Molecular Ecology 1996, 5, 427–436

Diversity and distribution of subterranean bacteria in groundwater at Oklo in Gabon, Africa, as determined by 16S rRNA gene sequencing

K. PEDERSEN, J. ARLINGER, L. HALLBECK and C. PETTERSSON*
*Department of General and Marine Microbiology, The Lundberg Institute, Göteborg University, Medicinaregatan 9C, S-413 90 Göteborg, Sweden and *Department of Water and Environmental Studies, Linköping University, S-581 83 Linköping, Sweden*

Abstract

This paper describes how groundwater was sampled, DNA extracted, amplified and cloned and how information available in the ribosomal 16S rRNA gene was used for mapping diversity and distribution of subterranean bacteria in groundwater at the Bangombé site in the Oklo region. The results showed that this site was inhabited by a diversified population of bacteria. Each borehole was dominated by species that did not dominate in any of the other boreholes; a result that probably reflects documented differences in the geochemical environment. Two of the sequences obtained were identified at genus level to represent *Acinetobacter* and *Zoogloea*, but most of the 44 sequences found were only distantly related to species in the DNA database. The deepest borehole, BAX01 (105 m), had the highest number of bacteria and also total organic carbon (TOC). This borehole harboured only Proteobacteria beta group sequences while sequences related to Proteobacteria beta, gamma and delta groups and Gram-positive bacteria were found in the other four boreholes. Two of the boreholes, BAX02 (34 m) and BAX04 (10 m) had many 16S rRNA gene sequences in common and they also had similar counts of bacteria, content of TOC, pH and equal conductivity, suggesting a hydraulic connection between them.

Keywords: 16S rRNA, bacteria, Bangombé, diversity, DNA sequencing, Oklo, phylogeny

Received 2 October 1995; revision accepted 12 December 1995

Introduction

The discovery in 1972 of isotopic anomalies in the Oklo uranium deposit in Gabon, Africa, was an important scientific event which revealed the existence of natural nuclear reactions in the distant past (Bodu *et al.* 1972). The Oklo reactors became critical about 2 billion years ago and operated intermittently for some 10^5–10^6 years. Roughly 1000–2000 t of uranium was initially present as fuel of which about 6–12 t consisted of ^{235}U that underwent fission and produced some 4 t of plutonium. The reactors operated at temperatures of up to 600 °C and pressures of 800–1000 bar (Miller *et al.* 1994).

Long-lived radioactive waste, produced by the nuclear energy industry and from other sources, must find safe disposal. Many countries are planning to build underground repositories for the disposal of such waste. Although it may be possible to make reasonably accurate predictions of the outcome of single processes in a repository, the long-term behaviour of the total geological environment is very complex. Therefore, natural analogues, or natural systems, are investigated for the understanding of processes that proceed over geological times, as a part of the task to develop safe disposal concepts. A generally accepted definition of the term 'natural analogue' is 'An occurrence of materials or processes which resembles those expected in a proposed geological radioactive waste repository' (Miller *et al.* 1994). The Oklo region contains the only known examples of natural fission reactors and is, therefore, perhaps the best known natural analogue for geological disposal of radioactive waste. It has been investigated by multidisciplinary consortia of researchers and new achievements were recently published (Blanc & Maravic 1995).

Understanding the microbial part of biogeochemical

Correspondence: K. Pedersen. Tel.: + 46-31-7732578. Fax: + 46-31-7732599. E-mail: pedersen@gmm.gu.se

Borehole	Screened section, m	Sample date	Temp. °C	E_h mV	pH	Conductivity mS/m
BAX01	96–105	06 Mar. 1993	28.6	74	7.0	272
		09 Jul. 1994	27.6	12	6.5	227
BAX02	27.2–33.9	03 Mar. 1993	28.3	35	6.2	89
		12 Jul. 1994	27.1	67	5.7	64
BAX03	11.9–12.5	03 Mar. 1993	29.0	26	6.7	162
		10 Jul. 1994	29.0	132	6.0	104
BAX04	8.9–10.2	03 Mar. 1993	26.1	130	5.9	67
		13 Jul. 1994	28.3	104	5.5	61
BAX05	23.8–31	06 Mar. 1993	26.1	147	6.6	89
		14 Jul. 1994	26.7	-36	6.0	112
BAX07	4.5–6.5	11 Jul. 1994	28.0	69	4.8	15

Table 1 Borehole information and groundwater chemistry data for the Bangombé nuclear reactor zone.

processes which may have influenced radionuclide migration from the Oklo natural analogue requires information about diversity and distribution of present bacteria. Traditional studies of diversity of microbial analogue communities have been incomplete because of an inability to identify and quantify all contributing populations. New methods in molecular biology are now being used to overcome this problem (Ekendahl *et al.* 1994; Amann *et al.* 1995). Genotypic identification of bacteria using sequence information available in ribosomal RNA genes has become a very important complement to phenotypic identification and classification techniques (Boivin-Jahns *et al.* 1995). This paper describes how groundwater was sampled, DNA extracted, amplified and cloned and how the information available in the ribosomal 16S rRNA gene was used for mapping the diversity and distribution of subterranean bacterial populations in the Oklo region. The total number of bacteria was also determined, concomitant with analysis of the concentration of total organic carbon (TOC).

Materials and methods

Site description

The Bangombé reactor was discovered in the 1980s and is located some 30 km south of the Oklo deposit. This reactor zone occurs at a depth of 16 m and the investigated boreholes were drilled during 1992 to evaluate the geological surroundings of the reactor as shown in Fig. 1. The boreholes were packed-off at different depth sections (Table 1) and a submersible pump was used to bring groundwater up to ground surface. Several borehole volumes were pumped during a day or two before sampling to ensure clean and representative aquifer water. Table 1 presents borehole information and groundwater chemistry data that provide a background for interpretation of achieved data.

Sampling for determination of the total number of bacteria

Samples of 40 mL from BAX01, BAX02, BAX03, BAX04 and BAX05 (only 1993) were preserved with formaldehyde (2% final concentration) and transported to the laboratory in Göteborg. During the July 1994 expedition, sampling was repeated three times with intervals of approximately 1 h.

Determination of the total number of bacteria

Acridine orange direct count (AODC) (Hobbie *et al.* 1977) was used to determine the total number of cells in the water. Nuclepore filters of 0.2 μm pore size and 13 mm in diameter were prestained with Sudan-black and air dried.

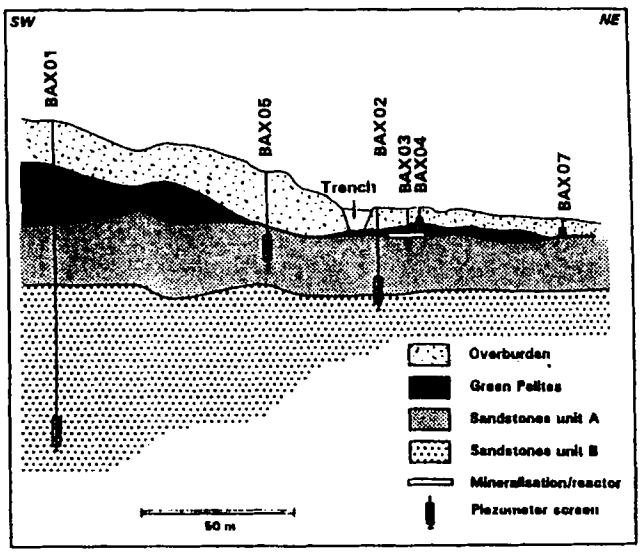

Fig. 1 Conceptual geological section of the Bangombé reactor showing the location of the boreholes and screen positions. The depths of the boreholes are as follows, BAX01, 105 m; BAX02, 33.9 m; BAX03, 12.5 m; BAX04, 10.2 m, BAX05, 31 m; BAX07, 6.5 m.

All solutions were filter sterilized (0.2 μm). The sample was diluted twice its volume with 0.1% oxalic acid and vigorously shaken to reduce clogging of the filters. A portion of the sample was filtered on to a prestained Nuclepore filter at – 20 kPa and stained for 6 min with acridine orange. The number of bacteria was counted using blue light (390–490 nm) under an epifluorescence microscope (Olympus BH-2). Two filters were counted for each water sample. Between 500 and 600 cells or a minimum of 15 fields (0.0064 mm^2 each) were counted on each filter. This procedure will predict a sample mean with a precision of 5% (Niemelä 1983; Hallbeck & Pedersen 1990). The results were calculated as mean values of two filters from each sample.

Determination of TOC

Samples for TOC determination were collected on two occasions, March 1993 and July 1994, concomitant with sampling for determination of total bacterial cell number and 16S rRNA gene sequencing. The concentration of TOC was analysed on a Shimadzu TOC-5000 analyser, utilizing catalytic combustion and IR-detection of produced carbon dioxide. The water samples were acidified to pH 3 and inorganic carbon was removed by flushing the sample with pure air prior to analysis.

16S rRNA gene extraction, PCR-amplification, cloning and sequencing

Groundwater samples from BAX01, BAX02, BAX03, BAX04 and BAX07 were collected directly from the boreholes using a syringe at the outflow of the tubing connected to the submersed borehole pump; 10-mL samples filtered on to sterilized 0.2-μm-pore Nuclepore filters. The filters were deep frozen with dry ice (approximately – 25 °C) and transported deep frozen to the laboratory in Göteborg. The sampling was repeated three times with intervals of approximately 1 h concurrent with sampling for total numbers of bacteria and TOC. BAX07 was sampled once only.

Our protocol for DNA extraction is based, with minor modification, on the procedure described by Marmur (1961) and Wallace (1987). The filters were resuspended in 760 μL of 20 mM Tris-HCl, pH 8.0; 20 mM EDTA; 0.35 M sucrose and incubated with 2 mg/ml lysozyme (Sigma) at 37 °C for 1 h to destroy cell walls. Thereafter the cells were lysed by adding 40 μL 20% sodium dodecyl sulphate (SDS) and proteins were digested with 200 μg/ml proteinase K (Sigma) during an additional incubation at 60 °C for 1 h. The mixture was extracted with an equal volume phenol : chloroform : isoamylalcohol (25 : 24.1), and thereafter three extractions with chloroform : isoamylalcohol (24 : 1, called chisam) so that no cell debris was visible. Obtained DNA was precipitated with 1/3 volume of 10 M NH_4Ac (final concentration 2.5 M) and 2.5 volumes of 99% ethanol. To ensure complete precipitation, 50 μg tRNA was added as a coprecipitant and the mixture was incubated at – 70 °C over night. The precipitate was washed with 100 μL 70% ethanol (v/v) and dried in vacuum for 30 s, dissolved in TE-buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) over night and stored at –20 °C.

One microlitre of the extracted DNA solution was added to a mixture of 10 μL of 10xPCR buffer (Stratagene), 0.2 mM of each nucleotide triphosphate, 0.25 μM of each primer and double-distilled water to a final volume of 100 μL. The samples were treated with 10 mg/ml of RNase A (Sigma) for 15 min at 37 °C and incubated at 95 °C for 5 min, before addition of 1 μL Pfu DNA polymerase (Stratagene) and coating with 100 μL mineral oil (Sigma). A total of 30 cycles were performed at 95 °C (30 s), 55 °C (1 min), 72 °C (2 min) followed by a final incubation at 72 °C for 10 min.

The 5′ and 3′ primers used matched the universally conserved positions 519–536 and 1392–1404, *Escherichia coli* Brosius numbering (Brosius *et al.* 1978). These were chosen to ensure that both bacterial, archaeal and eucaryal species could be amplified. The amplification products were purified with the QIAEX agarose extraction kit (Qiagen) following the specification of the manufacturer and were finally diluted in 20 μL TE-buffer and stored at –20 °C.

The purified samples were cloned with the pCR-Script SK(+) cloning kit (Stratagene) following the specification of the manufacturer. From each DNA extraction, a total of 10 white colonies containing the insert was randomly picked. The colonies were inoculated in 3.5 mL LB (Luria Broth) + Ampicillin overnight at 37 °C. From each culture 0.5 mL was suspended in 0.5 mL of concentrated glycerol and stored at –80 °C. The recombinant plasmids were extracted from the bacteria with the Wizard Magic miniprep kit (Promega).

The sequencing was accomplished by Autoread Sequencing kit (Pharmacia Biotech) following the manufactur's instructions. All clones were sequenced using the 3′ primer 907–926 (Ekendahl *et al.* 1994) labelled with fluorescein. Gel electrophoresis was performed on an ALF DNA Sequencer (Pharmacia Biotech). Sequences obtained from the three separate sampling occasions were pooled and compared directly on the monitor of the ALF manager OS/2 computer, using the ALF manager program, version 2.5. Clones that were identical were put in the same group.

Sequence analysis

The 16S rRNA gene clones were compared with sequences available in the European Molecular Biology Laboratory

(EMBL) database using the FastA procedure in the GCG program package (Genetic Computer Group, WN, USA). This procedure calculates identities between an unknown sequence (clone) and sequenced bacteria in the database. The phylogenetic analysis was performed by using the programs contained in the PHYLIP version 3.5c package (Felsenstein 1989) compiled for PC. Nucleotide positions that could be unambiguously aligned for all clones were included in the analysis. The final data set comprised 353 nucleotide positions, position no. 542–883 [*E. coli* Brosius numbering, but not including 12 inserts in the analysed data set (Brosius *et al.* 1978)]. The distances were calculated using the DNADIST program and a tree was built running the KITCH program with contemporary tips. The KITCH program was run with a randomized input order of data with 10 jumbles and during execution, 171 833 trees were examined. The tree was drawn by using a drawing program, DRAWTREE, also available in the PHYLIP package.

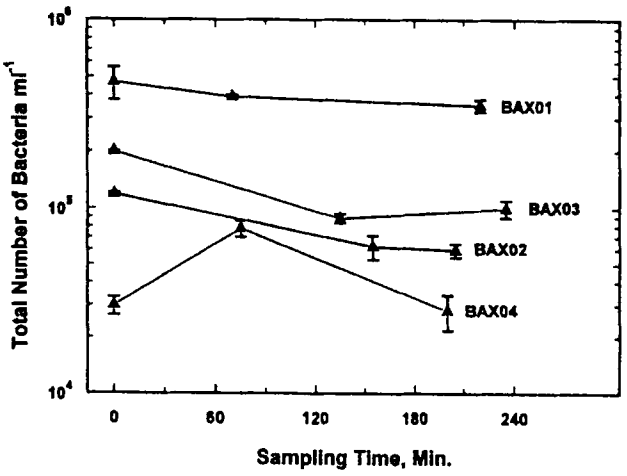

Fig. 2 The total number of bacteria in flowing (pumped) groundwater from the Bangombé boreholes BAX01 to BAX04, sampled in series at three different times during July 1994. Bars denote the standard deviation of the two filters counted for each water sample.

Results

Total number of bacteria and TOC

A comparison of the samples from both expeditions shows that between 1.4 and 5.3 times more bacteria were found in the samples from 1993 than in samples from July 1994 (Table 2). The opposite situation was found for the TOC values which had increased in July 1994 to between 1.2 and 2.9 times concentrations found in March 1993. Each of the boreholes BAX01–04 was sampled three times during approximately 3 h. Figure 2 shows a minor decrease in total numbers from the first to the last sampling time. The numbers were eventually stabilized at the end of the sampling period. The deepest borehole, BAX01, had the highest numbers of bacteria and also the highest concentrations of TOC during both sampling periods. There was a good correlation between the total number of bacteria assayed and the TOC (Fig. 3). The quotients of total bacterial counts measured from both the 1993 and 1994 dates, and TOC (Table 2) were fairly uniform with exception of the total number quotient in BAX01 and the TOC quotient in BAX02 and BAX04.

Distribution and diversity of bacteria analysed by 16S rRNA gene analysis

A total of 130 clones were sequenced, 30 from each borehole BAX01, BAX02, BAX03 and BAX04 and 10 from BAX07. The clones examined could be allocated into 44 specific clone groups; those that were identical were placed in the same clone group, and each such group, together with clone sequences appearing only once, was given a clone group name, ranging from G1 to G44. Figure 4 depicts a phylogenetic tree for 39 of the clone group

Borehole	Total number of bacteria x 10^5/mL			TOC mg/L		
	March 1993 (n = 1)	July 1994 (n = 3)	March 1993/ July 1994	March 1993 (n = 1)	July 1994 (n = 3)	July1994/ March 1993
BAX01	5.8	4.03 (0.73)	1.4	6.6	14.2 (0.4)	2.2
BAX02	2.9	0.80 (0.28)	5.3	1.7	4.9 (0.8)	2.9
BAX03	5.5	1.29 (0.5)	4.2	4.1	8.3 (2.1)	2.0
BAX04	2.4	0.45 (0.24)	5.3	1.2	1.4 (0.05)	1.2
BAX05	1.7	ND	-	ND	ND	-
BAX07	ND	ND	-	ND	4 (n = 1)	-

Table 2 The total numbers of bacteria and the concentration of total organic carbon (TOC) determined in Bangombé groundwater samples collected March 1993 and July 1994. The reduction in total numbers in between March 1993 and July 1994 and the increase in TOC for the same period are shown as quotients. N is the number of independent water samples analysed, numbers in brackets show standard deviations, ND = not determined

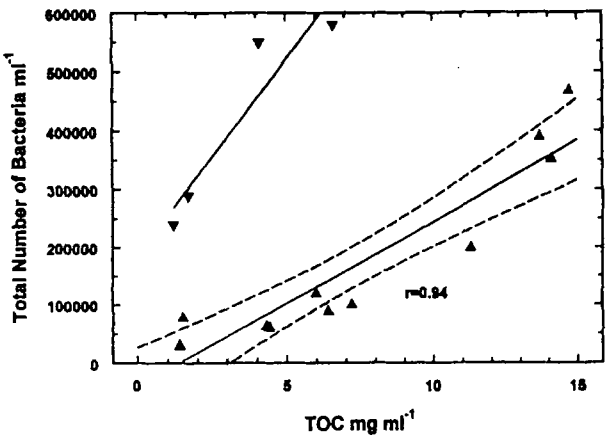

Fig. 3 The correlation between the total number of bacteria and the content of TOC in flowing (pumped) groundwater from the Bangombé boreholes BAX01 to BAX04, sampled March 1993 and in series at three different times during July 1994. ▼, March 1993; ▲ = July 1994. The dashed lines show the 95% confidence interval, the correlation coefficient, $r = 0.94$.

sequences. Four distinct groups of phylogenetically related bacteria were found (Woese 1987), the beta, gamma and delta groups of the Proteobacteria, Gram-positive bacteria and a group (denoted V) with sequences only very distantly related to known and sequenced bacteria reported to the EMBL database. The Proteobacteria beta-group was the largest phylogenetic branch, to which 65 of the found clones belonged (50%) and with the highest diversity, having 20 different clone groups as shown in Table 3. With one exception, only 16S rRNA genes belonging to the Proteobacteria beta group could be demonstrated in BAX01, while the other boreholes harboured 16S rRNA genes from three to five of the listed groups. Between 11 and 16 different 16S rRNA genes were detected in each borehole with this sequencing method. Figure 5 demonstrates the relationship between the number of specific clone groups obtained as the sequences from sample extraction no. 1, 2 and 3 are added. Generally, each set of 10 sequences added new clone groups to the results, but from two to six of the 10 sequences were common when different extractions were compared. There were only 10 clones sequenced from BAX07 and the number of six clone groups from this very shallow borehole can therefore not be directly compared with the other boreholes.

Some of the clone groups have several identical clone sequences, while others represent only a few or one clone sequence. The distribution of the seven most commonly occurring clone sequences for the sampled boreholes, and the number of identical clone sequences these clone groups represent, are shown in Table 4. Typically, each clone group dominates in a separate borehole. Five of them are represented in all three extractions in the borehole they dominate (underlined). The closest species in the database for the dominating clones in Table 4 are shown in Table 5 (cf. Fig. 4). They correspond to 56% of the sequenced clones (73 clones). Several identical clones appeared when different boreholes were compared (Table 6). The boreholes BAX01–04 shared between two and four

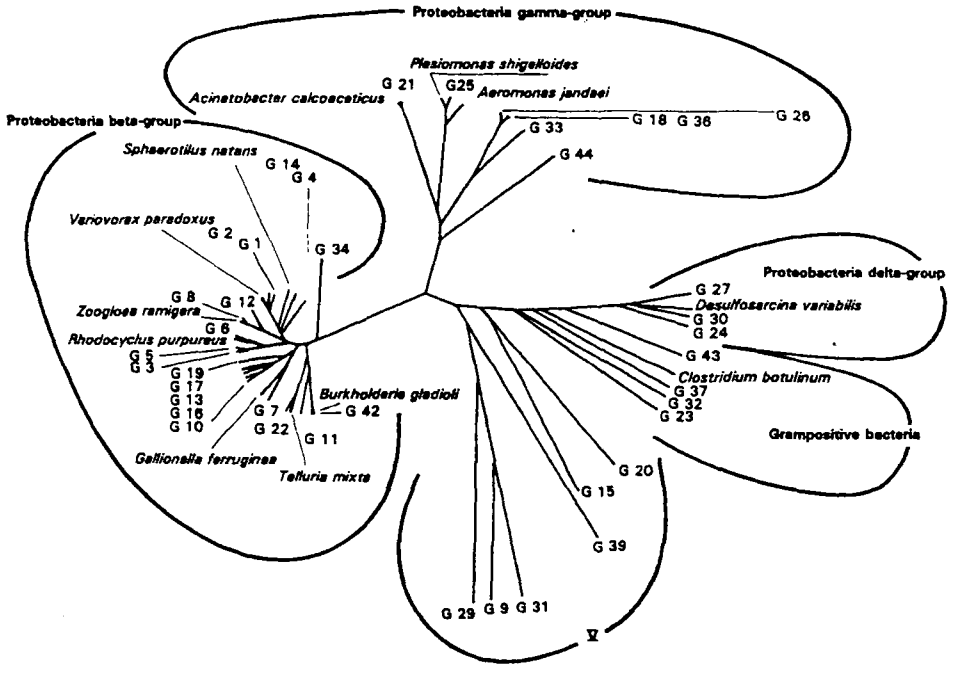

Fig. 4 Evolutionary distance tree based on the 16S rRNA gene sequences of clones from boreholes in the geological surroundings of the Bangombé site. Major phylogenetic groups of bacteria have been designated with their generally accepted names. Cluster V comprises clone sequences only very distantly related to known, sequenced and reported bacteria. As references, some 16S rRNA gene sequences of known bacteria from the EMBL database have been added to the tree and they are indicated with their Latin names. The tree shows 39 of the total of 44 specific clone groups found; five sequences were too short to be incorporated in this analysis. The branch lengths are proportional to calculated evolutionary distances.

Table 3 The numbers of clones that were affiliated in different major branches of the phylogenetic tree by PHYLIP as shown in Fig. 4, distributed over the sampled boreholes. The total number of sequenced clones as well as the number of clone groups with a specific sequence are shown for each borehole. Note that some or several of the specific clone group sequences reported for each borehole may occur in several boreholes, thereby resulting in a smaller sum of specific clone sequences, than the sum of specific clone group sequences for each borehole

Phylogenetic branch in Fig. 4	Number of clones and clone groups										Sum of clones and clone groups	
	BAX01		BAX02		BAX03		BAX04		BAX07			
	Clones	Groups	Clones	Groups	Clones	Groups	Clones	Groups	Clones	Groups	Clones	Groups
Proteobacteria beta group	29	13	14	4	4	2	11	7	7	3	65	20
Proteobacteria gamma group	0	0	7	4	5	2	9	5	2	2	23	9
Proteobacteria delta group	0	0	1	1	2	2	0	0	0	0	3	3
Gram-positive bacteria	0	0	1	1	15	2	1	1	1	1	18	4
Remaining sequences (V)	1	1	7	2	4	3	9	3	0	0	21	8
Sum of clones and groups	30	14	30	12	30	11	30	16	10	6	130	44

Clone group	BAX01	BAX02	BAX03	BAX04	BAX07
G4	2 (0 + 0 + 2)	4 (0 + 4 + 0)	0	1 (0 + 1 + 0)	0
G5	*8 (2 + 3 + 3)*	0	3 (2 + 1 + 0)	2 (0 + 1 + 1)	0
G8	*7 (1 + 3 + 3)*	0	0	0	0
G16	0	*8 (4 + 2 + 2)*	0	1 (0 + 1 + 0)	0
G21	0	2 (0 + 1 + 1)	4 (3 + 0 + 1)	2 (1 + 0 + 1)	1
G23	0	1 (0 + 0 + 1)	*14 (4 + 6 + 4)*	0	0
G15	0	6 (5 + 0 + 1)	0	*7 (1 + 4 + 2)*	0

Table 4 The dominating bacterial clone groups screened from Bangombé groundwater and their distribution over the sampled boreholes. The total number of clones in each clone group is followed, in parentheses, by the numbers achieved in the extractions from samples 1, 2 and 3, respectively. The numbers for clone groups that occur in all three extractions are in italic

identical clone sequences with the exception of the BAX02-BAX04 couple, which shared 11 identical clone sequences distributed over five different clone groups.

EMBL accession numbers

The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers: G1, X91173; G2, X91179; G3, X91174; G4, X91175; G5, X91176; G6, X91177; G7, X91178; G8, X91180; G9, X91181; G10, X91182; G11, X91183; G12, X91184; G13, X91185; G14, X91186; G15, X91187; G16, X91188; G19, X91271; G20, X91272; G21, X91273; G22, X91274; G23, X91275; G24, XX91276; G25, X91277; G26, X91278; G27, X91279; G28, X91280; G29, X91281; G30, X91282; G31, X91283; G32, X91284; G33, X91285; G34, X91286; G35, X91287; G36,

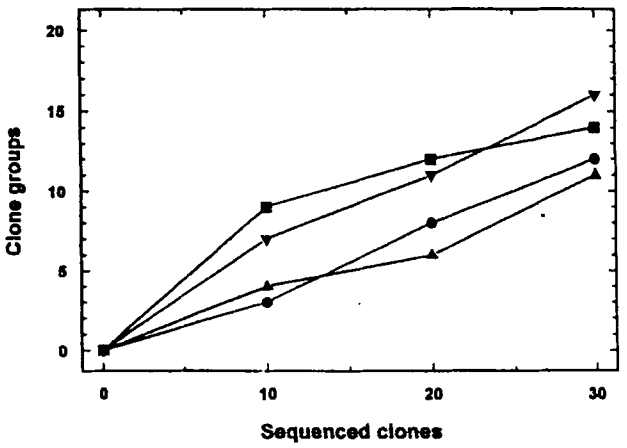

Fig. 5 The number of specific clone groups found as a function of the number of sequenced clones from different boreholes. ■, BAX01; ●, BAX02; ▼, BAX03; ▲, BAX04.

Table 5 Comparison of the identity between some of the sequenced and most commonly occurring Bangombé clone groups and 16S rRNA sequences in the EMBL database by September 1995. The percentage identity between the obtained sequence and the closest related bacterium in the database is listed

Phylogenetic group	Clone group	No. of identical clone sequences	EMBL Accession no. for clone sequences	Closest species in the EMBL database	EMBL Accession no. for closest species	Identity %
Proteobacteria beta group	G4	7	X91175	*Sphaerotilus natans*	Z18534	94.8
	G5	13	X91176	*Rhodocyclus purpureus*	M34132	94.5
	G8	7	X91180	*Zoogloea ramigera*	D14257	98.3
	G16	9	X91188	*Rhodocyclus tenuis*	D16209	91.7
Proteobacteria gamma group	G21	9	X91273	*Acinetobacter junii*	X81658	99.4
Gram-positive bacteria	G23	15	X91275	*Afipia felis*	M65248	91.6
Group V	G15	13	X91187	*Thiomicrospira denitrificans*	L40808	87.0

X91288; G37, X91289; G38, X91290; G39, X91291; G40, X91292; G41, X91293; G42, X91294; G43, X91295; G44, X91296.

Discussion

The total number of bacteria obtained in the Bangombé groundwater was in the range of other previously measured groundwater environments; from 10^3 up to 10^7 bacteria per ml (Pedersen & Ekendahl 1990; Pedersen 1993a,b, Pedersen & Karlsson 1995; Pedersen, 1996). Contribution of these bacteria to the TOC can be calculated assuming an average weight of a bacterium to be 2.8×10^{-13} g of which 50% is carbon (Neidhardt *et al.* 1990). This was carried out for BAX01; in March 1993, BAX01 water had 5.8×10^8 bacteria per litre (Table 2) which gives a bacterial contribution to TOC corresponding to 0.081 mg/L. Comparing this number with the TOC analysed on the same occasion, 6.6 mg/L, suggests that the bacteria are responsible for ≈ 1.25% of the TOC in this sample. This proportion was in agreement with the other samples; consequently, bacterial biomass constitutes only a minor part of the TOC content.

A positive correlation between the TOC and the total numbers of bacteria has been demonstrated for deep granitic groundwater (129 m down to 860 m) sampled during the preinvestigation phase of the Äspö Hard Rock laboratory (HRL) in SE Sweden (Pedersen & Ekendahl 1990). The Bangombé samples were from more shallow levels (4.5 m down to 105 m) and had higher TOC than those in the deep Äspö groundwater. Despite this, there was a very good correlation between TOC and total number of bacteria obtained during both sampling periods (Fig. 3). Increasing TOC seems to stimulate bacterial growth, which is a likely conclusion, as TOC may serve as a nutrient and a carbon source for heterotrophic bacteria. However, it is important to note that, at least in these cases, the relationship is not absolute, but relative (cf. Fig. 3). This relative correlation of TOC and bacterial number may be due to differences such as variations in degradability of TOC, species of bacteria present, the ratio between attached and unattached bacteria and also the rate with which the TOC is supplied to the ecosystem.

Nucleic acid probes for 16S rRNA sequences can be used for detection of bacteria and their distribution in various ecosystems (Amann *et al.* 1991; Braun-Howland *et al.* 1992; Risatti *et al.* 1994). Use of such probes requires genetic information about the 16S rRNA sequences of a bacterial population to be studied, or specific probes cannot be constructed. Therefore, when probes are to be used in new environments with unknown bacteria such as those in Bangombé, a first step must be to analyse for genetic diversity of the 16S rRNA gene of the inhabiting populations, as was done here.

We have applied the 16S rRNA gene-sequencing method to five very different subterranean habitats; the Stripa research mine and Äspö hard rock laboratory in Sweden, Maqarin groundwater in Jordan, sand/bentonite buffer clays from a full-scale nuclear waste canister experiment, performed at the Atomic Energy Canada Limited (AECL) underground research laboratory, North of Winnipeg, Canada, and finally the Bangombé site reported here. Descriptions and comparisons of these habitats as well as thorough discussions about the limitations and possibilities with the method are published elsewhere (Ekendahl *et al.* 1994; Pedersen & Karlsson 1995; Pedersen 1996). These investigations comprise more than 600 partially or totally sequenced 16S rRNA gene clones from 60 independent samples. More than 175 specific clone group

	No. of shared clone sequences and clone groups							
	BAX01		BAX02		BAX03		BAX04	
Borehole	Sequences	Groups	Sequences	Groups	Sequences	Groups	Sequences	Groups
BAX01	30	14	2	1	3	1	3	3
BAX02			30	12	4	3	11	5
BAX03					30	11	4	3
BAX04							30	16

Table 6 Numbers of identical clone sequences and clone groups shared between the different boreholes sampled from the subterranean groundwater environment of the Bangombé site

sequences have been found during the investigations and 44 of them came from Bangombé.

When PCR amplification is used for determination of species diversity, the result may be biased due to methodological problems, such as uneven extraction of DNA and biased PCR due to differences in genome size (Farrelly *et al.* 1995). One of the most important biases is that organisms belonging to the Archaean domain have been found to have only one or a few gene copies of the 16S rRNA gene while bacteria can have several copies, from five to seven or more, which will bias towards bacteria (Ward *et al.* 1992). The results presented in this paper were obtained using universal primers only and should, for reasons discussed above, be expected to reveal mainly bacterial diversity and distribution.

Ekendahl *et al.* (1994) has shown that there exists a relation between the numbers of clone groups obtained from a sample and the numbers of different species in the sample. The clone groups with many identical sequences that are arrayed in Table 4 then correspond to bacteria that occurred in relatively large numbers. We found different dominating clone groups in different boreholes (Table 4) indicating that one or a few specific species dominate among bacteria in each borehole. An extrapolation of Fig. 5 suggests that sequencing additional clones would have resulted in the discovery of additional clone groups. Therefore, the absolute number of different bacterial species in Bangombé must be larger. However, sequencing of 30 clones per boreholes certainly was enough to find the predominating bacterial sequence (Table 4). A similar observation was made earlier with this method on attached populations in two levels of the Stripa borehole V2 (Ekendahl *et al.* 1994). Differences in groundwater chemistry, geology, minerals, etc. in each borehole obviously favour different bacteria, a situation not unlike terranean ecosystems.

There is not an accepted value of percentage identity at which two 16S rRNA genes can be concluded to belong to the same genus or species. It can be quite different for different genera. It has been suggested, based on a comparison of rRNA sequences and on DNA–DNA reassociation, that a relationship at species level does not exist at less than 97.5% identity in 16S rRNA. At higher identity values, species identity must be confirmed with DNA–DNA hybridization (Fox *et al.* 1992). Accepting this level as identifying a sequence approximately on the genus level, some conclusions can be made about the sequences reported here. A comparison of obtained 16S rRNA gene sequences with EMBL sequences in the database, reveals 10 sequences that were related on genus level (not shown). The other 34 sequences had such a low identity (below 97.5%) that they should be regarded as either unknown or not reported to the database. This finding of many new and unknown bacterial 16S rRNA sequences in a natural environments is consistent with the results of others (Giovannoni *et al.* 1990; Ward *et al.* 1990, 1992)

Sphaerotilus (G4), *Rhodocyclus* (G5), *Zoogloea* (G8) and *Acinetobacter*-like (G21) sequences were predominating in one or several boreholes (Table 5). There were also three predominating sequences, G15, G16 and G23, that were so distantly related to known species that they must be regarded as new, at least when compared with the DNA database. Two other clones, G8 and G21, had an identity higher than 98% to 16S rRNA sequences in the database, and are therefore probably identified to genus level. Acinetobacters are ubiquitous organisms that are present in soils, water and sewage and it has been estimated that *Acinetobacter* may constitute as much as 0.001% of the total aerobic population of soil and water (Tower 1992). Finding *Acinetobacter*-like sequences (G21) in all boreholes investigated here, except the deep BAX01, is consequently in agreement with what is known about this group. The only other dominating clone which was matched to the database at an identity above 98% was G8, related to *Zoogloea*, and unlike G21, it was found only in BAX01. This genus is often found in environments enriched in organic nutrients (Dugan *et al.* 1992) and if *Zoogloea* should be expected at all in this investigation, borehole BAX01 would be the most the plausible one as it had the highest content of organic material detected in this investigation, 14.2 mg/L.

Hydrological investigations and modelling have shown that there is a regional flow from the recharge area SW of the Bangombé reactor zone (Fig. 1) to the deep sandstone penetrated by BAX01. The water then discharges upwards towards the reactor (Gurban *et al.* 1995). Locally, the reactor environment is mainly influenced by a shal-

low-derived lateral groundwater flow with a minor deep discharging component. This means that some deep groundwater from BAX01 reaches the shallow boreholes, but there is certainly no recharge of groundwater from BAX02–07 to BAX01. There was an absolute dominance of members of the Proteobacteria beta group in BAX01 (Table 3). With one exception, this was the only major phylogenetic branch found in BAX01. The other, more shallow boreholes, were populated by bacteria from three to five of the groups of the phylogenetic tree (Fig. 4) including the Proteobacteria beta group. Some of the beta group members found in the BAX01 borehole were also found in the shallow boreholes (Table 6). This implies that groundwater from BAX01 reaches and mixes with the groundwater in the shallow boreholes, but the shallow groundwater does not reach BAX01, as confirmed by hydrological modelling (Gurban *et al.* 1995). Details about the local flow around the reactor is more uncertain. Table 6 suggests that BAX02 and BAX04 have a hydraulic connection, as they share several common sequences and clone groups, but there is presently no direct hydrological evidence to support this suggestion. However, Tables 2 and 3 show that pH, conductivity, TOC and the total numbers of bacteria agree (with exception for the 1994 TOC values) between BAX02 and BAX04, further supporting the possible existence of a hydraulic connection between these two boreholes.

Conclusions

The results and conclusions achieved with this 16S rRNA sequencing technique demonstrate the strength of using molecular tools in microbial ecology. Much of the information achieved here would have been extremely difficult to obtain using traditional methods in microbiology, which to a very large extent were developed for bacteria within the medicinal and food industries. Most subterranean bacteria will not grow on media used for such purposes (Pedersen 1993a).

The 16S rRNA gene sequencing of DNA extracted from the subterranean environment at the Bangombé site, showed that it was inhabited by a diversified microbiota. Each borehole was dominated by species that did not dominate any of the other boreholes; a result that probably reflects documented differences in the geochemical environment. Two of the clones were identified at the genus level to represent species of *Acinetobacter* and *Zoogloea*, but most of the 44 sequences were only distantly related to species in the database. Therefore, they were regarded as new and unknown species, at least in relation to the database. The deepest borehole, BAX01, had the highest numbers of bacteria and also of TOC. This borehole penetrates a geological stratum that is influenced by a deep regional groundwater flow. The distribution of 16S rRNA genes indicates that a component of this groundwater reaches the shallow environment of the Bangombé reactor. Two of the boreholes, BAX02 and BAX04 contained many 16S rRNA gene sequences in common and they also had similar bacterial counts, TOC, pH and equal conductivity, suggesting that these boreholes are hydrologically connected.

The Bangombé natural analogue for a radioactive waste repository was inhabited by many different bacteria which support the idea that bacteria will also inhabit constructed repositories. The next step will be to gather information about the predominating species and their *in situ* activity. Media can now be directed towards these bacteria and the species of interest may be selected using nucleic acid probes and optical tweezers as described recently (Huber *et al.* 1995). Finally, the influence from the activity of these bacteria on their environment must be studied. Important parameters to study, which relate to radionuclide mobility, are redox effects due to microbial activity, production of complexing agents and mobility of the bacteria and uptake of radionuclides by the bacteria (Pedersen & Karlsson 1995; Pedersen, 1996).

Acknowledgements

This investigation was made possible by the Oklo natural analogue which is operated by Commissariat á l'Energie Atomique (CEA), France and supported by the European Commission (EC). We thank John Smellie for editorial and technical comments on the manuscript. It has been a pleasure to co-operate with Paul-Louis Blanc and Didier Louvat at CEA. This work was supported by Swedish Nuclear Fuel and Waste Management and the Swedish Natural Science Research Council, contract no. 08466.

References

Amann R, Ludwig W, Schleifer K-H (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological Reviews*, **59**, 143–169.

Amann R, Springer N, Ludwig W, Görtz H-D, Schleifer K-H (1991) Identification *in situ* and phylogeny of uncultured bacterial endosymbionts. *Nature*, **351**, 161–164.

Blanc PL, Maravic HV (1996) *OKLO working group Proceedings of the fourth joint EC-CEA progress and final meeting held in Saclay, France, on 22 and 23 June 1995. European Comission Report EUR 16704*, pp. 1–282. Office for official publications of the European Communities, Luxenbourg.

Bodu R, Bouzigues H, Murin N, Pfiffelmann JP (1972) Sur l'existence d'anomalies isotopicues rencontrées dans l'uranium du Gabon. *CR Acad Sci*, **275D**, 1731.

Boivin-Jahns V, Bianchi A, Ruimy R, Garcin J, Daumas S, Christen R (1995) Comparison of phenotypical and molecular methods for the identification of bacterial strains isolated from a deep subsurface environment. *Applied and Environmental Microbiology*, **61**, 3400–3406.

Braun-Howland EB, Danielsen SA, Nierzwicki-Bauer SA (1992) Development of a rapid method for detecting bacterial cells *in situ* using 16S RNA-targeted probes. *Biotechniques*, **13**, 928–933.

Brosius J, Palmer ML, Kennedy PJ, Noller HF (1978) Complete

nucleotide sequence of a ribosomal RNA gene from *Escherichia coli*. *Proceedings of the National Academy of Science*, **75**, 4801–4805.

Dugan PR, Stoner DL, Pickrum HM (1992) The genus *Zoogloea*. In: *The Prokaryotes* (eds Balows A, Trüper HG, Dworkin M, Harder W, Schleifer K-Z), pp. 3137–3143. Springer-Verlag, New York.

Ekendahl S, Arlinger J, Ståhl F, Pedersen K (1994) Characterization of attached bacterial populations in deep granitic groundwater from the Stripa research mine with 16S-rRNA gene sequencing technique and scanning electron microscopy. *Microbiology*, **140**, 1575–1583.

Farrelly V, Rainey FA, Stackebrandt E (1995) Effect of genome size and *rrn* gene copy number on PCR amplification of 16S RNA genes from a mixture of bacterial species. *Applied and Environmental Microbiology*, **61**, 2798–2801.

Felsenstein J (1989) PHYLIP – phylogeny inference package. *Cladistics*, **5**, 164–166.

Fox GE, Wisotzkey JD, Jurtshuk JR P (1992) How close is close: 16S RNA sequence identity may not be sufficient to guarantee species identity. *International Journal of Systematic Bacteriology*, **42**, 166–170.

Giovannoni SJ, Britschgi TB, Moyer CL, Field KG (1990) Genetic diversity in Sargasso sea bacterioplankton. *Nature*, **345**, 60–63.

Gurban I, Ledoux E, Made B, Salignac A-L, Winberg A (1996) Hydrogeology and modelling of flow and solute transport in the Okélobondo and Bangombé deposits. In: *OKLO Working group. Proceedings of the fourth joint EC–CEA progress and final meeting held in Saclay, France, on 22 and 23 June 1995. European Comission Report EUR 16704* (eds Blanc PL, von Maravic H), pp. 1–282. Office for official publications of the European Communities, Luxenbourg.

Hallbeck L, Pedersen K (1990) Culture parameters regulating stalk formation and growth rate of *Gallionella ferruginea*. *Journal of General Microbiology*, **136**, 1675–1680.

Hobbie JE, Daley RJ, Jasper S (1977) Use of nuclepore filter for counting bacteria by fluorescence microscopy. *Applied and Environmental Microbiology*, **33**, 1225–1228.

Huber R, Burggraf S, Mayer T, Barns SM, Rossnagel P, Stetter KO (1995) Isolation of a hyperthermophilic archaeum predicted by *in situ* RNA analysis. *Nature*, **376**, 57–58.

Marmur JA (1961) A procedure for the isolation of deoxyribonucleic acid from microorganisms. *Molecular Biology*, **3**, 208–218.

Miller W, Alexander R, Chapman N, McKinley I, Smellie J (1994) Natural analogue studies in the geological disposal of radioactive waste. *Studies in Environmental Science* **57**, Elsevier, London.

Neidhardt FC, Ingraham JL, Schaechter M (1990) *Physiology of the Bacterial Cell*. Sinauer Associates, Inc., Sunderland.

Niemelä S (1983) *Statistical evaluation of results from quantitative microbiological examinations. Nordic committee on food analysis.* Report no. 1, 2nd edn, ISSN 0281–5303.

Pedersen K (1993a) The deep subterranean biosphere. *Earth-Science Reviews*, **34**, 243–260.

Pedersen K (1993b) Bacterial processes in nuclear waste disposal. *Microbiology Europe*, **1**, 18–24.

Pedersen K (1996) Investigations of subterranean bacteria in deep crystalline bedrock and their importance for the disposal of nuclear waste. *Canadian Journal of Microbiology*, **42**, *in press*.

Pedersen K, Ekendahl S (1990) Distribution and activity of bacteria in deep granitic groundwaters of southeastern Sweden. *Microbial Ecology*, **20**, 37–52.

Pedersen K, Karlsson F (1995) *Investigations of subterranean bacteria – Their influence on performance assessment of radioactive waste disposal. SKB Technical Report* 95–10. Swedish Nuclear Fuel and Waste Co., Stockholm, pp. 1–222.

Risatti JB, Capman WC, Stahl DA (1994) Community structure of a microbial mat: The phylogenetic dimension. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 10173–10177.

Tower KJ (1992) The genus *Acinetobacter*. In: *The Prokaryotes* (eds Balows A, Trüper HG, Dworkin M, Harder W, Schleifer K-Z), pp. 3137–3143. Springer-Verlag, New York.

Wallace DM (1987) Large and small-scale phenol extractions. *Methods in Enzymology*, **152**, 33–41.

Ward DM, Bateson MM, Weller R, Ruff-Roberts AL (1992) Ribosomal RNA analysis of microorganisms as they occur in nature. In: *Advances in Microbial Ecology* (ed. Marshall KC), pp. 219–286. Plenum Press Inc., New York.

Ward DM, Weller R, Bateson MM (1990) 16S RNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature*, **345**, 63–65.

Woese C (1987) Bacterial evolution. *Microbiological Reviews*, **51**, 221–271.

This paper is a result of our interest in the deep subterranean biosphere, that has been studied at University of Göteborg under the direction of Karsten Pedersen since 1987. Much of this work relates to performance assessment of radioactive waste disposal. Johanna Arlinger is responsible for the technical performance of our DNA work and is now also an experienced African traveller. Lotta Hallbeck has produced a series of papers on *Gallionella ferruginea* and is responsible for the sequence analysis on GCG and with PHYLIP. Catharina Pettersson holds a position at Linköping University where she studies the relationship between organic molecules such as fulvic acids and metals in natural environments.
