Microb Ecol (2016) 71:604–615
DOI 10.1007/s00248-015-0659-7

ENVIRONMENTAL MICROBIOLOGY

The Diversity and Co-occurrence Patterns of N_2-Fixing Communities in a CO_2-Enriched Grassland Ecosystem

Qichao Tu[1,2] · Xishu Zhou[2,3] · Zhili He[3] · Kai Xue[3] · Liyou Wu[3] · Peter Reich[4,5] · Sarah Hobbie[4] · Jizhong Zhou[2,6,7]

Received: 8 February 2015 / Accepted: 3 August 2015 / Published online: 18 August 2015

Abstract Diazotrophs are the major organismal group responsible for atmospheric nitrogen (N_2) fixation in natural ecosystems. The extensive diversity and structure of N_2-fixing communities in grassland ecosystems and their responses to increasing atmospheric CO_2 remain to be further explored. Through pyrosequencing of *nifH* gene amplicons and extraction of *nifH* genes from shotgun metagenomes, coupled with co-occurrence ecological network analysis approaches, we comprehensively analyzed the diazotrophic community in a grassland ecosystem exposed to elevated CO_2 (eCO_2) for 12 years. Long-term eCO_2 increased the abundance of *nifH* genes but did not change the overall *nifH* diversity and diazotrophic community structure. Taxonomic and phylogenetic analysis of amplified *nifH* sequences suggested a high diversity of *nifH* genes in the soil ecosystem, the majority belonging to *nifH* clusters I and II. Co-occurrence ecological network analysis identified different co-occurrence patterns for different groups of diazotrophs, such as *Azospirillum*/Actinobacteria, *Mesorhizobium*/Conexibacter, and *Bradyrhizobium*/Acidobacteria. This indicated a potential attraction of non-N_2-fixers by diazotrophs in the soil ecosystem. Interestingly, more complex co-occurrence patterns were found for free-living diazotrophs than commonly known symbiotic diazotrophs, which is consistent with the physical isolation nature of symbiotic diazotrophs from the environment by root nodules. The study provides novel insights into our understanding of the microbial ecology of soil diazotrophs in natural ecosystems.

Keywords *nifH* · Soil diazotrophs · Community structure · Co-occurrence patterns · Elevated CO_2

Qichao Tu and Xishu Zhou contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00248-015-0659-7) contains supplementary material, which is available to authorized users.

✉ Jizhong Zhou
jzhou@ou.edu

1 Department of Marine Sciences, Ocean College, Zhejiang University, Hangzhou, Zhejiang 310058, China

2 Institute for Environmental Genomics and Department of Microbiology and Plant Biology, The University of Oklahoma, Norman, OK 73019, USA

3 School of Minerals Processing and Bioengineering, Central South University, Changsha, Hunan 410083, China

4 Department of Forest Resources, University of Minnesota, St. Paul, MN 55455, USA

5 Hawkesbury Institute for the Environment, University of Western Sydney, Richmond 2753, NSW, Australia

6 Earth Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

7 State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China

Introduction

Biological nitrogen fixation (BNF), the reduction of atmospheric N_2 to biologically available ammonium, is the major pathway that atmospheric N_2 enters the Earth's biosphere, contributing approximately 128 Tg nitrogen (N) per year in natural terrestrial ecosystems [1]. BNF is catalyzed by diverse but limited groups of nitrogenase-containing bacteria and archaea known as diazotrophs. Among the nitrogenase subunits, the nitrogenase reductase subunit, encoded by *nifH*, has the most reference sequences available and has become a promising marker for analyzing N_2 fixation of microbial communities in various environments [2, 3]. Although phylogenetic

analysis of taxonomically identified nitrogenase genes provided evidences of an ancient horizontal gene transfer between archaea and bacteria, recent events of horizontal gene transfer were not observed [4].

During the past decades, the *nifH* gene family has been widely analyzed to study diazotrophic microbial communities in various environments, especially in marine and soil ecosystems [5–13], resulting in novel insights into the diversity and structure of N_2-fixing communities. For example, even though *Trichodesmium* has been found as the major N_2-fixing cyanobacteria in marine ecosystems [6], a recent study suggested that the contributions of other N_2 fixers are far more significant than previously estimated [14], indicating the important roles of less dominant N_2 fixers. In soil, N_2 fixation is dominated by symbiotic bacteria that form root–nodule symbiotic relationships with plants [15]. Similar to that in ocean, it is believed that both symbiotic and free-living diazotrophs contribute significantly to the terrestrial N budget [15]. Although many previous studies focused on the relationship between symbiotic diazotrophs and plants, the diversity and community structure of diazotrophic communities were recently analyzed by several studies [7, 8, 11–13, 16]. It has also been pointed out that the N_2 fixation rates in soil were significantly affected by diazotrophic community structure [11]. However, the extensive diversity and complex community structure in soil ecosystems remain to be further explored.

Natural ecosystems under increased atmospheric CO_2 concentration are subjected to progressive N limitation [17–20] due to the stimulated plant growth rate and limited biologically available N in soil. Such progressive N limitation not only constrains the sustainability of ecosystem responses to elevated CO_2 (eCO_2) [18, 19] but may also suppress the microbial decomposition rate in soil [17]. As biologically available N mainly comes from the microbial decomposition of biomass and BNF, the stimulated plant growth and suppressed microbial decomposition should have proposed higher demand for BNF in soil. Thus, it is of crucial interest for ecologists to understand how the belowground diazotrophic microbial community responds to eCO_2. A recent study of N_2-fixing bacteria communities in forest ecosystems suggested that N fertilization has a stronger effect on the diazotrophic community than eCO_2. However, the responses of diazotrophic community diversity and structure in the grassland ecosystem, one of Earth's largest ecosystems, are still not clear yet, although a previous GeoChip survey suggested increased *nifH* abundance in this same BioCON experimental site [21, 22].

Similar to that in macroecosystems, interactive relationships should be considered when analyzing the community structure of microorganisms. Using ecological network approaches, co-occurrence ecological networks of microbial communities can be constructed and analyzed [23–27]. By implementing co-occurrence ecological network approaches, co-occurrence patterns can be identified for diazotrophic microorganisms, providing novel insights into how microorganisms potentially interact with diazotrophs. Since symbiotic diazotrophs can enter plant roots, forming nodules that physically isolate them from the environment, they are less likely to form complex relationships with free-living microorganisms, and co-occurrence networks may provide information to potentially identify free-living diazotrophs from symbiotic diazotrophs.

In this study, by coupling sequencing *nifH* amplicons with the extraction of shotgun metagenome sequencing data and co-occurrence ecological network analysis, we aimed to reveal the responses of soil diazotrophs to eCO_2 and to determine the diversity and structure of soil diazotrophs, as well as their co-occurrence patterns in the BioCON grassland ecosystem [28]. The following hypotheses were tested: (1) Increased plant growth would enhance the demand for biologically available N in soil, which may lead to increased *nifH* gene abundance as well as changed diazotrophic community diversity and structure; and (2) Free-living rather than symbiotic diazotrophs would form more complex co-occurrence ecological networks and that different co-occurrence patterns would be observed for different diazotrophs. Our results indicated that long-term eCO_2 significantly increased *nifH* gene abundance but did not change the overall *nifH* community structure. Ecological network analysis identified several types of co-occurrence patterns for soil diazotrophs, such as *Azospirillum*/Actinobacteria, *Mesorhizobium*/Conexibacter, and *Bradyrhizobium*/Acidobacteria, indicating a potential attraction of non-N_2-fixers by soil diazotrophs. The study provided valuable insights into our understanding of microbial ecology of diazotrophs in soil.

Materials and Methods

Site Description and Sample Collection

The study was conducted within the BioCON (Biodiversity, CO_2, and Nitrogen) experimental site located at the Cedar Creek Ecosystem Science Reserve in Minnesota, USA (45.4086° N, 93.2008° W). The long-term experiment was started in 1997 on a secondary successional grassland on a sandy outwash soil after removing the previous vegetation [28]. The main BioCON field experiment has 296 (of a total of 371) evenly distributed plots (2×2 m) in six 20-m-diameter FACE (free air CO_2 enrichment) rings, three with ambient CO_2 (aCO_2) concentrations, and three with CO_2 concentrations elevated by 180 μmol/mol [29]. In this study, 24 plots (12 from aCO_2, 12 from eCO_2, all with 16 species and without additional N supply) were used.

Bulk soil samples were taken in July 2009 under ambient and eCO_2 conditions for microbial community analysis, and each sample was composited from five soil cores at a depth of 0–15 cm. All samples were immediately transported to the

laboratory, frozen, and stored at −80 °C for DNA extraction, PCR amplification, and 454 pyrosequencing.

The aboveground and belowground (0–20 cm) biomass were measured as previously described (Reich et al. 2001; Reich et al. 2006). Net N mineralization rates were measured concurrently in each plot for 1-month in situ incubations with a semi-open core at 0–20-cm depth during midsummer of each year (Reich et al. 2001; Reich et al. 2006). Net N mineralization rates were determined by the difference between the final and initial NH_4^+–N+NO_3^-–N pool sizes determined with 1 M KCl extractions. Net NH_4^+ change was determined by the difference between the final and initial NH_4^+–N pool sizes. Net NO_3^- change was determined by the difference between the final and initial NO_3^--N pool sizes.

DNA Extraction, Purification, and Quantification

Soil DNA was extracted by freeze-grinding mechanical SDS-based lysis as described previously [30] and was purified using a low melting agarose gel followed by phenol extraction for all 24 soil samples collected. DNA quality was assessed by the ratios of 260/280 and 260/230 nm using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA), and final soil DNA concentrations were quantified with PicoGreen [31] using a FLUOstar Optima (BMG Labtech, Jena, Germany).

PCR Amplification and 454 Pyrosequencing

A total of 23 samples instead of 24 were subjected to 454 pyrosequencing due to insufficient remaining DNA and soil for one of the samples. Amplification was performed using the *nifH* PolF/PolR primers (PoLF: TGCGAYCCSAARGCBGACTC and PolR: ATSGCCATCATYTCRCCGGA), whose products are expected to be approximately 362 bp [32]. A unique 8-mer barcode was added for each sample at the 5′-end of the forward primer. The barcode primers were synthesized by Invitrogen (Carlsbad, CA, USA) and used for the generation of PCR amplicons. Quadruplicate 20-μl PCR reactions were performed as follows: 4 μl Promega GoTaq buffer, 1 U DNA polymerase, 1.5 μl Roche 25 mM $MgCl_2$, 1 μl Invitrogen 10 mM dNTP mix, 1 μl of each primer (10 μM), 0.2 μl New England BioLabs BSA (non-acetylated, 10 mg/ml), 10 ng template, and 9.8 μl H_2O. Cycling conditions were an initial denaturation of 94 °C for 3 min, 30 cycles of 94 °C for 1 min, 62 °C for 40 s, 72 °C for 1 min, and a final extension at 72 °C for 10 min. PCR products were gel-purified using the Qiagen Gel Purification Kit following band excision. Products were further purified using the Qiagen PCR purification kit. After adapter ligation, amplicons were sequenced on a FLX 454 system (454 Life Sciences, Branford, CT, USA) by Macrogen (Seoul, South Korea) using Lib-L kits and processed using the shotgun protocol.

Data Analysis

Raw pyrosequencing reads were extracted from the *sff* file using the *sffinfo* tool from Roche 454. Two files, a *fasta* file containing the sequence and a *qual* file containing the quality information, were generated and then converted into a *fastq* file using the python script "faqual2fastq2.py" that comes with the UPARSE pipeline [33]. Quality filtering, chimera removal, and OTU clustering were carried out using the UPARSE pipeline [33], which is a recently developed approach that identifies highly accurate OTUs from amplicon sequencing data. Only the reads with perfectly matched barcodes and maximum of two primer mismatches were kept for further analysis. Barcodes and primers were deleted from reads. The remaining reads were then truncated to 300 bp, and reads with expected errors >0.5 were discarded. The program FrameBot [12] was used to correct potential frame shifts caused by sequencing errors, and only reads whose translated proteins got mapped to reference *nifH* protein sequences with >30 % identity were kept. The reads were then dereplicated, sorted, and clustered into candidate OTUs with an identity cutoff of 0.94, which is the average nucleotide identity being used for microbial species definition in post-genomic era [34]. Chimeric OTUs were then identified and removed using uchime by searching against the *nifH* reference sequences maintained and curated by Zehr et al. (http://pmc.ucsc.edu/~wwwzehr/research/database/) [2]. Finally, qualified reads were mapped to representative OTU sequences for relative abundance calculation.

Taxonomic assignment for *nifH* OTUs was carried out by searching representative OTU sequences against reference *nifH* sequences with known taxonomic information. A minimum global identity cutoff of 80 % was used to filter searching results. A lowest common ancestor algorithm similar as described in MEGAN [35] was applied for taxonomic assignment based on the best hits with highest global identity. Taxonomic information at genus level or higher was assigned. For phylogenetic analysis, representative OTU sequences were aligned by the MUSCLE program [36]. The alignment was visually inspected, and no manual curation was needed based on the inspection. Phylogenetic tree was built by FASTTREE [37]. Significance tests for different taxonomic groups and OTUs were performed by response ratio analysis [38] at 95 % confidence interval level. UniFrac PCoA analysis was done by the FastUniFrac pipeline [39]. Species richness, evenness, and diversity indices were calculated by the Mothur package [40], with rarefaction analysis of 1000 bootstrap random sampling iterations and 0.1 % incremental resampling efforts.

Two additional datasets including one 16S amplicon sequencing and one shotgun metagenome dataset from the same 24 soil DNA were also used in this study. The 16S amplicon sequencing data were used for network construction, and the

shotgun metagenome data set was used for *nifH* gene abundance comparison purposes. The 16S amplicon data were processed in the same way as the *nifH* amplicon data, except that 97 % identity cutoff was used. The shotgun metagenome dataset was sequenced by Los Alamos National Lab with standard 454 sequencing protocols and was annotated against eggNOG database [41] for gene abundance profiles. All metagenome data used in this study were deposited in the NCBI database and can be found under accession number SRP034704.

Co-occurrence Ecological Network Construction

In order to identify co-occurrence relationships between diazotrophs and other microbial species, a 16S rRNA gene amplicon dataset from the same DNA was also included. Relative abundance profiles were generated for both *nifH* (random subsampling of 1200 reads) and 16S rRNA (random subsampling of 18,000 reads) OTUs. Co-occurrence ecological networks were constructed and analyzed using the online MENA pipeline, which implements random matrix theory (RMT) for threshold identification [42]. The RMT approach identifies the threshold by determining a transition point of nearest-neighbor spacing distribution of eigenvalues from Gaussian to Poisson distribution, which are two universal extreme distributions [27]. The RMT-based approach is a reliable and robust tool for network construction and has been successfully applied to construct various networks, including gene regulatory networks [43–47], functional molecular ecological networks [27], and phylogenetic molecular ecological networks [25]. To construct highly confident co-occurrence ecological networks, only OTUs presented in at least ten samples were used for calculating Pearson correlation. A Pearson correlation coefficient cutoff of 0.81 was determined for network construction. Ecological networks were then visualized by Cytoscape [48].

Results

Effects of eCO_2 on Plant Biomass and Soil N

The plant biomass (aboveground and root) and soil N levels (NO_3^- and NH_4^+) were collected and analyzed. Long-term eCO_2 significantly increased plant biomass and ammonium concentration in soil, but not nitrate concentration (Fig. 1). Consistent with previous observations [21, 28], significantly increased plant biomass was found for both aboveground (P= 0.01) and root (P=0.06) biomass (Fig. 1a). Such increased aboveground plant biomass and root biomass would have imposed higher demand for biologically available N (NO_3^- and NH_4^+) in soil. To analyze the changes of nitrate and ammonium availability in soil, NO_3^- and NH_4^+ concentrations were then measured using a semi-open core, 1-month in situ incubation approach. No significant differences were observed for

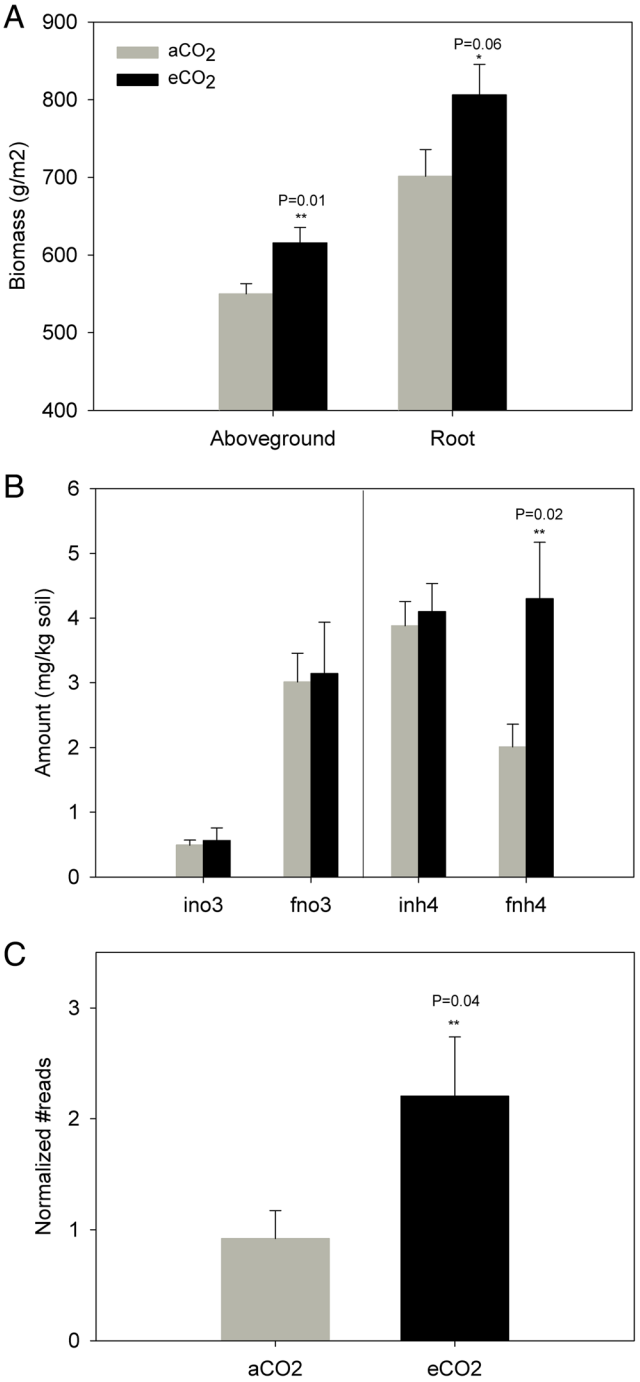

Fig. 1 Effects of long-term eCO_2 on plant biomass (**a**), soil NO_3^- and NH_4^+ (**b**), and *nifH* gene abundance (**c**). Both aboveground and root biomass were averaged from 5 years at the time of sampling, i.e., 2005–2009. Soil NO_3^- and NH_4^+ concentrations were then measured using a semi-open core, 1-month in situ incubation approach. The abundance of *nifH* genes was obtained from a shotgun metagenome datasets, by extracting sequences mapped to *nifH* genes. Statistical testing was performed by Student's *t* test. ino3 and fno3: initial and final NO_3^- concentration; inh4 and fnh4: initial and final NH_4^+ concentration

initial NO_3^- and NH_4^+ amount between aCO_2 and eCO_2 samples. After 1-month in situ incubation, the NO_3^- amount in the soil increased from 0.53 to 3.08 mg/kg soil (Fig. 1b). Similarly, no significant differences were obtained for the final NO_3^- amount between aCO_2 and eCO_2 samples, suggesting similar nitrate availability under aCO_2 and eCO_2. Interestingly, the NH_4^+ amount in aCO_2 samples decreased significantly after incubation, while the amount in eCO_2 samples remained almost unchanged at ~4.3 mg/kg soil, resulting in significantly higher final NH_4^+ amount in eCO_2 samples (Fig. 1b). This suggested that microbial communities under eCO_2 harbored a higher ammonium production rate, which might be caused by increased N_2 fixation and/or organic decomposition.

eCO2 Effects on nifH Gene Abundance

The abundance of *nifH* genes was assessed by extracting *nifH* sequences from a shotgun metagenome dataset targeting all 24 samples, which was annotated by searching against eggNOG database [41]. Comparisons were performed by randomly selecting 350,000 reads per sample from the shotgun metagenome. Seven out of 12 aCO_2 samples were detected with *nifH* sequences. The number for eCO_2 samples was ten. A total of 27 *nifH* sequences were identified. After normalization, an average of 0.92 and 2.2 sequences per sample was found for aCO_2 and eCO_2 samples, respectively. With 80 % identity as cutoff, 48 % of these sequences were assigned to *Bradyrhizobium* species, 26 % to *Rhodopseudomonas* species, 15 % to *Clostridium* species, and 11 % to *Anabaena*, *Roseilexus*, and *Roseobacter* species. Although detected with low sequence numbers, the *nifH* genes were twice more abundant in eCO_2 samples than those in aCO_2 samples (Fig. 1c), with significant Student's *t* test *p*-value of 0.04. Together with our previous evidence from GeoChip results [21, 22], this suggested that eCO_2 had an increased abundance of *nifH* genes in soil.

Amplicon Sequencing Summary

Using 454 pyrosequencing, a total of 102,679 raw reads targeting *nifH* gene amplicons were obtained for 23 samples with an average length of 338 bp. Four samples (two aCO_2 –SOI and SOJ and two eCO_2 –SOM and SON) were excluded from further data analysis for their having<500 reads. After quality trimming, frameshift correction, and chimera removal, 73,161 reads were clustered into 749 *nifH* OTUs at 94 % identity cutoff, of which 624 (a total of 73,036 reads with 42,725 from aCO_2 samples and 30,436 from eCO_2 samples) were non-singleton OTUs. The number of sequences in each sample ranged from 1184 to 7579 (3851 on average), resulting in 80 to 287 OTUs per sample. A random re-sampling effort of 1200 reads per sample was made for further statistical analysis.

No Significant Changes of eCO_2 Effects on Overall nifH-Community Diversity and Structure

To analyze the *nifH*-community diversity in the grassland soil ecosystem and their responses to eCO_2, the OTU richness (Chao1), evenness, taxonomic, and phylogenetic diversity indices were calculated (Fig. 2). A total of 633 and 616 OTUs were identified for aCO_2 and eCO_2 samples with the current sequencing effort, respectively. No significant differences between aCO_2 and eCO_2 samples were observed for the OTU richness as the 95 % confidence intervals were overlapped with any number of randomly sampled sequences (Fig. 2a). Similarly, no significant differences were observed for the evenness of the overall *nifH* community between aCO_2 and eCO_2 sites (Fig. 2b), resulting in insignificant changes of the taxonomic diversity (Fig. 2c). Consistently, the phylogenetic diversity, which considers the phylogenetic relationship among OTUs, did not significantly change in response to eCO_2 (Fig. 2d). All these results suggested that the diversity of *nifH*-community was not significantly affected by long-term eCO_2 in the grassland ecosystem.

Long-term eCO_2 did not significantly alter the overall *nifH*-community structure in the grassland soil ecosystem either (Fig. S1). The overall community structural differences among all samples were assessed by both UniFrac PCoA and non-metric multidimensional scaling (NMDS) analyses using Bray–Curtis distance. A separation trend of eCO_2 samples from aCO_2 samples could be observed by both unweighted UniFrac PCoA (Fig. S1A) and NMDS analysis (Fig. S1B). Further dissimilarity analysis that accounts for OTU abundance also suggested that the overall community structure between aCO_2 and eCO_2 samples was not significantly different (ADONIS: F=0.062, P=0.329; ANOSIM: R=0.035, P=0.249; MRPP: δ=0.531, P=0.252).

The Taxonomic and Phylogenetic Composition of *nifH* Genes

Unlike 16S rRNA genes, reference sequences for *nifH* genes from cultivated microbial strains/species are still very limited, making it difficult to classify *nifH* sequences into their taxonomic groups, especially at the species/strain level. We first tried a strict manner to only assign taxonomic information to OTUs having a minimum of 94 % sequence identity with references in the *nifH* database [2]. As a result, only 49 OTUs could be assigned to known taxonomic groups, among which six were assigned at 100 % identity. Even at 90 % sequence identity cutoff, this number only increased to 119, indicating a large diversified genetic pool of *nifH* gene variants in the soil microbial community. The taxonomic information for *nifH* OTUs was hence assigned as the lowest common ancestor of the best hits at a cutoff of 80 % minimum sequence identity with reference sequences. Genus or higher taxonomic

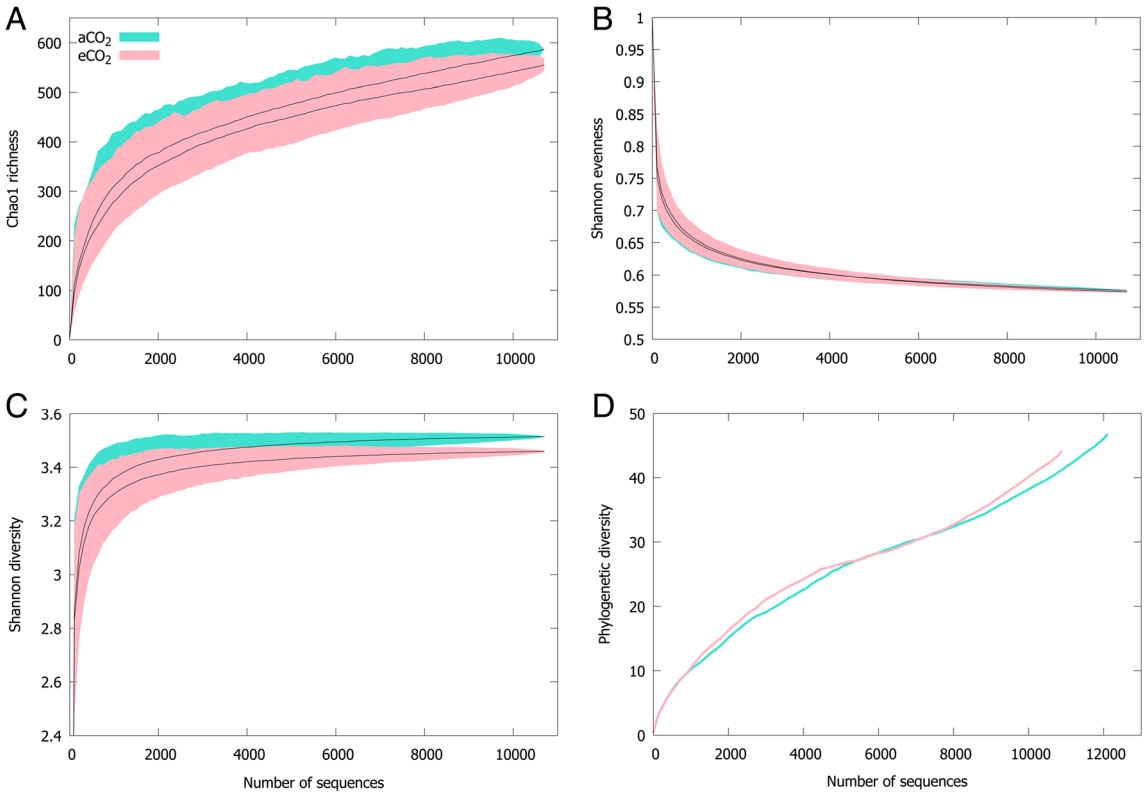

Fig. 2 The diversity of *nifH* genes in the grassland ecosystem under ambient CO_2 and elevated CO_2 conditions: **a** Chao1 richness; **b** Shannon evenness; **c** Shannon diversity; **d** phylogenetic diversity. *Black line* represents the averaged value for each diversity index. *Turquoise and light-pink* regions represent 95 % confidence intervals

information was then assigned to 478 *nifH* OTUs as their nearest taxonomic matches.

The *nifH* community was dominated by Alphaproteobacteria as viewed by both OTU number and relative abundance, followed by Betaproteobacteria, Actinobacteria, Delta-/Gamma-proteobacteria, and Bacilli (Fig. 3a). At the genus level, a total of 134 OTUs were assigned to *Bradyrhizobium* and accounted for 56.1 % of the total sequences in the community. Other abundant genera detected with >1 % relative abundance were *Mesorhizobium* (nine OTUs, 12.7 % relative abundance), *Azospirillum* (20 OTUs, 4.8 % relative abundance), *Azohydromonas* (four OTUs, 3.3 % relative abundance), *Frankia* (two OTUs, 1.9 % relative abundance), *Methylocystis* (12 OTUs, 1.45 % relative abundance), and *Sideroxydans* (18 OTUs, 1.4 % relative abundance). The most dominant OTU (OTU_1) belonged to *Bradyrhizobium* and accounted for 35.3 % of the *nifH*-containing community, followed by OTU_5 (*Mesorhizobium*, 8.34 % relative abundance), OTU_2 (*Bradyrhizobium*, 5.36 % relative abundance), and OTU_7 (*Azohydromonas*, 3.2 % relative abundance). No significant changes of relative abundance were observed for the majority of OTUs, although 34 OTUs did significantly change their abundances at eCO_2, which accounted for 13.15 % of the total captured sequences. Among these, 18 were enriched in eCO_2 samples, and 16 were enriched in aCO_2 samples. All significantly changed OTUs were found with>0.1 % relative abundance in aCO_2 or eCO_2 samples (Fig. S2). Among these, five were found with >1 % relative abundances and were assigned to *Bradyrhizobium* (OTU_450), *Mesorhizobium* (OTU_711), *Azospirillum* (OTU_13), *Sideroxydans* (OTU_30), and *Frankia* (OTU_206) (Fig. S2). Notably, OTU_206 detected in 18 samples, which was also the only *Frankia* OTU found in more than three samples, increased significantly ($p<0.05$) under eCO_2 (3.6 % relative abundance in eCO_2 vs. 0.5 % in aCO_2).

Phylogenetic clade assignment of *nifH* OTUs was performed at a lower cutoff (30 % sequence identity), and the clade information of best BLAST hits aligned by each OTU was selected. The *nifH* community was dominated by sequences belonging to group I and group II NifH clades, which encode Mo-dependent nitrogenase [2, 3, 16], and accounted for 93.1 and 6.76 % of the total captured sequences and 78.54 and 21.13 % total OTUs, respectively. Only two OTUs accounting for 0.14 % of total sequences were found to be group III Mo-independent nitrogenase (Fig. 3b). No significant changes of relative abundances for any *nifH* groups were observed between aCO_2 and eCO_2.

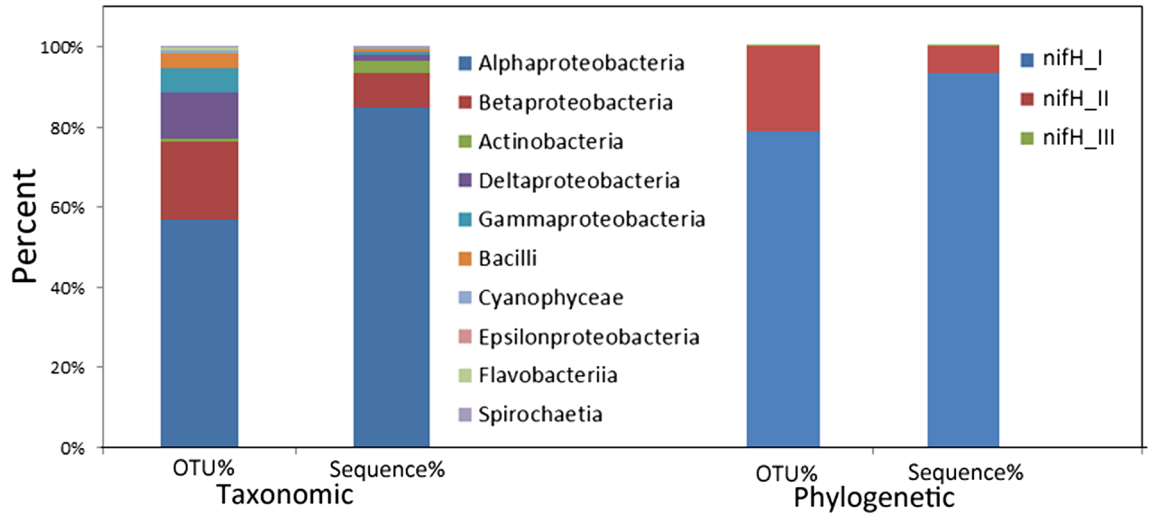

Fig. 3 Taxonomic (**a**) and phylogenetic (**b**) composition of *nifH* genes at both OTU and sequence level. Taxonomic groups were summarized at class level

Co-occurrence Ecological Networks of nifH Assemblages

To explore the co-occurrence patterns of *nifH*-containing microorganisms with other microbial groups, co-occurrence ecological network was constructed using 16S rRNA and *nifH* OTU profiles obtained from the experimental site. By using the random matrix theory approach, a Pearson correlation coefficient cutoff of 0.81 was determined for network construction. The first neighbor nodes connected with *nifH* OTUs were extracted for further analysis, with the purpose to identify *nifH* OTU-mediated co-occurrence patterns. As a result, 50 of the 84 *nifH* OTUs that present in≥10 samples were detected in the co-occurrence network. A total of 24 *nifH*-containing modules were identified. Six modules were found with more than five nodes, covering 154 nodes and 242 links in total, of which 29 nodes were *nifH* OTUs (Fig. S3).

Azospirillum Module

The most complex module (module I, Fig. S3) was centered by two OTUs (*nifH*_643 and *nifH*_31) belonging to *Azospirillum*. Extraction of their first neighbor nodes showed that these two *Azospirillum* OTUs were mainly connected by 17 Actinobacteria OTUs and 12 Proteobacteria OTUs, the latter of which included four *nifH* OTUs (Fig. 4a). Two Acidobacteria OTUs, two Chloroflexi OTUs, one Bacilli, one Sphingobacteria, and two unclassified *nifH* OTUs were also linked with the above *Azospirillum* OTUs. Among the 17 Actinobacteria OTUs connected with *Azospirillum* OTUs, nine were derived from Solirubrobacterales, six from Actinomycetales, and two from Acidimicrobiales. The connected Proteobacteria 16S rRNA OTUs were mainly dominated by four Polyangiaceae OTUs, followed by two Rhizobiales, one Syntrophobacteraceae, and one Oxalobacteraceae OTUs. These results suggested a high co-occurrence frequency of *Azospirillum* species with Actinobacteria species, especially those derived from Solirubrobacterales (Solirubrobacter and Conexibacter) and Actinomycetales.

Mesorhizobium Module

Module II was centered by a *nifH* OTU (*nifH*_325) belonging to *Mesorhizobium* (Fig. 4b). A total of seven Actinobacteria, two Acidobacteria, two Spartobacteria, and one Alphaproteobacteria OTUs were connected with the *Mesorhizobium* OTU. Notably, five of the Actinobacteria OTUs were assigned to Conexibacter, suggesting a high probability of co-occurrence relationship between *Mesorhizobium* and Conexibacter.

Bradyrhizobium Modules

Two modules were centered by *Bradyrhizobium* (modules III, Fig. S3). In both modules, *Bradyrhizobium nifH* OTUs were connected by a high number of Acidobacteria species belonging to multiple subgroups, such as *Gp3*, *Gp4*, *Gp7*, and *Gp17* (Fig. 4c). Specifically, the *nifH*_372 OTU was connected with three *Acidobacteria Gp4* species and one *Granulicella* species. Also, the *nifH*_450 OTU was connected with one *Acidobacteria Gp3*, two Gp4, one Gp7, and one Gp17 species. In addition, the *nifH*_450 OTU was connected with four Alphaproteobacteria 16S rRNA OTUs. These results suggested a high probability of co-occurrence patterns between *Bradyrhizobium* and Acidobacteria species.

Modules IV and V—Module IV was a relatively simple module centered by two *Burkholderiales nifH* OTUs, which were connected with two Acidobacteria, three Proteobacteria,

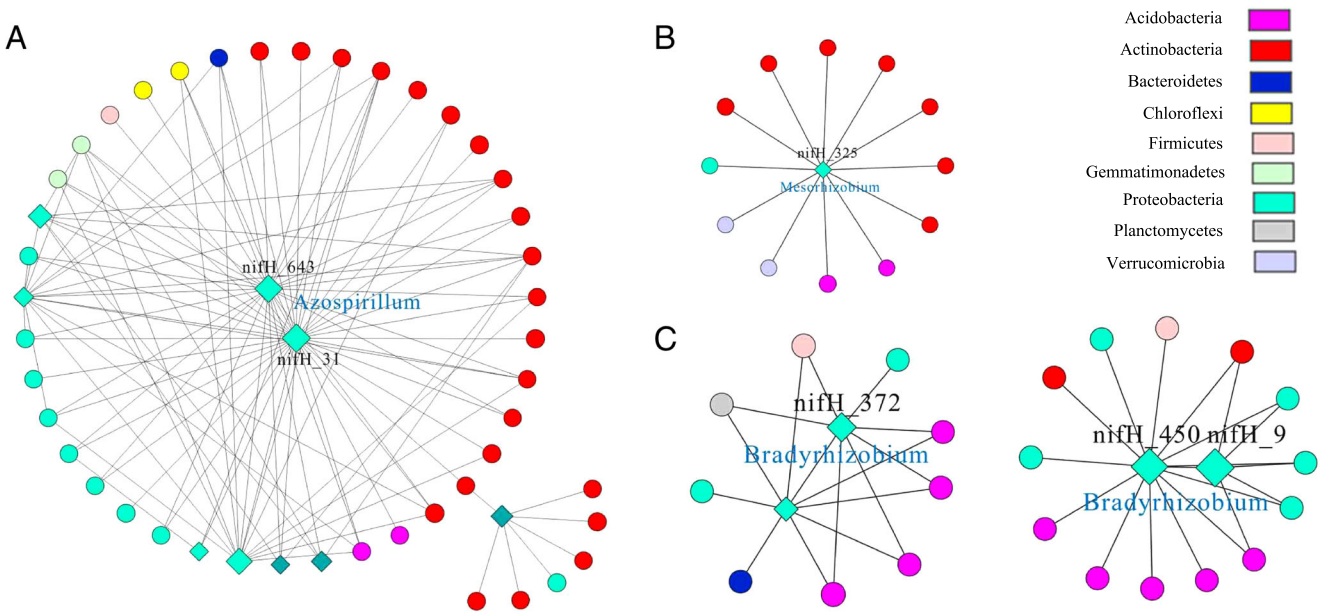

Fig. 4 Co-occurrence modules centered by *nifH* OTUs: **a** *Azospirillum* module; **b** *Mesorhizobium* module; **c** *Bradyrhizobium* module. *nifH* OTUs were represented by *diamond* shape. 16S OTUs were represented by *circular* shape. Different colors refer to different phyla

and one Actinobacteria species (Fig. S3). In contrast, module V was a relatively complex module centered by one *Bradyrhizobium*, one *Desulfovibrio*, and one unclassified *nifH* OTU (Fig. S3). The *Bradyrhizobium* OTU was connected with multiple 16S rRNA OTUs, without preferred co-occurrence patterns. However, the *Desulfovibrio* and unclassified *nifH* OTUs were connected with a high portion of Actinobacteria and Acidobacteria species.

Discussion

Understanding the diversity, composition, and structure of N_2-fixing communities and their interactions with other groups is essential for reliably assessing and predicting N dynamics in ecosystems. In this study, we used next-generation sequencing and co-occurrence ecological network approaches to analyze N_2-fixing communities from grassland soils subjected to 12-year eCO_2 exposure. Our results showed that long-term eCO_2 increased the abundance of *nifH* genes but did not change the overall *nifH* diversity and structure of N_2-fixing communities. Co-occurrence ecological networks were observed with other microbial groups as well as within *nifH*-containing microorganisms. The study provides novel insights into our understanding of the microbial ecology of N_2-fixing communities in grassland ecosystems.

The first question is how soil *nifH* assemblages respond to long-term eCO_2 in this grassland ecosystem. As expected, long-term eCO_2 stimulated the plant growth rate, resulting in increased aboveground and belowground plant biomass, which is consistent with our previous observations [21, 28, 49], as well as many other similar studies [50–52]. According to the increased plant biomass, it was expected that at least 15 % more biologically available N was required under eCO_2, which should be originated from N fixation because of identical initial soil conditions in the field and unchanged soil carbon and N in later years [21]. Meta-analysis indicated that such increased plant growth rate as a result of eCO_2 imposed a higher demand for biologically available N in soil in the form of increased NH_4^+. Since NH_4^+ in natural ecosystems mainly originates from microbial fixation of atmospheric N_2 and decomposition of soil biomass, it is expected that key functional genes involved in microbial N_2 fixation would be changed. The eCO_2 effect on *nifH* gene abundance was analyzed by shotgun metagenome sequencing approach in this study and GeoChip technology in our previous studies [21, 22]. A total of 27 and 147 *nifH* genes were detected in this and previous studies, respectively. In both studies, increased *nifH* gene abundance as a result of eCO_2 was observed. Such increased *nifH* abundance should be most obvious in the first few years of eCO_2 treatment but still be expected in later years to sustain stimulated plant growth under eCO_2, in which N limitation was observed in the ecosystem [19]. Strikingly, the overall *nifH* community structure and diversity did not change significantly in response to eCO_2. Although it is possible that the polF/R primer set may miss some taxa in PCR amplification, this may also suggest that long-term eCO_2 increased the overall *nifH* gene abundance, but not necessarily changed the *nifH* diversity or community structure. Although such results contradicted our hypothesis that the imposed demand for more N by increased plant growth in response to eCO_2 would change the *nifH* diversity and community structure in the soil

ecosystem, the observation was consistent with several recent studies [7, 53]. For example, Berthrong et al. found relatively small effects of CO_2 treatment on N-fixing bacterial community in the four long-term eCO_2 experimental sites they investigated, and no consistent differences were observed for *nifH* diversity between aCO_2 and eCO_2 soils [7]. Notably, although only two OTUs were found for the genus *Frankia*, they were about eightfold more abundant than in aCO_2, which is also consistent with a previous study that the activity of *Frankia* species increased with infertile soil in response to eCO_2 [54]. Taking all our current and previous observations together, the long-term treatment of eCO_2 in this grassland ecosystem has increased the overall abundance of *nifH* gene family, but not necessarily changed the diazotrophic community diversity and structure.

Another objective of this study is to determine the diversity of *nifH* community in the grassland ecosystem. Notably, although the PolF/PolR primer sets we used in this study were of relatively low coverage (~25 %) among all available *nifH* primers [55], they were still selected for their higher specificity with much fewer bands after PCR amplification, even though a high diversity of *nifH* community was observed, with OTUs from six major phyla (ten classes), among which Proteobacteria (Alpha- and Beta-) are the most dominant groups, which is generally consistent with several previous studies in soil [7, 8]. Although found with high diversity, 271 OTUs were still not classified to any taxonomic groups at 80 % identity cutoff, suggesting a highly diverse genetic pool for *nifH* genes. Contrasting with the studies of Berthrong et al. and Collavino et al. with *nifH* communities in forest [7] and pampas [8] soil ecosystems, low amounts of Deltaproteobacteria and cluster III/IV *nifH* genes were found in this study. Since the same PolF/PolR primer [32] was used for PCR amplification of the *nifH* genes, such differences might be due to the different plant composition in these ecosystems, which may favor distinct subsets of diazotroph communities [56]. Notably, *Bradyrhizobium* species, widely known as N_2-fixing bacteria forming symbiotic relationships with legume species through nodules [57], are obviously the most abundant and dominant *nifH*-containing microorganisms in this soil ecosystem at both sequence and OTU levels, suggesting that they play major roles in N fixation in the grassland soil. More interestingly, the most abundant OTU, OTU_1, is 100 % identical with *nifH* sequences from three *Bradyrhizobium* strains isolated from plant root nodules of *Centrosema virginianu*, *Centrosema virginianum*, and *Lupinus perennis* collected in New York and North Carolina [58], of which *L. perennis* is also planted in this BioCON experimental site [28]. This suggested potential existence of some *Bradyrhizobium* species associated with specific plants in different locations, although different geographical distance and soil properties these locations may inhabit.

The third interesting question we would like to address is how *nifH*-containing microorganisms interact with other microorganisms, i.e., who are the members included in the N fixation niche? Since the majority of soil microorganisms are still uncultivable [59], such co-occurrence relationships between microbial species can hardly be directly observed by currently available experimental procedure. In this study, we took advantage of amplicon sequencing of 16S rRNA genes and *nifH* genes, as well as random matrix theory-based co-occurrence ecological network approach, to predict potential neighbors that co-occur with these *nifH*-containing microbial species. Such a rational design provided us opportunities to identify potential microbial interactions not only among *nifH*-containing microorganisms but also between other microbial species and *nifH* assemblages. Although such approaches to identify *nifH* co-occurrence networks have not been previously performed, similar methods have been used to identify bacteria–diatom relationships [60] and bacteria–archaea–protist relationships [24]. As a result, several interesting messages have been brought to our attention. First, only a few OTUs belonging to *Bradyrhizobium* were included in the constructed co-occurrence ecological network, although *Bradyrhizobium* OTUs were most abundant at both OTU and sequence level. This is because most *Bradyrhizobium* species enter plant roots and form symbiotic relationships with plants in the form of root nodules [57]; these species are physically isolated from other free-living soil microorganisms, resulting in few co-occurrence patterns with other microbial species. Second, the taxonomic information of co-occurring 16S rRNA OTUs with *nifH* OTUs is distinctly different from these diazotrophs, which is contradictory to a previous observation based on a global co-existing network using whole-genome sequencing data that co-existing microorganisms are phylogenetically closely related and co-existing genomes tend to be more similar regarding pathway content and genome size [61]. This is possibly because the N_2-fixing ability of diazotrophs to produce NH_4^+ attracted more microbial species without such abilities, rather than other diazotrophs, which is also indirectly evidenced by a previous study that many bacteria prefer ammonia as N source [62].

Several co-occurrence modules centered by *nifH* OTUs were identified. Among them, the module centered by *Azospirillum* OTUs was the most complex one. Since *Azospirillum* species are usually isolated from the rhizosphere of various plant species and are free-living N-fixing bacteria closely associated with grasses [63–66], the complex module formed by *Azospirillum* OTUs confirmed our hypothesis that free-living diazotrophs tend to form more complex networks than symbiotic ones. Co-occurrence patterns were identified between diazotrophs and other microbial species, such as *Azospirillum*/Actinobacteria, *Azospirillum*/Proteobacteria, *Mesorhizobium*/Conexibacter, and *Bradyrhizobium*/Acidobacteria, indicating different co-occurrence patterns for

different diazotrophs. Such co-occurrence patterns indicated a potential attraction of Actinobacteria and Acidobacteria by ammonium produced by potentially free-living diazotrophs. However, the exact underlying mechanism can hardly be identified with current approaches and knowledge.

In conclusion, this study comprehensively analyzed the diversity, structure, and co-occurrence patterns of N_2-fixing microbial communities in a CO_2-enriched grassland ecosystem. Our results provided several valuable insights into the microbial ecology of N_2-fixing microorganisms and their responses to long-term eCO_2. First, this study was conducted in a grassland ecosystem subjected to >12 years of eCO_2 treatment using multiple complementary approaches. Together with our previous GeoChip-based surveys [21, 22], we provided reliable evidence that long-term eCO_2 affects microbial communities by increasing the abundance of *nifH*-containing microorganisms. Second, the diversity and community structure identified in this contributed to our better understanding of the soil diazotrophs. In addition, co-occurrence network analysis provided informative clues about how N_2-fixing microorganisms may interact with other species in the environment. Such information may also help to identify free-living N_2-fixers from symbiotic ones in a predictive manner. However, more experimental approaches are needed for accurate and reliable identification of species–species interactions.

Acknowledgments We thank James W. Voordeckers for editing this paper. This work is supported by the U.S. Department of Agriculture (project 2007-35319-18305) through the NSF-USDA Microbial Observatories Program, by the Department of Energy under contract DE-SC0004601 through Genomics: GTL Foundational Science, Office of Biological and Environmental Research, and by the National Science Foundation under grants DEB-0716587 and DEB-0620652 as well as grants DEB-0322057, DEB-0080382 (the Cedar Creek Long Term Ecological Research project), DEB-0218039, DEB-0219104, DEB-0217631, and DEB-0716587 (BioComplexity, LTER and LTREB projects), the DOE Program for Ecosystem Research, and the Minnesota Environment and Natural Resources Trust Fund.

Conflict of Interest None

References

1. Galloway JN, Dentener FJ, Capone DG, Boyer EW, Howarth RW, Seitzinger SP, Asner GP, Cleveland CC, Green PA, Holland EA, Karl DM, Michaels AF, Porter JH, Townsend AR, Vöosmarty CJ (2004) Nitrogen cycles: past, present, and future. Biogeochemistry 70:153–226. doi:10.1007/s10533-004-0370-0
2. Zehr JP, Jenkins BD, Short SM, Steward GF (2003) Nitrogenase gene diversity and microbial community structure: a cross-system comparison. Environ Microbiol 5:539–554
3. Raymond J, Siefert JL, Staples CR, Blankenship RE (2004) The natural history of nitrogen fixation. Mol Biol Evol 21:541–554. doi:10.1093/molbev/msh047
4. Gaby JC, Buckley DH (2014) A comprehensive aligned *nifH* gene database: a multipurpose tool for studies of nitrogen-fixing bacteria. Database 2014: bau001.
5. Moisander PH, Shiue L, Steward GF, Jenkins BD, Bebout BM, Zehr JP (2006) Application of a *nifH* oligonucleotide microarray for profiling diversity of N_2-fixing microorganisms in marine microbial mats. Environ Microbiol 8:1721–1735
6. Zehr JP (2011) Nitrogen fixation by marine cyanobacteria. Trends Microbiol 19:162–173
7. Berthrong S, Yeager CM, Gallegos-Graves L, Steven B, Eichorst SA, Jackson RB, Kuske CR (2014) Nitrogen fertilization has a stronger effect on soil nitrogen-fixing bacterial communities than elevated atmospheric CO_2. Appl Environ Microbiol 80(10):3103–3112
8. Collavino MM, Tripp HJ, Frank IE, Vidoz ML, Calderoli PA, Donato M, Zehr JP, Aguilar OM (2014) *nifH* pyrosequencing reveals the potential for location-specific soil chemistry to influence N_2-fixing community dynamics. Environ Microbiol 16(10):3211–3223
9. Mohamed NM, Colman AS, Tal Y, Hill RT (2008) Diversity and expression of nitrogen fixation genes in bacterial symbionts of marine sponges. Environ Microbiol 10:2910–2921. doi:10.1111/j.1462-2920.2008.01704.x
10. Großkopf T, Mohr W, Baustian T, Schunck H, Gill D, Kuypers MM, Lavik G, Schmitz RA, Wallace DW, LaRoche J (2012) Doubling of marine dinitrogen-fixation rates based on direct measurements. Nature 488:361–364
11. Hsu S-F, Buckley DH (2009) Evidence for the functional significance of diazotroph community structure in soil. ISME J 3:124–136
12. Wang Q, Quensen JF, Fish JA, Kwon Lee T, Sun Y, Tiedje JM, Cole JR (2013) Ecological patterns of *nifH* Genes in four terrestrial climatic zones explored with targeted metagenomics using FrameBot, a new informatics tool. mBio 4. doi:10.1128/mBio.00592-13
13. Izquierdo J, Nüsslein K (2006) Distribution of extensive *nifH* gene diversity across physical soil microenvironments. Microb Ecol 51:441–452
14. Groszkopf T, Mohr W, Baustian T, Schunck H, Gill D, Kuypers MMM, Lavik G, Schmitz RA, Wallace DWR, LaRoche J (2012) Doubling of marine dinitrogen-fixation rates based on direct measurements. Nature 488:361–364
15. Cleveland CC, Townsend AR, Schimel DS, Fisher H, Howarth RW, Hedin LO, Perakis SS, Latty EF, Von Fischer JC, Elseroad A, Wasson MF (1999) Global patterns of terrestrial biological nitrogen (N_2) fixation in natural ecosystems. Glob Biogeochem Cycles 13:623–645
16. Gaby JC, Buckley DH (2011) A global census of nitrogenase diversity. Environ Microbiol 13:1790–1799
17. Hu S, Chapin FS, Firestone MK, Field CB, Chiariello NR (2001) Nitrogen limitation of microbial decomposition in a grassland under elevated CO_2. Nature 409:188–191
18. Luo Y, Su BO, Currie WS, Dukes JS, Finzi A, Hartwig U, Hungate B, Murtrie REM, Oren RAM, Parton WJ, Pataki DE, Shaw MR, Zak DR, Field CB (2004) Progressive nitrogen limitation of ecosystem responses to rising atmospheric carbon dioxide. Bioscience 54:731–739
19. Reich PB, Hobbie SE, Lee T, Ellsworth DS, West JB, Tilman D, Knops JM, Naeem S, Trost J (2006) Nitrogen limitation constrains sustainability of ecosystem response to CO_2. Nature 440:922–925
20. Finzi AC, Moore DJ, DeLucia EH, Lichter J, Hofmockel KS, Jackson RB, Kim HS, Matamala R, McCarthy HR, Oren R, Pippen JS, Schlesinger WH (2006) Progressive nitrogen limitation of ecosystem processes under elevated CO_2 in a warm-temperate forest. Ecology 87:15–25
21. He Z, Xu M, Deng Y, Kang S, Kellogg L, Wu L, Van Nostrand JD, Hobbie SE, Reich PB, Zhou J (2010) Metagenomic analysis reveals a marked divergence in the structure of belowground microbial communities at elevated CO_2. Ecol Lett 13:564–575

22. Xu M, He Z, Deng Y, Wu L, Van Nostrand JD, Hobbie SE, Reich PB, Zhou J (2013) Elevated CO_2 influences microbial carbon and nitrogen cycling. BMC Microbiol 13:124
23. Barberan A, Bates ST, Casamayor EO, Fierer N (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J 6:343–351
24. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, Sachdeva R, Jones AC, Schwalbach MS, Rose JM, Hewson I, Patel A, Sun F, Caron DA, Fuhrman JA (2011) Marine bacterial, archaeal and protistan association networks reveal ecological linkages. ISME J 5:1414–1425
25. Zhou J, Deng Y, Luo F, He Z, Yang Y (2011) Phylogenetic Molecular Ecological Network of Soil Microbial Communities in Response to Elevated CO2. mBio 2. doi:10.1128/mBio.00122-11
26. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, Huttenhower C (2012) Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol 8:e1002606
27. Zhou J, Deng Y, Luo F, He Z, Tu Q, Zhi X (2010) Functional molecular ecological networks. mBio 1. doi:10.1128/mBio.00169-10
28. Reich PB, Knops J, Tilman D, Craine J, Ellsworth D, Tjoelker M, Lee T, Wedin D, Naeem S, Bahauddin D (2001) Plant diversity enhances ecosystem responses to elevated CO_2 and nitrogen deposition. Nature 410:809–810
29. Lewin KF, Hendrey GR, Nagy J, LaMorte RL (1994) Design and application of a free-air carbon dioxide enrichment facility. Agric For Meteorol 70:15–29
30. Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62:316–322
31. Ahn SJ, Costa J, Rettig Emanuel J (1996) PicoGreen quantitation of DNA: effective evaluation of samples pre-or psost-PCR. Nucleic Acids Res 24:2623–2625
32. Poly F, Monrozier LJ, Bally R (2001) Improvement in the RFLP procedure for studying the diversity of *nifH* genes in communities of nitrogen fixers in soil. Res Microbiol 152:95–103
33. Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods
34. Konstantinidis KT, Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A 102:2567–2572
35. Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data. Genome Res 17:377–386
36. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797
37. Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650
38. Hedges LV, Gurevitch J, Curtis PS (1999) The meta-analysis of response ratios in experimental ecology. Ecology 80:1150–1156
39. Hamady M, Lozupone C, Knight R (2009) Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. ISME J 4:17–27
40. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541
41. Muller J, Szklarczyk D, Julien P, Letunic I, Roth A, Kuhn M, Powell S, von Mering C, Doerks T, Jensen LJ, Bork P (2010) eggNOG v2.0: extending the evolutionary genealogy of genes with enhanced non-supervised orthologous groups, species and functional annotations. Nucleic Acids Res 38:9
42. Deng Y, Jiang Y-H, Yang Y, He Z, Luo F, Zhou J (2012) Molecular ecological network analyses. BMC Bioinforma 13:113
43. Lin L, Song H, Tu Q, Qin Y, Zhou A, Liu W, He Z, Zhou J, Xu J (2011) The thermoanaerobacter glycobiome reveals mechanisms of pentose and hexose co-utilization in bacteria. PLoS Genet 7: e1002318
44. Lin L, Ji Y, Tu Q, Huang R, Teng L, Zeng X, Song H, Wang K, Zhou Q, Li Y (2013) Microevolution from shock to adaptation revealed strategies improving ethanol tolerance and production in *Thermoanaerobacter*. Biotechnol Biofuels 6:103
45. Luo F, Yang Y, Zhong J, Gao H, Khan L, Thompson DK, Zhou J (2007) Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. BMC Bioinforma 8:299
46. Zhou A, He Z, Redding-Johanson AM, Mukhopadhyay A, Hemme CL, Joachimiak MP, Luo F, Deng Y, Bender KS, He Q (2010) Hydrogen peroxide-induced oxidative stress responses in *Desulfovibrio vulgaris* Hildenborough. Environ Microbiol 12: 2645–2657
47. Yang Y, Harris DP, Luo F, Wu L, Parsons AB, Palumbo AV, Zhou J (2008) Characterization of the *Shewanella oneidensis* Fur gene: roles in iron and acid tolerance response. BMC Genomics 9:S11
48. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27:431–432
49. Reich PB, Hobbie SE (2013) Decade-long soil nitrogen constraint on the CO_2 fertilization of plant biomass. Nat Clim Chang 3:278–282
50. Langley JA, Megonigal JP (2010) Ecosystem response to elevated CO_2 levels limited by nitrogen-induced plant species shift. Nature 466:96–99
51. Zak DR, Pregitzer KS, Kubiske ME, Burton AJ (2011) Forest productivity under elevated CO_2 and O_3: positive feedbacks to soil N cycling sustain decade-long net primary productivity enhancement by CO_2. Ecol Lett 14:1220–1226
52. Drake JE, Gallet-Budynek A, Hofmockel KS, Bernhardt ES, Billings SA, Jackson RB, Johnsen KS, Lichter J, McCarthy HR, McCormack ML (2011) Increases in the flux of carbon belowground stimulate nitrogen uptake and sustain the long-term enhancement of forest productivity under elevated CO_2. Ecol Lett 14:349–357
53. Law CS, Breitbarth E, Hoffmann LJ, McGraw CM, Langlois RJ, LaRoche J, Marriner A, Safi KA (2012) No stimulation of nitrogen fixation by non-filamentous diazotrophs under elevated CO_2 in the South Pacific. Glob Chang Biol 18:3004–3014
54. Koike T, Izuta T, Lei T, Kitao M, Asanuma SI (1997) Effects of high CO_2 on nodule formation in roots of Japanese mountain alder seedlings grown under two nutrient levels. In: Ando T, Fujita K, Mae T, Matsumoto H, Mori S, Sekiya J (eds) Plant nutrition for sustainable food production and environment. Springer, Netherlands, pp 887–888
55. Gaby JC, Buckley DH (2012) A comprehensive evaluation of PCR primers to amplify the nifH gene of nitrogenase. PLoS ONE 7(7): e42149–e42149
56. Mutch LA, Young JP (2004) Diversity and specificity of *Rhizobium leguminosarum* biovar viciae on wild and cultivated legumes. Mol Ecol 13:2435–2444
57. Stacey G (1995) *Bradyrhizobium japonicum* nodulation genetics. FEMS Microbiol Lett 127:1–9
58. Parker MA (2012) Legumes select symbiosis island sequence variants in Bradyrhizobium. Mol Ecol 21:1769–1778
59. Rappe MS, Giovannoni SJ (2003) The uncultured microbial majority. Annu Rev Microbiol 57:369–394
60. Stanish LF, O'Neill SP, Gonzalez A, Legg TM, Knelman J, McKnight DM, Spaulding S, Nemergut DR (2013) Bacteria and diatom co-occurrence patterns in microbial mats from polar desert streams. Environ Microbiol 15:1115–1131

61. Chaffron S, Rehrauer H, Pernthaler J, von Mering C (2010) A global network of coexisting microbes from environmental and whole-genome sequence data. Genome Res 20:947–959
62. Müller T, Walter B, Wirtz A, Burkovski A (2006) Ammonium toxicity in bacteria. Curr Microbiol 52:400–406
63. Steenhoudt O, Vanderleyden J (2000) *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. FEMS Microbiol Rev 24:487–506
64. Tien TM, Gaskins MH, Hubbell DH (1979) Plant growth substances produced by *Azospirillum brasilense* and their effect on the growth of pearl millet (*Pennisetum americanum* L.). Appl Environ Microbiol 37:1016–1024
65. Reinhold B, Hurek T, Fendrik I, Pot B, Gillis M, Kersters K, Thielemans S, De Ley J (1987) *Azospirillum halopraeferens* sp. nov., a nitrogen-fixing organism associated with roots of kallar grass (*Leptochloa fusca* (L.) Kunth). Int J Syst Bacteriol 37:43–51
66. Eckert B, Weber OB, Kirchhof G, Halbritter A, Stoffels M, Hartmann A (2001) *Azospirillum doebereinerae* sp. nov., a nitrogen-fixing bacterium associated with the C4-grass miscanthus. Int J Syst Evol Microbiol 51:17–26