**Response of an extremophile microbiome to a major climatic perturbation reveals distinct community adaptation strategies**

Gherman Uritskiy, Samantha Getsin, Adam Munn,

James Taylor\* and Jocelyne DiRuggiero\*

**INTRODUCTION**

Microbial communities harbor vast taxonomic and functional diversity and as such, they have a tremendous ability to adapt to, and recover from a wide variety of environmental changes [1, 2]. While the microbiome’s higher-order taxonomic composition (i.e., phylum, order) is often partially linked to its functional potential, the fine-scale individual membership of the communities has limited impact on their overall functioning, due to functional redundancy between closely related taxa [3]. This ensures the functional potential of the microbiome persists even after a major community rearrangement [4-6].

The resilience and adaptation of microbiomes to major perturbations such as temperature changes or antibiotic administration have been demonstrated in controlled environments [1, 7, 8]. However, due to compounding external factors such studies are much more difficult under natural environmental conditions. Previous longitudinal studies looked at adaptations of environmental microbiomes in response to gradually changing environmental conditions [9], but adaptations to acute changes remain largely unexplored, particularly in extreme environments.

The northern Atacama Desert is one of the harshest places on Earth, with an average annual precipitation of less than 1mm [10, 11]. Despite this, poly-extremophilic microbiota have evolved to exist in these extreme conditions by relying on the protection of various minerals [12]. One such endolithic (inside rock) community type is harbored within halites – salt rock nodule formations found on the surface salt flats in Salar Grande [13, 14]. A unique property of halite microbiomes that make them a compelling system for measuring responses to environmental stressors is their contained nature. Encased in rocks, halite communities have minimal biomass exchange, and have limited nutrient input beyond atmospheric water and gasses [13-15]. As such, each halite nodule represents a near-closed system, allowing us to track microbial community changes with limited external factors compounding the results.

Because of salt’s deliquescent properties, the halite endoliths are able to survive by receiving water almost exclusively from the humidity in the air [14-16]. While halite microbiomes are primarily comprised of *Archaea*, they rely on carbon fixation by halophilic algae and cyanobacteria, which support a number of heterotrophs [14, 15]. In particular, the majority of the biomass comes from *Halobacteria* and *Bacteroidetes* – two taxonomically diverse groups of hyper-halophilic salt-in strategists. This unique adaptation allows them to keep sodium ions out with an internal osmotic pressure from potassium ions, which they actively pump in and is less damaging to the cell [14, 17, 18]. This is energetically favorable to actively pumping out sodium [19], but required the cells’ proteomes to be have an extremely low isoelectric point (pI) to be able to function at high potassium concentrations [20-22].

The highly specialized nature of the halite microbial communities can make them more vulnerable to change compared to habitat generalists [18], particularly to sudden changes in water availability. In August 2015, Northern Atacama received its first major rain in 13 years [10, 23]. Combined with the isolated nature of halite microbiomes, this gave us a perfect opportunity to track the response of an environmental microbiome to a major natural perturbation. Our longitudinal study not only captured the microbiome’s immediate adaptations, but also its recovery in the subsequent year, revealing two strikingly different community adaptation strategies.

**RESULTS**

**Longitudinal sampling strategy and sequencing approach**

To investigate the temporal dynamics of halite microbiomes, samples of halite nodules from Salar Grande were harvested at regular intervals in a 4-year longitudinal study, capturing the rare rain events that occurred in 2015 throughout the desert [10]. A nearby weather station (Diego Aracena airport), located 47.7km North of the sampling site, recorded rainfalls on 2015-08 (4.1mm). The previous notable precipitation in the area occurred in 2002 (4.1mm) [23, 24].

The main sampling site was revisited four times during the study – twice before the rain (2014-09 and 2015-06), and twice after the rain (2016-02 and 2017-02). For each time point, 9-12 biological replicates were collected and their 16S rDNA sequenced, yielding 535,233 paired-end reads (insert size 419±7bp), which were used for taxonomic profiling of the microbiomes (Table S2). For 5 biological replicates at each time point, whole-metagenomic (WMG) sequencing was also done to determine the functional potential of the communities over time, yielding a total of 70,689,467 paired-end reads (insert size 277±217bp). In addition, a nearby site was also sampled after the rain at a higher temporal resolution (2016-02, 2016-07, 2016-10, and 2017-02), with 5-13 replicates per time point. The 16S rDNA amplicons from samples at this site were also sequenced, yielding 357,325 paired end 250bp reads (insert size 419±4bp).

**Taxonomic structure and functional potential of the community were temporarily perturbed, but were resilient long-term**

The halite community structure shifted from an *Archaea*-dominated community to a more balanced *Archaea-Bacteria* community following the rain. The relative abundance of *Archaea* dropped significantly (two-sided t-tests: pval<0.0001) in both rDNA amplicon (Figure 1B) and WMG sequencing (Figure S2), but recovered in the following year. We found that many Phyla also followed a similar trend of change and recovery - *Cyanobacteria*, Green algae (estimated by chloroplast rDNA abundance), and *Bacteroidetes* significantly increased in relative abundance following the rain, and gradually lowered back to baseline abundance in the following year. On the other hand, the abundance of *Halobacteria* (the major Archaea phylum in this community) significantly decreased and subsequently recovered following the rain (Figures S1B, S2). This trend in domain and phyla abundance recovery was also seen in the sequencing of the supplementary site (Site 2), with incremental shifts over 18 months after the rain. (Figure S3).

Analysis of the Weighted Unifrac dissimilarity matrix (constructed from OTUs clustered at 97% identiy) confirmed that the overall structure of the halite communities was different between time-points (permanova: pval<0.001), with the the halite microbial community shifting following the rain, but recovering in the following year (Figure 1A). Samples from before the rain (2014-09 and 2015-06) clustered together, providing a baseline to compare the post-rain samples against (Figure S1E). Strikingly, 2016-02 samples (6-months post-rain) were significantly more distant from the pre-rain samples than this baseline (two-sided t-test: pval<0.0001), but 2017-02 samples (18-months post-rain) were more similar to the pre-rain samples than to the 2016-02 samples (two-sided t-test: pval<0.0001) (Figure S1A).

The functional potential of the community, determined with annotation of KEGG functional pathways in the WMG co-assembly, also significantly changed after the rain. Consistent with the taxonomy-based clustering, samples from before the rain (2014-09 and 2015-06) were distinctly separate from samples collected shortly after the rain (2016-02) (Figure 1C). Indeed, 2014-09 samples had a significantly stronger Pearson correlation with 2015-06 samples than with 2016-02 samples (two-sided t-test: pval<0.0001). Furthermore, the 2014-09 samples correlated significantly better with 2017-02 samples than 2016-02 samples did (two-sided t-test: pval<0.0001), indicating a recovery in the functional potential of the microbiome.

While the majority community functions were present in similar abundances between replicates and time points, a number of gene functions were differentially represented between time points (ANOVA test, *pval*<0.01, FRD<1%) (Fig 1D). Of these, the majority were significantly over- or under-represented in the samples collected shortly after the rain (2016-02) (SigClust 2-group significance: pval<0.0001), but present in similar abundances in the pre-rain samples and the “recovered” samples. Cell motility, nucleic acid processing and repair, and lipid metabolism were some of the major functional categories that changed in relative abundance after the rain, indicating active adaptation to changes in the internal halite conditions.

**Differences in salt adaptations likely drove fitness of salt-in halophilic strategists**

The majority of this halophilic community is comprised of salt-in strategists (*Halobacteria* and *Bacteroidetes*), which actively import potassium ions, producing an internal osmotic pressure that keep out sodium ions [14, 17, 18]. As a result of this adaptation, their proteomes have unusually low isoelectric points (*pI*) to function at high potassium concentrations

[20-22]. We found that the distribution of *pI* of proteins encoded in community gene pool changed significantly after the rain (Kolmogorov-Smirnov 2-sample test: pval<0.0001) (Figure S5A). The average *pI* significantly increased after the rain (two-sided t-test: pval<0.01), and then recovered to pre-rain levels in the following year (Figure S5B). This is largely explained by the *pI* differences *Halobacteria* (*pI*=5.04) and *Bacteroidetes* (*pI*=5.80), which were the major taxonomic groups that changed in abundance after the rain (Kolmogorov-Smirnov 2-sample test: pval<0.0001) (Figure S5D). Consistent with salt-in adaptations, we found that the average potassium uptake potential (estimated from Trk gene abundances) of the communities significantly decreased after the rain, and then recovered to pre-rain levels in the following year (Figure S5C). The shift in gene pool *pI* and potassium uptake potential was also observed within the highly heterogeneous *Halobacteria* phylum (Figures S5E,F). Together, these results suggest adaptations to a temporary decrease in salt concentrations during the rains.

**Taxonomic composition of functional niches was permanently rearranged by the perturbation**

Analysis and hierarchical clustering of the Unweighted Unifrac dissimilarity matrix revealed that the halite communities harvested at different dates were also different in terms of presence or absence of OTUs (97% identity) (permanova test: pval<0.001). Strikingly, we found that samples harvested shortly after the rain (2016-02) were not only more distant from pre-rain samples than the baseline pre-rain variance (two-sided t-test: pval<0.0001), but also clustered together with the samples collected 18 months after the rain (2017-02) (Figure 2A). Indeed, we found that 2017-02 samples were less distant to 2016-02 samples than the pre-rain samples (two-sided t-test: pval<0.0001), suggesting that the community did not return to its initial state after the perturbation in terms of the presence or absence of individual OTUs. This contrasts the results from clustering the Weighted Unifrac dissimilarity matrix (Figure 1A), which suggester the community made a recovery.

This fine-scale composition of the microbiomes was also investigated through metagenome-assembled genomes (MAGs). With the use of metaWRAP [25], 94 high-quality MAGs (>70% completion, <5% contamination) were recovered from the WMG sequencing data (Table S1), and their abundances were tracked over time. Despite high heterogeneity in MAG abundances, hierarchical clustering of the abundance table revealed two significantly distinct groups of replicates (SigClust 2-group significance: pval<0.05) – pre-rain samples (2014-09 and 2015-06) and post-rain samples (2016-02 and 2017-02) (Figure 2B). While MAG abundances changed during the post-rain recovery (2016-02 to 2017-02), the resulting change was subtler when compared to the drastic MAG rearrangement immediately following the rain. Pearson correlation comparison (two-sided t-test: pval<0.0001) as well as group significance analysis (SigClust 2-group significance: pval<0.01) of the contig abundance table further illustrated that the community did not recover from the rain in terms of individual community member abundance, as 2017-02 samples were better correlated with 2016-02 than with the pre-rain samples (Figures 2C, S6). Additionally, two new *Cyanobacteria* MAGs – *Synechococcaceae* and *Chroococcales* – that were previously reported to appear in only a small fraction of the halites [15], appeared in high abundances in most of the samples after the rain (Figure S7). Taken together with the resilience of the higher-order community structure, these results suggest that while the abundances of higher-order taxonomic ranks recovered to the pre-rain state, the individual organism within those groups have been permanently reshuffled. Even more strikingly, we found that the functional potentials of the MAGs were not predictive of their survival after the rain, suggesting that this could have been a stochastic process.

To investigate changes in functional niche membership, a rearrangement index (*RI*) was calculated for each KEGG KO identifier, evaluating the degree of change over time in taxa that carry a given function (see Methods). The distribution in the *RIs* of all functions between two time points () allows us to visualize and quantify changes in niche representation between two time points (Figure 2D, S7). Compared to the baseline weighted average rearrangement index before the rain (=0.24±0.05), the rearrangement following the rain was significantly higher (=0.46±0.07) (Kolmogorov-Smirnov 2-sample test: pval<0.0001), indicating that the same community functions were being performed by a new set of organisms. However, the rearrangement of functional niche membership during the year following the rain (=0.32+/-0.04) was significantly lower than right after the rain event (Kolmogorov-Smirnov 2-sample test: pval<0.0001), and the final communities found in 2017-02 samples were still very distinct from pre-rain samples (= 0.43+/-0.09). Taken together, these results suggest that while the 2017 halite communities function the same as they did prior to the rain in 2014/15, their functional niches are constituted by a new set of microbial strains and species.

**The microbiome’s response and recovery from the rain reveal contrasting community adaptation strategies**

The two composition shifts that the halite microbiomes underwent following the rain – the initial response and subsequent recovery – resulted in a similar degree of change to the overall functional potential of the community, as indicated by the similarity between the pre-rain and post-recovery samples. However, there two shifts were fundamentally distinct. The first shift caused functional rearrangement not only at the community level by means of changing relative abundances of major taxa, but also within finer taxonomic ranks (Figure S4), and likely resulted from rapid changes in selective pressures following the rain. The second functional potential shift came largely from changes in relative abundance of diverse *Halobacteria* and *Bacteroidetes*, which themselves remained largely unchanged. This shift was subtler, possibly indicative of the gradual nature of the recovery phase.

Observing the responses of the halite microbial communities to the rain event and their subsequent recovery over the course of the following 18 months granted us a better understanding of microbiome dynamics in response to the changing environment. We postulate that microbial communities can undergo two fundamentally distinct types of taxonomic composition shifts in response to changing selective pressures – a rapid rearrangement of the community at higher and lower taxonomic ranks, or a gradual shift resulting from changes in abundances of higher-order taxa (Figure 3).

**DISCUSSION**

Our findings demonstrate that Atacama Desert’s extremophiles are hyper-sensitive to environmental climate changes, however their higher-order taxonomic structure and functional potential are resilient in the long-term. We found that just 4mm of rainfall induced a drastic temporary change in the higher-order community structure, and a permanent rearrangement of the individual composition of the community. The highly specialized nature of halite microbiomes renders them particularly sensitive to changes in water availability, as hyper-halophiles have a narrow range of tolerated salt concentrations [26]. Because the nodules are primarily comprised of porous salt [27], a major rain could at temporarily alter the internal osmotic conditions. The 2015 rain, which was the first major precipitation in 13 years, resulted in increased average isoelectric point (*pI*) of proteins encoded in the community’s genepool, and decrease in potassium uptake potential. Considering that low *pI* and potassium uptake are the hallmark of hyper-halophilic salt-in strategists [20-22], we deduce that the rain temporarily decreased the salt concentrations within the colonized pores of the halites, disrupting existing micro niches within.

Due to the rapid change in selective pressures, this acute perturbation resulted in composition adjustments at the higher-order taxonomic ranks, as well as a complete rearrangement of the finer taxa that constitute them. It is surprising that we observe these changes in samples harvested 6 months after the rain (2016-02 samples), and it suggests that the immediate effects of the rain may have been much more drastic. This also reveals the slow-growing nature of these microbiomes, which likely results from scarce resources and harsh environmental conditions [14]. A previous study documented great taxonomic diversity between microbiomes within individual halites, suggesting the presence of a seed bank – diverse collection of low-abundance organisms present in all halites, which opportunistically and stochastically occupy any given halite [15]. However, due to functional redundancy of closely-related community members [3-5], the functional landscape of the communities remains relatively robust. We infer that the significant perturbation from the rain created gaps in functional niches that new organisms from the seed bank opportunistically took over [28], “resetting” the colonization of each halite.

Despite the major shift however, the community was able to gradually recover functionally and taxonomically (at domain and phylum levels) over the course of 18 months. Contrasting the original shift, the community returned to the pre-rain functional potential almost exclusively via changes in relative abundance of higher-order taxonomic groups, which themselves remained largely unchanged functionally and taxonomically. The recovered community appears to be functionally equivalent to the pre-rain community, while being comprised of a new set of individual organisms.

The response and recovery of this system to the perturbation allowed for inference of two contrasting models of microbial community shift in response to changing environmental conditions. The first type of shift an adjustment in existing community structure, and results from gradual changes in environmental conditions, which gives the currently dominant organisms an opportunity to adjust without allowing new organism to take over. The functional potential changes in this type of shift are driven by changes in overall higher-order taxonomic composition. The second type of shift is a community rearrangement, resulting from adaptations to a sudden major perturbation, which creates a void in existing functional niches, presenting an opportunity for new organism from the seed bank to come in. In this type of response, the changes in the community’s overall functional potential are driven by changes in higher order taxonomic composition, as well as changes within the individual taxa.

**METHODS**

**Sample collection and DNA extraction**

Halites were harvested from three sites in Salar Grande, a Salar in the Northern part of the Atacama Desert (Robinson et al., 2014). All the sites were within 5 km of each other and, at each site, halite nodules were harvested within a 50m3 area. Sites were as follow: S1 was used for most of the analysis in this work, S2 was used for one analysis post-rain; and S3 was used to improve binning results but not for abundance calculation because too few samples and replicates were collected (See Table S2 for details on sampling sites and replication). Halite nodules were collected as in Robinson et al. (2014) and ground into a powder, pooling from 1-3 nodules until sufficient material was collected, and stored in dark in dry conditions until DNA extraction in the lab. gDNA was extracted as previously described [13, 14] with the DNAeasy Powersoil DNA extraction kit (QIAGEN).

**16S rDNA amplicon library preparation and sequencing**

The communities’ 16S rDNA was amplified with a 2-step amplification and barcoding PCR strategy as previously described [13] by amplifying the hypervariable V3-V4 region with 515F (ACACGACGCTCTTCCGATCTGTGYCAGCMGCCGCGGTAA) and 926R (CGGCATTCCTGCTGAACCGCTCTTCCGATCTCCGYCAATTYMTTTRAGTTT) primers. PCR was done with the Phusion High-Fidelity PCR kit (New England BioLabs) with 40ng of gDNA. Barcoded samples were quantified with the Qubit dsDNA HS Assay Kit (Invitrogen), pooled and sequenced on the Illumina MiSeq platform with 250 bp paired-end reads at the Johns Hopkins Genetic Resources Core Facility (GRCF).

**WMG library preparation and sequencing**

Whole genome sequencing libraries were prepared using the KAPA HyperPlus kit (Roche). The fragmentation was performed with 5ng of input gDNA for 6 minutes to achieve size peaks of 800bp. Library amplification was done with dual-index primers for a total of 7 cycles, and the product library was cleaned 3 times with XP AMPure Beads (New England BioLabs) to remove short fragments and primers (bead ratios 1X and 0.6X, keep beads) and long fragments (0.4X bead ratio, discard beads). Other steps followed the manufacturer’s recommendations. The final barcoded libraries were quantified with Qubit dsDNA HS kit, inspected on a dsDNA HS Bioanalyzer, pooled to equal molarity, and sequenced with paired 150bp reads on the HiSeq 2500 at GRCF.

**16S rDNA amplicon sequence analysis**

The de-multiplexed and quality trimmed 16S amplicon reads from the MiSeq sequencer were processed with MacQIIME v1.9.1 [29]. Samples from site 1 and 2 were processed separately. The reads were clustered into OTUs at a 97% similarity cutoff with the pick\_open\_reference\_otus.py function (with --suppress\_step4 option), using the SILVA 123 database [30] release as reference and USEARCH v6.1.554 [31]. The OTUs were filtered with filter\_otus\_from\_otu\_table.py (-n 2 option), resulting in a total of 472 OTUs for site 1 and 329 OTUs for site 2. The taxonomic composition of the samples was visualized with summarize\_taxa\_through\_plots.py (default options). The beta diversity metrics of samples from the two sites were compared by normalizing the OTU tables with normalize\_table.py (default options), and then running beta\_diversity.py (-m unweighted\_unifrac, weighted\_unifrac). The sample dissimilarity matrices were visualized on PCoA plots with principal\_coordinates.py (default parameters) and clustered heat maps with clustermap in Seaborn v0.8 [32] (method=‘average’, metric=‘correlation’). Group significance was determined with compare\_categories.py (--method=permanova). Relative similarity between metadata categories (harvest dates) was calculated with the make\_distance\_boxplots.py statistical package, which summarized the distances between pairs of sample groups (from Weighted or Unweighted Unifrac dissimilarity matrices), and then performed a two-sided Student's two-sample t-test to evaluate the significance of differences between the distances. Relative abundance of phyla and domain taxa were computed from the sum of abundances of OTUs with their respective taxonomy, and group significance calculated with a two-sided Student's two-sample t-test.

**WMG sequence processing**

The de-multiplexed WMG sequencing reads were processed with the complete metaWRAP v0.8.2 pipeline [25] with recommended databases on a UNIX cluster with 48 cores and 1024GB of RAM available. Read trimming and human contamination removal was done by the metaWRAP Read\_qc module (default parameters) on each separate sample. The taxonomic profiling was done on the trimmed reads with the metaWRAP Kraken module [33] (default parameters, standard KRAKEN database, 2017). The reads from all samples from the 3 sampling sites were individually assembled (for *pI* calculations) and co-assembled (for all other analysis) with the metaWRAP Assembly module (--use-metastades option) [34]. For improved assembly and binning of low-abundance organisms, reads from all samples were co-assembled, then binned with the metaWRAP Binning module (--maxbin2 --concoct --metabat2 options) while using all the available samples for differential coverage information. The resulting bins were then consolidated into a final bin set with metaWRAP’s Bin\_refinement module (-c 70 -x 5 options). The bins and the contig taxonomy were then visualized with the Blobology [35] module (--bins option specified), classified with the Classify\_bins module (default parameters), and quantified by Salmon [36] with the Quant\_bins module (default parameters). Contig read depth was estimated for each sample with the metaWARP’s Quant\_bins module, and the weighted contig abundance calculated by multiplying the contig’s depth by its length, and standardizing to the total contig abundance in each replicate.

**Functional annotation**

Gene prediction and functional annotation of the co-assembly was done with the JGI Integrated Microbial Genomes & Microbiomes (IMG) [37] annotation service. Gene abundances were calculated as the depths of the contigs carrying those genes. KEGG KO identifiers were linked to their respective functions using the KEGG BRITE pathway classification [38]. KEGG pathway abundances were calculated as the sum of depths of genes (estimated from the depths of the contigs carrying them) classified to be part of the pathway.

**Isoelectric point (*pI*) analysis**

The average isoelectric points of gene pools were calculated from individual replicate metagenomic assemblies. Open reading frames (ORFs) were predicted by PRODIGAL [39] with the use of metaWRAP [25], and the *pI* of each ORF was calculate with ProPAS [40]. The average *pI* of the entire gene pool as well as individual taxa were calculated from the average *pI* of proteins encoded on contigs of relevant (KRAKEN) taxonomy.

**Taxonomic rearrangement index (*RI*)**

The rearrangement indexes of gene functions represent the changes in co-assembled contigs carrying them. To calculate the *RI*, all contigs carrying genes of a given KEGG KO were identified, and the change in their read depths was calculated between two time-points of interest. Finally, *RI* for each KEGG KO identifier was calculate from the weighted average of the absolute values of these changes (Equation 1). Contig depths for entire time points were taken to be the sum of the contig depths in individual replicates.

***Equation 1:*** *Formula calculating one function’s rearrangement index RI, where T1 and T2 are standardized abundances of a contig carrying that function in two samples, and N is the number of contigs carrying that function.*

**WMG statistical analysis**

The significance in abundance changes of gene functions (i.e. KEGG KO identifiers), functional pathways (i.e. KEGG BRITE identifiers), and average *pI* of gene pools were estimated with a two-sided Student’s two-sample t-test. The relative similarity between groups of replicates (ordered by harvest dates) in terms of total pathway abundances (Figure 1C) and co-assembly contig abundances (Figure 2C) were computed by comparing Pearson correlations between samples. A Pearson correlation coefficient distance matrix was computed from all replicates, and a two-sided Student’s two-sample t-test was performed to evaluate the significance of the difference between the correlation distances. Differentially abundant KEGG (level 2) pathways were selected with a one-way ANOVA test (pval<0.01, FDR<1%), and hierarchically clustered with Seaborn v0.8 [32] (method=’average’, metric=’euclidean’). The significance of the differences in distributions of *RIs* between pairs of time points, as well as differences in *pI* distributions of gene pool proteins were calculated with the Kolmogorov-Smirnov 2-sample test. Significance of MAG abundance, contig abundance, and pathway abundance clustering was determined with SigClust (nsim=1000, icovest=3) [41]. For time considerations, the contig clustering test was limited to contigs over 5kbp in length, which were then subsampled randomly to 5000 contigs prior to the test.

**FIGURE LEGENDS**

**Figure 1:** Taxonomic composition and functional potential differences between halite samples harvested at different dates. (A) Heat map and hierarchical clustering (correlation metric) of a Weighted Unifrac dissimilarity matrix comparing taxonomic composition based on 16S rDNA sequences clustered into OTUs at 97% identity. (B) Average relative abundance of Archaea sequences in 16S rDNA sequences (significance calculated with two-sided t-test). (C) PCA of the microbial community functional potential based on the abundance of KEGG functions (1st level) inferred from WMG co-assembly quantitation. (D) Hierarchical clustering (Euclidean metric) and abundances of select KEGG pathways (1st level) that are differentially present between time points (ANOVA pval<0.01), standardized to the maximum value in each row (FDR=<1%). Bars represent group signifficance based on a two tail t-test, and stars denote the p-value thresholds (\*=0.01, \*\*=0.001, \*\*\*=0.0001).

**Figure 2:** Changes in fine-scale composition in halite samples harvested at different dates. (A) Heat map and hierarchical clustering (correlation metric) of an Unweighted Unifrac dissimilarity matrix comparing taxonomic composition based on 16S rDNA sequences clustered into OTUs at 97% identity. (B) Hierarchical clustering (Euclidean metric) of standardized MAG abundances using metaWRAP’s quant\_bins module. (C) PCA of standardized abundances of co-assembly contigs in different samples. (D) Weighted distributions of function Rearrangement Indexes of gene functions between pairs of time points, averaged between replicates.

**Figure 3:** Cartoon model of a microbial community (A) in response to (B) gradually changing environmental conditions through changes in the relative abundance of major taxa (i.e. domain or phylum); and (C) a major sudden perturbation, through changes in relative abundances of major taxa as well as rearrangement in the minor taxa (i.e. species or genus) that constitute them. Both taxonomic shifts result in a similar change in the community’s functional potential (represented to the right of each community).

**Figure S1:** The taxonomic composition differences between halite samples harvested from Site 1 at different dates, infered from 16S rDNA sequences clustered into OTUs at 97% identity and visualized through (A-D) relative abundance of major differentially abundant phyla and a (E) PCA of a Weighted Unifrac dissimilarity matrix comparing taxonomic composition based on. Bars represent group signifficance based on a two tail t-test, and stars denote the p-value thresholds (\*=0.01, \*\*=0.001, \*\*\*=0.0001).

**Figure S2:** Average taxonomic composition of halite microbial communities from Site 1 sampled at different dates, estimated from WMG reads with KRAKEN and visualized with KronaTools.

**Figure S3:** The taxonomic composition differences between halite samples harvested from Site 2 at different dates post-rain, infered from 16S rDNA sequences clustered into OTUs at 97% identity and visualized through (A-D) relative abundance of major differentially abundant phyla and (E) Archaea abundance. Bars represent group signifficance based on a two tail t-test, and stars denote the p-value thresholds (\*=0.01, \*\*=0.001, \*\*\*=0.0001).

**Figure S4:** Abundance of KEGG pathways (1st level) that are differentially abundant between time points in (A) the entire community, (B) only Archaeal contigs, and (C) only Bacterial contigs. Differential abundance of pathways estimated from quantitation of the WMG co-assembly, and significance based on two-sided ANOVA (pval<0.001, FDR<1%).

**Figure S5:** Analysis of the isoelectric points (*pI*) of proteins encoded in individual replicate WMG assemblies of samples harvested at different dates, showing (A) the overall weighted distribution of the protein *pIs*, and the weighted average *pI* of proteins encoded (B) all contigs and (E) only *Halobacteria* contigs. (D) *pI* distribution of proteins encoded on Bacteroidetes and Halobacteria contigs. Average potassium uptake potential across time point samples, inferred from Trk gene abundance and quantified in (C) all contigs and (F) only *Halobacteria* contigs.

**Figure S6:** Hierarchical clustering (Euclidean metric) of abundances of 5kbp+ contigs in the WMG co-assembly, quantified with reads from samples harvested at different dates, displayed on (A) a log scale and (B) standardized to the maximum abundance of each contig.

**Figure S7:** (B) Hierarchical clustering (Euclidean metric) of photosynthetic MAG abundances, quantified with metaWRAP’s quant\_bins module, showing the emergence of two new Cyanobacteria MAGs after the rain.

**BIBLIOGRAPHY**

1. Raymond F, Deraspe M, Boissinot M, Bergeron MG, Corbeil J: **Partial recovery of microbiomes after antibiotic treatment**. *Gut Microbes* 2016, **7**(5):428-434.

2. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA *et al*: **Diet rapidly and reproducibly alters the human gut microbiome**. *Nature* 2014, **505**(7484):559-563.

3. Goldford JE, Lu N, Bajic D, Estrela S, Tikhonov M, Sanchez-Gorostiaga A, Segre D, Mehta P, Sanchez A: **Emergent Simplicity in Microbial Community Assembly**. *bioRxiv* 2017, 10.1101/205831.

4. Eng A, Borenstein E: **Taxa-function robustness in microbial communities**. *Microbiome* 2018, **6**(1):45.

5. Louca S, Jacques SMS, Pires APF, Leal JS, Srivastava DS, Parfrey LW, Farjalla VF, Doebeli M: **High taxonomic variability despite stable functional structure across microbial communities**. *Nat Ecol Evol* 2016, **1**(1):15.

6. Nie Y, Zhao JY, Tang YQ, Guo P, Yang Y, Wu XL, Zhao F: **Species Divergence vs. Functional Convergence Characterizes Crude Oil Microbial Community Assembly**. *Front Microbiol* 2016, **7**:1254.

7. Jurburg SD, Nunes I, Brejnrod A, Jacquiod S, Prieme A, Sorensen SJ, Van Elsas JD, Salles JF: **Legacy Effects on the Recovery of Soil Bacterial Communities from Extreme Temperature Perturbation**. *Front Microbiol* 2017, **8**:1832.

8. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R: **Diversity, stability and resilience of the human gut microbiota**. *Nature* 2012, **489**(7415):220-230.

9. Haro-Moreno JM, Lopez-Perez M, de la Torre JR, Picazo A, Camacho A, Rodriguez-Valera F: **Fine metagenomic profile of the Mediterranean stratified and mixed water columns revealed by assembly and recruitment**. *Microbiome* 2018, **6**(1):128.

10. Bozkurt D, Rondanelli R, Garreaud R, Arriagada A: **Impact of Warmer Eastern Tropical Pacific SST on the March 2015 Atacama Floods**. *Monthly Weather Review* 2016, **144**(11):4441-4460.

11. Azua-Bustos A, Gonzalez-Silva C, Arenas-Fajardo C, Vicuna R: **Extreme environments as potential drivers of convergent evolution by exaptation: the Atacama Desert Coastal Range case**. *Front Microbiol* 2012, **3**:426.

12. Meslier V, Casero MC, Dailey M, Wierzchos J, Ascaso C, Artieda O, McCullough PR, DiRuggiero J: **Fundamental drivers for endolithic microbial community assemblies in the hyperarid Atacama Desert**. *Environ Microbiol* 2018, **20**(5):1765-1781.

13. Robinson CK, Wierzchos J, Black C, Crits-Christoph A, Ma B, Ravel J, Ascaso C, Artieda O, Valea S, Roldan M *et al*: **Microbial diversity and the presence of algae in halite endolithic communities are correlated to atmospheric moisture in the hyper-arid zone of the Atacama Desert**. *Environ Microbiol* 2015, **17**:299-315.

14. Crits-Christoph A, Gelsinger DR, Ma B, Wierzchos J, Ravel J, Davila A, Casero MC, DiRuggiero J: **Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community**. *Environ Microbiol* 2016, **18**(6):2064-2077.

15. Finstad KM, Probst AJ, Thomas BC, Andersen GL, Demergasso C, Echeverria A, Amundson RG, Banfield JF: **Microbial Community Structure and the Persistence of Cyanobacterial Populations in Salt Crusts of the Hyperarid Atacama Desert from Genome-Resolved Metagenomics**. *Front Microbiol* 2017, **8**:1435.

16. Davila AF, Hawes I, Araya JG, Gelsinger DR, DiRuggiero J, Ascaso C, Osano A, Wierzchos J: **In situ metabolism in halite endolithic microbial communities of the hyperarid Atacama Desert**. *Front Microbiol* 2015, **6**:1035.

17. Mongodin EF, Nelson KE, Daugherty S, DeBoy RT, Wister J, Khouri H, Weidman J, Walsh DA, Papke RT, Sanchez Perez G *et al*: **The genome of Salinibacter ruber: Convergence and gene exchange among hyperhalophilic bacteria and archaea**. *PNAS* 2005, <http://www.pnas.org/cgi/content/abstract/0509073102v1:0509073102>.

18. Monard C, Gantner S, Bertilsson S, Hallin S, Stenlid J: **Habitat generalists and specialists in microbial communities across a terrestrial-freshwater gradient**. *Sci Rep* 2016, **6**:37719.

19. Oren A: **Bioenergetic aspects of halophilism**. *Microbiol Mol Biol Rev* 1999, **63**(2):334-348.

20. Oren A: **Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes**. *Front Microbiol* 2013, **4**:315.

21. Paul S, Bag SK, Das S, Harvill ET, Dutta C: **Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes**. *Genome Biol* 2008, **9**(4):R70.

22. Mongodin EF, Nelson KE, Daugherty S, Deboy RT, Wister J, Khouri H, Weidman J, Walsh DA, Papke RT, Sanchez Perez G *et al*: **The genome of Salinibacter ruber: convergence and gene exchange among hyperhalophilic bacteria and archaea**. *Proc Natl Acad Sci U S A* 2005, **102**(50):18147-18152.

23. **Servicios Climáticos** [<https://climatologia.meteochile.gob.cl/application/index/productos/RE2009>]

24. N. S, P. BJ, P. A: **Climate change along the arid coast of northern Chile**. *International Journal of Climatology* 2012, **32**(12):1803-1814.

25. Uritskiy GV, DiRuggiero J, Taylor J: **MetaWRAP - a flexible pipeline for genome-resolved metagenomic data analysis**. *bioRxiv* 2018, 10.1101/277442.

26. Thombre RS, Shinde VD, Oke RS, Dhar SK, Shouche YS: **Biology and survival of extremely halophilic archaeon Haloarcula marismortui RR12 isolated from Mumbai salterns, India in response to salinity stress**. *Sci Rep* 2016, **6**:25642.

27. Davila AF, Hawes I, Ascaso C, Wierzchos J: **Salt deliquescence drives photosynthesis in the hyperarid Atacama Desert**. *Environ Microbial* 2013, (DOI: 10.1111/1758-2229.12050).

28. Locey KJ, Fisk MC, Lennon JT: **Microscale Insight into Microbial Seed Banks**. *Frontiers in Microbiology* 2017, **7**:2040.

29. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI *et al*: **QIIME allows analysis of high-throughput community sequencing data**. *Nat Methods* 2010, **7**(5):335-336.

30. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO: **The SILVA ribosomal RNA gene database project: improved data processing and web-based tools**. *Nucleic Acids Res* 2013, **41**(Database issue):D590-596.

31. Edgar RC: **Search and clustering orders of magnitude faster than BLAST**. *Bioinformatics* 2010, **26**:2460-2461.

32. Waskom M, Botvinnik O, O'Kane D, Hobson P, Lukauskas S, Gemperline DC, Augspurger T, Halchenko Y, Cole JB, Warmenhoven J *et al*: **Seaborn**. In*.*, 10.5281/zenodo.883859, 0.8.1 edn: GitHub; 2017: <https://github.com/mwaskom/seaborn>.

33. Wood DE, Salzberg SL: **Kraken: ultrafast metagenomic sequence classification using exact alignments**. *Genome Biol* 2014, **15**(3):R46.

34. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA: **metaSPAdes: a new versatile metagenomic assembler**. *Genome Res* 2017, **27**(5):824-834.

35. Kumar S, Jones M, Koutsovoulos G, Clarke M, Blaxter M: **Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots**. *Frontiers in Genetics* 2013, **4**:237.

36. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C: **Salmon provides fast and bias-aware quantification of transcript expression**. *Nat Methods* 2017, **14**(4):417-419.

37. Chen IA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, Ratner A, Huang J, Andersen E, Huntemann M *et al*: **IMG/M: integrated genome and metagenome comparative data analysis system**. *Nucleic Acids Res* 2017, **45**(D1):D507-D516.

38. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M: **KEGG as a reference resource for gene and protein annotation**. *Nucleic Acids Res* 2016, **44**(D1):D457-462.

39. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ: **Prodigal: prokaryotic gene recognition and translation initiation site identification**. *BMC Bioinformatics* 2010, **11**:119.

40. Wu S, Zhu Y: **ProPAS: standalone software to analyze protein properties**. *Bioinformation* 2012, **8**(3):167-169.

41. Liu Y, Hayes DN, Nobel A, Marron JS: **Statistical Significance of Clustering for High-Dimension, Low-Sample Size Data**. *Journal of the American Statistical Association* 2008, **103**(483):1281-1293.