Bacterial diversity in the active stage of a bioremediation system for mineral oil hydrocarbon-contaminated soils

Nicole Popp, Michael Schlömann and Margit Mau

Interdisziplinäres Ökologisches Zentrum, TU Bergakademie Freiberg, Leipziger Str. 29, D-09599 Freiberg, Germany

Soils contaminated with mineral oil hydrocarbons are often cleaned in off-site bioremediation systems. In order to find out which bacteria are active during the degradation phase in such systems, the diversity of the active microflora in a degrading soil remediation system was investigated by small-subunit (SSU) rRNA analysis. Two sequential RNA extracts from one soil sample were generated by a procedure incorporating bead beating. Both extracts were analysed separately by generating individual SSU rDNA clone libraries from cDNA of the two extracts. The sequencing results showed moderate diversity. The two clone libraries were dominated by Gammaproteobacteria, especially Pseudomonas spp. Alphaproteobacteria and Betaproteobacteria were two other large groups in the clone libraries. Actinobacteria, Firmicutes, Bacteroidetes and Epsilonproteobacteria were detected in lower numbers. The obtained sequences were predominantly related to genera for which cultivated representatives have been described, but were often clustered together in the phylogenetic tree, and the sequences that were most similar were originally obtained from soils and not from pure cultures. Most of the dominant genera in the clone libraries, e.g. Pseudomonas, Acinetobacter, Sphingomonas, Acidovorax and Thiobacillus, had already been detected in (mineral oil hydrocarbon) contaminated environmental samples. The occurrence of the genera Zymomonas and Rhodoferax was novel in mineral oil hydrocarbon-contaminated soil.

INTRODUCTION

The contamination of soils with mineral oil hydrocarbons is a widespread environmental problem. Strictly speaking, mineral oil hydrocarbons are not xenobiotic, but their large-scale use and various applications lead in many cases to environmental contamination (Gallego et al., 2001). Such contamination may be a consequence of petroleum transport, storage and refining, or accidents (Juteau et al., 2003). Soil is an environmental compartment which is renewable only within long time periods, or even not renewable at all. Therefore, it is important to protect soil from such contamination or to clean contaminated soils. After treatment, the soil should be usable, at least for applications with a low purity requirement, such as road construction or the recultivation of industrial areas.

One possibility for cleaning such soils is bioremediation (Juteau et al., 2003), which aims at the biological mineralization of organic compounds to CO₂ and water, or at least at transformation to less-toxic or innocuous forms (Plaza et al., 2003). The success of biodegradation depends on the predominant environmental conditions, on the chemical structure of the pollutants, on the biodegradability of the contaminating compounds, and thus on the interaction between pollutant, soil matrix and microorganisms (De’Arco & de França, 2001; Volkering et al., 1998). Another problem of contaminated sites is the wide variety of mineral oil hydrocarbons (Li et al., 2004).

There are many known consortia of micro-organisms which are able to degrade mineral oil hydrocarbons under laboratory or field conditions (Ratajczak et al., 1998; Wikström et al., 1996). However, a single bacterium usually has only a relatively small degradation range, and thus not all fractions of the mineral oil hydrocarbon can be degraded by a single species (Mishra et al., 2001). Consequently, the use of a laboratory strain to clean a contaminated soil is usually
not successful. In addition to the restricted substrate range, such a strain is not necessarily adapted to the particular environmental conditions, such as temperature, and the availability of water, O₂ and nutrients.

Rarely, contamination occurs suddenly (e.g. by tanker accidents or explosions), but more often creeping contamination occurs. A pipeline, or a storage tank at a petrol station, may leak for many years and contaminate the surrounding soil, and thus a degrading indigenous micro-flora can develop and adapt to the use of the pollutants as a carbon and energy source. For this reason, the naturally occurring microflora is often used for the degradation of mineral oil hydrocarbons (Watanabe, 2001).

In practice, a variety of methods is used for the bioremediation of hydrocarbon-contaminated soils, with off-site remediation being the most important remediation strategy. Often the soil is ventilated, and a mineral fertilizer may be added at the beginning of or during the treatment. During the bioremediation process in such systems, the temperature of the soil may increase (Condé & Hagedorn, 1997), which reflects the activity of the indigenous microflora using the mineral oil hydrocarbons as a carbon and energy source.

For better control of such a bioremediation system, it may be helpful to know which micro-organisms actually bring about the biodegradation. Various methods have been used to analyse the microbial diversity in mineral oil hydrocarbon-contaminated soils. One method is the cultivation of bacteria (Chao & Hsu, 2004; Greene et al., 2000; La Rosa et al., 2006; Plaza et al., 2003; Radwan et al., 1995; Saul et al., 2005), but it has been estimated that only ~1 % of soil micro-organisms can be cultivated by conventional methods (Torsvik et al., 1990). The established cultivation methods favour the isolation of fast-growing micro-organisms, and those that accept high nutrient concentrations (Felske et al., 1999). Thus, for diversity analyses, cultivation-independent methods are more reliable. Earlier studies have described the mineral oil hydrocarbon-degrading community at the small subunit (SSU) rDNA level in soils (Juck et al., 2000; Saul et al., 2005), in a biodegradation system (Kaplan & Kitts, 2004), in coastal material (Macnaughton et al., 1999) or in soils treated in bioreactors (Mills et al., 2003). SSU rDNA reflects both active and inactive micro-organisms, and thus the genetic potential rather than the active portion of a microbial community (Lüdemann et al., 2000). The active community in a bioremediation system may be characterized by investigation of SSU rRNA, because the latter is directly involved in active protein synthesis (Morgan et al., 2002). In addition, after release from lysed bacterial cells into the soil, chemical degradation of rRNA is relatively fast, and thus rRNA does not remain intact in the soil, while SSU rDNA is more stable (Trevors, 1996). An investigation has been reported of the metabolically active members of a bacterial community in a polychlorinated biphenyl-polluted soil with no degrading activity (Nogales et al., 1999). But, to our knowledge, SSU rRNA analyses of mineral oil hydrocarbon-degrading communities have so far not been described in the literature. Degrading micro-organisms should be particularly active in phases of temperature increase in the system. Thus, in the present study, a sample from the ‘hot’ stage of a bioremediation system was investigated by SSU rRNA analysis to show which micro-organisms were active in this degradation stage.

METHODS

Soil samples. Soil samples were obtained from a bioremediation system for material from a contaminated site at Rositz, a village in Thuringia, Germany, where a tar factory was processing brown coal tar between 1917 and 1990. The factory was destroyed in World War II, and the destruction of the factory accounts for much of the contamination of the soil in this area. The contamination consisted of alkanes, benzene, toluene, ethylbenzene and xylene (BTEX), poly-aromatic compounds, as well as phenols (Verwaltungs- und Verwertungsgesellschaft Industrieländle Rositz, 1998), which had remained on-site for more than 60 years.

The soil had a silty and partially loamy structure, and the pH was almost neutral. There were only marginal amounts of organic matter, as the excavated material contained only minor portions of the upper soil horizons. The bioremediation heap Rositz 3 was composed of 700 t mineral oil-contaminated soil material (8000 milligrams mineral oil per kilogram soil dry weight), with an area of 60 m² and a height of 3.5 m. Mineral fertilizer (nitrogen : phosphate : potassium, 15 :15 :15) had initially been added at 0.5 t per 1000 t soil. The bioremediation heap Rositz 3 was permanently ventilated by extracting soil air from the heap through a system of suction tubes and thereby providing fresh air from the environment (Condé & Hagedorn, 1997).

At the time point at which according to chemical analyses the concentration of mineral oil hydrocarbons had dropped to less than 6000 milligrams per kilogram soil dry weight, and at which the temperature inside the heap had increased from 12 to 21 °C within 43 days, the heap was considered to contain an actively degrading microflora, and a sample was taken for the analysis of bacterial diversity. The sampling of bioremediation system Rositz 3 took place on 24 April 2002. The temperature inside the heap (21 °C) was 10 °C above the ambient air temperature on that day. The sample was collected at a central point with a special drill for manual soil sampling from a height of 1.5 m and 0.75 m into the system. The soil material was put into a 50 ml reaction tube and frozen at ~80 °C.

Nucleic acid extraction. For the simultaneous extraction of RNA and DNA from the soil sample, the method of Görrès (2001) was used in a modified form. This nucleic acid extraction method allows the separation of the nucleic acid extract into sequential fractions, and thus provides the possibility of performing a differential analysis. The first extract may contain the nucleic acids of micro-organisms which are easy to lyse or located outside soil particles. The subsequent extracts may contain nucleic acids of the less-accessible organisms.

Soil (0.5–1 g) was transferred into a 2 ml reaction tube with 0.5 g glass beads of 0.1–0.25 mm diameter (Roth). Then, 750 μl 120 mM sodium phosphate buffer (pH 8) and 500 μl phenol/chloroform : isoamyl alcohol (25:24:1 by volume) were added. The tubes were placed into a bead beater (MM200; Retsch) for 2 min at 26-6 rounds s⁻¹. To separate the aqueous and organic phases, the tubes were centrifuged at 12 000 g for 5 min. The aqueous layers were transferred to a fresh 2 ml reaction tube and stored on ice. This tube contained the first nucleic acid extract. The brown phenol/chloroform layer was removed from
the tube with the soil/bead mixture and discarded. Five hundred microlitres of the sodium phosphate buffer and 500 μl fresh phenol/ chloroform:isomyl alcohol were added to the soil/bead mixture, and the bead-beating procedure was repeated. The aqueous phases were transferred into a fresh 2 ml reaction tube and stored on ice (second extract). The extraction step was repeated once again, and the aqueous phase from this step was combined with the aqueous phase from the second extraction. Further treatment of this second/third extract was separate from that of the first nucleic acid extract.

The preparations were extracted three times with 1 ml phenol/ chloroform:isomyl alcohol. Nucleic acids were precipitated by adding solutions of 5 M NaCl and 30 % (v/v) PEG 6000 at a ratio of sample: salt: PEG 6000 of 9:1:10 (by volume). The mixtures were left at room temperature for 2 h. To pellet the nucleic acids, the tubes were centrifuged at 12 000 g for 5 min, and the resulting supernatants were discarded. The nucleic acid pellets were washed with 500 μl 80 % (v/v) ethanol, centrifuged at 12 000 g for 5 min, and the ethanol was removed. This step was repeated once. The pellets were dried at room temperature for about 30 min, and the nucleic acids were resuspended in 50 μl deionized water and stored at −20 °C.

The nucleic acids were visualized by a urea denaturing 4·5 % polyacrylamide gel and ethidium bromide staining. Bands of DNA, as well as SSU rRNA and large subunit (LSU) rRNA, were visible when 2·5 μl of the extract was loaded onto the gel. Quantification of several extracts of the same material showed that there was about twice as much RNA in the first extract as in the second/third extract. The extraction step was repeated once again, and the nucleic acids were detanatured at 72 °C for about 15 min (depending on sample size). The nucleic acids were resuspended in 5 μl deionized water and stored at −20 °C.

Separation of RNA and DNA. Separation of RNA and DNA was performed with the nucleic-acid-binding silica spin-columns of an Invirob TwinSpin Cell Mini kit (Invitek), using the buffers supplied by the manufacturer. The manufacturer’s protocol was modified at some points: Lysis Solution T, which is meant to lyse cells, was added to the nucleic acid extracts from soil to obtain the optimal reaction conditions for the spin columns. At the end of the procedure, the elutions of RNA and DNA from the respective binding spin filters were carried out twice. The DNA extracts were stored at 4 °C for later use in another study.

To further purify the RNA extracts, traces of DNA were digested with 60 U RNase-free DNase I (Roche) in the enzyme-specific buffer (10 mM Tris/HC1, pH 7·5, 2·5 mM MgCl2, 0·1 mM CaCl2; Ambion) at 37 °C for 2 h. DNase I was removed by phenol extraction and ethanol precipitation (Moore, 1987). The RNA extracts were stored at −20 °C. The success of the DNase I digestion was checked by SSU rRNA PCR, from which no PCR product could be visualized by 1 % agarose gel electrophoresis.

Amplification of SSU rRNA by RT-PCR. The SSU rRNA was used as template for amplification of the corresponding cDNA by RT-PCR using a TGradient Thermocycler (Whatman Biometra). The primers 27F, 5′-AGAGTTTGATCCTGCGGCTCAG-3′, and 519R, 5′-G(C)ATTACCGCGGCTACG-3′ (obtained from MWG Biotech), target most bacterial SSU rRNAs (Lane, 1991). The RT-PCR reactions were prepared following the instructions of the GeneAmp EZ rTh RNA PCR kit (Perkin Elmer-Applied Biosystems). The 50 μl reaction mixture contained the kit components and final concentrations of 0·25 μM each primer, 0·1 μg BSA μl−1 (MBI Fermentas) and 1 μl RNA extract. Negative controls were prepared without RNA extract. The rTh DNA polymerase, which is able to perform both the reverse transcription step and the amplification step (PCR), was added to the reactions after 1 min at 60 °C. The following PCR programme was used: 30 min at 60 °C (reverse transcription), initial denaturation for 2 min at 94 °C, 40 cycles (1 min at 94 °C, 30 s at 55 °C, 45 s at 72 °C), terminal elongation for 7 min at 72 °C, indefinitely at 15 °C (modified from the instructions of the GeneAmp EZ rTh RNA PCR kit).

The RT-PCR products were visualized by 1 % agarose gel electrophoresis and ethidium bromide staining. Bands of RT-PCR products were excised from the agarose gels and DNA was isolated with an EasyPure DNA Purification kit (Biozym).

Cloning and screening of RT-PCR products. RT-PCR products were ligated into a T-vector (Marchuk et al., 1991), which was prepared from EcoRV-digested pBluescript II SK+ (+) (Stratagene). The transformation into competent Escherichia coli DH5α (Gibco-BRL) was performed by electroporation using an EquiBio device (peglab Biotechnologie) and the standard protocol for transformation of E. coli described in the EquiBio optimization guide. The cell suspensions were plated onto Luria–Bertani (LB) medium (1 % tryptone, 0·5 % yeast extract, 1 % NaCl, 1·5 % agar) with 50 μg ampicillin ml−1, and the clones were screened by blue–white detection (Sambrook & Russell, 2001).

White clones were screened for inserts by large-scale colony screening (Campbell & Choy, 2001). Extraction of the plasmid DNA from liquid cultures was performed using the FlexiPrep kit (Amersham Biosciences), and the bound plasmid DNA was eluted with deionized water and stored at −20 °C. Clones from the first bead treatment were named Rositz2 Rxx, and clones from the second/third bead treatment were named Rositz23 Rxx, where x indicates the clone number in the respective library.

The final test for inserts of expected size was by PnuIl digestion of the plasmid DNA.

DNA sequencing and phylogenetic analysis. Sequencing was done on a LI-COR 4200 IR sequencer using the CycleReader Auto DNA Sequencing kit (MBI Fermentas) with labelled primer T3 (5′-CGCGCAATTTAACCCTCCTAAG-3′; Stratagene). The following programme was employed: initial denaturation for 2 min at 92 °C, followed by 25 cycles (30 s at 92 °C, 30 s at 55 °C and 1 min at 72 °C).

Sequence analysis was performed using the programs e-Seq (LI-COR) and EditSeq (DNASTAR; Lasergene). The most similar SSU rRNA sequences were identified using BLAST (Altschul et al., 1990). Using the Chimera Check tool of the Ribosomal Database Project II, the obtained sequences were tested for potential chimeric sequences (Cole et al., 2003). The ARB software was applied for the phylogenetic analysis of sequences (Ludwig et al., 2004). Initial alignment was done by the ARB fast aligner, and then manual verifications and corrections were necessary. At first, a main tree including reference sequences of the taxonomic group of interest was calculated with the maximum-likelihood method and the respective taxonomic filter, which included only positions that are conserved among at least 50 % of the sequences in the respective taxonomic group. After this, the Rositz sequences were added to the tree with the parsimony method without changing the tree topology and applying the respective filter. Finally, an outgroup was defined and added to the tree with the parsimony method by applying the universal bacterial filter. In some cases, many sequences were very similar to each other, and they were therefore grouped together. Such a group was named with the representative species name, and the numbers of corresponding Rositz sequences were specified in parentheses.

Representative sequences obtained in this study were submitted to the EMBL database and can be retrieved under accession numbers AM237215–AM237284, AM279418, AM279419, and AM237445–AM237448 (partial sequences from chimeras).

Estimated phylotype diversity and comparison of clone libraries. Treeclimber (Schloss & Handelsman, 2006) was applied for a P test (Martin, 2002) of the two libraries. Operational taxonomic units (OTUs) were defined based on a 3 % distance cut-off,
and clone groups were calculated by DOTUR (Schloss & Handelsman, 2005). To avoid misassignments due to small overlaps of partial sequences in conserved regions, the sequence parts from chimeras were manually assigned to OTUs. The resulting OTU indices were transformed into abundance tables for $F_\text{ST}$-based comparison (Slatkin, 1991) in Arlequin 3.0 (Excoffer et al., 2005), and for calculation of richness estimators and assemblage comparison in SPADE (Chao & Shen, 2003).

**RESULTS**

**Construction of SSU rRNA library**

The aim of this study was the characterization of the active bacterial community in the bioremediation system Rositz 3 degrading mineral oil hydrocarbons. Therefore, a sample was taken during the active period of the heap and was subjected to rRNA-based sequencing analysis. The amplification of about 480 nucleotide fragments of bacterial SSU rRNA was performed for the extract of the first bead treatment and separately for the extract of the second/third bead treatment to investigate the effect of the sequential extraction. In total, 221 clones were analysed, 110 from the first and 111 from the second/third bead treatments, respectively. All clones were sequenced over their full lengths.

Eight sequences were assumed to be chimeras according to the results of the Chimera Check program (Cole et al., 2003). These sequences were split into two parts at the specified position, and the two parts were handled separately for the next steps of the analysis. In most cases, full-length sequences of at least one of the two sequences were available in other clones. Only those short sequences for which no full-length counterpart was available in the clone libraries were submitted to the databases. For example, the *Alcaligenes* sequence Rositz1 Ro44a (AM237448) formed a chimera with part of a *Methylobacterium* sequence in the full-length sequence Rositz1 Ro74. Furthermore, the *Hydrocarboniphaga*-related Rositz1 Ro78a (AM237447) formed a chimera with a portion of *Betaproteobacterium* Rositz1 Ro35 (AM237265). A total of 112 sequences from library Rositz1 and 110 sequences from Rositz23 were included in the analyses.

**Statistics**

Richness estimation resulted in an ACE (abundance-based coverage estimator) of 236 OTUs in the Rositz1 library and 160 OTUs in the Rositz23 library. The numbers of observed OTUs (cut-off 3 % distance) within the clone libraries were 67 in Rositz1 and 59 in Rositz23. Shannon–Weaver indices of the two libraries were calculated as 3.925 and 3.782, respectively. Sample coverage was 0.563 for Rositz1 and 0.636 for Rositz23.

A $P$ test (Martin, 2002) was done using a parsimony-calculated dendrogram of all Rositz sequences. The parsimony dendrogram had a score of 59, at which the cumulative frequency of the random trees was 0.031 (Supplementary Fig. S3). Analysis of differentiation, which was based on abundance data of OTUs at the 3 % cut-off, was also significant, although the $F_{\text{ST}}$ was quite low ($F_{\text{ST}}=0.00505$, $P<0.00001$). Abundance-based Jaccard ($J_{\text{abd}}$) and Sorensen ($S_{\text{abd}}$) estimates (Chao et al., 2005) were 0.7382 (variance 0.1900) and 0.8494 (0.1341), respectively.

**Phylogenetic distribution of clones at the phylum and subphylum level: predominance of Gammaproteobacteria**

For a first phylogenetic assignment, sequences were compared to the database with the BLAST program, which showed the following phylogenetic distribution (Table 1). *Proteobacteria* dominated in both clone libraries. *Gammaproteobacteria* represented the main group. *Alphaproteobacteria* and *Betaproteobacteria* occurred with about the same frequency in the clone libraries, but made up only one third of the numbers of *Gammaproteobacteria*. *Epsilonproteobacteria*, *Bacteroidetes*, *Actinobacteria* and *Firmicutes* were detected at lower levels. Such less-common sequences occurred at somewhat higher levels in the extract from the second/third bead treatment.

<table>
<thead>
<tr>
<th>Phylogenetic group</th>
<th>Absolute number of sequences in extracts from:</th>
<th>Relative number of sequences (%) in extracts from:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First bead treatment</td>
<td>Second/third bead treatment</td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>25</td>
<td>22</td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>61</td>
<td>53</td>
</tr>
<tr>
<td>Epsilonproteobacteria</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Unaffiliated</td>
<td>–</td>
<td>2</td>
</tr>
</tbody>
</table>

---

Table 1. Phylogenetic distribution of SSU rRNA sequences recovered from two sequential extractions of the Rositz samples
The dominating group among the Gammaproteobacteria was the genus Pseudomonas (Fig. 1a, b), but there was no clear association to one Pseudomonas species. The Rositz sequences were distributed rather evenly within the genus. One of the largest groups was the Pseudomonas anguilliseptica group within the P. aeruginosa lineage. Many of the other Rositz sequences clustered in the Pseudomonas syringae, Pseudomonas fluorescens and Pseudomonas putida lineages. The Rositz sequences within the P. anguilliseptica group originated from the extract after the first bead treatment, while members of the P. syringae and P. fluorescens lineages were more typical of the second/third bead treatment (see origin of the sequences in legend of Fig. 1b). An isolate similar to the Rositz sequences of the P. putida group was obtained in an earlier study of the microflora in biomass-recycle reactors (Morgan et al., 2002).

Another larger group of Rositz sequences within the Gammaproteobacteria showed high similarity to Alphaproteobacteria spp. (Fig. 1a). It could be discerned on the basis of the phylogenetic distribution that the Rositz sequences formed two clusters within the Moraxellaceae. A cluster of seven Rositz sequences of the Alphaproteobacteria group contained mostly sequences from the extract from the second/third bead treatment, and most of the other Moraxellaceae-like Rositz sequences were from the first extract.

Eight of the 11 Rositz sequences which belonged to the clone TRS13 group were obtained in the first bead treatment (Fig. 1a). Clone TRS13 was the closest relative of Rositz sequences and was originally found in an investigation of the phylogenetic diversity of bacterial communities in rhizosphere soil of Lolium perenne and Trifolium repens (Marilley & Aragno, 1999). Some individual sequences of the Rositz sample affiliated with the Xanthomonadaceae (Fig. 1a).

Diversity beyond Gammaproteobacteria

Alphaproteobacteria. Most clones related to Alphaproteobacteria were similar to Sphingomonas spp. (Fig. 2), and most of these sequences were obtained from the first bead treatment. The phylogenetic distribution demonstrated two main groups within the Sphingomonadaceae. One of the clusters affiliated to the Sphingobium lineage and the second to Zymomonas mobilis subsp. mobilis. The Sphingobium lineage also harbours the strain Sphingomonas sp. Dxa-95, which was originally isolated from hydrocarbon-contaminated Arctic soil and has the ability to degrade alkanes and jet fuel (Yu et al., 2000).

Six of the eight Rositz sequences which belonged to the Zymomonas group originated from the second/third bead treatment.

Betaproteobacteria. The Rositz clones showed a wide distribution in the phylogenetic tree of Betaproteobacteria (Fig. 3). However, three larger groups with eight to 12 sequences each were recognizable. The sequences of interest affiliated to the genera Rhodoferax, Acidovorax and Thiobacillus, respectively. The three larger sequence groups demonstrated an even distribution between the first and second/third bead treatments.

The sequences of the Rositz Rhodoferax group were related to sequences obtained from an aquatic system (clone GKS2-77; Glöckner et al., 2000) and from flooded paddy soil cores (clone oxSSC-37; Lüdemann et al., 2000). A similar sequence to the Rositz Thiobacillus group (clone RB7C6) was found in monochlorobenzene-contaminated groundwater (Alfreider et al., 2002). Some more sequences could be affiliated to an uncultured bacterium (clone 36-9) which dominated the clone library from an aquifer contaminated with coal tar waste (Bakermans & Madsen, 2002). This cluster consisted exclusively of Rositz sequences from the first bead treatment.

Epsilonproteobacteria. The four Rositz sequences from the Epsilonproteobacteria (Supplementary Fig. S2) were related to sequences from oil or chlorobenzene-contaminated groundwater (Alfreider et al., 2002; Kodama & Watanabe, 2003; Watanabe et al., 2000, 2002). The closest cultured relative Sulfitooccus yujinse [ATCC BAA-921 (T)] was isolated from oil-contaminated groundwater. It was reported to oxidize sulfur compounds in crude oil and to use the oxidation as an energy source under microaerobic or anaerobic conditions (Kodama & Watanabe, 2003).

Clones of Actinobacteria. The Actinobacteria-related Rositz sequences were phylogenetically very diverse (Fig. 4). Rositz1 Ro106 had a similar sequence to that of clone LTUGr00156 from a soil with aged hydrocarbon contamination (Iida et al., 2002; Kaplan & Kitts, 2004). However, others, such as Rositz23 Ros27 representatives, were similar to Propionibacterium acnes, which has also been found at the brine–sea water interface (clone library KT-2K; Eder et al., 2001). Rositz23 Ros6 representatives were most similar to the sequence of Atopobium riniae (ATCC 49626) from the human oral cavity (Dewhirst et al., 2001; Paster et al., 2001).

Bacteroidetes. The four Rositz clones which represented Bacteroidetes were affiliated to Flavobacterium spp. (Supplementary Fig. S1). The phylogenetic distribution obtained two branches for these Rositz sequences, which were related to different Flavobacterium species, i.e. Flavobacterium aquatile and Flavobacterium gondwanense.

Firmicutes. The only sequence related to Firmicutes, Rositz23 Ros53, was most similar to clone HuCB5 from human colonic samples (Holta et al., 2002; Fig. 4).

DISCUSSION

The Shannon diversity indices of 3·93 (library Rositz1) and 3·78 (library Rositz23) of the investigated sample seem low compared to those of other studies, which report Shannon diversity indices between 4 and 5 for soils (Schloss &
Fig. 2. Maximum-likelihood phylogenetic tree of Alphaproteobacteria sequences from the active stage of the Rositz samples with Betaproteobacteria sequences as outgroup. The total numbers of the respective Rositz sequence types in the two clone libraries are given in parentheses. The individual groups include the following representative sequences. Sphingomonas group: Sphingomonas sp. DhA-95 (hydrocarbon-contaminated soil; AF177917); Rositz1 Ro46 (AM237246). Zymomonas group: Sphingomonas sp. B18 (freshwater; AF410927); Rositz1 Ro31 (AM237248).

Fig. 1. (a) Maximum-likelihood phylogenetic tree of Gammaproteobacteria sequences from the active stage of the Rositz samples with Betaproteobacteria sequences as outgroup. The total numbers of the respective Rositz sequence types in the two clone libraries are given in parentheses. The individual groups include the following representative sequences. Acinetobacter group: Rositz23 Ros98 (AM237238); group clone TRS13, uncultured bacterium TRS13 (plant rhizosphere; AJ005875); Rositz1 Ro120 (AM237233; 8 clones library Rositz1; 2 clones library Rositz23); Rositz23 Ros119 (AM237241). The Pseudomonas group is illustrated in (b). (b) Maximum-likelihood phylogenetic tree of the genus Pseudomonas sequences from the active stage of the Rositz samples with the Alcanivorax borkumensis sequence as outgroup. The phylogenetic tree was calculated without a filter. The total numbers of the respective Rositz sequence types in the two clone libraries are given in parentheses. The individual groups include the following representative sequences. P. anguilliseptica group: P. anguilliseptica [DSM 12111 (T)]; Rositz1 Ro112 (AM237221; 12 clones library Rositz1; 1 clone library Rositz23); Rositz23 Ros110 (AM237215; 11 clones library Rositz1; 11 clones library Rositz23). P. putida group: Rositz1 Ro132 (AM237220). P. syringae group: Rositz1 Ro71 (AM237217; 3 clones library Rositz1; 6 clones library Rositz23); Rositz1 Ro90 (AM237216; 2 clones library Rositz1); Rositz23 Ros81 (AM237224); Rositz23 Ros94 (AM237223; 2 clones library Rositz23). P. fluorescens group: Rositz23 Ros4 (AM237227). *The classification of the Pseudomonas lineages followed that proposed by Kuske et al. (1999). (T), type strain.
One aspect might be that the Rositz material came from an industrial area and that significant portions of the material were from lower soil horizons, while most other studies deal with soil from the A and B horizons of biologically active areas. Furthermore, Shannon indices of $3 \pm 76$ for surface material and $3 \pm 70$ for subsurface material (1–7 cm) were determined in a study of stony, gravely sand sampled in Antarctica, and the diversities were significantly lower ($2 \pm 76$ and $2 \pm 93$, respectively) for hydrocarbon-contaminated soil from the same area (Saul et al., 2005).

The soil with aged petroleum contamination which was studied by Kaplan & Kitts (2004) showed a Shannon diversity index of $1 \pm 45$, which increased to $1 \pm 6$ during the first 3 months of treatment. The latter values were calculated based on terminal restriction fragment length polymorphism (TRFLP) analysis and are therefore not directly comparable to sequence-based calculations, but there is an obvious trend towards lower Shannon indices in soils with aged hydrocarbon contamination. The detected sequence diversities in the two extracts of this study were similar.

$F_{ST}$ regards only species (here, OTUs) which are common to both samples (here, libraries). Therefore the low $F_{ST}$ is due to the overlap between the two OTU collections. Such an overlap has to be expected when a fractionated analysis of the extracts of one sample is done. However, calculation of $F_{ST}$ does not regard abundance of individual OTUs in the two samples. The results of the $P$ and $F_{ST}$ tests do not support the null hypothesis that samples Rositz1 and Rositz2 are from the same population. The differences in diversity and composition between the two samples are significant.

**Fig. 3.** Maximum-likelihood phylogenetic tree of Betaproteobacteria sequences from the active stage of the Rositz samples with Gammaproteobacteria sequences as outgroup. The total numbers of the respective Rositz sequence types in the two clone libraries are given in parentheses. The individual groups include the following representative sequences. *Rhodobacter* group: uncultured bacterium GK52-77 (freshwater; AJ290040); uncultured bacterium oxSCC-37 (paddy soil; AJ387875); Rositz1 Ro29 (AM237267); Rositz1 Ro34 (AM237266; 3 clones library Rositz1; 7 clones library Rositz23); Rositz1 Ro68 (AM237262). *Thiobacillus* group: uncultured bacterium RB7C6 (groundwater-treating reactor; AF407385); Rositz23 Ros12 (AM237274); Rositz23 Ros69 (AM237271; 4 clones library Rositz1; 3 clones library Rositz23).
Rositz23 originate from the same community. Based on the existing dataset, the two clone libraries were considered to differ significantly. To further compare the two datasets, abundance-based Jaccard (\(J_{abd}\)) and Sorensen (\(S_{abd}\)) estimates were calculated, in which unseen shared species were also considered (Chao et al., 2005). The two estimators 
\[ J_{abd} = 0.7382 \text{ (variance 0.1900)} \]
\[ S_{abd} = 0.8494 \text{ (0.1341)} \] 
demonstrate that, if sample sizes were very large, one would expect most sequences in the two libraries to belong to OTUs which are shared between both libraries. As a transition between the fractionated extracts suggests itself, this result would be no surprise. The only difference which is expected to be preserved is the different abundance variation of certain OTUs in the individual libraries.

Predominance of Gammaproteobacteria

SSU rRNA sequencing reflected considerable diversity of the microflora, in which, however, Gammaproteobacteria clearly dominated. A similar phenomenon is known from cultivation-based analyses of environmental samples and has been termed ‘\(\gamma\)-shift’. It occurs under conditions of nutrient oversupply (Ammann et al., 1995). Given that in pristine soils Bacillus spp. (Hugenholtz et al., 1998), Bacteroidetes, Gram-positives with low GC content (Borneman et al., 1996), Acidobacteria (Saul et al., 2005) and Alphaproteobacteria (McCaig et al., 1999) are abundant, the dominance of Gammaproteobacteria may result from the degradation of high levels of contaminants. The ratio between Proteobacteria and the Acidobacterium division might also increase with the trophic level of soils (Smit et al., 2001), a tendency which is also visible when bacterial communities of hydrocarbon-contaminated and corresponding uncontaminated soils are compared (Saul et al., 2005). A \(\gamma\)-shift has also been reported to occur upon contamination of soil (Chao & Hsu, 2004) and Arctic sea ice (Gerdes et al., 2005). In the present case (Rositz), the contamination with 8000 milligrams mineral oil hydrocarbons per kilogram soil dry weight provided a high nutrient supply for degraders. This might have caused a \(\gamma\)-shift, although the ability to degrade mineral oil hydrocarbons is not limited to Gammaproteobacteria (Mills et al., 2003; Radwan et al., 1995; Vomberg & Klinner, 2000).

As a consequence of work with laboratory strains, the pseudomonads, which dominate in this study, are often considered as model organisms for pollutant degradation. In fact, several studies of contaminated environments have reported a dominance of pseudomonads. They are prevalent in mineral oil hydrocarbon-contaminated soil under bioremediation (Evans et al., 2004; Mills et al., 2003), but also in pristine rhizosphere soils (Kuske et al., 1999; Marilley & Aragno, 1999), which are commonly considered to be nutrient-rich habitats. Pseudomonas isolates from pinyon-juniper woodland soils have been found to affiliate with the P. syringae and P. fluorescens lineages (Kuske et al., 1999), although, in contrast to our study, no members of the P. aeruginosa and P. putida lineages were obtained in that study.

Pseudomonas spp. increased transiently in the early phase of a bioremediation test on soil with mineral oil contamination (Kaplan & Kitts, 2004). This implies that pseudomonads degrade mainly those pollutants which are fully bioavailable (Heiss-Blanquet et al., 2005). At the end of this process, when most of the contamination was gone, other microorganisms were most prevalent (Kaplan & Kitts, 2004).

Preliminary quantitative data on Rositz 3 provide a similar
The Moraxellaceae contributed another significant portion of the clone libraries (Fig. 1a). A high prevalence of Acinetobacter has been reported mostly in connection with diesel contamination. Thus, the Acinetobacter concentration was reported to increase in soil within a few days of contamination with diesel (Chao & Hsu, 2004). Acinetobacter were also the most abundant diesel-degrading isolates during a bioremediation test with diesel-contaminated sandy soil (Gallego et al., 2001), which was biostimulated with nutrients and periodically tilled for aeration. Alkanindiges illinoisensis is able to degrade linear alkane hydrocarbons with a chain length below C_{16} (Bogan et al., 2003).

Four Rositz sequence types affiliated with the Xanthomonadaceae (Fig. 1a). The genus Xanthomonas has been detected in petroleum hydrocarbon-contaminated soils (Juck et al., 2000). It has also been reported that soil isolates related to Xanthomonas spp. and Stenotrophomonas spp. are able to utilize alkanes (Tesar et al., 2002).

Another two Rositz sequences affiliated with the alkane degrader Hydrocarboniphaga effusa, a novel member of the Gammaproteobacteria from soil contaminated with fuel oil (Palleroni et al., 2004).

Diversity beyond Gammaproteobacteria: known degraders and unexpected genera

The detected Alphaproteobacteria were mostly sphingomonads, which are characteristic of a variety of subsurface environments, and which are known for their broad degradation potential (Thiel et al., 2005; Vomberg & Klínen, 2000). Consequently, such micro-organisms might have played a significant role in the bioremediation of the investigated soil. Similar to the present analysis, sphingomonads represented a dominant group besides the pseudomonads in a bioreactor treating petroleum-contaminated soil (Mills et al., 2003). However, Sphingomonas prevalence has been low in other bioremediation systems investigated with culture-independent methods (Kaplan & Kitts, 2004), or sphingomonads have not been included in corresponding hybridization analyses (Chao & Hsu, 2004).

The other main Rositz lineage of Alphaproteobacteria consisted of Zymomonas members, which are typically isolated from ethanol-fermenting samples (Dien et al., 2003), and not in degrading soil samples. Rositz sequences showed greater similarities to members of aquatic microflora, such as Sphingomonas sp. B18 (O’Sullivan et al., 2002; Urbach et al., 2001; Wolf et al., 2003) than to the typical ethanol-producing Zymomonas. The other Rositz sequences of this class showed a high similarity to the methyloathrophic Methylobacterium spp. and to Rhodobacter sphaeroides, a facultative phototrophic bacterium (Dryden & Kaplan, 1990). Sequences related to the Rhodobacteriaceae group have also been obtained from an oil-enriched microcosm (Yakimov et al., 2005), and the genus Methylobacterium has been detected in petroleum hydrocarbon-contaminated soils (Juck et al., 2000).

The Rositz sequences of Betaproteobacteria showed a striking relationship to sequences or strains from contaminated or pristine aquatic ecosystems, in which members of this class dominate. Rhodoferax and Acidovorax-related Rositz sequences were quite common among the Rositz clones. However, no degradation potential for mineral oil hydrocarbons has been described for the phototrophic Rhodoferax. The genus Acidovorax predominated in enrichments from northern soils with polycyclic aromatic hydrocarbons as carbon source (Eriksson et al., 2003). However, similar sequences to the Rositz clones have been found in an aquatic habitat (Huber et al., 2003). Another cluster of sequences was similar to those of Thiobacillus, a genus which is known for autotrophic oxidation of sulfur compounds, although some species can also grow heterotrophically (Kelly et al., 2005). A sequence similar to Thiobacillus (clone RB7C6) has also been found in monochlorobenzene-contaminated groundwater (Alfreider et al., 2002).

An uncultured bacterium (clone 36-9) dominated the clone library from an aquifer contaminated with coal tar waste (Bakermans & Madsen, 2002), a contamination similar to that of Rositz. Significant numbers of this kind of sequence were also found in the clone libraries of this study.

Actinobacteria are dominant in pristine soils and sediments (Macnaughton et al., 1999; Saul et al., 2005), and they grow slowly, in contrast to Pseudomonas spp. (Bouchez et al., 1997; Vildanova-Martsishin et al., 2002). Some Actinobacteria have been described as able to degrade mineral oil hydrocarbons or related substrates (Greene et al., 2000). Rositz1 Ro106 had a similar sequence to that of clone LTUGr00156 from soil with aged hydrocarbon contamination (Iida et al., 2002; Kaplan & Kitts, 2004). However, the closest relatives of all other Actinobacteria sequences in this study did not appear to be of importance with respect to mineral oil hydrocarbon contamination.

Flavobacteria are chemo-organotrophic, and are isolated from soil and aquatic samples, but also from diseased fish. Members of the genus Flavobacterium have previously been reported to increase in the mid phase of a bioremediation process on petroleum-contaminated soil (Kaplan & Kitts, 2004). Sequences similar to the Rositz sequences have been obtained from a well-ventilated agricultural soil (clone S026; Furlong et al., 2002), but also from groundwater with aged chlorophenol contamination (Männistö et al., 1999) and groundwater biofilms (Ross et al., 2001).

Comparison to other mineral oil hydrocarbon-degrading communities

Several investigations have accumulated knowledge on the microbial diversity of soils contaminated with mineral oil...
hydrocarbon (Evans et al., 2004; Kaplan & Kitts, 2004; Plaza et al., 2003; Radwan et al., 1995). On the one hand, it is difficult to draw general conclusions on the microflora in hydrocarbon-contaminated soil from a small number of individual analyses, considering the broad spectrum of possible pollutants and the varying conditions with respect to age of the contamination, humidity of the soil, land use, etc. On the other hand, rules regarding the occurrence of micro-organisms appear to exist; for example, Sphingomonas species have been detected in all investigated soils with polycyclic aromatic hydrocarbon contamination, independently of the geological and chemical properties of the soils and their geographic origin (Leys et al., 2004).

The analyses of microbial diversity in mineral oil hydrocarbon-contaminated soils at the SSU rDNA level have shown some differences to the SSU rRNA analysis of the Rositz heap in this study. Juck et al. (2000) found Actinobacteria to be dominant in a cold-adapted bacterial community in petroleum hydrocarbon-contaminated soils with no degradation activity. Also Proteobacteria were abundant, in particular Gammaproteobacteria and Alphaproteobacteria. A number of spongomonads were detected, but no pseudomonads, which were the dominant genus in the active stage of bioremediation of heap Rositz 3. The bioreactor samples investigated by Mills et al. (2003) were originally amended with inorganic nutrients (nitrogen : phosphate, 16 : 1) and had a similar phylogenetic distribution to that of the Rositz sample. Pseudomonads and spongomonads were most abundant. In contrast, Acinetobacter spp., which also have a known degradation potential for mineral oil hydrocarbons, Acidovorax spp. and Rhodoferax spp. were not detected. Kaplan & Kitts (2004) investigated a treatment unit for petroleum-contaminated land, in which the soil heap had been contaminated with 2440 milligrams petroleum hydrocarbon per kilogram soil, i.e. one-third of the Rositz contamination. The authors analysed clone libraries from samples taken at day 14 (initial fast-degradation phase) and at day 56 (slow-degradation phase). The clone library of day 14 was dominated by Flavobacterium spp., Sphingomonas spp. and Azocarcus spp., and it showed some differences to the Rositz analysis. Interestingly the analysis of the clone library of day 56 contained the same genera as found in Rositz soil, although the later, because of its higher contamination level, might still have been in an earlier degradation phase, perhaps one comparable with day 14 of the petroleum-contaminated land treatment.

Considering the differences in contamination, soil types and treatment procedures, it is interesting overall to note that the rRNA-based analysis in this study revealed several genera, e.g. Pseudomonas, Sphingomonas, Acidotrophicus, Acidovorax and Flavobacterium, as probably important that have previously been reported to be relevant in rDNA-based studies.

Overall, sequences obtained from the bioremediation system Rositz 3 had a wide phylogenetic distribution. Similar clones or strains have been found in a variety of habitats, some of which were contaminated with mineral oil hydrocarbons and others of which were pristine. According to the literature, representatives of nearly all related genera can degrade such contamination. Some genera, however, such as Sphingomonas and Rhodoferax, have so far not been found in soils with mineral oil hydrocarbon contamination.

Effects of sequential nucleic acid extraction

The initial hypothesis was that a greater number of sequences from bacteria which are hard to lyse would be found in extracts after more than one bead treatment. Because of their cell wall structure, Gram-positive bacteria are harder to lyse than Gram-negative bacteria (Bürgmann et al., 2001). However, sequences from Gram-positive bacteria were as infrequent in the extract after three bead treatments of the soil sample as in the extract after the first bead treatment. Furthermore, the sequences from the two sequential extractions showed no significant difference in phylogenetic distribution at the phylum or subphylum level (Table 1).

A more detailed analysis, however, demonstrated that certain sequences were specific to one of the sequential extractions. Some genera were only detected in one of the two sequential extracts. For example, Sphingomonas sequences were detectable predominantly in the first extract, and Zymomonas in the second/third extract (Fig. 2). Sequences of genera which were detected in both extracts often formed separate sequence clusters for the first bead treatment and the second/third treatment. For example, P. anguilliseptica was typical for the first, and P. syringae for the second/third extract (Fig. 1b). Acidovorax-related sequences also formed two clusters, in which the extracts were represented to different extents (Fig. 3).

Different frequencies of sequences in the sequential extracts do not necessarily result from differences in cell wall stability. An alternative assumption is that most of the nucleic acids extracted by the second bead treatment originated from micro-organisms located inside soil particles, and that nucleic acids from the first bead treatment were obtained mostly from micro-organisms located at the outside of the soil particles. Then, the extraction of nucleic acids by sequential bead beating would have resulted in the lysis of micro-organisms from different soil niches, i.e. inside or outside soil particles (Bürgmann et al., 2001; Ranjard & Richaume, 2001). Different degraders could coexist in contaminated soil, and selection might be based, among other factors, on the bioavailability of the contaminant (Friedrich et al., 2000; Grosser et al., 2000). Micro-organisms inside soil particles might be adapted to degrade mineral oil hydrocarbons, which are adsorbed to natural organic matter and consequently available to fewer micro-organisms than the free compounds. In addition, changes of redox potential have been reported to be the factor with the strongest effect on soil microflora subsequent to contamination (Fahy et al., 2005). The partial pressure of oxygen
inside soil particles will always be lower than that in the outer layers of the particles, even in aerated soil systems such as the one investigated here, and therefore related bacteria with different frequencies in the sequential extracts might also differ in their response to oxygen (e.g. preferring microaerobic to aerobic conditions).

Abundant sequences in the clone libraries of the present study probably included those micro-organisms which were the main degraders of mineral oil hydrocarbons at this stage of bioremediation, because only the active micro-organisms were detected.

There are many different niches available in soil, providing optimal conditions for many microbial consortia involved in degradation. Not all members of a microbial community will be active in degradation at every time point. A succession of different members of the community has been observed in other systems, even at the DNA level (Kaplan & Kitts, 2004). At a later point in the bioremediation process, other bacteria probably would have been detectable in the soil heap Rositz 3.

Further investigation of bioremediation system Rositz 3 at other time points of the process should verify the presence of the micro-organisms detected so far. The analysis of SSU rRNA clarified the identity of the micro-organisms, but provided no direct evidence of their function in the soil or their metabolic potential. Determination of marker genes for degradation of mineral oil hydrocarbons will be necessary to address this aspect.

ACKNOWLEDGEMENTS

We thank Mr M. Hagedorn (Dierichs & Hagedorn Consulting) and Mr J. Seelig (Bodensanierungsgesellschaft Seifersbach) for providing samples and the results of their chemical analyses. N. P. is a recipient of a PhD fellowship from the Deutsche Bundesstiftung Umwelt. M. S. acknowledges generous long-term support from the Deutsche Bundesstiftung Umwelt.

REFERENCES


Microbial diversity in a bioremediation system


