Chapter 4

Results

Results

4.1 Phylogeny of microbial communities from crude oil-polluted soil according to DGGE profiles

DGGE with the 16S rDNA PCR product from soil samples 1-8 (Table 1) resulted in a gel displaying a denser (closer spaced and higher number) banded fingerprint pattern with a higher colour intensity in unpolluted soil (lanes 5-7) than in the diesel-polluted soils (lanes1-4 and 8) (Fig. 6).

Each band on the gel is assumed to be representative of only one distinct species, which was proved upon sequencing of the band. The gel also revealed a decrease in microbial diversity, i.e. number of bands, in subsoil layers (lanes 2 and 8) in comparison with topsoil. This was to be expected since these two samples should have very low species diversity as few organisms are present 1m and 1.5m deep in soil (Zhou *et al.* 2002). Of the polluted topsoil samples (lanes 1, 3 and 4), lane 3 DNA, extracted from the rhizosphere of plants growing in the soil, showed the greatest array of bands. This could have been due to plant root exudates enriching the soil around the roots, thus providing a nutrient boost within the polluted soil for microbes growing in the immediate vicinity of the roots. However, lane 4, which was from barren soil, displayed only slightly fewer bands than lane 3.

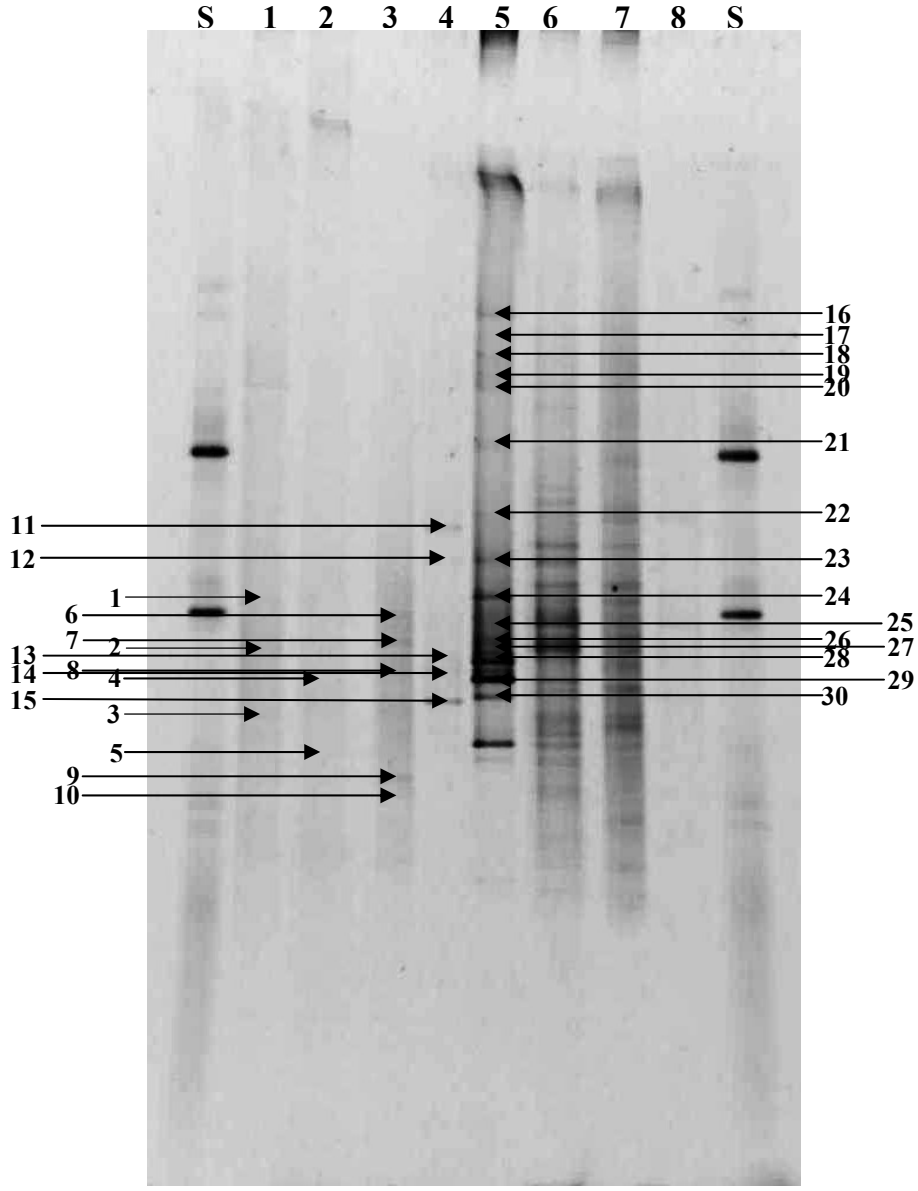
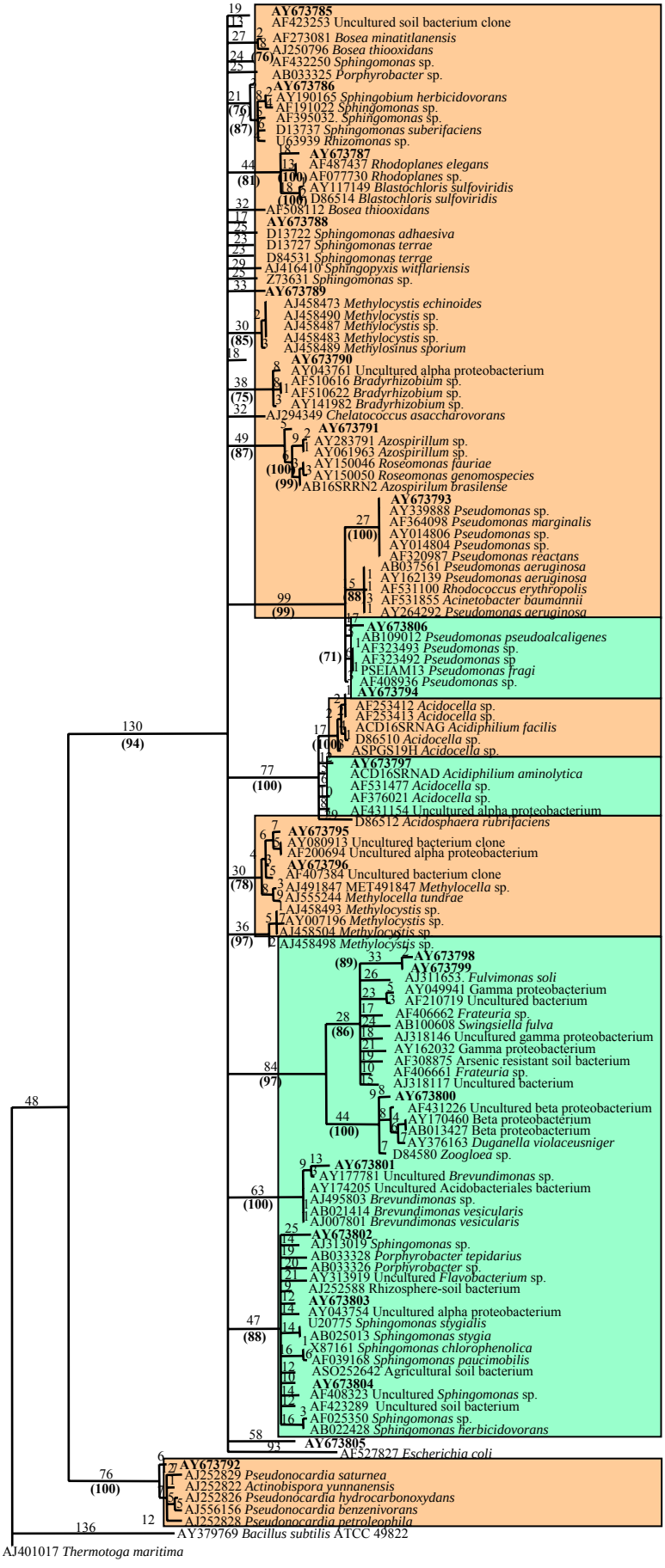


Figure 6: Denaturing gradient gel fingerprints resulting from assessing bacterial diversity between communities isolated from site 1 in Table 1, using a gradient of 15-55%. 16S rDNA PCR product separated according to base-pair sequence differences as a result of the denaturing gradient. Lanes 1-8 correspond with samples 1-8 in Table 1. Sequences 1-3 were taken from lane 1, 4-5 from lane 2, 6-10 from lane 3, 11-15 from lane 4 and 16-30 from lane 5. S represents the *Escherichia coli* standard used.

Distance (data not shown) and parsimony trees generated from sequence data displayed almost identical topography, within all of which distinct clade differentiation between polluted and unpolluted soil sequences was evident (Fig. 7). This is supported by the data

matrix comparing banding with lanes on the DGGE gel (Table 4), although, the matrix indicated that there were no sequences exclusively obtained from polluted soils. However several species were found solely in unpolluted soils viz. 16, 17, 19, 21, 22, 24, 26-30 (Table 4). Fig. 7 contains the sequences obtained from the gel with their corresponding BLAST results. Characters for this analysis were re-weighted to the consistency index, and only informative characters were included, missing and ambiguous characters as well as constant characters were excluded. The number of taxa included in this analysis was 134 thus making the required number of informative characters 264. The 271 informative characters presented in this phylogenetic analysis as well as the relatively high RI value obtained indicates significant signal within this data set. Sequences with e-values closest to 0.0 and displaying best alignments (data not shown) were compared with their most plausible matching organisms for tentative identification (Table 5). According to GenBank identification, many of these organisms are known to occur in hydrocarbon-polluted soils, e.g. *Sphingomonas adhaesiva*, *Sphingomonas terrae*, *Sphingopyxis witflariensis*, *Sphingomonas* sp., *Methylocystis* sp., *Pseudomonas* sp., *Pseudomonas marginalis*, *Acidocella* sp. and *Acidiphilium facilis*. The topography of the tree supports the tentative molecular identification of the unknown species isolated from polluted and unpolluted soils. The tentative identification of these species constitutes the first report of these microbes from polluted soils in South Africa.

Phylogenetic analysis of the sequences obtained from polluted soils (Fig. 8) showed a lower species diversity than those from unpolluted soils (Fig. 9). Species analysed from polluted soils appeared to be more closely related to each other as they grouped into one main clade, with AY673792 basal to that clade. Those from unpolluted soils showed a higher species diversity, i.e. genetically more distinct, as they grouped into three main clades, with AY673801 and AY673797 falling basal to these.



Parsimony informative characters 271
CI 0.231
RI 0.745
Number of trees 1
Tree length 2502
*g*1 -0.179919

Figure 7: Phylogeny of bacterial 16S rRNA gene amplicons recovered from the DGGE gel in Fig. 1 from soils collected at site 1 (Table 1). Sequences from polluted and unpolluted soils are indicated in orange and green areas, respectively (Distance values are indicated above branches and bootstrap values are indicated in brackets below).

Table 4: Data matrix of the banding pattern displayed on the DGGE gel (Fig. 6), showing the bands present in each lane and the identity of the most likely bacterial species that the band represents.

Sequence number	Lane number[g]								Most probable species
	1	2	3	4	5	6	7	8	
1	x		x	x	x	x	x	x	Uncultured soil bacterium clone
2	x	x	x	x	x	x	x		*Sphingomonas* sp.
3	x		x		x	x	x		*Sphingomonas* sp.
4		x		x	x	x	x	x	*Blastochloris sulfoviridis*
5		x		x	x	x	x		*Methylocystis* sp.
6	x		x		x	x	x	x	No identification available
7	x		x		x	x	x	x	*Bradyrhizobium* sp.
8	x		x		x	x	x		*Azospirillum* sp.
9			x			x	x		*Pseudonocardia saturnea*
10			x			x	x		*Pseudomonas* sp.
11				x	x	x	x	x	No identification available
12				x	x	x	x		*Acidocella* sp.
13			x	x	x	x	x	x	Uncultured alpha Proteobacterium
14	x			x	x	x	x	x	No identification available
15				x	x	x	x		*Methylocella* sp.
16					x	x	x		No identification available
17					x	x	x		*Acidocella* sp.
18	x				x	x	x		*Fulvimonas soli*
19					x	x	x		*Frateuria* sp.
20	x				x	x	x		*Duganella violaceusniger*
21					x	x	x		No identification available
22					x	x	x		*Brevundimonas* sp.
23	x				x	x	x		*Sphingomonas* sp.
24					x	x	x		*Sphingomonas* sp.
25			x		x	x	x	x	*Sphingomonas* sp.
26					x	x	x		No identification available
27					x	x	x		No identification available
28					x	x	x		*Pseudomonas aeruginosa*
29					x	x	x		*Pseudomonas pseudoalcaligenes*
30					x	x	x		No identification available

[g] Lanes from polluted and unpolluted soils are indicated in orange and green respectively.

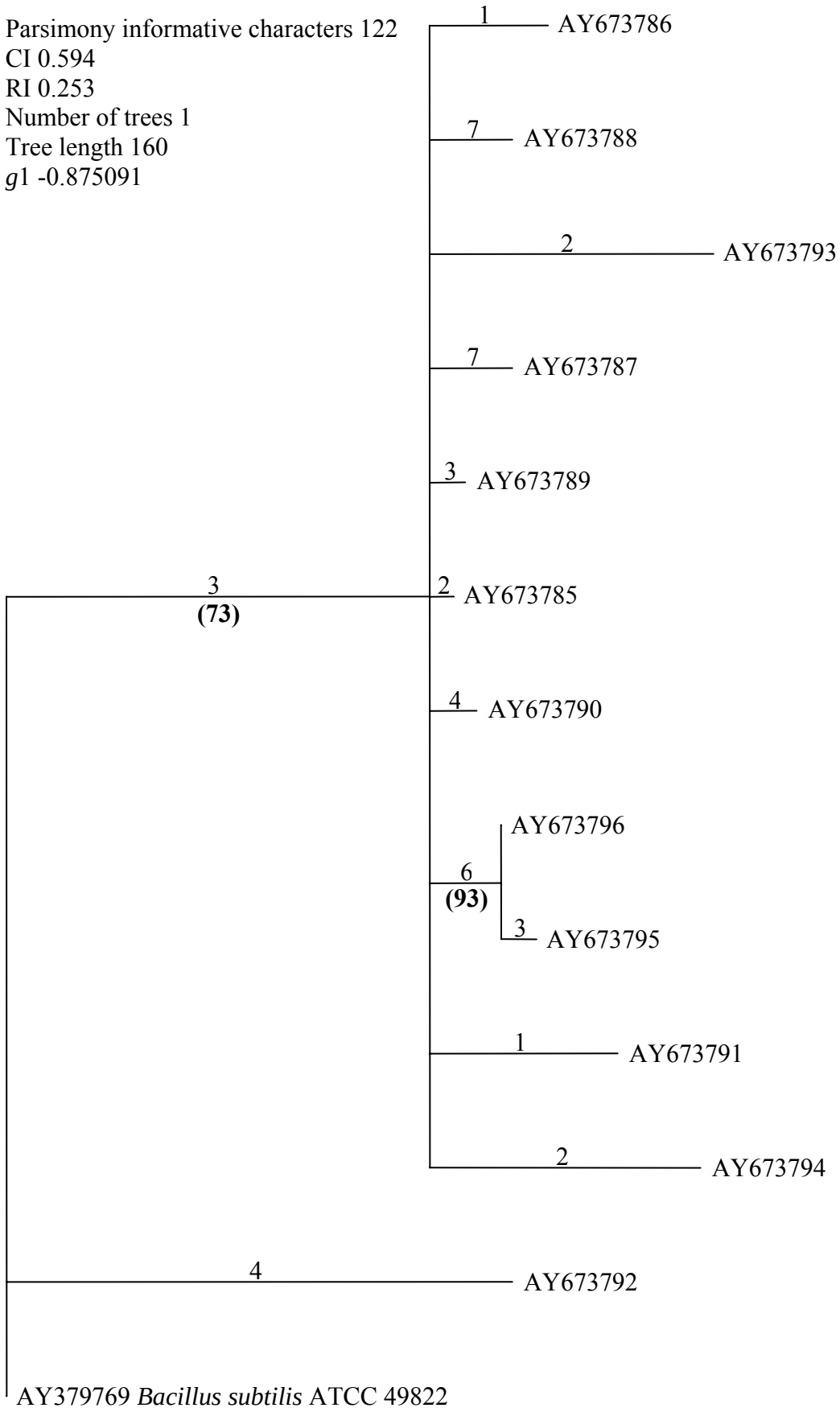


Figure 8: Phylogeny of the 16S eubacterial gene of rDNA from bands picked from the DGGE fingerprint in Fig. 1 of polluted soils sampled at site 1 (Table 1) (Distance values are indicated above branches and bootstrap values are indicated in brackets below).

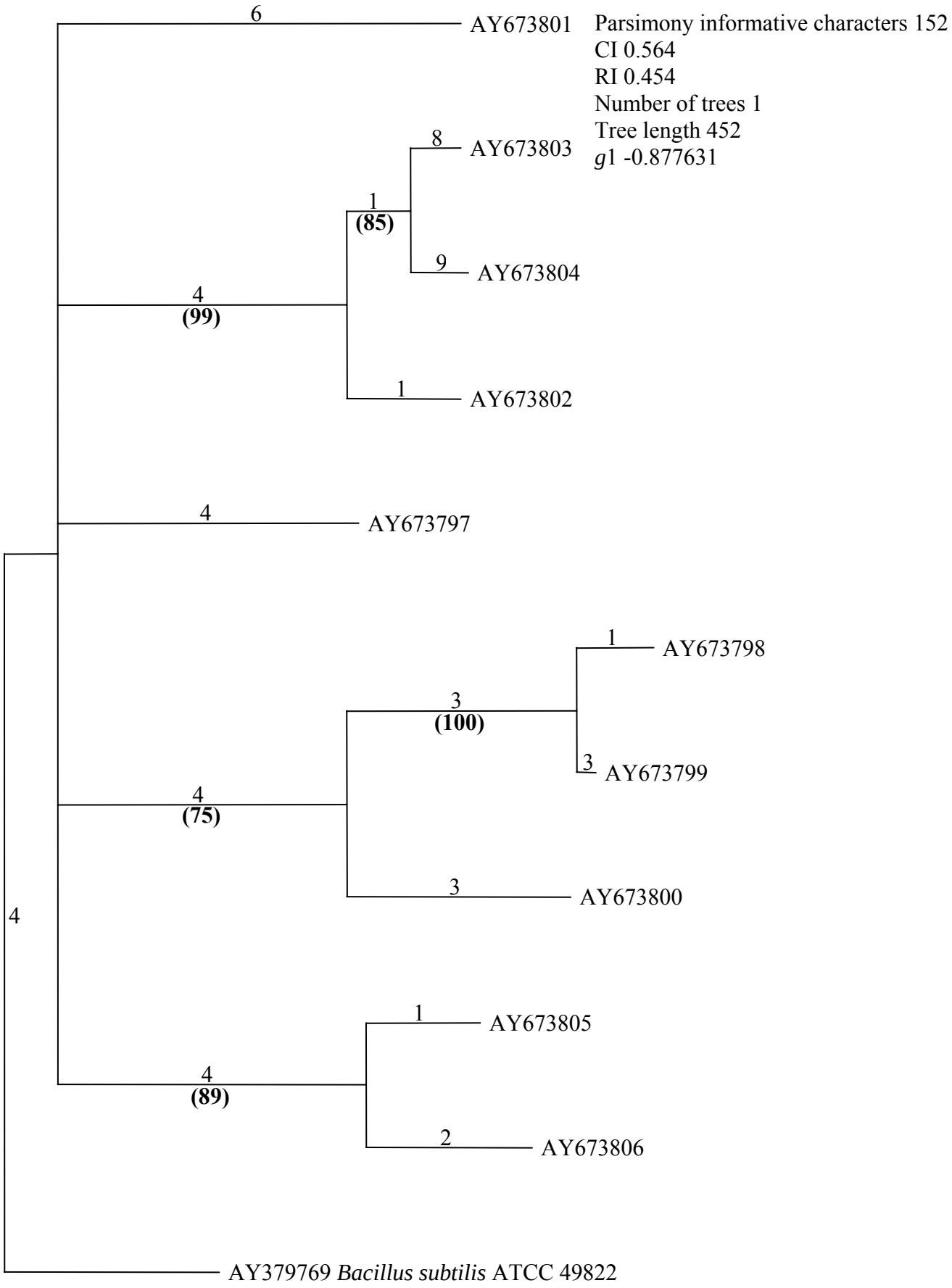


Figure 9: Phylogeny of the 16S eubacterial gene of rDNA from bands picked from the DGGE fingerprint in Fig. 1 of unpolluted soils sampled at site 1 (Table 1) (Distance values are indicated above branches and bootstrap values are indicated in brackets below).

Table 5: 16S eubacterial rDNA sequences from bands picked from the DGGE fingerprint in Fig. 1 of soils sampled at site 1 (Table 1). Bacterial species having the closest possible e-value for identification purposes and their descriptions are presented.

Deposited GenBank sequence number	DGGE gel band sequence number	Matching GenBank accession numbers	Closest species Identification	e-value	Description
AY673785	1	AF423253	Uncultured soil bacterium clone	-163	Bacteria; environmental samples.
AY673786	2	AF395032	*Sphingomonas* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingomonas*; from Lake Vostok accretion ice.
		AY190165	*Sphingobium herbicidovorans*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingobium*; from environmental samples.
		D13737	*Sphingomonas suberifaciens*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingomonas*; from corky root of lettuce.
		U63939	*Rhizomonas* sp.	0.0	Bacteria; Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingomonas*; from marine bacterioplankton.
AY673787	4	D86514	*Blastochloris sulfoviridis*	-124	Bacteria: Proteobacteria Alphaproteobacteria Rhizobiales Hyphomicrobiaceae *Blastochloris*.
AY673788	3	D13722	*Sphingomonas adhaesiva*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingomonas*; polyethylene glycol-utilising bacteria.
		D13727	*Sphingomonas terrae*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingopyxis*; polyethylene glycol-utilising bacteria.

Table 5 (continued)

Deposited GenBank sequence number	DGGE gel band sequence number	Matching GenBank accession numbers	Closest species Identification	e-value	Description
		D84531	*Sphingomonas terrae*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingopyxis*; from ears of plants in Gramineae.
		AJ416410	*Sphingopyxis witflariensis*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingopyxis*; from activated sludge.
		Z73631	*Sphingomonas* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingomonas*; marine ultramicrobacterium.
AY673789	5	AJ458490	*Methylocystis* sp.	-153	Bacteria: Proteobacteria Alphaproteobacteria Rhizobiales Methylocystaceae *Methylocystis*; type II methane-oxidising bacteria isolated from various environments.
AY673790	7	AY141982	*Bradyrhizobium* sp.	-122	Bacteria: Proteobacteria Alphaproteobacteria Rhizobiales Bradyrhizobiaceae *Bradyrhizobium*; rhizobia.
AY673791	8	AY283791	*Azospirillum* sp.	-99	Bacteria: Proteobacteria Alphaproteobacteria Rhodospirillales Rhodospirillaceae *Azospirillum*; from wheat rhizosphere.
AY673792	9	AJ252829	*Pseudonocardia saturnea*	-171	Bacteria: Actinobacteria Actinobacteridae Actinomycetales Pseudonocardineae Pseudonocardiaceae *Pseudonocardia*.
AY673793	10	AY339888	*Pseudomonas* sp.	0.0	Bacteria: Proteobacteria Gammaproteobacteria Pseudomonadales Pseudomonadaceae *Pseudomonas*; nitrobenzene biodegradation.
		AF364098	*Pseudomonas marginalis*	0.0	Bacteria: Proteobacteria Gammaproteobacteria Pseudomonadales Pseudomonadaceae *Pseudomonas*; cold tolerant; from carrot.
		AF320987	*Pseudomonas reactans*	0.0	Bacteria: Proteobacteria Gammaproteobacteria Pseudomonadales Pseudomonadaceae *Pseudomonas*; brown blotch disease of *Agaricus bisporus*.
AY673794	12	AF253412	*Acidocella* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Rhodospirillales Acetobacteraceae *Acidocella*; aromatic degradation.
		D30774	*Acidiphilium facilis*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Rhodospirillales Acetobacteraceae *Acidocella*.

Table 5 (continued)

Deposited GenBank sequence number	DGGE gel band sequence number	Matching GenBank accession numbers	Closest species Identification	e-value	Description
AY673795	13	AF200694	Uncultured alpha Proteobacterium	0.0	Bacteria: Proteobacteria Alphaproteobacteria; from environmental samples.
		AY080913	Uncultured bacterium clone	0.0	Bacteria; environmental samples; methylotroph communities in an acidic forest soil.
AY673796	15	AJ491847	*Methylocella* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Rhizobiales Beijerinckiaceae *Methylocella*; novel methanotroph isolated from an acidic forest cambisol.
		AJ555244	*Methylocella tundrae*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Rhizobiales Beijerinckiaceae *Methylocella*; novel methanotrophic bacterium from acidic tundra peatlands.
		AJ458498	*Methylocystis* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Rhizobiales Methylocystaceae *Methylocystis*; type II methane-oxidising bacteria isolated from various environments.
AY673797	17	D30771	*Acidiphilium aminolytica*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Rhodospirillales Acetobacteraceae *Acidocella*.
		AF376021	*Acidocella* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Rhodospirillales Acetobacteraceae *Acidocella*; microorganisms indigenous to acidic drainage waters at abandoned Norwegian copper mine.
AY673798	18	AJ311653	*Fulvimonas soli*	0.0	Bacteria: Proteobacteria Gammaproteobacteria Xanthomonadales Xanthomonadaceae *Fulvimonas*; from soil after enrichment on acetylated starch plastic.
		AF406662	*Frateuria* sp.	0.0	Bacteria: Proteobacteria Gammaproteobacteria Xanthomonadales Xanthomonadaceae *Frateuria*; from potato.
AY673799	19	AY162032	Gamma Proteobacterium	0.0	Bacteria: Proteobacteria Gammaproteobacteria.
		AF406661	*Frateuria* sp.	-175	Bacteria: Proteobacteria Gammaproteobacteria Xanthomonadales Xanthomonadaceae *Frateuria*; from potato.

Table 5 (continued)

Deposited GenBank sequence number	DGGE gel band sequence number	Matching GenBank accession numbers	Closest species Identification	e-value	Description
AY673800	20	AF431226	Uncultured beta Proteobacterium	-110	Bacteria: Proteobacteria Betaproteobacteria; environmental samples; lodgepole pine (*Pinus contorta*) rhizosphere soils from British Columbia forest soils.
		AY376163	*Duganella violaceusniger*	-106	Bacteria: Proteobacteria Betaproteobacteria Burkholderiales Oxalobacteraceae *Duganella*.
AY673801	22	AY177781	Uncultured *Brevundimonas* sp.	-129	Bacteria: Proteobacteria Alphaproteobacteria Caulobacterales Caulobacteraceae *Brevundimonas*; environmental samples; soil bacterial communities in California grassland.
AY673802	23	AJ313019	*Sphingomonas* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingomonas*; from hexane degrading biofilters.
		AB033328	*Porphyrobacter tepidarius*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Porphyrobacter*.
		AB033326	*Porphyrobacter* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Porphyrobacter*.
		AY313919	Uncultured *Flavobacterium* sp.	0.0	Bacteria: Bacteroidetes Flavobacteria Flavobacteriales Flavobacteriaceae *Flavobacterium* environmental samples; from Tibetan plateau deep glacier ice.
		AJ252588	Rhizosphere soil bacterium	0.0	Bacteria
AY673803	24	U20775	*Sphingomonas stygialis*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae Novosphingobium; aromatic-degrading *Sphingomonas* isolates from deep soil subsurface.
		AB025013	*Sphingomonas stygia*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Novosphingobium*; from deep-sea sediments.
		X87161	*Sphingomonas chlorophenolica*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingobium*; pentachlorophenol-degrading.

Table 5 (continued)

Deposited GenBank sequence number	DGGE gel band sequence number	Matching GenBank accession numbers	Closest species Identification	e-value	Description
		ASO252642	Agricultural soil bacterium	0.0	Bacteria; environmental samples.
		AY043754	Uncultured alpha Proteobacterium	0.0	Bacteria: Proteobacteria Alphaproteobacteria; from British Columbia forest soils subjected to disturbance.
AY673804	25	AF039168	*Sphingomonas paucimobilis*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingomonas*; lindane-degrading bacterium.
		AF025350	*Sphingomonas* sp.	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingomonas*; dicamba-degrading bacterium.
		AB022428	*Sphingomonas herbicidovorans*	0.0	Bacteria: Proteobacteria Alphaproteobacteria Sphingomonadales Sphingomonadaceae *Sphingobium*.
AY673805	28	AB037561	*Pseudomonas aeruginosa*	-107	Bacteria: Proteobacteria Gammaproteobacteria Pseudomonadales Pseudomonadaceae *Pseudomonas*.
AY673806	29	AB109012	*Pseudomonas pseudoalcaligenes*	-138	Bacteria: Proteobacteria Gammaproteobacteria Pseudomonadales Pseudomonadaceae *Pseudomonas*; naphthalene-degrading bacteria from tar plant.

4.2 Comparative DGGE fingerprint analysis of pro- and eukaryotes in PAH-polluted soil

DNA was successfully extracted from samples 9-18 (Table 1). No protein contamination in the wells of the gel, or RNA contamination smears below the DNA bands, was evident (Fig. 10). PCR of 16S prokaryotic rDNA and ITS eukaryotic DNA yielded a *ca.* 550bp fragment of PCR product on the 1.5% TAE agarose gel (Fig. 11).

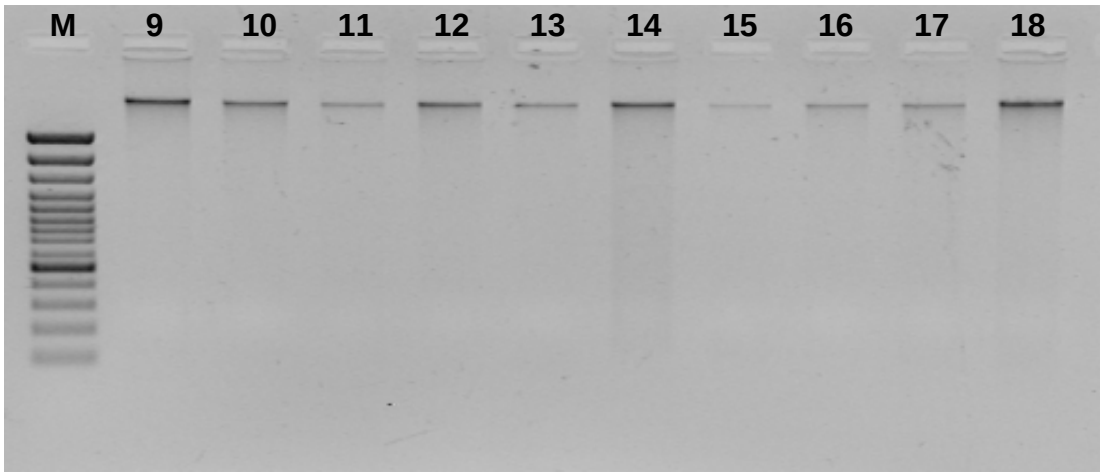


Figure 10: 1.5% TAE agarose gel showing high-quality, clean genomic DNA extracted from soil samples 9-18 at site 2 (Table 1) by means of the BIO101 Fast DNA Spin Kit for soil (M = 100bp marker, 9-18 = soil DNA from samples).

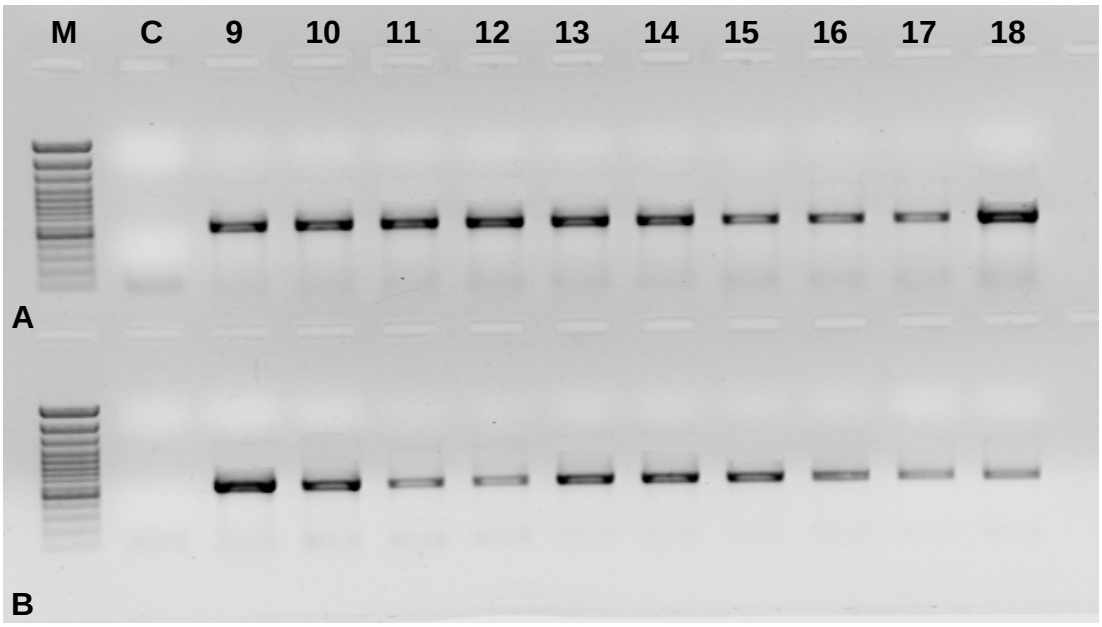


Figure 11: 1.5% TAE agarose gel showing (**A)** 16S rDNA PCR product of prokaryotic genomic DNA, and (**B)** partial internal transcribed spacer gene PCR product from eukaryotic genomic DNA, both *ca*. 550bp, from soil samples 9-18 at site 2 (Table 1) (M = 100bp marker, C = negative control, 9-18 = PCR product from samples).

DGGE yielded a gel showing clear multiple banding, forming a fingerprint in each lane (Fig. 12A and B). Due to the basis of DGGE, PCR fragments of the same size are separated into bands according to their sequence, the resulting fingerprint pattern being indicative of species diversity. From the gel a graphic cluster representation of the band pattern was drawn using Gel2K (Norland 2004). The programme does this by estimating band peak intensity along a lane. Peaks can be manipulated to ensure that, should more than one peak be registered per band, they can be grouped together. Dominant species per lane are indicated as dark prominent bands across the lane.

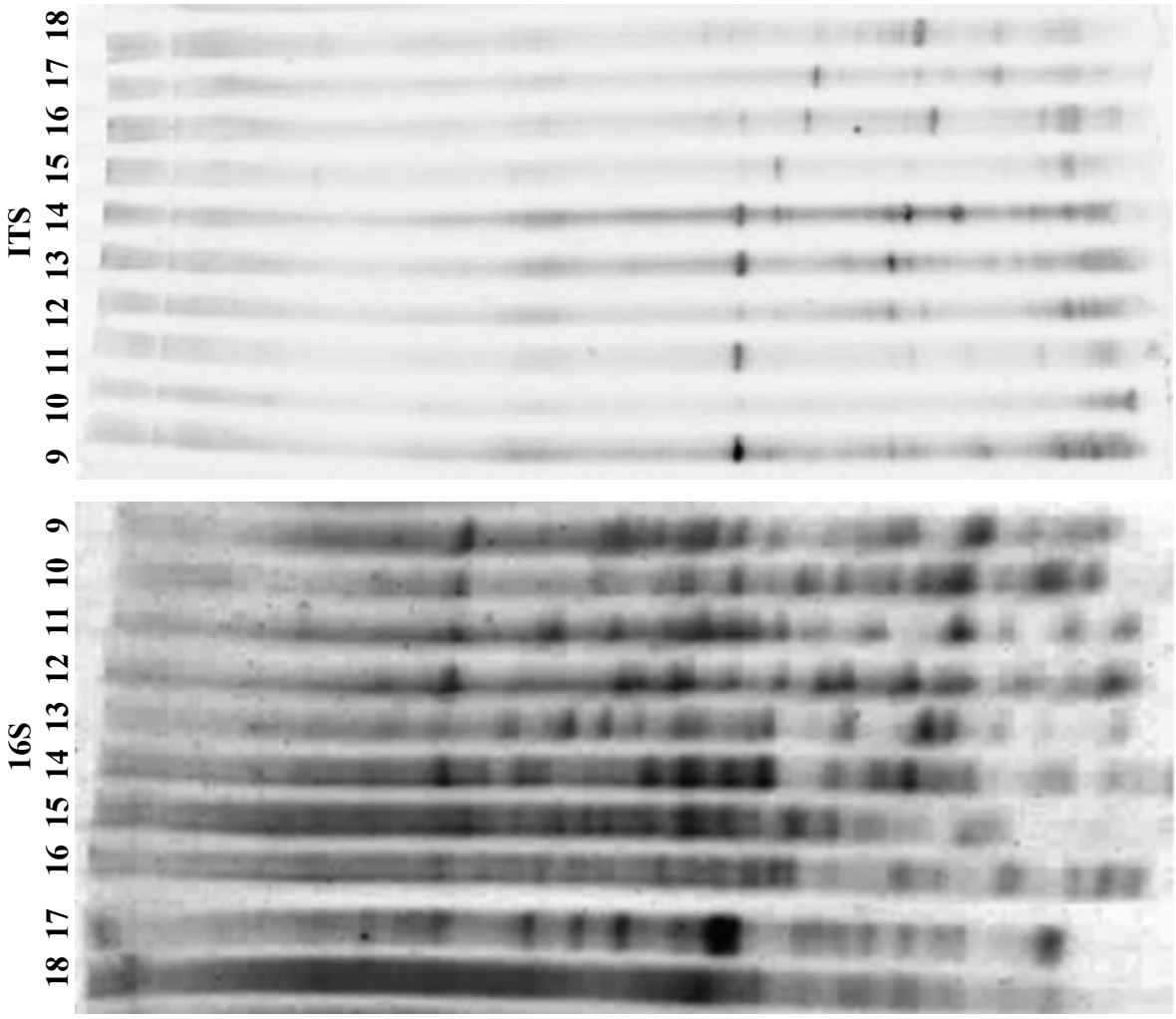


Figure 12A: DGGE gels, using a denaturant gradient of 15-55%. 16S rDNA (16S) and partial internal transcribed spacer sequence (ITS) PCR products separated according to base-pair sequence differences indicating community richness and diversity of pro- and eukaryotes in soils 9-18 (Table 1).

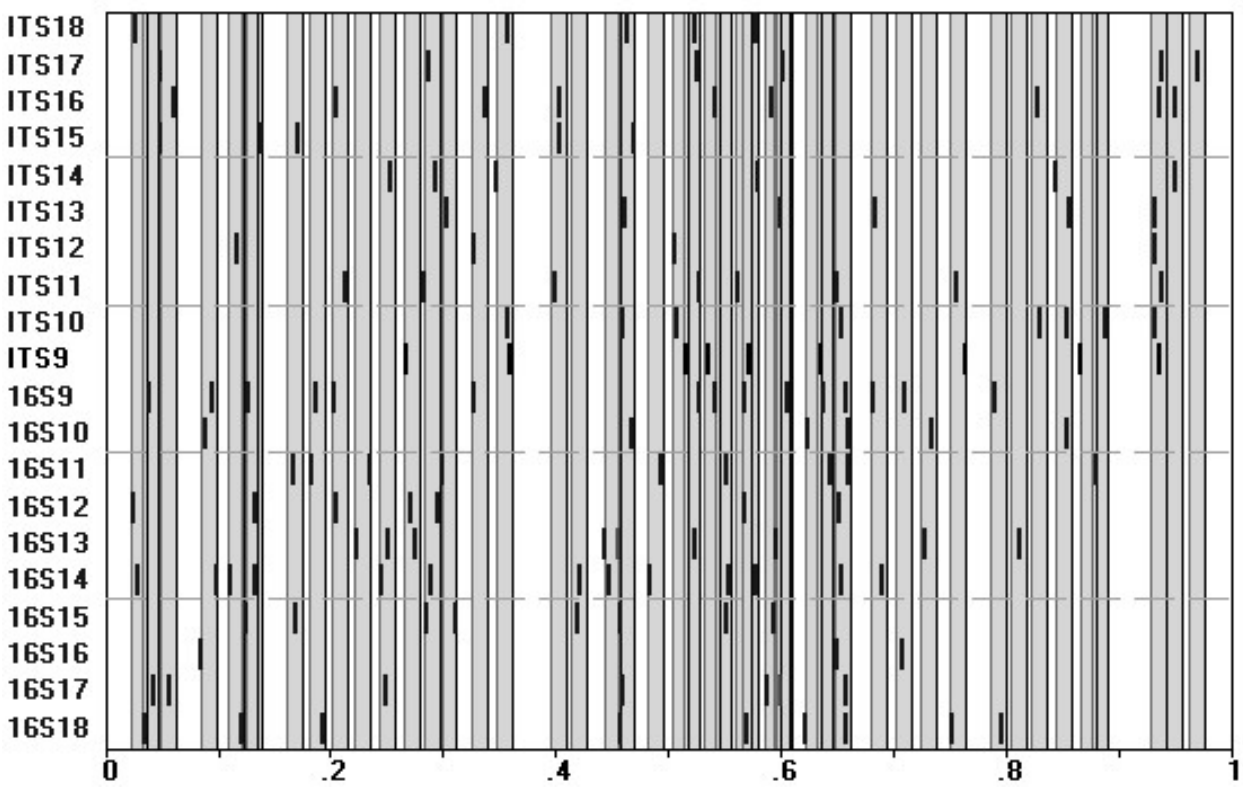


Figure 12B: Graphic representation of the DGGE gel, using a denaturant gradient of 15-55%. 16S rDNA (16S) and partial internal transcribed spacer sequence (ITS) PCR products separated according to base-pair sequence differences indicating community richness and diversity of pro- and eukaryotes in soils 9-18 (Table 1).

On average, more dominant bands were found for prokaryotes (mean 8.6 per sample) than for eukaryotes (mean 6.6 per sample) (Fig. 13), indicating that prokaryotic diversity was higher than that of eukaryotes. Species richness was also more evident within the prokaryotes, i.e. more dark/dominant species bands were observed within the prokaryotic gel lanes than in eukaryotic lanes.

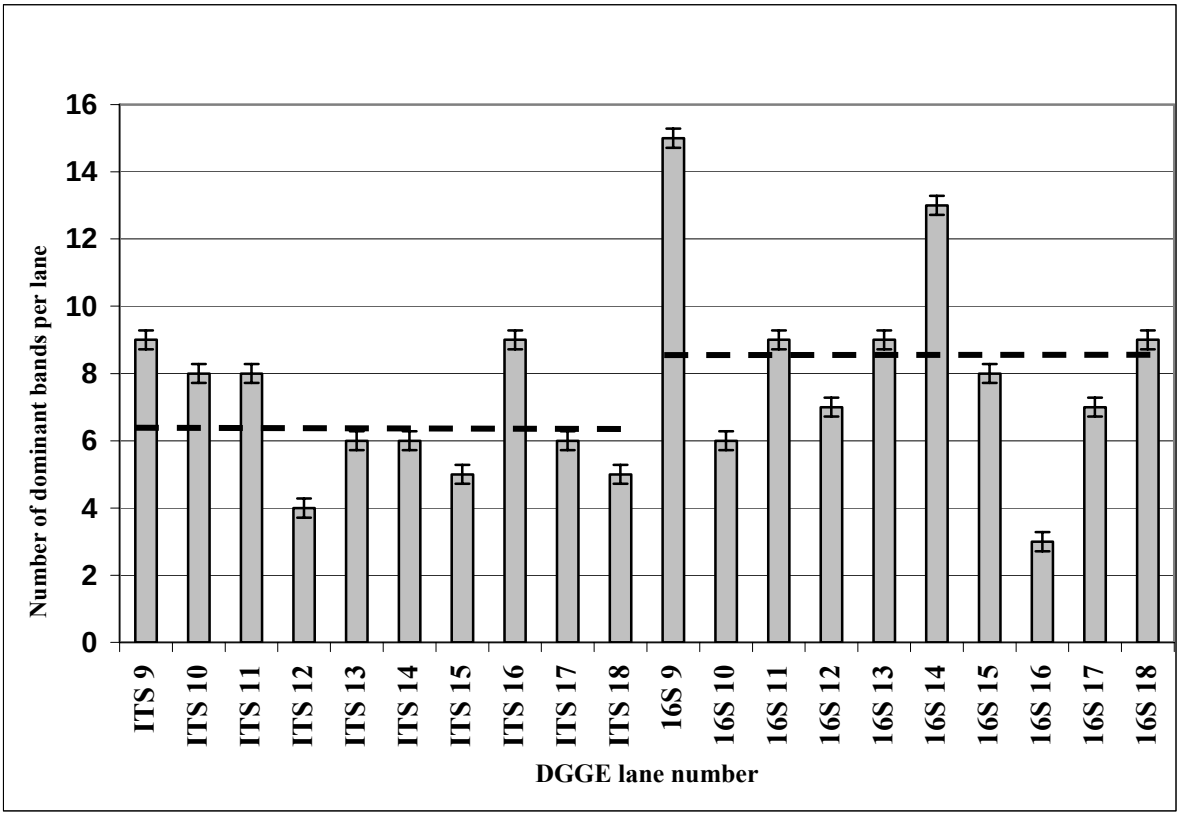


Figure 13: Number of dominant species per soil sample as indicated by DGGE of 16S rDNA (16S) and partial internal transcribed spacer sequence (ITS) PCR products. Note that the mean number of dominant species is 6.6 for eukaryotes and 8.6 for prokaryotes (dotted lines).

Following this, species diversity, and to certain extent species richness, were derived from the gel by compiling a dendogram (Fig. 14). The programme CLUST (Norland 2004) is based on Shannon index algorithms and groups the profiles of the species in each sample according to how similar in community composition the samples are. Thus, samples from similar environments would be expected to display analogous communities and group together in the CLUST dendogram.

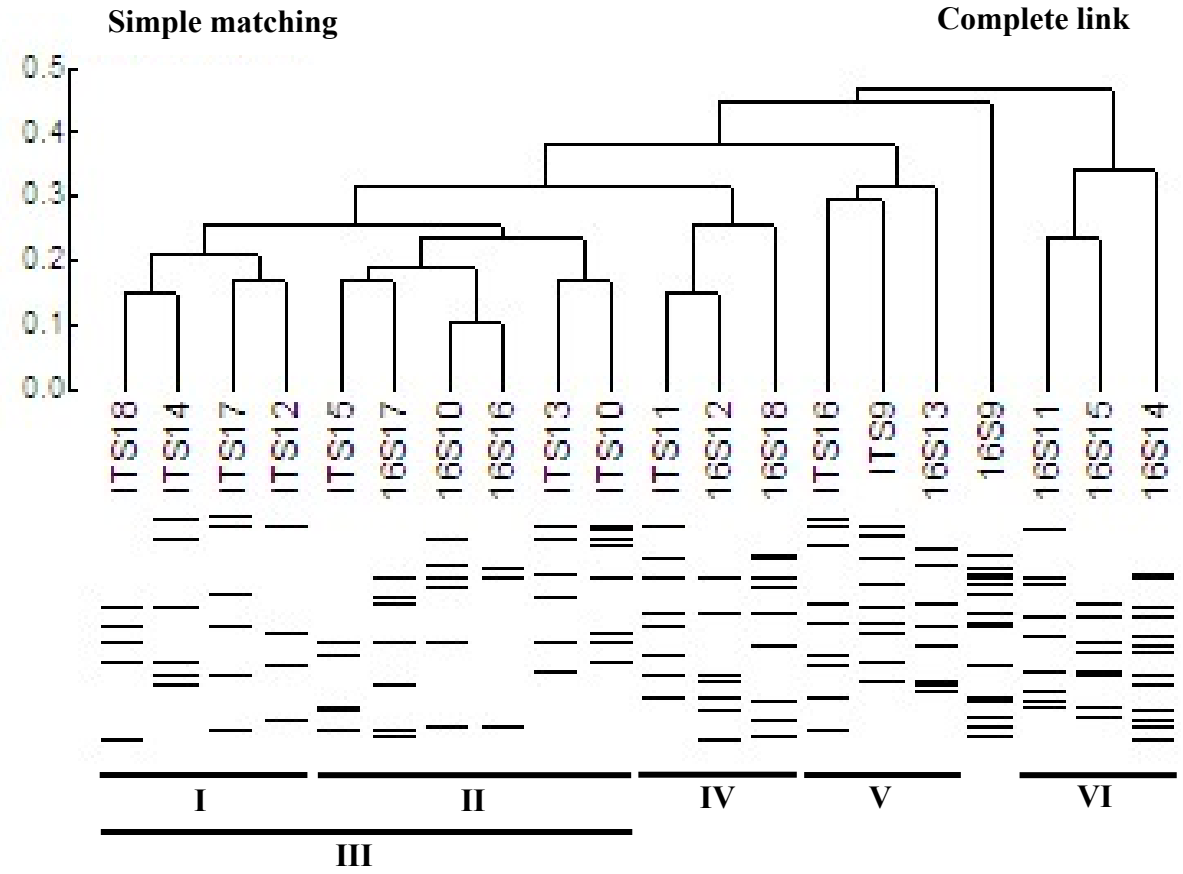


Figure 14: Cluster analysis of the banding pattern in Fig. 12, using a simple matching, complete link setting to separate communities according to species sequence differences (Thicker bands represent a brighter band on the gel, and therefore greater species richness).

4.3 DGGE analysis of toluene and naphthalene degraders in polluted soil

PCR of DNA extracted from samples 9-18 (Table 1, Fig. 11) with *ndoB* and *xylE* gene primers yielded a PCR fragment of *ca.* 650bp and 400bp, respectively on a 1.6% TBE agarose gel (data not shown). The DGGE gel showed clear multiple banding within each fingerprint in each lane (Fig. 15A and B). However, the gel had a great number of bands present and only dominant bands were therefore considered for diversity analysis. Control soil sample 9 harboured the most diverse *xylE* gene community, whereas the most diverse *ndoB* gene community was present in soils 17 and 18 polluted with machinery oil. On

average the gel showed a higher number of bands for the *xylE* gene (T) than for *ndoB* gene (N). The mean number of dominant species per soil sample was 13.6 among naphthalene degraders and 19.4 among toluene degraders. Dominant band numbers were compared across naphthalene and toluene degraders and, except for soil samples 17 and 18, all *xylE* gene products showed a considerably higher number of dominant species than *ndoB* gene products, possibly indicating a higher toluene than naphthalene degrading capacity across samples 9-16 (Fig. 16).

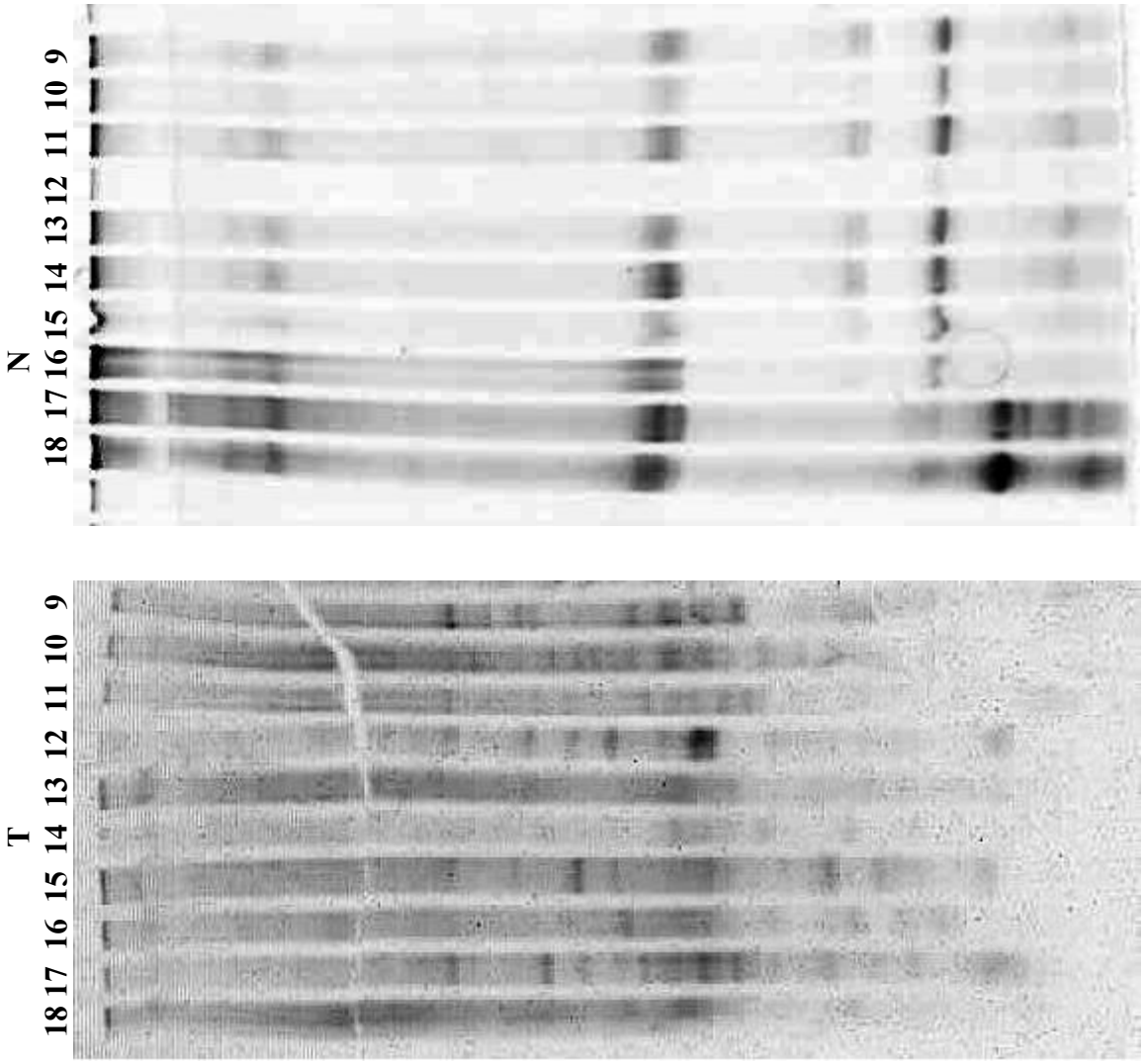


Figure 15A: DGGE gel, using a gradient of 30-65%, with *ndoB* gene (N) and *xylE* gene (T) PCR product amplified from DNA extracted from samples 9-18 (Table 1).

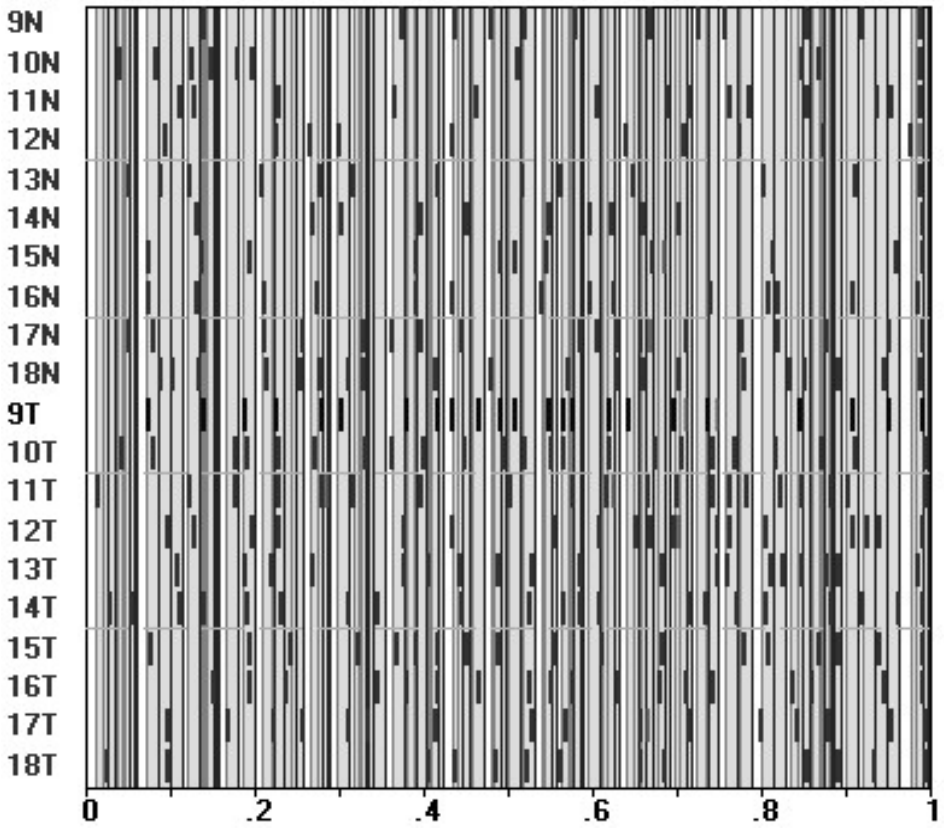


Figure 15B: Graphic representation of the DGGE gel, using a gradient of 30-65%, with *ndoB* gene (N) and *xylE* gene (T) PCR product amplified from DNA extracted from samples 9-18 (Table 1).

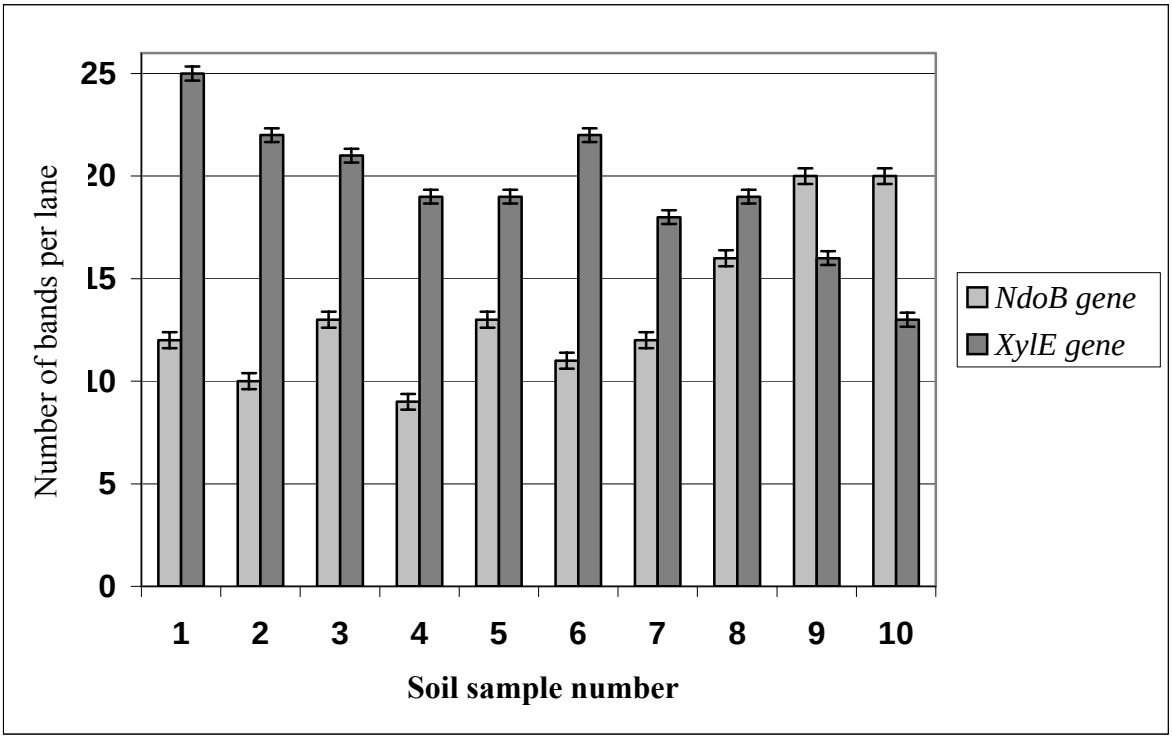


Figure 16: Number of dominant species per sample as indicated on the DGGE gel with *ndoB* gene and *xylE* gene PCR products from DNA extracted from samples 9-18 (Table 1) (Standard error for *ndoB* gene bands is 0.386 and 0.337 for *xylE* gene bands).

A combined dendogram of all communities across both *xylE* gene and *ndoB* gene diversity showed five main clades (Fig. 17). Clades I and V consisted mainly of *xylE* gene communities, while clade IV included mainly *ndoB* gene communities. Within clade IV there were two subclades (II and III), subclade II containing only *ndoB* gene communities and grouping three of the unpolluted soils and polluted soil 15.

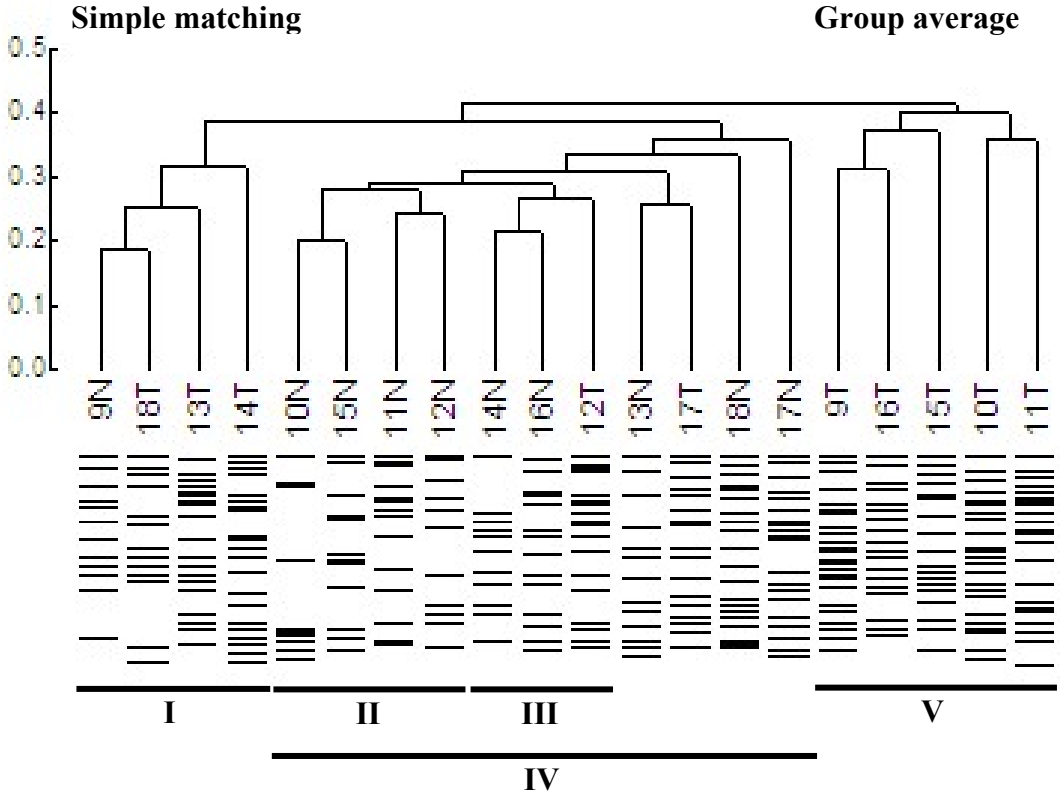


Figure 17: Cluster analysis of the banding pattern in Fig. 15 of *ndoB* and *xylE* gene PCR products, using a simple matching, group average setting to separate communities on the basis of species sequence differences from a multi-gene community (Thicker bands represent a brighter band on the gel, and therefore tentatively indicate a higher species richness).

The topography in Fig. 17 is also evident in clade I of the *ndoB* gene dendogram (Fig. 18). Within subclade III soils 14 and 16 grouped together as could be expected since they were both from rhizosphere soil (Table 1). The predominance of naphthalene degraders in clade IV indicates a similar genetic community profile across unpolluted and polluted soils. However, the majority of toluene degraders were split into two separate clades (I and V), those in clade I grouping with the naphthalene community from control soil 9.

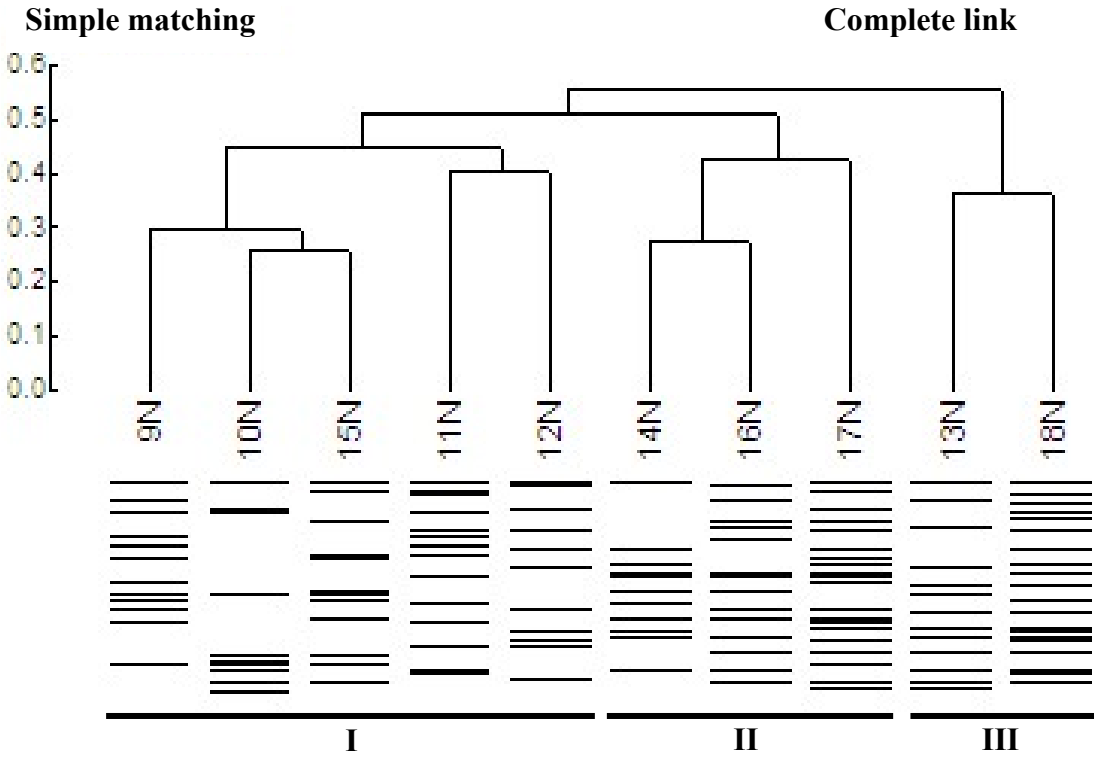


Figure 18: Cluster analysis of the banding pattern in Fig. 15 includes *ndoB* gene PCR products, using a simple matching, complete link setting to separate communities on the basis of species sequence differences from a single gene community (Thicker bands represent a brighter band on the gel, and therefore tentatively indicate a higher species richness).

The dendogram depicting *ndoB* gene diversity displayed three informative clades (Fig. 18). Clade I included mainly unpolluted soils grouping separately from polluted soils. Soils 11 and 12 clustered together as in clade II in Fig. 17, thus indicating that these soils contained very similar communities within the scope of the *ndoB* gene. Clades II and III contained polluted soils, with soils 14 and 16 grouping closer to each other than to soil 17, and soils 13 and 18 clustering together.

The final dendogram focusing on *xylE* gene diversity accommodated three informative

clades (Fig. 19). Clade I had soils 9, 13, 14, 17 and 18 assembling together. Except for soil 17 this is the same grouping found in clade I of Fig. 17, indicating a similar community profile. Clade II clustered soils 15 and 16 together, as was also evident in Fig. 17, clade V. Falling basal to the above clades was clade III containing the unpolluted soils 10, 11 and 12, hence corroborating the similarity of their community structures for *xylE* gene diversity.

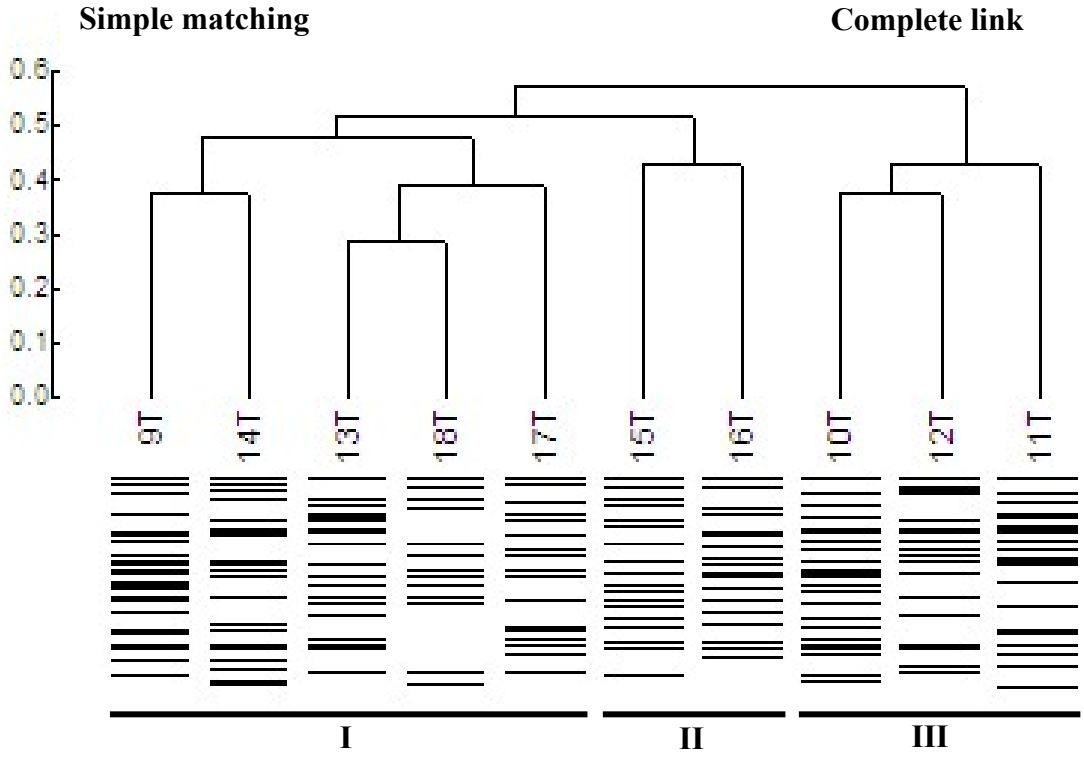


Figure 19: Cluster analysis of the banding pattern of the *xylE* gene PCR product in Fig. 15 using a complete link, simple matching setting to separate communities on the basis of species sequence differences from a single gene community (Thicker bands represent a brighter band on the gel, and therefore higher species richness).

4.4 Molecular identification of toluene and naphthalene degrading capacity of bacteria from PAH-polluted soil

All eight isolates yielded bands of *ca.* 400 (Fig. 20) and *ca.* 600bp when screened for the production of catechol 2,3-dioxygenase and naphthalene dioxygenase, respectively. The phylogenetic tree, containing these isolates, revealed two large clades rooted to *T. maritima* (Fig. 21). The larger of the two clades, containing various soil inhabitants and PAH degraders, housed six distinct sub-clades, viz. I (*Burkholderia*), II (*Sphingomonas*), III (*Pseudomonas*), IV (*Staphylococcus* and *Bacillus*), V (*Providencia*) and VI (*Burkholderia, Methylobacterium, Klebsiella, Rhodococcus* and *Pseudomonas*). Within this clade the genera *Burkholderia*, *Sphingomonas*, *Pseudomonas*, *Bacillus*, *Methylobacterium*, *Klebsiella* and *Rhodococcus* as well as *Vibrio* sp. are known to be involved in the metabolism of PAHs, particularly naphthalene and phenanthrene (Samanta *et al*. 2002; Kang *et al*. 2003).

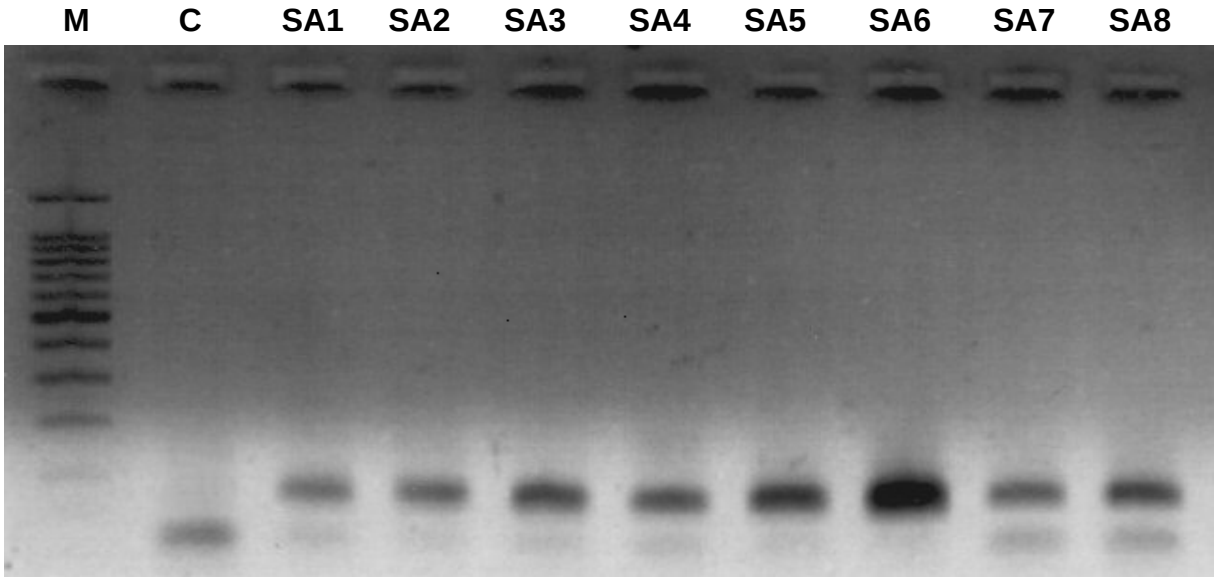


Figure 20: Bands of *ca.* 400bp on a 1.6% TBE gel showing production of catechol 2,3-dioxygenase (*xylE* gene) by bacterial isolates from a PAH-polluted soil in South Africa (M = λ marker, C = negative control, SA1-SA8 = isolates).

Five of the sequenced isolates (SA1, SA5, SA6, SA7 and SA8) clustered into the *Pseudomonas* sub-clade (III), one (SA4) into *Staphylococcus* (IV) and two (SA2 and SA3) into *Providencia* (V). None of the isolates clustered into the smaller of the two clades comprising mainly *Mycobacterium* species (Fig. 21). Most mycobacteria are capable of degrading PAHs but are not readily isolated and grow very slowly on artificial media (Prescott *et al*. 1999). There are, however, exceptions such as *M. tusciae*, *M. moriokaense*, *M. septicum* and *M. isoniacini*, which can grow rapidly in culture (Leys *et al.* 2005). The absence of *Mycobacterium* isolates in the present study was probably not due to the fact that they could not be isolated, but rather to the genus being very rare or not present in polluted soils in South Africa, as evident from a previous preliminary survey (Surridge *et al.* 2004).

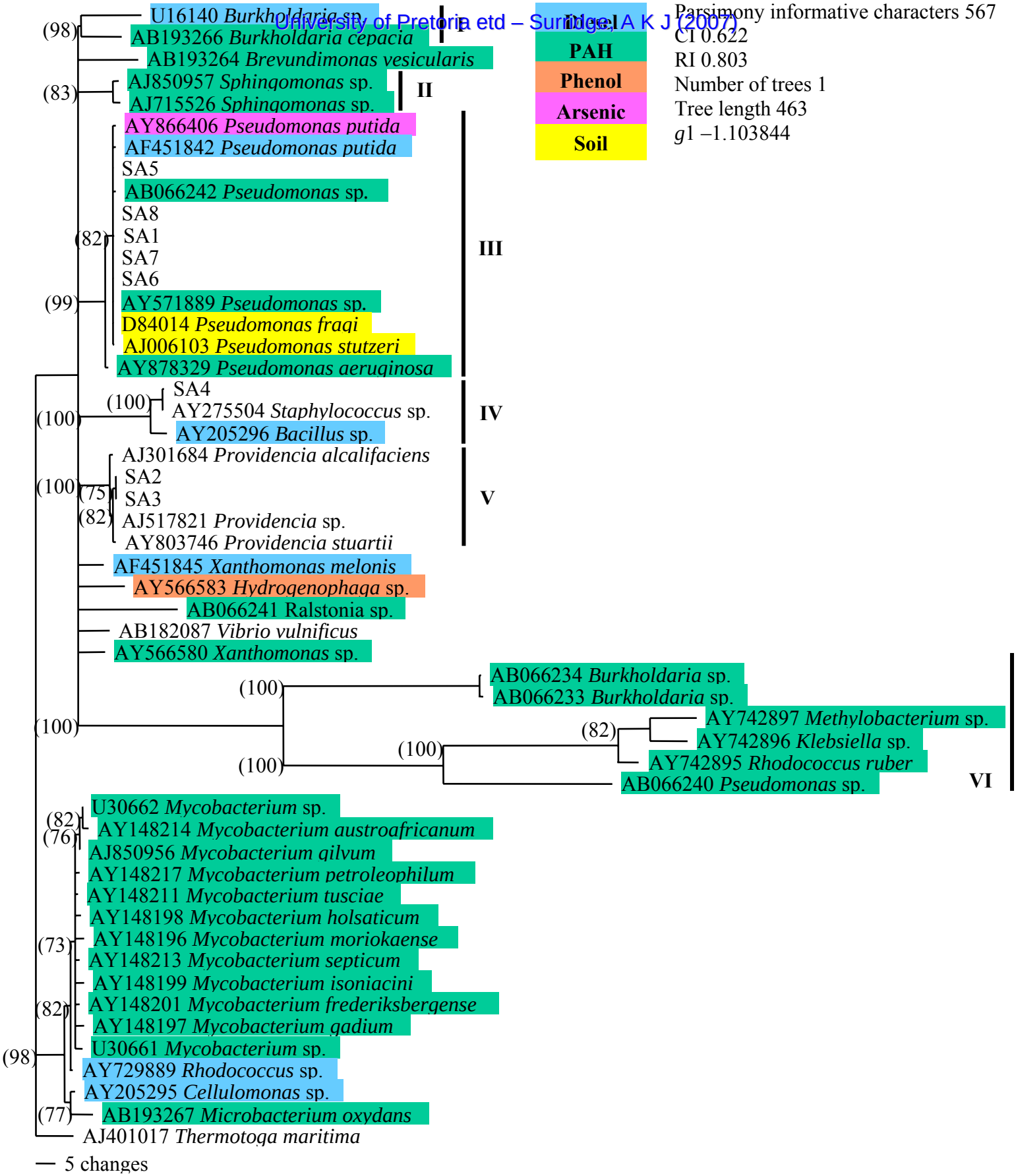


Figure 21: Phylogenetic tree (partial 16S eubacterial gene) of bacterial isolates (SA1-SA8) from a PAH-polluted soil in South Africa, their corresponding BLAST results and known PAH degraders.

4.5 DGGE community analysis of nitrogen-fixing bacteria in polluted soil

PCR was employed in a diagnostic capacity to determine the presence of nitrogen-fixing genes in the soil and pure culture DNA (Table 2). The presence of a PCR product was deemed a positive result. Positive results indicated that the *nifH* gene was present in the sample being tested. *nifH* nested-PCR performed on 10 soil and 8 pure culture (Fig. 22) DNA samples, yielded a *ca.* 370bp fragment on a 1.5% TAE agarose gel. All soil samples tested positive for the presence of *nifH* genes and thus nitrogen fixing capacity. Pure culture samples also tested positive for *nifH* genes, except for *nif*12 and *nif*18 (Fig. 22). *nif*12 was tentatively identified based on sequencing a portion of the 16S gene (4.4 above, sample SA2) as a member of the *Providencia* genus.

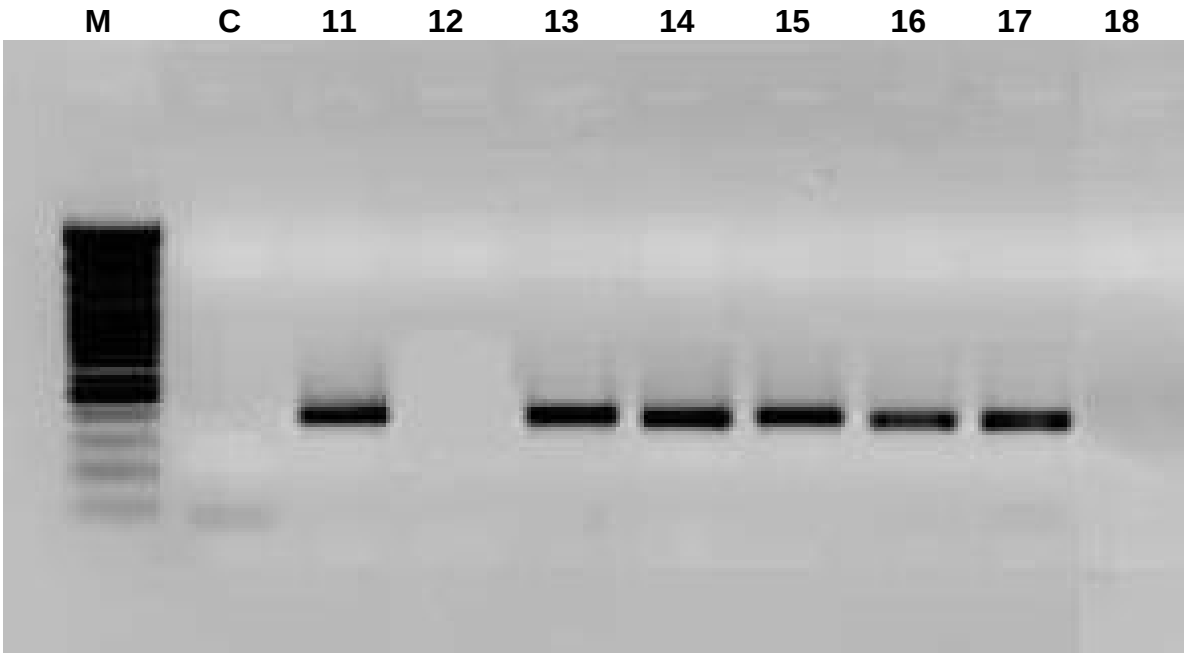


Figure 22: PCR product of *ca.* 370bp from nested-PCR using *nifH* degenerate primers (M = 100bp marker, C = negative control, 11-18 = *nif* PCR product of pure culture isolates (Table 2)).

DGGE fingerprint analysis revealed noteworthy species diversity, evident from a different dominant banding pattern within the *nifH* gene from the various soil samples screened (Fig. 23A and B). The mean number of dominant bands per lane for all soil samples tested was 7.2. Control soils sampled from unpolluted areas, as well as rhizosphere soils sampled from polluted areas, displayed on average a lower number of bands (mean 6.5 and 6.3 respectively) than the mean number per soil. All non-rhizosphere polluted soils displayed 8 or more bands (mean 9.0), indicating an above-average presence of nitrogen-fixing bacterial species.

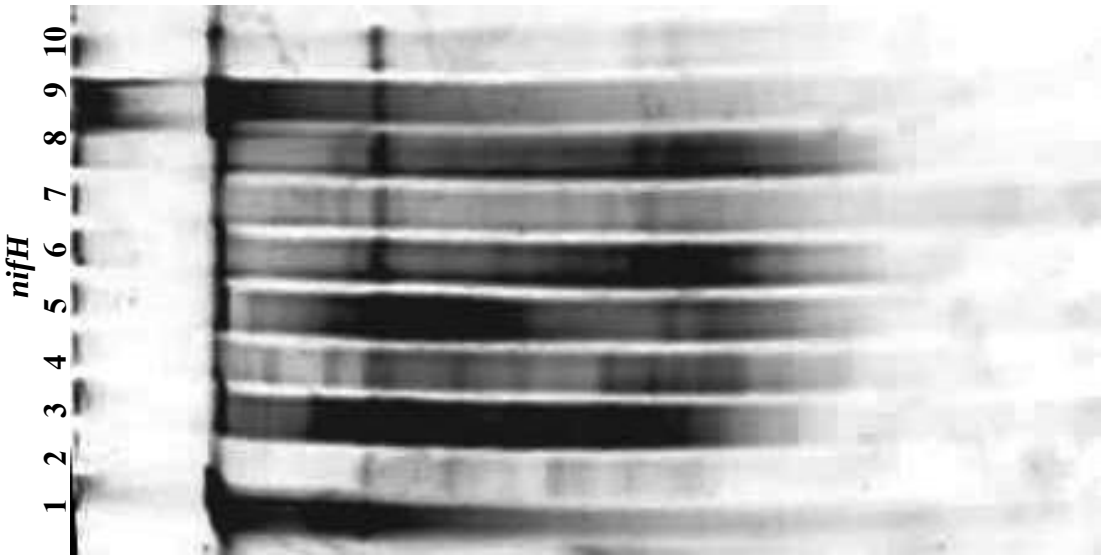


Figure 23A: DGGE gel of *nifH* PCR product from soil (*nifH* PCR product 1-10 Table 2) at 30-65% denaturant concentration.

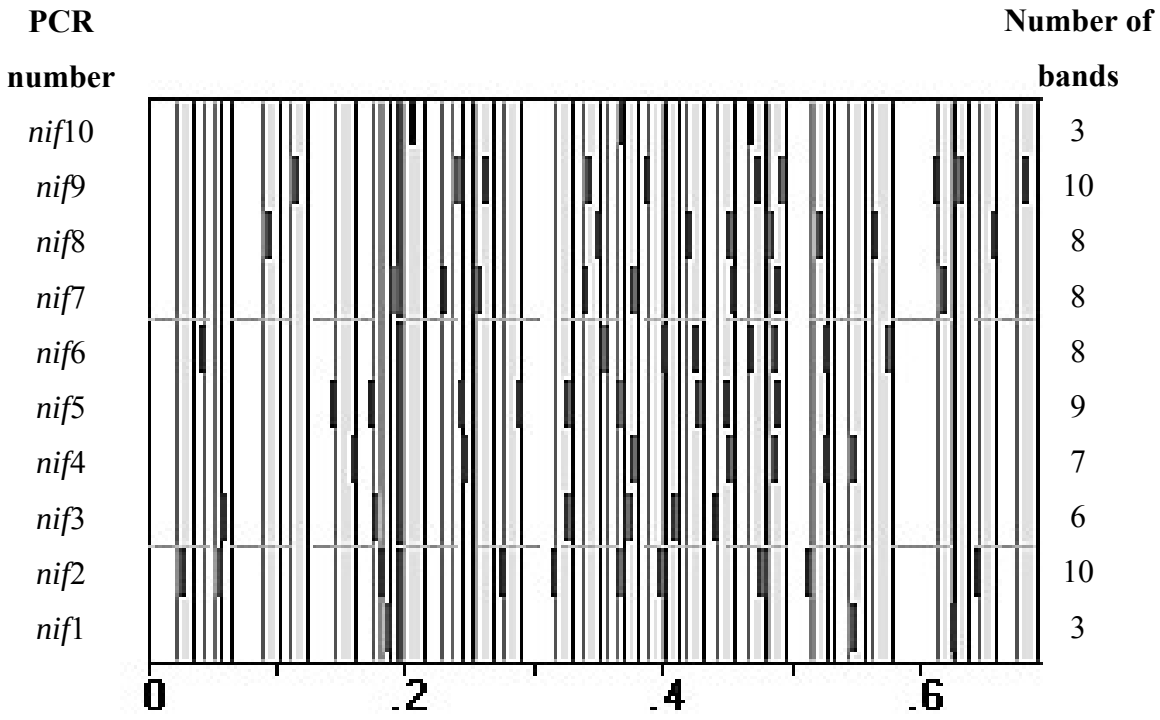


Figure 23B: Graphic representation of DGGE gel of *nifH* PCR product from soil at 30-65% denaturant concentration (Black bars represent dominant bands present in each sample).

Cluster analysis of the DGGE banding pattern revealed two main community sub-clades within the main cluster (I and II, Fig. 24), with *nif*2 and *nif*9 grouping outside. *nif*2 and *nif*9 displayed among the highest diversities of all the samples tested. *nif*9 also displayed the most dominant *nifH* gene presence as it had more intense banding. *nif*1, 3, 7 and 10 grouped in sub-clade I and *nif*4, 5, 6 and 8 in sub-clade II. Within sub-clade I, the community of *nif*3 and 10 from unpolluted and polluted *B. serratia* rhizosphere, respectively (Table 1), were the most similar. Also grouping in sub-clade I, but somewhat separated from *nif*3 and 10 were the *nif*1 (control) and *nif*7 (near *B. serratia* polluted rhizosphere). Sub-clade II included three samples (*nif*4, 5 and 6) from unpolluted and polluted soil under *C. esculentus*.

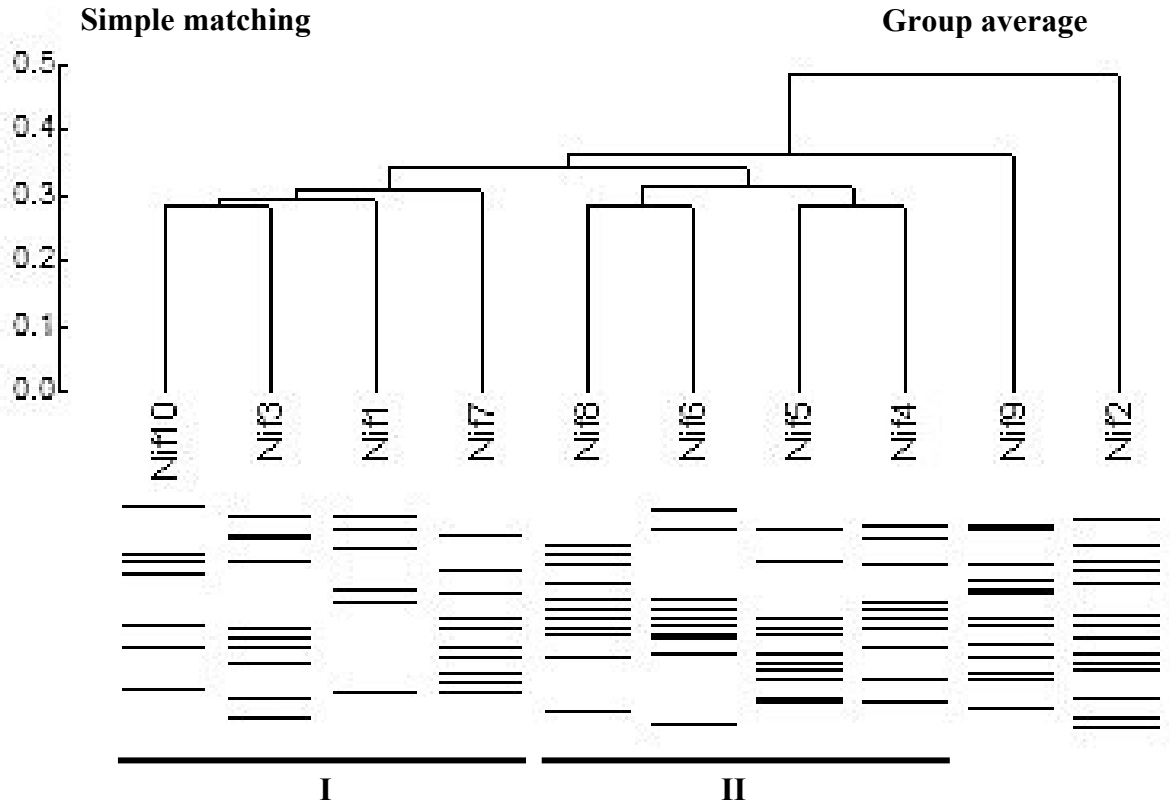


Figure 24: Cluster analysis of the banding pattern in Fig. 23, using a simple matching, group average setting to separate communities according to single gene species sequence differences (Thicker bands represent a brighter band on the gel, and therefore tentatively indicates a greater species richness).