Response of extremophile microbiome to major climatic perturbation reveals two distinct community adaptation strategies

Gherman Uritskiy, Samantha Getsin, Adam Munn,
James Taylor* and Jocelyne DiRuggiero*

INTRODUCTION

Due to the vast taxonomic and functional diversity found in microbial communities, they have an incredible ability to change, 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 can be inconsequential to its overall functioning [3]. Due to functional redundancy between taxa, the functional potential of the microbiome persists even after a major community rearrangement [4-6].

The resilience and adaptations 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 environmental factors such studies are much more difficult to perform in uncontrolled environmental conditions. Previous longitudinal studies looked at adaptations of environmental microbiomes in response to gradually changing environmental conditions [9], but adaptations of such communities to acute changes remain largely unexplored.

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. One such endolithic (inside rock) community type is harbored within halites – salt rock nodule formations found on the surface of Salar Grande. A unique property of halite microbiomes that make them a compelling system for measuring responses to environmental stressors is their isolated nature. Encased in rocks, halite communities receive very little biomass exchange, and have limited nutrient input beyond atmospheric gasses [12-14]. As such, each halite nodule represents a near-closed system, allowing us to track microbial community changes without unintended external factors compounding the results.

Because of salt's deliquescent properties, these *Archaea*-dominated microbiomes are able to survive by receiving water almost exclusively from the humidity in the air [12, 13, 15]. Due to the high-salt internal environment, the major taxonomic components of this community are *Halobacteria* and *Bacteroidetes* – two 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 [12, 16, 17]. This is energetically favorable to actively pumping out sodium [18], but required the cells' proteomes to be have an extremely low isoelectric point (pI) to be able to function at high potassium concentrations [19-21].

The highly specialized nature of the halite microbial communities can make them more vulnerable to change compared to habitat generalists [17], particularly to sudden changes in water availability. Because the nodules are primarily comprised of porous salt [22], a major rain could at temporarily alter the internal conditions and create a major osmotic shock to the microbial communities. The rain that northern Atacama received in August 2015 was its first major precipitation event 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

In order to investigate the temporal dynamics of halite microbiomes, samples 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 250bp reads (insert size 419±7bp), which were used for taxonomic profiling the microbiomes. For 5 biological replicates in each time point, whole-metagenomic (WMG) sequencing was also done for evaluating the functional potential of the communities over time, yielding a total of 70,689,467 paired-end 150bp 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

Ribosomal amplicon sequences were clustered into operational taxonomic units (OTUs) at 97% nucleotide identity to evaluate overall taxonomic structure of samples in the longitudinal study. Analysis of the Weighted Unifrac dissimilarity matrix revealed that samples harvested at different dates were significantly different (permanova test: statistic=38.69, p-value<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. Strikingly, 2016-02 samples (6 months post-rain) were significantly more distant from the pre-rain samples than this baseline (two-sided t-test: t-statistic=20.35, parametric p-value=1.23E-55), 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: t-statistic=21.17, parametric p-value=6.4E-63) (Fugure S1A).

We also found changes in the the abundances of a number of higher-order taxa following the rain. Most interestingly, the relative *Archaea* abundance in both the rDNA amplicon (Figure 1B) and WMG sequencing (Figure S2) significantly dropped following the rain (two-sided t-test: t-statistic=9.78, p-value=3.4e-10), but recovered in the following year (two-sided t-test: t-statistic=-2.14, pvalue=0.043). 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).

The functional potential of the community, determined with functional annotation of the WMG co-assembly, also changed after the rain. Principal component analysis as well as hierarchical clustering of Pearson correlation comparison matrix of functional category abundances revealed a significant shift in overall functional potential (Figure 1C). 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). Indeed, 2014-09 samples had a significantly stronger Pearson correlation with 2015-06 samples than with 2016-

02 samples (two-sided t-test: t-statistic=6.64, p-value=2.65e-08). Furthermore, the 2014-09 samples correlated significantly better with 2017-02 samples than 2016-02 samples did (two-sided t-test: t-statistic=9.17, p-value=4.06e-12), 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 that were differentially represented between time points (ANOVA test, *pval*<0.01) (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: p-value=0.000), but present in similar abundances in the pre-rain samples and the "recovered" samples. Cell motility, signaling pathways, and DNA replication and repair 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. Of these, many also changed in abundance after the rain within only Bacterial and only Archaeal contigs (Figure S4).

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

We found that the distribution of the isoelectric points (*pI*) of proteins encoded in community gene pool changed after the rain (Kolmogorov-Smirnov 2-sample test: statistic=0.061, p-value=0.0). The average *pI* significantly increased after the rain (two-sided t-test: t-statistic=3.46, p-value=0.0086), and then recovered to pre-rain levels in the the following year (two-sided t-test: t-statistic=0.086, pvalue=0.93) (Figures S5A,B). This major shift results from the *pI* differences of the two major salt-in strategist halophiles that changed in abundance after the rain – *Halobacteria* (*pI*=5.04) and *Bacteroidetes* (*pI*=5.80) (Kolmogorov-Smirnov 2-sample test: statistic=0.306, p-value=0.0) (Figure S5C). However, we also found a reduction in the average pI following the rain within the highly heterogeneous *Halobacteria* phylum (Figure S5D). Accordingly, we also found changes in the other major adaptation of salt-in strategists – potassium uptake. The average total potassium uptake potential (estimate from Trk gene abundances) of the communities significantly decreased after the rain, and then recovered to pre-rain levels in the following year (Figure S5E). This was also observed within the functional potential *Halobacteria* phylum (Figure S5F). Considering *Bacteroidetes* and *Halobacteria* have adapted to high external sodium concentrations by pumping in potassium and having low protein isoelectric points, these results suggest adaptations to a temporary decrease in salt concentrations during the rains.

Individual community members constituting functional niches were 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 also different in terms of presence or absence of OTUs (97% identity) (permanova test: statistic= 10.48, p-value<0.001). Strikingly, we found that samples harvested shortly after the rain (2016-02) not only more distant from pre-rain samples than the baseline pre-rain variance (two-sided t-test: t-statistic=17.28, parametric p-value=5.00E-45), but also clustering together with the samples collected 18 months after the rain (2017-02) (Figure 2A), contrasting the results of clustering the Weighted Unifrac dissimilarity matrix (Figure 1A). Indeed, we found that 2017-02 samples were less distant to 2016-02 samples than the pre-rain samples (two-sided t-test: t-statistic=12.87, parametric p-value=0), suggesting that the community did not recover from the perturbation in terms of the presence or absence of individual OTUs.

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, and their abundances were tracked across the longitudinal study. Despite high heterogeneity in MAG abundances, hierarchical clustering of the abundance table revealed two significantly

distinct groups of replicates (SigClust 2-group significance: p-value= 0.011) – 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: t-statistic=8.94 p-value=8.41e-12) as well as group significance analysis (SigClust 2-group significance: p-value=0.003) 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). Taken together with the resilience of the overall community structure, this result suggests 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.

To more directly investigate changes in functional niche membership, a rearrangement index (*RI*) was calculated for each KEGG function, evaluating the degree of change in taxa that carry a given function (see Methods). The distribution in the *RIs* of all functions between two time points (RI_{after}^{before}) 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 (RI_{2015}^{2014}=0.24±0.05), the rearrangement following the rain was significantly higher (RI_{2016}^{2015}=0.46±0.07) (Kolmogorov-Smirnov 2-sample test: 0.66, pvalue=0.0), indicating that the community functions were being performed by a new set of organisms. However, the rearrangement during the following year was significantly (Kolmogorov-Smirnov 2-sample test: statistic=0.55, pvalue=0.0) lower (RI_{2017}^{2016}=0.32+/-0.04), despite overall functional potential recovery (Figure 1C,D). Finally, the final communities found in 2017-02 samples were still very distinct from pre-rain samples (RI_{2017}^{2014}= 0.43+/-0.09). Taken together with the recovery in overall functional potential (Figure 1C,D), these results suggest that while the recovered halite communities function the same as they did prior to the rain, their functional niches are constituted by a new set of organisms.

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. However, there two shifts were fundamentally distinct. The first shift resulted from rapid changes in selective pressures following the rain, and required 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). The second functional potential shift occurred during the recovery phase, and was much more gradual in nature, resulting largely from changes in relative abundance of diverse *Halobacteria* and *Bacteroidetes*, which themselves remained largely unchanged.

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 functional potential 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

The highly specialized nature of halite microbiomes renders them sensitive to environmental changes, particularly to changes in water availability, which is a limiting factor for life in this desiccated environment. From adaptations in isoelectric point (*pI*) and potassium

uptake of the community's salt-in strategists, we deduce that the 2015 rain, which was the first major precipitation in 13 years, likely 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.

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. This makes sense in context of functional redundancy of closely-related community members, as previously low-abundance taxa were able to take advantage of the perturbed system and occupy existing functional niches without consequence to the functioning of the community as a whole.

Our findings show the that highly specialized extremophilic endoliths of Atacama Desert are hyper-sensitive to environmental 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 the higher-order community structure structure, and a permanent rearrangement of the individual composition of the community. 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 more gradual changes in environmental conditions. 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. 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 functional changes within individual taxa.

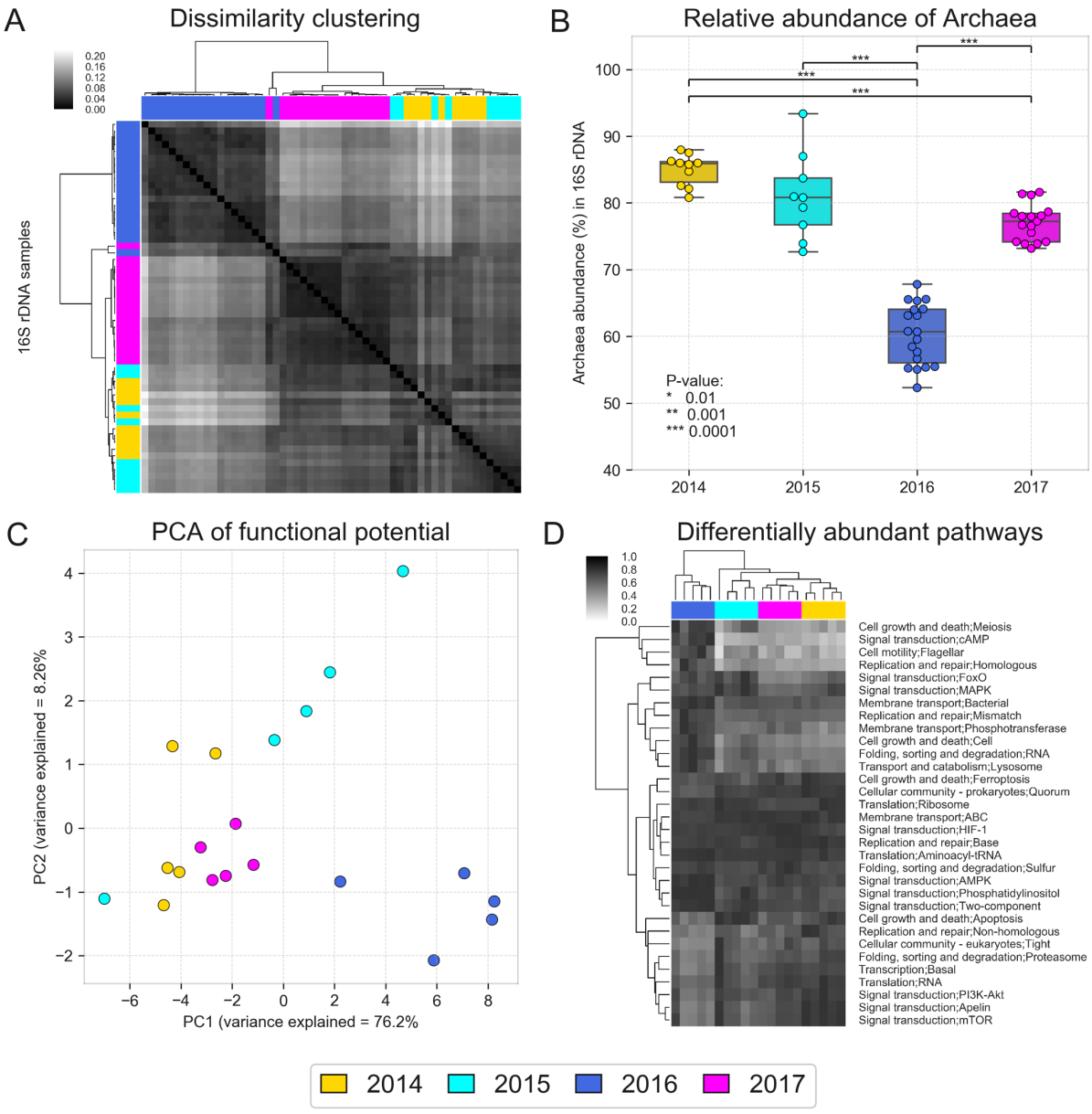


Figure 1: *The taxonomic composition and functional potential differences between halite samples harvested at different dates. (A) A heat map and hierarchical clustering (correlation metric) of a Weighted Unifrac dissimilarity matrix comparing taxonomic composition based on 16S rDNA sequences clustered 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). (D) Hierarchical clustering (Euclidean metric) and abundances of KEGG cellular, environmental and genetic processing functions (2nd level) that are differentially present between time points (ANOVA pval<0.01), standardized to the maximum value in each row.*

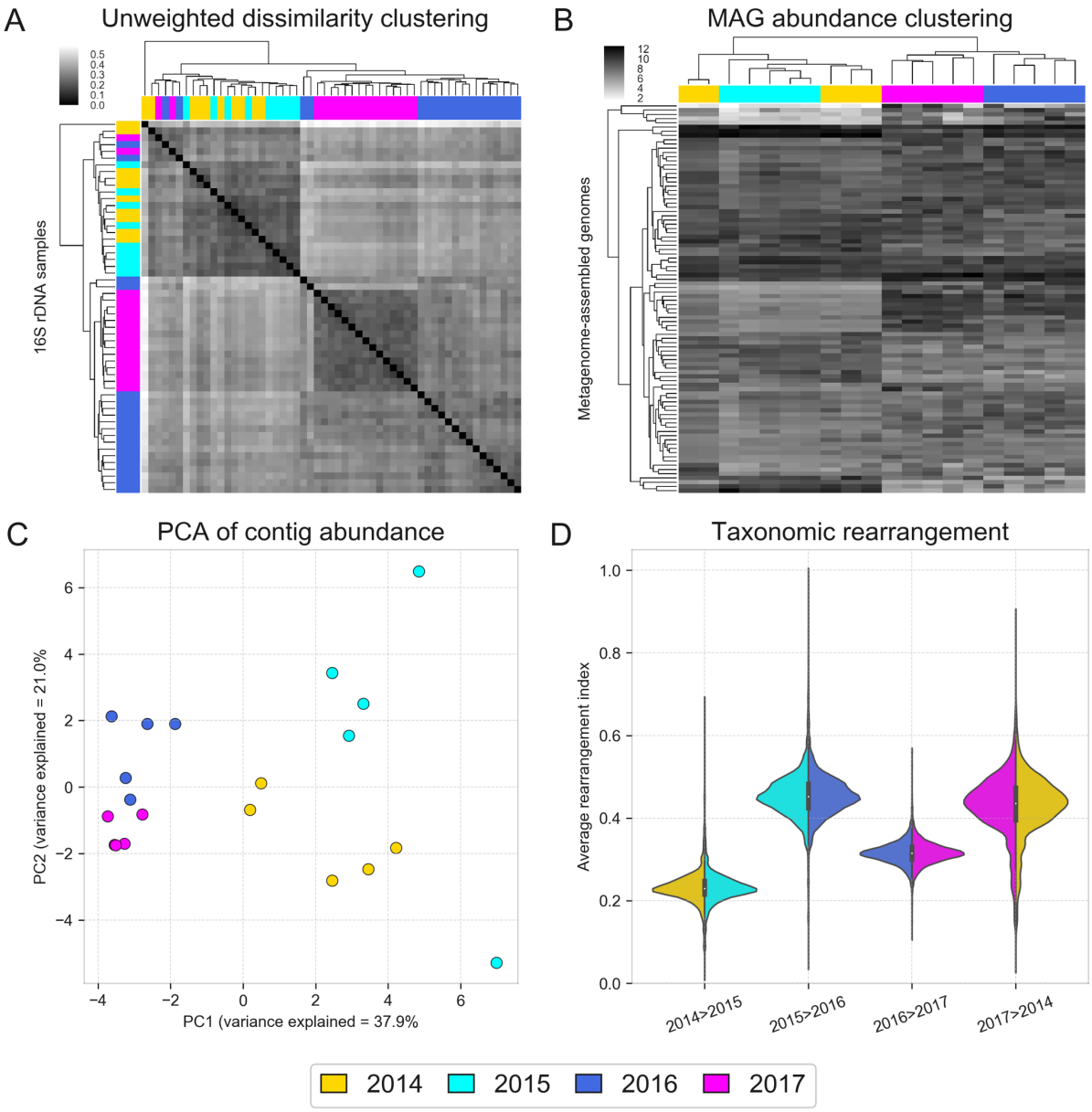


Figure 2: *Changes in fine-scale composition in halite samples harvested at different dates. (A) A heat map and hierarchical (correlation metric) clustering of an Unweighted Unifrac dissimilarity matrix comparing taxonomic composition based on 16S rDNA sequences clustered into OTUs at 97% identity. (B)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

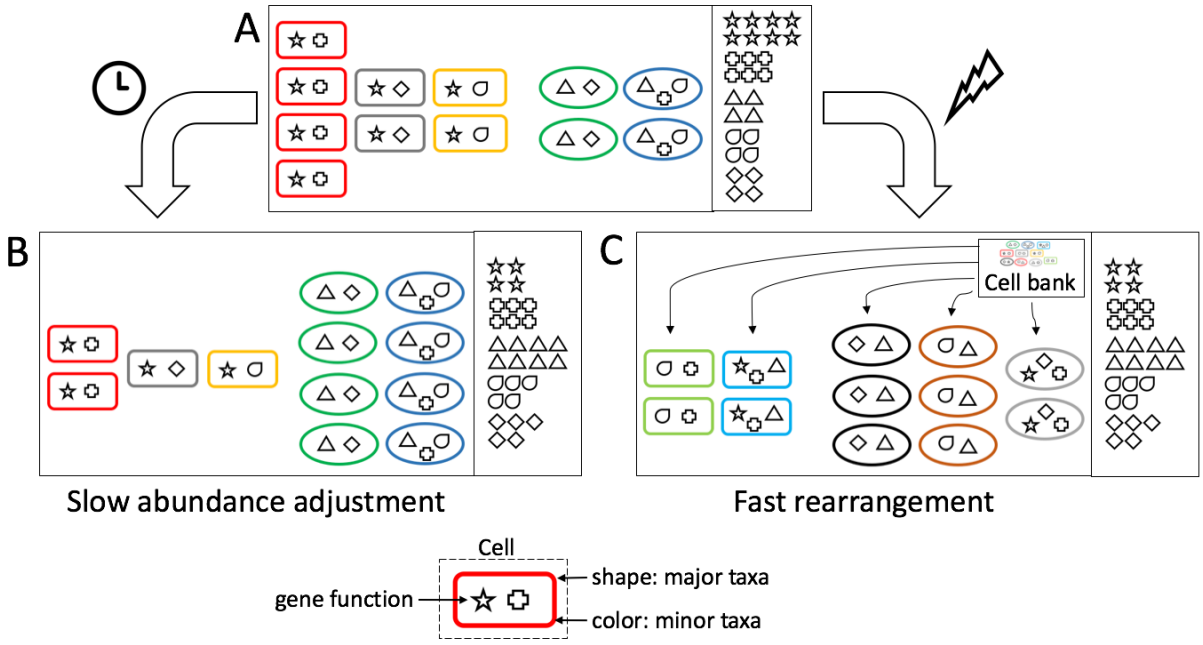
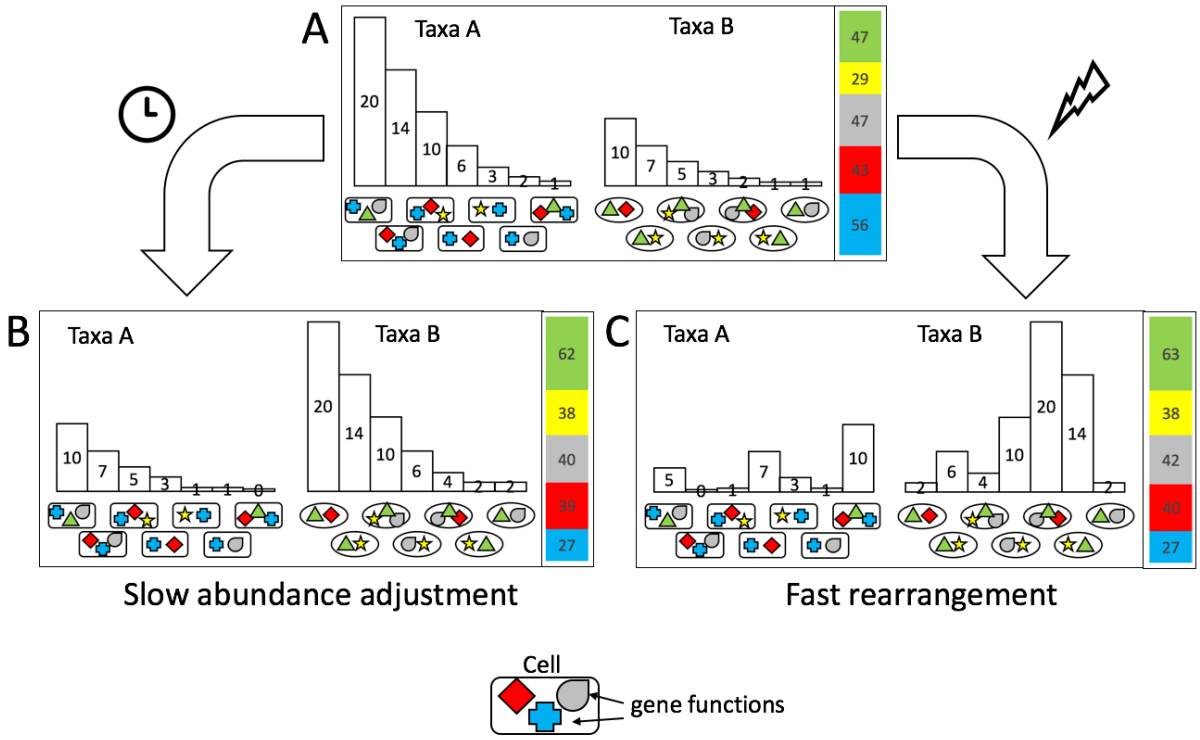
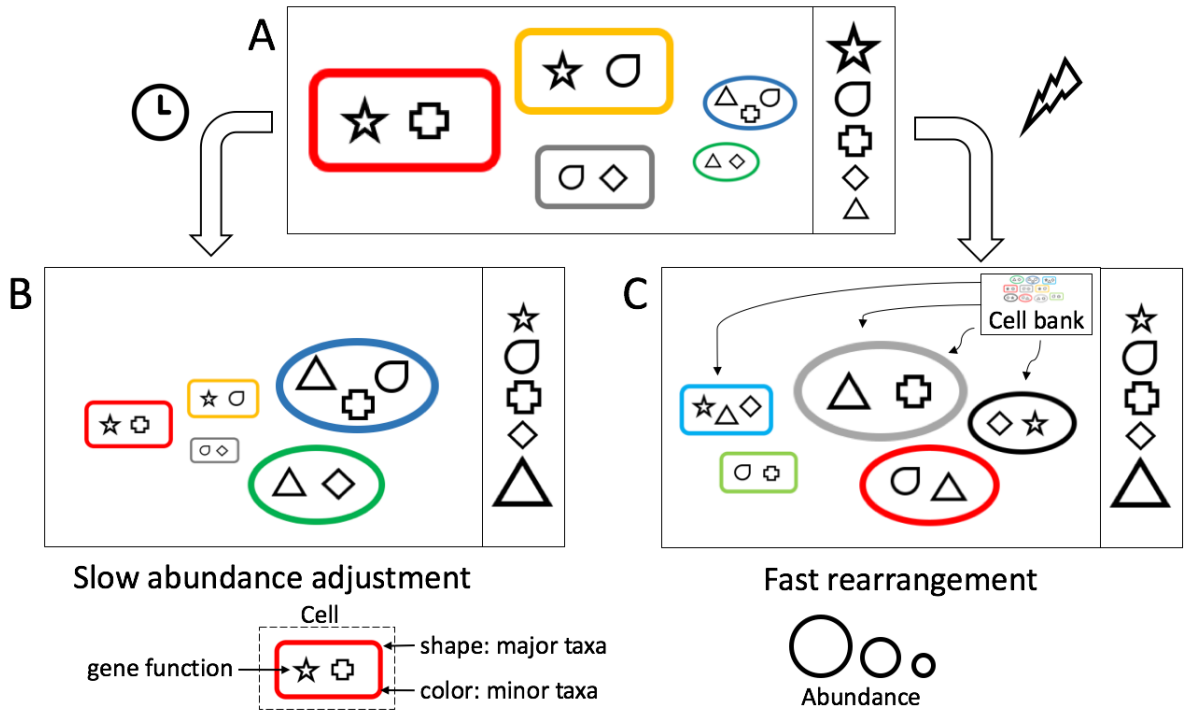


Figure 3 (Literal representation): *Cartoon model of a slow and rapid responses of a microbial community (A) to changes in selective pressures by changing relative abundances of major (i.e. domain or phylum) and minor (i.e. species or genus) taxa, resulting in identical change in functional potential (represented to the right of each community). (C) Gradual adjustment resulting in changes in the relative abundance of major taxa, which themselves remain unchanged. (D) Rapid adaptation through not only abundance changes in major taxa, but also rearrangement of minor taxa that constitute them, resulting in new groups of organism occupying existing functional niches.*

Alternative Figure 3 (barplot representation):



Alternative Figure 3 (size representation):



METHODS

Sample collection and DNA extraction

In this study, halites from three sites were harvested in Salar Grande in the Atacama Desert – site 1 (used for pre- and post-rain replicates) at the at 20°57'12.006"S 70°1'10.5996"W 680m above sea level, site 2 (used only for post-rain samples) at the bottom of that hill at 20°57'8.5212"S 70°1'1.2612"W 664m above sea level, and site 3 (for supplementary sequencing to improve binning results) at 20°55'48.18"S 70°0'49.32"W 676m above sea level. At each site, a 50^2m area was randomly sampled. Halite nodules were randomly broken with a geological hammer into smaller pieces (<20cm), and pieces of halite with visible green coloration (indicative of colonization), were ground into powder. Colonized halite powder was pooled from 1-3 nodules until sufficient material was collected, and stored in dark, dry conditions until DNA extraction in the lab. The halite powder was gradually dissolved in water and the cells spun down into a pellet as previously described [12, 14]. The gDNA was extracted from the cells by using the DNAeasy Powersoil DNA extraction kit (QIAGEN), and quantified by using the Qubit dsDNA HS Assay Kit from Invitrogen.

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 (ACACGACGCTCTTCCGATCTGTGYCAGCMGCCGCGGTAA) and 926R (CGGCATTCCTGCTGAACCGCTCTTCCGATCTCCGYCAATTYMTTTRAGTTT) universal primers, and then adding unique barcoding sequences. PCR was done with the Phusion High-Fidelity PCR kit (New England BioLabs) on 40ng of gDNA input. Barcoded samples were quantified with the Qubit dsDNA HS Assay Kit (Invitrogen), pooled together to equal molarity, and sequenced on Illumina's MiSeq platform with paired 250bp reads at Johns Hopkins Genetic Resources Core Facility (GRCF).

WMG library preparation

The whole genome sequencing libraries of the halite DNA were prepared using the KAPA HyperPlus kit (Roche). The fragmentation was performed on 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 up 3 times with XP AMPure Beads (New England BioLabs) with the following bead ratios: 1X ratio (discard unbound), 0.4X (discard beads), and 0.6 (discard unbound). 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 processing

The de-multiplexed and quality trimmed 16S amplicon reads from the MiSeq sequencer were processed with MacQIIME v1.9.1 [26]. 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 [27] release as reference and USEARCH v6.1.554 [28]. 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).

16S rDNA amplicon statistical analysis

The beta diversity metrics of samples in 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 [29] (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 matrixes), 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 [30] (default parameters, standard KRAKEN database, 2017). The reads from all samples were individually assembled (for *pI* calculations) and co-assembled (for all other analysis) with the metaWRAP Assembly module (--use-metastades option) [31]. For improved assembly and binning of low-abundance organisms, reads from all samples co-assembled and subsequently 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 [32] module (--bins option specified), classified with the Classify_bins module (default parameters), and quantified by Salmon [33] with the Quant_bins module (default parameters). Contig read depth was estimated 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) [34] annotation service. Gene abundances were assumed to be the depths of the contigs carrying those genes. KEGG KO identifiers were linked to their respective functions using the KEGG BRITE pathway classification [35]. 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 [36] with the use of metaWRAP [25], and the *pI* of each ORF was calculate with ProPAS [37]. 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*, we found all contigs carrying genes of a given KEGG KO identifier, calculated the change in their read depths between two time-points of interest, and

took 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.

$$RI = \frac{\sum_0^N |T2 - T1|}{\sum_0^N T1 + T2}$$

Equation 1: *Formula calculating one function's rearrangement index RI, where T1 and T2 are standardized abundances of a contigs 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 correlations between samples. A Pearson correlation coefficient distance matrix was computed from all replicates, and then performing a two-sided Student's two-sample t-test to evaluate the significance of the difference between the correlation distances. The relative distances were also evaluated visually with weighted Principal Component analysis and hierarchical clustering with Seaborn v0.8 [29] (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 and contig abundance clustering was determined with SigClust (nsim=1000, icovest=3) [38]. For time considerations, the contig clustering test was limited to contigs over 5kbp in lenegth, which were then subsampled randomly to 5000 contigs prior to the test.

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. 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.
13. 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.
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. 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.
16. 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.
17. 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.
18. Oren A: **Bioenergetic aspects of halophilism**. *Microbiol Mol Biol Rev* 1999, **63**(2):334-348.
19. Oren A: **Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes**. *Front Microbiol* 2013, **4**:315.
20. 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.
21. 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.
22. 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).
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. 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.
27. 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.
28. Edgar RC: **Search and clustering orders of magnitude faster than BLAST**. *Bioinformatics* 2010, **26**:2460-2461.
29. 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.
30. Wood DE, Salzberg SL: **Kraken: ultrafast metagenomic sequence classification using exact alignments**. *Genome Biol* 2014, **15**(3):R46.
31. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA: **metaSPAdes: a new versatile metagenomic assembler**. *Genome Res* 2017, **27**(5):824-834.
32. 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.
33. 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.

34. 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.
35. 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.
36. 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.
37. Wu S, Zhu Y: **ProPAS: standalone software to analyze protein properties**. *Bioinformation* 2012, **8**(3):167-169.
38. 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.