JMB Papers in Press. First Published online Jul 12, 2018
DOI: 10.4014.jmb/1804.0405

Manuscript Number: JMB18-04005

Title: Spatial physiochemical and metagenomic analysis of desert environment

Article Type: Research article

Keywords: Desert, arid soil, spatial heterogeneity, microbial diversity, functional mapping, Actinobacteria
Spatial physiochemical and metagenomic analysis of desert environment

Sivakala K.K.¹, Jose P.A.², Anandham R.², Thinesh, T.³, Jebakumar S.R.D.¹#, Sandipan Samadder⁴, Poulami Chatterjee⁴, Sivakumar N.¹* and Tongmin Sa⁴*

¹School of Biotechnology, Madurai Kamaraj University, Madurai, India

²Department of Agricultural Microbiology, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai, India.

³School of Life sciences, Pondicherry University, Pudhucherry, India

⁴Department of Environmental and Biological Chemistry, Chungbuk National University, Cheongju, Republic of Korea

⁵Current affiliation: The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Israel.

# Diseased

* Corresponding authors

Tongmin Sa,
Department of Environmental and Biological Chemistry,
Chungbuk National University,
Cheongju 28644, Republic of Korea
E-mail: tomsa@chungbuk.ac.kr

Dr. Sivakumar, N.
School of Biotechnology
Madurai Kamaraj University Madurai 625021,
India.
Email: microshivaak@gmail.com

Running title: Bacterial community in desert soil
Abstract

Investigating the bacterial diversity and their metabolic capabilities are crucial for interpreting ecological patterns in desert environment, and assessing the presence of exploitable microbial resources. In this study, we evaluated the spatial heterogeneity of physico-chemical parameters, soil bacterial diversity and metabolic adaptation at meter scale. Soil samples were collected from two quadrates a desert environment (Thar Desert, India) which face hot arid climate with very little rainfall and extreme temperatures. Analysis of physico-chemical parameters and subsequent variance analysis (p-values < 0.05) revealed that sulfate, potassium and magnesium ions were the most variable between the quadrates. Microbial diversity of the two quadrates was studied using Illumina bar coded sequencing by targeting V3-V4 regions of 16S rDNA. As the results, 702504 high-quality sequence reads, assigned to 173 operational taxonomic units (OTUs) at species level. The most abundant phyla in both quadrates were Actinobacteria (38.72%), Proteobacteria (32.94%), and Acidobacteria (9.24%). At genus level, Gaiella represented highest prevalence, followed by Streptomyces, Solirubrobacter, Aciditerrimonas, Geminicoccus, Geodermatophilus, Microvirga, and Rubrobacter. Between the quadrates, significant difference (p-values < 0.05) was found in the abundance of Aciditerrimonas, Geodermatophilus Geminicoccus, Ilumatobacter, Marmoricola, Nakamurella and Solirubrobacter. Metabolic functional mapping revealed diverse biological activities, and was significantly correlated with physico-chemical parameters. The results revealed spatial variation of ions, microbial abundance and functional attributes in the studied quadrates, and patchy nature in local scale. Interestingly, abundance of the biotechnologically important phylum Actinobacteria, with large proportion of unclassified species in the desert suggested that this arid environment is the promising site for bioprospection.

Key words

Desert, arid soil, spatial heterogeneity, microbial diversity, functional mapping, Actinobacteria

Introduction

Desert environments usually encounter water scarcity, high levels of temperature fluctuation, ultraviolet radiation, and oligotrophy [1, 2]. All these factors cause extreme abiotic stress to the ecosystem process, eventually limits the vegetation cover in the arid environment. Microorganism associated these environments drives ecosystem process and biogeochemical cycling [3, 4], with
a minimal contribution from plant systems. With the demonstrated microbial activities on various nutrients cycling, it is rational to a premise that there might be a relation between the microbial distribution and nutrients content across space. But, preceding studies have shown relatively less or no correlation between spatial distribution of microorganisms and nutrients [5, 6, 7]. However, physicochemical parameters associated with water availability are better correlated with microbial distribution [7, 8]. On the other hand, the diversity and biological activities of microbial communities, and their correlation with soil parameter have been mostly studied in large spatial scales [9] and smaller scales, especially in rhizospheres [10]. Nevertheless, little is known about the variation on the spatial patterns of microbial diversity and its interactions with soil abiotic properties in arid environments [11, 8], while having ecological importance. As an interesting observation, Pajares et al. [8] have reported the significant correlation between microbial heterogeneity and spatial variation of soluble ions in a desert ecosystem in the middle of the Chihuahuan desert in Mexico.

On the other hand, microorganisms associated with extreme environments including oligotrophic ecosystems are considered as exploitable resources for developing industrially important enzymes and clinically relevant natural compounds. Among different bacteria, the phylum *Actinobacteria* constitutes one of the largest and anciently divergent phyla, and has been extensively studied in various environments due to its unique biotechnological potential [12]. Understanding the diversity and occurrence of such potential microorganisms in previously unexplored arid habitats tend to support bioprospecting efforts. The high diversity of endemic cultivable *Actinobacteria* has recently been found in the samples collected from extremely oligotrophic desert Oasis [13]. Exploring bacterial diversity in similar arid environments, using metagenomic approaches, will be helpful in finding novel habitats for future ecological and bioprospecting studies.

The insights into bacterial diversity and abundance in spatial scale, their biological implications at a yet unexplored desert site may be useful when making ecological inferences and bioprospecting for potential microbes. We intended to understand spatial distribution and importance of microbial diversity of arid soils in the Thar Desert of India. This study was formulated with following objectives a) to reveal physiochemical characteristics and their variations in a spatial scale, b) to explore bacterial communities associated with the arid soil and their variation in spatial scale (Illumina based 16S PCR-amplicon metagenomics approach), and to envisage the metabolic implications in the oligotrophic system (functional mapping of bacterial communities).
Material and Methods

Sampling site

Sampling site of this study is Bikaner, an arid region which has a geographical location of East Longitude 28º1’ and North Latitude 73º19’, in Rajasthan (India). The Bikaner is surrounded by the Thar Desert, and possesses a hot arid climate (BWh) with very little rainfall and extreme temperatures, according to updated Köppen-Geiger climate classification [14]. Based on general elevation, slope and landscape configuration in the terrain, Bikaner lies under physiographic unit Wa - western sandy arid plain [15]. The mean annual temperature is 29°C, while the annual summer temperature approaches to 35.4°C. The annual rainfall ranges between 200-300 mm. The soils are calcareous or non calcareous, sandy in nature. The vegetations are sparse, comprising mainly the xerophytic plants such as *Acacia* sp., *Acacia senegal*, *Prosopis spicigera* and *Prosopis julifora* and some halophytes [15].

Sampling scheme and soil collection

With an objective to explore the spatial relationship between soil chemical distribution and microbial community structure (16S rDNA metagenome) at local scale, quadrate sampling scheme (Fig. 1) was adopted. Soil samples were collected at a depth of 5-10 cm, from two quadrates of 4x4 meter size, represented as quadrate A (Q-A) and quadrate B (Q-B). From the each one meter block, ~250 g soil sample was collected and pooled to get 4 samples from each quadrate. One part of each sample was stored at 4°C for physicochemical analysis and the remainder frozen at -20°C for subsequent DNA extraction and culture-independent microbial diversity analysis.

Physicochemical and biochemical analyses

Eight soil samples derived from Q-A and Q-B were air dried in laminar air hood, sieved through 2 mm mesh, and used for physico-chemical analysis following standard methods. Twenty grams of sieved soil was taken in a clean 100 ml beaker and 40 ml of distilled water added to make soil suspension (1: 2 ratio, w/v). Electrical conductivity was measured in terms of resistance offered to the flow of current using a conductivity meter and expressed in dSm⁻¹. The water-soluble anions (Cl⁻, SO₄²⁻) and cations (Ca²⁺, Mg²⁺, K⁺) in the soils were estimated from the filtrate collected after 19 h of shaking with deionised water, according to Bower et al. [16].
Organic carbon present in the soil samples was estimated by chromic acid wet digestion method [17]. Total nitrogen in soil was determined after the acid digestion by the micro kjeldahl method [18], whereas mineral N (NO$_3^-$-N and NH$_4^+$-N) were extracted using 2M KCl and determined using the spectrophotometer [19]. Total P was estimated according to Pemberton [20].

16S rRNA Gene amplicon sequencing

Soil DNA was isolated using 500 mg of each soil ample, using HiPurA Soil DNA Purification Kit (HiMedia, India) according the manufacturer’s protocol. The isolated soil DNA was visualized by agarose gel (0.8%, w/v) electrophoresis and quantified using nanodrop spectrophotometer.

16S metagenomic sequencing library preparation involved two PCR reactions, an amplicon PCR and an indexing PCR. In amplicon PCR, V3 region of 16S was targeted using primers 341F, 5”-CCTACGGGGNGGCWGCAG-3” and 785R, 5”-GACTACHVGGGTATCTAATCC-3”. Reaction mixture contained 2.5 μl of soil DNA (5 ng/μl in 10 mM Tris pH 8.5), 5 μl of 341F primer (1 μM), 5 μl of 785R primer (1 μM), and 12.5 μl of KAPA HiFi HotStart ReadyMix (2x) in total volume of 25 μl. Thermal cycle conditions were an initial denaturation at 95°C for 3min, followed by 25 cycles at 95°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec, and a final extension step at 72°C for 5 min. One micro litter of the PCR product was analyzed on Bioanalyzer DNA 1000 chip (Agilent Technologies) to verify the expected size of amplicons (~ 550 bp). The amplified product was purified using AMPure XP beads (Agilent Technologies) to remove free primers and primer dimer species, according to manufacturer’s protocol. In a subsequent index PCR, dual indices and Illumina sequencing adapters were attached to product of amplicon PCR, using the Nextera XT Index Kit (Illumina) according to manufacturer’s protocol. The resulted 16S metagenomic sequencing library was cleaned up using AMPure XP beads (Agilent Technologies). One micro litter of 1:50 dilution of the final library was analyzed on a Bioanalyzer DNA 1000 chip to verify the expected size ~630 bp.

Sequence processing and bioinformatics analysis

The final 16S metagenomic library was diluted to 4nM using resuspension buffer. An aliquot 5 μl from each library with unique indices was pooled and sequenced with standard sequencing oligos on the Illumina MiSeq platform. The raw Illumina reads were subjected to stringent quality control (QC) before considering for bacterial community analysis. Illumina reads with more than 70% of
high quality bases with Phred Score greater than 20 were selected using FASTQC Toolkit (Version 0.11.5). A Phred quality score is logarithmically related to the base-calling error probabilities, and being used as a measure to estimate the quality of nucleobases generated by automated DNA sequencing. For instance, Phred Score 20 indicates the probability of one incorrect base call in 100 and represents 99% accuracy of base call in automated sequencing. Subsequently, adapter sequences, primer sequences, the barcode and were removed to generate the QC passed reads.

The QC passed reads were subjected to de-replication (merging of identical reads), de-noising, (trimming to a fixed length, optionally discarding singleton reads) and chimera filtering using UPARSE pipeline [21]. The processed reads were clustered into non-redundant, representative OTUs using UPARSE-OTU algorithm [21], using a 97% cutoff. Sequences which could not be classified to at least a kingdom level were excluded from subsequent analysis. Representatives of operational taxonomic units (OTUs) were determined at a phylum, class, order, genus, and species level. Rarefaction curves were generated using MEGAN5 [22] and compared among the samples. Richness (Chao1), coverage-based richness (ICE and ACE), diversity (Shannon) and evenness were determined using EstimateS [23].

**Taxonomy to functional mapping of microbial community**

With intention to probe the microbial metabolic and functional pathways in the microbial community, taxonomic data was automatically mapped with phenotype categories in METAGENassist [24] using default settings. Percentage of relative abundance of different phenotypes was calculated and a bar chart was plotted.

**Statistical analysis**

All quantitative data obtained on soil were analysed using Minitab 6.0 (Minitab, Inc., State College, Pennsylvania). Non-normal data were normalized by log-transformation. The correlations between each pair of variables were calculated using Pearson's correlation coefficient. Soil physico-chemical parameters were then studied by principal components analysis (PCA), visualized with the two principal components. A multivariate data analysis of the OTUs was performed using METAGENassist, a web server for comparative metagenomics [24]. All the OTUs found in two samples were subjected to normalization based on interquantile range (IQR) [25] and log2-
transformation. Both hierarchical clustering and PLS-DA plotting (Partial least squares Discriminant Analysis) of samples were also performed using METAGENassist.

**Results and Discussion**

**Heterogeneity in physicochemical parameters**

Spatial heterogeneity in physicochemical characteristics was revealed in the arid sampling sites, quadrate A (Q-A) and quadrate B (Q-B) (Table 1). Overall vegetation cover was very less in the experimental area, and observed as 12% and 9% in Q-A and Q-B, respectively. Soil pH and EC varied from 7.4 to 7.7 and 0.14 to 0.47 dS m$^{-1}$, respectively. As expected and reported elsewhere (Pajares et al., 2016), the pH was positively correlated with Mg$^{2+}$. Correlation among different soil parameters are summarized in Supplementary Table S1.

The OC concentration was insignificantly higher in Q-A (5.4 ± 2.8 g kg$^{-1}$) than Q-B (3.9 ± 1.4 g kg$^{-1}$). This difference roughly correlates with relatively dense vegetation cover in Q-A than Q-B, and suggested greater substrate availability for microbial metabolism in this quadrat Q-A. Total forms of nutrients such as total nitrogen, total phosphorous, ammonical N (A-N) and nitrates N (N-N) were showed no significant differences among the two quadrats. The nitrogen content is less in both quadrates (0.18 to 0.21 g kg$^{-1}$) with negligible difference between Q-A and Q-B. This deficiency is, however made up to an extent by presence of both Ammonical-N and Nitrate-N. The total phosphorus content was 0.02 ± 0.008 and 0.042 ± 0.012 g kg$^{-1}$ in Q-A and Q-B, respectively.

The overall values of C and P in the experimental site were higher than the values reported for other deserts, as well as for soil, sand dune and rock substrates in other sites of Thar Desert [26]. Therefore, this type of desert soil is reasonably fertile if moisture is not deficient [27].

In the experimental site, K$^{+}$, SO$_{4}^{2-}$ and Mg$^{2+}$ were the most variable parameters, showed significant differences between the two quadrats, and suggested their contribution to soil heterogeneity in local scale. It agrees well with a recent report described contribution of soluble anions (Cl$^{-}$, SO$_{4}^{2-}$) and cations (Ca$^{2+}$, K$^{+}$) in the heterogeneity of different quadrates studied at Cuatro Cienegas, Mexico [8]. However, significant variation was not observed for Ca$^{2+}$ in the current study at Thar Desert. Among the quadrates, Q-A exhibited the high value of Cl$^{-}$, K$^{+}$ and Mg$^{2+}$ (20 ± 0.4, 13.7 ± 2.9 and 369.7 ± 76.8 mmol kg$^{-1}$, respectively), while Q-B showed high value of SO$_{4}^{2-}$ and Ca$^{2+}$ (72 ± 7.5 and 7.4 ± 1.7 mmol kg$^{-1}$, respectively).
The principal component analysis (PCA) of the physicochemical parameters was performed to further explore their spatial heterogeneity among the two quadrates (Fig. 2). The first principal component (PC 1) explained 36.74%, while the second principal component (PC 2) explained 28.30% of the total variation in the soil parameters among quadrats. The variables associated with the PC1 were pH, Mg\(^{2+}\), K\(^{+}\), Cl\(^{-}\), N-N, OC and EC. The PC 2 was mainly related to soil nutrients such as TN, and TP, along with SO\(^{2-}\).4.

**Sequence data and diversity estimates**

In this study, a total of 702,504 16S rDNA high-quality reads with average read length 301 bp were generated from the desert DNA samples. Rarefaction curves (Supplementary Fig. S1) of both samples were tend to approach saturation plateau, revealed deepness of the experiments and a greater coverage of the bacterial community associated with the desert environment [28]. After the stringent quality filtration using UPARSE metagenome pipeline, 61,585 (8.76%) reads were assigned to 173 OTUs. Samples wise sequence reads, and OTUs along with various diversity estimates are summarized in Table 2. The studied desert samples showed relatively higher and similar diversity, with closer Shannon values for both Q-A and Q-B. These diversity estimates are consistent with previous studies of Thar Desert at different locations [26].

**Bacterial communities in the Desert environment**

Bacterial community profiling of extreme and stressful environments in local scales is crucial for interpreting broader ecological patterns. Especially, extreme hot and cold deserts characterized by high spatial heterogeneity [8], possess endemic microbial communities [29, 30, 13] that are thought to play key roles in ecosystem processes [31]. To test the bacterial heterogeneity in a local scale, as it represents excellent model systems for investigating microbial adaptation and responses to environmental stressors [13], bacterial community data of Thar Desert soils was generated through 16S rRNA deep sequencing using Illumina platform.

The taxonomic composition of Q-A and Q-B was collectively studied first to understand the general bacterial diversity at in the studied desert environment (Fig. 3a). Our analysis indicated that two phyla, namely *Actinobacteria* (38.7%), and *Proteobacteria* (32.9%) are most abundant, followed by moderately abundant *Acidobacteria* (9.24%), and less abundant *Bacteroidetes* (3.46%), *Firmicutes* (3.46%) and *Planctomycetes* (3.46%). At the class level, 38.15% and 26.01%
of the total sequences represented the *Actinobacteria* and *Alphaproteobacteria*, respectively, followed by *Acidobacteria* Gp16 (3.46%). At the bacterial family level, an unclassified family affiliated to *Actinobacteria* was most abundant (15.6%), followed by *Gaelliaceae* (6.35%) and *Streptomycesaceae* (5.78%). At genus level, the *Gaiella* (6.35%), *Streptomyces* (5.2%) and *Solirubrobacter* (4.62%) showed relatively higher abundance, followed by array genera with 3.4 to 0.5% among the population.

Abundance of different bacterial phyla in the studied sample is consistent with previous studies of both hot and cold desert soils [31, 13]. Particularly, abundance of *Actinobacteria* and *Proteobacteria* has been widely reported, with limited variations according to climatic difference and specific niches [26, 31, 32]. A recent real-time quantitative PCR based approach by Rao et al. [26], revealed abundance of Actinobacteria and Proteobacteria in semi-arid soils collected from the other sites of Thar Desert. Similarly, metagenome analysis [31] of hot Namib Desert (Namibia) and cold Miers Valley (Antarctica) also witnessed the abundance of *Actinobacteria* and *Proteobacteria*, while some hypoliths are dominated by *Cyanobacteria* as well. In the present study, *Cyanobacteria* are relatively minor contributor to total bacterial diversity, as visible hypoliths were not found in study site.

The phylum *Actinobacteria* constitutes one of the largest phyla and systematically studied due to its unique biotechnological potential [13], especially in production of drug leads [33]. Consequently, environments that have abundance of *Actinobacteria* capture the attention of natural products researchers [12, 34-36]. In the present study, abundance of *Actinobacteria* is 38.7%, and is comprised of several species affiliated to both unclassified classes and well-known biotechnologically important genus *Streptomyces*. The relative abundance of different genera affiliated to *Actinobacteria* was showed in Supplementary Fig. S2. Ultimately, it is a suggestive sign of suitability of desert environment for bioprospection studies, notably focused on *Actinobacteria*, as witnessed by recent reports [13].

**Spatial heterogeneity of bacterial community in the Desert environment**

Our knowledge of the structure, composition and physiology of biotic components in arid soils has improved in recent past [37, 2, 8]. But, the spatial heterogeneity of bacteria at local scales is largely unaddressed; however, often has ecological significance. In this study heterogeneity in diversity...
of bacteria was studied at local scale, and the notable difference found in the abundance rather than the diversity (Fig. 3b).

Partial least squares discriminant analysis (PLS-DC) was performed and the plot (Supplementary Fig. S3) was generated to illustrate bacterial community variation between the quadrates. First component accounted for 93.1% of variation, and indicated that overall differences were related to the abundance of specific OTUs rather than their presence or absence, as observed in a different study on a soil-borne bacterial structure [38]. This observation suggested that adjacent experimental quadrates harbor similar (but not equal) bacterial communities and the variations were controlled by many bacterial taxa. This result was also complemented by the heat map analysis (Supplementary Fig. S4). An underlying reason behind this difference may be the spatial heterogeneity observed in soil physicochemical parameters at local scale as found in a different desert experimental site [8].

**Physiological and biological mapping of desert bacterial community**

In order to evaluate the physical and biological attributes of the desert bacterial community, automated taxonomy-to-phenotype mapping was done. Among the observed bacterial community, 20 to 32% were belong to mesophiles (Fig. 4a), followed by 7 to 10% of thermophiles, and 6 to 8% of psychrophiles. Abundance of mesophilic bacteria may be attributed to vast temperature fluctuations in the Bikaner [15], and correlates well with different reports of mesophilic bacterial load in arid environments [39, 26]. These mesophilic bacteria were represented mainly by aerobic (Fig. 4b) and free living bacteria (Fig. 4c), as vegetation scarce and sparsely found in the current study sites similar to other arid environments [8].

In the further mapping for biological attributes, predominance of Gram-positive (Fig. 5a) and spore forming bacteria (Fig. 5b) was observed in the quadrates Q-A and Q-B. It reveals the specific bacterial adaptation according to the stress conditions in desert. Indeed, desert bacteria were resisting the extreme conditions either by forming desiccation-resistant spores or tolerating low water potential as vegetative cells [40].

Bacterial communities associated with nutrients limited arid soils play key roles as the drivers of energy input and biogeochemical processes [11, 6, 41, 8]. Based on energy and carbon requirements (mode of nutrition), bacterial taxa associated with the experimental sites were categorized as many as 11 bacterial groups (Fig. 5c). The major nutritional types were heterotroph,
autotroph, and chemoorganotroph in both quadrates. Despite this, significant variations were found in the relative abundance of minor nutritional groups, and strongly attested the nutritional heterogeneity in the desert soil in the local scale.

**Metabolic implications**

Metabolic activities were predicted based on the taxonomy-to-phenotype mapping results. A total of 27 types of bacterial metabolic activities (Fig. 6) were identified among the bacterial communities of the study site. This number includes nitrogen fixation, ammonia oxidizer, dehalogenation, sulfate and sulfur reduction, nitrite reduction, sulfide oxidation and several bioremediation activities (aromatic hydrocarbons degrader, xylan degrader, chitin degrader, sulfur oxidizer, cellulose degrader, and lignin degrader). Similar taxonomy-to-phenotype mapping has recently been performed for gut microbial communities of termites [42]. The bacteria mineralizing organic nutrients were very rich in the desert soil, and capable to run key biogeochemical cycles.

Recently, Vikram et al. [32] described schemes of biogeochemical cycling pathways based on functional capacity dictated by functional marker genes [43] in a hyper arid desert soil niche community. Even though the present study is based on automated taxonomy-to-phenotype mapping, bacterial types were found to be accorded with that deduced by functional genes based approach [32].

**Conclusions**

The microbial communities are key drivers of ecosystem services in arid environments. On the other hand, the arid ecosystems are characterized by high spatial heterogeneity. Together, these features of arid ecosystem result complex ecological events which are largely unaddressed, while considered being under the control of microbial communities and their association with nutrients and other physicochemical gradients in arid soils. The experimental quadrates of the current study showed statistically significant variation of anionic and cationic content and ultimately the microbial community distribution and relative abundance rather than the diversity. The functional implications of the bacterial communities associated with the desert soil were studied though automated taxonomy-to-phenotype mapping approach, results revealed selective abundance of bacterial taxa relevant to physicochemical abiotic conditions. Overall, the variability of physicochemical properties and microbial relative abundance and the metabolic predicted implications of this arid ecosystem are likely to exist in most soils ecosystems, and needs to be
considered when making ecological inferences. Interestingly, the relative abundance of *Actinobacteria*, including large proportion of unclassified species in the Thar Desert suggested that this arid environment is suitable for bioprospection of bioactive natural products. Moreover, these actinobacteria are the suitable models to study and understand abiotic stress alleviating mechanisms in arid ecosystems.

**Acknowledgements**

The authors would like to thank the Basic Science Research Program of the National Research Foundation (NRF) under the Ministry of Education, Science and Technology (2015R1A2A1A05001885), South Korea for providing funding support toward the completion of this study. KKSK gratefully acknowledge UGC (Govt. of India) for the financial support (Research Fellowship), and Tamil Nadu Agricultural University for soil analysis facility. We thank M/s Bionivid Technology Private Limited (Bangalore, India), more specifically Mr. Subodh Tambat for extending NGS sequencing services.

**Conflict of Interest**

The authors disclose no potential conflicts of interest associated with this manuscript.

**References**


**Main text Figure Legends**

**Fig. 1** Map showing sampling point and design. (Map was generated using Diva –GIS (version 7.5))

**Fig. 2**: PCA analysis: loading plot of soil physicochemical parameters (EC = electrical conductivity, OC = organic carbon, TP = total phosphorus, TN = total nitrogen, AN = ammoniacal nitrogen, NN = nitrate-nitrogen).

**Fig. 3**: Diversity and relative abundance of different bacteria in the arid experimental site (a), and relative difference in Q-A and Q-B (b)

**Fig. 4** Mapping of desert bacterial community to different physiological characteristics, growth temperature (a), respiration (b) and interactions with other organisms (c). Inner and outer ripples represent Q-A and Q-B, respectively.

**Fig. 5**: Mapping of desert bacterial community to different biological characteristics, cell membrane type (a), sporulation (b) and nutritional types (c). Inner and outer ripples represent Q-A and Q-B, respectively.

**Fig. 6**: Microbial metabolic functions encoded in bacterial communities from experimental sites.
**Supplementary Table and Figure Legends**

**Table S1**: Correlation coefficients among different soil parameters studied in the Thar Desert (India).

**Fig. S1** Rarefaction curves of quadrates, Q-A and Q-B shows the coverage of the bacterial community in the present study.

**Fig. S2** Relative abundance of different genera affiliated to *Actinobacteria*.

**Fig. S3** PLS-DA plot of bacterial community associated with quadrates, Q-A and Q-B.

**Fig. S4** Heat-map showing correlation between bacterial communities associated with quadrates, Q-A and Q-B.
Table 1: Different physicochemical parameters of arid soils averaged over sampling quadrates, Q-A and Q-B.

<table>
<thead>
<tr>
<th>Quadrates</th>
<th>Physicochemical parameters</th>
<th>pH</th>
<th>EC (dSm⁻¹)</th>
<th>OC (g kg⁻¹)</th>
<th>TN (g kg⁻¹)</th>
<th>TP (g kg⁻¹)</th>
<th>Cl⁻ (mmol kg⁻¹)</th>
<th>SO₄²⁻ (mmol kg⁻¹)</th>
<th>K⁺ (mmol kg⁻¹)</th>
<th>Ca²⁺ (mmol kg⁻¹)</th>
<th>Mg²⁺ (mmol kg⁻¹)</th>
<th>A-N (mg kg⁻¹)</th>
<th>N-N (mg kg⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q-A</td>
<td></td>
<td>7.58 ±0.11</td>
<td>0.28 ±0.06</td>
<td>5.4 ±2.8</td>
<td>0.21 ±0.12</td>
<td>0.02 ±0.008</td>
<td>20 ±4.0</td>
<td>44 ±17.4</td>
<td>13.7 ±2.9</td>
<td>6.8 ±1.8</td>
<td>369.7 ±20</td>
<td>1.40 ±0.32</td>
<td>3.01 ±1.40</td>
</tr>
<tr>
<td>Q-B</td>
<td></td>
<td>7.54 ±0.03</td>
<td>0.3 ±0.14</td>
<td>3.9 ±1.4</td>
<td>0.18 ±0.15</td>
<td>0.042 ±0.012</td>
<td>15 ±4.0</td>
<td>72 ±7.5</td>
<td>9.25 ±1.9</td>
<td>7.4 ±1.7</td>
<td>240.2 ±15</td>
<td>1.19 ±0.14</td>
<td>2.73 ±0.26</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>7.56 ±0.08</td>
<td>0.29 ±0.10</td>
<td>4.7 ±2.2</td>
<td>0.20 ±0.13</td>
<td>0.031 ±0.015</td>
<td>17.5 ±4.6</td>
<td>58 ±19.4</td>
<td>11.5 ±3.3</td>
<td>7.1 ±1.66</td>
<td>305 ±8.7</td>
<td>1.29 ±0.25</td>
<td>2.87 ±0.46</td>
</tr>
</tbody>
</table>

*Significant difference (p<0.05) among quadrates

**EC:** electrical conductivity, **OC:** organic carbon, **TN:** total nitrogen, **TP:** total phosphorus, **A-N:** ammonia-nitrogen, **N-N:** nitrate-nitrogen.
<table>
<thead>
<tr>
<th>Richness estimates and diversity indices</th>
<th>Q-A</th>
<th>Q-B</th>
</tr>
</thead>
<tbody>
<tr>
<td>QC passed reads</td>
<td>3,98,170</td>
<td>3,04,334</td>
</tr>
<tr>
<td>Processed reads</td>
<td>33418</td>
<td>28168</td>
</tr>
<tr>
<td>Number of OTUs</td>
<td>173</td>
<td>170</td>
</tr>
<tr>
<td>Chao1 (richness)</td>
<td>174.03</td>
<td>173</td>
</tr>
<tr>
<td>ACE (coverage-based richness)</td>
<td>178.54</td>
<td>173</td>
</tr>
<tr>
<td>Shannon (Diversity)</td>
<td>4.33</td>
<td>4.43</td>
</tr>
</tbody>
</table>
Fig. 1. Map showing sampling point and design. (Map was generated using Diva-GIS (version 7.5).)
Fig. 2. PCA analysis: loading plot of soil physicochemical parameters (EC = electrical conductivity, OC = organic carbon, TP = total phosphorus, TN = total nitrogen, AN = ammoniacal nitrogen, NN = nitrate-nitrogen).
Fig. 3. Diversity and relative abundance of different bacteria in the arid experimental site (a), and relative difference in Q-A and Q-B (b)
Fig. 4. Mapping of desert bacterial community to different physiological characteristics, growth temperature (a), respiration (b) and interactions with other organisms (c). Inner and outer ripples represent Q-A and Q-B, respectively.
Fig. 5. Mapping of desert bacterial community to different biological characteristics, cell membrane type (a), sporulation (b) and nutritional types (c). Inner and outer ripples represent Q-A and Q-B, respectively.
Fig. 6. Microbial metabolic functions encoded in bacterial communities from experimental sites.