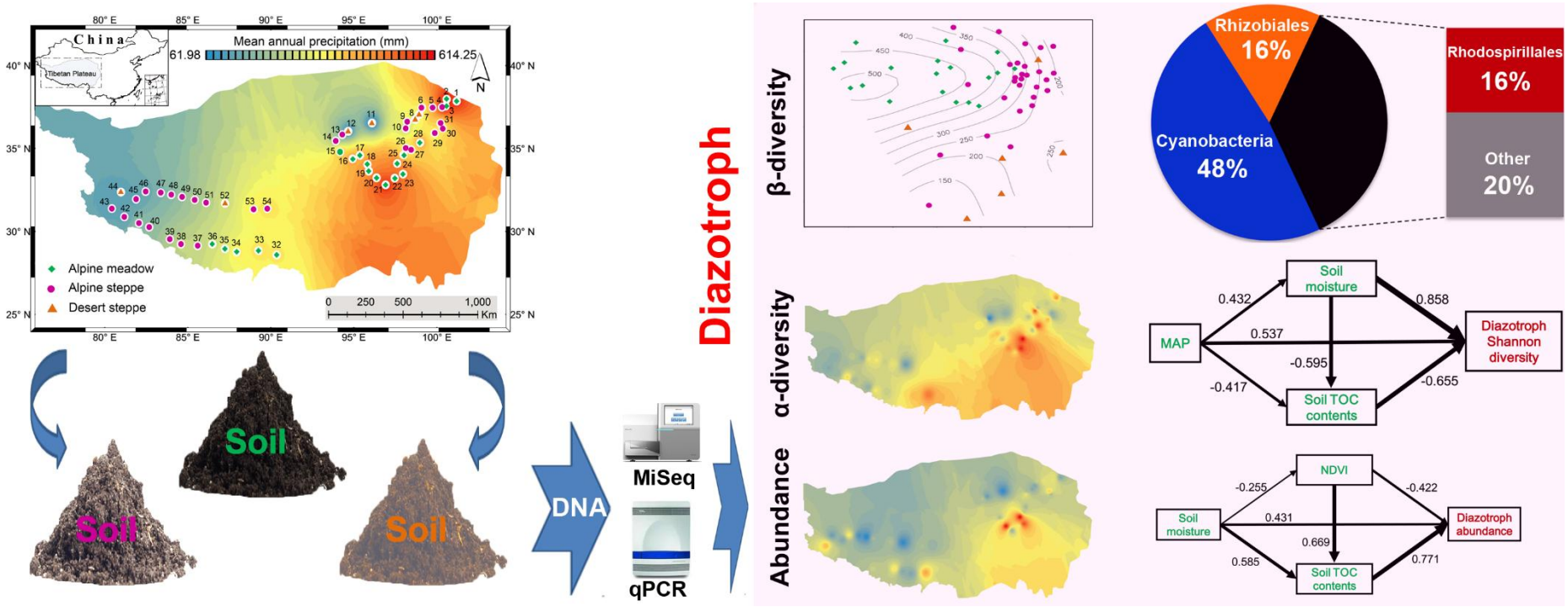
***Graphical Abstract**



***Highlights (for review : 3 to 5 bullet points (maximum 85 characters including spaces per bullet point)**

Highlights

Soil diazotrophs were studied across 54 grasslands on the Tibetan Plateau

Cyanobacteria and Proteobacteria dominated the diazotroph communities

Biological crust and symbiotic nitrogen fixation may serve as vital N sources

Diazotroph abundance and diversity distributed according to mean annual precipitation

Diazotroph distribution may be mainly driven by soil moisture and nutrient contents

Title: Autotrophic and symbiotic diazotrophs dominate nitrogen-fixing communities in Tibetan grassland soils

Authors: Rongxiao Che[a, b], Yongcui Deng[c, d], Fang Wang[a, b], Weijin Wang[b], Zhihong Xu[b], Yanbin Hao[a], Kai Xue[a], Biao Zhang[a], Li Tang[a, b], Huakun Zhou[e], Xiaoyong Cui[a]

a College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

b Environmental Futures Research Institute, School of Environment and Science, Griffith University, Brisbane 4111, Australia

c Nanjing Normal University, Nanjing 210097, China

d Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China

e Key Laboratory of Restoration Ecology of Cold Area in Qinghai Province, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810008, China

Corresponding authors:

Xiaoyong Cui

Telephone: +86 10 88256497

Fax: +86 10 88256497

Email: cuixy@ucas.ac.cn

Abstract

Biological nitrogen fixation, conducted by soil diazotrophs, is the primary nitrogen source for natural grasslands. However, the diazotrophs in grassland soils are still far from fully investigated. Particularly, their regional-scale distribution patterns have never been systematically examined. Here, soils (0–5 cm) were sampled from 54 grasslands on the Tibetan Plateau to examine the diazotroph abundance, diversity, and community composition, as well as their distribution patterns and driving factors. The diazotroph abundance was expressed as *nifH* gene copies, measured using real-time PCR. The diversity and community composition of diazotrophs were analyzed through MiSeq sequencing of *nifH* genes. The results showed that Cyanobacteria (47.94%) and Proteobacteria (45.20%) dominated the soil diazotroph communities. Most Cyanobacteria were classified as Nostocales which are main components of biological crusts. Rhizobiales, most of which were identified as potential symbiotic diazotrophs, were also abundant in approximately half of the soil samples. The soil diazotroph abundance, diversity, and community composition followed the distribution patterns in line with mean annual precipitation. Moreover, they also showed significant correlations with prokaryotic abundance, plant biomass, vegetation cover, soil pH values, and soil nutrient contents. Among these environmental factors, the soil moisture, organic carbon, available phosphorus, and inorganic nitrogen contents could be the main drivers of diazotroph distribution due to their strong correlations with diazotroph indices. These findings suggest that autotrophic and symbiotic diazotrophs are the predominant nitrogen fixers in Tibetan grassland soils, and highlight the key roles of water and nutrient availability in determining the soil diazotroph distribution on the Tibetan Plateau.

Keywords: Tibetan Plateau; Soil microbes; Biogeography; MiSeq sequencing; Biological nitrogen fixation; Grassland

1. Introduction

Nitrogen availability predominantly limits ecosystem primary productivity globally (LeBauer and Treseder, 2008; Vitousek and Howarth, 1991). Biological nitrogen fixation (BNF) contributes 40 ~ 100 Tg nitrogen to the terrestrial ecosystem annually (Vitousek et al., 2013), being recognized as the largest source of reactive nitrogen for our planet (Cleveland et al., 1999; Galloway et al., 2004). BNF is performed by a small fraction of prokaryotes which are collectively known as diazotrophs (Gaby and Buckley, 2011; Raymond et al., 2004). Partly due to the high lateral transfer frequency of nitrogen-fixing genes (Menna and Hungria, 2011; Raymond et al., 2004), diazotrophs are highly diverse in phylogeny (Dos Santos et al., 2012; Gaby and Buckley, 2011; Zehr et al., 2003) and life strategies (e.g., autotrophic *vs* heterotrophic and symbiotic *vs* free-living). The high diversity results in a wide distribution of diazotrophs in both aquatic and terrestrial ecosystems (Gaby and Buckley, 2011; Luo et al., 2012; Munoz-Martin et al., 2014; Nelson et al., 2016). Among all the environmental media, soils have been identified to harbor the most diverse diazotrophs (Gaby and Buckley, 2011). Moreover, the abundance, diversity, and community composition of diazotrophs have been recognized as key biotic factors determining soil nitrogen-fixing capacity (Hsu and Buckley, 2009; Lindsay et al., 2010; Reed et al., 2010; Stewart et al., 2011). Therefore, determining the abundance, diversity, and community composition of soil diazotrophs can substantially contribute to the understanding of terrestrial ecosystem nitrogen cycling.

In recent decades, diazotroph abundance, diversity, and community composition have been extensively investigated employing *nifH* gene as a molecular marker (Che et al., 2017a; Che et al., 2018b; Wang et al., 2017a). A global census based on *nifH* gene revealed that Proteobacteria (α, β, and γ; 56%) and Cyanobacteria (10%) were the most abundant soil nitrogen-fixers (Gaby and Buckley, 2011). However, the soil diazotroph abundance, diversity, and community composition are highly variable across study sites, even within the same type of ecosystem (Che et al., 2017b; Tu et al., 2016; Wang et al., 2017b). Thus, illustrating soil diazotroph distribution pattern at a regional scale requires systematic and multiple-site investigations. Nevertheless, unlike the distribution patterns of marine diazotrophs or rhizobia which have been well documented (Luo et al., 2012; Wang et al., 2016; Zhang et al., 2018a), the distribution of soil diazotroph has been rarely examined across multiple sites, until recently (Shay et al., 2015; Tu et al., 2016; Wang et al., 2017b). This situation is even worse for grassland soils, in which the distribution pattern of diazotroph abundance, diversity, and community composition has never been examined at a regional scale.

Tibetan Plateau is the highest plateau in the world. Its total area is about 2.6×10^6 km^2, on which 60% is covered by natural grasslands (i.e., alpine meadow, alpine steppe, and desert steppe). Similar to other natural ecosystems, the primary productivity of the Tibetan grassland is limited by nitrogen availability (Dong et al., 2016; Mishra and Mainali, 2017; Zhang et al., 2018b). Recently, the Tibetan Plateau has been experiencing a significant increase in atmospheric nitrogen deposition (Liu et al., 2013; Lü and Tian, 2007), which may partially alleviate nitrogen limitation for plants. However, on the Tibetan Plateau, the overall nitrogen deposition rate is still much lower than those in other regions (Galloway et al., 2004; Liu et al., 2013). Moreover, anthropogenic nitrogen inputs on the plateau are almost negligible. Consequently, BNF is probably the main source of available nitrogen in this region. In the past decades, soil microbes in Tibetan soils have been extensively studied (e.g., Che et al., 2015; Deng et al., 2013; Zhang et al., 2016b). In particular, the distribution patterns of soil bacteria, archaea, fungi, and even invertebrates have been well investigated on the Tibetan Plateau (Chu et al., 2016; Shi et al., 2016; Yang et al., 2017; Zhang et al., 2016a; Zhao et al., 2017). Nevertheless, soil diazotrophs have only been examined in a few Tibetan alpine meadows (Che et al., 2017b; Wang et al., 2017b; Zhang et al., 2006), which seriously limits our understanding of the nitrogen inputs in this region. Therefore, systematically determining the distribution pattern of diazotroph in the Tibetan grassland soils can not only improve our knowledge on soil diazotroph biogeography, but also provide essential bases for understanding and managing the Tibetan grassland nitrogen cycling.

In this study, soil samples (0–5 cm) were collected from 54 grasslands on the Tibetan Plateau to investigate the abundance, diversity, and community composition of diazotrophs, and to determine their distribution patterns and main driving environmental factors. We also aimed to examine the differences in diazotroph abundance, diversity, and community compositions among alpine meadows, alpine steppes, and desert steppes.

2. Materials and methods

2.1. Study sites and soil sampling

Soils were sampled from 54 grasslands (19 alpine meadows, 29 alpine steppes, and 6 desert steppes) on the Tibetan Plateau (Fig. 1) in August 2014. As illustrated in Table S1, environmental conditions among the sampling sites differed dramatically. In general, the sampling sites were randomly selected along highways with 60–100 km intervals. The perpendicular distances between the sampling sites and the highways were more than 500 m to attenuate potential disturbances from traffics (Ackermann et al., 2012). At each site, we first surveyed

the plant community compositions in five randomly selected quadrats. The aboveground part of each plant species was collected to determine the aboveground biomass and plant community composition. Subsequently, soil samples (0–5 cm) were collected with a 7 cm auger in each quadrat, thoroughly homogenized, and sieved to ≤ 2 mm. The plant roots in the remaining soils (> 2 mm) were collected to measure the belowground biomass. Then, each soil sample was divided into two subsamples (200 g each). One subsample was air-dried, while another was kept at -20 °C during transportation and stored at -80 °C in the laboratory. We also recorded the elevation and coordinates at each site using GPS. In addition, mean annual temperature and precipitation (MAT and MAP; 1980–2014) were obtained from China Meteorological Data Sharing Service System (http://data.cma.cn/site/index.html). The Normalized Difference Vegetation Index (NDVI) at each site around the sampling time was collected from Moderate Resolution Imaging Spectroradiometer (MODIS).

2.2. Soil and plant characterization

Plant biomass was measured after oven drying at 65 °C for 72 hours, and soil moisture contents were determined by drying at 105 °C for 24 hours. Soil NH_4^+-N and NO_3^--N contents were determined using KCl solution extraction (soil mass to solution ratio of 1:5) followed by colorimetry with an auto analyzer system (SEAL Analytical GmbH, Norderstedt, Germany). Dissolved organic carbon (DOC) contents in soils were analyzed with the same extraction method and a TOC Analyser (Liqui TOC II; Elementar Analysensysteme GmbH, Hanau, Germany). Soil available phosphorus (AP) contents were measured as depicted by Olsen (1954). Soil pH was determined with the soil mass to water ratio of 1:5, using a pH meter (STARTER 3100, Ohaus Instruments Co., Ltd., Shanghai, China). The contents of soil gravimetric moisture, NH_4^+-N, NO_3^--N, and DOC were analyzed with field moist soils, while soil pH and available phosphorus contents were determined using air-dried soils.

2.3. Soil DNA extraction and real-time PCR

Soil DNA was extracted from 0.30 g of fresh soil using a PowerSoil™ DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). The extraction was conducted following the instructions from the manufacturer.

The copies of prokaryotic 16S rRNA and *nifH* gene were quantified using a 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). The quantification was respectively conducted with universal

primer sets for *nifH* (PolF: TGC GAY CCS AAR GCB GAC TC; Poly et al. 2001; PolR: ATS GCC ATC ATY TCR CCG GA; Poly et al. 2001) and 16S rRNA gene (515F: GTG YCA GCM GCC GCG GTA; Caporaso et al. 2011; 909R: CCC CGY CAA TTC MTT TRA GT; Wang and Qian 2009). The 20 μL reaction systems contained: 10 μL of SYBR Green and ROX mixture (2 ×, Takara Bio Inc., Shiga, Japan), 0.5 μL of forward primer (20 μmol L^{-1}), 0.5 μL of reverse primer (20 μmol L^{-1}), 1 μL of template DNA, and 8 μL of nuclease-free water. Standard curves were separately constructed using plasmids harboring the *nifH* or 16S rRNA gene fragments, as detailed in our previous paper (Che et al., 2018b). The PCR runs started with an initial denaturation at 95 °C for 30 s, followed by 40 cycles of 5 s at 95 °C, 30 s at 56 °C, 40 s at 72 °C, and 30 s at 80 °C. The fluorescence signals were collected at 80 °C to attenuate influences of primer dimers. The specificities of PCR products were checked by melting curve analysis.

2.4. MiSeq sequencing and bioinformatic analyses

The PCR amplifications of 16S rRNA and *nifH* genes were conducted with barcoded universal primer sets 515F-909R and PolF-PolR, respectively. The barcode sequences (12 bp) were added to the 5'-end of the forward primers. The PCR reaction mixture (50 μL) contained: 1 μL of template DNA, 25 μL of Premix Taq™ Hot Start Version (Takara Bio Inc., Shiga, Japan), 1 μL of each primer (20 μmol L^{-1}), and 22 μL of water. The PCR runs started with an initial denaturation and enzyme activation step at 95 °C for 10 minutes, followed by 30 (16S rRNA gene) or 40 (*nifH* gene) cycles of 30 s at 95 °C, 30 s at 56 °C, and 45 s at 72 °C, ending with a final extension at 72 °C for 10 minutes. Three technical replicates were conducted for each sample, and the technical replicates were subsequently pooled into one tube.

These PCR products were purified using the GeneJET Gel Extraction Kit (Thermo Scientific, Lithuania), after which all the samples were pooled together with equal molar for each sample. The sequencing samples were prepared using the TruSeq DNA kit, as described by suppliers. The purified products were diluted, denatured, re-diluted, mixed with PhiX (equal to 30% of final DNA amount), following the Illumina library preparation protocols, and were sequenced with the Reagent Kit v3 (2 × 300 bp) in an Illumina MiSeq system at the Chengdu Institute of Biology, Chinese Academy of Sciences.

The *nifH* sequence data were mainly processed following the UPARSE analysis pipeline as recommended by Edgar (2013) with a series of tools including Qiime (v1.9.1; Caporaso et al., 2010), USEARCH (v8.0.1623;

Edgar, 2010), Mothur (v1.27; Schloss et al., 2009), ARB (v5.5; Ludwig et al., 2004), and MEGA (v7.0.14; Kumar et al., 2016b). Specifically, the paired-end reads were first spliced to produce contigs which were then assigned to each sample based on their barcodes using Qiime. Subsequently, all the contigs were merged using mothur, and were filtered using USEARCH. We removed all the contigs with a maximum expected error greater than 1.0 or length shorter than 329 bp, after which the qualified contigs were dereplicated and sorted with the removal of singletons. Then, the *nifH* operational taxonomic unit (OTU) representative sequences were first picked out at 98% sequence similarity. Chimeras were checked using both the inherent de novo chimera detection in USEARCH and the uchime_ref command with more than 20 000 *nifH* sequences (Gaby and Buckley, 2014) as references. We further picked out OTU representatives at 94% level based on the aforementioned OTU representative sequences (Tu et al., 2016), after which the 94%-level OTU representatives were aligned using Mothur against the *nifH* database from Gaby and Buckley (2014). The representative sequences with poor alignment performance were removed, and the remaining OTU representative sequences were imported into ARB. In ARB, the *nifH* gene sequences were translated to peptides to further checkout chimeras and sequence mistakes (e.g., frameshifting). Then, a preliminary OTU table was generated by mapping all *nifH* gene sequences to the high-quality OTU representatives in USEARCH with a 94% threshold. Subsequently, a phylogenetic tree, including all the high-quality *nifH* representative peptide sequences were constructed in MEGA using the neighbor-joining algorithm. After dereplicating the sequences grouping together (with similarities higher than 97%) in the phylogenetic tree, we then manually picked out a new group of *nifH* OTU representatives, and the OTU table was revised accordingly. After the sequence number of each sample being rarefied to 6788, the Shannon diversity index was calculated in R with the vegan package.

The taxonomic assignment for *nifH* OTUs was conducted similarly as described by Tu et al. (2016). In detail, we first downloaded the well-annotated *nifH* peptide sequences (from cultured diazotrophs) with highest similarity against each *nifH* representative peptide sequence using NCBI blastp (https://blast.ncbi.nlm.nih.gov). Then, a neighbor-joining phylogenetic tree was constructed with the OTU representative and downloaded peptide sequences, using MEGA. The taxonomies of the OTU representative peptide sequences were assigned using the lowest common ancestor method. The peptide sequences with similarities lower than 20% against the nearest downloaded sequences were defined as unclassified diazotrophs. The 16S rRNA gene sequences were processed following the UPARSE analysis pipeline as detailed in our previous study (Che et al., 2018a). All the raw sequencing data were deposited in the NCBI Sequence Read Archive under BioProject PRJNA417160.

In total, 1 726 912 *nifH* gene sequences were obtained for the 54 soil samples, and the total sequence number was reduced to 1 345 445 after the quality filtering. The qualified sequence number of each sample ranged from 6788 to 43 686, and thus the sequence number of each sample was rarefied to 6 788 for further analysis. With an identity cut-off of 94%, we obtained 1 020 OTUs which were then manually cluttered into 350 OTUs with the phylogenetic tree constructed in MEGA. As for the 16S rRNA gene sequencing, we obtained 339 714 qualified sequences for 53 soil samples (sample S24 failed in the 16S rRNA gene sequencing), with 9 175 to 71 661 reads per sample. Similar to the *nifH* gene, a rarefaction effort of 9 175 reads per sample was conducted for further analysis. At 97% similarity level, we obtained 6 290 OTUs.

2.5. Statistical analyses

Pearson correlation analysis was used to examine the relationships between abundance or Shannon diversity of diazotrophs and environmental factors. One-way analysis of variations (ANOVA) and Duncan's post hoc test were conducted to reveal the differences in the abundance and Shannon diversity of diazotrophs among different ecosystems. The geographical distribution pattern of diazotroph abundance and Shannon diversity were illustrated through anti-distance-weighted interpolation methods in ArcMap (v10.2.2). We performed non-metric multidimensional scaling (NMDS) and permutational multivariate analysis of variance (PERMANOVA) to reveal the differences in soil diazotroph community compositions among different ecosystems. The diazotroph community composition differences between alpine meadow and steppe were further examined through the "linear discriminant analysis effect size" (LEfSe) method (Segata et al., 2011). The relationships between soil diazotroph community compositions and environmental factors were determined using the "envfit" based on NMDS, Mantel test, and multiple regression tree. The LEfSe analysis was performed using the online Huttenhower Galaxy server (huttenhower.sph.harvard.edu/galaxy) with default pipelines. All the statistical analyses were performed using R with vegan, mvpart, and ggplot2 packages.

3. Results

3.1. Soil *nifH* gene abundances

The soil *nifH* gene copies were in the range of 2.69×10^6 to 5.39×10^9 copies g^{-1} dry soil, and were significantly different among the ecosystem types ($F = 12.41$, $P < 0.001$; Fig. 2a). Specifically, the *nifH* gene copies in the alpine meadow and steppe soils were significantly more than those in the desert steppe soils (Fig. 2a). Among

the alpine meadow soils, the *nifH* gene copies at three wetland-like sites (S20, S23, S24) were significantly more than at the other study sites (Fig. 2a). The geographic distribution of soil diazotroph *nifH* gene copies followed a similar pattern to that of MAP which decreased from the southeast to northwest of the Tibetan Plateau (Figs. 1 and 2a).

As revealed by the Pearson correlations (Table 1), *nifH* gene copies showed significant positive correlations with MAP, plant biomass, NDVI, vegetation cover, soil nutrient contents (e.g., carbon, nitrogen, and phosphorus), and prokaryotic 16S rRNA gene copies (Fig. 3a), whereas they were significantly negatively correlated with soil pH values. Through the pathway analysis based on stepwise regression, we found that the best regression model included soil moisture contents, soil TOC contents, and NDVI, and could explain 66.0% of the variations in *nifH* gene copies (Fig. 4a). In this model, the direct effects of soil moisture and TOC contents on the *nifH* gene copies were positive, while NDVI negatively affected the *nifH* gene copies. The *nifH* gene copies were mostly affected by the direct effect of soil TOC contents, and the indirect effects of soil moisture contents and NDVI through soil TOC on soil *nifH* gene copies were also high.

3.2. Soil diazotroph Shannon α-diversity

The Shannon diversity of soil diazotrophs ranged from 0.196 to 3.79, and was also significantly different among the ecosystem types ($F = 12.41$, $P < 0.001$; Fig. 2b). As revealed by Duncan's test, the diazotroph Shannon diversity in the alpine meadow soils was significantly higher than in the alpine and desert steppes (Fig. 2b). It also distributed in positive correlation with MAP (Table 1).

The soil diazotroph Shannon diversity was significantly correlated with a series of the environmental factors (Table 1). Interestingly, the diazotroph Shannon diversity was significantly correlated with the abundance of soil diazotrophs ($r = 0.439$, $P = 0.001$) and prokaryotes ($r = 0.509$, $P < 0.001$) but not with the prokaryotic Shannon diversity ($r = 0.010$, $P = 0.945$). As suggested by the pathway analysis, the best regression model of soil diazotroph Shannon diversity consisted of soil moisture contents, soil TOC contents, and MAP, which could explain 41.7% of the total variations (Fig. 4b). In this model, soil moisture and TOC contents showed the strongest positive and negative effect on the soil diazotroph Shannon diversity, respectively (Fig. 4b). The MAP also showed a positive direct effect on soil diazotroph diversity (Fig. 4b). However, the indirect effects of the soil moisture contents and MAP on diazotroph diversity through soil TOC contents were negative (Fig. 4b).

3.3. Soil diazotroph community composition

Overall, the Tibetan soil diazotroph communities were dominated by Cyanobacteria (47.94%) and Proteobacteria (45.2%) at the phylum level. In addition, there were some diazotrophs being identified as Verrucomicrobia (1.63%), Nitrospirae (0.64%), Actinobacteria (0.39%), Firmicutes (0.04%), and Bacteroidetes (0.03%). Approximately 4% of the sequences failed to be assigned into a phylum-level taxonomy. As shown in Fig. 5, the nitrogen-fixing Cyanobacteria were mainly identified as Nostocales (45.23%) which could be further classified as Scytonemataceae (29.88%), Nostocaceae (8.60%), Stigonemataceae (4.95%), and Tolypothrichaceae (1.43%). The relative abundance of Oscillatoriales was also high in some soil samples (Fig. 5). Most of the nitrogen-fixing Proteobacteria were classified as α-Proteobacteria (33.06%) which were mainly composed of Rhizobiales (16.02%) and Rhodospirillales (16.09%). The Rhizobiales could be further identified as *Methylobacterium*, *Bradyrhizobium*, and *Rhizobiaceae*, while most of Rhodospirillales were *Skermanella*. There were also some nitrogen-fixing Proteobacteria being classified as β-Proteobacteria (1.40%), γ-Proteobacteria (3.35%), and δ-Proteobacteria (3.86%). The β-Proteobacteria were mainly Rhodocyclales (0.75%) and Burkholderiales (0.51%); the γ-Proteobacteria were mainly classified as Enterobacterales (1.78%) and Pseudomonadales (1.36%); and the δ-Proteobacteria were mainly Desulfuromonadales (3.45%). Additionally, the nitrogen-fixing Verrucomicrobia were mainly classified as Opitutaceae (1.60%).

The community compositions of soil diazotrophs were significantly different among the ecosystem types ($F = 3.23$, $P < 0.001$), and also followed a similar distribution pattern with MAP (Fig. 6a). Compared to the alpine steppe soils, the alpine meadow soils were significantly richer in nitrogen-fixing Proteobacteria, while less enriched in nitrogen-fixing Nostocales (Fig. 6b). In particular, the relative abundance of Desulfuromonadales, Bradyrhizobiaceae, Geobacteraceae, and *Skermanella* were significantly higher than in the alpine steppe soils (Fig. 6b). Additionally, the relative abundance of nitrogen-fixing Verrucomicrobia in some desert steppe soils (i.e., S7, S8, and S12) was much higher than in the other grassland soils (Fig. 5).

Similar to diazotroph abundance and Shannon diversity, the diazotroph community structures were significantly correlated with MAP, plant biomass, NDVI, vegetation cover, soil pH values, as well as soil nutrient and moisture contents (Table 1). In addition, they were significantly correlated with both the plant and prokaryotic community compositions (Fig. 3b), and only weakly correlated with geographical distance ($r = 0.079$, $P =$

0.036). Among these environmental factors, soil moisture, TOC, and available phosphorus contents showed the strongest correlations. As revealed by the multiple regression tree analysis, the plant aboveground biomass, NDVI, as well as the contents of soil available phosphorus, inorganic N, and moisture also played vital roles in determining the soil diazotroph community compositions (Fig. 7).

4. Discussions

This study showed that nitrogen-fixing Cyanobacteria were the most abundant diazotrophs in the Tibetan grassland soils (Fig. 5). Interestingly, these nitrogen-fixing Cyanobacteria were mainly classified as four families under the Nostocales: Nostocaceae, Scytonemataceae, Stigonemataceae, and Tolypothrichaceae, which were widely recognized as the main components of biological crusts (Dojani et al., 2014; Kumar et al., 2016a; Yeager et al., 2007; Yeager et al., 2012). Biological crusts have been shown to own the ability of light-dependent nitrogen fixation (Johnson et al., 2005; Zhou et al., 2016), and account for approximately 30% of the global nitrogen input through BNF (Elbert et al., 2012). Moreover, the ability of Tibetan biological crusts to fix nitrogen has been observed, and approximately 40% of the soil surface in the alpine and subnival areas on the Tibetan Plateau was covered by biological crusts (Janatkova et al., 2013). Therefore, the nitrogen-fixing Nostocales probably play a key role in the nitrogen input of the Tibetan Plateau.

The dominance of Nostocales in the Tibetan diazotroph communities could be attributed to a number of reasons. First, nitrogen-fixing Nostocales can survive in many extreme situations (Dodds et al., 1995; Potts, 1994), and dominate the diazotroph communities in the barren soils under arid and semi-arid climates (Bhatnagar et al., 2008; Dodds et al., 1995; Dojani et al., 2014; Yeager et al., 2012). This may explain why highly abundant nitrogen-fixing Nostocales were identified in the alpine steppe and desert soils on the Tibetan Plateau, and why they were more abundant in alpine steppe soils than in alpine meadow soils (Figs. 5 and 6). In addition, biological crusts were frequently observed in the alpine meadows that we investigated, especially when they were degraded (Che et al., 2017b). Thus, the high proportions of nitrogen-fixing Nostocales detected in some alpine meadows might be owing to the formation of the biological crust in response to grassland degradations (He and Richards, 2015; Zhang et al., 2017). As Cyanobacteria are autotrophic, these findings indicate that the soil nitrogen-fixing communities are dominated by autotrophs at almost half of the Tibetan grasslands. This is further supported by several studies which found that autotrophic microbial communities in some Tibetan soils were dominated by diazotrophs (Guo et al., 2015; Liu et al., 2016a). However, this study only targeted on the

topsoils usually with high light availability, which should be another reason for the dominance of autotrophic diazotrophs. Thus, the diazotroph community compositions in deep soils (e.g., 5–20 cm) should be examined in future studies to obtain a more systemic understanding of Tibetan soil diazotrophs.

Being the second largest diazotroph phylum in the Tibetan soils, Proteobacteria mainly consisted of Rhizobiales and Rhodospirillales (Fig. 5). The Rhizobiales were mainly identified as *Methylobacterium*, *Bradyrhizobium*, and *Rhizobiaceae*, which are well known for forming root nodules with legumes and conducting symbiotic nitrogen fixation (Oldroyd and Downie, 2008; Schmidt et al., 2017; Spaink, 2000). Thus, the high proportions of Rhizobiales indicate a potentially vital role of symbiotic nitrogen fixation in the nitrogen input on the Tibetan Plateau. This is in line with the plant community survey which observed that legumes could occupy approximately 10% of the aboveground biomass. Most Rhodospirillales detected in this study were classified as *Skermanella* which belongs to a separated subline of descent in *Azospirillum* branch (An et al., 2009). Interestingly, none of the six validated species in this genus are able to fix nitrogen, although nitrogen-fixing genes have been detected in their genomes (Subhash and Lee, 2016; Subhash et al., 2017; Zhang et al., 2015; Zhu et al., 2014). In addition, we also detected high proportions of nitrogen-fixing Verrucomicrobia in some desert soils, which is likely attributed to the outstanding survivability of Verrucomicrobia in extreme environments. These nitrogen-fixing Verrucomicrobia showed high similarity with the *Opitutaceae* which were separated from the anorectum of white ant (Kotak et al., 2015; Stevenson et al., 2004). However, due to the paucity of knowledge on the nitrogen-fixing *Opitutaceae*, it is still difficult to judge their contributions to nitrogen fixation.

The soil diazotroph abundance, Shannon diversity, and community composition were all significantly correlated with MAP, and showed the strongest correlations with soil moisture among all the environmental factors included in this study (Table 1). This highlights the vital role of water availability in driving the distribution of diazotrophs in Tibetan soils. First, water is widely recognized as a main limiting factor for soil microbes, and plays a key role in structuring soil microbial communities (Fierer, 2017; Heděnec et al., 2017; Huang et al., 2018; Liu et al., 2016b). The MAP and soil moisture contents differed greatly among the study sites (Fig. 1; Table S1). Hence, it is persuasive to assert that water availability may directly drive the diazotroph distribution on the Tibetan Plateau. Second, water availability is also a determinative factor for plant growth and community composition (Liu et al., 2018; Xu et al., 2018; Zhang et al., 2016c), affecting the availability and quality of

organic carbon and light, and thus indirectly drives the distribution of soil diazotrophs. Collectively, water content might influent the distribution of Tibetan soil diazotrophs in both direct and indirect manners, which is also supported by the pathway analysis (Fig. 4).

The other environmental factors, such as phosphorus and nitrogen were also widely recognized as crucial factors affecting soil diazotrophs (Che et al., 2017a; Reed et al., 2011). In particular, phosphorus is one of the key factors limiting nitrogen fixation due to the high adenosine triphosphate (ATP) demand of BNF (Hill, 1988; Reed et al., 2011; Vitousek et al., 2002), which is a possible explanation for the significant positive correlations between soil phosphorus and diazotroph abundance. Furthermore, atmospheric nitrogen deposition on the Tibetan Plateau has significantly increased over the past decades (Liu et al., 2013; Lü and Tian, 2007), which could shift the grasslands to less nitrogen-limited but more phosphorus-limited status. In this case, the phosphorus availability may also indirectly driving diazotroph distribution through affecting plant growth. Interestingly, although higher nitrogen availability usually negatively affects BNF and soil diazotrophs (Che et al., 2017a; Reed et al., 2011), in this study, soil total and inorganic nitrogen contents showed strong positive correlations with diazotroph abundance (Table 1). On the one hand, this could be partly attributed to the collinearity between soil nitrogen contents and other soil properties (e.g., moisture, organic carbon, and available phosphorus), which could offset and even reverse the correlations between soil nitrogen contents and diazotroph abundance. On the other hand, the positive correlation between soil nitrogen contents and diazotroph abundance may indicate that BNF contributed substantially to the nitrogen inputs on the Tibetan Plateau.

In a small scale study on the Tibetan Plateau, Zhang et al. (2006) found that the distribution of soil diazotroph community compositions was mainly affected by elevations. However, in this study, the elevations only showed weak correlations with the diazotroph community compositions. A recent investigation conducted at nine Tibetan alpine meadows showed that the soil diazotroph community compositions were mainly affected by soil pH and MAT (Wang et al., 2017b). Nevertheless, we found that the MAT showed no significant correlations with soil diazotroph community composition at the regional scale. Moreover, in this study, the correlation between soil pH values and diazotroph Shannon diversity was negative rather than positive as described in the aforementioned investigation at Tibetan alpine meadows (Wang et al., 2017b). These inconsistencies could result from a number of reasons. For instance, the Tibetan alpine meadows are rainy and cold. The diazotrophs in alpine meadow soils were likely limited by temperature rather than water availability. MAT is usually

negatively correlated with elevation, and positively correlated with soil pH when the precipitation is similar (Slessarev et al., 2016). Thus, MAT, pH, and elevation were recognized as main factors driving the distribution of diazotrophs in alpine meadow soils (Wang et al., 2017b; Zhang et al., 2006). However, at the regional scale, the precipitation is highly variable across the study sites (Fig. 1). For soil microbes, water availability may become the predominant limiting factor over temperature, as observed in multitude investigations (Fierer, 2017). Moreover, precipitation is generally negatively correlated with soil pH (Slessarev et al., 2016). Hence, the roles of MAT, pH, and elevation in driving the distribution of soil diazotrophs at the regional scale on the Tibetan Plateau were totally different compared to those in alpine meadows. Collectively, these findings suggest that the driving mechanisms of the diazotroph distribution in the Tibetan grassland soils can differ at different scales.

In addition, the abundance and community composition of soil diazotrophs also showed significant correlations with prokaryotic abundance and community composition (Fig. 3), which indicate that soil diazotrophs may follow a similar distribution pattern with the whole prokaryote community. A possible explanation to this phenomenon is that the distribution of Tibetan soil diazotrophs is mainly driven by the environmental factors affecting microbial basal physiological metabolism rather than specific functional traits. Indeed, most of the environmental factors (e.g., water, carbon, nitrogen, and phosphorus availability) that showed significant correlations with diazotroph abundance and community composition are closely correlated to microbial basal metabolisms. Therefore, assessing the roles of microbial common traits (e.g., respiration, gene transcription, and mRNA translation) and specific functional traits (e.g., nitrogen fixation) in determining the microbial distribution should be of great interest for future studies.

In this study, 350 OTUs were identified across all the sampling sites. This is far less than the previous investigation at Tibetan alpine meadows, which identified 4 162 OTUs across just nine alpine meadows (Wang et al., 2017b). There were a number of potential reasons resulting in this dramatic inconsistency. First, we used a lower identity cut-off of 94%, which should classify less OTUs than at 96%. Moreover, the dereplication of representative sequences according to the phylogenetic tree would further reduce the OTU number. Second, in this study, the OTUs were classified using UPARSE pipeline which is a highly accurate method with strict sequence quality controls, and usually generates less OTUs than other strategies such as UCLAST (Edgar, 2013). Finally, the PCR amplification biases, which are almost inevitable in all PCR based microbial studies, might also contribute to this inconsistency (Diallo et al., 2008; Gaby and Buckley, 2012). In particular, the primer set

employed in this study may not cover all the diazotroph species (Gaby and Buckley, 2012), and thus future studies should combine different primer sets to confirm the findings in this study. Nevertheless, we assumed that 350 OTUs should be reasonable for the soil diazotroph communities in this region. First, diazotrophs usually occupy less than 1% of total prokaryote population. We only identified 6 290 prokaryote OTUs across all the soil samples, and thus 350 OTUs of diazotroph should be acceptable. Moreover, the global census of nitrogenase diversity only generated 3 358 diazotroph OTUs with a similarity threshold of 95% (Gaby and Buckley, 2011), which further supports that there should be less than a few thousands of diazotroph OTUs in the Tibetan grassland soils. Finally, the excellent OTU annotation performance also indirectly verified the sequence analysis strategy used in this study.

In conclusion, this study showed that the nitrogen-fixing communities in the Tibetan grassland soils were dominated by autotrophic and symbiotic diazotrophs, highlighting the potentially primary roles of biological crusts and symbiotic BNF in the nitrogen inputs on the Tibetan Plateau. This study also showed that the distribution patterns of soil diazotroph abundance, diversity, and community composition on the Tibetan Plateau were similar to that of the MAP, and might be mainly driven by the availability of soil water and nutrients. This investigation not only provides the first insight into the spatial distribution of grassland soil diazotrophs at the regional scale, but also contributes substantially to understanding the nitrogen input pattern on the Tibetan Plateau. However, in this study, the diazotrophs were analyzed using DNA-based methods which failed to identify the nitrogen-fixing activity and the diazotrophs actively fixing nitrogen. Therefore, future study should combine acetylene reduction assay, stable isotope labeling, and RNA-based methods to systematically examine the nitrogen-fixing activity and active diazotrophs in soils.

Acknowledgments

This work was supported by the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB15010201), National Natural Science Foundation of China (31570518), National Key Research and Development Program of China (2016YFC0501800), and Qinghai Innovation Platform Construction Project (2017-ZJ-Y20). Rongxiao Che was also supported by the Griffith University Postgraduate Research Scholarships. We sincerely appreciate the sampling and experiment assistance from Prof. Yanfen Wang, Prof. Xiangzhen Li, A/Prof. Lili Jiang, A/Prof. Minjie Yao, Jing Zhang, Wenjun Liu, Zhe Pang, Wenyu Fan, Shutong Zhou, Anquan Xia, Di Wang, Hanke Liu, Chaoting Zhou, Yuan Zhang, and Qingzhou Zhao.

References

Ackermann, K., Zackrisson, O., Rousk, J., Jones, D.L., DeLuca, T.H., 2012. N_2 fixation in feather mosses is a sensitive indicator of N deposition in boreal forests. Ecosystems 15, 986–998.

An, H., Zhang, L., Tang, Y., Luo, X., Sun, T., Li, Y., Wang, Y., Dai, J., Fang, C., 2009. *Skermanella xinjiangensis* sp. nov., isolated from the desert of Xinjiang, China. Int. J. Syst. Evol. Microbiol. 59, 1531–1534.

Bhatnagar, A., Makandar, M.B., Garg, M.K., Bhatnagar, M., 2008. Community structure and diversity of cyanobacteria and green algae in the soils of Thar Desert (India). J. Arid. Environ. 72, 73–83.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Tumbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight, R., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. U.S.A. 108, 4516–4522.

Che, R.X., Deng, Y.C., Wang, F., Wang, W.J., Xu, Z.H., Wang, Y.F., Cui, X.Y., 2015. 16S rRNA-based bacterial community structure is a sensitive indicator of soil respiration activity. J. Soils Sediments 15, 1987–1990.

Che, R.X., Deng, Y.C., Wang, W.J., Rui, Y.C., Zhang, J., Tahmasbian, I., Tang, L., Wang, S.P., Wang, Y.F., Xu, Z.H., Cui, X.Y., 2018a. Long-term warming rather than grazing significantly changed total and active soil procaryotic community structures. Geoderma 316, 1–10.

Che, R.X., Deng, Y.C., Wu, Y.B., Zhang, J., Wang, F., Tang, L., Ma, S., Liu, H.K., Zhao, X., Wang, Y.F., Hao, Y.B., Cui, X.Y., 2017a. Relationships between biological nitrogen fixation and available nitrogen at scales from molecular to community level. Chinese J. Ecol. 36, 224–232.

Che, R.X., Qin, J.L., Tahmasbian, I., Wang, F., Zhou, S.T., Xu, Z.H., Cui, X.Y., 2018b. Litter amendment rather than phosphorus can dramatically change inorganic nitrogen pools in a degraded grassland soil by affecting nitrogen-cycling microbes. Soil Biol. Biochem. 120, 145–152.

Che, R.X., Wang, F., Wang, W.J., Zhang, J., Zhao, X., Rui, Y.C., Xu, Z.H., Wang, Y.F., Hao, Y.B., Cui, X.Y., 2017b. Increase in ammonia-oxidizing microbe abundance during degradation of alpine meadows may lead to greater soil nitrogen loss. Biogeochemistry 136, 341–352.

Chu, H., Sun, H., Tripathi, B.M., Adams, J.M., Huang, R., Zhang, Y., Shi, Y., 2016. Bacterial community dissimilarity between the surface and subsurface soils equals horizontal differences over several kilometers in the western Tibetan Plateau. Environ. Microbiol. 18, 1523–1533.

Cleveland, C.C., Townsend, A.R., Schimel, D.S., Fisher, H., Howarth, R.W., Hedin, L.O., Perakis, S.S., Latty, E.F., Von Fischer, J.C., Elseroad, A., Wasson, M.F., 1999. Global patterns of terrestrial biological nitrogen (N_2) fixation in natural ecosystems. Glob. Biogeochem. Cycles 13, 623–645.

Deng, Y., Cui, X., Lüke, C., Dumont, M.G., 2013. Aerobic methanotroph diversity in Riganqiao peatlands on the Qinghai–Tibetan Plateau. Environ. Microbiol. Rep. 5, 566–574.

Diallo, M.D., Reinhold-Hurek, B., Hurek, T., 2008. Evaluation of PCR primers for universal *nifH* gene targeting and for assessment of transcribed *nifH* pools in roots of *Oryza longistaminata* with and without low nitrogen input. FEMS Microbiol. Ecol. 65, 220–228.

Dodds, W.K., Gudder, D.A., Mollenhauer, D., 1995. The ecology of *nostoc*. J. Phycol. 31, 2–18.

Dojani, S., Kauff, F., Weber, B., Budel, B., 2014. Genotypic and phenotypic diversity of Cyanobacteria in biological soil crusts of the Succulent Karoo and Nama Karoo of Southern Africa. Microb. Ecol. 67, 286–301.

Dong, J.F., Cui, X.Y., Wang, S.P., Wang, F., Pang, Z., Xu, N., Zhao, G.Q., Wang, S.P., 2016. Changes in biomass and quality of alpine steppe in response to N & P fertilization in the Tibetan Plateau. PLoS One 11, e0156146.

Dos Santos, P.C., Fang, Z., Mason, S.W., Setubal, J.C., Dixon, R., 2012. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. BMC genomics 13, 162–162.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460-2461.

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–998.

Elbert, W., Weber, B., Burrows, S., Steinkamp, J., Budel, B., Andreae, M.O., Poschl, U., 2012. Contribution of cryptogamic covers to the global cycles of carbon and nitrogen. Nat. Geosci. 5, 459–462.

Fierer, N., 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. Nat. Rev. Microbiol. 15, 579–590.

Gaby, J.C., Buckley, D.H., 2011. A global census of nitrogenase diversity. Environ. Microbiol. 13, 1790–1799.

Gaby, J.C., Buckley, D.H., 2012. A comprehensive evaluation of PCR primers to amplify the *nifH* gene of nitrogenase. PLoS One 7, e42149.

Gaby, J.C., Buckley, D.H., 2014. A comprehensive aligned *nifH* gene database: a multipurpose tool for studies of nitrogen-fixing bacteria. Database 2014, bau001.

Galloway, J.N., Dentener, F.J., Capone, D.G., Boyer, E.W., Howarth, R.W., Seitzinger, S.P., Asner, G.P., Cleveland, C.C., Green, P.A., Holland, E.A., Karl, D.M., Michaels, A.F., Porter, J.H., Townsend, A.R., Vorosmarty, C.J., 2004. Nitrogen cycles: past, present, and future. Biogeochemistry 70, 153–226.

Guo, G., Kong, W., Liu, J., Zhao, J., Du, H., Zhang, X., Xia, P., 2015. Diversity and distribution of autotrophic microbial community along environmental gradients in grassland soils on the Tibetan Plateau. Appl. Microbiol. Biotechnol. 99, 8765–8776.

He, S.Y., Richards, K., 2015. Impact of meadow degradation on soil water status and pasture management a case study in Tibet. Land Degrad. Dev. 26, 468–479.

Heděnec, P., Rui, J., Lin, Q., Yao, M., Li, J., Li, H., Frouz, J., Li, X., 2017. Functional and phylogenetic response of soil prokaryotic community under an artificial moisture gradient. Applied Soil Ecology.

Hill, S., 1988. How is nitrogenase regulated by oxygen? FEMS Microbiol. Lett. 54, 111–129.

Hsu, S.F., Buckley, D.H., 2009. Evidence for the functional significance of diazotroph community structure in soil. ISME J. 3, 124–136.

Huang, G., Li, L., Su, Y.G., Li, Y., 2018. Differential seasonal effects of water addition and nitrogen fertilization on microbial biomass and diversity in a temperate desert. Catena 161, 27–36.

Janatkova, K., Rehakova, K., Dolezal, J., Simek, M., Chlumska, Z., Dvorsky, M., Kopecky, M., 2013. Community structure of soil phototrophs along environmental gradients in arid Himalaya. Environ. Microbiol. 15, 2505–2516.

Johnson, S.L., Budinoff, C.R., Belnap, J., Garcia-Pichel, F., 2005. Relevance of ammonium oxidation within biological soil crust communities. Environ. Microbiol. 7, 1–12.

Kotak, M., Isanapong, J., Goodwin, L., Bruce, D., Chen, A., Han, C.S., Huntemann, M., Ivanova, N., Land, M.L., Nolan, M., Pati, A., Woyke, T., Rodrigues, J.L.M., 2015. Complete genome sequence of the

Opitutaceae bacterium strain TAV5, a potential facultative methylotroph of the wood-feeding termite *Reticulitermes flavipes*. Genome Announcements 3, e00060-15.

Kumar, D., Kastanek, P., Adhikary, S.P., 2016a. Diversity of cyanobacteria in biological crusts on arid soils in the Eastern region of India and their molecular phylogeny. Curr. Sci. 110, 1999–2004.

Kumar, S., Stecher, G., Tamura, K., 2016b. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

LeBauer, D.S., Treseder, K.K., 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. Ecology 89, 371–379.

Lindsay, E.A., Colloff, M.J., Gibb, N.L., Wakelin, S.A., 2010. The abundance of microbial functional genes in grassy woodlands is influenced more by soil nutrient enrichment than by recent weed invasion or livestock exclusion. Applied Environ. Microbiol. 76, 5547–5555.

Liu, J., Kong, W., Zhang, G., Khan, A., Guo, G., Zhu, C., Wei, X., Kang, S., Morgan-Kiss, R.M., 2016a. Diversity and succession of autotrophic microbial community in high-elevation soils along deglaciation chronosequence. FEMS Microbiol. Ecol. 92, fiw160.

Liu, W., Allison, S.D., Xia, J., Liu, L., Wan, S., 2016b. Precipitation regime drives warming responses of microbial biomass and activity in temperate steppe soils. Biol. Fert. soils 52, 469–477.

Liu, W., Lu, X.T., Xu, W.F., Shi, H.Q., Hou, L.Y., Li, L.H., Yuan, W.P., 2018. Effects of water and nitrogen addition on ecosystem respiration across three types of steppe: The role of plant and microbial biomass. Sci. Total Environ. 619, 103–111.

Liu, X., Zhang, Y., Han, W., Tang, A., Shen, J., Cui, Z., Vitousek, P., Erisman, J.W., Goulding, K., Christie, P., 2013. Enhanced nitrogen deposition over China. Nature 494, 459.

Lü C., Tian H., 2007. Spatial and temporal patterns of nitrogen deposition in China: synthesis of observational data. J. Geophys. Res. Atmos. 112, D22S05.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., Schleifer, K.H., 2004. ARB: a software environment for sequence data. Nucleic Acids Res. 32, 1363–1371.

Luo, Y.W., Doney, S.C., Anderson, L.A., Benavides, M., Berman-Frank, I., Bode, A., Bonnet, S., Bostrom, K.H., Bottjer, D., Capone, D.G., Carpenter, E.J., Chen, Y.L., Church, M.J., Dore, J.E., Falcon, L.I.,

551 Fernandez, A., Foster, R.A., Furuya, K., Gomez, F., Gundersen, K., Hynes, A.M., Karl, D.M.,
552 Kitajima, S., Langlois, R.J., LaRoche, J., Letelier, R.M., Maranon, E., McGillicuddy, D.J., Moisander,
553 P.H., Moore, C.M., Mourino-Carballido, B., Mulholland, M.R., Needoba, J.A., Orcutt, K.M., Poulton,
554 A.J., Rahav, E., Raimbault, P., Rees, A.P., Riemann, L., Shiozaki, T., Subramaniam, A., Tyrrell, T.,
555 Turk-Kubo, K.A., Varela, M., Villareal, T.A., Webb, E.A., White, A.E., Wu, J., Zehr, J.P., 2012.
556 Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates. Earth Syst.
557 Sci. Data 4, 47–73.
558 Menna, P., Hungria, M., 2011. Phylogeny of nodulation and nitrogen-fixation genes in Bradyrhizobium:
559 supporting evidence for the theory of monophyletic origin, and spread and maintenance by both
560 horizontal and vertical transfer. Int. J. Syst. Evol. Microbiol. 61, 3052–3067.
561 Mishra, N.B., Mainali, K.P., 2017. Greening and browning of the Himalaya: Spatial patterns and the role of
562 climatic change and human drivers. Sci. Total Environ. 587, 326–339.
563 Munoz-Martin, M.A., Mateo, P., Leganes, F., Fernandez-Pinas, F., 2014. A battery of bioreporters of nitrogen
564 bioavailability in aquatic ecosystems based on cyanobacteria. Sci. Total Environ. 475, 169–179.
565 Nelson, M.B., Martiny, A.C., Martiny, J.B.H., 2016. Global biogeography of microbial nitrogen-cycling traits in
566 soil. Proc. Natl. Acad. Sci. U.S.A. 113, 8033-8040.
567 Oldroyd, G.E.D., Downie, J.M., 2008. Coordinating nodule morphogenesis with rhizobial infection in legumes.
568 Annu. Rev. Plant. Biol. 59, 519–546.
569 Olsen, S.R., 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. United
570 States Department Of Agriculture, Washington.
571 Poly, F., Monrozier, L.J., Bally, R., 2001. Improvement in the RFLP procedure for studying the diversity of
572 *nifH* genes in communities of nitrogen fixers in soil. Res. Microbiol. 152, 95–103.
573 Potts, M., 1994. Desiccation tolerance of prokaryotes. Microbiol. Rev. 58, 755–805.
574 Raymond, J., Siefert, J.L., Staples, C.R., Blankenship, R.E., 2004. The natural history of nitrogen fixation. Mol.
575 Biol. Evol. 21, 541–554.
576 Reed, S.C., Cleveland, C.C., Townsend, A.R., 2011. Functional Ecology of Free-Living Nitrogen Fixation: A
577 Contemporary Perspective. Annu. Rev. Ecol. Evol. Syst. 42, 489–512
578 Reed, S.C., Townsend, A.R., Cleveland, C.C., Nemergut, D.R., 2010. Microbial community shifts influence
579 patterns in tropical forest nitrogen fixation. Oecologia 164, 521–531.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.

Schmidt, J., Fester, T., Schulz, E., Michalzik, B., Buscot, F., Gutknecht, J., 2017. Effects of plant-symbiotic relationships on the living soil microbial community and microbial necromass in a long-term agro-ecosystem. Sci. Total Environ. 581, 756–765.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., Huttenhower, C., 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12, R60.

Shay, P.E., Winder, R.S., Trofymow, J.A., 2015. Nutrient-cycling microbes in coastal Douglas-fir forests: regional-scale correlation between communities, in situ climate, and other factors. Front. Microbiol. 6, 1097.

Shi, Y., Adams, J.M., Ni, Y., Yang, T., Jing, X., Chen, L., He, J.-S., Chu, H., 2016. The biogeography of soil archaeal communities on the eastern Tibetan Plateau. Scientific Reports 6, 38893.

Slessarev, E.W., Lin, Y., Bingham, N.L., Johnson, J.E., Dai, Y., Schimel, J.P., Chadwick, O.A., 2016. Water balance creates a threshold in soil pH at the global scale. Nature 540, 567–569.

Spaink, H.P., 2000. Root nodulation and infection factors produced by rhizobial bacteria. Annu. Rev. Microbiol. 54, 257–288.

Stevenson, B.S., Eichorst, S.A., Wertz, J.T., Schmidt, T.M., Breznak, J.A., 2004. New strategies for cultivation and detection of previously uncultured microbes. Appl. Environ. Microbiol. 70, 4748–4755.

Stewart, K.J., Coxson, D., Siciliano, S.D., 2011. Small-scale spatial patterns in N_2-fixation and nutrient availability in an arctic hummock-hollow ecosystem. Soil Biol. Biochem. 43, 133–140.

Subhash, Y., Lee, S.-S., 2016. *Skermanella rosea* sp. nov., isolated from hydrocarbon-contaminated desert sands. Int. J. Syst. Evol. Microbiol. 66, 3951–3956.

Subhash, Y., Yoon, D.-E., Lee, S.-S., 2017. Skermanella mucosa sp. nov., isolated from crude oil contaminated soil. Antonie van Leeuwenhoek, 1–8.

Tu, Q.C., Deng, Y., Yan, Q.Y., Shen, L.N., Lin, L., He, Z.L., Wu, L.Y., Van Nostrand, J.D., Buzzard, V., Michaletz, S.T., Enquist, B.J., Weiser, M.D., Kaspari, M., Waide, R.B., Brown, J.H., Zhou, J.Z., 2016. Biogeographic patterns of soil diazotrophic communities across six forests in the North America. Mol. Ecol. 25, 2937–2948.

610 Vitousek, P.M., Cassman, K., Cleveland, C., Crews, T., Field, C.B., Grimm, N.B., Howarth, R.W., Marino, R.,
611 Martinelli, L., Rastetter, E.B., Sprent, J.I., 2002. Towards an ecological understanding of biological
612 nitrogen fixation. Biogeochemistry 57, 1–45.
613 Vitousek, P.M., Howarth, R.W., 1991. Nitrogen limitation on land and in the sea : how can it occur? .
614 Biogeochemistry 13, 87–115.
615 Vitousek, P.M., Menge, D.N.L., Reed, S.C., Cleveland, C.C., 2013. Biological nitrogen fixation: rates, patterns
616 and ecological controls in terrestrial ecosystems. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368,
617 20130119.
618 Wang, H.L., Deng, N., Wu, D.Y., Hu, S., Kou, M., 2017a. Long-termnet transformation and quantitative
619 molecular mechanisms of soil nitrogen during natural vegetation recovery of abandoned farmland on
620 the Loess Plateau of China. Sci. Total Environ. 607, 152–159.
621 Wang, L., Cao, Y., Wang, E.T., Qiao, Y.J., Jiao, S., Liu, Z.S., Zhao, L., Wei, G.H., 2016. Biodiversity and
622 biogeography of rhizobia associated with common bean (*Phaseolus vulgaris L.*) in Shaanxi Province.
623 Syst. Appl. Microbiol. 39, 211–219.
624 Wang, Y., Li, C., Kou, Y., Wang, J., Tu, B., Li, H., Li, X., Wang, C., Yao, M., 2017b. Soil pH is a major driver
625 of soil diazotrophic community assembly in Qinghai-Tibet alpine meadows. Soil Biol. Biochem. 115,
626 547–555.
627 Wang, Y., Qian, P.-Y., 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S
628 ribosomal DNA amplicons in metagenomic studies. PLoS One 4, e7401.
629 Xu, W., Zhu, M., Zhang, Z., Ma, Z., Liu, H., Chen, L., Cao, G., Zhao, X., Schmid, B., He, J.S., 2018.
630 Experimentally simulating warmer and wetter climate additively improves rangeland quality on the
631 Tibetan Plateau. J. Appl. Ecol..
632 Yang, T., Adams, J.M., Shi, Y., He, J.s., Jing, X., Chen, L., Tedersoo, L., Chu, H., 2017. Soil fungal diversity in
633 natural grasslands of the Tibetan Plateau: associations with plant diversity and productivity. New
634 Phytol. 215, 756–765.
635 Yeager, C.M., Kornosky, J.L., Morgan, R.E., Cain, E.C., Garcia-Pichel, F., Housman, D.C., Belnap, J., Kuske,
636 C.R., 2007. Three distinct clades of cultured heterocystous cyanobacteria constitute the dominant N_2-
637 fixing members of biological soil crusts of the Colorado Plateau, USA. FEMS Microbiol. Ecol. 60, 85–
638 97.

639 Yeager, C.M., Kuske, C.R., Carney, T.D., Johnson, S.L., Ticknor, L.O., Belnap, J., 2012. Response of
640 biological soil crust diazotrophs to season, altered summer precipitation, and year-round increased
641 temperature in an arid grassland of the Colorado Plateau, USA. Front. Microbiol. 3, 358.
642 Zehr, J.P., Jenkins, B.D., Short, S.M., Steward, G.F., 2003. Nitrogenase gene diversity and microbial
643 community structure: a cross-system comparison. Environ. Microbiol. 5, 539–554.
644 Zhang, B., Zhang, J., Liu, Y., Guo, Y., Shi, P., Wei, G., 2018a. Biogeography and ecological processes affecting
645 root-associated bacterial communities in soybean fields across China. Sci. Total Environ. 627, 20–27.
646 Zhang, J., Wang, F., Che, R.X., Wang, P., Liu, H.K., Ji, B.M., Cui, X.Y., 2016a. Precipitation shapes
647 communities of arbuscular mycorrhizal fungi in Tibetan alpine steppe. Sci. Rep. 6, 23488.
648 Zhang, J., Yan, X., Su, F., Li, Z., Wang, Y., Wei, Y., Ji, Y., Yang, Y., Zhou, X., Guo, H., Hu, S., 2018b. Long-
649 term N and P additions alter the scaling of plant nitrogen to phosphorus in a Tibetan alpine meadow.
650 Sci. Total Environ. 625, 440–448.
651 Zhang, L., Unteregelsbacher, S., Hafner, S., Xu, X., Schleuss, P.-M., Miehe, G., Kuzyakov, Y., 2017. Fate of
652 organic and inorganic nitrogen in crusted and non-crusted *Kobresia* grasslands. Land Degrad. Dev. 28,
653 166–174.
654 Zhang, Y., Dong, S.K., Gao, Q.Z., Liu, S.L., Zhou, H.K., Ganjurjav, H., Wang, X.X., 2016b. Climate change
655 and human activities altered the diversity and composition of soil microbial community in alpine
656 grasslands of the Qinghai-Tibetan Plateau. Sci. Total Environ. 562, 353–363.
657 Zhang, Y., Zhang, C.B., Wang, Z.Q., Chen, Y.Z., Gang, C.C., An, R., Li, J.L., 2016c. Vegetation dynamics and
658 its driving forces from climate change and human activities in the Three-River Source Region, China
659 from 1982 to 2012. Sci. Total Environ. 563, 210–220.
660 Zhang, Y.G., Li, D.Q., Wang, H.M., Xiao, Q.M., Liu, X.D., 2006. Molecular diversity of nitrogen-fixing
661 bacteria from the Tibetan Plateau, China. FEMS Microbiol. Lett. 260, 134–142.
662 Zhang, Z.-Y., Gao, X.-H., Zhang, Y.-J., Jia, M., Lu, X.-J., Ma, Y.-C., Tian, F., Xie, Q., Tang, S.-K., 2015.
663 *Skermanella rubra* sp. nov., a bacterium isolated from the desert of Xinjiang, China. Antonie van
664 Leeuwenhoek 108, 627–632.
665 Zhao, K., Jing, X., Sanders, N.J., Chen, L., Shi, Y., Flynn, D.F., Wang, Y., Chu, H., Liang, W., He, J.S., 2017.
666 On the controls of abundance for soil- dwelling organisms on the Tibetan Plateau. Ecosphere 8,
667 e01901.

668 Zhou, X.B., Smith, H., Silva, A.G., Belnap, J., Garcia-Pichel, F., 2016. Differential responses of dinitrogen
669 fixation, diazotrophic Cyanobacteria and ammonia oxidation reveal a potential warming-induced
670 imbalance of the N-cycle in biological soil crusts. PLoS One 11, e0164932.
671 Zhu, W., Huang, J., Li, M., Li, X., Wang, G., 2014. Genomic analysis of *Skermanella stibiiresistens* type strain
672 SB22 T. Stand. Genomic Sci. 9, 1211.
673

Figure legends

Fig. 1. The distribution of the sampling sites.

Fig. 2. The distribution of soil diazotroph abundance (a) and Shannon diversity (b) on the Tibetan Plateau. The maps of diazotroph abundance and Shannon diversity were illustrated through anti-distance-weighted interpolation methods in ArcMap (v10.2.2).

Fig. 3. The relationships among the abundance (a) and community composition (b) of diazotrophs, prokaryotes, and plants. AGB: plant aboveground biomass; PCCD: plant community composition dissimilarity; gene copy numbers were presented in g^{-1} dry soil; and AGB were presented in $g.m^{-2}$. The community composition dissimilarities were calculated based on Bray-Curtis methods at the OTU level.

Fig. 4. The relationships among the abundance and diversity of soil diazotrophs and environmental factors. NDVI: normalized difference vegetation index; TOC: total organic carbon; and MAP: mean annual precipitation. The models were constructed through path way analysis based on stepwise regressions. The vectors between the environmental factors and the diazotroph indices represented the direct effects; and the vectors among the environmental factors represented the indirect effects.

Fig. 5. Soil diazotroph community compositions at the order level. Only the orders with average relative abundance higher than 0.5% were shown separately; "Other" represents the sum of the proportions of the less abundant orders (lower than 0.5%).

Fig. 6. The NMDS ordination of diazotroph community compositions (a) and the differences in soil diazotroph community compositions between alpine meadow and steppe (b). The line in the NMDS ordination plot represents the mean annual precipitation isocline (mm). The community composition dissimilarities were calculated based on Bray-Curtis methods at OTU level. The diazotroph lineages in green were significantly more abundant in alpine meadow soils, while the diazotroph lineages in pink were significantly more abundant in alpine steppe soils. These differences were determined using the LEfSe analysis, and the threshold on the logarithmic LDA score was 2.0.

Fig. 7. The relationships between the soil diazotroph community compositions and environmental factors. AGB: plant aboveground biomass ($g.m^{-2}$); AP: soil available P contents ($mg.kg^{-1}$); NO_3^--N: soil nitrate N contents (mg kg^{-1}); NH_4^+-N; soil ammonium N contents ($mg.kg^{-1}$); SM: soil moisture contents (%); BGB: plant belowground biomass ($kg.m^{-2}$); TOC: soil total organic carbon contents ($g.kg^{-1}$); ELE: elevation (m); TN: soil total nitrogen contents ($g.kg^{-1}$); DOC: soil dissolved organic carbon contents ($mg.kg^{-1}$); MAT: mean annual temperature (°C); and NDVI: normalized difference vegetation index. The multiple regression tree analysis was conducted at the OTU level.

Table 1
Click here to download Table: Table 1.docx

Table 1 The relationships between soil diazotroph indices and environmental factors.

	nifH gene copies*		Shannon diversity		Community composition	
	r	*P*	r	*P*	r^2	*P*
Elevation	0.072	0.606	0.125	0.368	0.020	0.592
NDVI	0.505	< **0.001**	0.393	**0.003**	0.264	< **0.001**
Mean annual precipitation	0.421	**0.002**	0.554	< **0.001**	0.335	< **0.001**
Mean annual temperature	-0.150	0.281	-0.052	0.711	0.008	0.818
Plant cover	0.469	< **0.001**	0.306	**0.024**	0.263	**0.001**
Plant richness	0.225	0.101	0.204	0.140	0.100	0.067
Plant aboveground biomass	0.494	< **0.001**	0.306	**0.024**	0.168	**0.009**
Soil pH values	-0.400	**0.003**	-0.388	**0.004**	0.182	**0.006**
Soil moisture contents	0.765	< **0.001**	0.535	< **0.001**	0.504	< **0.001**
Soil total N contents (TN)	0.648	< **0.001**	0.410	**0.002**	0.286	< **0.001**
Soil total organic C contents (TOC)	0.733	< **0.001**	0.467	< **0.001**	0.402	< **0.001**
Soil dissolved organic C contents	0.562	< **0.001**	0.474	< **0.001**	0.355	< **0.001**
Soil available P contents (AP)	0.674	< **0.001**	0.358	**0.008**	0.406	< **0.001**
Soil NH_4^+-N contents	0.521	< **0.001**	0.468	< **0.001**	0.330	< **0.001**
Soil NO_3^--N contents	0.063	0.649	0.013	0.924	0.059	0.211
Soil inorganic N contents (IN)	0.474	< **0.001**	0.436	< **0.001**	0.337	< **0.001**
Soil TOC/TN	0.592	< **0.001**	0.292	**0.032**	0.246	**0.001**
Soil AP/IN	0.227	0.099	-0.071	0.611	0.034	0.412

* The *nifH* gene copies and the corresponding environmental factors were logarithmically transformed. NDVI: normalized difference vegetation index. The relationships between the *nifH* gene copies or diazotroph Shannon diversity and the environmental factors were determined using Pearson correlation; and the relationships between the diazotroph community composition and the environmental factors were analyzed using the "envfit" based on NMDS ordination.

Figure 1

Click here to download high resolution image

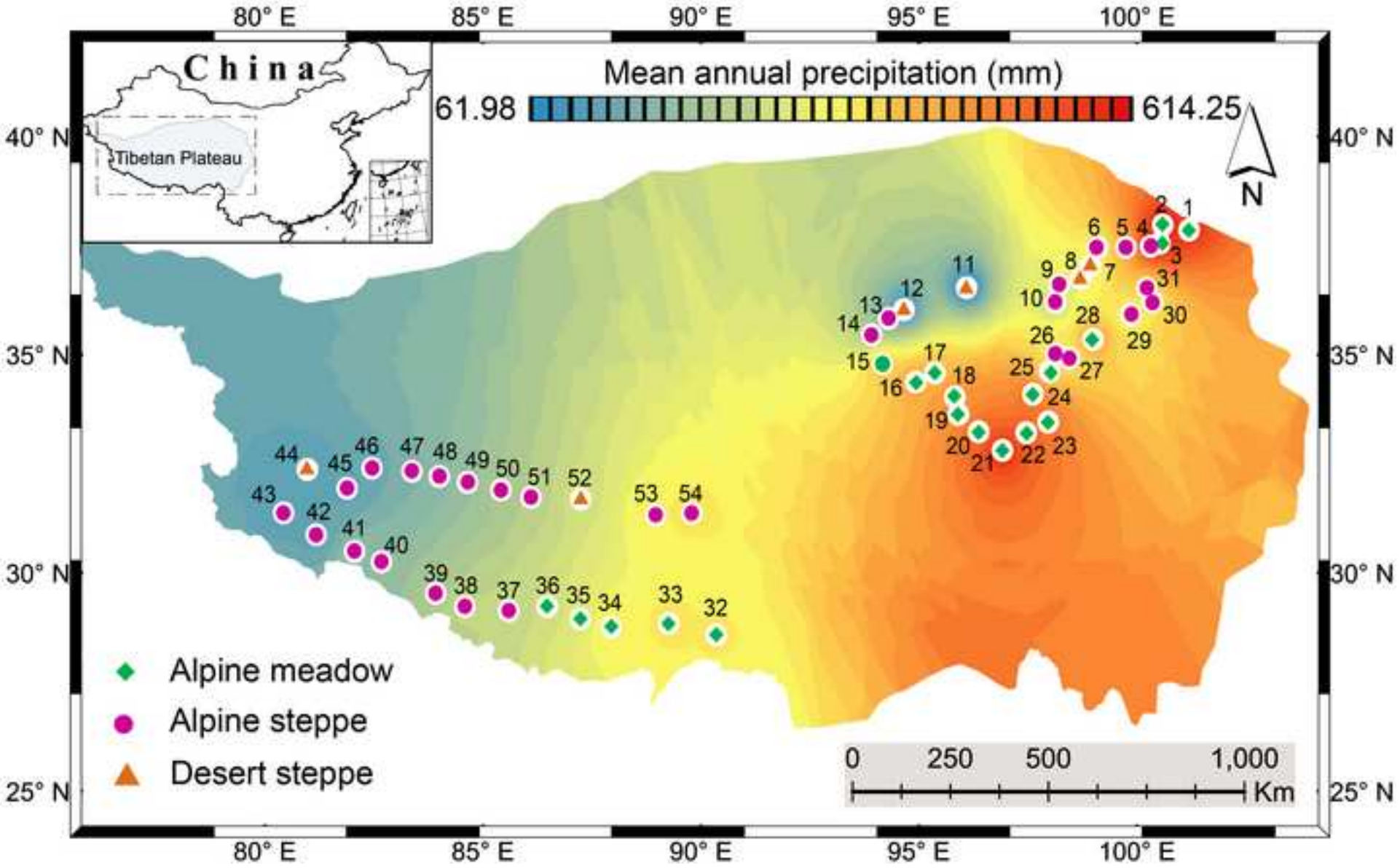


Figure 2

Click here to download high resolution image

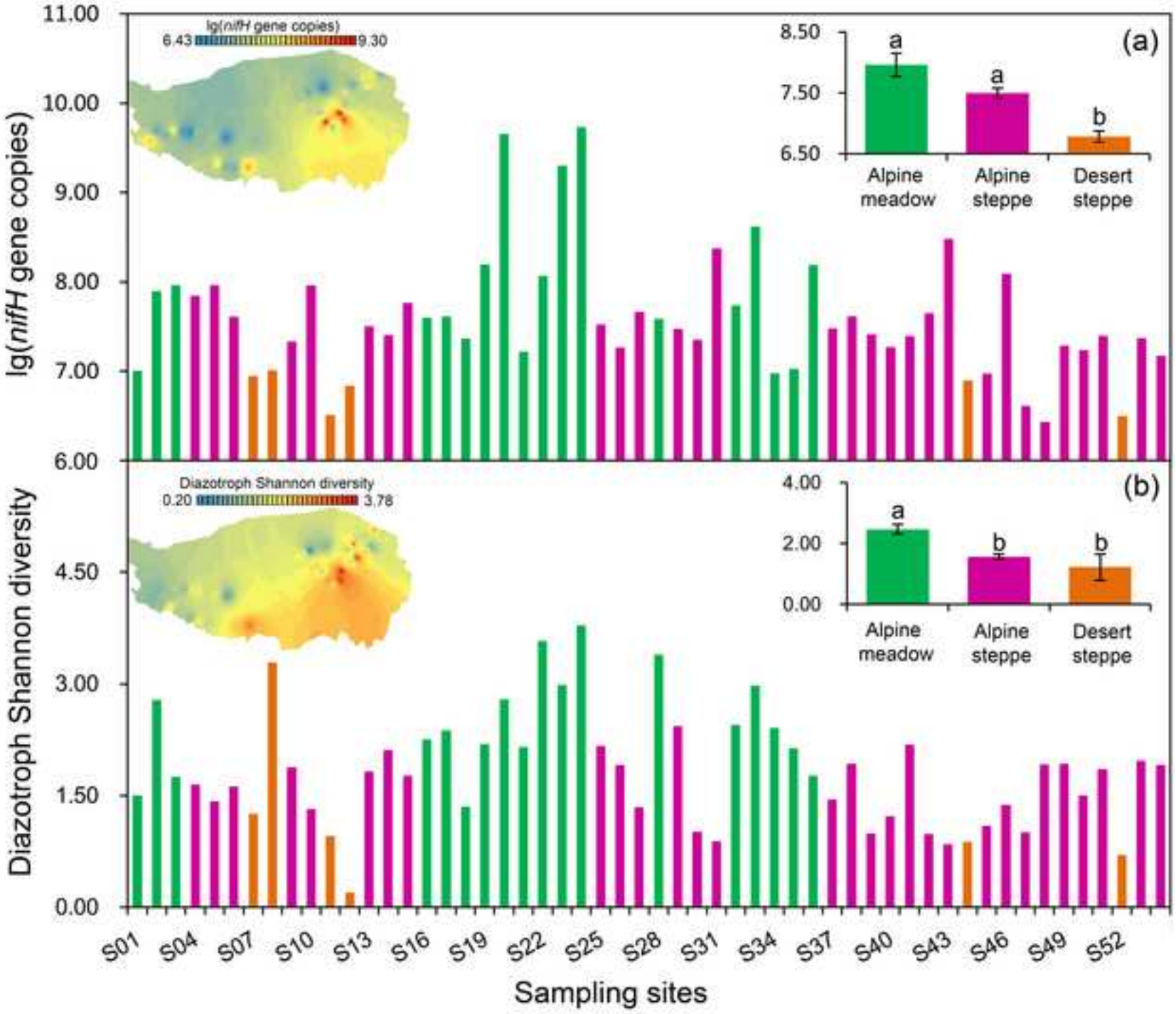


Figure 3

Click here to download high resolution image

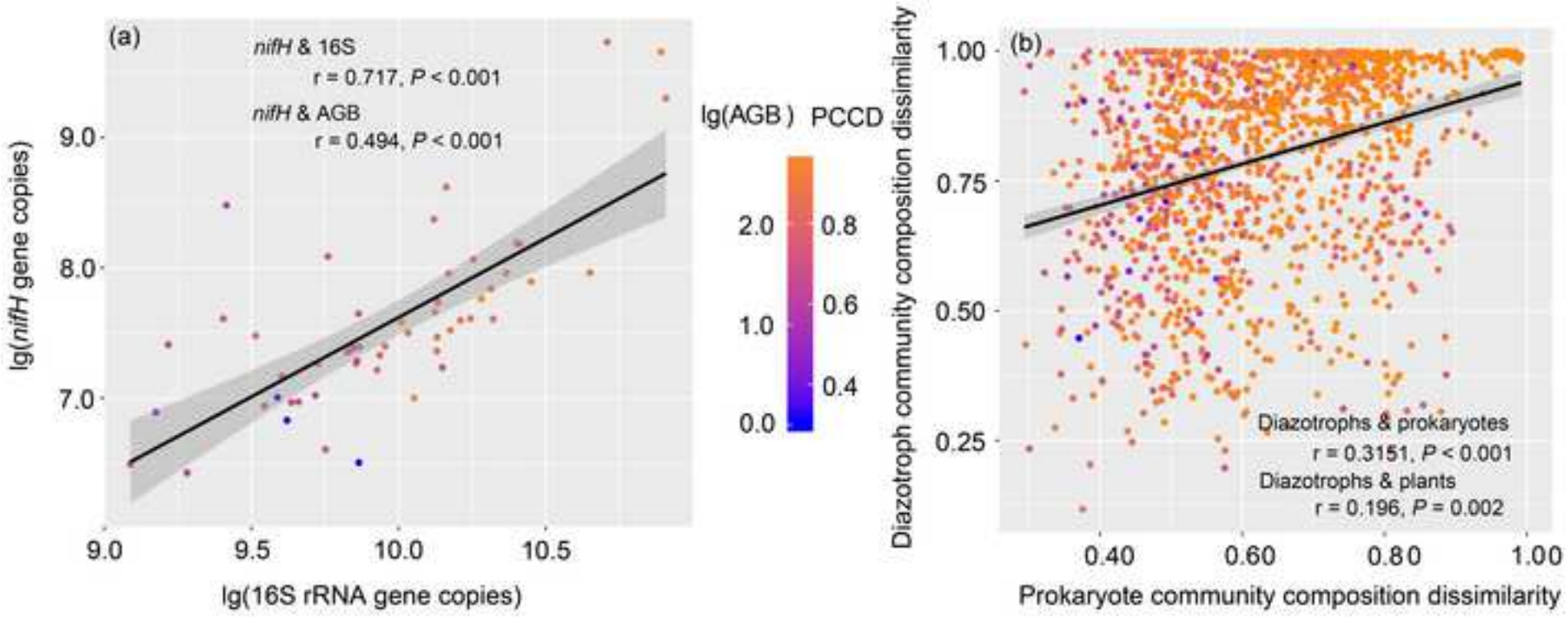


Figure 4

Click here to download high resolution image

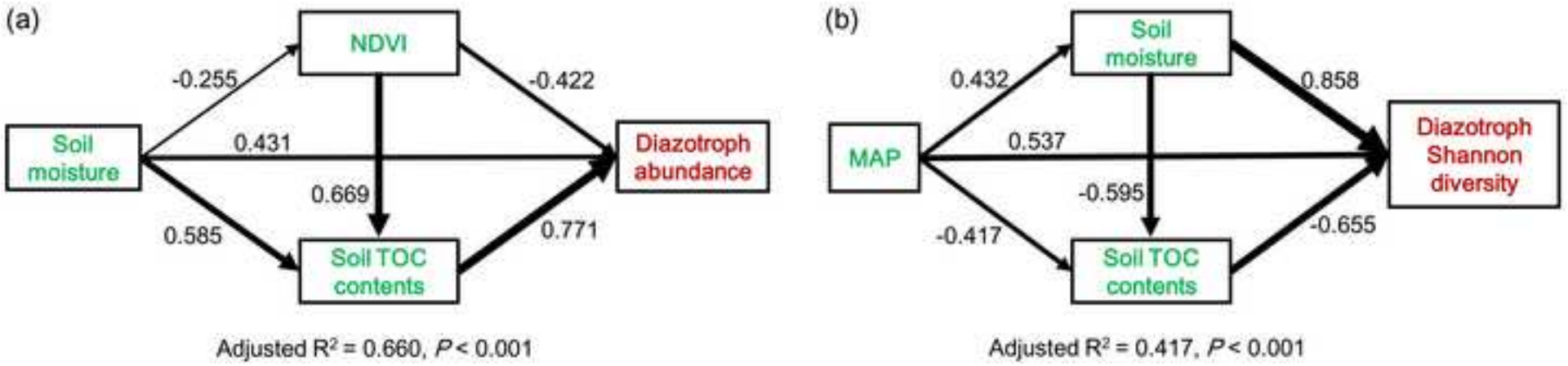


Figure 5

Click here to download high resolution image

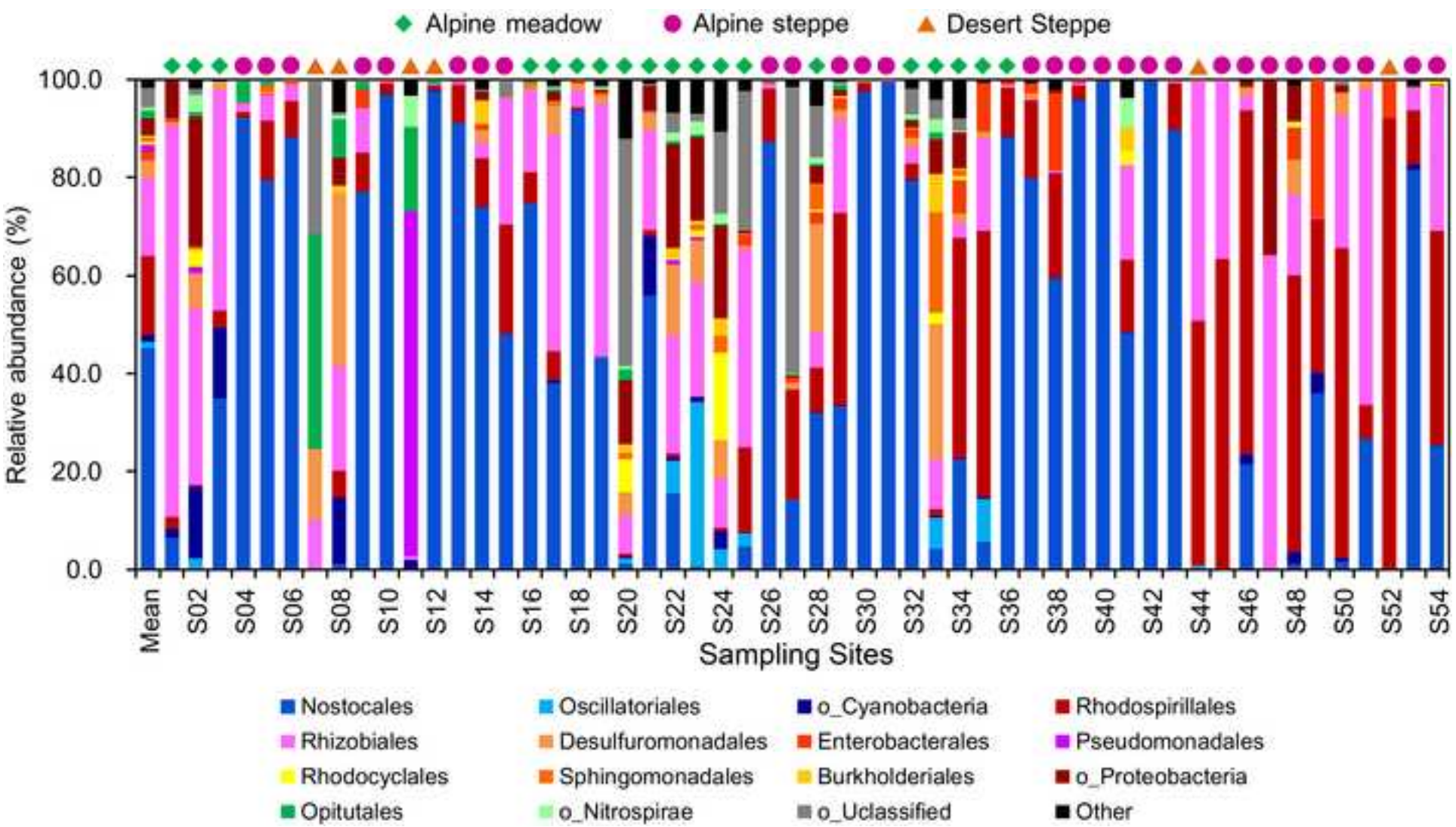


Figure 6

Click here to download high resolution image

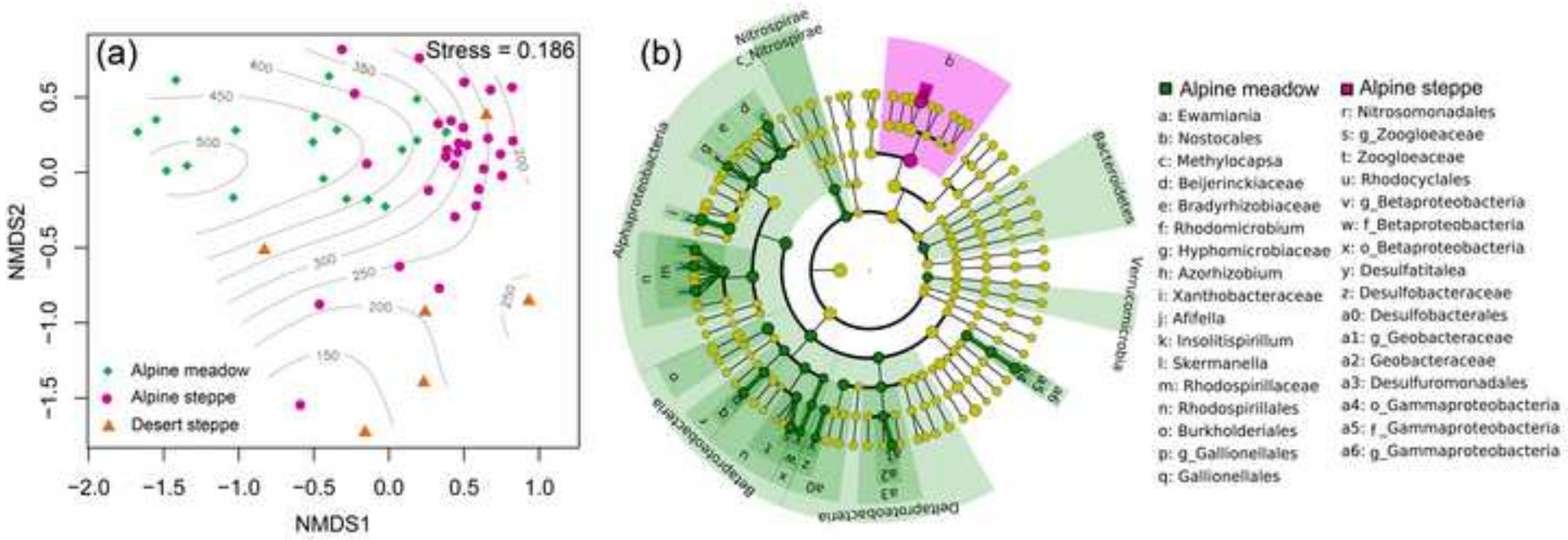
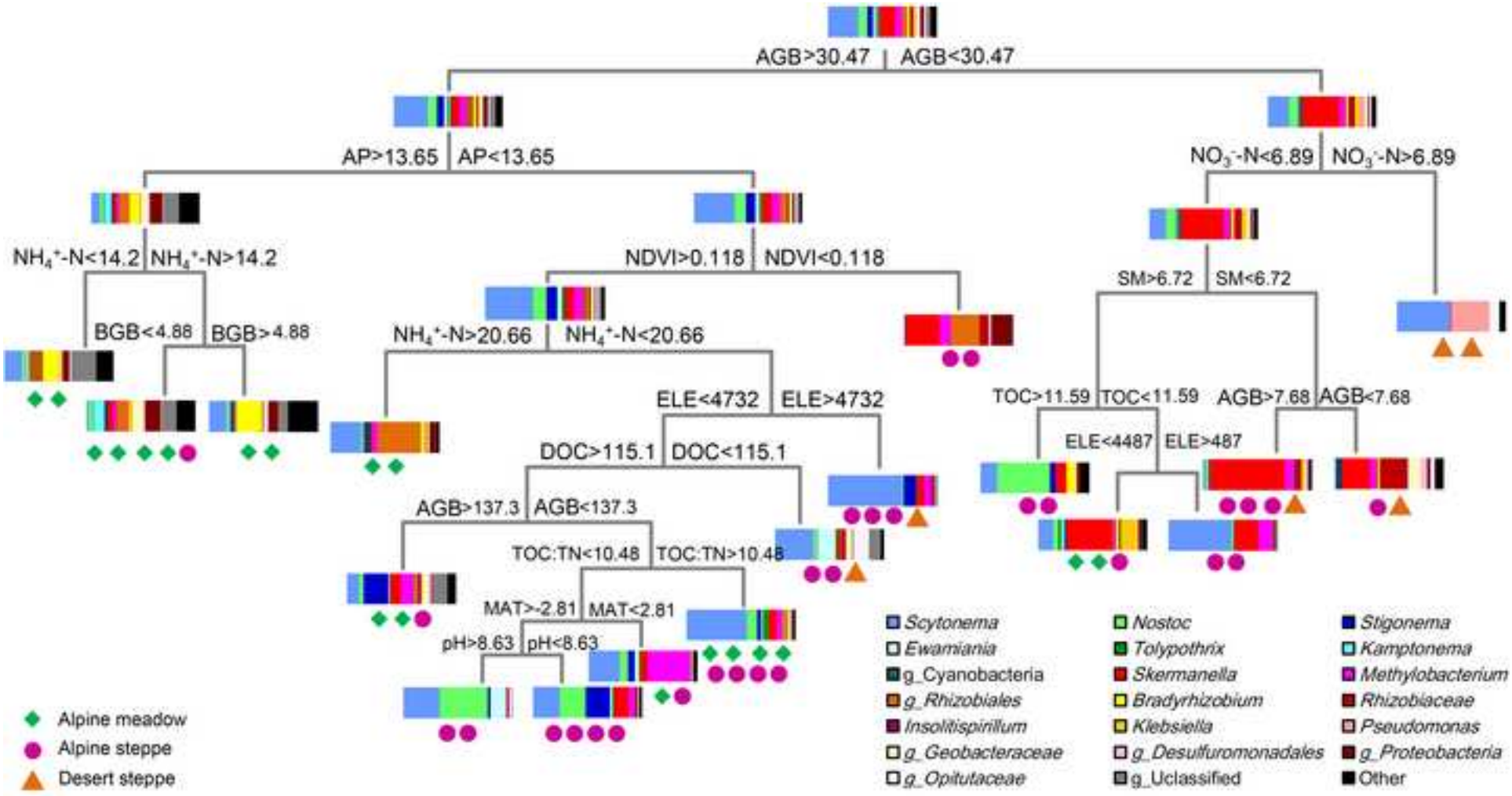


Figure 7

Click here to download high resolution image



Supplementary material for on-line publication only

Click here to download Supplementary material for on-line publication only: Supplementary materials.xlsx