

Chapter 5

Discussion

Discussion

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

It is evident from the density of the banding patterns observed on the gel (Fig. 6) that there was a higher microbial diversity in unpolluted than in polluted soils, as well as a decrease in microbial diversity in subsoil layers in comparison with topsoil. This was expected, since microorganisms present in polluted soil must be able to survive under PAH/PCB-induced stress conditions, with their growth furthermore limited by the availability of essential elements such as nitrogen (Röling *et al.* 2002). It is also known that few microbes are found deeper than 300mm into soil, especially when no plants are present (Zhou *et al.* 2002). It can be concluded that these microbes were either tolerating the presence of excess PAH/PCB or were in fact utilising them as a source of carbon. Utilisation of PAH/PCB requires specific metabolic pathways for obtaining carbon in an exploitable form. Competent microbial communities surviving and growing in oil-polluted soils can be considered metabolic generalists, differences in these community structures can be noted decades after an initial pollution event (Lindstrom *et al.* 1999). Bacteria capable of such growth have multi-component enzyme systems comprising of aromatic hydrocarbon dioxygenases (Gibson and Parales 2000). For example, toluene dioxygenase adds a dioxygen to the aromatic nucleus of benzene forming an arene-*cis*-diol-*cis*-1,2-dihydroxycyclohexa-3,5-diene (benzene-*cis*-diol) (Gibson and Parales 2000). These

specific enzymes enable growth, as opposed to survival, under aromatic hydrocarbon stressed conditions.

Based on phylogenetic relationships as well as distance analysis it was clear that the predominant representatives from the two soil types (polluted vs. unpolluted) grouped separately (Fig. 7). Supporting this were the two phylogenetic trees (Figs 8 and 9), indicating a closer relatedness and thus lower species diversity within the species isolated from polluted than from unpolluted soils. Topography of the beta and gamma bacteria shows them grouping in separate clades, this supports their classification as different evolutionary lineages. However, proteobacteria as whole seem to have a rather diverse phylogenetic grouping especially among the alpha and gamma groups (Purkhold *et al.* 2000; Sjöling and Cowan 2003; Davidov 2006). In these studies the phylogenetic groupings of the proteobacteria targeted show marked diversity from each other within clades and between clades. Purkhold *et al.* (2000) and Sjöling and Cowan (2003) both found that gamma clades could be very diverse forming bootstrap supported branches within clades. Purkhold *et al.* (2000) and Davidov *et al.* (2006) also show these definite branches forming almost completely separate clades within the alpha clade. The alpha bacteria are spread across clades in Fig. 7, possibly as a result of the lack of resolution for this specific taxonomic group, this could be due to limited number of specimen sequences or related species included in the analysis. However, different taxonomic contexts such as proteobacteria require different degrees of phylogenetic resolution (Woese 2000), thus perhaps the phylogeny within this tree is too broad for proteobacteria species diversity and more base pairs could be included to elucidate this. Nevertheless, Fig. 7 is showing that bacteria found in either polluted or unpolluted soils seem to group

together on the whole leading to “polluted” and “unpolluted” clades of organisms within the phylogram.

Preliminary tentative identification of the species representatives demonstrated that microbial equilibrium in polluted soils shifts to a different predominant group of closely-related organisms. This shift towards phylogenetic groupings displaying very similar consortia within polluted and unpolluted soils is in accordance with previous reports (Leys *et al.* 2004). Sequences obtained from DNA in lanes 1-3 (Fig. 6, samples 1-3 in Table 1) grouped closely together, except for sequence AY673793, a nitrobenzene-degrading *Pseudomonas* sp., that grouped in a clade with AY673806, a naphthalene-degrading *Pseudomonas* sp. Grouping together of nitrobenzene- and naphthalene-degrading bacteria by DGGE band sequencing has not been reported before.

Two main clades formed in both phylogenetic and distance trees and all sequences fell within one main clade, except for AY673792, which formed a clade on its own in both cases (Fig. 7; data not shown). AY673792 also fell basal to the main clade within the phylogenetic tree depicting species from polluted soils (Fig. 8). Its sequence grouped with *Pseudonocardia hydrocarbonoxydans*, *P. benzenivorans*, *P. petroleophila*, *P. saturnea* and *Actinobispora yunnanensis*. Many *Pseudonocardia* species are capable of oxidising complex hydrocarbons (Warwick *et al.* 1994; Lee *et al.* 2000). *P. hydrocarbonoxydans* degrades general hydrocarbon compounds, *P. benzenivorans* benzene, and *P. petroleophila* petroleum. *P. saturnea* and *A. yunnanensis* are both known to degrade aromatic hydrocarbons. From this information it can be concluded that this sequence (AY673792), in all likelihood,

represents an actinomycete that is capable of degrading polyaromatic hydrocarbons of the BTEX isomer group.

Other sequences from polluted soil samples that are also likely to be those of BTEX hydrocarbon degraders are AY673788, AY673789, AY673793, AY673794, AY673795 and AY673796. The hits that were obtained on GenBank BLAST searches indicate that, in accordance with Guerin (1999) and Tesar *et al.* (2002), there were many organisms utilising PAHs and PCBs in the polluted samples. In addition, alpha-proteobacteria were found to cluster with the above sequences. This is consistent with findings by MacNaughton *et al.* (1999) who detected alpha-proteobacteria only in crude oil-polluted soils. It is then plausible to assume that the microorganisms inherently present in these soils are capable of soil rehabilitation, and that pollution levels within the soils should be decreasing as PAHs/PCBs are being utilised. This remediation could occur until a certain point is reached when, perhaps, limiting factors such as seasonal climate change or exhaustion of an essential micronutrient will slow down the process (Ferris and Ward 1997; Ward *et al.* 1998; Muyzer 1999; Smalla *et al.* 2001; Koizumi *et al.* 2002). As the soil begins to return to a more natural, unpolluted state due to pollutant utilisation, new microbes will probably colonise the soil leading to an exponentially faster remediation process (White *et al.* 1998). Analysis of microbial communities has proved to be a far more comprehensive indicator of residual pollutants. For example, O'Sullivan and Mahenthiralingam (2005) found *Burkholderia* to be the predominant genus isolated from PAH-polluted soils capable of toluene degradation, but the population diminished as the pollutant level was reduced. Monitoring the return of a baseline community known to colonise unpolluted soil, e.g. several *Pseudomonas* spp. and

Sphingomonas spp., is used to indicate that the biological community of a soil is returning to normal (White *et al.* 1998).

In conclusion, it is evident that DGGE has indicated a decrease in microbial diversity in PAH/PCB polluted soils, as has been observed in other studies (Kozdrój and Van Elsas 2001). Andreoni *et al.* (2004) found that microbial soil biodiversity decreased with longer-term exposure to PAHs. However, soil pollution also selects for microorganisms with the ability to activate metabolic pathways, thereby allowing them to utilise the pollutants as an alternative source of carbon. Soils that become polluted with complex industrial hydrocarbons undergo a microbial community change that allows for a natural bioremediation process to be initiated.

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

It is estimated that there are more than 10^{16} prokaryotic cells in a tonne of soil. This is according to recent research which indicated that prokaryotic diversity is higher than previously thought by almost three orders of magnitude (Curtis and Sloan 2005). Measuring reservoirs of prokaryotic diversity, such as in soil, is challenging and it is widely accepted that disregarding the organisms themselves and focusing on their DNA is the most plausible option. Currently, diversity is estimated by targeting particular genes that occur in all the organisms being screened. Diversity is inferred by the number of different variants of these particular genes that can be cloned from a sample of environmental DNA (Curtis and Sloan 2005). However, unfortunately

the number of clones analysed provides only a small indication of the number of individual microbes present that could possibly be investigated.

Despite the possible major role of eukaryotes in bioaugmentation of soil, very little is known about which species are present (Meyer and Schmidt 2003). Literature contains studies of cultured soil eukaryotes, but these are of limited value because of flaws in methodology, inadequacy of surveys, technical problems with isolation and culture, and controversy of identification (Meyer and Schmidt 2003). DGGE is thus a more appropriate method to employ in diversity studies as it facilitates fingerprinting of communities at species level, hence allowing specific species to be targeted for diversity estimation. In view of this, prokaryotic and eukaryotic species diversity was compared and analysed concurrently in this study.

Prokaryotic species diversity and richness was notably higher than that of eukaryotes in the samples examined (Fig. 12). On average a 30% greater diversity was evident for prokaryotes than for eukaryotes (Fig. 13). Thus, at diversity level, a larger portion of prokaryotes appeared to be better able to adapt quickly to abrupt changes in the environment and/or in carbon source. The exception to the prokaryotic diversity dominance was sample 16, which had the lowest prokaryotic but nevertheless a high eukaryotic diversity overall, this may have been due to the soil pH being 3.9, by far the lowest of all the soil samples. The reason for the low pH is unclear, but most likely was not due to pollution since petrol, oil and diesel have a

pH of 5.7-6.0 (Jacobs^h, personal communication). Physico-chemical soil characteristics peculiar to the soil seems to be a more feasible explanation.

Although eukaryotes were found to be less diverse in this environment, their richness (indicated to an extent by band intensity) appeared to be very similar to that of prokaryotes. It is possible that low numbers of tolerant eukaryotic species were present in high numbers, creating a lower diversity but similar richness. This is supported by the work done by Lara (2005) who found, through 18S rRNA gene cloning, DGGE screening and artificial cultivation, that eukaryotes (protists) were present in higher numbers than prokaryotes in PAH-polluted soils. He attributed this to the fact that trophic webs are less complex in PAH-polluted soils due to a decrease in species diversity. However, it is also interesting to note that in the present study the prokaryote:eukaryote ratio in polluted and unpolluted soil was approximately the same. Many eukaryotes, such as various fungi, are capable of tolerating harsh environmental conditions than bacteria and are involved in the degradation of PAHs/PCBs in soil (Prenafeta-Boldú *et al.* 2002). Da Silva *et al.* (2003b) isolated filamentous fungi from estuarine sediments in Brazil and monitored their ability to degrade PAHs. They found a *Cyclothyrium* sp. to be the most efficient, simultaneously degrading 74, 70, 59 and 38% of pyrene, phenanthrene, anthracene and benzo[a]pyrene, respectively. Additionally, toluene, ethylbenzene and xylene have been shown to be degraded by a *Cladophialophora* sp. (Prenafeta-Boldú *et al.* 2002).

^h A. Jacobs, ARC-Plant Protection Research Institute, Private Bag x134, Queenswood, 0121 Pretoria.

Diversity profiles for pro- and eukaryotes demonstrated a distinct cluster pattern evident in the dendrogram (Fig. 14). Generally, prokaryotic profiles grouped together, probably due to their higher diversity and possibly closer relations (as shown in section 5.1). 16S9 was the most diverse community according to banding in the dendrogram and this was expected since it is the unpolluted control soil. It was also the only community not clustering in a clade. Clade I contained only ITS communities with a low number of bands, indicating low diversity. Clade II contained three ITS and three 16S sequences, indicating very similar community diversity across these communities. Clade III displayed a sparse banding throughout the communities and had the majority of ITS communities represented. This clade had the lowest diversity of all the clades, due to most of the eukaryotes grouping within it. There was an increase in community diversity as clades branched off from Clade III. Clades IV, V and VI contained mainly 16S and this grouping was supported by an evident increase in diversity. Diversities within clade VI were the most similar to each other as these were basal to all other communities, although diversity was not as high as in 16S9. 16S9 is similar in community to clade V and VI, but grouped out due to a much higher species diversity.

The PAH breakdown potential of prokaryotic vs. eukaryotic communities was not determined. However, extrapolation of the DGGE data indicates that, due to the significantly higher diversity within prokaryotic communities, and their better adaptation potential within pollution stressed environments, the prokaryotic component of these samples would have the greatest PAH metabolism potential. This is in accordance with studies conducted by Atlas (1981) and Ahn *et al.* (1999) who showed that PAH/PCB degradation is not restricted to only a few bacterial

genera, but that many are in fact capable of survival and growth in a petroleum-hydrocarbon polluted environment. However, there have been cases where more fungal and protist genera than bacterial genera were found to degrade PAHs/PCBs in soil samples (Atlas 1981; Lara 2005). In many cases such as these, DGGE has proved to be the most valuable molecular tool in screening many diverse samples for different microorganisms.

5.3 DGGE analysis of toluene and naphthalene degraders in polluted soil

Different soils contain different microbial communities that have adapted to survive and develop within the specific environment in which they exist. However, Gelsomino *et al.* (1999) found after extensive molecular fingerprinting that similar soil types tend to contain comparable dominant bacteria. This was also the case in the present study, as unpolluted soils tended to group together, as did soil from the rhizosphere of the same species of plant.

The mean number of bands on the DGGE gel between the *xylE* and *ndoB* genes was higher for *xylE*, thus indicating a higher toluene degrading capacity, except in soils 17 and 18 (Fig. 16). Samples 17 and 18 were from soil mulched with wood chips and polluted with workshop oil, which is derived from tar and thus contains naphthalene, hence accounting for the higher naphthalene degrading capacity. Milcic-Terzic *et al.* (2001) found that consortia isolated from diesel-polluted soils possessed both catabolic genes. They determined that toluene-degrading consortia contained only the *xylE* gene, whereas the naphthalene-degrading consortia

possessed only the *ndoB* gene. Results of the current study demonstrated that both catabolic genes were present in all soil samples. However, a higher level of diversity among *xylE*⁺ organisms indicated that there could have been a greater amount of toluene in samples 9-16. Similarly in samples 17 and 18 higher diversity among *ndoB*⁺ organisms indicated the possibility of a predominantly naphthalene polluted environment. It was interesting to note a high toluene and naphthalene degrading capacity in soils 9-12 that contained no PAH/PCB pollution. This indicates that many soil dwellers contain the capacity for toluene and naphthalene degradation but that these genes are not always active. Siciliano *et al.* (2003) similarly found that there were often more *ndoB*⁺ and *xylE*⁺ organisms present in their unpolluted control soils than in the polluted soils being studied. Furthermore, Schneegurt and Kulpa (1998) reported that exposure to aromatic substrates caused an increase in inherent *ndoB* and *xylE* carrying plasmid levels in indigenous soil bacteria, leading to enzymatically catabolised degradation of naphthalene and toluene, respectively.

The grouping together of soil communities in clades according to the gene amplified was anticipated, due to the similar community structure expected when one gene is targeted for fingerprinting. However, soil community fingerprints showed some mixing of polluted and unpolluted profiles within single clades. This could have been due to different gene loci exhibiting similar gel profiles as a result of genetic similarity with respect to, for instance, G-C content. Similar electrophoretic mobilities of phylogenetically related species sharing analogous sequences could cause co-migration across samples within the amplified area, whereas a similar melting behaviour between phylogenetically unrelated species could cause bands pattern to show some level of homology (Smalla *et al.* 2001).

The *ndoB* gene dendrogram (Fig. 18) showed a clustering of control soils 9, 10, 11 and 12 within clade I, and a cluster of polluted soils 14, 16 and 17 in clade II, in both cases possibly due to similar community structures within each clade. The assembly of 14 and 16 closer to each other probably can be ascribed to the fact that they were both from rhizosphere soil. Samples 13 and 18 grouped together in the dendrograms of both *xylE* and *ndoB* genes. This grouping was also evident in clades I and IV of the dual-gene dendrogram (Fig. 17). In this respect the community similarity seems to be stronger within the *xylE* gene (Fig. 19) as this grouping was reflected closest within its dendrogram and the dual-gene dendrogram. Samples 13 and 18 did not have common a rhizosphere or plant derivation, but shared a common pollutant, namely oil. It can thus be concluded that the *xylE* and *ndoB* genes are both needed for oil degradation, confirming the studies by Siciliano *et al.* (2003) and Luz *et al.* (2004). The common grouping of samples 15 and 16 in the *xylE* gene dendrogram was also reflected in the dual-gene dendrogram, which could be accounted for by the fact that these samples were from non-rhizosphere and rhizosphere soil under *B. serratia* in the same area, respectively. This clustering of samples (15 and 16) as a result of the *xylE* gene is in accordance with Siciliano *et al.* (2003), who found that change in the functional genetic composition of a community was not detectable by 16S rDNA, but was linked to specific functional genotypes (*ndoB*, *alkB* and *xylE*) with relevance to PAH degradation. In the context of the present study this implies that functional gene differences elucidated community dynamics, and response to PAH/PCB pollution, when targeting the *ndoB* and *xylE* genes. It is thus in agreement with Siciliano *et al.* (2003) who reported that phytoremediation systems increase the

catabolic potential of rhizosphere soils by influentially altering the functional composition of the microbial community.

On basis of the above results, it can be concluded that the polluted soil communities in South Africa have similar microbial communities, depending on their physical location, plant species proximity and soil conditions, viz. type of pollutants to which they are exposed. Examining metabolic gene diversity *in situ* has indicated that soils with one or more of the above factors in common tend to have a comparable community diversity within their functional genes (Milcic-Terzic *et al.* 2001; Siciliano *et al.* 2003). It was also evident that, even when not existing under PAH-polluted conditions, local soil microbial communities have the capacity to break down PAHs should pollution occur. This supports the finding by Wünsche *et al.* (1995) that substrate utilisation patterns in the Biolog system changed upon addition of hydrocarbons. According to them, previously pristine soil bacterial communities shifted to a predominantly *Pseudomonas* population with hydrocarbon degradation adaptation capacity.

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

The genera *Staphylococcus* and *Providencia* have not previously been reported to degrade PAHs and are also not common rhizosphere colonisers. Members of the *Providencia* genus can, however, completely break down hexahydro-1,3,5-trinitro-1,3,5-tiazine (RDX) and nitroso-RDX, and have been used to achieve such in

bioremediation (Kitts *et al.* 1994). In the present study it was initially assumed that *Staphylococcus* and *Providencia* were most likely present and surviving in polluted soil rather than performing PAH-degradation, but both have since been found to degrade PAH in culture (Molobela 2005, unpublished data). Members of these genera clustering with known PAH-degraders such as *Bacillus* in the present study, supports this. Furthermore, the presence of the *xylE* and *ndoB* genes within *Staphylococcus* and *Providencia* isolates has been reported (Kitts *et al.* 1994), which indicates that they may be opportunistic PAH degraders in environments providing none or very little of an alternative carbon source.

Pseudomonas, *Providencia* and *Staphylococcus* constituted a very narrow spectrum of taxa compared with the previous investigation of crude oil-polluted soil (5.1), albeit from a different site in South Africa, which indicated the presence of the following genera (in order of prevalence): *Pseudomonas*, *Sphingomonas*, *Methylocystis*, *Pseudonocardia*, *Acidocella*, *Azospirillum*, *Bradyrhizobium*, *Bosea*, *Roseomonas*, *Rhodoplanes*, *Blastochloris*, *Porphyrobacter*, *Sphingobium*, *Rhizomonas*, *Sphingopyxis*, *Acinetobacter*, *Rhodococcus*, *Acidiphilium*, *Methylocella*, *Actinobispora* and *Acidospheara* (Fig. 7). Except for *Pseudomonas*, none of these genera was isolated from any of the samples in the present study, either because they were not culturable or not present in sufficient numbers to allow isolation. *Pseudomonas* therefore appears to be the genus with the greatest potential as bioaugmentation agent, since it grows rapidly in culture and is a robust soil, rhizosphere and PAH-polluted site coloniser (Tesar *et al.* 2002; Kuiper *et al.* 2004). *Pseudomonas* dominated the bacterial population in the rhizosphere of *C. esculentus* and also occupied non-rhizosphere soil. *Cyperus* and *Eleusine* are associated with

phytoremediation of oil-polluted soils (Merkl *et al.* 2005), and species such as *C. esculentus* and *E. corocana*, as well as *B. pilosa*, are common invaders of PAH-polluted sites in South Africa (Maila *et al.* 2005a; Molobela *et al.* 2005). Of these three weed species, only *E. corocana* has thus far been found to be actively involved in phytoremediation of PAH-polluted soils in the country (Maila *et al.* 2005a). Although probably purely coincidental, it is interesting to note that *E. corocana* was also the only species from which *Staphylococcus* could be isolated.

Indigenous microbial communities inhabiting polluted sites are known to include species capable of bioaugmentation of the site, after being targeted, isolated and identified. The organisms within these communities are affected by biotic and abiotic factors that influence the ability of microbes to mineralise PAH and aliphatic compounds, but have proven to be effective in bioremediation at field scale (Samanta *et al.* 2002). Samanta *et al.* (2002) reported isolating a large number of naphthalene degrading microorganisms, including *Alcaligenes denitrificans*, *Mycobacterium* sp., *Pseudomonas putida*, *P. fluorescens*, *P. paucimobilis*, *P. vesicularis*, *P. cepacia*, *P. testosteroni*, *Rhodococcus* sp., *Corynebacterium venale*, *Bacillus cereus*, *Moraxella* sp., *Streptomyces* sp., *Vibrio* sp. and *Cyclotrophicus* sp. The identification of so many *Pseudomonas* species being capable of naphthalene degradation is consistent with earlier literature. Davies and Evans (1964) penned the first report of the catabolic pathway for naphthalene degradation by *Pseudomonas* spp. Consistent with the studies of Davies and Evans (1964) and Samanta *et al.* (2002), the present study found *Pseudomonas* to be the most common PAH-degrading genus, in this case identified and isolated from the rhizosphere of *C. esculentus*. However, it was established that, although less common, *Providencia* and *Staphylococcus* possess the

ability to mobilise PAH-degrading enzymes. Considering the well-established rhizosphere competence and PAH-degrading capacity of *Pseudomonas*, this genus nevertheless seems to be the best suited for bioaugmentation purposes in South Africa.

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

PCR revealed all soil samples to include *nifH* genes, regardless of whether they were from rhizosphere soil or polluted sites or not. This was expected, as nitrogen-fixing organisms are not only common in unpolluted soils but also able to survive and grow in pollutant-stressed soils (K.A. Reynoldsⁱ, personal communication). Diesel and other fuel spills result in a very high carbon to nitrogen ratio within affected soils (Eckford *et al.* 2002). Initially it was believed that such a situation could limit microbial degradation of PAHs. Eckford *et al.* (2002) found several consortia capable of PAH-degradation and nitrogen fixation under nitrogen-limited conditions in polluted Antarctic soils, whereas Rahman *et al.* (2002) observed a high level of inorganic nitrogen utilisers in samples taken from soils at petrol and diesel stations. Thus, from a pollution perspective, these studies support the presence of *nifH* genes in samples screened in the present study. However, isolates *nif12* and *nif18* did not respond to the *nifH* PCR (Fig. 22), hence supporting the apparent absence of this nitrogen fixation gene in the *Providencia* genus. *Providencia* belongs to the family *Enterobacteriaceae* and when subjected to metabolic tests is positive for the reaction nitrate to nitrite, this is the same reaction that is catalysed by denitrifying bacteria

ⁱ K.A. Reynolds, Eskom resources and strategy, CR &D department, Private Bag x40175, Cleveland, 2022.

(Manos and Belas 2006). Thus *Providencia* is capable of performing the same functions as those of denitrifying bacteria. It is interesting to note that, although the remaining isolates were colony PCR'd from pure culture on agar plates, they still maintained the ability for nitrogen fixation.

According to DGGE profiles (Fig. 23), diversity differed between the various soil samples, evidently because of the different soil types and plant rhizospheres being sampled. The control soil (*nif1*) showed low nitrogen-fixing bacterial diversity, as could be predicted since most *nifH*-containing microorganisms are associated with plant roots, which were absent in this soil. Polluted, non-rhizosphere soils, by contrast, showed a fairly high diversity of nitrogen-fixing organisms. This is attributed to the survival of these microbes under stress, where they have no plants with which to associate and from which to gain habitat and nutrients. However, Muratova *et al.* (2003) found that the growth of microbes associated with the roots of lucerne (*Medicago sativa* L.) in PAH-polluted soil was greatly improved by addition of nitrogen. Fewer organisms are present in non-rhizosphere than in rhizosphere soils (Blagodatskaya *et al.* 2004), thus it follows that there will be less competition in non-rhizosphere soil despite more stressed living conditions. This cumulatively leads to an increased diversity, but not necessarily to increased species richness, when compared to rhizosphere soils. Also, microbes existing in a less nutrient-stressed environment, such as the plant rhizosphere, may have less need for nitrogen fixation due to indigenous root-associated microbes fixing nitrogen on their behalf. Conversely, microbes living some distance from plants will require an active *nifH* gene for nitrogen-fixation, as plant root nodules and associated bacteria are inaccessible.

Cluster analysis (Fig. 24) showed *nif2* and *nif9* falling basal to the other samples tested. This seems feasible as these two samples displayed the highest species diversity. However, it is interesting to note that these samples are exact opposites, *nif2* was sampled from *B. pilosa* rhizosphere in unpolluted soil and *nif9* from non-rhizosphere polluted soil. *B. pilosa* is known to have a rhizosphere rich in nitrogen-fixers and it therefore is plausible that species diversity may be high (Wolfe and Klironomos 2005). However, *nif9* showed a much higher species richness than *nif2*, as indicated by the darker banding. This implies that the diversity of *nif2* and *nif9* was similar but that richness was higher in the non-rhizosphere polluted soil, possibly due to the reason stated above that more free-living microbes are forced to fix nitrogen under stressed living conditions.

Clade I of the cluster analysis showed *nif3* and *nif10* grouping together. They were both from *B. serratia* rhizosphere, which could explain the similarity in their rhizosphere communities. *nif3* and *nif10* were low in diversity, possibly due to nitrogen-fixing microbes being present mainly inside the plant roots. In the same clade were *nif1* and *nif7*, both sampled from non-rhizosphere soils. *nif7* showed a higher diversity, possibly because of its pollutant-stressed habitat, whereas *nif1* was a control soil that is expected to have a high nitrogen-fixing capacity. Clade II grouped together *nif6* and *nif8*, possibly because they were both rhizosphere soils and have similar diversity, their community fingerprints were sufficiently similar to conclude that the species within their respective communities are alike. The grouping together of *nif4*, *nif5* and *nif6* was expected as they were all associated with *C. esculentus*.

Nitrogen-fixing organisms were present in all the soils sampled. This was anticipated since free, available nitrogen is a limited resource in the soil environment as it easily returns to the atmosphere (Kaye and Hart 1997). Nitrogen fixation was also found to be present in 80% of the pure cultures isolated from polluted and unpolluted soils and rhizospheres. Although different rhizospheres and pollutants were examined, it was found that the highest *nifH* gene diversity of polluted soils existed within machinery oil-polluted, wood-chip-mulched, non-rhizosphere soil. This was also the most visibly and olfactorily highest level of polluted soil that according to site knowledge had been deliberately polluted severely for 10 years. It therefore appears that the more polluted the soil, the higher the free microbial nitrogen fixation diversity due to environmental stress, which is in accordance with Zepp *et al.* (2003) who found that PAH pollution of the environment alters the nitrogen cycling therein indirectly effecting the microbial communities.