Seasonal variations in the diversity and abundance of diazotrophic communities across soils

Michele C. Pereira e Silva, Alexander V. Semenov, Jan Dirk van Elsas & Joana Falcão Salles

Department of Microbial Ecology, Centre for Life Sciences, University of Groningen, Groningen, The Netherlands

Correspondence: Joana Falcão Salles, Department of Microbial Ecology, Centre for Life Sciences, University of Groningen, PO Box 11103, 9700 CC Groningen, The Netherlands. Tel.: +31 50 363 2162; fax: +31 50 363 2348; e-mail: j.falcao.salles@rug.nl

Received 30 September 2010; revised 1 March 2011; accepted 1 March 2011. Final version published online 1 April 2011. DOI:10.1111/j.1574-6941.2011.01081.x

Editor: Christoph Tebbe

Keywords
nifH; soil; diversity; abundance; baseline; soil type.

Abstract
The nitrogen (N)-fixing community is a key functional community in soil, as it replenishes the pool of biologically available N that is lost to the atmosphere via anaerobic ammonium oxidation and denitrification. We characterized the structure and dynamic changes in diazotrophic communities, based on the nifH gene, across eight different representative Dutch soils during one complete growing season, to evaluate the amplitude of the natural variation in abundance and diversity, and identify possible relationships with abiotic factors. Overall, our results indicate that soil type is the main factor influencing the N-fixing communities, which were more abundant and diverse in the clay soils (n = 4) than in the sandy soils (n = 4). On average, the amplitude of variation in community size as well as the range-weighted richness were also found to be higher in the clay soils. These results indicate that N-fixing communities associated with sandy and clay soil show a distinct amplitude of variation under field conditions, and suggest that the diazotrophic communities associated with clay soil might be more sensitive to fluctuations associated with the season and agricultural practices. Moreover, soil characteristics such as ammonium content, pH and texture most strongly correlated with the variations observed in the diversity, size and structure of N-fixing communities, whose relative importance was determined across a temporal and spatial scale.

Introduction
Biological nitrogen fixation (BNF) is the reduction of atmospheric N2 gas to biologically available ammonium, catalyzed by the enzyme nitrogenase (Postgate, 1998). It is performed by phylogenetically diverse groups of prokaryotic microorganisms that harbor the nifH gene, one of the genes coding for the structural part of nitrogenase. BNF is important in terrestrial ecosystems as it replenishes the pool of biologically available N that is lost to the atmosphere via anaerobic ammonium oxidation and denitrification (Capone & Knapp, 2007). N-fixing bacteria are important regulators of plant productivity because plants cannot fix atmospheric N and because N is, together with phosphorus and potassium, the main element that limits plant productivity (Chapin, 1980). Although the majority of N fixation is carried out by bacteria that live in association with plants, free-living diazotrophs in soils have been shown to be important contributors to the N budgets in several ecosystems, reaching up to 60 kg ha⁻¹ year⁻¹ (Cleveland et al., 1999). Duc et al. (2009) studied the diversity of free-living diazotrophs in the soils of the Damma glacier (Swiss Central Alps), and also measured the potential asymbiotic activity using the acetylene reduction assay, revealing a high diversity of nifH gene sequences and illustrating the importance of free-living diazotrophs and their potential contribution to the N input in this environment.

Symbiotic N-fixing bacteria are known to be highly sensitive to perturbation (Doran & Safley, 1997). Indeed, several environmental factors have been suggested to influence N fixation in soils, including soil moisture, oxygen, pH, carbon quantity and quality, N availability (Hsu & Buckley, 2009), soil texture and aggregate size (Poly et al., 2001b), and clay content (Roper & Smith, 1991). For instance, Wakelin et al. (2007) showed that the retention of stubble in soil approximately doubled the abundance of nifH genes. This was not affected by the application of an N source such as urea. In another experiment, the addition of N fertilizers...
led to rapid changes (within 15 days) in the community structure of \textit{nifH}-harboring bacteria in association with rice roots (Tan \textit{et al.}, 2003). Plants are known to affect free-living N-fixing activities in soil, which was correlated with the grass (Patra \textit{et al.}, 2006), as well as the diazotrophic community structure, as it was shown that different cultivars of sorghum (\textit{Sorghum bicolor}) select for different \textit{nifH} communities (Coelho \textit{et al.}, 2008). The community structure of diazotrophic bacteria is also likely to be affected by soil type, as different soil fractions, such as sand-sized and silt-sized ones, harbored different N-fixing populations (Gros \textit{et al.}, 2006). Moreover, this was also true among microenvironments (Poly \textit{et al.}, 2001b), defined on the basis of the size of constituent particles, organic carbon and clay content. Temporal variations were shown to affect the level of cultivable diazotrophic population in soils, with the highest numbers in autumn/winter/early spring and with low counts in summer (Mergel \textit{et al.}, 2001). However, the low number of cultured bacteria found did not allow further generalization of these data.

In general, soil microbial communities are affected by a magnitude of natural fluctuations (such as temperature, plant growth or rainfall), and also agricultural practices, such as plowing (Buckley & Schmidt, 2001; Clegg \textit{et al.}, 2003). The most important factors that determine the diversity of microorganisms in soil are soil type and soil edaphic factors, plant type and soil management practices (van Overbeek & van Elsas, 2008). Nevertheless, the extent to which these factors affect the normal operating range (amplitude of variation under natural field conditions) of microbial communities remains unclear. In order to understand the functioning of microbial communities and their resilience to external changes, a key issue is to assess the community composition, quantify individual microbial population sizes and study fluctuations thereof (van Elsas \textit{et al.}, 2000; Hartmann & Widmer, 2006), as these fluctuations will form the reference against which external disturbances can be reflected upon. However, a thorough description of the natural variation of the N-fixing community in a wide range of soils is still missing.

In the present study, we characterized the dynamics and variation of the diazotrophic community structure using denaturing gradient gel electrophoresis (DGGE) and clone libraries. Furthermore, we used real-time PCR to assess community size. Eight different representative soils (both in terms of soil properties and biological factors) were used. Our main hypotheses were: (1) diazotrophic assemblages are responsive to biotic and abiotic parameters, ultimately affecting functional aspects and (2) soil type is the major factor influencing the structure and abundance of the N-fixing community. This study will provide insights into the basic parameters that influence the N-fixing community in our soils, which will allow us to set the baseline of this important functional group, across soil types, land use types and seasonal changes.

**Materials and methods**

**Experimental sites and soil sampling**

Eight fields across the Netherlands were selected based on their soil properties and cropping systems: four sandy (B, V, D, W) and four clay soils (S, K, G, L) (Table 1). These are all potato fields, subjected to crop rotation with nonleguminous plants, except for W, which is a permanent grassland soil. Bulk soil samples were collected four times over a growth season plus fall: April, June, September and November 2009. April and June are associated with the growing season, September represents the end of the growing season and November is associated with the raining season and no plants were present in the field. All eight soils were sampled on the same day at each sampling time. For each soil, four replicates were taken. Each replicate consisted of 10 subsamples (15–20 cm deep) collected between plots, away from the roots, with a spade. A total of 2 kg of each soil (0.5 kg per replicate) were thus collected in plastic bags and thoroughly homogenized before further processing in the lab.

**Soil chemical analysis**

Dried (24 h at 40 °C) soil samples were ground, sieved through a 2-mm mesh and analyzed. The pH was measured in a water suspension 1: 4.5 (g/v) using an Inolab Level 1 pH-meter (WTW GmbH, Weilheim, Germany). Nitrate (NO$_3^-$) and ammonium (NH$_4^+$) were determined calorimetrically in a solution of 0.01 M CaCl$_2$ using an Autoanalyzer II (Technicon Instrument Corporation, Tarrytown, NY). Dissolved organic carbon (DOC) was measured using a carbon analyzer in a soil extract of 0.01 M CaCl$_2$. Water content was measured by comparison of fresh and dried (48 h at 65 °C) weight of samples. Organic matter (OM) content is calculated as the difference between the initial and the final sample weights measured after 2 h at 550 °C, divided by the initial sample weight times 100%.

**Nucleic acid extraction**

For extraction of soil DNA, the PowerSoil DNA extraction kit (Mo Bio Laboratories Inc., New York) was used with 0.5 g of soil, according to the manufacturer’s instructions, with the exception that glass beads (diameter 0.1 mm; 0.25 g) were added to the soil slurries. The cells were disrupted by bead beating (mini-bead beater; BioSpec Products) three times for 60 s. To assess the quantity and purity, the crude DNA extracts were run on 1.5% agarose gels at 90 V for 1 h in 0.5 × TAE buffer (20 mM Tris, 10 mM acetate, 0.5 mM...
EDTA; pH 8.0) using a fixed amount (5 μL) of a 1-kb DNA ladder (Promega, Leiden, the Netherlands) as the molecular size and quantity marker. After staining with ethidium bromide, DNA quality was determined based on the degree of DNA shearing (average molecular size) as well as the amounts of coextracted compounds. Good-quality DNA was obtained from all soils and yields were between 0.5 and 3.2 μg DNA g⁻¹ soil.

**PCR amplification of the nifH gene for DGGE analysis**

PCR of nifH genes was conducted using a nested PCR according to Diallo et al. (2004), where a detailed protocol is described. The primers used in the first PCR reaction were FPGH19 (5'-TACGGCAARGGTGGNATHG-3'; Simonet et al., 1991) and PolR (5'-ATSGCCCATCAYTCTCCCGGA-3'; Poly et al., 2001a), and 2 μL of the first PCR product was used as a template in the second reaction with primers PolF containing a GC clamp (5'-TGCGAYCCSAARGCBGACTC-3') and AQER (5'-GACGATGTAGATYTCCTG-3') (both Poly et al., 2001a). The concentration of the PCR products was determined by a 1.5% (w/v) agarose TAE gel, staining with ethidium bromide to confirm product integrity and size by comparing with a molecular weight marker (Smart ladder; Eurogentec). All soils were PCR-amplifiable, resulting in one single band of the expected size (360 bp).

**DGGE systems for the detection of nifH**

DGGE profiles were generated using the Ingeny Phor-U system (Ingeny International, Goes, the Netherlands). PCR products (250–300 ng per lane) were loaded onto 6% (w/v) polyacrylamide gels, 1 mm thick, in 0.5 × TAE buffer with a 40–65% denaturant gradient [100% denaturant corresponded to 7 M urea and 40% (v/v) deionized formamide] to separate the generated amplicons. Electrophoresis was performed at a constant voltage of 100 V for 16 h at 60 °C. The gels were stained for 60 min in 0.5 × TAE buffer with SYBR Gold (final concentration 0.5 μg L⁻¹; Invitrogen, Breda, the Netherlands). Images of the gels were obtained with Imagemaster VDS (Amersham Biosciences, Buckinghamshire, UK) and stored as TIFF files.

**Computer-assisted analysis of DGGE fingerprinting**

DGGE patterns were compared by clustering the different lanes by Pearson’s correlation coefficient implemented in the GELCOMPAR II software (Applied Maths, Sint-Martens Latem, Belgium), using the unweighted-pair group method with arithmetic mean, rolling-disk background subtraction and no optimization (Rademaker et al., 1999; Kropf, 2004). Range-weighted richness (Rw) values (Marzorati et al.,

<table>
<thead>
<tr>
<th>Soillocation</th>
<th>pH</th>
<th>OM(%)</th>
<th>N-NH₄</th>
<th>N-NO₃</th>
<th>DOC(mg kg⁻¹)</th>
<th>N-NHa(mg kg⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandy</td>
<td>4.4</td>
<td>4.3</td>
<td>3.6</td>
<td>4.1</td>
<td>3.0</td>
<td>52.3</td>
</tr>
<tr>
<td>Buinen (B)</td>
<td>4.4</td>
<td>4.8</td>
<td>5.1</td>
<td>12.97</td>
<td>4.2</td>
<td>3.6</td>
</tr>
<tr>
<td>Valthermond (V)</td>
<td>4.4</td>
<td>4.3</td>
<td>5.1</td>
<td>5.9</td>
<td>3.4</td>
<td>70.9</td>
</tr>
<tr>
<td>Droevendaal (D)</td>
<td>4.5</td>
<td>5.1</td>
<td>5.5</td>
<td>2.85</td>
<td>3.0</td>
<td>65.6</td>
</tr>
<tr>
<td>Wildenja (W)</td>
<td>4.6</td>
<td>4.6</td>
<td>5.0</td>
<td>3.3</td>
<td>3.7</td>
<td>68.0</td>
</tr>
<tr>
<td>Clay</td>
<td>4.7</td>
<td>7.6</td>
<td>7.5</td>
<td>2.6</td>
<td>2.6</td>
<td>52.8</td>
</tr>
<tr>
<td>Kollumerward (K)</td>
<td>5.7</td>
<td>5.7</td>
<td>5.5</td>
<td>5.9</td>
<td>2.6</td>
<td>3.3</td>
</tr>
<tr>
<td>Steenharst (S)</td>
<td>5.7</td>
<td>7.5</td>
<td>7.6</td>
<td>5.4</td>
<td>4.8</td>
<td>108.5</td>
</tr>
<tr>
<td>Grebbendonk (G)</td>
<td>7.5</td>
<td>7.5</td>
<td>7.6</td>
<td>5.4</td>
<td>5.6</td>
<td>101.0</td>
</tr>
<tr>
<td>Lelystad (L)</td>
<td>7.5</td>
<td>7.4</td>
<td>7.7</td>
<td>2.8</td>
<td>2.6</td>
<td>57.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Soillocation</th>
<th>Soiltype</th>
<th>April</th>
<th>June</th>
<th>September</th>
<th>November</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandy</td>
<td>Sandy</td>
<td>4.7</td>
<td>4.6</td>
<td>4.7</td>
<td>4.6</td>
</tr>
<tr>
<td>Buinen (B)</td>
<td>Sandy</td>
<td>4.4</td>
<td>4.8</td>
<td>4.4</td>
<td>4.8</td>
</tr>
<tr>
<td>Valthermond (V)</td>
<td>Sandy</td>
<td>4.4</td>
<td>4.3</td>
<td>4.4</td>
<td>4.3</td>
</tr>
<tr>
<td>Droevendaal (D)</td>
<td>Sandy</td>
<td>4.5</td>
<td>5.1</td>
<td>4.5</td>
<td>5.1</td>
</tr>
<tr>
<td>Wildenja (W)</td>
<td>Clay</td>
<td>4.6</td>
<td>4.6</td>
<td>4.6</td>
<td>4.6</td>
</tr>
<tr>
<td>Clay</td>
<td>Clay</td>
<td>4.7</td>
<td>7.6</td>
<td>4.7</td>
<td>7.6</td>
</tr>
<tr>
<td>Kollumerward (K)</td>
<td>Clay</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
</tr>
<tr>
<td>Steenharst (S)</td>
<td>Clay</td>
<td>5.7</td>
<td>7.5</td>
<td>5.7</td>
<td>7.5</td>
</tr>
<tr>
<td>Grebbendonk (G)</td>
<td>Clay</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Lelystad (L)</td>
<td>Clay</td>
<td>7.5</td>
<td>7.4</td>
<td>7.5</td>
<td>7.4</td>
</tr>
</tbody>
</table>

Table 1. Soil characteristics measured in this study.

Values are averages of three replicates.
2008) were calculated based on the total number of bands (N), and the denaturing gradient comprised between the first and the last band of the pattern (Dg), according to the following equation: \( R_s = N^2 \cdot D_g \). To assess the interspecies abundance ratios, Pareto–Lorenz (Lorenz, 1905) curve distribution patterns of the \( \text{nifH} \) DGGE profiles were plotted based on the numbers of bands and their intensities to visualize the functional organization (Fo) of the diazotrophic community over time (Mertens et al., 2005; Marzorati et al., 2008). Data derived on the basis of Jaccard correlation (a band-based analysis) were used for principal component analysis (PCA) using CANOCO (version 4.52, Wageningen, the Netherlands). The matrix of similarities was also used to perform moving window analysis (Wittebolle et al., 2005) and to calculate the percentage of change, as change\% = 100 – similarity\%.

**Quantification of the N-fixing community**

To quantify the number of copies of the \( \text{nifH} \) gene, the primers FPQH19 (Simonet et al., 1991) and PoIR (Poly et al., 2001a) were used. The PCR mixture and thermal cycling conditions are described in Taketani et al. (2009). Absolute quantification of the \( \text{nifH} \) gene was carried out in four replicates on the ABI Prism 7300 Cycler (Applied Biosystems, Germany). The specificity of the amplification products was confirmed by melting curve analysis, and the expected size (450 bp) of the amplified fragments was checked in a 1.5% agarose gel stained with ethidium bromide. Standard curves were obtained using serial dilutions of the Escherichia coli-derived vector plasmid JM 109 (Promega, Madison, WI) containing a cloned \( \text{nifH} \) gene from Bradyrhizobium liaoningense, using 10\(^{-2}\)–10\(^{2}\) gene copy numbers \( \mu \text{L}^{-1} \). The efficiency was calculated using the formula Eff = \[ 10^{1/C_{0}} - 1 \]. Two independent qPCRs were performed for all samples and the results were similar. To test for inhibition in the PCR reactions, DNA extracted from soil was diluted and mixed with a known amount of standard DNA before qPCR. The \( C_t \) values obtained for the standard DNA did not change in the presence of diluted soil DNA, indicating the absence of severe inhibition.

**Construction of \( \text{nifH} \) clone libraries**

DNA extracted from soil collected in June was used to construct \( \text{nifH} \)-based clone libraries for each of the eight soils. The four replicates of each soil were used in PCR amplifications with the primers PoLF and AQER (both Poly et al., 2001a). A detailed protocol is described in Diallo et al. (2004). The products (size 320 bp) were pooled per soil and gel-purified using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI). The purified PCR products were then inserted into the vector pGEM-T-Easy (Promega, Madison, WI) and introduced into competent E. coli JM 109 cells in accordance with the manufacturer’s instructions (Promega, Madison, WI). Clones containing the insert were sequenced using an Applied Biosystems 3730 XL DNA Analyzer at LGC Genomics GmbH (Berlin, Germany).

**Phylogenetic analysis**

The \( \text{nifH} \) gene sequences were compared with sequences in GenBank using nucleotide–nucleotide BLAST (BLAST-N) to obtain the nearest phylogenetic neighbors (http://www.ncbi.nlm.nih.gov/BLAST). Sequences were checked for chimera using BELLEROPHON v.3 (http://greengenes.lbl.gov) (Huber et al., 2004). Sequences were then processed in MEGA 4 (Tamura et al., 2007), translated and the deduced amino acid sequences were aligned using CLUSTAL W (Jeanmougin, 1998). Phylogenetic trees were constructed using the neighbor-joining method with the Jones–Taylor–Thornton model for amino acid substitution. In addition, we performed maximum-likelihood analyses (Olken et al., 1994). Both trees showed similar topologies. Bootstrapping (500 replicates) was used to estimate the reliability of the phylogenetic reconstructions. The program DOTUR [DISTANCE-BASED OTU AND RICHNESS (Schloss & Handelsman, 2008); http://www.schloss.micro.umass.edu/software] was used to create rarefaction curves, where a conservative operational taxonomic units cutoff of 97% similarity was used to determine the Shannon diversity index, as well as the bias-corrected Chao1 estimator of richness. Differences in the community structures of \( \text{nifH} \) clone libraries were analyzed with UNIFRAC (Lozupone et al., 2006), using maximum-likelihood-based trees.

**Statistical analysis**

Physicochemical and biological variables were checked for normality and transformed when necessary. Differences in these variables between sandy and clay soils, among all eight soils, and over time, were assessed using independent sample \( t \)-tests. In order to check for type I error, corrections methods (Bonferroni, Hochberg and false discovery rate) were also implemented in multiple \( t \)-tests. Correction methods did not lead to changes in the results obtained, and therefore were not used in the statistical analyses based on \( t \)-test. Multiple regression analyses were conducted on normalized data to avoid possible nonlinear relations (SAS\textsuperscript{®} system for Windows version 8.02, SAS Institute Inc, Cary, NC, 2001). The following parameters were included in the analysis: diversity (Shannon diversity index \( H' \) of clone libraries) and abundance (\( A \)) of N-fixing community; nitrate in mg kg\(^{-1} \) (\( \text{NO}_3 \)); ammonium in mg kg\(^{-1} \) (\( \text{NH}_4^+ \)); pH; and DOC in mg kg\(^{-1} \) and OM in %. Variables in the regression models were significant at the 0.1 level. Models
were restricted to a maximum of two parameters. Multiple regressions were conducted for each soil separately.

**Nucleotide sequence accession numbers**

The sequences generated in this study have been deposited in the GenBank database under accession numbers HQ335394–HQ336041.

**Results**

**Physicochemical characteristics of soils**

Soil chemical analyses were performed across a season and all soils revealed significant differences between various parameters (Table 1). Regarding seasonal variations, there were significant \( P < 0.05 \) differences in the levels of \( \text{NO}_3^- \), from June to September, and in the levels of \( \text{NH}_4^+ \), from June to September and September to November. Small, but nonsignificant, variations of \( \text{pH} \), \( \text{OM} \) and \( \text{DOC} \) were found over time (Table 1). Considering soil type, \( \text{pH} \) was significantly higher \( \left( P < 0.05 \right) \) in the clay \((6.94 \pm 0.11, n = 12)\) than in the sandy soils \((4.78 \pm 0.08, n = 12)\) during the entire period. The \( \text{NO}_3^- \) and \( \text{NH}_4^+ \) concentrations \( (\text{mg kg}^{-1}) \) were significantly higher in the sandy soils \((\text{NO}_3^-: 75.14 \pm 11.86; \text{NH}_4^+: 7.96 \pm 0.66, n = 12) \) than the clay soils \((\text{NO}_3^-: 21.31 \pm 3.87; \text{NH}_4^+: 5.66 \pm 0.22, n = 12) \) in June and September only. In April and November, the differences were not significant. Although bulk soil was always analyzed, the crops present in the field over the season were barley (soils B and V), triticale (cross between wheat and rye; soil D), potato (soils K, G and L) and grass (soils W and S).

**Abundance of \textit{nifH}-harboring bacteria**

Overall, the population sizes of N-fixing bacteria, quantified by real-time PCR targeting the \textit{nifH} gene, varied from \(10^5\) gene copies g\(^{-1}\) dw in April to \(10^7\) gene copies g\(^{-1}\) dw in November (Fig. 1a). Grouping of the soils according to their texture revealed that, on average, clay soils had higher \textit{nifH} abundances \(\left( \log 6.4 \text{ gene copies g}^{-1} \text{ dw} \right)\) than sandy ones \(\left( \log 5.6 \text{ gene copies g}^{-1} \text{ dw} \right) \) \(\left( P < 0.05, n = 16 \right)\) (Fig. 1b). Analyses across the season revealed that the \textit{nifH} gene abundance in the sandy soils was significantly lower in April \(\left( P < 0.05 \right)\) than in November, but no significant changes were found from June to September. The \textit{nifH} gene abundance in the clay soils showed a higher variation, increasing significantly from April to June, and decreasing significantly from June to September (Fig. 1b).

![Fig. 1](https://academic.oup.com/femsec/article-abstract/77/1/57/528994) Real-time quantification of \textit{nifH} gene across eight soils. Absolute numbers for each sampling time (a) and average abundance values between sandy and clay soils (b). Soils used in this work: V, Valthermond; B, Buiten; W, Wildekamp; D, Droevendaal; K, Kollumerwaard; S, Steenharst; G, Grebbedijk; L, Lelystad. Bars are SEs (A, \( n = 4 \); B, \( n = 16 \)).
Community structure analysis based on the nifH gene

In order to characterize the structure and dynamics of the N-fixing communities through time across different soils, we performed DGGE analyses based on the nifH gene. Analysis of the DGGE gels revealed that the numbers of bands per sample varied between 35 ± 1 and 14 ± 1, and decreased from April to November for both sandy and clay soils. The range-weighted richness decreased significantly from April (122 ± 10.4, n = 12) to November (26.5 ± 8.2, n = 12) (Supporting Information, Fig. S1), being significantly higher in the sandy soil in June (Fig. 2). It is worthwhile mentioning that W soil (grassland) had, on average, the lowest \( R_r \) value. Multivariate analyses (PCA) of the nifH DGGE profiles showed that the profiles tend to cluster by soil type in April and June, along the first and second axes, respectively (Fig. S2a and b). In September and November, the profiles formed a diffuse cluster, and no clear trend was observed (Fig. S2c and d).

The dynamics of N-fixing community were established with the moving window analysis and showed that all eight soils had pronounced changes through time, with the amplitude of variations being higher in the clay soils (\( \Delta t_s = 65.19 ± 14.69\% \)) attaining significance in November, compared with the sandy ones (\( \Delta t_s = 61.64 ± 6.20\% \)) (Fig. 3). The functional organization was comparable, as 20% of the bands (number based) corresponded to, on average, 45% of the cumulative band intensities (Fig. S3).

Diversity of nifH-harboring bacteria

In order to identify and characterize the dominant types of N-fixing bacteria across different soils, we constructed one clone library per soil, based on the nifH gene, using the samples collected in June 2009, by pooling the nifH products obtained from four replicates before cloning. The results first showed that the similarity of the deduced sequences to known nifH sequences (database) varied from 81% to 100%. Around 11% of the sequences remained unclassified (< 89% similarity), and these numbers were higher (15.11%) in the clay soils than in the sandy ones (6.02%) (Table 2). Among the sequences that could be classified, the two most dominant classes were Alphaproteobacteria and Betaproteobacteria (respectively 62.5% and 16.2% on average), the Betaproteobacteria class being more abundant in the clay soils (29.5%) than in the sandy ones (3.0%). The other classes that were found varied enormously depending on the type of soil that was analyzed (see Table 2). Sequences affiliated with nifH genes of Azospirillum and Bradyrhizobium were found in all eight soils, whereas Rhizobium nifH-related sequences were found only in two of the eight soils (Table 2 and Table S1).

No clear trend could be observed between soil texture and the diversity of N fixers based on both rarefaction and the
Chao1 estimator of richness (Fig. 4a and b). However, the Shannon diversity index based on the same data set revealed a significantly higher diversity in the clay ($H^0 = 2.98$) than in the sandy ($H^0 = 2.43$) soils. Results from UNIFRAC analysis revealed that the composition of the $nifH$-containing community in the four sandy soils did not differ significantly among each other, nor did the community in the clay soils ($P = 0.06$). However, when all environments were included in one comparison, the differences between sequences’ composition in sandy vs. clay soils were significant ($P < 0.01$, after Bonferroni correction). This was confirmed by jack knife analysis of the clustering (Fig. 5).

### Influence of physicochemical and biological soil characteristics on the N-fixing community

Stepwise multiple regressions were performed to identify the factors that could explain the observed variation in gene abundance (qPCR) and the Shannon diversity index (clone library). Both were normally distributed (Shapiro–Wilk test: $P = 0.53$ and $P = 0.83$, respectively). The best predictor ($P < 0.05$) for the abundance of the N-fixing community for all eight soils during the entire vegetation period was soil pH, explaining 20.7% of the variation (Table 3). The remaining variation in the abundance of the N-fixing community over all soils was best explained by DOC ($P < 0.05$). Regarding the sandy soils, the variation in the abundance of N-fixing community was best explained by pH and OM content. In the case of clay soils, the level of $NH_4^+$ was the main predictor of $nifH$ gene abundance (Table 3). The variation observed in $nifH$ gene diversity was best predicted by $NH_4^+$ and DOC, with a negative ($P < 0.05$) influence of $NH_4^+$ that explained 66.5% of the variation.

### Discussion

Considering the importance of N fixation for N cycling and crop production, we focused our study on the soil
diazotrophic communities, which were explored in respect of their natural fluctuations across eight selected Dutch soils, throughout a growing season. We used direct molecular methods, based on soil DNA, focusing on the \emph{nif}H gene. In order to target the \emph{nif}H gene, we used sets of primers, based on those developed by Poly \textit{et al.} (2001a), that amplified fragments of different sizes. More specifically, the primer sets used for quantitative PCR amplified a longer fragment (450 bp), whereas the fragment amplified for cloning and DGGE was shorter (320 bp), the sequence of the latter being nested in the former, which reduces potential pitfalls when comparing the data.

Variations in the abundance of diazotrophic communities

Clay fractions in soils are important in imparting specific physical properties, forming micro- and macroaggregates (Gupta & Roper, 2010), and providing microaerophilic or anaerobic conditions that are favorable to N fixation. This knowledge is consistent with our findings, which showed higher abundances of \emph{nif}H genes in the clay soils compared with the sandy ones. Indeed, it has been shown that 70\% of the free-living N-fixing bacteria are located in the clay fraction (Chotte \textit{et al.}, 2002). Alternatively, the higher \emph{nif}H gene abundance in the clay soils could be explained by the higher pH of these soils, which has been shown to favor the potential for BNF (Roper & Smith, 1991; Nelson & Mele, 2006). The results from our multiple regression analysis supported this latter conclusion, as soil pH was positively correlated with the overall \emph{nif}H gene abundance at all four sampling times. Interestingly, the effect of soil pH on the abundance of diazotrophs varied across the season, being stronger in June, when it explained almost half of the variation in \emph{nif}H gene copy numbers. The availability of carbon is another factor that affects the level of N fixation in soils, the carbon becoming available from rhizodeposition or via retention of crop residues (Gupta & Roper, 2010). Indeed, in our work, soil DOC had an overall positive effect on \emph{nif}H gene abundance, albeit to a lesser extent than pH. Similarly, Morales \textit{et al.} (2010) found that the abundance of \emph{nif}H genes was positively correlated with organic carbon levels in agricultural soils.
It is important to realize that the primers used in this study are degenerated, and thus represent a mixture of primers with different binding affinities for different templates. Although degeneracy was necessary, as it allows for greater coverage of genes that are highly polymorphic, amplified DNA can only reflect the quantitative abundance of the species if the amplification efficiencies are the same for all molecules (von Wintzingerode et al., 1997), and degeneracy may influence the formation of primer–template hybrids. Thus, although it is possible that our approach underestimates the actual gene copy numbers, it provides an indication of the size of the nifH-harboring communities. Appropriate controls to assess this bias are rarely performed due to feasibility and costs in environmental studies, but it is an important concern regarding the validity of the qPCR results.

In summary, our results showed that the abundance of diazotrophic bacteria was highly correlated with soil pH, the more neutral pH being most favorable for nifH gene abundance. Moreover, it appeared that soil texture was important. As the concentration of clay in soil can be correlated with nitrogenase activity, we hypothesized that this is due to a higher abundance of N-fixing bacteria (Roper & Smith, 1991). However, determination of whether N fixation rates correlate with nifH-containing community size will require further analyses, as we did not measure nitrogenase activities in our soils. Future microcosm experiments, aimed to establish the level of correlation between nifH gene abundance and potential N fixation and to explore the effect of soil type (either due to pH or texture), will provide more concrete evidence.

Structure of N-fixing communities

Analysis of the N-fixing communities in the eight soils revealed diverse and dynamic diazotrophic assemblages. Soil type had a greater effect in the beginning than later on in the growing season. It has been shown that agricultural practices, such as fertilization and ploughing, play a major role as determinants of bacterial community structure in soil (Patra et al., 2006; Salles et al., 2006; Wakelin et al., 2009) and these might explain the stronger cluster in the beginning of the growing season.

Using a conceptual framework proposed by Marzorati et al. (2008), we explored possible ecological interpretations based on the structure of diazotrophic communities as determined by DGGE. We observed that the carrying capacity of the soils was significantly higher in the beginning of the season, as determined by the range-weighted richness (Marzorati et al., 2008), indicating that the environmental conditions in this period were more favorable to diazotrophic communities than at the end of the year. Despite the changes in richness observed over time, the functional organization of the community, as analyzed by Pareto–Lorenz curves, showed small seasonal variation, indicating communities that were highly structured over time (Marzorati et al., 2008). The reduction in richness towards the end of the growing season and the increase in community size indicate that these communities (November) are dominated by diazotrophic species that are better adapted to the higher concentrations of NH$_4^+$, regardless of soil type.

Regarding the use of DGGE-based approaches, the number and intensity of bands in a gel do not necessarily provide an accurate picture of the microbial community due to the fact that one organism may produce more than one band. Because of these drawbacks, the parameters calculated from DGGE fingerprints should be interpreted as indications and not absolute measurements of microbial community structure and diversity. Nevertheless, the ecological parameters used to analyze diazotrophic community structure in this study were able to provide a better characterization of the N-fixing community in our soils.

Table 3. Stepwise multiple regressions showing to what extent the variation of biological/chemical parameters can explain the variation of nifH gene abundance over time and the diversity of N-fixing communities, as determined by clone library based on samples collected in June.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>nifH gene Abundance</th>
<th>pH</th>
<th>NO$_3$</th>
<th>NH$_4^+$</th>
<th>DOC</th>
<th>OM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall</strong></td>
<td>+0.2077**</td>
<td>+0.1037**</td>
<td>+0.2230**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sandy</strong></td>
<td>+0.0396*</td>
<td>+0.1676**</td>
<td>-0.2157**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clay</strong></td>
<td>+0.2512*</td>
<td>+0.3852**</td>
<td>-0.6645**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>April</strong></td>
<td>+0.4191**</td>
<td>+0.2556**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>June</strong></td>
<td>+0.2535**</td>
<td>+0.1925**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>September</strong></td>
<td>+0.0953*</td>
<td>+0.1450**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>November</strong></td>
<td>-0.6645**</td>
<td>+0.2157**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

No value: no significant relation.
*Significant at the 0.15 level (two tailed).
**Significant at the 0.05 level (two tailed).
+ represents a positive influence on the diversity and abundance of N-fixing community.
Diversity of N-fixing communities

In order to identify and characterize dominant nifH types, we selected soil samples from June, when significant correlations had been found based on nifH gene abundance, to determine the diversity of diazotrophic communities across the eight soils. The sequences recovered from the clone libraries were affiliated with several major groups of bacteria. Not surprisingly, the class Alphaproteobacteria was the most dominant class found in our soils, as has been found before (Smit et al., 2001; Buckley & Schmidt, 2003). In our case, it was mostly represented by the order Rhizobiales, in particular, the genus Bradyrhizobium. UNIFRAC analyses showed two clear clusters formed by sand- and clay-derived sequences, confirming the results obtained by PCR analysis of DGGE patterns. Moreover, the soils under study differed considerably in terms of diversity, higher in clay than in sandy soils, and that could be largely explained by the level of NH$_4^+$+. The negative effect of ammonia on N fixation has been known for a long time, as the expression of nitrogenase is very often inhibited by the presence of NH$_4^+$+ (Brotonegro, 1974; Houwaard, 1978; Christiansen-Weninger & van Veen, 1991). The decrease in nifH gene diversity could thus be explained by assuming the selection of less sensitive types under ammonia pressure. As mentioned previously, the differential clustering of the soils by texture (sand and clay) could be explained not just by assuming a role for texture, but, alternatively, by invoking one for pH. Additional experiments are, however, needed to differentiate the effects of texture and pH on diazotrophic communities.

Normal operating range of diazotrophic communities

A high turnover in community composition was observed for N-fixing communities, varying from 60% to 65% throughout the season. This internal structuring of the highly dynamic diazotrophic community indicated that at each season, different species attained significant dominance. Furthermore, the structure and size of diazotrophic communities differed considerably between clay and sandy soils, the former displaying higher abundances and diversity of nifH genes across a range of Dutch soils. Moreover, clay soil showed higher amplitude of variation in community size, which indicates that the communities associated with this soil type might be more sensitive to fluctuations associated with the season and agricultural practices. Thus, diazotrophic communities associated with sandy and clay soil were shown to have distinct normal operating ranges. Further investigation is necessary to unravel to what extent the amplitude of variation in community size and structure observed for clay soils affects the functioning of these communities and their ability to cope with environmental stresses.

In conclusion, our results indicate a highly dynamic diazotrophic community, whose structure, size and diversity were mainly driven by soil type, pH and NH$_4^+$+, respectively. Although these parameters were known to affect N-fixing communities, our analyses allowed us to quantify their relevance for different components of diazotrophic communities across both a temporal and a spatial scale.

Acknowledgements

This work was supported by the NWO-ERGO Program and was part of a collaborative project with Utrecht University, Utrecht, the Netherlands. We would like to thank our Utrecht partners, Heike Schmitt and Agnieszka Szturc, for their help with sampling and interesting discussions; and Cyrus A. Mallon for his comments on the manuscript. The authors thank Stefanie Töwe, from the Institute of Soil Ecology, German Research Center for Environmental Health (GmbH), Neuherberg, for support and help with real-time PCR data analysis. The authors also thank the reviewers for their valuable comments.

References

Coelho MRR, de Vas M, Carneiro NP, Marriel IE, Paiva E & Seldin L (2008) Diversity of nifH gene pools in the rhizosphere of two cultivars of sorghum (Sorghum bicolor) treated with...


**Supporting information**

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Range-weighted richness (*R*) of the eight soils over time.

Fig. S2. Ordination biplot diagrams generated by PCA of N-fixing communities in the eight soils across the four sampling times: April (a), June (b), September (c) and November (d).

Fig. S3. Pareto–Lorenz distribution curves based on DGGE profiles of the N-fixing community over time, based on average values of the sandy and clay soils: April (a), June (b), September (c) and November (d).

Table S1. Distribution of percentages of *nif*H sequences of some species found in the soil clone libraries.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.