

Diversity of Hydrocarbon-Related Catabolic Genes in Oil Samples from Potiguar Basin (RN, Brazil)

Leandro Costa Lima Verde¹, Tiago Rodrigues e Silva¹, Bruna Martins Dellagnezze¹, Eugênio Vaz dos Santos Neto² and Valéria Maia de Oliveira^{1*}

¹Research Center for Chemistry, Biology and Agriculture (CPQBA), Campinas University - UNICAMP, CP 6171, CEP 13081-970, Campinas, SP, Brazil

²Petrobras R&D Center, Cidade Universitária, Quadra 7, Rio de Janeiro, RJ, CEP: 21949-900, Brazil

Abstract

Biodegradation may result in physicochemical changes in crude oil and natural gas properties, being responsible for the decrease of saturated hydrocarbons and yielding heavy oil with low economic value. Studies on the diversity of microbial catabolic genes in oil reservoirs are scarce and could help to predict the potential of a petroleum sample to be biodegraded. The aim of this study was to evaluate the diversity of genes involved in hydrocarbon degradation in Brazilian petroleum samples (biodegraded and non-biodegraded) through the construction and analysis of gene libraries (alkane monooxygenase – *alk*, dioxygenase – ARHDs and 6-oxocyclohex-1-ene-1-carbonyl-CoA hydrolase – *bamA*). The results showed a differential distribution of catabolic genes between the sites, being the biodegraded oil more diverse for the *alk* and *bamA* genes. Sequences were similar to the *alkB* genes from *Geobacillus thermoleovorans* and several species of *Acinetobacter*, to ARHD genes from *Pseudomonas* spp. and two species of *Burkholderia*, and to *bamA* genes from deltaproteobacteria. Interestingly, most of the catabolic sequences recovered from both petroleum reservoirs grouped together forming distinct clusters in the phylogenetic tree reconstruction and may correspond to potentially new genes, possibly harbored by yet uncultivated microorganisms. This is the first report on the detection of *alk*, ARHD and *bamA* genes in petroleum reservoir environments, demonstrating the genetic potential of such microbial communities to biodegrade the oil.

Keywords: Petroleum microbiology, Microbial diversity; Oil degradation; Catabolic genes; Biodegradation, Hydrocarbons, Geobacillus, Oil

Introduction

Most of the world's oil is biodegraded and while the effects of biodegradation on the molecular composition and physical properties of crude oil and natural gas are relatively well known empirically, the actual processes that occur during the biodegradation of oil in deep reservoirs (below some hundred meters) remain unclear [1]. During biodegradation, the hydrocarbon content is transformed, with a consequent increase in oil density, sulfur content, acidity and viscosity. All these factors interfere with the extraction and refining operations, resulting in significant economic losses [2]. On the other hand, the mechanisms of oil degradation in petroleum reservoirs, as well as the microorganisms involved are still poorly understood.

Oil is a complex mixture of hydrocarbons such as saturated, unsaturated, linear, monoaromatic and polycyclic hydrocarbons [3,4]. Each of these compounds is biodegraded through different routes of several steps. Many studies have already demonstrated the existence of large and diverse populations of microbes with different metabolic activities in petroleum systems [5,6].

Nonetheless, knowledge of the diversity of catabolic genes involved in the processes of degradation of oil is still scarce. Information on community composition of bacteria and genes involved in the biodegradation process can bring light on the understanding of microbial metabolic pathway preferences, as well as a better application of specific microorganisms in biodegradation or bioremediation processes. Techniques able to identify the genes involved in the catabolic degradation of hydrocarbons are valuable tools for elucidating the structure of the microbial community that is truly functional in the environment.

This study examined the diversity of the functional gene involved in the degradation of hydrocarbons, the alkane hydroxylase (*alk*) gene,

a membrane-bound enzyme that catalyzes the terminal or sub terminal oxidation of n-alkanes (C₅-C₁₆) for primary or secondary alcohols, which is the initial step in the aerobic degradation of n-alkanes [7]. Alkane monooxygenase proteins have had their genetics deciphered in a wide variety of bacteria, including Gram positive and negative such as *Rhodococcus erythropolis* NRRL B-16531, *Rhodococcus* sp. Q15, *Mycobacterium tuberculosis* H37Rv, *Pseudomonas putida* GPO1, *Acinetobacter* sp. ADP1 and *Acinetobacter* sp. M-1 [8-11]. Alkane monooxygenases were found in bacteria that inhabit diverse environments, including sediments in Alaska [12], contaminated soils [13,14], cold ecosystems [15,16], shallow aquifer [17] and soils of the Arctic and Antarctica [18-21].

The functional gene involved in the aerobic degradation of aromatic hydrocarbons, the dioxygenase protein coding gene, was also analyzed in this study. The first step in the degradation of aromatic hydrocarbons usually occurs through the incorporation of molecular oxygen in aromatic nucleus by a multicomponent enzyme system, forming cis-dihydrodiol. In this system, the terminal dioxygenase is composed of one large α subunit and one small β subunit [22]. The α -subunit contains two conserved regions: the Rieske center [Fe₂-S₂] and the

***Corresponding author:** Valéria Maia de Oliveira, Research Center for Chemistry, Biology and Agriculture (CPQBA), Campinas University - UNICAMP, CP 6171, CEP 13081-970, Campinas, SP, Brazil, Tel. : +55-19-2139 2874; E-mail: vmaia@cpqba.unicamp.br

Received November 26, 2012; Accepted March 07, 2013; Published March 15, 2013

Citation: Verde LCL, e Silva TR, Dellagnezze BM, Santos Neto EVd, de Oliveira VM (2013) Diversity of Hydrocarbon-Related Catabolic Genes in Oil Samples from Potiguar Basin (Rn, Brazil). J Pet Environ Biotechnol 4: 138. doi:10.4172/2157-7463.1000138

Copyright: © 2013 Verde LCL, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

catalytic domain containing mononuclear iron [23]. The dioxygenases belong to a large family known as aromatic-ring-hydroxylating dioxygenases (ARHDs) [24]. These genes are located in chromosomal or plasmid DNA and were identified in bacterial strains belonging to α -Proteobacteria (*Sphingomonas*) [25], β -Proteobacteria (*Alcaligenes*, *Burkholderia*, *Comamonas*, *Polaromonas*, *Ralstonia*) [26-28] and γ -Proteobacteria (*Pseudomonas*) [29].

The process of biodegradation in subsurface reservoirs is often anaerobic in nature [2,6]. Most degraded oils contain metabolic markers for the anaerobic degradation, such as specific reduced naftoic acids, indicating that the anaerobic metabolism of oil is the mechanism by which most biodegraded oil in the world were produced [30].

Therefore, research on the anaerobic degradation of hydrocarbons is also necessary for the understanding of the whole process of oil degradation in reservoirs. In this sense, this study also examined *bamA*, the functional gene involved in the degradation route of benzyl-CoA, by which most aromatic substrates (benzene, toluene, ethylbenzene, xylenes and also other aromatic compounds) are anaerobically degraded [31]. *bamA* gene encodes the protein 6-oxociclohex-1-ene-1-carboxyl-CoA hydrolase (6-OCH-CoA) and has been used for the detection of a wide range of anaerobic microorganisms able to degrade hydrocarbons [32]. This functional marker was recently used to detect anaerobic hydrocarbon-degrading micro-organisms in environmental samples [32], as well as to study the microbial community structure of sulfate-reducing enrichment cultures growing on petroleum hydrocarbons, where *Desulfosarcina ovata* was showed to be dominant [33]. *bamA* was also used as a biomarker to investigate the diversity of hydrocarbon-degrading bacteria under iron-reducing conditions in a leachate contaminated aquifer, with *bamA* sequences found to be closely related to the ones from *Geobacter* species and *Georgfuchsia toluolica* [31].

This study aimed at investigating the presence and diversity of catabolic genes involved in the hydrocarbon degradation process in biodegraded and non-biodegraded oil samples from petroleum reservoirs through the construction and analyses of *alk*, ARHD and *bamA* gene libraries. The results allowed us to evaluate the effectiveness of the primer sets used as a tool to analyze the community structure of bacteria that have the genetic potential to degrade hydrocarbons by aerobic and anaerobic pathways.

Materials and Methods

Sampling

Oil samples (naturally mixed with formation water) were collected

in July 2008 from two production wells, named GMR75 and PTS1, at the onshore Potiguar Basin (Northeast, Brazil), with logistic support from CENPES/Petrobras. Sampling details as well as oil characteristics and geological settings are given in Silva et al. [34].

Total DNA extraction and PCR amplification *alk*, ARHD and *bamA* genes

DNA extraction from oil/formation water samples was carried out using the PowerSoil™ DNA Isolation Kit (MoBio Laboratories, California), according to the manufacturer's instructions. Five tubes from the kit were used to start with the DNA extraction. At the end of the procedure, DNA samples obtained from the five tubes were pooled, concentrated and used for subsequent PCR amplification. Crude oil/formation water samples were not filtered or centrifuged before DNA extraction. The PCR amplifications of the three functional genes were performed independently using the respective set of primers (Table 1) [35]. Twenty five microliter reaction mixtures contained 5 μ l of total DNA, 1 U *Platinum Taq* DNA polymerase (Invitrogen), 0.2 mM dNTP mix (GE Healthcare), 1.2 μ M of each primer, 1X *Taq* buffer and 1.5 mM MgCl₂. PCR amplifications were done using an Eppendorf Mastercycler Gradient (Eppendorf Scientific, New York, USA) with the programs shown in table 2. The amplification products were checked by electrophoresis in 1.2% (wt vol⁻¹) Agarose gels. Four replicate PCR reactions were done for each sample with each pair of primers. The corresponding PCR replicates were then pooled and concentrated in a speed vacuum concentrator 5301 Eppendorf, A-2-VC rotor.

Clone library construction and sequencing

The PCR products amplified with functional gene-targeted primer pairs were cloned using the pGEM-T Easy Vector (Promega), according to manufacturer's guidelines, and transformed into *Escherichia coli* JM 109 cells. The functional gene-targeted inserts were amplified from plasmid DNA of selected clones using the universal the M13 forward (5'-CGCCAG GGT TTT CCC AGT CAC GAC-3') and reverse (5'-TTT CAC ACA GGA AAC AGC TAT GAC-3') primers. PCR was performed in a 50 μ L-reaction volume, containing 1 μ L of an overnight clone culture, 0.4 μ M each primer, 0.2 mM dNTP mix, 2 U *Taq* DNA polymerase (Invitrogen), 1X *Taq* buffer, and 1.5 mM MgCl₂. The amplification program consisted of an initial denaturation step at 94°C for 3 min, followed by 30 cycles of 94°C/20 s, 60°C/20 s, and 72°C/90 s. PCR products were purified as previously described for automated sequencing in the MegaBace DNA Analysis System 500 (GE Healthcare) and ABI Prism 377 DNA Sequencer™ (Applied Biosystems). The sequencing was carried out with the corresponding primers used for PCR amplification. The sequencing programs consisted of 30 cycles of

Primer	Gene target	Sequence (5'→3')	Direction	Expected Amplicon size (bp)	Reference
alkF	<i>alk</i>	GCICAIGARITIRKICAYAA	forward	524	[19]
alkR	<i>alk</i>	GCITGITGTCISWRTGICGYTG	reverse		
ARHdf	α -ARHD	TTYRYITGYAIITAYCAYGGITGGG	forward	329	[35]
ARHDr	α -ARHD	AAITKYTCIGCIGSIRMYTTCCA	reverse		
SP9f	<i>bamA</i>	CAGTACAAATTCTACACVACBG	forward	300	[32]
ASP1r	<i>bamA</i>	CMATGCCGATYTCCTGRC	reverse		

Table 1: Oligonucleotide primers used for PCR assays.

Primer set	Initial denaturation	Cycle number	Denaturation	Annealing	Extension (72°C)	Final extension (72°C)
alkF/alkR	97°C, 3 min	35	94°C, 1 min	48°C, 1 min	1 min	5 min
ARHdf/ARHDr	97°C, 3 min	35	94°C, 1 min	58°C, 1 min	1 min	5 min
SP9f/ASP1r	94°C, 10 min	35	94°C, 30 s	59°C, 45 s	1 min	10 min

Table 2: Polymerase chain reaction programs used in this study.

94°C/20 s, 43°C/20 s, and 72°C/90 s for *alk* gene, 30 cycles of 94°C/20 s, 53°C/20 s, and 72°C/90 s for ARHD gene and 30 cycles of 94°C/20 s, 54°C/20 s, and 72°C/90 s for *bamA* gene.

Phylogenetic analysis

The partial sequences from GMR75 and PTS01 clones were assembled in a *contig* (consensus sequence) using the phred/Phrap/CONSED program [36]. Nucleotide sequences were translated into protein sequences using the translate tool in the ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (<http://us.expasy.org/tools/dna.html>). Deduced protein sequences from each library were grouped into Operational Protein Families (OPFs) using MOTHUR version 1.14.0 [37]. The cutoff value was selected after initial examination of phylogenetic trees that included all protein sequences generated by neighbor-joining analysis [38] using MEGA5 software [39], with 1,000 bootstrap replicates. Deduced protein sequences were aligned using ClustalX, version 1.83 [40], with related reference sequences identified from BLASTp searches, in addition to other sequences of several representative *alkB*, ARHD and *bamA* sequences identified from the literature.

Diversity indices and statistics

Diversity index calculations (α -diversity measures) were performed individually for each catabolic gene library with the program MOTHUR [37]. Rarefaction curves were calculated using the OPF cutoff values established based on MOTHUR analyses in combination with phylogenetic clustering of the catabolic sequences. In addition, the Shannon (H') diversity index, Simpson index [41], and the nonparametric richness estimators ACE (abundance-based coverage estimator) [42] and Chao1 [43] were calculated for each community based on each target gene.

Nucleotide sequences accession numbers

The partial sequences of *alk*, ARHD and *bamA* genes determined in this study for the environmental clones were deposited at the Genbank database under the accession numbers: *alk* - JX171303-JX171314, ARHD - JX512729-JX512738 and *bamA* - JX512739-JX512747. It is worth to note that only one representative sequence of each OPF was deposited at the database.

Results and Discussion

The composition of microbial communities in samples from Brazilian petroleum fields was determined by the construction and analysis of three functional gene libraries: *alk*, ARHDs and *bamA*. In the case of *alk* gene libraries, 63 clones from GMR75 oil sample (biodegraded) and 87 clones from PTS01 oil sample (non-biodegraded) were obtained. For the ARHDs gene libraries, 88 clones from GMR75 sample and 80 clones from PTS01 sample were obtained, and about 90 clones from GMR75 library and 93 clones from PTS01 library were obtained for the *bamA* gene.

Operational Protein Family (OPF) was defined based on cutoff values at 93% dissimilarity level between sequences [44] in the case of *alk* gene, and 96% and 97% for the ARHDs and *bamA* genes, respectively. These cutoff values were selected based on the analysis of the phylogenetic trees of total clones in combination with MOTHUR analysis. Shannon, ACE and CHAO1 scores indicated that the GMR75 library is more diverse than the PTS01 library for the two functional genes analyzed (Table 3).

In both libraries of the *alk* gene, all 150 sequences showed the

Library (n ^a)	Richennes ^b	ACE ^c	Chao1 ^d	Shannon ^e	1/Simpson ^f
<i>alk</i> (cut-off 93%)					
GMR75 (63)	10	14	12	1.69	4.52
PTS01 (87)	5	9.21	6.5	0.80	1.60
ARHD (cut-off 96%)					
GMR75 (88)	4	4	4	1.15	2.90
PTS01 (80)	7	13	9	1.34	3.48
<i>bamA</i> (cut-off 97%)					
GMR75 (90)	7	20	9	1.16	2.66
PTS01 (93)	6	7	6	0.97	1.99

^a Number of gene sequences analyzed.

^b Richness is based on observed unique OTUs.

^c Nonparametric statistical prediction of total richness of different OTUs and OPFs based on distribution of abundant (10) and rare (10) OTUs.

^d Nonparametric statistical predictions of total richness of OTUs and OPFs based on distribution of singletons and doubletons.

^e Shannon diversity index. A higher number represents more diversity.

^f Reciprocal of Simpson's diversity index. A higher number represents more diversity.

Table 3: Diversity indices and richness estimators of the functional gene clone libraries.

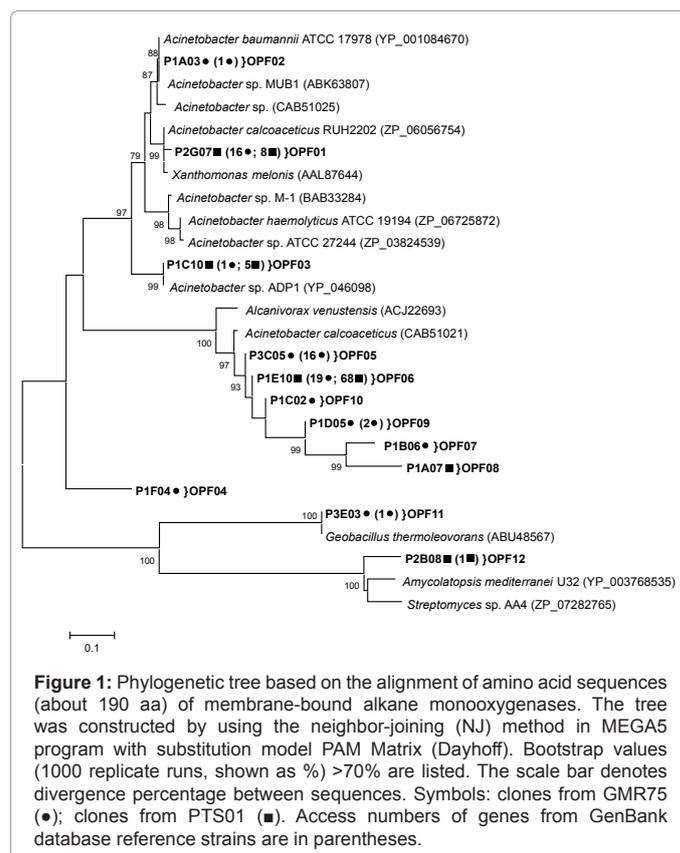


Figure 1: Phylogenetic tree based on the alignment of amino acid sequences (about 190 aa) of membrane-bound alkane monooxygenases. The tree was constructed by using the neighbor-joining (NJ) method in MEGA5 program with substitution model PAM Matrix (Dayhoff). Bootstrap values (1000 replicate runs, shown as %) >70% are listed. The scale bar denotes divergence percentage between sequences. Symbols: clones from GMR75 (●); clones from PTS01 (■). Access numbers of genes from GenBank database reference strains are in parentheses.

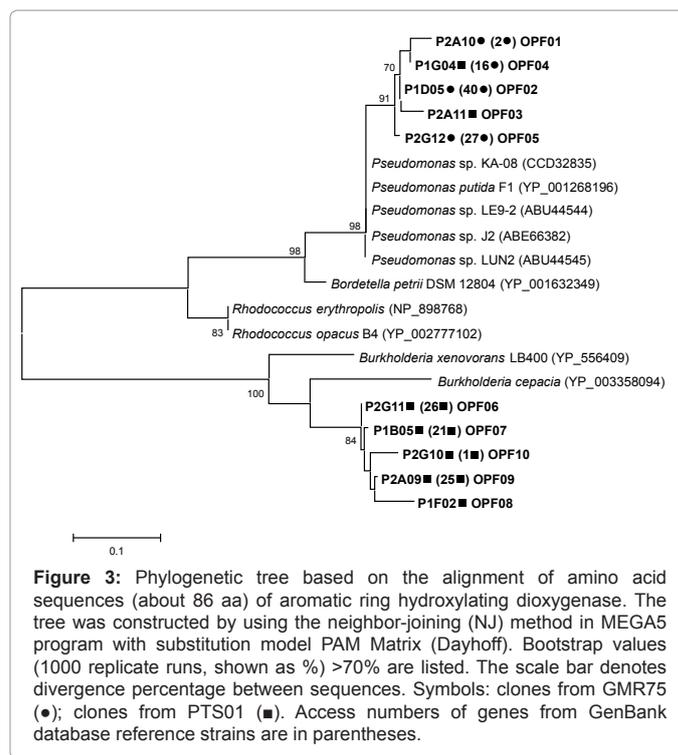
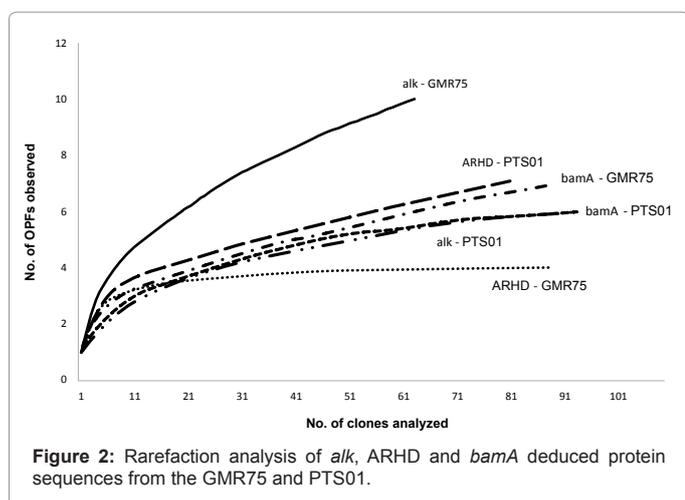
presence of the two internal conserved regions (motif B, EHXXGHH and motif C, NYXEHYG) [45,46] and matched with *alk* genes from GenBank using the BLASTX program. Seven OPFs in GMR75, 2 OPFs in PTS01 and 3 OPFs composed by sequences from both *alk* gene libraries were found (Figure 1).

One of these OPFs (OPF 11) matched with the *alkB* gene from *Geobacillus thermoleovorans*, which has the ability to grow in oil and its

hydrocarbon degradation capacity was reported in an extensive study of the genus *Geobacillus* [47]. This bacterium has been commonly isolated from petroleum reservoirs, especially high temperature oil fields [48,49]. *Geobacillus* spp. has attracted industrial interest due to their potential applications in biotechnological processes as sources of various thermostable enzymes [50]. Shestakova and co-authors [51] also detected the presence of *Geobacillus*-like organisms in the *alk* gene RNA library and in the 16S rRNA gene library from hydrocarbon-oxidizing aerobic enrichments from a high-temperature petroleum reservoir in China, showing the predominance and functional activity of *geobacilli*.

Three of the *alk* OPFs (01, 02 and 03) matched with genes from *Acinetobacter baumannii*, *Acinetobacter calcoaceticus* and some strains of *Acinetobacter* spp. *Acinetobacter* spp. are widespread in nature and can be obtained from water, soil and living organisms. They can use various carbon sources for growth [52]. Sette et al. [53] characterized and compared the bacterial community structure of two distinct samples (non-biodegraded and highly biodegraded oils) from a petroleum field in Brazil and showed that the genus *Acinetobacter* was exclusive to and predominant in the biodegraded oil sample. Bacteria belonging to this genus are known to be involved in biodegradation and are considered one of the most efficient oil degraders [9,11,52]. The OPF04 was represented by only one sequence, which was recovered in a clearly separate cluster, not related to any previously identified *alk* sequences present in the public database (Figure 1). This OPF is a great candidate for a putative new *alk* gene from a bacterium indigenous to the oil reservoir. A great number of OPFs (05, 06, 07, 08, 09 and 10) were grouped together and were somewhat distantly related to sequences of organisms represented in the databases. These results indicate the presence of novel putative alkane monooxygenase genes in petroleum reservoirs not previously identified in any other environment or isolate. Kuhn and colleagues [19] found a differential distribution of *alk* genes between two sites of Antarctic marine environments, and the predominant presence of new *alk* genes, mainly in the pristine site. Recently, Miquelletto et al. [54] also observed a higher diversity of an alkane hydroxylase gene in petroliferous soil sample in comparison to a non-petroliferous soil. The authors showed that most of the alkane hydroxylase genes corresponded to potentially new genes.

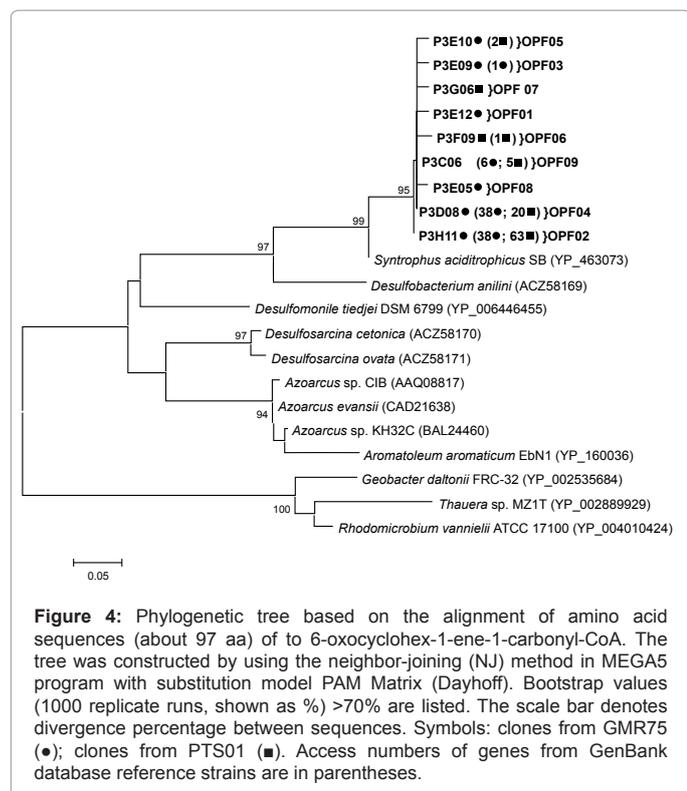
Diversity analyses clearly showed profound differences between GMR75 and PTS01 libraries. ACE and Chao estimators yielded higher richness values of *alk* genes in the GMR75 library (Table 3). Both



Shannon and Simpson diversity indices were consistent with rarefaction analyses, showing higher diversity of *alk* sequences in GMR75 library than in the PTS01 library (Table 3 and Figure 2).

In the ARHD gene libraries, the majority of the sequences from both libraries matched with ARHD genes from GenBank using the BLASTx program. Phylogenetic analysis revealed 3 OPFs exclusive to the GMR75 library, 6 OPFs exclusive to the PTS01 library and 1 OPF composed by sequences from both libraries (Figure 3). Five OPFs (OPFs 01, 02, 03, 04 and 05) were related with dioxygenase genes from several species of *Pseudomonas*. The enzymes for the degradation of aromatic hydrocarbons have been reported to be encoded by genes located on plasmids (e.g., TOL, NAH, SAL, OCT, and CAM) in *Pseudomonas* spp. [55]. A new genetic organization and co-regulation of a cluster of genes involved in the first steps of phenol and benzene catabolic pathways was described in *Pseudomonas* spp. M1 [56]. The authors showed that, differently from the established models for *Pseudomonas* upper pathway, the *Pseudomonas* spp. M1 exhibited exceptional biodegradation ability towards a wide range of toxic and/or recalcitrant compounds. On the other hand, 5 OPFs (OPFs 06, 07, 08, 09 and 10) were related with dioxygenase genes from *Burkholderia cepacia* and *Burkholderia xenovorans*. These two species have been extensively studied in terms of degradation of the low-molecular-weight PAHs naphthalene and phenanthrene [57] and detected in many crude-oil samples [34,58,59].

Cunha et al. [60] isolated bacteria of the genus *Bacillus* from the rock of an oil reservoir in Brazil and detected the presence of catechol dioxygenase genes in 4 *Bacillus* isolates by PCR. Interestingly, in the present study none of the ARHD gene sequences analyzed matched with ARHD genes from *Bacillus* strains represented in the databases. In addition, rarefaction analysis of ARHD genes (Figure 2) yielded plateau-shaped curves, indicating that the ARHD gene diversity in the environment was close to saturation and the sampling effort was satisfactory. These results might suggest that ARHD genes belonging to the genus *Bacillus* are simply not present in Potiguar oil samples,



or that they could not be detected due to extremely low abundance or to bias inherent to the DNA extraction (differential lysis) and/or PCR techniques (preferential amplification).

Phylogenetic analysis of *bamA* gene libraries obtained from samples GMR75 and PTS01 (Figure 4) showed that nine OPFs were grouped together and were somewhat distantly related to sequences of organisms represented in the databases, being *bamA* genes from *Synthrophus aciditrophicus* and *Desulfobacterium anilini* the closest described sequences. *Synthrophus aciditrophicus* is a bacterium that degrades a wide range of organic compounds, including alcohols, fatty acids, and hydrocarbons, including methane, which are degraded syntrophically under anaerobic conditions [61]. *Desulfobacterium anilini* is a sulfate-reducing bacterium known to be involved in phenol degradation [62]. Kuntze et al. [32] observed that the *bamA* gene was highly conserved among organisms studied by the group. In the present study, most of the clone sequences from the oil samples GMR75 and PTS01 were clustered into OPF02 and OPF04 (Figure 4), demonstrating that the gene under study has indeed a conserved region. In addition, all of the 183 sequences obtained from GMR75 and PTS01, showed none or low similarity level to the sequences from reference organisms. This suggests that the sequences corresponding to the *bamA* gene recovered from the petroleum reservoirs are potentially new sequences.

In spite of the higher diversity indices observed for the GMR75 library in the case of *alk* and *bamA* genes, the same was not observed for the ARHD gene. ACE and Chao indices indicated that the GMR75 library contained more distinct OPFs (more richness) than the PTS01 library for both catabolic genes (Table 3). Shannon and Simpson diversity indices also showed higher diversity of *alk* and *bamA* sequences in library GMR75 than in library PTS01. These results suggested that the bacterial communities of the biodegraded and non-biodegraded oil reservoirs are different also on the distribution of such catabolic genes.

A recent paper published by our research team [34] reported the microbial communities present in the same oil samples used in the present work, originated from the biodegraded (GMR75) and non-biodegraded (PTS1) terrestrial reservoirs from Potiguar Basin, by using 16S rRNA gene libraries. The authors demonstrated likewise the presence of the genera *Acinetobacter*, *Geobacillus*, *Pseudomonas* and *Syntrophus* in the microbial community. These results are in agreement with the ones found in the present work, confirming the presence of bacteria bearing catabolic genes involved in the process of biodegradation. However, Silva et al. [34] found the predominance of *Geobacillus* in the microbial community of PTS1 sample, but no representative of this genus was found in the GMR75 sample. These results are in contrast with those found in the present work, since *alk* genes similar to *Geobacillus* alkane monooxygenases were found only in the GMR75 oil. These contrasting results could be explained by the fact that the sampling effort was not sufficient to cover both phylogenetic [34] and functional diversity (Figure 2) present in these oil samples. On the other hand, Silva et al. [34] suggested that anaerobic metabolism is the dominant process for hydrocarbon degradation in GMR75 oil field, corroborating with the result that the biodegraded oil sample was more diverse for the *bamA* gene.

In conclusion, the combined results obtained from the catabolic gene libraries showed that both oil samples contained *alk*, ARHD and *bamA* genes related to genes present in bacterial species known to be hydrocarbon-degraders, such as *Acinetobacter*, *Pseudomonas*, *Geobacillus*, *Burkholderia* and deltaproteobacteria. Nonetheless, subsets of the catabolic sequences recovered from both petroleum reservoirs may correspond to new genes endemic to such environments, possibly harbored by yet uncultivated and unidentified microorganisms. Moreover, it is important to point out that this was the first work using degenerate primers to detect the presence of functional genes involved in the degradation of hydrocarbons in petroleum reservoirs, demonstrating the genetic potential of such microbial communities to biodegrade the oil. However, such biodegradation will only occur if the environmental conditions are appropriate for the metabolism of the degrading microorganisms. Curiously, the non-biodegraded oil sample showed to be more diverse in terms of the catabolic genes ARHD, suggesting that the environmental parameters that control biodegradation are playing a major role in differentiating both biodegraded and non-biodegraded oil reservoirs. As previously detailed in Silva et al. [34], the oil sample GMR75 has been classified as moderately biodegraded, produced from a reservoir with an *in situ* temperature of 42.2°C at a depth of 535.5-540.5 m. On the other hand, the PTS1 sample has been classified as non-biodegraded, collected from a deeper and hotter reservoir, with an *in situ* temperature of 48.3°C at a depth of 801-803 m [63]. Both oils were collected in the same basin, but they are related to different petroleum systems, so different source rocks and reservoirs. These discrepancies *in situ* temperature, depth and, possibly, in source rock composition and water salinity might explain the functional differences of the hydrocarbon degrading populations between both reservoirs. It is worth to mention that studies carried out with DNA sequences shows only the bacteria with potential to degrade oil, requiring further studies at the level of gene expression and GC-MS assays for the precise evaluation of the degradation process.

Finally, the results obtained allowed us to confirm the effectiveness of the primer sets used for the analysis of the community structure of bacteria that have the genetic potential to degrade hydrocarbons by aerobic and anaerobic pathways. The application of molecular methods for the rapid detection of specific microorganisms or genes in environmental samples is a valuable tool for studies focusing on

diversity, abundance and distribution of hydrocarbon degradation populations, providing new insights into the diversity of catabolic genes in the oil reservoir environment and allowing a better understanding of the role of microorganisms on the biodegradation process.

Acknowledgements

We thank PETROBRAS for the financial support, donation of samples and authorization to publish this work. L. C. L. Verde was supported by grants from FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo). The authors are grateful to Rubens T. D. Duarte for technical assistance.

References

- Röling WF, Head IM, Larter SR (2003) The microbiology of hydrocarbon degradation in subsurface petroleum reservoirs: perspectives and prospects. *Res Microbiol* 154: 321-328.
- Connan J (1984) Biodegradation of crude oils in reservoirs. *Advances in Petroleum Geochemistry*. Academic Press, London, UK.
- Van Hamme JD, Singh A, Ward OP (2003) Recent advances in petroleum microbiology. *Microbiol Mol Biol Rev* 67: 503-549.
- Higashioka Y, Kojima H, Sato S, Fukui M (2009) Microbial community analysis at crude oil-contaminated soils targeting the 16S ribosomal RNA, *xyIM*, *C23O*, and *bcr* genes. *J Appl Microbiol* 107: 126-135.
- Magot M, Ollivier B, Patel BK (2000) Microbiology of petroleum reservoirs. *Antonie Van Leeuwenhoek* 77: 103-116.
- Aitken CM, Jones DM, Larter SR (2004) Anaerobic hydrocarbon biodegradation in deep subsurface oil reservoirs. *Nature* 431: 291-294.
- van Beilen JB, Li Z, Duetz WA, Smits THM, Witholt B (2003) Diversity of Alkane Hydroxylase Systems in the Environment. *Oil & Gas Sci Technol – Rev IFP* 58: 427-440.
- Watkinson RJ, Morgan P (1990) Physiology of aliphatic hydrocarbon-degrading microorganisms. *Biodegradation* 1: 79-92.
- Smits TH, Balada SB, Witholt B, van Beilen JB (2002) Functional analysis of alkane hydroxylases from gram-negative and gram-positive bacteria. *J Bacteriol* 184: 1733-1742.
- van Beilen JB, Smits TH, Whyte LG, Schorch S, Röthlisberger M, et al. (2002) Alkane hydroxylase homologues in Gram-positive strains. *Environ Microbiol* 4: 676-682.
- van Beilen JB, Funhoff EG (2007) Alkane hydroxylases involved in microbial alkane degradation. *Appl Microbiol Biotechnol* 74: 13-21.
- Sotsky JB, Greer CW, Atlas RM (1994) Frequency of genes in aromatic and aliphatic hydrocarbon biodegradation pathways within bacterial populations from Alaskan sediments. *Can J Microbiol* 40: 981-985.
- Knaebel DB, Crawford RL (1995) Extraction and purification of microbial DNA from petroleum-contaminated soils and detection of low numbers of toluene, octane and pesticide degraders by multiplex polymerase chain reaction and Southern analysis. *Mol Ecol* 4: 579-591.
- Guo C, Sun W, Harsh JB, Ogram A (1997) Hybridization Analysis of Microbial DNA from Fuel Oil-Contaminated and Noncontaminated Soil *Microb Ecol* 34: 178-187.
- Whyte LG, Greer CW, Inniss WE (1996) Assessment of the biodegradation potential of psychrotrophic microorganisms. *Can J Microbiol* 42: 99-106.
- Yergeau E, Arbour M, Brousseau R, Juck D, Lawrence JR, et al. (2009) Microarray and real-time PCR analyses of the responses of high-arctic soil bacteria to hydrocarbon pollution and bioremediation treatments. *Appl Environ Microbiol* 75: 6258-6267.
- Stapleton RD, Saylor GS (1998) Assessment of the Microbiological Potential for the Natural Attenuation of Petroleum Hydrocarbons in a Shallow Aquifer System. *Microb Ecol* 36: 349-361.
- Whyte LG, Smits TH, Labbé D, Witholt B, Greer CW, et al. (2002) Gene cloning and characterization of multiple alkane hydroxylase systems in *Rhodococcus* strains Q15 and NRRL B-16531. *Appl Environ Microbiol* 68: 5933-5942.
- Kuhn E, Bellicanta GS, Pellizari VH (2009) New alk genes detected in Antarctic marine sediments. *Environ Microbiol* 11: 669-673.
- Panicker G, Mojib N, Aislabie J, Bej AK (2010) Detection, expression and quantitation of the biodegradative genes in Antarctic microorganisms using PCR. *Antonie Van Leeuwenhoek* 97: 275-287.
- Bell TH, Yergeau E, Martineau C, Juck D, Whyte LG, et al. (2011) Identification of nitrogen-incorporating bacteria in petroleum-contaminated arctic soils by using [¹⁵N]DNA-based stable isotope probing and pyrosequencing. *Appl Environ Microbiol* 77: 4163-4171.
- Kauppi B, Lee K, Carredano E, Parales RE, Gibson DT, et al. (1998) Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene 1,2-dioxygenase. *Structure* 6: 571-586.
- Cébron A, Norini MP, Beguiristain T, Leyval C (2008) Real-Time PCR quantification of PAH-ring hydroxylating dioxygenase (PAH-RHD α) genes from Gram positive and Gram negative bacteria in soil and sediment samples. *J Microbiol Methods* 73: 148-159.
- Gibson DT, Parales RE (2000) Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Curr Opin Biotechnol* 11: 236-243.
- Romine MF, Stillwell LC, Wong KK, Thurston SJ, Sisk EC, et al. (1999) Complete sequence of a 184-kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199. *J Bacteriol* 181: 1585-1602.
- Goyal AK, Zylstra GJ (1996) Molecular cloning of novel genes for polycyclic aromatic hydrocarbon degradation from *Comamonas testosteroni* GZ39. *Appl Environ Microbiol* 62: 230-236.
- Jeon CO, Park M, Ro HS, Park W, Madsen EL (2006) The naphthalene catabolic (*nag*) genes of *Polaromonas naphthalenivorans* CJ2: evolutionary implications for two gene clusters and novel regulatory control. *Appl Environ Microbiol* 72: 1086-1095.
- Laurie AD, Lloyd-Jones G (1999) The *phn* genes of *Burkholderia* sp. strain RP007 constitute a divergent gene cluster for polycyclic aromatic hydrocarbon catabolism. *J Bacteriol* 181: 531-540.
- Simon MJ, Oslund TD, Saunders R, Ensley BD, Suggs S, et al. (1993) Sequences of genes encoding naphthalene dioxygenase in *Pseudomonas putida* strains G7 and NCIB 9816-4. *Gene* 127: 31-37.
- Head IM, Jones DM, Larter SR (2003) Biological activity in the deep subsurface and the origin of heavy oil. *Nature* 426: 344-352.
- Staats M, Braster M, Röling WF (2011) Molecular diversity and distribution of aromatic hydrocarbon-degrading anaerobes across a landfill leachate plume. *Environ Microbiol* 13: 1216-1227.
- Kuntze K, Shinoda Y, Moutakki H, McInerney MJ, Vogt C, et al. (2008) 6-Oxocyclohex-1-ene-1-carboxyl-coenzyme A hydrolases from obligately anaerobic bacteria: characterization and identification of its gene as a functional marker for aromatic compounds degrading anaerobes. *Environ Microbiol* 10: 1547-1556.
- Higashioka Y, Kojima H, Fukui M (2011) Temperature-dependent differences in community structure of bacteria involved in degradation of petroleum hydrocarbons under sulfate-reducing conditions. *J Appl Microbiol* 110: 314-322.
- Silva TR, Verde LCL, Santos Neto EV, Oliveira VM (2012) Diversity analyses of microbial communities in petroleum samples from Brazilian oil fields. *Int Biodeter Biodegr*.
- Bellicanta GS (2005) Diversity of catabolic genes in sediment samples of Santos and São Vicente Estuarine Systems, SP. *Biota Neotropica* 5.
- Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8: 175-185.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75: 7537-7541.
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425.
- Hall BG (2013) Building Phylogenetic Trees from Molecular Data with MEGA. *Mol Biol Evol*.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876-4882.
- Magurran AE (2004) *Measuring Biological Diversity*. Blackwell Publishing, Oxford, UK.

42. Chao A, Ma MC, Yang MCK (1993) Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika* 80: 193-201.
43. Chao A (1984) Non-parametric estimation of the number of classes in a population. *Scand J Stat* 11: 265-270.
44. Schloss PD, Handelsman J (2008) A statistical toolbox for metagenomics: assessing functional diversity in microbial communities. *BMC Bioinformatics* 9: 34.
45. Smits TH, Röthlisberger M, Witholt B, van Beilen JB (1999) Molecular screening for alkane hydroxylase genes in Gram-negative and Gram-positive strains. *Environ Microbiol* 1: 307-317.
46. van Beilen JB, Smits TH, Roos FF, Brunner T, Balada SB, et al. (2005) Identification of an amino acid position that determines the substrate range of integral membrane alkane hydroxylases. *J Bacteriol* 187: 85-91.
47. Turova TP, Nazina TN, Mikhaĭlova EM, Rodionova TA, Ekimov AN, et al. (2008) AlkB homologues in thermophilic bacteria of the genus *Geobacillus*. *Mol Biol (Mosk)* 42: 247-257.
48. Onstott TC, Phelps TJ, Colwell FS, Ringelberg D, White DC, et al. (1998) Observations pertaining to the origin and ecology of microorganisms recovered from deep subsurface Taylorsville Basin, Virginia. *Geomicrobiol J* 15: 353-385.
49. Kato T, Miyayama A, Kanaya S, Morikawa M (2009) Alkane inducible proteins in *Geobacillus thermoleovorans* B23. *BMC Microbiol* 9: 60.
50. McMullan G, Christie JM, Rahman TJ, Banat IM, Ternan NG, et al. (2004) Habitat, applications and genomics of the aerobic, thermophilic genus *Geobacillus*. *Biochem Soc Trans* 32: 214-217.
51. Shestakova NM, Korshunova AV, Mikhailova EM, Sokolova DSH, Tourova TP, et al. (2011) Characterization of the aerobic hydrocarbon-oxidizing enrichments from a high-temperature petroleum reservoir by comparative analysis of DNA and RNA derived clone libraries. *Microbiology* 80: 60-69.
52. Barbe V, Vallenet D, Fonknechten N, Kreimeyer A, Oztas S, et al. (2004) Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Res* 32: 5766-5779.
53. Sette LD, Simioni KC, Vasconcelos SP, Dussan LJ, Neto EV, et al. (2007) Analysis of the composition of bacterial communities in oil reservoirs from a southern offshore Brazilian basin. *Antonie Van Leeuwenhoek* 91: 253-266.
54. Miqueletto PB, Andreote FD, Dias AC, Ferreira JC, Dos Santos Neto EV, et al. (2011) Cultivation-independent methods applied to the microbial prospection of oil and gas in soil from a sedimentary basin in Brazil. *AMB Express* 1: 35.
55. Irie S, Doi S, Yorifuji T, Takagi M, Yano K (1987) Nucleotide sequencing and characterization of the genes encoding benzene oxidation enzymes of *Pseudomonas putida*. *J Bacteriol* 169: 5174-5179.
56. Santos PM, Sá-Correia I (2007) Characterization of the unique organization and co-regulation of a gene cluster required for phenol and benzene catabolism in *Pseudomonas* sp. M1. *J Biotechnol* 131: 371-378.
57. Laurie AD, Lloyd-Jones G (2000) Quantification of *phnAc* and *nahAc* in contaminated New Zealand soils by competitive PCR. *Appl Environ Microbiol* 66: 1814-1817.
58. Yoshida N, Yagi K, Sato D, Watanabe N, Kuroishi T, et al. (2005) Bacterial communities in petroleum oil in stockpiles. *J Biosci Bioeng* 99: 143-149.
59. Korenblum E, Souza DB, Penna M, Seldin L (2012) Molecular analysis of the bacterial communities in crude oil samples from two Brazilian offshore petroleum platforms. *Int J Microbiol* 2012: 156537.
60. da Cunha CD, Rosado AS, Sebastián GV, Seldin L, von der Weid I (2006) Oil biodegradation by *Bacillus* strains isolated from the rock of an oil reservoir located in a deep-water production basin in Brazil. *Appl Microbiol Biotechnol* 73: 949-959.
61. McInerney MJ, Rohlin L, Mouttaki H, Kim U, Krupp RS, et al. (2007) The genome of *Syntrophus aciditrophicus*: life at the thermodynamic limit of microbial growth. *Proc Natl Acad Sci U S A* 104: 7600-7605.
62. Ahn YB, Chae JC, Zylstra GJ, Häggblom MM (2009) Degradation of phenol via phenylphosphate and carboxylation to 4-hydroxybenzoate by a newly isolated strain of the sulfate-reducing bacterium *Desulfobacterium anilini*. *Appl Environ Microbiol* 75: 4248-4253.
63. dos Santos Neto EV, Hayes JM (1999) Use of hydrogen and carbon stable isotopes characterizing oils from the Potiguar Basin (Onshore), Northeastern Brazil. *AAPG Bulletin* 83: 496-518.