**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]. Functional adaptations to such changes are often accomplished through adjustments in relative abundance of higher-order taxa (i.e., phylum, order), however 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].

Understanding adaptation strategies of environmental microbial communities is critical for learning more about diversification of microbial life, and the evolution of more translational microbiomes following stress. 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 more difficult under natural environmental conditions. While previous longitudinal studies addressed adaptations of environmental microbiomes in response to long-term environmental changes including humidity and temperature [9, 10], specifics of adaptation strategies to more acute changes remain largely unexplored, particularly in extreme environments.

One such extreme environment is the northern Atacama Desert, which is one of the harshest places on Earth with an average annual precipitation of less than 1mm [11, 12]. Despite this, poly-extremophilic microbiota have evolved to survive in these extreme conditions by relying on the protection of various mineral substrates [13]. One such endolithic (inside rock) community type is harbored within halite – salt rock – nodules found in the Neogene evaporitic basin of Salar Grande [14, 15]. 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 with the outside, and have limited nutrient input beyond atmospheric water and gasses [14-16]. As such, each halite nodule represents a near-closed system, allowing the tracking of 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 atmospheric moisture [15-17]. While halite microbiomes are primarily comprised of *Archaea*, they rely on carbon fixation by halophilic algae and cyanobacteria, which support a number of heterotrophs [15, 16]. In particular, the majority of the biomass comes from *Halobacteria* and *Bacteroidetes* – two taxonomically diverse groups of extreme halophilic salt-in strategists that accumulate potassium ions to match the external osmotic pressure from sodium ions [15, 18, 19]. This allows them to survive in extremely high-salt environment, but restricts their fitness to a narrow range of external salt concentration [20].

The highly specialized nature of the halite microbial communities can make them more vulnerable to change compared to habitat generalists [19], particularly to sudden changes in external osmotic pressure from increased water availability. In August 2015, Northern Atacama received its first major rain in 13 years [11, 21]. Combined with the isolated nature of halite microbiomes, this major climatic event provided the opportunity to track the response of an environmental microbiome to a major natural perturbation. Our longitudinal study over 4 years not only captured the microbiome’s short-term adaptations to the perturbation, but also its recovery in the subsequent years, 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 [11]. 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) [21, 22].

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. This yielded 535,233 paired-end reads (insert size 419±7bp), which were used for taxonomic profiling of the microbiomes (Table S1). For 5 biological replicates at each of the 4 time points, whole-metagenomic (WMG) sequencing was performed 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 shifted significantly following the rain**

Following the rain, the halite community structure shifted from an *Archaea*-dominated community in 2014-09 and 2015-06 to a more balanced *Archaea-Bacteria* community in 2016-02 (6-months post-rain). The relative abundance of *Archaea* dropped significantly (two-sided t-tests: *p*<0.0001) in both 16S rDNA (Figure 1B) and WMG sequencing (Figure S2). We found that many Phyla also shifted in abundance: *Cyanobacteria*, green algae (estimated by chloroplast rDNA abundance), and *Bacteroidetes* significantly increased in relative abundance following the rain, while the abundance of *Halobacteria* (the major archaea phylum in this community) significantly decreased (Figures S1A-D, S2, two-sided t-tests: *p*<0.01).

Analysis of the Weighted Unifrac dissimilarity matrix (constructed from OTUs clustered at 97% identity) confirmed that the overall structure of the halite communities was different between time-points (PERMANOVA: *p*<0.001), with the halite microbial community shifting in composition following the rain (Figure S1E). Indeed, samples from before the rain (2014-09 and 2015-06) clustered together, while 2016-02 samples were significantly more distant (two-sided t-test: *p*<0.0001) (Figure 1A).

The functional potential of the community, determined by 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, the functional potential of 2014-09 samples were better correlated with that of 2015-06 samples than 2016-02 samples (two-sided t-test of Pearson correlations: *p*<0.0001). While the majority of functional pathways were present in similar abundances between replicates and time points, a number of pathways were differentially represented between time points (ANOVA test, *p*<0.01, FDR<1%) (Fig 1D). Of these, the majority was significantly over- or under-represented in the samples collected shortly after the rain (2016-02) (SigClust 2-group significance: *p*<0.0001), Major functional categories such as cell motility, lipid metabolism, and carbohydrate metabolism increased in relative abundance after the rain, while translation, transcription, and nitrogen metabolism decreased in abundance, indicating active adaptations to the post-rain internal halite environment.

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

The majority of the halite community is comprised of salt-in strategists (*Halobacteria* and *Bacteroidetes*), which actively import potassium ions, producing an internal osmotic pressure [15, 18, 19]. While energetically favorable compared to producing compatible solutes [23], this requires the cells’ proteomes to have an extremely low isoelectric point (pI) to be able to function at high potassium concentrations [24-26]. We found that the *pI* of proteins encoded in community gene pool shifted significantly after the rain, favoring higher *pI* composition (Kolmogorov-Smirnov 2-sample test: *p*<0.0001) (Figure S5A). Due to significantly different *pI* distributions in the proteomes of *Halobacteria* (*pI*=5.04) and *Bacteroidetes* (*pI*=5.80) (Kolmogorov-Smirnov 2-sample test: *p*<0.0001) (Figure S5D), the shift in their relative abundances resulted in the average *pI* of the community to significantly increase after the rain (two-sided t-test: *p*<0.01) (Figure S5B). Consistent with salt-in adaptations, we also found that the average potassium uptake potential (estimated from Trk gene abundances) of the communities significantly decreased after the rain (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 within the halite during the rains.

**Individual strain composition of the communities was rearranged after the rain**

The halite communities harvested at different dates were also different in terms of presence or absence of OTUs (unweighted Unifrac) (PERMANOVA: *p*<0.001). We found that samples harvested shortly after the rain (2016-02) were more distant from pre-rain samples than the pre-rain samples were from each other (two-sided t-test: *p*<0.0001) (Figure 2A). This fine-scale composition of the halite microbiomes was also investigated through metagenome-assembled genomes (MAGs). With the use of metaWRAP [27], 91 high-quality MAGs (>70% completion, <5% contamination) were recovered from the WMG sequencing data (Table S2), and their abundances were tracked over time. Hierarchical clustering of the samples based on MAG abundances confirmed a significant shift in individual strain composition after the rain (SigClust 2-group significance: *p*<0.05) (Figure 2B). Additionally, two new *Cyanobacteria* MAGs – *Synechococcaceae* and *Chroococcales* – that were previously reported in only a small fraction of halite nodules[16], were found in high abundances in most of the samples after the rain (Figure S7). Strikingly, we did not find a correlation between the functional potentials of the MAGs and their survival after the rain, suggesting 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 contigs 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: *p*<0.0001), indicating that the same community functions were being performed by a new set of organisms.

**The community higher-order taxonomic structure and functional potential are resilient long-term**

Despite major shifts after the rain (2016-02), we found that the higher-order taxonomic structure of the halite communities recovered in the following year (2017-02). The relative abundance of Archaea and the major phyla shifted significantly again (two-sided t-tests: *p*<0.0001), partially or completely returning to the pre-rain levels (Figures 1B, S1A-D). This trend in domain and phyla abundance recovery was also observed in the sequencing of the supplementary site (Site 2), with incremental shifts over 18 months after the rain. (Figure S3). The overall taxonomic structure as seen in the Weighted Unifrac dissimilarity matrix also confirmed this recovery, as 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: *p*<0.0001) (Figure 1A).

Finally, the functional potential was also restored during this period (Figure 1C). When looking at pathways differentially represented between time points (ANOVA test, *p*<0.01, FDR<1%), we found that 2017-02 samples correlated significantly better with the pre-rain samples than 2016-02 samples did (two-sided t-test: *p*<0.0001), indicating that most pathways that changed in relative abundance in response to the rain recovered to their normal levels in the following year (Figure 1D).

**The strain-level community composition and niche membership were permanently altered by the perturbation**

When considering presence or absence of OTUs with an Unweighted Unifrac dissimilarity matrix (Figure 2A), we found that 2017-02 samples were less distant to 2016-02 samples than the pre-rain samples (two-sided t-test: *p*<0.0001), suggesting that the fine-scale (strain) community composition did not return to its initial state after the perturbation. Hierarchical clustering of the MAG abundances also confirmed that 2017-02 samples formed a separate cluster together with 2016-02 samples and away from the pre-rain samples (SigClust 2-group significance: *p*<0.05) (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. At the contig scale, Pearson correlation comparison (two-sided t-test: *p*<0.0001) as well as group significance analysis (SigClust 2-group significance: *p*<0.01) of the 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). Taken together with the resilience of the higher-order community structure, these results indicate 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.

When looking at strains contributing to KEGG functional pathways, we found that the rearrangement of functional niche membership during the recovery period (=0.32+/-0.04) was significantly lower than right after the rain event (=0.46±0.07) (Figure 2D). Additionally, the niche membership in the recovered communities was still very distinct from pre-rain samples (= 0.43+/-0.09) ) (Kolmogorov-Smirnov 2-sample tests: *p*<0.0001). Taken together with the recovery of the overall functional potential, this could mean that while the 2017 halite communities function the same as they did prior to the rain, their functional niches are constituted by a new set of microbial strains and species.

**DISCUSSION**

Consistent with previous desert microbiome studies [9], we found that desert extremophiles are highly sensitive to environmental climate changes, particularly to changes in water availability. The 2015 rain, which was the first major precipitation in Northern Atacama in 13 years, induced a drastic response in functional potential and taxonomic structure of the poly-extremophilic halite microbiomes. It is remarkable that we observed these changes in samples harvested 6 months after the rain, as a similar study of desert soils reported a recovery in higher-order taxonomy structure just one-month post-rain [9]. This suggests that the immediate effects of the rain may have been more drastic, and reveals the slow-growing nature of these microbiomes, which likely results from scarce resources and harsh climatic conditions [15, 28].

We observed an increase in average isoelectric point (*pI*) of proteins encoded in the community’s gene pool after the rain, and a decrease in potassium uptake potential, represented by overall abundance of *Trk* genes. Considering that low proteome *pI* and high potassium uptake are the hallmarks of halophilic salt-in strategists [24-26], we suggest that the rain temporarily decreased the salt concentrations within the colonized pores of the halite nodules [29], rapidly changing the osmotic conditions within. Extreme halophiles such as those residing within halite nodules have a narrow range of tolerated salt concentrations [20], which likely rendered this specialized microbiome more sensitive to the rain [23]. This could have led to a mass death event of poorly adapted organisms immediately following the rain, while giving others a significant advantage. In particular, organisms with a higher *pI* of their proteomes (*Bacteroidetes*), which were better adapted to the osmotic conditions following the rain, significantly increased in relative abundance.

The shift in the halite microbiome’s functional potential after the rain was achieved though composition adjustments at the higher-order taxonomic ranks (i.e. domains and phyla), as well as a complete rearrangement of the finer taxa (i.e. strains) that constitute them. We infer that the rain caused a significant perturbation that resulted in gaps in functional niches that new organisms opportunistically took over. Despite this major compositional shift however, the community was able to gradually recover functionally and taxonomically (at domain and phylum levels) over the course of 18 months. Unlike the original shift, the community adjusted its functional potential almost exclusively via changes in relative abundance of higher-order taxonomic groups, the fine-scale composition of which remained largely unchanged throughout the recovery. The recovered community is to be functionally equivalent to the pre-rain community, while being comprised of a new set of individual organisms carrying the same functions as at pre-rain. These results are conceptually similar to those observed in gut microbiome studies, where microbial communities are able to recover functionally following antibiotic administration [30], but with some loss of former taxonomic diversity and acquisition of new species through niche intrusion [31].

While the composition changes at the domain and phylum level resulted from functional adaptations, the stochastic rearrangement of individual strains contributing to functional pathways was likely driven by processes similar to those governing original colonization of halite nodules. Individual nodule display a great degree of inter-nodule taxonomic diversity at the strain level [16], suggesting that halite colonization is at least partially driven by neutral (i.e. random) processes [32]. Each niche is stochastically colonized by competitively equivalent organisms by random draw from the seed bank [33] – a diverse genetic reservoir consisting of a large collection of low-abundance organisms. Seed banks are critical in microbiomes as they conserve genetic and functional diversity during prolonged unchanging environmental conditions (such as the past 13 years prior to the rain in Northern Atacama), which allows for rapid adaptation and restructuring of microbial communities following drastic perturbations [34], as seen following the rain.

The functional redundancy of community members ensured a robust functional landscape in the halite microbial communities despite a turn-over of individual strains after the rain [3, 4]. This functional consistency despite taxonomic variance has been documented in a variety of microbiomes [4-6, 8]. In particular, isolated microbiomes such as miniature aquatic ecosystems found in bromeliad rosettes (similarly isolated like the halite nodules) appear to converge on an identical functional landscapes given similar environmental conditions despite great inter-community taxonomic diversity. The convergence of community functional potential occurs through mechanisms such as stoichiometric balancing between metabolic pathways, which are largely decoupled from taxonomic lineage, particularly at the strain level, resulting in divergent taxonomic composition [5]. While the halite nodules support a taxonomically diverse arrays of *Bacteroidetes* and *Halobacteria* strains, they are largely functionally redundant, which allowed the consistent functional potential landscape following the rearrangement.

In this study, 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. Mechanistically however, the two shifts were fundamentally distinct. The first shift caused functional rearrangement at the community level not only by means of changing relative abundances of major taxa (i.e domains, phyla), but also within finer taxonomic ranks (i.e strains) (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 vivid example of the de-coupling of function and taxonomy in microbial communities allowed for inference of two contrasting models by which a given microbial community can respond to changing environmental conditions, resulting in a similar functional potential (Figure 3A). The first type of shift is a community rearrangement, resulting from adaptations to a sudden major perturbation, which creates gaps in existing functional niches, presenting an opportunity for new organism from the seed bank to come in through niche intrusion [31] (Figure 3C). The change in fine-scale (i.e. strains) taxonomic composition in this type of response is likely driven by neutral processes. The second type of shift is an adjustment in existing community structure, and results from gradual changes in environmental conditions (Figure 3B). The taxonomic composition in this type of response is more deterministic, as currently dominant organisms have the opportunity to adjust their relative abundances without allowing new organism to take over.

*In one sentence, what is this paper about?*

This paper is about the response and recovery of an extreme microbiome after a rare climatic perturbation, which served as an illustrative example of two possible mechanisms that a community’s structure can undergo to respond to changing environmental conditions.

*What are the main findings of this paper?*

* The halite extremophile microbiomes are hyper-sensitive to climatic perturbations, and take an unusually long time to recover.
* Major perturbation can result in a permanent turn-over in the fine-scale taxonomic composition of microbiomes, resulting in a community that functions the same as before, but is comprised of a different set of organisms.
* The two observed community shifts – the response and the recovery – have the same magnitude of functional change, but have drastically different mechanisms at the taxonomic scale.
* We propose a model to explain the two different modes of community taxonomic response to achieve a functional potential shift.

*What is the impact of this paper?*

* We observe two modes of functional response *in the same community*, demonstrating that a given microbiome can achieve the same magnitude of functional shit via different mechanisms. Our model of two adaptation strategies is potentially applicable to other microbiomes.
* This is the first study (as far as I could find) to investigate both the response and recovery of a microbiome after a perturbation with WMG sequencing, looking at the MAG distribution and functional potential (similar studies only use amplicon sequencing as proxies for functional potential).
* The taxonomic rearrangement we observed following the rain highlights the role of seedbanks and neutral processes in microbial community assembly.
* This is the first documentation of any temporal dynamics of lithic microbiomes, showing that even these communities are dynamic and responsive. [*Jocelyne: I know you want to bring in biomarker detection on Mars/etc. to give even more relevance, but I could not find a way to smoothly bring that in. It feels forced and unrelated to everything else in the MS. Unless you want to show me how, I will leave it out for now.*]
* This is the only study describing the impact of such a rare climatic event on a microbiome, which foreshadows the potential impact of climate change on microbiota in desiccated environments.

**METHODS**

**Sample collection and DNA extraction**

Halite nodules 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 S1 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 [14, 15] 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 [14] by amplifying the hypervariable V3-V4 region with 515F and 926R primers [35]. 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**

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 2000 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 [36]. 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 [37] release as reference and USEARCH v6.1.554 [38]. 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 [39] (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 [27] 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 [40] (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) [41]. 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 [42] module (--bins option specified), classified with the Classify\_bins module (default parameters), and quantified by Salmon [43] 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) [44] 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 [45]. 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 [46] with the use of metaWRAP [27], and the *pI* of each ORF was calculate with ProPAS [47]. 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 (KO IDs) 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 (*p*<0.01, FDR<1%), and hierarchically clustered with Seaborn v0.8 [39] (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) [48]. 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 *p*<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 (KO IDs) between pairs of time points, averaged between replicates.

**Figure 3:** Model of two different microbiome composition adaptation mechanisms, both of which result in an identical functional potential change (A) following a change in environmental conditions. The first functional shift (B) happens via changes in the relative abundance of major taxa already present in the community, and typically occurs in response to gradually changing environmental conditions. The second shift (C) occurs via a rearrangement of the community’s taxonomic structure, resulting in new organisms from the seed bank displacing most previously dominant taxa through niche intrusion. This type of shift follows a significant perturbation that destabilizes the existing niche membership.

**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 (*p*<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. Armstrong A, Valverde A, Ramond JB, Makhalanyane TP, Jansson JK, Hopkins DW, Aspray TJ, Seely M, Trindade MI, Cowan DA: **Temporal dynamics of hot desert microbial communities reveal structural and functional responses to water input**. *Sci Rep* 2016, **6**:34434.

10. 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.

11. 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.

12. 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.

13. Wierzchos J, Casero MC, Artieda O, Ascaso C: **Endolithic microbial habitats as refuges for life in polyextreme environment of the Atacama Desert**. *Current Opinion in Microbiology* 2018, **43**:124-131.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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>.

19. 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.

20. 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.

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

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

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

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

25. 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.

26. 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.

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

28. Ziolkowski LA, Wierzchos J, Davila AF, Slater GF: **Radiocarbon evidence of active endolithic microbial communities in the hyper-arid core of the Atacama Desert,**. *Astrobiology* 2013, **13**:607-616.

29. 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).

30. Sommer MO, Dantas G: **Antibiotics and the resistant microbiome**. *Curr Opin Microbiol* 2011, **14**(5):556-563.

31. Modi SR, Collins JJ, Relman DA: **Antibiotics and the gut microbiota**. *J Clin Invest* 2014, **124**(10):4212-4218.

32. Hubbell SP: **The Unified Neutral Theory of Biodiversity and Biogeography.** . Princeton: New Jersey: Princeton Univ Press; 2001.

33. Li L, Ma ZS: **Testing the Neutral Theory of Biodiversity with Human Microbiome Datasets**. *Sci Rep* 2016, **6**:31448.

34. Lennon JT, Jones SE: **Microbial seed banks: the ecological and evolutionary implications of dormancy**. *Nat Rev Microbiol* 2011, **9**(2):119-130.

35. Needham DM, Fuhrman JA: **Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom**. *Nat Microbiol* 2016, **1**:16005.

36. 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.

37. 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.

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

39. 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>.

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

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

42. 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.

43. 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.

44. 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.

45. 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.

46. 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.

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

48. 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.