



HAL
open science

The bacterial communities of surface soils from desert sites in the eastern Utah (USA) portion of the Colorado Plateau

Jorge R. Osman, Yang Wang, Chloé Jaubert, Tuyet-Nga Nguyen, Gustavo R. Fernandes, Michael S. Dubow

► To cite this version:

Jorge R. Osman, Yang Wang, Chloé Jaubert, Tuyet-Nga Nguyen, Gustavo R. Fernandes, et al.. The bacterial communities of surface soils from desert sites in the eastern Utah (USA) portion of the Colorado Plateau. *Microbiological Research*, 2021, 244, pp.126664 -. 10.1016/j.micres.2020.126664 . hal-03493419

HAL Id: hal-03493419

<https://hal.science/hal-03493419>

Submitted on 2 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **The bacterial communities of surface soils from desert sites in the eastern Utah**
2 **(USA) portion of the Colorado Plateau**

3 Jorge R. Osman^{1#}, Yang Wang[#], Chloé Jaubert, Tuyet-Nga Nguyen, Gustavo R.
4 Fernandes and Michael S. DuBow^{1*}

5 Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Saclay,
6 Univ Paris-Sud, Bâtiment 409, 91405 Orsay, France

7 ¹ Current address: Laboratoire de Géologie de Lyon, Terres, Planètes, Environnement
8 (LGL-TPE) Université Claude Bernard Lyon-1, UMR5276, 69622 Villeurbanne, France.

9 [#] These authors contributed equally to this work.

10 **Corresponding author:* Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS,
11 Université Paris-Saclay, Univ Paris-Sud, Bâtiment 409, 91405 Orsay, France; Tel:
12 +33(0)169826149; e-mail: michael.dubow@universite-paris-saclay.fr

13 **Running Head:** Soil bacterial communities from the Colorado Plateau (Utah, USA)
14 desert

15 Declarations of conflicts of interest: none

16 **Abstract**

17 Desert-like areas located in the eastern portion of the state of Utah (USA) have
18 geographic features that can resemble the surface of the planet Mars, characterized
19 by red-colored hills, soils and sandstones. We examined the bacterial diversity of
20 surface soil samples from several sites from the Colorado Plateau Desert in eastern
21 Utah using pyrosequencing of PCR amplified bacterial 16S rRNA genes from total
22 extracted soil DNA. The sample sites cover the Great Basin, Goblin Valley State Park
23 and nearby regions on the Colorado Plateau. We also examined several
24 physicochemical parameters of the soil samples to investigate any possible
25 correlations between bacterial community structure and environmental drivers. The
26 predominant bacterial phyla present in the samples were found to belong to
27 members of the *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, and
28 *Gemmatimonadetes*. The most abundant genera in our samples were found to
29 belong to the *Cesiribacter*, *Lysobacter*, *Adhaeribacter*, *Microvirga* and *Pontibacter*
30 genera. We found that the relative abundance of *Proteobacteria* and
31 *Gemmatimonadetes* were significantly correlated with soil pH and a low
32 concentration of organic matter, suggesting that, in these relatively high-altitude
33 desert soils, these two parameters may be of primary importance to influence
34 bacterial community composition.

35 Key Words: High Altitude Desert, Bacteria, Biodiversity, 16S rDNA, Pyrosequencing

36 **Introduction**

37 Deserts are defined as regions of land that have ≤ 25 cm of annual precipitation, and
38 where a concomitant reduced biodiversity of plants and animals, compared to that
39 found in more temperate zones, is generally observed. Arid lands represent the
40 largest terrestrial biome (Pointing and Belnap 2012). In these environments, stresses
41 such as drought, large daily temperature fluctuations and relatively high UV radiation
42 exposure are believed to limit biodiversity (Ward 2009). Research concerning desert
43 soil microbial populations and their potential wind dispersion are being increasingly
44 pursued to estimate this microbial biodiversity that may, in fact, play an important
45 role in desert reclamation, soil stability, nutrient cycles and overall environmental
46 health (An et al. 2013; 2015). Microbial community composition and function in
47 desert biomes have been shown to be more diverse than previously assumed and
48 remarkably different from other biomes (Fierer et al. 2012; Leung et al. 2020).
49 However, this biodiversity of microbial communities, and their functional structure,
50 in deserts are still not well understood, and many unanswered questions remain
51 regarding their biology, physiology and ecology (Pointing and Belnap 2012). With the
52 development of high-throughput DNA sequence and analysis technologies, more
53 knowledge is able to be accumulated to fully understand the composition and
54 functioning of desert microbial communities (van Belkum et al. 2001).

55 Amplification and analysis of bacterial (and archaeal) 16S rRNA genes have been
56 widely used as a culture-independent method to document the evolutionary history
57 and taxonomic assignment of individual micro-organisms, as well as to characterize
58 the composition of microbial communities (Head et al. 1998). More recently,
59 meta-barcode methods, using amplicon-based PCR-targeted sequencing of selected
60 hypervariable regions of the 16S rRNA gene, have become the method of choice to
61 determine the composition of microbial populations (Bruno et al. 2015). There exist

62 several relatively small deserts on the Colorado Plateau (USA), which stretch from
63 southeastern California to western Texas, and from Nevada and Utah to the northern
64 Mexican states of Sonora, Chihuahua, and Coahuila and also include much of the
65 peninsula of Baja California in western North America (Wickens 1998). The boundary
66 between what are considered relatively cold and warm deserts lies across southern
67 Nevada and Utah. The Great Basin Desert in Utah has geographic features that
68 resemble the surface of the planet Mars, and is frequently used for geology and
69 astrobiology studies (Chan et al. 2011). These features include extensive wind
70 erosion, moisture deficits, the frequent absence of vegetation, and high daytime UV
71 radiation exposure, among others (Makhalanyane et al. 2015; Leung et al. 2020). Also,
72 the iron oxide prevalent on the surface of the Mars is similar to that found in the
73 Great Basin Desert in Utah, characterized by red-colored hills, soils and sandstones.
74 The Mars Society (<http://www.marssociety.org>) set up the Mars Desert Research
75 Station (MDRS) in the area to act as a Mars analog for these reasons.

76 Bacteria constitute the largest microbial group in soils, and play an important role in
77 maintaining soil processes and thus the functioning and resilience of terrestrial
78 ecosystems (Epp et al. 2012; Lee et al. 2016). Here, we examined the bacterial
79 biodiversity found on several surface soil samples from deserts on the Colorado
80 Plateau. The samples were collected in Goblin Valley (GV), Little Wild Horse Canyon
81 (LWH) located just south of Goblin Valley, to the south of Utah State Road 24 (SR) and
82 Temple Junction Road (TJR) near Goblin Valley State Park, and 200 km southwest of
83 the park near Muddy Creek Road (MCR).

84 Goblin Valley is a State Park in the USA, and lies in the southeastern portion of the
85 eastern edge of the San Rafael Swell, which itself is part of the Colorado Plateau
86 physiographic region. The park is characterized by the presence of hoodoo rocks,
87 referred to as "goblins". These rock sculptures result from the weathering of Entrada

88 Sandstone which was deposited in the Jurassic Period around 170 million years ago,
89 and consists of debris eroded from former highlands and redeposited as alternating
90 layers of sandstone, siltstone and shale (Milligan 1999). The distinct shape of these
91 rocks comes from an erosion-resistant layer of hard rock nestled atop more quickly
92 eroded softer sandstone. The average daytime temperature in summer is between
93 32°C and 41°C, and in winter, daytime temperatures are above freezing most days
94 but can drop to as low as -12°C at night, with occasional snow. The average annual
95 precipitation is < 20cm
96 (<http://www.usgs.gov/media/images/map-annual-average-precipitation-us-1981-20>
97 10). During the summer, monsoons can arrive in some regions between July and
98 mid-September, and can comprise $\geq 50\%$ of the annual precipitation.

99

100 The richness and structure of microbial populations in soils can be affected by many
101 different environmental factors. Studies of microbial biogeography using
102 metagenomics has provided insights into the physiology, environmental tolerances,
103 and ecological strategies of certain microbial taxa, particularly those difficult to
104 culture taxa that can frequently dominate in natural environments (Lauber et al.
105 2009). Hot (or cold) desert soil environments can be regarded as novel ecosystems
106 for fundamental research on ecological and evolutionary processes, as their harsh
107 and oligotrophic environment has nonetheless allowed a rich bacterial biodiversity to
108 exist (Cary et al. 2010). In this study, we examined the bacterial biodiversity and
109 community structures of surface soil from the Colorado Plateau (in the state of Utah,
110 USA) using pyrosequencing of 16S rRNA gene PCR amplicons. We also searched for
111 correlations between selected environmental parameters and bacterial biodiversity
112 in these desert soils, and found that physicochemical factors such as low pH and a
113 relatively low concentration of organic material appears to play an important role in
114 the prevalence of specific bacterial groups in these relatively high-altitude desert
115 soils.

116 **Materials and Methods**

117 **Sampling.** To examine the bacterial biodiversity of surface soils from the Colorado
118 Plateau in Utah (USA), we analyzed 18 samples of surface soil (sand and soft rock
119 particles) from 4 different regions (**Fig. 1**): 1) Dry river bed in Little Wild Horse Canyon
120 (LWH), 2) Goblin Valley State Park (GVA, GVB), 3) near State Road 24 (SRA)/I70
121 Highway and 4) near Muddy Creek Road (MCR). All samples were taken at least 91
122 meters (100 yards) apart. The sampling site GPS coordinates and descriptions are
123 shown in **Table 1**. The samples were collected by scooping surface (top 5 cm) sand
124 into 50 ml sterile polyethylene conical centrifuge tubes at the different locations,
125 during the month of September (2011). After collection, the samples were treated as
126 previously described ([An et al. 2013](#)). The analyses of selected physicochemical
127 parameters of the soil samples was performed, using standard methods, by the
128 Laboratoire d'Analyses de Sols (INRA Arras, France), as previously described ([An et al.](#)
129 [2013](#)).

130 **Sample Preparation.** Total DNA was extracted from each sample using the protocol
131 described in [An et al. \(2013\)](#). Then, aliquots of extracted DNA were adjusted to a final
132 DNA concentration of 15 ng/ μ l in 1/10 TE buffer (1 mM Tris pH 8; 0.1 mM EDTA)
133 using a NanoVue spectrophotometer (GE Healthcare, Little Chalfont,
134 Buckinghamshire, UK), and the concentration verified by ethidium bromide
135 fluorescence after electrophoresis through a 1% agarose gel in TAE buffer (2 mM
136 Tris-acetate pH 8; 5 mM Na-EDTA). PCR reactions were performed in 25 μ l reaction
137 volumes. Each reaction contained one of two different thermostable DNA
138 polymerases and their corresponding reaction buffers, 200 μ M of each dNTP, 0.5 μ M
139 of each primer and 1 to 10 ng of extracted DNA. The 16S rRNA genes were amplified
140 using the universal bacterial primers for pyrosequencing covering hypervariable

141 regions V1-V3: primer 27F (A adaptor + GAGTTTGATCMTGGCTCAG) and primer 518R
142 (B adaptor + Mid + WTTACCGCGGCTGCTGG) (Lee et al. 2010), where A and B
143 represent the adaptors using the 454 Roche FLX Titanium pyrosequencing reaction
144 platform. The Mid sequences are eight nucleotide tags designed for sample
145 identification barcoding according to the 454 pyrosequencing protocol, which
146 allowed long read lengths and allowed the detection of the bacterial biodiversity
147 present in our samples down to the genus level. PCR amplification conditions were
148 adapted for the use of two different thermostable DNA polymerases: A) Phusion
149 High-Fidelity DNA Polymerase (Finnzymes, Finland): 98°C for 2 min, followed by 28
150 cycles of 98°C for 30 secs, 54°C for 20 secs and 72°C for 15 secs, and a final
151 elongation step at 72°C for 5 min; B) Pfu DNA Polymerase (Fermentas, Canada): 95°C
152 for 2 min, followed by 30 cycles of 95°C for 30 secs, 48°C for 30 secs and 72°C for 1
153 min, and a final elongation step at 72°C for 5 min. Each DNA sample was subjected to
154 3-5 different PCR reactions per DNA polymerase to minimize PCR amplification bias.
155 The PCR products were pooled and subjected to electrophoresis through a 1%
156 agarose gel in TAE buffer. After electrophoresis and visualization of the PCR products
157 by ethidium bromide staining and long wave UV light illumination, NucleoSpin
158 Extract II kits (Macherey-Nagel, Germany) were used to purify the 16S rDNA
159 amplicons. Then, 40 ng of PCR products from each sample were mixed for
160 pyrosequencing, which was performed using a 454 Roche FLX Titanium
161 Pyrosequencer (Microsynth AG, Switzerland).

162 **DNA Sequence Data Processing.** The raw DNA sequences were first assigned to each
163 sample via their Mid tag using MOTHUR version 1.33 (Schloss et al. 2009), and reads
164 were removed if at least one of the following criteria was found: (i) length <200 nt
165 or >600 nt, (ii) one mismatch to the barcode sequences or ≥ 1 mismatch to the primer,
166 and (iii) the presence of homopolymers of >8 bp in length. Adaptor sequences were
167 removed from the sequences using the "Cutadapt" tool version 1.8 (Martin 2011)

168 implemented in the Galaxy server of the Institut de Génétique et Microbiologie (IGM)
169 of the Université Paris-Sud (<http://galaxy.igmors.u-psud.fr>) and now available at:
170 <https://github.com/galaxyproject/tools-iuc/tree/master/tools/cutadapt>. Then, the
171 sequences were checked for quality scores by ConDeTri version 2.2 ([Smeds and](#)
172 [Kunstner 2011](#)), using a 90% quality score and those below this score were removed.
173 We used UCHIME ([Edgar et al. 2011](#)) and Decipher ([Wright et al. 2012](#)) to detect
174 chimeric sequences. Sequences detected as chimeras by either program were
175 removed from the data sets. The raw sequences have been deposited in the
176 GenBank sequence read archive (SRA), under the accession number SRP063276.
177 Taxonomy assignments of the remaining 16S rDNA reads were conducted using
178 SILVAngs classifier (online server) with a similarity threshold of 90, using the Silva
179 database release 123 ([Cole et al. 2014](#); [Yilmaz et al. 2014](#)). Sequences classified as
180 chloroplast, eukaryotes and/or mitochondria, or that could not be classified as
181 belonging to the Bacteria Kingdom, were removed. Diversity analyses were
182 performed using the software package MOTHUR. Normalization of 16S rDNA
183 sequences based on the expected gene copy number of different species were
184 performed using the Tax4Fun package in R ([Aßhauer et al. 2015](#)).

185 **Statistical Analyses.** The clean reads were clustered into operational taxonomic units
186 (OTUs) using UPARSE version 7 at a cutoff value of 97% sequence identity ([Edgar](#)
187 [2010](#); [Edgar 2013](#)). The Chao1 and Shannon indices were calculated to estimate
188 taxon richness and diversity ([Schloss and Handelsman 2008](#)). Bray-Curtis distance
189 was used for calculating the distances of multiple samples. The Principal Coordinate
190 Analysis (PCoA) was conducted in the “ade4” R package (version 3.1.2) ([Chessel et al.](#)
191 [2004](#)) to ascertain groups in the microbial communities of the different samples. The
192 `beta_diversity.py` command was used to estimate the beta diversity of bacterial
193 communities using Qiime version 1.80 ([Caporaso et al. 2010](#)). Pearson correlation
194 coefficients between relative abundances of sequences classified to specific phyla of

195 different sample sites and environmental parameters were calculated using SPSS
196 version 22.

197 **Results**

198 **Physical and chemical characteristics of the sand samples**

199 Selected physical and chemical properties of the sand samples are shown in [Table 2](#).
200 For all sample sites, the pH values of the sand were found to be > 8, indicating an
201 alkaline soil environment, which is generally observed in surface sand samples from
202 hot deserts ([Ward 2009](#)). The pH values of the soil at the site of MCR and SRA (see
203 [Table 1](#) for abbreviation name details) are 10.1, higher than the other sampled sites
204 (pH = 8.2 to 8.8), however the pH of all studied sites are in the basic range. The
205 chemical component results also show that the concentrations of sodium (Na),
206 potassium (K), iron (Fe) and aluminum (Al) were higher in these two sites (MCR and
207 SRA) than the others. These results suggest that the high concentrations of these
208 metal ions, which generally arise from chemical decomposition of alkaline salts in the
209 soil, may occur due to the relatively high pH of these soils. In contrast, the
210 concentration of organic matter and organic carbon were relatively lower in these
211 two sites (MCR and SRA), but higher in TJR, which appears to be inversely correlated
212 with the concentrations of Na, K, Fe and Al.

213 **DNA sequencing data analyses**

214 We obtained 483,364 sequences of raw reads in total from the different samples
215 after 454 pyrosequencing (Microsynth AG, Switzerland), with an average read length
216 of 480 nt. After bioinformatic cleaning with Mothur, approximately 95% of the
217 sequences remained (460,475 reads). The sequences were further filtered for quality
218 score and chimeric sequence removal, leaving 65% of the total reads (312,358 reads).
219 The average length of the sequences after processing was approximately 400 nt. The

220 number of remaining reads for each sample are presented in [Table 3](#).

221 **Bacterial community composition**

222 In total, 305,167 sequences from the 18 samples were able to be classified as
223 belonging to the Bacteria domain (sequences classified as chloroplast, eukaryotes
224 and/or mitochondria were removed from the DNA sequence dataset). The sequences
225 were classified at 6 taxonomic levels using SILVAngs, resulting in the delineation of
226 sequences belonging to 37 Phyla, 136 Classes, 136 Orders, 254 Families and 565
227 Genera. The sequences (97%), classified to 37 phyla, were found to represent 13
228 phyla found to be present in all the samples ([Fig. 2a](#)). The predominant phyla of the
229 samples represent members belonging to the *Proteobacteria*, *Actinobacteria*,
230 *Bacteroidetes*, and *Gemmatimonadetes*, at 28.3%, 25.7%, 19.7%, and 6.5%, as the
231 respective averages for the 18 samples, followed by members of the *Acidobacteria*
232 (3.9%), *Chloroflexi* (3.8%), *Firmicutes* (2.6%), *Armatimonadetes* (2.3%),
233 *Deinococcus-Thermus* (1.9%), *TM7* (1.3%), *Saccharibacteria* (1.3%), *Planctomycetes*
234 (1.2%), *Cyanobacteria* (0.9%), and *Verrucomicrobia* (0.5%). Other phyla, with < 0.5%
235 of bacterial sequences in total, were grouped together as “Rare phyla”.

236 We found that the relative abundance of *Gemmatimonadetes* members varies
237 among our sample sites, with the highest proportion in the SR and MCR sites and
238 lowest in the TJR site. Based on an analysis of environmental factors, the SR and MCR
239 sites were found to have the lowest concentration of organic materials and highest
240 concentration of mineral elements, as well as the highest pH value (10.1) of all the
241 sites we examined. The relations between environmental factors and
242 *Gemmatimonadetes* relative abundances were performed using a Pearson’s
243 correlation analysis (Table 4). We found that the abundance of *Gemmatimonadetes*
244 members was significantly correlated ($P < 0.001$) with soil pH ($r^2=0.854$) in the pH

245 ranges from 8.2 to 10.1. We also found that the relative abundance of sequences
246 representing members of this phylum was significantly correlated with the
247 concentration of Na, Fe, K and Al, as well as the concentration of organic carbon and
248 organic materials ($P < 0.05$).

249 The sequences assigned to the *Proteobacteria* phylum, representing 28.3% of the
250 total, could be assigned to the *Alpha*-, *Beta*-, and *Gamma*- classes. We found that the
251 relative abundance of sequences belonging to the *Alphaproteobacteria* varied across
252 the different sites, with the highest percentage in sample TJR (23.8% on average) and
253 the lowest percentage in the SR sample (5.9% on average). Members belonging to
254 the *Proteobacteria* were the most abundant in TJR_3 and TJR_4 (56.5% and 61.2%,
255 respectively), but the primary families in the two sites are different, with
256 *Methylobacteriaceae* from the *Alphaproteobacteria* dominant in TJR_3 (27.8%) and
257 *Xanthomonadaceae*, from *Gammaproteobacteria*, found to be dominant in TJR_4
258 (30.4%). Using a Pearson Correlation test, it was found that the relative abundance of
259 *Proteobacteria* members was inversely correlated ($P < 0.05$) with the soil pH value
260 (over pH values ranging from 8.2 to 10.1), as well as with the concentrations of Na, K,
261 Fe, and Al. We also found that the relative abundance of members of this phylum
262 was positively correlated with the concentration of organic carbon and organic
263 materials ($P < 0.05$).

264 Within the *Betaproteobacteria* class, members of the *Burkholderiales* family were
265 found to be the most abundant in all our samples. A summary of the most abundant
266 sequences classified at the Class level is shown in [Fig. 3](#).

267 The Beta-diversity of bacterial community structures among our different samples
268 was performed to determine if the variability of microbial communities among the
269 different sites was greater than the variability within sites. We found that the
270 bacterial community in our samples was different among the different sites (Fig.2).
271 However, the bacterial communities of the SR, LWH and GV sites were found to be
272 more closely related, and different from those present at the MCR and TJR sites.

273 Using Pearson correlation analyses, an assessment of the physicochemical factors
274 suggests that the concentration of selected mineral components appears to correlate
275 with bacterial community composition (Table 4). In the TJR site, the concentrations of
276 organic carbon (much greater) and mineral salts (much less) were significantly
277 different from those observed at the other sites ($P < 0.05$). Moreover, this site
278 displays the largest dissimilarity in OTU distribution and relative abundance of
279 specific taxonomic groups (*Alphaproteobacteria*, *Acidimicrobiia*, *Gemmatimonadetes*
280 and *Deinococci* (P value < 0.05) versus those observed in the other sites.

281 **Genera composition and prevalence among the samples**

282 Among the 565 different identified genera, 242 genera were found to belong to the
283 phylum *Proteobacteria*, 125 genera to belong to the phylum *Actinobacteria*, 24
284 genera to belong to the phylum *Bacteroidetes*, and 9 genera belonging to the phylum
285 *Acidobacteria*. The most abundant genera found to be present in the sample sites are
286 shown in Table 5. We found sequences classified to certain genera can be very
287 abundant, accounting for $>10\%$ of the total sequences in some samples. Members
288 belonging to the genus *Microvirga* are most abundant in sample TJR_3, and were
289 found to represent almost one-third (32.2%) of the total sequences. Members of this
290 genus have been reported to be thermophiles and have been isolated from a
291 semi-arid site in Brazil, as well as a thermal aquifer (Radl et al. 2014). Sequences
292 belonging to the *Cesiribacter* genus were predominant in the MCR_3 and SRB_2
293 samples, accounting for 30.8% and 19.9% of the total sequences, respectively. Strains
294 of this genus have previously been isolated from desert sand in China (Liu et al. 2012),
295 and may represent an important factor in explaining why these 2 sites cluster closely
296 together by PCoA (Fig. 4). Other abundant genera in the samples, representing $>10\%$
297 of the total sequences, belong to the *Lysobacter* (25.5% in TJR_4), *Salinimicrobium*
298 (12.8% in TJR_2) and *Achromobacter* (11.7% in TJR_4) genera.

299 **Normalization of sequences**

300 To better approximate the composition of the bacterial population in the samples,
301 sequences were normalized based on the estimated 16S rRNA gene copy numbers of
302 different species using the program Tax4fun (Aßhauer et al. 2015). We found that,
303 using this program, the bacterial community composition was altered at all
304 taxonomic levels. At the phylum level (Fig. 2b), the relative abundance of most phyla
305 was found to be reduced, as sequences that could not be aligned to the reference
306 database of 16S rRNA gene copy number were assigned to unclassified sequences
307 (15.5% of total sequences at the phylum level). However, the relative abundance of
308 the *Actinobacteria* increased from 25.7% before normalization to 35.3% after
309 normalization, suggesting that the *Actinobacteria* population was likely
310 underestimated by simply counting classified 16S rRNA genes.

311 **OTU-based analyses**

312 The resultant (after various cleaning steps) sequences were clustered into OTUs at 97%
313 sequence similarity, excluding the unclassified sequences at the phylum level and
314 sequences classified as chloroplast, eukaryote or mitochondria. In total, 9,908 OTUs
315 were discernable in all the samples, ranging from 778 in the Drb_1 sample to 3,595
316 in the Tjr_1 sample. The OTU numbers observed in each sample are shown in Table 3.
317 Eleven core OTU taxons (most abundant OTUs comprising > 1% of total sequences)
318 were found, comprising 37% of the population of all samples. The number of
319 singletons (OTUs with only one sequence) was found to be 3,926, comprising 37% of
320 the total OTU numbers. It is interesting to note that 49 total OTUs were found to be
321 present in all 18 samples. Among the different samples from each Colorado Plateau
322 region, the numbers of common OTUs within a given site were: 599 for DRB (Drb_1
323 and Drb_2), 323 for GV (GV_1 to GV_5), 170 for MCR (Mcr_1 to Mcr_4), 323 for SR
324 (SR_1 to SR_3), and 238 for TJR (Tjr_1 to Tjr_4).

325 The relative similarity of the microbial communities among the samples was

326 calculated based on OTU abundance using the Bray-Curtis measure, and Principal
327 Component Analyses (PCoA) were used to identify the relative population distances
328 of the 18 samples (Fig. 4). It is interesting to note that the Goblin Valley (GB), State
329 Road 24 (SR) and Temple Junction (TJ) sites were able to be grouped into three
330 distinct clusters, while we were unable to discern any distinct grouping of the
331 populations from the Little Wild Horse Canyon (LWH) and Muddy Creek Road (MCR)
332 samples (Fig. 4). The bacterial population distance between different sites, based on
333 the OTU relative abundances, were also illustrated by a UPGMA tree (Fig. 5). The tree
334 was generated using the same number of reads (12,501) from each location by
335 normalization of the sequences using Qiime. The bacterial populations in the sample
336 sites GV and LWH are more closely related, compared with the other sites, and are
337 also the closest in geographic distance. The microbial population of the TJR site is the
338 most distant from the other sites, and may correspond to differences in certain
339 environmental parameters, such as the concentration of organic carbon (much
340 greater) and mineral salts (much lower), as these are significantly different at the TJR
341 site ($P < 0.05$). We used Pearson correlation analyses to identify possible correlations
342 between bacterial diversity and selected environmental factors (Table 6). We found
343 that the most significant correlation with bacterial richness (Chao 1) was the
344 concentration of organic carbon ($P < 0.01$). Moreover, the correlation between
345 community richness and the concentration of mineral elements (Na, K, Al, Fe) was
346 also found to be significant ($P < 0.05$). In addition, we found significant correlations
347 between the Shannon diversity index and the pH value of soils, along with the
348 concentration of sodium ($P < 0.05$). The concentrations of Ca, P, Mg and Mn were not
349 found to be significantly correlated with soil bacterial biodiversity. However, due to
350 the limited number of geographically different sites in our study ($n=5$), any
351 correlation between these environmental factors and bacterial biodiversity will
352 require further examination.

353 **Discussion**

354 The desert regions in the relatively high altitude eastern portion of the state of Utah
355 on the Colorado Plateau present a range of arid conditions, with high salinity and low
356 concentrations of carbon, providing distinctive ecosystems for microbial communities
357 to inhabit. Additionally, these types of environments also have a unique combination
358 of geochemistry and biodiversity that can be useful to study environmental factors
359 that may shape microbial soil community structures ([Crits-Christoph et al. 2013](#);
360 [Osman et al. 2018a](#)). In this study, we used pyrosequencing of 16S rDNA amplicons to
361 identify bacterial populations present on the surface soil in samples of surface soils
362 taken from selected Colorado Plateau desert region sites. We did attempt to quantify
363 the numbers of archaea in our samples using qPCR, but found that they were below
364 our threshold limit of detection (<100 16S rDNA equivalents/g of soil) (Not shown).
365 We also identified possible physicochemical parameters of the soil from these sites
366 that allowed us to begin to investigate possible correlations between bacterial
367 community structure and physical and chemical environmental parameters. [Fierer et](#)
368 [al. \(2012\)](#) reported that soils close to neutral pH had the highest bacterial diversity
369 levels, compared with very basic (most deserts) or acidic (rainforests and Arctic
370 tundra) soils, using metagenomic sequencing of total DNA extracted from soils from
371 a wide variety of ecosystems. However, the contribution of pH to soil bacterial
372 biodiversity may be limited depending upon the contribution of other soil physical
373 and chemical parameters ([Finkel et al. 2012](#); [Osman et al. 2018b](#)).
374 Not surprisingly, given the differences in physical and chemical parameters, previous

375 bacterial community studies of surface soils from selected deserts found that the
376 geographical location can play an important role in the composition of the microbial
377 community (Finkel et al. 2011; Finkel et al. 2012; Nemergut et al. 2013; Wang et al.
378 2020).

379 Non-Arctic desert environments, and their associated soil chemistries, have been
380 reported to present unique bacterial combinations even at higher taxonomic levels
381 (phylum to family) (Tamames et al. 2010). In this study, we found that the relative
382 abundance of sequences assigned to the phyla *Proteobacteria* and
383 *Gemmatimonadetes* can be significantly related to certain environmental parameters,
384 as discerned using a Pearson's correlation analysis (Table 4). For example, salinity has
385 been shown to be an important factor in shaping prokaryotic diversity (Li et al. 2013;
386 Canfora et al. 2014; Geyer et al. 2014; Osman et al. 2018a).

387 Members of the *Actinobacteria* were found to be one of the most abundant phyla
388 found in our samples, accounting for 25.7% of the total sequences. The three
389 predominant classes within this phylum found in our samples belonged to the
390 *Frankiales*, *Micrococcales* and *Acidimicrobiales* classes. It has been reported that
391 sequences assigned to *Actinobacteria* were the most abundant (>70%) in soil
392 samples from the Atacama Desert (Drees et al. 2006). Members of the
393 *Actinobacteria* are often found to be predominant in desert soils (de la Torre et al.
394 2003; Neilson et al. 2012), and often comprise members belonging to the
395 *Rubrobacter*, *Arthrobacter*, and *Streptomyces* genera, as we also observed here.
396 However, many taxa, belonging to *Actinobacteria* isolated from desert soils, appear
397 to represent new species (Mayilraj et al. 2006; Yung et al. 2007; Qin et al. 2009;
398 Makhalanyaane et al. 2015). In our study, we found a large proportion (> 50%) of
399 sequences assigned to this phylum could not be classified to the genus level within

400 the *Frankiales* and *Acidimicrobiales* orders.

401 Members belonging to the *Bacteroidetes* phylum were also found to be abundant in
402 our samples, representing 19.7% of total sequences, and commonly present in other
403 desert soils. Members of this phylum show optimum growth at high pH values, which
404 is consistent with the alkaline pH observed in many hot deserts (Lauber et al. 2009).
405 Members of four families were found to be the predominant *Bacteroidetes* families
406 in our samples: *Cytophagaceae*, *Flammeovirgaceae*, *Chitinophagaceae* and
407 *Flavobacteriaceae*. *Bacteroidetes* phylum members were most abundant in MCR_3,
408 representing 64.3% of the total sequences. *Firmicutes* are often observed to be very
409 abundant in extreme environments, especially in hot deserts (Andrew et al. 2012;
410 Marasco et al. 2012; Crits-Christoph et al. 2013). An et al. (2013) reported that
411 sequences classified to the phylum *Firmicutes* accounted for > 60% of total
412 sequences in a study of surface hot desert soils from the Gobi and Taklamaken
413 deserts.

414 Sequences belonging to members of the *Gemmatimonadetes* phylum are commonly
415 found in many soil types, including hot deserts (Azua-Bustos et al. 2012), pasture
416 land (Chim Chan et al. 2008), crop agriculture soils (Montecchia et al. 2015; Osman et
417 al. 2017), forest soils (Li et al. 2014) and freshwater sediments (Zhang et al. 2015),
418 although only six cultured isolates have been reported to date (Fawaz 2013).
419 DeBruyn et al. (2011) reported that the relative abundance of members of the
420 phylum *Gemmatimonadetes* was inversely correlated to the moisture content in soils,
421 and that many *Gemmatimonadetes* phylotypes have higher relative abundances in
422 semi-arid and arid soils and deserts (DeBruyn et al. 2011). This result was reinforced
423 by a study on the distribution of the *Gemmatimonadetes* in agricultural soils, where
424 members of this phylum were found to be more common in dryer soils with a pH
425 close to neutral (DeBruyn et al. 2011; Fawaz 2013). These results suggest that
426 members of the *Gemmatimonadetes* may be able to act as potential biomarkers of

427 bacterial communities of arid soil environments characterized by relatively high
428 salinity and low levels of organic matter.

429 At the genus level, sequences belonging to members of the *Acidobacteria* phylum
430 were also rarely able to be classified to the genus level in our data. These results
431 suggest that sequences belonging to the phyla *Actinobacteria* and *Acidobacteria* in
432 desert soils may represent new and unknown genera and species. [Direito et al. \(2011\)](#)
433 studied microbial biodiversity near the MDRS site in the Utah portion of the Colorado
434 Plateau using DGGE profiles of ribosomal RNA genes, and identified 239 different
435 clones of bacteria. All phyla observed in their study were also detected in our
436 samples though, not surprisingly, we discerned many more different bacterial taxa
437 using high-throughput pyrosequencing to identify and classify rDNA genes.
438 Sequences assigned to genera that are often called extremophiles, or previously
439 isolated from desert-like environments, were also observed in our samples, including
440 radio-resistant, halophilic, thermophilic and endolithic bacteria. Members of the
441 genus *Pontibacter* (*Bacteroidetes* phylum) were found to be abundant in some hot
442 deserts ([Zhou et al. 2007](#); [Subhash et al. 2014](#)), and may possess unique abilities to
443 adapt to these types of environments ([Makhalanyane et al. 2015](#)). In our samples, we
444 observed that classifiable sequences belonging to this genus are very abundant (2.2%
445 of total sequences), and most abundant in the SRB_2 sample (7.7% of total
446 sequences). Sequences classifiable to the genera *Rubrobacter*, *Truepera* and
447 *Deinococcus* have been shown to be very radiotolerant ([Cox and Battista 2005](#)), and
448 we found many sequences belonging to these genera to be present in our samples.
449 Other genera that have been reported to represent extremophilic bacteria were
450 observed, including: halophilic bacteria such as *Salegentibacter*, *Halomonas* and
451 *Salinimicrobium* ([Rossello-Mora et al. 2003](#); [Oren 2015](#)); thermophilic bacteria such
452 as *Thermoleophilum* and *Thermosporothrix* ([Andrade et al. 1999](#)); and *Rhodobacter*,
453 known as an endolithic bacteria ([Stivaletta et al. 2010](#)). Another group of bacteria
454 that we found belong to the *Pseudomonas* family, members of which have been

455 found to play a protective role for bacterial communities in many extreme
456 environments due to their ability for prolific biofilm formation ([Drenkard and](#)
457 [Ausubel 2002](#); [Selenska-Pobell et al. 2002](#)).

458 Oligotrophic soil ecosystems, such as sandy soils, can be a strong niche filter for the
459 selection of specific bacterial taxa ([Osman et al. 2018c](#)). For example, [Lee et al. \(2016\)](#)
460 recovered samples from the Colorado Plateau Desert (Utah, USA) in the vicinity of
461 Moab (Utah, USA). They showed that colonization in sandstone supported almost
462 exclusively members belonging to a single genus, *Chroococidiopsis*. In our samples,
463 members belonging to the *Proteobacteria* and *Actinobacteria* phyla were found to
464 comprise > 50% of the bacterial communities from the desert sites in the eastern
465 Utah portion of the Colorado Plateau.

466 In this study, we used pyrosequencing of PCR amplified bacterial 16S rRNA genes to
467 reveal the bacterial diversity and community structure in surface soil samples from
468 desert sites in the Utah portion of the Colorado Plateau, and examined their possible
469 correlations with selected environmental factors, revealing that bacterial richness
470 was significantly influenced by the concentration of organic carbon ($P < 0.01$) and
471 mineral elements (Na, K, Al, Fe). It was also observed that bacterial communities
472 from the studied sites were taxonomically different. The research of bacterial
473 communities in desert soils of the Utah portion of the Colorado Plateau presents a
474 unique opportunity to understand bacterial communities in arid soils, characterized
475 by high salinity and low concentrations of organic matter, and at relatively high
476 altitudes.

477

478 **Sample CRediT author statement**

479 **Jorge R. Osman:** Methodology; Resources; Writing - review & editing. **Yang Wang:**
480 Investigation; Methodology; Roles/Writing - original draft. **Chloé Jaubert:**
481 Methodology; Resources. **Tuyet-Nga Nguyen:** Methodology; Software. **Gustavo R.**
482 **Fernandes:** Data curation; Software; Validation; Visualization. **Michael S. DuBow:**
483 Conceptualization; Funding acquisition; Supervision; Project administration; Writing -
484 review & editing.

485 **Acknowledgements**

486 We thank all the members of the Laboratoire de Génomique et Biodiversité
487 Microbienne des Biofilms (LGBMB) of the Institute for Integrative Biology of the Cell
488 (I2BC) and the Editor and two anonymous Reviewers for their interesting discussions,
489 comments and suggestions. JO was a doctoral scholarship recipient from CONICYT
490 Becas Chile. YW was a doctoral scholarship recipient from the China Scholarship
491 Council (CSC). This work was supported by the Centre National de la Recherche
492 Scientifique (CNRS), France.

493 **References**

- 494 Aßhauer, K.P., Wemheuer, B., Daniel, R., Meinicke, P. 2015. Tax4Fun: predicting
495 functional profiles from metagenomic 16S rRNA data. *Bioinformatics*. 31,
496 2882-2884. doi:10.1093/bioinformatics/btv287.
- 497 An, S., Couteau, C., Luo, F., Neveu, J., DuBow, M.S. 2013. Bacterial diversity of surface
498 sand samples from the Gobi and Taklamaken deserts. *Microb. Ecol.* 66,
499 850-860. doi:10.1007/s00248-013-0276-2.
- 500 An, S., Sin, H.H., DuBow, M.S. 2015. Modification of atmospheric sand-associated
501 bacterial communities during Asian sandstorms in China and South Korea.
502 *Heredity*. 114, 460-467. doi:10.1038/hdy.2014.102.
- 503 Andrade, C.M.M.C., Pereira Jr, N., Antranikian, G. 1999. Extremely thermophilic
504 microorganisms and their polymer-hydrolytic enzymes. *Revista de*
505 *Microbiologia*. 30, 287-298. doi: 10.1590/S0001-37141999000400001.
- 506 Andrew, D.R., Fitak, R.R., Munguia-Vega, A., Racolta, A., Martinson, V.G., Dontsova, K.
507 2012. Abiotic factors shape microbial diversity in Sonoran Desert soils. *Appl.*
508 *Environ. Microbiol.* 78, 7527-7537. doi:10.1128/AEM.01459-12.
- 509 Azua-Bustos, A., Urrejola, C., Vicuña, R. 2012. Life at the dry edge: Microorganisms of
510 the Atacama Desert. *FEBS Lett.* 586, 2939-2945. doi:
511 10.1016/j.febslet.2012.07.025.
- 512 Bruno, F., Marinella, M., Santamaria, M. 2015. e-DNA Meta-Barcoding: From NGS
513 Raw Data to Taxonomic Profiling. In: Picardi E (ed) *RNA Bioinformatics*. vol
514 1269. *Methods in Molecular Biology*. Springer New York. pp 257-278.
515 doi:10.1007/978-1-4939-2291-8_16.
- 516 Canfora, L., Bacci, G., Pinzari, F., Lo Papa, G., Dazzi, C., Benedetti, A. 2014. Salinity and
517 bacterial diversity: to what extent does the concentration of salt affect the
518 bacterial community in a saline soil? *PLoS One*. 9, e106662.
519 doi:10.1371/journal.pone.0106662.
- 520 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K.,
521 Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T.,
522 Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D.,
523 Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann,
524 J., Yatsunenko, T., Zaneveld, J., Knight, R. 2010 QIIME allows analysis of high
525 throughput community sequencing data. *Nat. Methods*. 7, 335-336. doi:
526 10.1038/nmeth.f.303.
- 527 Cary, S.C., McDonald, I.R., Barrett, J.E., Cowan, D.A. 2010. On the rocks: the

528 microbiology of Antarctic Dry Valley soils. *Nat. Rev. Micro.* 8, 129-138. doi:
529 0.1038/nrmicro2281.

530 Chan, M.A., Nicoll, K., Ormö, J., Okubo, C., and Komatsu, G. 2011. Utah's geologic and
531 geomorphic analogs to Mars—An overview for planetary exploration, *in* Garry,
532 W.B., and Bleacher, J.E., eds., *Analogues for Planetary Exploration: Geological*
533 *Society of America Special Paper 483*, p. 349–375, doi:10.1130/2011.2483.

534 Chessel, D., Dufour, A.B., Thioulouse, J. 2004. The ade4 package – I: One-table
535 Methods. *R News*. ISSN 1609-3631 4:5-10.

536 Chim Chan, O., Casper, P., Sha, L.Q., Feng, Z.L., Fu, Y., Yang, X.D., Ulrich, A., Zou, X.M.
537 2008. Vegetation cover of forest, shrub and pasture strongly influences soil
538 bacterial community structure as revealed by 16S rRNA gene T-RFLP
539 analysis, *FEMS Microbiol. Ecol.* 64, 449–458. doi:
540 10.1111/j.1574-6941.2008.00488.x.

541 Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T.,
542 Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M. 2014. Ribosomal Database Project:
543 data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42,
544 D633-D642. doi:10.1093/nar/gkt1244.

545 Cox, M.M., Battista, J.R. 2005. *Deinococcus radiodurans* - the consummate survivor.
546 *Nat. Rev. Micro.* 3, 882-892. doi: 10.1038/nrmicro1264

547 Crits-Christoph, A., Robinson, C.K., Barnum, T., Fricke, W.F., Davila, A.F., Jedynek, B.,
548 McKay, C.P., DiRuggiero, J. 2013. Colonization patterns of soil microbial
549 communities in the Atacama Desert. *Microbiome.* 1, 28.
550 doi:10.1186/2049-2618-1-28.

551 de la Torre, J.R., Goebel, B.M., Friedmann, E.I., Pace, N.R. 2003. Microbial diversity of
552 cryptoendolithic communities from the McMurdo Dry Valleys, Antarctica.
553 *Appl. Environ. Microbiol.* 69, 3858-3867.
554 doi:10.1128/AEM.69..7.3858-3867.2003.

555 DeBruyn, J.M., Nixon, L.T., Fawaz, M.N., Johnson, A.M., Radosevich, M. 2011. Global
556 biogeography and quantitative seasonal dynamics of *Gemmatimonadetes* in
557 soil. *Appl. Environ. Microbiol.* 77, 6295-6300. doi:10.1128/AEM.05005-11.

558 Direito, S.O.L., Ehrenfreund, P., Marees, A., Staats, M., Foing, B., Röling, W.F.M. 2011.
559 A wide variety of putative extremophiles and large beta-diversity at the Mars
560 Desert Research Station (Utah). *Int. J. Astrobiol.* 10, 191-207. doi:
561 10.1017/S1473550411000012.

562 Drees, K.P., Neilson, J.W., Betancourt, J.L., Quade, J., Henderson, D.A., Pryor, B.M.,
563 Maier, R.M. 2006. Bacterial community structure in the hyperarid core of the
564 Atacama Desert, Chile. *Appl. Environ. Microbiol.* 72, 7902-7908.

565 doi:10.1128/AEM.01305-06.

566 Drenkard, E., Ausubel, F.M. 2002. *Pseudomonas* biofilm formation and antibiotic
567 resistance are linked to phenotypic variation. *Nature*. 416, 740-743. doi:
568 10.1038/416740a.

569 Edgar, R.C. 2010. Search and clustering orders of magnitude faster than BLAST.
570 *Bioinformatics*. 26, 2460-2461. doi:10.1093/bioinformatics/btq461.

571 Edgar, R.C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon
572 reads. *Nat Methods*. 10, 996-998. doi:10.1038/nmeth.2604.

573 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R. 2011. UCHIME improves
574 sensitivity and speed of chimera detection. *Bioinformatics*. 27, 2194-2200.
575 doi:10.1093/bioinformatics/btr381.

576 Epp, L.S., Boessenkool, S., Bellemain, E.P., Haile, J., Esposito, A., Riaz, T., Erseus, C.,
577 Gusarov, V.I., Edwards, M.E., Johnsen, A., Stenøien, H.K., Hassel, K., Kauserud,
578 H., Yoccoz, N.G., Bråthen, K.A., Willerslev, E., Taberlet, P., Coissac, E.,
579 Brochmann, C. 2012, New environmental metabarcodes for analysing soil
580 DNA: potential for studying past and present ecosystems. *Mol. Ecol*.
581 21:1821-1833. doi:10.1111/j.1365-294X.2012.05537.x.

582 Fawaz M N. 2013. Revealing the ecological role of Gemmatimonadetes through
583 cultivation and molecular analysis of agricultural soils. MSc thesis, University
584 of Tennessee, USA.

585 Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S.,
586 Gilbert, J.A., Wall, D.H., Caporaso, J.G. 2012. Cross-biome metagenomic
587 analyses of soil microbial communities and their functional attributes. *Proc.*
588 *Natl. Acad. Sci. USA*. 109, 21390-21395. doi:10.1073/pnas.1215210110.

589 Finkel, O.M., Burch, A.Y., Elad, T., Huse, S.M., Lindow, S.E., Post, A.F., Belkin, S. 2012.
590 Distance-decay relationships partially determine diversity patterns of
591 phyllosphere bacteria on tamrix trees across the Sonoran Desert. *Appl.*
592 *Environ. Microbiol*. 78, 6187-6193. doi:10.1128/AEM.00888-12.

593 Finkel, O.M., Burch, A.Y., Lindow, S.E., Post, A.F., Belkin, S. 2011. Geographical
594 location determines the population structure in phyllosphere microbial
595 communities of a salt-excreting desert tree. *Appl. Environ. Microbiol*. 77,
596 7647-7655. doi:10.1128/AEM.05565-11.

597 Geyer, K.M., Altrichter, A.E., Takacs-Vesbach, C.D., Van Horn, D.J., Gooseff, M.N.,
598 Barrett, J.E. 2014. Bacterial community composition of divergent soil habitats
599 in a polar desert. *FEMS Microbiol. Ecol*. 89, 490-494.
600 doi:10.1111/1574-6941.12306.

601 Head, I.M., Saunders, J.R., Pickup, R.W. 1998. Microbial evolution, diversity, and

602 ecology: A decade of ribosomal RNA analysis of uncultivated microorganisms.
603 Microb. Ecol. 35, 1-21. doi: 10.1007/s002489900056.

604 Lauber, C.L., Hamady, M., Knight, R., Fierer, N. 2009. Pyrosequencing-based
605 assessment of soil pH as a predictor of soil bacterial community structure at
606 the continental scale. Appl. Environ. Microbiol. 75, 5111-5120.
607 doi:10.1128/aem.00335-09.

608 Lee, T.K., Van Doan, T., Yoo, K., Choi, S., Kim, C., Park, J. 2010. Discovery of commonly
609 existing anode biofilm microbes in two different wastewater treatment MFCs
610 using FLX Titanium pyrosequencing. Appl Microbiol Biotechnol 87, 2335–
611 2343 . <https://doi-org.docelec.univ-lyon1.fr/10.1007/s00253-010-2680-6>.

612 Lee, K.C., Archer, S.D.J., Boyle, R.H., Lacap-Bugler, D.C., Belnap.J., Pointing, S.B. 2016.
613 Niche filtering of bacteria in soil and rock habitats of the Colorado Plateau
614 Desert, Utah, USA. Front. Microbiol. 7, 1489. doi: 10.3389/fmicb.2016.01489.

615 Leung, P.M., Bay, S.K., Meier, D.V., Chiri, E., Cowan, D.A., Gillor, O., Woebken, D.,
616 Greening, C. 2020. Energetic basis of microbial growth and persistence in
617 desert ecosystems. mSystems. 5, e00495. doi: 10.1128/mSystems.00495-19.

618 Li, H., Ye, D., Wang, X., Settles, M., Wang, J., Hao, Z., Zhou, L., Dong, P. Jiang, Y., Ma, Z.
619 2014. Soil bacterial communities of different natural forest types in Northeast
620 China. Plant Soil. 383, 203-216. doi:10.1007/s11104-014-2165-y.

621 Li, K., Liu, R., Zhang, H., Yun, J. 2013. The diversity and abundance of bacteria and
622 oxygenic phototrophs in saline biological desert crusts in Xinjiang, Northwest
623 China. Microb. Ecol. 66:40-48. doi:10.1007/s00248-012-0164-1.

624 Liu, M., Qi, H., Luo, X., Dai, J., Peng, F., Fang, C. 2012. *Cesiribacter roseus* sp. nov., a
625 pink-pigmented bacterium isolated from desert sand. Int. J. Syst. Evol.
626 Microbiol. 62, 96-99. doi:10.1099/ijss.0.028423-0.

627 Makhalanyane, T.P., Valverde, A., Gunnigle, E., Frossard, A., Ramond, J.B., Cowan, D.A.
628 2015. Microbial ecology of hot desert edaphic systems. FEMS Microbiol. Rev.
629 39, 203-221. doi: 10.1093/femsre/fuu011.

630 Marasco, R., Rolli, E., Ettoumi, B., Vigani, G., Mapelli, F., Borin, S., Abou-Hadid, A.F.,
631 El-Behairy, U.A., Sorlini, C., Cherif, A., Zocchi, G., Daffonchio, D. 2012. A
632 drought resistance-promoting microbiome is selected by root system under
633 desert farming. PLoS ONE 7, e48479. doi:10.1371/journal.pone.0048479.

634 Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput
635 sequencing reads. EMBnet.journal. 17, 10-12. doi:10.14806/ej.17.1.200.

636 Mayilraj S, Kroppenstedt R.M., Suresh K., Saini H.S. 2006. *Kocuria himachalensis* sp.
637 nov., an Actinobacterium isolated from the Indian Himalayas. Int. J. Syst. Evol.

638 Micr. 56, 1971-1975. doi:doi:10.1099/ijs.0.63915-0.

639 Milligan, M. 1999. The geology of Goblin Valley State Park. Utah Geological Survey.
640 24p.

641 Montecchia, M.S., Tosi, M., Soria, M.A., Vogrig, J.A., Sydorenko, O., Correa, O.S. 2015.
642 Pyrosequencing reveals changes in soil bacterial communities after
643 conversion of Yungas forests to agriculture. PLoS One. 10, e0119426.
644 doi:10.1371/journal.pone.0119426.

645 Neilson, J., Quade, J., Ortiz, M., Nelson, W., Legatzki, A., Tian, F., LaComb, M.,
646 Betancourt, J.L., Wing, R.A., Soderlund, C.A., Maier, R.M. 2012. Life at the
647 hyperarid margin: novel bacterial diversity in arid soils of the Atacama Desert,
648 Chile. Extremophiles. 16, 553-566. doi:10.1007/s00792-012-0454-z.

649 Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F.,
650 Knelman, J.E., Darcy, J.L., Lynch, C., Wickey, P., Ferrenberg, S. 2013. Patterns
651 and processes of microbial community assembly. Microbiol. Mol. Biol. R. 77,
652 342-356. doi:10.1128/MMBR.00051-12.

653 Oren, A. 2015. Halophilic microbial communities and their environments. Curr. Opin.
654 Biotechnol. 33, 119-124. doi:10.1016/j.copbio.2015.02.005.

655 Osman, J.R, Fernandes, G., Regeard, C., DuBow, M.S. 2017. Bacterial diversity of the
656 rhizosphere and nearby surface soil of rice (*Oryza sativa*) growing in the
657 Camargue (France). Rhizosphere. 3, 112-122. doi:
658 10.1016/j.rhisph.2017.03.002.

659 Osman, J.R, Fernandes, G., Regeard, C., DuBow, M. 2018a. Variation of bacterial
660 biodiversity from saline soils and estuary sediments present near the
661 Mediterranean Sea coast of Camargue (France). Anton. Leeuw. Int. J. G. 112,
662 351-365. doi:10.1007/s10482-018-1164-z.

663 Osman, J.R, Fernandes, G., Regeard C., Jaubert, C., DuBow, M.S. 2018b. Examination
664 of the bacterial biodiversity of coastal eroded surface soils from the Padza de
665 Dapani (Mayotte Island). Geomicrobiol. J. 35, 355-365. doi:
666 10.1080/01490451.2017.1368740.

667 Osman, J.R and DuBow, M.S. 2018c. Bacterial communities on the surface of
668 oligotrophic (nutrient-poor) soils. Curr. Top. Biotechnol. 9, 31-44.

669 Pointing, S.B., Belnap, J. 2012. Microbial colonization and controls in dryland systems.
670 Nat. Rev. Microbiol. 10, 551-562. doi:10.1038/nrmicro2831.

671 Qin, S., Li, J., Chen, H.H., Zhao, G.Z., Zhu, W.Y., Jiang, C.L., Xu, L.H., Li, W.J. 2009.
672 Isolation, diversity, and antimicrobial activity of rare *Actinobacteria* from
673 medicinal plants of tropical rain forests in Xishuangbanna, China. Appl.
674 Environ. Microbiol. 75, 6176-6186. doi:10.1128/AEM.01034-09.

675 Radl, V., Simoes-Araujo, J.L., Leite, J., Passos, S.R., Martins, L.M., Xavier, G.R.,
676 Rumjanek, N.G., Baldani, J.I., Zilli, J.E. 2014. *Microvirga vignae* sp. nov., a root
677 nodule symbiotic bacterium isolated from cowpea grown in semi-arid Brazil.
678 Int. J. Syst. Evol. Microbiol. 64, 725-730. doi:10.1099/ijs.0.053082-0.

679 Rossello-Mora, R., Lee, N., Anton, J., Wagner, M. 2003. Substrate uptake in extremely
680 halophilic microbial communities revealed by microautoradiography and
681 fluorescence *in situ* hybridization. Extremophiles 7, 409-413.
682 doi:10.1007/s00792-003-0336-5.

683 Schloss, P.D., Handelsman, J. 2008. A statistical toolbox for metagenomics: assessing
684 functional diversity in microbial communities. BMC Bioinformatics. 9, 34.
685 doi:10.1186/1471-2105-9-34.

686 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B.,
687 Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B.,
688 Thallinger, G.G., Van Horn, D.J., Weber, C.F. 2009. Introducing MOTHUR:
689 open-source, platform-independent, community-supported software for
690 describing and comparing microbial communities. Appl. Environ. Microbiol.
691 75, 7537-7541. doi:10.1128/AEM.01541-09.

692 Selenska-Pobell, S., Flemming, K., Tzvetkova, T., Raff, J., Schnorpfeil, M., Geißler, A.
693 2002. Bacterial communities in uranium mining waste piles and their
694 interaction with heavy metals. In: Merkel B, Planer-Friedrich B, Wolkersdorfer
695 C (eds) Uranium in the Aquatic Environment. Springer Berlin Heidelberg. pp
696 455-464. doi:10.1007/978-3-642-55668-5_53.

697 Smeds, L., Kunstner, A. 2011. ConDeTri--a content dependent read trimmer for
698 Illumina data. PLoS One. 6, e26314. doi:10.1371/journal.pone.0026314.

699 Stivaletta, N., López-García, P., Boihem, L., Millie, D.F., Barbieri, R. 2010. Biomarkers
700 of endolithic communities within gypsum crusts (Southern Tunisia).
701 Geomicrobiol. J. 27:101-110. doi:10.1080/01490450903410431.

702 Subhash, Y., Sasikala, C., Ramana, Ch.V. 2014. *Pontibacter ruber* sp. nov. and
703 *Pontibacter deserti* sp. nov., isolated from the desert. Int. J. Syst. Evol.
704 Microbiol. 64, 1006-1011. doi:10.1099/ijs.0.058842-0.

705 Tamames, J., Abellan, J.J., Pignatelli, M., Camacho, A., Moya, A. 2010. Environmental
706 distribution of prokaryotic taxa. BMC Microbiol. 10, 85.
707 doi:10.1186/1471-2180-10-85.

708 van Belkum, A., Struelens, M., de Visser, A., Verbrugh, H., Tibayrenc, M. 2001. Role of
709 genomic typing in taxonomy, evolutionary genetics, and microbial
710 epidemiology. Clin. Microbiol. Rev. 14, 547-560.
711 doi:10.1128/CMR.14.3.547-560.2001.

712 Wang Y., Osman J., DuBow, M.S. 2020. Bacterial communities on the surface of the
713 mineral sandy soil from the Desert of Maine (USA). *Curr. Microbiol.* 77,
714 1439-1437. doi:10.1007/s00284-020-01946-z.

715 Ward, D. 2009. *The Biology of Deserts*. London: Oxford Univ. Press. 339 pp. ISBN:
716 978-0-19-921147-0.

717 Wickens, G. 1998. Arid and Semi-arid Environments of the World. In: *Ecophysiology*
718 *of Economic Plants in Arid and Semi-Arid Lands. Adaptations of Desert*
719 *Organisms.* Springer Berlin Heidelberg. pp 5-15.
720 doi:10.1007/978-3-662-03700-3_2.

721 Wright, E.S., Yilmaz, L.S., Noguera, D.R. 2012. DECIPHER, a search-based approach to
722 chimera identification for 16S rRNA sequences. *Appl. Environ. Microbiol.* 78,
723 717-725. doi:10.1128/AEM.06516-11.

724 Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies,
725 J. Ludwig, W., Glöckner, F.O. 2014. The SILVA and "All-species Living Tree
726 Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* 42, D643-D648.
727 doi:10.1093/nar/gkt1209.

728 Yung, P.T., Lester, E.D., Bearman, G., Ponce, A. 2007. An automated front-end monitor
729 for anthrax surveillance systems based on the rapid detection of airborne
730 endospores. *Biotechnol. Bioeng.* 98, 864-871. doi:10.1002/bit.21466.

731 Zhang, J., Yang, Y., Zhao, L., Li, Y., Xie, S., Liu, Y. (2015) Distribution of sediment
732 bacterial and archaeal communities in plateau freshwater lakes. *Appl.*
733 *Microbiol. Biotechnol.* 99, 3291-3302. doi:10.1007/s00253-014-6262-x.

734 Zhou, Y., Wang, X., Liu, H., Zhang, K.Y., Zhang, Y.Q., Lai, R., Li, W.J. 2007. *Pontibacter*
735 *akesuensis* sp. nov., isolated from a desert soil in China. *Int. J. Syst. Evol.*
736 *Microbiol.* 57, 321-325. doi:10.1099/ijs.0.64716.

Table 1. Locations of sample sites

Region	Location Description	Lat/Lon	Sample Name
LWH	Dry river bed in Little Wild Horse Canyon	38°34'32.18" N 110°46'55.06" W	LHW_1 LHW_2
GVA	Goblin Valley State Park	38°34'30.73" N 110°43'47.24" W	GVA_1 GVA_2 GVA_3
GVB		38°34'12.96" N 110°44'31.89" W	GVB_1 GVB_2
MCR	Near Muddy Creek Road	37°19'57.49" N 112°41'42.87" W	MCR_1 MCR_2 MCR_3 MCR_4
SRA	Near State Road 24 / I 70	38°49'51.46" N 110°22'52.35" W	SRA_1
SRB		38°52'53.72" N 110°20'8.87" W	SRB_1 SRB_2
TJR	Near Temple Junction Road / Goblin Valley Road	38°39'38.18" N 110°38'46.79" W	TJR_1 TJR_2 TJR_3 TJR_4

Table 2. Selected physicochemical properties of the soil samples

	LWH	GVA	MCR_1	SRA	TJR
pH	8.2	8.84	10.1	10.1	8.23
Ca (g/100g)	11.6	4.14	4.29	0.59	0.32
Fe (g/100g)	1.49	0.85	1.81	1.85	0.2
Mg (g/100g)	2.42	0.91	1.53	0.96	0.11
P (g/100g)	0.17	0.08	0.14	0.1	0.02
Na (g/100g)	0.37	0.55	0.72	0.81	0.08
K (g/100g)	1.46	1.49	2.26	2.14	0.92
Mn (mg/kg)	511	272	359	172	29.3
Al (g/100g)	3.19	3.11	4.8	4.8	1.11
Organic materials (g/kg)	2.61	1.17	1.15	0.21	3.49
Organic carbon (g/kg)	1.51	0.68	0.67	0.12	11.7

Table 3. Number of sequences and diversity indices of the soil samples

	Sequence numbers		Diversity indices		
	Raw data	After cleaning	OTUs	Chao1	Shannon
LWH_1	11,632	7,572	1,191	1,764	6.11
LWH_2	20,221	13,515	1,570	2,372	5.88
GVA_1	28,382	18,721	1,247	1,924	5.07
GVA_2	30,479	19,357	1,541	2,431	5.56
GVA_3	28,744	18,533	1,694	2,605	5.83
GVB_1	25,342	17,198	1,063	1,763	4.25
GVB_2	27,282	17,258	1,638	2,756	5.79
MCR_1	21,955	14,976	1,459	2,349	5.75
MCR_2	25,839	15,756	1,305	2,216	5.59
MCR_3	45,693	31,143	778	1,307	4.34
MCR_4	19,433	13,165	1,628	2,684	6.08
SRA_1	20,331	14,028	851	1,532	4.85
SRB_1	33,407	22,134	972	1,665	4.59
SRB_2	37,157	25,574	1,170	2,089	4.72
TJR_1	25,380	16,546	3,595	6,669	7.18
TJR_2	28,836	15,976	1,320	2,552	5.42
TJR_3	19,900	12,750	2,196	4,709	4.94
TJR_4	33,351	16,549	2,015	5,383	4.40

Table 4. Pearson's correlation between the relative abundance of sequences of *Proteobacteria* and *Gemmatimonadetes* and selected environmental factors from all the samples.

		pH	Organic carbon	Organic materials	Na	Ca	Fe	Al	P	K	Mg	Mn
<i>Proteobacteria</i>	<i>R</i> ²	-.523*	.595*	.585*	-.623*	-.249	-.569*	-.627*	-.464	-.617*	-.390	-.423
	<i>P</i>	.049	.027	.029	.020	.231	.034	.020	.075	.022	.118	.098
<i>Gemmatimonadetes</i>	<i>R</i> ²	.854**	-.751**	-.896**	.899**	.057	.695**	.844**	.426	.856**	.293	.341
	<i>P</i>	<.001	.004	<.001	<.001	.434	.009	.001	.095	<.001	.191	.152

* with P value < 0.05

** with P value < 0.01

Table 5. Assignments of the most abundant sequences at the genus level in the soil samples. The number of each location represent the average of relative abundance from each numbered site (see Material and Methods for details)

	GVA	GVB	LWH	MCR	SRA	SRB	TJR
<i>Cesiribacter</i>	0.1%	0.6%	1.0%	6.4%	6.0%	7.4%	0.2%
<i>Lysobacter</i>	0.1%	0.2%	0.5%	0.3%	0.2%	0.2%	10.1%
<i>Pontibacter</i>	0.9%	1.2%	1.1%	3.0%	3.5%	5.9%	1.2%
<i>Adhaeribacter</i>	2.8%	1.6%	1.9%	3.1%	1.6%	2.0%	0.8%
<i>Microvirga</i>	0.2%	0.2%	1.0%	0.4%	0.6%	0.2%	7.5%
<i>Sphingomonas</i>	1.2%	1.3%	2.4%	2.0%	0.1%	0.4%	3.3%
<i>Segetibacter</i>	4.0%	2.0%	1.2%	2.3%	0.1%	0.5%	0.8%
<i>Flavisolibacter</i>	3.9%	1.3%	0.8%	1.8%	0.2%	0.3%	1.8%
<i>Massilia</i>	0.9%	3.2%	1.9%	1.4%	5.7%	0.6%	0.9%
<i>Cytophaga</i>	0.3%	0.2%	0.8%	3.1%	3.2%	2.2%	0.1%
<i>Bacillus</i>	1.2%	0.5%	0.9%	2.4%	1.1%	0.2%	1.2%
<i>Truepera</i>	0.5%	0.5%	1.9%	0.6%	4.2%	4.0%	0.2%
<i>Rubellimicrobium</i>	0.9%	0.7%	2.4%	0.9%	1.3%	1.2%	1.1%
<i>Arthrobacter</i>	0.6%	1.5%	1.9%	0.7%	3.3%	1.3%	0.8%
<i>Hymenobacter</i>	1.9%	1.8%	0.8%	1.2%	0.1%	1.2%	0.4%
<i>Blastocatella</i>	0.8%	0.3%	0.6%	1.2%	0.0%	0.1%	2.4%
<i>Salinimicrobium</i>	0.1%	0.1%	0.5%	0.5%	0.7%	0.0%	3.2%
<i>Stenotrophomonas</i>	3.3%	0.1%	0.3%	0.0%	0.0%	1.4%	0.8%
<i>Achromobacter</i>	0.0%	0.1%	0.3%	0.0%	0.0%	0.0%	3.4%
<i>Acinetobacter</i>	3.0%	0.2%	1.0%	0.1%	0.0%	0.9%	0.2%
<i>Rubrobacter</i>	1.8%	0.1%	0.4%	0.5%	0.9%	1.8%	0.3%
<i>Euzebya</i>	0.4%	0.5%	0.3%	1.0%	0.2%	2.4%	0.3%
<i>Deinococcus</i>	0.3%	0.3%	0.8%	1.8%	0.6%	0.8%	0.1%
<i>Flavobacterium</i>	0.2%	4.0%	0.5%	0.0%	0.0%	0.0%	0.6%
<i>Nocardioides</i>	0.6%	0.6%	1.3%	0.3%	1.1%	0.3%	0.7%
<i>Streptomyces</i>	0.3%	0.3%	1.1%	0.2%	6.6%	0.1%	0.1%
<i>Altererythrobacter</i>	0.2%	0.9%	1.5%	0.2%	2.0%	0.1%	0.4%
<i>Pseudomonas</i>	0.6%	1.5%	0.5%	0.1%	0.9%	0.3%	0.6%
<i>Blastococcus</i>	0.3%	0.2%	1.1%	0.3%	0.7%	0.5%	1.0%
<i>Solirubrobacter</i>	1.8%	0.2%	0.3%	0.3%	0.1%	0.3%	0.4%
<i>Gillisia</i>	0.0%	0.1%	0.3%	0.0%	0.0%	0.0%	1.8%
<i>Nibribacter</i>	0.1%	0.4%	0.4%	0.8%	0.5%	1.0%	0.1%
<i>Patulibacter</i>	1.2%	0.5%	0.3%	0.5%	0.0%	0.1%	0.2%
<i>Gemmatimonas</i>	0.7%	0.6%	0.4%	0.6%	0.1%	0.2%	0.3%
<i>Noviherbaspirillum</i>	0.3%	1.0%	0.9%	0.5%	0.1%	0.2%	0.2%
<i>Planomicrobium</i>	0.0%	0.1%	0.3%	0.1%	0.0%	0.1%	1.5%

Table 6. Pearson's correlations between OTU richness and environmental factors

		PH	Organic carbon	Organic materials	Na	Ca	Fe	Al	P	K	Mg	Mn
OTUs	R ²	-.781	.912*	.956**	-.966**	-.028	-.852*	-.930*	-.529	-.875*	-.366	-.340
	P	.060	.016	.005	.004	.482	.033	.011	.180	.026	.272	.288
Chao1	R ²	-.657	.972**	.815*	-.895*	-.356	-.938**	-.938**	-.794	-.865*	-.667	-.644
	P	.114	.003	.047	.020	.278	.009	.009	.054	.029	.110	.120
Shannon	R ²	-.843*	.547	.879*	-.837*	.449	-.648	-.755	-.102	-.757	.087	.175
	P	.036	.170	.025	.039	.224	.119	.070	.435	.069	.445	.389

* with P value < 0.05

** with P value < 0.01

737 **Figure Legends**

738 Fig. 1 Location of the sampling sites in eastern Utah (USA). See Material and Methods
739 for the GPS coordinates.

740 Fig. 2. Relative abundance of bacterial 16S rDNA gene sequences from our samples
741 at the Phylum level. See the Materials and Methods for details.

742 Fig. 3. Relative abundance of bacterial 16S rDNA gene sequences from our samples
743 at the Class level. See the Materials and Methods for details.

744 Fig. 4. PCoA plot of distance of the microbial communities among the 18 samples
745 based on OTU relative abundance using the Bray-Curtis measure. See the Materials
746 and Methods for details.

747 Fig. 5. The bacterial population distance between different sites based on OTU
748 relative abundance illustrated using a UPGMA tree. See the Materials and Methods
749 for details.