die aus Agarose-Nudeln stammte. Ein Grund dafür könnten leichte Kontaminationen der verwendeten Puffer oder Reagenzien sein. Dafür spricht, dass diese Beobachtungen auch bei DNAs von zeitgleich präparierten anderen Umweltproben gemacht wurden. Um die DNA-Fragmentierung zu minimieren, konnte das Verfahren dahingehend verbessert werden, dass die DNA nach der 1. Größenselektion im 2-Phasengel sofort unter hoher Spannung und für kurze Zeit elektroeluiert wurde und die eventuell vorhandenen DNAsen durch eine sofortige Phenol-Chloroform-Isoamylalkohol Extraktion inaktiviert wurden.

### **Herstellung der Umwelt-Genbanken**

Ein wichtiger Schritt der Umweltgenomik ist die Konservierung von möglichst grossen Mengen genomischer Information. Dabei spielen zwei Punkte eine wichtige Rolle: Einerseits die Größe der klonierten Inserts und andererseits die Größe der Genbanken. Zwei Vektoren, BAC- und Fosmid-Vektoren, bieten sich zur Klonierung hochmolekularer DNA besonders an. Während BAC-Vektoren DNA-Inserts variierender Größe und sogar

bis zu 300 kb Fremd-DNA aufnehmen können, besteht der Vorteil von Fosmiden darin, dass über die Verpackung in  $\lambda$ -Phageköpfe eine Größenselektion (32,5-43,5 kb) stattfindet. Bei Versuchen, DNA aus Bodenproben in BAC-Vektoren zu klonieren, hatten sich mehrere entscheidende Nachteile gezeigt. Die Ausbeute der klonierbaren DNA war aufgrund der hydrolytischen Spaltung zur Klonierung über klebrige Enden und der notwendigen Größenselektion sehr gering. Die erreichten Insert-Größen waren mit durchschnittlich nur ca. 15 kb sehr klein. Auch die Klonierungseffizienz, wahrscheinlich als Folge der geringen DNA-Ausbeute, war ebenfalls sehr gering, wodurch die erhaltenen Klonzahlen nur relativ klein waren.

Andererseits hatten sich bei der Klonierung in das hier verwendete Fosmid mehrere Vorteile herauskristallisiert: Durch die Klonierung über glatte Enden konnte eine zusätzliche Hydrolyse und die damit einhergehende Fragmentierung der DNA sowie eine weitere Größenselektion umgangen werden. Dadurch waren die Ausbeuten an hochmolekularer, klonierbarer DNA mit 3  $\mu$ g aus ca. 50 g Boden (RUD) ausreichend groß. Die Folge war eine recht hohe Klonierungseffizienz. Unter Verwendung des Fosmids wurden ca. 25000 Klone in der Genbank Rud001.AQ, die aus der Prokaryontenfraktion des Bodenhabitats RUD erstellt wurde, und ca. 30000 Klone in der Genbank Apae002.AQ, die aus der Prokaryontenfraktion des marinen Schwammes *Aplysina aerophoba* erstellt wurde, konserviert. Es hatte sich gezeigt, dass die Kombination der Präparationsmethode mit der Verwendung des Fosmids eine beliebige Erweiterung der Genbanken möglich machte. Letztendlich machen die Anzahl der erhaltenen Klonzahlen und die verlässliche Größenselektion durch die Verpackung in  $\lambda$ -Phagen die Fosmide zu den prädestinierten Vektoren für das Ablegen von Umweltgenbanken.

Durch die Restriktionsanalyse von einigen zufällig ausgewählten Fosmid-Klonen von beiden Genbanken mit dem Restriktionsenzym *NotI* wurden ausschließlich unterschiedliche Restriktionsmuster erhalten, woraus hervorging, dass weitgehend unterschiedliche Inserts in den Klonen vorliegen sollten, und damit eine große Diversität in den Genbanken abgelegt wurde. Aufgrund der hohen Anzahl an Restriktionsfragmenten der Inserts nach Spaltung mit *NotI*, konnte ein hoher GC-Gehalt der klonierten Genomfragmente abgeleitet werden. Die durchschnittlichen Insert-Größen der analysierten Stichprobe von Fosmid-Klonen lagen bei 38,5 kb der Genbank Rud001.AQ und bei 38 kb der Genbank Apae002.AQ. Somit wurden in der Genbank Rud001.AQ ca. 960 Mio bp und in der Genbank Apae002.AQ 1200 Mio bp des Metagenoms abgelegt. Unter der Annahme, dass ein prokaryotisches Genom aus ca. 4 Mbp besteht, entsprechen 960 Mio bp ca. 200-250 prokaryotischen Genomen. Schätzungen gehen davon aus, dass sich die Gesamtheit aller in Böden vorkommenden unterschiedlichen Genome (Community Genome complexity) auf 15 - 35 Gbp beläuft, was bei einer durchschnittlichen Genomgröße von 4 Mbp ca. 3700 -

8800 Genomäquivalenten entspricht (Torsvik *et al.*, 2002). Das bedeutet, dass in der Genbank Rud001.AQ ca. 3-6% der vorhandenen gesamten genomischen Information des Bodenhabitats RUD ohne Redundanz vorliegen. Wollte man das Metagenom dieses Bodenhabitats komplett sequenzieren oder ein spezifisches Genom aus der Genbank extrahieren und komplett durch Sequenzierung schließen, so bräuchte man bei 6-facher Abdeckung das 100-200 fache der vorliegenden Fosmid-Klone, also ca. 2,5-5 Mio Klone. Dennoch konnten spezifische Genomfragmente von mesophilen nicht-kultivierten Crenarchaeota und von Acidobacteria identifiziert und charakterisiert werden.

### **Suche nach archaealen Genomfragmenten in den Genbanken**

Nach Erstellung von Klonpools, in denen die DNA von jeweils 384 Fosmid-Klone vereinigt vorlagen, wurde mittels Multiplex-PCR die Genbank Rud001.AQ, die aus insgesamt 66 Klonpools bestand, mit archaealen 16S rDNA-spezifischen Oligonukleotiden durchgemustert. Dabei konnten drei positive Klonpools identifiziert werden, d.h. aus 3 der 66 Ansätze der Multiplex-PCR wurde ein Amplifikat der erwarteten Größe von ca. 940 bp erhalten. Unter der Annahme, dass tatsächlich drei Fosmid-Klone in der Genbank Rud001.AQ vorhanden sind, die ein archaeales 16S rRNA-Gen tragen, kann unter Berücksichtigung der theoretisch abgelegten Gesamtgenomzahl von ca. 200 - 250 prokaryotischen Genomäquivalenten der relative archaeale 16S rDNA-Anteil in dieser Genbank auf knapp über 1% abgeschätzt werden. In einer quantitativen Analyse durch *Real-Time* PCR wurde der relative archaeale 16S rDNA-Anteil im Habitat RUD auf 1,7 % beziffert (Ochsenreiter *et al.*, eingereicht), was dem abgeschätzten archaealen Anteil in der Genbank nahe kommt.

In der Genbank Apae002.AQ, die aus der Prokaryontenfraktion von *Aplysina aerophoba* erstellt wurde, konnten überraschenderweise 7 positive Pools mit Archaea-spezifischen Oligonukleotiden erhalten werden, die allerdings noch nicht durch Sequenzierung bestätigt wurden. Aufgrund von Fluoreszenz-*in situ*-Hybridisierungs-Analysen (FISH) wurde bisher davon ausgegangen, dass in diesem Schwamm keine Archaea vorhanden sind (Hentschel, persönliche Mitteilung).

### **Suche nach acidobakteriellen Genomfragmenten**

Mit Acidobacteria-spezifischen Oligonukleotiden konnten in den vier Genbanken aus Bodenhabitaten (Rud001.AQ, Rud003.AT, S247 und S248) 54 positive Klonpools, die aus 384 Fosmid-Klonen bestanden, identifiziert werden. In den Genbanken Rud001.AQ und Rud003.AT, die aus demselben Habitat (RUD) erstellt wurden, konnten aus 22 Klonpools Amplifikate der erwarteten Größe erhalten werden. Bezogen auf die Gesamtzahl der in

den beiden Genbanken vorhandenen Klonpools von 145 entspricht das ca. 15%. Eine quantitative Analyse hatte gezeigt, dass der relative Anteil der 16S rRNA-Operons nur der Gruppe V der Acidobacteria (vgl. Abbildung 26, S.47) im Habitat RUD 8% ausmachten (Ochsenreiter, 2002). Demnach müssten in den beiden Genbanken Rud001.AQ und Rud003.AT, in denen ca. 520 klonierte Genomäquivalente vorliegen (4 Mb), ca. 40 Klone ein acidobakterielles 16S rRNA-Gen der Gruppe V tragen (8% von 520 Genome = 41,6). Auffallend ist der große Unterschied der Anzahl der positiven Pools in den beiden Genbanken. Während in der Genbank Rud001.AQ 26% der Klonpools ein Fragment der erwarteten Größe von ca. 500 bp erhalten wurde (Tabelle 6), waren es in der Genbank Rud003.AT nur 6%. Dies deutete an, dass tatsächlich durch die beiden unterschiedlichen Zellysemethoden, direkte Zellyse bei Rud003.AT und Lyse nach Einbettung der Biomasse in Agarose-Nudeln bei Rud001.AQ, eine Anreicherung oder Abreicherung für bestimmte Mikroorganismen stattgefunden hatte.

Da eine quantitative Analyse bezüglich der Acidobacteria aus den Bodenhabitaten S247 und S248 nicht vorlag, konnten die positiven Pools nicht in Relation mit einer Quantifizierung gebracht werden. Allerdings schien der Anteil an Acidobacteria in diesen Genbanken noch höher zu liegen, wenn man das Verhältnis der positiven Pools zur Gesamtzahl betrachtete. In der Genbank S247 waren 50% und in der Genbank S248 waren 33% der Pools positiv.

In der Genbank Apae002.AQ konnten mit *Acidobacteria*-spezifischen Oligonukleotiden 27 Amplifikate aus 77 Klonpools erhalten werden, das entspricht 35% der Pools. In einer 16S rDNA Studie der Schwamm-assoziierten-Prokaryonten des Schwammes *Aplysina aerophoba* waren von 160 Sequenzen 44 acidobakterieller Herkunft, was 23% entspricht (Fieseler, persönliche Mitteilung).

Tabelle 6: Zusammenfassung der Klonpools, in denen bei der Duchmusterung der Genbanken Rud001.AQ, Rud003.AT, S247 und S248, sowie Apae002.AQ mit Acidobacteria-spezifischen Oligonukleotiden ein Amplifikat der erwarteten Größe von ca. 500 bp erhalten wurde.

Genbank	Verhältnis positive zu vorhandenen Pools	Habitat
Rud001.AQ	26%	Sandökosystem (RUD)
Rud003.AT	6%	Sandökosystem (RUD)
S247	50%	Graslandboden
S248	33%	Mischwaldboden
Apae002.AQ	35%	<i>Aplysina aerophoba</i>

### **Charakterisierung eines Genomfragmentes von nicht-kultivierten Crenarchaeoten**

Aus der Genbank Rud001.AQ konnte mit einer Archaea-spezifischen Sonde der Fosmid-Klon 2914 identifiziert werden. Die Sequenzanalyse zeigte, dass ein zusammenhängendes Genomfragment von 33925 bp eines nicht-kultivierten Boden Crenarchaeota isoliert wurde und zwar aus folgenden Gründen: I) Die phylogenetische Analyse basierend auf kompletten, ribosomalen 16S und 23S RNA Genen zeigte die spezifische Zugehörigkeit zur Gruppe der mesophilen Crenarchaeota; II) Gene mit hoher Ähnlichkeit zu Archaea konnten über das gesamte Fragment identifiziert werden (siehe Tabelle 3); III) der GC-Gehalt und die *codon usage* der vorgeschlagenen Proteine waren über das ganze Insert ähnlich; IV) die abgeleitete Aminosäuresequenz eines DNA-Polymerase Gens zeigte große Ähnlichkeit mit der homologen DNA-Polymerase des unkultivierten marinen Symbionten *Cenarchaeum symbiosum*, dessen funktionelle und biochemische Analyse den nicht-thermophilen Phänotyp dieses Crenarchaeots bestätigt hatte (Schleper *et al.*, 1997b).

Andererseits wurden große Abweichungen von den marinen Crenarchaeota gefunden. Die ungewöhnlich große Lücke im 16S-23S Operon und das Fehlen des GSAT Gens (Glutamat Semialdehyd-Aminotransferase), das in allen marinen Genomfragmenten mit dem ribosomalen Operon verknüpft ist (Abbildung 34; Beja *et al.*, 2002b), deuten die Unterschiede zwischen Crenarchaeota aus Bodenhabitaten und den nahe-verwandten aus marinen Habitaten an. Die für Crenarchaeota untypisch große Lücke der intergenen Region des 16S/23S rRNA-Operons betrug 830 bp und ist damit um das 3-4 fache größer als bei den verwandten marinen Vertretern, die eine intergene Region von 130 - 300 bp aufweisen. Der direkte Vergleich mit marinen planktonischen Crenarchaeota und des Symbionten *Cenarchaeum symbiosum* brachte nur eine Übereinstimmung, nämlich das vorhergesagte DNA-Polymerase Gen, hervor. Auch die Kodierungsdichte des Klons 2914, nur 69% der Sequenz codierten für RNA oder Proteine, war wesentlich geringer als bei anderen Archaea, wo die Kodierungsdichte in der Regel deutlich über 80% liegt. Eine weitere große, nicht-kodierende Region von 2787 bp wurde zwischen ORF 10 und fixA (ORF 11) identifiziert. Weder Gene noch strukturelle Elemente, wie repetitive Sequenzen, konnten in dieser Region detektiert werden.

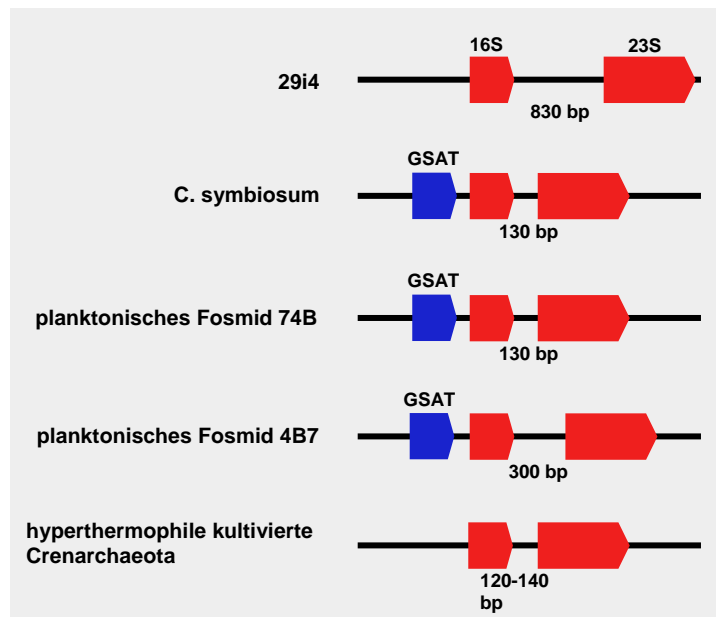


Abbildung 34: Ribosomale Genorganisation in 4 mesophilen Crenarchaeota im Vergleich mit dem ribosomalen Operon von kultivierten, hyperthermophilen Crenarchaeota

Die genomischen Informationen von Genen auf Klon 2914 geben erste Hinweise über die metabolischen Eigenschaften von mesophilen Crenarchaeota aus Böden und könnten als Basis für künftige funktionelle Genomstudien dienen. Neben Genen, die für Proteine mit "Haushalts"-Funktion codieren (Replikation, Aminosäuremetabolismus), waren zwei  $\alpha/\beta$  Hydrolasen vorhanden, die spezifisch für Crenarchaeota aus Böden zu sein scheinen. Eine davon codiert für ein Protein, das die Polymerisierung von Hydroxyacyl Coenzym A Polyester zu Polyhydroxyalkanoat unter Freisetzung von Coenzym A katalysiert. Es handelt sich hierbei um ein PhaC-homologes Protein der Typ III Polyhydroxyalkanoat-Synthasen (Lütke-Eversloh und Steinbüchel, persönliche Mitteilung), das 41,8% Sequenzidentität mit einer homologen Sequenz eines phototrophen Bakteriums (*Thiocystis violacea*) zeigte. Mitglieder der Typ III PHA-Synthasen Subfamilie bestehen im Unterschied zu Typ I und Typ II PHA Synthasen aus zwei Untereinheiten, die von phaC und phaE codiert werden (Liebergesell *et al.*, 1994). Überraschenderweise war das Gen für die zweite Untereinheit phaE nicht auf dem Klon 2914 codiert. Expressionsversuche in *E. coli* sowie in den PHA-Mutanten, *Ralstonia eutropha* PHB<sup>-4</sup> und *Pseudomonas putida* Gpp104, die u.a. mit phaE von *Thiocapsa pfennigii* co-exprimiert werden sollten, waren bislang nicht erfolgreich (Lütke-Eversloh und Steinbüchel, persönliche Mitteilung). Eine Komplementierung der Mutationen, sowie Enzymaktivität in Rohextrakten konnte nicht erhalten werden. Obwohl bereits PHA-Akkumulierung in Halobacteria nachgewiesen wur-

den (Kirk und Ginzburg, 1972), ist dies die erste bekannte archaeale Sequenz einer PHA-Synthase.

Das Operon, das für FixABCX codierte, zeigte große Sequenzähnlichkeiten mit Flavoprotein-haltigen Elektronentransportketten, wie sie häufig in symbiotischen, stickstofffixierenden Bacteria gefunden wurden. Eine detaillierte phylogenetische Analyse zeigte, dass das vorgeschlagene FixA Protein von 2914 einen größeren Verwandtschaftsgrad mit archaealen FixA Homologen und nicht mit bakteriellen FixA Proteinen oder paralogen der Elektronentransport Flavoproteine (ETF) besaß. Aufgrund dessen ist es unwahrscheinlich, dass die *fixABCX*-Gene von 2914 über horizontalen Gentransfer von z.B. stickstofffixierenden Bacteria, die auch in demselben Boden vorkommen könnten, akquiriert wurden. Sie scheinen eher von einem gemeinsamen archaealen Vorfahren abzustammen.

### **Charakterisierung eines Genomfragmentes von nicht-kultivierten Acidobacteria**

Mit *Acidobacteria*-spezifischen Sonden konnten in den beiden Genbanken aus dem Bodenhabitat RUD bislang 83 Fosmid-Klone identifiziert werden, die in einer PCR-Analyse mit *Acidobacteria*-spezifischen Oligonukleotiden ein Fragment der erwarteten Größe ergaben. Einer der 9 bisher durch phylogenetische Analyse auf 16S rRNA Basis eindeutig identifizierten Klone, der Klon 32D1, wurde vollständig sequenziert und analysiert.

Der Klon enthielt ein Insert von 35152 bp. Die Analyse des GC-Gehaltes zeigte einen konstanten Wert von 63,7% bis zum ribosomalen Operon, wo er auf 55,5% abfällt. Diese Daten deuten daraufhin, dass das Genomfragment einen zusammenhängenden Abschnitt aus einem einzigen Organismus darstellte. Größere Lücken von nicht-codierenden Bereichen wurden nicht identifiziert. Die phylogenetische Analyse des 16S rRNA-Gens konnte die spezifische Zugehörigkeit zur Gruppe V der Acidobacteria zeigen. Eine phylogenetische Analyse der Formamidopyrimidin-DNA Glycosylase (FAPY) konnte die nahe Verwandtschaft mit einer homologen Proteinsequenz eines weiteren acidobakteriellen Umweltklons 41B15 bestätigen (Ochsenreiter, 2002). Auf dem Klon 32D1 ist ein vollständiges ribosomales Operon, bestehend aus 16S rRNA-Gen, tRNA<sup>Ile</sup>, tRNA<sup>Ala</sup>, 23S rRNA-Gen und 5S rRNA-Gen codiert. Diese Operonstruktur ist häufig unter den Bacteria zu finden (Chen *et al.*, 2000).

Etwa 77% der Proteingene, die über das ganze Insert verbreitet vorlagen, zeigten große Ähnlichkeit mit Proteinen von Proteobacteria. Durch vergleichende Genomanalyse wurde ein kolinearer Bereich mit *Rhodopseudomonas palustris* identifiziert, der 10 kb überspannte und 8 offene Leserahmen enthielt. Die abgeleiteten Aminosäuresequenzen aller 8 ORFs zeigten in paarweisem Alignment mit der Sequenz von *Rhodopseudomonas*

*palustris* 63,9% identische Positionen über eine Länge von 2843 AS. Neben Genen mit "Haushalts"-Funktion (PII Uridyltransferase ORF 1, *Penicillin-binding* Protein ORF 2) wurde auf dem kolinearen Bereich eine Zink-Metallopeptidase und eine intrazelluläre Polyhydroxybutyrat-Depolymerase (iPHB-Depolymerase) identifiziert. Nur 22 Homologe der iPHB-Depolymerase konnten in den Datenbanken identifiziert werden, wobei sie ausschließlich von Proteobacteria stammten. Auffallend war das häufige paarweise Auftreten dieses Gens mit einem Gen für eine Zink-Metallopeptidase, wie das auch bei 32D1 vorgefunden wurde. Diese strukturelle Auffälligkeit könnte eine zusammenhängende Funktion implizieren. Eine gemeinsame Expression in einem adäquaten Modellorganismus könnte diese Hypothese untermauern.

Das ebenfalls im kolinearen Bereich liegende Zwei-Komponentensystem (ORF 8 und ORF 9) zeigte große Ähnlichkeit mit Zwei-Komponentensystemen der  $\alpha$ -Proteobacteria. Durch Alignierung der konservierten Motive, die für die C-terminalen Bereiche der Histidinkinase-typischen Motive beschreiben, konnte die auf 32D1 vorhandene Histidinkinase eindeutig der Subfamilie HPK 3e zugeordnet werden. Die potentielle Phosphorylierungsstelle und das typische Helix-Turn-Helix Motiv des *Response*-Regulators konnte durch Alignierung und Sequenzvergleich identifiziert werden. Die phylogenetische Untersuchung des *Response*-Regulators zeigte deutlich die Zugehörigkeit zur Gruppe I der homologen *Response*-Regulatoren, die ausschließlich von *Rhizobiaceae* der  $\alpha$ -Proteobacteria stammten (siehe Abbildung 35). Einige homologe *Response*-Regulatoren wurden funktionell bereits gut charakterisiert. Es konnte gezeigt werden, dass diese an der Regulation unterschiedlicher Operons in verschiedenen Spezies der  $\alpha$ -Proteobacteria beteiligt sind, wie z.B. an der Stickstofffixierung in *Bradyrhizobium japonicum* (Bauer *et al.*, 1998), der CO<sub>2</sub>-Fixierung in *Sinorhizobium meliloti* (Emmerich *et al.*, 2000a) und der Photosynthese in *Roseobacter denitrificans* (Sganga und Bauer, 1992; Masuda *et al.*, 1999). Ausserdem konnte gezeigt werden, dass die homologen *Response*-Regulatoren von *Rhodobacter sphaeroides* und von *Sinorhizobium meliloti* *in vivo* die Expression des *fixR-nifA* Operons (Stickstofffixierung) in *Bradyrhizobium japonicum* RegR Mutanten komplementieren (Emmerich *et al.*, 2000a). Aufgrund dieser Aktivitäten wird angenommen, dass es bei diesem *Response*-Regulator um eine Art "Master"-Regulator handelt, der in  $\alpha$ -Proteobacteria unterschiedliche Stoffwechselwege reguliert. Die große Sequenzähnlichkeit des *Response*-Regulators des acidobacteriellen Klons 32D1 deutet auf eine ähnliche Funktion in Acidobacteria hin und könnte durch Expression und funktionelle Analyse in einem genetisch zugänglichen System, wie *Bradyrhizobium japonicum* untersucht werden.



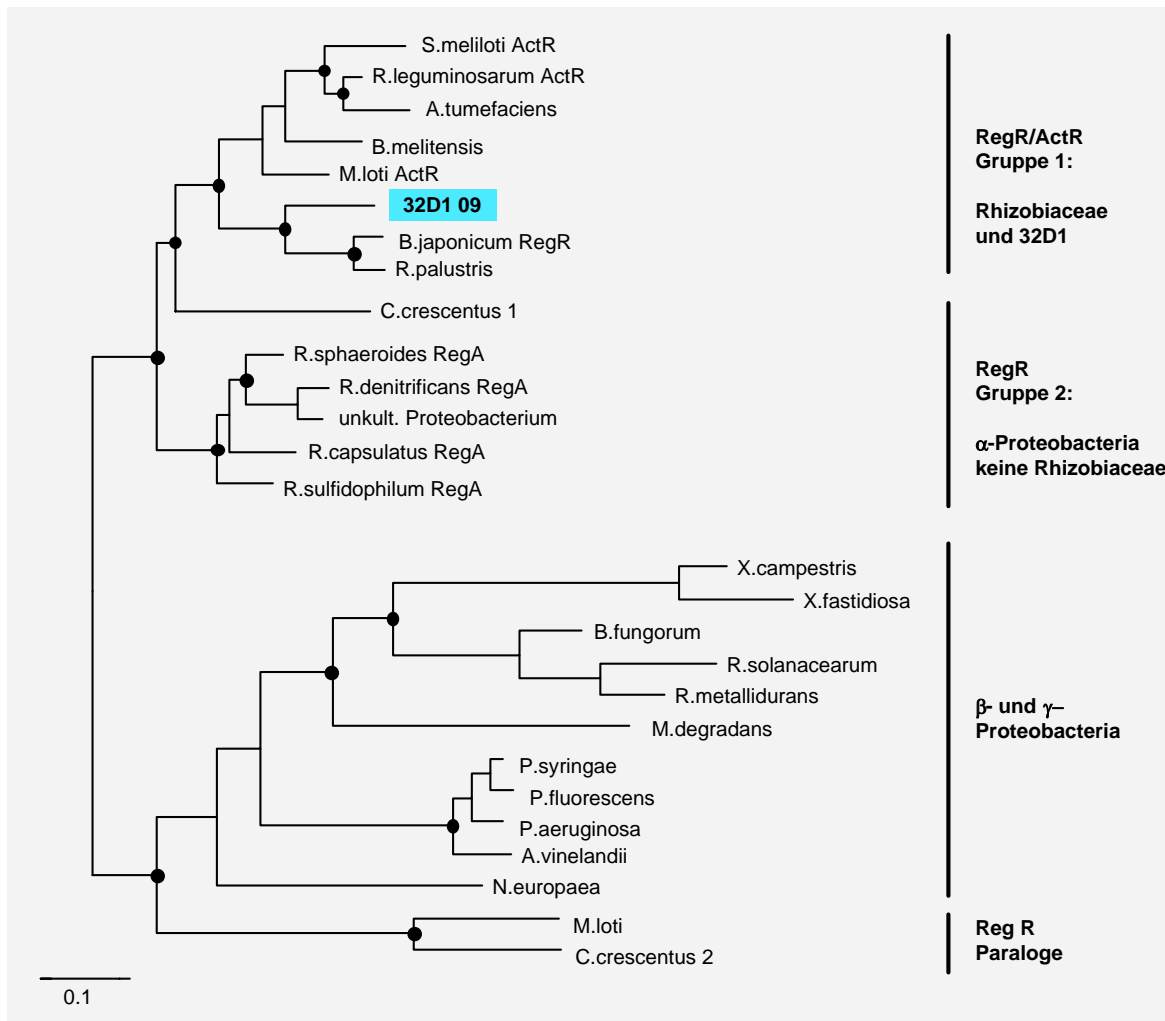


Abbildung 35: Phylogenetische Analyse des *Response*-Regulator Proteins von 32D1 (ORF 9) sowie den 26 ähnlichsten Proteinen (blastp) aus dem Genom anderer Mikroorganismen. Schwarze Punkte markieren Verzweigungsordnungen, die mit Maximum Likelihood und Neighbour Joining Methoden erhalten wurden. Die Homologen der "Master" Regulatoren bilden eine monophyletische Gruppe und sind ausschließlich unter den  $\alpha$ -Proteobacteria (Gruppe 1 und Gruppe 2) zu finden. Die phylogenetische Rekonstruktion wurde von Torsten Ochsenreiter durchgeführt. Abkürzungen: *R.palustris*, *Rhodopseudomonas palustris*; *B.japonicum*, *Bradyrhizobium japonicum*; *M.loti*, *Mesorhizobium loti*; *B.melitensis*, *Brucella melitensis*; *A.tumefaciens*, *Agrobacterium tumefaciens*; *R.leguminosarum* ActR, *Rhizobium leguminosarum* bv. viciae; *S.meliloti* ActR, *Sinorhizobium meliloti*; *R.denitrificans* RegA, *Roseobacter denitrificans*; *R.sphaeroides* RegA, *Rhodobacter sphaeroides*; *R.capsulatus* RegA, *Rhodobacter capsulatus*; *R.sulfidophilum* RegA, *Rhodovulum sulfidophilum*; *C.crescentens*, *Caulobacter crescentens* CB15; unkult. Proteobacterium, unkultiviertes Proteobacterium - mariner Umweltklon; *P.aeruginosa*, *Pseudomonas aeruginosa* PAO1; *B.fungorum*, *Burkholderia fungorum*; *R.solanacearum*, *Ralstonia metallidurans*; *M.degradans*, *Micrlobulbifer degradans* 2-40; *X.fastidiosa*, *Xylolla fastidiosa* Ann-1; *X.campestris*, *Xanthomonas campestris* pv. campestris str. ATCC 33913; *P.syringae*, *Pseudomonas syringae* pv. syringae B728a; *N.europaea*, *Nitrosomonas europaea*; *P.fluorescens*, *Pseudomonas fluorescens*; *R.sulfidophilum*, *Rhodovulum sulfidophilum*.

Die vergleichende Genomanalyse brachte ein weiteres interessantes Phänomen zu Tage, das in der Literatur noch nicht beschrieben wurde. Auffallend war das häufige Vorkommen des konserviert hypothetischen Proteins (ORF 10) in Verknüpfung mit dem Zwei-Komponentensystem, das ausschließlich in der Gruppe der Rhizobiaceae und auf dem acidobakteriellen Klon 32D1 zu finden ist. Diese Regelmäßigkeit gewinnt unter Einbeziehung der phylogenetischen Analyse der homologen *Response*-Regulatoren, insofern an Bedeutung, als dass das konserviert hypothetische Protein ausschließlich mit Homologen der *Response*-Regulatoren der Gruppe I verknüpft vorkommt. Eine solche Konservierung macht einen funktionellen Zusammenhang wahrscheinlich. Diese Studie könnte einen Anstoß bieten, bei der Analyse der "Master"-Regulatoren dieses hypothetische Protein miteinzubeziehen und dessen Funktion festzustellen.

### **Horizontaler Gentransfer zwischen Rhizobiaceae und Acidobacteria ?**

Bei Betrachtung des kolinearen Bereichs, der ausschließlich bei 32D1 und *Rhodopseudomonas palustris* zu finden ist, drängt sich die Annahme auf, dass hier ein horizontaler Gentransfer zwischen Acidobacteria und Rhizobiaceae stattgefunden hat. Phylogenetisch konnte 32D1 auf 16S rDNA Basis eindeutig den Acidobacteria, die nur ca. 80% identische Positionen mit  $\alpha$ -Proteobacteria teilen, zugeordnet werden (Abbildung 36). Andererseits zeigte die phylogenetische Analyse des abgeleiteten *Response*-Regulators eindeutig die Zugehörigkeit zu den Homologen der Rhizobiaceae ( $\alpha$ -Proteobacteria; siehe Abbildung 35) und auch die weiteren Gene des kolinearen Bereichs zeigten größte Ähnlichkeiten mit  $\alpha$ -Proteobacteria. Insbesondere die Sensor-Histidinkinase (ORF 8) und das konserviert hypothetische Protein (ORF 10), das ausschließlich bei  $\alpha$ -Proteobacteria vorkommt und nur bei den Rhizobiaceae und 32D1 verknüpft mit dem *Response*-Regulator vorliegt, deuten auf eine nahe Verwandtschaft hin. Auch Homologe der iPHB-Depolymerase (ORF 5) sind ausschließlich bei Proteobacteria zu finden. Einen weiteren Hinweis auf einen horizontalen Gentransfer gibt die Konservierung des kolinearen Bereichs bei 32D1 und *Rhodopseudomonas palustris*. Andere  $\alpha$ -Proteobacteria, die in der 16S rDNA Phylogenie tiefer als *Rhodopseudomonas palustris* abzweigen, besitzen zwar das konservierte Zwei-Komponentensystem, aber weitere Gene dieses Bereichs sind in ihrer Anordnung nicht konserviert (Abbildung 36). Alternativ zu horizontalem Gentransfer könnte man auch eine Koevolution des kolinearen Bereichs in Acidobacteria und Proteobacteria in Betracht ziehen. Allerdings müsste man dann eine sehr langsame Evolution von RegR Homologen annehmen, um die in Abbildung 36 gezeigte Struktur erklären zu können. Der kolineare Bereich auf 32D1 und *Rhodopseudomonas palustris* würde dann von einem gemeinsamen Vorfahren abstammen

und die erwähnten nahen Verwandten hätten durch Genverlust oder Genomumlagerungen diese Struktur verloren. Gefestigtere Aussagen werden aber erst gemacht werden können, wenn weitere Sequenzen von verwandten Stämmen von 32D1 und von *Rhodospseudomonas palustris* verfügbar sind.

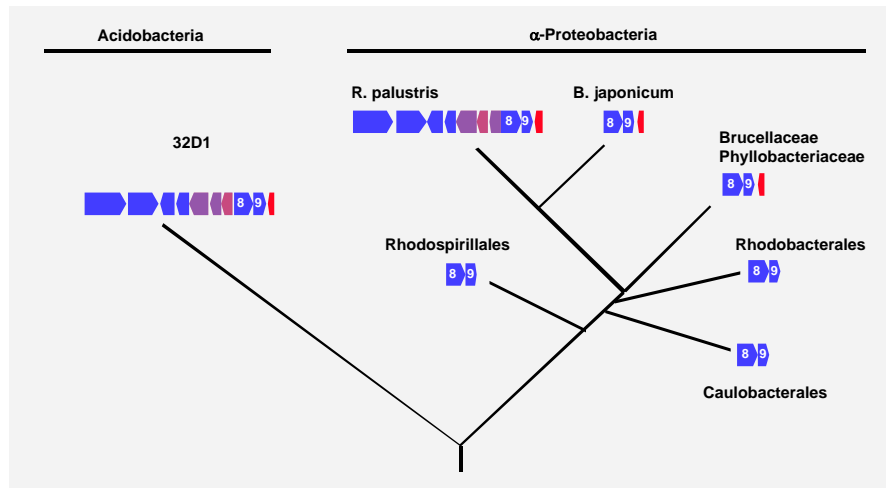


Abbildung 36: Schematische Darstellung der 16S rDNA Phylogenie von 32D1 der Acidobacteria und einiger Gruppen der α-Proteobacteria mit schematischer Darstellung des Zwei-Komponentensystems (8 und 9) und weiteren kolinearen Genen soweit vorkommend.

## Ausblick

Techniken zur Klonierung von hochmolekularer DNA direkt aus marinen Umweltproben wurden vor kurzem bereits beschrieben und stellte eine wichtige Basis für erste umweltgenomische Studien an marinen Mikroorganismen dar (Preston *et al.*, 1996; Schleper *et al.*, 1998; Beja *et al.*, 2000b). Das in dieser Arbeit entwickelte Verfahren zur Isolierung von hochmolekularer DNA ermöglicht nun die Ausdehnung der Umweltgenomik auf weitere Habitate, wie Böden und Sedimente, die durch Kontamination mit Huminsäuren und verwandten Substanzen bislang nicht zugänglich waren.

Die bereits identifizierten Genomfragmente können zum Design von weiteren Sonden verwendet werden, um in den vorliegenden Genbanken zusätzliche genomische Informationen von mesophilen nicht-kultivierten Archaea und Acidobacteria und natürlich auch von anderen Mikroorganismengruppen zu erhalten. Es ist zu erwarten, dass durch die Analyse weiterer Genomfragmente physiologische Eigenschaften und Metabolismen der Organismen erhalten werden. Einerseits können so Rückschlüsse auf die Funktion dieser Mikroorganismen in einem Habitat gezogen werden und andererseits könnten daraus Substratspezifitäten, die zur Kultivierung dieser Mikroorganismen im Labor dienen

identifiziert werden. Tatsächlich wurde bereits vor kurzem bei der Analyse eines zweiten Genomfragments von nicht-thermophilen Crenarchaeota ein Gen identifiziert, das auf eine Rolle dieser Archaea in der Umgebung von Stickstoffverbindungen im Boden hinweist (Treusch und Schleper, persönliche Mitteilung).

Neben der Erforschung von Genomstruktur und Physiologie von unkultivierten Mikroorganismen liegt noch ein weiteres interessantes Potential in der Umweltgenomik. Über Aktivitäts-Assays oder durch Screening mit geeigneten Gensonden können in den Genbanken biotechnologisch relevante Proteingene identifiziert werden (Schulze et al., 2002). Die in dieser Arbeit erstellten Genbanken haben bereits eine Vielzahl neuer Protein-Kandidaten für biotechnologische Anwendungen hervorgebracht (Eck, persönliche Mitteilung, BRAIN AG).

Wenn man berücksichtigt, dass ca. 5000 unterschiedliche Spezies in einem Bodenhabitat leben, was ca. 20 Gbp an genomischem Gehalt entspricht (Torsvik *et al.*, 2002), so erscheint die vorliegende Genbank mit ca. 0,96 Gbp nur als kleine Stichprobe. Noch scheint es unvorstellbar, dass das Metagenom eines Habitats mit ca. 5000 unterschiedlichen Spezies insgesamt erfasst und sequenziert werden könnte. Wenn man allerdings den rasanten Fortschritt der letzten 10 Jahre in der molekularen Biologie betrachtet, so relativiert sich die enorme Anzahl von 400000 Fosmid-Klonen, die für eine einfache Abdeckung eines Bodenhabitats benötigt werden. In 400000 Fosmid-Klonen sind ca. 15 Gbp konserviert und das entspricht nur ca. dem 5-fachen des haploiden menschlichen Genoms (3 Gbp).

Durch die zunehmende Akkumulation und Analyse von weiteren genomischen Informationen werden kurzfristig physiologische Eigenschaften von bestimmten Mikroorganismen erhalten werden und längerfristig können ökologische Zusammenhänge und spezielle Rollen von Mikroorganismen in komplexen Populationen aufgeklärt werden.

## E. Material und Methoden

### 1. Material

#### 1.1. Chemikalien und Hilfsmittel

Agar	Difco, Detroit
Agarose (SeaKem ME)	FMC Bio Products, Rochland
Agarose, Low melting (SeaPlaque GTG)	FMC Bio Products, Rochland
Ampicillin	Roche Diagnostics
Ammoniumperoxodisulfat (APS)	Sigma, St. Louis
Chloramphenicol	Roche Diagnostics
dNTP's	Roche Diagnostics
EpiFOS Kit	Epicentre Technologies, Madison
Ethidiumbromid	Sigma, St. Louis
Hefe-Extrakt	Difco, Detroit
IPTG (Isopropylthiogalactopyranosid)	Eurogentec
Phenol (Roti-)	Roth
Polyvinylpyrrolidon (PVP)	Sigma, St. Louis
QIAGEN Plasmid Kits	QIAGEN, Düsseldorf
Concert™ Rapid Plasmid	Gibco
SequiTherm EXCEL II Long-Read DNA Sequencing Kits-LC	Epicentre Technologies, Madison
Sodiumdodecylsulfat (SDS)	Roth
TOPO™ Blunt Cloning Kit	Invitrogen, Carlsbad CA
Trypton	Oxoid, Hameln
Ultrafree-DA	Millipore, Bedford
X-Gal (5-Bromo-4-Chloro-3-Indoyl- β-D-galactopyranosid)	GERBU, Gaiberg

Die hier nicht aufgeführten Chemikalien stammten von AppliChem, Roche Diagnostics, Merck, Roth und Sigma. Nährmedienbestandteile wurden von den Firmen Bacto, Difco, Fluka, Merck, Oxoid, Sigma und Serva bezogen.

## 1.2. Synthetische Oligonukleotide

Die Oligonukleotide wurden bei ARK (Darmstadt) hergestellt. Sie wurden in lyophilisierter Form geliefert, anschließend in Wasser aufgenommen und bei -20 °C gelagert. Oligonukleotide für Sequenzreaktionen wurden HPLC-gereinigt von MWG (Ebersberg) bezogen und bei -20 °C gelagert.

Tabelle 7: Verwendete synthetische Oligonukleotide für PCR-Amplifikationen und Sequenzierung.

Name	Sequenz 5'→3' *	Länge (bp)	Anlagerung (°C)	Zweck
Arch 20F	TTCCGGTGGATCC <b>Y</b> GCCGGA	20	56	PCR
Arch 958R	<b>Y</b> CCGGCGTTGAMTCCAATT	19	56	PCR
Ac31F	GATCCTGGGCTCAGAATC	17	50	PCR
Univ533R	TTACCGCGGCKGCTGGCAC	19	50	PCR
Univ1492R	GGTTACCTTGTACGACTT	19	54	PCR
M-20-universal	TTGTAAAACGACGGCCAGT	19	55	Seq. **

\* Fett gedruckte Buchstaben zeigen degenerierte Positionen, die nach der Standardabkürzung benannt wurden. Weitere Angaben beziehen sich auf die Längen der Oligonukleotide in Basenpaaren (bp) und die verwendeten Anlagerungstemperaturen (°C).

\*\* Das Oligonukleotid für Sequenzierungen war am 5'-Ende IRD 800 Fluoreszenz-markiert.

## 1.3. Enzyme

DAP <sub>GoldStar</sub> Polymerase	Eurogentec, Seraing
GELase	Epicentre Technologies, Madison
Proteinase K	Merck
Restriktionsendonukleasen	MBI Fermentas
SequiTherm EXCEL II DNA Polymerase	Biozym, Oldendorf
T4 DNA-Polymerase	Epicentre Technologies, Madison
T4 Polynukleotid-Kinase	Epicentre Technologies, Madison
T4 DNA-Ligase	Epicentre Technologies, Madison
DIG DNA Labeling Kit	Roche Diagnostics, Mannheim

## 1.4. DNA-Größenstandards

100 bp DNA Ladder, MBI Fermentas, DNA-Fragmente der Größen: 1031, 900, 800, 700, 600, 500, 400, 300, 200, 100 und 80 bp; Konzentration: 500 µg/ml in 10 mM Tris/HCl (pH 7,5) und 1 mM EDTA.

1 kb DNA Ladder, New England Biolabs, DNA-Fragmente der Größen: 10002, 8001, 6001, 5001, 4001, 3001, 2000, 1500, 1000, 517 und 500 bp; Konzentration: 500 µg/ml in 10 mM Tris/HCl (pH 8,0) und 1 mM EDTA.

Lambda DNA/*EcoRI*+*HindIII* Marker, MBI Fermentas, DNA-Fragmente der Größen: 21226, 5148, 4973, 4268, 3530, 2027, 1904, 1584, 1375, 947, 831 und 564; Konzentration: 500 µg/ml in 10 mM Tris/HCl (pH 7,6) und 1 mM EDTA.

MidRange PFG Marker I, New England Biolabs, DNA-Fragmente der Größen: 291, 276, 257,5, 242,5, 227,5, 209, 194, 179, 160,5, 145,5, 130,5, 112, 97, 82, 63,5, 48,5, 33,5, 15 kb; Konzentration: 50 µg/ml in 10 mM Tris/HCl (pH 8,0), 1 mM EDTA und 50% Glycerin.

Yeast Chromosome Marker, New England Biolabs, DNA Fragmente der Größen: 1900, 1640, 1120+1100, 945, 915, 815, 785, 745, 680, 610, 555, 450, 375, 295, 225.

### 1.5. Medien

Luria Bertani- (LB-) Medium

für <i>E. coli</i> (pro Liter):	10 g Trypton; 5 g Hefeextrakt; 10 g NaCl
für Festmedium:	Agar 15 g/l
Antibiotikum:	Ampicillin 100 mg/l
für Blau/Weiß-Selektion:	X-Gal 40 mg/l (Substrat für $\beta$ -Galaktosidase), IPTG 40 mg/l (Induktion der $\beta$ -Galaktosidase)
für Glycerinkulturen:	7,5% Glycerin
SOC-Medium für <i>E. coli</i> :	2% Bacto Trypton; 0,5% Hefeextrakt; 10 mM NaCl; 2,5 mM KCl; 20 mM Glucose

### 1.6. Puffer und Lösungen

DNA-Probenpuffer:	50% Saccharose w/v; 0,25% Bromphenolblau; 0,1 M EDTA, pH 8,0
SM-Phagen- Verdünnungspuffer:	50 mM Tris/HCl, pH 7,5; 10 mM MgSO <sub>4</sub> ; 100 mM NaCl 0,01% Gelatine
TAE-Puffer:	40 mM Tris/Acetat pH 8,0; 2 mM EDTA
TE-Puffer:	20 mM Tris/HCl pH 8,0; 1 mM EDTA
10x TBE-Puffer:	890 mM Tris; 889 mM Borsäure; 25 mM EDTA
Filter- Hybridisierungslösung:	50% Formamid; 5x SSC; 2% Blockierungslösung; 0,02% SDS; 0,1% Lauroylsarcosin
Depurinierungslösung:	0,25 M HCl
Denaturierungslösung:	1,5 M NaCl; 0,5 M NaOH
Neutralisierungslösung:	1,5 M NaCl; 0,5 M Tris/HCl pH 7,5

20x SSC:	3 M NaCl; 0,3 M Natriumcitrat pH 7,0
DIG-Waschpuffer 1:	0,1 M Maleinsäure; 0,15 M NaCl auf pH 7,0 mit NaOH-Plätzchen
DIG-Waschpuffer:	0,3% Tween-20 in DIG-Puffer 1
DIG-Puffer 2:	1% Blockierungsreagens in 1x DIG-Puffer 1
DIG-Puffer 3:	100 mM Tris/HCl pH 9,5; 100 mM NaCl; 50 mM MgCl <sub>2</sub>

## 2. Methoden

### 2.1. Beschreibung der Bodenprobe

Die Bodenprobe RUD wurde aus Rotböhl, einem Standort bei Darmstadt, der zum Naturraum Hessische Rheinebene gehört, mit einem Folienprobennehmer entnommen (Abbildung 4). Der Boden gehört zu einem Naturschutzgebiet und wurde bereits unter geobotanischen Gesichtspunkten analysiert (Storm *et al.*, 1998). Es handelt sich hierbei um einen Boden eines Sandökosystems, der sich durch schwaches Nährstoffbindevermögen auszeichnet. Lehm- und Schlickbestandteile, sowie organische Substanzen (2,6% loss-on-ignition) sind nur zu einem geringen Anteil vorhanden. Charakteristisch für diesen Boden ist die 50-100%ige Deckung mit *Stipia capillata* und *Festuca ovina s.l.*. Der Stickstoffgehalt (1,2 mg/kg) und der Phosphatgehalt (10 mg/kg) ist für einen kalkreichen Wiesenboden erwartungsgemäß niedrig. Da für die Untersuchung der im Boden vorkommenden nicht-kultivierten *Crenarchaeota* nur die aeroben Bodenschichten von Interesse waren, wurden nur die oberen 10 cm des Bohrkerns verwendet.

### 2.2. Isolierung von hochmolekularer, klonierbarer DNA aus Bodenproben und marinen Invertebraten

#### 2.2.1. Gewinnung der mikrobiellen Fraktion

Um Mikroorganismen von Bodenpartikeln zu trennen, wurden 50 g Boden in 6-fachem Volumen (w/v) 1x Bodenpuffer (20 mM Tris-HCl, pH 8; 10 mM  $\epsilon$ -Aminocaprinsäure; 10mM EDTA, pH 8) 15 Stunden bei 200 rpm in einem 1l Erlenmeyerkolben aufgeschlämmt.

Der aufgeschlämmt Boden wurde durch einen Faltenfilter (Schleicher und Schuell; Faltenfilter 591 1/2) filtriert, um grobe Bodenbestandteile zu eliminieren. Das Filtrat wurde durch Zentrifugation in 500 ml BSA-Zentrifugenröhrchen (Sorvall) bei 11000 rpm für 30 min (GS-3 Sorvall Rotor) bei 4°C sedimentiert. Das Sediment wurde anschließend in 5 ml 1x Bodenpuffer resuspendiert und auf Eis gestellt.



### 2.2.2. Herstellung von Agarose-Nudeln

Es wurden 5 ml einer 4%-igen Agaroselösung (low-melting GTG SEAPlaque, FMC Bioproducts) in bi-deionisiertem H<sub>2</sub>O hergestellt und auf 45°C äquilibriert. Die im vorigen Schritt aus Boden erhaltene Zellsuspension wurde auf 45°C erwärmt. Die Agaroselösung und die Zellsuspension wurden im Verhältnis 1:7 gemischt und durch kurzes starkes Schütteln homogenisiert. Das Gemisch wurde dann in sterile 1 ml Einwegspritzen (Omni-fix F, B. Braun) aufgezogen, die dann auf Eis gelegt wurden, damit die Agarose polymerisieren konnte. Die Agarosekonzentration in den Plugs betrug 0,5%.

Agarose-Nudeln der Schwamm-assoziierten-Prokaryonten von *Aplysina aerophoba* und *Axinella polypoides* wurden durch das Mischen unterschiedlicher Zellkonzentrationen ( $2 \times 10^9$ ,  $1 \times 10^9$ ,  $0,5 \times 10^9$ ,  $0,2 \times 10^9$ ,  $0,04 \times 10^9$  Zellen/ml) mit 1% LM Agarose im Verhältnis 1:1 hergestellt.

### 2.2.3. Zellyse

Die erhaltenen Agarose-Nudeln wurden bei 37°C für 1 Stunde in 50 ml Lysozympuffer (100 mM EDTA pH 8, 10 mM Tris-HCl pH 8, 50 mM NaCl, 0,2% Desoxycholate, 1% Lauroylsarcosin und 1 mg/ml Lysozym) inkubiert. Anschließend wurden die Agarose-Nudeln in 50 ml ESP Puffer (0.5 M EDTA pH 8, 1% Lauroylsarcosin, 2 mg/ml Proteinase K) transferiert und bei 50°C für 24 Stunden rotierend inkubiert. Nach 24 Stunden wurde der ESP Puffer erneuert und die Agarose-Nudeln nochmals für 24 Stunden inkubiert. Die Proteinase K wurde durch dreimaliges Inkubieren für jeweils eine Stunde in TE plus 1 mM PMSF (Phenylmethansulfonylfluorid, Merck) inaktiviert. Es folgte das Auswaschen des PMSF durch dreimaliges Dialyse in TE für jeweils eine Stunde.

### 2.2.4. Gelelektrophoretische Reinigung der hochmolekularen DNA aus Boden:

#### 2-Phasengel

Das 2-Phasengel bestand aus zwei unterschiedlichen Phasen. Das erste Drittel des Gels enthielt 1% PFGE Agarose (Sigma A-2929) und 2% PVP (Polyvinylpyrrolidon VP-360, Sigma). Die zwei unteren Drittel enthielten nur 1% PFGE Agarose.

Drei Agarose-Nudeln wurden nebeneinander gegen den Kamm in einer PFGE (Pulsfeldgelelektrophorese) Gelgießkammer gelegt. Das obere Drittel der Gelgießkammer wurde durch einen Spacer abgetrennt. Die Agarose-Nudeln wurden mit einer 0,5x TAE Lösung, die 1% Agarose und 2% PVP enthielt, übergossen. Nach Polymerisierung der oberen Phase wurde der Spacer entfernt und der Rest der Kammer mit einer 0,5x TAE Lösung, die nur 1% Agarose enthielt, gefüllt.

Die hochmolekulare DNA in den Agarose-Nudeln wurde in einer PFGE (CHEF-DR™, BIORAD) unter unterschiedlichen Bedingungen aufgetrennt. Die Bedingungen für die PFGE wurden im 2-Phasengel so gewählt, dass die gesamte DNA durch die obere Phase migriert und die DNA  $\geq 90$  Kb (Kompressionszone) in der unteren Hälfte des Gels zu finden war (4,5 V/cm, 1-4 s, 36 h, 10°C). Die Ränder des Gels, auf denen jeweils der Größenstandard (PFG MidRange Marker, NEB) aufgetragen war, wurden so abgeschnitten, dass jeweils ein kleiner Teil der DNA enthalten war. Nach Ethidiumbromid Färbung der äußeren Ränder wurde die Kompressionszone mit einem Skalpell markiert und das Gesamtgel wieder zusammengesetzt. An der markierten Stelle wurde die ungefärbte und unbestrahlte DNA ausgeschnitten. Danach folgte je nach Umweltprobe die 2. Größenselektion oder die Elektroelution.

Die 2. Größenselektion wurde unter den Bedingungen 6 V/cm, 0,1-1s, 9 h, 10°C durchgeführt, wodurch sich Fragmente von  $\geq 50$  kb in der Kompressionszone befanden.

### **2.2.5. Restriktion von hochmolekularer DNA in Agarose-Nudeln**

Die Restriktion in Agarose-Nudeln wurde mit ca. 20 mg der Agarose-Nudeln, die mit einem sterilen Skalpell abgeschnitten wurden, 1h mit Restriktionspuffer auf Eis gestellt und nach Zugabe des Restriktionsenzym in einem Gesamtvolumen von 100  $\mu$ l bei 37°C 1 h inkubiert. Die Reaktion wurde mit 1/10 Volumen 0,5 M EDTA gestoppt und das Gelstück in die Taschen eines PFGE Gels gelegt und mit 1% LM Agarose übergossen.

### **2.2.6. Elektroelution von hochmolekularer DNA**

Das ausgeschnittene Gelstück wurde unter Vermeidung von Luftblasenbildung in mit 2 ml 0,5x TAE gefüllte Dialyseschläuche (Dialysis Tubing- 3/4 in. Diameter, GIBCO) verschlossen, und die DNA in 0,5x TAE bei 1,5 V/cm und Raumtemperatur für 15 Stunden oder bei 6 V/cm bei 4°C für 2 Stunden elektroeluiert. Nach der Elektroelution wurde die Spannung für 1 min umgekehrt, damit die am Schlauch haftende DNA abgelöst wurde. Die DNA wurde dann mit einer abgeschnittenen Pipettenspitze mit großer Öffnung aus dem Schlauch aufgenommen.

### **2.2.7. Konzentrierung der DNA**

Um die Konzentration zu erhöhen, wurde die DNA in 500  $\mu$ l Zentrifugenröhrchen mit einer Ausschlussgröße von 30000 Da (Vivaspin, Vivascience) gegeben und bei 5000 rpm zentrifugiert bis das gewünschte Volumen erreicht war. Um das Pipettieren der DNA zu vermeiden wurde der obere Teil des Zentrifugenröhrchens mit einem abgeflamten

Skalpell abgeschnitten und kopfüber in ein Eppendorfgefäß gestellt. Nach Zentrifugation bei 1000 rpm für 30s, lag die DNA im Eppendorfgefäß vor.

### **2.3. Klonierung hochmolekularer DNA in Fosmide**

Die aufgearbeitete DNA aus Bodenproben wurde in einen Fosmid-Vektor (Abbildung 37; EpiFOS-5, Epicentre) mit der Größe von 7,5 kb und einer Aufnahmekapazität von 32,5 – 43,5 kb, kloniert. Das Grundgerüst dieses Fosmids basiert auf dem F-Faktor (O'Connor *et al.*, 1989) von *E. coli* und ähnelt stark *Bacterial Artificial Chromosome* (BAC) Vektoren. Die Eigenschaften, die F-Faktoren als Vektoren so interessant machen, sind vor allem die niedrige Kopienzahl (1-2 pro Zelle) und die u.a. dadurch gegebene Stabilität der Replikation von hochmolekularer Fremd-DNA (Shizuya *et al.*, 1992). Das lacZ Gen ermöglicht eine blau/weiß Selektion von rekombinanten Klonen. Außerdem ist der Vektor mit einer Klonierungskassette, die eine glatte Schnittstelle *Eco72I* enthält und einem Chloramphenicol-Resistenzgen (ChIR) als Selektionsmarker ausgestattet. Die vorhandene cos-Sequenz ist für die Verpackung von Fremd-DNA in vorgefertigte Phagenköpfe notwendig.

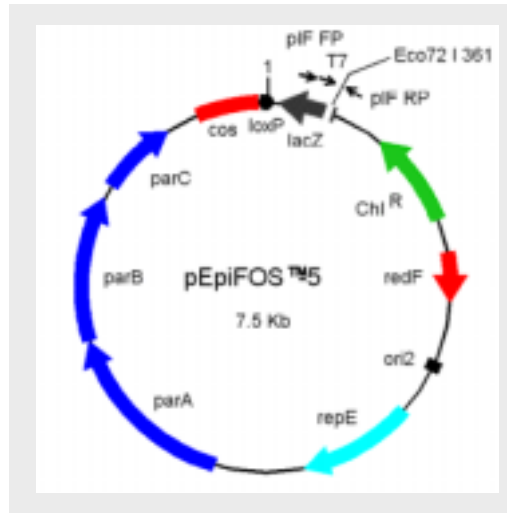


Abbildung 37: Diagramm des Fosmid-Vektors: Die Grundstruktur enthält vier essentielle Bereiche, die für die Plasmidstabilität und die Kopienzahl notwendig sind (Willets und Skurray, 1987) parA, parB und parC koordinieren die genaue Partitionierung von F-Faktor DNA Molekülen auf die Tochterzellen während der Zellteilung. parB und parC machen die Zelle inkompatibel für andere F-Plasmide. Ori2 ist der Replikationsursprung und ist unidirektional. Das repE Gen codiert für ein Protein, das essentiell für die Replikation von ori2 aus ist.

### 2.3.1. Manipulation der DNA Enden

Damit die DNA in den Fosmid-Vektor kloniert werden konnte, wurde sie an ihren Enden geglättet und phosphoryliert werden. Dies wurde in einem Reaktionsansatz mit Hilfe eines Enzym-Gemisches, indem eine T4 DNA-Polymerase und eine T4 Polynukleotid-Kinase vorhanden sind, durchgeführt. Es wurden ca. 3 µg DNA in die Reaktion eingesetzt. Hinzugefügt wurden 6 µl dNTP Mix (2,5 mM), 6 µl ATP (10mM), 6 µl 10x Puffer und 2 µl des Enzym-Gemisches (EpiFOS Library Production Kit, Epicentre). Der Reaktionsansatz von 60 µl Gesamtvolumen wurde bei Raumtemperatur für 1,5 Stunden inkubiert. Die Abtrennung der Proteine erfolgte durch eine Phenol-Chloroform-Isoamylalkohol Extraktion. Die DNA Sedimente der Boden- und der Schwamm-assoziierten-Prokaryonten-DNA wurden in 200 µl bi-deionisiertem Wasser resuspendiert und anschließend einer 2. Größenselektion unterzogen.

### 2.3.2. Ligation

Es wurden 500 ng modifizierte Insert-DNA aus Bodenproben mit 250 ng Fosmid-Vektor eingesetzt. Dies entspricht einem molaren Verhältnis von Vektor zu Insert von 3,5:1. Hinzugefügt wurden außerdem 2 µl 10x Ligationspuffer, 1µl ATP (10mM), 1 µl (2U) Fast-Link DNA Ligase und 13 µl bi-deionisiertes Wasser. Nach 2 stündiger Inkubation bei Zimmertemperatur wurde die Ligase 10 min bei 70°C hitzeinaktiviert. Danach wurde die Ligation bei -70°C eingefroren.

In die Ligation der Schwamm-assoziierten-Prokaryonten wurden 300 ng (SAPs *Aplysina aerophoba*) und 350 ng (*Axinella polypoides*) modifizierte Insert DNA und 500 ng Vektor eingesetzt. Das entspricht einem molaren Vektor zu Insert Verhältnis von 11:1 und 10:1. Nach Zugabe von 2 µl Ligationspuffer, 1 µl ATP (10 mM), 1 µl (2U) Fast-Link DNA Ligase und bi-deionisiertes Wasser wurde der Ligationsansatz von 20 µl Gesamtvolumen bei 12 °C für 13 h inkubiert. Bei 70°C wurde die Ligase für 10 min inaktiviert und der Ligationsansatz bei -70°C gelagert.

### 2.3.3. Verpackung des Ligationsansatzes

Die Verpackung der Ligationsansätze erfolgte *in vitro* in vorgefertigte Verpackungsextrakte (Epicentre). Jeweils 5 µl des Ligationsansatzes wurden zu einem Verpackungsextrakt gegeben, durch Rühren mit einer Pipettenspitze vermischt und bei 30°C 90min inkubiert. Hierbei wurde die DNA in λ-Phagenköpfe verpackt.

Die Reaktion wurde mit 300 µl SM-Phagen-Verdünnungspuffer (50 mM Tris/HCl, pH 7,5; 10 mM MgSO<sub>4</sub>; 100 mM NaCl; 0,01% Gelatine) und 20 µl Chloroform abgestoppt. Durch invertieren wurden die Reaktionsansätze vorsichtig gemischt und 30 s bei 15000 rpm zentrifugiert. Die Phagenüberstände wurden bis zur Infektion bei 4°C gelagert.

### 2.3.4. Infektion und Ausplattierung

"Kompetente" *E. coli* (EPI100, Epicentre) Zellen wurden durch Wachstum in 200 ml LB-Medium mit 10 mM MgSO<sub>4</sub> und 0,2% Maltose hergestellt und die Konzentration wurde entsprechend einer OD von 1,8, abweichend vom Standardprotokoll, eingestellt. Bei der Infektion wurden in einem 2 ml Reaktionsgefäß jeweils 50 µl Phagenüberstand und 200 µl "kompetente" *E. coli* Zellen gemischt und 30 min bei Zimmertemperatur inkubiert. Es wurde 1 ml LB-Flüssigmedium zugegeben und zur Regeneration 1 Stunde bei 37°C inkubiert. Nach jeweils 15 min wurden die Ansätze invertiert. Anschließend wurden die Zellen bei 13000 rpm 1 min sedimentiert und der Überstand verworfen. Das Sediment wurde in 200 µl frischem LB-Flüssigmedium resuspendiert. Die Plattierung erfolgte auf

LB-Chloramphenicol Agarplatten. Um Einzelkolonien zu erhalten, wurde so ausplattiert, dass ca. 200 Kolonien pro Petrischale gewachsen waren. Die Agarplatten wurden bei 37°C für 18 h inkubiert, dann ausgewertet und bei 4°C gelagert.

#### **2.4. Herstellung von Glycerinkonserven der Genbanken**

Zur Konservierung der auf den LB-Chloramphenicol-Agarplatten gewachsenen Einzelkolonien wurden Glycerinkonserven in Mikrotiterplatten angelegt.

In 384er Mikrotiterplatten (Genetics) wurde LB-Chloramphenicol-Medium, das 7,5% Glycerin enthielt, vorgelegt, und die Einzelkolonien wurden mit einem Zahnstocher in die 384 Kavitäten pro Platte angeimpft. Die Mikrotiterplatten wurden bei 37°C 18-20 h stehend inkubiert. Mit Hilfe eines 384er Stempels wurden Kopien hergestellt und nach der Inkubation wurden die Mikrotiterplatten bis zur weiteren Verwendung bei -70°C gelagert.

#### **2.5. Herstellen von Matrizen-DNA für Multiplex-Polymerase-Kettenreaktion**

Das Durchmustern der Genbanken wurde durch das Herstellen von DNA-Pools, die DNAs von jeweils 384 Klonen einer Mikrotiterplatte stammten, vereinfacht. Mit Hilfe eines 384er Stempel wurden Klone der Mikrotiterplatten auf Chloramphenicol-Agarplatten gestempelt. Nach Wachstum wurden jeweils 384 der auf LB-Chloramphenicol-Agarplatten gewachsenen Klone mit Puffer P1 (Qiagen, Plasmid Midi Protocol) abgeschwemmt und vereinigt. Anschließend wurde die Fosmid-DNA isoliert (Plasmid Midi Protocol Qiagen) und in TE resuspendiert. Somit war es möglich mittels Multiplex-PCR-Technik die Genbanken mit nur wenigen PCR Ansätzen – pro Ansatz 384 Klone - nach bestimmten Genfragmenten zu durchsuchen.

#### **2.6. Isolierung von rekombinanten Fosmiden**

In der Regel wurde Fosmid-DNA nach der herkömmlichen *E. coli* Plasmid Extraktionsmethode über Schleudersäulchen (GibcoBRL) durchgeführt.

Aufgrund der geringen Kopienzahl der Fosmid-Vektoren in *E. coli* Zellen stellt die chromosomale *E. coli* DNA, von der ein Teil, bei herkömmlichen Plasmid Extraktionsmethoden, mitextrahiert wird, bei Sequenzierungen ein häufiges Problem dar. Zur Erstellung von "shot-gun" Genbanken und zur Sequenzierung von DNA-Fragmenten direkt von Fosmid-Vektoren wurde deshalb eine Methode angewendet – "Doppel-Acetat Präparation"-, die den Hintergrund an chromosomaler DNA absenkt und somit das Risiko der Sequenzierung von chromosomaler *E. coli* DNA reduziert.

Die Fosmid-Präparation erfolgte in Anlehnung an das Protokoll der Plasmid Midi-Extraktion (Qiagen). Die Fosmid-Einzelklone wurden in zweimal 250 ml LB-Chloramphenicol-Flüssigmedium in 1 l Erlenmeyerkolben angeimpft und bei 37°C über Nacht bei 200 rpm inkubiert. Die Zellen wurden in drei 250 ml Sorvall GSA-Zentrifugenbecher sedimentiert. Die Sedimente wurden in 20 ml P1 Puffer (Qiagen-Kit) ohne RNase resuspendiert und in ein 250 ml GSA-Polykarbonatbecher überführt. Es wurden 20 ml Puffer P2 zugegeben und 5 min bei Zimmertemperatur stengelassen. Dann wurden 20 ml P3 Puffer zugegeben, gemischt und 30 min auf Eis gestellt. Der Überstand wurde zweimal abzentrifugiert und anschließend mit 42 ml Isopropanol bei Zimmertemperatur präzipitiert. Die DNA wurde bei 10000 rpm 30 min bei 4°C (Sorvall GSA) sedimentiert und der Überstand verworfen. Nach kurzem Antrocknen wurde das Sediment in 2 ml TE (10mM Tris-HCl, 50 mM EDTA pH 8) aufgenommen und auf zwei 2 ml Reaktionsgefäße verteilt. Jeweils 0,5 ml 7,5 M KOAc-Lösung wurden zugegeben, durch invertieren gemischt und 30 min bei -70°C eingefroren.

Durch Einführung der Kaliumacetat-Präzipitation vor der Säule wird diese nicht mit RNA vollgeladen, was zu erhöhten Ausbeuten führt.

Beim Abzentrifugieren der DNA-Lösung nach Zugabe von Kaliumacetat und Einfrieren bei -70°C fallen RNA und lineare DNA aus, wodurch der Anteil an chromosomaler DNA von *E. coli* verringert wird. Im Überstand befindet sich relativ reine Plasmid-DNA.

Nach dem Auftauen wurde der Überstand 10 min mit 6000 rpm bei 4°C abzentrifugiert und der Überstand in ein 15 ml Schraubkappenröhrchen überführt. Die DNA-Lösung wurde mit QBT Puffer (Qiagen-Kit) auf 10 ml aufgefüllt. Danach erfolgen die Schritte wie im Qiagen Plasmid Midi-Kit-100 angegeben.

Nach Präzipitation der DNA wurde das Sediment in 100 µl TE mit RNase (10mM Tris-HCl, 1mM EDTA, 10µg/ml RNase A) aufgenommen. Die Fosmid-Lösung wurde dann bei 4°C gelagert.

### **2.7. Restriktionsanalyse von isolierten Fosmid-DNAs**

Um die klonierte DNA von der Fosmid-Vektor DNA zu trennen und ihre Größe bestimmen zu können, wurde eine Restriktion mit *Not* I durchgeführt. *Not* I besitzt zwei das Insert-flankierende Schnittstellen im Fosmid-Vektor. Durch einen Schnitt mit *Not* I wird das Insert und zusätzlich 641bp des Vektors ausgeschnitten, so dass nach elektrophoretischer Trennung eine Vektorbande von ca. 6900 bp und eine oder mehrere Banden des Insert vorhanden sind. Pro Restriktionsansatz wurden 1 µg Plasmid DNA eingesetzt. Hinzugefügt wurden 2,5 µl 10x O<sup>+</sup> Puffer (MBI/Fermentas), 0,5 µl *Not* I (5U; MBI/Fermentas) und bi-deionisiertes Wasser, so dass das Gesamtvolumen eines Restrik-

tionsansatzes 25 µl betrug. Die Restriktion wurde bei 37°C für 1 h inkubiert. Das Restriktionsmuster wurde auf einem 1%-igen TAE-Agarosegel in einer Pulsfeldgelelektrophorese mit den Bedingungen 200V, 0,5 – 2s, 10 h bei 10°C analysiert.

## 2.8. Polymerase-Kettenreaktion zur Amplifizierung von 16S rRNA-Genen

16S rDNA-Fragmente wurden mit verschiedenen Oligonukleotid-Paaren amplifiziert. Zur Identifizierung von archaealen 16S rDNA-Fragmenten wurden zwei spezifische Oligonukleotide Arch20F und Arch958R (Arch20F: 5'- TTCCGGTTGATCCYGCCGG - 3'; Arch958R: 5'-YCCGGCGTTGAMTCCAATT- 3'; DeLong, 1992) verwendet. Um 16S rDNA-Fragmente von Acidobacteria zu identifizieren wurde ein spezifisches Oligonukleotid Ac31F (Ac31F:5'- GATCCTGGCTCAGAATC- 3'; Barns *et al.*, 1999) und ein universelles Oligonukleotid Univ533R (Univ533R: 5'- TTACCGCGGCKGCTGGCAC- 3'; Reysenbach und Pace, 1995; modifiziert von Torsten Ochsenreiter) eingesetzt.

Im Allgemeinen wurde folgender PCR Ansatz verwendet:

5 µl 10-fach *Opti*-Puffer  
2 mM MgCl<sub>2</sub>  
0,2 mM Desoxynukleosidtriphosphate  
20 pmol jedes Oligonukleotids  
1 U *DAP<sub>GoldStar</sub>*-Polymerase  
1 µl DNA-Matrize  
ad 50 µl bi-deionisiertes Wasser

Die Bedingungen für die Amplifizierung in der PCR-Maschine T-Gradient (Biometra) wurde in der Regel unter folgenden Bedingungen durchgeführt:

1. Zyklus:           96 °C für 7 min  
                          54 °C für 30 sec  
                          72 °C für 1 min  
2.-36. Zyklus:     96 °C für 30 sec  
                          54 °C für 30 sec  
                          72 °C für 1 min  
37. Zyklus:         96 °C für 30 sec  
                          54 °C für 30 sec  
                          72 °C für 10 min

Bei der Amplifizierung von acidobakteriellen 16S rRNA-Genfragmenten wurde außerdem 5x GC-Solution (Roche; Fast Start DNA Polymerase Kit Komponente) zugesetzt, da dadurch ausschließlich spezifische PCR-Produkte erhalten wurden. Die Anlagerungstemperatur betrug im Unterschied zu hier angegebenen Bedingungen 50°C. Die PCR-Produkte wurden auf einem 1%-igen oder 2%-igen TAE-Agarosegel analysiert.



### **2.8.1. Isolierung von PCR-Produkten aus Agarosegelen**

Die PCR Produkte wurden mit einer sterilen Klinge aus dem Ethidiumbromid gefärbten Gel ausgeschnitten und in Schleudersäulchen (Ultrafree-DA, Millipore) gegeben. Das Gelstück wurde bei 5000 g für 10 min zentrifugiert, wodurch die Gelbestandteile in der Säule zurückgehalten wurde und sich die DNA im Durchfluß befand. Um die Reinheit der DNA zu erhöhen wurde eine Phenol/Chloroform/Isoamylalkohol (PCI/25:24:1) Extraktion angeschlossen. 1/10 Volumen Natriumacetat und 1 Volumen PCI wurden zugegeben, durch Vortexen gemischt und 10 min bei 15000 rpm zentrifugiert. Die wässrige Phase wurde abgenommen und die DNA mit 0,7 Volumen Isopropanol präzipitiert und anschließend das DNA Sediment mit 70%-igem EtOH gewaschen. Das DNA-Sediment wurde in bi-deionisiertem Wasser resuspendiert und in die Ligationsreaktion eingesetzt.

### **2.8.2. Klonierung von PCR-Produkten**

Die Klonierung von PCR Produkten wurde nach den Angaben des Herstellers mit pCR-Blunt II-TOPO Vektor (Invitrogen) durchgeführt. Dieser Vektor besitzt den ColE1 ori, das lacZ $\alpha$ -Gen und das Kanamycin-Resistenzgen. Außerdem besitzt der Vektor ein "Selbstmordgen" ("suicide gene"), das durch Insertionen unterbrochen wird, wodurch die Expression des toxischen Proteins verhindert wird und nur die Zellen mit Insert im Vektor wachsen können.

## **2.9. DNA-Sequenzbestimmung**

Die Bestimmung von DNA-Sequenzen erfolgte mit der Kettenabbruchmethode (Sanger *et al.*, 1977). Das verwendete Gerät zur DNA-Sequenzierung war ein Li-Cor DNA Sequencer 4000. Die Markierung der DNA-Fragmente erfolgte mit Hilfe der Sequenzierungs-Oligonukleotide, die am 5´-Ende mit dem Fluoreszenz-Farbstoff IRD800 (Lmax bei 795 nm) versehen waren. Das Oligonukleotid M13 Universal hat eine Bindestelle zu einer flankierenden Region im Vektor pCR II TOPO-Blunt und auch im Fosmid-Vektor. Dadurch war es möglich die Sequenzen der klonierten PCR Produkte und auch ein Ende der in das Fosmid klonierten DNA-Fragmente zu bestimmen.

Die Sequenzreaktion wurde nach dem Protokoll des Herstellers (EXCEL II Long-Read DNA Sequencing Kit, Epicentre) durchgeführt.

Es wurden unterschiedliche Reaktionsmischungen und auch Sequenzreaktionsbedingungen für Plasmide und Fosmide verwendet:

Reaktionsmischung für Plasmide:

3,6 µl 3,5-fach Sequenzierungspuffer  
2 pmol markiertes Oligonukleotid M-20-universal  
3,4 µl DNA  
0,5 µl *SequiTherm EXCEL II DNA Polymerase* (5 U/µl)

Reaktionsmischung für Fosmide:

3,6 µl 3,5-fach Sequenzierungspuffer  
20 pmol markiertes Oligonukleotid M-20-universal  
3,4 µl DNA  
0,5 µl *SequiTherm EXCEL II DNA Polymerase* (5 U/µl)

Zu je 1 µl vorgelegter Terminationsmischung A, C, G und T, die neben den vier Desoxynukleotiden eins der vier Didesoxynukleotide enthielten, wurden je 2 µl der Reaktionsmischung zugegeben. Nach dem Überschichten der Ansätze mit etwas Mineralöl wurden die Reaktionen in der PCR-Maschine T-Gradient (Biometra) durchgeführt. Auch hier wurden unterschiedliche Bedingungen für die Plasmide und für die Fosmide verwendet.

Sequenzreaktions-Bedingungen für Plasmide:

1. Zyklus:           95 °C für 5 min  
                  54 °C für 30 sec  
                  70 °C für 1 min 30 sec  
2.-36. Zyklus:       95 °C für 30 sec  
                  54 °C für 30 sec  
                  70 °C für 1 min 30 sec

Sequenzreaktions-Bedingungen für Fosmide:

1. Zyklus:           95 °C für 5 min  
                  60 °C für 30 sec  
                  70 °C für 1 min 30 sec  
2.-50. Zyklus:       95 °C für 30 sec  
                  60 °C für 30 sec  
                  70 °C für 1 min 30 sec

Nach Beendigung der Sequenzierungsreaktion wurden je 1,5 µl Stoppuffer zugegeben und die Proben bis zur Elektrophorese bei -20°C gelagert. Nach Denaturierung der Proben bei 90 °C für 5 min erfolgte die elektrophoretische Auftrennung der DNA-Fragmente in einem 6%-igen denaturierenden Polyacrylamidgel nach den Angaben des Herstellers (Licor).

## 2.10. Herstellung DIG-markierter DNA-Sonden

Die Markierung von DNA-Fragmenten mit Digoxigenin (DIG) erfolgte nach dem Protokoll des Herstellers (DNA Labeling Kit, Böhlinger Mannheim). Die DNA wurde dabei mit sondenmarkierten Hexanukleotiden als Primer inkubiert.

Der Nachweis der an komplementäre Sequenzen hybridisierten Sonde findet dann über einen Immunoassay mit einem Anti-DIG-Antikörper statt, der an alkalische Phosphatase gekoppelt ist. Bei Zugabe des Substrats (CSPD) wird die Bindung in Form einer Lichtreaktion nachgewiesen.

Die Markierung wurde unter folgenden Bedingungen durchgeführt:

10 µl DNA (25-50 ng)  
5 µl H<sub>2</sub>O bi-deionisiert  
2 µl Hexanukleotide  
2 µl dNTP-DIG-Mix  
1 µl Klenow-Enzym (2U)

Das Reaktionsgemisch wurde 10 bis 12 h bei 37°C inkubiert, anschließend wurde die Reaktion mit 1 µl 0,5 M EDTA (pH 8) abgestoppt und die Sonde mit 2,5 µl 4 M LiCl und 75 µl 100% EtOH über 30 min bei -70°C präzipitiert. Nach 30 minütiger Zentrifugation bei 15000 rpm und 4°C wurde die Sonde mit 70% Ethanol gewaschen und schließlich in 100 µl TE-Puffer aufgenommen.

## 2.11. Herstellung von Koloniefiltern

Um spezifische Einzelklone auf den Mikrotiterplatten zu identifizieren, wurden die Klone auf eine Membran (Hybond-N, Amersham) übertragen. Die gewachsenen Kolonien wurden lysiert und die spezifischen Einzelklone mit einer DIG-markierten Sonde identifiziert.

Die Membran wurde auf in LB-Medium getränktes Whatman-Papier unter Vermeidung von Luftblasenbildung gelegt. Nach Durchfeuchtung wurde die Membran auf LB-Chloramphenicol-Medium (12,5 µg/ml) in eine Petrischale gelegt. Mit dem 384er Stempel wurden die Kulturen aus den Mikrotiterplatten auf die Membran übertragen und bei 37°C 18 h inkubiert. Auf der Membran waren definierte Kolonien gewachsen, die im nächsten Schritt auf der Membran lysiert wurden.

Das Protokoll wurde in Anlehnung an eine Technik, die vom Ressourcenzentrum in Berlin angewendet wurde, entwickelt (Schleper, persönliche Mitteilung).

Der bewachsene Filter wurde vorsichtig aus der Petrischale entfernt und auf Whatman Papier gelegt, das mit Denaturierungslösung (0,5 M NaOH, 1,5 M NaCl) getränkt war und 4 min bei Raumtemperatur inkubiert. Zur Denaturierung wurde der Filter auf

mit Denaturierungslösung getränktes Whatman-Papier, das auf einer Glasplatte lag, überführt und die Glasplatte über ein Wasserbad von 95°C gelegt und 4 min inkubiert. Anschließend wurde der Filter auf Whatman-Papier, das mit Neutralisierungslösung (1 M Tris-HCl, 1 M Tris-Base; 1,5M NaCl, pH 7,4) getränkt war, überführt und erneut für 4 min bei Zimmertemperatur inkubiert. Im nächsten Schritt wurde der Filter in Proteinase K-Puffer-Lösung (50 mM Tris-HCl pH 7,8, 1% Laurylsarkosin, 20 mg/ml Proteinase K) gegeben und 45 min bei 37°C inkubiert. Der Filter wurde in 50 mM Natriumphosphat, pH 7,2 10 min bei Zimmertemperatur gewaschen und anschließend auf Blotting Papier über Nacht getrocknet. Am nächsten Tag wurde die DNA durch UV-Cross-Linking (Stratagene) auf der Membran fixiert.

### **2.12. Filterhybridisierung**

#### **2.12.1. Hybridisierung**

Die Membran wurde 5 min im Wasserbad von Salzresten befreit und in einer quadratischen Petrischale in 100 ml Hybridisierungslösung gelegt. Die Prähybridisierung erfolgte 5 h bei 42°C. Die Sonde wurde 5 Minuten bei 100°C denaturiert und nach dem Abkühlen auf Eis mit den 100 ml Hybridisierungslösung aus der Prähybridisierung in einem gesonderten Gefäß gemischt und auf den Filter in der Petrischale gegeben. Die Hybridisierung erfolgte bei 42°C über Nacht.

#### **2.12.2. Waschen und Exposition:**

Im Anschluss an die Hybridisierung wurde die Membran zweimal für je 15 min mit 2x SSC / 0,1% SDS bei RT und zweimal mit 0,2x SSC / 0,1% SDS bei 68°C gewaschen. Durch das Erhöhen der Temperatur von 60°C auf 68°C – abweichend vom Standardprotokoll - wurde die Stringenz für die Bindung der Sonde erhöht, so dass ein größerer Anteil unspezifisch gebundener Sondenmoleküle abgewaschen wurde. Anschließend wurde die Membran bei RT 5 Minuten in Waschpuffer und 30 min in Puffer 2 gewaschen. Die Bindung des Konjugats aus Schafs-Anti-DIG-Antikörpern und alkalischer Phosphatase (1:10.000 in Puffer 2 verdünnt) an die Sonde erfolgte 30 min bei RT. Nach der Antigen-Antikörper-Reaktion wurden die ungebundenen Antikörper-Konjugate zweimal für 15 min mit Waschpuffer abgewaschen. Nach dem letzten Waschschrift für 5 Minuten in Puffer 3 wurde das Substrat der alkalischen Phosphatase, CSPD (1:100 in Puffer 3 verdünnt), zugegeben und 5 min bei Zimmertemperatur inkubiert. Die Membran wurde dann kurz zwischen Whatmanpapier getrocknet. Anschließend wurde ein Röntgenfilm auf die Memb-

ran gelegt und für mindestens 10 Minuten bei 37°C im Dunkeln inkubiert. Die Lichtreaktion verursacht an Stellen, an der die Sonde spezifisch an DNA gebunden hat, auf dem Röntgenfilm eine Schwärzung. Der Film wurde je nach Signalstärke zwischen 10 Minuten und 2 Stunden exponiert und im Anschluss in der Dunkelkammer entwickelt.

### **2.13. Subklonierung und Sequenzierung**

Die DNA der Einzelklone mit hochmolekularem Insert archaealer und acidobakterieller Herkunft wurden nach mechanischer Fragmentierung auf 2-3 kb in den Vektor pGEM-T-easy (Promega) subkloniert. Die Sequenzen der Enden der klonierten DNA Fragmente wurde mit Vektor-Oligonukleotiden in einem ABI3700 Kapillarsequenzierer bestimmt. Die verbleibenden Sequenzlücken wurden durch Oligonukleotid-Wanderung geschlossen. Der archaeale Klon 2914 wurde von Arne Zibat subkloniert und sequenziert (Arne Zibat ist ein Mitarbeiter unseres Kooperationspartners Hans-Peter Klenk, Epidaurus). Die Sequenzen des Klons 2914 wurden in EMBL unter der Nummer AJ496176 abgelegt. Die Einzelklone mit acidobakteriellem Insert 32D1, 39P18, 41b15 wurden von Torsten Ochsenreiter und die Klone 23K22, 38C7 und 3/46H5 von Christa Lanz subkloniert und sequenziert, sowie von Günther Raddatz assembliert (Christa Lanz und Günther Raddatz arbeiten in der Arbeitsgruppe unseres Kooperationspartners Stephan Schuster am MPI in Tübingen).

### **2.14. Phylogenetische Analyse**

Das ARB-software package (Ludwig *et al.*, 1998) wurde zur Alignierung und zur phylogenetischen Analyse von 16S und 23S rRNA-Genen verwendet. Die Topologien des 23S rDNA Baums wurden durch maximum parsimony (Parsimony Interactive) und die Distanzmatrix Methode (Felsenstein Korrektur) mit unterschiedlichen Alignment-Filtern (Gap-Filter, Positional Variability Filter, Maximum Frequency Filter) evaluiert. Die Topologien des entsprechenden 16S rDNA Baumes wurde nach denselben Methoden und zusätzlich mit der Maximum Likelihood Methode (fastDNAmI) mit den oben angegebenen Filtern evaluiert. Die phylogenetischen Analysen der rRNA-Gene wurden von Torsten Ochsenreiter durchgeführt.

Die Phylogenetische Analyse des vorgeschlagenen FixA Gens (ORF 12) wurde mit Protein Parsimony (PROTPARS) und Neighbor-Joining (NEIGHBOR) Programmen von PHYLIP 3.6 und PAUP 3.1.1. (Swofford, 2000) durchgeführt. Dieselbe Gesamttopologie wurde mit beiden Methoden basierend auf 211 konservierten Positionen aus einem Alignment von 33 FixA/ETF $\beta$  (flavoprotein containing electron transport chain) Homologen identifiziert. Die phylogenetische Analyse von FixA wurden von unserem Kooperationspartner Hans-Peter Klenk durchgeführt.

Paarweise Alignments wurden mit GAP (Husar) und Best Fit durchgeführt.

### **2.15. Sequenzannotation**

Die Identifizierung der ORFs und die automatische Genannotation wurde mit Hilfe des Magpie program packages (Gaasterl und Sensen, 1996) für Klon 2914 und mit MacPlasmap für den Klon 32D1 durchgeführt. Das Wisconsin Package (Heidelberg Unix Sequence Analysis Server, HUSAR) wurde zur Identifizierung von Proteinfamilien (PFAMs) und Transmembrandomänen, sowie für Sekundärstrukturvorhersagen und Peptidmotividentifizierungen verwendet. Die tRNA Gene wurden mit Hilfe des tRNAscans (<http://www.genetics.wustl.edu/eddy/tRNAscan-SE>) identifiziert. Multiple Alignments wurden mit PILUP, CLUSTAL und MULTALIN (Corpet, 1988) durchgeführt.

### **2.16. Schwammmaterial**

Die Mittelmeerschwämme *Aplysina aerophoba* und *Axinella polypoides* wurden von Sporttauchern aus einer Tiefe von 5 – 15 m bei Banyuls sur Mer (Frankreich) im Frühling 2002 gesammelt. Die Individuen wurden in Plastiksäcken unter Luftausschluß an die Oberfläche gebracht und innerhalb von 1-5 Tagen weiterverarbeitet.

### **2.17. Trennung der prokaryotischen Zellen von Schwammzellen**

Im ersten Schritt der DNA Aufbereitung wurden die prokaryotischen Zellen von Schwammzellen durch differentielle Zentrifugation (Lars Fieseler) getrennt. Dabei wurden ca. 25 g Schwamm (Frischgewicht) in ca. 5 mm<sup>3</sup> große Schwammstücke mit einem Skalpell in Artificial Seawater (ASW) zugeschnitten. Die Blöcke wurden 2x in doppeltem Volumen ASW für 1h aufgeschlämmt und anschließend mit einem Mörser unter ASW zermalmt. Anschließend folgte eine zweimalige differentielle Zentrifugation des Überstandes bei 1000 g für 15 min und ein Sedimentierungsschritt bei 5000 g für 15 min. Das Zellsediment wurde in ASW aufgenommen und die Zellkonzentration mikroskopisch in einer Thoma Kammer auf  $8,8 \times 10^9$  Zellen/ml bestimmt. Die Zellkonzentration wurde auf  $2 \times 10^9$  Zellen/ml eingestellt.

## F. Literatur

- Barns, S.M., Takala, S.L., and Kuske, C.R. (1999) Wide distribution and diversity of members of the bacterial kingdom Acidobacterium in the environment. *Appl Environ Microbiol* **65**: 1731-1737.
- Bassler, B.L. (2002) Small talk. Cell-to-cell communication in bacteria. *Cell* **109**: 421-424.
- Bauer, E., Kaspar, T., Fischer, H.M., and Hennecke, H. (1998) Expression of the fixR-nifA operon in Bradyrhizobium japonicum depends on a new response regulator, RegR. *J Bacteriol* **180**: 3853-3863.
- Beja, O., Spudich, E.N., Spudich, J.L., Leclerc, M., and DeLong, E.F. (2001) Proteorhodopsin phototrophy in the ocean. *Nature* **411**: 786-789.
- Beja, O., Suzuki, M.T., Heidelberg, J.F., Nelson, W.C., Preston, C.M., Hamada, T. et al. (2002a) Unsuspected diversity among marine aerobic anoxygenic phototrophs. *Nature* **415**: 630-633.
- Beja, O., Koonin, E.V., Aravind, L., Taylor, L.T., Seitz, H., Stein, J.L. et al. (2002b) Comparative genomic analysis of archaeal genotypic variants in a single population and in two different oceanic provinces. *Appl Environ Microbiol* **68**: 335-345.
- Beja, O., Suzuki, M.T., Koonin, E.V., Aravind, L., Hadd, A., Nguyen, L.P. et al. (2000a) Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ Microbiol* **2**: 516-529.
- Beja, O., Aravind, L., Koonin, E.V., Suzuki, M.T., Hadd, A., Nguyen, L.P. et al. (2000b) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**: 1902-1906.
- Bintrim, S.B., Donohue, T.J., Handelsman, J., Roberts, G.P., and Goodman, R.M. (1997) Molecular phylogeny of Archaea from soil. *Proc Natl Acad Sci U S A* **94**: 277-282.
- Blanco, L., Bernad, A., Blasco, M.A., and Salas, M. (1991) A general structure for DNA-dependent DNA polymerases. *Gene* **100**: 27-38.
- Bowman, W.C., Du, S., Bauer, C.E., and Kranz, R.G. (1999) In vitro activation and repression of photosynthesis gene transcription in Rhodospirillum rubrum. *Mol Microbiol* **33**: 429-437.
- Brady, S.F., Chao, C.J., Handelsman, J., and Clardy, J. (2001) Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA. *Org Lett* **3**: 1981-1984.
- Braithwaite, D.K., and Ito, J. (1993) Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res* **21**: 787-802.
- Briese, B.H., Jendrossek, D., and Schlegel, H.G. (1994) Degradation of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) by aerobic sewage sludge. *FEMS Microbiol Lett* **117**: 107-111.
- Chandler, D.P., Schreckhise, R.W., Smith, J.L., and Bolton Jr, H. (1997) Electroelution to Remove Humic Compounds from Soil DNA and RNA Extracts. *Journal of Microbiological Methods* **28**: 11-19.
- Charles, T.C., Cai, G.Q., and Aneja, P. (1997) Megaplasmid and chromosomal loci for the PHB degradation pathway in Rhizobium (Sinorhizobium) meliloti. *Genetics* **146**: 1211-1220.
- Chen, J., Banks, D., Jarret, R.L., and Jones, J.B. (2000) Evidence for conserved tRNA genes in the 16S-23S rDNA spacer sequence and two rrn operons of Xylella fastidiosa. *Can J Microbiol* **46**: 1171-1175.

- Choi, S.a.W., R. A. (2000) The Construction of Bacterial Artificial Chromosome (BAC) Libraries. In *Plant Molecular Biology Manual 2nd Edition*. Schilperoort, S.G.a.R. (ed): Kluwer Academic Publishers, pp. 1-32.
- Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**: 10881-10890.
- DeLong, E.F. (1992) Archaea in coastal marine environments. *Proc Natl Acad Sci U S A* **89**: 5685-5689.
- DeLong, E.F., Wu, K.Y., Prezelin, B.B., and Jovine, R.V. (1994) High abundance of Archaea in Antarctic marine picoplankton. *Nature* **371**: 695-697.
- Desaint, S., Hartmann, A., Parekh, N.R., and Fournier, J. (2000) Genetic diversity of carbofuran-degrading soil bacteria. *FEMS Microbiol Ecol* **34**: 173-180.
- Dunbar, J., Takala, S., Barns, S.M., Davis, J.A., and Kuske, C.R. (1999) Levels of bacterial community diversity in four arid soils compared by cultivation and 16S rRNA gene cloning. *Appl Environ Microbiol* **65**: 1662-1669.
- Emmerich, R., Hennecke, H., and Fischer, H.M. (2000a) Evidence for a functional similarity between the two-component regulatory systems RegSR, ActSR, and RegBA (PrrBA) in alpha-Proteobacteria. *Arch Microbiol* **174**: 307-313.
- Emmerich, R., Strehler, P., Hennecke, H., and Fischer, H.M. (2000b) An imperfect inverted repeat is critical for DNA binding of the response regulator RegR of Bradyrhizobium japonicum. *Nucleic Acids Res* **28**: 4166-4171.
- Emmerich, R., Panglungtshang, K., Strehler, P., Hennecke, H., and Fischer, H.M. (1999) Phosphorylation, dephosphorylation and DNA-binding of the Bradyrhizobium japonicum RegSR two-component regulatory proteins. *Eur J Biochem* **263**: 455-463.
- Fernandez-Castillo, R., Rodriguez-Valera, F., Gonzalez-Ramos, J., and Ruiz-Berraquero, F. (1986) Accumulation of poly( $\beta$ -hydroxybutyrate) by Halobacteria. *Appl Environ Microbiol* **51**: 214-216.
- Fjellbirkeland, A., Torsvik, V., and Ovreas, L. (2001) Methanotrophic diversity in an agricultural soil as evaluated by denaturing gradient gel electrophoresis profiles of pmoA, mxaF and 16S rDNA sequences. *Antonie Van Leeuwenhoek* **79**: 209-217.
- Fuhrman, J.A., and Davis, A.A. (1997) Widespread archaea and novel bacteria from the deep sea as shown by 16S rRNA gene sequences. *Mar Ecol Prog Ser* **150**: 275-285.
- Fuhrman, J.A., McCallum, K., and Davis, A.A. (1992) Novel major archaeobacterial group from marine plankton. *Nature* **356**: 148-149.
- Gaasterland, T., and Sensen, C.W. (1996) Fully automated genome analysis that reflects user needs and preferences. A detailed introduction to the MAGPIE system architecture. *Biochimie* **78**: 302-310.
- Gao, D., Maehara, A., Yamane, T., and Ueda, S. (2001) Identification of the intracellular polyhydroxyalkanoate depolymerase gene of Paracoccus denitrificans and some properties of the gene product. *FEMS Microbiol Lett* **196**: 159-164.
- Graham, D.E., Overbeek, R., Olsen, G.J., and Woese, C.R. (2000) An archaeal genomic signature. *Proc Natl Acad Sci U S A* **97**: 3304-3308
- Grebe, T.W., and Stock, J.B. (1999) The histidine protein kinase superfamily. *Adv Microb Physiol* **41**: 139-227.
- Gubler, M., and Hennecke, H. (1988) Regulation of the fixA gene and fixBC operon in Bradyrhizobium japonicum. *J Bacteriol* **170**: 1205-1214.
- Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., and Goodman, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* **5**: R245-249.



- Hentschel, U., Schmid, M., Wagner, M., Fieseler, L., Gernert, C., and Hacker, J. (2001) Isolation and phylogenetic analysis of bacteria with antimicrobial activities from the Mediterranean sponges *Aplysina aerophoba* and *Aplysina cavernicola*. *FEMS Microbiol Ecol* **35**: 305-312.
- Hermansson, A., and Lindgren, P.E. (2001) Quantification of ammonia-oxidizing bacteria in arable soil by real-time PCR. *Appl Environ Microbiol* **67**: 972-976.
- Hershberger, K.L., Barns, S.M., Reysenbach, A.L., Dawson, S.C., and Pace, N.R. (1996) Wide diversity of Crenarchaeota. *Nature* **384**: 420.
- Heywood, V.H. (1995) *Global Biodiversity Assessment*. Cambridge: Cambridge University Press.
- Hugenholtz, P., Goebel, B.M., and Pace, N.R. (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* **180**: 4765-4774.
- Inagaki, K., Nakahira, K., Mukai, K., Tamura, T., and Tanaka, H. (1998) Gene cloning and characterization of an acidic xylanase from *Acidobacterium capsulatum*. *Biosci Biotechnol Biochem* **62**: 1061-1067.
- Janssen, P.H., Yates, P.S., Grinton, B.E., Taylor, P.M., and Sait, M. (2002) Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Appl Environ Microbiol* **68**: 2391-2396.
- Jurgens, G., and Saano, A. (1999) Diversity of soil Archaea in boreal forest before, and after clear-cutting and prescribed burning. *FEMS Microbiology Ecology* **29**: 205-213.
- Kaeberlein, T., Lewis, K., and Epstein, S.S. (2002) Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296**: 1127-1129.
- Kirk, R.G., and Ginzburg, M. (1972) Ultrastructure of two species of halobacterium. *J Ultrastruct Res* **41**: 80-94.
- Krsek, M., and Wellington, E.M. (1999) Comparison of different methods for the isolation and purification of total community DNA from soil. *J Microbiol Methods* **39**: 1-16.
- Kuske, C.R., Barns, S.M., and Busch, J.D. (1997) Diverse uncultivated bacterial groups from soils of the arid southwestern United States that are present in many geographic regions. *Appl Environ Microbiol* **63**: 3614-3621.
- Lane, D.J. (1991) 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*. Stackebrandt, E., and Goodfellow, M. (eds). New York, N.Y: John Wiley and Sons, p. 115-175.
- Lee, N., Nielsen, P.H., Andreasen, K.H., Juretschko, S., Nielsen, J.L., Schleifer, K.H., and Wagner, M. (1999) Combination of fluorescent in situ hybridization and microautoradiography-a new tool for structure-function analyses in microbial ecology. *Appl Environ Microbiol* **65**: 1289-1297.
- Lottspeich, F., and Zorbas, H. (1998) *Bioanalytik*. Heidelberg; Berlin: Spektrum, Akademischer Verlag.
- Ludwig, W., and Klenk, H.P. (2001) Overview: A Phylogenetic Backbone and Taxonomic Framework for Prokaryotic Systematics. In *Bergey's Manual of Systematic Bacteriology*. Boone, D.R., and Castenholz, R.W. (eds): Springer.
- Liebergessell, M., Sonomoto, K., Madkour, M., Mayer, F., and Steinbuchel, A. (1994) Purification and characterization of the poly(hydroxyalkanoic acid) synthase from *Chromatium vinosum* and localization of the enzyme at the surface of poly(hydroxyalkanoic acid) granules. *Eur J Biochem* **226**: 71-80.

- Ludwig, W., Strunk, O., Klugbauer, S., Klugbauer, N., Weizenegger, M., Neumaier, J. et al. (1998) Bacterial phylogeny based on comparative sequence analysis. *Electrophoresis* **19**: 554-568.
- Lütke-Eversloh, T., and Steinbüchel, A. (2002) The first cloned polyhydroxyalkanoate gene from an uncultured Archaeon is homologous to phaC genes of type-III PHA synthases. In: Münster, Deutschland: Westfälische Wilhelms Universität Münster, Institut für Mikrobiologie, p. 15.
- MacGregor, B.J., Moser, D.P., Alm, E.W., Nealson, K.H., and Stahl, D.A. (1997) Crenarchaeota in Lake Michigan sediment. *Appl Environ Microbiol* **63**: 1178-1181.
- MacNeil, I.A., Tiong, C.L., Minor, C., August, P.R., Grossman, T.H., Loiacono, K.A. et al. (2001) Expression and isolation of antimicrobial small molecules from soil DNA libraries. *J Mol Microbiol Biotechnol* **3**: 301-308.
- Massana, R., Murray, A.E., Preston, C.M., and DeLong, E.F. (1997) Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Appl Environ Microbiol* **63**: 50-56.
- Masuda, S., Matsumoto, Y., Nagashima, K.V., Shimada, K., Inoue, K., Bauer, C.E., and Matsuura, K. (1999) Structural and functional analyses of photosynthetic regulatory genes regA and regB from *Rhodovulum sulfidophilum*, *Roseobacter denitrificans*, and *Rhodobacter capsulatus*. *J Bacteriol* **181**: 4205-4215.
- McCaig, A.E., Glover, L.A., and Prosser, J.I. (1999) Molecular analysis of bacterial community structure and diversity in unimproved and improved upland grass pastures. *Appl Environ Microbiol* **65**: 1721-1730.
- McInerney, J.O., Wilkinson, M., Patching, J.W., Embley, T.M., and Powell, R. (1995) Recovery and phylogenetic analysis of novel archaeal rRNA sequences from a deep-sea deposit feeder. *Appl Environ Microbiol* **61**: 1646-1648.
- Miller, D.N., Bryant, J.E., Madsen, E.L., and Ghiorse, W.C. (1999) Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Appl Environ Microbiol* **65**: 4715-4724.
- Nesbo, C.L., Boucher, Y., and Doolittle, W.F. (2001) Defining the core of nontransferable prokaryotic genes: the euryarchaeal core. *J Mol Evol* **53**: 340-350
- O'Connor, M.B., Peifer, M., and Bender, W. (1989) Construction of large DNA segments in *Escherichia coli*. *Science* **244**: 1307-1312.
- Ochsenreiter, T. (2002) Mikrobielle Ökologie von Archaea und Acidobacteria: von 16S-Diversitätsstudien zur Genomanalyse. In *Institut für Mikrobiologie und Genetik*. Darmstadt: Technische Universität Darmstadt.
- Ochsenreiter, T., Quaiser, A., Lanz, C., Schuster, S.C., and Schleper, C. (in Vorbereitung) Analysis of Genomic Fragments from Uncultivated Acidobacteria.
- Ochsenreiter, T., Selezi, D., Quaiser, A., Bonch-Osmolovskaya, E.A., and Schleper, C. (2002; eingereicht) Diversity and Abundance of Crenarchaeota in Terrestrial Habitats studied by 16S RNA Surveys and Real Time PCR. *Environ Microbiol*.
- Olsen, G.J., and Woese, C.R. (1997) Archaeal genomics: an overview. *Cell* **89**: 991-994.
- Olsen, G.J., Lane, S.J., Giovannoni, S.J., and Pace, N.R. (1986) Microbial Ecology and Evolution: A ribosomal RNA Approach. *Ann. Rev. Microbiol.* **40**: 337-365.
- Ovreas, L., Forney, L., Daae, F.L., and Torsvik, V. (1997) Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. *Appl Environ Microbiol* **63**: 3367-3373.

- Pace, N.R. (1997) A molecular view of microbial diversity and the biosphere. *Science* **276**: 734-740.
- Palleroni, N.J. (1997) Prokaryotic diversity and the importance of culturing. *Antonie Van Leeuwenhoek* **72**: 3-19.
- Pernthaler, A., Pernthaler, J., Schattenhofer, M., and Amann, R. (2002) Identification of DNA-synthesizing bacterial cells in coastal North Sea plankton. *Appl Environ Microbiol* **68**: 5728-5736.
- Preston, C.M., Wu, K.Y., Molinski, T.F., and DeLong, E.F. (1996) A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc Natl Acad Sci U S A* **93**: 6241-6246.
- Prieme, A., Braker, G., and Tiedje, J.M. (2002) Diversity of nitrite reductase (*nirK* and *nirS*) gene fragments in forested upland and wetland soils. *Appl Environ Microbiol* **68**: 1893-1900.
- Quaiser, A., Ochsenreiter, T., Klenk, H.P., Kletzin, A., Treusch, A.H., Meurer, G. et al. (2002a) First insight into the genome of an uncultivated crenarchaeote from soil. *Environ Microbiol* **4**: 603-611.
- Quaiser, A., Ochsenreiter, T., Treusch, A.H., Schleper, C., Lorenz, P., and Eck, J. (2002b) Isolation and Cloning of DNA from Uncultivated Organisms. In *European Patent*. Germany: BRAIN AG.
- Quaiser, A., Ochsenreiter, T., Raddatz, G., Schuster, S.C., and Schleper, C. (in Vorbereitung) Acidobacteria form a coherent group within the bacterial domain, but exchange genes with Rhizobiaceae: insights from environmental genomic studies.
- Reysenbach, A.L., Pace, N.R (1995) Reliable Amplification of Hyperthermophilic Archaeal 16S rRNA genes by Polymerase Chain Reaction. In *Archaea, A Laboratory Manual*. Robb, F.T.P., A.R., Sowers, K.R., Schreier H.J., DasSarma S., Fleischmann E.M. (ed). Rogers, S.L., Colloff, M., and Gomez, D. (2002). Detection and Expression of *nifH* Gene Sequences in Nucleic Acid Extracts from Australian Agricultural Soils [Abstract]. URL [http://www.ciw.csiro.au/conferences/nitrogen/abstracts/A\\_N\\_Talk14\\_Rogers&Colloff.pdf](http://www.ciw.csiro.au/conferences/nitrogen/abstracts/A_N_Talk14_Rogers&Colloff.pdf)
- Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R. et al. (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* **66**: 2541-2547.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**: 5463-5467.
- Schleper, C., and Eck, J. (2000) Umweltgenomik: Charakterisierung und Nutzung nicht-kultivierter Mikroorganismen. *Biospektrum* **6**: 449-451.
- Schleper, C., Holben, W., and Klenk, H.P. (1997a) Recovery of crenarchaeotal ribosomal DNA sequences from freshwater-lake sediments. *Appl Environ Microbiol* **63**: 321-323.
- Schleper, C., Swanson, R.V., Mathur, E.J., and DeLong, E.F. (1997b) Characterization of a DNA polymerase from the uncultivated psychrophilic archaeon *Cenarchaeum symbiosum*. *J Bacteriol* **179**: 7803-7811.
- Schleper, C., DeLong, E.F., Preston, C.M., Feldman, R.A., Wu, K.Y., and Swanson, R.V. (1998) Genomic analysis reveals chromosomal variation in natural populations of the uncultured psychrophilic archaeon *Cenarchaeum symbiosum*. *J Bacteriol* **180**: 5003-5009.
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**: 5857-5864.

- Schulze, R., Meurer, G., and Schleper, C. (2002) Das Metagenom als Quelle neuartiger, rekombinanter Wirkstoffe und Enzyme. *BIOspektrum* **8**: 498-501.
- Selesi, D. (2000) Phylogenetische Diversität nicht kultivierter Archaea und Bacteria in hessischen Sandökosystemen. In *Institut Mikrobiologie und Genetik*. Darmstadt: Technische Universität.
- Sganga, M.W., and Bauer, C.E. (1992) Regulatory factors controlling photosynthetic reaction center and light-harvesting gene expression in *Rhodobacter capsulatus*. *Cell* **68**: 945-954.
- Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* **89**: 8794-8797.
- Simon, H.M., Dodsworth, J.A., and Goodman, R.M. (2000) Crenarchaeota colonize terrestrial plant roots. *Environ Microbiol* **2**: 495-505.
- Smit, E., Leeflang, P., Gommans, S., van den Broek, J., van Mil, S., and Wernars, K. (2001) Diversity and seasonal fluctuations of the dominant members of the bacterial soil community in a wheat field as determined by cultivation and molecular methods. *Appl Environ Microbiol* **67**: 2284-2291.
- Sonnhammer, E.L., Eddy, S.R., Birney, E., Bateman, A., and Durbin, R. (1998) Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res* **26**: 320-322.
- Stein, J.L., Marsh, T.L., Wu, K.Y., Shizuya, H., and DeLong, E.F. (1996) Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *J Bacteriol* **178**: 591-599.
- Stevenson (1982). Properties of humic substances. URL [www.ar.wroc.pl/~weber/kwasy2.htm](http://www.ar.wroc.pl/~weber/kwasy2.htm)
- Storm, C., Herget, I., Kappes, J., and Vormwald, B. (1998) Nährstoffökologische Untersuchungen im Darmstadt-Dieburger Sandgebiet in (teilweise ruderalisierten) Sandpionierfluren und -rasen. *Botanik und Naturschutz in Hessen* **10**: 41-85.
- Swofford, D.L. (2000) PAUP\*, Phylogenetic Analysis Using Parsimony (\*and other methods). *Sinauer Associates* **V4.0b4**.
- Torsvik, V., Ovreas, L., and Thingstad, T.F. (2002) Prokaryotic diversity--magnitude, dynamics, and controlling factors. *Science* **296**: 1064-1066.
- Trevors, J.T., and van Elsas, J.D. (1995) *Nucleic Acids in the Environment: Methods and Application*. Berlin: Springer.
- Ueda, T., Suga, Y., and Matsuguchi, T. (1995) Molecular phylogenetic analysis of a soil microbial community in a soybean field. *European Journal of Soil Science* **46**: 415-421.
- Valinsky, L., Della Vedova, G., Scupham, A.J., Alvey, S., Figueroa, A., Yin, B. et al. (2002) Analysis of bacterial community composition by oligonucleotide fingerprinting of rRNA genes. *Appl Environ Microbiol* **68**: 3243-3250.
- Vetriani, C., Jannasch, H.W., MacGregor, B.J., Stahl, D.A., and Reysenbach, A.L. (1999) Population structure and phylogenetic characterization of marine benthic Archaea in deep-sea sediments. *Appl Environ Microbiol* **65**: 4375-4384.
- Wagner, M., Amann, R., Lemmer, H., and Schleifer, K.H. (1993) Probing activated sludge with oligonucleotides specific for proteobacteria: inadequacy of culture-dependent methods for describing microbial community structure. *Appl Environ Microbiol* **59**: 1520-1525.

- Weidenhaupt, M., Rossi, P., Beck, C., Fischer, H.M., and Hennecke, H. (1996) Bradyrhizobium japonicum possesses two discrete sets of electron transfer flavoprotein genes: fixA, fixB and etfS, etfL. *Arch Microbiol* **165**: 169-178.
- Whitman, W.B., C., C.D., and Wiebe W. J. (1998) Prokaryotes: the unseen majority. *Proc Natl Acad Sci* **95**: 6578-6583.
- Willems, N., and Skurray, R. (1987) Structure and function of the F-factor and mechanism of conjugation. In: Escherichia coli and Salmonella typhimurium. *Cellular and Molecular Biology Neidhardt, F. C. ASM, Washington* **2**: 1110-1133.
- Wise, M.G., McArthur, J.V., and Shimkets, L.J. (1997) Bacterial diversity of a Carolina bay as determined by 16S rRNA gene analysis: confirmation of novel taxa. *Appl Environ Microbiol* **63**: 1505-1514.
- Woese, C.R., Kandler, O., and Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **87**: 4576-4579.
- Wong, S.W., Wahl, A.F., Yuan, P.M., Arai, N., Pearson, B.E., Arai, K. et al. (1988) Human DNA polymerase alpha gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases. *Embo J* **7**: 37-47.
- Yakimov, M.M., Giuliano, L., Timmis, K.N., and Golyshin, P.N. (2001) Upstream-independent ribosomal RNA amplification analysis (URA): a new approach to characterizing the diversity of natural microbial communities. *Environ Microbiol* **3**: 662-666.
- Young, C.C., Burghoff, R.L., Keim, L.G., Minak-Bernero, V., Lute, J.R., and Hinton, S.M. (1993) Polyvinylpyrrolidone-agarose gel electrophoresis purification of polymerase chain reaction-amplifiable DNA from soils. *Appl Environ Microbiol* **59**: 1972-1974.
- Zengler, K., Toledo, G., Rappe, M., Elkins, J., Mathur, E.J., Short, J.M., and Keller, M. (2002) Cultivating the uncultured. *Proc Natl Acad Sci U S A* **99**: 15681-15686.
- Zhou, J., Bruns, M.A., and Tiedje, J.M. (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62**: 316-322.