

# Molecular diversity and characterization of nitrite reductase gene fragments (*nirK* and *nirS*) from nitrate- and uranium-contaminated groundwater

Tingfen Yan,<sup>1,2,3†</sup> Matthew W. Fields,<sup>1†\*</sup> Liyou Wu,<sup>1</sup> Yuangang Zu,<sup>3</sup> James M. Tiedje<sup>2</sup> and Jizhong Zhou<sup>1,2\*</sup>

<sup>1</sup>Environmental Sciences Division, Oak Ridge National Laboratory, P.O. Box 2008, MS-6038, Bldg. 1505, Oak Ridge, TN 37831–6036, USA.

<sup>2</sup>Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824, USA.

<sup>3</sup>Open Research Laboratory of Forest Plant Ecology, Northeast Forestry University, Harbin 150040, P. R. China.

## Summary

Nitrate-contaminated groundwater samples were analysed for *nirK* and *nirS* gene diversity. The samples differed with respect to nitrate, uranium, heavy metals, organic carbon content, pH and dissolved oxygen levels. A total of 958 *nirK* and 1162 *nirS* clones were screened by restriction fragment length polymorphism (RFLP) analysis: 48 and 143 distinct *nirK* and *nirS* clones, respectively, were obtained. A single dominant *nirK* restriction pattern was observed for all six samples and was 83% identical to the *Hyphomicrobium zavarzinii* *nirK* gene. A dominant *nirS* pattern was observed for four of the samples, including the background sample, and was 95% identical to the *nirS* of *Alcaligenes faecalis*. Diversity indices for *nirK* and *nirS* sequences were not related to any single geochemical characteristic, but results suggested that the diversity of *nirK* genes was inversely proportional to the diversity of *nirS*. Principal component analysis (PCA) of the sites based on geochemistry grouped the samples by low, moderate and high nitrate but PCA of the unique operational taxonomic units (OTUs) distributions grouped the samples differently. Many of the sequences were not closely related to previously observed genes and some phylogenetically related sequences were obtained from similar samples. The results indicated that the contaminated groundwater contained novel *nirK* and *nirS* sequences, functional diversity of both genes changed in relation to the contaminant gradient, but

the *nirK* and *nirS* functional diversity was affected differently.

## Introduction

The reduction of nitrate to dinitrogen gas, denitrification, is part of the global nitrogen cycle and is primarily a bacterial respiratory process. Besides the importance in global nitrogen balance, more recent concerns have arisen as a result of nitrate- and nitrogen oxide-polluted groundwater, as well as the emergence of dinitrogen oxide as a potent greenhouse gas (Dickinson and Cicerone, 1986). Different anthropogenic processes can be a major source for nitrogen oxides, including agriculture, wastewater treatment and industry (Thiemens and Trogler, 1991; Zumft, 1997).

The products of partial nitrate reduction can inhibit bacterial growth and also be problematic in the reduction of mobile metal species, such as uranium (Tiedje, 1988; Zumft, 1992; Senko *et al.*, 2002). To understand how biogeochemical processes affect microbial community structure and bioremediation, the U.S. Department of Energy's (DOE) Natural and Accelerated Bioremediation Research (NABIR) program has established a Field Research Center (FRC) on the DOE Oak Ridge Reservation (<http://www.esd.ornl.gov/nabirfrc/>) in eastern Tennessee. The FRC is heavily contaminated with nitrate and heavy metals, and is a designated site for research associated with the NABIR program.

Dissimilatory nitrite reduction is catalysed by two types of nitrite reductases: the copper-containing *nirK* and the cytochrome *cd*, *nirS* gene products (Braker *et al.*, 1998). Previous work has indicated that the *nirK* and *nirS* sequences are useful for understanding the composition of denitrifier communities (Braker *et al.*, 1998; Hallin and Lindgren, 1999; Braker *et al.*, 2000; Prieme *et al.*, 2002). To investigate how nitrate and other contaminants can affect denitrifying communities, the diversity of *nirK* and *nirS* genes was determined for groundwater samples contaminated with varying levels of nitrate and heavy metals. Our results indicate that many of the *nirK* and *nirS* clones represent novel sequences and that the environmental conditions appear to have affected presumptive, denitrifying communities differently along a contaminant gradient. To our knowledge, this is the first assessment of nitrite

Received 23 August, 2002; accepted 18 October, 2002 \*For correspondence. E-mail ytz@ornl.gov or fieldsm1@ornl.gov; Tel. (+1) 865 573 7544 or (+1) 865 241 3775. †Contributed equally to the presented research

reductase community structure and diversity from nitrate- and uranium-contaminated groundwater.

## Results

### Site characteristics

The chemical characteristics were significantly different among the six samples. Wells FW-005, -010 and -015 in Area 1 and 3 had low pH values, and background (FW-300), TPB-16 and FW-003 had circum neutral pH values (Table 1). All wells except background had nitrate levels over the drinking water standards of 10 p.p.m. or 0.16 mM (<http://www.epa.gov/safewater/dwh/c-ioc/nitrates.html>).

The nitrate concentrations ranged from 0.02 mM (~1 p.p.m.) at the background site and 0.67 M (~42 000 p.p.m.) at FW-010. Samples from FW-005 and FW-015 had increased nitrate, low pH, and high uranium and heavy metals. TPB-16 and FW-300 samples had circum neutral pH, low heavy metals and low nitrate. Samples FW-010 and FW-003 had increased nitrate and low uranium, and low and circum neutral pH values respectively. The TOC (total organic carbon) levels were similar for FW-005, FW-015 and TPB-16 were approximately two-fold above background. FW-003 and FW-010 had approximately threefold and sixfold more TOC than background. The substratum of the sampling sites was similar and consisted of shale with inner-embedded limestone that has been weathered to clay, silty saprolite (<http://www.esd.ornl.gov/nabirfrc/>). The flow was in a south-south-west direction from the source ponds towards FW-005, FW-010 and FW-015.

### RFLP analysis of *nirK* gene fragments

All six samples yielded PCR products with the *nirK*-specific primer pairs and a total of 958 *nirK* clones were screened from the six sites. Rarefaction and regression analyses indicated that the majority of recovered diversity

**Table 1.** Geochemical data for the groundwater samples used for molecular analysis.

	pH	Nitrate (mM)	U ( $\mu$ M)	Ni ( $\mu$ M)	TOC (mM)	DO (mM)
FW-300 <sup>a</sup>	6.1	0.02	0.004	0.09	2.50	0.17
FW-003	6.0	17.0	0.042	0.26	8.30	0.03
FW-005	3.9	2.80	27.0	85.2	5.80	0.05
FW-010	3.5	677	0.71	307	14.6	0.05
FW-015	3.4	135	32.4	150	5.40	nd
TPB-16	6.3	0.50	4.62	ND	5.40	0.03

**a.** FW-300 is the background area spatially isolated from the contaminated sites.

Nitrate, total organic carbon (TOC) and dissolved oxygen (DO) values are reported as mM, uranium (U) and nickel (Ni) as  $\mu$ M were determined by the FRC management team at Oak Ridge National Laboratory and the data are maintained at <http://www.esd.ornl.gov/nabirfrc/>. ND, not detected; nd, not determined.

was sampled within 20 analysed clones for all the *nirK* libraries except TPB-16, which required approximately 40 clones (data not shown).

The same *nirK* restriction pattern was observed for all six wells as the dominant recovered clone (C01-03-6) accounted for 78% of the total *nirK* library and comprised 40–91% of the respective sample libraries (Fig. 1). The wells FW-300, TPB-16 and FW-015 each had approximately 25% unique RFLP patterns that were not observed in the other samples, but only 5% of the clones from wells FW-003, FW-005 and FW-010 had unique patterns. A total of nine restriction patterns were shared among the samples and the predominant pattern was the only RFLP pattern observed at all sites. A total of 48 operational taxonomic units (OTUs, i.e. unique patterns) were unique to the different samples.

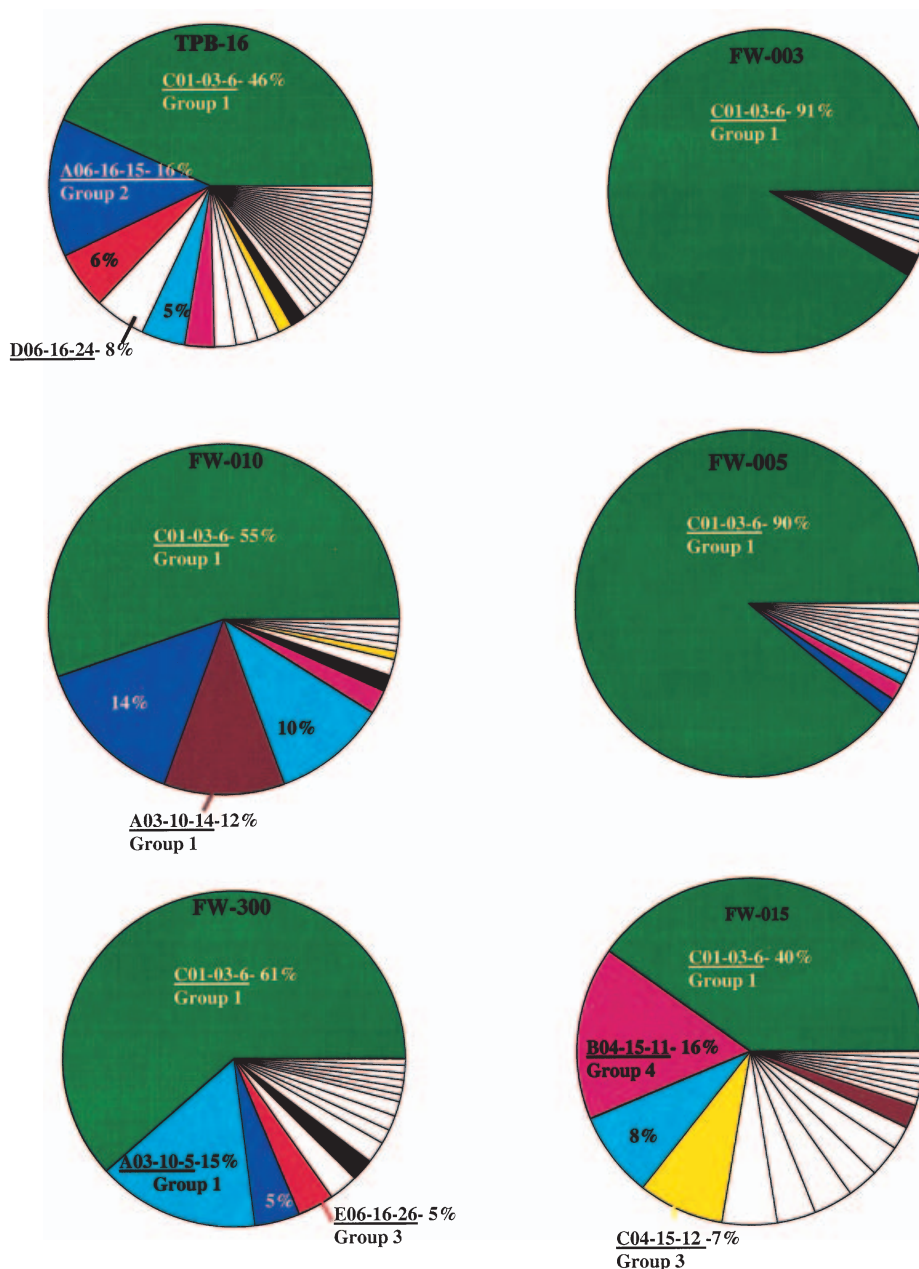
The clone libraries from FW-003 and FW-005 displayed one predominant pattern and had low diversity. The other four samples displayed higher diversity and evenness (Table 2). Samples FW-300 and FW-010 had similar diversity, evenness and richness indices, even though FW-010 had 34 000-fold higher nitrate levels. The background sample had the highest similarity value with FW-015 (0.49) and the clone similarity of the three acidic sites ranged from 0.33 to 0.36 (Table 3).

### RFLP analysis of *nirS* gene fragments

All six samples yielded PCR products with the *nirS*-specific primers and a total of 1162 *nirS* clones were screened from the six sites. Rarefaction and regression analyses indicated that the majority of diversity was sampled within 30 analysed clones from FW-300 and TPB-16, 40 clones for FW-015, 55 clones for FW-005 and FW-010, and 60 clones for FW-003 (data not shown).

A single predominant OTU (s-10-4, clone Group 1) was obtained from wells FW-300, FW-005, FW-010 and FW-015 (Fig. 2), and the wells FW-003 and TPB-16 were predominated by different OTUs (clone Groups 5 and 10 respectively). The samples FW-300, FW-005 and FW-003 had between 40 and 60% OTUs which were not observed in the other samples, seven *nirS* patterns were observed in more than one sample and 40% of the entire library was unique to at least one groundwater sample. The diversity of *nirS* varied from well to well and FW-015 and TPB-16 had decreased diversity indices compared to FW-300 (Table 2). The samples, FW-003 and FW-005, displayed the highest *nirS* diversity, and also had the lowest *nirK* diversity. The FW-003 sample had the greatest diversity of both the *nirK* or *nirS* genes from all of the wells.

The *nirS* libraries from each site had little overlap and, except for the predominant patterns, most *nirS* clones were unique to the respective well. The background well shared the most *nirS* clones with wells FW-005 and FW-



**Fig. 1.** The clone distribution and overlap among sites for the *nirK* gene sequences. Each colour is a unique sequence, and the same colour shared between sites represents sequence overlap and has the same clone number. Pie segments depicted as white represent unique sequences observed only at the respective site. The sequence names are denoted for predominant clones as well as the respective percentage for that site.

010 and the similarity values were 0.25 and 0.20 respectively (Table 3). The samples FW-003 and TPB-16 each had a unique, predominant pattern, one of the lowest similarity values (Table 3), and the next most observed OTU was the dominant clone surveyed in the other wells (Fig. 2).

#### Sequence analysis of *nirK* clones

Partial sequences were determined for all clones having

distinctive restriction patterns. When *nirK* sequences from representatives of the  $\alpha$ -,  $\beta$ - and  $\gamma$ -*Proteobacteria* were compared to the unique *nirK* clones ( $n = 48$ ), many of the FRC sequences were more closely related to one another than to previously reported genes (Fig. 3). Only one clone, D05-015-127, had higher than 90% nucleotide identity with a known *nirK* gene (*Bradyrhizobium japonicum*,  $\alpha$ ). The *nirK* sequence from *Hyphomicrobium zavarzinni* ( $\alpha$ ) was the most closely related known sequence to almost half of the FRC *nirK* OTUs and approximately 60% of the

**Table 2.** Characteristics and diversity estimates for the *nirK* and *nirS* gene clones from six FRC groundwater samples.

Well	# of clones <sup>a</sup>	OTU <sup>b</sup>	H <sup>c</sup>	1/D <sup>d</sup>	Evenness <sup>e</sup>	Richness <sup>f</sup>
<i>nirK</i>						
FW-300	143	16	2.33	3.30	0.58	22 ± 5
FW-003	229	12	0.70	1.20	0.20	37 ± 19
FW-005	185	12	0.78	1.23	0.22	18 ± 5
FW-010	134	12	2.13	2.90	0.60	18 ± 6
FW-015	140	20	3.10	5.11	0.71	71 ± 33
TPB-16	127	31	2.93	4.07	0.67	35 ± 9
<i>nirS</i>						
FW-300	173	36	2.92	3.10	0.55	84 ± 18
FW-003	210	56	3.86	6.51	0.75	39 ± 3
FW-005	253	59	3.60	4.01	0.61	64 ± 8
FW-010	174	54	2.15	2.59	0.52	38 ± 14
FW-015	175	39	1.33	1.47	0.32	27 ± 7
TPB-16	177	37	1.91	2.07	0.45	28 ± 7

a. Number of clones in each sublibrary analysed by RFLP.

b. Operational taxonomic units based on unique RFLP.

c. Shannon-Wiener index, higher number represents more diversity.

d. Reciprocal of Simpson's index, higher number represents more diversity.

e. As evenness approaches 1, the population is more evenly distributed.

f. Statistical prediction of the number of different species or genes.

FW-300 represents the background area, FW-003 has circum neutral pH but high nitrate (1000 p.p.m.), TPB-16 is circum neutral pH, low nitrate (0.5 mM), and the acidic sites, FW-005, FW-010, and FW-015, have low pH (3.5–3.9), moderate to high nitrate (2.8 mM to 0.7 M), high nickel and high uranium.

OTUs had 80–85% identity with known sequences from the  $\alpha$ -*Proteobacteria*, *H. zavarzinni*, *Rhizobium melitoti* or *Blastobacter denitrificans*. The remaining unique OTUs had 70–78% nucleotide identity with the known *nirK* genes from *H. zavarzinni* or *Blastobacter denitrificans*.

The *nirK* clone Group 1 sequences in Cluster I were 98–99% identical, contained the predominant, recovered OTU from every site and had 81–83% sequence identity with the *nirK* from *H. zavarzinni* (Fig. 3). The clone Group 2, which contained seven sequences with 97–99% sequence identity, had 75% nucleotide identity with the *H. zavarzinni nirK* gene (Fig. 3). This group also contained the second most frequently recovered sequence from TPB-16 and was well represented by clones from TPB-16 and background (e.g. low nitrate, circum neutral pH). The unique sequences C05-015–113 and F04-015–43 were predominantly recovered from the similar sites, FW-005 and FW-015. Cluster III contained the second most recovered sequence from FW-015 (B04-15–11, Group 4). Clone B04-15–11 formed a cluster with three other sequences (Group 4) that were 91–99% identical to each other (Fig. 3) and were predominately (87%) recovered from the acidic sites (e.g. high contaminants). The clone Group 3 contained four clones that were slightly more prevalent (64%) at TPB-16 and the background sites. The two clones in Cluster IV were grouped with *Nitrosomonas* sp. C-45 with the similarity of 61%.

#### Sequence analysis of *nirS* clones

When the *nirS* OTUs ( $n = 143$ ) were compared with pre-

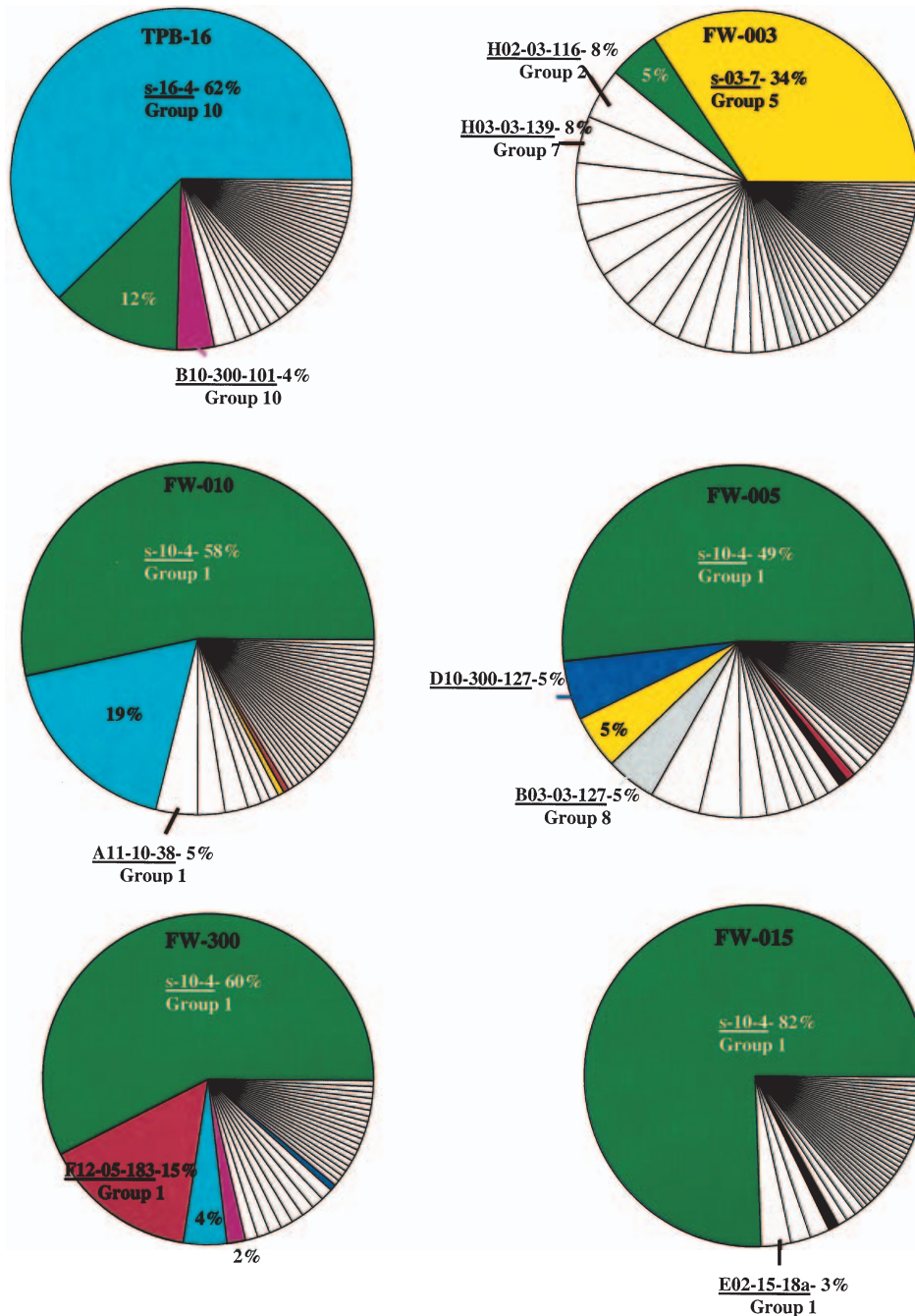
viously known genes, the identity values ranged from 59 to 100%. Almost 45% of the OTUs had  $\geq 95\%$  sequence identity with previously known genes, approximately 30% of the OTUs were 80–85% identical to previously known genes, and the remaining 25% were 70–80% identical to previously reported sequences (Fig. 4).

The *nirS* dendrogram could be divided into five clusters (I–V) and Cluster I contained a majority of the FRC clones as well as the predominant sequences for samples FW-300, -005, -010 and -015. The clone Group 1 in Cluster I accounted for over half of the entire library and had 95% identity with the *nirS* from *Alcaligenes faecalis* ( $\beta$ ) and approximately 92% identity with the *nirS* of *Pseudomonas stutzeri* ( $\gamma$ ). Clone Group 1 was dominated by recovered clones from low pH samples, whereas almost all recovered sequences in clone Group 2 were obtained from circum neutral pH samples. The *nirS* of *Azospirillum brasilense* clustered with clone Groups 3 and 4 with approximately 83% similarity for each group.

The clone Group 5 (Cluster II) contained the predominant, recovered clone from FW-003, was only 75% iden-

**Table 3.** Pair-wise similarity coefficients of *nirK* (top) and *nirS* (bottom) operational taxonomic units (OTUs) based on RFLP analysis.

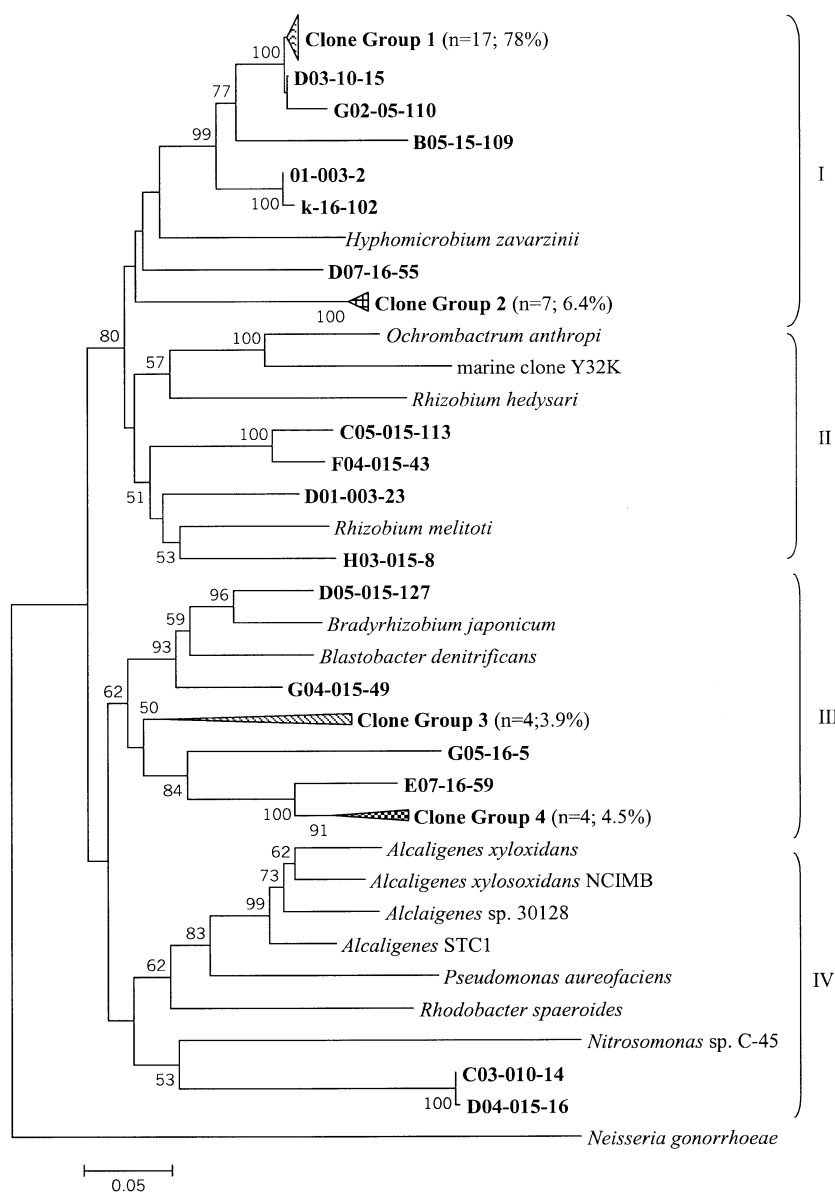
	FW-003	FW-005	FW-010	FW-015	TPB-16	FW-300
FW-300	0.36	0.36	0.36	0.49	0.30	
TPB-16	0.19	0.23	0.28	0.38		0.16
FW-015	0.24	0.36	0.36		0.03	0.13
FW-010	0.25	0.33		0.11	0.09	0.20
FW-005	0.42		0.16	0.12	0.06	0.25
FW-003		0.10	0.05	0.04	0.04	0.11



**Fig. 2.** The clone distribution and overlap among sites for the *nirS* gene sequences. Each colour is a unique sequence the same colour shared between sites is the same sequence. Pie segments depicted as white represent unique sequences observed only at the respective site. The sequence names are denoted for predominant clones as well as the respective percentage for that site.

tical to the closest known reference (*A. faecalis*), and a majority (85%) of the recovered sequences were from FW-003 (e.g. circum neutral pH and high nitrate). The clone Groups 6 and 7 were almost exclusively observed from the FW-003 sample and Group 7 had at least 99% sequence similarity with the *nirS* from *Pseudomonas aeruginosa* (Fig. 4). The clone Group 8 (Cluster IV) and

clone Group 9 (Cluster V) were not closely related to previously observed sequences and were predominantly obtained from the acidic sites FW-005 and FW-010 respectively. The clone Group 10 (Cluster IV) had the predominant clone from TPB-16 and the majority of recovered clones (75%) were from the similar sites, TPB-16 and FW-300. Also, Cluster V contained only one previously



**Fig. 3.** Phylogenetic relationship of all unique cloned *nirK* sequences from the FRC and reference sequences from the GenBank database. The tree is based on a neighbor-joining method and pairwise deletion. The four clone groups represent closely related FRC sequences the number (n) depicts the number of sequences in the group and the percentage (%) represents the total of the entire *nirK* library. The percentage of 500 bootstrap values that supported is branch is shown bootstrap values below 50% are not shown. The classification of the reference microorganisms is denoted with  $\alpha$ ,  $\beta$ , or  $\gamma$  for *Proteobacteria* affiliation. The accession numbers for reference sequences are as follows: *H. microbium* (AJ224902), *S. meliloti* (AJ224909), *R. hedysari* (U65658), *O. anthropi* (AJ224907), marine clone Y32K (AJ248448), *B. japonicum* (AJ002516), *B. denitrificans* (AJ224906), *R. spaeroides* (U62291), *P. aureofaciens* (Z21945), *Alcaligenes* STC1 (AB046603), *Alcaligenes* 30128 (AJ224903), *A. xylooxidans* (AJ224905), *A. xylooxidans* (AJ224904) and the *aniA* gene from *Neisseria gonorrhoeae* (M97926) as an outgroup.

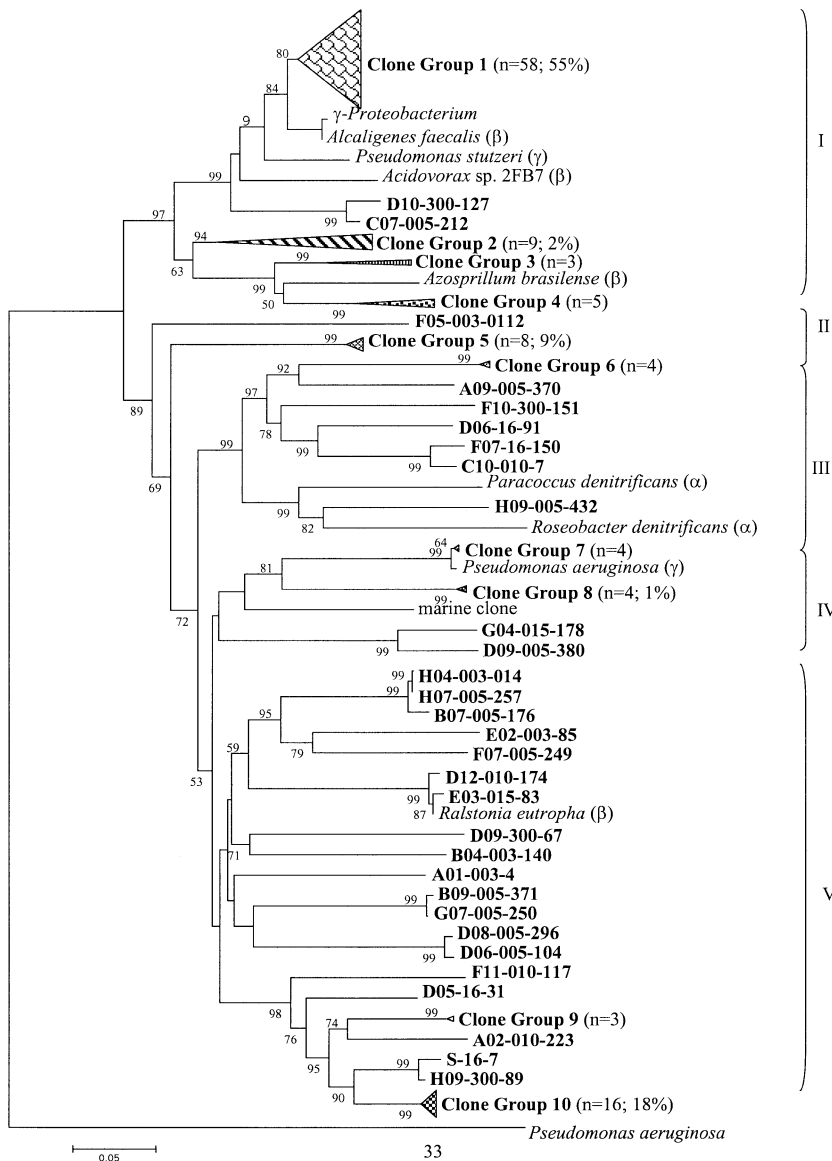
known *nirS* sequence (*R. eutropha*), a majority (95%) of the recovered OTUs appeared to be novel sequences (<75% nucleotide identity) and was predominated by sequences from circum neutral samples.

#### Principal component analysis (PCA)

The seven geochemical analytes, nitrate, uranium, pH, nickel, total organic carbon, dissolved oxygen (DO) and non-purgable organic carbon were used for PCA (Fig. 5). The background site (FW-300) and TPB-16 were distinctly different from the other sites and these two sites had the lowest nitrate levels. The well FW-005 had relatively moderate nitrate levels and appeared to be affected differently than the other low pH sites. The wells, FW-003, -010 and

-015 were grouped and all had high nitrate levels. When the distribution of unique sequences and the geochemical measurements were simultaneously analysed with PCA, the samples displayed similar results as with the physical measurements alone (data not shown). For both *nirK* and *nirS*, FW-300, TPB-16 and FW-005 appeared to be affected differently, and FW-003, FW-010 and FW-015 were closely grouped.

However, when the OTU distributions from the samples were analysed with PCA different sample associations were observed. For the *nirK* library, site FW-015 did not associate with the other samples and the FW-300, FW-010, TPB-16, FW-005 and FW-003 samples had more similar OTU distributions (Fig. 6A). The sites, FW-003, TPB-16, FW-010 and FW-005 were more closely grouped

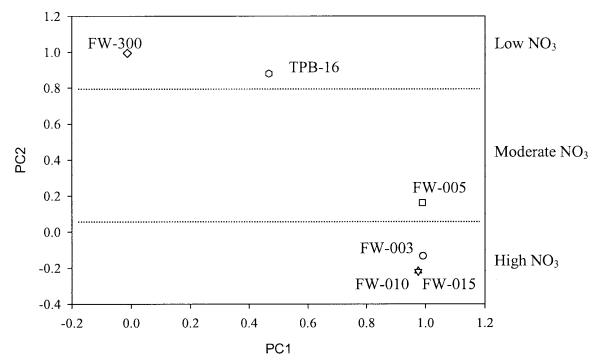


**Fig. 4.** Phylogenetic relationship of all unique cloned *nirS* sequences from the FRC and reference sequences from the GenBank database. The tree is based on a neighbor-joining method and pairwise deletion. The 10 clone groups represent closely related FRC sequences the number (n) depicts the number of sequences in the group and the percentage (%) represents the total of the entire *nirS* library. The percentage of 500 bootstrap values that supported is branch is shown bootstrap values below 50% are not shown. The classification of the reference microorganisms is denoted with  $\alpha$ ,  $\beta$ , or  $\gamma$  for *Proteobacteria* affiliation. The accession numbers for the reference sequences are as follows:  $\gamma$ -*Proteobacterium* D7-6 (AJ248396), *A. faecalis* (AJ224913), *P. stutzeri* (X56813), *Acidovorax* 2FB7 (AY078273), *A. brasilense* (AJ224912), *P. denitrificans* (U05002), *R. denitrificans* (AJ224911), *P. aeruginosa* (X16452), marine clone pA94 (AJ248416), *R. eutropha* (X91394) the *nirN* gene from *Pseudomonas aeruginosa* (D84475) as an outgroup.

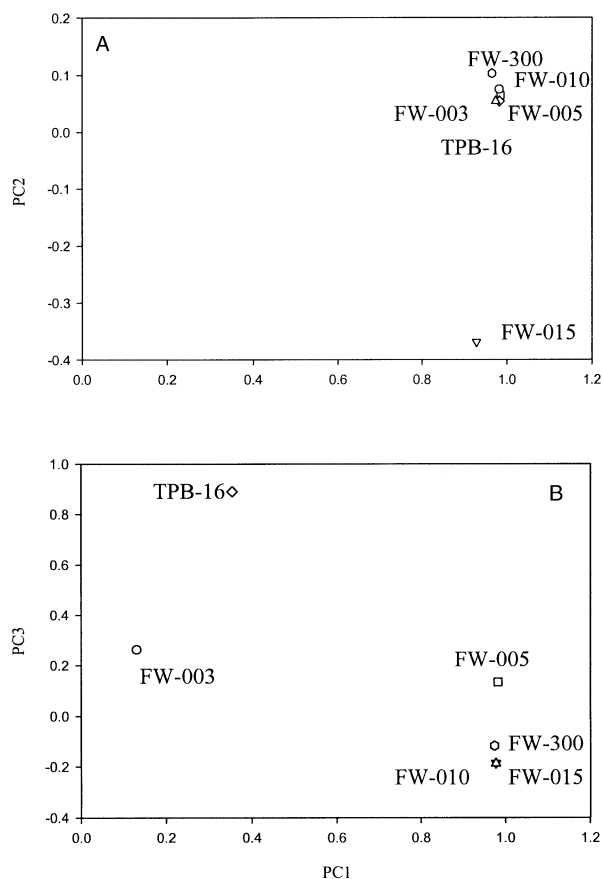
away from FW-300. For the *nirS* library, FW-003 and TPB-16 were distinct from FW-005, FW-010, FW-015 and FW-300, and the background site appeared to be more similar to FW-010 and FW-015 (Fig. 6B).

## Discussion

Polymerase chain reaction-based cloning approaches are commonly used to assess diversity and community structure and can be an informative tool for environmental studies. However, the abundance, dominance or diversity may not be completely epitomized by clone distribution, therefore, the measurements and indices are used for relative comparisons. The formation of chimeras, heteroduplexes and mutations can be intrinsic to PCR ampli-



**Fig. 5.** Principal components analysis (PCA) of the six sites based on the geochemical parameters (nitrate, pH, nickel, uranium, total organic carbon, non-purgable organic carbon and dissolved oxygen).



**Fig. 6.** PCA of the unique (A) *nirK* OTUs or (B) unique *nirS* OTUs characterized by site. The symbols are labelled for each site.

cation (Qiu *et al.*, 2001), and protocols were followed to minimize PCR amplification-induced artifacts as described in the *Experimental procedures* section. Our previous study suggested that the chimeric and heteroduplex frequency was approximately 10% based on 16S rRNA genes (Zhou *et al.*, 2002). Because similar protocols were used for these studies, we expect that PCR generated artifacts were not a significant problem.

Wells with varying geochemical characteristics were sampled from a field site with a single contamination source and the wells FW-005, FW-010 and FW-015 had low pH values and increased nitrate, FW-003 had circum neutral pH and high nitrate levels, and FW-300 and TPB-16 had circum neutral pH and relatively low nitrate levels. Approximately 130–230 clones were evaluated from each of the samples. Distinctive clones based on RFLP analysis were used for sequence comparisons, and RFLP analysis with a combination of tetrameric restriction endonucleases can be a sensitive method for clone differentiation. Sequence divergence may exist within RFLP groups, but is most likely not significant.

Previously reported phylogenetic relationships based on the 5S rDNA were highly similar with the type of *nir*

gene for known microorganisms (Ohkubo *et al.*, 1986), and recent work with *amoA* suggested that environmental sequences with less than 80% nucleotide identity were indicative of previously unobserved species (Purkhold *et al.*, 2000). Such a relationship between denitrifiers and the *nirK* and *nirS* genes has not yet been elucidated, but the *nirK* and *nirS* genes within the  $\alpha$ -,  $\beta$ - and  $\gamma$ -*Proteobacteria* subdivisions have approximately 75% or greater nucleotide sequence identity. In this study, we used this value as a threshold to evaluate whether the recovered sequences were considerably different.

The dominant *nirK* sequence for all sites might represent novel nitrite reductases distantly related to *H. zavarzinii* ( $\alpha$ ). *Hyphomicrobium* species are methylotrophic bacteria able to degrade C-1 compounds, and recently, Layton *et al.* (2000) detected *Hyphomicrobium* spp. as a dominant population in a municipal wastewater treatment system in Kingsport, TN. Because the same FRC *nirK* sequence predominated all sites under a wide range of geochemical conditions, the host microorganisms could play an important role in denitrification at FRC sites and not be significantly impacted by the contaminants.

The novel *nirK* clone Groups 2, 3 and 4 comprised almost 15% of the entire *nirK* clonal library. The novel clone Group 4 sequences were predominantly from acidic samples and may represent presumptive *nirK*-containing denitrifiers that can survive in low pH environments. Also, the novel clones, C05-015–113 and F04-015–43, were predominantly from sites FW-005 and -015, and this result suggested that these sequences might represent previously uncultivated denitrifiers acclimated to such environments (low pH, high U and Ni). The copper-containing dissimilatory nitrite reductase was recently reported in ammonia-oxidizing, nitrifying bacteria (Casciotti and Ward, 2001), and the sequences, C03-10–14 and D04-15–16 (both obtained from low pH sites), clustered with the presumptive *nirK* gene of *Nitrosomonas* C-45. These two sequences represented 3% of the total *nirK* library and corresponded to 11% or 8% of the respective libraries.

The predominant *nirS* sequences from FW-003 (s-03–7; Group5) and TPB-16 (s-16–4; Group 10) appeared to be novel, and Groups 5 and 10 were predominated by sequences from the circum neutral sites. When only the OTU distribution was analysed with PCA, the results suggested that FW-003 and TPB-16 were affected differently than the other sites. Both FW-003 and TPB-16 are circum neutral sites with similar TOC and low nickel levels, but have different nitrate and uranium levels. These results suggested that the novel sequences represented by Groups 5 and 10 might be in response to nitrate and/or uranium, but other factors most likely are involved. These two novel sequence groups had approximately 73% sequence identity.



Groups 8 and 9 (*nirS*) were almost exclusive to the acidic sites, FW-005 and FW-010, respectively, and might represent presumptive denitrifiers acclimated to low pH environments. However, low pH can have a negative impact on denitrification (Prieme *et al.*, 2002), and the occurrence of these sequences might be the result of other physiological attributes. The identification of divergent gene sequences underscores the need for further isolation, identification, and characterization of the microorganisms with pertinent genotypic and phenotypic traits from sites of interest.

Hallin and Lindgren (1999) observed *nirK* sequences in all 13 samples of wastewater sludge tested, but only five samples were positive for *nirS*. On the other hand, Braker *et al.* (2000) identified *nirK* sequences in Washington state margin sediments only but observed *nirS* sequences in both Washington margin and Puget Sound sediments. In a more recent study, *nirK* sequences were amplified from both forested upland and wetland soils but *nirS* was only observed in the marsh samples (Prieme *et al.*, 2002). In our study, both *nirK* and *nirS* sequences were detected in all groundwater samples tested, even though the contaminant levels differed drastically. However, the *nirK* and *nirS* diversity appeared to be different within the same microbial communities.

Previous studies with marine sediments and terrestrial soils reported higher *nirS* diversity than *nirK* (Braker *et al.*, 2000; Prieme *et al.*, 2002) and the FRC *nirS* groundwater libraries were generally more diverse than the *nirK*. Interestingly, TPB-16 and FW-015 had the lowest *nirS* diversity, but had the highest *nirK* diversity. Also, the sites FW-003 and FW-005 had the lowest *nirK* diversity but displayed the highest *nirS* diversity. When the different diversity indices for *nirK* and *nirS* sequences were compared from the same site, the two gene pools were inversely proportional ( $r^2 = 0.88$ ; data not shown).

Increased *nirS* diversity was noted in environments with relatively moderate levels of nitrate (~3 mM and 16 mM), and higher *nirK* diversity was observed in groundwater samples with extremes in nitrate levels (< 0.5 mM or > 16 mM). These results suggested that particular environmental conditions might alter the proportions of denitrifiers with a *nirK* or *nirS*, and that the community dynamics for *nirK* gene microorganisms might affect the community structure of *nirS* gene microorganisms and vice versa. Little is known regarding environmental impacts on *nirK*- and *nirS*-denitrifiers or communities, but oxygen sensitivity, kinetic characteristics and substrate requirements can vary between microorganisms as well as communities (Tiedje, 1988; Zumft, 1997; Cavigelli and Robertson, 2000). The apparent differences between and within the samples may represent unique properties of an active denitrifying community under respective conditions, but other physiological traits of the community most likely play

a role. Additional sampling and characterization are needed to explore the possible biochemical and ecological associations.

The major difference between and among the sites appeared to be the nitrate levels based on PCA of geochemical measurements and a similar relationship was observed when OTU distribution was included. These results suggested that the nitrate levels might be a major factor on presumptive denitrifiers with a *nirK* and *nirS* gene. However, the PCA of OTU distribution alone indicated different site relationships and these results indicated that OTU distribution was not always a sole function of nitrate levels. These data suggested that simple relationships were difficult to delineate from environmental parameters, ecological indices and sequence data, particularly when the datasets were combined and complexity increased. Also, microbial community structure is most likely affected by many different abiotic and biotic variables in a non-linear fashion.

The *nirK* FW-015 library had different RFLP patterns, a high occurrence of Group 4 sequences, and PCA did not group FW-015 with the other sites. For the *nirS* library, the circum neutral sites, FW-003 and TPB-16, were not grouped with the other sites. The comparisons of the unique *nirS* sequences indicated that Groups 5, 6 and 7 were phylogenetically distinct and appeared to predominate FW-003. The novel Group 8 (*nirS*) sequences were a significant portion of only the FW-005 library and PCA suggested that FW-005 was affected differently than the other low pH samples. The *nirS* Groups 5 and 8 sequences might represent unique denitrifiers acclimated to increased nitrate levels and low pH environments respectively.

In general, less *nirK* diversity was observed among the samples and the same sequence appeared too predominant every site. The high nitrate, low pH sites (-005, -010, -015) were predominated by group 1 sequences, but FW-015 was dominated to a lesser extent and also displayed more diversity despite low pH values. Although all samples were predominated by the same sequence, PCA distinguished sample FW-015 from the others based on unique sequences.

The same *nirS* sequence predominated background and the acidic samples, irrespective of nitrate and DO levels. Perhaps, the low pH environment of the acidic wells has constrained the *nirS* population in relation to nitrate levels in a similar fashion to a low-nitrate, increased-oxygen environment (background). The other circum neutral pH sites, FW-003 and TPB-16, appeared to be predominated by different, previously uncultivated *nirS* denitrifiers and PCA distinguished these circum neutral pH sites from the low pH sites as well as background.

In summary, the results indicated that: (i) novel *nirK* and *nirS* sequences were identified along a single contamina-

tion gradient (acidic to circum neutral pH and low, moderate and high nitrate levels), and, in some cases, similar sequences could be identified at sites with similar geochemistry; (ii) *nirK* and *nirS* community structure was affected differently for the five contaminated samples whereas diversity at the background site was similar for both genes; and (iii) *nirK* and *nirS* diversity indices were inversely proportional at the contaminated sites. Principal component analysis indicated that nitrate was a major factor, but sequence occurrence and distribution was not always easily explained in relation to nitrate, pH, or DO levels. Future work is needed to characterize the associations between environmental conditions and *nirK* and *nirS* communities and better understand biotic and abiotic affects on functional dynamics. We are currently investigating the use of functional gene microarrays (Wu *et al.*, 2001) and neural network analysis tools, as well as the isolation of bacterial isolates with novel *nir* genes.

## Experimental procedures

### Sampling sites

The FRC site includes three areas of contaminated soil and groundwater and an uncontaminated background area that contains soils similar to those found in the contaminated areas. The site contained four unlined ponds that received approximately 10 million litres of liquid nitric acid and uranium bearing wastes per year for approximately 30 years until closure in 1984. The waste ponds contribute to both sediment and groundwater contamination with nitrate, uranium and heavy metals. The site is also contaminated with a variety of organic contaminants at varying levels and a full description can be found at the FRC website (<http://www.esd.ornl.gov/nabirfrc/>). The source ponds were presumptively neutralized and denitrified before being capped, but the efficacy of the treatment was not monitored.

Groundwater samples were collected from six wells at the FRC (Table 1). The wells FW-005 and FW-010, located in Area 3, are 32.5 m apart and are approximately 20 m from the former waste pond. Well FW-015 is 27 m from the waste pond embankment in Area 1 and is approximately 130 m from the wells in Area 3. The wells FW-003 and TPB-16, located in Area 2, are separated by 40 m. Area 2 is approximately 275 m down gradient from the waste ponds, Area 1 and Area 3. The well, FW-300, is located in the uncontaminated background area, approximately 6 km northwest of the source ponds. The water table was approximately 6 m and water was collected from a screened interval below the water table at each of the six wells on the same day (December 2000).

### DNA extraction and purification

Groundwater samples (1–2 l) were collected and transported to the laboratory in amber glass bottles. Bacteria were harvested by centrifugation (10 000 *g*, 4°C for 30 min) and the pellets were stored at –80°C until used for DNA extraction.

The cell pellet was resuspended in a lysis buffer and the cells disrupted with a previously described grinding method (Zhou *et al.*, 1996). The DNA was extracted as previously described (Zhou *et al.*, 1996; 1997) and the precipitated DNA was purified by gel electrophoresis plus mini-column preparation (Wizard DNA Clean-Up system, Promega, Madison, WI).

### PCR amplification and cloning

The fragments of the *nirK* and *nirS* genes were amplified in a 9700 Thermal Cycler (Perkin-Elmer) with the redesigned primer pairs copper 583 F (5'-TCA TGG TGC TGC CGC GKG ACG G-3') and copper 909R (5'-GAA CTT GCC GGT KGC CCA GAC-3') for the *nirK* gene, and heme 832 F (5'-TCA CAC CCC GAG CCG CGC GT-3') and heme 1606R (5'-AGK CGT TGA ACT TKC CGG TCG G-3') for the *nirS* gene. The PCR reactions (20 µl) contained 2 µl 10 × buffer (500 mM KCl, 100 mM Tris HCl pH 9.0 and 1% Triton X-100), 1.5 µl 25 mM MgCl<sub>2</sub>, 0.2 µl 400 ng µl<sup>-1</sup>. Bovine serum albumin (Boehringer Mannheim, Indianapolis, IN), 0.2 µl 25 mM 4 × dNTPs (USB, Cleveland, OH), 10 pmol each primer, 2.5 U Taq polymerase and 1 µl purified DNA (5–10 ng). To minimize PCR-induced artifacts, the optimal number of cycles was determined and five PCR reactions were combined before cloning as described previously (Qiu *et al.*, 2001). The PCR parameters were as follows: 80°C for 30 s, 94°C for 2 min; 94°C for 30 s, 60°C for 1 min, 72°C for 1 min, 30 cycles; 72°C for 7 min. Polymerase chain reaction products were analysed on 1.5% (wt/vol) TAE agarose gels.

The combined PCR products (326 bp and 774 bp, respectively) were separated by electrophoresis in a low-melting point agarose gel (0.8%), the appropriate band excised and the DNA extracted with a Wizard Prep Kit (Promega, Madison, WI) according to manufacture's instructions. Recovered DNA was resuspended in 6 µl ddH<sub>2</sub>O, 2 µl was ligated with pCR2.1 vector from a TA-cloning kit and competent *Escherichia coli* cells were transformed according to the provided protocol (Invitrogen, San Diego, CA).

### RFLP analysis

Approximately 200 white colonies from each sample were randomly selected and screened for *nirK* and *nirS* inserts, which were detected with the primers specific to the polylinker of the vector pCR<sup>TM</sup>II as described previously (Qiu *et al.*, 2001). Each single white colony was picked, resuspended in 60 µl 0.1 × TE buffer (1 mM Tris, 0.1 mM EDTA), boiled for 5 min at 100°C and then stored at –40°C. The inserts were amplified (20 µl reactions) with the TA primers specific for the pCR2.1 vector as described previously (Qiu *et al.*, 2001). The inserts were visualized on a 1.5% TAE agarose gel stained with ethidium bromide and clones with the expected size were used for further analysis.

Amplified PCR products (6 µl) were digested with the tetrameric enzymes *MspI* and *RsaI* (Gibo-BRL, overnight at 37°C, 0.1 U each). Digested fragments were separated by electrophoresis (7 V cm<sup>-1</sup>, 4 h) in 3.5% Metaphor agarose (BMA, Rockland, ME) gels with 10 µl of 10 mg l<sup>-1</sup> ethidium bromide in 1 × TBE buffer. The RFLP patterns were visualized with ultraviolet radiation and saved as TIFF images. The

RFLP patterns were analysed and clustered with the MOLECULAR ANALYST 1.6 software (Applied Math, Kortrijk, Belgium) using the unweighted pair group method of arithmetic averages and the Jaccard algorithm. The clusters were validated by comparison with the gel images.

### Sequence and phylogenetic analysis

Unique *nirK* and *nirS* clones from each site were selected for sequence analysis based on differences in RFLP patterns and are designated as OTUs (operational taxonomic unit). The number of sequences determined for each site is listed in Table 2. Polymerase chain reaction products (100 µl) amplified with vector-specific primers were purified with the ArrayIt™ PCR Purification Kit (TeleChem International, Sunnyvale, CA) or treated with ExoSAP-IT™ (USB, Cleveland, OH) according to manufacturer instructions. The DNA sequences were determined with a BigDye Terminator kit (Applied Biosystem, Foster City, CA) using a 3700 DNA analyser (Perkin-Elmer, Wellesley, MA) according to the manufacturer instructions at a 5 : 1 dilution. The DNA sequences were assembled and edited using the SEQUENCHER™ program (v.4.0, Gene Codes, Ann Arbor, MI).

The sequences obtained with the vector-specific primers were compared with *nirK* and *nirS* sequences from GenBank (March 2002). The sequences were aligned with CLUSTALW AND alignments compared with reference sequences from the database (Thompson *et al.*, 1994). Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 2.1 (Kumar *et al.*, 2001) and phylogenetic trees were constructed with distance matrices and the neighbor-joining method within MEGA. Trees constructed with maximum-likelihood and neighbor-joining methods were not significantly different. The nucleic acid sequence accession numbers in GenBank are as follows: AF548909-AF548940 and AF548942-AF548958 for *nirK* clones and AF548959-AF549061, AF549063-AF549065, and AF549066-AF549105 for *nirS*.

### Data analysis

The diversity (Shannon-Wiener and Simpson's index) and evenness (equitability) indices were calculated with RFLP data obtained from the clonal libraries using equations from Krebs (Pielou, 1969; Krebs, 1989; Brown and Bowman, 2001). The reciprocal of Simpson's index ensures a better relationship between an increasing value of the index and an increased diversity (Magurran, 1988). The evenness index (E) measures the equitability of species abundance for a sample and a sample contains many clones with similar representation as E approaches a value of 1 (Krebs, 1989; Brown and Bowman, 2001).

Clone richness was estimated using the non-parametric Chao model (Chao, 1984) and gives a statistical prediction of the total number of different clones in a given sample. Similarity coefficients were calculated for each pair of samples using the following equation (Odum, 1971; Brown and Bowman, 2001):  $S = 2C/A + B$ , where A and B are the number of RFLP patterns in libraries A and B, respectively and C is the number of shared patterns. Principal components

analysis used the programs in the SYSTAT statistical computing package (v. 9.0, SPSS, Chicago, IL).

### Acknowledgements

The authors thank Drs C. S. Criddle and C. E. Bagwell for helpful discussions. The authors would also like to thank B. L. Kinsall and K. A. Lowe for sample collection, as well as Dr P. M. Jardine, T. L. Mehlhorn and D. B. Watson for field sample analysis. This research was supported by The United States Department of Energy under the Natural and Accelerated Bioremediation Research Program of the Office of Biological and Environmental Research, Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for DOE under contract # DE-AC05-96OR22464.

### References

- Braker, G., Fesefeldt, A., and Witzel, K.-P. (1998) Development of PCR primer systems for amplification of nitrite reductase genes (*nirK* and *nirS*) to detect denitrifying bacteria in environmental samples. *Appl Environ Microbiol* **64**: 3769–3775.
- Braker, G., Zhou, J., Wu, L., Devol, A.H., and Tiedje, J.M. (2000) Nitrite reductase genes (*nirK* and *nirS*) as functional markers to investigate diversity of denitrifying bacteria in Pacific Northwest marine sediment communities. *Appl Environ Microbiol* **66**: 2096–2104.
- Brown, M.V., and Bowman, J.P. (2001) A molecular phylogenetic survey of sea-ice microbial communities (SIMCO). *FEMS Microbiol Ecol* **35**: 267–275.
- Casciotti, K.L., and Ward, B.B. (2001) Dissimilatory nitrite reductase genes from autotrophic ammonia-oxidizing bacteria. *Appl Environ Microbiol* **67**: 2213–2221.
- Cavigelli, M., and Robertson, G.P. (2000) The functional significance of denitrifier community composition in a terrestrial ecosystem. *Ecology* **81**: 1402–1414.
- Chao, A. (1984) Nonparametric estimation of the number of classes in a population. *Scand J Stat* **11**: 265–270.
- Dickinson, R.E., and Cicerone, R.J. (1986) Future global warming from atmospheric trace gases. *Nature* **319**: 109–115.
- Hallin, S., and Lindgren, P. (1999) PCR detection of genes encoding nitrite reductase in denitrifying bacteria. *Appl Environ Microbiol* **65**: 1652–1657.
- Krebs, C.J. (1989) *Ecological Methodology*. New York: Harper and Row.
- Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. (2001) *MEGA2: Molecular Evolutionary Genetics Analysis Software*. Arizona State University, Tempe, Arizona, USA.
- Layton, A.C., Karanth, P.N., Lajoie, C.A., Meyers, A.J., Gregory, I.R., Stapleton, R.D., Taylor, D.E., and Saylor, G.S. (2000) Quantification of *Hyphomicrobium* populations in activated sludge from an industrial wastewater treatment system as determined by 16S rRNA analysis. *Appl Environ Microbiol* **66**: 1167–1174.
- Magurran, E. (1988) *Ecological Diversity and its Measurement*. Princeton, NJ: Princeton University Press.
- Odum, E.P. (1971) Principles and concepts pertaining to organization at the community level. In *Fundamentals of*

- Ecology*. Philadelphia, PA : Saunders College Publishing, pp. 140–161.
- Ohkubo, S., Iwasaki, H., Hori, H., and Osawa, S. (1986) Evolutionary relationships of denitrifying bacteria as deduced from 5S rRNA sequences. *J Biochem* **100**: 1261–1267.
- Pielou, E.C. (1969) *An Introduction to Mathematical Ecology*. New York: Wiley and Sons.
- Prieme, A., Braker, G., and Tiedje, J.M. (2002) Diversity of nitrite reductase (*nirK* and *nirS*) gene fragments in forested upland and wetland soils. *Appl Environ Microbiol* **68**: 1893–1900.
- Purkhold, U., Pommerening-Roser, A., Juretschko, S., Schmid, M.C., Koops, H.-P., and Wagner, M. (2000) Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and *amoA* sequence analysis: implications for molecular diversity surveys. *Appl Environ Microbiol* **66**: 5368–5382.
- Qiu, X., Wu, L., Huang, H., McDonel, P.E., Palumbo, A.V., Tiedje, J.M., and Zhou, J. (2001) Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. *Appl Environ Microbiol* **67**: 880–887.
- Senko, J.M., Istok, J.D., Suflita, J.M., and Krumholz, L.R. (2002) *In-situ* evidence for uranium immobilization and remobilization. *Environ Sci Technol* **36**: 1491–1496.
- Thiemens, M.H., and Trogler, W.C. (1991) Nylon production: an unknown source of atmospheric nitrous oxide. *Science* **251**: 932–934.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
- Tiedje, J.M. (1988) Ecology of denitrification and dissimilatory nitrate reduction to ammonium. In *Biology of Anaerobic Microorganisms* Zehnder, A.J.B., (ed.). New York: John Wiley and Sons, pp. 179–244.
- Wu, L., Thompson, D.K., Li, G., Hurt, R.A., Tiedje, J.M., and Zhou, J. (2001) Development and evaluation of functional gene arrays for detection of selected genes in the environment. *Appl Environ Microbiol* **67**: 578–5790.
- Zhou, J., Burns, M.A., and Tiedje, J.M. (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62**: 316–322.
- Zhou, J., Davey, M.E., Figueras, J.B., Rivkina, E., Gilichinsky, D., and Tiedje, J.M. (1997) Phylogenetic diversity of a bacterial community determined from Siberian tundra soil DNA. *Microbiol* **143**: 3913–3919.
- Zhou, J., Xia, B., Treves, D.S., Wu, L., Marsh, T.L., O'Neill, R.V., Palumbo, A.V., and Tiedje, J.M. (2002) Spatial and resource factors influencing high microbial diversity in soil. *Appl Environ Microbiol* **68**: 326–334.
- Zumft, W.G. (1992) The denitrifying prokaryotes. In *The Prokaryotes. A Handbook on the Biology of Bacteria*. Balows, A., Truper H.G., Dworkin M., Harder W. and Schleifer K.H., (eds). New York: Springer-Verlag, pp. 554–582.
- Zumft, W.G. (1997) Cell biology and molecular basis of denitrification. *Microbiol Mol Biol Rev* **61**: 533–616.