

Classification and Identification of Petroleum Microorganisms by MALDI-TOF Mass Spectrometry

Célio F. F. Angolini,^a Eduardo J. Pilau,^a Patrícia F. Lopes-Oliveira,^b Isabel N. S. Garcia,^b
Fábio C. Gozzo,^a Valéria M. de Oliveira^b and Anita J. Marsaioli^{*a}

^aChemistry Institute, University of Campinas (Unicamp), CP 6154, 13083-970 Campinas-SP, Brazil

^bMicrobial Resource Division, Research Center of Chemistry, Biology and Agriculture (CPQBA),
University of Campinas (Unicamp), CP 6171, 13081-970 Campinas-SP, Brazil

Indigenous bacteria isolated from a crude oil sample from a deep water reservoir in the Pampo Sul Oilfield (Campos Basin-RJ, Brazil) were previously classified as strains of *B. pumilus*. However, their enzymatic activities with fluorogenic probes and rates of petroleum biodegradation were completely different. Some of the bacteria depleted *n*-alkanes, whereas others did not. Aromatic compounds reported to be recalcitrant were also biodegraded by some of these *Bacillus* strains, revealing their outstanding ability to deplete petroleum. Further classification using matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) followed by statistical analysis revealed that these strains could be clustered into three different groups, consistent with their enzymatic activity evaluation. A more accurate phylogenetic analysis using *gyrB* gene sequences confirmed the MALDI-TOF MS classification of three groups of strains and identified them as *Bacillus safensis*, *B. cereus* and *B. thuringiensis*.

Keywords: petroleum microorganisms, *Bacillus sp.*, 16S rRNA, microorganisms identification

Introduction

Petroleum is a complex mixture of hydrocarbons (aliphatic and aromatic) and compounds containing oxygen (carboxylic acids, phenols, ethers, etc.), nitrogen (imidazoles, carbazoles, etc.) and sulfur.¹ Hydrocarbons are economically important and are preferentially consumed during biodegradation, affording carboxylic acids and phenols, a process that adds new constituents to the petroleum (*de novo* synthesis).²

Petroleum biodegradation can occur aerobically, anaerobically or both, as recently suggested by Cruz *et al.*³ The mechanisms of the biodegradation processes are complex,⁴ and understanding the role of each member of a microbial consortium during petroleum depletion is important, with particular emphasis on *Bacillus* strains, which are often present in petroleum samples. Thus, identification of the microorganisms is of importance, which is cumbersome and requires digital genomic information for the detection of 16S rRNA genes to provide specific bacteria classifications.^{5,6}

Alternatively, the use of matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) protein profiling of whole cells has been successfully applied for the classification of bacteria.⁷⁻¹⁰ However, this methodology fails to identify some environmental microorganisms because the current data libraries were mainly constructed using clinical microorganisms.¹¹ Additionally, some cases of the misidentification of microorganisms by phylogenetic techniques, such as the *Bacillus safensis* that were classified as *B. pumilus* by 16S rRNA,¹² prompted us to verify the potential of this technique to classify and identify petroleum microorganisms.

Therefore, the aim of this work was to evaluate the reliability of MALDI-TOF MS in classifying petroleum microorganisms and its versatility in identifying petroleum microorganisms to provide the basis for constructing a data library with petroleum bacteria.

Experimental

Microbial strains and identification

The seven strains used in this work (Table 1) were isolated from water and oil samples by Cruz *et al.*¹³ After

*e-mail: anita@iqm.unicamp.br

growth on nutrient agar (Difco) plates, the genomic DNA of pure cultures were isolated according to the protocol described by Pitcher *et al.*¹⁴ Primers *gyrB* UP-1 and UP-2r were used for the amplification of DNA gyrase subunit B genes,¹⁴ aiming at the identification of the isolates. The reaction mixtures (25 μ L) contained 50 ng of genomic DNA, 2 U of Taq DNA polymerase (Invitrogen), 1X Taq buffer, 1.5 mmol L⁻¹ MgCl₂, 0.2 mmol L⁻¹ dNTP mix (GE Healthcare) and 0.4 μ mol L⁻¹ of each primer. The polymerase chain reaction (PCR) amplification program consisted of 1 cycle at 94 °C for 5 min, 30 cycles at 94 °C for 1 min, 60 °C for 1 min, 72 °C for 2 min and 1 cycle of final extension at 72 °C for 7 min in an Eppendorf thermal cycler. The PCR amplification of *gyrB* gene fragments was confirmed on 1% agarose gel stained with Sybr Safe (Invitrogen). DNA fragments corresponding to gyrase gene sequences were purified using mini-columns (GFX PCR DNA and Gel Band Purification Kit, GE Healthcare) and subjected to sequencing in an automated sequencer (ABI 3500 XL Genetic Analyzer, Applied Biosystems). The sequencing reactions were performed with the BigDye Terminator Cycle Sequencing Standard Kit V3.1 (Applied Biosystems), according to the manufacturer's specifications. The primers for sequencing were UP-1 and UP-2r.¹⁵ Identification was achieved by comparing the sequences with sequence data from type strains available in the GenBank public database (<http://www.ncbi.nlm.nih.gov>). The sequences were aligned using the CLUSTAL X program¹⁶ and analyzed using MEGA software v. 2.1.¹⁷ Evolutionary distance was derived from sequence-pair dissimilarity, which was calculated as implemented in MEGA using the DNA substitution model.¹⁸ The phylogenetic tree was prepared using the neighbor-joining algorithm,¹⁹ with bootstrap values calculated from 1,000 replicate runs.

Aerobic degradation assays

The biodegradation assays were conducted as described by Cruz *et al.*¹³ with isolated microorganisms rather than a consortium. The experiments were performed in duplicate

via incubation in a shaker (150 rpm, 28 °C) and monitoring at 30 and 60 days. The biodegraded oil was extracted with CH₂Cl₂ (3 \times 20 mL). The extracts were combined and then dried over anhydrous Mg₂SO₄, and the solvent was evaporated under reduced pressure. The oily residue was submitted to silica gel column chromatography (60 Å, 70-230 mesh, Sigma-Aldrich), which afforded fractions of saturated constituents (F1, eluted with hexane); heavy aromatic constituents (F2, hexane:toluene, 1:1 v/v); and resins and asphaltenes (F3, CHCl₃:MeOH, 95:5 v/v).

Gas chromatography-mass spectrometry (GC-MS)

The F1 fractions were analyzed using gas chromatography-mass spectrometry (GC-MS) with a Hewlett-Packard 6890 instrument connected to a Hewlett-Packard 5970-MSD mass spectrometer. The GC conditions were as follows: split injection (10:1) with He as a carrier gas at a flow rate of 1 mL min⁻¹. The chromatographic column was a MDN5S (30 m \times 0.25 mm \times 0.25 mm, Supelco). Data were obtained in both full scan and single ion monitoring (SIM) modes by electron ionization at 70 eV. The injector temperature was 300 °C. For the analysis of *n*-alkanes (reconstructed ion chromatogram (RIC) *m/z* 71), the scan mode and the temperature program were used as follows: 80 °C (2 min) to 270 °C at 4 °C min⁻¹ and to 300 °C at 10 °C min⁻¹ (hold for 25 min). Hopanes, homohopanes (*m/z* 191) and steranes (*m/z* 217)¹ were analyzed using the SIM mode and the following temperature program: 70 °C (2 min) to 190 °C at 30 °C min⁻¹, to 250 °C at 1.5 °C min⁻¹ and to 300 °C at 2 °C min⁻¹ (hold 20 min). All samples were analyzed using 0.03 mg mL⁻¹ 5- α -cholestan-3-one as an internal standard for biomarker quantitation.

Biomarker ratios

Biomarker ratios were calculated using the peak areas from RICs. For quantitative analysis, the response factor for the internal standard was calculated by dividing the concentration by the respective peak area. The concentrations of all compounds were determined

Table 1. Hydrolase enzymatic activities of bacteria isolated from the petroleum sample

MO	Fluorogenic probes					MO	Fluorogenic probes				
	EP1	EP2	ES1	ES2	LIP		EP1	EP2	ES1	ES2	LIP
<i>B. safensis</i> SG 10	2%	7%	0.5%	2%	17%	<i>B. safensis</i> SG 26	0.4%	2%	1%	5%	77%
<i>B. thuringiensis</i> SG12-1	2%	6%	2%	3%	46%	<i>B. safensis</i> SG 30	2%	8%	1%	7%	6%
<i>B. safensis</i> SG13	1%	6%	3%	18%	51%	<i>B. safensis</i> SG 42-1	2%	5%	0.9%	2%	1%
<i>B. cereus</i> SG16	0.6%	7%	0.3%	1%	11%						

MO: microorganisms; LIP: R = octanoyl; EP1: R = H; EP2: R = ethyl; ES1: R = acetyl; ES2: R = propanoyl.

by multiplying the respective peak area of the mass chromatograms by the response factor for the internal standard.

Preparation of bacterial cell samples for matrix-assisted laser desorption ionization-MS (MALDI-MS)

Approximately 2-5 mg of bacterial cells were suspended in 1 mL of 0.1% trifluoroacetic acid (TFA), vortexed for 3 min and centrifuged at 11000 rpm. After centrifugation, the supernatant was removed. The process was repeated 3× *per* sample. After extraction, 50 μ L of 0.1% TFA were added, and then the cells were cooled and stored in a refrigerator.

All measurements were performed using a Synapt HDMS (Waters, Manchester, UK) mass spectrometer. MALDI(+)-quadrupole time-of-flight (QTOF) MS of cells was performed using the reflectron V-mode, a 200 Hz solid state (Nd:YAG) laser and a positive ion mode. An aliquot (4 μ L) of a cell suspension was added to 10 μ L of acetonitrile (MeCN) with 0.1% HCO₂H and mixed (1:1) with a 10 mg mL⁻¹ solution of α -cyano-4-hydroxycinnamic acid (α -CHCA) matrix in 50:50 H₂O/MeCN with 0.1% HCO₂H. The mixture was then directly spotted (1.5 μ L) onto MALDI stainless steel target plates and allowed to dry. The typical operating conditions were as follows: laser energy of 250 a.u., sample plate of 20 V, and trap energy and transfer collision energy of 6 and 4 V, respectively. Calibration was performed using polyethylene glycol (PEG) 600/1000/2000 sodium adducts between *m/z* 300 and 3000.

Results and Discussion

The seven microorganism strains were isolated from a biodegraded petroleum sample (P2) from the Pampo

Sul Oilfield (Campos Basin), and they exhibited different hydrolase activities (Table 1) against five fluorogenic probes (Figure 1).²⁰ Enzymatic hydrolyses of epoxides (EP1 and EP2) produced diols (**1**) via epoxide ring opening, which were subsequently oxidized by the sodium periodate present in the experimental solution, yielding the aldehyde (**2**) that, by β -elimination, produced the fluorescent signal of the umberiferyl anion (**3**). The intensity of this signal was used to monitor the enzymatic activity, i.e., a higher fluorescent signal compared to a standard reaction provided the enzymatic reaction yield as a percentage. Thus, more intense fluorescent signals correspond to more active enzymes. Similar processes occurred with the esters ES1, ES2 and ES3, in which the enzymatic hydrolysis of the esters produced fluorescent signals (Figure 1).

Monitoring of the RIC of *m/z* 71 in petroleum biodegradation assays (Figure 2) revealed that four strains preferentially depleted short unbranched aliphatic hydrocarbons (SG 10, 42-1 and 13); however, some strains (SG 12-1) depleted the unbranched and branched [pristane (Pr) and phytane (Ph)] aliphatic hydrocarbons at similar rates. The other strains did not significantly degrade linear hydrocarbons (SG 16, 26 and 30, Table 2). Additionally SG 10, 42-1 and 12-1 showed a preference for even-numbered *n*-alkanes (Table 2, carbon preference index (CPI)). The *n*-alkyl cyclohexanes were also degraded by these strains (monitored RIC *m/z* 82).

Aromatic compounds, such as phenanthrene and methylphenanthrenes, were also biodegraded. These classes of aromatic compounds were used in organic geochemistry to obtain information regarding the maturity of source rock.¹ The methylphenanthrene index (MPI-1) was evaluated in biodegradation assays, and we observed

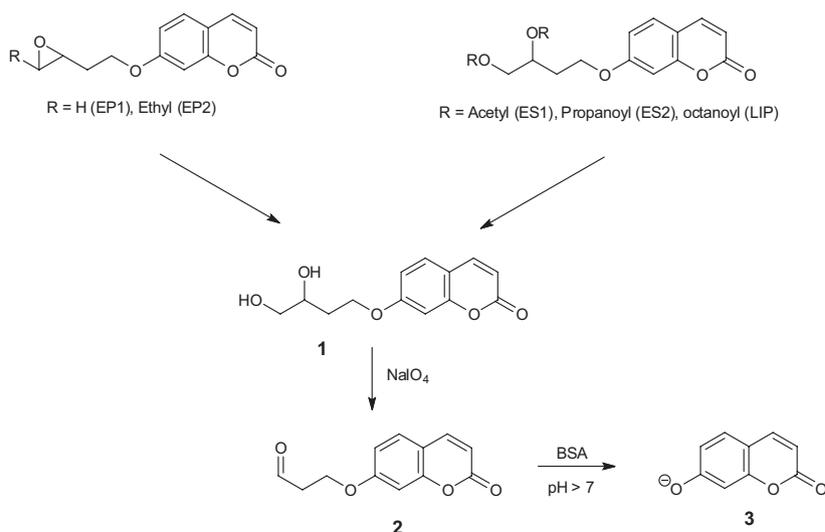


Figure 1. Fluorogenic probes used in miniaturized assays.

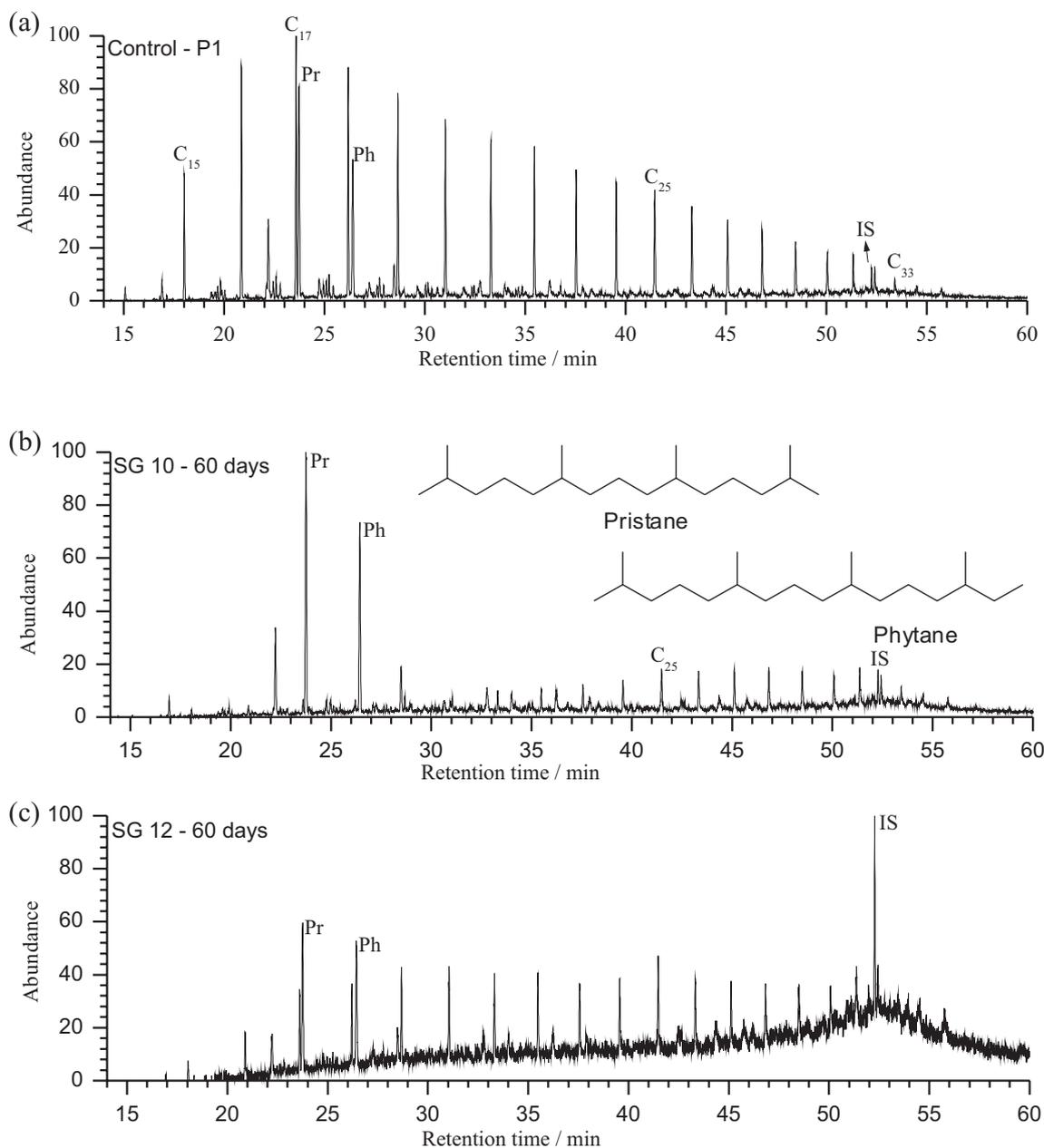


Figure 2. GC-MS chromatograms (RIC m/z 71) for monitoring the biodegradation of oil P1: (a) control sample; (b) by *B. safensis* SG 10 - 60 days; and (c) by *B. thuringiensis* SG 12-1 - 60 days. IS: internal standard (5 α -cholestan-3-one).

different trends with decreasing values for SG 12-1 and 26 and with increasing values for SG 13 and 30 (Table 3).

To identify these strains a 16S rRNA phylogenetic dendrogram of these microorganisms was generated using their specific gene sequences (Figure 3), grouping them as the *B. pumilus/B. safensis* group. However, all seven strains of the *B. pumilus/B. safensis* group were enzymatically different, showing different preferences for the biodegradation of the petroleum constituents. These results required additional evidence to confirm that these microorganisms were distinct, and thus, we performed whole cell MALDI-TOF-MS analyses.

All samples were prepared in triplicate and the MS data were automatically acquired for 1.7 min using a spiral laser pattern over the sample. All of the acquired spectra showed a good reproducibility (see Supplementary Information), so spectrum data were smoothed (Savitsky Golay) using MassLynx® and bin sized using MStable_v2 (homemade software) to proceed with the statistical analysis. The bin sized data were mean centered and submitted to statistical analysis. The hierarchically clustered mass spectra (HCA) using the software Pirouette® v. 3.11 revealed three different groups with ca. 50% similarity (Figure 4).

Table 2. Biodegradation parameters (*m/z* 71) for the biodegradation assays

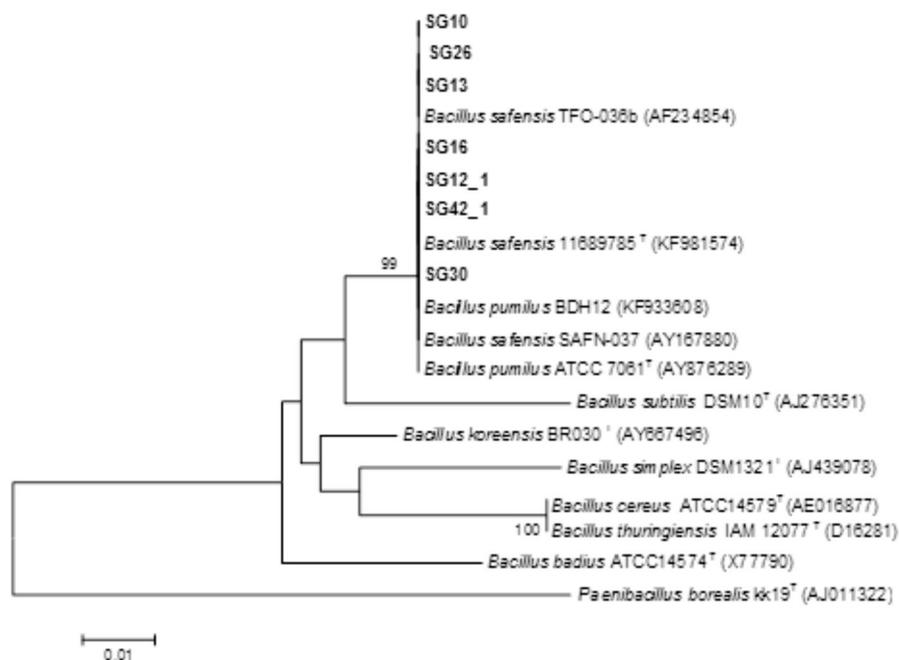
Microorganism	Period / day	Parameter			
		CPI ^a	Pr/ <i>n</i> -C ₁₇	Pr/ <i>n</i> -C ₁₈	Pr/Ph
Control	30	1.12	1.12	1.12	1.12
	60	1.12	1.12	0.93	1.16
<i>B. safensis</i> SG 10	30	1.25	1.51	1.28	1.34
	60	1.41	n.d. ^b	n.d. ^b	1.32
<i>B. thuringiensis</i> SG12-1	30	1.18	1.41	1.19	1.27
	60	1.85	2.64	2.11	1.13
<i>B. safensis</i> SG13	30	1.13	1.08	0.88	1.33
	60	n.d. ^b	n.d. ^b	n.d. ^b	1.41
<i>B. cereus</i> SG16	30	1.16	1.36	1.13	1.36
	60	1.11	1.74	1.53	1.43
<i>B. safensis</i> SG 26	30	1.12	1.10	0.86	1.44
	60	1.18	1.12	0.91	1.43
<i>B. safensis</i> SG 30	30	1.08	1.09	0.88	1.50
	60	1.14	1.10	0.89	1.30
<i>B. safensis</i> SG 42-1	30	1.20	15.63	7.94	1.38
	60	n.d. ^b	n.d. ^b	n.d. ^b	2.10

^aCPI: carbon preference index = $[2 \times \sum(n\text{-Hydrocarbon}_{\text{odd}})] / [n\text{-hydrocarbon}_{\text{even}} + \sum(n\text{-hydrocarbon}_{\text{even}} \text{ intermediate}) + n\text{-hydrocarbon}_{\text{even}}]$; ^bn.d.: not determined - all linear hydrocarbons were completely degraded.

Table 3. MPI-1 index values for the biodegradation assays

Sample	MPI-1 ^a		Sample	MPI-1 ^a	
	30 days	60 days		30 days	60 days
Control	0.77	0.78	<i>B. safensis</i> SG 10	0.77	0.74
<i>B. thuringiensis</i> SG 12-1	0.73	0.67	<i>B. safensis</i> SG 13	0.83	t.d. ^b
<i>B. cereus</i> SG 16	0.78	0.78	<i>B. safensis</i> SG 26	0.59	0.73
<i>B. safensis</i> SG 30	0.76	0.82	<i>B. safensis</i> SG 42-1	0.76	0.74

^aMPI-1 = $1.5 \times \{[(2\text{-MP}) + (3\text{-MP})] / [(P) + (1\text{-MP}) + (9\text{-MP})]\}$; where P = phenanthrene, 2-MP = 2-methylphenanthrene, 3-MP = 3-methylphenanthrene, 1-MP = 1-methylphenanthrene and 9-MP = 9-methylphenanthrene; ^bt.d.: completely degraded.

**Figure 3.** Phylogenetic analysis based on partial 16S rRNA sequences obtained from the bacterial isolates belonging to the *Bacillus safensis*/*B. pumilus* group.

Similar groups were also discriminated via principal component analysis (PCA) (Figure 5, score 1 vs. score 2). The principal component (PC) numbers were set to 2 explaining 65.49% of variance, cross-validation was also applied (Table S1). Additionally, the loadings vs. variables plot from the PCA analysis showed that all of the intense mass signals contributed to the sample classification (Figure 6).

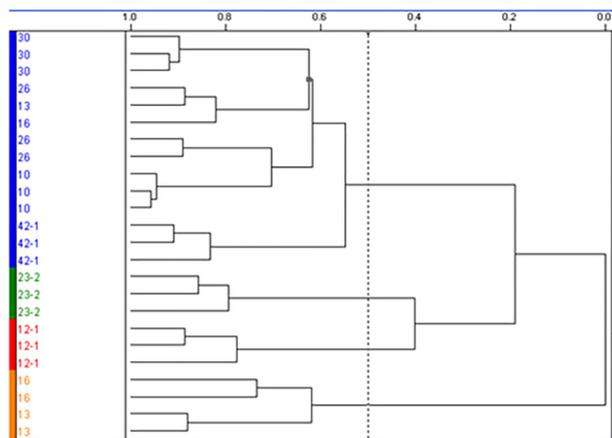


Figure 4. Dendrogram from HCA (mean center) of the MALDI-TOF MS analysis of whole cells from seven isolated strains preliminarily identified as *Bacillus pumilus*.

These results were inconsistent with the previous identification by 16S rRNA gene sequencing. Consequently, *gyrB* gene sequencing and a new phylogenetic analysis were undertaken, which allowed the discrimination of these bacterial isolates into distinct *Bacillus* species (Figure 7). These use of the *gyrB* gene as a phylogenetic marker confirmed the MALDI-TOF MS analysis discrimination

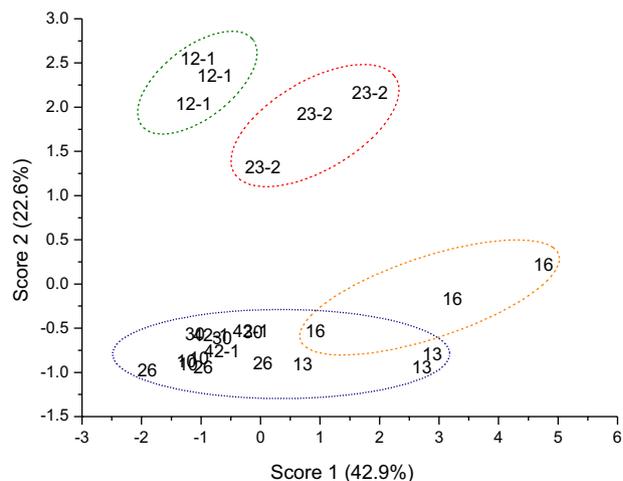


Figure 5. PCA (mean center) scores plot from the MALDI-TOF MS analysis of SG-X whole cells from seven isolated strains preliminarily identified as *Bacillus pumilus*.

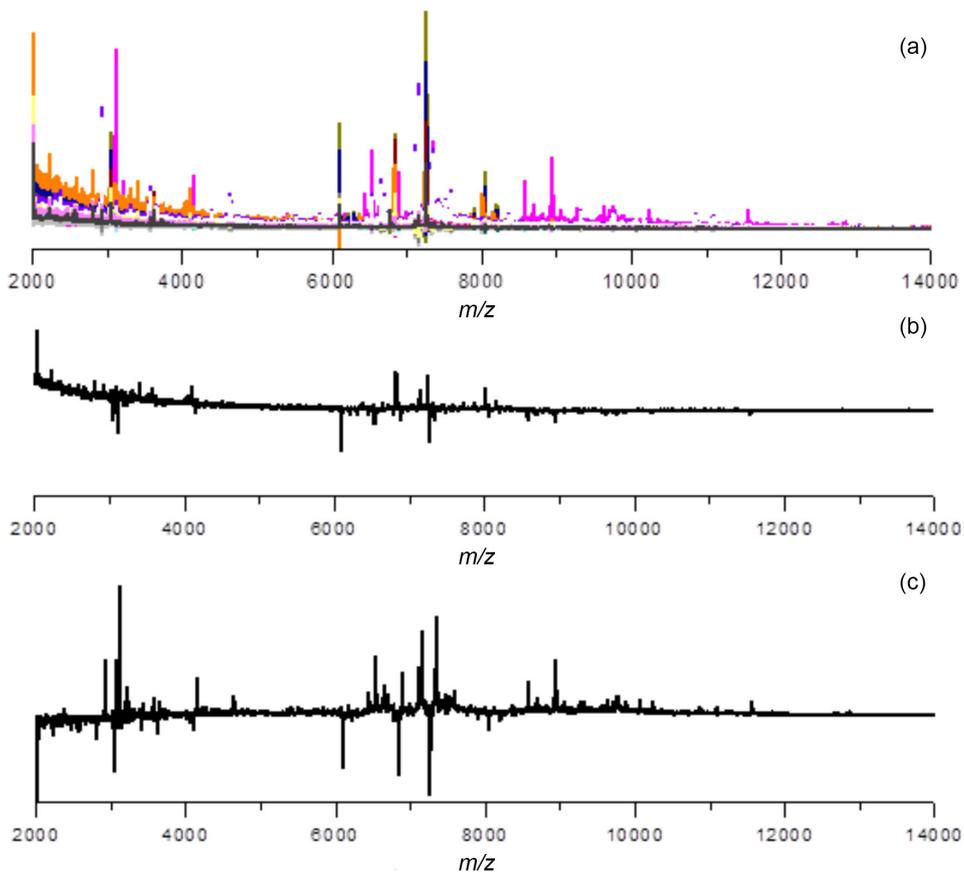


Figure 6. (a) Reconstructed chromatograms and loadings vs. variables: (b) Score 1 and (c) Score 2.

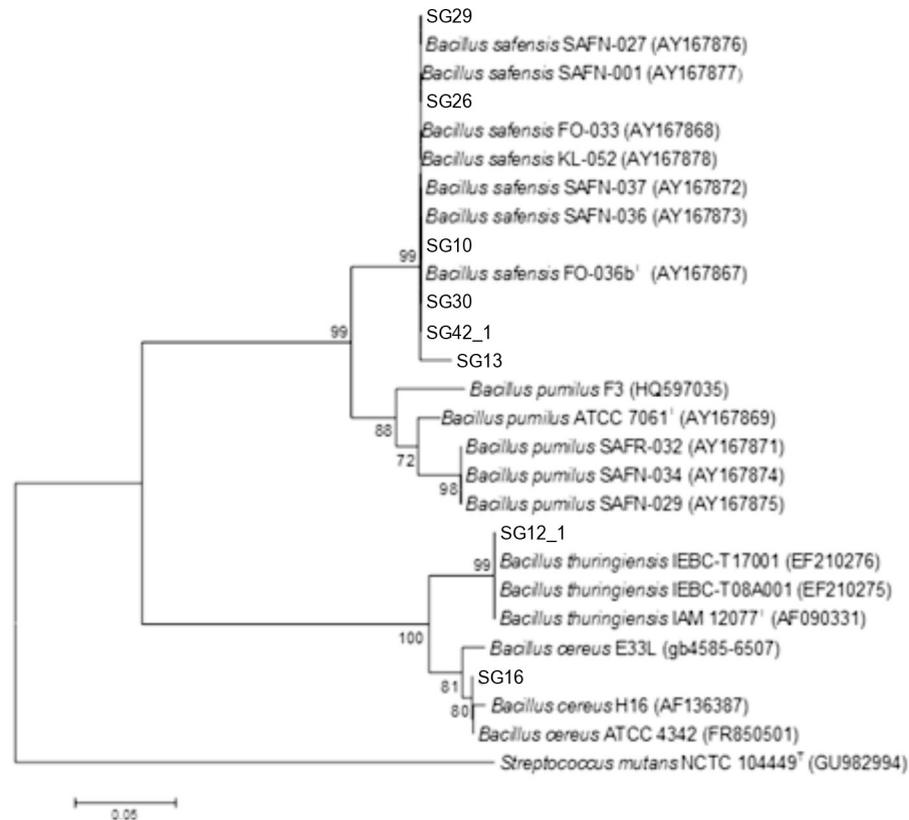


Figure 7. Phylogenetic analysis based on partial gyrase B gene sequences obtained from bacterial isolates belonging to *Bacillus safensis*/*B. pumilus* and *B. cereus*/*B. thuringiensis* groups.

of the strains, indicating the effectiveness of this technique even for strains unable to be differentiated by the 16S rRNA gene.

Conclusions

MALDI-TOF MS was successfully applied for the characterization of seven petroleum bacteria based on whole cell MS analyses. These seven petroleum strains were clustered into three groups, consistent with the molecular identification using the *gyrB* gene sequence as the phylogenetic marker. These results show the technical extent, as the usual phylogenetic marker (16S rRNA) was not able to distinguish these strains due to their similarities.

The intrinsic characteristics of the technique, such as its analysis speed, low sample requirement and relative ease of use, make MALDI-TOF MS a promising approach for the identification of petroleum bacteria, which motivates the assembly of a data library from petroleum microorganisms obtained using MALDI-TOF MS equipment. Additionally, the commercially available programs for bacterial identification are limited when the identification of bacteria is running with a consortium, which indicates the need for the development of robust software to address mixtures of microorganisms, such as the one found in petroleum.

Supplementary Information

Supplementary information is available free of charge at <http://jbcns.sbq.org.br> as PDF file.

Acknowledgements

We are grateful to Petrobras, CAPES and FAPESP for financial support.

References

1. Peters, K.; Moldowan, M.; *The Biomarker Guide. Interpreting Molecular Fossils in Petroleum and Ancient Sediments*; Prentice-Hall: Englewood Cliffs, 1993.
2. Head, I. M.; Jones, M.; Larter, S. R.; *Nature* **2003**, *426*, 344.
3. Cruz, G. F.; Vasconcellos, S. P.; Angolini, C. F. F.; Dellagnezze, B. M.; Garcia, I. N. S.; Oliveira, V. M.; Neto, E. V. S.; Marsaioli, A.; *AMB Express* **2011**, *1*, 47.
4. Mukred, A. M.; Hamid, A. A.; Hamzah, A.; Yusoff, W. M. W.; *OnLine J. Biol. Sci.* **2008**, *8*, 73.
5. Sauer, S.; Freiwald, A.; Maier, T.; Kube, M.; Reinhardt, R.; Kostrzewa, M.; Geider, K.; *PLoS One* **2008**, *3*, 1.
6. Janda, J. M.; Abbott, S. L.; *J. Clin. Microbiol.* **2007**, *45*, 2761.

7. Claydon, M. A.; Davcy, S. N.; Edward-Jones, V.; Gordon, D. B.; *Nat. Biotechnol.* **1996**, *14*, 1584.
8. Demirev, P. A.; Fenselau, C.; *Annu. Rev. Anal. Chem.* **2008**, *1*, 71.
9. Lovecka, P.; Pacovska, I.; Stursa, P.; Vrchotova, B.; Kochankova, L.; Demnerova, K.; *New Biotechnol.* **2014**, *32*, 26.
10. Han, S.-K.; Hong, Y.; Kwak, H.-L.; Kim, E.-S.; Kim, M.-J.; Srivastav, A.; Oh, M.-H.; Kim, H.-Y.; *J. Food Saf.* **2014**, *34*, 224.
11. Barberis, C.; Almuzara, M.; Join-Lambert, O.; Ramirez, M. S.; Famiglietti, A.; Vay, C.; *PLoS One* **2014**, *9*, e106303.
12. Satomi, M.; La Duc, M. T.; Venkateswaran, K.; *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 1735.
13. Cruz, G. F.; Neto, E. V. S.; Marsaioli, A.; *Org. Geochem.* **2008**, *39*, 1204.
14. Pitcher, D. G.; Saunders, N. A.; Owen, R. J.; *Appl. Microbiol.* **1989**, *8*, 151.
15. Yamada, S.; Ohashi, E.; Agata, N.; Venkateswaran, K.; *Appl. Environ. Microbiol.* **1999**, *65*, 1483.
16. Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F.; Higgins, D. G.; *Nucleic Acids Res.* **1997**, *24*, 4876.
17. Kumar, S.; Tamura, K.; Jakobsen, I. B.; Nei, M.; *Bioinformatics* **2001**, *17*, 1244.
18. Kimura, M.; *J. Mol. Biol.* **1980**, *16*, 111.
19. Saitou, N.; Nei, M.; *Mol. Biol. Evol.* **1987**, *4*, 406.
20. Vasconcellos, S. P.; Crespim, E.; Cruz, G. F.; Senatore, D. B.; Simioni, K. C. M.; Neto, E. V. S.; Marsaioli, A. J.; Oliveira, V. M.; *Org. Geochem.* **2009**, *40*, 574.

Submitted: August 17, 2014

Published online: January 20, 2015

FAPESP has sponsored the publication of this article.