

# Niche partitioning of bacterial communities in biological crusts and soils under grasses, shrubs and trees in the Kalahari

David R Elliott<sup>1\*</sup>, Andrew D Thomas<sup>2</sup>, Stephen R Hoon<sup>1</sup>, Robin Sen<sup>1</sup>

1. School of Science & the Environment, Manchester Metropolitan University, Manchester, M1 5GD, U.K.
2. Department of Geography and Earth Science, Aberystwyth University, Llandinam Building, Penglais, Aberystwyth, SY23 3DB, U.K.

\* Corresponding author [d.elliott@mmu.ac.uk](mailto:d.elliott@mmu.ac.uk)

Keywords: biological soil crust, 454 pyrosequencing, bacterial community, Kalahari Sand, carbon, vegetation

## Abstract

The Kalahari of southern Africa is characterised by sparse vegetation interspersed with microbe-dominated biological soil crusts (BSC) which deliver a range of ecosystem services including soil stabilisation and carbon fixation. We characterised the bacterial communities of BSCs (0-1 cm depth) and the subsurface soil (1-2 cm depth) in an area typical of lightly grazed Kalahari rangelands, composed of grasses, shrubs, and trees. Our data add substantially to the limited amount of existing knowledge concerning BSC microbial community structure, by providing the first bacterial community analyses of both BSCs and subsurface soils of the Kalahari region based on a high throughput 16S ribosomal RNA gene sequencing approach. BSC bacterial communities were distinct with respect to vegetation type and soil depth, and varied in relation to soil carbon, nitrogen, and surface temperature. *Cyanobacteria* were predominant in the grass interspaces at the soil surface (0-1 cm) but rare in subsurface soils (1-2 cm depth) and under the shrubs and trees. *Bacteroidetes* were significantly more abundant in surface soils of all areas even in the absence of a consolidated crust, whilst subsurface soils yielded more sequences affiliated to *Acidobacteria*, *Actinobacteria*,

*Chloroflexi*, and *Firmicutes*. The common detection of vertical stratification, even in disturbed sites, suggests a strong potential for BSC recovery after physical disruption, however severe depletion of *Cyanobacteria* near trees and shrubs may limit the potential for natural BSC regeneration in heavily shrub-encroached areas.

## Introduction

Soils are vital for agricultural productivity, biodiversity and carbon storage (Stringer 2008), properties that are dependent upon the presence and activities of soil microbial communities (Brussard 2012). Most studies on soil microbes have focused on productive agricultural systems and typically target bulk soil or the plant-associated rhizosphere (e.g. Rousk et al. 2010; Pereira e Silva et al. 2012; Fierer et al. 2012a; Phosri et al. 2012). Because drylands typically support a patchy vegetation cover, light reaches the soil surface in plant interspaces and facilitates the formation of a complex autotrophic and heterotrophic microbial community, which binds soil into a crust. These biological soil crusts (BSC) are a major component of dryland biodiversity (Büdel et al. 2009) and of global carbon and nitrogen cycles (Elbert et al. 2012). However, we have very little information on the microbial content of dryland soils and BSCs, and even less on how the soil microbes affect dryland ecosystem function. A key factor is the lack of a detailed characterisation of the microbial composition of BSCs, with the exception of a few important groups such as the *Cyanobacteria* (e.g. Dojani et al. 2013). Addressing this research gap is essential if future management and conservation of drylands is to be effective in the face of increasing climatic and anthropogenic pressures (Stringer et al. 2012). This is a major area of concern because drylands cover approximately 40 % of the global land surface and support 38 % of the human population (Reynolds et al. 2007).

Many of the ecosystem functions of BSCs are attributed to *Cyanobacteria*, which comprise a large fraction of BSC biomass (Gundlapally and Garcia-Pichel 2006). *Cyanobacteria* can sequester carbon through photosynthesis and fix nitrogen (Elbert et al. 2012), enabling BSCs to perform similar ecosystem functions to plants (Bowker et al. 2010). BSCs also facilitate numerous ecosystem services of importance to land management, conservation and productivity (Thomas 2012). These include soil

stabilisation (Thomas and Dougill 2007), nitrogen fixation (Aranibar et al. 2003), moisture retention (Menon et al. 2011), and modulation of surface runoff (Eldridge and Greene 1994; Belnap 2006).

Despite the recognition of a significant non-phototrophic component of BSC communities (Garcia-Pichel et al. 2003, Bowker et al. 2010) there have been relatively few studies characterising functionally important bacterial heterotrophs in BSCs. Most studies of dryland BSC communities have used a combination of cultivation and molecular 16S rRNA gene fingerprinting techniques followed by identification of isolates and molecular types by sequencing. These methods have revealed much about BSC community structure, including relationship to ecosystem functioning (Castillo-Monroy et al. 2011), demonstration of vertical stratification (Garcia-Pichel et al. 2003), similarity of diversity regardless of nearby plants (Nagy et al. 2005), and identification of numerically dominant phyla (e.g. Gundlapally and Garcia-Pichel 2006). However, the methodology employed in these studies severely limits microbial identification to no more than a few hundred species or operational taxonomic units (OTUs). This is insufficient to make rigorous community comparisons, particularly in a well-replicated study where more samples must be identified, thus yielding a lower number of sequences per sample. Current high-throughput sequencing technologies such as 454 pyrosequencing overcome this practical limitation and have recently been used by Steven et al. (2013) to reveal differences in BSC and soil bacterial community structure in relation to parent material in Colorado, USA. In a broader study, Fierer et al. (2012b) went a step further by performing a cross-biome metagenomic survey of soil (0-5 cm depth) microbial communities in tandem with a high-throughput phylogenetic survey, enabling a comparison of functional gene frequencies as well as microbial taxa. This study showed that hot and cold desert communities are quite distinct from forest, grassland and tundra on both a taxonomic and functional level.

In this paper we present the first description of bacterial community structure within BSCs and soils of the Kalahari in the south west of Botswana, and assess niche partitioning with respect to depth and nearby vegetation based on high-throughput sequencing of the bacterial 16S rRNA gene. Weakly developed BSCs in the Kalahari Sand soils are found in large areas that are subjected to disturbance by livestock and wildlife activity, which prevents succession into more developed stages (Thomas and

Dougill 2006). Grazing also selectively removes palatable grasses and eventually leads to woody shrub encroachment which renders the land useless for continued grazing (Thomas 2012). Thus, both vegetation and co-occurring BSC cover in Kalahari rangelands are strongly affected by human activity and can be influenced by land management decisions relating to factors such as animal stocking density, fencing, and water access.

Our hypothesis is that landscape impacts related to grazing pressure such as disturbance and shrub encroachment drives functionally significant soil surface bacterial community changes and thus should be included in land management decision-making. Our specific objectives were to determine whether there are significant differences in bacterial populations associated with: i) soils under tree, shrub and grasses, ii) subsurface soils and BSCs.

## Materials and Methods

### Study site

Samples were collected from a long-term research site near Tsabong in south west Botswana (25° 56' 51"S, 22° 25' 40"E) at the end of the dry season in November 2011 and during the wet season in March 2012. Soils are formed on Kalahari Sands and are weakly acidic, fine sand-sized Arenosols (FAO 1990), with little or no horizon development. In lightly grazed areas, around 80 % of the surface is covered in a 3-4 mm deep BSC, which has been described in detail elsewhere (Thomas and Dougill 2007; Thomas 2012). BSC cover is inversely related to grazing intensity and in frequently grazed areas, cover is typically < 10 % of the surface. The organic matter, carbon and nitrogen content of the Kalahari Sand soils is low, reflecting the limited biological productivity and highly oxidising nature of the soils (Thomas et al. 2012). BSCs, however, are enriched in ammonium, total N and organic C compared to the mineral soil (Thomas and Dougill, 2007; Thomas 2012).

Vegetation cover is typical of an open-canopy, fine-leaf savannah, with a mix of perennial (*Eragrostis lehmanniana*) and annual (*Schmidtia kalahariensis*) grasses, woody shrubs (*Grewia flava* and *Acacia mellifera*) and trees, predominantly *Acacia erioloba* (Figure 1). Mean annual precipitation is 331 mm (1996 to 2013), with a low of 114 mm in 2006-07 and a high of 532 mm in 2001-02. Seasonal

variations in air temperature are extreme, with summer maxima frequently in excess of 40°C and winter below 0°C.

### **Sample collection and preparation**

Soils and BSCs were collected from within a fenced 800 m x 500 m paddock where grazing animals had been excluded since the previous year. Sites were selected according to the overlying vegetation type, with three replicate sites under trees (*A. erioloba*), woody shrubs (*G. flava*), perennial (*Eragrostis lehmanniana*) and annual (*Schmidtia kalahariensis*) grasses (Figure 1).

Soils at all sites, except those under trees, were covered in a BSC, previously classified as type 1 (weakly consolidated with no surface discolouration) and 2 (more consolidated with black or brown speckled surface) by Thomas and Dougill (2006; 2007). Soils underneath trees were not crusted, but unconsolidated and slightly darkened by fragmented litter. There was evidence of severe animal disturbance under all tree canopies, a legacy of cattle seeking shade in the year prior to sampling. Soils under *G. flava* canopies were covered in a dry layer of leaf litter but were also well crusted.

Samples were collected using aseptic techniques from 0 – 1 cm (incorporating the BSC) and 1 – 2 cm (the soil immediately below the BSC). The same sites were sampled in November 2011 and March 2012, giving a total of 96 samples (48 in each season). Soils were dry at the time of sampling.

### **Physico-chemical analyses**

Total carbon and nitrogen content of BSC and subsurface soils were determined using a CN element analyser (Leco TruSpec). Soil surface temperature was measured at approximately 2-hour intervals during the day at each sampling site for the duration of each field campaign, using an infrared thermometer.

### **Molecular analyses of bacterial community composition**

DNA was extracted from soil samples within 18 hours of sampling using a Powersoil DNA extraction kit (MoBio Inc.). Prior to extraction, samples of approximately 20 g were homogenised by shaking followed by cutting with a scalpel to disaggregate. Extractions were performed according to the manufacturer's instructions except that the soil mass was increased slightly from 0.25 to 0.4 g based

on laboratory extraction tests and consultation with the manufacturer. DNA was eluted into 50 µl of buffer (10 mM Tris). Phylogenetically informative DNA sequences were obtained from each sample by tag-encoded FLX amplicon pyrosequencing targeting the bacterial 16S rRNA gene (Dowd et al. 2008). This analysis was performed by Research and Testing Laboratory (Lubbock, TX, USA), using a Roche 454 FLX instrument with Titanium reagents.

DNA was amplified for pyrosequencing using forward and reverse fusion primers. The forward primer was constructed with the Roche A linker (CCATCTCATCCCTGCGTGTCTCCGACTCAG), an 8-10bp barcode (see Online Resource 1), and the 341F primer (CCTACGGGAGGCAGCAG) (Muyzer et al. 1993). The reverse fusion primer was constructed with a biotin molecule, the Roche B linker (CCTATCCCCTGTGTGCCTTGGCAGTCTCAG), and the 907R primer (CCGTCAATTCMTTGGAGTTT) (Muyzer et al. 1998). Amplifications were performed in 25 µl reactions with Qiagen HotStar Taq master mix (Qiagen Inc, Valencia, California), 1 µl of each 5 µM primer, and 1 µl of template. Reactions were performed on ABI Veriti thermocyclers (Applied Biosystems, Carlsbad, California) under the following thermal profile: 95°C for 5 min, then 35 cycles of 94°C for 30 sec, 54°C for 40 sec, 72°C for 1 min, followed by one cycle of 72°C for 10 min and 4°C hold. PCR products were visualized with eGels (Life Technologies, Grand Island, New York), pooled equimolar, and size selected before sequencing following manufacturer protocols.

### **Bioinformatics and Statistical Analyses**

Sequence data were processed through the QIIME pipeline (Caporaso et al. 2010) and further analyses were performed using R (R Core Team 2012). Denoising, quality filtering, OTU assignment, OTU table generation, and phylogenetic determinations were all performed in QIIME. Sequences shorter than 200 bp or having an average quality score <25 within a 50 base pair window were discarded. A 97% sequence similarity was used to define OTUs (approximately species level; Stackebrandt and Goebel 1994) which were assigned using UCLUST (Edgar 2010), and chimeras were removed using chimeraSlayer (Haas et al. 2011). After quality control there were on average 1,004 sequences per sample and a total of 2,705 OTUs (further details are provided in Online Resource 1). OTUs were identified through the RDP classifier (Wang et al. 2007) using the Greengenes database release of

October 2012 (DeSantis et al. 2006). Identified OTUs were assembled into an OTU table summarising the frequency of observation in each sample. These results formed the basis for the determination and comparison of community structure.

Thousands of different taxa were detected, so constrained correspondence analysis was used to discern community features which specifically relate to vegetation zone or depth. Rare species comprising <0.01% of the sequences detected in the study were excluded from correspondence analysis because rare species can obscure community patterns and may be differentially detected depending on sample sequencing depth. Correspondence analysis was based on the Bray-Curtis distance measure and performed using the Phyloseq (McMurdie and Holmes 2013) wrapper to the Vegan package (Oksanen et al. 2012) for R. Unconstrained correspondence analysis (CA) was performed to visualise the overall community structure (shown in supplementary data Online Resource 2 only) whilst constrained correspondence analysis (CCA) was employed to discern community features which specifically relate to vegetation zone or depth. Permutation tests ( $n=1000$ ) were used to test the significance of measured environmental variables to the ordination.

Differences in community structure were assessed using ADONIS, a permutational multivariate analysis of variance test in the R package Vegan (Oksanen et al. 2012), to determine whether communities differ with respect to vegetation zone, depth, or sampling month. We used the Bray-Curtis distance measure and performed the test at all taxonomic levels from phylum to species. OTU richness and community diversity were estimated using the Chao1 and Shannon methods respectively, implemented in the Phyloseq package. Richness and diversity calculations used the full data set.

Kruskal-Wallis tests were used to determine whether each OTU relative abundance differed between vegetation zone, depth, or sampling month. The significance of OTU abundance correlations with continuous variables (e.g carbon or richness) was assessed using Spearman's test. P-values were corrected to account for multiple comparisons using the false discovery rate method (Benjamini and Hochberg 1995). Significant findings were tested further using post-hoc tests to identify the changes responsible. We regarded results with corrected  $p<0.05$  as being significant.

### **Availability of sequence data**

Sequence data and metadata are available on the MG-RAST metagenomics analysis server (Meyer et al. 2008) at <http://metagenomics.anl.gov/linkin.cgi?project=6691>.

## **Results**

### **Soil Chemistry**

Soil carbon was significantly higher in the soil surface ( $0.7\% \pm 0.1$  SE) compared to the subsurface soil ( $0.4\% \pm 0.1$  SE), and also differed significantly between vegetation classifications (Figure 2) but not by month. Total soil carbon and nitrogen were closely correlated and the mean C:N ratio was 9:1. The C:N ratio was significantly higher in BSCs, (ANOVA  $F=8.91$ ,  $df=1$ ,  $p=0.0047$ ), but did not vary significantly with respect to vegetation zones.

### **Bacterial diversity**

The Chao1 richness estimate and Shannon diversity index (Figure 3) provide an indication of the total number of species and the microbial diversity (taking account of number of species and evenness) in each niche. Richness and diversity differed significantly with respect to vegetation zone (Kruskal-Wallis chi-squared = 22.91,  $df=3$ ,  $p\text{-value} = 4.22 \times 10^{-5}$ ) but not depth, although diversity was close to our significance threshold for depth (Kruskal-Wallis chi-squared = 2.93,  $df=1$ ,  $p\text{-value} = 0.087$ ). Both diversity (Spearman's rank correlation  $\rho = 0.5$ ,  $p\text{-value} = 3 \times 10^{-4}$ ) and richness ( $\rho = 0.5$ ,  $p\text{-value} = 3 \times 10^{-4}$ ) were positively correlated with carbon content of the soil. Diversity measures for each sample are provided in Online Resource 1.

### **Bacterial community structure**

ADONIS showed that the bacterial communities differed by depth and by vegetation zone ( $p < 0.05$ , see Online Resource 3 for test statistics) at all taxonomic ranks from phylum to species, but did not differ in relation to sampling month except for at the rank of Family. Interactions were detected ( $p < 0.05$ ) between vegetation zone and depth at all taxonomic ranks except for Genus ( $p=0.051$ ), and also between month and vegetation zones at family level and higher taxa. A total of 28 bacterial phyla were detected, and the top 9 shown in Figure 4 account for 99% of sequences. Sequences from

*Actinobacteria* and *Proteobacteria* numerically dominated the samples, together representing 63% of sequences. Most phyla detection frequencies differed significantly ( $p < 0.05$ , see Online Resource 4 for test statistics) with respect to depth or vegetation zone, or both (Table 1). *Cyanobacteria* and *Chloroflexi* were more abundant in grass areas, especially the *Cyanobacteria* which were only rarely detected in tree and shrub areas. *Cyanobacteria* were also very rare in all subsurface soils (1-2 cm depth), only being found in large number in the BSCs of grass interspaces. *Bacteroidetes* and *Cyanobacteria* were significantly associated with BSCs (0-1 cm depth), whilst *Acidobacteria*, *Actinobacteria*, *Chloroflexi*, and *Firmicutes* were significantly associated with subsurface soil (1-2 cm depth). Phylum composition did not vary significantly in relation to the sampling season for any of the top 9 phyla. Data for rarer phyla is also provided in Online Resource 4.

Constrained correspondence analysis was used to generate a visual representation of the microbial community structure differences between depths and vegetation zones. Sequences accounting for  $< 0.01\%$  of the library were excluded from this analysis, leaving 934 OTUs. The microbial communities were separated by vegetation type (tree/shrub or grasses) on axis 1, and by depth on axis 2 (Figure 5). Unconstrained correspondence analysis yielded similar but less well defined patterns (Online Resource 2). Soil carbon (and nitrogen) content and soil temperature were significant to the ordination as determined by permutation tests ( $n=1000$ ). Vectors show that soil carbon and nitrogen increase with axis 1 (direction of tree and shrub samples), whilst soil surface temperature increases in the opposite direction (i.e. in the direction of grass interspace samples).

### **Abundant taxa (OTU level)**

The detection frequencies of the most abundant 9 OTUs are shown in Figure 6. Together these 9 OTUs accounted for 27% of sequences and all of them had different detection frequencies ( $p < 0.05$ , see Online Resource 4 for test statistics) with respect to depth or vegetation zone (Table 2). Similar plots and tables for the top 9 *Bacteroidetes* and *Cyanobacteria* are shown in Figure 7 and Table 3, which account for 10% of sequences. The OTUs of these two phyla were selected because they were detected more often in BSC samples compared to subsurface soil. OTU composition did not vary significantly in relation to the sampling month for any of the top 9 OTUs (Online Resource 4). In

addition to the overall top OTUs shown in Figure 6, we present in Online Resource 4 similar plots and tables for the most abundant OTUs belonging to each phylum. These are included to permit readers interested in particular taxa to easily investigate these in our data, however space and time do not permit detailed presentation of more than a few OTUs here.

The OTUs shown in figures 6 and 7 are additionally plotted on the ordination in Figure 5, illustrating their contributions to the community structure of samples in the ordination. Cyanobacterial OTUs are all clustered near the grass interspace BSC samples (0-1 cm depth) whilst *Bacteroidetes* OTUs are more spread out but tending towards the tree and shrub soil surface samples (0-1 cm depth). The overall most abundant OTUs are spread out on the ordination but with more near the grass area samples.

## Discussion

### **Phylum level community structure**

The distribution of the top 9 bacterial phyla accounting for 99% of sequences (Figure 4, Table 1) indicate that *Bacteroidetes* and *Cyanobacteria* are significantly associated with BSCs (Figure 7, Table 3). The crucial role of *Cyanobacteria* in BSC carbon and nitrogen cycling is already widely recognised, whereas *Bacteroidetes*, although ubiquitous in soil, are not commonly regarded as key BSC community members in the current literature.

Using similar methods to this study, Steven et al. (2013) aimed to determine BSC microbial community differences with respect to soil type (sandstone, shale, and gypsum). Their samples were numerically dominated by the same top 6 phyla found in this study: *Proteobacteria*, *Cyanobacteria*, *Chloroflexi*, *Bacteroidetes*, *Actinobacteria*, and *Acidobacteria*. They found *Cyanobacteria* and *Proteobacteria* to be associated with BSCs whilst we found *Cyanobacteria* and *Bacteroidetes* to be associated with BSCs. The difference may be because *Bacteroidetes* are particularly dominant in the soil surface under trees and shrubs (e.g. Figure 5) which were not a factor in the work of Steven et al. (2013). In the underlying soil they found enrichment of *Chloroflexi*, which we also found in addition to *Acidobacteria*, *Actinobacteria*, and *Firmicutes*. Steven et al. (2013) suggested that the *Chloroflexi*

might be involved in anaerobic processes including photoheterotrophy and chemoheterotrophy, which would increase the productivity of BSCs and enable them to continue functioning under a wider range of environmental conditions.

*Cyanobacteria* comprised 8.1% of our sequences as 38 different OTUs. They were found predominantly in BSCs of the grass areas and are dominated by a single *Phormidium* species (OTU 1912). Other typical cyanobacterial genera found in BSCs include *Microcoleus*, *Leptolyngbya*, *Nostoc*, and *Scytonema* species (Büdel et al. 2009), but these were not identified in our samples. The majority of cyanobacterial sequences obtained could not be classified to genus level however and it is known that molecular identification of *Cyanobacteria* is problematic (Dojani et al. 2013), so it is possible that these genera were present but not detected or identified. A higher frequency of *Cyanobacteria* was detected in the perennial grass compared to the annual grass, probably reflecting the more stabilised interspaces of perennial grass being slightly more developed. Due to the dominance of *Cyanobacteria* in grass interspace BSCs, it is likely that some of the phyla or OTUs found to be more abundant in subsurface soil are not specifically adapted for the soil niche, but may be excluded from the surface by competition.

Like the *Cyanobacteria*, *Bacteroidetes* were also significantly more abundant in the soil surface compared to the subsurface soil, however in this case the relationship held also for soils under tree and shrub canopies. Furthermore, whereas the cyanobacterial abundance was mostly achieved through a single OTU being very abundant in grass area BSCs, the *Bacteroidetes* phylum abundance is due to the collective abundance of many OTUs across all of the vegetation zones (Figure 7). Overall the *Bacteroidetes* were slightly less abundant than the *Cyanobacteria*, comprising 5.1% of sequences, but there were many more OTUs (297).

*Bacteroidetes* are ubiquitous in the soil and have a vast catabolic repertoire, particularly in the breakdown of complex carbohydrates. Thomas et al. (2011) point out that many carbohydrates are niche specific, and this seems a very likely scenario for BSCs because their constituent organisms are subject to unique survival challenges which will select for production of specialist biomolecules.

*Bacteroidetes* may therefore play a vital role in the recycling of niche specific carbohydrates and associated molecules. It is clear that the soil surface in grass and tree/shrub areas will be fundamentally different in terms of carbohydrate composition because only the grass interspace areas receive significant primary production from *Cyanobacteria* and the stress of midday sun, therefore different *Bacteroidetes* OTUs may be providing specialist degradation roles in these distinct niches. A key role for BSC *Bacteroidetes* in carbon cycling is further supported by the work of Bailey et al. (2013), which demonstrated a link between *Chitinophagaceae* family *Bacteroidetes* abundance with  $\beta$ -glucosidase activity in soil. Eight of the top 9 *Bacteroidetes* in our samples belonged to the *Chitinophagaceae* family.

In an attempt to ecologically classify soil bacteria, Fierer et al. (2007) have shown in a meta-analysis that phylum abundances in soil are correlated with carbon availability. This enabled them to broadly classify *Bacteroidetes* and  $\beta$ -*Proteobacteria* as copiotrophs, and *Acidobacteria* as oligotrophs based on whether phylum abundance is positively or negatively correlated with carbon availability respectively. Previously Smit et al. (2001) had made a similar suggestion: that the ratio of  $\alpha$ - and  $\gamma$ -*Proteobacteria* to *Acidobacteria* could provide an indication as to the nutritional status of the soil. Copiotrophic and oligotrophic groupings are similar to the r- and K-strategists recognised in macroscopic ecology (MacArthur and Wilson 1967). Copiotrophs therefore can be expected to maximise their growth rates when resources are plentiful whereas oligotrophs are adapted for maximal efficiency in the use of rare resources.

Based on correlation with total soil carbon (Figure 4b, Online Resource 4) our data identify *Bacteroidetes*, *Gemmatimonadetes* and all proteobacterial classes ( $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\gamma$ ) as potential copiotrophs, and *Acidobacteria* and candidate division WPS-2 as potential oligotrophs in Kalahari soils.

*Chloroflexi* abundances were also negatively correlated with carbon, however due to their carbon fixing abilities soil carbon data are not suitable to attempt classification. In general, oligotrophs may be expected to function as primary colonisers in the development of BSCs on nutrient-poor soils such as Kalahari Sand, and copiotrophs may become more dominant as the crust develops and becomes more productive. The mean carbon content in our samples was  $0.5\% \pm 0.1$  SE with a maximum of

2.2%, which is low compared to typical mesic soils (Lal 2004) so an abundance of oligotrophs might be expected, especially in the grass interspaces where there was least carbon. However, the oligotroph/copiotroph classification is a continuum and may relate not only to soil total carbon, but more specifically to available carbon and carbon turnover rate. From a microbial growth perspective the available carbon fraction is determined by the physiological capabilities of the community and this also defines the amount of soil carbon which can be regarded as recalcitrant (Schmidt et al. 2011). High carbon turnover rate has previously been calculated for cryptogamic crusts (Elbert et al. 2012). This could help copiotrophs to maintain high catabolic rates even when soil carbon is low, so long as there is primary production or other nutrient input available.

### **Bacterial community diversity**

Previous studies have found that dryland BSC and soil bacterial richness and diversity do not vary with respect to depth or presence of plants, although the community composition does vary (Garcia-Pichel et al. 2003; Saul-Tcherkas and Steinberger 2011; Steven et al. 2013). Furthermore Nagy et al. (2005) found no difference in community composition between plant canopies and interspaces, interpreting this to indicate probable independence of BSC communities for vascular plant resources. Conversely, our results show that samples from different depths and near different plants have distinct microbial communities (Figure 4) which also differ in richness and diversity (Figure 3). This is not necessarily a contradiction, however, due to differences in methodology, edaphic factors and disturbance regime in the previous and current studies. In our samples it seems most likely that the tree and shrub bacterial communities are at least partly, if not significantly, dependent on vascular plant and animal-derived resources because they lack phototrophs. Animal disturbance could be a key factor in differentiating soil microbial communities because severe animal disturbance would lead to burial of phototrophs, selectively disadvantaging these organisms but not significantly disadvantaging heterotrophs which can function at any depth. In the case of our grass interspaces, animal disturbance is less concentrated and plant and animal derived inputs are expected to be less, suggesting that the phototrophic component of BSCs will be more resilient compared to those near trees and shrubs.

### **Niche partitioning**

Our results clearly demonstrate niche partitioning of the microbial community between BSCs and the subsurface soil, and between soils under different vegetation types. This is evident from the phylum level breakdown (Figure 4), the most abundant OTUs overall (Figure 6), the OTUs within each phylum (Figure 7, Online Resource 4), and the overall community structure (Figure 5). Although in principle different microbial communities can be functionally identical, we expect that the differences observed between sites and depths are at least partly driven by different environmental conditions requiring a different functional response.

From our field observations and soil chemistry results (Figure 2), we would expect the grass area BSCs to be functionally similar on the macroscopic scale because the conditions are similar. One difference is that the BSC patches in perennial grass interspaces are likely to be older than the patches in annual grass interspaces, because the interspaces have become stabilised near perennial grasses.

A remarkable result was the detection of a clear difference between the soil bacterial community in the surface and below the surface under trees. At the time of sampling, it appeared that the homogenisation of the soil resulting from animal disturbances, had removed all possibility of retaining any depth dependent structure in the biological community. In fact it can be seen clearly in Figure 5 and other figures that the surface (0-1 cm depth) and subsurface (1-2 cm depth) communities from under the trees were quite different, and also that they cluster closely with the same respective communities under the shrub which were not subject to disturbance by cattle. The separation of BSC and subsurface soil communities for both tree/shrub and grass locations on axis 2 of the ordination in Figure 5 suggests that there are community structure patterns characterising surface and sub-surface microbial communities, which apply regardless of the presence of BSC or nearby vegetation.

The clear identification of niche partitioning leads to questions about the functional significance of different communities, and the ecosystem services delivered by BSCs compared to the subsurface soil. The experimental design of this study means that we cannot thoroughly address these questions but it does highlight the need for future targeted studies to do so. We can use the taxonomic

identification of sequences to infer possible functions based on existing knowledge about microbial function in the environment, but proof of function needs to be addressed separately.

Although we have taken microbial community niche partitioning as an indicator of functional differences between soil depths and vegetation zones, this does not imply that BSC community structure is shaped only by deterministic factors. There is a growing recognition of the applicability of neutral models in determining microbial community structure. These models confer key roles to the stochastic processes of birth, death and immigration (Sloan et al. 2006) whilst ignoring deterministic factors such as a species adaptation to drought. Caruso et al. (2011) have suggested that both stochastic and deterministic processes interact in the assembly of dryland microbial communities and that the role of niche partitioning at fine scales in dryland ecosystems has been previously underestimated. They suggest it is incorrect to assume that extreme conditions are the main determinant of species distribution, because this would lead to the conclusion that communities in extreme environments should converge towards stable low diversity communities. The most extreme micro-environments in our study were the BSCs in grassed areas, but they should not necessarily be regarded extreme for adapted microbes. These areas did have the lowest diversity in our study, although the community structure varied greatly between individual samples. Whilst this variability in BSC community structure may be explained in the context of classical ecology as resulting from unobserved environmental heterogeneity and community succession, it is also quite plausible that stochastic processes may be playing a major role in community assembly. As BSCs have been recognised as ecosystem engineers (Bowker 2007), the prospect of stochastic community assembly is very important because it may directly drive diversity of BSC function and soil properties, thus diversifying the landscape itself.

### **Possible ecological significance of the most abundant OTUs**

Three of the top 9 OTUs (Figure 6, Table 2) were significantly associated with BSCs (0-1 cm depth) based on frequency of sequence detection. These were a cyanobacterium of the genus *Phormidium* (OTU 1912), a *Methylobacterium* species (OTU 1064) belonging to the phylum *Proteobacteria*, and a *Balneimonas* species (OTU 1150) of the phylum *Proteobacteria*. In addition, sequences of

proteobacterial OTU 2850 and actinobacterial OTU 3323 were found to be significantly more abundant in subsurface soil compared to BSC. It can be seen in Figure 6 that in several cases differential abundance between BSC and subsurface soil appears to be affected by the vegetation classification, particularly grasses vs. tree or shrub (e.g. OTU 1912). This is supported by the ADONIS results (Online Resource 3) which suggest that there is a significant interaction between depth and vegetation zone driving microbial community structure.

*Phormidium* species have previously been described as typical *Cyanobacteria* of early successional stage BSCs by Büdel et al. (2009), consistent with our present observation in weakly developed type 1 and type 2 BSCs (Thomas and Dougill 2006; 2007). They are widespread filamentous *Cyanobacteria* which are found commonly in hot and cold arid soils but are also found in aquatic habitats including ultra-oligotrophic Antarctic seasonal lakes (Keskitalo et al. 2013), demonstrating a great plasticity for the contrasting environmental conditions of relevance to the Kalahari Sand soil surface.

Chen et al. (2012) found *P. tenue* to be potentially useful for the stabilisation of sand dunes by inoculation due to the desiccation tolerance afforded by its extracellular polysaccharide (EPS). The EPS of *P. tenue* has also been shown to promote the germination of seeds and the fitness of seedlings (Xu et al. 2013), which may have been related to numerous mechanisms including water retention, provision of nutrients, and protection from oxidative stress. Furthermore, Boopathi et al. (2013) showed that a *Phormidium* species associated with mangroves produces indole-3-acetic acid (IAA) which is an important plant hormone associated with diverse responses including enhanced germination and root growth.

BSCs have been associated with both enhancement and inhibition of plant growth through a variety of mechanisms and with some controversy (e.g. Beyschlag et al. 2008; Prasse and Bornkamm 2000) . We tentatively suggest that microbial secretion of plant hormones may be an as-yet unrecognised process of relevance to BSC-plant interactions, and that *Phormidium* species could influence plant cover in the Kalahari if they are secreting IAA similarly to the species studied by Boopathi et al. (2013). The recruitment of plants is potentially at odds with the maintenance of a photosynthetic BSC

due to competition, so one might expect the exclusion of plants rather than their promotion to be a more successful ecological strategy for BSC communities. Another effect of IAA on plant roots is to reduce cell wall integrity, causing the release of nutrients such as sugars which can promote microbial populations near the root. Thus it is possible that in addition to being a significant primary producer in BSCs, *Phormidium* species may also be able to derive carbon from plants, reducing reliance on photosynthesis and potentially promoting further plant-microbe interactions. Again, this supposition is dependent upon the production of plant hormones by the *Phormidium* species in the BSC which has not been tested. The dominance of sequences assigned to this *Phormidium* OTU in our samples strongly suggests a numerical dominance of the BSC bacterial community, and leads to the expectation of a large amount of *Phormidium*-derived EPS in the BSCs and soils examined in this study. The phyla *Bacteroidetes* and *Cyanobacteria* were both more abundant in BSC compared to subsurface soil, however unlike the *Bacteroidetes* which were represented by many OTUs, the *Cyanobacteria* were dominated by the single *Phormidium* OTU 1912. This may be a reflection of the different life strategies of the phyla. The *Bacteroidetes* likely being specialist degraders in BSCs as discussed earlier, are strongly dependent on the soil makeup as determined by the life history and current activity in the soil they inhabit, so a large genetic (functional) diversity is called for. On the other hand *Cyanobacteria* being principally photosynthetic are probably less sensitive to soil life history, but their distribution is likely to be controlled more by abiotic factors such as weather and hydrology. Since the Kalahari Sand substratum is so uniform over very large areas this suggests a possibility for one or a few *Phormidium* species to be strongly influencing the Kalahari soil surface on a very large scale far exceeding our study site.

The *Balneimonas* (OTU 1150) abundance was positively correlated to soil carbon, suggesting it could be functioning as a copiotroph. Six of the other abundant OTUs which were not significantly associated with BSCs were negatively correlated with soil carbon, so although potential oligotrophs are not significantly increased in crusts, they do appear to form a large fraction of the BSC and subsurface soil communities.

*Balneomonas* was recognised as a new genus in 2004 (Takeda et al. 2004), however it has since been proposed for reclassification as *Microvirga* by Weon et al. (2010). In any case the original description of the type species is of a thermophilic (40-45 °C optimum) cellulose producing species isolated from a bath fed by a hot spring in Japan and related to *Methylobacterium* which is the genus of our other OTU of interest here. Cellulose production at high temperature was identified as a remarkable property of the type species *B. flocculans*, and noted for causing flocculation of the cells via adhesive cellulose fibrils similar to those thought to be involved in nonspecific binding of other rhizobial bacteria to plant host cells. The production of EPS is recognised as an important factor in the development and survival of BSCs, and is normally attributed to cyanobacteria (Mager and Thomas 2011). This result suggests that BSC stabilisation by EPS may also be facilitated by heterotrophic bacteria, and a role for *Balneomonas* species as BSC pioneers in advance of cyanobacterial establishment seems quite plausible.

Our *Methylobacterium* OTU was significantly associated with BSCs but not with any particular vegetation zone, suggesting a role in the soil surface which is not related to the presence of a particular vegetation type. The type species *M. organophilum* (Patt et al. 1976) and others are claimed to be facultative methylotrophs – capable of growth on methane or other C1 carbon sources, however although facultative methylotrophy has now been proved for some organisms there has been some controversy over this (Thiesen and Murrell 2005). It is thought that 50-90% of methane released in soils is oxidised by methylotrophs before reaching the atmosphere (Nazaries et al. 2013), therefore it is an important ecosystem function with relevance to climate change. None of the known 21 obligate methane oxidising genera (Nazaries et al. 2013) were detected in our samples.

The most widely recognised role for *Methylobacterium* species is as epiphytes which consume methanol emitted from stomata and secrete plant growth promoting hormones (Lidstrom and Chistoserdova 2002). Several *Methylobacterium* isolates have been confirmed as aerobic anoxygenic phototrophs in BSCs (Csotonyi et al. 2010), meaning that their photosynthetic pathway does not produce oxygen, and that they are obligate aerobes (in contrast to most other anoxygenic phototrophs). This is an important finding because it increases the potential light harvesting efficiency

of BSCs. Csotonyi et al. (2010) found that anoxygenic phototrophs represented up to 5.9% of the cultivable BSC bacterial community, and we found *Methylobacterium* OTU 1064 alone represented 3% of sequences in annual grass zone BSCs (Online Resource 4).

### **Methodological limitations**

The determination of bacterial community structure through analysis of ribosomal RNA genes as carried out in this study has some limitations which should be kept in mind when interpreting results and planning future studies. Methodological aspects including DNA extraction efficiency differences, PCR bias, and primer specificity can affect observed sequence frequencies. Furthermore, natural rRNA gene copy number variation will affect results. For instance, it is known that oligotrophs in general carry fewer rRNA gene copies than copiotrophs (Klappenbach et al. 2000), which could lead to under-representation of oligotrophs in our study.

### **Implications for land management**

The management of BSCs through protection, restoration, or engineering has the potential to deliver environmental benefits from local to global scales, and is relevant to numerous land management challenges. Landscape changes associated with moderate grazing in the Kalahari include a shift from perennial grass to annual grass species and the promotion of thorny shrubs (Skarpe 1990, Ward 2009). We characterised the microbial communities from soils near these plants and found that in all cases a surface specific community can be detected, even beneath the *Acacia* canopy where there was no consolidated BSC. This indicates a strong tendency of the soil community to become vertically stratified and suggests that once grazing pressure is removed a rapid recovery of the BSC is likely. Soil communities beneath tree/shrub canopies were distinct from those in grass interspaces as has been shown previously (Saul-Tcherkas and Steinberger 2011), showing that typical grazing induced vegetation change can be associated with a change in soil microbiota, notably a loss of *Cyanobacteria*. The extreme reduction in cyanobacterial inoculum may retard BSC establishment in heavily shrub-encroached areas, for instance after fire or attempts to reclaim the land by removal of shrubs.

The deliberate rehabilitation of BSCs to restore ecosystem function has been discussed by Bowker (2007) and recently confirmed by cyanobacterial inoculation of shifting sand dunes to establish BSCs which facilitated vascular plant succession from early nitrogen fixing legumes to latter successional grasses (Lan et al. 2014). A problem with the inoculation approach is that it usually relies upon a sacrifice zone, and may not work if there are nutrient or stability limitations. These problems can potentially be overcome by the industrial preparation of a designed mixed inoculum. Candidates for a BSC inoculum in the Kalahari rangelands could be identified for testing based upon our data, which is ideal for this purpose because early stage BSCs are likely to still contain in large number the pioneer species which helped them to become established. The *Phormidium* genus of *Cyanobacteria* (*P. tenue* specifically) has been shown by Chen et al. (2012) to possess excellent qualities for the stabilisation of sand, and its abundance in our study suggests that deliberate establishment on degraded land in the Kalahari could be feasible. A suitable inoculum to help the establishment of *Cyanobacteria* in general might include oligotrophic organisms such as *Balneomonas* species which are able to utilise recalcitrant carbon and quickly stabilise the soil by release of EPS, including beneath the photic zone. The idea that the early stages of BSC recovery or development could be helped by oligotrophic non-photosynthetic bacteria has been confirmed by Wu et al. (2010) in laboratory and field experiments, however the identity of organisms involved was not known.

## Conclusions

We found that the bacterial OTU (approximate species level) diversity was greater in subsurface soil (1-2 cm depth) compared to the BSC at the surface (0-1 cm depth), and community composition exhibited clear spatial patterns. Crusted grass interspaces were dominated by a single cyanobacterial OTU from the genus *Phormidium*, but *Cyanobacteria* were very rare in tree and shrub areas. The BSC community structure was not defined by the presence of *Cyanobacteria* alone nor by other phototrophs as is often presumed. In all areas a characteristic but variable BSC bacterial community was present, and this was defined in part by non-cyanobacterial OTUs which were associated with the soil surface, and especially the *Bacteroidetes* phylum. The functions of these BSC specific bacteria are unknown, but we speculated possible roles for the most abundant ones in soil stabilisation,

carbohydrate catabolism, photosynthesis, and plant interactions based on similar species reported in the literature. Our data suggest that shrub encroachment in Kalahari rangelands can almost eliminate *Cyanobacteria* from soil surfaces in some circumstances, depriving soils of the ecosystem services delivered by well-developed phototrophic crusts and limiting the potential for natural BSC regeneration.

## **Acknowledgements**

Financial support was provided by the Leverhulme Trust (F/00 426/H) and Manchester Metropolitan University. Research in Botswana was conducted with the Republic of Botswana Research Permit No. EWT8/36/4 VIII(4). We would like to thank Jill Thomas of Berrybush farm for her considerable support and access to her land. We also thank Peter Harding for assistance with computer resources and configuration. The constructive feedback of an anonymous reviewer and Matthew Bowker is greatly appreciated.

## **References**

- Aranibar JN, Anderson IC, Ringrose S, Macko SA (2003) Importance of nitrogen fixation in soil crusts of southern African arid ecosystems: acetylene reduction and stable isotope studies. *Journal of Arid Environments* 54: 345-358
- Bailey VL, Fansler SJ, Stegen JC, McCue LA (2013) Linking microbial community structure to  $\beta$ -glucosidic function in soil aggregates. *ISME Journal* 7(10):2044-53. doi: 10.1038/ismej.2013.87.
- Belnap J (2006) The potential roles of biological soil crusts in dryland hydrologic cycles. *Hydrological Processes* 20(15): 3159–3178
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* 57: 289–300
- Beyschlag W, Wittland M, Jentsch A, Steinlein T (2008) Soil crusts and disturbance benefit plant germination, establishment and growth on nutrient deficient sand. *Basic and Applied Ecology* 9(3): 243-252.

- Boopathi T, Balamurugan V, Gopinath S, Sundararaman M (2013) Characterization of IAA Production by the Mangrove Cyanobacterium *Phormidium* sp. MI405019 and Its Influence on Tobacco Seed Germination and Organogenesis. *Journal of Plant Growth Regulation* 32(4): 758-766
- Bowker MA (2007) Biological soil crust rehabilitation in theory and practice: an underexploited opportunity. *Restoration Ecology* 15(1): 13-23.
- Bowker MA, Maestre FT, Escolar C (2010) Biological crusts as a model system for examining the biodiversity–ecosystem function relationship in soils. *Soil Biology and Biochemistry* 42(3): 405-417
- Brussard L (2012) Ecosystem Services provided by soil biota. In: Diana H Wall (ed) *Soil ecology and ecosystem services*. Oxford University Press, Oxford pp.45-58.
- Büdel B, Darienko T, Deutschewitz K, Dojani S, Friedl T, Mohr KI, Salisch M, Reisser W, Weber B. (2009) Southern African biological soil crusts are ubiquitous and highly diverse in drylands, being restricted by rainfall frequency. *Microbial Ecology* 57(2):229-47. doi: 10.1007/s00248-008-9449-9.
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7: 335–336.
- Caruso T, Chan Y, Lacap DC, Lau MCY, McKay CP, Pointing SB (2011) Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *ISME Journal* 5: 1406–1413
- Castillo-Monroy AP, Bowker MA, Maestre FT, Rodríguez-Echeverría S, Martínez I, Barraza-Zepeda CE, Escolar C (2011) Relationships between biological soil crusts, bacterial diversity and abundance, and ecosystem functioning: Insights from a semi-arid Mediterranean environment. *Journal of Vegetation Science* 22: 165–174

Chen L, Yang Y, Deng S, Xu Y, Wang G, Liu Y (2012) The response of carbohydrate metabolism to the fluctuation of relative humidity (RH) in the desert soil cyanobacterium *Phormidium tenue*. *European Journal of Soil Biology* 48: 11-16

Csotonyi JT, Swiderski J, Stackebrandt E and Yurkov V (2010) A new environment for aerobic anoxygenic phototrophic bacteria: biological soil crusts. *Environmental microbiology reports* 2(5): 651-656

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL (2006) Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl Environ Microbiol* 72:5069-72

Dojani S, Kauff F, Weber B, Büdel B (2013) Genotypic and Phenotypic Diversity of Cyanobacteria in Biological Soil Crusts of the Succulent Karoo and Nama Karoo of Southern Africa. *Microbial Ecology*. doi: 10.1007/s00248-013-0301-5

Dowd S, Callaway T, Wolcott R, Sun Y, McKeethan T, Hagevoort R, Edrington T (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC microbiology* 8(1): 125

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19): 2460-2461. doi: 10.1093/bioinformatics/btq461

Elbert W, Weber B, Burrows S, Steinkamp J, Büdel B, Andreae MO, Pöschl U (2012) Contribution of cryptogamic covers to the global cycles of carbon and nitrogen. *Nature Geoscience* 5(7): 459-462

Eldridge DJ, Greene RSB (1994) Microbiotic Soil Crusts: A Review of their Roles in Soil and Ecological Processes in the Rangelands of Australia. *Aust. J. Soil Res* 32: 389-415

FAO, 1990. Guidelines for soil description. 3<sup>rd</sup> Ed. (revised). Soil Resources, Management and Conservation Service, Land and Water Development Division. FAO, Rome. 70 pp

- Fierer N, Bradford MA, Jackson RB (2007) Toward an ecological classification of soil bacteria. *Ecology* 88(6): 1354-1364
- Fierer N, Lauber CL, Ramirez, KS, Zaneveld, J, Bradford MA, Knight R (2012a). Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME Journal* 6: 1007-1017
- Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL, Owens S, Gilbert JA, Wall DA, Caporaso JG (2012b) Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl. Acad. Sci. USA.* 109: 21390-21395
- Garcia-Pichel F, Johnson SL, Youngkin D, Belnap J (2003) Small-scale vertical distribution of bacterial biomass and diversity in biological soil crusts from arid lands in the Colorado plateau. *Microbial Ecology* 46(3): 312-21
- Gundlappally SR and Garcia-Pichel F (2006) The community and phylogenetic diversity of biological soil crusts in the Colorado Plateau studied by molecular fingerprinting and intensive cultivation. *Microbial Ecology* 52(2): 345-357
- Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E, Methé B, DeSantis TZ; Human Microbiome Consortium, Petrosino JF, Knight R, Birren BW (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 21(3):494-504. doi: 10.1101/gr.112730.110.
- Keskitalo J, Leppäranta M, Arvola L (2013) First records of primary producers of epiglacial and supraglacial lakes in western Dronning Maud Land, Antarctica. *Polar biology* 36(10): 1441-1450
- Klappenbach JA, Dunbar JM, Schmidt TM (2000) rRNA operon copy number reflects ecological strategies of bacteria. *Applied and Environmental Microbiology* 66(4): 1328-1333

- Lan S, Zhang Q, Wu L, Liu Y, Zhang D, Hu C (2014) Artificially accelerating the reversal of desertification: cyanobacterial inoculation facilitates the succession of vegetation communities. *Environmental science & technology*. doi: 10.1021/es403785j
- Lal R (2004) Carbon Sequestration in Dryland Ecosystems. *Environmental Management* 33 (4): 528 – 544
- Lidstrom ME, Chistoseerdova L (2002) Plants in the pink: Cytokinin production by *Methylobacterium*. *Journal of Bacteriology* 184: 1818
- MacArthur RH, Wilson EO (1967) *The theory of island biogeography*. Princeton NJ: Princeton Univ. Press. 203 pp
- Mager DM, Thomas AD (2011) Extracellular polysaccharides from cyanobacterial soil crusts: A review of their role in dryland soil processes. *Journal of Arid Environments* 75(2): 91-97
- McMurdie PJ, Holmes S (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*. 8(4):e61217
- Menon M, Yuan Q, Jia X, Dougill AJ, Hoon SR, Thomas AD, Williams RA (2011) Assessment of physical and hydrological properties of biological soil crusts using X-ray microtomography and modeling. *Journal of hydrology* 397(1): 47-54
- Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M et al. (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics* 9(1): 386
- Muyzer G, Brinkhoff T, Nübel U, Santegoeds C, Schäfer H, Waver C. (1998). Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. In: Akkermans ADL, van Elsas JD, de Bruijn FJ (eds). *Molecular Microbial Ecology Manual*. Kluwer Academic Publishers: Dordrecht, The Netherlands. pp 1–27

Muyzer G, De Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* 59: 695–700

Nagy ML, Pérez A, Garcia-Pichel F (2005) The prokaryotic diversity of biological soil crusts in the Sonoran Desert (Organ Pipe Cactus National Monument, AZ). *FEMS microbiology ecology* 54(2): 233-245.

Nazaries L, Murrell JC, Millard P, Baggs L, Singh BK (2013) Methane, microbes and models: fundamental understanding of the soil methane cycle for future predictions. *Environ Microbiol.* 15(9):2395-2417. doi: 10.1111/1462-2920.12149

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens HMH, Wagner H (2013) *vegan: Community Ecology Package*. R package version 2.0-7. <http://CRAN.R-project.org/package=vegan>

Patt TE, Cole GC, Hanson RS (1976) *Methylobacterium*, a new genus of facultatively methylotrophic bacteria. *International Journal of Systematic Bacteriology* 26: 226–229.

Phosri C, Polme S, Taylor AFS, Koljalg U, Suwannasai N, Tedersoo L (2012) Diversity and community composition of ectomycorrhizal fungi in a dry deciduous dipterocarp forest in Thailand. *Biodiversity and Conservation* 21: 2287-2298

Pereira e Silva MC, Dias ACF, vanElsas JD, Salles JF (2012) Spatial and Temporal Variation of Archaeal, Bacterial and Fungal Communities in Agricultural Soils. *PLoS One* 7: e51554 doi: 10.1371/journal.pone.0051554

Prasse R, Bornkamm R (2000) Effect of microbiotic soil surface crusts on emergence of vascular plants. *Plant Ecology* 150(1-2): 65-75

R Core Team (2012) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>

Reynolds JF, Smith DMS, Lambin EF, Turner BL, Mortimore M, Batterbury SP et al. (2007) Global desertification: building a science for dryland development. *Science* 316(5826): 847-851

Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, Knight R, Fierer N. (2010) Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME Journal* 4: 1340-1351

Saul-Tcherkas V, Steinberger Y (2011) Soil microbial diversity in the vicinity of a Negev Desert shrub--*Reaumuria negevensis*. *Microb Ecol.* 61(1):64-81. doi: 10.1007/s00248-010-9763-x.

Schmidt MW, Torn MS, Abiven S, Dittmar T, Guggenberger G, Janssens IA, Kleber M, Kögel-Knabner I, Lehmann J, Manning DA, Nannipieri P, Rasse DP, Weiner S, Trumbore SE (2011) Persistence of soil organic matter as an ecosystem property. *Nature*. doi: 10.1038/nature10386.

Skarpe C (1990) Structure of the woody vegetation in disturbed and undisturbed arid savanna, Botswana. *Vegetatio* 87(1): 11-18.

Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology* 8(4): 732–740

Smit E, Leeftang P, Gommans S, van den Broek J, van Mil S, Wernars K (2001) Diversity and seasonal fluctuations of the dominant members of the bacterial soil community in a wheat field as determined by cultivation and molecular methods. *Applied and Environmental Microbiology* 67(5): 2284-2291

Stackebrandt E, Goebel BM (1994) A place for DNA–DNA reassociation and 16S ribosomal-RNA sequence-analysis in the present species definition in bacteriology. *Int J Syst Evol Microbiol* 44: 846–849

Steven B, Gallegos-Graves LV, Belnap J, Kuske CR (2013) Dryland soil microbial communities display spatial biogeographic patterns associated with soil depth and soil parent material. *FEMS Microbiol Ecol* 86(1):101-113. doi: 10.1111/1574-6941.12143.

Stringer L (2008) Can the UN Convention to Combat Desertification guide sustainable use of the world's soils? *Frontiers in Ecology and the Environment* 6(3): 138-144

Stringer LC, Dougill AJ, Thomas AD, Spracklen DV, Chesterman S, Speranza CI, Rueff H et al. (2012) Challenges and opportunities in linking carbon sequestration, livelihoods and ecosystem service provision in drylands. *Environmental Science & Policy* 19: 121-135

Takeda M, Suzuki I, Koizumi JI (2004) *Balneomonas flocculans* gen. nov., sp. nov., A New Cellulose-producing Member of the  $\alpha$ -2 subclass of *Proteobacteria*. *Systematic and applied microbiology* 27(2): 139-145

Theisen AR, Murrell JC (2005) Facultative methanotrophs revisited. *J Bacteriol.* 187(13):4303-5

Thomas AD (2012) Impact of grazing intensity on seasonal variations of soil organic carbon and soil CO<sub>2</sub> efflux in two semi-arid grasslands in southern Botswana. *Philosophical Transactions of the Royal Society B* 367: 3076-3086

Thomas AD, Dougill AJ (2006) Distribution and characteristics of cyanobacterial soil crusts in the Molopo Basin, South Africa. *Journal of Arid Environments* 64: 270–283

Thomas AD, Dougill AJ (2007) Spatial and temporal distribution of cyanobacterial soil crusts in the Kalahari: implications for soil surface properties. *Geomorphology* 85: 17-29

Thomas AD, Hoon SR, Dougill AJ, Mairs H (2012) Soil Organic Carbon in Deserts: Examples from the Kalahari. In, Mol, L., and Sternberg, T., Eds. *Changing Deserts: Integrating Environments, People and Challenges*. Pp. 40-60. White Horse Press, Cambridge.

Thomas F, Hehemann J, Rebuffet E, Czjzek M, Michel G (2011) Environmental and gut Bacteroidetes: the food connection. *Frontiers in Microbiology* 2. doi: 10.3049/fmicb.2011.00093

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73: 5261–5267

Ward D (2009) *The Biology of Deserts*. Oxford University Press, Oxford

Weon HY, Kwon SW, Son JA, Jo EH, Kim SJ, Kim YS, Kim BY, Ka JO (2010) Description of *Microvirga aerophila* sp. nov. and *Microvirga aerilata* sp. nov., isolated from air, reclassification of *Balneimonas flocculans* Takeda et al. 2004 as *Microvirga flocculans* comb. nov. and emended description of the genus *Microvirga*. *International journal of systematic and evolutionary microbiology* 60(11): 2596-2600

Wu N, Zhang YM, Pan HX, Zhang J (2010) The Role of Nonphotosynthetic Microbes in the Recovery of Biological Soil Crusts in the Gurbantunggut Desert, Northwestern China. *Arid Land Research and Management* 24(1): 42-56

Xu Y, Rossi F, Colica G, Deng S, De Philippis R, Chen L (2013) Use of cyanobacterial polysaccharides to promote shrub performances in desert soils: a potential approach for the restoration of desertified areas. *Biology and fertility of soils* 49(2): 143-152.

## Figure legends

**Fig. 1.** Photographs showing examples of the four sampling zones. Letters in square brackets indicate abbreviations used throughout the manuscript. (a) Annual grass [AG] *Schmidtia kalahariensis* interspace. (b) Perennial grass [PG] *Eragrostis lehmanniana* interspace. (c) Shrub [S] *Grewia flava*. (d) Tree [T] *Acacia erioloba*.

**Fig. 2.** Total carbon and nitrogen in BSC (0-1 cm depth) and soil (1-2 cm depth) samples at each site (n=6). Boxes represent the interquartile range (IQR), and error bars extend to the most extreme values within 1.5 \* IQR of the box. Median values are shown as a line within the box and outliers are shown as black spots. Sample coding: AG=annual grass, PG=perennial grass, S=shrub, T=tree. Note that y axes differ.

**Fig. 3.** OTU richness estimation (Chao1) and diversity index (Shannon). Results are plotted with respect to (a) sampling site; and (b) sample carbon content. Boxes represent the interquartile range (IQR), and error bars extend to the most extreme values within 1.5 \* IQR of the box (n=5 or 6). Median values are shown as a line within the box and outliers are shown as black spots. Error bars in (b) indicate the standard error of the individual Chao1 estimations. Sample coding: AG=annual grass, PG=perennial grass, S=shrub, T=tree. Source data were rarefied to an even sampling depth of 437 sequences per sample, causing loss of 2 samples and leaving 46 in the analysis.

**Fig. 4.** Phylum abundance by (a) site and (b) soil carbon content. Boxes represent the interquartile range (IQR), and error bars extend to the most extreme values within 1.5 \* IQR of the box (n=6). Median values are shown as a line within the box and outliers are shown as black spots. Sample coding: AG=annual grass, PG=perennial grass, S=shrub, T=tree. Significance and direction of correlation between phylum abundance and soil carbon is indicated by + or – (determined by Spearman test). Significance codes for positive correlation: +++ < 0.001; ++ < 0.01; + < 0.05. Similar plots for less abundant phyla are included in Online Resource 4.

**Fig. 5. Constrained** correspondence analysis of the microbial community. Coloured markers indicate individual samples, and dispersion ellipses show the 95 % standard deviation confidence interval for

different depths and tree/shrub vs grass interspace classifications. OTU identification numbers are shown in different colours for the 9 most abundant OTUs belonging to the following groups: full dataset (black), phylum *Cyanobacteria* (green), and phylum *Bacteroidetes* (blue). Environmental variables with significance  $p < 0.05$ , are shown as biplotted vectors (based on permutation tests;  $n=1000$ ).

**Fig. 6.** Relative abundance of the top 9 OTUs detected in the study, by (a) site and (b) soil carbon content. Boxes represent the interquartile range (IQR), and error bars extend to the most extreme values within  $1.5 * IQR$  of the box ( $n=6$ ). Median values are shown as a line within the box and outliers are shown as black spots. Significance and direction of correlation between OTU abundance and soil carbon is indicated by + or - (determined by Spearman test). Significance codes for positive correlation: +++  $< 0.001$ ; ++  $< 0.01$ ; +  $< 0.05$ . Sample coding: AG=annual grass, PG=perennial grass, S=shrub, T=tree.

**Fig. 7.** Relative abundance of the top 9 OTUs from the phyla (a) *Bacteroidetes*; (b) *Cyanobacteria*. Boxes represent the interquartile range (IQR), and error bars extend to the most extreme values within  $1.5 * IQR$  of the box ( $n=6$ ). Median values are shown as a line within the box and outliers are shown as black spots. Sample coding: AG=annual grass, PG=perennial grass, S=shrub, T=tree. Similar plots for other phyla are included in in Online Resource 4.

## Tables

**Table 1.** Significance of factors depth and vegetation on phylum abundance, as determined by Kruskal Wallis test. Significant effects were further tested by post-hoc analyses which are provided in Online Resource 4. Significance codes: \*\*\*<0.001; \*\*<0.01; \*<0.05; .<0.1.

<b>Phylum</b>	<b>Depth</b>	<b>Veg.</b>
Chloroflexi	*	***
Proteobacteria		***
Bacteroidetes	**	***
Acidobacteria	**	**
Firmicutes	**	.
Actinobacteria	**	
Gemmatimonadetes		***
Cyanobacteria	**	**
WPS-2		***

**Table 2.** Taxonomic classification of the most abundant 9 OTUs found in the study. Significance of depth and vegetation on OTU abundance is indicated by \*\*\*<0.001; \*\*<0.01; \*<0.05; .<0.1., as determined by Kruskal Wallis test. Significant effects were further tested by post-hoc analyses which are provided in Online Resource 4.

	Phylum	Class	Order	Family	Genus	Depth	Veg.
654	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae			***
761	Acidobacteria	Solibacteres	Solibacterales	Solibacteraceae	Candidatus Solibacter	.	***
1064	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Methylobacterium	***	
1150	Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Balneimonas	*	***
1912	Cyanobacteria	Oscillatoriophyceae	Oscillatoriales	Phormidiaceae	Phormidium	***	*
2850	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae		***	**
3067	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae			*
3302	Actinobacteria	Actinobacteria	Actinomycetales	Kineosporiaceae			***
3323	Actinobacteria	Actinobacteria	Actinomycetales	Pseudonocardiaceae		***	

**Table 3.** Taxonomic classification of the most abundant 9 *Bacteroidetes* and *Cyanobacteria* OTUs found in the study. Significance of depth and vegetation on OTU abundance is indicated by \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ,  $. < 0.1$ ., as determined by Kruskal Wallis test. Significant effects were further tested by post-hoc analyses which are provided in Online Resource 4.

	Phylum	Class	Order	Family	Genus	Depth	Veg.
177	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Chitinophagaceae			**
1461	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Chitinophagaceae	Flavisolibacter	.	
1508	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Chitinophagaceae	Flavisolibacter		
2853	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Chitinophagaceae			
3081	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Flexibacteraceae		**	*
3166	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Chitinophagaceae	Segetibacter		
3388	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Chitinophagaceae	Segetibacter		
3476	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Chitinophagaceae			
3635	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Chitinophagaceae			
686	Cyanobacteria	Oscillatoriothycideae				*	*
1110	Cyanobacteria	Oscillatoriothycideae				*	
1198	Cyanobacteria	Oscillatoriothycideae					
1499	Cyanobacteria	4C0d-2	MLE1-12				
1626	Cyanobacteria	Oscillatoriothycideae					
1912	Cyanobacteria	Oscillatoriothycideae	Oscillatoriales	Phormidiaceae	Phormidium	**	*
1977	Cyanobacteria	Oscillatoriothycideae					.
3313	Cyanobacteria	Oscillatoriothycideae				.	*
3453	Cyanobacteria	Oscillatoriothycideae	Oscillatoriales	Phormidiaceae			

## Online resources

**OR1.** Sample metadata and summary statistics, including pyrosequencing barcodes, number of sequences obtained, richness estimates, and diversity measures. This is a tab separated text file based on the sample map used in the QIIME pipeline and other downstream analyses.

**OR2:** Constrained and unconstrained correspondence analyses of the microbial community including scree plots.

**OR3.** Text file containing results from the ADONIS test comparing community structure with respect to vegetation, depth, and month. The test was performed at all taxonomic ranks from phylum to species.

**OR4.** Zip file containing plots, statistical tests, and result tables of OTU abundances for the top 9 OTUs in each phylum, and the proteobacterial classes. Files are descriptively named and include for each group of OTUs a zone plot and a carbon plot similar to those shown in Figure 4a and 4b respectively. Boxes represent the interquartile range (IQR), and error bars extend to the most extreme values within  $1.5 * IQR$  of the box ( $n=6$ ). Median values are shown as a line within the box and outliers are shown as black spots. Sample coding: AG=annual grass, PG=perennial grass, S=shrub, T=tree. Significance and direction of correlation between OTU abundance and soil carbon is indicated by + or - (determined by Spearman test). Significance codes for positive correlation: +++ < 0.001; ++ < 0.01; + < 0.05.