Antonie van Leeuwenhoek **86:** 317–328, 2004.
© 2004 Kluwer Academic Publishers. Printed in the Netherlands.

Investigation of bacterial diversity in Brazilian tropical estuarine sediments reveals high actinobacterial diversity

Fernanda Francischetti Piza[1,*], Paulo Inácio Prado[2] and Gilson Paulo Manfio[1]
*[1]Microbial Resources Division, CPQBA, State University of Campinas (UNICAMP), PO Box 6171, CEP 13081-970, Campinas, SP, Brazil; [2]Environmental Research Centre NEPAM, State University of Campinas (UNICAMP), PO Box 6166, CEP 13084-971, Campinas, SP, Brazil; *Author for correspondence (email: fepiza@uol.com.br; phone: +55 19 3884-7500 ext 261; fax: +55 19 3884-7811)*

Received 1 March 2004; accepted in revised form 19 April 2004

Key words: 16S rRNA gene, *Actinobacteria*, Estuary, Microbial diversity, PAH, Sediment

Abstract

Phylogenetic and statistical analyses of 16S rRNA gene libraries were used for the investigation of actinobacterial communities present in two tropical estuarine sediments (Santos-São Vicente estuary, Brazil). The libraries were constructed from samples collected at the brackish end of the estuary, highly hydrocarbon-contaminated, and at the marine end, uncontaminated. Clones from the marine end of the estuary were all related to sequences from non-cultured *Actinobacteria* and unidentified bacteria recovered from a wide range of environmental samples, whereas clones from the brackish end were mainly related to sequences from cultured *Actinobacteria*. Statistical analyses showed that the community recovered from the hydrocarbon-contaminated sediment sample, at the brackish end, was less diverse than the uncontaminated one, at the marine end, and that the communities from the two libraries were differently structured, suggesting that these may have not originated from the same community. The recognition of the spatial pattern of actinobacterial distribution in a natural environment is a first step towards understanding the way these communities are organized, providing valuable data for further investigations of their taxonomic and functional diversity.

Introduction

Natural habitats usually exhibit highly diversified microbial communities, with both high taxonomic and functional diversity (Fulthorpe et al. 1998; Torsvik et al. 1990). The composition and the relative abundance of species in these communities are determined by the environmental conditions to which they are exposed, including physical and chemical factors, as well as biotic conditions, such as organism's interactions. Consequently, environmental changes due to biogenic or anthropogenic activities, such as farming, agriculture and industrial activities are likely to impose changes in the structure of microbial communities, giving rise to communities better adapted to the new conditions (Fulthorpe et al. 1998; Lindstrom et al. 1999).

Investigations of the temporal and spatial distribution of species in such environments, i.e., the patterns of species distribution, as well as their causes and consequences, are essential for understanding evolutionary processes, such as speciation, and ecological processes, such as succession, community development, and the spread and persistence of species (Levin 1992). Efforts to understand the way communities are organised must, therefore, attempt to discover patterns within systems and to compare them across systems.

The Santos-São Vicente estuarine system, located between the Atlantic Ocean and the Serra do Mar mountain range, in the State of São Paulo, Brazil

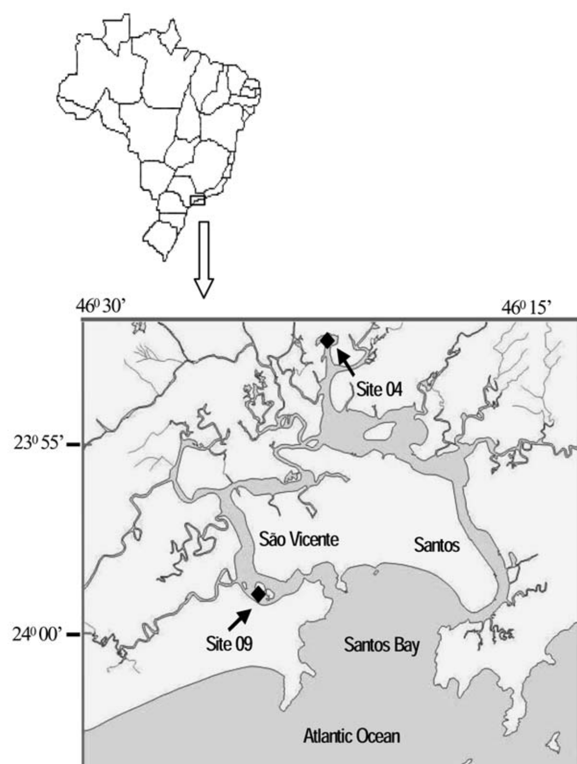

Figure 1. The map of Brazil showing the localization of the Santos-São Vicente estuary, in São Paulo state, and the location of sampling sites 4 and 9. The width of the picture showing the estuary corresponds to 30km.

(Figure 1), exhibit particular environmental characteristics that make it suitable for a wide range of microbiological investigations. Since the 1950s, a wide variety of industries, including petrochemical, steel works and fertiliser production plants, have been installed in this area, amongst a complex network of estuarine channels and mangroves. Industrial pollutants, together with domestic sewage from the nearby cities and effluent from the port of Santos, the major port in Latin America, have caused serious environmental degradation. Despite efforts by the state environmental agency (CETESB – Technology Company for Environmental Sanitation) to prevent the disposal of contaminated effluents into the estuary, recalcitrant and toxic compounds, including polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs) and heavy metals are still detected above the probable effect level (PEL) (Anonymous 1999) in water columns, sediments and aquatic organisms at various sites within this estuary (Anonymous 2001).

In order to initiate the characterisation of the microbial diversity present in the sediments from this estuarine system, two contrasting sites were chosen: one from the brackish end of the estuary that is highly PAH-contaminated and the other from the marine end, uncontaminated. Phylogenetic analysis of the clones recovered in bacterial 16S rRNA gene libraries constructed from these sediments revealed a single group with interesting peculiarities: *Actinobacteria*. The representatives of this group were the only ones, among the various phylogenetic groups recovered, which exhibited a clear differentiation between the two sediment samples analysed and were the major components in the library from the hydrocarbon-contaminated sediment. These peculiarities prompted the separate analysis of this group. Phylogenetic reconstruction was used to assess the taxonomic diversity of *Actinobacteria* present in these libraries and statistical analyses were used to compare the genetic diversity of the actinobacterial communities recovered and to assess the degree of differentiation between them.

Materials and Methods

Study area and sampling procedures

Figure 1 shows the location of the sampling sites in the Santos–São Vicente estuary. Sampling was done in October 2000 at sites 4 (23º52′31,1″ S, 46º22′ 35,9″ W) and 9 (23º58′28,4″ S, 46º24′48,7″ W). Site 4, located at the brackish end of the estuary, receives effluents from steel works and petrochemical and fertilizer industries, whereas site 9, at the marine end of the estuary, receives effluents from domestic sewage. Sediment samples (one for each site) were collected using a Van-veen grab at a maximum depth of 1.3 m below the water surface. The samples were homogenized and sub-samples were transferred: (i) to aluminium boxes, which were kept at 4 °C, for use in physical and chemical analyses and (ii) to sterile tubes, which were immediately frozen on dry ice and kept at − 80 °C, for use in DNA extraction.

Physical and chemical characterization of sediment samples

Physical and chemical data for the sediment samples were obtained from Rodrigues (Rodrigues, 2002), by analyzing aliquots of the same samples used to con-

Table 1. Chemical data derived from estuarine sediment samples from sites 4 and 9

Site	Total n-alkanes ($\mu g.g^{-1}$) [d]	Total aliphatics ($\mu g.g^{-1}$) [d]	Σ PAHs [a] ($ng.g^{-1}$) [d]	CPI [b]	UMC [c]
4	113.9	2,508	407,560	1.83	2,015
9	4.1	4.1	14	5.96	n.d. [e]

[a] Σ PAHs = total polycyclic aromatic hydrocarbons. [b] CPI = carbon preferential index. [c] UMC = unresolved complex mixture. [d] Dry weight of sediments. [e] n.d. = not detected. Data obtained from Rodrigues (Rodrigues, 2002).

struct the 16S rRNA gene libraries. Sediment from site 9 had higher salinity (23.2‰) than sediment from site 4 (16.9‰), due to the proximity of the former to the open sea (see Figure 1), and pH values were 7.4 and 8.2, respectively. The redox potentials were below −300 mV, indicating anoxic environments in both sediments. Sediment from site 4 showed high concentrations of n-alkanes, aliphatic hydrocarbons and PAHs (Table 1), whereas sediment from site 9 had concentration of these compounds just above the threshold of detection (Anonymous 1991). The pollution bio-indicators CPI (Anonymous 1985; Volkman et al. 1992) and UCM (Farrington and Tripp 1975) indicated a predominance of hydrocarbons of anthropogenic origin, mainly from petroleum, in sediment from site 4, while sediment from site 9 showed a predominance of hydrocarbons of biogenic origin.

DNA extraction and purification from sediment samples

The DNA extraction method described by van Elsas and Smalla (van Elsas and Smalla 1995) was used with some modifications: (i) the volumes were adapted to extraction of DNA from 0.5 g of sediment (from 5 to 9 replicates), (ii) lyozyme incubation for 2 hs, (iii) SDS addition before the bead-beating step and (iv) three bead-beating cycles at 3800 rpm for 90 s. DNA extracts were pooled and purified by using the Wizard DNA clean-up system (Promega), according to the manufacturer's recommendations.

PCR amplification of 16S rRNA genes

The primers used for amplification of the bacterial 16S rRNA genes were 27f (Lane 1991) and 1401r (Heuer et al. 1997). The reaction mixtures were as follows: 0.2 μM of each primer, 200 μM of deoxynucleoside triphosphates (Amersham Biosciences), 2 units of Taq DNA polymerase (Amersham Biosciences) and approximately 100 ng of total genomic DNA extracted from the sediment samples. The PCR reactions also contained 0.5 $ng.\mu L^{-1}$ of T4g32p protein, in order to increase the efficiency of Taq DNA polymerase (Schwarz et al. 1990). The amplification program was optimised for bacterial domain-specific 16S rRNA gene as follows: an initial denaturation at 94 °C for 2 min, 10 touchdown cycles of 1 min denaturation at 94 °C, 30 s of primer annealing, beginning with 69 °C and decreasing 0.5 °C per cycle, and 3 min at 72 °C for primer extension, followed by 10 more cycles of 1 min denaturation at 94 °C, 30 s of primer annealing at 63 °C and 3 min at 72 °C for primer extension. For optimization of PCR protocol, a diverse set of reference bacterial and archaeal strains were used as controls. The amplification reactions were done in an MJ DNA thermal cycler (MJ Research, USA).

To overcome known biases introduced by PCR amplification, some procedures reported in previous studies were adopted. The number of PCR cycles was limited to 20 to avoid a reduction in the microbial diversity represented in the libraries (Bonnet et al. 2002; Polz and Cavanaugh 1998) and to minimize alterations in community composition caused by "bias 1:1" (Suzuki and Giovannoni 1996) which can result in higher equitability in the final amplified community. In addition, the DNA template concentration was increased to 100 ng per 25 μL reaction volume in order to overcome stochastic variations (Polz and Cavanaugh, 1998), and the products of five PCR amplification replicates were pooled prior to cloning in order to minimize PCR drift (Polz and Cavanaugh 1998; Wagner et al. 1994).

Cloning and sequencing of 16S rRNA gene PCR products

A 16S rRNA gene library was constructed for each of the sediment samples (from site 4 and site 9). Five PCR replicates were pooled, purified on Sephacryl S200-HR columns (Amersham Biosciences), concentrated and quantified in 1% agarose gels. Bands with the expected molecular weight were excised from the

gel using a concert nucleic acid purification system (Life Technologies) and cloned into the plasmid vector pGEM-T (pGEM-T Easy Vector System, Promega), according to the manufacturer's recommendations, and transformed into DH-10b *Escherichia coli* cells, resulting in two libraries with approximately 400 clones each. Clones containing correct-length inserts were checked by PCR amplification using M13 forward and reverse primers. 16S rRNA gene sequences were determined from plasmid DNA prepared according to a standard mini-prep protocol (Sugar Cane EST Genome Project, http://www.fapesp.br), by using the M13 forward and reverse primers and a BigDye terminator cycle sequencing kit (Applied Biosystems) in an ABI 3100 sequencer (Applied Biosystems), according to the manufacturer's recommendations. Since pGEM-T does not allow for directional cloning, both ends of the inserts were sequenced to guarantee that the 5′-end of the 16S rRNA gene was sequenced.

Sequence trimming and analysis

The 16S rRNA gene sequences were checked for chimeras by using the CHIMERA_CHECK program of the Ribosomal Database Project (RDP) (Maidak et al. 2001) and sequences considered as chimeras were excluded from subsequent analyses. The remaining sequences were trimmed for quality using a trimming procedure (Telles and Silva, 2001) modified to process rRNA genes: plasmid vector sequences and regions with low sequence quality (average quality below 20, according to Phred scores (Ewing et al. 1998)) at the 3′- and 5′-ends of the 16S rRNA gene inserts were removed. Thus, only high quality sequences were used in the analyses. Sequences with less than 500 bp were excluded from subsequent analyses.

The sequences were aligned with reference taxa and available environmental clones within the genetic database GenBank (Benson et al. 2002) using the CLUSTAL X program (Thompson et al. 1994). The regions with ambiguous alignment were removed from the subsequent analyses. The alignable nucleotides used in the phylogenetic analysis corresponded to the 5′-terminal region of the 16S rRNA gene, from positions 102 to 562, relative to a corresponding *Escherichia coli* 16S rRNA gene, encompassing the variable regions V2 and most of V3. Evolutionary distances were derived from sequence-pair dissimilarities calculated using the Kimura 2-parameters algorithm, as implemented in the PAUP program, version 4 beta 10 (Swofford 2000). The reconstruction of tree topologies was done by using the neighbour-joining (NJ) algorithm and the consensus tree was calculated after bootstrapping (10^5 replicates).

Statistical analyses

The genetic diversity of the actinobacterial communities recovered in libraries constructed from sediment samples from sites 4 (Bac4) and 9 (Bac9) was determined by calculating the mean taxonomic distance between all pairs of sequences within each community as a diversity index (Rao, 1980), by using the equation: $D\text{mean} = \Sigma\ d(i,j) / S(S + 1)/2$, where d(i,j) is the taxonomic distance between the ith and the jth sequence and S is the total number of sequences in the community. The statistical significance of the difference between the *D*mean calculated for the two communities was obtained by randomization analysis (Manly, 1997), in which sequences were reassigned at random to libraries of the same size. This assemblage was the expected under the null hypothesis that both libraries were samples of the same community. This permutation process was repeated 4000 times and, for each randomization, *D*mean was calculated for both libraries, as well as for the difference between the two values obtained ($\Delta D\text{mean} = D\text{mean}_{Bac9} - D\text{mean}_{Bac4}$). The *P* value was determined by ranking the observed ΔDmean amongst the values obtained for the randomized communities.

The degree of differentiation between the actinobacterial communities recovered in libraries Bac4 and Bac9 was assessed by calculating a modified F_{ST} (Martin 2002; Wright 1978), which compares the genetic diversity within each community to the total genetic diversity of both communities combined, using the equation: $F_{ST} = (\theta_T - \theta_W) / \theta_T$, where θ_T is the genetic diversity (*D*mean) for both communities combined and θ_W is the genetic diversity (*D*mean) within each community averaged over both communities being compared. Statistical significance for the real F_{ST} was obtained by randomization analysis, as described above. All randomization analyses were done by using the program Resampling Stats 2.0 (Resampling Stats Inc., http://www.resample.com/), Microsoft Excel version.

The comparison between the estimated coverage of the source community obtained in the two libraries was calculated by using the LIBSHUFF program (Singleton et al. 2001). Taxonomic distances were

obtained by using the Jukes-Cantor algorithm as implemented in DNADIST (PHYLIP package, J. Felsenstein, University of Washington, Seattle).

Nucleotide sequence accession numbers.

The nucleotide sequences determined in this study have been deposited in the Genbank under the accession numbers AY307856 to AY307896 and AY310345.

Results

Analyses of 16S rRNA gene libraries

In total, 160 clones were randomly picked from each of the two libraries constructed: the 16S rRNA gene library of sediment sample from site 4 (Bac4) and the one from site 9 (Bac9). These were submitted to preliminary analyses where they were checked for chimeras and that they were bacterial sequences by analysis in Blastn. From these, 4 eukaryotic sequences and 15 chimeras were detected in library Bac4 and 6 eukaryotic sequences and 23 chimeras, in Bac9. These results showed that the PCR protocol used had a high specificity for the domain Bacteria. The chimeras and the eukaryotic sequences, as well as the sequences with poor sequencing quality (see Materials and Methods) were excluded from further statistical and phylogenetic analyses.

Phylogenetic analysis of clones from the two bacterial 16S rRNA gene libraries revealed representatives of various phylogenetic divisions. Of 208 bacterial clones analysed, 79 were from library Bac4 and 129 from the library Bac9. Library Bac9 recovered sequences from 11 phylogenetic divisions, of which *Chloroflexi* was the most abundant, representing 23.3% of the clones whereas the sub-divisions alpha-, delta- and gamma-*Proteobacteria* and the *Bacteroides* were represented each by 12.4, 16.5, 15.5 and 15.5% of the total clones in the library, respectivery. The class *Actinobacteria* accounted for 9.3% of the total number of clones. The other five phylogenetic divisions (clostridia, *Spirochaetales*, *Firmicutes* and the Candidate Divisions TM7 and WS3) accounted each for 0.8–3.1% of the total number of clones. Library Bac4 recovered sequences representative of nine phylogenetic divisions, with a high predominance of *Actinobacteria*, which accounted for 39% of all clones. The other more abundant groups were represented by clostridia, *Bacteroides*, *Chloroflexi* and the Candidate Division TM7 sequences, corresponding to 18, 14, 13 and 11% of the clones, respectively. The minor divisions (*Spirochaetales*, *Nitrospirae* and the sub-divisions alpha- and beta-*Proteobacteria*) each represented 1% of the clones.

Phylogenetic analysis showed that, among the various divisions, only the representatives of *Actinobacteria* displayed a tree topology in which the sequences from Bac4 were completely separated from the sequences from Bac9 (Figure 2). This tree topology and the high abundance of *Actinobacteria* in the library Bac4 prompted for separate and detailed analyses of representatives of this group. Thus, phylogenetic and statistical analyses were applied to the actinobacterial data in order to better characterise the observed distribution pattern.

All actinobacterial sequences from library Bac9 (clone names beginning with ‘B’ in Figure 2, corresponding to accession numbers AY307856 to AY307865, and AY310345) were related to uncultured *Actinobacteria* and to unidentified bacteria grouped with actinobacterial taxa recovered from various environments in different studies. The clusters were supported by high bootstrap values. The sequence similarity values between the Bac9 clones and related database sequences ranged from 89% to 95%.

A very different situation was seen for Bac4 clones (clone names beginning with ‘A’ in Figure 2, corresponding to accession numbers AY307866 to AY307896). These were mainly related to sequences from cultivated *Actinobacteria* recovered from a wide range of isolates, including representatives of five *Actinobacteria* taxa: the families *Cellulomonadaceae, Intrasporangiaceae, Microbacteriaceae* and *Propionibacteriaceae* and the suborder *Corynebacterineae*. Only two clones from library Bac4 (AR-B12 and AR-G06), were related to a group comprised exclusively of uncultured *Actinobacteria*. The overall sequence similarity between the clones mentioned above and these uncultured *Actinobacteria* was 91–92%.

Statistical analysis

To address and compare the genetic coverage obtained in the libraries, 16S rRNA gene sequence data were analysed using the LIBSHUFF program (Singleton et al. 2001). These analyses showed statistically significant differences between the genetic

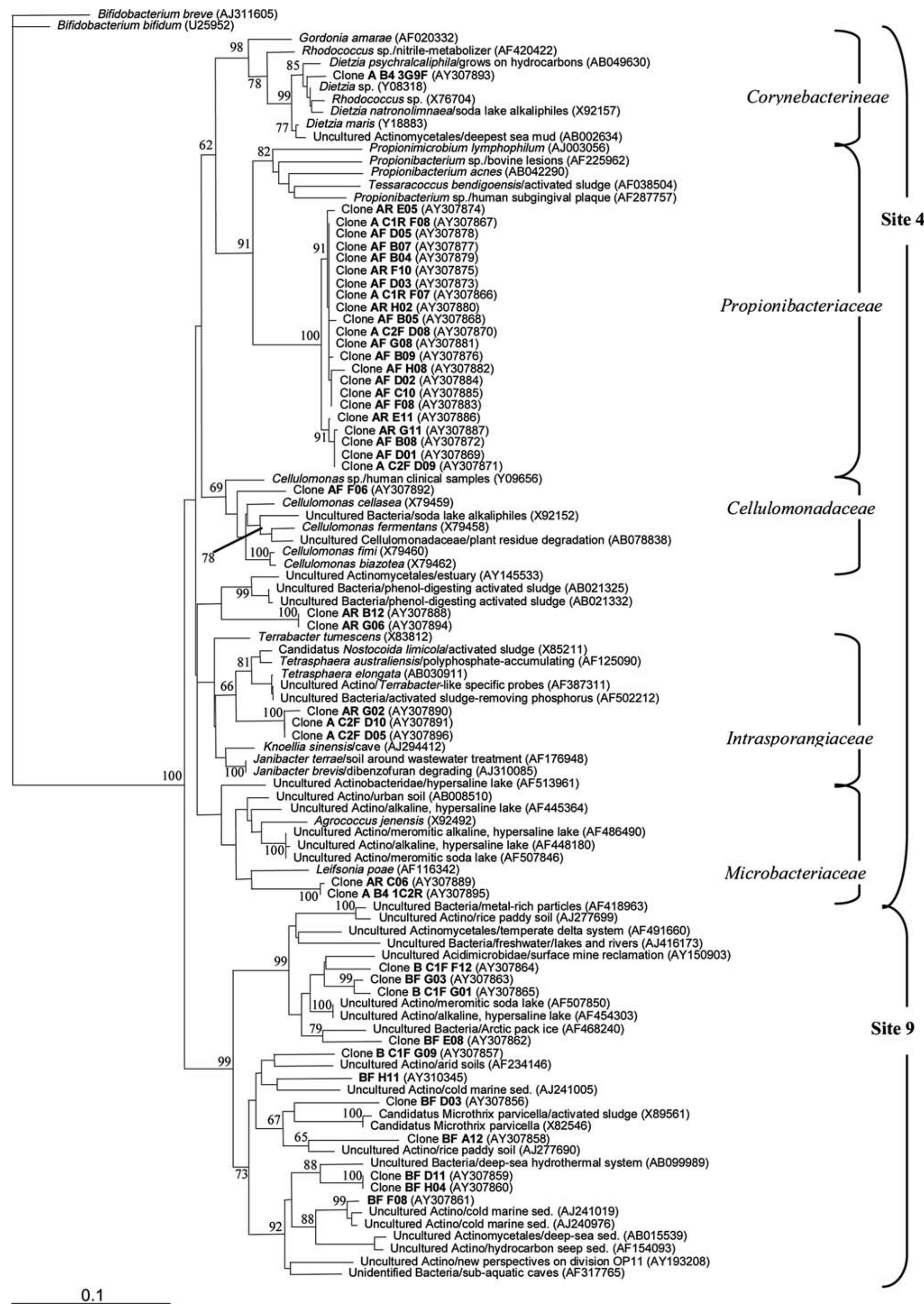

Figure 2. Neighbor-joining tree showing the relationships between the actinobacterial sequences recovered from estuarine sediment samples and reference sequences. Phylogenetic reconstruction was based on 460 alignable bases of 16S rRNA gene sequences. Clones beginning with "A" are from library Bac4 and those with "B" from library Bac9. The GenBank accession numbers are indicated in parentheses. 'Actino' means *Actinobacteria* and 'sed.' means sediments. Sequences with accession numbers AB062696, AF547209 and X77789 were used as out-groups. The bootstrap values from 10^5 randomizations are indicated near the branch nodes.

coverage obtained in libraries Bac4 and Bac9, with low *P* values ($P = 0.01$) for both analyses (Figures. 3A and 3B). The distribution of $(C_X–C_{XY})^2$ with D showed that the values calculated for the real data exceeded the values calculated by using the randomized data (at a *P* value of 0.05), at taxonomic distances lower than 0.22. These results indicated that libraries Bac4 and Bac9 differed greatly at the level of closely related sequences and that both libraries had high coverage estimates at taxonomic distances correspondent to deeply branched taxa.

To statistically evaluate the actinobacterial diversity present in the libraries Bac4 and Bac9, the *D*mean index was calculated for both communities, providing a measure of the mean taxonomic distance within each community. This index is a good measure of diversity since it takes into account the richness and the equitability present within the communities, independent of the operational taxonomic unit (OTU) used, as well as the distance between each pair of sequences, thus reflecting the three important dimensions of diversity (Watve and Gangal 1996). The actinobacterial diversity in library Bac4 (*D*mean = 0.085) was lower than that in library Bac9 (*D*mean = 0.145). Only 52 out of 4000 randomizations had a Δ*D*mean value equal to or lower than the observed value ($P = 0.013$), thus indicating that the actinobacterial diversity was higher in Bac9.

To measure the degree of differentiation between the actinobacterial communities present in the libraries, we calculated a modified F_{ST} index (Martin 2002) for these communities. Although originally devised to measure the reduction in heterozygosity of sub-populations resulting solely from population subdivision and random genetic drift (Wright 1978), according to Martin (Martin 2002), this index can also be used to assess the degree of differentiation between microbial communities. If the diversity within communities is almost the same of that of the two communities combined, the F_{ST} will be very close to zero (see equation in Materials and Methods). The F_{ST} calculated from the original actinobacterial data was 0.159, and was significantly higher than that expected to occur by chance ($P = 0.034$). Hence, the level of diversity within each community was different from the level of diversity from the two communities combined.

Discussion

In microbial ecology studies, the use of molecular methods has greatly increased the possibility of recognising spatial and temporal patterns in natural environments, and comparing them. Although DNA-based methods applied to studies of microbial diversity have some limitations, mainly associated with the efficiency of DNA extraction and with PCR amplification biases (Bonnet et al. 2002; Farreley et al. 1995; Polz and Cavanaugh 1998; Suzuki and Giovannoni 1996; Wintzingerode et al. 1997), these techniques provide a less biased picture of community composition compared to currently available cultivation-based procedures and thus give better estimates of the most abundant organisms in environmental samples (Amann et al. 1995; Head et al. 1998; Hugenholtz et al. 1998). Moreover, it is worth noting that when relative comparisons are established, the problems associated with biases can be avoided, as long as the biases affect both samples (Hughes et al. 2001).

The estimated coverage values obtained by using the LIBSHUFF program (Singleton et al. 2001) showed that libraries Bac4 and Bac9 presented high coverage of the source communities (from sites 4 and 9, respectively) at the level of sequence similarity corresponding to deeply branched taxa, suggesting that the sample sizes analysed were sufficient to recover the most abundant deep taxa. At the low values of taxonomic distance (from $0 \leq D \leq 0.05$), including the one suggested to represent prokaryotic species (97% overall sequence similarity (Stackebrandt and Goebel 1994)), the two libraries presented very different coverage estimates. Library Bac9 showed a very limited coverage at these D values (Figure 3A), as expected for highly diversified communities, such as the ones found in soils and sediments. Similar results were obtained in studies of bacterial diversity from soil samples, which, similarly to the present report, analysed partial sequences of approximately 100 clones per library (Bowman and McCuaig 2003; Singleton et al. 2001). On the contrary, despite the small number of clones analysed, library Bac4 revealed high coverage estimates of the source community at the same D values (Figure 3B). This result suggests that the sediment sample from site 4 may harbour lower bacterial diversity, such as those found in bioreactors. Indeed, the estimated coverage values obtained for library Bac4 were similar to the ones obtained for an activated sludge by LIBSHUFF analyses using a similar number of partial

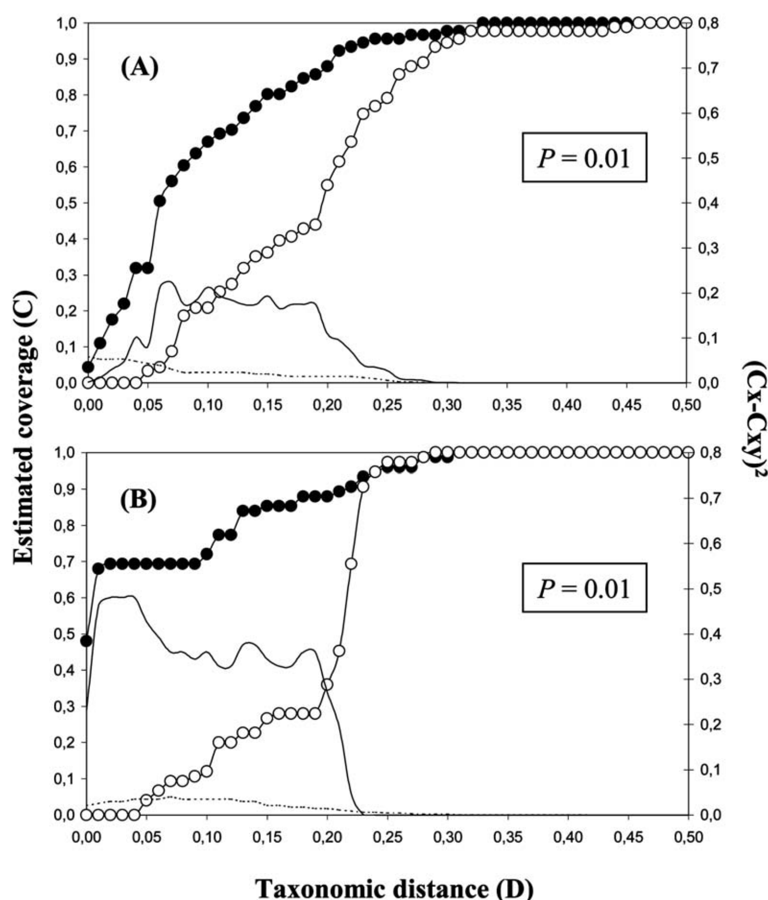

Figure 3. Genetic coverage comparison between libraries Bac4 and Bac9. Filled and opened circles represent the homologous and heterologous coverage curves, respectively. Solid lines indicate the value of $(C_X - C_{XY})^2$ for the original samples and broken lines show $P = 0.05$ value of $(C_X - C_{XY})^2$ for the randomized samples. (A) Bac9 as homologous coverage and Bac4 as heterologous coverage, and (B) Bac4 as homologous coverage and Bac9 as heterologous coverage. Taxonomic distances (D) were calculated using the Jukes-Cantor correction.

16S rRNA gene sequences (Bond et al. 1995; Singleton et al. 2001).

In this study, we recognized a spatial pattern of actinobacterial distribution, detected by the analyses of 16S rRNA gene libraries constructed from estuarine sediment samples. The phylogenetic reconstruction of actinobacterial sequences from the two libraries revealed a clear difference between the recovered communities, since the sequences from the library Bac9 formed a distinct clade, not related to the one formed by sequences from Bac4 (Figure 2). Statistical analysis using F_{ST} index also indicated differentiation between the two actinobacterial communities. If the F_{ST} interpretation used to evaluate population structure (Wright 1978) is also valid for communities (Martin 2002), the F_{ST} value of 0.159 for the observed data and the *P* value lower than 0.05 provide statistical support for the conclusions that the actinobacterial communities recovered in libraries Bac4 and Bac9 were highly structured and may not have originated from the same community.

Despite the higher number of actinobacterial sequences recovered in the library Bac4, diversity analyses showed that the actinobacterial community recovered in this library was less diverse than that recovered in the library constructed from the uncontaminated sediment (Bac9). This result, as well as the indication of low bacterial diversity suggested by the high coverage estimates obtained for library Bac4 (Figure 3B), agrees with previous studies, which reported that highly impacted environments had lower microbial diversity than less impacted or non-impacted ones (Atlas et al. 1991; Cheung and Kinkle 2001).

Until recently, *Actinobacteria* encountered in marine environments were considered to be the result of dispersion from adjacent terrestrial habitats. Nonetheless, recent studies have reported a widespread and persistent occurrence of indigenous actinobacterial populations in marine sediments (Colquhoun et al. 1998; Jensen et al. 1991; Mincer et al. 2002; Moran et al. 1995; Stach et al. 2003; Urakawa et al. 1999), revealing their presence as metabolically active members of microbial communities (Moran et al. 1995) and physiological adaptations for growth in seawater (Colquhoun et al. 1998; Jensen et al. 1991; Mincer et al. 2002). Although none of these studies were conducted in estuarine sediments, the evidence of metabolic activity and physiological adaptations to seawater support the hypothesis that the sequences recovered in libraries Bac4 and Bac9 could be part of the indigenous actinobacterial communities present in the original samples, and not derived from adjacent environments.

The sequences recovered in the library constructed from the uncontaminated sediment sample (Bac9) were related to sequences retrieved from a wide range of habitats, including arid soil (AF234146), paddy soil (AJ277690 and AJ277699), marine sediments (AJ241005, AJ241019, 240976 and AB015539), an alkaline hypersaline lake (AF454303), freshwater (AJ416173), hydrocarbon seep sediment (AF154093), a temperate estuarine delta system (AF491660), metal-rich particles (AF418963), Arctic pack ice (AF468240), deep-sea hydrothermal systems (AB099989), and a sludge system (X89561). Thus, the organisms represented by these sequences may be widely distributed and may survive under very different environmental conditions.

In library Bac4, constructed from the contaminated sediment, the sequences were related to a considerable range of known cultivated actinobacterial taxa, although some clades were not highly supported by bootstrap analysis. Despite the high level of hydrocarbon contamination at site 4, none of the sequences recovered in the Bac4 were from actinobacterial taxa known to be xenobiotic degraders or closely related to these, such as some *Arthrobacter, Gordonia, Mycobacterium, Nocardioides, Rhodococcus* and *Terrabacter* species (Bell et al. 1998; Cheung and Kinkle 2001; Grifoll et al. 1992; Kanaly and Harayama 2000; Kästner et al. 1994; Kelley and Cerniglia 1995; Lang 1996). Only four clones were moderately related to some members of these genera: (i) clones AR-G02, A-C2F-D10 and A-C2F-D05, which shared 92, 94 and 94% overall similarity, respectively, to *Terrabacter tumescens* (X83812), and (ii) clone A-B4-3G9, with 96% overall similarity to a non-characterized *Rhodococcus* species (X76704,) and 93% similarity to a nitrile-metabolizing *Rhodococcus* sp. (AF420422). The latter clone was also related (95% sequence similarity) to *Dietzia psychroalkaliphila* (AB049630), which is able to grow on hydrocarbons (Yumoto et al. 2002).

The absence of sequences related to known xenobiotic–degrading *Actinobacteria* was not unexpected since the latter have been isolated mainly from selective experiments *ex situ*. Classical enrichment procedures and selective isolation conditions usually favour strains with high growth rates under specific nutrient and incubation parameters, which are often not representative of conditions *in situ*. Thus, cultured microorganisms account for only a very small percentage of the microbial diversity in the environment, and do not necessarily represent the most abundant and/or ecologically important fraction of the microbial community (Felske et al. 1997; Miskin et al. 1999; Nogales et al. 1999).

Actinobacterial communities are a good example of a very diverse and ecologically important but still poorly understood bacterial group. Although *Actinobacteria* play important roles in the degradation of xenobiotic compounds, including hydrocarbons and chlorinated-compounds, in cultivation and in microcosms experiments, as well as in contamination controlled-experiments *in situ* (Bell et al. 1998; Greene et al. 2000; Ilori and Amund 2000; Pellizari et al. 1996; Shi et al. 1999; Wagner-Dobler et al. 1998), the representatives of *Actinobacteria* have been reported as minor components of bacterial communities in xenobiotic-contaminated natural environments assessed by molecular methods. In a hydrocarbon- and chlorinated-solvent-contaminated aquifer undergoing intrinsic bioremediation, an assessment of microbial diversity along the main redox zones recovered actinobacterial 16S rRNA gene sequences from soil samples only from methanogenic and iron/sulfate reducing zones, and in very low numbers (Dojka et al. 1998). In another study, comparison between 16S rRNA and 16S rRNA gene libraries constructed from polychlorinated biphenyl (PCB)-polluted soil revealed a good correspondence between these libraries, which showed similar bacterial community compositions in terms of major constituents, including alpha-, beta- and gamma-Proteobacteria (Nogales et al. 2001). Sequences from *Actinobacteria* were re-

covered only in the rRNA library, indicating the occurrence of organisms belonging to this class as an active population in the contaminated site, despite possible molecular biases (Nogales et al. 2001). A study done on a marine nearshore sediment highly contaminated with PAHs, in which 16S rRNA gene amplification and cloning were used, revealed the predominance of γ- and δ-Proteobacteria, and clostridia, whereas sequences assigned to high G+C Gram-positive bacteria were minor constituents of the library (Gray and Herwig 1996).

The only study reporting *Actinobacteria* as major constituents of a microbial community in xenobiotic-contaminated sediments was conducted in the Sagami and Tokyo Bays in Japan (Urakawa et al. 1999). Tokyo Bay receives industrial and domestic waste discharges from a large human population via four main rivers, whereas Sagami Bay is less impacted and has a good water exchange with the Pacific Ocean. Libraries constructed from sediments from both bays revealed γ-Proteobacteria and the class *Actinobacteria* as the major constituents, the latter representing the dominant clones. Nevertheless, although these bays received very different inputs, *Actinobacteria* from both libraries were grouped in the same cluster in the phylogenetic reconstruction, and were related to *Microbacterium arborescens*, *Clavibacter michiganensis* and to an unidentified actinobacterial taxon.

The present study revealed a different pattern of actinobacterial distribution in the two sediment samples analysed, providing valuable information for further investigations on microbial estuarine processes in the Santos-São Vicente estuary. This pattern, supported by statistical analyses, suggests that distinct actinobacterial communities might have adapted to different environmental conditions, resulting in the differently structured communities recovered from the samples analysed. Since the comprehension of ecological processes, including the generation and/or the maintenance of species diversity, depends on efforts to discover patterns within systems and compare them (Levin 1992), the recognition of the spatial pattern of actinobacterial distribution can be considered as the first step towards understanding the way these communities are organized. Investigations of the possible causes and consequences of this pattern will depend on the analysis of molecular and physicochemical data from replicate sediment samples from sites 4 and 9.

Acknowledgments

The authors thank Dr. Valeria Maia de Oliveira (Microbial Resource Division, CPQBA – UNICAMP), Dr. Vivian H. Pellizari and Dr. Rosana Vazoller (Environmental Microbiology Laboratory, Institute of Biological Sciences, USP), and all colleagues in their laboratories for the discussions, criticisms and suggestions during this study. We also thank Dr. Anete Pereira de Souza and Dr. Gonçalo A. G. Pereira (IB-UNICAMP) for providing access to laboratory facilities, and Dr. Marcos Antônio de Oliveira for helping with the 16S rRNA gene libraries. We are also thankful to researchers and technicians from CETESB for providing infra-structure for the sample collections. We also thank Renato Santos Vicentini for the computer expertise in modifying the Perl script for 16S rRNA gene sequence trimming. We are grateful to Dr. Klaus Nüsslein (University of Massachesetts, USA) and to Dr. Jem Stach (University of Kent, UK) for critically reviewing this manuscript. This work was funded by the FAPESP Biota Program (grant 98/05068-0) and F.F.Piza was supported by a FAPESP scholarship (grant 99/10739-4).

References

Amann R.I., Ludwig W. and Schleifer K. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59: 143–169.

Anon. 1991. Determinations of petroleum hydrocarbons in sediments. Reference methods for marine pollution studies no 20. UNEP (United Nations Environment Programme).

Anon. 1999. Canadian Sediment Quality Guidelines for the protection of Aquatic Life. Summary Tables., Environment Canada.

Anon. 2001. Sistema Estuarino Santos-São Vicente. Relatório Técnico. CETESB, São Paulo.

Anon. 1985. NRC (National Research Council). Oil in the sea, inputs, fates and effects. National Academy Press, Washington D.C., USA.

Atlas R.M., Horowitz A., Krichevsky M. and Bej A.K. 1991. Response of microbial populations to environmental disturbance. Microb. Ecol. 22: 249–256.

Bell K.S., Philp J.C., Aw D.W. and Christofi N. 1998. The genus *Rhodococcus*. J. Appl. Microbiol. 85(2): 195–210.

Benson D.A. et al. 2002. GenBank. Nucleic Acids Res. 30: 17–20.

Bond P.L., Hugenholtz P., Keller J. and Blackall L.L. 1995. Bacterial community structures of phosphate-removing and non-phosphate-removing activated sludges from sequencing batch reactors. Appl. Environ. Microbiol. 61: 1910–1916.

Bonnet R., Suau A., Dore J., Gibson G.R. and Collins M.D., 2002. Differences in rDNA libraries of faecal bacteria derived from 10- and 25-cycle PCRs. Int. J. Syst. Evol. Microbiol. 52(Pt3): 757–763.

Bowman J.P. and McCuaig R.D. 2003. Biodiversity, community structural shifts, and biogeography of prokaryotes within antarctic continental shelf sediments. Appl. Environ. Microbiol. 69: 2463–2483.

Cheung P.Y. and Kinkle B.K., 2001. Mycobacterium diversity and pyrene mineralization in petroleum-contaminated soils. Appl. Environ. Microbiol. 67(5): 2222–2229.

Colquhoun J.A. et al. 1998. Taxonomy and biotransformation activities of some deep-sea actinomycetes. Extremophiles 2(3): 269–277.

Dojka M.A., Hugenholtz P., Haack S.K. and Pace N.R., 1998. Microbial diversity in a hydrocarbon- and chlorinated-solvent-contaminated aquifer undergoing intrinsic bioremediation. Appl. Environ. Microbiol. 64(10): 3869–3877.

Ewing B., Hillier L., Wendl M. and Green P. 1998. Basecalling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8: 175–185.

Farreley V., Rainey F.A. and Stackebrandt E., 1995. Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. Appl. Environ. Microbiol. 61: 2798–2801.

Farrington J. and Tripp R.W., 1975. Hydrocarbons in cores of northwestern Atlantic coastal and continental margin sediments. Estuar. Coast. Mar. Sci. 5: 793–808.

Felske A., Rheims H., Wolterink A., Stackebrandt E. and Akkermans A.D.L. 1997. Ribosome analysis reveals prominent activity of an uncultured member of the class Actinobacteria in grassland soils. Microbiology 143: 2983–2989.

Fulthorpe R.R., Rhodes A.N. and Tiedje J.M. 1998. High levels of endemicity of 3-chlorobenzoate-degrading soil bacteria. Appl. Environ. Microbiol. 64: 1620–1627.

Gray J.P. and Herwig R.P. 1996. Phylogenetic analysis of the bacterial communities in marine sediments. Appl. Environ. Microbiol. 62(11): 4049–4059.

Greene E.A., Kay J.G., Jaber K., Stehmeier L.G. and Voordouw G. 2000. Composition of soil microbial communities enriched on a mixture of aromatic hydrocarbons. Appl. Environ. Microbiol. 66(12): 5282–5289.

Grifoll M., Casellas M., Bayona J.M. and Solanas A.M. 1992. Isolation and characterization of a fluorene-degrading bacterium: identification of ring oxidation and ring fission products. Appl. Environ. Microbiol. 58(9): 2910–2917.

Head I.M., Saunders J.R. and Pickup R.W. 1998. Microbial Evolution, Diversity, and Ecology: A Decade of Ribosomal RNA Analysis of Uncultivated Microorganisms. Microb. Ecol. 35(1): 1–21.

Heuer H., Krsek M., Baker P., Smalla K. and Wellington E.M. 1997. Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. Appl. Environ. Microbiol. 63(8): 3233–3241.

Hugenholtz P., Goebel B.M. and Pace N.R., 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. 180(18): 4765–4774.

Hughes J.B., Hellmann J.J., Ricketts T.H. and Bohannan B.J.M. 2001. Counting the uncountable: statistical approaches to estimating microbial diversity. Appl. Environ. Microbiol. 67: 4399–4406.

Ilori M.O.N. and Amund D.- I. 2000. Degradation of anthracene by bacteria isolated from oil polluted tropical soils. Z. Naturforsch. 55: 890–897.

Jensen P.R., Dwight R. and Fenical W. 1991. Distribution of actinomycetes in near-shore tropical marine sediments. Appl. Environ. Microbiol. 57(4): 1102–1108.

Kanaly R.A. and Harayama S. 2000. Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria. J. Bacteriol. 182: 2059–2067.

Kästner M., Breuer-Jammali M. and Mahro B. 1994. Enumeration and characterization of the soil microflora from hydrocarbon-contaminated soil sites able to mineralize polycyclic aromatic hydrocarbons (PHA). Appl. Microbiol. Biotechnol. 41: 267–273.

Kelley I. and Cerniglia C.E., 1995. Degradation of a mixture of high-molecular-weight polycyclic aromatic hydrocarbons by a *Mycobacterium* strain PYR-1. J. Soil Contam. 74: 77–91.

Lane D. 1991. 16S/23S rRNA sequencing.. In: Stackebrandt E. and Goodfellow M. (eds), Nucleic Acid Techniques in Bacterial Systematics. West Sussex, England, pp. 115-175.

Lang E. 1996. Diversity of Bacterial capabilities in utilizing alkylated benzenes and other aromatic compounds. Lett. Appl. Microbiol. 23: 257–260.

Levin S.A. 1992. The problem of pattern and scale in ecology. Ecology 73: 1943–1967.

Lindstrom J.E., Barry R.P. and Braddock J.F. 1999. Long-term effect on microbial communities after a subartic oil spill. Soil Biol. Biochem. 31: 1677–1689.

Maidak B.L. et al. 2001. The RDP-II (Ribosomal Database Project). Nucleic Acids Res 29(1): 173–174.

Manly B.F.J. 1997. Randomization. In Randomization, Bootstrap and Monte Carlo Methods in Biology. Chapman and Hall, Glasglow, UK, pp. 1–23.

Martin A.P. 2002. Phylogenetic approaches for describing and comparing the diversity of microbial communities. Appl. Environ. Microbiol. 68: 3673–3682.

Mincer T.J., Jensen P.R., Kauffman C.A. and Fenical W. 2002. Widespread and persistent populations of a major new marine actinomycete taxon in ocean sediments. Appl. Environ. Microbiol. 68(10): 5005–5011.

Miskin I.P., Farrimond P. and Head I.M. 1999. Identification of novel bacterial lineages as active members of microbial populations in a freshwater sediment using a rapid RNA extraction procedure and RT-PCR. Microbiology 145 (Pt8): 1977–1987.

Moran M.A., Rutherford L.T. and Hodson R.E. 1995. Evidence for indigenous *Streptomyces* populations in a marine environment determined with a 16S rRNA probe. Appl. Environ. Microbiol. 61(10): 3695–3700.

Nogales B., Moore E.R., Abraham W.R. and Timmis K.N., 1999. Identification of the metabolically active members of a bacterial community in a polychlorinated biphenyl-polluted moorland soil. Environ. Microbiol. 1(3): 199–212.

Nogales B. et al. 2001. Combined use of 16S ribosomal DNA and 16S rRNA to study the bacterial community of polychlorinated biphenyl-polluted soil. Appl. Environ. Microbiol. 67(4): 1874–1884.

Pellizari V.H., Bezborodnikov S., Quensen J.F. 3rd and Tiedje J.M. 1996. Evaluation of strains isolated by growth on naphthalene and biphenyl for hybridization of genes to dioxygenase probes and polychlorinated biphenyl-degrading ability. Appl. Environ. Microbiol. 62(6): 2053–2058.

Polz M.F. and Cavanaugh C.M. 1998. Bias in template-to-product ratios in multitemplate PCR. Appl. Environ. Microbiol. 64(10): 3724–3730.

Rao C.R. 1980. Diversity and dissimilarity coefficients: a unified approach. Theor. Pop. Biol. 21: 24–43.

Rodrigues D.F. 2002. Caracterização polifásica da biodiversidade de isolados degradadores de poluentes xenobióticos na baixada santista. MSc Thesis Thesis, University of São Paulo (USP), São Paulo.

Schwarz K., Hansen-Hagge T. and Batram C. 1990. Improved yields of long PCR products using gene 32 protein. Nucleic Acids Res. 18: 1079.

Shi S. et al. 1999. Molecular analysis of microbial community structures in pristine and contaminated aquifers: field and laboratory microcosms experiments. Appl. Environ. Microbiol. 65: 2143–2150.

Singleton D.R., Furlong M.A., Rathbun S.L. and Whitman W.B. 2001. Quantitative comparisons of 16S rRNA gene sequence libraries form environmental samples. Appl. Environ. Microbiol. 67: 4374–4376.

Stach J.E.M. et al. 2003. Statistical approaches for estimating actinobacterial diversity in marine sediments. Appl. Environ. Microbiol. 69: 6189–6200.

Stackebrandt E. and Goebel B.M. 1994. Taxonomic note: a place for DNA:DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int. J. System. Bacteriol. 44: 846–849.

Suzuki M.T. and Giovannoni S.J. 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 62: 625–630.

Swofford D.L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sunderland, Massachusetts, USA.

Telles G.P. and Silva F.R, 2001. Trimming and clustering sugarcane ESTs. Gen. Mol. Biol. 24: 17–23.

Thompson J.D., Higgins D.J. and Gibson T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673–4680.

Torsvik V., Goksoyr J. and Daae F.L. 1990. High diversity in DNA of soil bacteria. Appl. Environ. Microbiol. 56: 782–787.

Urakawa H., Kita-Tsukamoto K. and Ohwada K. 1999. Microbial diversity in marine sediments from Sagami Bay and Tokyo Bay, Japan, as determined by 16S rRNA gene analysis. Microbiology 145: 3305–3015.

van Elsas J.D. and Smalla K. 1995. Extraction of microbial community DNA from soils.. In: Akkerman A.D., van Elsas J.D. and Bruijn F.J. (eds), Molecular Microbial Ecology Manual. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 1-11.

Volkman J.K., Holdsworth G.D., Neill G.P. and Bavor H.J.J. 1992. Identification of natural, anthropogene and petroleum hydrocarbons in aquatic sediments. Sci. Total Environ. 112: 203–219.

Wagner A. et al. 1994. Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Syst. Biol. 43: 250–261.

Wagner-Dobler I. et al. 1998. Microcosm enrichment of biphenyl-degrading microbial communities from soils and sediments. Appl. Environ. Microbiol. 64(8): 3014–3022.

Watve M.G. and Gangal R.M. 1996. Problems in measuring bacterial diversity and a possible solution. Appl. Environ. Microbiol. 62: 4299–4301.

Wintzingerode F.V., Göbel U.B. and Stackebrandt E. 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol. Rev. 21: 213–229.

Wright S. 1978. Genetic variability in natural populations: methods. In Evolution and the genetics of populations. The University of Chicago Press, Chicago, USA, , pp. 79-103.

Yumoto I. et al. 2002. *Dietzia psychralcaliphila* sp. nov., a novel, facultatively psychrophilic alkaliphile that grows on hydrocarbons. Int. J. Syst. Evol. Microbiol. 52(Pt 1): 85–90.