Journal of Industrial Microbiology (1996) 17, 159–169

A molecular approach to search for diversity among bacteria in the environment

H Rheims, FA Rainey and E Stackebrandt

DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Germany

Molecular 16S rDNA-based techniques were applied to a peat sample from northern Germany in order to investigate the bacterial diversity present and compare the clone sequences with those obtained from similar studies on other terrestrial samples. Genomic DNA was extracted from the peat matrix by a direct lysis procedure. 16S rRNA genes were amplified using PCR primers targeting conserved regions of bacterial 16S rDNA. 16S rDNA fragments were blunt end cloned into a plasmid vector and the resulting clone library of 262 sequences was screened by hybridization with different oligonucleotide probes and sequence analysis of randomly selected clones. The 16S rDNA insert of 76 clones was partially sequenced. Clones identified either by hybridization or by sequence analysis fell into three phyla. As judged by hybridization with a specific oligonucleotide probe, 42% of the clones represented members of the alpha subclass of Proteobacteria. Twenty-five of these clones were selected randomly for sequence analysis; none could be assigned to any of the known genera of this subclass. The second largest clone group comprises 15% of the clones and clusters around *Acidimicrobium ferrooxidans* and *Rubrobacter radiotolerans*, both of which are remotely related to members of the order Actinomycetales. The third major clone cluster (10%) was moderately to remotely related to the *Acidobacterium capsulatum* phylum. Of the additional clones sequenced, a few could be assigned to other subclasses of *Proteobacteria*, the *Verrucomicrobium* phylum and the phylum of spirochetes. Comparison of the results presented here with those from other environments reveals a significant number of common clone clusters. As the vast majority of sequences retrieved from any of the marine and terrestrial samples investigated so far by molecular methods indicate the presence of novel bacterial species it can be assumed that a huge, as yet untapped biotechnological potential is present in the environment.

Keywords: molecular ecology; 16S rDNA; gene library; oligonucleotide probes; uncultured organisms; phylogeny

Introduction

Most of our knowledge of bacterial communities is based on studies of pure cultures that were isolated by selective methods for cultivating different groups of microorganisms. The majority of these enrichment and cultivation methods have a long history and so it is not surprising that the application of these growth conditions will lead to successful isolation of those taxa that have been found before. Small changes in the composition of cultivation media can lead to the recognition of a significantly broader spectrum of taxa than those described. This strategy has recently been demonstrated by the isolation of novel types of planctomycetes [22], the phylogenetic diversity of which was unraveled by 16S rDNA sequence analysis of the isolates [28]. It can be assumed that only a fraction of the phylogenetic diversity of culturable organisms has been explored. Novel strategies, including innovative isolation and cultivation approaches as well as new concepts for handling huge numbers of strains must be developed.

One new discipline in microbiology which may have a significant impact on ecological studies is that which has been termed 'molecular microbial ecology'. Nucleic acids are isolated directly from an environmental sample and the genomic diversity is assessed either by sequence analysis of cloned PCR products of 16S rDNA or by electrophoretic separation of PCR products via temperature gradient or denaturing gel electrophoresis [16,17]. Results are available for a wide range of environments, including the open ocean and coastal regions of the Pacific and Atlantic, soil, hot springs, activated sludge, river sediment, a bioleach site, rhizosphere and human tissue.

Molecular studies on soil have included the elucidation of the microbial population in the rhizosphere of rice [13,14], and moderately acidic soil (pH 4–4.5), from a forested Australian site [11,12,23], forest soils from England (M Embley, personal communication) and Finland [21], a thermal soil in New Zealand [18], a peat bog in Germany [19], and a bioleach sample from Queensland, Australia [8]. The main goal of these studies was the assessment of molecular tools for the determination of natural diversity, omitting the cultivation step. Since a gram of soil contains millions of cells and only a few hundred clones were ever investigated by oligonucleotide probing and/or sequence analysis, the results could not be interpreted to reflect the actual distribution of taxa in such environments, nor did they allow speculation about species richness and species abundance. The results of environmental studies revealed the presence of novel ubiquitous and broad lineages, but also single or small numbers of geographically restricted but phylogenetically diverse lineages. These include a new and very deep branching lineage within the actinomycete line of descent, sequences which are related to the orders

Correspondence: E Stackebrandt, DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Germany
Received 22 March 1996; accepted 30 August 1996

Planctomycetales and *Verrucomicrobiales*, and several new lines within the alpha subgroup of Proteobacteria. Only a few isolated clone sequences were related to those organisms routinely isolated from soil, eg of bacilli and spore-forming actinomycetes, including streptomycetes.

In this communication we report the molecular analysis of a low pH peat bog sample from northern Germany. One aspect of this work is the recovery of clone groups previously obtained from other environments. In the majority of cases, those clone groups can now be associated at the phylum level to cultured bacteria. Since the sequence similarities can only be identified at the level of higher taxa, it can be assumed that the cloned sequences originate from strains that represent novel genera and their species, corresponding to new biotechnological potential.

Materials and methods

Sampling site and procedure

Samples were taken in April 1994 from a peat bog (depth: 40 cm) near Gifhorn, Lower Saxony, Germany (10° 33′ E, 52° 30′ N). A core drill was driven horizontally into the side wall of a freshly cut drainage ditch for 50 cm. The first and the last 5 cm of the peat core were discarded while the rest of the peat was put into sterile tubes and immediately placed on dry ice until storage at –70°C. All material used to take the samples was either autoclaved or flamed prior to use. Acetate and propionate were determined by HPLC according to Wagner *et al* [26].

Extraction and purification of total DNA from peat

Genomic DNA was extracted using a modified direct lysis technique. Five grams of peat were taken from a sample and thawed. A solution containing 20 ml of saline-EDTA-buffer (0.15 M NaCl, 0.01 M EDTA – disodiumsalt, pH 8.0) was added. Following the addition of 2 g PVPP (Polyvinylpolypyrrolidone, Sigma, St Louis, MO, USA), and 150 mg lysozyme (in 5 ml of saline-EDTA-buffer), the solution was mixed and then treated with a homogenizer (Ultra-Turrax, Janke and Kunkel, Staufen, Germany) for 45 s. After incubation at 37°C for 30 min, 200 μl of Proteinase K (1% w/v) and 500 μl of 25% (w/v) SDS were added, homogenized by vigorous shaking and the homogenate was further incubated at 70°C for 45 min. A centrifugation step was performed at $5800 \times g$ for 20 min to separate soil debris and PVPP with bound co-extracted humic acids from the nucleic acids. The supernatant was transferred to a fresh tube and subsequently extracted with an equal volume of TE-saturated phenol (TE: 0.01 U Tris-Cl, 0.001 U EDTA-Na_2), phenol/chloroform (1 : 1 vol : vol) and chloroform, each step followed by centrifugation as indicated above. Following the last extraction step, 1/10 volume of 3 M sodium acetate (pH 5.2) was added and nucleic acids were precipitated with absolute ethanol. The DNA pellet was dissolved in sterile water and further purified with the Prep-A-Gene™ DNA Purification Kit (Bio-Rad, Hercules, CA, USA). The presence of humic acids in the sample was verified by electrophoresis of a humic acid standard (H1, 675-2; Aldrich, Steinheim, Germany) alongside the environmental DNA. The significant reduction of humic acids within the solution of genomic DNA after the purification procedure was checked by performing electrophoresis of a DNA aliquot on a 1% (w/v) agarose gel.

Construction of a 16S rDNA clone library

The 16S ribosomal DNA was amplified from the total DNA using Ultma™ DNA Polymerase (available with the specific reaction buffer from Perkin Elmer, Weiterstadt, Germany) and bacterial 16S rDNA oligonucleotide primers 27f and 1385r (Table 1). The PCR protocol is a modification of the protocol recommended by Perkin Elmer for the use of Ultma™ Polymerase, in that 4 μl of the 25 mM $MgCl_2$ solution were used. PCR was performed on a DNA thermal cycler (PE 9600, Perkin Elmer), amplification was for 28 cycles (annealing for 1 min at 52°C, elongation for 2 min at 72°C, denaturation for 1 min at 93°C). PCR products were purified with the Prep-A-Gene™ DNA Purification Kit (Biorad).

Cloning was performed with Epicurian Coli® XL1-Blue MRF'KAN competent cells (Stratagene, Heidelberg, Germany) and the pCR-Script™ SK(+) Cloning Kit (Stratagene). Modifications from the manufacturer's protocol included incubation of the ligation mix for at least 2 h and heat pulsing of the competent cells for 50 s at 42°C, for which best results were observed.

Plasmids were extracted by boiling the bacterial cells (10 min at 98°C) and pelleting the debris by centrifugation (1 min at $15\,000 \times g$). The supernatant was transferred into fresh tubes. The inserts were further amplified using Taq Polymerase and M13(–20) and M13rev Primers, applying the PCR cycles as indicated above.

Determination and evaluation of the nucleotide sequences

Sequencing reactions of the 16S rDNA fragments were performed with the Taq DyeDeoxy™ Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and electrophoresed using an automated sequence analyzer (ABI model 373A). The sequencing result files were manually aligned with ESEE (The Eyeball Sequence Editor) [3] and then included in the alignment editor ae2 [14]. The lengths of the analyzed 16S rDNA sequences range from about 300–400 nucleotides for partially sequenced clones to around 1370 nucleotides for fully sequenced clones. All of the sequences were included in the CHECK_CHIMERA program [14] to investigate the presence of chimeric structures.

Sequences from the major lines of descent [14] were used for comparison with the 16S rDNA clone sequences. Similarity values were transformed into phylogenetic distance values that compensate for multiple substitutions at any given site in the sequence [10]. The neighbor joining method contained in the PHYLIP package [6] was used in the construction of phylogenetic dendrograms.

DNA blotting and hybridization with oligonucleotide probes

Three microlitres of the 16S rDNA PCR product were blotted onto a membrane (Hybond N, Amersham International, Little Chalfont, UK), using a dot blot apparatus (Bio-Rad Bio-Dot® Microfiltration Apparatus) and fixed by UV (260 nm) for 1 min 40 s. Hybridization and detection were

Table 1 Sequences of the oligonucleotide primers and probes used in this study

Name	Sequence (5′–3′)	
27f	GAGTTTGATCCTGGCTCAG	16S rDNA amplification
1385r	CGGTGTGT(A/G)CAAGGCCC	16S rDNA amplification
519r	G(T/A)ATTACCGCGGC(T/G)GCTG	universal 16S rDNA probe
Alf 1b	CTGGCTCAGA(A/G)CGAACG	Taxon-specific probe
Gram-positive	TCATCATGCCCCTTATG	Taxon-specific probe
Planctomycete	GGC(A/G)TGGATTAGGCATGC	Taxon-specific probe
Spirochete	CGTTTTAAGCATGCAAGT	Taxon-specific probe
Probe TM2 group	CAAAGCAGCAATGCGCTT	Taxon-specific probe
Probe TM10 group	GATACCCTAGAGAAATTTAG	Taxon-specific probe

performed according to 'The DIG System User's Guide for Filter Hybridization' (Boehringer, Mannheim, Germany), using the DIG Luminescent Detection Kit (Boehringer). Oligonucleotide probes were synthesized by the Pharmacia Biotech Oligo-Synthesis-Service (Pharmacia, Roosendaal, The Netherlands) and labeled using the DIG Oligonucleotide 3′-End Labeling Kit (Boehringer).

Nucleotide sequence accession numbers

The TM sequences are available at the EMBL Nucleotide Sequence Database (Cambridge, UK) under accession numbers X97076 to X97117.

Results

Chemical compounds of the peat sample

The peat sample had a pH of 2.7. Acetate and propionic acids were not detected (H Drake, University of Erlangen, personal communication). The dry substance was 16.8% of which 29.7% were dry organic substances. Extraction with nitrohydrochloric acid revealed substantial amounts of Al (490 mg kg^{-1} of dry substance), Ca (690), Fe (550), Mg (370) and S (930). Main hydrocarbons (100–1000 mg kg^{-1} of dry substance) were long chain non-polar alkanes (docosane and hexatricontane).

Isolation of genomic DNA

As peat is a degradation product of plant material, substantial amounts of humic acids, known to interfere with DNA isolation, need to be removed during the procedure of DNA extraction from peat. Here, polyvinylpolypyrrolidone (PVPP) was used to bind co-extracted humic acids. Comparison of the isolated DNA by agarose gel electrophoresis before and after PVPP treatment showed a significant difference in the amount of humic acids. Without PVPP treatment humic acids appear as a bright band in an ethidium bromide-stained agarose gel under UV light, migrating faster than the genomic DNA obtained after gentle cell lysis. The 16S rDNA from this DNA preparation could not be amplified. After purification with PVPP, the amount of humic acids was reduced. Following further purification by a sodium acetate/ethanol precipitation step, humic acids were virtually absent and 16S rDNA PCR amplification products could be obtained from DNA.

Screening of the clone library with taxon-specific oligonucleotides

The clone library comprised 262 clones, named TM clones for '**T**orf, **M**ittlere Schicht' (= peat, middle layer). As judged from the electrophoretic migration of reamplified 16S rDNA inserts on an agarose gel, all cloned 16S rDNA fragments had a similar size (about 1400 nucleotides). Using the 16S rDNA oligonucleotide probe 519r (Table 1), which is specific for members of the domain Bacteria, the bacterial origin of all but three inserts could be verified (not shown). The identity of the fragments was further investigated by hybridizing a dot blot of all 259 16S rDNA inserts and 26 reference 16S rDNA PCR products of species of known phylogenetic affiliation with different taxon-specific oligonucleotide probes. The reference DNA represented 13 different bacterial genera, seven isolates from the same peat sample [25] and two plant cell cultures (Table 2). The latter samples were included to determine whether plant DNA had been co-extracted and plant chloroplast 16S rDNA would crossreact with any probes applied to the clone library. Using an oligonucleotide probe that is specific for Gram-positive bacteria (Table 1), 37 clones (15% of total) were identified as phylogenetic members of the Gram-positive phylum. One hundred and ten 16S rDNA PCR clones (43% of total) were identified as members of the alpha-Proteobacteria using probe Alf 1b (Table 1) (Figure 1). This probe was not absolutely specific for members of the alpha proteobacterial subclass as it also hybridized with the reference DNA of *Spirochaeta africana* and *Gemmata obscuriglobus*, a member of the order Planctomycetales. Application of a spirochete-specific probe and sequence analysis of clone TM3 confirmed its affiliation to the order *Spirochaetales*. No signals other than with the two reference planctomycete DNAs were obtained when a planctomycete-specific probe was used. It was thus concluded that no 16S rDNA clones affiliated to the family *Planctomycetaceae* were present amongst the clones analyzed. A probe designed to detect more clone sequences related to *Acidobacterium capsulatum*, some of which were found by sequence analysis of the clone library (probeTM2; see below, Table 1), indicated the presence of an additional 13 16S rDNA sequences originating from organisms of this subcluster. Another subcluster-specific probe (probeTM10, Table 1) was designed on the basis of the sequence of TM10 to detect additional relatives of this clone. Two

Table 2 List of 16S rDNA PCR products and their origin used for dot blot hybridization with different oligonucleotide probes

Position on membrane[a]	Name, collection number[b]
I A1	*Bacillus cereus*, DSM 31^T
I A2	*Bacillus mycoides*, DSM 307
I A3	*Bacillus mycoides*, DSM 2048^T
I A4	*Bacillus thuringiensis*, DSM 2046^T
I A5	*Bacillus subtilis*, DSM 10^T
I A6	Peat isolate 1, Gram-positive
I A7	Peat isolate 2, Gram-negative
I A8	Peat isolate 3, Gram-positive
I A9	Peat isolate 4, Gram-positive
I A10	Peat isolate 5, Gram-positive
I A11	Peat isolate 6, Gram-positive
I A12	Peat isolate 7, Gram-positive
I B1	*Comamonas acidovorans*, DSM 39^T
I B2	*Janthinobacterium lividum*, DSM 1522^T
I B3	*Streptomyces albidoflavus*, DSM 40455^T
I B4	*Methylobacterium fujisawaense*, DSM 5686
I B5	*Rhodococcus rhodochrous*, DSM 43241^T
I B6	*Desulfotomaculum orientis*, DSM 765^T
I B7	*Desulfovibrio africanus*, DSM 2603^T
I B8	*Spirochaeta africana*, DSM 8902^T
I B9	*Arthrobacter globiformis*, DSM 20124^T
I B10	*Escherichia coli*, DSM 498
I B11	*Planctomyces limnophilus*, DSM 3776^T
I B12	*Gemmata obscuriglobus*, DSM 5831^T
I C1	*Equisetum giganteum*, DSM PC 592
I C2	*Marchantia polymorpha*, DSM PC 713

[a]Position on membrane as shown in Figure 1. The Roman number stands for the membrane number, the letter and the Arabic number indicate the row and column on the membrane.
[b]T = Type strain.

additional clones were identified, which were shown to be highly related to clone TM10 following sequence analysis (TM21 and TM72 in Figure 2).

Identification of 16S rDNA clone inserts by sequence analysis

Following hybridization, selected clones from the two major hybridization clusters, (ie alpha-proteobacteria and Gram-positive bacteria), were sequenced. The stretches sequenced comprised about 350 nucleotides from the 5′ terminus of the 16S rDNA. The stretch was extended to 1000 nucleotides for representatives of the emerging phylogenetic groups. The similarity between the clone sequences and database entries was determined and the clone sequences were further evaluated by comparing them specifically to members of those phyla and subclasses to which they belonged phylogenetically. None of the 76 sequences analyzed were identical to a sequence of a database entry of cultured strains or clones obtained from investigations of other environmental habitats.

Figure 3 is a neighbor-joining tree based on the analysis of 350 5′ terminal nucleotides using a selection of sequences from cultured strains and from the clone clusters identified. The majority of clone sequences fall into the radiation of the alpha subclass of Proteobacteria, the *Acidobacterium capsulatum* phylum and the *Acidimicrobium ferrooxidans/Rubrobacter radiotolerans* cluster of the actinomycete subphylum of Gram-positive bacteria. Clones

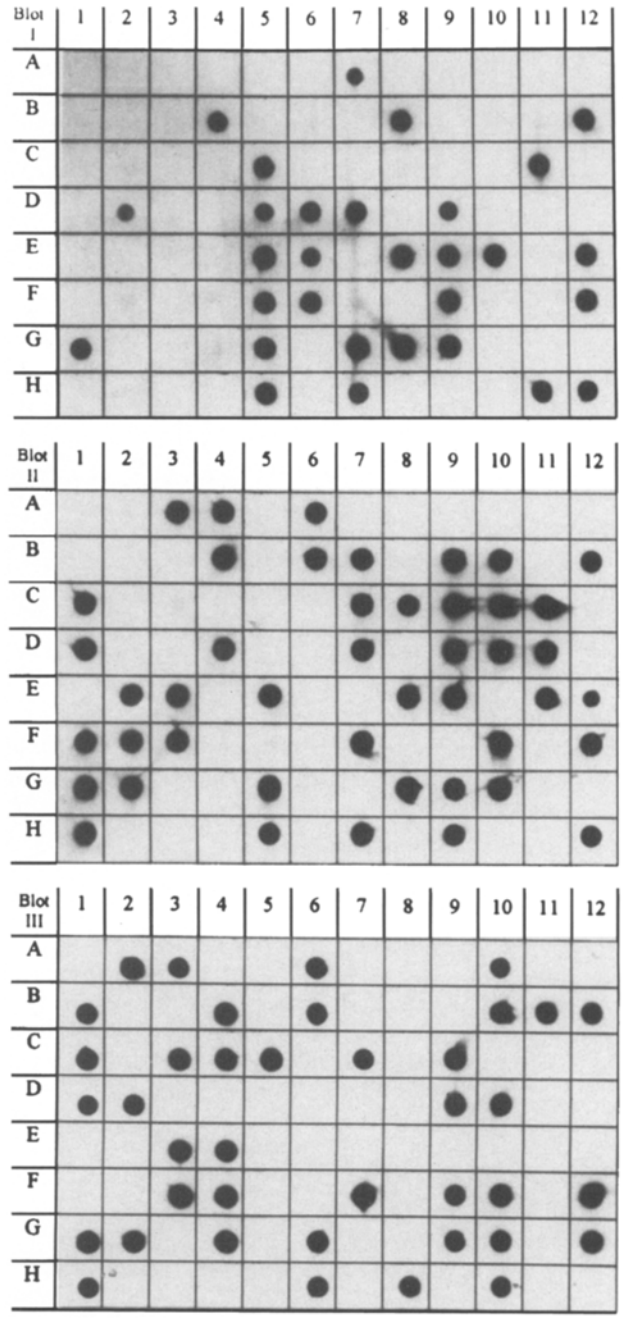

Figure 1 Hybridization of the TM clones with probe 'Alf 1b'. Each membrane (named 'Blot I' to 'Blot III') carries 96 16S rDNA PCR products. As there are for 'Blot I': controls 1–26 (see Table 2 for details) from position A1 to C2 and clones TM1–TM70 from position C3 to H12. For 'Blot II': clones TM71–TM166 from position A1 to H12. For 'Blot III': clones TM167–TM262 from position A1 to H12. All the three membranes were processed simultaneously in the same roller tube to assure reproducible hybridization and detection conditions. Hybridization with probe 'Alf 1b' was carried out at room temperature for 4 h. The temperature of the stringency wash in 6 × SSC, 0.1% SDS buffer was 54°C. Exposure to an X-ray film was for 2 h.

TM24, TM221, and TM252 are members of the beta subclass of Proteobacteria (1% of total), representing phylogenetic neighbors of the genus *Telluria*. Single clone representatives were clones TM18 and TM3, which are members of the *Verrucomicrobiales* subphylum and the

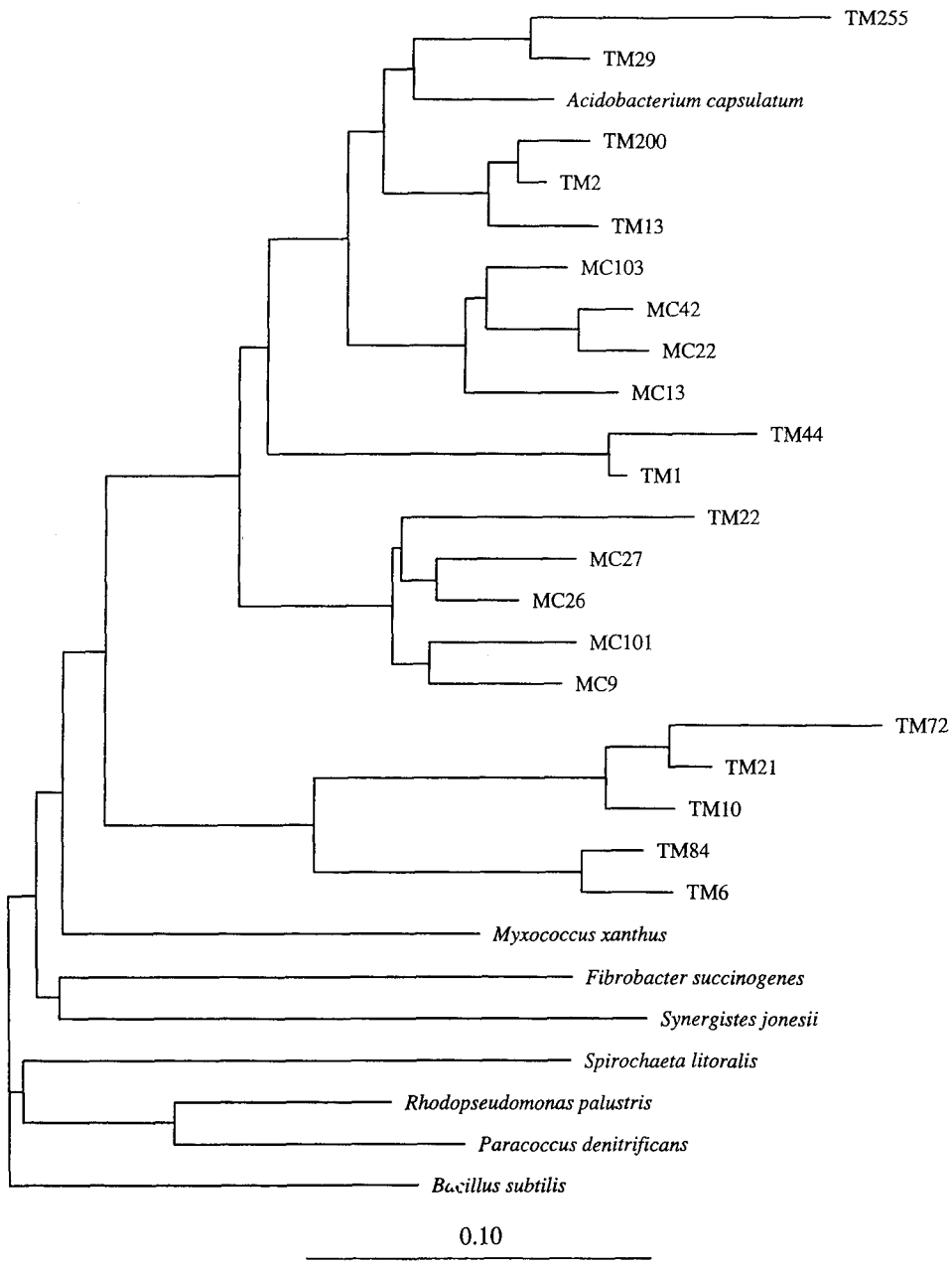

Figure 2 Dendrogram showing the phylogenetic position of the TM and related MC clone sequences, which are most closely related to the *Acidobacterium capsulatum* phylum. Analysis is based on the 5′ 350 nucleotides of the 16S rDNA sequences. The bar represents 10 nucleotide substitutions per 100 nucleotides.

spirochetes, respectively. Clones TM5 and TM7 form phylogenetically deep rooting lines of their own. Bootstrap analysis, however, indicates that the deeply branching lineages, which are defined by less than 75% 16S rDNA similarity, have no statistical significance (not shown).

Clone sequences belonging to the alpha-Proteobacteria

Sequence analysis of a stretch of 350 nucleotides was performed on 25 of the 110 16S rDNA PCR clones which were detected with probe Alf 1b, and 17 were found to represent distinct lineages (six clone sequences were duplicates). Eleven sequences were closely related to each other and to members of the *Beijerinckia/Bradyrhizobium* cluster which also contains *Rhodopseudomonas palustris, Azorhizobium, Rhodopseudomonas acidophila* and others (>96% similarity) (Figure 4). The clone sequences of TM9 and TM51 are most similar to *Beijerinckia indica* (>98% similarity). Clone sequences, like TM19, TM91 and TM41 form an individual cluster (>94% similarity) while clone sequence TM113 stands isolated within this cluster. Interestingly, several clone sequences (abbreviated MC) retrieved from a clone library of DNA that was isolated from a forest soil at Mt Coot-tha, Brisbane, Queensland, Australia (Liesack and Stackebrandt, 1992) are intermixed with the peat clone sequences of this cluster. Two of these

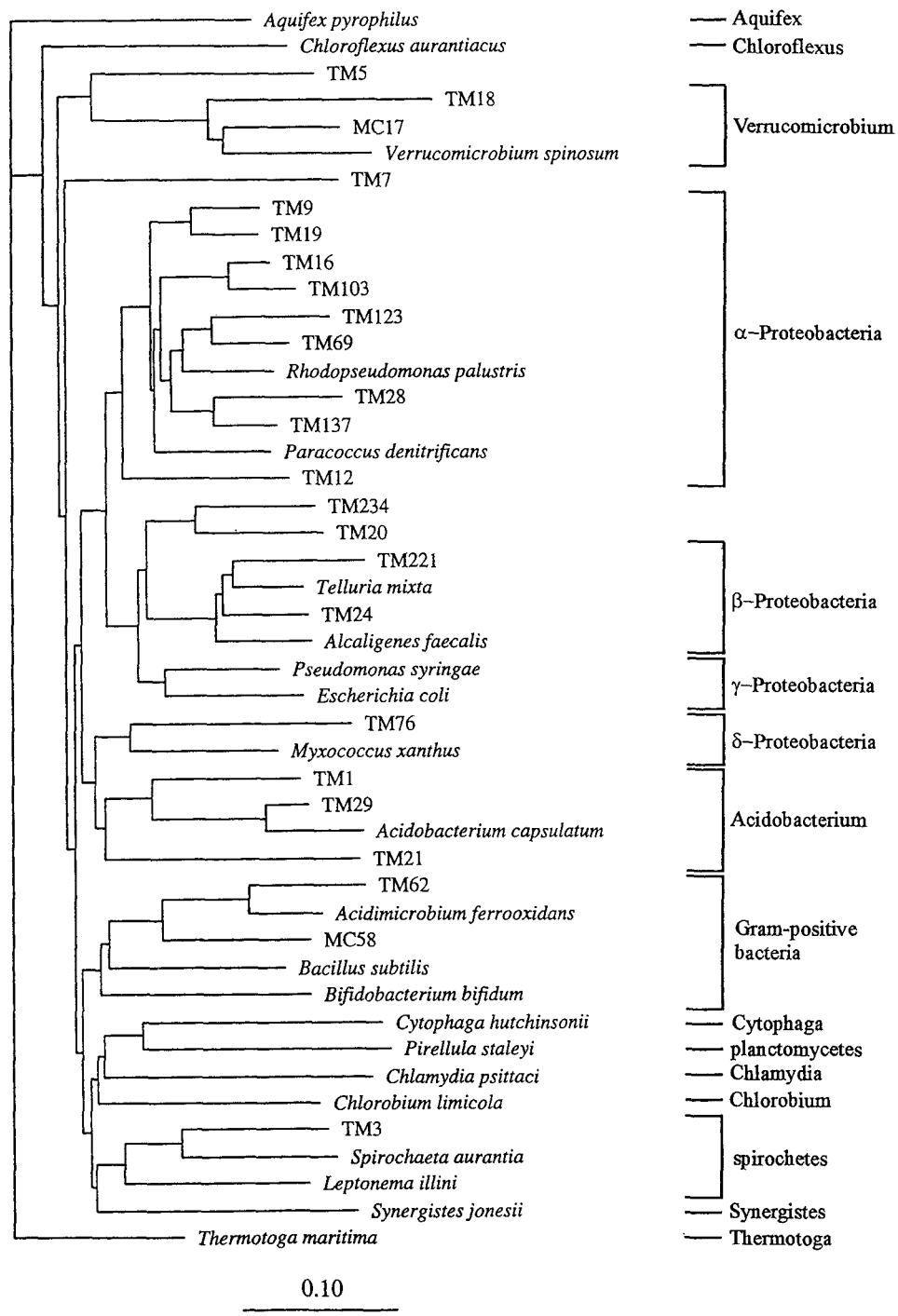

Figure 3 Dendrogram based upon the distance method of Jukes and Cantor, showing the phylogenetic position of representatives of the TM clones and related MC clone sequences. The bacterial main lines of descent are indicated on the right side. Analysis is based on the 5′ 350 nucleotides of the 16S rDNA sequences. The 16S rDNA sequence of *Thermotoga maritima* served as an outgroup sequence. The bar represents 10 nucleotide substitutions per 100 nucleotides.

MC clones group with clone TM113 (>95% similarity). These MC clones are not shown in Figure 3 because a common stretch of only 200 nucleotides of the 16S rDNA is available for comparison.

The members of another cluster, containing TM123, TM69 and related clones, are moderately to highly related to each other and the cluster branches deeply to the *Bradyrhizobium/Beijerinckia* cluster. These sequences do not show high 16S rDNA similarity to any sequence of alpha-Proteobacteria currently available. As judged from the similarity values, they appear to be equidistantly related to bradyrhizobia and relatives, and to caulobacters. Clones TM103 and TM69 are highly related to *Caulobacter* species (up to 98% similarity).

Clone sequences TM120 and TM212 are highly similar and belong to the *Rhodopila/Acidiphilium* group. This

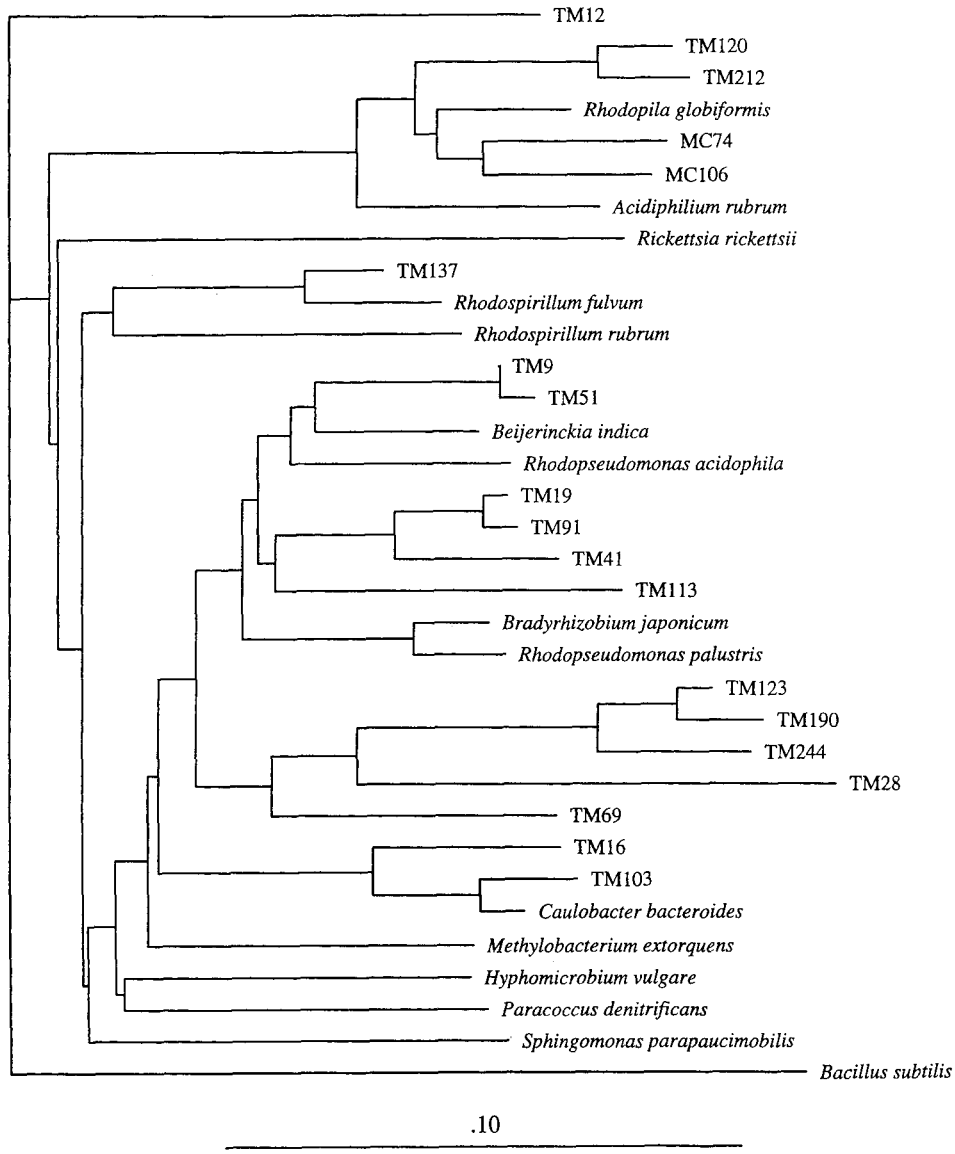

Figure 4 Dendrogram showing the phylogenetic position of alpha-proteobacterial TM clone sequences and related organisms. The 16S rDNA sequence of *Bacillus subtilis* served as an outgroup sequence. Analysis is based on the 5′ 300 nucleotides of the 16S rDNA sequences. The bar represents 10 nucleotide substitutions per 100 nucleotides.

group also contains two sequences that were retrieved from the MC clone library, ie MC74 and MC106 (Figure 4).

Clone sequences belonging to the Acidobacterium capsulatum phylum

The cultivated organism *Acidobacterium capsulatum* was described as phylogenetically belonging to 'a unique lineage deeply branching from the *Chlamydia-Planctomyces* group or from the Gram-positive line' [9]. As these higher taxa represent individual phyla we consider *A. capsulatum* to represent a monospecific phylum as well.

A total of 26 peat clone sequences are members of the *Acidobacterium* phylum (13 each from sequence analysis and hybridization), the phylogenetic depth is considerably broadened by the addition of these uncultured representatives (Figure 2). The two clusters represented by TM29 and TM2 are phylogenetic neighbors of *A. capsulatum*, with 16S rDNA similarity values around 90%. An additional group of four clone sequences that originates from the MC clone library of Australian soil DNA (MC42 and relatives) branches slightly deeper than the TM29/TM2 clone clusters. The MC sequences of the *Acidobacterium* phylum have previously been assigned with no statistical significance to a novel, deeply branching lineage within the actinomycete line of descent [23]. The sequence of *Acidobacterium* was not available at the time of that publication. The lineage comprising five clone sequences (TM10 lineage, Figure 2) is linked with no statistical significance to the *Acidobacterium capsulatum* phylum.

Clone sequences belonging to the beta-Proteobacteria

Three of the clone sequences which were randomly selected for sequencing from those clones that did not hybridize

with any of the taxon-specific probes, are members of the beta-Proteobacteria. Clones TM24, TM221 and TM252 are related to the *Telluria/Burkholderia* cluster. While *Alcaligenes eutrophus* shares 96% sequence similarity to TM24, clones TM221 and TM252 show moderate sequence similarities to *Burkholderia solanacearum* (92% and 94%, respectively).

Clone sequences belonging to the Acidimicrobium ferrooxidans and the Rubrobacter radiotolerans sublines of the actinomycete phylum

A broad group of moderately similar 16S rDNA sequences clusters around the iron-oxidizing species *Acidimicrobium ferrooxidans* [5] and '*Candidatus* Microthrix parvicella', from activated sludge [1,2] (Figure 5). The relationship between these clone sequences has been discussed in detail by Rheims *et al* [19]. The diversity of this group increases when published sequences from clones are included in the analysis that have been retrieved from other environmental habitats. These include clones from a paddy field [14] and geothermically heated soil from New Zealand [18]. Unfortunately the phylogenetic analysis is hampered by the fact that these sequences are short and originate from different parts of the 16S rDNA sequence. A complete sequence comparison that would include all these sequences is not possible.

A second lineage represented by a cultured species *Rubrobacter radiotolerans* contains three TM clone sequences represented by TM36 and TM146. These clones are related to clones MC4 and MC19 from forested soil from Australia [23] (Figure 5). Although not shown in Figure 5 this lineage also contains moderately related clone sequences from Finland [21], soil from a paddy field [14] and a soybean field from Japan [24], and the marine environment [7].

Discussion

The results presented in this study, obtained from the molecular analysis of a peat bog sample, support the general statements made in previously published studies of different terrestrial and marine environments [27]. These general remarks refer to the potential and limitations of the methods used to assess microbial diversity in environmental samples, some of which will be discussed briefly.

The rationale for assessing the phylogenetic diversity of prokaryotes originates firstly from the finding that only a

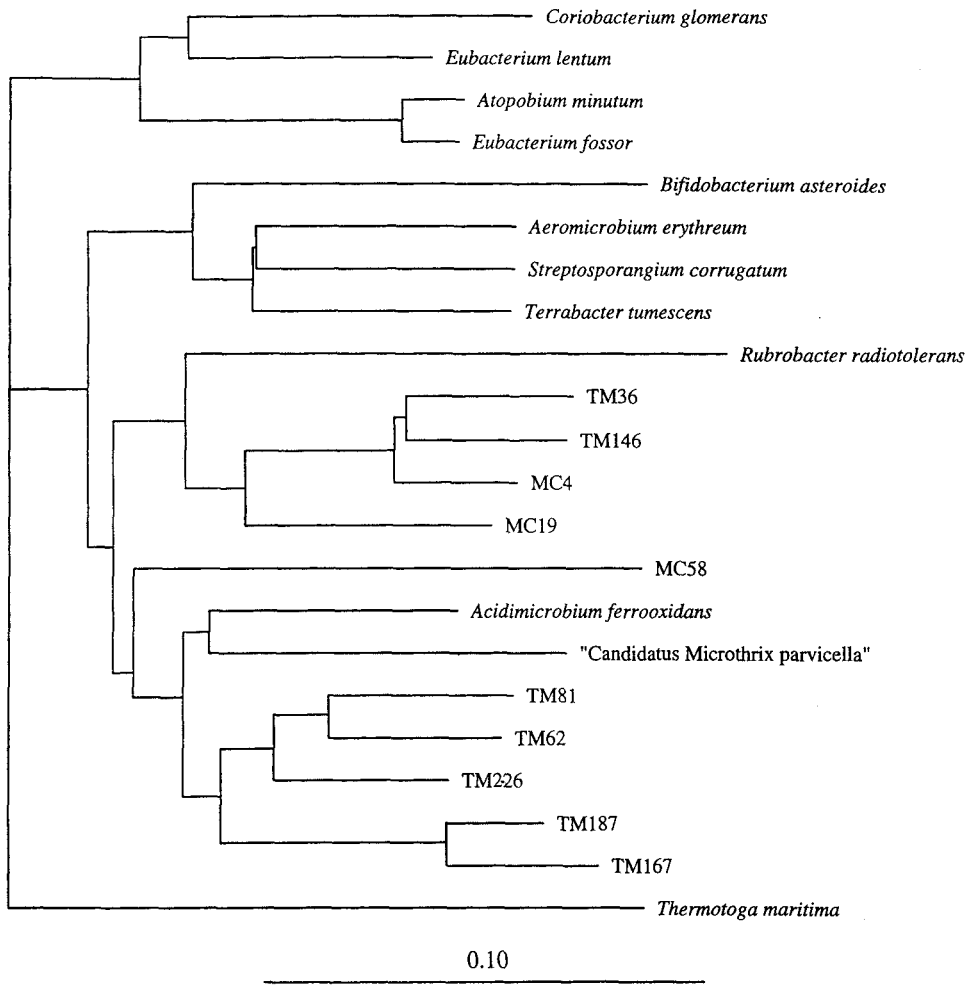

Figure 5 Dendrogram showing the phylogenetic position of representatives of the TM clones which are related to *Acidimicrobium ferrooxidans* and related MC clone sequences. Analysis is based on the 5′ 350 nucleotides of the 16S rDNA sequences. The bar represents 10 nucleotide substitutions per 100 nucleotides.

fraction of the morphologically diverse organisms detected by light microscopy could be cultured. Secondly, the diversity among prokaryotic species is presently represented by about 4000 valid species which is only 0.3% of the 1.45 million described biological species. Clearly, this small fraction does not reflect the natural diversity of prokaryotic species considering the existence of prokaryotes on this planet for more than 3.5 billion years. It can be assumed that the actual number of prokaryotic species is higher than current estimates by several powers of ten but this number must remain a mere guess.

Certainly, one reason for the failure to estimate the number of species reliably is due to the inability to provide optimal cultivation conditions, and this in turn is due to the fact that for the majority of environmental strains almost nothing is known about the biochemistry, physiology, and associations with organic (including hosts in the case of symbionts) and inorganic matter. Traditional enrichment and isolation procedures appear to be inadequate to match the conditions needed to allow resting and dormant cells to grow. It must be stressed that the analysis of clone libraries of environmental DNA will not change this situation since only in rare cases does the phylogenetic position of a hitherto uncultured strain reveal the culture conditions needed for its cultivation. The strength of the molecular approach to microbial ecology needs to be seen in a different light.

One advantage of using 16S ribosomal DNA clone libraries is the possibility of determining the presence or absence of similarities in the phylogenetic patterns of environmental samples of different origin. The main results are summarized as: (i) In contrast to the expectation of novel diversity, only a few archaeal and bacterial sequences have been retrieved that point towards the occurrence of novel kingdoms; the existence of a fourth domain, as defined by sequence analysis of an as yet uncultured organism, is under discussion. Most bacterial sequences can be assigned to known phyla and representative environmental sequences have been found that fall into the radiation of almost all known bacterial phyla. Sequences originally described as representing such novel phyla [12,13,23] were later found to be moderately similar to those of cultured strains. Examples are clone MC58, related to *Acidimicrobium ferrooxidans*, clones MC4 and MC19, related to *Rubrobacter radiotolerans*, clones MC22 and MC103, related to *Acidobacterium capsulatum*, and the MC18 clone group, related to *Verrucomicrobium spinosum* [29]. (ii) Sequences of strains isolated from the same locations from which a clone library has been generated are rarely identical to the clone sequences, or to those available from the 16S rDNA database; examples have been published for treponemas [4] and cyanobacteria [20] and are known for *Paenibacillus* (L Sly, personal communication). (iii) It has recently been reported that investigations of marine samples from the Atlantic and Pacific oceans revealed similarities in the occurrence of novel phylogenetic groups [15]. We can demonstrate that the same phenomenon occurs in terrestrial environments. Of the three major groups determined here to be present in peat, ie the *Acidobacterium* group, the *Acidimicrobium/Rubrobacter* group and the *Bradyrhizobium/Rhodopseudomonas palustris* group, all of them have been found to be present in a second well-characterized environment, ie soil of Mount Coot-tha, Queensland, Australia. The *Acidimicrobium/Rubrobacter* group was also present in rice paddy field, soybean field, geothermally heated soil from New Zealand, and in a marine environment. It can therefore be concluded, that the phylogenetic diversity of those phyla, which are currently represented by only a single or a few cultured species, actually constitute phenotypically broad groups in the environment, containing a potentially rich structure of novel taxa. As pointed out above, no conclusions about their physiological properties can be made and hence the phenotypic diversity of such lineages remains unexplored.

Some sequences obtained from samples of geographically widely separated locations may show surprisingly high similarities, ie from marine and terrestrial origin, or from New Zealand and Germany. It has been assumed that the relative abundance of sequences retrieved from the environment in a clone library does not indicate anything about the quantitative distribution of species in the environment. Reasons for excluding such quantification are the many methodological biases (cell lysis, recovery of DNA, PCR selectivity, genomic properties) that are introduced at various steps in the procedure of generating a clone library. Another problem is the identification of minority populations occurring only in numbers of 10^2–10^4 cells within a population of 10^8 cells per gram of soil. The chance of recovering a single representative clone of such a minority population from a clone library would require analysis of 10^4–10^6 clones, a costly and laborious procedure. Nevertheless, if a group of related sequences is found to constitute a major fraction of clone libraries that have been generated using quite different approaches, including different cell lysis procedures, PCR primers, and polymerases, as well as different cloning systems, one would assume that these sequences originate from organisms that are well-represented in the environmental sample. Such a group is the *Acidomicrobium/Rubrobacter* clone cluster that is so well-represented in at least five different environmental clone libraries [19].

Besides the aspect of phylogenetic diversity, the question of physiological diversity is equally important. As these properties cannot be determined by the analysis of ribosomal DNA, probes specific for other genes must be applied by *in situ* hybridization or dot blot hybridization, or alternatively, PCR primers used to detect these genes by amplification.

Although a few functional probes have been described for the detection of certain genes involved in anabolic and catabolic processes, significantly more data need to be generated. The probes/primers are needed because it is almost impossible to predict physiological properties (and hence culture conditions) of a strain only from its phylogenetic position within the 16S rDNA dendrogram (except in those cases where the organism groups within the radiation of a physiologically well-defined genus). For isolation work, the main obstacle appears to be the recovery of minority populations, and of those strains that are over-grown by cultures better adapted to the growth media. It appears important to match in the laboratory the physico-chemical conditions of the habitat, which could be investigated by using sophisticated microprobes that are designed to measure various

physical parameters, gases and organic compounds. Direct proof of the existence of an organism determined to be present in a sample by any of the molecular approaches is by detection *in situ*, ie directly within the environmental sample by 16S rDNA oligonucleotide probes. This approach has reached a high level of sophistication using confocal laser microscopy and the simultaneous application of multiple probes. In future studies, genes other than rRNA genes will be targeted by this approach.

The question from the pharmaceutical and biotechnology companies is how to make use of the novel insights into microbial diversity. The results indicate that phylogenetic novelty, hence biotechnological novelty is ubiquitous and present in samples from the Southern and the Northern hemisphere. Two strategies can be followed to obtain information about novel biotechnological properties of microorganisms in the environment. One approach would include an isolation strategy which, however, is extremely time-consuming and does not guarantee success. The industry, more than any other group, is extremely experienced in searching for microbial novelty and the occurrence of rare species. It appears likely that large screening programs that are developed by the pharmaceutical industry with the goal to specifically retrieve novel bacteria will continue to be successful, considering the extensive diversity in the environment as demonstrated by the molecular techniques. The second strategy could be circumscribed by the genetic approach. Here, the isolation of strains is a secondary priority to a molecular genetic analysis of an environmental DNA, which would include cloning of DNA fragments in expression vectors and the search for industrial enzymes (such as glycosidases, esterases/lipases, aminotransferases, or phosphatases), secondary metabolites, or therapeutic agents. This approach has now been taken by the chemical industry as advertized in commercial journals. This strategy is certainly justified from a commercial point of view but experiments can only be designed to obtain information about those properties which are targeted, while novel properties will most likely escape detection. Microbiologists welcome this strategy and consider this approach as an addition to their desire to obtain knowledge about genotypic and phenotypic properties.

Acknowledgements

This study was supported by a grant from the German Research Council (DFG-Sta 184/13-1) to ES.

References

1 Blackall LL, EM Seviour, MA Cunningham, RJ Seviour and P Hugenholtz. 1994. '*Microthrix parvicella*' is a novel, deep branching member of the *Actinomyces* subphylum. Syst Appl Microbiol 17: 513–518.
2 Blackall LL, H Stratton, D Bradford, T Del Dot, C Sjörup, EM Seviour and RJ Seviour. 1996. '*Candidatus* Microthrix parvicella', a filamentous bacterium from activated sludge sewage treatment plants. Int J Syst Bacteriol 46: 344–346.
3 Cabot EL and AT Beckenbach. 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comp Appl Biosci 5: 233–234.
4 Choi KB, BJ Paster, FE Dewhirst and UB Göbel. 1994. Diversity of cultivable and uncultivable oral spirochetes from a patient with severe destructive periodontitis. Infect Immun 62: 1889–1895.
5 Clark DA and PR Norris. 1996. *Acidimicrobium ferrooxidans* gen nov, sp nov: mixed culture ferrous iron oxidation with *Sulfobacillus* species. Microbiology 142: 785–790.
6 Felsenstein J. 1993. PHYLIP Phylogeny Inference Package, version 3.5.1. Seattle: Department of Genetics, University of Washington.
7 Fuhrman JA, K McCallum and AA Davis. 1993. Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific oceans. Appl Environ Microbiol 59: 1294–1302.
8 Goebel B and E Stackebrandt. 1994. Cultural and phylogenetic analysis of mixed microbial populations found in natural and commercial bioleach environments. Appl Environ Microbiol 60: 1614–1621.
9 Hiraishi A, N Kishimoto, Y Kosako, N Wakao and T Tano. 1995. Phylogenetic position of the menaquinone-containing acidophilic chemo-organotroph *Acidobacterium capsulatum*. FEMS Microbiol Lett 132: 91–94.
10 Jukes TH and CR Cantor. 1969. Evolution of protein molecules. In: Mammalian protein metabolism (Munro HN, ed), pp 21–132, Academic Press, New York.
11 Liesack W and E Stackebrandt. 1992. Occurence of novel groups of the domain Bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. J Bacteriol 174: 5072–5078.
12 Liesack W and E Stackebrandt. 1992. Unculturable microbes detected by molecular sequences and probes. Biodivers Conser 1: 250–262.
13 Liesack W, K-J Chin, U Hengstmann, R Groβkopf and PH Janssen. 1995. Characterization of microbial populations of rice microcosms using traditional cultivation and molecular techniques. 7th International Symposium on Microbial Ecology, Santos, Brazil. Abstract S4-5.1. Brazilian Society for Microbiology, Santos.
14 Maidak BL, N Larsen, MJ McCaughey, R Overbeck, GJ Olsen, K Fogel, J Blandy and CR Woese. 1994. The ribosomal database project. Nucleic Acids Res 22: 3483–3487.
15 Mullins TD, TB Britschgi, RL Krest and SJ Giovannoni. 1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. Limnol Oceanogr 40: 148–158.
16 Muyzer G, E de Waal and A Uitterlinde. 1993. Profiling of complex microbial populations using denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 59: 695–700.
17 Muyzer G and EC de Waal. 1994. Determination of the genetic diversity of microbial communities using DGGE analysis of PCR-amplified 16S rDNA. In: Structure, Development and Environmental Significance of Microbial Mats (Stal LJ and P Caumette, eds), pp 207–214, NATO ASI Series, vol G35, Springer-Verlag, Berlin.
18 Rainey FA, NL Ward and E Stackebrandt. 1993. Molecular ecology study of a New Zealand acidothermal soil. Thermophiles '93, Hamilton, New Zealand. Abstract A7. University of Waikato, Hamilton.
19 Rheims H, C Spröer, FA Rainey and E Stackebrandt. 1996. Molecular biological evidence for the occurrence of uncultured members of the actinomycete line of descent in different environments and geographical locations. Microbiology 142: 2863–2870.
20 Ruff-Roberts AL, GJ Kuenen and DM Ward. 1994. Distribution of cultivated and uncultivated cyanobacteria and *Chloroflexus*-like bacteria in hot spring microbial mats. Appl Environ Microbiol 60: 697–704.
21 Saano A, K Lindström and JD Van Elsas. 1995. Eubacterial diversity in Finnish forest soil. 7th International Symposium on Microbial Ecology, Santos, Brazil. Abstract P1-3.9. Brazilian Society for Microbiology, Santos.
22 Schlesner H. 1994. The development of media suitable for the microorganisms morphologically resembling *Planctomyces* spp, *Pirellula* spp, and other *Planctomycetales* from various aquatic habitats using dilute media. Syst Appl Microbiol 17: 135–145.
23 Stackebrandt E, W Liesack and BM Goebel. 1993. Bacterial diversity in a soil sample from a subtropical Australian environment as determined by 16S rDNA analysis. FASEB 7: 232–236.
24 Ueda T, Y Suga and T Matsuguchi. 1995. Molecular phylogenetic analysis of a soil microbial community in a soybean field. Eur J Soil Sci 46: 415–421.
25 Von Wintzingerode F. 1994. Phylogenetische und chemotaxonomische Hinweise auf die Anwesenheit neuer Bakterienarten in einer Torflagerstätte in Ost-Niedersachsen. Diploma Thesis. Technical University, Braunschweig.

26 Wagner C, A Grieβhammer and HL Drake. 1996. Acetogenic capacities and the anaerobic turnover of carbon in a Kansas prairie soil. Appl Environ Microbiol 62: 494–500.
27 Ward N, FA Rainey, B Goebel and E Stackebrandt. 1995. Identifying and culturing the 'unculturables'—a challenge for microbiologists. In: Microbial Diversity and Ecosystem Function (Allsopp D, RR Colwell and DL Hawksworth, eds), pp 89–110, CAB International, Wallingford.
28 Ward N, FA Rainey, E Stackebrandt and H Schlesner. 1995. Unraveling the extent of diversity within the order *Planctomycetales*. Appl Environ Microbiol 61: 2270–2275.
29 Ward-Rainey N, FA Rainey, H Schlesner and E Stackebrandt. 1995. Assignment of hitherto unidentified 16S rDNA species to a main line of descent within the domain Bacteria. Microbiology 141: 3247–3250.