

**Hochparallele Analyse
natürlicher mikrobieller Lebensgemeinschaften
durch oberflächengebundene Oligonukleotidsonden**

**Highly parallel analysis
of natural microbial communities
by surface-bound oligonucleotide probes**

Dissertation

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

- Dr. rer. nat. -

dem Fachbereich Biologie/Chemie der

Universität Bremen

vorgelegt von

Jörg Peplies

im Juni 2003

Diese Arbeit wurde in der Zeit von Juli 2000 bis Juni 2003 am Max-Planck-Institut für marine Mikrobiologie in Bremen angefertigt, im Rahmen des Forschungs- & Entwicklungsverbundes Gensensorik der Universität Bremen, seit August 2001 Centrum für Angewandte Gensensorik (CAG).

1. Gutachter: Prof. Dr. Rudolf Amann
2. Gutachter: Prof. Dr. Dietmar Blohm

Tag des Promotionskolloquiums: 4. Juli 2003

"We wish to suggest a structure for the salt
of deoxyribose nucleic acid (D.N.A.).
This structure has novel features which are
of considerable biological interest."

James Watson and Francis Crick,
in a brief letter to the journal Nature, April 1953

Zusammenfassung

DNA-Microarrays erlauben durch reverse Hybridisierung die Anwendung von Nukleinsäuresonden im Hochdurchsatz und stellen somit ein leistungsstarkes Werkzeug für den parallelen Nachweis von Mikroorganismen dar. Ihre Anwendbarkeit in der Umweltmikrobiologie wird jedoch durch eine Reihe von methodischen Defiziten eingeschränkt, welche in erster Linie durch die Komplexität dieser Lebensgemeinschaften hervorgerufen werden.

Verschiedene Parameter, die eine oberflächenvermittelte Hybridisierung beeinflussen, wurden anhand eines Modellsystems aus sechs Umweltisolaten und zwanzig gegen die 16S rRNA-gerichtete Oligonukleotidsonden systematisch untersucht. Es zeigte sich, dass Spezifität und Sensitivität der Hybridisierung direkt miteinander verknüpft sind. Unter geeigneten Hybridisierungsbedingungen konnten falsch-positive Signale fast vollständig eliminiert und somit alle sechs Stämme eindeutig differenziert werden. Die Zahl der gehäuft aufgetretenen falsch-negativen Ereignisse konnte durch eine gerichtete Immobilisierung der Fängersonden reduziert werden.

Für die Microarray-Analyse von komplexen natürlichen Gemeinschaften wurde ein Protokoll etabliert, welches die direkte chemische Markierung von extrahierten rRNA-Molekülen erlaubt. Auf diese Weise wird eine vorgeschaltete PCR-Amplifikation der Zielmoleküle und somit mögliche Populationsverschiebungen vermieden. Die Ergebnisse aus der Untersuchung von Proben unbekannter Zusammensetzung wurden durch quantitative Fluoreszenz-*in situ*-Hybridisierung (FISH) validiert. Bakterielle Gemeinschaften ansteigender Komplexität wurden untersucht und häufige Vertreter konnten spezifisch nachgewiesen werden. Die parallel durchgeführte FISH-Analyse deutete aber darauf hin, dass Microarrays gegenwärtig bestenfalls semiquantitative Daten bereitstellen. Insgesamt wurde damit ein einfaches und schnelles Protokoll etabliert, das den Microarray-vermittelten Nachweis von großen mikrobiellen Populationen (ca. >5%) auch in komplexen und nährstoffarmen Systemen erlaubt.

Die Anwendung von phylogenetisch redundanten und/oder hierarchisch geordneten Sonden ist eine Grundvoraussetzung für den zuverlässigen Nachweis von Mikroorganismen durch Hybridisierungstechniken. Mögliche weitere Sondenbindungsstellen können durch die zusätzliche Verwendung der 23S rRNA als Zielmolekül bereitgestellt werden. Gegenwärtig sind allerdings nur wenige Vollsequenzen dieses Markers in den öffentlichen Datenbanken hinterlegt. Die nahezu vollständigen 23S rRNA-Gene von 11 aus der Nordsee isolierten alpha-Proteobakterien wurden in dieser Arbeit sequenziert. Fünf dieser Stämme gehören der marinen alpha-Gruppe an, deren Vertreter im marinen Freiwasser häufig sind. Die vergleichende Sequenzanalyse von 16S und 23S rRNA-Genen ergab Stammbäume mit ähnlicher Topologie innerhalb der alpha-Proteobakterien. Die Eignung beider Marker für die Entwicklung von spezifischen Oligonukleotidsonden wurde *in silico* systematisch untersucht und verglichen. In einem Modellsystem konnte die Gesamtzahl der verfügbaren phylogenetisch redundanten Sonden durch die zusätzliche Verwendung der 23S rRNA deutlich erhöht werden.

Summary

DNA microarrays represent a high throughput format for the application of nucleic acid probes by reverse hybridization and, therefore, provide a powerful tool for the highly parallel detection of microorganisms. However, their application in molecular ecology is hampered by different methodological limitations mainly because distinct nucleic acids have to be identified against an often large and unknown genetic background.

Various parameters affecting a surface-mediated hybridization were systematically evaluated with a model system composed of six environmental strains and 20 oligonucleotide probes targeting the 16S rRNA. These investigations revealed that specificity and sensitivity of microarray hybridizations are directly linked. With adequate hybridization conditions, false-positive signals could be almost completely prevented, resulting in a clear identification of the six strains analyzed. False-negative results that were more common within the data sets could be decreased by directed application of the capture oligonucleotides.

For microarray analysis of mixed microbial communities, a protocol for direct chemical labeling of extracted rRNA was established to avoid potential biases introduced by a prior PCR step. Results obtained for samples of unknown composition were validated by fluorescence *in situ* hybridization (FISH). Samples of increasing complexity were analyzed and abundant community members could specifically be detected by microarray hybridization. Quantitative data from parallel FISH experiments suggested that current microarray data provide, at best, semiquantitative information on community composition. In summary, a fast and easy protocol for the microarray-mediated identification of major bacterial populations (approx. >5%), even in complex and oligotrophic systems is presented.

Application of phylogenetically redundant and/or hierarchical structured probe sets is a prerequisite for the reliable hybridization-based identification of microorganisms. A potential reservoir of additional probe target sites is represented by the 23S rRNA. Currently, only a small number of sequences is available for this marker in the public databases. In this study, almost complete 23S rRNA gene sequences were obtained from eleven alpha-proteobacteria isolated from North Sea bacterioplankton, including five members of the marine alpha cluster, which is abundant in marine surface waters. Comparative sequence analysis revealed consistent tree topologies for the alpha-proteobacteria based on their 16S and 23S rRNA genes. The suitability of both markers for the design of specific oligonucleotide probes was systematically evaluated and compared *in silico*. The overall number of phylogenetic redundant probes available could be significantly increased by also targeting the 23S rRNA.

Inhaltsverzeichnis

Zusammenfassung

Summary

Teil I: Darstellung der Ergebnisse im Gesamtzusammenhang

A	Einleitung	1
1	Gängige molekularbiologische Methoden für die Analyse von Diversität und Struktur natürlicher mikrobieller Gemeinschaften	5
1.1	Methoden für die Erfassung mikrobieller Diversität	6
1.1.1	Der klassische Ansatz: Klonierung und Sequenzierung von 16S rRNA-Genen aus der Umwelt	6
1.1.2	Fingerprinting-Techniken	7
1.2	Direkter Nachweis von ausgewählten Zielsequenzen mittels Hybridisierungstechniken	9
1.2.1	Quantitative "dot blot"-Hybridisierung.....	10
1.2.2	Ganzzell- und <i>in situ</i> -Hybridisierung.....	10
2	Ein leistungsstarkes molekularbiologisches Werkzeug: DNA-Microarrays	13
2.1	Ursprung, Entwicklung und gängige Anwendungen	14
2.2	Nachweis von natürlichen mikrobiellen Populationen und hieraus erwachsende Anforderungen	17
3	Zielsetzung dieser Arbeit.....	20
B	Ergebnisse und Diskussion	25
1	DNA-Microarrays für den hochparallelen Nachweis von 16S rRNA-Zielsequenzen in Proben unterschiedlicher Komplexität	25
1.1	Identifizierung von bakteriellen Reinkulturen und systematische Untersuchungen an definierten Modellsystemen	25
1.1.1	Parameteranalyse	26
1.1.2	Optimierungsstrategien	30
1.2	Analyse von bakteriellen Gemeinschaften unbekannter Zusammensetzung	33
1.2.1	Direkter Nachweis von 16S rRNA-Molekülen.....	33
1.2.2	FISH-Validierung der Ergebnisse aus Mischkulturen und Umweltproben	35
2	Vergleichende Sequenzanalyse und Sondenentwicklung basierend auf 23S rRNA-Genen.....	37
3	Ausblick.....	40
C	Literaturverzeichnis	47

Teil II: Publikationen

Publikationsliste mit Erläuterungen	63
1 Optimization strategies for DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes	67
2 Application and validation of DNA microarrays for the identification of 16S rRNA targets in marine surface water	81
3 Comparative sequence analysis and oligonucleotide probe design based on the 23S rRNA genes of alpha-proteobacteria from North Sea bacterioplankton	107

Teil I
Darstellung der Ergebnisse im
Gesamtzusammenhang

A Einleitung

Mit der Einführung von molekularbiologischen Methoden, die auf der vergleichenden Sequenzanalyse von genetischen Markern beruhen, hat die Disziplin der mikrobiellen Ökologie Mitte der 80er Jahre des vergangenen Jahrhunderts einen Quantensprung vollzogen. Die Arbeiten der Gruppe um Norman Pace ermöglichten es erstmals, Mikroorganismen ohne deren vorherige Kultivierung in Umweltproben nachzuweisen (Olsen *et al.*, 1986; Pace *et al.*, 1986). Hierfür wurde das etwa 10 Jahre zuvor von Carl Woese entwickelte Gerüst zur phylogenetischen Klassifizierung von ribosomalen RNA-Molekülen (Woese, 1987) mit Techniken verbunden, welche die Isolierung, Vereinzeln, Identifizierung und Analyse dieser molekularen Marker aus komplexen Proben erlauben. Bedenkt man, dass mehr als 90% der in der Umwelt vorkommenden Bakterien mit den gängigen Methoden nicht kultivierbar sind (Amann *et al.*, 1995), wird der Stellenwert dieser Errungenschaft für die mikrobielle Umweltforschung deutlich.

Während die Pionierstudien noch auf die Analyse der nur etwa 120 Nukleotide langen 5S rRNA beschränkt waren, setzte sich, bedingt durch verbesserte Techniken für die DNA-Sequenzierung, bald die 16S rRNA als phylogenetischer Marker durch. Mit einer Länge von etwa 1500 Nukleotiden weist diese einen deutlich höheren Informationsgehalt als die 5S rRNA auf und erlaubt somit eine zuverlässigere Klassifikation (Olsen *et al.*, 1986). Erste umfangreiche Studien, die sich einer PCR-basierten Variante des oben beschriebenen Ansatzes bedienten, lieferten unerwartete Ergebnisse. In Genbanken aus marinem Freiwasser bzw. einer heißen Quelle konnten ausschließlich Sequenzen gefunden werden, die nicht mit den in den Datenbanken hinterlegten übereinstimmten (Giovannoni *et al.*, 1990; Ward *et al.*, 1990). In den folgenden Jahren schlossen sich zahlreiche Studien an, die diese Ergebnisse bestätigten. Es zeigte sich zunehmend, dass die bakterielle Diversität zuvor erheblich unterschätzt worden war, und dass molekulare Techniken, verglichen mit den klassischen Kultivierungsansätzen, einen erweiterten Einblick in die Zusammensetzung natürlicher mikrobieller Gemeinschaften geben können. Ein prominentes Beispiel hierfür ist der Nachweis von Mitgliedern der Domäne *Archaea* im marinen "Bakterioplankton" (DeLong, 1992; Fuhrman *et al.*, 1992), da man zuvor davon ausgegangen war, dass das Vorkommen dieser Organismen auf "extreme" Standorte beschränkt sei.

Heute (April 2003), etwa 15 Jahre später, sind fast 70.000 16S rRNA Sequenzen in den öffentlichen Datenbanken hinterlegt (Cole *et al.*, 2003) und ihre Zahl steigt weiter. Auch das Spektrum der molekularbiologischen Methoden, die auf der vergleichenden Sequenzanalyse von ribosomalen Genen beruhen, wurde seit ihrer Einführung stetig erweitert. Eine zentrale Rolle kommt hierbei den Nukleinsäure-Hybridisierungstechniken zu (Amann *et al.*, 1995; Head *et al.*, 1998). Sie erlauben den gezielten Nachweis ausgewählter Populationen und deren Quantifizierung, auch in ihrem natürlichen Habitat (Amann and Schleifer, 2001). Ermöglicht wird dies durch den Einsatz von kurzen Oligonukleotidsonden, die gegen Sequenzabschnitte unterschiedlicher evolutionärer Konserviertheit auf der 16S rRNA gerichtet sind. Auf diese Weise lassen sich neben Sonden, die ausgewählte phylogenetische Gruppen unterscheiden, auch solche konstruieren, die gattungs-, art- oder sogar unterartspezifisch sind (Göbel *et al.*, 1987; Giovannoni *et al.*, 1988; Stahl and Amann, 1991). Eine einschneidende Erkenntnis der quantitativen Hybridisierungstechniken war, dass nur wenige Mikroorganismen mit hohen *in situ*-Häufigkeiten in den existierenden Stammsammlungen vertreten sind. Durch die Kombination von molekulargenetischen und klassischen mikrobiologischen Techniken konnten kürzlich erste Erfolge bei der gezielten Kultivierung abundanter Organismen erzielt werden (Pinhassi *et al.*, 1997; Eilers *et al.*, 2001; Rappe *et al.*, 2002).

Neben der Beschreibung von Diversität und Struktur natürlicher mikrobieller Gemeinschaften tritt zunehmend die Frage nach der Funktion ihrer einzelnen Vertreter in den Vordergrund (DeLong, 2002; Fuhrman, 2002). Das Wissen über den Beitrag verschiedener Mikroorganismen an globalen Stoffumsetzungen, beispielsweise der jährlichen Menge an oxidiertem Methan oder fixiertem Stickstoff, ist nicht nur für die Grundlagenforschung von großem Interesse. In der klassischen Mikrobiologie werden Reinkulturen für die Bestimmung von physiologischen Eigenschaften ausgewählter Spezies benötigt. Dies ist zum Einen sehr zeitaufwendig, zum Anderen führen die niedrigen Kultivierungseffizienzen zu einer geringen Zahl an Isolaten, deren ökologische Relevanz zusätzlich unklar ist. Aus den genannten Gründen haben sich die Mikrobiologen bei der Untersuchung funktioneller Aspekte von Ökosystemen lange Zeit auf eine Betrachtung der *in situ*-Umsatzraten beschränkt, die vorgefundenenen mikrobiellen Gemeinschaften also als funktionale Einheiten oder "black box" betrachtet. Auch wenn die grundlegenden Wechselwirkungen zwischen Bakteriengemeinschaft und anderen Elementen der untersuchten Nahrungsnetze auf diese Weise aufgeklärt werden konnten, erlaubt dieses Vorgehen keinen detaillierten Einblick in die Interaktion einzelner Organismen mit ihrer Umwelt - ein unbefriedigender Zustand, bedenkt man, dass

beispielsweise im marinen Freiwasser, verschiedene Bakterien unterschiedliche Bestandteile des verfügbaren organischen Materials abbauen (Martinez *et al.*, 1996).

Aus Verwandtschaftsverhältnissen, basierend auf der Sequenzanalyse von rRNA-Molekülen, können aufgrund der physiologischen Vielfalt der Bakterien, selbst unter nahverwandten Arten, selten Rückschlüsse auf das metabolische Potenzial oder die spezifische Aktivität einzelner Spezies abgeleitet werden (Fuhrman, 2002). Ein prominentes Beispiel sind in diesem Zusammenhang die Vertreter der marinen alpha-Proteobakterien, bei denen es sich um die Gattung *Roseobacter* und Nahverwandte handelt. Diese Gruppe weist hohe *in situ*-Abundanzen im marinen Freiwasser und zugleich eine Reihe von kultivierten Vertretern auf (Gonzalez and Moran, 1997; Eilers *et al.*, 2001). Anhand der physiologischen Eigenschaften von charakterisierten Isolaten wurde eine Beteiligung dieser Bakterien an der Umsetzung von schwefelhaltigen Verbindungen postuliert, welche an der globalen Klimaregulation beteiligt sind (Gonzalez *et al.*, 1999). Allerdings konnten phylogenetisch hochaufgelöste *in situ*-Studien zeigen, dass die marine alpha-Gruppe in der Umwelt durch bisher nichtkultivierte Vertreter dominiert wird (Eilers *et al.*, 2001; Zubkov *et al.*, 2001). Dass diese Untergruppe entsprechende metabolische Fähigkeiten aufweist, kann nur vermutet werden.

Auch hier bieten molekularbiologische Ansätze erweiterte Möglichkeiten, da sie neben dem reinen Nachweis von Organismen, auch die kultivierungsunabhängige Untersuchung von funktionellen Aspekten ausgewählter Populationen oder Spezies erlauben. Beispiele für entsprechende Methoden sind die Kombination von *in situ*-Hybridisierungstechniken mit mikroaudioradiographischen Verfahren für die parallele Visualisierung von phylogenetischer Identität und metabolischer Leistung auf Einzelzellebene (Lee *et al.*, 1999; Ouverney and Fuhrman, 1999) oder die Klonierung und Analyse von großen genomischen DNA-Fragmenten aus der Umwelt (Beja *et al.*, 2000; Rondon *et al.*, 2000). Letztere erlaubt einen Einblick in die genetische Struktur unkultivierter Mikroorganismen und, durch den Gebrauch von bioinformatischen Werkzeugen, die Identifizierung von potenziellen physiologischen Eigenschaften dieser "genomischen Einheiten".

So, wie die Nutzung von molekulargenetischen Werkzeugen der mikrobiellen Ökologie neue, ungeahnte Möglichkeiten eröffnet hat, wurde vor etwa einem Jahrzehnt das Feld der molekularbiologischen Forschung in seiner ganzen Breite durch das Aufkommen eines Werkzeuges revolutioniert, welches als DNA-Chip- oder DNA-Microarray-Technologie bezeichnet wird (Ramsay, 1998). In diesem reversen Hybridisierungsformat werden oberflächenimmobilisierte Nukleinsäuresonden für das spezifische Einfangen von markierten

Zielmolekülen eingesetzt. Dieser Ansatz ermöglicht bei ausreichender Miniaturisierung die parallele Analyse einer nahezu unbegrenzten Anzahl von Zielsequenzen in einem einzigen Experiment (Lipshutz *et al.*, 1999). Ursprünglich für die *de novo*-Sequenzierung genomischer DNA-Fragmente konzipiert (Bains and Smith, 1988; Drmanac *et al.*, 1989; Kharpko *et al.*, 1989), zeigte sich schnell das Potenzial dieser Technik für die funktionelle Genomanalyse (Skena *et al.*, 1995). Heute besteht die Hauptanwendung von DNA-Microarrays - wie sie in Folge bezeichnet werden - in der Suche nach differentiell exprimierten Genen in eukaryontischen Transkriptomen, wobei mehrere tausend Gene gleichzeitig analysiert werden können (Lander, 1999).

Auch für Anwendungen in der mikrobiellen Ökologie ist die DNA-Microarray-Technologie von großem Interesse. Hier könnte sie den Einsatz von Nukleinsäuresonden für den Nachweis von Zielsequenzen in Umweltproben durch eine massive Parallelisierung der Analyse erheblich beschleunigen, und damit, wie in der Genomforschung, eine völlig neue Dimension von Untersuchungen ermöglichen. Ein Beispiel für die Relevanz dieser Entwicklung läßt sich aus Arbeiten ableiten, die eine ausgeprägte zeitliche Dynamik für einzelne Populationen des marinen Picoplanktons zeigen konnten (Eilers *et al.*, 2001; Karner *et al.*, 2001). Mit Hilfe von DNA-Microarrays, die auf dem Einsatz von gegen rRNA-gerichteten Sonden beruhen, sollte sich innerhalb kürzester Zeit eine Vielzahl von Taxa auf ihre Anwesenheit in einer Probe komplexer Zusammensetzung testen lassen (Guschin *et al.*, 1997). Eine grundsätzliche Eignung der Microarray-Technologie für den Nachweis von funktionellen Genen in Nukleinsäuregemischen, die aus komplexen Proben extrahiert wurden, konnte ebenfalls gezeigt werden (Wu *et al.*, 2001; Dennis *et al.*, 2003; Taroncher-Oldenburg *et al.*, 2003).

Im Blickpunkt dieser Arbeit steht die Anwendung von DNA-Microarrays für die Untersuchung von Diversität und Struktur natürlicher mikrobieller Gemeinschaften. In der Folge sollen zunächst die gängigen molekularbiologischen Verfahren für eine kultivierungsunabhängige Analyse aufgezeigt und ihre jeweiligen Vor- und Nachteile erörtert werden. Es schließt sich dann eine nähere Betrachtung der DNA-Microarray-Methodik an - von ihren Ursprüngen, über gängige Anwendungen bis hin zu den besonderen Anforderungen, die sich aus der Untersuchung von komplexen mikrobiellen Gemeinschaften ergeben.

1 Gängige molekularbiologische Methoden für die Analyse von Diversität und Struktur natürlicher mikrobieller Gemeinschaften

Der mikrobiellen Ökologie stehen eine Reihe von Standardtechniken für die kultivierungsunabhängige Untersuchung der Diversität und Struktur natürlicher mikrobieller Populationen zur Verfügung. In der heute üblichen Form des als "rRNA-Ansatz" bezeichneten Verfahrens wird idealerweise eine Kombination aus verschiedenen molekularbiologischen Methoden für die Untersuchung von komplexen Proben unbekannter Zusammensetzung herangezogen (Amann *et al.*, 1995), da diese ihre volle Leistungsfähigkeit erst in ihrem Zusammenspiel entfalten. Durch das zyklische Durchlaufen von Sequenzierungs- und Hybridisierungsphasen wird gleichzeitig eine interne Fehleranalyse ermöglicht. Eine Übersicht ist in Abb. 1 gegeben.

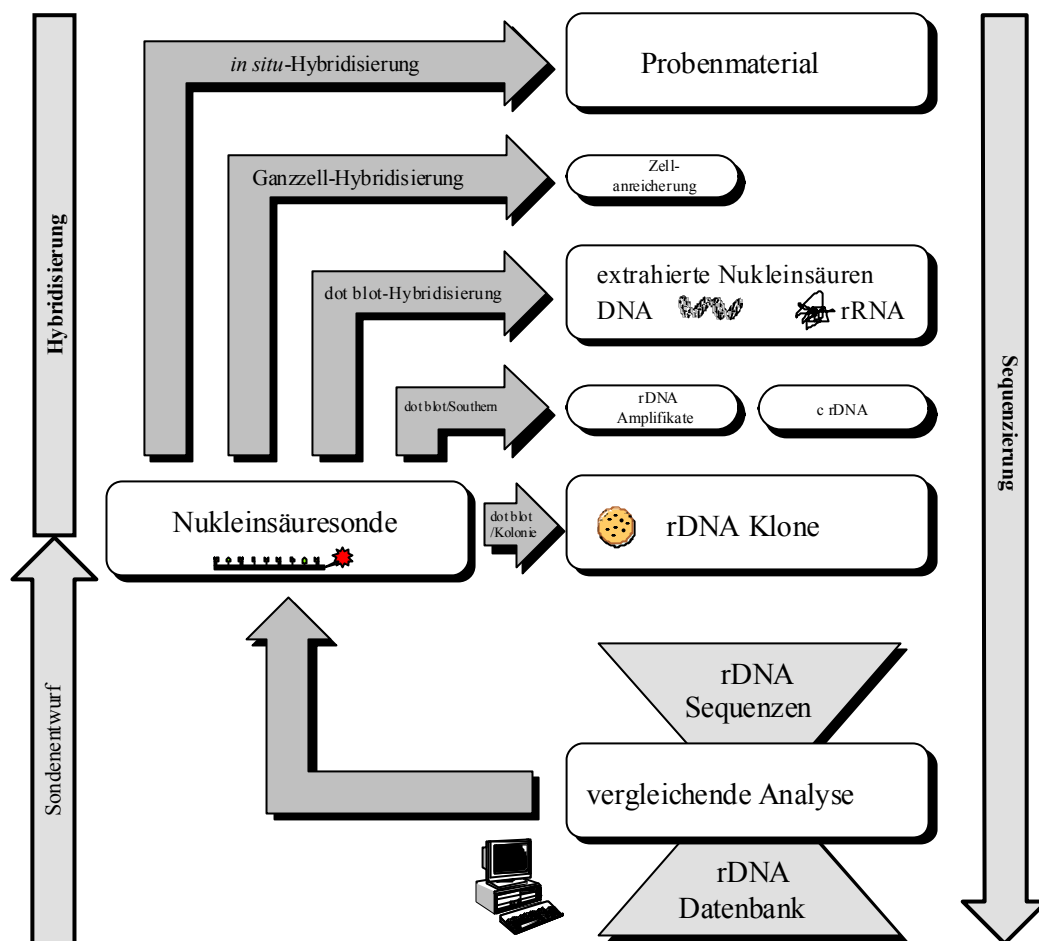


Abbildung 1: Der rRNA-Ansatz (modifiziert aus Glöckner, 1999)

1.1 Methoden für die Erfassung mikrobieller Diversität

1.1.1 Der klassische Ansatz: Klonierung und Sequenzierung von 16S rRNA-Genen aus der Umwelt

Die ersten Versuche, natürliche mikrobielle Gemeinschaften mit Hilfe von molekularbiologischen Methoden zu charakterisieren, basierten auf der direkten Extraktion, Vereinzelung und Sequenzanalyse von 5S rRNA-Molekülen aus Umweltproben (Stahl *et al.*, 1984; Stahl *et al.*, 1985). Dieser Ansatz war aber durch den geringen Informationsgehalt der 5S rRNA und die Notwendigkeit einer elektrophoretischen Auftrennung der verschiedenen 5S rRNA-Moleküle auf Ökosysteme geringer Komplexität beschränkt. Fast zeitgleich wurde daher die Betrachtung der längeren 16S rRNA vorgeschlagen (Olsen *et al.*, 1986). In ersten Anwendungen wurde die aus der Probe extrahierte Gesamt-DNA fragmentiert und kloniert. Es schloss sich dann ein Screening der Genbank mit Hilfe von 16S rDNA-spezifischen Sonden an (der Begriff 16S rDNA beschreibt das Gen der 16S rRNA oder dessen Amplifikat), um entsprechende Klone ausfindig zu machen (Schmidt *et al.*, 1991). Alternativ kann die exprimierte 16S rRNA direkt aus Umweltproben extrahiert, in cDNA umgeschrieben und anschließend kloniert werden (Weller and Ward, 1989). Hierdurch entfällt das aufwendige Screening der Genbanken. Auch ergeben sich Vorteile für die initiale Nukleinsäureextraktion, da eine größere Zahl von Zielmolekülen in der untersuchten Probe vorliegt.

Mit der Einführung der Polymerase-Kettenreaktion (PCR) (Saiki *et al.*, 1988) ließ sich der Ansatz zur Klonierung und Analyse von Umweltsequenzen deutlich vereinfachen. In seiner heute gängigen Form, werden allgemeine oder für eine Gruppe von Organismen spezifische Primerpaare für die selektive Amplifizierung von 16S rRNA-Genfragmenten aus komplexen DNA-Proben eingesetzt. Die verschiedenen Klone der auszuwertenden Genbank tragen dann ausschließlich definierte Fragmente, welche über die bekannten Primer-Bindungsstellen einfach und schnell sequenziert werden können (Amann *et al.*, 1995; Head *et al.*, 1998).

Die PCR-basierten 16S rDNA-Genbanken geben zwar einen Einblick in die prokaryontische Artenvielfalt eines Lebensraumes, von einer Korrelation der Häufigkeit eines rDNA-Klones in einer Genbank und der Häufigkeit des zugrundeliegenden Mikroorganismus im beprobtem Habitat darf aber nicht ausgegangen werden (Wintzingerode *et al.*, 1997). Verschiedene Arbeiten haben gezeigt, dass die 16S rRNA-Gene der unterschiedlichen Organismen in einer komplexen Probe nicht quantitativ amplifiziert werden

(Reysenbach *et al.*, 1992; Suzuki and Giovannoni, 1996). Des Weiteren können Populationen völlig von der Amplifikation ausgeschlossen werden, da selbst hochkonservierte Primer-Bindungsstellen eine gewisse Variabilität aufweisen (Marchesi *et al.*, 1998; Schmalenberger *et al.*, 2001).

Neben den genannten technischen Aspekten, wird der Rückschluss von Klon-Häufigkeiten auf zelluläre Abundanzen durch die variierende Anzahl von rRNA-Gen-Operonen auf bakteriellen Genomen limitiert. Beschrieben wurde das Vorkommen von einem Operon für eine Vielzahl von Organismen bis hin zu 15 Operonen für *Clostridium paradoxum* (Klappenbach *et al.*, 2001). Für quantitative Betrachtungen in natürlichen mikrobiellen Systemen sind die in Abschnitt 1.2 aufgeführten Hybridisierungstechniken besser geeignet.

Weitere Nachteile einer PCR-basierten Untersuchung von komplexen Gemeinschaften sind die grundsätzliche Empfindlichkeit dieser Methodik gegenüber Kontaminationen, der Falscheinbau von Nukleotiden durch die DNA-Polymerase (Chen *et al.*, 1991) und die mögliche Ausbildung von Chimären, also von PCR-Produkten, die sich aus den 16S rRNA Genen verschiedener Arten zusammensetzen (Liesack *et al.*, 1991). Dies kann, wie die Operondiversität einzelner Stämme (Clayton *et al.*, 1995), zu einer Überschätzung der wirklich vorhandenen mikrobiellen Diversität in einem untersuchten Habitat führen.

Neben möglichen Artefakten, die durch einen vorgeschalteten PCR-Schritt verursacht werden, existieren bei der Erstellung von 16S rRNA-Genbanken weitere Fehlerquellen, die zu quantitativen Verschiebungen oder dem Ausschluss einzelner Populationen führen können. Zu nennen sind hier die unterschiedlich effiziente Zellyse von Mitgliedern verschiedener Taxa, sowie abweichende Klonierungseffizienzen für verschiedene 16S rDNA-Moleküle (Head *et al.*, 1998).

Als Konsequenz aus den aufgeführten Limitationen bei der Klonierung und Analyse von 16S rRNA-Sequenzen aus der Umwelt ergibt sich, dass die in Abschnitt 1.2 aufgeführten, PCR-unabhängigen Methoden für eine Validierung der erhaltenen Ergebnisse unverzichtbar sind.

1.1.2 Fingerprinting-Techniken

Im Gegensatz zu Verfahren, die auf der Sequenzanalyse zuvor klonierter DNA-Fragmente beruhen, erlauben sogenannte Fingerprinting-Techniken einen schnellen Einblick in die Komplexität einer Umweltprobe. In ihrer grundsätzlichen Funktionsweise gleichen sich die

verschiedenen Ansätze: mittels PCR wird unter Verwendung von universellen, gegen die 16S rRNA-gerichteten Primern ein Gemisch unterschiedlicher DNA-Fragmente erzeugt, gegebenenfalls nachbehandelt, und abschließend in einem Gel elektrophoretisch aufgetrennt. Das erhaltene Bandenmuster liefert den sogenannten Fingerabdruck, der charakteristisch für eine spezifische Gemeinschaftsstruktur sein sollte.

Die gängigste Fingerprinting-Technik in der mikrobiellen Umweltforschung ist die von Muyzer *et al.* (1993) eingeführte "Denaturierende Gradienten Gelelektrophorese" (DGGE). Es werden durch PCR generierte DNA-Fragmente identischer Länge anhand von Sequenzunterschieden in einem denaturierenden Gradienten aufgetrennt. Einzelne Banden im Gel, also einzelne DNA-Fragmente, können dann ausgeschnitten, reamplifiziert und sequenziert werden. In der Folge lassen sich ausgewählte Populationen in verschiedenen Proben schnell auf ihre Anwesenheit testen. Die Auflösung dieses Ansatzes ist allerdings beschränkt, da nicht alle Sequenzunterschiede aufgetrennt werden können.

Die SSCP-Technik (Single Strand Conformation Polymorphism) ähnelt in ihrer Funktionsweise der DGGE. Die Auftrennung der erzeugten Fragmente anhand von Sequenzunterschieden wird hier durch die Sekundärstruktur einzelsträngiger DNA-Moleküle vermittelt (Lee *et al.*, 1996). Allerdings weist diese Technik eine geringere Reproduzierbarkeit auf. Auch werden sehr kurze DNA-Fragmente (etwa 200 bp) für eine saubere Auftrennung im Gel benötigt, wodurch die nach Sequenzierung erhaltene phylogenetische Information noch weiter reduziert wird.

Einem abweichenden Funktionsprinzip unterliegt die ARDRA (Amplified Ribosomal DNA Restriction Analysis). Hier werden die erzeugten DNA-Fragmente mit Restriktionsenzymen geschnitten, um dann in Abhängigkeit ihrer Länge in einem Gel aufgetrennt zu werden. Auf diese Weise läßt sich eine Vielzahl von Proben schnell und einfach miteinander vergleichen (Martinez-Murcia *et al.*, 1995). Allerdings lassen sich ausgewählte Fragmente nicht ohne weiteres sequenzieren, da definierte Primerbindungsstellen fehlen.

Eine weitere häufig genutzte Fingerprinting-Technik ist die T-RFLP-Technik (Terminal Restriction Fragment Length Polymorphism). Die PCR-Amplifikation der 16S rRNA-Gene wird unter Verwendung eines fluoreszenzmarkierten Primers durchgeführt. Anschließend werden die Produkte mit Restriktionsenzymen geschnitten und in einem DNA-Sequenzierer aufgetrennt (Liu *et al.*, 1997). Dieser Ansatz erlaubt bei einem hohen Probendurchsatz eine schnelle und sehr genaue Auftrennung. Die Sequenzanalyse von ausgewählten Fragmenten ist aber nicht möglich.

Die Extraktion und elektrophoretische Auftrennung von niedermolekularer RNA wie der 5S rRNA kann ebenfalls als Fingerprinting-Technik betrachtet werden (Höfle and Brettar, 1996). Die Limitationen dieses Ansatzes wurden bereits beschrieben.

Im Vergleich zu Methoden, die auf der Klonierung von 16S rDNA-Molekülen beruhen, erlauben die Fingerprinting-Techniken keinen detaillierten Einblick in die 16S rRNA-Diversität einer Probe, u.a. bedingt durch ihre limitierte Auflösung bei der Erfassung von DNA-Sequenzunterschieden. Diese Einschränkung wird durch die Analyse von vergleichsweise kurzen Sequenzabschnitten noch verstärkt. Da in allen Fällen ein PCR-Schritt vorgeschaltet ist, unterliegen diese Techniken ebenfalls den hieraus erwachsenden Limitationen. Allerdings ermöglichen sie die schnelle und einfache Analyse einer Vielzahl von Proben und eignen sich daher gut für die Untersuchung von räumlichen oder zeitlichen Populationsverschiebungen.

1.2 Direkter Nachweis von ausgewählten Zielsequenzen mittels Hybridisierungstechniken

Für eine Verifizierung der Ergebnisse aus den bisher beschriebenen Methoden des rRNA-Ansatzes bieten sich Hybridisierungstechniken an, die auf dem Einsatz von markierten Nukleinsäuresonden beruhen. Der Nachweis von erhaltenen Umweltsequenzen kann dann auf Ebene der anfänglich extrahierten Nukleinsäuren, der PCR-Amplifikate oder idealerweise in den Zellen der ursprünglichen Probe durchgeführt werden (vgl. Abb. 1).

Aus verschiedenen Gründen hat sich die 16S rRNA als geeignetes Zielmolekül für die Identifizierung von Mikroorganismen durch Nukleinsäure-Hybridisierung erwiesen (Stahl and Amann, 1991). (i) Ribosomale RNA-Moleküle kommen in jeder lebenden Zelle mit einer Kopienzahl von mindestens 10^3 vor, (ii) sie sind aufgrund ihrer Stabilität mit den gängigen Methoden leicht zu extrahieren, und (iii) anhand der umfangreichen Datensätze kann die Spezifität der Sonden *in silico* abgeschätzt werden. Zudem ermöglichen Bereiche unterschiedlicher evolutionärer Konserviertheit entlang des rRNA-Moleküls die Analyse auf verschiedenen phylogenetischen Ebenen.

Mit Hilfe der im Folgenden vorgestellten Methoden lassen sich auch die relativen Häufigkeiten der detektierten molekularen Marker (Abschnitt A1.2.1) und sogar die zellulären Abundanzen von ausgewählten Populationen (Abschnitt 1.2.2) in einer Probe bestimmen.

Weitere Hybridisierungstechniken, die auf dem Einsatz von längeren DNA- oder RNA-Sonden beruhen, werden aufgrund ihrer geringen Verbreitung bei dieser Betrachtung nicht berücksichtigt.

1.2.1 Quantitative "dot blot"-Hybridisierung

Grundsätzlich existieren zwei verschiedene Ansätze, Oligonukleotidsonden, die gegen die 16S rRNA gerichtet sind, für den quantitativen Nachweis von bakteriellen Populationen aus der Umwelt einzusetzen. Die dot blot-Hybridisierung (Kafatos *et al.*, 1979) wurde 1988 von Stahl *et al.* an die Fragestellungen der mikrobiellen Ökologie angepasst und in der Folge vielfach angewandt (Raskin *et al.*, 1994; Gonzalez and Moran, 1997; MacGregor *et al.*, 1997; Sahn *et al.*, 1999; Ravensschlag *et al.*, 2000). Dabei werden die Gesamtnukleinsäuren aus einer Probe extrahiert und in einer runden Aussparung (dot) direkt auf einer Membran immobilisiert. Mit Hilfe einer spezifischen und einer allgemeinen Oligonukleotidsonde, die unter geeigneten Bedingungen gegen diese "dots" hybridisiert werden, kann die rRNA ausgewählter Populationen im Verhältnis zur extrahierten Gesamt-rRNA auch parallel für unterschiedliche Proben detektiert und quantifiziert werden. Aus den erhaltenen Werten lassen sich allerdings keine relativen zellulären Abundanzen ableiten, da die Zahl der Ribosomen pro Zelle für verschiedene bakterielle Spezies schätzungsweise zwischen 10^3 und 10^5 schwankt (Amann *et al.*, 1995). Zusätzlich variiert der zelluläre rRNA-Gehalt einzelner Stämme in Abhängigkeit der Wachstumsrate deutlich (Kemp *et al.*, 1993; Wallner *et al.*, 1993). Die relative Häufigkeit ausgewählter ribosomaler RNA-Moleküle sollte aber eine brauchbare Kenngröße für die relative physiologische Aktivität der entsprechenden Population darstellen, da sie sich aus dem Produkt aus der Anzahl an detektierten Zellen und deren mittleren rRNA-Gehalt ergibt. Durch die vorherige Freisetzung der zu detektierenden Zielmoleküle unterliegt die dot blot-Hybridisierung ebenfalls den bereits aufgezeigten Limitationen entsprechender Ansätze.

1.2.2 Ganzzell- und *in situ*-Hybridisierung

Im Gegensatz zu Verfahren wie der Southern-, dot blot- oder Koloniehybridisierung beruht die Ganzzellhybridisierung nicht auf einer Freisetzung der Zielnukleinsäuren aus dem Zellmaterial. Die markierten Sonden werden in fixierte, morphologisch aber intakte Zellen eingeschleust und binden dort spezifisch an ihre Zielsequenzen. Nichtgebundene Sonden

können durch stringentes Waschen anschließend wieder selektiv aus den Zellen entfernt werden. Diese Technik ermöglicht es, Organismen in ihrem natürlichen Habitat zu identifizieren, gebräuchlich ist daher die Bezeichnung *in situ*-Hybridisierung. Durch die Verwendung von rRNA-spezifischen Sonden können auch bisher nicht kultivierte Organismen visualisiert werden (Amann *et al.*, 1995). Für Fragestellungen der mikrobiellen Ökologie wurde diese Methode erstmals in der Gruppe von Norman Pace in den späten 1980er Jahren eingesetzt (Giovannoni *et al.*, 1988; DeLong *et al.*, 1989). In der folgenden Zeit hat sich die *in situ*-Hybridisierung zu einem integralen Teil des rRNA-Ansatzes entwickelt, durch deren Anwendung die in den Genbanken vorgefundene Diversität auch *in situ* gezeigt werden konnte (Amann *et al.*, 1996).

Werden fluoreszenzmarkierte Sonden eingesetzt, spricht man von Fluoreszenz-*in situ*-Hybridisierung (FISH). Aufgrund der einfachen Handhabbarkeit dieser Farbstoffe, verbunden mit einer hohen bildlichen Auflösung, hat sich dieses spezifische Format weitestgehend durchgesetzt. Mit der Einführung von neuen, helleren Farbstoffen, besseren Mikroskopen und optimierten Hybridisierungsprotokollen konnten ab 1996 auch erstmals oligotrophe, marine Gewässer mittels FISH untersucht werden (Glöckner *et al.*, 1996). Dies war zuvor aufgrund der geringen Größe und des geringen rRNA-Gehalts der dort vorkommenden Mikroorganismen schwierig (Hicks *et al.*, 1992).

Die FISH weist im Vergleich zu anderen molekulargenetischen Techniken für die Untersuchung von mikrobiellen Gemeinschaften einige besondere Charakteristika auf. Durch die Visualisierung einzelner intakter Zellen, lassen sich nicht nur Informationen über die Morphologie der detektierten Zellen erhalten, sondern auch zelluläre Abundanzen ausgewählter Populationen ableiten. Da der Nachweis *in situ*, also im natürlichen Mikrohabitat erfolgt, kann zudem die räumliche Verteilung der Organismen erfasst werden.

Wie alle vorgestellten molekularen Techniken, weist aber auch die FISH spezifische Beschränkungen auf. Aus verschiedenen Gründen können falsch-negative Ergebnisse bei der Analyse auftreten. Zunächst müssen die Sonden in die Zelle eindringen können. Dies zu gewährleisten, ist insbesondere für Gram-positive Bakterien schwierig (Roller *et al.*, 1994). Eine weitere Beschränkung ist durch die Zugänglichkeit einzelner Sondenbindungsstellen auf der rRNA im fixierten Ribosom selber gegeben (Fuchs *et al.*, 1998). Wie bei Methoden, die auf einer Freisetzung der nachzuweisenden Zielmoleküle beruhen, ist auch die zelluläre Kopienzahl der rRNA entscheidend für die Detektierbarkeit einzelner Populationen. Als Folge der genannten Einschränkungen, können in der Regel nicht alle Mitglieder einer

prokaryontischen Gemeinschaft durch FISH detektiert werden. Neben der Verwendung von helleren Fluoreszenzfarbstoffen und leistungsstärkeren Mikroskopen, existieren eine Reihe von weiteren Ansätzen für die Steigerung der Sensitivität. Hierzu gehören die Auswahl geeigneter Sondenbindungsstellen auf der 16S rRNA (Fuchs *et al.*, 1998; Fuchs *et al.*, 2001; Behrens *et al.*, 2003), die Verwendung mehrfachmarkierter (Wallner *et al.*, 1993) oder mehrerer einfachmarkierter Sonden (Lee *et al.*, 1993), der Einsatz von Helfer-Oligonukleotiden (Fuchs *et al.*, 2000) und die enzymatische Signalverstärkung (Amann *et al.*, 1992).

Hochrelevant für alle Hybridisierungsformate ist die Vermeidung von falsch-positiven Ergebnissen, hervorgerufen durch die unspezifische Bindung von Sonden an Nicht-Zielsequenzen, die einzelne Fehlpaarungspositionen aufweisen. Um dies zu vermeiden wird jede Sonde unter optimierten Bedingungen bezüglich Temperatur, Ionenkonzentration und denaturierenden Agenzien (z.B. Formamid) angewandt. Diese Bedingungen leiten sich aus der Dissoziationstemperatur T_d einer Sonde ab, welche durch die Nukleotidsequenz des Oligomers determiniert wird (Breslauer *et al.*, 1986), und müssen für jede neuentwickelte Sonde zunächst empirisch bestimmt werden (Amann and Schleifer, 2001). Da die Abtrennung einzelner, randständiger Fehlpaarungspositionen selbst unter optimierten Bedingungen kaum möglich ist, sollte für den eindeutigen Nachweis ausgewählter Populationen stets ein Satz phylogenetisch redundanter oder hierarchisch geordneter Sonden eingesetzt werden (Amann and Schleifer, 2001). Verstärkt wird die Notwendigkeit dieses "Mehrfachsondenkonzeptes" durch den Umstand, dass ausgewählte Sequenzsignaturen in der Regel nicht auf der 16S rRNA aller Mitglieder der zu detektierenden phylogenetischen Gruppe vorkommen. Aufgrund der hohen mikrobiellen Diversität können des Weiteren auch Organismen existieren, die nicht zu dieser Zielgruppe gehören und trotzdem eine identische Sondenbindungsstelle aufweisen.

Grundsätzlich ist auch die Detektion von funktionellen Genen auf mRNA-Ebene oder sogar auf dem bakteriellen Chromosom mittels *in situ*-Hybridisierung vorstellbar. Allerdings werden hierfür wesentlich empfindlichere Nachweisverfahren benötigt, als momentan verfügbar sind, da die "natürliche" Signalamplifikation, wie sie durch die hohe Kopienzahl der rRNA-Moleküle in jeder Zelle gegeben ist, entfällt. Vielversprechende Ansätze wurden bereits mit der Etablierung von hochsensitiven *in situ*-Hybridisierungstechniken gemacht, die auf einer enzymatischen Signalamplifikation beruhen (Pernthaler *et al.*, 2002a).

2 Ein leistungsstarkes molekularbiologisches Werkzeug: DNA-Microarrays

Der DNA-Microarray-Technologie kommt ein besonderer Stellenwert unter den molekularbiologischen Methoden zu, da ihre Einführung eine völlig neue Dimension von Experimenten ermöglichte. Wo noch wenige Jahre zuvor in mühsamer Arbeit die Funktion und die Sequenz einzelner Gene bestimmt wurden, kann nun das komplette Transkriptom eines Organismus in einem einzigen Experiment analysiert werden.

Seit dem Jahr 1998 steigt die Zahl der Studien, in denen DNA-Microarrays eingesetzt werden, exponentiell an (Abb. 2). Ein Ende dieser Entwicklung ist derzeit nicht abzusehen. Dies darf allerdings nicht darüber hinwegtäuschen, dass der derzeitige Entwicklungsstand der Methode nicht allen denkbaren Anwendungsbereichen gewachsen ist.

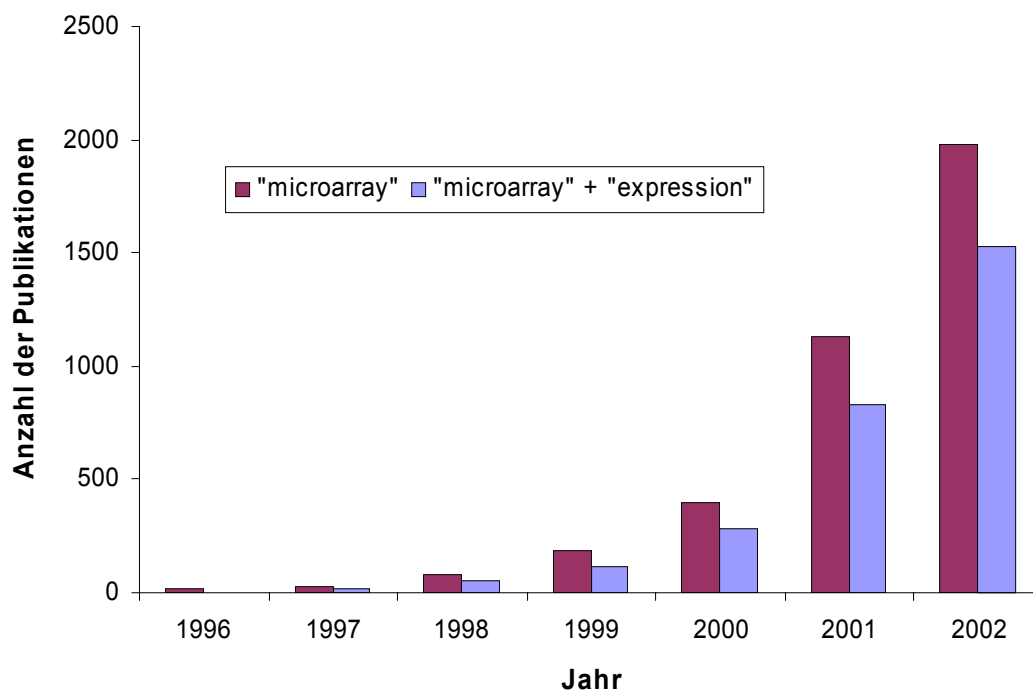


Abbildung 2: Die zunehmende Bedeutung von DNA-Microarrays. Dargestellt ist für die Jahre 1996-2002 die Anzahl der im ISI Web of Science (<http://www.isiwebofknowledge.com/>) hinterlegten Publikationen, die die angegebenen Stichwörter im Titel oder der Kurzfassung enthalten.

2.1 Ursprung, Entwicklung und gängige Anwendungen

Das Konzept der oberflächenvermittelten Nukleinsäurehybridisierung geht bis in das Jahr 1975 zurück, als Edwin Southern erstmals in einem Gel aufgetrennte DNA-Fragmente auf eine poröse Oberfläche "blottete", sie dort anschließend immobilisierte und mit einer radioaktiv markierten DNA-Sonde hybridisierte (Southern, 1975). Auf diese Weise konnten DNA-Fragmente identifiziert werden, die eine ausgewählte Zielsequenz trugen. Dieses als Southern-Hybridisierung bezeichnete Verfahren stellt noch heute eine der Standardmethoden der Molekularbiologie dar und findet in Abwandlung auch Anwendung bei dem Screening von klonierten DNA-Fragmenten ("Kolonie-Hybridisierung"; für eine Übersicht siehe Amann and Schleifer, 2001). Für diese ursprünglichen oberflächenvermittelten Hybridisierungsformate, oft auch als "Standard"-Formate bezeichnet, ist die Immobilisierung der zu analysierenden Nukleinsäuregemische charakteristisch. Die nachzuweisenden Zielmoleküle sind also unmarkiert und oberflächengebunden. In der Regel wird mit einer einzelnen markierten Nukleinsäuresonde gegen sie hybridisiert (weitere Beispiele sind die dot blot-Hybridisierung und die FISH).

Im Gegensatz hierzu, zeichnen sich die "reversen" Hybridisierungsformate durch die Analyse von markierten Zielmolekülen aus, die sich in Lösung befinden und gegen eine Reihe von unmarkierten, oberflächengebundenen Sonden hybridisiert werden. Nichtgebundene Zielmoleküle werden anschließend von der Oberfläche gewaschen. Da sich die verschiedenen "Fänger"-Moleküle an definierten Positionen der Matrix befinden, ermöglicht dieser Ansatz den hochparallelen Einsatz von Nukleinsäuresonden. Es besteht keine Notwendigkeit mehr, verschiedene Sonden mit unterschiedlichen Reportermolekülen auszustatten, um die erfolgten Hybridisierungsereignisse differenzieren zu können. Ursprünglich für die genetische Analyse von PCR-amplifizierter DNA mittels Oligonukleotidsonden eingeführt (Saiki *et al.*, 1989), fand dieses Format in der Folge eine Vielzahl von weiteren Anwendungsgebieten. Beispiele hierfür sind die Untersuchung der Genexpression von *Escherichia coli* (Chuang *et al.*, 1993) und die Identifizierung von Bakterien (Ehrmann *et al.*, 1994).

DNA-Microarrays stellen ebenfalls ein reverses Hybridisierungsformat dar, unterscheiden sich aber von den klassischen Ansätzen durch eine erheblich höhere Dichte der immobilisierten Fängersonden (Abb. 3). Heute kann eine Fläche von 1,28 x 1,28 cm ohne weiteres bis zu 300.000 verschiedene Oligonukleotidsonden tragen (Lipshutz *et al.*, 1999), womit eine massive Parallelisierung der Analyse ermöglicht wird.

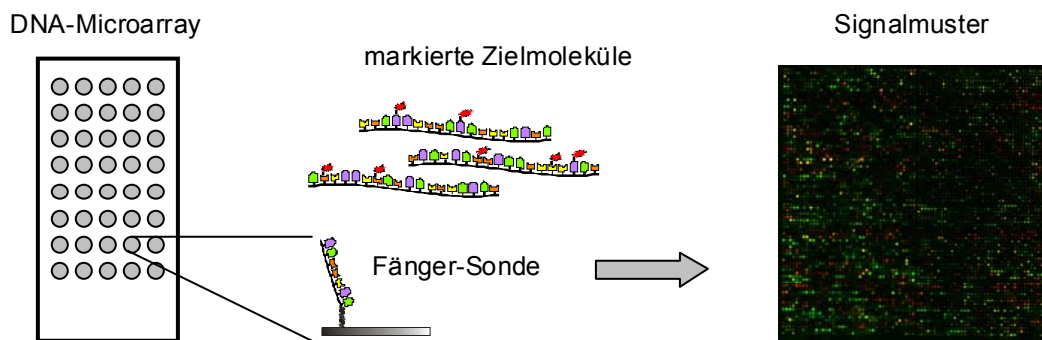


Abbildung 3: Funktionsprinzip von DNA-Microarrays. Verschiedene Fängersonden (Poly- oder Oligonukleotide) werden mit hoher Dichte an definierten Positionen auf einer Matrix immobilisiert (Links schematisch gezeigt). Die markierten Zielmoleküle befinden sich in Lösung (Mitte) und stehen somit jeder Sonde als potentieller Bindungspartner zur Verfügung. Nach erfolgter spezifischer Hybridisierung werden nichtgebundene Zielmoleküle von der Oberfläche gewaschen. Rechts ist ein typisches Bild für einen "High-Density"-Microarray gezeigt, wie es nach Hybridisierung und Detektion der Fluoreszenzsignale erhalten wird. Der Bildausschnitt entspricht einer Fläche von ca. 1,3 x 1,3 cm auf der hybridisierten Oberfläche und zeigt etwa 6400 Bindungsstellen, sogenannte "Spots".

Entscheidend für die Etablierung der DNA-Microarray-Technologie war die Einführung von nicht-porösen Oberflächen für die Immobilisierung der Fängernukleinsäuren. Dieser Schritt ermöglichte eine Miniaturisierung der reversen Hybridisierungsformate und erlaubte gleichzeitig den hierfür benötigten Einsatz von Fluoreszenzfarbstoffen für die Markierung der Zielmoleküle. Heute beschäftigen sich eine Vielzahl von Studien ausschließlich mit der Optimierung der genutzten Oberflächen (Beier and Hoheisel, 1999; Diehl *et al.*, 2001; Benters *et al.*, 2002). Gleichzeitig waren Verfahren für die Generierung von abgegrenzten Feldern (Spots) aus verschiedenen Fängernukleinsäuren im μm -Bereich unabdingbar. Einen Meilenstein stellte diesbezüglich die Einführung von Techniken für die hochaufgelöste "on-chip"-Synthese von organischen Oligomeren dar. Fodor *et al.* (1991) adaptierten hierfür photolithographische Maskierungstechniken aus der Halbleiterproduktion. Maskos und Southern (1992) bedienten sich eines Ansatzes, der auf einer hochaufgelösten Verschiebung von Reaktionskammern beruhte. Alternativ können die Fängermoleküle aber auch extern generiert und mit Hilfe von "Spotting"-Robotern auf die Oberflächen aufgebracht werden. Dies

kann kontaktfrei geschehen, durch Piezo- oder Ink-Jet-Systeme, oder mit Hilfe von Stahlnadeln, die bei Oberflächenkontakt kleinste Volumina ablegen. Auch wenn bei diesen apparativen Varianten, bedingt durch steuerungstechnische Grenzen, nur mittlere Sondendichten erzielt werden können, sind entsprechende Verfahren insbesondere in der Grundlagenforschung weit verbreitet.

Die ursprüngliche Triebkraft der Microarray-Entwicklung entstammte dem Bestreben des öffentlichen Human-Genom-Projektes, neue, leistungsfähige Techniken für die *de novo*-Sequenzierung von DNA-Fragmenten zu etablieren. Allerdings konnte sich das "sequencing by hybridization" (SBH) (Bains and Smith, 1988; Drmanac *et al.*, 1989) aufgrund unerwarteter praktischer Probleme bis heute nicht durchsetzen. Repetitive Sequenzen und unspezifische Bindungen führen zu Artefakten, welche eine eindeutige Sequenzbestimmung nicht zulassen. Insbesondere erschwert die oft unzureichende Spezifität von Microarray-Hybridisierungen auch in anderen Bereichen die potentielle Anwendung dieser Technik. Hervorgerufen wird dieses Problem durch die parallele Anwendung von Sonden mit unterschiedlichen Charakteristika, unter identischen Bedingungen bezüglich Temperatur, Ionenkonzentration und denaturierenden Agenzien. Dieser Aspekt ist seit Jahren Gegenstand intensiver Untersuchungen (Weiler *et al.*, 1997; Nguyen *et al.*, 1999; Liu *et al.*, 2001; Dai *et al.*, 2002; Religio *et al.*, 2002; Urakawa *et al.*, 2002).

Heute existieren zwei grundsätzliche Formen von DNA-Microarrays, die entweder durch die Verwendung von längeren PCR-Produkten (100 bis etwa 2.500 Nukleotide bzw. Basenpaare) als Fängernukleinsäuren charakterisiert sind oder auf dem Einsatz von kürzeren synthetischen Oligonukleotidsonden (<10 Nukleotide bis 70mere) beruhen. Beide Typen werden für die am weitesten verbreitete Anwendung von DNA-Microarrays eingesetzt, die Analyse der Genexpression von ausgewählten Organismen. In der gängigen Form dieses Ansatzes, der Identifizierung von differentiell exprimierten Genen durch eine Zweifarbenhybridisierung (Skena *et al.*, 1995), werden zunächst die mRNA-Populationen aus zwei Linien einer Zellkultur extrahiert, die zuvor unterschiedlichen Wachstumsbedingungen ausgesetzt wurden. Die Extrakte werden dann in cDNA umgeschrieben, mit unterschiedlichen Fluoreszenzfarbstoffen markiert und als Gemisch gegen ausgewählte, immobilisierte Gene (Heller *et al.*, 1997) oder gegen alle Gene des untersuchten Genomes (Sudarsanam *et al.*, 2000) hybridisiert. Auf diese Weise kann die Auf- oder Abregulation einzelner Gene zuverlässig identifiziert und einer weiteren Untersuchung zugänglich gemacht werden. Für eine detaillierte

Betrachtung soll an dieser Stelle auf ausgewählte Übersichtsartikel verwiesen werden (Duggan *et al.*, 1999; Lucchini *et al.*, 2001; Churchill, 2002; Stears *et al.*, 2003).

Der zweite große Anwendungsbereich von DNA-Microarrays ist der Nachweis von DNA-Sequenzvarianten, der sich bis in die klinische Diagnostik hinein erstreckt. Als Fängersonden werden kurze Oligomere mit einer Länge von unter 30 Nukleotiden eingesetzt. Hierdurch wird die Spezifität der Hybridisierung erhöht, da einzelne Basen-Fehlpaarungspositionen das Gleichgewicht der Hybridisierung stärker beeinflussen als dies für längere Fängersonden der Fall ist. Durch das Aufbringen von zusätzlichen Kontrollsonden, die einzelne Fehlpaarungspositionen gegenüber den nachzuweisenden Zielmolekülen aufweisen, lassen sich charakteristische Signalmuster für einzelne genetische Varianten erzeugen, auch wenn nicht jede vorliegende Fehlpaarung vollständig diskriminiert werden kann. Eine gängige Anwendung ist der Nachweis von Punktmutationen (single nucleotide polymorphisms, SNP's), der beispielsweise für die genetische Diagnostik (Guo *et al.*, 1994; Hacia *et al.*, 1996), populationsgenetische Untersuchungen (Chakravarti, 1999) oder die pharmazeutische Industrie (Debouck and Goodfellow, 1999) von hoher Relevanz ist.

Neben den klassischen Microarray-Formaten für die Analyse von Nukleinsäuren, gewinnen zunehmend sogenannte Protein-Microarrays an Bedeutung, bei denen Protein-Protein- oder Protein-Nichtprotein-Interaktionen auf einer Oberfläche nachgewiesen werden (MacBeath, 2002). Bis zu 5.800 verschiedene Genprodukte konnten auf diese Weise bereits parallel auf ihre Fähigkeit, einen ausgewählten Liganden zu binden, getestet werden (Zhu *et al.*, 2001).

2.2 Nachweis von natürlichen mikrobiellen Populationen und hieraus erwachsende Anforderungen

Die Standard-Hybridisierungstechniken stellen einen zentralen Aspekt des rRNA-Ansatzes in der mikrobiellen Ökologie dar. Trotz ihrer Leistungsfähigkeit, ist eine deutliche Einschränkung in der Anwendbarkeit dieser Methoden durch die limitierte Anzahl von Sonden gegeben, die in einem einzelnen Experiment angewandt werden können. Diese Limitation kommt insbesondere dann zum Tragen, wenn diese Techniken für eine umfangreiche Charakterisierung von komplexen Probe herangezogen werden sollen. Durch den Einsatz von reversen Hybridisierungsformaten könnten diese Einschränkungen aufgehoben und damit eine Vielzahl von Varianten eines genetischen Markers, beispielsweise der 16S rRNA, oder sogar

verschiedener genetischer Marker parallel auf Anwesenheit in einer Probe geprüft werden. Gleichzeitig vereinfacht sich die Anwendung von komplexen Sätzen aus phylogenetisch redundanten und hierarchisch geordneten Sonden, die für einen eindeutigen Nachweis von Zielsequenzen in komplexen Proben notwendig sind. Auf diese Weise können falsch-positive und falsch-negative Signale nach dem Ausschlussprinzip als solche erkannt werden. Im Idealfall wird die Gesamt-RNA aus einer Probe extrahiert, chemisch markiert und gegen einen Satz von Oligonukleotidsonden, der alle in der Probe zu erwartenden Taxa abdeckt, hybridisiert. Eine auf diese Weise vorcharakterisierte Probe könnte durch weitere Techniken, wie beispielsweise die FISH, gezielt analysiert werden. Dieses Vorgehen würde die molekulargenetische Analyse von komplexen Proben unbekannter Zusammensetzung erheblich beschleunigen. In der Praxis erschweren aber eine Reihe von Anforderungen, die sich insbesondere aus der Analyse von natürlichen mikrobiellen Populationen ergeben, den oben skizzierten Einsatz von DNA-Microarrays.

Eine grundsätzliche methodische Herausforderung bei der hochparallelen Anwendung von Oligonukleotidsonden im Microarray-Format, stellt die Gewährleistung einer ausreichenden Spezifität der Hybridisierung dar. Wie in Abschnitt 1.2 beschrieben, wird in den konventionellen Hybridisierungsformaten jede Sonde unter optimierten Bedingungen eingesetzt, um die Wahrscheinlichkeit für das Auftreten von unspezifischen Bindungsereignissen zu minimieren. Im Microarray-Format muss aufgrund der massiven Miniaturisierung eine Vielzahl von Sonden unterschiedlicher Charakteristika unter einheitlichen Bedingungen angewandt werden. Somit ist eine umfassende Spezifität der Hybridisierung nicht ohne weiteres zu gewährleisten. Als Beispiel soll der Datensatz von Loy *et al.* (2002) angeführt werden, der eine signifikante Anzahl von falsch-positiven Hybridisierungsereignissen beinhaltet, die auf zwei oder mehr Fehlpaarungspositionen beruhen. In den gängigen Anwendungsbereichen der DNA-Microarray-Technologie spielt diese Einschränkung nur eine untergeordnete Rolle, da hier entweder sehr "unähnliche" Zielmoleküle (differentielle Genexpression) oder eine begrenzte Anzahl bekannter Varianten (z.B. Allel-Detektion) untersucht werden. Sollen aber einzelne Zielmoleküle gegen einen sehr vielfältigen und in Teilen unbekanntem genetischen Hintergrund abgetrennt werden, wie es bei der spezifischen Detektion von 16S rRNA-Molekülen in komplexen Gemeinschaften der Fall ist, werden Microarray-Formate benötigt, welche die Diskriminierung einzelner Fehlpaarungspositionen erlauben (Liu *et al.*, 2001; Urakawa *et al.*, 2002; Urakawa *et al.*, 2003). Vereinzelt auftretende falsch-positive oder falsch-negative Ereignisse können dann mit

Hilfe einer Software-gestützten Auswertung der Signalmuster als solche erkannt werden (Loy *et al.*, 2002). Ein naheliegender Lösungsansatz, die Schmelzpunktangleichung der eingesetzten Sonden durch Sequenzmodifikation, ist nur sehr eingeschränkt praktikabel, da letztere oft zu einer veränderten Spezifität der Sonden führt. Alternativ vorgeschlagene Verfahren, die auf dem Einsatz von Additiven zur Schmelzpunktangleichung beruhen (Maskos and Southern, 1993; Rees *et al.*, 1993), konnten sich aufgrund eines ungenügenden Wirkungsgrades nicht durchsetzen. Aussichtsreiche Ansätze sind durch die Aufnahme von Schmelzkurven oder Hybridisierungskinetiken gegeben (Liu *et al.*, 2001; Dai *et al.*, 2002), erfordern aber einen hohen experimentellen und apparativen, und damit auch finanziellen Aufwand.

Eine weitere Einschränkung für den Nachweis von Bakterien in der Umwelt mittels Microarray-Hybridisierung ist oftmals durch eine unzureichende Sensitivität der Analyse gegeben. Bedingt durch niedrige Extraktionseffizienzen liegt sie insbesondere dann vor, wenn auf eine vorherige enzymatische Amplifizierung der Zielmoleküle verzichtet wird (Koizumi *et al.*, 2002). Eben diese soll aber vermieden werden, da ihre Selektivität, wie erläutert, zu Populationsverschiebungen führen kann. Aus diesem Grund werden Protokolle benötigt, die eine effiziente Extraktion, chemische Markierung und Hybridisierung der Zielmoleküle aus komplexen Umweltproben ermöglichen. In der Folge kann dann systematisch untersucht werden, ob DNA-Microarrays einen quantitativen oder zumindest semi-quantitativen Nachweis von 16S rRNA-Molekülen aus der Umwelt erlauben.

Bis zum jetzigen Zeitpunkt, wurde nur eine kleine Anzahl von Studien veröffentlicht, in denen die Zusammensetzung natürlicher mikrobieller Gemeinschaften mit Hilfe von DNA-Microarrays untersucht wurde (siehe Tab. 1). In der Mehrzahl handelt es sich hierbei um einen grundsätzlichen Eignungsnachweis. Lediglich die Arbeiten von Wilson *et al.* (2002) und Loy *et al.* (2002) können als Anwendung der Microarray-Technologie auf komplexe Bakteriengemeinschaften betrachtet werden, da nur hier die Anzahl der eingesetzten Sonden das Ausmaß gegenwärtiger Standard-Hybridisierungsverfahren überschreitet. Beide Untersuchungen beruhen allerdings auf einer zuvorigen Amplifikation der nachzuweisenden Zielmoleküle. Mit einer Ausnahme (Koizumi *et al.*, 2002) wurden bisher keine Arbeiten veröffentlicht, in denen Microarray-Daten mit Ergebnissen aus einer parallel angewandten, unabhängigen und quantitativen Methode verglichen wurden. Dies ist für eine zuverlässige Bewertung der Spezifität und Sensivität von "Umwelt-Microarrays" aber zwingend notwendig.

Tabelle 1: Bisherige Anwendungen von DNA-Microarrays für die Untersuchung von komplexen mikrobiellen Gemeinschaften

Publikation	untersuchte Probe	hybridisiertes Zielmolekül	Anzahl der Sonden	Validierung der Microarray-Daten	Einschränkungen
Rudi et al., 2000	8 versch. Seen	PCR-Amplicon der 16S rDNA (Länge?)	10	Mikroskopie	PCR-bias, kleiner Sondensatz
Small et al., 2001	Boden (Weizenfeld)	extrahierte Gesamt-RNA, angereichert mit Gesamt-RNA aus Reinkulturen (nur diese konnten nachgewiesen werden)	4 versch. Bindungsstellen	keine	sehr kleiner Sondensatz
Wilson et al., 2002	Freiluft	PCR-Amplicon der 16S rDNA (100 bp Fragment)	31179	Sequenzierung von 16S rDNA-Klonen	PCR-bias
Koizumi et al., 2002	Öl-kontaminiertes Sediment (Kuwait)	in vitro transkribierte RNA	8 versch. Bindungsstellen	DGGE, quantitative Membran-Hybridisierung	PCR-bias, sehr kleiner Sondensatz
Loy et al., 2002	Solar Lake Matte, Backentasche (Mensch)	PCR-Amplicon der 16S rDNA	132	Sequenzierung von 16S rDNA-Klonen	PCR-bias

3 Zielsetzung dieser Arbeit

Obwohl die grundsätzliche Eignung der DNA-Microarray-Technologie für die Differenzierung und Identifizierung von Bakterien bereits 1997 gezeigt werden konnte (Guschin *et al.*, 1997) und nachfolgende Arbeiten auch die Analyse von Umweltproben einbezogen haben (vgl. Tabelle 1), ist dieses Hybridisierungsformat für die Charakterisierung von komplexen mikrobiellen Gemeinschaften als noch nicht etabliert zu bewerten. Die Gründe sind in einer Reihe von methodischen Defiziten zu finden, die insbesondere bei der Differenzierung von natürlichen prokaryontischen Populationen zum Tragen kommen. Hierzu gehören aufwendige Protokolle für die Aufbereitung der zu hybridisierenden Zielmoleküle, unzureichende Detektionsgrenzen und ein hoher apparativer und zeitlicher Aufwand, um die Spezifität der erhaltenen Signale zu gewährleisten. Derzeit wiegen die genannten Limitationen in ihrer Summe schwerer als die Vorteile, die sich aus der Anwendung von DNA-Microarrays für den Nachweis von 16S rRNA-Sequenzen in der Umwelt ergeben.

Ein Ziel dieser Arbeit war es, ein Protokoll zu etablieren, das die Anwendbarkeit von DNA-Microarrays für die Untersuchung von Diversität und Struktur natürlicher prokaryontischer Gemeinschaften erhöht und somit die Basis für eine breitere Nutzung dieser Technik in der mikrobiellen Ökologie schafft. Hierfür wurden zunächst systematische Untersuchungen an definierten Modellsystemen aus marinen Isolaten und einer Reihe von

Sonden durchgeführt, die bereits für den Nachweis von bakteriellen Populationen genutzt wurden. Aus den gewonnenen Erkenntnissen sollten neue Optimierungsstrategien abgeleitet werden, um eine hochspezifische Hybridisierung der eingesetzten Oligonukleotidsonden auch ohne aufwendige Meßtechniken zu gewährleisten. In der Folge wurde die Komplexität der untersuchten Proben sukzessive erhöht. Ein Schwerpunkt lag hierbei auf der Entwicklung einer Markierungsmethode, die auf einer direkten, chemischen Fluoreszenzmarkierung von extrahierter Gesamt-RNA aus den Proben beruht. Auf diese Weise sollten Populationsverschiebungen, bedingt durch enzymatische Markierungsansätze, vermieden und damit die Genauigkeit der Analyse erhöht werden. Als Modellsystem dienten Mischkulturen der zuvor untersuchten Isolate, sowie marine Picoplankton-Gemeinschaften.

Die vorliegende Arbeit wurde im Rahmen der Arbeiten des Centrums für Angewandte Gensensorik (CAG) der Universität Bremen durchgeführt. In diesem Teilprojekt sollte unter anderem ein Validierungsverfahren für die DNA-Microarray-Analyse von natürlichen mikrobiellen Populationen ausgearbeitet werden. Da sich die grundlegenden Mechanismen aller Microarray-Hybridisierungen gleichen, sollten die erhaltenen Ergebnisse auch relevant für andere Anwendungsgebiete sein. Ein weiterer Schwerpunkt dieser Arbeit lag daher auf dem Abgleich der Microarray-Daten mit den Ergebnissen aus einer parallel angewandten, unabhängigen Methode. Als geeignetes Verfahren bot sich die Fluoreszenz-*in situ*-Hybridisierung an. Dieses quantitative Hybridisierungsformat basiert nicht auf einem PCR-Schritt, ist nicht oberflächengebunden und wird seit Jahren in der mikrobiellen Umweltforschung eingesetzt (Amann and Schleifer, 2001). Im Rahmen dieser Untersuchungen sollte auch die Nachweisgrenze des entwickelten Microarray-Protokolls für bakterielle Populationen in den Umweltproben bestimmt, sowie die Eignung von DNA-Microarrays für einen quantitativen/semi-quantitativen Nachweis dieser Populationen abgeschätzt werden.

Ihre volle Leistungsfähigkeit entfaltet die DNA-Microarray-Technologie erst bei der Nutzung von Sondensätzen, die mehrere Dutzend verschiedene Fängernukleinsäuren umfassen. Andernfalls bieten sich eher membrangebundene Formate an, da diese "Macroarrays" keinen gesonderten apparativen Aufwand für das Aufbringen der Fängermoleküle benötigen. Auf Basis der 16S rRNA lassen sich Oligonukleotidsonden für eine Vielzahl von Taxa auf verschiedenen phylogentischen Ebenen konstruieren (Amann *et al.*, 1995). Allerdings ist die Entwicklung von phylogenetisch redundanten Sonden durch den limitierten Informationsgehalt der 16S rRNA oftmals beschränkt. Auch kann nicht jedes beliebige Verwandtschaftsverhältnis mit Hilfe dieses Markers differenziert werden. Als mögliche Alternative bietet sich die fast

doppelt so lange 23S rRNA an. Ein prominentes Beispiel ist die Differenzierung der beta- und gamma-Proteobakterien anhand kurzer diagnostischer Sequenzabschnitte, wie sie nur auf der 23S rRNA möglich ist (Manz *et al.*, 1992).

In einem weiteren Teil dieser Arbeit wurden die nahezu vollständigen 23S rRNA-Gene einer Reihe von aus der Nordsee isolierten alpha-Proteobakterien sequenziert, darunter auch Vertreter der marinen alpha-Gruppe (Gonzalez *et al.*, 1996). Bisher liegen keine 23S rRNA-Sequenzen aus dieser Gruppe in den öffentlichen Datenbanken vor, obwohl ihren Vertretern an verschiedenen Standorten hohe *in situ*-Abundanzen im marinen Picoplankton zugeordnet werden konnten (Gonzalez and Moran, 1997; Eilers *et al.*, 2001). Die erhaltenen Sequenzen sollten phylogenetisch analysiert und auf die Möglichkeit der Entwicklung neuer Sonden getestet werden.

B

Ergebnisse und Diskussion

Im Folgenden werden die Ergebnisse der im zweiten Teil dieser Arbeit aufgeführten Publikationen zusammengefasst und übergreifend diskutiert. Dieser Abschnitt kann nicht die detaillierten Diskussionen der einzelnen Publikationen ersetzen, auf die hiermit verwiesen sei. Ausführlicher werden dafür Aspekte behandelt, die in den Publikationen nicht oder nur kurz angesprochen werden.

B Ergebnisse und Diskussion

1 DNA-Microarrays für den hochparallelen Nachweis von 16S rRNA-Zielsequenzen in Proben unterschiedlicher Komplexität

Der Einsatz molekularbiologischer Techniken in der mikrobiellen Umweltforschung hat nicht nur die unerwartet hohe Komplexität von natürlichen prokaryontischen Lebensgemeinschaften aufgezeigt, sondern weist auch zunehmend auf eine ausgeprägte Dynamik in deren räumlicher und zeitlicher Zusammensetzung hin. Um diese in Zukunft detailliert analysieren zu können, werden neue und leistungsstarke molekularbiologische Werkzeuge benötigt. Ein geeignetes Format für den hochparallelen Nachweis von Mikroorganismen stellt die DNA-Microarray-Methodik dar, deren Anwendbarkeit auf Umweltgemeinschaften durch verschiedene methodische Defizite bisher eingeschränkt wurde. Letztere traten speziell bei der Untersuchung komplexer mikrobieller Gemeinschaften in nährstoffarmen Systemen auf.

1.1 Identifizierung von bakteriellen Reinkulturen und systematische Untersuchungen an definierten Modellsystemen

Neben der Untersuchung von komplexen Gemeinschaften ist ein weiteres mögliches Anwendungsgebiet von DNA-Microarrays in der Mikrobiologie durch die Identifizierung von bakteriellen Reinkulturen gegeben, welche aus Umweltstudien oder in der klinischen Diagnostik in großer Zahl gewonnen werden. Durch die PCR-Amplifikation von 16S rRNA-Genen unter der Verwendung von fluoreszenzmarkierten Primern und die anschließende Hybridisierung der Produkte gegen geeignete Sondenätze, sollten diese einfach und schnell phylogenetisch zugeordnet werden können. Damit bietet die Microarray-Methodik eine Alternative zu der insbesondere in der Umweltmikrobiologie herangezogenen Sequenzanalyse von 16S rRNA-Genen. Allerdings lässt sich aus dem Nachweis einzelner Sequenzsignaturen auf einem genetischen Marker keine hochaufgelöste phylogenetische Einordnung vollziehen, wie sie durch die vergleichende Sequenzanalyse ermöglicht wird. Für einen zuverlässigen Schnellnachweis von pathogenen Mikroorganismen ist dies aber auch nicht erforderlich.

Des Weiteren ermöglicht die Hybridisierung einzelner PCR-amplifizierter 16S rRNA-Gene aus bakteriellen Reinkulturen, aufgrund der minimalen Komplexität der untersuchten

Probe, die Durchführung von Microarray-Experimenten unter vollständig definierten Bedingungen. Hierdurch können ausgewählte, die Reaktion beeinflussende Parameter gezielt analysiert werden. Auch können nach dem Baukastenprinzip Modellsysteme erstellt werden, die sich aus ausgewählten Reinkulturen zusammensetzen, welche in hoher Zahl und Variation in den Stammsammlungen verfügbar sind, und aus bereits publizierten Oligonukleotidsonden, die für den spezifischen Nachweis einer Vielzahl von Arten, Gattungen oder Gruppen anhand kurzer 16S rRNA-Sequenzsignaturen entwickelt wurden. Auf diesen Ansatz wurde in der vorliegenden Arbeit für die initiale Charakterisierung des zugrundeliegenden Systems zurückgegriffen (Publikation 1). Anhand der erhaltenen Ergebnisse wurden Ansatzpunkte für die Entwicklung neuer Optimierungsstrategien abgeleitet, die eine ausreichende Spezifität der Hybridisierung bei nur einer einzelnen durchgeführten Messung gewährleisten. Verbleibende falsch-positive und falsch-negative Ereignisse können dann mittels komplexer SONDENSÄTZE nach dem Ausschlussprinzip als solche erkannt werden.

1.1.1 Parameteranalyse

Für die Microarray-Analyse von bakteriellen Reinkulturen wurde ein Modellsystem aus sechs aus dem Oberflächenwasser der Nordsee isolierten Stämmen und zwanzig Oligonukleotidsonden etabliert (Publikation 1, Tab. 1). Letztere erlauben bei einer hinreichenden Spezifität der Hybridisierung die eindeutige Differenzierung der ausgewählten Organismen. Vier der sechs Isolate gehören zur Gruppe der alpha-Proteobakterien und jeweils ein Isolat ist der gamma-Gruppe der *Proteobacteria* und der *Cytophaga-Flavobacterium*-Gruppe der *Bacteroidetes* zuzuordnen. Bei der Auswahl der Sonden wurde weder deren G+C-Gehalt noch deren Länge berücksichtigt, so dass ersterer zwischen 40 und 78% und letztere zwischen 15 und 20 Nukleotiden variierte. Erste systematische Untersuchungen wurden an Stamm KT0202a (alpha-Proteobakterien) durchgeführt, da dessen 16S rRNA zu den meisten Sonden (11) in diesem Modellsystem komplementär ist. Ein Teil der analysierten Parameter wurde bereits in früheren Studien eingehend untersucht. Hierzu gehören die Sekundärstruktur der Zielmoleküle (z.B. Mir and Southern, 1999) und die oberflächenvermittelte sterische Hinderung (z.B. Shchepinov *et al.*, 1997). Allerdings ist bisher nicht bekannt, inwiefern sich diese Faktoren auf die Spezifität einer reversen Hybridisierung auswirken. Aus diesem Grund wurde der angewandte SONDENSATZ durch hochauflösende Kontrollen ergänzt. Für jede der

zwanzig Sonden wurde ein zusätzliches Kontrolloligonukleotid mitgeführt, das eine mittelständige Fehlpaarungsposition zu der entsprechenden Sondenzielsequenz aufwies.

In einer ersten Versuchsreihe wurde ein fast vollständiger, einfach-fluoreszenzmarkierter DNA-Einzelstrang, der die entsprechende Sequenz der exprimierten 16S rRNA des Referenstammes aufwies, gegen den kompletten SONDENSATZ unter zunehmender Stringenz hybridisiert (Publikation 1, Abb. 1). Alle aufgebrauchten Sonden trugen einen 5'-Aminolinker, der mit aktivierten Gruppen an der Microarray-Oberfläche reagiert (Benters *et al.*, 2002). Somit ist von einer überwiegend definierten Orientierung der immobilisierten Fängeroligonukleotide auszugehen. Da aus verschiedenen Studien ein negativer Einfluss der Microarray-Oberfläche auf die Hybridisierungskinetik abzuleiten ist (Guo *et al.*, 1994; Dai *et al.*, 2002), wurde die Inkubationszeit im Vergleich zu einer Standard-*in situ*-Hybridisierung (Pernthaler *et al.*, 2001) von 1,5 h auf 3 h erhöht. Die besten Hybridisierungsergebnisse bezüglich der Anzahl von falsch-positiven Signalen (0) und falsch-negativen Ereignissen (2) wurden unter moderaten Bedingungen erhalten. Diese Beobachtung deckt sich gut mit denen aus FISH-Experimenten (Fuchs *et al.*, 1998). Gleichzeitig konnte die Übertragbarkeit eines Standard-FISH-Protokolls (Pernthaler *et al.*, 2001) auf das genutzte Microarray-Format demonstriert werden. Diese ist von Bedeutung für eine spätere FISH-Validierung der Microarray-Daten, da hierdurch die Vergleichbarkeit der Ergebnisse erhöht wird.

Verschiedene Studien haben gezeigt, dass die Sekundärstrukturen der nachzuweisenden Zielmoleküle massiv limitierend auf die Effizienz einer Hybridisierung wirken können (Mir and Southern, 1999; Nguyen and Southern, 2000). Auch selbstkomplementäre Sequenzbereiche der Fängernukleinsäuren können zu entsprechenden Effekten führen. In der vorliegenden Arbeit sollte der Einfluss der Sekundärstruktur eines vollständigen, einzelsträngigen 16S rDNA-Zielmoleküls auf die Hybridisierung durch den Einsatz von sogenannten "Helfer"-Oligonukleotiden abgeschätzt werden. Hierbei handelt es sich um kurze, unmarkierte DNA-Stränge, die gegen die flankierenden Sequenzbereiche einer Sondenbindungsstelle gerichtet sind, um deren Zugänglichkeit zu erhöhen (Niemeyer *et al.*, 1998; O'Meara *et al.*, 1998; Fuchs *et al.*, 2000). Durch diesen Ansatz konnten die Signalintensitäten ausgewählter Sonden selektiv gesteigert werden (Publikation 1, Abb. 3A).¹ Gleichzeitig ließen sich unter sonst identischen Bedingungen zusätzliche Signale unter den Kontrollen der entsprechenden Sonden beobachten (siehe auch Publikation 1, Abb. 3B). Die bereitgestellte Stringenz der Hybridisierung (46°C, keine denaturierenden Additive) liegt also

¹ Für dieses Experiment ist exemplarisch das Bild eines ausgelesenen Microarrays in Abb. 1 dargestellt (S. 28).

nicht über dem Schmelzpunkt der zuvor diskriminierten Kontrolloligonukleotide. Des Weiteren zeigten sich auch nach dem Einsatz nur einzelner Helfer-Oligonukleotide deutliche Verschiebungen in den Intensitäten und Verhältnissen aller gemessenen Signale. Offenbar führt das partielle Aufbrechen von Sekundärstrukturen zu deren Reorganisation im Gesamtmolekül, wodurch die Zugänglichkeit zuvor "offener" Bindungsstellen herabgesetzt werden kann. Es muss daher im Umkehrschluss angenommen werden, dass sich die variierenden 16S rRNA-Primärstrukturen einer phylogenetisch breiten Zielgruppe in unterschiedlicher Weise auf die Hybridisierungseffizienz einer Sonde auswirken können. Aus dem vorliegenden Datensatz (Publikation 1, Abb. 2) sollten aufgrund der geringen Zahl an analysierten Reinkulturen diesbezüglich aber keine Rückschlüsse gezogen werden.

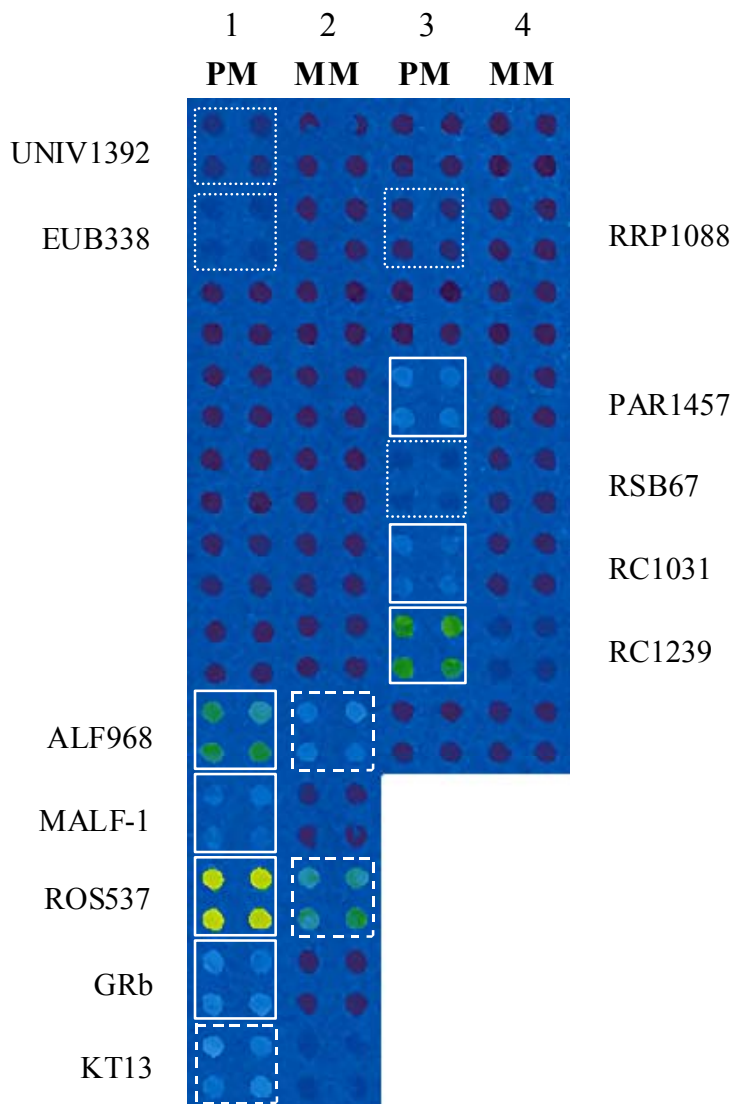


Abbildung 1:

Exemplarische Darstellung eines ausgelesenen Microarrays in Falschfarben. Es wurde die amplifizierte und Cy3-markierte 16S rDNA aus Stamm KT0202a als Einzelstrang unter Anwesenheit von Helfer-Oligonukleotiden für die Sonden ALF968, ROS537 und RC1239 gegen den kompletten Sondensatz hybridisiert. Alle Oligonukleotide wurden als Blöcke aus vier Replikaten aufgebracht. In den Spalten 1 + 3 befinden sich die Sonden (PM, "perfect match") und in den Spalten 2 + 4 die jeweiligen Kontrollen (MM, "mismatch"). Markierte Blöcke sind korrekt-positive Signale (durchgezogene Linie), falsch-positive Signale (gestrichelte Linie), oder falsch-negative Ereignisse (gepunktete Linie).

Ein weiterer Parameter, der die Effizienz einer Microarray-Hybridisierung herabsetzen kann, ist die oberflächenvermittelte sterische Hinderung (Southern *et al.*, 1999). Um deren Einfluss zu mindern, können Spacer-Moleküle zwischen der aktivierten Microarray-Oberfläche und dem Fängeroligonukleotid eingeführt werden. Hierbei handelt es sich idealerweise um ungeladene, auf Kohlenwasserstoffketten basierende Verbindungen (Shchepinov *et al.*, 1997). Oftmals wird aus Gründen der Verfügbarkeit und Praktikabilität auch auf Oligomere aus identischen Nukleotiden zurückgegriffen (Guo *et al.*, 1994; Loy *et al.*, 2002). In der vorliegenden Arbeit wurden Polyadenosinmonophosphat-Spacer variierender Länge (6, 12, 18 und 24mere) eingesetzt, um den Einfluss der oberflächenvermittelten sterischen Hinderung auf die Hybridisierungseffizienz und die Spezifität der Reaktion zu untersuchen. Da sich die Dichte der immobilisierten Fängernukleinsäuren ebenfalls auf die sterische Situation auswirkt, (Peterson *et al.*, 2001) wurde im Vorfeld eine geeignete Konzentration der aufzubringenden Sonden empirisch ermittelt (Daten nicht gezeigt). Es zeigte sich eine lineare Abhängigkeit von Signalintensität zu Spacer-Länge, wobei eine Sättigung des Signals für verschiedene Sonden mit unterschiedlichen Spacer-Längen erreicht wurde (Publikation 1, Abb. 4). Erneut ließen sich durch die Eindämmung eines signallimitierenden Parameters zusätzliche Signale unter den Kontrollen der untersuchten Sonden beobachten.

Das Ausmaß der sterischen Hinderung sollte nicht nur von der Länge einer nachzuweisenden Nukleinsäure und der Distanz zwischen Oberfläche und Sonde abhängig sein, sondern auch von der Lage der Sondenbindungsstelle auf dem Zielmolekül. Letzteres bindet in einer definierten Orientierung an die revers-komplementäre Sondensequenz, womit in Abhängigkeit der Position der Sondenbindungsstelle unterschiedlich lange Reste gegen die Oberfläche gerichtet sind (Abb. 2). Dieser Effekt wurde bisher kaum in der Literatur berücksichtigt, zumal meist fragmentierte Zielmoleküle in die Hybridisierung eingesetzt werden, um den Einfluss von Sekundärstrukturen zu minimieren. Lediglich Mir and Southern (1999) haben den Einfluss von 5'- und 3'-immobilisierten Sonden auf deren Fähigkeit, eine tRNA "einzufangen", vergleichend untersucht. Ausgehend von einer 5'-Immobilisierung, sollte die gezielte Invertierung von Fängersonden zu Signalsteigerungen führen, insbesondere wenn sich deren Bindungsstellen nahe am 5'-Ende des Zielmoleküls befinden. Diese Annahme konnte in der vorliegenden Arbeit ansatzweise für eine ausgewählte 16S rRNA bestätigt werden (Publikation 1, Abb. 4B). Allerdings lieferten nicht alle diesbezüglich durchgeführten Untersuchungen die erwarteten Ergebnisse (Daten nicht gezeigt). Offenbar verhindern die Sekundär- und Tertiärstrukturen gebundener, einzelsträngiger 16S rDNA-Zielmoleküle eine einfache Extrapolation.

Insgesamt weisen die vorliegenden Daten auf eine komplexe Interaktion der untersuchten Parameter hin. Es konnte gezeigt werden, dass die Spezifität und Sensitivität der Hybridisierung eng miteinander verknüpft sind, und dass insbesondere die Sekundärstrukturen von vollständigen, einzelsträngigen 16S rDNA-Molekülen massiv limitierend auf die Effizienz der Hybridisierung wirken. Aus den erfolgten Beobachtungen konnten gezielte Optimierungstrategien abgeleitet werden, die im folgenden Abschnitt aufgezeigt und erörtert werden sollen.

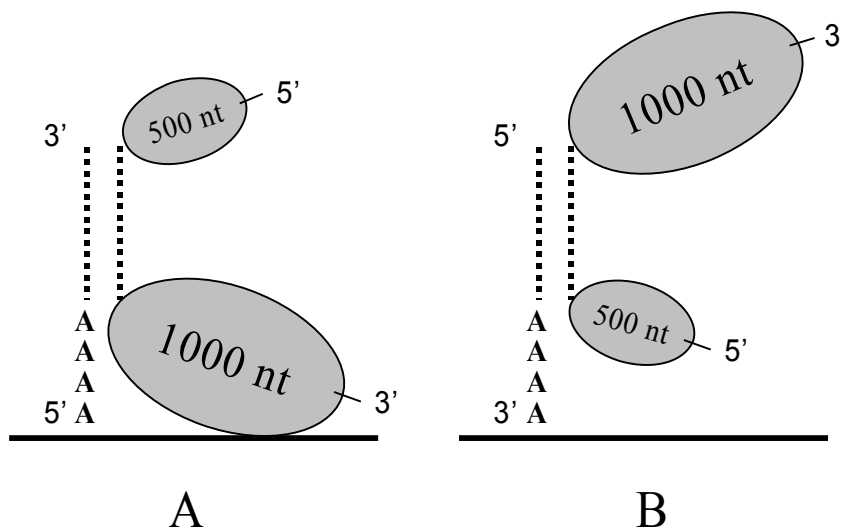


Abbildung 2: Vereinfachte Darstellung der sterischen Situation an der Microarray-Oberfläche nach Hybridisierung einer 16S rRNA gegen eine Sonde, deren Bindungsstelle näher am 5'-Ende des Zielmoleküls lokalisiert ist. (A) Sonde ist 5'-immobilisiert, (B) Sonde ist 3'-immobilisiert. Die Lage der Sondenbindungsstelle auf dem Zielmolekül bestimmt, neben dessen Sekundär- und Tertiärstruktur, das Ausmaß der sterischen Hinderung.

1.1.2 Optimierungsstrategien

In einer Versuchsreihe wurden die vollständigen 16S rRNA-Gene aller sechs Referenzstämme in unabhängigen Experimenten gegen den kompletten Sondensatz unter Standardbedingungen (46°C; keine Additive; alle Sonden gleichförmig und ohne Spacer aufgebracht) hybridisiert (Publikation 1, Abb. 2). In dem resultierenden Datensatz konnte unter 199 potenziellen unspezifischen Hybridisierungsereignissen nur ein einziges falsch-positives Signal (0,5%) detektiert und somit alle sechs Stämme eindeutig differenziert werden.

Wie konnte dieses hohe Maß an Spezifität gewährleistet werden? Wie in Abb. 3 schematisch dargestellt, verursacht die Hybridisierung unterschiedlicher Sonden unter einheitlichen Bedingungen variierende Signalintensitäten (Guschin *et al.*, 1997; Southern *et al.*, 1999). Die verminderte Affinität eines Kontrolloligonukleotids, das eine einzelne Fehlpaarungsposition zu einer Zielsequenz aufweist, sollte darüber hinaus stets zu einer niedrigeren Signalintensität führen als die der entsprechend vollständig komplementären Sonde (Publikation 1, Abb. 3B). Es ist aber nicht selten, dass das Kontrollsignal einer Sonde stärker ist als das Signal einer anderen, vollständig komplementären Sonde (Publikation 1, Abb. 1-3). Diese Beobachtungen können in ihrer Summe unter nicht optimierten Bedingungen zu einer deutlichen Verzerrung des detektierten Signalmusters führen und somit eine sinnvolle Datenauswertung verhindern.

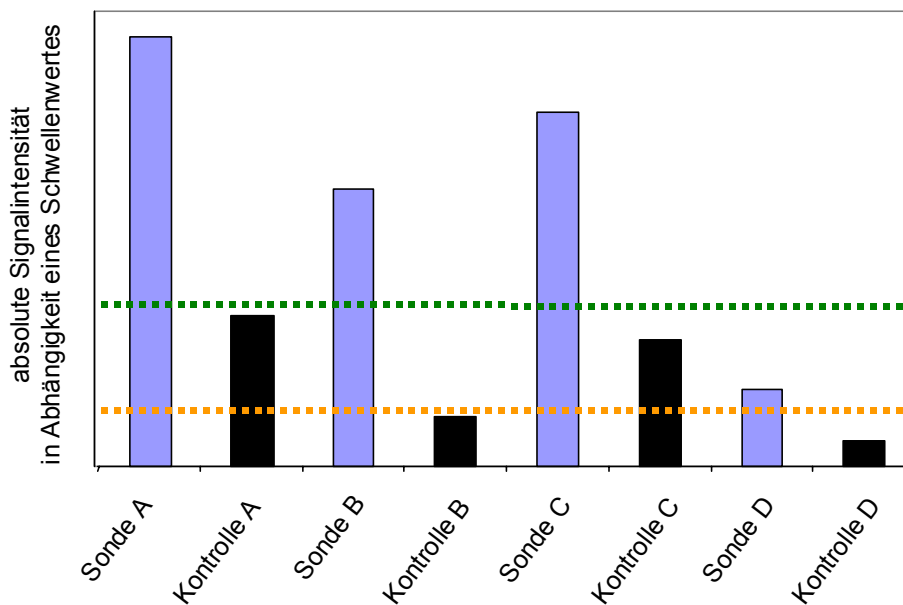


Abbildung 3: Schematische Darstellung der Konsequenzen, die sich aus einer parallelen Anwendung von Sonden mit unterschiedlichen Charakteristiken ergeben. Die Sonden A-D sind gegen verschiedene Sequenzabschnitte eines Zielmoleküls gerichtet (phylogenetisch redundante Sonden). Durch die Auswahl geeigneter Bedingungen (grüne Linie), kann die Hybridisierungseffizienz aller Sonden soweit herabgesetzt werden, dass die schwächeren falsch-positiven "Signale" nicht mehr zu detektieren sind. Dies führt zu einem gehäuften Auftreten von falsch-negativen Ereignissen (hier eins, Sonde D). Vermindert man gezielt den Einfluss eines die Hybridisierungseffizienz limitierenden Parameters (beispielsweise durch den Einsatz von Spacer-Molekülen für alle vier Sonden), werden zwar die Signalintensitäten aller Sonden gesteigert (orange Linie), doch können zusätzliche falsch-positive Signale auftreten (hier zwei, Kontrollen A und C), da die entsprechenden Bindungsereignisse nun über der Nachweisgrenze liegen.

Um dem Ausfall einzelner Fängeroligonukleotide vorzubeugen, der durch Hybridisierungstemperaturen verursacht wird, die über dem Schmelzpunkt der Sonden liegen, werden größere Sondensätze oft unter wenig stringenten Bedingungen angewandt (Guo *et al.*, 1994; Guschin *et al.*, 1997; Mir and Southern, 1999; Rudi *et al.*, 2000; Loy *et al.*, 2002; Wilson *et al.*, 2002). Gleichzeitig werden die Zielmoleküle in der Regel fragmentiert (Guschin *et al.*, 1997; Small *et al.*, 2001; Koizumi *et al.*, 2002) oder nur kurze PCR-Amplifikate in die Hybridisierung eingesetzt (Rudi *et al.*, 2000; Wilson *et al.*, 2002), um die Sensitivität der Analyse zu erhöhen. Hierdurch wird die Wahrscheinlichkeit für das Auftreten von falsch-positiven Signalen erheblich gesteigert. In der vorliegenden Arbeit wurde auf eine Fragmentierung der Zielmoleküle verzichtet. In Kombination mit geeigneten Hybridisierungsbedingungen konnten so hochspezifische Signalmuster erhalten werden. Die Hybridisierung kürzerer 16S rDNA-Fragmente führte zu zusätzlichen unspezifischen Signalen (Daten nicht gezeigt). Offenbar stellt eine vollständige, einzelsträngige 16S rDNA einen geeigneten Mediator für eine globale Signalsuppression dar, die im untersuchten Modellsystem zu einer selektiven Verschiebung der falsch-positiver Signale unter die Nachweisgrenze geführt hat (vgl. Abb. 3).

Eine zwangsläufige Konsequenz aus dem beschriebenen Optimierungsansatz ist die erhöhte Anzahl von falsch-negativen Ereignissen (Publikation 1, Abb. 2), deren Auftreten auch schon an anderer Stelle beschrieben wurden (Loy *et al.*, 2002). Da diese Ereignisse durch signallimitierende Parameter verursacht werden, die in ihrem Ausmaß spezifisch für jede Sonde wirken, bietet sich ein gerichteter Einsatz von Fängernukleinsäuren für deren Vermeidung an. Wie beschrieben, kann der Einfluss von signallimitierenden Parametern durch die Art des Aufbringens der Sonde gezielt herabgesetzt werden. Die "Signale" falsch-negativer Sonden sollten also selektiv über die Nachweisgrenze "zu heben" sein. Ein ähnlicher Ansatz wurde bereits 1997 von Guschin *et al.* für die Normalisierung der variierenden Signalintensitäten verschiedener Sonden vorgeschlagen. Hier wurden in Gelstücken befindliche Oligonukleotidsonden in unterschiedlichen Konzentrationen angeboten. In der vorliegenden Arbeit wurde erfolgreich auf den gerichteten Einsatz von Spacern unterschiedlicher Länge zurückgegriffen (Publikation 1, Abb. 5). Allerdings konnten unter den getesteten Bedingungen nicht alle falsch-negativen Ereignisse vermieden werden, ohne dass unspezifische Signale in den Kontrollen auftraten. Ein Nachteil dieses Ansatzes ist, dass die Optimierung gegenwärtig für jede einzelne Sonde empirisch vollzogen werden muss. Daher ist diese Strategie vielmehr als Option zu betrachten, sollten phylogentisch redundante Sonden vereinzelt nicht verfügbar sein (vgl. Abschnitt 3). Grundsätzlich sollten Einschränkungen, die

sich aus falsch-positiven und falsch-negativen Ereignissen für die Dateninterpretation ergeben, durch die Anwendung von komplexen Sondersätzen ausgeschlossen werden, zumal die Microarray-Methodik hierfür eine ideale Plattform darstellt (Loy *et al.*, 2002). Für die Auswertung der erhaltenen Signalmuster sollte dann auf eine computergestützte Dateninterpretation zurückgegriffen werde (Loy *et al.*, 2002; Urakawa *et al.*, 2002).

1.2 Analyse von bakteriellen Gemeinschaften unbekannter Zusammensetzung

Aus der Anpassung der DNA-Microarray-Methodik für den parallelen Nachweis von ausgewählten Populationen in komplexen mikrobiellen Lebensgemeinschaften (Publikation 2) ergeben sich zwei spezifische Anforderungen, die bei der Entwicklung entsprechender Protokolle, insbesondere in ihrer Kombination, oftmals vernachlässigt werden. Hierbei handelt es sich um (i) den direkten Nachweis von transkribierten 16S rRNA-Molekülen aus der untersuchten Probe und (ii) die Validierung der erhaltenen Microarray-Ergebnisse durch eine unabhängige und möglichst quantitative Methode. Die Gründe für diese Notwendigkeit wurden bereits erläutert (vgl. Teil I, Abschnitt 2.2).

1.2.1 Direkter Nachweis von 16S rRNA-Molekülen

Ein Protokoll, das den direkten Nachweis von 16S rRNA-Molekülen durch Microarray-Analyse erlaubt, wurde anhand von bakteriellen Reinkulturen etabliert. Die zelluläre Gesamt-RNA wurde extrahiert und für deren direkte chemische Markierung das "Universal Linkage System" (van Gijlswijk *et al.*, 2001) eingesetzt. Diese Methode erlaubt eine äußerst schnelle (10 min) und einfache Markierung von Nukleinsäuren auf Basis eines Platin-Komplexes, der beliebige Fluoreszenzfarbstoffe tragen kann und bevorzugt an Guanin-Reste der Zielmoleküle kovalent bindet. Die Eignung dieser Markierungsmethode für den Nachweis von RNA-Molekülen durch Microarray-Hybridisierung konnte auch an anderer Stelle gezeigt werden (Gupta *et al.*, 2003). Aufgrund des geringen mRNA-Anteils in den Extrakten war eine zuvorige Abtrennung der rRNA nicht erforderlich. Insgesamt wurde ein Protokoll entwickelt, das zelluläre rRNA in ausreichender Menge und Qualität bereitstellt, deren nicht-enzymatische Fluoreszenzmarkierung ermöglicht und durch einen verhältnismäßig geringen Zeit- (ca. 4 h) und Arbeitsaufwand gekennzeichnet ist.

Die extrahierte und markierte Gesamt-RNA aus zwei der in Publikation 1 eingeführten Referenzstämmen wurde gegen den bekannten SONDENSATZ unter Standardbedingungen hybridisiert (Publikation 2, Abb. 1). Die erhaltenen, hochspezifischen Signale erlaubten auch hier eine eindeutige Differenzierung der untersuchten Stämme. Mit diesen Experimenten konnte die Übertragbarkeit des für die Microarray-Analyse von amplifizierten 16S rRNA-Genen optimierten Protokolls (Publikation 1) auf die Untersuchung markierter zellulärer Gesamt-RNA gezeigt werden. Damit sind auch die zuvor identifizierten Hybridisierungsmechanismen und eingeführten Optimierungsansätze als übertragbar zu bewerten. In Abb. 4 findet sich eine vergleichende Darstellung der erhaltenen Signalmuster nach Hybridisierung der PCR-amplifizierten 16S rDNA und der Gesamt-RNA aus Stamm KT0202a. Während die Qualität des Signalmusters eine hohe Ähnlichkeit aufweist, zeigen sich deutliche Abweichungen in den Verhältnissen der gemessenen Signalintensitäten. Ob diese primär durch Unterschiede in der Sekundär-, bzw. Tertiärstruktur der analysierten ssDNA- und RNA-Moleküle oder durch die höhere Komplexität der Gesamt-RNA-Probe verursacht werden, bleibt offen.

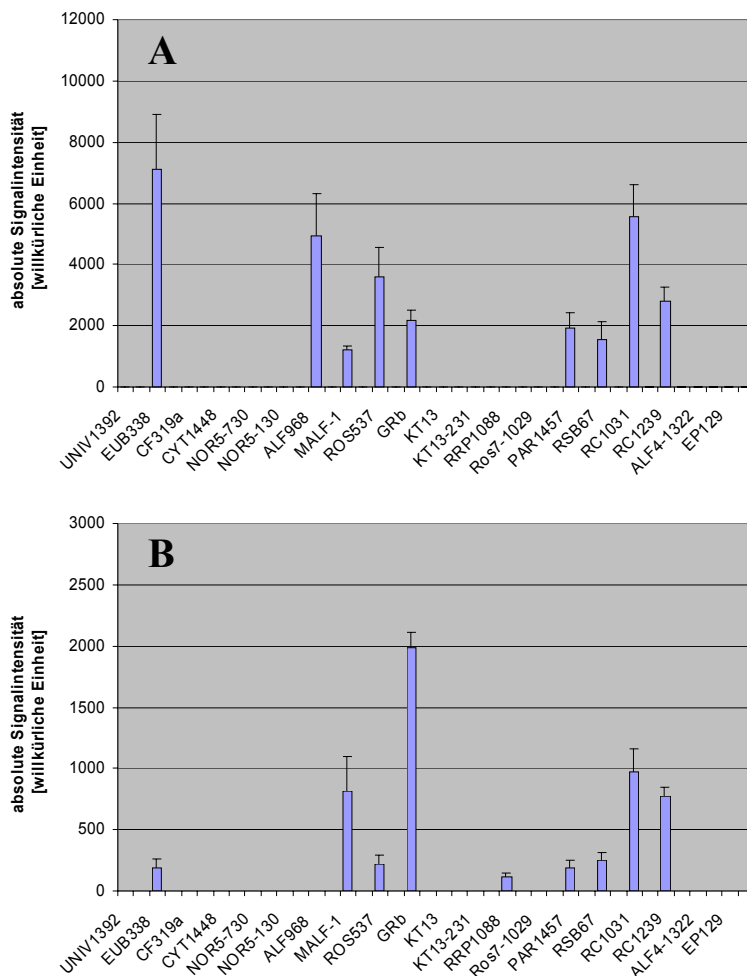


Abbildung 4:
Erhaltene Signalmuster nach Microarray-Hybridisierung der (A) Gesamt-RNA und der (B) amplifizierten 16S rDNA aus Stamm KT0202a gegen den vollständigen SONDENSATZ. Die zusätzlich mitgeführten Kontroll-Oligonukleotide sind nicht gezeigt, da keine falsch-positiven Signale detektiert wurden. In beiden Datensätzen traten zwei falsch-negative Ereignisse auf (A: UNIV1392 und RRP1088; B: UNIV1392 und ALF968).

1.2.2 FISH-Validierung der Ergebnisse aus Mischkulturen und Umweltproben

Die Validierung von Microarray-Daten durch eine geeignete unabhängige Methode wurde in der mikrobiellen Ökologie bisher weitestgehend vernachlässigt. In der vorliegenden Arbeit wurde auf die quantitative Fluoreszenz-*in situ*-Hybridisierung (FISH) zurückgegriffen, um die Microarray-Analyse von komplexen Gemeinschaften unbekannter Zusammensetzung bezüglich ihrer Spezifität, Sensitivität und Quantitativität zu bewerten.

Erste Experimente erfolgten an einer Mischkultur geringer Komplexität, die durch das Animpfen eines synthetischen Mediums mit allen sechs Referenzstämmen aus Publikation 1 und eine anschließende, einwöchige Inkubation etabliert wurde. Die Gesamt-RNA aus einer Probe dieser Kultur wurde extrahiert und gegen den vollständigen SONDENSATZ in einem Microarray-Experiment hybridisiert (Publikation 2, Abb. 3A). Das erhaltene Signalmuster deutete lediglich auf die Anwesenheit der Stämme KT11ds2, KT1117 und JP13.1 in der Kultur hin. Die parallel durchgeführte FISH-Analyse ergab für diese drei Stämme relative zelluläre Abundanzen von 60, 16 und 7% zum Zeitpunkt der Beprobung. Die übrigen Stämme konnten entweder nicht oder mit nur geringen relativen Abundanzen von maximal 2% detektiert werden. Die vorliegenden FISH-Daten zeigen auf, dass durch die Hybridisierung von 0,45 µg intakter zellulärer Gesamt-RNA bei einer Konzentration von 15 ng/µl für 3 h, Populationen mit einer relativen zellulären Abundanz von 7% oder mehr, hochspezifisch durch Microarray-Analyse nachgewiesen werden konnten. Den Nachweis der weniger abundanten Stämme wurde unter den angewandten Bedingungen offenbar durch die Verfügbarkeit der entsprechenden Zielmoleküle limitiert. Im Vergleich hierzu benötigten Koizumi *et al.* (2002) 10 µg (250 ng/µl) fragmentierte Gesamt-RNA und eine Hybridisierungsdauer von 16 h, um Populationen mit relativen Abundanzen von mindestens 16% in einer Anreicherung detektieren zu können. Weitere Vergleichswerte liegen nicht vor.

Eine Korrelation zwischen den relativen zellulären Abundanzen der durch Microarray-Hybridisierung nachgewiesenen Stämme und den Intensitätsverhältnissen der entsprechenden SONDENSIGNALE ließ sich für die analysierte Mischkultur nicht beobachten. Dies war auch nicht zu erwarten, da durch die vorausgegangenen Untersuchungen an Reinkulturen gezeigt werden konnte, dass Sonden, die gegen unterschiedliche Bereiche eines Zielmoleküls gerichtet sind, deutliche Schwankungen in ihrer Signalstärke aufweisen können. Wie bereits in Abschnitt 1.1 erörtert, verhindern eine Reihe von Parametern, welche die Hybridisierungseffizienz verschiedener Sonden selektiv beeinflussen, eine ausschließliche Abhängigkeit der detektierten Signalintensitäten von den Häufigkeiten der entsprechenden Zielmoleküle. Weiterhin ist zu

berücksichtigen, dass durch Microarray-Analyse keine zellulären Abundanzen bestimmt werden können, da dieses Format auf dem Nachweis von freigesetzten Zielmolekülen beruht. Zwar deuteten die durchgeführten FISH-Experimente auf einen vergleichbaren rRNA-Gehalt der drei dominanten Stämme hin, doch auch hier basieren die gemessenen Signalintensitäten auf einer Kombination aus zellulärem Ribosomengehalt und einer für verschiedene Sondenbindungsstellen variierenden *in situ*-Zugänglichkeit (Fuchs *et al.*, 1998; Fuchs *et al.*, 2001; Behrens *et al.*, 2003).

Eine Eignung des in Abschnitt 1.1.2 vorgestellten Optimierungsansatzes zur gezielten Vermeidung von falsch-negativen Signalen konnte anhand der Mischkultur auch für den direkten Nachweis von 16S rRNA Molekülen gezeigt werden (Publikation 2, Abb. 3B). In einem weiteren Experiment wurde der Einfluss einer erhöhten Hybridisierungsdauer (18 statt 3 h) auf die Sensitivität und Spezifität der Analyse untersucht, da zuvor durchgeführte Zeitreihen an Stamm KT0202a unter Verwendung ausgewählter Sonden gezeigt hatten, dass das Verhältnis von Signal zu Kontrollsignal auf diese Weise erhöht werden kann (Publikation 2, Abb. 2). Dieser Ansatz führte zwar zu deutlich erhöhten Signalintensitäten für alle zuvor positiven Sonden, doch wie auch schon für den Einsatz von Helfer-Oligonukleotiden oder Spacer-Molekülen beobachtet (Abschnitt 1.1.1), traten zusätzliche unspezifische Signale auf, die unter diesen Bedingungen nicht mehr zu diskriminieren waren.

Für die Ausweitung der Microarray-Analyse auf eine komplexe natürliche Lebensgemeinschaft wurde das Oberflächenwasser der Nordsee beprobt. Hierbei handelt es sich um ein vergleichsweise nährstoffarmes Habitat, dessen mikrobiellen Vertreter oftmals durch einen geringen Ribosomengehalt gekennzeichnet sind (Pernthaler *et al.*, 2002b). Die Bereitstellung einer ausreichenden Menge an Gesamt-RNA für eine einzelne Hybridisierung erforderte die Filtration von 375 - 750 ml Wasser. Im Vergleich hierzu konnten aus 100 ml der zuvor untersuchten Mischkultur eine Menge an Gesamt-RNA gewonnen werden, die rechnerisch bis zu 100 Microarray-Hybridisierungen ermöglichen würde. Zusätzlich mussten für die Analyse der Umweltprobe Spacer-Moleküle eingesetzt und die Hybridisierungsdauer auf 18 h angehoben werden, um eindeutige Signale zu erhalten (Publikation 2, Abb. 4A). Legt man die Ergebnisse der parallel durchgeführten FISH-Analyse zu Grunde (Publikation 2, Tab. 2), konnten unter diesen Bedingungen alle Populationen mit relativen zellulären Abundanzen von mindestens 5% spezifisch nachgewiesen werden. Um auch kleinere Populationen detektieren zu können, wurde in einem weiteren Experiment die 7-fache Menge an Gesamt-RNA aus der Umweltprobe in die Hybridisierung eingesetzt. Dies führte erneut zu einer allgemeinen Steigerung der Signalintensitäten in Kombination mit einer Reihe von

Signalen, die auf Basis der vorliegenden FISH-Daten als unspezifisch bewertet werden müssen (Publikation 2, Abb. 4B). Auch in dieser Versuchsreihe zeigte sich damit eine enge Verknüpfung der Sensitivität und Spezifität der Analyse. Offenbar ist es unter den in dieser Studie eingesetzten Bedingungen nicht möglich, sehr große und sehr kleine Populationen parallel spezifisch nachzuweisen, da die Spezifität der Analyse nicht über eine temperaturvermittelte Stringenz der Hybridisierung bereitgestellt wurde. Stattdessen basierte die Diskriminierung von unspezifischen Bindungsereignissen auf einer selektiven Signalsuppression, welche unter anderem durch die jeweiligen Konzentrationen der nachzuweisenden Zielmolekül-Populationen beeinflusst wird. Wie aus den vorliegenden Daten ersichtlich ist, führt dieser Ansatz zu einer semi-Quantitativität der Analyse. Unter geeigneten Bedingungen konnten dominante Populationen selektiv nachgewiesen werden.

2 Vergleichende Sequenzanalyse und Sondenentwicklung basierend auf 23S rRNA-Genen

Durch die gegenwärtigen Anstrengungen, die Microarray-Technologie für den Nachweis von ausgewählten Zielsequenzen in Umweltproben komplexer Zusammensetzung verfügbar zu machen, werden in naher Zukunft umfangreiche SONDENSÄTZE schnell, einfach und zuverlässig für die Bearbeitung entsprechender Fragestellungen anwendbar sein. Damit können dann nicht nur eine Vielzahl von Gruppen, Gattungen oder Arten von Mikroorganismen parallel in einer Probe detektiert, sondern auch eine Vielzahl phylogenetisch redundanter und hierarchisch geordneter Sonden eingesetzt werden. Die Verwendung solcher SONDENSÄTZE ist für den eindeutigen Nachweis von ausgewählten Zielsequenzen mittels rRNA-gerichteten Oligonukleotidsonden, auch bei der Anwendung von Standard-Hybridisierungsverfahren, zwingend notwendig (Amann and Schleifer, 2001), wobei mit zunehmender Zahl der Sonden die Zuverlässigkeit der erhaltenen Ergebnisse steigt. Allerdings ist die Entwicklung von entsprechenden SONDENSÄTZEN durch den limitierten Informationsgehalt der etwa 1500 Nukleotide langen 16S rRNA beschränkt. Oftmals finden sich auf diesem Molekül nur einzelne Bindungsstellen für den spezifischen Nachweis ausgewählter phylogenetischer Gruppen. Auch lassen sich nicht alle Verwandtschaftsverhältnisse anhand von kurzen 16S rRNA-Signaturen differenzieren. In diesem Zusammenhang ist eine Erweiterung der vorliegenden 23S rRNA-Datensätze, die momentan nur wenige hundert Vollsequenzen umfassen (Wuyts *et al.*, 2001), von großem Interesse. Die rRNA der großen ribosomalen

Untereinheit weist ebenfalls alle Charakteristika auf, welche die 16S rRNA zu einem idealen Zielmolekül der *in situ*-Hybridisierungstechniken machen (Ludwig and Schleifer, 1994). Weiterhin ist sie durch eine doppelte Länge von etwa 2900 Nukleotiden und einen erhöhten Anteil an variablen Sequenzpositionen gekennzeichnet. Damit stellt dieser Marker ein potenziell umfangreiches Reservoir für zusätzliche Sondenbindungsstellen dar. Die grundsätzliche Eignung der 23S rRNA als Zielmolekül für die dot blot- und *in situ*-Hybridisierung konnte bereits gezeigt werden (Ehrmann *et al.*, 1992; Manz *et al.*, 1992; Fuchs *et al.*, 2001).

In der vorliegenden Arbeit wurden die nahezu vollständigen 23S rRNA-Gene von 11, kürzlich aus der Nordsee isolierten alpha-Proteobakterien (Eilers *et al.*, 2000; Eilers *et al.*, 2001) sequenziert (Publikation 3). Ursprünglich sollten auch marine Vertreter anderer phylogenetischer Gruppen in die Untersuchung miteinbezogen werden. Es zeigte sich allerdings, dass die 23S rDNA einer Reihe von Stämmen nicht mit den verfügbaren "allgemeinen" PCR-Primern zu vervielfältigen sind. Dieser Befund spiegelt die im Vergleich zur 16S rRNA erhöhte Variabilität (Ludwig *et al.*, 1992; Ludwig *et al.*, 1995) speziell in der 5'-Region des Moleküls wider. Um die hieraus erwachsenden Limitationen bei einer PCR-basierten Analyse zu umgehen, wird alternativ auch auf die Amplifikation des gesamten rRNA-Operons zurückgegriffen (Suzuki *et al.*, 2001), zumal auf diese Weise gleichzeitig die 16S rDNA und der intergenische Bereich (internal transcribed spacer, ITS) einer direkten Sequenzierung zugänglich gemacht werden. Auch letztgenannter Marker wird für die vergleichende Sequenzanalyse herangezogen (Scheinert *et al.*, 1996; Söller *et al.*, 2000; Hassan *et al.*, 2003). In der vorliegenden Studie sollten die anfänglichen Schwierigkeiten bei der PCR-Amplifikation von 23S rRNA-Genen in der Folge durch die Fokussierung auf eine einzelne phylogenetische Gruppe behoben werden. Da Vertreter der alpha-Proteobakterien hohe *in situ*-Abundanzen in marinen Freiwasser-Systemen aufweisen (Gonzalez and Moran, 1997; Glöckner *et al.*, 1999; Eilers *et al.*, 2001; Morris *et al.*, 2002), ist dieses Phylum als interessante Zielgruppe zu bewerten. Trotz Ähnlichkeiten von mindestens 82,3% unter den 16S rRNA-Sequenzen der analysierten Stämme erlaubte aber auch hier keine der getesteten Kombination aus bereits veröffentlichten und/oder in dieser Studie entwickelten Primern die Amplifikation der nahezu vollständigen 23S rDNA aller 11 Isolate (Publikation 3, Tab.1).

Die Sequenz der 23S rRNA setzt sich wie die der 16S rRNA aus alternierenden Bereichen geringer und hoher evolutionärer Konserviertheit zusammen und eignet sich daher ebenfalls für die Rekonstruktion von Verwandtschaftsverhältnissen auf verschiedenen phylogenetischen Ebenen (Ludwig and Schleifer, 1994). Zwei Fallstudien, die für die Gruppe

der Gram-positiven Bakterien mit niedrigem GC-Gehalt (Ludwig *et al.*, 1992) und der *Proteobacteria* (Ludwig *et al.*, 1995) durchgeführt wurden, konnten des Weiteren eine hohe Ähnlichkeit für die Topologie von 16S und 23S rRNA-Stammbäumen zeigen. Diese Beobachtungen wurden in der vorliegenden Arbeit für die alpha-Gruppe der *Proteobacteria* anhand eines ausgeweiteten 23S rRNA-Datensatzes bestätigt (Publikation 3, Abb. 1). Da die verfügbaren Modelle für die Rekonstruktion von Stammbäumen in ihren Ergebnissen voneinander abweichen können, sollten idealerweise verschiedene Verfahren herangezogen und sogenannte "Konsensbäume" aus der Summe der gewichteten Ergebnisse erstellt werden, in denen unklare Verwandtschaftsverhältnisse durch Mehrfachverzweigungen gekennzeichnet sind. Signifikante Unterschiede in der lokalen Topologie von 16S und 23S rRNA-Bäumen sind nicht als Widersprüche der beiden Marker zu betrachten, sondern auf deren unterschiedliches phylogenetisches Auflösungsvermögen zurückzuführen (Ludwig *et al.*, 1995). Dieses sollte bei der 23S rRNA aufgrund der größeren absoluten Zahl, sowie des größeren Anteils an variablen Sequenzpositionen höher sein. Es stellt sich daher die Frage, ob die 23S rRNA auch eine erweiterte Differenzierung von Umweltisolaten ermöglicht. Für eine systematische Bearbeitung dieser Fragestellung müssten zunächst Isolate, die verschiedenen Stämmen angehören aber identische 16S rRNA-Gene tragen, durch DNA-DNA-Hybridisierung als solche identifiziert werden. In der Folge könnten dann die entsprechenden 23S rRNA-Gene gezielt analysiert werden. Das Screening von verschiedenen, 16S rRNA-sequenzidentischen Nordsee-Isolaten durch das Ansequenzieren von weniger konservierten Abschnitten auf deren 23S rDNA erlaubte keine zusätzliche Differenzierung (Daten nicht gezeigt). Es ist jedoch wahrscheinlich, dass hier einzelne Stämme mehrfach isoliert wurden. Alternativ wird häufig die hochvariable ITS für die Differenzierung von nahverwandten Organismen herangezogen (Scheinert *et al.*, 1996; Hassan *et al.*, 2003). Mit der zunehmenden Verfügbarkeit von hochsensitiven Hybridisierungsformaten (Pernthaler *et al.*, 2002a) wird dieser zwischen 16S und 23S rDNA gelegene, transkribierte Abschnitt des rRNA-Operons in naher Zukunft auch ein potenzielles Zielmolekül für die *in situ*-Detektion darstellen. Allerdings sind die rRNA-Gene nicht in allen Mikroorganismen in Operonen organisiert (Ward *et al.*, 2000).

Abschließend sollte die Eignung der 16S und 23S rRNA für die Entwicklung von spezifischen Oligonukleotidsonden *in silico* evaluiert und verglichen werden. Hierfür wurden Sequenzgruppen unterschiedlicher phylogenetischer Tiefe, sowie einzelne Stämme innerhalb der alpha-Proteobakterien ausgewählt (Publikation 3, Abb. 1). Aus dem zugrundeliegenden 16S rRNA-Datensatz wurden alle Sequenzen ohne Equivalent im verfügbaren 23S rRNA-Datensatz entfernt, um eine direkte Vergleichbarkeit der Ergebnisse zu gewährleisten.

Erwartungsgemäß stellte die 23S rRNA in fast allen Fällen eine Reihe von zusätzlichen, phylogenetisch redundanten Sonden bereit (Publikation 3, Tab. 3). Insbesondere für die Differenzierung von nahverwandten Organismen wurde die Zahl der verfügbaren Sonden erheblich gesteigert. Auch deutete sich an, dass die 23S rRNA eine erweiterte Differenzierung ermöglicht, wie es bereits für die gamma- und beta-Proteobakterien gezeigt werden konnte (Manz *et al.*, 1992). Offenbar werden die Möglichkeiten einer 16S rRNA-gerichteten Sondenentwicklung durch die zusätzliche Berücksichtigung der 23S rRNA nicht nur ergänzt, sondern deutlich erweitert. Eine Ausweitung der existierenden 23S rRNA-Datensätze ist daher als wünschenswert zu bezeichnen. Auch wenn die systematische Sequenzierung dieses Markers aufgrund des relativ hohen experimentellen Aufwandes nicht absehbar ist, wird die Zahl der verfügbaren 23S rDNA-Sequenzen in Zukunft deutlich ansteigen. Momentan (Mai 2003) befinden sich die vollständigen Genome von 342 Prokaryonten in Sequenzierung (Genome OnLine Database; <http://wit.integratedgenomics.com/GOLD/>) und die Zahl dieser Projekte wird weiter steigen.

3 Ausblick

In der vorliegenden Arbeit konnte gezeigt werden, dass DNA-Microarrays, die auf der Anwendung von 16S rRNA-gerichteten Oligonukleotidsonden basieren, für den spezifischen Nachweis von abundanten Populationen in komplexen mikrobiellen Gemeinschaften aus der Umwelt eingesetzt werden können. Das hier entwickelte Protokoll erlaubt eine direkte chemische Markierung der rRNA-Zielmoleküle und ist für die Gewährleistung einer spezifischen Hybridisierung der 16S rRNA bzw. den Erhalt zuverlässiger Ergebnisse nicht an die zeitlich und momentan noch apparativ aufwendige Aufnahme von Schmelzkurven oder Hybridisierungskinetiken gebunden. Kürzlich wurde der direkte Nachweis von rRNA-Zielmolekülen und die erfolgreiche Anwendung auf Umweltproben in einer unabhängigen Arbeit ebenfalls beschrieben (El Fantroussi *et al.*, 2003). Allerdings wurde hier die Spezifität der einzelnen Bindungsereignisse durch die Aufnahme von Schmelzkurven sichergestellt und auf eine Validierung der Ergebnisse durch die parallele Anwendung einer unabhängigen Methode verzichtet.

Betrachtet man die erzielten Fortschritte bei der Anpassung der Microarray-Technologie an den Nachweis von rRNA-Zielmolekülen in Umweltproben komplexer Zusammensetzung, zeigt sich, dass dieses Hybridisierungsformat in Kürze für entsprechende

Untersuchungen verfügbar sein wird. Hierdurch kann bei einem geeigneten Umfang der angewandten Sondensätze eine erhebliche Beschleunigung der molekularbiologischen Analyse von natürlichen mikrobiellen Gemeinschaften erreicht werden. Nichtsdestotrotz lassen sich aus dem gegenwärtigen Entwicklungsstand der Methodik weitere Optimierungsansätze ableiten. Wie sehen diese im Einzelnen aus? Grundsätzlich sind eine "maximale" Spezifität, Sensitivität und Quantitativität der Analyse anzustreben. Die Ergebnisse der vorliegenden Arbeit weisen darauf hin, dass insbesondere für eine weitere Steigerung der Sensitivität ein ausreichend methodisches und apparatives Potenzial vorhanden ist. Es konnte gezeigt werden, dass die Sensitivität des etablierten Protokolls weniger durch die Verfügbarkeit der Zielmoleküle, die verwendete Markierungsstrategie oder die genutzte technische Plattform limitiert wurde, sondern vielmehr durch die Schwierigkeit, Populationen, die sich deutlich in ihren Abundanzen unterscheiden, parallel und gleichzeitig spezifisch nachzuweisen. Die aufgeführten Aspekte lassen sich also nicht unabhängig voneinander optimieren. An dieser Stelle soll darauf hingewiesen werden, dass die Nachweisgrenze der durchgeführten Untersuchungen definiert war durch die minimale Anzahl an gebundenen Zielmolekülen, die gerade noch detektiert werden konnten. Dieser Wert war abhängig von den eingesetzten Fluoreszenzfarbstoffen und der verwendeten Apparatur zum Auslesen der Signale, und ist somit als feste Größe zu betrachten. Dagegen soll die Sensitivität der Analyse hier definiert werden als die Fähigkeit, Zielmoleküle in Abhängigkeit ihrer Konzentration zu detektieren. Letztere kann somit durch eine Verschiebung von Hybridisierungseffizienzen beeinflusst werden.

Die größte Schwierigkeit bei der parallelen Hybridisierung von Oligonukleotidsonden wird durch deren unterschiedliche Nukleotidsequenzen hervorgerufen. Diese determinieren unter einheitlichen Hybridisierungsbedingungen für jede Sonde sowohl die Spezifität der jeweiligen Reaktion, als auch die Affinität der Sonde zum jeweiligen Zielmolekül und damit die gemessene Signalstärke. Hierbei kann es zu deutlichen Schwankungen kommen. Die Basenabfolge einer Oligonukleotidsonde wirkt sich also unter definierten Bedingungen nicht nur auf die Genauigkeit des Nachweises aus, sondern schränkt auch die Quantifizierung von Zielmolekülen anhand von gemessenen Signalintensitäten ein. Im Folgenden soll erörtert werden, wie das vollständige "Sensitivitätspotenzial" der Analyse ausgeschöpft und eine zuverlässige Quantifizierung der Zielmoleküle erreicht werden könnte, ohne eine Eichung jeder einzelnen Sonde durchführen zu müssen. Letztere ist für Gruppensonden ohnehin nur bedingt anwendbar, da nicht auszuschliessen ist, dass Variationen in den flankierenden

Sequenzbereichen der Zielmoleküle einen unterschiedlichen Einfluss auf die Hybridisierung ausüben.

In einem idealen Microarray-System sollte die Hybridisierungseffizienz einer Sonde primär von der Konzentration des entsprechenden Zielmoleküls abhängig sein. Um diese Anforderung erfüllen zu können, müssen zunächst alle sondenspezifischen Störgrößen eliminiert werden, die nicht auf die Sequenz einer Sonde zurückzuführen sind. Dies sollte durch eine ungerichtete und vollständige Fragmentierung der Zielmoleküle ansatzweise erreichbar sein. Liegen ausschließlich Fragmente vor, deren Länge die der Fängeroligonukleotide nur unwesentlich überschreiten, ist davon auszugehen, dass Parameter wie die Sekundärstruktur der Zielmoleküle und die oberflächenvermittelte sterische Hinderung, deren Ausmaß, wie gezeigt, von der Lage der Sondenbindungsstelle auf dem Zielmolekül abhängig ist, keinen entscheidenden Einfluss mehr auf die Hybridisierungseffizienz haben. Weiterhin können aber Effekte, die beispielsweise durch überstehende einzelsträngige Sequenzabschnitte hervorgerufen werden (Williams *et al.*, 1994), einen sondenselektiven Einfluss auf die gemessenen Signalintensitäten ausüben. Ungeachtet dessen, ergibt sich die eigentliche Herausforderung bei der Optimierung von reversen Hybridisierungsformaten aus den unterschiedlichen Bindungsaffinitäten verschiedener Sonden. Deren Höhe wird durch die jeweilige Sequenz einer Sonde determiniert (Breslauer *et al.*, 1986). Ein Angleich dieser Affinitäten durch Sequenzmodifikation kann für kurze, gegen rRNA-gerichtete Oligonukleotidsonden, nur schwerlich vollzogen werden, da in der Regel bereits einzelne Nukleotidaustausche zu einer veränderten phylogenetischen Zielgruppe der Sonden führen. Bereits früh wurde der Einsatz von chaotropen Additiven wie Tetramethylammoniumchlorid (TMACl) vorgeschlagen (Maskos and Southern, 1993). Diese gleichen aber lediglich grob die Stabilität von GC und AT-Basenpaaren an und eignen sich daher nur eingeschränkt für eine Anpassung der Bindungsaffinitäten von kurzen Oligonukleotidsonden (Jacobs *et al.*, 1988), welche hier primär durch Nearest-Neighbor-Wechselwirkungen bestimmt werden. Vielversprechendere Ansätze ergeben sich aus dem Einsatz von modifizierten Nucleinsäuren. Auf diese Weise können gezielt die Bindungseigenschaften einzelner Nucleotide verändert werden. Ein Beispiel ist der Einsatz von N-4-ethyl-2'-deoxycytidin (dC^{4Et}), das an natürliche Guanosin-Reste bindet (Nguyen *et al.*, 1999). Diese Basenpaare weisen eine sehr ähnliche Stabilität zu natürlichen AT-Basenpaaren auf.

Ein zusätzliche Quelle für das Auftreten von Artefakten, die eine direkte Quantifizierung von Zielmolekülen erschweren können, sind schwankende und/oder sequenzabhängige Markierungseffizienzen. Diese Einschränkung könnte durch einen

markierungsfreien Nachweis der Zielmoleküle umgangen werden. Gleichzeitig würde das Protokoll für die Aufbereitung der Zielmoleküle weiter vereinfacht werden. Entsprechende Methoden sind bereits verfügbar (Haake *et al.*, 2000; Nelson *et al.*, 2002), weisen aber im Vergleich zu einem fluoreszenzbasierten Nachweis eine geringere Sensitivität auf.

Ein weiterer Ansatz für die Optimierung von Microarray-Experimenten beruht auf dem Einsatz von Nukleinsäure-Analoga wie PNAs (Peptide Nucleic Acids) oder LNAs (Locked Nucleic Acids). Diese weisen nicht nur eine erhöhte Bindungsaffinität auf, sondern können auch die Spezifität des Nachweises steigern (Weiler *et al.*, 1997; Simeonov and Nikiforov, 2002). Während die Verwendung von PNAs aufgrund verschiedener Schwierigkeiten bei deren Handhabung, wie beispielsweise deren schlechte Löslichkeit, gegenwärtig nicht von Relevanz ist, haben die LNAs ein breiteres Anwendungsgebiet gefunden (Braasch *et al.*, 2002; Jacobsen *et al.*, 2002; Parekh-Olmedo *et al.*, 2002; Vester *et al.*, 2002).

Eine Nivellierung der Bindungsaffinitäten von verschiedenen Sonden würde nicht nur die Quantitativität einer Microarray-Analyse erhöhen, sondern auch zu einem Angleich der Schmelztemperaturen der angewandten Sonden führen. Durch eine geeignete Stringenz der Hybridisierung sollten dann unspezifische Bindungen mehrheitlich abzutrennen sein, selbst bei einer gesteigerten Sensitivität der Analyse. Letzteres war in den hier durchgeführten Untersuchungen nicht möglich. Sollten entsprechende Entwicklungen in naher Zukunft nicht vollzogen werden können, bieten sich alternative Ansätze an, die eine Abschätzung der Hybridisierungsspezifität für jedes einzelne Bindungsereignis durch die Aufnahme von Hybridisierungskinetiken (Dai *et al.*, 2002) oder Schmelzkurven (Liu *et al.*, 2001) ermöglichen könnten. In diesem Zusammenhang ist die Entwicklung von vollständig integrierten Systemen, welche eine online-Detektion der Hybridisierungsereignisse im Durchfluss erlauben, von großem Interesse.

Grundsätzlich existieren eine Vielzahl von Möglichkeiten, die Sensitivität einer Microarray-Analyse zu erhöhen. In der vorliegenden Arbeit wurden verschiedene Ansätze systematisch untersucht. Hierzu gehören der Einsatz von Helfer-Oligonukleotiden und Spacer-Molekülen, die Verkürzung der zu hybridisierenden Fragmente und die Erhöhung der Zielmolekül-Konzentration. Bereits durch eine Kombination dieser Ansätze sollten die Hybridisierungseffizienzen deutlich zu steigern sein. Weitere Möglichkeiten, die Sensitivität der Analyse zu erhöhen, sind durch Oberflächen-Modifikationen (Benters *et al.*, 2002) oder eine Signalamplifikation gegeben. Letztere kann auf einer "rolling circle"-Amplifikation (Nallur *et al.*, 2001) oder einer enzymatischen Substratumsetzung beruhen (Karsten *et al.*, 2002).

Eine hohe Sensitivität der Analyse ist die Grundvoraussetzung für einen weiteren Anwendungsbereich der DNA-Microarray-Methodik in der mikrobiellen Ökologie, sofern die Analyse nicht auf einem vorgeschalteten PCR-Schritt beruhen soll. Hierbei handelt es sich um den Nachweis von funktionellen Genen aus Umweltproben. Grundsätzlich kann dieser auf Ebene des Genoms oder des Transkriptoms erfolgen. Für beide Ansätze liegt im Falle eines direkten Nachweises eine stark reduzierte Verfügbarkeit der Zielmoleküle vor. Bisher wurde nur der PCR-basierte Nachweis von Umweltgenen (Wu *et al.*, 2001; Cho and Tiedje, 2002; Taroncher-Oldenburg *et al.*, 2003) bzw. der cDNA-basierte Nachweis von mRNA-Populationen (Dennis *et al.*, 2003) beschrieben.

Wie erläutert, existieren verschiedene Strategien für die Optimierung von Microarray-Experimenten, deren Eignung bereits gezeigt werden konnte. Es werden weitere Studien benötigt, um diese Ansätze auf die in der mikrobiellen Ökologie eingesetzten Systeme zu übertragen. Gleichzeitig sollte die Zahl an Umweltstudien und ihr Umfang ausgedehnt werden, um eine breite Basis für die Etablierung der DNA-Microarray-Methodik als Routine-Werkzeug in der mikrobiellen Umweltforschung zu schaffen. In diesem Zusammenhang wird eine begleitende Validierung der erhaltenen Microarray-Daten unabdingbar sein. Welchen Weg die künftigen Entwicklungen einschlagen werden und mit welcher Geschwindigkeit sie vollzogen werden können, bleibt abzuwarten.

C

Literaturverzeichnis

C Literaturverzeichnis

- Amann, R.I., Snaidr, J., Wagner, M., Ludwig, W., and Schleifer, K.-H. (1996) *In situ* visualization of high genetic diversity in a natural microbial community. *J. Bacteriol.* **178**: 3496-3500.
- Amann, R.I., and Schleifer, K.-H. (2001) Nucleic acid probes and their application in environmental microbiology. p. 67-82. In G. Garrity, D.R. Boone, and R.W. Castenholz (ed.), *Bergey's Manual of Systematic Bacteriology, 2nd ed., vol 1., Springer Verlag, New York.*
- Amann, R.I., Ludwig, W., and Schleifer, K.-H. (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**: 143-169.
- Amann, R.I., Zarda, B., Stahl, D.A., and Schleifer, K.-H. (1992) Identification of individual prokaryotic cells by using enzyme-labeled, rRNA-targeted oligonucleotide probes. *Appl. Environ. Microbiol.* **58**: 3007-3011.
- Bains, W., and Smith, G.C. (1988) A novel method for nucleic acid sequence determination. *J. Theor. Biol.* **135**: 303-308.
- Behrens, S., Ruhland, C., Inacio, J., Huber, H., Fonseca, A., Spencer-Martins, I. *et al.* (2003) *In situ* accessibility of small-subunit rRNA of members of the domains *Bacteria*, *archaea*, and *Eucarya* to Cy3-labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **69**: 1748-1758.
- Beier, M., and Hoheisel J.D. (1999) Versatile derivatisation of solid support media for covalent bonding on DNA microchips. *Nucleic Acids Res.* **27**: 1970-1977.
- Beja, O., Suzuki, M.T., Koonin, E.V., Aravind, L., Hadd, A., Nguyen, L.P. *et al.* (2000) Construction and analysis of bacterial artificial chromosome libraries from a marine microbial assemblage. *Environ. Microbiol.* **2**: 516-529.
- Benters, R., Niemeyer, C.M., Drutschmann, D., Blohm, D., and Wöhrle, D. (2002) DNA microarrays with PAMAM dendritic linker systems. *Nucleic Acids Res.* **30**: e10.
- Braasch, D.A., Liu, Y., and Corey, D.R. (2002) Antisense inhibition of gene expression in cells by oligonucleotides incorporating locked nucleic acids: effect of mRNA target sequence and chimera design. *Nucleic Acids Res.* **30**: 5160-5167.
- Breslauer, K.J., Frank, R., Blocker, H., and Marky, L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83**: 3746-3750.

- Chakravarti, A. (1999) Population genetics - making sense out of sequence. *Nat. Genet.* **21**: 56-60.
- Chen, J., Sahota, A., Stambrook, P.J., and Tischfield, J.A. (1991) Polymerase Chain Reaction amplification and sequence analysis of human mutant adenine phosphoribosyltransferase genes: the nature and frequency of errors caused by *Taq* DNA polymerase. *Mutat. Res.* **249**: 169-176.
- Cho, J.C., and Tiedje, J.M. (2002) Quantitative detection of microbial genes by using DNA microarrays. *Appl. Environ. Microbiol.* **68**: 1425-1430.
- Chuang, S.-E., Daniels, D.L., and Blattner, F.R. (1993) Global regulation of gene expression in *Escherichia coli*. *J. Bacteriol.* **175**: 2026-2036.
- Churchill, G.A. (2002) Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.* **32**: 490-495.
- Clayton, R.A., Sutton, G., Hinkle, P.S., Bult, C., and Fields, C. (1995) Intraspecific variation in small-subunit rRNA sequences in Genbank - why single sequences may not adequately represent prokaryotic taxa. *Int. J. Syst. Bacteriol.* **45**: 595-599.
- Cole, J.R., Chai, B., Marsh, T.L., Farris, R.J., Wang, Q., Kulam, S.A. *et al.* (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.* **31**: 442-443.
- Dai, H.Y., Meyer, M., Stepaniants, S., Ziman, M., and Stoughton, R. (2002) Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays. *Nucleic Acids Res.* **30**: e86.
- Debouck, C., and Goodfellow, P.N. (1999) DNA microarrays in drug discovery and development. *Nat. Genet.* **21**: 48-50.
- DeLong, E.F. (1992) *Archaea* in coastal marine environments. *Proc. Natl. Acad. Sci. USA* **89**: 5685-5689.
- DeLong, E.F. (2002) Microbial population genomics and ecology. *Curr. Opin. Microbiol.* **5**: 520-524.
- DeLong, E.F., Wickham, G.S., and Pace, N.R. (1989) Phylogenetic stains: ribosomal RNA-based probes for the identification of single microbial cells. *Science* **243**: 1360-1363.
- Dennis, P., Edwards, E.A., Liss, S.N., and Fulthorpe, R. (2003) Monitoring gene expression in mixed microbial communities by using DNA microarrays. *Appl. Environ. Microbiol.* **69**: 769-778.

- Diehl, F., Grahlmann, S., Beier, M., and Hoheisel, J.D. (2001) Manufacturing DNA microarrays of high spot homogeneity and reduced background signal. *Nucleic Acids Res* **29**: e38.
- Drmanac, R., Labat, I., Brukner, I., and Crkvenjakov, R. (1989) Sequencing of megabase plus DNA by hybridization: theory of the method. *Genomics* **4**: 114-128.
- Duggan, D.J., Bittner, M., Chen, Y.D., Meltzer, P., and Trent, J.M. (1999) Expression profiling using cDNA microarrays. *Nat. Genet.* **21**: 10-14.
- Ehrmann, M., Ludwig, W., and Schleifer, K.-H. (1992) Species specific oligonucleotide probe for the identification of *Streptococcus thermophilus*. *Syst. Appl. Microbiol.* **15**: 453-455.
- Ehrmann, M., Ludwig, W., and Schleifer, K.-H. (1994) Reverse dot blot hybridization: a useful method for the direct identification of lactic acid bacteria in fermented food. *FEMS Microbiol. Lett.* **117**: 143-149.
- Eilers, H., Pernthaler, J., Glöckner, F.O., and Amann, R. (2000) Culturability and *in situ* abundance of pelagic bacteria from the North Sea. *Appl. Environ. Microbiol.* **66**: 3044-3051.
- Eilers, H., Pernthaler, J., Peplies, J., Glöckner, F.O., Gerds, G., and Amann, R. (2001) Isolation of novel pelagic bacteria from the German Bight and their seasonal contributions to surface picoplankton. *Appl. Environ. Microbiol.* **67**: 5134-5142.
- El Fantroussi, S., Urakawa, H., Bernhard, A.E., Kelly, J.J., Noble, P.A., Smidt, H. *et al.* (2003) Direct profiling of environmental microbial populations by thermal dissociation analysis of native rRNAs hybridized to oligonucleotide microarrays. *Appl. Environ. Microbiol.* **69**: 2377-2382.
- Fodor, S.P.A., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T., and Solas, D. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science* **251**: 767-773.
- Fuchs, B.M., Glöckner, F.O., Wulf, J., and Amann, R. (2000) Unlabeled helper oligonucleotides increase the *in situ* accessibility of 16S rRNA for fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **66**: 3603-3607.
- Fuchs, B.M., Sytsubo, K., Ludwig, W., and Amann, R. (2001) *In situ* accessibility of *Escherichia coli* 23S rRNA to fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **67**: 961-968.
- Fuchs, B.M., Wallner, G., Beisker, W., Schwippl, I., Ludwig, W., and Amann, R. (1998) Flow cytometric analysis of the *in situ* accessibility of *Escherichia coli* 16S rRNA for fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **64**: 4973-4982.

- Fuhrman, J.A. (2002) Community structure and function in prokaryotic marine plankton. *Anton. Leeuw. Int. J. G.* **81**: 521-527.
- Fuhrman, J.A., McCallum, K., and Davis, A.A. (1992) Novel major archaeobacterial group from marine plankton. *Nature* **356**: 148-149.
- Giovannoni, S.J., DeLong, E.F., Olsen, G.J., and Pace, N.R. (1988) Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. *J. of Bacteriol.* **170**: 720-726.
- Giovannoni, S.J., Britschgi, T.B., Moyer, C.L., and Field, K.G. (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60-63.
- Glöckner, F.O. (1999) Phylogenie und *in situ*-Identifizierung von Prokaryonten in limnischen und marinen Ökosystemen. Dissertation. Technische Universität, München, Deutschland
- Glöckner, F.O., Fuchs, B.M., and Amann, R. (1999) Bacterioplankton composition in lakes and oceans: a first comparison based on fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **65**: 3721-3726.
- Glöckner, F.O., Amann, R., Alfreider, A., Pernthaler, J., Psenner, R., Trebesius, K., and Schleifer, K.-H. (1996) An *in situ* hybridization protocol for detection and identification of planktonic bacteria. *Syst. Appl. Microbiol.* **19**: 403-406.
- Göbel, U.B., Geiser, A., and Stanbridge, E.J. (1987) Oligonucleotide probes complementary to variable regions of ribosomal RNA discriminate between *Mycoplasma* species. *J. Gen. Microbiol.* **133**: 1969-1974.
- Gonzalez, J.M., and Moran, M.A. (1997) Numerical dominance of a group of marine bacteria in the alpha-subclass of the class *Proteobacteria* in coastal seawater. *Appl. Environ. Microbiol.* **63**: 4237-4242.
- Gonzalez, J.M., Kiene, R.P., and Moran, M.A. (1999) Transformation of sulfur compounds by an abundant lineage of marine bacteria in the alpha-subclass of the class *Proteobacteria*. *Appl. Environ. Microbiol.* **65**: 3810-3819.
- Gonzalez, J.M., Whitman, W.B., Hodson, R.E., and Moran, M.A. (1996) Identifying numerically abundant culturable bacteria from complex communities: an example from a lignin enrichment culture. *Appl. Environ. Microbiol.* **62**: 4433-4440.
- Guo, Z., Guilfoyle, R.A., Thiel, A.J., Wang, R., and Smith, L.M. (1994) Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.* **22**: 5456-5465.

- Gupta, V., Cherkassky, A., Chatis, P., Joseph, R., Johnson, A.L., Broadbent, J. *et al.* (2003) Directly labeled mRNA produces highly precise and unbiased differential gene expression data. *Nucleic Acids Res.* **31**: e13.
- Guschin, D.Y., Mobarry, B.K., Proudnikov, D., Stahl, D.A., Rittmann, B.E., and Mirzabekov, A.D. (1997) Oligonucleotide microchips as genosensors for determinative and environmental studies in microbiology. *Appl. Environ. Microbiol.* **63**: 2397-2402.
- Haake, H.M., Schutz, A., and Gauglitz, G. (2000) Label-free detection of biomolecular interaction by optical sensors. *Fresen. J. Anal. Chem.* **366**: 576-585.
- Hacia, J.G., Brody, L.C., Chee, M.S., Fodor, S.P.A., and Collins, F.S. (1996) Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-colour fluorescence analysis. *Nat. Genet.* **14**: 441-447.
- Hassan, A.A., Khan, I.U., Abdulmawjood, A., and Lämmle, C. (2003) Inter- and intraspecies variations of the 16S-23S rDNA intergenic spacer region of various Streptococcal species. *Syst. Appl. Microbiol.* **26**: 97-103.
- Head, I.M., Saunders, J.R., and Pickup, R.W. (1998) Microbial evolution, diversity, and ecology: A decade of ribosomal RNA analysis of uncultivated microorganisms. *Microb. Ecol.* **35**: 1-21.
- Heller, R., A., Schena, M., Chai, A., Shalon, D., Bedilion, T., Gilmore, J. *et al.* (1997) Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc. Natl. Acad. Sci. USA* **94**: 2150-2155.
- Hicks, R.E., Amann, R.I., and Stahl, D.A. (1992) Dual staining of natural bacterioplankton with 4',6-diamino-2-phenylindole and fluorescent oligonucleotide probes targeting kingdom level 16S rRNA sequences. *Appl. Environ. Microbiol.* **58**: 2158-2163.
- Höfle, M.G., and Brettar, I. (1996) Genotyping of heterotrophic bacteria from the central Baltic Sea by use of low-molecular-weight RNA profiles. *Appl. Environ. Microbiol.* **62**: 1383-1390.
- Jacobs, K., Rudersdorf, R., Neill, S., Dougherty, J., Brown, E., and Fritsch, E. (1988) The thermal stability of oligonucleotide duplexes is sequence independent in tetraalkylammonium salt solutions: application to identifying recombinant DNA clones. *Nucleic Acids Res.* **16**: 4637-4650.
- Jacobsen, N., Bentzen, J., Meldgaard, M., Jakobsen, M.H., Fenger, M., Kauppinen, S., and Skouv, J. (2002) LNA-enhanced detection of single nucleotide polymorphisms in the apolipoprotein E. *Nucleic Acids Res.* **30**: e100.

- Kafatos, F.C., Jones, C.W., and Efstratiadis, A. (1979) Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. *Nucleic Acids Res.* **7**: 1541-1552.
- Karner, M.B., DeLong, E.F., and Karl, D.M. (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507-510.
- Karsten, S.L., Van Deerlin, V.M.D., Sabatti, C., H. Gill, L., and Geschwind, D.H. (2002) An evaluation of tyramide signal amplification and archived fixed and frozen tissue in microarray gene expression analysis. *Nucleic Acids Res.* **30**: e4.
- Kemp, P.F., Lee, S., and LaRoche, J. (1993) Estimating the growth rate of slowly growing marine bacteria from RNA content. *Appl. Environ. Microbiol.* **59**: 2594-2601.
- Kharpko, K.R., Lysov Yu, P., Khorlyn, A.A., Shick, V.V., Florentiev, V.L., and Mirzabekov, A.D. (1989) An oligonucleotide hybridization approach to DNA sequencing. *FEBS Lett.* **256**: 118-122.
- Klappenbach, J.A., Saxman, P.R., Cole, J.R., and Schmidt, T.M. (2001) rrndb: the Ribosomal RNA Operon Copy Number Database. *Nucleic Acids Res.* **29**: 181-184.
- Koizumi, Y., Kelly John, J., Nakagawa, T., Urakawa, H., El Fantroussi, S., Al Muzaini, S. *et al.* (2002) Parallel characterization of anaerobic toluene- and ethylbenzene-degrading microbial consortia by PCR-denaturing gradient gel electrophoresis, RNA-DNA membrane hybridization, and DNA microarray technology. *Appl. Environ. Microbiol.* **68**: 3215-3225.
- Lander, E.S. (1999) Array of hope. *Nat. Genet.* **21**: 3-4.
- Lee, D., Zo, Y., and Kim, S. (1996) Nonradioactive method to study genetic profiles of natural bacterial communities by PCR-single-strand-conformation polymorphism. *Appl. Environ. Microbiol.* **62**: 3112-3120.
- Lee, N., Nielsen, P., Andreasen, K., Juretschko, S., Nielsen, J., 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.
- Lee, S.H., Malone, C., and Kemp, P.F. (1993) Use of multiple 16S rRNA-targeted fluorescent probes to increase signal strength and measure cellular RNA from natural planktonic bacteria. *Mar. Ecol. Prog. Ser.* **101**: 193-201.

- Liesack, W., Weyland, H., and Stackebrandt, E. (1991) Potential risks of gene amplification by PCR as determined by 16S rDNA analysis of a mixed-culture of strict barophilic bacteria. *Microb. Ecol.* **21**: 191-198.
- Lipshutz, R.J., Fodor, S.P.A., Gingeras, T.R., and Lockhart, D.J. (1999) High density synthetic oligonucleotide arrays. *Nat. Genet.* **21**: 20-24.
- Liu, W.T., Marsh, T.L., Cheng, H., and Forney, L.J. (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl. Environ. Microbiol.* **63**: 4516-4522.
- Liu, W.T., Mirzabekov, A.D., and Stahl, D.A. (2001) Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. *Environ. Microbiol.* **3**: 619-629.
- Loy, A., Lehner, A., Lee, N., Adamczyk, J., Meier, H., Ernst, J. *et al.* (2002) Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl. Environ. Microbiol.* **68**: 5064-5081.
- Lucchini, S., Thompson, A., and Hinton, J.C.D. (2001) Microarrays for microbiologists. *Microbiology* **147**: 1403-1414.
- Ludwig, W., and Schleifer, K.H. (1994) Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiol. Rev.* **15**: 155-173.
- Ludwig, W., Kirchhof, G., Klugbauer, N., Weizenegger, M., Betzl, D., Ehrmann, M. *et al.* (1992) Complete 23S ribosomal RNA sequences of gram-positive bacteria with a low DNA G plus C content. *Syst. Appl. Microbiol.* **15**: 487-501.
- Ludwig, W., Rossello-Mora, R., Aznar, R., Klugbauer, S., Spring, S., Reetz, K. *et al.* (1995) Comparative sequence analysis of 23S rRNA from *Proteobacteria*. *Syst. Appl. Microbiol.* **18**: 164-188.
- MacBeath, G. (2002) Protein microarrays and proteomics. *Nat. Genet.* **32**: 526-532.
- 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.
- Manz, W., Amann, R., Ludwig, W., Wagner, M., and Schleifer, K.-H. (1992) Phylogenetic oligodeoxynucleotide probes for the major subclasses of *Proteobacteria*: problems and solutions. *Syst. Appl. Microbiol.* **15**: 593-600.
- Marchesi, J.R., Sato, T., Weightman, A.J., Martin, T.A., Fry, J.C., Hiom, S.J., and Wade, W.G. (1998) Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Appl. Environ. Microbiol.* **64**: 795-799.

- Martinez, J., Smith, D.C., Steward, F., and Azam, F. (1996) Variability in ectohydrolytic enzyme activities of pelagic marine bacteria and its significance for substrate processing in the sea. *Aquat. Microb. Ecol.* **10**: 223-230.
- Martinez-Murcia, A.J., Acinas, S.G., and Rodriguez-Valera, F. (1995) Evaluation of prokaryotic diversity by restrictase digestion of 16S rDNA directly amplified from hypersaline environments. *FEMS Microbiol. Ecol.* **17**: 247-256.
- Maskos, U., and Southern, E.M. (1992) Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised *in situ*. *Nucleic Acids Res.* **20**: 1679-1684.
- Maskos, U., and Southern Edwin, E.M. (1993) A study of oligonucleotide reassociation using large arrays of oligonucleotides synthesised on a glass support. *Nucleic Acids Res.* **21**: 4663-4669.
- Mir, K.U., and Southern, E.M. (1999) Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat. Biotechnol.* **17**: 788-792.
- Morris, R.M., Rappe, M.S., Connon, S.A., Vergin, K.L., Siebold, W.A., Carlson, C.A., and Giovannoni, S.J. (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**: 806-810.
- Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **59**: 695-700.
- Nallur, G., Luo, C.H., Fang, L.H., Cooley, S., Dave, V., Lambert, J. *et al.* (2001) Signal amplification by rolling circle amplification on DNA microarrays. *Nucleic Acids Res.* **29**: e118.
- Nelson, B.P., Liles, M.R., Frederick, K.B., Corn, R.M., and Goodman, R.M. (2002) Label-free detection of 16S ribosomal RNA hybridization on reusable DNA arrays using surface plasmon resonance imaging. *Environ. Microbiol.* **4**: 735-743.
- Nguyen, H.-K., Fournier, O., Asseline, U., Dupret, D., and Thuong, N. (1999) Smoothing of the thermal stability of DNA duplexes by using modified nucleosides and chaotropic agents. *Nucleic Acids Res.* **27**: 1492-1498.
- Nguyen, H.-K., and Southern E.M. (2000) Minimising the secondary structure of DNA targets by incorporation of a modified deoxynucleoside: Implications for nucleic acid analysis by hybridisation. *Nucleic Acids Res.* **28**: 3904-3909.

- Niemeyer, C.M., Bürger, W., and Peplies, J. (1998) Covalent DNA-streptavidin conjugates as building blocks for novel biometallic nanostructures. *Angew. Chem. Int. Ed.* **37**: 2265-2268.
- Olsen, G.J., Lane, D.J., Giovannoni, S.J., Pace, N.R., and Stahl, D.A. (1986) Microbial ecology and evolution: a ribosomal rRNA approach. *Annu. Rev. Microbiol.* **40**: 337-365.
- O'Meara, D., Nilsson, P., Nygren, P.-A., Uhlen, M., and Lundeberg, J. (1998) Capture of single-stranded DNA assisted by oligonucleotide modules. *Anal. Biochem.* **255**: 195-203.
- Ouverney, C.C., and Fuhrman, J.A. (1999) Combined microautoradiography-16S rRNA probe technique for determination of radioisotope uptake by specific microbial cell types in situ. *Appl. Environ. Microbiol.* **65**: 1746-1752.
- Pace, N.R., Stahl, D.A., Lane, D.J., and Olsen, G.J. (1986) The analysis of natural microbial populations by ribosomal RNA sequences. *Adv. Microb. Ecol.* **9**: 1-55.
- Parekh-Olmedo, H., Drury, M., and Kmiec, E.B. (2002) Targeted nucleotide exchange in *Saccharomyces cerevisiae* directed by short oligonucleotides containing locked nucleic acids. *Chem. Biol.* **9**: 1073-1084.
- Pernthaler, A., Pernthaler, J., and Amann, R. (2002a) Fluorescence *in situ* hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl. Environ. Microbiol.* **68**: 3094-3101.
- Pernthaler, A., Preston C.M., Pernthaler, J., DeLong E.F., and Amann, R. (2002b) Comparison of fluorescently labeled oligonucleotide and polynucleotide probes for the detection of pelagic marine bacteria and archaea. *Appl. Environ. Microbiol.* **68**: 661-667.
- Pernthaler, J., Glöckner, F.O., Schönhuber, W., and Amann, R. (2001) Fluorescence *in situ* hybridization (FISH) with rRNA-targeted oligonucleotide probes. *Methods Microbiol.* **30**: 207-226.
- Peterson, A.W., Heaton, R.J., and Georgiadis, R.M. (2001) The effect of surface probe density on DNA hybridization. *Nucleic Acids Res.* **29**: 5163-5168.
- Pinhassi, J., Zweifel, U.L., and Hagström, A. (1997) Dominant marine bacterioplankton species found among colony-forming bacteria. *Appl. Environ. Microbiol.* **63**: 3359-3366.
- Ramsay, G. (1998) DNA Chips: State-of-the-art. *Nat. Biotechnol.* **16**: 40-44.
- Rappe, M.S., Connon, S.A., Vergin, K.L., and Giovannoni, S.J. (2002) Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* **418**: 630-633.

- Raskin, L., Poulsen, L.K., Noguera, D.R., Rittmann, B.E., and Stahl, D.A. (1994) Quantification of methanogenic groups in anaerobic biological reactors by oligonucleotide probe hybridization. *Appl. Environ. Microbiol.* **60**: 1241-1248.
- Ravenschlag, K., Sahn, K., Knoblauch, C., Jørgensen, B.B., and Amann, R. (2000) Community structure, cellular rRNA content and activity of sulfate-reducing bacteria in marine Arctic sediments. *Appl. Environ. Microbiol.* **66**: 3592-3602.
- Rees, W.A., Yager, T.D., Korte, J., and Von Hippel, P.H. (1993) Betaine can eliminate the base pair composition dependence of DNA melting. *Biochemistry* **32**: 137-144.
- Religio, A., Schwager, C., Richter, A., Ansorge, W., and Valcarcel, J. (2002) Optimization of oligonucleotide-based DNA microarrays. *Nucleic Acids Res.* **30**: e51.
- Reysenbach, A.L., Giver, L.J., Wickham, G.S., and Pace, N.R. (1992) Differential amplification of rRNA genes by polymerase chain reaction. *Appl. Environ. Microbiol.* **58**: 3417-3418.
- Roller, C., Wagner, M., Amann, R., Ludwig, W., and Schleifer, K.-H. (1994) *In situ* probing of Gram-positive bacteria with high DNA G+C content using 23S rRNA-targeted oligonucleotides. *Microbiology* **140**: 2849-2858.
- 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.
- Rudi, K., Skulberg, O.M., Skulberg, R., and Jakobsen, K.S. (2000) Application of sequence-specific labeled 16S rRNA gene oligonucleotide probes for genetic profiling of cyanobacterial abundance and diversity by array hybridization. *Appl. Environ. Microbiol.* **66**: 4004-4011.
- Sahn, K., MacGregor, B.J., Jørgensen, B.B., and Stahl, D.A. (1999) Sulphate reduction and vertical distribution of sulphate-reducing bacteria quantified by rRNA slot-blot hybridization in a coastal marine sediment. *Environ. Microbiol.* **1**: 65-74.
- Saiki, R.K., Walsh, P.S., Levenson, C.H., and Erlich, H.A. (1989) Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proc. Natl. Acad. Sci. USA* **86**: 6230-6234.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T. *et al.* (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-491.

- Scheinert, P., Krausse, R., Ullmann, U., Söller, R., and Krupp, G. (1996) Molecular differentiation of bacteria by PCR amplification of the 16S-23S rRNA spacer. *J. Microbiol. Methods* **26**: 103-117.
- Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**: 467-470.
- Schmalenberger, A., Schwieger, F., and Tebbe C.C. (2001) Effect of primers hybridizing to different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling. *Appl. Environ. Microbiol.* **67**: 3557-3563.
- Schmidt, T.M., DeLong, E.F., and Pace, N.R. (1991) Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J. Bacteriol.* **173**: 4371-4378.
- Shchepinov, M.S., Case-Green, S.C., and Southern, E.M. (1997) Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. *Nucleic Acids Res.* **25**: 1155-1161.
- Simeonov, A., and Nikiforov, T.T. (2002) Single nucleotide polymorphism genotyping using short, fluorescently labeled locked nucleic acid (LNA) probes and fluorescence polarization detection. *Nucleic Acids Res.* **30**: e91
- Small, J., Call, D.R., Brockman, F.J., Straub, T.M., and Chandler, D.P. (2001) Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. *Appl. Environ. Microbiol.* **67**: 4708-4716.
- Söller, R., Hirsch, P., Blohm, D., and Labrenz, M. (2000) Differentiation of newly described Antarctic bacterial isolates related to *Roseobacter* species based on 16S-23S rDNA internal transcribed spacer sequences. *Int. J. Syst. Evol. Micr.* **50**: 909-915.
- Southern, E.M., Mir, K., and Shchepinov, M. (1999) Molecular interactions on microarrays. *Nat. Genet.* **21**: 5-9.
- Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**: 503-517.
- Stahl, D.A., and Amann, R. (1991) Development and application of nucleic acid probes. In *E. Stackebrandt, and M. Goodfellow (eds), Nucleic Acid Techniques in Bacterial Systematics, John Wiley&Sons Ltd., Chichester, UK.*
- Stahl, D.A., Lane, D.J., Olsen, G.J., and Pace, N.R. (1984) Analysis of hydrothermal vent associated symbionts by ribosomal RNA sequences. *Science* **224**: 409-411.

- Stahl, D.A., Lane, D.J., Olsen, G.J., and Pace, N.R. (1985) Characterization of a Yellowstone hot spring microbial community by 5S rRNA sequences. *Appl. Environ. Microbiol.* **49**: 1379-1384.
- Stahl, D.A., Flesher, B., Mansfield, H.R., and Montgomery, L. (1988) Use of phylogenetically based hybridization probes for studies of ruminal microbial ecology. *Appl. Environ. Microbiol.* **54**: 1079-1084.
- Stears, R.L., Martinsky, T., and Schena, M. (2003) Trends in microarray analysis. *Nat. Med.* **9**: 140-145.
- Sudarsanam, P., Iyer, V.R., Brown, P.O., and Winston, F. (2000) Whole-genome expression analysis of *snf/swi* mutants of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **97**: 3364-3369.
- Suzuki, M.T., and Giovannoni, S.J. (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol.* **62**: 625-630.
- Suzuki, M.T., Beja, O., Taylor, L.T., and DeLong, E.F. (2001) Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton. *Environ. Microbiol.* **3**: 323-331.
- Taroncher-Oldenburg, G., Griner, E.M., Francis, C.A., and Ward, B.B. (2003) Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment. *Appl. Environ. Microbiol.* **69**: 1159-1171.
- Urakawa, H., Noble P.A., El Fantroussi, S., Kelly J.J., and Stahl D.A. (2002) Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses. *Appl. Environ. Microbiol.* **68**: 235-244.
- Urakawa, H., El Fantroussi, S., Smidt, H., Smoot, J.C., Tribou, E.H., Kelly, J.J. *et al.* (2003) Optimization of single-base-pair mismatch discrimination in oligonucleotide microarrays. *Appl. Environ. Microbiol.* **69**: 2848-2856.
- van Gijlswijk, R.P.M., Talman, E.G., Janssen, P.J.A., Snoeijers, S.S., Killian, J., Tanke, H.J., and Heetebrij, R.J. (2001) Universal Linkage System: versatile nucleic acid labeling technique. *Expert Rev. Mol. Diagn.* **1**: 81-91.
- Vester, B., Lundberg, L.B., Sorensen, M.D., Babu, B.R., Douthwaite, S., and Wengel, J. (2002) LNAzymes: Incorporation of LNA-type monomers into DNAzymes markedly increases RNA cleavage. *J. Am. Chem. Soc.* **124**: 13682-13683.

- Wallner, G., Amann, R., and Beisker, W. (1993) Optimizing fluorescent *in situ* hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* **14**: 136-143.
- Ward, D.M., Weller, R., and Bateson, M.M. (1990) 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**: 63-65.
- Ward, N.L., Rainey, F.A., Hedlund, B.P., Staley, J.T., Ludwig, W., and Stackebrandt, E. (2000) Comparative phylogenetic analyses of members of the order *Planctomycetales* and the division *Verrucomicrobia*: 23S rRNA gene sequence analysis supports the 16S RNA gene sequence-derived phylogeny. *Int. J. Syst. Evol. Micr.* **50**: 1965-1972.
- Weiler, J., Gausepohl, H., Hauser, N., Jensen O.N., and Hoheisel, J.D. (1997) Hybridisation based DNA screening on peptide nucleic acid (PNA) oligomer arrays. *Nucleic Acids Res.* **25**: 2792-2799.
- Weller, R., and Ward, D.M. (1989) Selective recovery of 16S rRNA sequences from natural microbial communities in the form of cDNA. *Appl. Environ. Microbiol.* **55**: 1818-1822.
- Williams, J.C., Case-Green, S.C., Mir K.U., and Southern, E.M. (1994) Studies of oligonucleotide interactions by hybridisation to arrays: the influence of dangling ends on duplex yield. *Nucleic Acids Res.* **22**: 1365-1367.
- Wilson, K.H., Wilson, W.J., Radosevich, J.L., DeSantis, T.Z., Viswanathan, V.S., Kuczmariski, T.A., and Andersen, G.L. (2002) High-density microarray of small-subunit ribosomal DNA probes. *Appl. Environ. Microbiol.* **68**: 2535-2541.
- Wintzingerode, F.V., Goebel, U.B., and Stackebrandt, E. (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* **21**: 213-229.
- Woese, C.R. (1987) Bacterial evolution. *Microbiol. Rev.* **51**: 221-271.
- Wu, L., Thompson D.K., Li, G., Hurt, R.A., Tiedje, J.M., and Zhou, J. (2001) Development and evaluation of functional gene arrays for detection of selected genes in the environment. *Appl. Environ. Microbiol.* **67**: 5780-5790.
- Wuyts, J., De Rijk, P., Van de Peer, Y., Winkelmans, T., and De Wachter, R. (2001) The European Large Subunit Ribosomal RNA Database. *Nucl. Acids. Res.* **29**: 175-177.
- Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P. *et al.* (2001) Global analysis of protein activities using proteome chips. *Science* **293**: 2101-2105.

- Zubkov, M.V., Fuchs, B.M., Archer, S.D., Kiene, R.P., Amann, R., and Burkill, P.H. (2001)
Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ. Microbiol.* **3**: 304-311.

Teil II
Publikationen

Publikationsliste mit Erläuterungen

Diese Dissertation basiert zu einem großen Teil auf drei Publikationen. Die Beiträge der verschiedenen Autoren zu den einzelnen Publikationen sind im Folgenden aufgeführt.

1. **Optimization strategies for DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes.**

Jörg Peplies, Frank Oliver Glöckner, Rudolf Amann.

Applied and Environmental Microbiology 69(3):1397-1407 (2003)

Entwicklung des Konzeptes und Durchführung der Experimente von Jörg Peplies. Erstellung des Manuskriptes von Jörg Peplies unter der redaktionellen Mitarbeit aller Koautoren.

2. **Application and validation of DNA microarrays for the identification of 16S rRNA targets in marine surface water.**

Jörg Peplies, Rudolf Amann, Stanley Chun Kwan Lau, Jakob Pernthaler,
Frank Oliver Glöckner

Submitted to Environmental Microbiology

Entwicklung des Konzeptes von Jörg Peplies und Frank Oliver Glöckner. Etablierung des Extraktions- und Markierungsprotokolles von Jörg Peplies und Stanley Lau. Durchführung aller weiteren Experimente von Jörg Peplies. Bildanalyse und statistische Auswertung von Jörg Peplies und Jakob Pernthaler. Erstellung des Manuskriptes von Jörg Peplies unter der redaktionellen Mitarbeit von Rudolf Amann und Frank Oliver Glöckner.

3. **Comparative sequence analysis and oligonucleotide probe design based on the 23S rRNA genes of alpha-proteobacteria from North Sea bacterioplankton.**

Jörg Peplies, Frank Oliver Glöckner, Rudolf Amann, Wolfgang Ludwig

In preparation

Entwicklung des Konzeptes von Jörg Peplies, Frank Oliver Glöckner und Rudolf Amann. Sequenzierung der 23S rRNA-Gene von Jörg Peplies. Bereitstellung des 23S rRNA-Datensatzes und -Alignments durch Wolfgang Ludwig. Rekonstruktion der Stammbäume von Jörg Peplies unter Mitarbeit von Frank Oliver Glöckner. Sondenentwicklung von Jörg Peplies. Erstellung des Manuskriptes von Jörg Peplies unter der redaktionellen Mitarbeit von Frank Oliver Glöckner und Rudolf Amann.

1

**Optimization strategies for DNA microarray-based detection of
bacteria
with 16S rRNA-targeting oligonucleotide probes**

Jörg Peplies, Frank Oliver Glöckner, Rudolf Amann

Applied and Environmental Microbiology 69(3):1397-1407 (2003)

Optimization Strategies for DNA Microarray-Based Detection of Bacteria with 16S rRNA-Targeting Oligonucleotide Probes

Jörg Peplies, Frank Oliver Glöckner,* and Rudolf Amann

*Department of Molecular Ecology, Max Planck Institute for Marine Microbiology,
28359 Bremen, Germany*

Received 13 September 2002/Accepted 10 December 2002

The usability of the DNA microarray format for the specific detection of bacteria based on their 16S rRNA genes was systematically evaluated with a model system composed of six environmental strains and 20 oligonucleotide probes. Parameters such as secondary structures of the target molecules and steric hindrance were investigated to better understand the mechanisms underlying a microarray hybridization reaction, with focus on their influence on the specificity of hybridization. With adequate hybridization conditions, false-positive signals could be almost completely prevented, resulting in clear data interpretation. Among 199 potential nonspecific hybridization events, only 1 false-positive signal was observed, whereas false-negative results were more common (17 of 41). Subsequent parameter analysis revealed that this was mainly an effect of reduced accessibility of probe binding sites caused by the secondary structures of the target molecules. False-negative results could be prevented and the overall signal intensities could be adjusted by introducing a new optimization strategy called directed application of capture oligonucleotides. The small number of false-positive signals in our data set is discussed, and a general optimization approach is suggested. Our results show that, compared to standard hybridization formats such as fluorescence in situ hybridization, a large number of oligonucleotide probes with different characteristics can be applied in parallel in a highly specific way without extensive experimental effort.

During the last decade, molecular methods based on comparative analysis of 16S rRNA sequences have yielded new and unexpected insights into the diversity of microbial communities (13, 19, 48). Although diversity can be readily studied by PCR-based 16S rRNA gene libraries, with this approach it is not possible to deduce the composition of the analyzed communities quantitatively (2). With the introduction of fluorescence in situ hybridization (FISH), it first became possible to identify single bacterial cells with labeled oligonucleotide probes targeting the rRNA of selected phylogenetic groups. This allows the quantitative analysis of the spatiotemporal composition of microbial communities (3).

A drawback of the FISH method is the limited number of probes which can be applied in one hybridization experiment. This limitation becomes a distinct bottleneck when this technique is used for community analysis on a high level of phylogenetic resolution. Additionally, there is a need for multiple probes to check for false-positive and false-negative results caused by individual probes used for identification of selected target organisms (2). A highly parallel application of multiple probe sets is facilitated by DNA chips or DNA microarrays. In this reverse hybridization format, matrix-immobilized oligo- or polynucleotides (probes) are used for the specific capture of labeled target molecules. This, in principle, allows the simultaneous application of a nearly unlimited number of probes in a single hybridization experiment (23, 26).

The concept of DNA microarray hybridization was introduced more than a decade ago (5, 14, 29, 31, 45), and the

technique has been applied for many years in the pharmaceutical industry (12), clinical diagnostics (25), and many fields of research, e.g., functional genomics (28) and genetic analysis (22). In contrast, DNA microarrays are still not common in microbial ecology. Only a limited number of reports have been published, which mainly showed the “proof of principle” of the method in this field of research (24, 41, 43). What are the reasons? In contrast to “standard” applications in the medical field, in microbial ecology defined nucleic acids have to be identified against an often large and partly unknown genetic background. For this kind of analysis, the discrimination of single mismatches is crucial (27), but the specific hybridization of target molecules to immobilized capture oligonucleotides is one of the major challenges of the DNA microarray approach because large sets of probes with different characteristics are applied under identical hybridization conditions.

Several approaches exist to overcome this problem. An example is the addition of tetramethylammonium chloride (32) or betaine (39) to the hybridization buffer. Both compounds equalize the melting points of different oligonucleotides by stabilization of AT base pairs. Another strategy is based on the acquisition of complete melting curves for every single probe spotted on the chip surface (27), which is a promising but methodically and technically demanding approach. Further methodological investigations and optimizations are needed before DNA microarrays can be used for the fast, simple, and efficient screening of microbial communities. Currently, the usefulness of standard microarray formats is often limited by hard to interpret signal patterns caused by an accumulated number of false-positive signals.

In 1994, Williams et al. (49) pointed out the complexity of a microarray hybridization event and the shortcoming of the

* Corresponding author. Mailing address: Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, 28359 Bremen, Germany. Phone: 49 421 2028 938. Fax: 49 421 2028 580. E-mail: fog@mpi-bremen.de.

TABLE 1. Oligonucleotide probes used in this study and their specificities^a

Probe	Specificity	Sequence (5'-3')	16S rRNA binding site ^b and length (nt)	GC content (%)	No. of mismatches to target ^c										Reference
					KT0202a	KT1117	JP7.1	JP13.1	KT11ds2	KT71					
UNIV 1392	Universal	ACGGGCGGTGTGTAC	1,392 (15)	67	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	37
EUB338	Bacteria	GCTGCCCTCCCGTAGAGT	338 (18)	67	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	3
CF319a	Cytophaga-Flavobacterium group of Bacteroidetes	TGTCGGTGTCTCAGTAG	319 (18)	56	2 (4,17)	2 (4,17)	2 (4,17)	2 (4,17)	2 (4,17)	2 (4,17)	2 (4,17)	2 (4,17)	2 (4,17)	2 (4,17)	30
CYT1448	KT11ds2	CTAGCGGCTCCCTACGG	1,448 (18)	67	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4	16
NORS-730	NORS cluster of gamma-proteobacteria	TCGAGCCAGGAGCGCC	730 (18)	78	3	3	2 (10,12)	>4	>4	>4	>4	>4	>4	>4	16
NORS-130	KT71	CCGACCTACTGGATAGAT	130 (18)	50	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4	16
ALP968	Alpha-proteobacteria	GTTAAGGTTCTGGCGTGT	968 (18)	56	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	34
MALF-1	Marine alpha cluster	GCGGGGTTCTTTACCA	488 (18)	56	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	21
ROSS37	Marine alpha cluster	CAAGCTTACCCCTCC	537 (17)	65	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	16
GRb	Marine alpha cluster	GTGAGTATCGAGCCAGTAG	735 (20)	55	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	20
RSB67	Subgroup of marine alpha cluster	CGCTCCACCGAAGGTAG	67 (18)	67	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	50
RC1031	KT0202a	ACCTGCACTATGTCGCG	1,031 (18)	56	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	16
RC1239	KT0202a	TAAGTCACTGTAGTTGCCAT	1,239 (20)	40	PM	1 (14)	1 (14)	4	2 (12,18)	4	4	4	4	4	16
KT13	Subgroup of marine alpha cluster	TAAGTCACTGTAGTTGCCAT	1,239 (20)	40	1 (14)	1 (14)	1 (14)	1 (14)	1 (14)	1 (14)	1 (14)	1 (14)	1 (14)	1 (14)	This study
KT13-231	KT1117	ATCTAATCAAGCGGGCC	231 (19)	53	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	16
Ros7-1029	JP7.1	CTGTCACTGTCTCTGTG	1,029 (18)	50	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4	16
RRP1088	Subgroup of alpha-proteobacteria	CGTTCGGGRCCTAACC	1,088 (17)	59	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	34
PARI457	Subgroup of alpha-proteobacteria	TACCGTGGTCCCGTCC	1,457 (18)	72	PM	PM	PM	PM	PM	PM	PM	PM	PM	PM	34
ALF4-1322	Subgroup of alpha-proteobacteria	TCGCGCTTCACTGCTCG	1,322 (18)	61	4	4	1 (2)	3	4	4	4	4	4	4	34
EPI29	Subgroup of alpha-proteobacteria	CGAACCTTAAAGGCGAGTT	129 (18)	50	>4	>4	>4	>4	>4	>4	>4	>4	>4	>4	34

^a For every probe, a control containing one central mismatch was spotted additionally, i.e., the whole probe set includes a total of 40 different oligonucleotides.

^b E. coli numbering according to Brosius et al. (10); position of 3' nucleotide of probe is stated, nt, nucleotide.

^c For up to two mismatches, the position of the mismatches is also stated (numbers in parentheses counted from the 3' end), PM, perfect match.

existing rules to describe it. In recent years, several systematic studies were conducted on different aspects, such as the secondary structures of target molecules and steric hindrance mediated by the solid support (8, 22, 33, 40, 42, 49), but the complex interaction of all these mechanisms and their influence on hybridization specificity are still mostly unknown. In this study, we systematically evaluated the suitability of the microarray format for 16S rRNA-based detection of bacteria by analyzing a model system including single-mismatch controls. Our objective was to get new insights into the mechanisms underlying a microarray hybridization reaction for the deduction of novel optimization strategies leading to signal patterns with a reduced number of false-positive and false-negative results.

MATERIALS AND METHODS

Bacterial strains. For the analysis of 16S rRNA gene (16S rDNA) amplicons by DNA microarray hybridization, six bacterial strains (KT0202a, KT1117, JP7.1, JP13.1, KT11ds2, and KT71) were randomly selected from a pool of environmental strains isolated from the surface water of the German Bight by Eilers et al. (15, 16). The accession numbers of the corresponding sequences are AF173971, AF235111, AF305498, AY007676, AY007679, and AY007680.

Preparation of fluorescently labeled target single-stranded DNA. Labeled target molecules were prepared by PCR amplification of the nearly complete 16S rRNA genes of the reference strains with the general bacterial primers 8F and 1492R (11). Living cells were picked from culture plates, resuspended in 65 µl of PCR-water (Sigma, Deisenhofen, Germany) and treated three times by freeze-thawing in liquid nitrogen for cell lysis. For PCR amplification, 10× PCR buffer (Eppendorf, Hamburg, Germany), 30 µg of bovine serum albumin, deoxyribonucleoside triphosphates (final concentration, 250 µM each), primers (final concentration, 150 nM each), and 2.5 U of MasterTaq polymerase (Eppendorf, Hamburg, Germany) were added to a final volume of 100 µl.

For subsequent preparation of fluorescently labeled, single-stranded DNA targets, 5'-indocarbocyanine-labeled forward primer 8f and 5'-biotin-labeled reverse primer 1492R were applied, both purchased from Thermo Hybaid (Interactiva Division, Ulm, Germany). Amplification was done according to the protocol of Buchholz-Cleven et al. (11) except that a temperature of 48°C was used for annealing. Successful amplification was confirmed by analyzing PCR products on 1.5% agarose gels. PCR products were purified with the QIAquick PCR purification kit (Qiagen, Hilden, Germany). Subsequently, the biotin-labeled DNA strands were extracted with streptavidin-coated paramagnetic beads (Boehringer, Mannheim, Germany) as described by Niemeyer et al. (36). The remaining, fluorescently labeled single-stranded DNAs are equivalent to the 16S rRNAs of the corresponding strains. Finally, the concentration of the amplicons was determined by UV spectrometry. The labeled single-stranded DNA was stored at -18°C.

Oligonucleotide probe set. Twenty oligonucleotide probes of 15 to 20 nucleotides in length, targeting the bacterial 16S rRNA and originally designed for FISH, were selected from the literature for reliable differentiation of the six model strains based on the multiple-probe approach (4). The probes and their characteristics are listed in Table 1. One of the probes (KT13) was designed within this study with the software package ARB (<http://www.arb-home.de>). Current specificity of all probes was evaluated by ARB with the rRNA database of the Technical University of Munich (<http://www.arb-home.de>, release 08/01). For all 20 probes, a control containing one single central mismatch based on a transversion was included.

Preparation of glass slides and spotting. The DNA microarray format used in this study is based on standard microscopic glass slides (Menzel, Braunschweig, Germany). Slides were activated by 1,4-phenylenediisothiocyanate treatment for covalent binding of either 5'- or 3'-amino-modified capture oligonucleotides (Thermo Hybaid, Ulm, Germany) according to Benters et al. (7).

Probes were spotted onto the activated slide surface with the piezo-driven spotting device Robodrop (BIAS, Bremen, Germany). The concentration of the amino-modified oligonucleotides in PCR-water was 10 µM with 1% glycerol. Volumes of deposited probe solutions were about 250 pl, resulting in spots with a diameter of approximately 200 µm.

To complete covalent binding, after being spotted with probe solutions, slides were incubated overnight at room temperature in a wet chamber to restrict evaporation of the spots. Blocking of the slides was performed in 6-amino-1-

hexanol (50 mM) and diisopropylethylamine (150 mM) in dimethylformamide as described by Beier et al. (6). Finally, the slides were washed with deionized, particle-free water, air dried, and stored under nitrogen at 4°C.

DNA microarray hybridization. For hybridization and washing of the microarrays, a standard FISH protocol according to Perntaler et al. (38) was used. Hybridization time was extended from 1.5 to 3 h. Incubation was performed either at room temperature, at 46°C without formamide, or at 46°C with 20% formamide. In all cases, 3.5 pmol of labeled target single-stranded DNA in a total volume of 200 μ l of hybridization buffer was applied to the microarrays. To guarantee a uniform moistening of the slide surface, the sample was covered with a coverslip. Unless otherwise specified, probes were applied for hybridization experiments in the standard way, i.e., all probes immobilized via their 5' end, applied without spacer, and at the same concentration (10 μ M).

To investigate the impact of secondary structures, unlabeled "helper" oligonucleotides (18), purchased from Thermo Hybaid (Ulm, Germany), were applied to the hybridization solution at a final concentration of 5 nM. For all probe binding sites analyzed by helper oligonucleotides, a pair of two perfectly matching 20-mer oligonucleotides was applied, one binding adjacent to the 5' end of the corresponding probe and one to the 3' end.

For reducing steric hindrance, capture probes carrying polyadenosine triphosphate spacers of 6, 12, 18, or 24 nucleotides, located at either the 5' end or the 3' end of the capture oligonucleotide, depending on selected probe orientation, were used.

Signal detection and data analysis. Air-dried slides were imaged at a resolution of 10 μ m with a GenePix4000 microarray scanner (Axon, Union City, Calif.) at the same laser power and sensitivity level of the photomultiplier for each slide. Therefore, absolute signal intensities (arbitrary units) presented for independent experiments should be directly comparable.

For automatic spot detection and signal quantification, the image analysis software MetaMorph Offline version 4.6 (Universal Imaging Corporation, West Chester, Pa.) was used. Filters for the automatic removal of spots with poor circularity or low uniformity of pixel intensities were established. Signals were considered positive if values were above zero after local background correction. This strategy for assessing the reliability of positive spots was possible because nonspecific binding of labeled target molecules to the slide surface was not observed for regions where oligonucleotides were deposited. This leads to "black holes" for spots where no detectable hybridization occurred, as shown in Fig. 6, and ultimately results in negative measurements for these spots, which were considered zero. We foresee no biases on individual spots by this method because in general we had homogeneous background noise. Otherwise, experiments were not considered.

All probes were deposited in eight replicate spots arranged in two blocks of four parallels. Each data point shown represents the arithmetic mean of the replicates analyzed for one probe. Error bars indicate the standard deviation for the analyzed replicates.

RESULTS

Specific detection of environmental bacterial strains with a redundant and hierarchically structured set of probes. For optimization of global hybridization conditions, different combinations of temperature and formamide concentration were tested with strain KT0202a: hybridization and washing at room temperature without formamide (Fig. 1A), at 46°C without formamide (Fig. 1B), and at 46°C with 20% formamide (Fig. 1C). Hybridization at room temperature led only to one weak true-positive signal for probe RC1031. Additionally, nonspecific (false-positive) signals for probes KT13-231 (one mismatch), ALF4-1322 (four mismatches), and ALF4-1322 M1c (three mismatches) could be observed. After hybridization at 46°C without formamide, 9 of 11 probes targeting strain KT0202a showed a true-positive signal, and two spots were false-negative. Signal intensities ranged from 117 to 1,982 arbitrary units (a.u.) for probes RRP1088 and GRb, respectively. Hybridization at 46°C and 20% formamide led to four true-positive signals, ranging from 107 a.u. for probe RSb67 to 674 a.u. for probe GRb, and seven false-negative results. All further experiments were done at 46°C without formamide.

A comprehensive overview of the independent analysis of all six target strains is shown in Fig. 2. The isolates could be clearly differentiated by the signal patterns obtained. For all probes, false-positive signals could be almost completely prevented, even within the controls containing only one central mismatch. Among 199 potential nonspecific hybridization events (20 probes plus 20 controls times 6 targets minus 41 perfect-match situations), only one false-positive signal (0.5%) was observed for probe GRb in combination with target KT1117. However, 17 of 41 expected signals were missing (41.5%). For example, the universal probe UNIV1392 was false-negative with all six targets. Four of six signals for probe EUB338 were just slightly above the detection limit, and the other two were false-negative.

Furthermore, we observed profound differences in the signal intensities of different probes hybridized to the same target. For example, with probe GRb, strain KT0202a had a signal that was 17 times higher than that with probe RRP1088.

Analysis of parameters influencing a microarray hybridization reaction. Strain KT0202a was used for all further systematic evaluations because it was targeted by the most probes of the probe set applied (compare Table 1).

(i) Secondary structures. With pairs of helper oligonucleotides, the signals for probes ALF968, ROS537, and RC1239 could be specifically increased (Fig. 3A; see also Fig. 6). Compared to the corresponding data generated without helper oligonucleotides (Fig. 1B), the probe ALF968 signal was raised from 0 to 1,126 a.u. The signal for probe ROS537 was increased 16-fold, from 219 to 3,511 a.u., whereas the signal of probe RC1239 only doubled from 775 to 1,670 a.u. Interestingly, application of helper pairs also resulted in signals within the mismatch controls ALF968M1c and ROS537M1c. Also probe KT13, which has the same binding site as probe RC1239 and shows only one central mismatch to strain KT0202a, could no longer be discriminated under these conditions. Signals of probes MALF-1, GRb, and RC1031 decreased from 819 to 160 a.u., from 1,982 to 394 a.u., and from 973 to 197 a.u., respectively, compared to the corresponding experiments without helpers (Fig. 1B).

Subsequently, for all 11 probes targeting KT0202a, helper pairs were applied in parallel (Fig. 3B). For all the probes, absolute signal intensities could be increased, ranging from a factor of 3.8 (from 775 to 2,968 a.u.) for probe RC1239 to a factor of 46.5 (from 184 to 8,565 a.u.) for probe PAR1457. With three exceptions (UNIV1392, RRP1088, and RC1239), all probes showed nonspecific signals within the corresponding mismatch controls. Additionally, probes KT13 and KT13-231 featured weak nonspecific signals (one mismatch for both).

(ii) General steric hindrance. Poly(A) spacers of different lengths (none, 6-mer, and 12-mer) were tested on probes EUB338 and RRP1088. Both probes showed a linear increase of signal intensities with increasing spacer length (Fig. 4A). For probe EUB338, we observed signals within the corresponding mismatch controls when applying a 6- or 12-mer spacer, whereas for probe RRP1088, the hybridization efficiency could be improved by spacer molecules without getting nonspecific signals.

Figure 4B shows the hybridization of strain KT0202a to probes ROS537 and ALF968 provided with no, 6-, 12-, 18-, and 24-mer spacers and under the presence of helper pairs for both

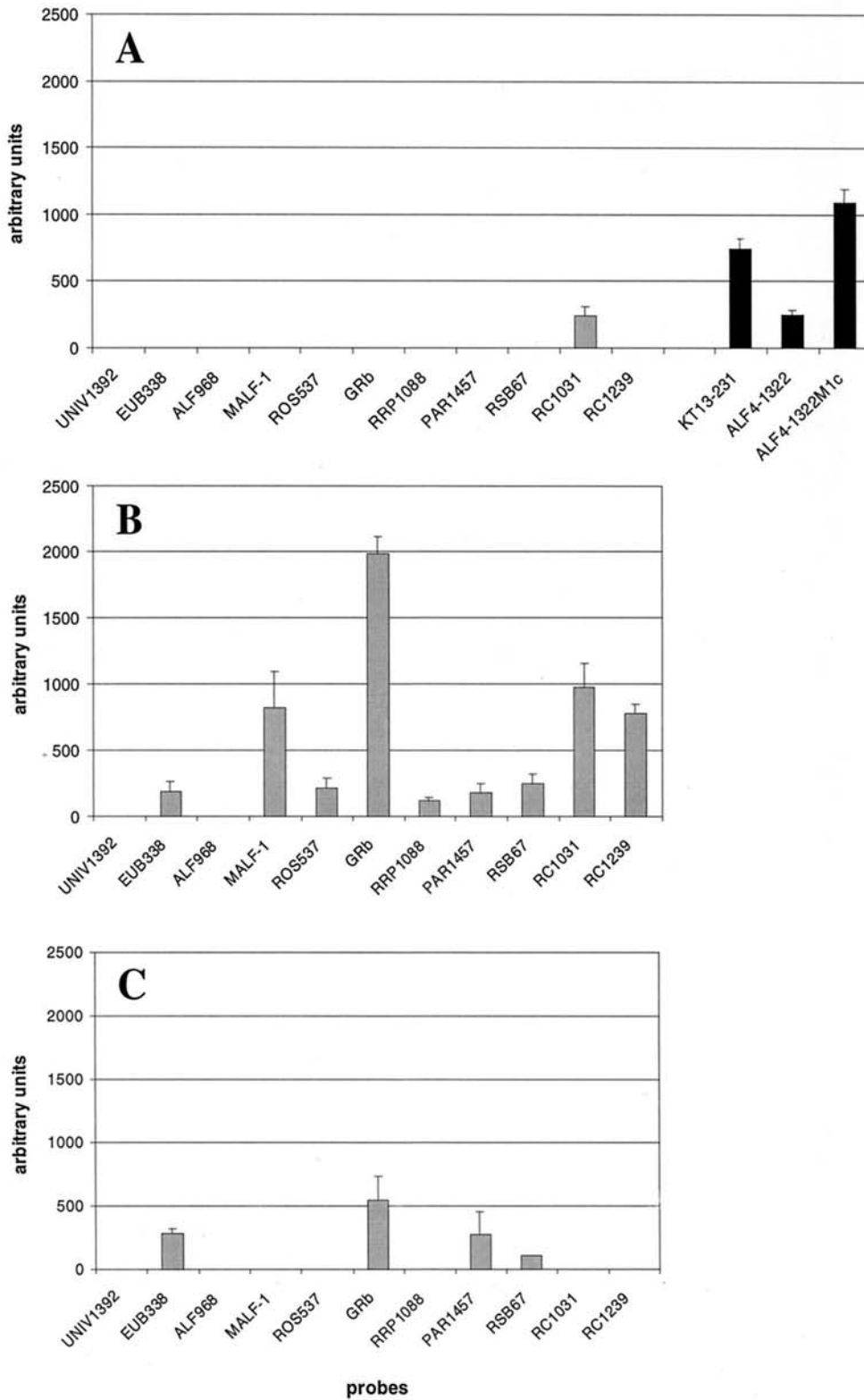


FIG. 1. Signal patterns after microarray hybridization of amplified 16S rDNA of strain KT0202a to the complete set of probes shown in Table 1 at room temperature without formamide (A), at 46°C without formamide (B), and at 46°C with 20% formamide (C). Only probes targeting strain KT0202a are shown. Black bars indicate false-positive signals (controls are only shown if signals were detected).

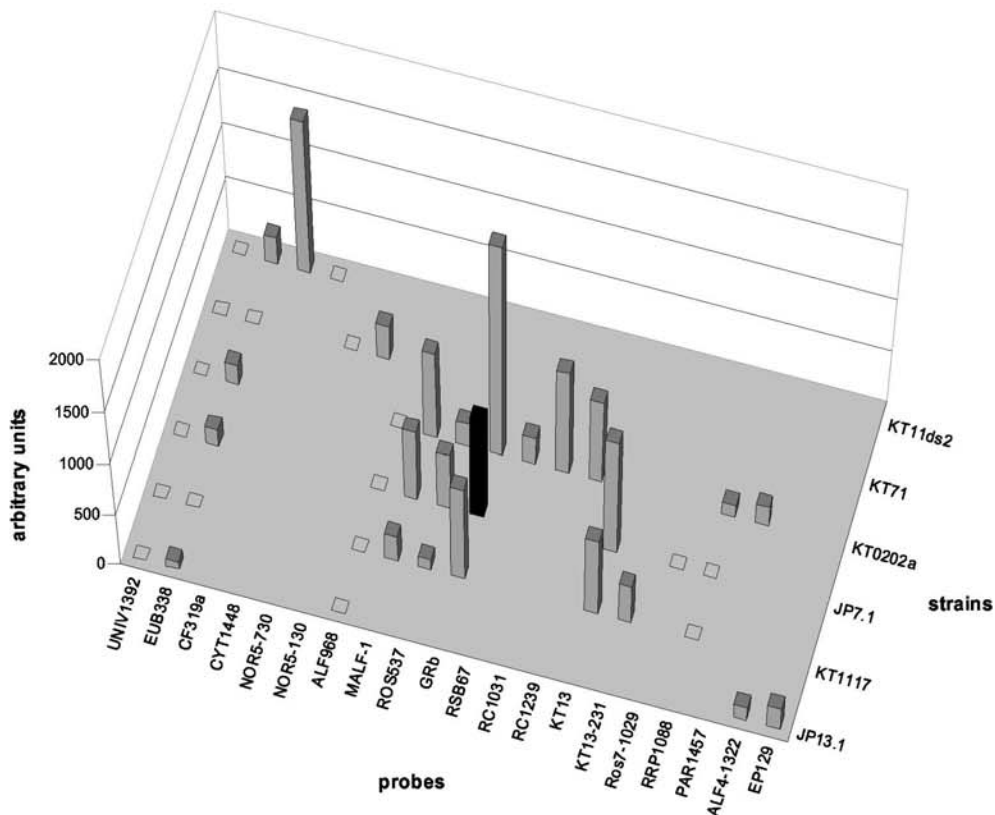


FIG. 2. Comprehensive overview of signal patterns after microarray hybridization of all six strains analyzed in this study under standard conditions (at 46°C without formamide, all probes 5' immobilized and applied without spacer). False-positive signals are indicated by black bars; false-negative signals are indicated by open rectangles. For clarity, the single-mismatch controls spotted for all 20 probes are not shown (here, no signals could be observed).

probes. Here, for probe ALF968, a saturation was reached with the 18-mer spacer and for probe ROS537 with the 6-mer spacer.

(iii) **Orientation of immobilized probe.** Probe orientation experiments with strain KT0202a (Fig. 4B) showed an increase in signal intensities after inversion (3' instead of 5' immobilized) from 3,716 to 6,763 a.u. (1.8-fold) and from 2,105 to 2,877 a.u. (1.4-fold) for probes ROS537 and ALF968, respectively. For the 12-mer spacer variants of the same two probes, increases of only about 1.1-fold were observed.

Optimization of signal patterns by directed application of capture oligonucleotides. By a directed modification of spacer length based on the signal intensities shown in Fig. 1B for strain KT0202a, probe-specific differences in signal intensity and number of false-negative results should be minimized. For example, the two probes featuring former false-negative signals, UNIV1392 and ALF968, were both tested with 12- and 18-mer spacers, whereas probe GRb was not modified. The results for the hybridization of the 16S rDNA of strain KT0202a with this approach are given in Fig. 5. By elongation of probe UNIV1392 with an 18-mer spacer, the earlier observed false-negative signal could be restored without getting a signal for the corresponding mismatch control. In addition,

signal intensities could be increased for probes MALF-1 (with a 6-mer spacer from 41 to 116%, normalized to the GRb signal), RRP1088 (12-mer, from 6 to 50%), PAR1457 (6-mer, from 9 to 25%), and RC1239 (6-mer, from 13 to 68%) without losing the specificity of the hybridization reaction. The sensitivity of probes EUB338, ALF968, ROS537, RSB67, and RC1031 could not be improved without the simultaneous appearance of signals within the corresponding controls.

DISCUSSION

In our initial experiments, we determined the optimal hybridization conditions for our microarray format in terms of a reduced number of false-positive and false-negative hybridization events. Best results were achieved for reference strain KT0202a at 46°C without formamide (Fig. 1B). Only under these conditions could a sufficient number of true-positive signals in combination with no false-positive signals be detected. This situation allows clear data interpretation. Concurrently, we demonstrated the general transferability of a standard FISH protocol to the DNA microarray format, which enables the easy validation of microarray data by parallel FISH analysis. This is crucial for the further application of microarrays

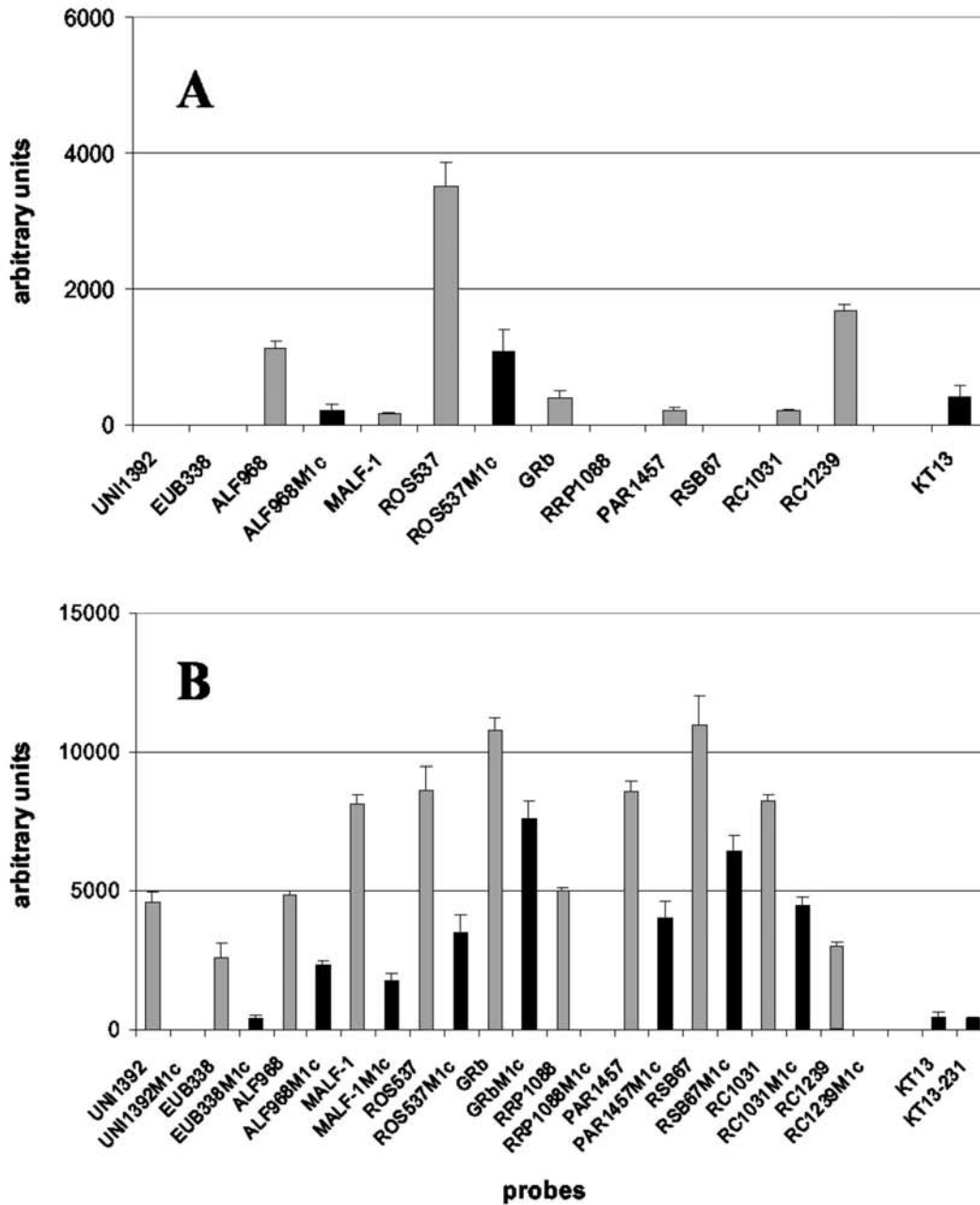


FIG. 3. Impact of secondary structures of target molecules on hybridization efficiency. Hybridizations were performed with amplified 16S rDNA of strain KT0202a under standard conditions. (A) For probes ALF968, ROS537, and RC1239, helper oligonucleotides were added. Controls are only shown if signals were detected. (B) For all 11 probes targeting strain KT0202a, helpers were added. All corresponding controls plus two additional probes also featuring false-positive signals are shown.

for environmental studies. The fact that the best hybridization results were achieved under moderate stringency is in good agreement with data obtained from FISH experiments (17). Presumably, the lack of hybridization signals at room temperature (Fig. 1A) is due to a more stable secondary structure of the target molecule at lower temperatures, resulting in reduced

accessibility of the probe binding sites. For hybridization at 46°C with 20% formamide (Fig. 1C), which equals a hybridization temperature of 56°C (47), we observed signals for only 4 of 11 perfect-match probes, presumably based on melting effects.

The results shown in Fig. 2 indicate that highly specific

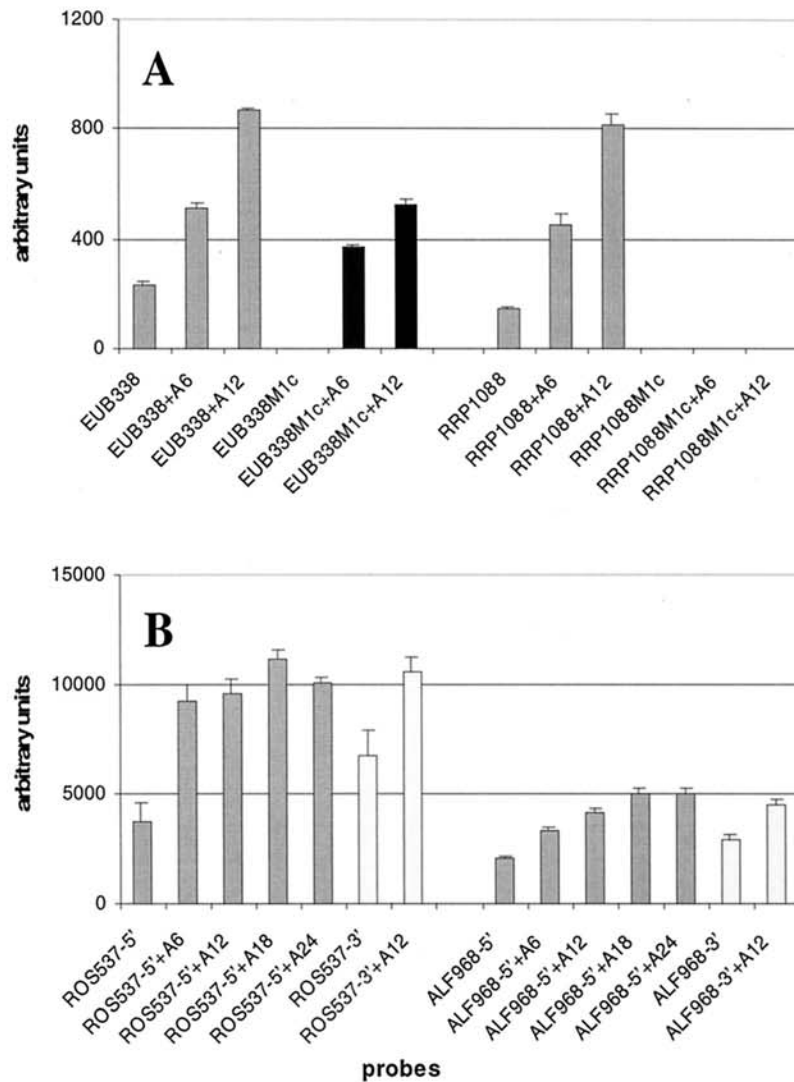


FIG. 4. Impact of steric hindrance. (A) Hybridization of the amplified 16S rDNA of strain KT0202a at 46°C without formamide to probes EUB338 and RRP1088, both provided with poly(A) spacers of multiple lengths (none, 6-mer, and 12-mer spacers). The corresponding controls are also shown. (B) Hybridization of strain KT0202a at 46°C without formamide to probes ROS537 and ALF968 in the presence of helper pairs for both probes. Probes were provided with spacers of multiple lengths (none, 6-mer, 12-mer, 18-mer, and 24-mer spacers). Probe variants without a spacer and with a 12-mer spacer were also provided by immobilization via the 3' end of the probe instead of the 5' end.

hybridization of a large number of oligonucleotide probes is feasible in a standard microarray format without any effort to adjust the melting points of the probes. Only one false-positive signal was observed among 199 potential nonspecific hybridization events (0.5%). This particular mismatch was one of the two weakest analyzed in this experimental series, a G:T mismatch three nucleotides away from the 3' end of the probe (see Table 1). Discrimination of a mismatch near the terminus of a short duplex is hard to achieve (46) even under optimized conditions. The fact that we did not find a single nonspecific hybridization event within the central mismatch controls spotted for all 20 probes seems to be in complete contradiction

with experiences from other hybridization formats such as FISH, where every probe has to be applied at appropriate stringency to ensure a specific hybridization reaction (4). Considering the results of the systematic investigations, the underlying mechanisms become more evident.

Parameter analysis. The addition of helper oligonucleotides had a clear impact on the signal intensities of particular probes (Fig. 3). Helpers are unlabeled oligonucleotides designed to bind adjacent to a probe binding site, resulting in an increase in the accessibility of the corresponding target site (18, 35). As shown for probes ALF968, ROS537, and RC1239, they can resolve the secondary structures of the target rDNAs in a

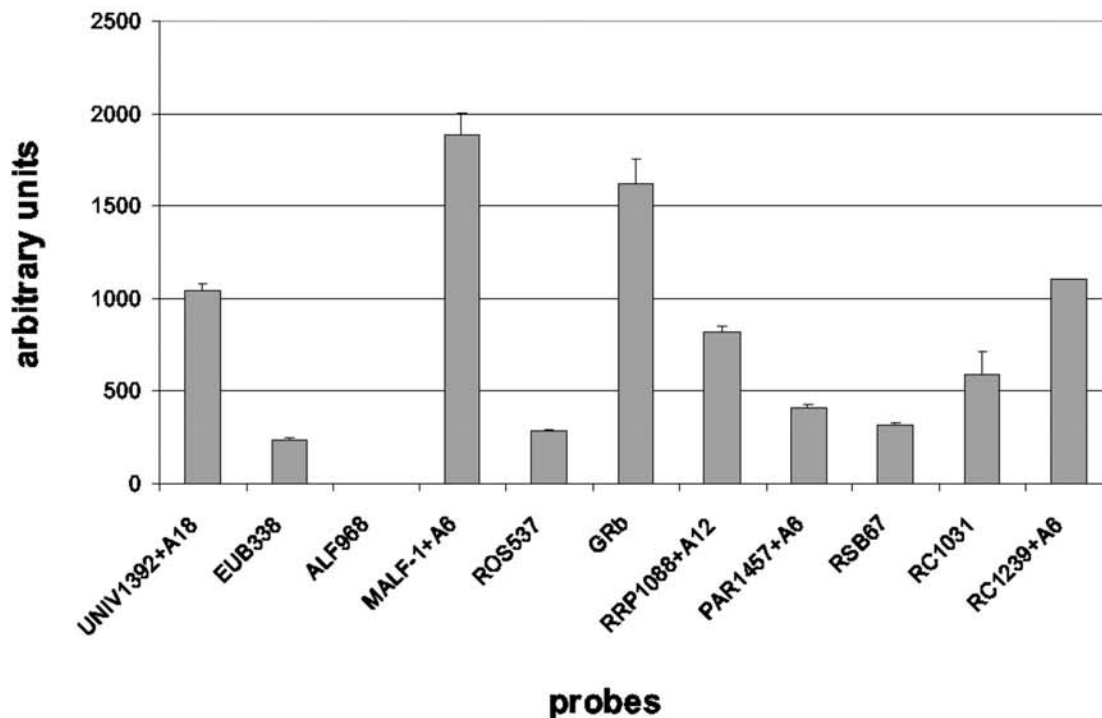


FIG. 5. Signal pattern optimization by directed application of capture oligonucleotides. Hybridization of the amplified 16S rDNA of strain KT0202a at 46°C without formamide to the 11 probes of the complete set targeting this strain. Probes were provided with spacers of different lengths, depending on former observed signal intensities (Fig. 1B, reference experiment). Only probe variants where no signals could be observed for the corresponding controls are shown.

directed and selective way for particular probe binding sites (Fig. 3A and 6). Interestingly, the former strong signals of probes MALF-1, GRb, and RC1031 decreased dramatically in this experiment, suggesting that the opening of selected probe binding sites leads to a reorganization of secondary structures in other target regions. The influence of secondary structures on hybridization efficiency becomes much clearer when we applied helper pairs for all 11 probes targeting strain KT0202a (Fig. 3B): former false-negative signals could be restored, overall signal intensities were dramatically increased, and most of the single mismatch controls now showed strong signals. We do not assume that this is based on a change of T_d values by base stacking of probe and helper oligonucleotides because no evidences for such an effect was found by Fuchs et al. (18).

Another outstanding feature of the data set shown in Fig. 2 is the distinct signal heterogeneity of different probes targeting the same preparation of 16S rDNA during parallel hybridization. This is a known phenomenon (44) and may even lead to the situation that signals of mismatch controls for particular probes were significantly higher than perfect match signals of other probes, as shown in Fig. 1A, 2, and 3. Considering the principles of DNA duplex stability prediction (9), measured signal intensities should be primarily connected to the nucleotide sequence of a probe. When applying helper oligonucleotides for all probes targeting an analyzed 16S rDNA, the distribution of measured signal intensities changes significantly (compare Fig. 1B and 3B). This observation is in agreement

with the common view that secondary structures of target molecules affect the duplex yield of particular probes selectively (44). In conclusion, the measured signal intensity for a particular probe does not directly correlate with the amount of the corresponding target molecule in solution.

Another parameter that has to be considered in the context of signal limitation is steric hindrance. When hybridizing on a solid support, the binding efficiency of target molecules may be reduced by unfavorable steric interactions mediated by the solid matrix (42). The size of this effect should be correlated with the size of the target nucleic acid and the distance between the support and the capture probe. To investigate the influence of steric hindrance in our system, we applied poly(A) spacers of different lengths (6-, 12-, 18-, and 24-mers) between the slide surface and the probe sequence. The disadvantage of this kind of spacer is of course the potential Watson-Crick interaction between the spacer and the target molecule, but in our system, we did not observe any hybridization events indicating such an effect. Without the addition of helper oligonucleotides, in most cases we found a linear correlation between spacer length and measured signal intensity (Fig. 4A; supplementary data are not shown). When adding helper oligonucleotides, signal saturation was reached even with the shorter spacer variants (Fig. 4B, probe ROS537 [6-mer] and probe ALF968 [18-mer]), indicating a complex interaction of the signal-determining parameters. Interestingly, signal enhance-

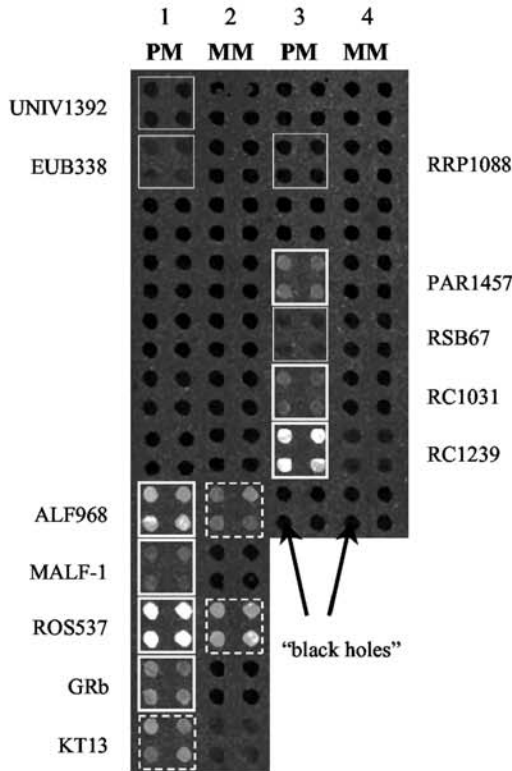


FIG. 6. Image of a scanned microarray taken from the experiment shown in Fig. 3A. Only four of the eight replicates spotted for each probe are shown (arranged in blocks). Columns 1 and 3 represent the 20 perfect-match probes, columns 2 and 4 show the corresponding mismatch controls. Highlighted spots are true-positive (solid line), false-positive (dashed line), or false-negative (dotted line) results.

ment by reduction of steric hindrance may also lead to signals within the corresponding controls (Fig. 4A, probe EUB338).

Furthermore, it must be assumed that the position of the probe binding site also has an effect on the extent of steric hindrance because every nucleic acid binds in a defined orientation to its complement. To address the significance of this parameter, we tested two variants of probes immobilized via the 5' and 3' end. Probe ROS537, whose binding site is located for target KT0202a at nucleotide position 474 of 1428, starting from the 5' end, shows a higher signal for the 3'-immobilized probe. This is in good agreement with our initial assumption that a 5'-immobilized probe binding to the 5' end of the target molecule is leading to an unfavorable steric situation because the longer tail of the target nucleic acid is oriented towards the slide surface and vice versa. Nevertheless, probe ALF968 also shows a slightly increased signal for the 3'-immobilized probe, although its binding site for target KT0202a is located at nucleotide position 907 of 1428. Apparently, the secondary and tertiary structures of the target molecules interfere with simple spatial extrapolations. As for application of spacer molecules, we found little influence on signal intensities for reverse probe immobilization compared to the application of helper oligonucleotides. In conclusion, the secondary structure of a target

molecule is the main limiting factor for duplex yield considering the parameters investigated in this study.

Mechanisms enabling observed hybridization specificity. As shown in Fig. 3 and 4A, signals for the single-mismatch controls can be "generated" under identical hybridization conditions if the influence of signal-limiting parameters such as secondary structures and steric hindrance is reduced. Therefore, we conclude that for the data set shown in Fig. 2, mismatch discrimination was mainly achieved by global signal suppression. This is possible because perfect match and mismatch signals react different on changes of signal limiting parameters as shown for secondary structures: without helpers (Fig. 1B), one mismatch signals were selectively shifted below the detection limit compared to the perfect-match signals.

A full-length 16S rDNA molecule and the resulting secondary structure appear to represent a well-suited "calibrator" for global signal suppression under optimized hybridization conditions. When shorter fragments were hybridized under identical hybridization conditions (data not shown), we also observed an overall increase of signal intensities in combination with an accumulated number of false-positive signals. Here, the reduced influence of two factors becomes evident: (i) shorter fragments lead to reduced steric hindrance and (ii) secondary structures of shorter targets are known to be less stable than those of longer fragments (44). Therefore, our results suggest that fragmentation of target molecules can be disadvantageous in terms of preventing false-positive signals.

In summary, we report on a new concept for improving the specificity of microarray hybridizations. It is based on a converse approach compared to current signal enhancement attempts because we "utilize" signal-limiting parameters to selectively shift nonspecific signals below the detection limit. We would like to point out that the overall sensitivity of the detection system is not affected by this strategy because the detection limit is not shifted but hybridization signals are specifically suppressed.

Prevention of false-negative results by a new optimization approach. As shown in Fig. 2, we found a large number of false-negative results in our data set. Subsequent application of helper pairs for all 11 probes targeting strain KT0202a under identical hybridization conditions led to signals for all 11 probes and nearly all corresponding controls (Fig. 3B). This indicates that false-negative results and discriminated single-mismatch controls were not based on melting effects. We suppose that nonoccurring signals are mainly caused by signal-suppressing parameters such as secondary structure of the target molecules and steric hindrance, selectively acting for different probe binding sites. Considering the fact that the number of 16S rRNA sequence positions is limited, a false-negative results for a particular probe can be a profound restriction for phylogenetic affiliation because additional probes featuring a redundant target specificity are often not available. For example, strain JP7.1 could be clearly assigned to the marine alpha group by the phylogenetically redundant probes MALF-1 and ROS537 (Fig. 2), but the single strain-specific probe available (Ros7-1029) featured a false-negative signal and therefore inhibited a more exact phylogenetic affiliation of the analyzed 16S rDNA.

Based on our systematic investigations we have deduced an approach for further signal pattern optimization called di-

rected application of capture oligonucleotides. Our study shows that (i) single-mismatch discrimination can also be achieved by "proper calibration" of signal intensities for particular probes and (ii) different parameters are available for a probe-specific signal adjustment. As shown in Fig. 3A and 4A, signals can be selectively enhanced by reducing the influence of signal suppressing factors such as secondary structures and steric hindrance. Our intention was to restore former false-negative signals and to adjust overall signal intensities with these options without losing the status of complete discrimination for the single mismatch controls, as shown in Fig. 4A for probe RRP1088. The usability of this approach, based on a directed variation of spacer length, was successfully tested for strain KT0202a, as shown in Fig. 5. This parameter was chosen because it allows a graduated signal adjustment. Another potential mediator would be the amount of immobilized probe, because a direct correlation to duplex yield was also observed for this parameter (data not shown), in agreement with the results of Guo et al. (22). In the context of directed application, the utilization of helper oligonucleotides is not practical for two reasons: (i) the strong signal increase caused by helpers often also results in signals for the corresponding mismatch controls (Fig. 3) and (ii) design of helpers showing the same specificity as the corresponding probe is often not possible (18).

Data normalization. An important issue when analyzing microarray experiments is data normalization. In the field of microbial ecology published data were either not normalized (24, 43) or domain specific probes were used for normalization (27). For this study we decided to work with absolute signal intensities (arbitrary units) when comparing data generated under different hybridization conditions or analyzing different target molecules. We assume that by using domain-specific probes as internal standards, an additional error will be introduced, because probe binding site accessibility should not be completely comparable for distantly related targets.

Why use a model system based on 16S rDNA targets? This study was performed with 16S rDNA targets generated by PCR amplification. We are well aware of the possible biases introduced by PCR during analysis of complex communities (3). For the analysis of complex samples, direct extraction and labeling of the rRNA are essential to circumvent possible artifacts. Furthermore, for environmental studies, aspects such as sensitivity and quantification have to be considered more comprehensively. However, we decided to conduct this initial systematic investigation with exactly defined 16S rDNA targets instead of extracted RNA because only single targets were analyzed in our model system. Our results should help to consider new kinds of optimization strategies for the specific detection of 16S rRNA/rDNA molecules by microarray hybridization and for environmental studies. We foresee PCR-based microarray systems in the (i) fast and low-cost screening of isolates for their phylogenetic affiliation and (ii) prescreening of PCR products for diversity or the presence of particular subpopulations (e.g., for generating gene libraries).

ACKNOWLEDGMENTS

We thank D. Blohm for providing the spotting device, D. Drutschmann for help in handling the device, and D. Wöhrle and B. Meyer-Schlosser for providing us with activated glass slides. The generous

support of J. Pernthaler by setting up the image and data analysis pipeline is gratefully acknowledged.

This work was done in the framework of the Centre of Applied Gensensorik (CAG) at the University of Bremen and supported by the German Ministry of Education and Research (BMBF, contract 0311833A) and by the Max Planck Society.

REFERENCES

- Amann, R. I., L. Krumholz, and D. A. Stahl. 1990. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *J. Bacteriol.* **172**:762–770.
- Amann, R. I. 1995. Fluorescently labeled, rRNA-targeted oligonucleotide probes in the study of microbial ecology. *Mol. Ecol.* **4**:543–554.
- Amann, R. I., W. Ludwig, and K.-H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**:143–169.
- Amann, R. I., and K.-H. Schleifer. 2001. Nucleic acid probes and their application in environmental microbiology, p. 67–82. In G. Garrity, D. R. Boone, and R. W. Castenholz (ed.), *Bergey's manual of systematic bacteriology*, 2nd ed., vol. 1. Springer-Verlag, New York, N.Y.
- Bains, W., and G. C. Smith. 1988. A novel method for nucleic acid sequence determination. *J. Theor. Biol.* **135**:303–308.
- Beier, M., and D. Hoheisel Jörg. 1999. Versatile derivatisation of solid support media for covalent bonding on DNA-microchips. *Nucleic Acids Res.* **27**:1970–1977.
- Benters, R., C. M. Niemeyer, D. Drutschmann, D. Blohm, and D. Wöhrle. 2002. DNA microarrays with PAMAM dendritic linker systems. *Nucleic Acids Res.* **30**:e10.
- Blake, R. D., and S. G. Delcourt. 1996. Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Res.* **24**:2095–2103.
- Breslauer, K. J., R. Frank, H. Blocker, and L. A. Marky. 1986. Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83**:3746–3750.
- Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller. 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J. Mol. Biol.* **148**:107–127.
- Buchholz-Cleven, B., B. Rattunde, and K. Straub. 1997. Screening for genetic diversity of isolates of anaerobic Fe(II)-oxidizing bacteria with DGGE and whole-cell hybridization. *Syst. Appl. Microbiol.* **20**:301–309.
- Debouck, C., and P. N. Goodfellow. 1999. DNA microarrays in drug discovery and development. *Nat. Genet.* **21**:48–50.
- DeLong, E. F. 1992. *Archaea* in coastal marine environments. *Proc. Natl. Acad. Sci. USA* **89**:5685–5689.
- Drmanac, R., I. Labat, I. Brukner, and R. Crkvenjakov. 1989. Sequencing of megabase plus DNA by hybridization theory of the method. *Genomics* **4**:114–128.
- Eilers, H., J. Pernthaler, F. O. Glöckner, and R. Amann. 2000. Culturability and in situ abundance of pelagic bacteria from the North Sea. *Appl. Environ. Microbiol.* **66**:3044–3051.
- Eilers, H., J. Pernthaler, J. Peplies, F. O. Glöckner, G. Gerdt, and R. Amann. 2001. Isolation of novel pelagic bacteria from the German bight and their seasonal contributions to surface picoplankton. *Appl. Environ. Microbiol.* **67**:5134–5142.
- Fuchs, B. M., G. Wallner, W. Beisker, I. Schwiippel, W. Ludwig, and R. Amann. 1998. Flow cytometric analysis of the in situ accessibility of *Escherichia coli* 16S rRNA for fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **64**:4973–4982.
- Fuchs, B. M., F. O. Glöckner, J. Wulf, and R. Amann. 2000. Unlabeled helper oligonucleotides increase the in situ accessibility of 16S rRNA for fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **66**:3603–3607.
- Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field. 1990. Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**:60–63.
- Giuliano, L., M. De Domenico, E. De Domenico, M. G. Höfle, and M. M. Yakimov. 1999. Identification of culturable oligotrophic bacteria within naturally occurring bacterioplankton communities of the Ligurian Sea by 16S rRNA sequencing and probing. *Microb. Ecol.* **37**:77–85.
- Gonzalez, J. M., W. B. Whitman, R. E. Hodson, and M. A. Moran. 1996. Identifying numerically abundant culturable bacteria from complex communities: an example from a lignin enrichment culture. *Appl. Environ. Microbiol.* **62**:4433–4440.
- Guo, Z., R. A. Guilloyle, A. J. Thiel, R. Wang, and L. M. Smith. 1994. Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.* **22**:5456–5465.
- Gupta, P. K., J. K. Roy, and M. Prasad. 1999. DNA chips, microarrays and genomics. *Curr. Sci.* **77**:875–884.
- Guschin, D. Y., B. K. Mobarry, D. Proudnikov, D. A. Stahl, B. E. Rittmann, and A. D. Mirzabekov. 1997. Oligonucleotide microchips as biosensors for determinative and environmental studies in microbiology. *Appl. Environ. Microbiol.* **63**:2397–2402.
- Hacia, J. G., L. C. Brody, M. S. Chee, S. P. A. Fodor, and F. S. Collins. 1996.

- Detection of heterozygous mutations in BRCA1 with high density oligonucleotide arrays and two-colour fluorescence analysis. *Nat. Genet.* **14**:441–447.
26. Hoheisel, J. D. 1997. Oligomer-chip technology. *Trends Biotechnol.* **15**:465–469.
 27. Liu, W.-T., A. D. Mirzabekov, and D. A. Stahl. 2001. Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. *Environ. Microbiol.* **3**:619–629.
 28. Lockhart, D. J., H. Dong, M. C. Byrne, M. T. Follettie, M. V. Gallo, M. S. Chee, M. Mittmann, C. Wang, M. Kobayashi, H. Horton, and E. L. Brown. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14**:1675–1680.
 29. Lysov, Y. P., V. L. Florentev, A. A. Khorlin, K. R. Khrapko, V. V. Shik, and A. D. Mirzabekov. 1988. Determining the nucleotide sequence of DNA by hybridization with oligonucleotides a new method. *Dokl. Akad. Nauk. SSSR* **303**:1508–1511. (In Russian.)
 30. Manz, W., R. Amann, W. Ludwig, M. Vancanneyt, and K.-H. Schleifer. 1996. Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum *Cytophaga-Flavobacter-Bacteroides* in the natural environment. *Microbiology* **142**:1097–1106.
 31. Maskos, U., and E. M. Southern. 1992. Parallel analysis of oligodeoxyribonucleotide-oligonucleotide interactions. I. Analysis of factors influencing oligonucleotide duplex formation. *Nucleic Acids Res.* **20**:1675–1678.
 32. Maskos, U., and E. M. Southern. 1993. A study of oligonucleotide reassociation with large arrays of oligonucleotides synthesised on a glass support. *Nucleic Acids Res.* **21**:4663–4669.
 33. Mir, K. U., and E. M. Southern. 1999. Determining the influence of structure on hybridization with oligonucleotide arrays. *Nat. Biotechnol.* **17**:788–792.
 34. Neef, A. 1997. Ph.D. thesis. Technical University, Munich, Germany.
 35. Niemeyer, C. M., W. Bürger, and J. Peplies. 1998. Covalent DNA-streptavidin conjugates as building blocks for novel biometallic nanostructures. *Angew. Chem. Int. Ed.* **37**:2265–2268.
 36. Niemeyer, C. M., L. Boldt, B. Ceyhan, and D. Blohm. 1999. Evaluation of single-stranded nucleic acids as carriers in the DNA-directed assembly of macromolecules. *J. Biomol. Struct. Dyn.* **17**:527–538.
 37. Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen. 1986. The use of rRNA sequences to characterize natural microbial populations. *Adv. Microb. Ecol.* **9**:1–55.
 38. Pernthaler, J., F. O. Glöckner, W. Schönhuber, and R. Amann. 2001. Fluorescence in situ hybridization (FISH) with rRNA-targeted oligonucleotide probes. *Methods Microbiol.* **30**:207–226.
 39. Rees, W. A., T. D. Yager, J. Korte, and P. H. von Hippel. 1993. Betaine can eliminate the base pair composition dependence of DNA melting. *Biochemistry* **32**:137–144.
 40. Riccelli, P. V., F. Merante, K. T. Leung, S. Bortolin, R. L. Zastawny, R. Janeczko, and A. S. Benight. 2001. Hybridization of single-stranded DNA targets to immobilized complementary DNA probes: comparison of hairpin versus linear capture probes. *Nucleic Acids Res.* **29**:996–1004.
 41. Rudi, K., O. M. Skulberg, R. Skulberg, and K. S. Jakobsen. 2000. Application of sequence-specific labeled 16S rRNA gene oligonucleotide probes for genetic profiling of cyanobacterial abundance and diversity by array hybridization. *Appl. Environ. Microbiol.* **66**:4004–4011.
 42. Shechinov, M. S., S. C. Case-Green, and E. M. Southern. 1997. Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. *Nucleic Acids Res.* **25**:1155–1161.
 43. Small, J., D. R. Call, F. J. Brockman, T. M. Straub, and D. P. Chandler. 2001. Direct detection of 16S rRNA in soil extracts by with oligonucleotide microarrays. *Appl. Environ. Microbiol.* **67**:4708–4716.
 44. Southern, E., K. Mir, and M. Shechinov. 1999. Molecular interactions on microarrays. *Nat. Genet.* **21**:5–9.
 45. Southern, E. M., U. Maskos, and J. K. Elder. 1992. Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides evaluation with experimental model. *Genomics* **13**:1008–1017.
 46. Stahl, D. A., and R. Amann. 1991. Development and application of nucleic acid probes, p. 205–248. *In* E. Stackebrandt and M. Goodfellow (ed.), *Nucleic acid techniques in bacterial systematics*. John Wiley & Sons Ltd., Chichester, United Kingdom.
 47. Wahl, G. M., S. L. Berger, and A. R. Kimmel. 1987. Molecular hybridization of immobilized nucleic acids: theoretical concepts and practical considerations. *Methods Enzymol.* **152**:399–407.
 48. Ward, D. M., R. Weller, and M. M. Bateson. 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**:63–65.
 49. Williams, J. C., S. C. Case-Green, K. U. Mir, and E. M. Southern. 1994. Studies of oligonucleotide interactions by hybridisation to arrays: the influence of dangling ends on duplex yield. *Nucleic Acids Res.* **22**:1365–1367.
 50. Zubkov, M. V., B. M. Fuchs, S. D. Archer, R. P. Kiene, R. Amann, and P. H. Burkhill. 2001. Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ. Microbiol.* **3**:304–311.

2

**Application and validation of DNA microarrays
for the identification of 16S rRNA targets
in marine surface waters**

Jörg Peplies, Rudolf Amann, Stanley Chun Kwan Lau,
Jakob Pernthaler, Frank Oliver Glöckner

submitted to Environmental Microbiology

Application and validation of DNA microarrays for the identification of 16S rRNA targets in marine surface water

JÖRG PEPLIES,¹ RUDOLF AMANN,¹ STANLEY CHUN KWAN LAU,²
JAKOB PERNTHALER,¹ and FRANK OLIVER GLÖCKNER^{1*}

Max Planck Institute for Marine Microbiology, Department of Molecular Ecology,

28359 Bremen, Germany¹

Hong Kong University of Science and Technology, Department of Biology,

Clear Water Bay, Kowloon, SAR Hong Kong²

The applicability of an oligonucleotide probe-based DNA microarray for the direct detection of 16S rRNA in samples of different complexity was systematically evaluated in terms of specificity, sensitivity, and quantification. Results obtained for samples of unknown composition were validated by quantitative fluorescence in situ hybridization (FISH). For extraction and direct labeling of total RNA, a fast and efficient protocol was established. Two pure cultures could be clearly differentiated and phylogenetically affiliated by the signal patterns obtained. Among 65 potential nonspecific hybridization events, no false-positive signal was observed. In a co-culture, abundant members could specifically be detected by microarray hybridization. Signal ratios of positive probes did not correlate well with quantitative data from FISH experiments, suggesting that microarray data provide, at best, semiquantitative information on community composition. False-negative results could be prevented by the directed application of capture oligonucleotides. In a marine picoplankton sample, bacterial subpopulations with relative abundances of at least 5% could specifically be detected by hybridizing 0.1 µg of total RNA extracted from a sample of 375 ml with a total number of 4.1×10^8 cells. Our results show that bacterial communities can be analyzed on DNA microarrays in a fast and easy way.

With the introduction of molecular methods based on comparative analysis of 16S rRNA gene sequences in the late 1980s (Olsen *et al.*, 1986), environmental microbiology has entered a new age. The characterization of microbial communities has overcome the requirement for prior cultivation and PCR-based methods have enormously affected our understanding of microbial diversity. One of the latest and most powerful tools in molecular biology is the DNA microarray technology which represents a high throughput format for the application of nucleic acid probes by reverse hybridization. Although, the DNA microarray technique is routinely applied since many years in clinical diagnostics, functional genomics and genetic analysis (Lander, 1999), and the general transferability of this format to the detection and identification of microorganisms was already shown in 1997 (Guschin *et al.*, 1997), there are still various issues hampering the broad application of DNA microarrays in the field of microbial community analysis. So far, only a limited number of studies have been published, in which microarrays were used for the parallel detection of 16S rRNAs or amplified 16S rRNA genes from the environment (Rudi *et al.*, 2000; Small *et al.*, 2001; Koizumi *et al.*, 2002; Loy *et al.*, 2002; Wilson *et al.*, 2002).

The major challenge of the microarray approach is to prevent nonspecific duplex formation leading to biased signal patterns (Hacia, 1999). This issue becomes potentially limiting when microarrays are used for community studies in environmental microbiology where distinct nucleic acids have to be identified against an often large and unknown genetic background. Therefore, single mismatch discrimination is required (Liu *et al.*, 2001). Various studies have focused on this aspect, but hybridization protocols and optimization strategies are still quite diverse (Liu *et al.*, 2001; Dai *et al.*, 2002; Religio *et al.*, 2002; Urakawa *et al.*, 2002; Peplies *et al.*, 2003).

Sensitivity of hybridization becomes crucial when analyzing communities of high complexity. Abundances of single target species or target groups of interest can be quite low and RNA extraction yields from environmental samples are often low, hampering microarray analysis of environmental microbial communities (Koizumi *et al.*, 2002). This limitation can be circumvented by the introduction of an initial PCR amplification step with, e.g., general 16S rRNA primer. Concomitantly, it provides an easy way for labeling the target molecules. However, community shifts must be expected by a prior PCR step (Suzuki and Giovannoni, 1996; Wintzingerode *et al.*, 1997; Marchesi *et al.*, 1998; Schmalenberger *et al.*, 2001). Consequently, for an unbiased view on community composition, direct labeling and hybridization of extracted cellular rRNA is a prerequisite (Small *et al.*, 2001).

To assess the reliability of microarray data obtained from complex microbial communities, convenient methods and strategies for their validation are needed. Whereas inclusion of internal controls for compensating spot-to-spot and slide-to-slide variations is common (Cho and Tiedje, 2002; Relogio *et al.*, 2002; Dennis *et al.*, 2003) and various studies have focused on processing of raw data (Brody *et al.*, 2002; Tran *et al.*, 2002), the validation of the obtained microarray results by independent techniques is often neglected.

In this study, we systematically evaluated and validated results obtained from DNA microarray analysis based on direct detection of bacterial 16S rRNA for samples of increasing complexity. Total RNAs from pure cultures, a co-culture of six environmental strains, and a picoplankton sample from German Bight surface waters were hybridized against an oligonucleotide probe set. Microarray data obtained from samples with unknown community composition were compared to parallel quantitative FISH counts.

RESULTS

Extraction and labeling of total RNA from samples of different complexity. The applied RNAwiz extraction protocol provided intact 16S rRNA from 50 ml of the pure cultures and 100 ml of the co-culture with extraction yields ranging from 30 to 75 μg of total RNA, allowing 20 to 100 hybridization reactions in our microarray format after labeling. Extraction yields from 750 ml of the picoplankton sample were significantly lower, restricting the number of microarray experiments to one or two hybridizations per preparation.

Fractions of the 16S rRNA within the different total RNA preparations varied between 15 and 45% for all types of samples. Only extractions with a 16S rRNA fraction between 20 and 30% were further processed, to guarantee comparability of the data obtained under quantitative aspects. For RNA extraction from the picoplankton sample, also a standard protocol according to Oelmüller *et al.* (1990) was tested, but no differences in total extraction yield, quality, or fraction size of the 16S rRNA were observed.

The relative labeling efficiency of the target molecules was similar for total RNA from the different types of samples. Spectroscopic data suggested one label every 26 to 41 nucleotides in all extracts analyzed.

Identification of pure cultures by microarray hybridization. The labeled total rRNA of two environmental strains, KT0202a (alpha-proteobacteria) and KT11ds2 (*Cytophaga-Flavobacterium* group of *Bacteroidetes*), was hybridized against a hierarchical and redundant

set of probes targeting the 16S rRNA (Tab. 1) to verify the suitability of our microarray format for a direct detection of extracted ribosomal RNA from bacterial cells. Specific signals were observed and both targets could be clearly differentiated by the signal patterns obtained. Strain KT11ds2 shows 3 out of 4 expected signals (1 false-negative result for probe UNIV1302) (Fig. 1A) and strain KT0202a shows 9 out of 11 expected signals (two false-negative results for probes UNIV1392 and RRP1088) (Fig. 1B). Among 36 potential unspecific hybridization events for strain KT11ds2 and 29 for strain KT0202a, respectively, not a single false-positive signal could be observed including all controls with only one central mismatch.

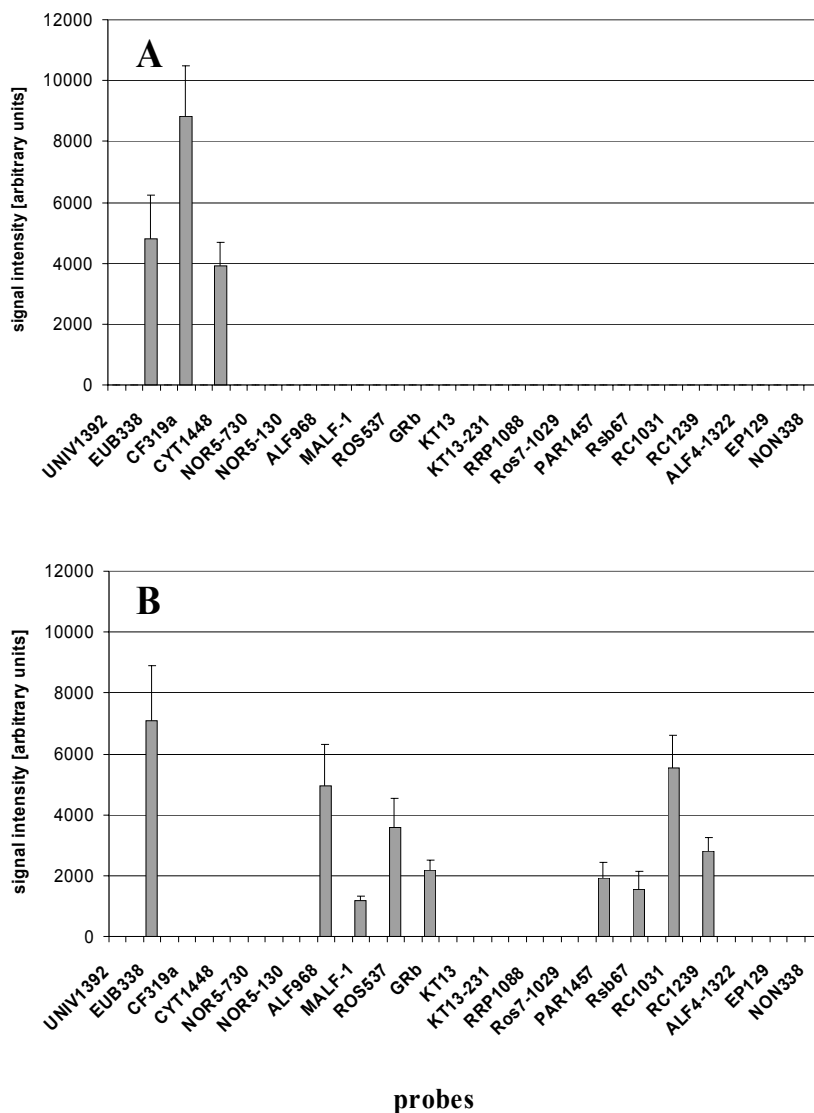


Figure 1: Signal patterns obtained by microarray hybridization of extracted total RNAs from strains KT11ds2 (A) and KT0202a (B), to the complete set of probes shown in Table 1 for 3 h at a target concentration of 8 ng/ μ l. Both experiments were scanned at a laser power of 70% to ensure comparability of the data sets. For clarity, single mismatch controls are not shown (here, no unspecific signals occurred).

TABLE 1. Oligonucleotide probes used in this study and their specificities^a

Probe	Specificity	Sequence (5'-3')	16S rRNA binding site ^b and length (nt)	GC content (%)	No. of mismatches to target ^c								Reference
					KT0202a	KT1117	JP7.1	JP13.1	KT11ds2	KT71			
UNIV1392	universal	ACGGGGGTGTGTAC	1392 (15)	67	PM	PM	PM	PM	PM	PM	PM	PM	Pace <i>et al.</i> (1986)
EUB338	Bacteria	GCTGCCTCCCGTAGGAGT	338 (18)	67	PM	PM	PM	PM	PM	PM	PM	PM	Amann <i>et al.</i> (1995)
CF319a	<i>Cytophaga-Flavobacterium</i> group of <i>Bacteroidetes</i>	TGGTCCGTGCTCAGTAC	319 (18)	56	2 (4,17)	2 (4,17)	2 (4,17)	2 (4,17)	PM	3	PM	3	Manz <i>et al.</i> (1996)
CYT1448	KT11ds2	CTAGGCCGCTCTTACGG	1448 (18)	67	>4	>4	>4	>4	PM	>4	PM	>4	Eilers <i>et al.</i> (2001)
NOR5-730	NOR5 cluster of gamma-proteobacteria	TCGAGCCAGGAGGCCGCC	730 (18)	78	3	3	2 (10,12)	>4	>4	PM	>4	PM	Eilers <i>et al.</i> (2001)
NOR5-130	KT71	CCCCACTACTGGATAGAT	130 (18)	50	>4	>4	>4	>4	PM	3	1 (12)	PM	Eilers <i>et al.</i> (2001)
ALF968	alpha-proteobacteria	GGTAAGTTCTGCGGTT	968 (18)	56	PM	PM	PM	PM	PM	3	1 (12)	PM	Neef (1994)
MALF-1	marine alpha cluster	GCCGGGTTTCTTTACCA	488 (18)	56	PM	PM	PM	PM	>4	>4	>4	PM	Gonzalez <i>et al.</i> (1996)
ROS537	marine alpha cluster	CAACGCTAACCCCTCC	537 (17)	65	PM	PM	PM	PM	2 (9,11)	4	4	PM	Eilers <i>et al.</i> (2001)
GRB	marine alpha cluster	GTCAGTATCGAGCCAGTGAG	735 (20)	55	PM	PM	PM	PM	>4	>4	>4	PM	Giuliano <i>et al.</i> (1999)
RSB67	subgroup of marine alpha cluster	CGCTCCACCCGAAAGTAG	67 (18)	67	PM	PM	>4	>4	>4	>4	>4	PM	Zubkov <i>et al.</i> (2001)
RC1031	KT0202a	ACCTGTCACTATGTCCCG	1031 (18)	56	PM	PM	PM	PM	2 (12,18)	>4	4	PM	Eilers <i>et al.</i> (2001)
RC1239	KT0202a	TAACTCACTGTAGTTGCCAT	1239 (20)	40	PM	1 (14)	1 (14)	4	>4	>4	>4	PM	Eilers <i>et al.</i> (2001)
KT13	subgroup of marine alpha cluster	TAACTCACTGTAGTGCAT	1239 (20)	40	1 (14)	PM	PM	3	>4	>4	>4	PM	Peplies <i>et al.</i> (2003)
KT13-231	KT1117	ATCTAATCAAAACGCGGGCC	231 (19)	53	1 (9)	PM	1 (9)	1 (9)	PM	4	4	PM	Eilers <i>et al.</i> (2001)
Ros7-1029	JP7.1	CTGTCACTTGGTCTCTTG	1029 (18)	50	>4	>4	PM	>4	PM	>4	>4	PM	Eilers <i>et al.</i> (2001)
RRP1088	subgroup of alpha-proteobacteria	CGTTGGCGGACTTAACC	1088 (17)	59	PM	PM	PM	PM	1 (7)	3	2 (5,7)	PM	Neef (1994)
PAR1457	subgroup of alpha-proteobacteria	CTACCGTGGTCCCGTGCC	1457 (18)	72	PM	3	1 (2)	3	>4	>4	>4	PM	Neef (1994)
ALF4-1322	subgroup of alpha-proteobacteria	TCCGCCCTTCATGCTCTCG	1322 (18)	61	4	4	4	PM	PM	>4	>4	PM	Neef (1994)
EP129	subgroup of alpha-proteobacteria	CGAACCTAAAGGCAGTGT	129 (18)	50	>4	>4	>4	PM	PM	>4	>4	PM	Neef (1994)

^a For every probe, a control containing one central mismatch was spotted, additionally, i.e. the whole probe set includes a total of 40 different oligonucleotides.

^b E. coli numbering according to Brosi *et al.* (1981); position of 3' nucleotide of probe is stated; nt, nucleotide.

^c For up to two mismatches, the position of the mismatches is also stated (numbers in parentheses counted from the 3' end); PM, perfect match.

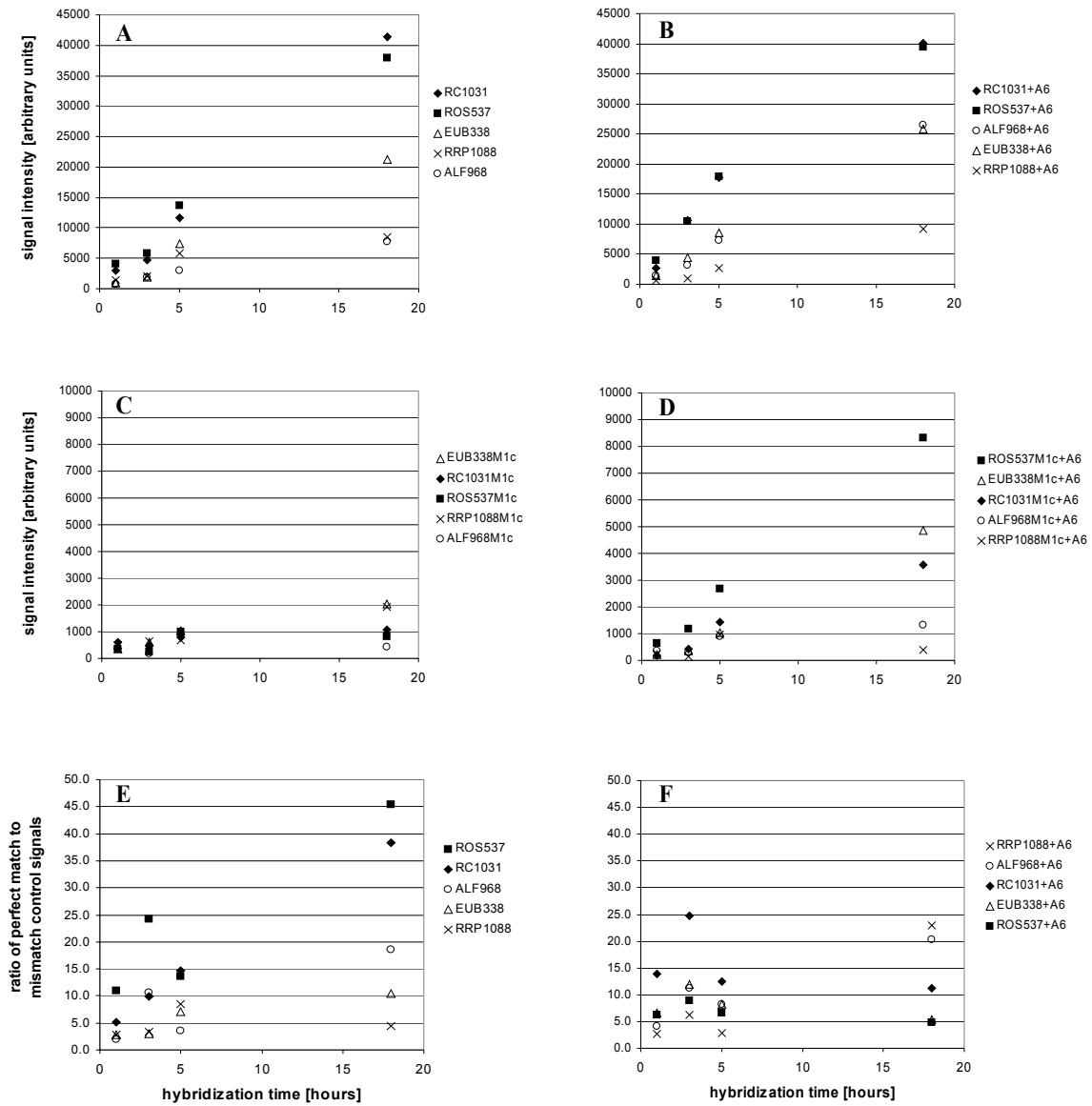


Figure 2: Influence of hybridization time on signal intensities of selected probes. Total RNA (final concentration 8 ng/ μ l) from strain KT0202a was hybridized for 1, 3, 5, and 18 h to five selected probes targeting this strain. (A) perfect match capture oligonucleotides without spacer, (B) perfect match oligonucleotides with a 6-mer spacer, (C) single mismatch controls without spacer, and (D) single mismatch controls with a 6-mer spacer. Panels (E) and (F) show ratios of perfect match to mismatch control signals for probes without and with a 6-mer spacer, respectively.

Parameter analysis. The influence of hybridization time on sensitivity and specificity of hybridization was evaluated by a time series (1, 3, 5, and 18 h) with labeled total RNA from strain KT0202a on probes EUB338, ALF968, ROS537, RRP1088, and RC1031. Comparing 3 and 18 h of hybridization time, absolute signal intensities for perfect matched probes could be increased by a factor of 4.0 for probe ALF968 to a factor of 11.6 for probe EUB338 (Fig. 2A). Under the hybridization conditions applied, the influence of a 6-mer spacer was negligible after 18 h except for probe ALF968 which showed a 3.4 times stronger signal (Fig. 2B).

Controls with one central mismatch also showed increased signal intensities with longer hybridization time (Fig. 2C and D) and here a pronounced effect of the 6-mer spacer on hybridization efficiency was observed (Fig. 2D).

Ratios of perfect match to corresponding single mismatch signals were increased for all probes tested when comparing 3 and 18 h of hybridization time without spacer, ranging from a factor of 1.3 for probe RRP1088 to a factor of 3.9 for probe RC1031 (Fig. 2E). For probe variants with 6-mer spacers, ratios were increased for probes ALF968 and RRP1088 (factor 1.8 and 3.7, respectively) but reduced to 50% for probes EUB338, ROS537, and RC1031 (Fig. 2F).

Microarray analysis of a co-culture. The signal pattern obtained after microarray hybridization of 0.45 μ g of total RNA (15 ng/ μ l) from the co-culture for 3 h (Fig. 3A), suggests the presence of strains KT11ds2, KT1117 and/or JP7.1, and JP13.1 because positive signals were only found for probes EUB338 (targets all six strains), CF319a, CYT1448 (both only targeting strain KT11ds2), ALF968 (all four alpha-proteobacteria), ROS537, GRb (all three members of the marine alpha cluster), KT13 (only strains KT1117 and JP7.1), and ALF4-1322 (only strain JP13.1). This combination excludes strains KT0202a and KT71 and suggests false-negative results for probes UNIV1392, MALF-1, KT13-231, RRP1088, and EP129 (compare Table 1).

Absolute signal intensities ranged from $1,175 \pm 243$ (mean \pm standard deviation of spot replicates) arbitrary units (a.u.) for probe ALF968 to $7,632 \pm 1,903$ a.u. for probe CF319a. Probes EUB338 and ROS537 yielded signals that were significantly above the background according to our criteria whereas the probes UNIV1392 and MALF-1 did not, although targeting exactly the same strains within the pool of reference strains, respectively.

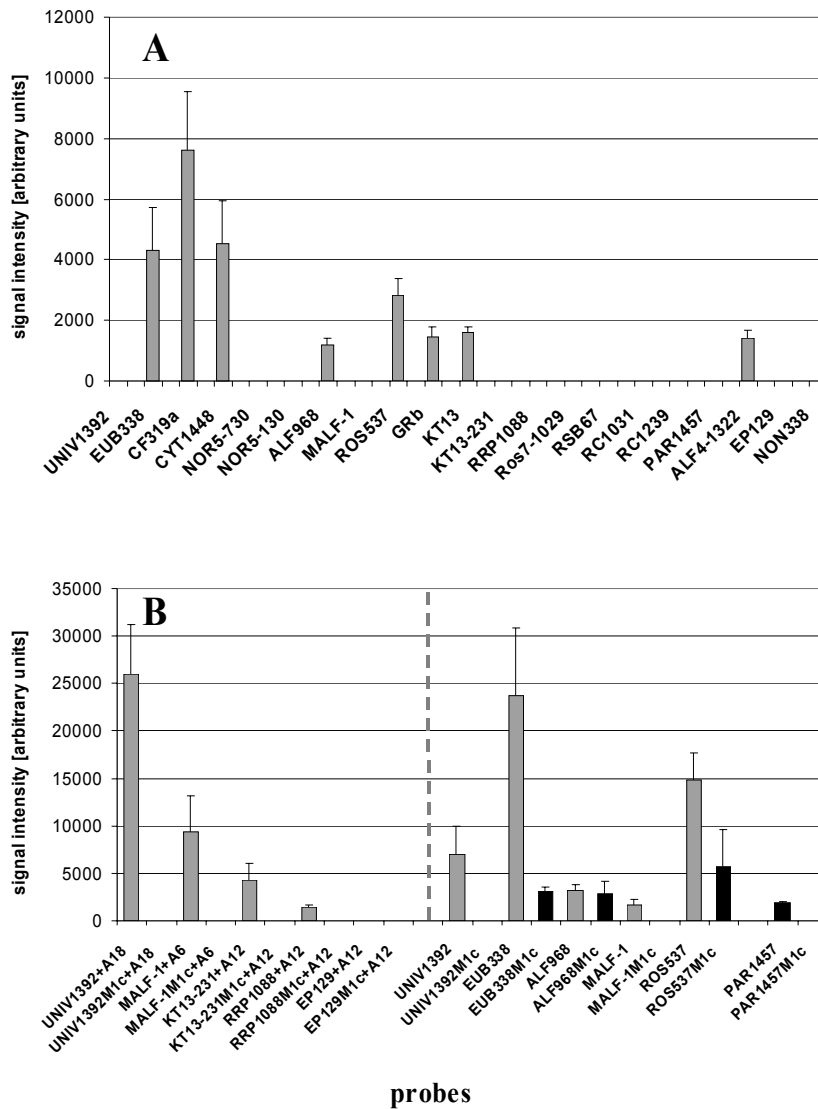


Figure 3: Microarray analysis of extracted total RNA from a co-culture after one week of incubation, originally inoculated with all six reference strains. (A) Hybridization was performed under standard conditions. (B) Changes in signal pattern after hybridization of the same RNA extract for 18 h to a modified set of probes (directed application). Capture oligonucleotides UNIV1392, MALF-1, KT13-231, RRP1088, and EP129 were additionally tested with poly(A) spacers of different lengths (left from separator). Former negative probes featuring signals after 18 h of hybridization are shown, additionally (right). Black bars indicate positive results within the single mismatch controls. Laser power for scanning was 70% for both experiments.

Microarray analysis of a picoplankton sample. For initial microarray analysis of the picoplankton sample, hybridization was performed under standard conditions, except that only a reduced amount of RNA was available, resulting in a final concentration of 3 ng/μl (0.1 μg in total). In this experiment, no signals could be detected. Therefore, hybridization time was

extended to 18 h and probes were equipped with 6-mer spacers (except the two general probes) to increase sensitivity. Now, microarray analysis gave signals for the general probes UNIV1392 (+A18) and EUB338, and suggested presence of members of the *Cytophaga-Flavobacterium* group of *Bacteroidetes*, the alpha-proteobacteria, and the marine alpha cluster (Fig. 4A). Absolute signal intensities ranged from $2,752 \pm 1089$ a.u. for probe ROS537+A6 to $12,805 \pm 2738$ a.u. for probe UNIV1392+A18. A single microarray hybridization required extraction of a 375 ml subsample.

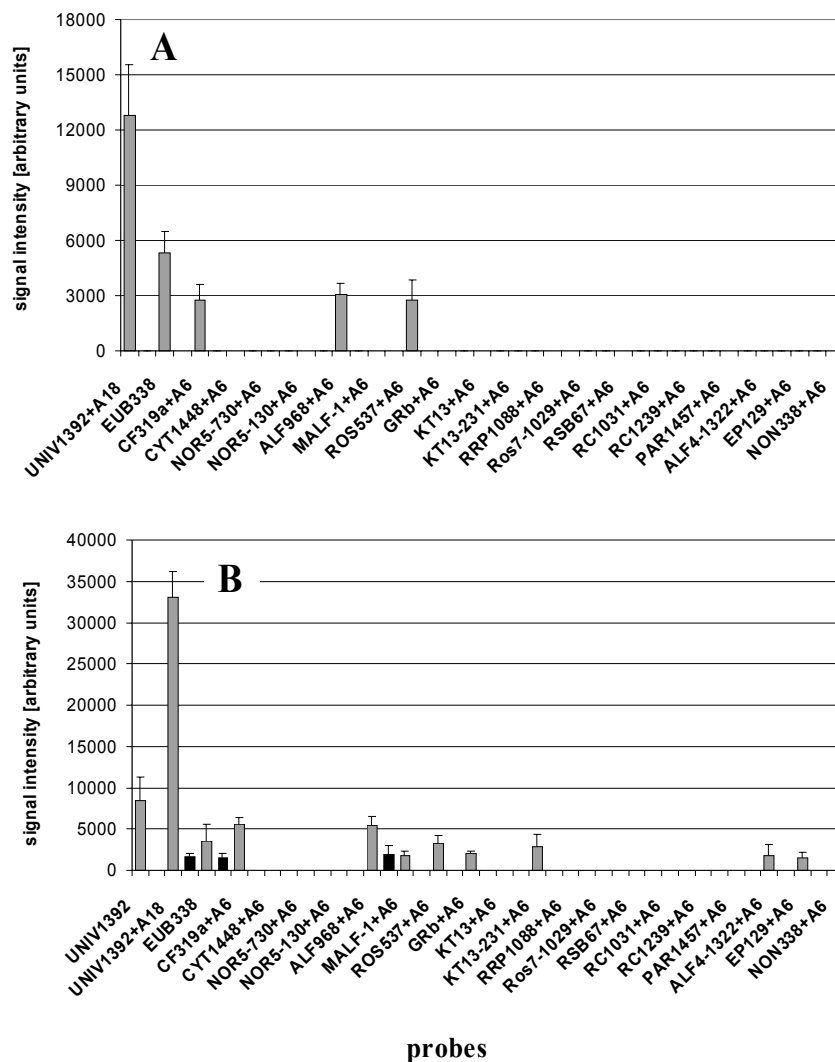


Figure 4: Microarray analysis of extracted total RNA from the environmental sample. (A) Hybridization was performed for 18 h at a target concentration of 3 ng/μl (70% laser power). For all probes a 6-mer spacer was applied, except for general probes UNIV1302 (A18) and EUB338 (none). For clarity, single mismatch controls are not shown (no signals). (B) Total RNA from an independent extraction was hybridized for 18 h at a target concentration of 23 ng/μl (40% laser power). Black bars indicate signals for single mismatch controls, for the probe on the left.

Parallel FISH analysis of the co-culture and the picoplankton sample. For validation of the microarray results obtained from the co-culture, probes EUB338, ALF968, MALF-1, RC1031, KT13-231, Ros7-1029, ALF4-1322, CF319a, and NOR5-730 were chosen for parallel FISH analysis. Quantitative FISH data are shown in Table 2. In summary, presence of strains KT11ds2, JP13.1, and KT1117 could be confirmed for the co-culture. Only these strains were detected at high relative abundances (60%, 16%, and 7% of total DAPI counts, respectively). Strains KT71 and KT0202a accounted for 2 and < 1% of the bacterial community, respectively. Absolute cell number of the analyzed co-culture determined by DAPI-staining was $5.0 \times 10^7 \text{ ml}^{-1}$. Within the microarray data set shown in Fig. 3A, distribution of absolute signal intensities for positive probes does not correlate with the relative abundances of the corresponding target strains in the co-culture. For example, in the microarray experiment, probes ALF968 and ALF4-1322 yielded signal intensities that were statistically indistinguishable (Student's t-test; $T = 1.967$, $n = 8$, $P = 0.069$), although probe ALF4-1322 targets only one (JP13.1) of the two abundant strains targeted by probe ALF968 (JP13.1 and KT1117).

Quantitative data of parallel FISH analysis (Tab. 2), also supported the signal pattern obtained by microarray analysis of the picoplankton sample (Fig. 4A). All target subpopulations with a relative abundance of at least 5% of total DAPI-counts could be detected by microarray hybridization, and no indications for false-positive microarray signals were found. The absolute cell number of the analyzed water sample was 1.1×10^6 cells/ml.

TABLE 2. Quantitative FISH analysis of samples with unknown community composition

Probe	co-culture ^a			environmental sample ^b	
	Targeted strains	Hybridized cells (%) of all DAPI-stained objects	Total cell counts (cells/ml)	Hybridized cells (%) of all DAPI-stained objects	Total cell counts (cells/ml)
UNIV1392	n.a.	-	-	n.d.	n.d.
EUB338	all six reference strains	80	4.0×10^7	56	6.2×10^5
CF319a	KT11ds2	60	3.0×10^7	17	1.9×10^5
CYT1448	n.a.	-	-	n.d.	n.d.
NOR5-730	KT71	2	1.0×10^6	n.d.	n.d.
ALF968	KT1117, KT0202a, JP13.1, and JP7.1	25	1.3×10^7	10	1.1×10^5
MALF-1	KT1117, KT0202a, and JP7.1	8	4.0×10^6	2	0.2×10^5
ROSS37	KT1117, KT0202a, and JP7.1	7	3.5×10^6	5	0.6×10^5
GRb	n.a.	-	-	n.d.	n.d.
RSB67	n.a.	-	-	2	0.2×10^5
RC1031	KT0202a	<1	$<0.5 \times 10^6$	-	-
KT13	n.a.	-	-	2	0.2×10^5
KT13-231	KT1117	7	3.5×10^6	n.d.	n.d.
Ros7-1029	JP7.1	n.d.	n.d.	-	-
RRP1088	n.a.	-	-	4	0.4×10^5
PAR1457	n.a.	-	-	1	0.1×10^5
ALF4-1322	JP13.1	16	8.0×10^6	n.d.	n.d.
EP129	n.a.	-	-	n.d.	n.d.
NON338	negative control	n.d.	n.d.	n.d.	n.d.

^a For all probes applied for FISH validation, reference strains targeted are also given.

^b All results for probes with a positive microarray signal in Fig. 4A are highlighted.

n.a. and minus, probe was not applied for FISH validation. n.d., no signal was detected.

Optimization attempts of microarray signal patterns for complex target mixtures. Based on the FISH data shown in Table 2, false-negative microarray results were obtained for the co-culture with probes UNIV1392, MALF-1, KT13-231, RRP1088, and EP129. Subsequently, these probes were applied with poly(A) spacers of different lengths (6, 12, and 18-mers) to increase their hybridization yields by reducing the influence of steric hindrance. Under standard hybridization conditions (46°C, 3 h), for probes UNIV1392 and MALF-1 specific signals of $25,886 \pm 7372$ a.u. and $1,129 \pm 118$ a.u., respectively, could be restored using an 18-mer spacer for probe UNIV1392 (+A18) and a 6-mer spacer for MALF-1 (+A6) spacer (data not shown). Specific signals for probes KT13-231+A12, and RRP1088+A12 were also suggested by the raw image but signals were not significantly above the background according to our criteria. The experiment was repeated with an incubation time of 18 h, since parameter analysis suggested that hybridization efficiency of particular probes can be improved by increasing the hybridization time without loss of hybridization specificity. Except for probes KT13 and ALF4-1322, signal intensities of unmodified probe variants were significantly increased (analysis of variance, followed by Scheffé post-hoc comparisons, $P < 0.001$) up to a factor of 6.3 for probe CYT1448. In Fig. 3B, data are given for the four modified probes (left part of the figure) and all unmodified probes showing qualitative differences after 18 h of hybridization. Except for probe EP129, all former false-negative results could be restored specifically by applying spacers. However, unmodified variants of probes EUB338, ALF968, and ROS537 now showed signals within the corresponding mismatch controls, i.e. hybridization specificity was reduced for these probes when incubating for 18 h. Furthermore, a signal for probe PAR1457 could be observed and probes UNIV1392 and MALF-1 now showed signals also without spacers.

For the picoplankton sample, another hybridization was performed for 18 h with an increased RNA concentration of 23 ng/ μ l (0.7 μ g total), to generate signals also for community members of low relative abundances (< 5%). The larger amount of RNA was provided by use of a reduced amount of elution buffer after the final precipitation step. Results are shown in Fig. 4B. Now, also for probes MALF-1, GRb, KT13-231, ALF4-1322, and EP129 signals were detected (all applied with a 6-mer spacer), although for probes GRb, KT13-231, ALF4-1322, and EP129 no hybridized cells could be observed after FISH analysis. Furthermore, signals for the mismatch controls of probes UNIV1392+A18, EUB338, and ALF968+A6 were observed.

DISCUSSION

Compared to other molecular methods used for microbial community analysis such as standard hybridization (FISH, dot blot hybridization, etc.) or fingerprinting techniques (DGGE, ARDRA, etc.), the major advantage of the DNA microarray approach is its potential to detect multiple variants of a genetic marker or even of multiple genetic markers in a single experiment. However, current application of DNA microarrays in the field of environmental microbiology is hampered by various methodical deficiencies, such as laborious protocols for preparation of target molecules, insufficient detection limits, or requirement for real time recording of dissociation rates or hybridization kinetics to assess specificity of obtained signals. In this work, we present a fast and easy protocol for 16S rRNA-based microarray analysis of bacterial communities. Results were validated by parallel quantitative FISH analysis to assess them under the aspects of (i) specificity, (ii) sensitivity, and (iii) quantification.

Methodical aspects of the microarray format applied. The initial microarray experiments with pure cultures showed that by using a protocol optimized for hybridization of corresponding PCR-amplified 16S rRNA genes (Peplies *et al.*, 2003), highly specific signal patterns could be obtained also with labeled total cellular RNAs. A direct comparison of the results in the present and our previous study for the pure cultures analyzed revealed nearly identical signal patterns, suggesting that former identified hybridization mechanisms for the corresponding 16S rRNA gene amplicons and optimization strategies are transferable to the direct detection of 16S rRNA molecules within a pool of total RNA. An additional DNA removal step after RNA extraction showed no effect on signal composition and strength (data not shown).

In contrast to common strategies, in our study, specificity of hybridization was assured by applying appropriate conditions in terms of spacer length, hybridization time, and target concentration (Peplies *et al.*, 2003). No effort on melting point adjustment of capture probes (Loy *et al.*, 2002) was made and signal detection was based on a single fluorescence measurement of the hybridized arrays. Various microarray studies are depending on multiple or on-line measurements for recording melting curves (Liu *et al.*, 2001) or hybridization kinetics (Dai *et al.*, 2002). Compared to other protocols for a direct detection of extracted 16S rRNA on microarrays (Small *et al.*, 2001; Koizumi *et al.*, 2002), no fragmentation of target molecules was necessary to obtain clear hybridization signals. In terms of sensitivity, we

assume no drawback by hybridizing intact RNA because hybridization efficiency of target molecules is increased by fragmentation (Small *et al.*, 2001) but, concomitantly, the number of reporter groups per captured target molecule is reduced. For extraction and direct labeling of total RNAs a simple and fast protocol was developed based on commercially available kits. Total time exposure for sample preprocessing, including quality checks after extraction and labeling, respectively, was only 4 hours. For hybridization and washing a straightforward protocol was applied with a hands-on time of only half an hour and a hybridization of 3 to 18 hours. In comparison, Loy *et al.* (2002) specified a total time exposure of 48 hours for their protocol.

Since also false-negative results within microarray signal patterns are a common phenomenon (Koizumi *et al.*, 2002; Loy *et al.*, 2002), lack of multiple redundant probes for strains or groups of interest can become a strong limitation. To minimize false-negative results an optimization strategy was introduced based on the directed application of capture oligonucleotides (Peplies *et al.*, 2003), which was successfully applied in the present study for probes UNIV1392 and MALF-1. However, optimization was difficult for probes KT13-231 and RRP1088, and not possible for probe EP129 under the conditions tested (Fig. 3B). Basically, an overall adjustment of signal intensities should also be feasible by this approach (Guschin *et al.*, 1997) but would be time consuming for large probe sets due to empirically probe calibration.

Validation of microarray data obtained from samples of unknown composition and sensitivity of our format. Considering current limitations of DNA microarrays for hybridization specificity (Liu *et al.*, 2001; Urakawa *et al.*, 2002) and target quantification (Koizumi *et al.*, 2002), powerful validation techniques are required to assess reliability of microarray results obtained from complex microbial communities. Common strategies are based on a PCR-mediated analysis, such as fingerprinting techniques (Koizumi *et al.*, 2002) or gene libraries (Loy *et al.*, 2002; Wilson *et al.*, 2002). Also quantitative membrane hybridization was applied for data comparison (Koizumi *et al.*, 2002).

In this study, FISH was used for validation of microarray data. This hybridization format does not rely on PCR, is surface independent, and quantitative. In the present study, cell numbers obtained by FISH analysis were used for validation of microarray data. This restricts direct comparability of the results since variations in cellular rRNA content of different bacterial populations are common (Amann *et al.*, 1995) and DNA microarrays are basically characterized by the detection of extracted target molecules.

FISH validation for the co-culture revealed that the abundant strains down to 7% of total DAPI counts could specifically be detected by microarray hybridization of 0.45 μg (15 $\text{ng}/\mu\text{l}$) intact total cellular RNA for 3 h. This is below the detection limit given by Koizumi and coworkers, who were able to detect 16S rRNA targets which accounted for at least 16% of total 16S rRNA in a toluene-degrading consortium by microarray hybridization of 10 μg (250 $\text{ng}/\mu\text{l}$) of fragmented total community RNA for 16 h (Koizumi *et al.*, 2002). In a study of Small *et al.* (2001), the detection limit for fragmented total RNA from pure cultures added to soil extracts was given with 0.5 μg of the targeted RNAs (representing approximately 7.5×10^6 cell equivalents), based on enzymatic signal amplification.

For the two probes NOR5-730 and NOR5-130, targeting only strain KT71 in the co-culture analyzed, total RNA isolated from approx. 3.0×10^6 cells gave no positive microarray signals under the conditions tested. Obviously, target concentration was the limiting factor for microarray detection of the less abundant strains KT71 and KT0202a.

Results from FISH validation for the picoplankton sample confirmed presence of the populations suggested by microarray analysis, indicating that under appropriate conditions, a specific hybridization of multiple probes can also be assured for complex 16S rRNA target mixtures with our protocol. Populations down to 5% of total DAPI counts could specifically be detected.

Additional detection of community members with relative cellular abundances below 5% by increasing the overall target concentration was not successful, since hybridization specificity for probes targeting more abundant populations was lost as shown in Fig. 4B. An increased overall target concentration positively affects hybridization efficiencies also of mismatched duplexes, resulting in detectable hybridizations yields for the mismatch controls "targeting" the dominant populations. This observation is in good agreement with results from our previous study and demonstrates that specificity and sensitivity of hybridization are directly linked.

Quantitative aspects of 16S rRNA detection by microarray analysis.

Accurate quantification of target molecules by microarray hybridization is often hampered by variations in spot quality and hybridization homogeneity for a single array. This limitation can be circumvented by the application of internal controls (Cho and Tiedje, 2002) or sophisticated statistical analysis (Brody *et al.*, 2002; Tran *et al.*, 2002). Additional restrictions for quantitative microarray studies may occur due to interaction and characteristics of the involved molecular components itself, such as the influence of probe sequence on hybridization

efficiency, secondary structures of the target molecules, and relative position of probe binding site on the target molecules. These parameters can interfere with a simple correlation of probe signal ratios and relative abundances of the corresponding target molecules (Guschin *et al.*, 1997; Peplies *et al.*, 2003). A comprehensive overview on parameters determining hybridization yield can be found at Southern *et al.* (1999). Furthermore, DNA-microarrays can not provide cellular abundances, since extracted RNA is analyzed. Composition of the extracted community rRNA pool is affected by both, relative abundance and activity of the particular populations (Amann, 1995).

A very limited number of studies was published in the field of microbial community studies comparing microarray data with results from independent methods under the aspect of quantification. Our results clearly confirm the finding of Koizumi *et al.* (2002), who pointed out the current limitation of DNA microarrays for quantitative microbial community analysis, even when hybridizing extracted total RNA. Signal ratios of positive probes obtained after microarray analysis of the co-culture and the environmental sample, respectively, did not correlate with quantitative data from FISH experiments (Fig. 3A, 4A, and Tab. 2) and several probes, applied in parallel and targeting the same 16S rRNA at different sites, show pronounced variations in hybridization yield (Fig. 1). Basically, inconsistency of probe signal intensities can also be caused by hybridization kinetics, but time series conducted (Fig. 2) gave no evidences for such an effect. For all probes tested, no equalization of signal intensities was observed by increasing hybridization time from 1 to 18 h.

Is a general application of DNA microarrays for quantitative microbial community studies possible? Various studies have shown that the absolute signal intensity of a particular probe increases with the concentration of the targeted nucleic acids (Guo *et al.*, 1994; Riccelli *et al.*, 2001; Wu *et al.*, 2001; Religio *et al.*, 2002). Therefore, empirical signal calibration should enable a direct quantification of selected target molecules. However, such an approach would be extremely laborious and time consuming for large probe sets. Furthermore, each probe is only targeting a defined binding site within the 16S rRNA. For probes, targeting larger phylogenetic clusters, this particular site can be found on a number of similar but not identical and partly unknown 16S rRNAs. The influence of sequences flanking the probe binding site on hybridization efficiency is still mostly unknown. Clearly, further optimizations are needed to reduce problems caused by differences in the probe specific hybridizations efficiencies.

Remarkably, the microarray data obtained in this study appear to be at least "semi-quantitative". Applying appropriate hybridization conditions (Fig. 3A and 4A), all populations

targeted and with a relative abundance below a threshold value of approximately 5% could specifically be detected.

In summary, a protocol is presented enabling the identification of major bacterial populations even in complex and oligotrophic systems such as marine surface waters.

EXPERIMENTAL PROCEDURES

Bacterial strains, co-culture and environmental samples. For microarray analysis of extracted RNA from pure cultures and inoculation of a co-culture, six reference strains (KT0202a, KT1117, JP7.1, JP13.1 [alpha-proteobacteria], KT11ds2 [*Cytophaga-Flavobacterium* group of *Bacteroidetes*], and KT71 [gamma-proteobacteria]), isolated from the surface water of the German Bight by Eilers *et al.* (2000 and 2001), were used. Accession numbers of the 16S rRNA gene sequences of the six strains are AF173971, AF235111, AF305498, AY007676, AY007679, and AY007680.

For setting up a co-culture, 250 ml of the synthetic seawater medium MPM-m according to Eilers *et al.* (2001) with 0.01% yeast extract were inoculated with all six reference strains using an inoculating loop. Sampling was carried out for parallel microarray and FISH analysis after one week of incubation at room temperature on a shaker.

For parallel microarray and FISH analysis of a marine picoplankton sample, surface water was collected on 23 September 2002 at the sampling site Helgoland Roads (54°9'N, 7°52'E) near the island of Helgoland, which is situated 23 miles offshore in the German Bight. After transfer to Bremen (approx. 10 h), the water was prefiltered (pore size 3 µm) and processed for RNA extraction and FISH analysis.

RNA extraction and labeling. Total RNA was extracted from pure cultures, the co-culture and the environmental sample using the RNAwiz reagent from Ambion (Austin, USA), which accords to a classical phenol-chloroform extraction but provides a quick and easy procedure based on a single reagent. Briefly, cells of pure cultures (50 ml) were collected by centrifugation, resuspended in 1 ml of RNAwiz reagent, and transferred to a FastRNA Blue Tube (Bio101, Carlsbad, USA). Cells from the co-culture (100 ml) and the environmental sample (750 ml) were collected on 0.2 µm white polycarbonate membrane filters (Millipore, Eschborn, Germany). The filters were chopped and transferred into 1 ml of the RNAwiz reagent in a Fast RNA Blue Tube. For improved cell lysis, tubes were placed into a FP 120 FastPrep cell disruptor (Savant Instruments, Holbrook, USA) and processed at a speed rating of 6 for 2 x 20 seconds. Further steps were conducted according to the RNAwiz manufacturer

protocol. After the final precipitation step, the RNA was resuspended in 50 µl of PCR-water (Sigma, Deisenhofen, Germany). Optionally, potential DNA contaminations of the RNA extracts were removed by the DNA-free kit from Ambion (Austin, USA).

Alternatively, total RNA from the environmental sample was extracted based on a hot phenol-chloroform treatment according to Oelmüller *et al.* (1990).

Quality and quantity of the extracted RNA was checked by capillary electrophoresis using an Agilent 2100 bioanalyzer with the RNA 6000 Nano Assay kit (Agilent Technologies, Palo Alto, USA).

Total RNA was labeled using an Alexa Fluor 488 ULYSIS labeling kit from Molecular Probes (Eugene, USA) according to the manufacturer protocol, except that for purification a second precipitation step was applied and precipitation was conducted at room temperature. This kit provides a rapidly (10 min), nonenzymatic method for chemically labeling of nucleic acids without fragmentation based on a platinum dye complex, resulting in stable coordination complexes with guanine residues. The labeled RNA was eluted in 25 µl of Tris-EDTA buffer pH 7.5. Finally, labeling efficiency and quantity of the RNA was determined by UV-spectrometry according to the manufacturer protocol. Labeled RNA was stored at -18°C.

Oligonucleotide probe set. The same set of 20 hierarchical and redundant oligonucleotide probes as published, recently (Peplies *et al.*, 2003), was used for differentiation of the six reference strains and analysis of the picoplankton sample. Part of the probes are known to target abundant bacterial groups which can be found in the German Bight surface waters (Eilers *et al.*, 2001). The probes and their characteristics are listed in Tab. 1. For all twenty probes a control containing one single central mismatch was included. A probe reverse complementary to EUB338 termed NON338 was spotted, additionally (not listed in Tab.1), to serve as a general negative control when analyzing samples of unknown composition.

Preparation of glass slides and spotting. For covalent immobilization of 5'-aminomodified capture oligonucleotides (Thermo Hybaid, Ulm, Germany), standard microscopic glass slides (Menzel, Braunschweig, Germany) were activated by 1,4-phenylenediisothiocyanate (PDITC) treatment according to Benters *et al.* (2002).

Probes were spotted onto the activated slide surface using a SpotArray24 spotting device with TeleChem Stealth pins (Packard BioChip Technologies, Billerica, USA). Concentration of the aminomodified oligonucleotides in PCR-water was 10 µM with 1% glycerol. Normal spot diameter was approx. 100 µm under the spotting conditions applied. Postprocessing of the spotted slides (binding, blocking, washing, and storing) was conducted as described, recently (Peplies *et al.*, 2003).

DNA-microarray hybridization. For hybridization and washing of spotted slides a standard FISH protocol according to Pernthaler *et al.* (2001) was used, as described recently (Peplies *et al.*, 2003). In all cases, hybridization was performed at 46°C without formamide. Standard hybridization time was extended to 3 h compared to the FISH protocol. For a time series with RNA from a pure culture hybridization times of 1, 5, and 18 h were applied, additionally, and total RNAs from the co-culture and the picoplankton sample were also hybridized for 18 h. Unless otherwise specified, a total amount of 0.23 µg of extracted and labeled total RNA of either a pure culture, the co-culture, or the picoplankton sample was applied in a volume of 30 µl of hybridization buffer to the microarrays (final concentration 8 ng/µl) under a 22 x 25 mm LifterSlip (Erie Scientific Company, Portsmouth, USA). Hybridization was conducted in a Corning Hybridization Chamber (Corning, Acton, USA). Unless otherwise specified, probes were applied for microarray analysis under standard conditions, i.e. immobilized via 5'-end, without spacer, and at a concentration of 10 µM. After washing, slides were dried by centrifugation at low speed.

For reduction of steric hindrance, probes were used with poly adenosin triphosphate spacers of 6, 12, and 18 nucleotides in length, located at the 5'-end of the capture oligonucleotides.

Signal detection and data analysis. Slides were imaged at a resolution of 10 µm using a ScanArray Express microarray scanner (Packard BioChip Technologies, Billerica, USA) at same sensitivity level of the photomultiplier (75%) and appropriate intensity of laser power for each individual slide.

For spot detection and signal quantification the microarray analysis software QuantArray 3.0 (Packard BioChip Technologies, Billerica, USA) was used. For spot quantification the fixed circle method was used. Signals were considered as positive if mean spot pixel intensity was higher than mean local background pixel intensity plus two times the standard deviation of the local background. This range decreases the probability of erroneous false-positive signals to < 2.5%. All probes were deposited in 10 replicate spots. Each datapoint shown represents the arithmetic mean of the positive replicates for a particular probe. Error bars indicate the standard deviation of the positive replicates.

Fluorescence in situ hybridization. Microarray analysis of samples with an unknown community composition was validated by fluorescence in situ hybridization (FISH) using the standard protocol (Pernthaler *et al.*, 2001). Probes UNIV1392, EUB338, CF319a, CYT1448, NOR5-730, ALF968, MALF-1, ROS537, GRb, KT13, RSB67, KT13-231, RRP1088, PAR1457, ALF4-1322, EP129, RC1031, and Ros7-1029 were used at 0, 35, 35, 30, 30, 20, 20,

35, 30, 20, 20, 30, 0, 20, 35, 0, 20, and 35% formamide, respectively. Briefly, 1 ml of the co-culture and 5 ml of the water sample were fixed with 2% (vol/vol) formaldehyde for 1 h at 4°C and filtered on 0.2 µm white polycarbonate membrane filters (Millipore, Eschborn, Germany). Cells on filter sections were hybridized, counterstained with 4,6-diamidino-2-phenylindole (DAPI), mounted, and microscopically evaluated as described previously (Pernthaler *et al.*, 2001).

ACKNOWLEDGEMENTS

We thank D. Wöhrle and B. Meyer-Schlosser for generous supply of activated glass slides. A. Ellrott is gratefully acknowledged for excellent help on setting up and maintaining technical equipment. This work was done in the framework of the Centre of Applied Gensensorik (CAG) at the University of Bremen and supported by the German Ministry of Education and Research (BMBF, contract 0311833A) and by the Max Planck Society.

REFERENCES

- Amann, R. (1995) Fluorescently labelled, rRNA-targeted oligonucleotide probes in the study of microbial ecology. *Mol. Ecol.* **4**: 543-554.
- Amann, R., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**: 143-169.
- Benters, R., Niemeyer, C.M., Drutschmann, D., Blohm, D., and Wöhrle, D. (2002) DNA microarrays with PAMAM dendritic linker systems. *Nucleic Acids Res.* **30**: e10.
- Brody, J.P., Williams, B.A., Wold, B.J., and Quake, S.R. (2002) Significance and statistical errors in the analysis of DNA microarray data. *Proc. Natl. Acad. Sci. USA* **99**: 12975-12978.
- Brosius, J., Dull, T.J., Sleeter, D.D., and Noller, H.F. (1981) Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J. Mol. Biol.* **148**: 107-127.
- Cho, J.C., and Tiedje J.M. (2002) Quantitative detection of microbial genes by using DNA microarrays. *Appl. Environ. Microbiol.* **68**: 1425-1430.
- Dai, H.Y., Meyer, M., Stepaniants, S., Ziman, M., and Stoughton, R. (2002) Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays. *Nucleic Acids Res.* **30**: e86.
- Dennis, P., Edwards, E.A., Liss, S.N., and Fulthorpe, R. (2003) Monitoring gene expression in mixed microbial communities by using DNA microarrays. *Appl. Environ. Microbiol.* **69**: 769-778.
- Eilers, H., Pernthaler, J., Glöckner, F.O., and Amann, R. (2000) Culturability and in situ abundance of pelagic bacteria from the North Sea. *Appl. Environ. Microbiol.* **66**: 3044-3051.
- Eilers, H., Pernthaler, J., Peplies, J., Glöckner, F.O., Gerds, G., and Amann, R. (2001) Isolation of novel pelagic bacteria from the German Bight and their seasonal contributions to surface picoplankton. *Appl. Environ. Microbiol.* **67**: 5134-5142.
- Giuliano, L., De Domenico, M., De Domenico, E., Höfle, M.G., and Yakimov, M.M. (1999) Identification of culturable oligotrophic bacteria within naturally occurring bacterioplankton communities of the Ligurian Sea by 16S rRNA sequencing and probing. *Microb. Ecol.* **37**: 77-85.
- Gonzalez, J.M., Whitman, W.B., Hodson, R.E., and Moran, M.A. (1996) Identifying numerically abundant culturable bacteria from complex communities: an example from a lignin enrichment culture. *Appl. Environ. Microbiol.* **62**: 4433-4440.

- Guo, Z., Guilfoyle, R.A., Thiel, A.J., Wang, R., and Smith, L.M. (1994) Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.* **22**: 5456-5465.
- Guschin, D.Y., Mobarry, B.K., Proudnikov, D., Stahl, D. A., Rittmann, B.E., and Mirzabekov, A.D. (1997) Oligonucleotide microchips as genosensors for determinative and environmental studies in microbiology. *Appl. Environ. Microbiol.* **63**: 2397-2402.
- Hacia, J.G. (1999) Resequencing and mutational analysis using oligonucleotide microarrays. *Nat. Genet.* **21**: 42-47.
- Koizumi, Y., Kelly, J.J., Nakagawa, T., Urakawa, H., El Fantroussi, S., Al Muzaini, S. et al. (2002) Parallel characterization of anaerobic toluene- and ethylbenzene-degrading microbial consortia by PCR-denaturing gradient gel electrophoresis, RNA-DNA membrane hybridization, and DNA microarray technology. *Appl. Environ. Microbiol.* **68**: 3215-3225.
- Lander, E.S. (1999) Array of hope. *Nat. Genet.* **21**: 3-4
- Liu, W.T., Mirzabekov, A.D., and Stahl, D.A. (2001) Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. *Environ. Microbiol.* **3**: 619-629.
- Loy, A., Lehner, A., Lee, N., Adamczyk, J., Meier, H., Ernst, J. et al. (2002) Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl. Environ. Microbiol.* **68**: 5064-5081.
- Manz, W., Amann, R., Ludwig, W., Vancanneyt, M., and Schleifer, K.-H. (1996) Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. *Microbiology* **142**: 1097-1106.
- Marchesi, J.R., Sato, T., Weightman, A. J., Martin, T.A., Fry, J.C., Hiom, S.J., and Wade, W.G. (1998) Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Appl. Environ. Microbiol.* **64**: 795-799.
- Neef, A. (1997) Ph.D thesis. Technical University, Munich, Germany.
- Oelmüller, U., Krüger, N., Steinbüchel, A., and Cornelius, G. (1990) Isolation of prokaryotic RNA and detection of specific mRNA with biotinylated probes. *J. Microb. Methods* **11**: 73-84.
- Olsen, G.J., Lane, D.J., Giovannoni, S.J., Pace, N.R., and Stahl, D.A. (1986) Microbial ecology and evolution: a ribosomal rRNA approach. *Annu. Rev. Microbiol.* **40**: 337-365.

- Pace, N.R., Stahl, D.A., Lane, D.J., and Olsen, G.J. (1986) The use of rRNA sequences to characterize natural microbial populations. *Adv. Microb. Ecol.* **9**: 1-55.
- Peplies, J., Glöckner, F.O., and Amann, R. (2003) Optimization strategies for the DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes. *Appl. Environ. Microbiol.* **69**: 1397-1407.
- Pernthaler, J., Glöckner, F.O., Schönhuber, W., and Amann, R. (2001) Fluorescence in situ hybridization (FISH) with rRNA-targeted oligonucleotide probes. *Methods Microbiol.* **30**: 207-226.
- Religio, A., Schwager, C., Richter, A., Ansorge, W., and Valcarcel, J. (2002) Optimization of oligonucleotide-based DNA microarrays. *Nucleic Acids Res.* **30**: e51.
- Riccelli, P.V., Merante, F., Leung, K.T., Bortolin, S., Zastawny, R.L., Janeczko, R., and Benight, A.S. (2001) Hybridization of single-stranded DNA targets to immobilized complementary DNA probes: comparison of hairpin versus linear capture probes. *Nucleic Acids Res.* **29**: 996-1004.
- Rudi, K., Skulberg, O.M., Skulberg, R., and Jakobsen, K.S. (2000) Application of sequence-specific labeled 16S rRNA gene oligonucleotide probes for genetic profiling of cyanobacterial abundance and diversity by array hybridization. *Appl. Environ. Microbiol.* **66**: 4004-4011.
- Schmalenberger, A., Schwieger, F., and Tebbe, C.C. (2001) Effect of primers hybridizing to different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling. *Appl. Environ. Microbiol.* **67**: 3557-3563.
- Small, J., Call, D.R., Brockman, F.J., Straub, T.M., and Chandler, D.P. (2001) Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. *Appl. Environ. Microbiol.* **67**: 4708-4716.
- Southern, E., Mir, K., and Shchepinov, M. (1999) Molecular interactions on microarrays. *Nat. Genet.* **21**: 5-9.
- Suzuki, M.T., and Giovannoni, S.J. (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol.* **62**: 625-630.
- Tran, P.H., Peiffer, D.A., Shin, Y., Meek, L.M., Brody, J.P., and Cho, K.W.Y. (2002) Microarray optimizations: increasing spot accuracy and automated identification of true microarray signals. *Nucleic Acids Res.* **30**: e54.

- Urakawa, H., Noble, P.A., El Fantroussi, S., Kelly John, J., and Stahl, D.A. (2002) Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses. *Appl. Environ. Microbiol.* **68**: 235-244.
- Wilson, K.H., Wilson, W.J., Radosevich, J.L., DeSantis, T.Z., Viswanathan, V.S., Kuczmariski, T.A., and Andersen, G.L. (2002) High-density microarray of small-subunit ribosomal DNA probes. *Appl. Environ. Microbiol.* **68**: 2535-2541.
- Wintzingerode, F.V., Goebel, U.B., and Stackebrandt, E. (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* **21**: 213-229.
- Wu, L., Thompson, D.K., Li, G., Hurt, R.A., Tiedje, J.M., and Zhou, J. (2001) Development and evaluation of functional gene arrays for detection of selected genes in the environment. *Appl. Environ. Microbiol.* **67**: 5780-5790.
- Zubkov, M.V., Fuchs, B.M., Archer, S.D., Kiene, R.P., Amann, R., and Burkill, P.H. (2001) Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ. Microbiol.* **3**: 304-311.

3

**Comparative sequence analysis and
oligonucleotide probe design on the 23S rRNA genes
of alpha-proteobacteria from North Sea bacterioplankton**

Jörg Peplies, Frank Oliver Glöckner, Rudolf Amann, Wolfgang Ludwig

in preparation

**Comparative sequence analysis and oligonucleotide probe design
based on the 23S rRNA gene of alpha-proteobacteria
from North Sea bacterioplankton**

JÖRG PEPLIES,¹ FRANK OLIVER GLÖCKNER,^{1*} RUDOLF AMANN,¹
and WOLFGANG LUDWIG²

*Max Planck Institute for Marine Microbiology, Department of Molecular Ecology,
28359 Bremen, Germany¹*

*Technische Universität München, Lehrstuhl für Mikrobiologie,
85350 München, Germany²*

Almost complete 23S rRNA gene sequences were obtained from 11 alpha-proteobacteria isolated from marine surface water of the German Bight. Five of the strains belong to the "marine alpha" group, a phylogenetic cluster which encompasses members of the genus *Roseobacter* and closely related bacteria. For PCR amplification and cycle sequencing of the 23S rDNA of alpha-proteobacteria, new primers were designed and applied with the focus on achieving maximum group coverage rather than maximum specificity. Comparative sequence analysis based on different tree reconstruction methods was included with additional 41 published as well as unpublished complete 23S rDNA sequences from alpha-proteobacteria. The basic topology of a consensus tree was consistent with the 16S rRNA gene sequence-derived phylogeny. Strains associated with the "marine alpha" group form a distinct cluster also by comparative 23S rRNA sequence analysis. The suitability of both markers for the design of oligonucleotide probes targeting selected groups of alpha proteobacteria was systematically evaluated and compared *in silico*. Six clusters of sequences covering different phylogenetic levels as well as two strains were selected as a model system for further analysis. To allow comparability of the results obtained, only 16S rRNA sequences were considered which also had a 23S rRNA equivalent in our data sets. The overall number of phylogenetic redundant probes available could be more than doubled by also targeting the 23S rRNA. For small clusters of high sequences similarity and single strains, a significantly higher number of discriminating binding sites was provided by the 23S rRNA.

Molecular methods based on the comparative sequence analysis of small subunit rRNA molecules have enormously affected our understanding of natural microbial communities (Amann *et al.*, 1995). With the introduction of techniques for the cultivation independent retrieval of rRNA sequence from environmental samples in the mid 1980s (Olsen *et al.*, 1986; Pace *et al.*, 1986), it became possible to address the diversity of natural microbial communities more directly. In the following years, powerful PCR-based methods were established for the amplification, cloning, and sequence analysis of 16S rRNA genes from the environment, which massively speed up the initial procedure (Giovannoni *et al.*, 1990; Ward *et al.*, 1990). These efforts have created one of the largest data sets for any individual gene (Maidak *et al.*, 2001). Currently, nearly 70,000 full or partial 16S rRNA sequences are deposited in the RDP-II database (Cole *et al.*, 2003).

The large subunit rRNA was shown to be useful for phylogenetic analysis of bacteria (Ludwig *et al.*, 1992; Ludwig *et al.*, 1995). Nevertheless, only a relatively low number of 23S rRNA sequences (less than one thousand) is available (Wuyts *et al.*, 2001). Apparently, 23S rDNA sequencing on a regular base is hampered by the two-fold length of this marker compared to the 16S rRNA, resulting in significantly higher efforts. Current interests in large subunit rRNA genes as a phylogenetic marker is mainly motivated by specific tasks (Ward *et al.*, 2000). However, there is a general interest in more extensive 23S rRNA gene sequence data sets, e.g., concerning the design of oligonucleotide probes used for the specific detection of selected groups or species by hybridization techniques, which are an integral part of the rRNA approach. An example for the significance of the 23S rRNA as a target molecule for oligonucleotide-based hybridization studies is the discrimination of the beta and the gamma subclass of the *Proteobacteria* which is possible with the 23S rRNA but not with the 16S rRNA (Manz *et al.*, 1992). Moreover, this marker provides additional probe target sites. Application of redundant and hierarchically structured sets of probes allow a more reliable identification of selected target sequences in the environment (Amann and Schleifer, 2001). With the current attempts to adapt high-throughput reverse hybridization formats (DNA-microarrays) to the field of microbial ecology (Guschin *et al.*, 1997; Liu *et al.*, 2001; Small *et al.*, 2001; Koizumi *et al.*, 2002; Loy *et al.*, 2002; Peplies *et al.*, 2003), a fast and easy application of large probe sets will become available in the near future.

In this study, we focused on 23S rRNA gene sequencing of alpha-proteobacteria, recently isolated from marine surface water (Eilers *et al.*, 2000; Eilers *et al.*, 2001). Members of the alpha-proteobacteria were found to be numerical important in marine picoplankton (Glöckner *et al.*, 1999). In particular, different studies have shown that the alpha-proteobacteria

are dominated by members of the "marine alpha" group (Gonzalez and Moran, 1997) in the picoplankton (Gonzalez and Moran, 1997; Eilers *et al.*, 2001). This cluster encompasses the genus *Roseobacter* and closely relatives. These bacteria are supposed to be major catalysts of the transformation of sulfur compounds, involved in global climate regulation (Gonzalez *et al.*, 1999). Currently, for this group no 23S rRNA sequences are available in the public databases. Five of the alpha-proteobacteria analyzed in this study belong to the marine alpha cluster. In a case study, comparative *in silico* design of 16S rRNA and 23S rRNA-targeting oligonucleotide probes for sequence clusters within the alpha-proteobacteria was conducted.

MATERIAL & METHODS

Bacterial strains and culture conditions. For 23S rRNA gene sequencing of alpha-proteobacteria from a marine pelagic environment, eleven strains isolated from the surface water of the German Bight by Eilers *et al.* (2000; 2001) were selected. The strains and their next relatives according to their 16S rRNA gene sequences are listed in Table 1. Accession numbers of the corresponding 16S rDNAs are AF173971-73, AF305498, AY007677-80, and AY007682-84. Cultures were maintained on solid agar plates with a synthetic seawater medium (MPM or MPM-m) according to Eilers *et al.* (2000; 2001).

PCR amplification. Amplification of the nearly full length 23S rRNA genes was performed with the modified primer 987RII (Ludwig *et al.*, 1992) (5'-CTTAKATGCYTTCAGC-3'; binding position 2744-2759 according to *E. coli* numbering by Brosius *et al.* (1981) and the newly designed primers LS48f (5'-AGGCGATGAARGACGT-3'; 44-59) and LS2669r (5'-CCGGTCCTCTCGTAC-3'; 2655-2669) using primer combinations listed in Table 1. Briefly, living cells were picked from culture plates, resuspended in 65 μ l of PCR-water (Sigma, Deisenhofen, Germany) and 3 times treated by freeze-thawing in liquid nitrogen for cell lysis. For PCR amplification 10x *Ex Taq* buffer (Takara, Otsu, Japan), 30 μ g of bovine serum albumin, deoxyribonucleoside triphosphates (final concentration 250 μ M of each), primer (final concentration 150 nM of each), and 2.5 units of *TaKaRa Ex Taq*-polymerase (Takara, Otsu, Japan) were added to a final volume of 100 μ l. The DNA-polymerase was added after an initial denaturing step at 94°C for 5 min, followed by 36 cycles at 94°C for 1 min, at 52°C for 1 min, and at 72°C for 6 min, followed by a final elongation step at 72°C for 12 min. Successful amplification was confirmed by analyzing PCR products on 1.5% agarose gels. PCR products were purified by the QIAquick PCR purification kit (Qiagen, Hilden, Germany).

The new primers were designed with the PROBE_DESIGN tool of the ARB package (<http://www.arb-home.de/>) based on a data set of 860 published as well as unpublished nearly full length 23S rDNA sequences (>2500 nt).

TABLE 1. Strains of which 23S rRNA genes were sequenced in this study

Strain	Next described relative ^a	% 16S rRNA similarity	Reference	Primer pair used for the amplification of the 23S rDNA and binding site ^b	Accession number of the 23S rDNA sequence
KT1117	<i>Sulfitobacter pontiacus</i>	96.3	Eilers <i>et al.</i> (2000)	LS48f - 987RfII	AY309493
KT0202a	<i>Roseobacter algicola</i>	94.1	Eilers <i>et al.</i> (2000)	LS48f - 987RfII	AY309490
JP7.1	<i>Roseobacter galaeiensis</i>	95.8	Eilers <i>et al.</i> (2001)	LS48f - 987RfII	AY309488
KT0917	<i>Sulfitobacter pontiacus</i>	96.1	Eilers <i>et al.</i> (2000)	LS48f - 987RfII	AY309492
JP88.1	<i>Roseobacter galaeiensis</i>	96.8	Eilers <i>et al.</i> (2001)	LS48f - 987RfII	AY309489
JP57	<i>Tetracoccus cechii</i>	87.9	Eilers <i>et al.</i> (2001)	LS48f - 987RfII	AY309483
JP22.1	<i>Mesorhizobium tianshanense</i>	90.8	Eilers <i>et al.</i> (2001)	LS48f - LS2669r (44-2669)	AY309485
JP66.1	<i>Mesorhizobium loti</i>	92.7	Eilers <i>et al.</i> (2001)	LS48f - LS2669r	AY309486
JP68	<i>Caulobacter maris</i>	96.3	Eilers <i>et al.</i> (2001)	LS48f - LS2669r	AY309487
JP13.1	<i>Erythrobacter litoralis</i>	97.3	Eilers <i>et al.</i> (2001)	LS48f - LS2669r	AY309484
KT0216	<i>Sphingomonas macrogoltabidus</i>	94.1	Eilers <i>et al.</i> (2000)	LS48f - LS2669r	AY309491

^a Based on 16S rRNA gene sequence similarity.

^b *E. coli* numbering according to Brosius *et al.* (1981).

Sequence determination. The 23S rRNA gene amplicons were directly sequenced using the eight newly designed primers listed in Table 2 in combination with the Big Dye v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, USA). For cycle sequencing a primer concentration of 5 pmol/μl, a total amount of 5 ng of DNA, and the following protocol was applied: 40 cycles at 96°C for 10 sec, at 50/52°C (see Table 2) for 5 sec, and at 60°C for 4 min. For data collection an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, USA) was used.

Data analysis and probe design.

The retrieved 23S rDNA sequences were added to a data set of 860 23S rDNA sequences, including 41 affiliated with the alpha-proteobacteria, using the program package ARB (see above). The tool ARB_EDIT was used for automatic sequence alignment. The alignment was checked by eye and corrected manually. The presented phylogenetic tree was reconstructed based on maximum-likelihood analysis of all 52 sequences affiliated with the alpha-proteobacteria and 7 sequences of the gamma-proteobacteria. The tree topology was evaluated by

maximum-parsimony and neighbor-joining analysis of the full data set in combination with different filters excluding highly variable positions. A consensus tree was constructed with topology corrected by consideration of the results of the various tree reconstruction algorithms.

The 16S rRNA gene sequences of the eleven strains were added to the 16S rDNA sequence database of the technical University of Munich (release June 2001). All sequences shorter than 1000 nucleotides were removed. Tree construction was done as described for the 23S rRNA sequences.

Oligonucleotide probes targeting the 23S rRNA and 16S rRNA were developed using the PROBE_DESIGN tool of the ARB package. Probes targeting the 16S rRNA were designed on a reduced data set from which all 16S rRNA alpha-proteobacterial sequences without a 23S rRNA equivalent were removed. Computed probes were manually rechecked for their group coverage and specificity using the tools ARB_EDIT and PROBE_MATCH on both data sets. Probes were not considered if a number of non-target sequences larger than 15% of the target sequences features a single terminal mismatch (less than five positions from the end) or if based on single nucleotide insertions compared to the next relatives of the targeted sequences.

Nucleotide sequence accession numbers. The almost full length 23S rRNA sequences determined in this study were deposited in Genbank under accession numbers AY309483-93 (Table 1).

TABLE 2. Oligonucleotide primers used for cycle sequencing of the 23S rDNA

Primer	Sequence (5'-3')	23S rDNA binding site ^a	Annealing temperature ^b
LS476r	CACCTTTCCCTCACGG	461 - 476	52°C
LS476f	CCGTGAGGGAAAGGTG	461 - 476	52°C
LS792r	TGATTGGCCTTTCACCC	776 - 792	52°C
LS792f	GGTGAAAGGCCAATCA	776 - 792	52°C
LS1902r	ATAGTTACGGCCGCCG	1902 - 1917	52°C
LS2242r	ACCGCCCCAGTCAAAC	2242 - 2257	52°C
LS2385f	CATAGTGATCCGGTGG	2385 - 2400	50°C
LS2669r	CCGGTCCTCTCGTAC	2655 - 2669	50°C

^a *E. coli* numbering according to Brosius *et al.* (1981).

^b Calculated according to the 4+2 rule of Suggs *et al.* (1981).

RESULTS

PCR amplification and sequencing of 23S rRNA genes. Nearly complete nucleotide sequences of 23S rRNA genes from 11 alpha-proteobacteria were determined using the PCR-primer combinations listed in Table 1 and the cycle sequencing primers listed in Table 2. All primers, designed in this study, were developed with the focus on targeting a maximum number of alpha-proteobacterial sequences within our data set. These primers are not "specific" for the alpha-proteobacteria, since they are targeting various sequences outside the group. Highly specific primers covering the major part of the alpha-group of the *Proteobacteria*, could not be derived from the underlying 23S rDNA data set.

Amplification of the 23S rRNA genes was initially tested with the primer pair LS48f - 987RII for all strains but was successful only for strains KT1117, KT0202a, JP7.1, KT0917, and JP88.1 (all affiliated with the "marine alpha" group) and strain JP57. For 23S rRNA gene amplification of strains JP22.1, JP66.1, JP68, JP13.1, and KT0216, primer pair LS48f - LS2669r was used, resulting in amplicons that were approx. 100 nucleotides shorter at the 3'-end of the gene compared to primer pair LS48f - 987RII.

Comparative sequence analysis. The phylogenetic trees reconstructed for 23S rRNA and 16S rRNA gene sequences of alpha-proteobacteria showed a consistent topology (Fig. 1). Different species of the genera *Acetobacter*, *Rickettsia*, *Rhodobacter*, and *Bartonella* as well as sequences affiliated with the marine alpha group formed clusters which branchings appeared stable and well separated from neighboring branchings based on the results of all tree reconstruction algorithms tested. However, various branching points could not be resolved on the underlying data sets. The resulting multifurcations are indicated by open circles in Fig. 1.

Comparative probe design. Both sequence collections, the 16S rRNA and the 23S rRNA data set, were used for a comparative *in silico* design of highly specific oligonucleotide probes targeting six sequence clusters (A-F, see also Fig. 1) and two single sequences affiliated with the alpha-proteobacteria. Only those probes were accepted which targeted all members of a particular cluster or sequence in combination with no nongroup hits.

Despite the two-fold length of the 23S rDNA compared to the 16S rDNA, for clusters B and C an identical number of 7 and 4 binding sites, respectively, were found on the two markers according to our criteria as shown in Table 1. For the clusters A and D, only one and none 16S rRNA specific probe, respectively, could be designed. In comparison, the 23S rRNA provided 8 and 7 available binding sites, respectively. These two clusters are characterized by a

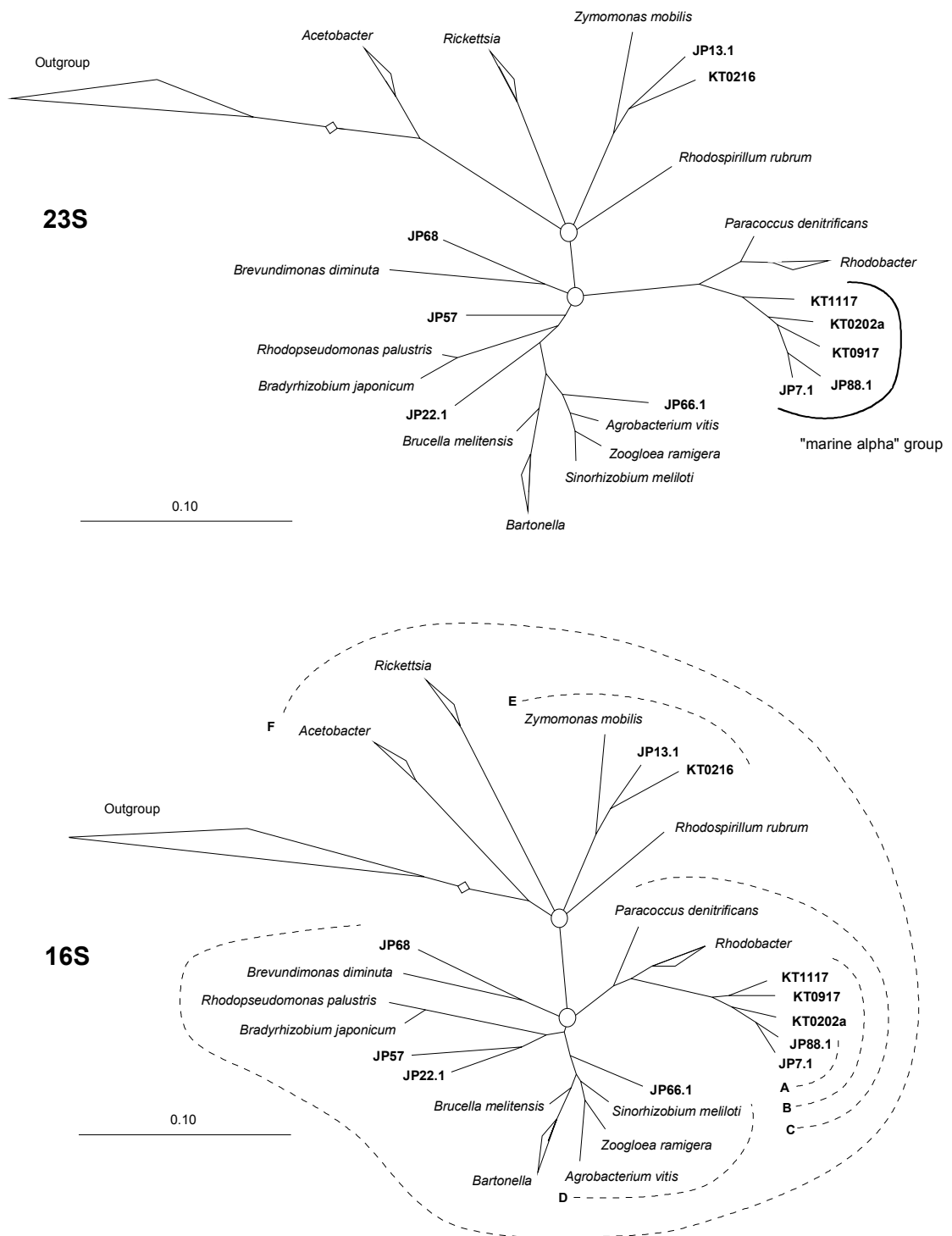


Fig. 1. Phylogenetic trees for the alpha-proteobacteria based on comparative sequence analysis of 16S rRNA and 23S rRNA genes. The bars indicate 10% estimated sequence divergence and the open boxes the roots. In the 16S RNA tree, clusters are indicated (A-F) which were selected for oligonucleotide probe design.

TABLE 3. Comparative oligonucleotide probe design for the 16S and 23S rRNA of alpha-proteobacteria

Cluster	Number of sequences	Minimum 16S rRNA similarity	Minimum 23S rRNA similarity	Probes targeting the 16S rRNA		Probes targeting the 23S rRNA		
				Total number	Binding site ^a	Total number	Sequence (5'-3')	Binding site ^a
A JP7.1 + JP88.1	2	96.7%	92.3%	1	AGACCCAAATCCGAAC	8	CCGGGTTTGGCTCTTATA GGCCATGCTTCCCAACAT GTGAGACCAAAGTCGATGA CCATGCCTAGATCACTCG CCTTCAACATCGTTCTCT CAGAATCCTGATACCAGT GGTTTCGATGCCACCAC CAGTGGTTCTTCAACCC	172 - 188 302 - 319 580 - 597 679 - 696 1650 - 1667 2090 - 2107 2121 - 2138 2668 - 2685
B marine alpha group	5	92.8%	88.6%	7	TATCATCTCTGGCGAAAG GGGGTTCTTTACCAGAT GCTAACCCCTCCGTATT CGTCAGATCGAGCCAGT GGAATGCTTAATCCGTTA CCGCTTATCAGGGCCAGT CCGCGTCACTGCTGTTACG	7	GCCCTACTTAATACGTCC GGTCAGTAAGGAGTACTT GTCGATGACCCATTATAC CATGCCATGATCACTCGG TCGGCTCATGGCTTGAAGC ACTCATGTCATCTCTC TATCTCAAGGGAGGCTCC	355 - 378 428 - 445 570 - 587 678 - 695 1139 - 1155 1250 - 1267 2156 - 2173
C marine alpha group, <i>Paracoccus denitrificans</i> <i>Rhodobacter</i> species	8	87.7%	84.2%	4	CAGCGTCAGTATCGAGCC CGGAATGCTTAATCCGTT AGGGTTGGTAAGGTTCTG CGTGTGTAGCCCCAACCCG	4	AGGACTACTTAGCCTTCG GAGCTGTACGCTATCTT TTACCCTGCTCAGATTAG GGCCTCCTTCTCGCGAAC	420 - 437 1085 - 1102 1325 - 1342 1691 - 1708
D <i>Sinarhizobium melleioli</i> , <i>Zoogloea ramifera</i> , <i>Agrobacterium vitis</i>	3	93.7%	92.2%	0	-	7	GAGTGGTCTCCCCATGTT GATGCGGGTTTCGGTCCCTC ATCCTCAGTTCGGTGCA CACTACCACAGGCCCCAC GTCTATACGGTGGAGCTA GGTTAGACATCCATGACG TTCCATATATAGCTACCC	401 - 418 726 - 743 1148 - 1165 1399 - 1426 1572 - 1599 2184 - 2201 2717 - 2734
E JP13.1 + KT0216, <i>Zymomonas mobilis</i>	3	88.8%	87.3%	3	AGCCGGAGCTTATTCTCC GCTAGCTCCCTCCGTATT TCGGGTGAATCCAACCTCC	8	CGGGTTGCTTACACAAA TCCCGGTTGATGGGCT AGACCCCGCTGTGCT TAGCCCTTCGGGCTGTG GCTACCCCTGGCCTGTG TCGTCTTGAAGCCGACTT CCACTCTGAAGCTGCAGT CTACCCGCCACAGAGGGT	170 - 186 783 - 800 941 - 958 1356 - 1373 1754 - 1771 1802 - 1815 2077 - 2094 2225 - 2242
F alpha-proteobacteria	52	80.5%	80.5%	0	-	0	-	-

TABLE 3. Continued

Cluster	Number of sequences	Minimum 16S rRNA similarity	Minimum 23S rRNA similarity	Probes targeting the 16S rRNA			Probes targeting the 23S rRNA		
				Total number	Sequence (5'-3')	Binding site ^a	Total number	Sequence (5'-3')	Binding site ^a
JP7.1	1	-	-	6	TCACTCCGAGAGTGGCT	65 - 101	10	CAGTTCATAAGTACCTGG	143 - 159
					CCTTAAAGAAAGGCACGT	130 - 147		CTCACACTCAAGGCTCGG	267 - 295
					ACACCAACAAGCATGCTT	838 - 856		ATACGTCATCATGCTT	368 - 365
					CTGGCGGTCCAAGGATGT	994 - 1011		TCAGGCGCAAGGGCCCT	538 - 556
					GATGAGGTTGCCCTCGT	1102 - 1119		CGGGGTATTCCGACGACG	815 - 832
					GCCCCCTCCGAAAAAGGTT	1446 - 1459		GGACATCCCACCTGCTT	1027 - 1044
KT0917	1	-	-	5	TCCCGAAGGACTGGGCTC	64 - 89	12	AAAGATATAGGCTGCCCT	140 - 141
					CGTAGATCTGGGTACGTT	129 - 146		ATACGTCCCATTGAGCTT	346 - 365
					ACTTAGACCAGGAGTTT	640 - 657		TCTAACTCAACGCCCTAT	644 - 661
					TAAAGGAACGCCGATCT	1014 - 1031		AGGCTGTAGGCCCCCAT	877 - 894
					CGCAGCGTCGTCAGGTAG	1426 - 1443		GCATCTAAACTTCGGCT	1150 - 1167
								ATGCCITFACAGCCGGCT	1214 - 1229
								TAGTCGCCAGATCCGT	1427 - 1444
								CCAGGAACCGTCTTGAAT	1548 - 1568
								GGTCTCATGAAGGGCTAT	1571 - 1600
								CAAAAGCCGGCCCTT	1700 - 1730
								CATCAGGCAAAATAAGGG	2177 - 2194
								TTGGCATACCAACTGGTC	2689 - 2706

^a *E. coli* numbering according to Brosius *et al.* (1981)

small size (2 and 3 sequences) and a high minimum sequence similarity (96.7 and 93.7% for the 16S rRNA and 92.3 and 92.2% for the 23S rRNA) compared to the other 4 clusters (Tab. 3). A similar situation was observed for the probes targeting single sequences. For strains JP7.1 and KT0917, 6 and 5 probes targeting the 16S rRNA and 10 and 12 probes targeting 23S rRNA were obtained, respectively. For cluster E (3 sequences, 88.8 and 87.3% minimum sequence similarity), 3 probes could be designed based on the 16S rRNA data set and 8 sequences based on the 23S rRNA data set.

It was not possible to design an appropriate probe for the alpha-proteobacteria (52 sequences, 80.5% minimum sequence similarity) based on both data sets, even when relaxing criteria for probe quality to a minimum value of 70% for target group coverage and a maximum number of 75 nongroup hits.

DISCUSSION

Primer design, PCR amplification and primary structure of 23S rRNA genes. With the "general" primer 987RII we could not amplify the 23S rRNA genes of strains JP22.1, JP66.1, JP68, JP13.1, and KT0216 which are widely spread in the group of alpha-proteobacteria (81.1% minimum 23S rRNA sequence similarity). Apparently, this reflects the relatively high degree of variability among 23S rRNA sequences which was already pointed out by Ludwig *et al.* (1995). Construction of primers, (i) specific for the alpha-group of *Proteobacteria*, (ii) with an acceptable group coverage, and (iii) flanking the nearly full length 23S rRNA was not feasible based on our 23S rRNA data set. In conclusion, for further primer design, we focused on achieving maximum group coverage for the alpha-proteobacteria, resulting in primer LS48f which targets all alpha-proteobacteria and 219 sequences outside this group in our data set, and primer LS2669r which targets all alpha-proteobacteria excluding the genus *Rickettsia* and 642 sequences outside this group in our data set. For construction of internal sequencing primers the same strategy was applied. All six 23S rRNA binding positions targeted, allowed sequencing of the corresponding gene sections of all 11 strains using identical primer sequences, respectively.

In comparison to available 23S rRNA gene sequences of alpha-proteobacteria, the new sequences display an overall structure similarity. However, within the variable regions of the gene substantial differences can be seen concerning the presence, length, and shape of primary and higher order structure elements. In general, these differences correlate well with the

phylogenetic relationships. Although a considerable degree of overall sequence lengths variation was found within the *Proteobacteria* in the study of Ludwig *et al.* (1995), similar lengths for the members of the alpha-proteobacteria were observed (2727-2755 nucleotides). In our study, lengths of the 23S rRNA gene fragments sequenced varied for strains KT1117, KT0202a, JP7.1, KT0917, JP88.1, and JP57 (PCR primer pair LS48f and 987RII) between 2568 and 2736 nucleotides, and for strains JP22.1, JP66.1, JP68, JP13.1, and KT0216 (PCR primer pair LS48f and LS2669r) between 2463 and 2541 nucleotides. In general, the 16S rRNA sequence similarities of *Proteobacteria* are higher than the corresponding 23S rRNA similarities (Ludwig *et al.*, 1995). This finding is supported by our data (compare Tab. 3), also indicating a higher number and fraction of variable residues within the 23S rRNA.

We also tried to amplify the 16S-23S rRNA intergenic spacer region of strains KT0202a, JP22.1, JP13.1, and KT0216 using the primer pair 1492f (reverse complementary to the 3'-16S rRNA primer 1492r [Buchholz-Cleven *et al.*, 1997]) and LS48r (reverse complementary to the 5'-23S rRNA primer LS48f). Although testing various PCR-conditions, no amplicons could be obtained in the range of length to be expected (data not shown). Scheinert *et al.* (1996) found that spacer lengths varied between 280 and 1300 bp for a wide range of microorganisms. Our results suggest that the 16S rRNA and 23S rRNA genes of the strains listed above are not linked.

Phylogenetic analysis. Comparison of the 16S rRNA and 23S rRNA-derived phylogeny of the alpha-proteobacteria analyzed in this study revealed a basic consistency. This finding is in good agreement with the results of previous studies (Leffers *et al.*, 1987; Ludwig *et al.*, 1992; Ludwig *et al.*, 1995; Ward *et al.*, 2000). The clustering of different sequences including the marine alpha group as shown in Fig. 1, was supported by both phylogenetic markers. In an earlier study of Ludwig *et al.* (1995), phylogenetic analysis based on 23S rRNA gene sequences was conducted for 8 different species affiliated with the alpha-group of *Proteobacteria*. The close relationship of *Rhodobacter* and *Paracoccus* as well as of *Bradyrhizobium* and *Rhodopseudomonas* could be confirmed by analysis of our extended data set comprising 52 23S rRNA sequences of alpha-proteobacteria. Also the phylogenetic distinctness of the marine alpha group suggested by 16S rRNA sequence analysis, previously (Gonzalez and Moran, 1997), is supported by the 23S rRNA-derived phylogeny of this study.

The branching points of most of these stable sequence clusters could not be resolved by the different treeing methods and filters applied in case of the 23S rRNA as well as the 16S rRNA data set. Such findings may reflect the limitation of the ribosomal RNA as a

phylogenetic marker (Ludwig *et al.*, 1995). Also within the stable sequence clusters observed, differences in the local tree topologies were found comparing the results of the different tree reconstruction algorithms applied. In these cases, topologies according to the maximum-likelihood analysis are shown for both markers. A stable positioning of single sequences such as strains JP57, JP22.1, JP66.1, or *Paracoccus denetrificans* in our 23S rRNA data set is generally hampered by the lack of sufficient sequence information needed for reliable tree reconstruction. However, also significant differences in local branching patterns observed previously, are supposed to reflect different resolution power rather than conflicting information content of these two types of rRNA (Ludwig *et al.*, 1995). Generally, the 23S rRNA should provide a better phylogenetic resolution due to its higher number and fraction of variable residues compared to the 16S rRNA.

Oligonucleotide probe design. It has been demonstrated that the large subunit RNA is not only suited for phylogenetic analysis but also provides good target sites for specific hybridization probes (Stackebrandt *et al.*, 1991; Ehrmann *et al.*, 1992; Manz *et al.*, 1992; Spring *et al.*, 1992; Fuchs *et al.*, 2001). Oligonucleotide probes can be designed for groups of different phylogenetic depths but the sequence variation of rRNA molecules is usually not sufficient to design strain-specific probes, even in case of the 23S rRNA with its higher degree of variability compared to the 16S rRNA (Ludwig and Schleifer, 1994). However, we also targeted single strains with relatives of high sequence similarity in our data set. Here, the suitability of the 16S rRNA and 23S rRNA gene for discrimination of selected groups of sequences (or single sequences) with signature regions of 18 nucleotides in length should be systematically evaluated and compared. Since probe design was conducted on a 23S rRNA data set comprising 871 nearly full length sequences and a 16S rRNA data set substantially reduced to guarantee comparability of the results, all probes listed in Table 3 should be reconsidered before experimental application.

As expected, the 23S rRNA provided significantly more probe binding sites for most of the sequence clusters or single sequences targeted. Whereas probe design for larger clusters such as the complete group of alpha-proteobacteria failed on both markers, or showed similar results as for cluster C, the 23S rRNA provided a higher number of probes for single sequences or clusters of high sequence similarity targeted. This finding is in good agreement with the higher degree of variability in 23S rRNA genes and their two-fold length and number of higher order structure elements compared to the 16S rRNA (Ludwig and Schleifer, 1994). However, no clear correlation for phylogenetic depth and number of available probe binding sites was

observed for both types of rRNA. For example, a single specific probe was available on the 16S rRNA for cluster A including strains JP7.1 and JP88.1, but 6 binding sites could be used to discriminate strain JP7.1 from all other sequences in the 16S rRNA data set, and 7 probes specific for the whole marine alpha group were obtained.

Since the first complete 23S rRNA sequence, that of *Escherichia coli*, has been published in 1980 (Brosius *et al.*, 1980), the large subunit rRNAs did not attract major interest as phylogenetic marker or reservoir of additional probe target sites. Our data indicate that not only additional phylogenetic probes are available by also targeting the 23S rRNA but potentially also better discrimination of closely related bacteria can be achieved compared to 16S rRNA-based probe design as for cluster D in our case study. However, clustering of few sequences may also be an artifact due to the low number of sequences processed for tree reconstruction. More extensive 23S rRNA data sets are needed to resolve these uncertainties and to allow for a reliable assessment of probe quality *in silico*. Even without systematic 23S rRNA gene sequencing these data will become available from the numerous genome sequencing projects.

ACKNOWLEDGEMENTS

We thank Silke Wetzels for expert technical assistance. This work was supported by the German Ministry of Education and Research (BMBF, contract 0311833A) and by the Max Planck Society.

REFERENCES

- Amann, R.I., and Schleifer, K.-H. (2001) Nucleic acid probes and their application in environmental microbiology. p. 67-82. In *G. Garrity, D.R. Boone, and R.W. Castenholz (ed.), Bergey's Manual of Systematic Bacteriology, 2nd ed., vol 1., Springer Verlag, New York.*
- Amann, R.I., Ludwig, W., and Schleifer, K.-H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**: 143-169.
- Brosius, J., Dull, T.J., and Noller, H.F. (1980) Complete nucleotide sequence of a 23S ribosomal RNA gene from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **77**: 201-204.
- Brosius, J., Dull, T.J., Sleeter, D.D., and Noller, H.F. (1981) Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J. Mol. Biol.* **148**: 107-127.
- Buchholz-Cleven, B., Rattunde, B., and Straub, K. (1997) Screening for genetic diversity of isolates of anaerobic Fe(II)-oxidizing bacteria using DGGE and whole-cell hybridization. *Syst. Appl. Microbiol.* **20**: 301-309.
- Cole, J.R., Chai, B., Marsh, T.L., Farris, R.J., Wang, Q., Kulam, S.A. et al. (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.* **31**: 442-443.
- Ehrmann, M., Ludwig, W., and Schleifer, K.H. (1992) Species specific oligonucleotide probe for the identification of *Streptococcus thermophilus*. *Syst. Appl. Microbiol.* **15**: 453-455.
- Eilers, H., Pernthaler, J., Glöckner, F.O., and Amann, R. (2000) Culturability and in situ abundance of pelagic bacteria from the North Sea. *Appl. Environ. Microbiol.* **66**: 3044-3051.
- Eilers, H., Pernthaler, J., Peplies, J., Glöckner, F.O., Gerds, G., and Amann, R. (2001) Isolation of novel pelagic bacteria from the German Bight and their seasonal contributions to surface picoplankton. *Appl. Environ. Microbiol.* **67**: 5134-5142.
- Fuchs, B.M., Syutsubo, K., Ludwig, W., and Amann, R. (2001) In situ accessibility of *Escherichia coli* 23S rRNA to fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **67**: 961-968.
- Giovannoni, S.J., Britschgi, T.B., Moyer, C.L., and Field, K.G. (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60-63.
- Glöckner, F.O., Fuchs, B.M., and Amann, R. (1999) Bacterioplankton composition in lakes and oceans: a first comparison based on fluorescence in situ hybridization. *Appl. Environ. Microbiol.* **65**: 3721-3726.

- Gonzalez, J.M., and Moran, M.A. (1997) Numerical dominance of a group of marine bacteria in the alpha-subclass of the class Proteobacteria in coastal seawater. *Appl. Environ. Microbiol.* **63**: 4237-4242.
- Gonzalez, J.M., Kiene, R.P., and Moran, M.A. (1999) Transformation of sulfur compounds by an abundant lineage of marine bacteria in the alpha-subclass of the class Proteobacteria. *Appl. Environ. Microbiol.* **65**: 3810-3819.
- Guschin, D., Yershov, G., Zaslavsky, A., Gemell, A., Shick, V., Proudnikov, D. et al. (1997) Manual manufacturing of oligonucleotide, DNA, and protein microchips. *Anal. Biochem.* **250**: 203-211.
- Koizumi, Y., Kelly John, J., Nakagawa, T., Urakawa, H., El Fantroussi, S., Al Muzaini, S. et al. (2002) Parallel characterization of anaerobic toluene- and ethylbenzene-degrading microbial consortia by PCR-denaturing gradient gel electrophoresis, RNA-DNA membrane hybridization, and DNA microarray technology. *Appl. Environ. Microbiol.* **68**: 3215-3225.
- Leffers, H., Kjems, J., Ostergaard, L., Larsen, N., and Garrett, R.A. (1987) Evolutionary relationships among archaeobacteria: a comparative study of 23S ribosomal RNA of a sulfur-dependent extreme thermophile, an extreme halophile, and a thermophilic methanogen. *J. Mol. Biol.* **195**: 43-62.
- Liu, W.T., Mirzabekov, A.D., and Stahl, D.A. (2001) Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. *Environ. Microbiol.* **3**: 619-629.
- Loy, A., Lehner, A., Lee, N., Adamczyk, J., Meier, H., Ernst, J. et al. (2002) Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl. Environ. Microbiol.* **68**: 5064-5081.
- Ludwig, W., and Schleifer, K.H. (1994) Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiol. Rev.* **15**: 155-173.
- Ludwig, W., Kirchhof, G., Klugbauer, N., Weizenegger, M., Betzl, D., Ehrmann, M. et al. (1992) Complete 23S ribosomal RNA sequences of gram-positive bacteria with a low DNA G + C content. *Syst. Appl. Microbiol.* **15**: 487-501.
- Ludwig, W., Rossello-Mora, R., Aznar, R., Klugbauer, S., Spring, S., Reetz, K. et al. (1995) Comparative sequence analysis of 23S rRNA from Proteobacteria. *Syst. Appl. Microbiol.* **18**: 164-188.
- Maidak, B., L., Cole, J., R., Lilburn, T., G., Parker, C.T., , Jr., Saxman, P., R., Farris, R., J. et al. (2001) The RDP-II (Ribosomal Database Project). *Nucleic Acids Res.* **29**: 173-174.

- Manz, W., Amann, R., Ludwig, W., Wagner, M., and Schleifer, K.-H. (1992) Phylogenetic oligodeoxynucleotide probes for the major subclasses of Proteobacteria: Problems and solutions. *Syst. Appl. Microbiol.* **15**: 593-600.
- Olsen, G.J., Lane, D.J., Giovannoni, S.J., Pace, N.R., and Stahl, D.A. (1986) Microbial ecology and evolution: a ribosomal rRNA approach. *Annu. Rev. Microbiol.* **40**: 337-365.
- Pace, N.R., Stahl, D.A., Lane, D.J., and Olsen, G.J. (1986) The Analysis of natural microbial populations by ribosomal RNA sequences. *Adv. Microb. Ecol.* **9**: 1-55.
- Peplies, J., Glöckner, F.O., and Amann, R. (2003) Optimization strategies for the DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes. *Appl. Environ. Microbiol.* **69**: 1397-1407.
- Scheinert, P., Krausse, R., Ullmann, U., Söller, R., and Krupp, G. (1996) Molecular differentiation of bacteria by PCR amplification of the 16S-23S rRNA spacer. *J. Microbiol. Meth.* **26**: 103-117.
- Small, J., Call, D.R., Brockman, F.J., Straub, T.M., and Chandler, D.P. (2001) Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. *Appl. Environ. Microbiol.* **67**: 4708-4716.
- Spring, S., Amann, R., Ludwig, W., Schleifer, K.-H., and Petersen, N. (1992) Phylogenetic diversity and identification of nonculturable magnetotactic bacteria. *Syst. Appl. Microbiol.* **15**: 116-122.
- Stackebrandt, E., Witt, D., Kemmerling, C., Kroppenstedt, R., and Liesack, W. (1991) Designation of Streptomyces 16s and 23s Ribosomal Rna-Based Target Regions for Oligonucleotide Probes. *Appl. Environ. Microbiol.* **57**: 1468-1477.
- Suggs, S.V., Hirose, T., Miyaki, T., Kawashima, E.H., Johnson, M.J., Itakura, K. et al. (1981) Use of synthetic oligodeoxyribonucleotides for the isolation of specific cloned DNA sequences. p. 683-693. In D. Brown and C.F. Fox (ed.), *Developmental biology using purified genes*, Academic press, New York.
- Ward, D.M., Weller, R., and Bateson, M.M. (1990) 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**: 63-65.
- Ward, N.L., Rainey, F.A., Hedlund, B.P., Staley, J.T., Ludwig, W., and Stackebrandt, E. (2000) Comparative phylogenetic analyses of members of the order Planctomycetales and the division Verrucomicrobia: 23S rRNA gene sequence analysis supports the 16S rRNA gene sequence-derived phylogeny. *Int. J. Syst. Evol. Microbiol.* **50**: 1965-1972.
- Wuyts, J., De Rijk, P., Van de Peer, Y., Winkelmans, T., and De Wachter, R. (2001) The European Large Subunit Ribosomal RNA Database. *Nucleic Acids Res.* **29**: 175-177.

Danksagung

Abschließend möchte ich einer Reihe von Personen danken, ohne deren Ein- und Mitwirken diese Arbeit in der vorliegenden Form nicht hätte entstehen können.

Beginnen möchte ich mit Herrn Prof. Dr. Rudolf Amann. Er hat mir nicht nur die Möglichkeit gegeben, in seinem Labor zu arbeiten und mein Promotionsvorhaben betreut, sondern stand auch jederzeit für anregende und hilfreiche Diskussionen zur Verfügung. Hierfür, und für das große Maß an Freiheit, das er mir gelassen hat, möchte ich ihm sehr herzlich danken.

Mein besonderer Dank gilt auch Herrn Prof. Dr. Dietmar Blohm. Mit dem Gensensorik-Verbund hat er ein Projekt initiiert, das meine Arbeit in der durchgeführten Form ermöglicht und in ihrem Verlauf ein vielschichtiges Forum geboten hat. Auch danken möchte ich ihm für die Möglichkeit, die Gerätschaften in seinem Labor zu nutzen, das Interesse an meiner Arbeit und eine stets offene Tür.

Dr. Frank Oliver Glöckner hat als mein direkter Betreuer maßgeblich zum Gelingen dieser Arbeit beigetragen. Er hat mir die Möglichkeit gegeben, mich zu entfalten, stand mir aber immer mit Rat und Tat zur Seite. Seine Sicht der Dinge hat sehr geholfen, niemals das Wesentliche aus den Augen zu verlieren. Hierfür danke ich ihm außerordentlich.

Unverzichtbar für diese Arbeit war die Vielzahl von Slides, mit denen mich Frau Britta Meyer-Schlosser aus der Arbeitsgruppe von Herrn Prof. Dr. Dieter Wöhrle versorgt hat. Vielen Dank hierfür.

Denja Drutschmann hat mich über viele Monate hinweg bei der Nutzung der Spotting-Roboter unterstützt. Hierfür, und für viele weitere "Kleinigkeiten", möchte ich mich ganz besonders bedanken.

Als von unschätzbarem Wert hat sich Andreas Ellrott erwiesen, der jedes Hardware- und Software-assoziierte Problem, das ich an ihn herangetragen habe, zuverlässig lösen konnte (nur wie man pdf-files erfolgreich in Word-Dokumente einbindet, bleibt weiterhin ein Mysterium). Silke Wetzels sei für die tatkräftige Unterstützung im 23S-Projekt gedankt.

An dieser Stelle soll auch die gesamte Arbeitsgruppe "Molekulare Ökologie" erwähnt sein. Das angenehme Arbeitsklima war eine zusätzliche Quelle von Energie. Außerdem freue mich sehr, im privaten Sektor einen "Trend angeschoben" zu haben ;o) der uns viele interessante Erlebnisberichte beschert hat.

Zu guter letzt, der größte Dank an Jenny und Emma, die insbesondere in der (doch recht langwierigen) "Endphase" viel Geduld mit mir haben mussten und dies tapfer ertragen haben. Ihren Rückhalt möchte ich nicht missen. Auch nicht vergessen werden dürfen alle Eltern und Schwiegereltern, die immer zur Stelle waren, wenn unsere Kleinfamilie organisatorisch aufgefangen werden musste. Vielen Dank!