Non-cyanobacterial diazotrophs dominate dinitrogen fixation in biological soil crusts at the early stage of crust formation.

Charles Pepe-Ranney [1], Chantal Koechli [1], Ruth Potrafka [2], Ferran Garcia-Pichel [2], Daniel H Buckley [1,*]

[1] *Cornell University, Department of Crop and Soil Sciences, Ithaca, NY, USA*
[2] *Arizona State University, School of Life Sciences, Tempe, AZ 85287, USA.*

Correspondence*:
Daniel H Buckley
Cornell University, Department of Crop and Soil Sciences, Ithaca, NY, USA,

1 ABSTRACT

Biological soil crusts (BSC) cover a vast global area and are key components of ecosystem productivity in arid soils. In particular, BSC contribute significantly to the nitrogen (N) budget via N_2-fixation. N_2-fixation in mature crusts is largely attributed to heterocystous cyanobacteria, however, early successional crusts possess few N-fixing cyanobacteria and this suggests that microorganisms other than cyanobacteria mediate N_2-fixation during the early stages of BSC development. DNA stable isotope probing (DNA-SIP) with $^{15}N_2$ revealed that *Clostridiaceae* and *Proteobacteria* are the most common microorganisms to assimilate ^{15}N in early successional crusts. The maximum $^{15}N_2$-assimilating *Clostridiaceae* and *Proteobacteria* Operational Taxonomic Unit (OTU, all sequences at least 97% sequence identity to OTU seed) relative abundance in environmental BSC SSU rRNA gene sequence surveys was 0.00225% and 0.00127% in any single sample, respectively. Their low abundance may explain why these heterotrophic diazotrophs have not previously been characterized in BSC. Diazotrophs play a critical role in BSC formation and characterization of these organisms represents a crucial step towards understanding how antropogenic change will effect the formation and ecological function of BSC in arid ecosystems.

2 INTRODUCTION

Biological soil crusts (BSC) are specialized microbial mat communitites that form at the soil surface in arid environments and fill a variety of important ecological functions. BSC occupy plant interspaces and cover a wide, global geographic range (Garcia-Pichel et al., 2003b). The ground cover of BSC on the Colorado Plateau has been measured as high as 80% by remote sensing (Karnieli et al., 2003). The global biomass of BSC *Cyanobacteria* alone is estimated at 54 x 10^{12} g C (Garcia-Pichel et al., 2003b). BSC play are responsible for significant nitrogen (N) flux (for review of BSC N_2-fixation see Belnap (2003)). N_2-fixation represents the dominant source of new ecosystem N in more than 80% of BSC from diverse sites across North America, Africa, and Australia (Evans and Belnap, 1999), while atmospheric N deposition was a dominant source of N in only a minority of sites. The presence of BSC is positively correlated with vascular plant survival due in part to BSC ecosystem N contributions (for review of BSC-vascular plant interactions see Belnap et al. (2003)). Climate change and disturbance could alter BSC microbial community structure/membership and possibly BSC diazotroph diversity and N_2-fixation.

BSC N_2-fixation rate studies (typically employing the acetylene reduction assay (ARA)) have explored BSC diazotroph activity across various ecological gradients. Reported BSC N_2-fixation rates vary significantly across samples and studies (Evans and Lange, 2001). The reasons for inter-site and inter-study variability are complex and likely include the spatial heterogeneity of BSC (Evans and Lange, 2001) and the impact of recent environmental change on N_2-fixation rates (see Belnap (2001) for discussion). Moreover, the ARA assay is subject to methodological artifacts that can complicate making robust comparisons across sample types that differ in physical and biological characteristics (see Belnap (2001) for review). Nonetheless, N_2-fixation rates are consistently higher in mature BSC than in early successional BSC (Belnap, 2002; Yeager et al., 2004). This difference may be due to the proliferation of heterocystous *Cyanobacteria* in mature BSC and is consistent with the theory that heterocystous *Cyanobacteria* provide the main source of fixed-N. Alternatively, the N_2-fixation rate differences between early successional and mature BSC might be attributable to methodological artifacts. For instance, N_2-fixation in mature BSC is maximal at the crust surface (coincident with heterocystous cyanobacteria) while it is maximal below the crust surface in early successional BSC (Johnson et al., 2005). Diffusional limitation can cause ARA to underestimate N_2-fixation which occurs below the crust surface and as a result ARA can systematically underestimate rates of N_2-fixation in early successional BSC (Johnson et al., 2005). Diffusion would not be an issue when measuring N_2-fixation rates in mature crusts as nitrogenase activity peaks near the surface. The difference between N_2 fixation rates of early successional and mature BSC were not statistically significant when N_2-fixation rates were estimated by integrating ARA rates from thin (1-3mm) slices along BSC depth profiles (as opposed to intact cores) (Johnson et al., 2005).

Molecular studies of BSC microbial diversity include explorations of the BSC microbial community vertical profile (Garcia-Pichel et al., 2003a), BSC *nifH* gene content surveys (e.g. Yeager et al. (2004), Yeager et al. (2012), Yeager et al. (2006) and Steppe et al. (1996)), and next-generation-sequencing (NGS) enabled studies of BSC SSU rRNA genes across wide geographic ranges (Garcia-Pichel et al., 2013; Steven et al., 2013). Early successional BSC are often described as ”light” in appearance relative to ”dark” mature BSC (Belnap, 2002; Yeager et al., 2004). Mature BSC possess greater numbers of heterocystous *Cyanobacteria* (i.e *Scytonema*, *Spirirestis*, and *Nostoc* (Yeager et al., 2006, 2012)) than developing BSC but both early successional and mature BSC are dominated by non-heterocystous *Cyanobacteria* (*Microcoleus vaginatus* or *M. steenstrupii*) (Yeager et al., 2004; Garcia-Pichel et al., 2013). Heterocystous *Cyanobacteria* are the numerically dominant BSC diazotrophs in *nifH* clone libraries (Yeager et al., 2006, 2004, 2012). Eighty-nine perent of 693 *nifH* sequences derived from Colorado Plateau and New Mexico BSC samples were described as heterocystous cyanobacterial (non-cyanobacterial *nifH* sequences were largely attributed to alpha- and beta- *proteobacteria*) (Yeager et al., 2006). However, an early survey of Colorado Plateau BSC *nifH* diversity recovered *nifH* genes related to *Gammaproteobacteria* as well as a clade that included *nifH* genes from the anaerobes *Clostridium pssteurianum*, *Desulfovibrio gigas* and *Chromatium buderi*,

The influence of microbial community membership and structure on BSC N_2-fixation is an ongoing research question (Belnap, 2013). While the presence/abundance of heterocystous *Cyanobacteria* has been proposed as the mechanism behind increased N_2-fixation in mature BSC, it is unclear if mature BSC actually fix more N than early successional BSC (see Johnson et al. (2005)). More studies are necessary to elucidate the microbial membership influence on BSC N_2-fixation and to determine if heterocystous *Cyanobacteria* are the only keystone diazotrophs. The first step in defining structure function relationships with respect to N_2-fixation is a full accounting of BSC diazotrophs. Towards this end we conducted $^{15}N_2$ DNA stable isotope probing (DNA-SIP) experiments with early successional Colorado Plateau BSC. DNA-SIP with $^{15}N_2$ has not been attempted with BSC. DNA-SIP provides an accounting of *active* diazotrophs whereas *nifH* clone libraries account for microbes with the genomic potential for N_2-fixation. Further, we investiage the distribution of these active diazotrophs through collections of SSU rRNA gene sequences from BSC NGS microbial diversity surveys over a range of spatial scales and soil types (Garcia-Pichel et al., 2013; Steven et al., 2013).

3 RESULTS

3.1 ORDINATION OF CSCL GRADIENT FRACTION SSU RRNA SEQUENCE COLLECTIONS SHOWS HEAVY FRACTIONS FROM CONTROL AND LABELED CSCL GRADIENTS ARE DIFFERENT

BSC were incubated for 4 days in the presence or absence of $^{15}N_2$ and DNA was extracted for DNA-SIP at 2 and 4 days. Fractionation of CsCl gradients permitted separation of DNA on the basis of buoyant density. Ordination of Bray-Curtis (Bray and Curtis, 1957) distances between gradient fractions (based on OTU abundance within each fraction) reveals that labeled gradient fractions (i.e. gradient fractions from CsCl gradients with $^{15}N_2$ labeled DNA) diverge from control (i.e. DNA from incubations without $^{15}N_2$) at the heavy end of the CsCl gradients (Figure 1 and Figure S1). Bray-Curtis distances between heavy gradient fractions are consistent label/control groups (p-value: 0.001, r^2: 0.18, Adonis test (Anderson, 2001)).

3.2 OTUS RESPONSIVE TO $^{15}N_2$ ARE PRIMARILY *PROTEOBACTERIA* AND *CLOSTRIDIACEAE*

A statistically significant increase in OTU abundance in heavy fractions of $^{15}N_2$ labeled samples relative to corresponding control fractions provides evidence for OTUs that have incorporated ^{15}N into their DNA. Specifically, we compared OTU proportion means between labeled and control samples from heavy gradient fractions using statistics developed to find differentially expressed genes with RNASeq data (McMurdie and Holmes, 2014; Love et al., 2014). OTUs that incorporated ^{15}N into DNA and increased in bouyant density were identified by rejecting the the null hypothesis that the labeled versus control proportion mean ratio for an OTU (considering only heavy fractions) was below a chosen threshold (see methods). p-values were adjusted by the BH method (Benjamini and Hochberg, 1995) and we used a false discovery rate (FDR) cutoff of 0.10 (typical FDR threshold in gene expression data analysis). A total of 2,127 and 2,160 OTUs were detected in days 2 and 4, respectively, and interrogated for evidence of $^{15}N_2$-labelling. Of these OTUs, only 208 and 233, respectively, passed a sparsity threshold we applied as an independent filtering step to pre-screen out OTUs not likely to produce significant p-values (see Love et al. (2014) for discussion of independent filtering). Of OTUs passing sparsity criteria 38 were found to be enriched significantly in heavy fractions relative to control. These OTUs likely incorporated ^{15}N into DNA ($^{15}N_2$ "responders"). Of these 38, 26 are annotated as *Firmicutes*, 9 as *Proteobacteria*, 2 as *Acidobacteria* and 1 as *Actinobacteria* (Figure 3, Figure 2). If the responder OTUs are ranked by descending, moderated proportion mean labeled:control ratios, the top 10 ratios (i.e. the 10 OTUs that were most enriched in the labeled gradients considering only heavy fractions) are either *Firmicutes* (6 OTUs) or *Proteobacteria* (4 OTUs) (Figure 4). Centroid or seed sequences for strongly responding *Proteobacteria* OTUs all share high sequence identity (>98.48%, Table 1) with cultivars from genera known to possess diazotrophs including *Klebsiella*, *Shigella*, *Acinetobacter*, and *Ideonella*. None of the *Firmicutes* OTU centroids in the top 10 responders share greater than 97% sequence identity with sequences in the LTP database (release 115) (see Table 1). OTUs that passed the sparsity threshold but were not classified as ^{15}N-responsive were subsequently tested with the null hypothesis that the OTU proportion mean ratio was above the selected threshold. Rejecting the second null would indicate an OTU did *not* incorporate ^{15}N into biomass. There were 58 and 70 "non-responders" at days 2 and 4, respectively. OTUs that did not pass sparsity or could not be classified as either a responder or non-responder are simply ambiguous with respect to ^{15}N labelling.

3.3 ^{15}N-RESPONSIVE OTUS IN ENVIRONMENTAL SAMPLES

Five of the 6 *Firmicutes* with the strongest response to ^{15}N-labelling (Table 1) belong in the *Clostridiacea*. We only observed one of these strongly responding *Clostridiaceae* in the data presented by Garcia-Pichel et al. (2013), "OTU.108" (closest BLAST hit in LTP Release 115 – *Caloramotor proteoclasticus*, BLAST

%ID 96.94, Accession X90488). OTU.108 was found in two samples both characterized as ”light” (i.e. early successional) crust. One other *Clostridiaceae* OTU with a proportion mean ratio (labeled:control) p-value less than 0.10 but outside the top 10 responders was found in the Garcia-Pichel et al. (2013) data (a ”light” crust sample) (Figure 2). None of the strongly responding *Clostridiacea* were found in the sequences in Steven et al. (2013) (Figure 2). *Clostridiaceae* ^{15}N-responder OTUs are not closely related to cultivars. (Table 1, Figure 5). One of the proteobacterial OTUs with the strongest ^{15}N response (Table 1) was found in Garcia-Pichel et al. (2013) samples (closest BLAST hit in LTP Release 115, BLAST %ID 100, Accession ZD3440, *Acinetobacter johnsonii*). None of the strongly responding *Protebacteria* OTUs were found in the Steven et al. (2013) samples. A responder OTU was found in a Steven et al. (2013) sample 133 times. Eighty-three times in sub-biocrust samples, 50 times in crust samples (see Figure 2). Two ^{15}N-responsive OTUs were found in an extensive number of environmental samples (61 of 65 samples from the combined data sets of Garcia-Pichel et al. (2013) and Steven et al. (2013)). Both OTUs were annotated as *Acidobacteria* but shared little sequence identity to any cultivar SSU rRNA gene sequences in the LTP (Release 115), with best LTP BLAST hits of 81.91 and 81.32% identity (Table 1). Additionally, the magnitude of the ^{15}N-response for each OTU was weak relative to other putative responders (3. Of the remaining 36 stable isotope responder OTUs, only 14 were observed in the environmental data (Figure 2, Figure S5).

3.4 COMPARING SEQUENCE COLLECTIONS AT ”STUDY”-LEVEL

We compared the sequences this DNA-SIP experiment to two previous surveys of SSU rRNA amplicons from BSC communities: the Garcia-Pichel et al. (2013) and Steven et al. (2013) study. There were 3,079 OTUs (209,354 total sequences after quality control) in the DNA-SIP data, 3,203 OTUs (129,033 total sequences after quality control) in the Garcia-Pichel et al. (2013) study, and 2,481 OTUs (129,358 total sequences after quality control) in the Steven et al. (2013) study. There were a total of 4,340 OTUs in all three datasets. Of the total 4,340 OTU centroids established for this study, 445 have matches in the Living Tree Project (LTP) (a collection of 16S gene sequences for all sequenced type strains (Yarza et al., 2008)) at greater or equal than 97% sequence identity (LTP version 115). That is, 445 of 4,340 OTUs are closely related to cultivars. The DNA-SIP data set shares 56% OTUs with the Steven et al. (2013) data and 46% of OTUs with the Garcia-Pichel et al. (2013) data (where total OTUs are from the combined data for each pairwise comparison). The Steven et al. (2013) and Garcia-Pichel et al. (2013) studies share 46% of OTUs. *Cyanobacteria* and *Proteobacteria* were the top two phylum-level sequence annotations for all three studies of BSC. Only the DNA-SIP data had more *Proteobacteria* annotations than *Cyanobacteria*. *Proteobacteria* represented the 29.8% of sequence annotations in DNA-SIP data as opposed to 17.8% and 19.2% for the Garcia-Pichel et al. (2013) and Steven et al. (2013) data, respectively. There is a contrast in the total percentage of sequences annotated as *Firmicutes* between the raw environmental samples and the DNA-SIP data. *Firmicutes* represent only 0.21% and 0.23% of total phylum level sequence annotations in the Steven et al. (2013) and Garcia-Pichel et al. (2013) studies, respectively (Figure S2). In the DNA-SIP sequence collection *Firmicutes* make up 19% of phylum level sequence annotations. SIP places focus upon organisms based on isotope incorporation and has the ability to detect activity by low abundance members of the community. DNA from OTUs that incopororate ^{15}N into their biomass moves towards the heavy end of the CsCl gradient and therefore OTUs in “labeled” DNA are enriched in the full data pool relative to bulk DNA. Phylum-level taxonomic annotations of ^{15}N-responsive OTUs (i.e. *Firmicutes* and *Proteobacteria*) are enriched in the DNA-SIP data relative to environmental data (Figure S2). Also in contrast for the DNA-SIP versus environmental data is the number of putative heterocystous *Cyanobacteria* sequences. Only 0.29% of *Cyanobacteria* sequences in the DNA-SIP data are annotated as belonging to ”Subsection IV” which is the heterocystous order of *Cyanobacteria* in the Silva taxonomic nomenclature (Pruesse et al., 2007). In the Steven et al. (2013) and Garcia-Pichel et al. (2013) studies 15% and 23%, respectively, of *Cyanobacteria* sequences are annotated as belonging to ”Subsection IV”.

4 DISCUSSION

BSC N-fixation has long been attributed to heterocystous *Cyanobacteria* and molecular surveys of BSC *nifH* gene content have been consistent with this hypothesis finding cyanobacterial *nifH* types to be numerically dominant in *nifH* gene clone libraries from BSC (Yeager et al., 2006, 2004, 2012). However, $^{15}N_2$ DNA-SIP revealed non-cyanobacterial microorganisms fixed N_2 in early successional BSC samples. DNA from early successional BSC samples–incubated for 2 and 4 days in the presence or absence of $^{15}N_2$–was collected and separated by bouyant density in CsCl density gradients. Heavy CsCl gradient fractions from gradients with ^{15}N-labelled DNA were different in phylogenetic membership/structure than heavy fractions with unlabeled DNA (Figure 1 and Figure S1)). Further, heavy gradient fractions clustered by DNA type (labeled or control) (Figure 1). Therefore, headspace $^{15}N_2$ in early successional BSC microcosms was incorporated into DNA. Further, OTUs enriched in labeled gradient heavy fractions relative to control are the specific taxa that incorporated ^{15}N (from $^{15}N_2$) into biomass. *Proteobacteria* and *Clostridiaceae* represented most OTUs enriched in DNA from labeled gradient heavy fractions relative to control as revealed by a robust statistical framework for quantifying and evaluating differential OTU abundance in microbiome studies (McMurdie and Holmes, 2014; Love et al., 2014). Additionally, *Proteobacteria* and *Clostridiaceae* represented OTUs that most strongly responded to ^{15}N.

We propose three mechanisms that could bias *nifH* clone libraries against heteroptophic diazotrophs. First, polyploidy in *Cyanobacteria* (Griese et al., 2011) would inflate the representation of cyanobacteria in community DNA (beyond a cell ratio). Second, *nifH* PCR primers could be biased against heterotrophic diazotrophs. In general the nifH PCR primers used by Yeager et al. (2006, 2004, 2012) ("19F" and "nifH3") for the first round of nested PCR have broad specificity and display at least 86% *in silico* coverage for *Proteobacteria, Cyanobacteria* and "Cluster III" (which includes clostridial *nifH*) reference *nifH* sequences (Gaby and Buckley, 2012). In the second round of the nested PCR protocol used by Yeager et al. (2006, 2004, 2012), primer "nifH11" is' biased against "Cluster III" (50% *in silico* coverage of reference *nifH* sequences) relative to *Proteobacteria* (79% coverage) and *Cyanobacteria* (67% coverage), and, primer "nifH22" matches *Proteobacteria*, *Cyanobacteria* and "Cluster III" reference sequences poorly (16%, 23% and 21% *in silico* coverage, respectively) (Gaby and Buckley, 2012). Unfortunately, it is difficult to assess or quantify this bias (in either direction) without knowing the *nifH* gene content *de novo*. Third, heterocysts (the specialized N-fixing cells along the trichome of filamentous heterocystous *Cyanobacteria* such as *Nostoc* and *Scytonema*) may be overrepresented (above a cell ratio) in *nifH* clone libraries. Heterocysts make up a fraction of cells along a trichome and even the non-heterocyst (non-N-fixing) cells in the trichome will possess the *nifH* gene. As a result of polyploidy and the frequency of heterocysts in a cyanobacterial filament, the ratio of cyanobacterial to heterotroph *nifH* gene copies may be 10^2-10^3 higher than the ratio of heterocysts to heterotrophic diazotroph cells. Regardless, our results suggest that BSC N-fixation may include a significant non-cyanobacterial component that requires further assessment across a more comprehensive sampling of BSC types.

We did not observe evidence for N-fixation by heterocystous *Cyanobacteria* in the early successional BSC samples used in this study. One possible explanation for our results is that the early successional BSC samples used in this study possessed too few heterocystous *Cyanobacteria* to statistically evaluate their ^{15}N-incorporation. Indeed, only 0.29% of sequences from this study's DNA-SIP 16S rRNA gene sequence collections were from heterocystous *Cyanobacteria* (see results) as opposed to 15% and 23% of total sequences in the Steven et al. (2013) and Garcia-Pichel et al. (2013) data, respectively. Nonetheless, we would still expect even low abundance diazotrophs to show evidence for ^{15}N-incorporation, provided sequence counts were not too sparse in heavy fractions. The OTUs defined by selected heterocystous *Cyanobacteria* sequences presented in Yeager et al. (2006), however, all fall below the sparsity threshold used in our analysis (see methods). Given the sparsity of heterocystous *Cyanobacteria* sequences in the DNA-SIP data set, it is not possible to assess whether heterocystous *Cyanobacteria* incorporated ^{15}N during the incubation. It should be noted that early successional BSC samples possess much less heterocystous *Cyanobacteria* in general (Figure S3) so the samples used in this study are not necessarily unrepresentative of typical early successional BSC simply because they are lacking heterocystous *Cyanobacteria*.

The OTUs that did appear to incorporate ^{15}N during the incubation were predominantly *Proteobacteria* and *Firmicutes*. The *Proteobacteria* OTUs for which ^{15}N-incorporation signal was strongest all shared high sequence identity (>=98.48%) with 16S sequences from cultivars in genera with known diazotrophs (Table 1). The *Firmicutes* that displayed signal for ^{15}N-incorporation (predominantly *Clostridiaceae*) were not closely related to any cultivars (Table 1, Figure 5). These BSC *Clostrodiaceae* diazotrophs represent a gap in culture collections. Assessing the physiological response of these diazotrophic *Clostridiaceae* to temperature would be useful for predicting how climate change will affect the BSC nitrogen budget.

Although too undersampled in the environmental data sets to reach statistical conclusions, ^{15}N-responsive OTUs were found more often in sub-crust or or early successional BSC samples as opposed to crust or mature samples (Figure 2 and Figure S5). This result generates some hypotheses that are counter to prior discussions regarding BSC diazotroph temporal dynamics. Specifically, the succession of BSC may not mark the *emergence* of diazotrophs in BSC but rather the *transition* of the diazotroph community from heterotroph to heterocystous *Cyanobacteria* dominance. Additionally, sub-biocrust soil may contribute significantly to the arid ecosystem N budget.

We propose that fast-growing heterotrophic diazotrophs such as *Clostridiaceae* are likely BSC ecosystem pioneers. *M. vaginatus* accumulates compatible solutes such as trehalose and sucrose as osmoprotectants during dessication (Rajeev et al., 2013). Additionally, although not demonstrated specifically with *M. vaginatus*, microorganisms can rapidly excrete compatible solutes upon wetting (Poolman and Glaasker, 1998). Many *Clostridiaceae* have a saccharolytic metabolism (Wiegel et al., 2006) and *Clostridiaceae* isolates have been shown to utilize trehalose and/or sucrose (summarized in Wiegel et al. (2006)). Further, *Clostridiaceae* isolates are fast-growing (doubling times typically between 30 min and 3 hr when grown on monosaccharides in culture (Wiegel et al., 2006)). Upon wetting, the early successional BSC environment may become rapidly rich in compatible solutes excreted by *M. vaginatus*. This boom-bust cycle would favor fast-growing microorganisms such as *Clostridiaceae* that can double rapidly and also fix N.

Rarefaction curves of all samples from Steven et al. (2013) and Garcia-Pichel et al. (2013) are still sharply increasing especially for sub-crust samples (Figure S4). Parametric richness estimates of BSC diversity indicate the Steven et al. (2013) and Garcia-Pichel et al. (2013) sequencing efforts recovered on average 40.5% (sd. 9.99%) and 45.5% (sd. 11.6%) of predicted 16S OTUs from samples (inset Figure S4), respectively. Therefore, is not alarming that few of the ^{15}N-responsive OTUS were found by Garcia-Pichel et al. (2013) and Steven et al. (2013). Even next-generation sequencing efforts of BSC 16S rRNA genes have only shallowly sampled the full diversity of BSC microbes.

4.1 CONCLUSION

Heterocystous *Cyanobacteria* are key contributors to the BSC N-budget, but, the ^{15}N-responsive OTUs found in this study and the *nifH* gene sequences from Steppe et al. (1996) in addition to the N-fixation rate data presented by Johnson et al. (2005) suggest there may be significant non-cyanobacterial BSC diazotrophs specifically within the *Clostrideaceae* and *Proteobacteria*. It seems clear that heterocystous *Cyanobacteria* increase in abundance with BSC age (Yeager et al., 2004). It is less clear if this transition marks the emergence of diazotrophs versus a re-structuring of the BSC diazotroph community from one dominated by *Clostridiaceae* and *Proteobacteria* to one predominantly heterocystous *Cyanobacteria*. DNA-SIP is a valuable tool in the molecular microbial ecologist's toolbox for identifying members of microbial community functional guilds (Neufeld et al., 2007). PCR-based surveys of diagnostic marker genes and DNA-SIP are both used to connect microbial phylogenetic types to microbial activities, but they occupy a non-overlapping set of strengths and weaknesses. DNA-SIP does not focus on a specific diagnostic marker but does identify *active* players in the studied process (i.e. N-fixation). Combined these tools can powerfully reveal connections between ecosystem membership/structure and function. Here we supplement previous surveys of BSC *nifH* diversity, a diagnostic marker PCR-driven approach, with $^{15}N_2$ DNA-SIP, While we do not confirm previous results, we expand knowledge of BSC diazotroph diversity. Predicting BSC N-fixation with respect to climate change, althered precipitation regimes and physical disturbance requires a careful accounting of diazotrophs including non-cyanobacterial types.

5 MATERIALS AND METHODS

5.1 BSC SAMPLING AND INCUBATION CONDITIONS

DNA was extracted from 1 g of BSC. Samples were taken from Green Butte, Arizona as previously described (site CP3, Beraldi-Campesi et al. (2009)). All samples were from early successional crusts as described by Johnson et al. (2005). Early successional BSC samples (37.5 cm^2, average mass 35 g) were incubated in sealed chambers under controlled atmosphere and in the light for 4 days. Crusts were sampled and transported while dry and wetted at initiation of the experiment. Treatments included an unlabeled control air headspace and $^{15}N_2$ enriched air (>98% atom $^{15}N_2$) headspace. Samples were taken at 2 days and 4 days incubation. Acetylene Reduction rates were measured daily. Acetylene reduction rates increased over the course of the experiment (0.8, 4.8, 8.8, and 14.5 μm m-2 hr-1 ethylene for days 1 through 4, respectively).

5.2 DNA EXTRACTION

DNA from each sample was extracted using a MoBio PowerSoil DNA Isolation Kit (following manufacturers protocol, but substituting a 2 minute bead beating for the vortexing step), and then gel purified to select high molecular weight DNA (>4 kb) using a 1% low melt agarose gel and β-agarase I for digestion (manufacturer's protocol, NEB, M0392S). Extracts were quantified using PicoGreen nucleic acid quantification dyes (Molecular Probes).

5.3 DNA-SIP

CsCl density gradients were formed in 4.7 mL polyallomer centrifuge tubes filled with gradient buffer (15mM Tris-HCl, pH 8; 15mM EDTA; 15mM KCl) which contained 1.725 g ml-1 CsCl. CsCl density was checked with a digital refractometer as described below. A total of 2.5-5 ug of DNA was added to each tube, and the tubes mixed, prior to centrifugation. Centrifugation was performed in a TLA-110 fixed angle rotor (Beckman Coulter) at 20C for 67 hours at 55,000 rpm. (Buckley et al., 2007). Centrifuged gradients were fractionated from bottom to top in 36 equal fractions of 100 μL, using a by syringe pump as described Manefield et al. (2002). The density of each fraction was determined using using an AR200 refractometer modified to accomidate 5ul samples as described previously (Buckley et al., 2007). DNA in each fraction was desalted on a filter plate (PALL, AcroPrep Advance 96 Filter Plate, Product Number 8035), using four washes with 300μL TE per fraction. After each wash, the filter plate was centrifuges at 500 g for 10 minutes, with a final spin of 20 minutes. Fractions were resuspended in 50 uL of TE buffer.

5.4 PCR, LIBRARY NORMALIZATION AND DNA SEQUENCING

Barcoded PCR of bacterial and archaeal 16S rRNA genes, in preparation for 454 Pyrosequencing, was carried out using primer set 515F/806R (Walters et al., 2011) (primers purchased from Integrated DNA Technologies). The primer 806R contained an 8 bp barcode sequence, a "TC" linker, and a Roche 454 B sequencing adaptor, while the primer 515F contained the Roche 454 A sequencing adapter. Each 25 μL reaction contained 1x PCR Gold Buffer (Roche), 2.5 mM $MgCl_2$, 200 μM of each of the four dNTPs (Promega), 0.5 mg/mL BSA (New England Biolabs), 0.3 μM of each primers, 1.25 U of Amplitaq Gold (Roche), and 8 μL of template. Each sample was amplified in triplicate. Thermal cycling occurred with an initial denaturation step of 5 minutes at 95C, followed by 40 cycles of amplification (20s at 95C, 20s at 53C, 30s at 72C), and a final extension step of 5 min at 72C. Triplicate amplicons were pooled and purified using Agencourt AMPure PCR purification beads, following manufacturers protocol. Once purified, amplicons were quantified using PicoGreen nucleic acid quantification dyes (Molecular Probes) and pooled together in equimolar amounts. Samples were sent to the Environmental Genomics Core Facility at the University of South Carolina (now Selah Genomics) to be run on a Roche FLX 454 pyrosequencing machine (FLX-Titanium platform).

5.5 DATA ANALYSIS

5.5.1 Sequence quality control Sequences were initially screened by maximum expected errors at a specific read length threshold (Edgar, 2013) which has been shown to be as effective as denoising with respect to removing pyrosequencing errors. Specifically, reads were first truncated to 230 nucleotides (nt) (all reads shorter than 230 nt were discarded) and any read that exceeded a maximum expected error threshold of 1.0 was removed. After truncation and max expected error trimming, 91% of original reads remained. Forward primer and barcode was then removed from the high quality, truncated reads. Remaining reads were taxonomically annotated using the "UClust" taxonomic annotation framework in the QIIME software package (Caporaso et al., 2010; Edgar, 2010) with cluster seeds from Silva SSU rRNA database (Pruesse et al., 2007) 97% sequence identity OTUs as reference (release 111Ref). Reads annotated as "Chloroplast", "Eukaryota", "Archaea", "Unassigned" or "mitochondria" were culled from the dataset. Finally, reads were aligned to the Silva reference alignment provided by the Mothur software package (Schloss et al., 2009) using the Mothur NAST aligner (DeSantis et al., 2006). All reads that did not align to the expected amplicon region of the SSU rRNA gene were discarded. Quality control parameters removed 34,716 of 258,763 raw reads.

5.5.2 Sequence clustering Sequences were distributed into OTUs using the UParse methodology (Edgar, 2013). Specifically, cluster seeds were identified using USearch on non-redundant reads sorted by count. The sequence identity threshold for establishing a new OTU centroid was 97%. After initial cluster centroid selection, select 16S rRNA gene sequences from Yeager et al. (2006) were added to the centroid collection. Specifically, Yeager et al. (2006) Colorado Plateau or Moab, Utah sequences were added which included the 16S rRNA gene sequences for *Calothrix* MCC-3A (accession DQ531700.1), *Nostoc commune* MCT-1 (accession DQ531903), *Nostoc commune* MFG-1 (accession DQ531699.1), *Scytonema hyalinum* DC-A (accession DQ531701.1), *Scytonema hyalinum* FGP-7A (accession DQ531697.1), *Spirirestis rafaelensis* LQ-10 (accession DQ531696.1). Centroid sequences that matched selected Yeager et al. (2006) sequences with greater than to 97% sequence identity were subsequently removed from the centroid collection. With USearch/UParse, potential chimeras are identified during OTU centroid selection and are not allowed to become cluster centroids effectively removing chimeras from the read pool. All quality controlled reads were then mapped to cluster centroids at an identity threshold of 97% again using USearch. A total of 95.6% of quality controlled reads could be mapped to centroids. Unmapped reads do not count towards sample counts and are removed from downstream analyses. The USearch software version for cluster generation was 7.0.1090. Garcia-Pichel et al. (2013) and Steven et al. (2013)) sequences were quality screened by determining if they covered the expected region of the 16S rRNA gene (described above) and included as input to USearch for OTU centroid selection and subsequent mapping to OTU centroids.

5.5.3 Phylogenetic tree The alignment for the "*Clostridiaceae*" phylogeny was created using SSU-Align which is based on Infernal (Nawrocki and Eddy, 2013; Nawrocki et al., 2009). Columns in the alignment that were not included in the SSU-Align covariance models or were aligned with poor confidence (less than 95% of characters in a position had posterior probability alignment scores of at least 95%) were masked for phylogenetic reconstruction. Additionally, the alignment was trimmed to coordinates such that all sequences in the alignment began and ended at the same positions. The

"*Clostridiaceae*" tree included all top BLAST hits (parameters below) for ^{15}N *Clostridiaceae* responders in the Living Tree Project database (Yarza et al., 2008) in addition to BLAST hits within a sequence identity threshold of 97% to ^{15}N responders from the Silva SSURef_NR SSU rRNA database (Pruesse et al., 2007). Only one SSURef_NR115 hit per study per OTU ("study" was determined by "title" field) was selected for the tree. FastTree (Price et al., 2010) was used to build the tree and support values are SH-like scores reported by FastTree. Short sequences were mapped to the reference backbone using pplacer (Matsen et al., 2010) (default parameters). pplacer finds the edge placements that maximize phylogenetic likelihood. Prior to being mapped to the reference tree, short sequences were aligned to the reference alignment using Infernal (Nawrocki et al., 2009) against the same SSU-Align covariance model used to align reference sequences.

5.5.4 Identifying OTUs that incorporated ^{15}N into their DNA DNA-SIP is a culture-independent approach towards defining identity-function connections in microbial communities (Buckley, 2011; Neufeld et al., 2007; Radajewski and Murrell, 2001). Microbes participating in a specific process are identified on the basis of isotope assimilation into DNA. Isotopically labeled nucleic acids can be separated from unlabeled by buoyant density in a CsCl gradient. As the buoyant density of a macromolecule is dependent on many factors in addition to stable isotope incorporation (e.g. GC-content in nucleic acids (Youngblut and Buckley, 2014)), labeled nucleic acids from one microbial population may have the same buoyant density of unlabeled nucleic acids from another. Therefore it is imperative to compare density gradients with nucleic acids from heavy stable isotope incubations to gradients from "control" incubations where everything mimics the experimental conditions except that unlabeled substrates are used. By contrasting heavy density gradient fractions in experimental density gradients (hereafter referred to as "labeled" gradients) against heavy fractions in control gradients, the identities of microbes with labeled nucleic acids can be determined

We used an RNA-Seq differential expression statistical framework (Love et al., 2014) to find OTUs enriched in heavy fractions of labeled gradients relative to corresponding density fractions in control gradients (for review of RNA-Seq differential expression statistics applied to microbiome OTU count data see McMurdie and Holmes (2014)). We use the term differential abundance (coined by McMurdie and Holmes (2014)) to denote OTUs that have different proportion means across sample classes (in this case the only sample class is labeled/control). CsCl gradient fractions were categorized as "heavy" or "light". The heavy category denotes fractions with density values above 1.725 g/mL. Since we are only interested in enriched OTUs (labeled versus control), we used a one-sided z-test for differential abundance (the null hypothesis is the labeled:control proportion mean ratio for an OTU is less than a selected threshold). P-values were corrected with the Benjamini and Hochberg method (Benjamini and Hochberg, 1995). We selected a $\log_2$ fold change null threshold of 0.25 (or a labeled:control proportion mean ratio of 1.19). DESeq2 was used to calculate the moderated $\log_2$ fold change of labeled:control proportion mean ratios and corresponding standard errors. Mean ratio moderation allows for reliable ratio ranking such that high variance and likely statistically insignificant mean ratios are appropriately shrunk and subsequently ranked lower than they would be as raw ratios. Those OTUs that exhibit a statistically significant increase in proportion in heavy fractions from $^{15}N_2$-labeled samples relative to corresponding controls have increased significantly in bouyant density in response to $^{15}N_2$ treatment; a response that is expected for N_2-fixing organisms.

5.5.5 Community and Sequence Analysis BLAST searches were done with the "blastn" program from BLAST+ toolkit (Camacho et al., 2009) version 2.2.29+. Default parameters were always employed and the BioPython (Cock et al., 2009) BLAST+ wrapper was used to invoke the blastn program. Pandas (McKinney, 2012) and dplyr (Wickham and Francois, 2014) were used to parse and munge BLAST output tables.

Principal coordinate ordinations depict the relationship between samples at each time point (day 2 and 4). Bray-Curtis distances were used as the sample distance metric for ordination. The Phyloseq

(McMurdie and Holmes, 2014) wrapper for Vegan (Oksanen et al., 2013) (both R packages) was used to compute sample values along principal coordinate axes. GGplot2 (Wickham, 2009) was used to display sample points along the first and second principal axes. Adonis tests Anderson (2001) were done with default number of permutations (1000).

Rarefaction curves were created using bioinformatics modules in the PyCogent Python package (Knight et al., 2007). Parametric richness estimates were made with CatchAll using only the best model for total OTU estimates (Bunge, 2010).

All code to take raw sequencing data through the presented figures can be found at:

`http://nbviewer.ipython.org/github/chuckpr/NSIP_data_analysis`

REFERENCES

Marti J. Anderson. A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26(1):32–46, 2001. doi: 10.1111/j.1442-9993.2001.01070.pp.x. URL `http://dx.doi.org/10.1111/j.1442-9993.2001.01070.pp.x`.

J. Belnap. Factors Influencing Nitrogen Fixation and Nitrogen Release in Biological Soil Crusts. In *Biological Soil Crusts: Structure Function, and Management*, pages 241–261. Springer Science, 2001. doi: 10.1007/978-3-642-56475-8_19. URL `http://dx.doi.org/10.1007/978-3-642-56475-8_19`.

J. Belnap. Factors Influencing Nitrogen Fixation and Nitrogen Release in Biological Soil Crusts. In Jayne Belnap and OttoL. Lange, editors, *Biological Soil Crusts: Structure, Function, and Management*, volume 150 of *Ecological Studies*, pages 241–261. Springer Berlin Heidelberg, 2003. ISBN 978-3-540-43757-4. doi: 10.1007/978-3-642-56475-8_19. URL `http://dx.doi.org/10.1007/978-3-642-56475-8_19`.

J. Belnap, R. Prasse, and K.T. Harper. Influence of Biological Soil Crusts on Soil Environments and Vascular Plants. In Jayne Belnap and OttoL. Lange, editors, *Biological Soil Crusts: Structure, Function, and Management*, volume 150 of *Ecological Studies*, pages 281–300. Springer Berlin Heidelberg, 2003. ISBN 978-3-540-43757-4. doi: 10.1007/978-3-642-56475-8_21. URL `http://dx.doi.org/10.1007/978-3-642-56475-8_21`.

Jayne Belnap. Nitrogen fixation in biological soil crusts from southeast Utah USA. *Biology and Fertility of Soils*, 35(2):128–135, 2002. doi: 10.1007/s00374-002-0452-x. URL `http://dx.doi.org/10.1007/s00374-002-0452-x`.

Jayne Belnap. Some Like It Hot, Some Not. *Science*, 340(6140):1533–1534, 2013. doi: 10.1126/science.1240318. URL `http://www.sciencemag.org/content/340/6140/1533.short`.

Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57 (1):289–300, 1995. ISSN 00359246. doi: 10.2307/2346101. URL `http://dx.doi.org/10.2307/2346101`.

H. Beraldi-Campesi, H. E. Hartnett, A. Anbar, G. W. Gordon, and F. Garcia-Pichel. Effect of biological soil crusts on soil elemental concentrations: implications for biogeochemistry and as traceable biosignatures of ancient life on land. *Geobiology*, 7(3):348–359, 2009. doi: 10.1111/j.1472-4669.2009.00204.x. URL `http://dx.doi.org/10.1111/j.1472-4669.2009.00204.x`.

J. Roger Bray and J. T. Curtis. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs*, 27(4):325, 1957. doi: 10.2307/1942268. URL `http://dx.doi.org/10.2307/1942268`.

Daniel H. Buckley. Stable Isotope Probing Techniques Using 15N. In *Stable Isotope Probing and Related Technologies*, pages 129–147. American Society of Microbiology, 2011. doi: 10.1128/9781555816896.ch7. URL `http://dx.doi.org/10.1128/9781555816896.ch7`.

DH Buckley, V Huangyutitham, SF Hsu, and TA Nelson. Stable isotope probing with 15N2 reveals novel noncultivated diazotrophs in soil. *Appl Environ Microbiol*, 73:3196–204, 2007.

John Bunge. Estimating the Number of Species with Catchall. In *Biocomputing 2011*, pages 121–130. World Scientific, 2010. doi: 10.1142/9789814335058_0014. URL `http://dx.doi.org/10.1142/9789814335058_0014`.

C Camacho, G Coulouris, V Avagyan, N Ma, J Papadopoulos, K Bealer, and TL Madden. BLAST+: architecture and applications. 10:421, 2009.

JG Caporaso, J Kuczynski, J Stombaugh, K Bittinger, FD Bushman, EK Costello, N Fierer, AG Pea, JK Goodrich, JI Gordon, GA Huttley, ST Kelley, D Knights, JE Koenig, RE Ley, CA Lozupone, D McDonald, BD Muegge, M Pirrung, J Reeder, JR Sevinsky, PJ Turnbaugh, WA Walters, J Widmann, T Yatsunenko, J Zaneveld, and R Knight. QIIME allows analysis of high-throughput community sequencing data. 7:335–6, 2010.

PJ Cock, T Antao, JT Chang, BA Chapman, CJ Cox, A Dalke, I Friedberg, T Hamelryck, F Kauff, B Wilczynski, and Hoon MJ de. Biopython: freely available Python tools for computational molecular biology and bioinformatics. 25:1422–3, 2009.

TZ Jr DeSantis, P Hugenholtz, K Keller, EL Brodie, N Larsen, YM Piceno, R Phan, and GL Andersen. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. 34:W394–9, 2006.

RC Edgar. Search and clustering orders of magnitude faster than BLAST. 26:2460–1, 2010.

RC Edgar. UPARSE: highly accurate OTU sequences from microbial amplicon reads. 10:996–8, 2013.

R. D. Evans and J. Belnap. Long-Term Consequences of Disturbance on Nitrogen Dynamics in an Arid Ecosystem. *Ecology*, 80(1):150–160, 1999. doi: 10.1890/0012-9658(1999)080[0150:ltcodo]2.0.co;2. URL `http://dx.doi.org/10.1890/0012-9658(1999)080[0150:LTCODO]2.0.CO;2`.

R. D. Evans and O. L. Lange. Biological Soil Crusts and Ecosystem Nitrogen and Carbon Dynamics. In *Biological Soil Crusts: Structure Function, and Management*, pages 263–279. Springer Science, 2001. doi: 10.1007/978-3-642-56475-8_20. URL `http://dx.doi.org/10.1007/978-3-642-56475-8_20`.

John Christian Gaby and Daniel H. Buckley. A Comprehensive Evaluation of PCR Primers to Amplify the nifH Gene of Nitrogenase. *PLoS ONE*, 7(7):e42149, 2012. doi: 10.1371/journal.pone.0042149. URL `http://dx.doi.org/10.1371/journal.pone.0042149`.

F. Garcia-Pichel, S. L. Johnson, D. Youngkin, and J. Belnap. Small-Scale Vertical Distribution of Bacterial Biomass and Diversity in Biological Soil Crusts from Arid Lands in the Colorado Plateau. *Microbial Ecology*, 46(3):312–321, 2003a. doi: 10.1007/s00248-003-1004-0. URL `http://dx.doi.org/10.1007/s00248-003-1004-0`.

F. Garcia-Pichel, V. Loza, Y. Marusenko, P. Mateo, and R. M. Potrafka. Temperature Drives the Continental-Scale Distribution of Key Microbes in Topsoil Communities. *Science*, 340(6140):1574–1577, 2013. doi: 10.1126/science.1236404. URL `http://dx.doi.org/10.1126/science.1236404`.

Ferran Garcia-Pichel, Jayne Belnap, Susanne Neuer, and Ferdinand Schanz. Estimates of global cyanobacterial biomass and its distribution. *Algological Studies*, 109(1):213–227, 2003b.

Marco Griese, Christian Lange, and Jrg Soppa. Ploidy in cyanobacteria. *FEMS Microbiology Letters*, 323(2):124–131, 2011. doi: 10.1111/j.1574-6968.2011.02368.x. URL `http://dx.doi.org/10.1111/j.1574-6968.2011.02368.x`.

SL Johnson, CR Budinoff, J Belnap, and F Garcia-Pichel. Relevance of ammonium oxidation within biological soil crust communities. 7:1–12, 2005.

A. Karnieli, R.F. Kokaly, N.E. West, and R.N. Clark. Remote Sensing of Biological Soil Crusts. In Jayne Belnap and OttoL. Lange, editors, *Biological Soil Crusts: Structure, Function, and Management*, volume 150 of *Ecological Studies*, pages 431–455. Springer Berlin Heidelberg, 2003. ISBN 978-3-540-43757-4. doi: 10.1007/978-3-642-56475-8_31. URL `http://dx.doi.org/10.1007/978-3-642-56475-8_31`.

Rob Knight, Peter Maxwell, Amanda Birmingham, Jason Carnes, J Gregory Caporaso, Brett C Easton, Michael Eaton, Micah Hamady, Helen Lindsay, Zongzhi Liu, Catherine Lozupone, Daniel McDonald, Michael Robeson, Raymond Sammut, Sandra Smit, Matthew J Wakefield, Jeremy Widmann, Shandy Wikman, Stephanie Wilson, Hua Ying, and Gavin A Huttley. {PyCogent}: a toolkit for making sense from sequence. *Genome Biol*, 8(8):R171, 2007. doi: 10.1186/gb-2007-8-8-r171. URL http://dx.doi.org/10.1186/gb-2007-8-8-r171.

M. I. Love, W. Huber, and S. Anders. Moderated estimation of fold change and dispersion for {RNA}-Seq data with DESeq2. Technical report, 2014. URL http://dx.doi.org/10.1101/002832.

Frederick A Matsen, Robin B Kodner, and E Virginia Armbrust. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics*, 11(1):538, 2010. doi: 10.1186/1471-2105-11-538. URL http://dx.doi.org/10.1186/1471-2105-11-538.

Wes McKinney. pandas: Python Data Analysis Library. Online, 2012. URL http://pandas.pydata.org/.

PJ McMurdie and S Holmes. Waste not, want not: why rarefying microbiome data is inadmissible. 10: e1003531, 2014.

EP Nawrocki and SR Eddy. Infernal 1.1: 100-fold faster RNA homology searches. 29:2933–5, 2013.

EP Nawrocki, DL Kolbe, and SR Eddy. Infernal 1.0: inference of RNA alignments. 25:1335–7, 2009.

JD Neufeld, J Vohra, MG Dumont, T Lueders, M Manefield, MW Friedrich, and JC Murrell. DNA stable-isotope probing. 2:860–6, 2007.

Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, and Helene Wagner. *vegan: Community Ecology Package*, 2013. URL http://CRAN.R-project.org/package=vegan. R package version 2.0-10.

Bert Poolman and Erwin Glaasker. Regulation of compatible solute accumulation in bacteria. *Molecular Microbiology*, 29(2):397–407, 1998. doi: 10.1046/j.1365-2958.1998.00875.x. URL http://dx.doi.org/10.1046/j.1365-2958.1998.00875.x.

MN Price, PS Dehal, and AP Arkin. FastTree 2–approximately maximum-likelihood trees for large alignments. 5:e9490, 2010.

E Pruesse, C Quast, K Knittel, BM Fuchs, W Ludwig, J Peplies, and FO Glckner. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. 35:7188–96, 2007.

Stefan Radajewski and J Colin Murrell. Stable Isotope Probing for Detection of Methanotrophs After Enrichment with 13CH4. In *Gene Probes*, pages 149–157. Humana Press, sep 2001. doi: 10.1385/1-59259-238-4:149. URL http://dx.doi.org/10.1385/1-59259-238-4:149.

Lara Rajeev, Ulisses Nunes da Rocha, Niels Klitgord, Eric G Luning, Julian Fortney, Seth D Axen, Patrick M Shih, Nicholas J Bouskill, Benjamin P Bowen, Cheryl A Kerfeld, Ferran Garcia-Pichel, Eoin L Brodie, Trent R Northen, and Aindrila Mukhopadhyay. Dynamic cyanobacterial response to hydration and dehydration in a desert biological soil crust. *The ISME Journal*, 7(11):2178–2191, 2013. doi: 10.1038/ismej.2013.83. URL http://dx.doi.org/10.1038/ismej.2013.83.

PD Schloss, SL Westcott, T Ryabin, JR Hall, M Hartmann, EB Hollister, RA Lesniewski, BB Oakley, DH Parks, CJ Robinson, JW Sahl, B Stres, GG Thallinger, Horn DJ Van, and CF Weber. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. 75:7537–41, 2009.

T.F. Steppe, J.B. Olson, H.W. Paerl, R.W. Litaker, and J. Belnap. Consortial N2 fixation: a strategy for meeting nitrogen requirements of marine and terrestrial cyanobacterial mats. *FEMS Microbiology Ecology*, 21(3):149–156, 1996. doi: 10.1111/j.1574-6941.1996.tb00342.x. URL http://dx.doi.org/10.1111/j.1574-6941.1996.tb00342.x.

Blaire Steven, La Verne Gallegos-Graves, Jayne Belnap, and Cheryl R. Kuske. Dryland soil microbial communities display spatial biogeographic patterns associated with soil depth and soil parent material. *FEMS Microbiol Ecol*, 86(1):101–113, 2013. doi: 10.1111/1574-6941.12143. URL http://dx.doi.org/10.1111/1574-6941.12143.

WA Walters, JG Caporaso, CL Lauber, D Berg-Lyons, N Fierer, and R Knight. PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. 27:1159–61, 2011.

Hadley Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009. ISBN 978-0-387-98140-6. URL http://had.co.nz/ggplot2/book.

Hadley Wickham and Romain Francois. *dplyr: dplyr: a grammar of data manipulation*, 2014. URL http://CRAN.R-project.org/package=dplyr. R package version 0.2.
Juergen Wiegel, Ralph Tanner, and Fred A. Rainey. An Introduction to the Family Clostridiaceae. In *The Prokaryotes*, pages 654–678. Springer US, 2006. doi: 10.1007/0-387-30744-3_20. URL http://dx.doi.org/10.1007/0-387-30744-3_20.
Pablo Yarza, Michael Richter, Jörg Peplies, Jean Euzeby, Rudolf Amann, Karl-Heinz Schleifer, Wolfgang Ludwig, Frank Oliver Glöckner, and Ramon Rosselló-Móra. The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains. *Systematic and Applied Microbiology*, 31(4):241–250, 2008. doi: 10.1016/j.syapm.2008.07.001. URL http://dx.doi.org/10.1016/j.syapm.2008.07.001.
Chris M. Yeager, Jennifer L. Kornosky, Rachael E. Morgan, Elizabeth C. Cain, Ferran Garcia-Pichel, David C. Housman, Jayne Belnap, and Cheryl R. Kuske. Three distinct clades of cultured heterocystous cyanobacteria constitute the dominant N2-fixing members of biological soil crusts of the Colorado Plateau USA. *FEMS Microbiology Ecology*, 60(1):85–97, 2006. doi: 10.1111/j.1574-6941.2006.00265.x. URL http://dx.doi.org/10.1111/j.1574-6941.2006.00265.x.
Chris M. Yeager, Cheryl R. Kuske, Travis D. Carney, Shannon L. Johnson, Lawrence O. Ticknor, and Jayne Belnap. Response of Biological Soil Crust Diazotrophs to Season Altered Summer Precipitation, and Year-Round Increased Temperature in an Arid Grassland of the Colorado Plateau, USA. *Front. Microbio.*, 3, 2012. doi: 10.3389/fmicb.2012.00358. URL http://dx.doi.org/10.3389/fmicb.2012.00358.
CM Yeager, JL Kornosky, DC Housman, EE Grote, J Belnap, and CR Kuske. Diazotrophic community structure and function in two successional stages of biological soil crusts from the Colorado Plateau and Chihuahuan Desert. 70:973–83, 2004.
ND Youngblut and DH Buckley. Intra-genomic variation in G+C content and its implications for DNA stable isotope probing (DNA-SIP). 2014.

6 FIGURES AND LONG TABLES

Table 1. ^{15}N responders BLAST search against Living Tree Project

OTU ID	*Genera*	BLAST %ID	*Phylum*
OTU.78	*Desulfocella, Bryobacter*	80.31	*Acidobacteria*
OTU.55	*Desulfocella, Bryobacter*	81.03	*Acidobacteria*
OTU.116	*Streptomyces*	100.0	*Actinobacteria*
OTU.140	*Bacillus*	100.0	*Firmicutes*
OTU.3	*Bacillus*	100.0	*Firmicutes*
OTU.1747	*Clostridium*	94.36	*Firmicutes*
OTU.327	*Clostridium*	94.92	*Firmicutes*
OTU.17	*Clostridium*	95.45	*Firmicutes*
OTU.61	*Clostridium*	95.92	*Firmicutes*
OTU.11	*Clostridium*	95.94	*Firmicutes*
OTU.2175	*Clostridium*	95.96	*Firmicutes*
OTU.1673	*Clostridium*	95.9	*Firmicutes*
OTU.330	*Clostridium*	96.94	*Firmicutes*
OTU.75	*Clostridium*	96.97	*Firmicutes*
OTU.643	*Clostridium*	97.45	*Firmicutes*
OTU.3712	*Clostridium, Eubacterium*	96.43	*Firmicutes*
OTU.37	*Clostridium, Ilyobacter*	96.43	*Firmicutes*
OTU.2404	*Domibacillus*	99.49	*Firmicutes*
OTU.4167	*Fonticella*	93.43	*Firmicutes*
OTU.4037	*Fonticella*	93.85	*Firmicutes*
OTU.57	*Fonticella, Caloramator*	93.88	*Firmicutes*
OTU.575	*Gracilibacter*	94.42	*Firmicutes*
OTU.259	*Parasporobacterium*	98.47	*Firmicutes*
OTU.278	*Symbiobacterium*	90.62	*Firmicutes*
OTU.88	*Symbiobacterium*	92.86	*Firmicutes*
OTU.342	*Acinetobacter*	100.0	*Proteobacteria*
OTU.263	*Azospirillum*	98.48	*Proteobacteria*
OTU.137	*Azospirillum*	98.98	*Proteobacteria*
OTU.176	*Delftia*	100.0	*Proteobacteria*
OTU.14	*Klebsiella, Pantoea, Erwinia, Enterobacter, Kluyvera, Buttiauxella*	99.49	*Proteobacteria*
OTU.586	*Ottowia, Diaphorobacter, Ideonella, Vitreoscilla, Comamonas*	98.48	*Proteobacteria*

Table 1 – continued from previous page

OTU ID	*Genera*	BLAST %ID	*Phylum*
OTU.321	*Pseudomonas*	100.0	*Proteobacteria*
OTU.54	*Shigella, Escherichia*	100.0	*Proteobacteria*

Figure 1. Ordination of heavy gradient fractions by Bray-Curtis distances on the basis of OTU content.

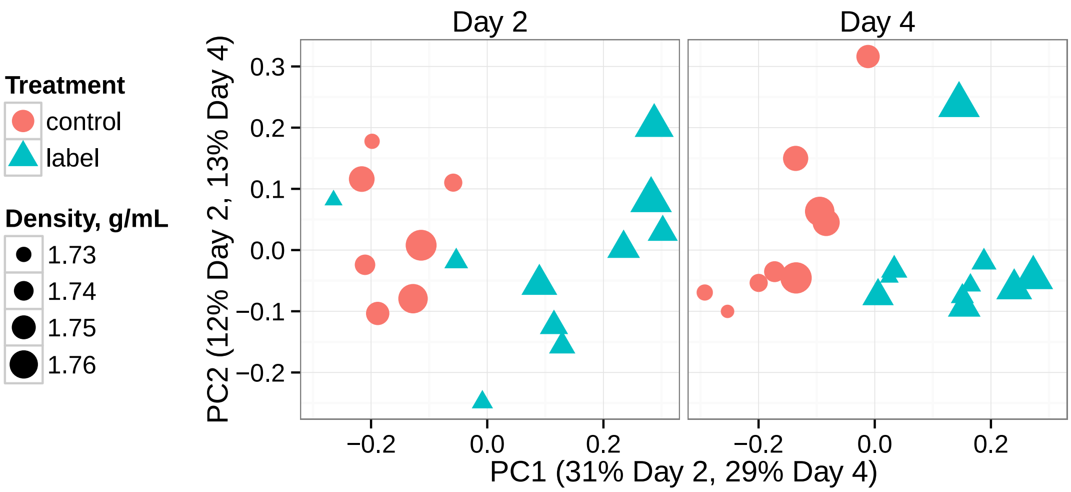


Figure 2. Phylogenetic trees of OTUs passing sparsity threshold for **A** *Proteobacteria*, **B** *Acidobacteria* and **C** *Firmicutes*. ^{15}N-responders are identified by dots present in column **i**. Log$_2$ of proportion mean ratios (labeled:control samples) for each OTU are presented as a heatmap in column **ii** with results from days 2 and 4 on the left and right sides of the column respectively. High values indicate ^{15}N incorporation. Presence/absence of OTUs (black indicates presence) in lichen, light, or dark environmental samples (Garcia-Pichel et al., 2013) is shown in column **iii**. Presence/absence of OTUs (black indicates presence) in crust and below crust samples (Steven et al., 2013) is shown in column **iv**

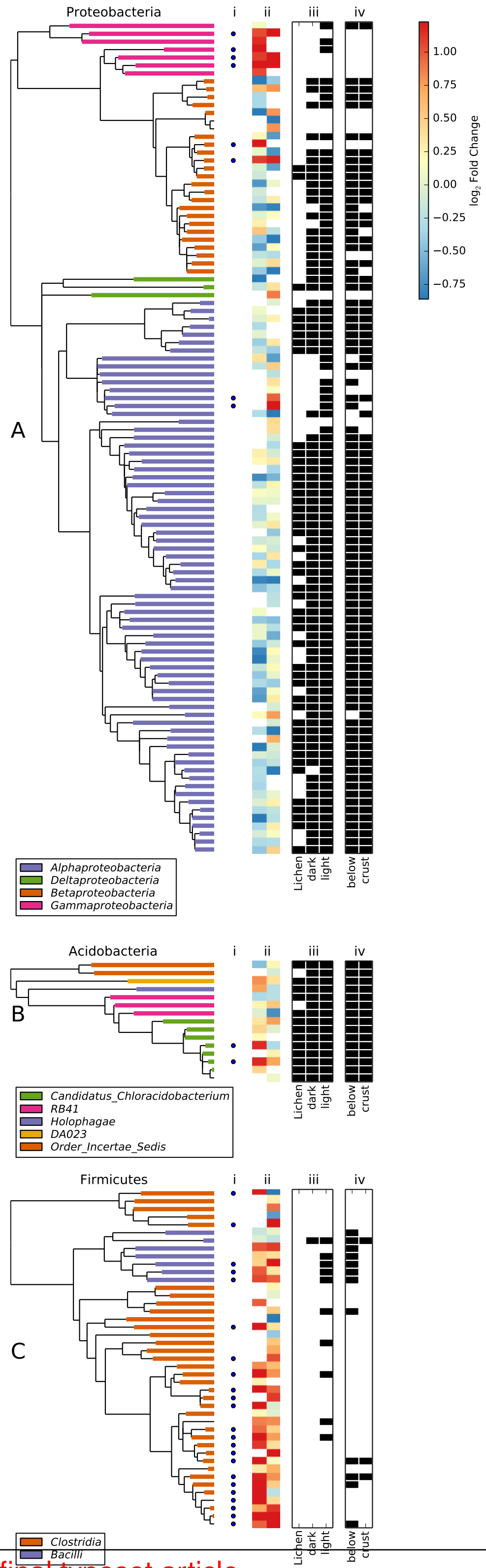


Figure 3. Moderated $\log_2$ of proportion mean ratios for labeled versus control gradients (heavy fractions only, densities >1.725 g/mL). All OTUs passing the sparsity treshold (see methods) at a specific incubation day are shown. Red color denotes a proportion mean ratio that has a corresponding adjusted p-value below a false discovery rate of 10% (the null model is that the proportion mean is ratio is below 0.25). The horizontal line is the proportion mean threshold for the null model, 0.25. The inset figure summarizes the taxonomy of OTUs that with proportion mean ratio p-vaules under 0.10 for at least one time point.

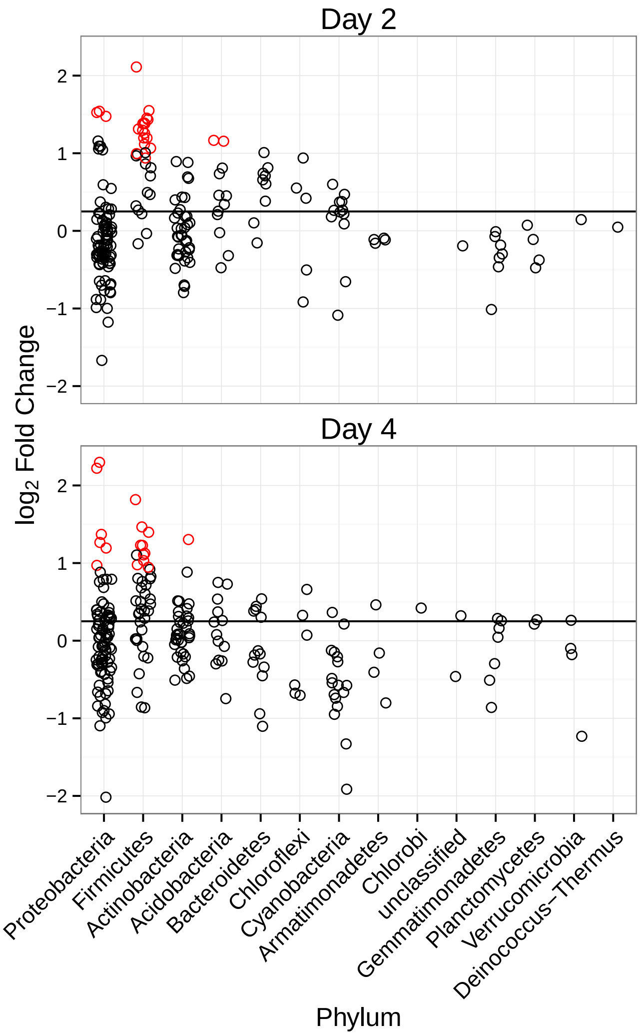


Figure 4. Relative abundance values in heavy fractions (density greater or equal to 1.725 g/mL) for the top 10 ^{15}N "responders" (putative diazotrophs, see results for selection criteria of top 10) at each incubation day. See Table 1 for BLAST results against the LTP database (release 115). Point area is proportional to CsCl gradient fraction density, and color signifies control (red) or labeled (blue) treatment.

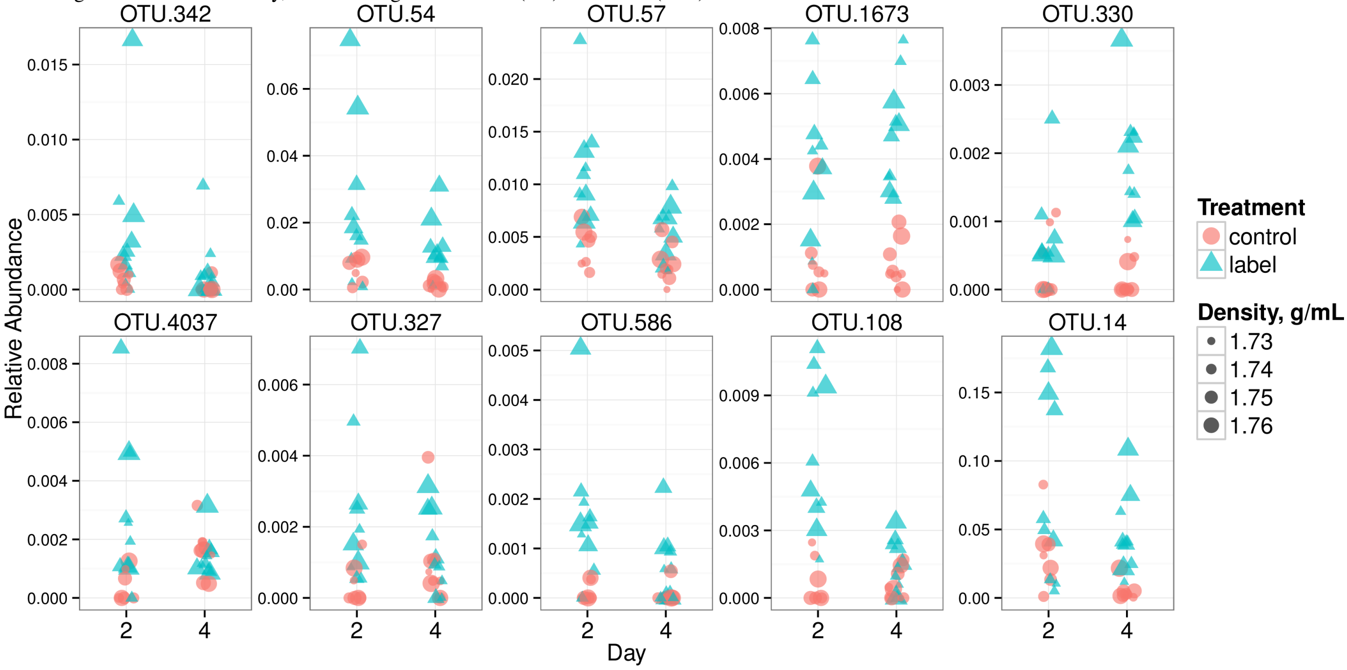
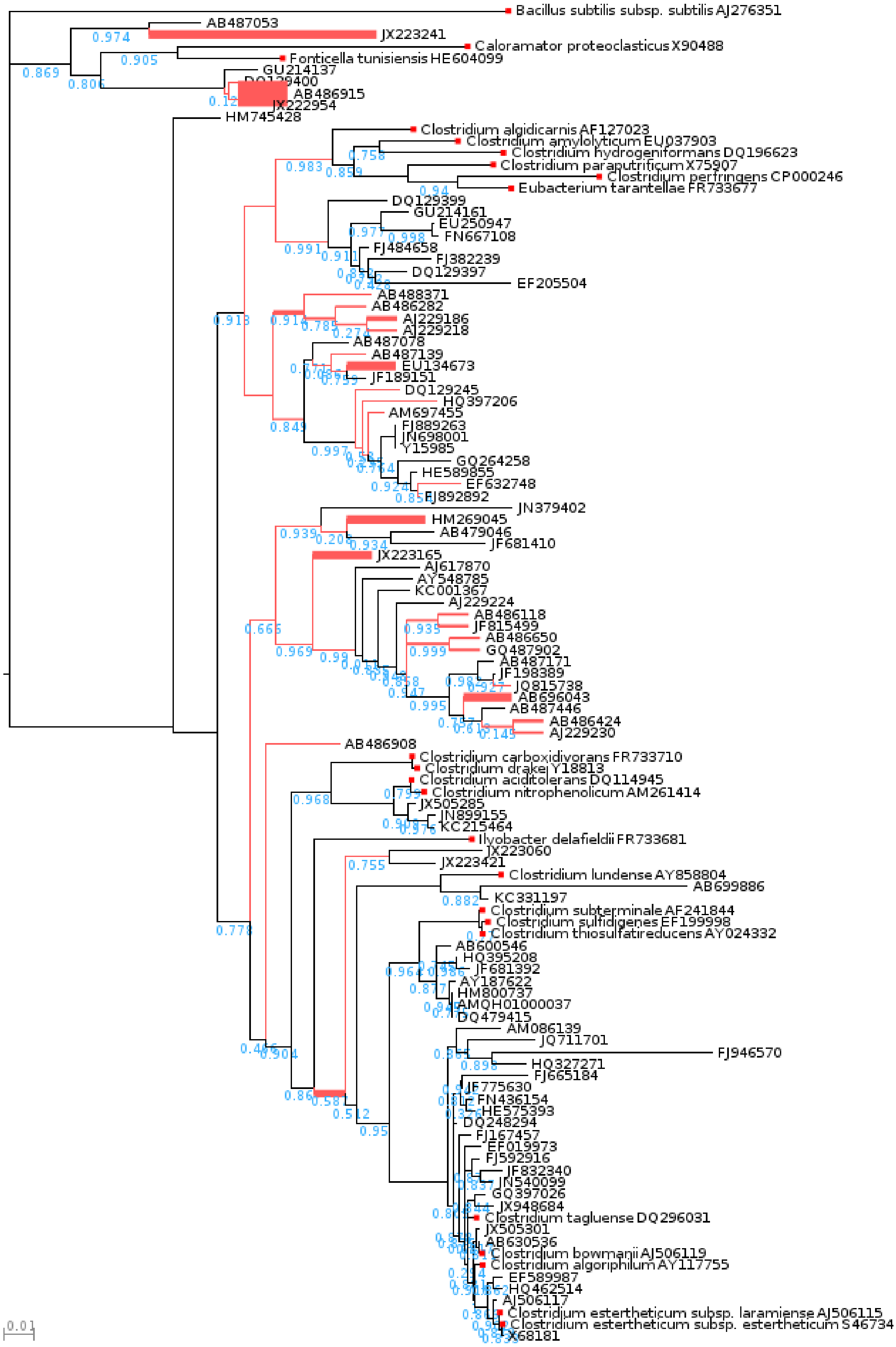


Figure 5. See methods for selection criteria for sequences in backbone tree. Edge width is proportional to number of short putative *Clostridiaceae* diazotroph sequences placed at that position. Placement of short sequences can be spread across multiple edges Matsen et al. (2010). Reference sequences from cultivars have boxes at tips and full species names. Tips with only accession annotations are from environmental reference sequences.



7 SUPPLEMENTAL FIGURES

Figure S1. Ordination of Bray-Curtis sample pairwise distances for each incubation time. Point area is proportional to the density of the CsCl gradient fraction for each sequence library, and color/shape reflects control (red triangles) or labeled (blue circles) treatment. Inset shows Bray-Curtis distances for paired control versus labeled CsCl gradient fractions (i.e. fractions from the same incubation day and same density) against the density of the pair (p-value: $4.526e^{-5}$, r^2: 0.434).

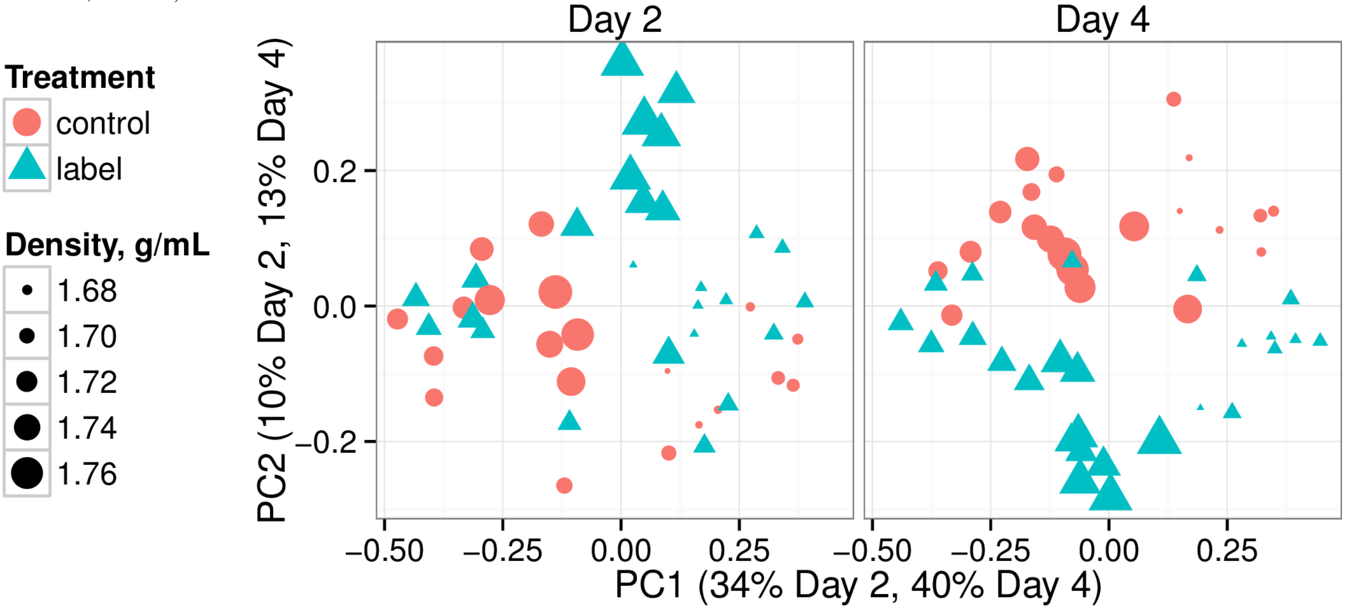
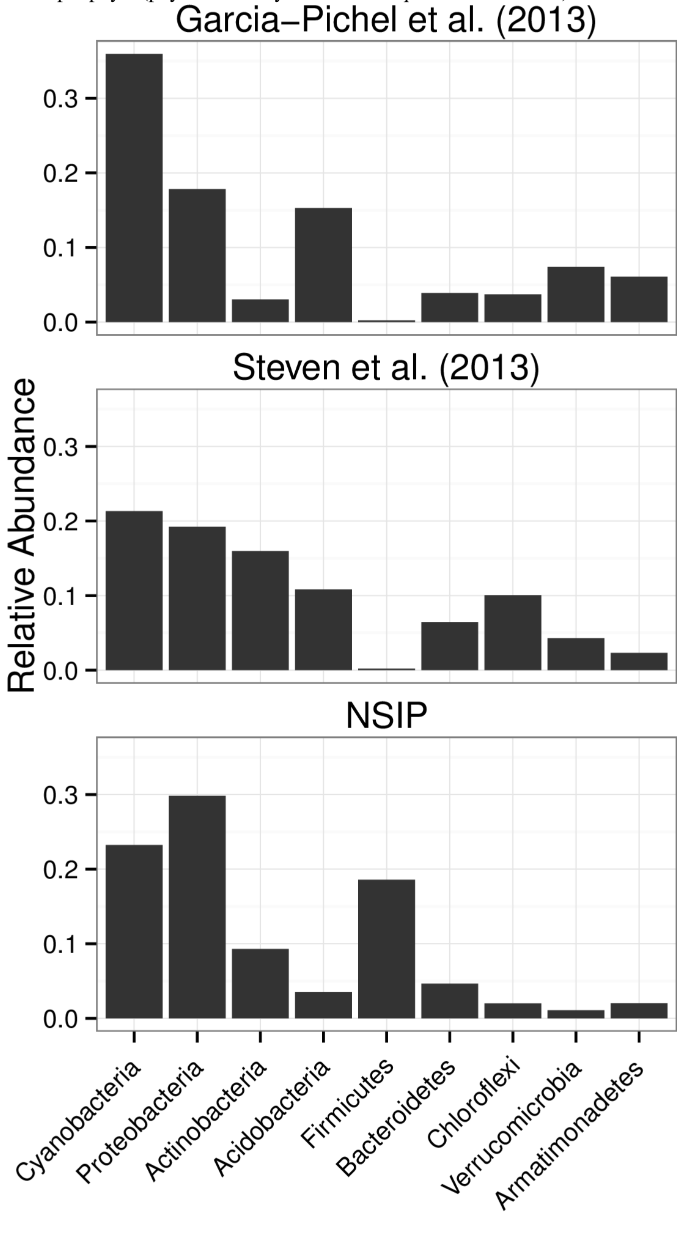


Figure S2. Distribution of sequences into top 9 phyla (phyla ranked by sum of all sequence annotations).



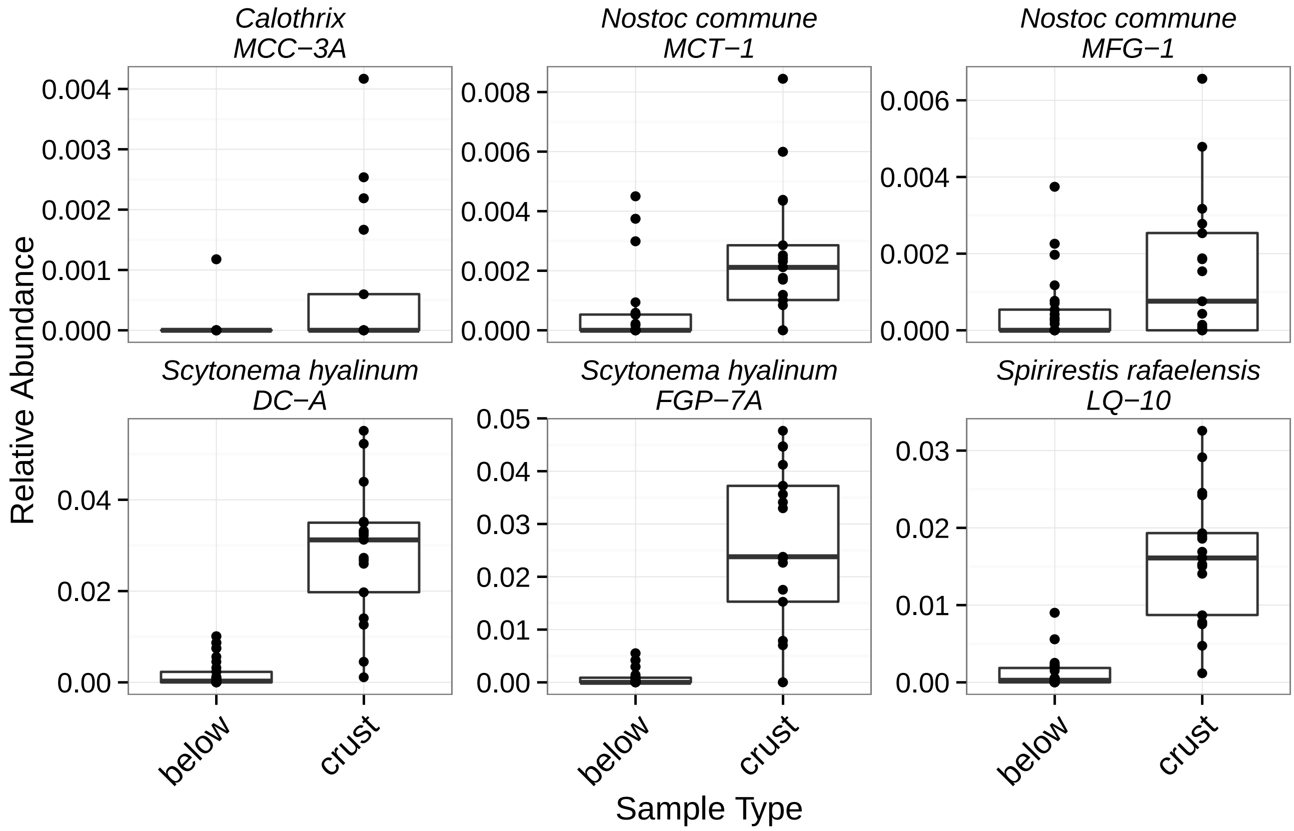


Figure S3. Relative abundance of selected heterocystous cyanobacterial OTUs with centroids from sequences described in Yeager et al. (2006) (see methods for selection criteria) in Steven et al. (2013) data set.

Figure S4. Rarefaction curves for all samples presented by Garcia-Pichel et al. (2013) and Steven et al. (2013). Inset is boxplot of estimated sampling effort for all samples in Garcia-Pichel et al. (2013) and Steven et al. (2013) (number of observed OTUs divided by number of CatchAll Bunge (2010) estimated total OTUs)

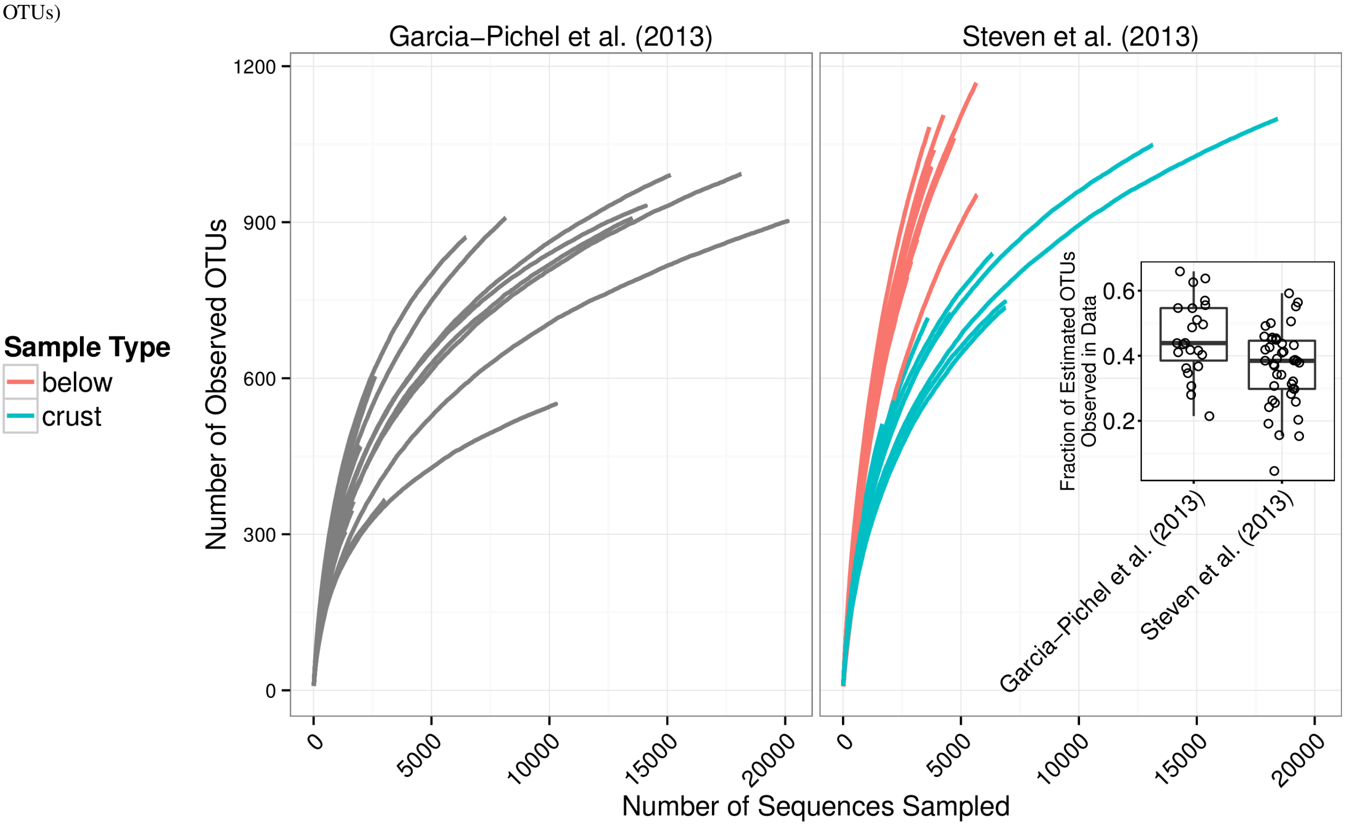


Figure S5. Counts of "responder" OTU occurrences in samples from Steven et al. (2013) and Garcia-Pichel et al. (2013). Steven et al. (2013) collected BSC samples (25 samples total) and samples from soil beneath BSC (17 samples total, "below" column in figure). Garcia-Pichel et al. (2013) collected samples from "dark" (9 samples total) and "light" (12 samples total) crusts in addition to "lichen" (2 samples total) dominated crusts.