Response of extremophile microbiome to a rare rainfall reveals a two-step adaptation mechanism

Gherman Uritskiy[1], Samantha Getsin[1], Adam Munn[1], Benito Gómez-Silva[2], Alfonso Davila3[3], Brian Glass[3], James Taylor[1,4,†] and Jocelyne DiRuggiero[1,†]

[1]Department of Biology, Johns Hopkins University, Baltimore MD 21218. [2]Biomedical Department, University of Antofagasta, Antofagasta, Chile. [3]NASA Ames Research Center, Mountain View CA 94035 [3]NASA Ames Research Center, Mountain View CA 94035 [4]Department of Computer Science, Johns Hopkins University, Baltimore MD 21218.

[†]Correspondence to JT (james@taylorlab.org) and JDR (jdiruggiero@jhu.edu)

Abstract

Microbiomes play a predominant role in the biosphere and, as such, understanding the mechanisms underlying their resistance and resilience is essential to predict the impact of climate change on Earth's ecosystems. The response of a highly specialized extremophile community from the Atacama Desert to a catastrophic rainfall provided the perfect opportunity to characterize microbial resilience to extreme perturbations. Our multi-year study revealed the contrasting dynamics of the initial response and recovery of this unique community after the rain, which resulted in a transient proteome adjustment and permanent taxonomic rearrangement. The recovered microbiome was comprised of an entirely new set of organisms, despite returning to the pre-rain functional potential. The response dynamics of this community allowed for the inference of a general model, outlining two differing modes of microbiome adaptation.

Keywords

microbiome, dynamics, perturbation, resilience, extremophile, desert

Microbial communities are essential to ecosystem functioning on our planet and their wide taxonomic and functional diversity allows them to respond rapidly to acute environmental changes (1). Because such response can greatly affect process rates of entire ecosystems, it is essential to understand the mechanisms underlying community resistance and resilience, in particular in the context of global climate change, as extreme weather events are becoming more frequent (1). Previous studies have shown that acute disturbances can push a community's taxonomic structure toward alternative equilibrium states, while retaining the preexisting functional potential (2). The dynamics of such changes have been addressed with experimental perturbations in a number of ecosystems, including soil, aquatic, engineered, and human-associated systems (1). However, far less is known about the temporal dynamics of microbial communities under natural environmental conditions in part due to difficulty in avoiding compounding environmental factors (3, 4). To address this, we examined the temporal dynamics of a unique microbial system at the dry limit for life in response to a major climate perturbation.

The hyper-arid core of the Atacama Desert, Chile is one of the harshest extreme environments on Earth, with an average annual precipitation of less than 1mm (5, 6). Despite this, extremophilic microbiota have evolved strategies to survive and grow within mineral substrates (7). One such community inhabits halite (salt rock) nodules found in evaporitic basins of the desert, including the Salar Grande basin (8). Encased in rocks, halite communities receive water almost exclusively from deliquescence, the ability of salt to produce concentrated brine when atmospheric relative humidity raises above 75% (9). The majority of the biomass is constituted of salt-in strategists (8, 10) that can only tolerate a narrow range of external salt concentrations (11). As such, these highly specialized communities are more vulnerable to change compared to habitat generalists, particularly to sudden changes in external osmotic pressure from increased water availability. Additionally, each halite nodule represents a near-closed miniature ecosystem, allowing community changes to be tracked without external factors compounding the results. Thus, halite microbiomes are ideal for studying temporal dynamics of microbial communities and their ability to adapt to major environmental changes.

An unusual rainfall in August 2015 provided the opportunity to track the response of an environmental microbiome to a major natural perturbation. A weather station (Diego Aracena, ID85418) located 40km north of Salar Grande recorded 4.1mm of rainfall over two days – the first major precipitation in Northern Atacama since 2002 (6, 12, 13). Dozens of halite nodules from the basin were sampled over four years and interrogated with ribosomal amplicon (16S rDNA) and whole metagenomic (WMG) sequencing (Table S1), twice pre-rain (Sep 2014, Jun 2015) and twice post-rain (Feb 2016, Feb 2017).

Our study captured not only the microbiome's short-term adaptations to the rain perturbation, but also its recovery in the subsequent years. The halite communities were found to be highly sensitive to acute perturbations in environmental conditions, as the rain induced a drastic change in their taxonomic structure. Weighted Unifrac analysis, which compares the dissimilarity of communities based on weighted taxonomic composition, revealed that the pre-rain (2014 and 2015) and post-recovery communities (2017) clustered together, and away from 2016 post-rain communities (two-sided t-tests: $p<0.0001$; Fig. 1A, S1E). The overall halite community structure (Fig. S2) shifted from a pre-rain *Archaea*-dominated community to a more balanced *Archaea-Bacteria* community 6-months post-rain and subsequently recovered to the pre-rain state 18-months after the rain (Fig. 1B; two-sided t-tests: $p<0.001$). Major Phyla also followed this trend in relative abundance changes, as *Cyanobacteria*, green algae, and *Bacteroidetes* significantly increased in relative abundance following the rain and then decreased to pre-rain levels. In contrast, the

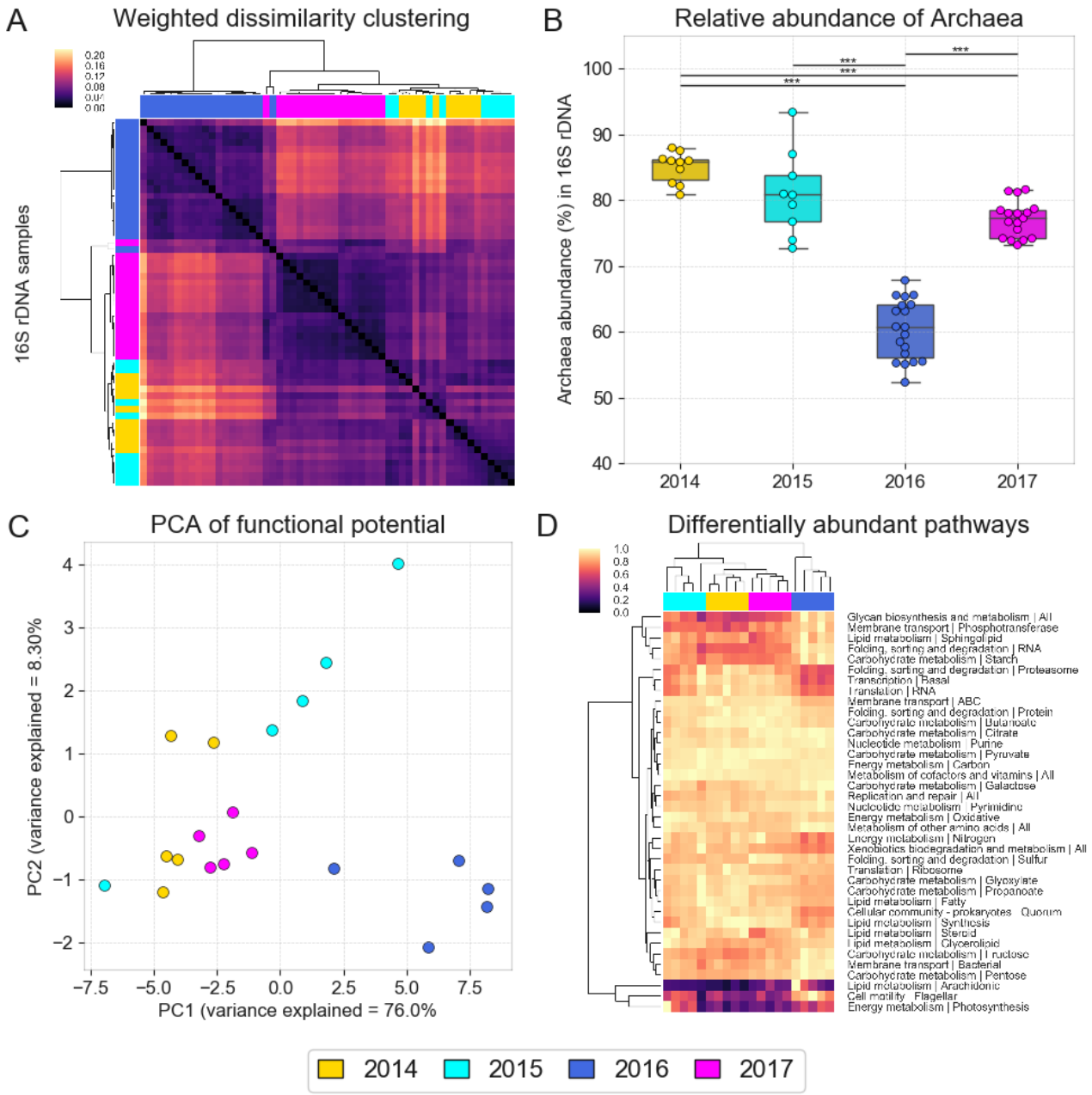

Figure 1: Taxonomic composition of halite microbiomes over time, shown through (A) hierarchical clustering (correlation metric) of the Weighted Unifrac dissimilarity matrix and (B) the average relative abundance of archaeal sequences, based on 16S rDNA amplicon sequencing. The changes in functional potential of the halite communities shown with (C) a PCA of the abundance of KEGG pathways inferred from WMG co-assembly quantitation and (D) hierarchical clustering (Euclidean metric) of differentially abundant pathways (ANOVA p<0.01, FDR=<1%), standardized to the maximum value in each row. Significance bars denote two tail t-test p-val<0.0001.

relative abundance of *Halobacteria* (the major archaea phylum in this community) dropped after the rain and then recovered (Fig. S1A-D, S2, two-sided t-tests: $p<0.01$). The incremental recovery of domain and phyla relative abundances was also observed over an 18-month period at a nearby supplementary Site 2, supporting our findings from Site 1 (Fig. S3). However, the remarkably long recovery period we observed is in contrast with studies in other desert systems where much quicker recoveries were documented (4). Our findings suggest that the immediate effects of the rain on the halite communities may have been more drastic than what we observed 6-months post-rain, and highlight the slow-growing nature of these microbiomes (8, 14).

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, communities from after the rain (2016) were functionally distinct from the pre-rain (2014 and 2015) and recovered (2017) communities (Fig. 1C; two-sided t-tests 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), particularly in post-rain samples (SigClust 2-group significance: $p<0.0001$). The most notable changes in the functional composition of the community was a shift to preferring proteins with a higher isoelectric point (pI), and a decrease in the potassium uptake potential (Trk genes; Fig. S4A-C; two-sided t-test: $p<0.01$). These changes were observed in the whole community as well as within the highly heterogeneous Halobacteria phylum (Fig. S4E, F). Low proteome pI and high potassium uptake are hallmarks of halophilic salt-in strategists, which actively import potassium ions to counteract the high external osmotic pressure, which in turn requires their proteomes to have an extremely low pI to function in the high potassium environment (11, 15). The increased pI and the decreased potassium import potential suggest that the rain temporarily decreased the salt concentrations within the colonized pores (9, 16), rapidly changing the osmotic conditions within. We hypothesize that this could have led to a mass death event of organisms poorly adapted to large osmotic changes immediately following the rain, while giving others a significant advantage. In particular, organisms with a higher pI of their proteomes (Bacteroidetes; Fig. S4D), were likely better adapted to such conditions and significantly increased in relative abundance.

The rain also drastically rearranged the fine-scale (strain level) composition of the halite communities. When interrogating the community in terms of presence or absence of microbial strains (unweighted Unifrac), we found that the community did not return to its initial state after the perturbation, as post-rain communities from 2016 and 2017 clustered together and away from the pre-rain communities (Fig. 2A; two-sided t-tests: $p<0.0001$). These unexpected results have been validated with WMG sequencing at the scale of individual contig abundances (Fig. 2C, S5) and with 91 high-quality metagenome-assembled genomes (MAGs), reconstructed with metaWRAP (Table S2) (17). Pearson correlation comparison (two-sided t-test: $p<0.0001$) as well as group significance analysis (SigClust 2-group significance: $p<0.01$) have confirmed the permanent shift in individual MAGs composition after the rain (Fig. 2B, 2C, S5). While the strain composition did change during the post-rain recovery between 2016 and 2017, the resulting shift was subtler when compared to the drastic rearrangement immediately following the rain. Additionally, two conditionally rare taxa (18) of *Cyanobacteria* that were previously reported in only a small fraction of halite nodules (10), were found in high abundances in most of the samples after the rain (Fig. S6). Surprisingly, we found no correlation between the functional potentials of the MAGs and their survival after the rain, suggesting that this rearrangement was a stochastic process. These results indicate that while

the abundances of higher-order taxonomic ranks recovered to the pre-rain state, the individual strains within those groups have been permanently reshuffled.

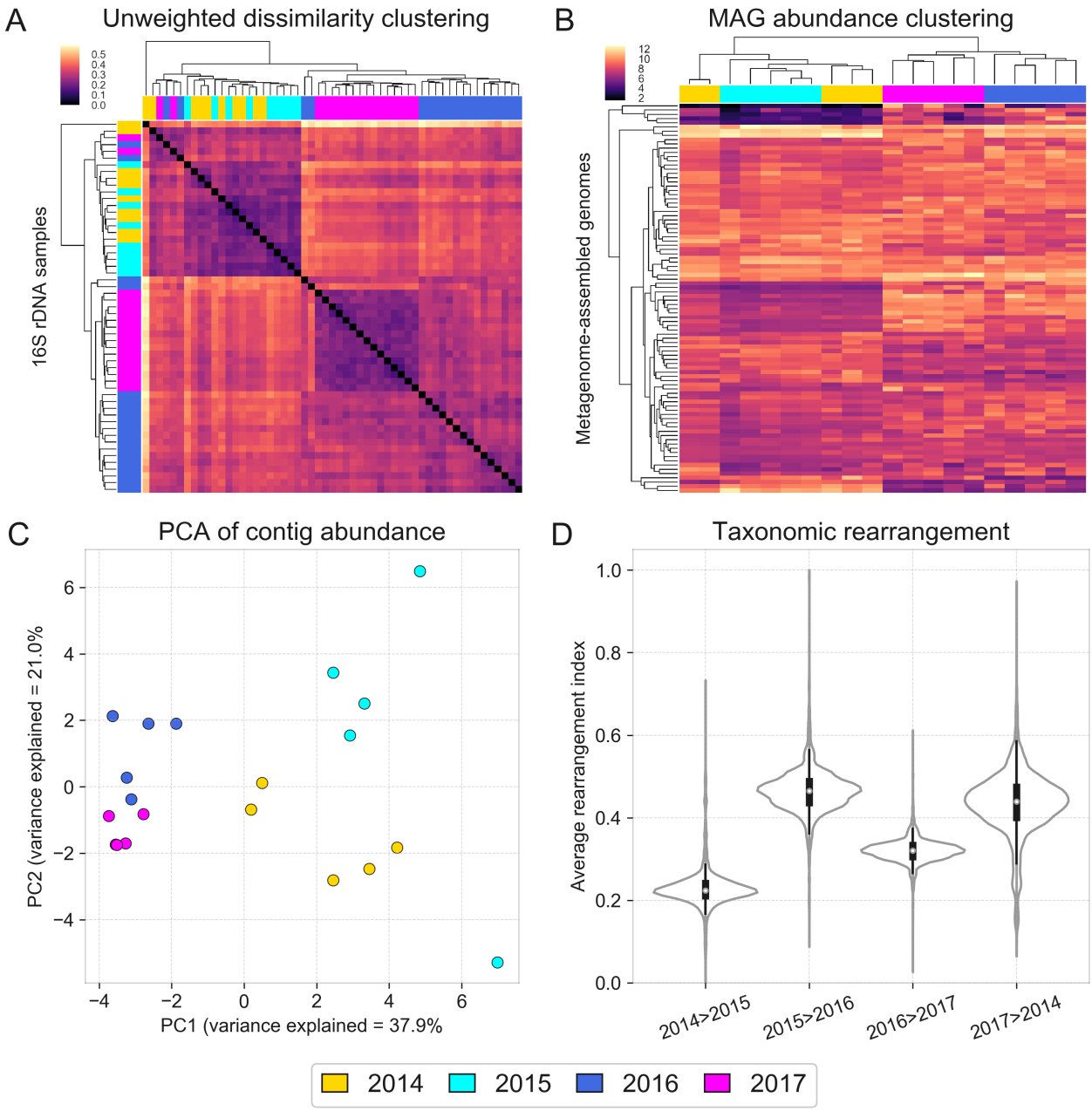

Figure 2: Changes in fine-scale composition of halite communities over time shown with (A) hierarchical clustering (correlation metric) of an Unweighted Unifrac dissimilarity matrix (based on 16S rDNA amplicon sequencing), (B) hierarchical clustering (Euclidean metric) of standardized MAG abundances, (C) PCA of co-assembly contig abundances, and (D) weighted distributions of strain rearrangement (RI) of functional niches between time points.

To investigate the impact of this rearrangement on the halite community functional niches, we introduced a strain rearrangement index (RI), which quantifies the turn-over of strains contributing to each community function. The RI was calculated for each KEGG Orthology identifier, evaluating the degree of change over time in abundances of contigs that carry a given function (Supp. Methods). The distribution in RIs for all functions between any two time points allowed us to quantify changes in niche representation over that time (Fig. 2D). The rearrangement following the rain (2015 to 2016) was significantly higher than the baseline strain rearrangement prior to the rain (2014 to 2015; KS 2-sample test: $p<0.0001$), indicating that the same functional pathways were being carried on a different set of contigs. However, the rearrangement of functional niche membership during the recovery phase (2016 to 2017) was low compared to the post-rain shift, indicating that the strain membership did not recover. The functional redundancy of community members ensured a robust functional landscape in the halite microbial communities despite change in individual strains. This functional consistency, disconnected from taxonomic variance, has been documented in a variety of microbiomes (19-22). In particular, isolated microbiomes such as miniature aquatic ecosystems found in bromeliad rosettes (similarly isolated like the halite nodules) appear to converge on identical functional landscapes through mechanisms such as stoichiometric balancing between metabolic pathways, despite great inter-community taxonomic diversity (23, 24).

The stochastic rearrangement of individual strains contributing to functional pathways was likely driven by neutral (i.e. random) processes (25, 26) similar to those governing original colonization of halite nodules, which results in high inter-nodule taxonomic diversity at the strain level (10), while the functional states are equivalent. Each nodule is stochastically colonized by competitively equivalent organisms by random draw from the seed bank – a diverse genetic reservoir consisting of a large collection of low-abundance organisms (1, 27). 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. The resulting diversity allows for rapid adaptation and restructuring of microbial communities following drastic perturbations, as seen after the 2015 rain.

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. Mechanistically however, the two shifts are fundamentally distinct and allow for the inference of two contrasting modes by which a microbiome responds to changing environmental conditions while resulting in a similar functional potential (Fig. 3). The first mode (Type I) is a community rearrangement, resulting from adaptations to an acute major perturbation that creates gaps in existing functional niches and presents an opportunity for new organisms from the seed bank to come in through niche intrusion (28) (Fig. 3A). The Type I shift is characterized by a high RI, with changes in fine-scale (i.e. strains) taxonomic composition driven by neutral processes. The second mode (Type II) is an adjustment in existing community structure, and results from gradual changes in environmental conditions (Fig. 3B). The taxonomic composition in this type of response is relatively deterministic, as currently dominant organisms have the opportunity to adjust their relative abundances without allowing new organisms to take over, resulting in a low RI.

In conclusion, we found the halite microbiomes to be both sensitive (taxonomically) and resilient (functionally) to major climate perturbations. On a short time-scale following the rainfall, the halite communities rearranged taxonomically and, on a long time-scale, entered an alternative equilibrium state characterized by a recovery of the initial functional potential (2). This, in effect, is similar to what is observed in other

ecosystems after disaster events (29) or antibiotic administration (28, 30), where communities are able to recover functionally, but with loss of former taxonomic diversity and acquisition of new species through

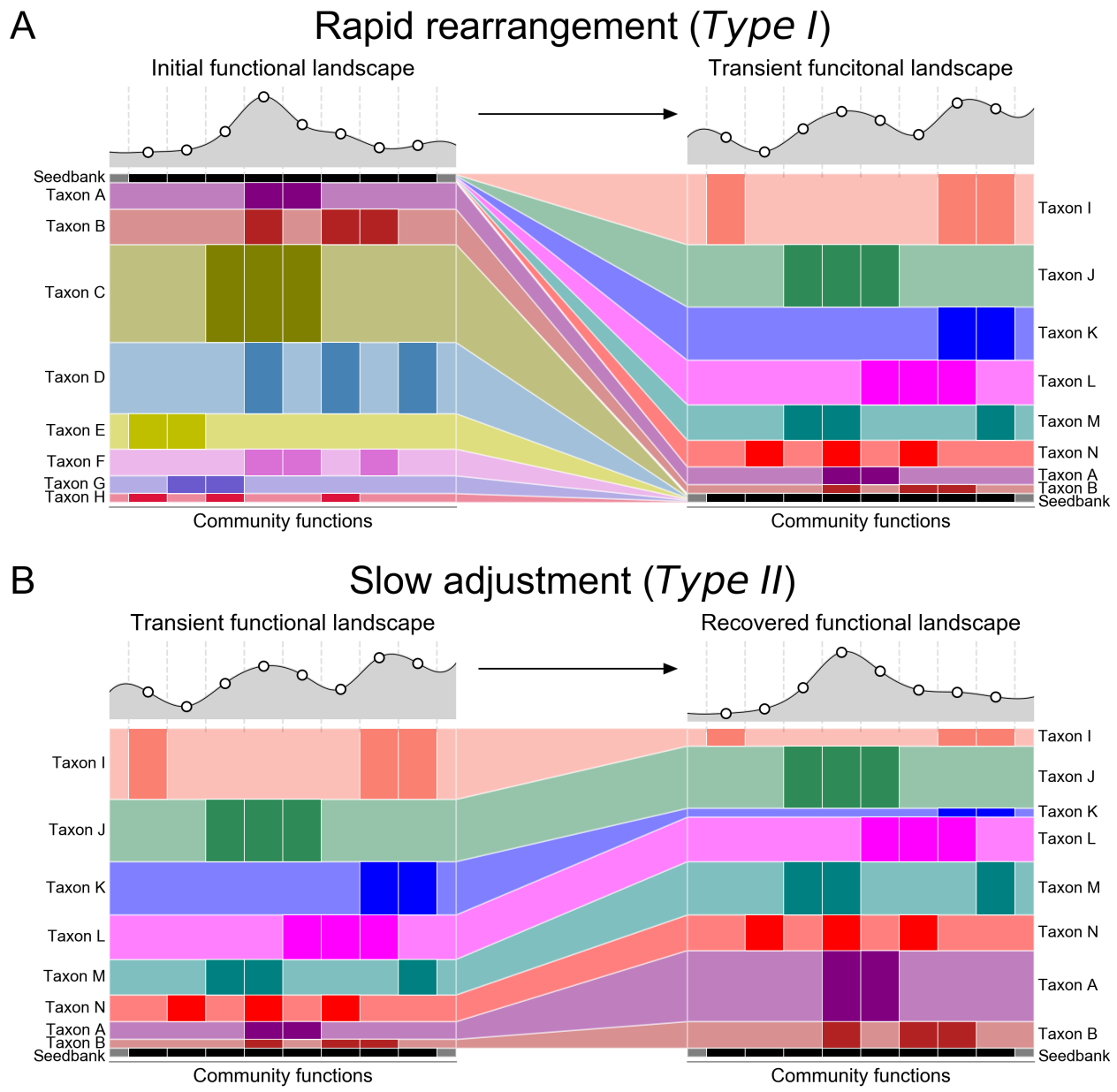

Figure 3: Model of a microbiome adapting its functional potential in response to changing environmental conditions either through (A) a rapid rearrangement of the community's taxonomic structure resulting from new organisms from the seed bank displacing most previously dominant taxa through niche intrusion (as seen in the initial shock from the rainfall), or through (B) a gradual adjustment in relative abundance of major taxa (as seen in the halite community recovery).

niche intrusion. However, the tractable nature of the halite microbiomes allowed us to extrapolate a general two-step mechanism of these communities' responses. The initial shock from the perturbation produces gaps in community's functional niches, resulting in a Type I rearrangement, which is followed by a recovery period in which the relative abundances of the newly dominant taxa adjust through a Type II shift, reproducing the initial functional landscape. This two-step model can potentially be used to explain some of the compositional and functional dynamics observed in other ecosystems.

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 (31). All the sites were within 5 km of each other and, at each site, halite nodules were harvested within a 50m2 area. Sites were as follow: S1 was used for the analysis in this work comparing pre- and post-rain samples, S2 was used for validating the post-rain recovery, and S3 was used to improve binning results but not for relative 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 previously described (31) 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. Genomic DNA was extracted as previously described (8, 31) 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 (31) by amplifying the hypervariable V3-V4 region with 515F and 926R primers (32). 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 platform at GRCF.

16S rDNA amplicon sequence analysis

The de-multiplexed and quality trimmed 16S rDNA amplicon reads from the MiSeq sequencer were processed with MacQIIME v1.9.1 (33). 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 (34) release as reference and USEARCH v6.1.554 (35). 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 (36) (method="average", metric="correlation"). Group significance was determined with compare_categories.py (–methodpermanova). 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. Detailed scripts for the entire analysis pipeline can be found at https://github.com/ursky/timeline_paper

WMG sequence processing

The de-multiplexed WMG sequencing reads were processed with the complete metaWRAP v0.8.2 pipeline (18) 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 (37) (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) (38). 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 (39) module (–bins option specified), classified with the Classify_bins module (default parameters), and quantified by Salmon (40) 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 Detailed scripts for the entire analysis pipeline can be found at https://github.com/ursky/timeline_paper

Functional annotation

Gene prediction and functional annotation of the co-assembly was done with the JGI Integrated Microbial Genomes & Microbiomes (IMG) (41) annotation service. Gene relative abundances were taken as the average read depth of the contigs carrying those genes (estimated with Salmon (40). KEGG KO identifiers were linked to their respective functions using the KEGG BRITE pathway classification (42). KEGG pathway relative abundances were calculated as the sum of read depths of genes (estimated from the read depths of the contigs carrying them) classified to be part of the pathway.

Isoelectric point (pI) analysis

The average pI of gene pools were calculated from individual replicate metagenomic assemblies. Open reading frames (ORFs) were predicted by PRODIGAL (43) with the use of metaWRAP (18), and the pI of each ORF was calculate with ProPAS (44). 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 (RI) of each gene function (KO ID) represent the changes in relative abundances of the contigs carrying them. To calculate the RI, all contigs carrying genes of a given KEGG KO were identified, and the change in their relative abundances was calculated between two time-points of interest. Contig abundances from individual replicates were added up for each time point Then the RI for each KEGG KO identifier was calculated from the weighted average of the absolute values of these changes (Equation 1). The RIs from all the KEGG functions were plotted and the difference in their distributions between the time points was computed with the Kolmogorov-Smirnov 2-sample test.

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

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

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 (36) (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) (45). 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.

Availability of data and materials

Raw sequencing data is available from the National Centre for Biotechnology Information under project ID PRJNA484015. All analysis pipelines, processed data, analysis and visualization scripts, and reconstructed MAGs are available at `https://github.com/ursky/timeline_paper`. The metagenome co-assembly and functional annotation are available from the JGI Genome Portal under IMG taxon OID 3300027982.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grants NNX15AP18G and NNX15AK57G from NASA, DEB1556574 from the NSF, FB-0001 from CONICYT, 5305 from Universidad de Antofagasta, Chile, and HG006620 from NIH/NHGRI.

Author's contributions

GU, JT, and JD conceptualized and designed the study; GU, JD, BGS and AD collected in-field samples; BG organized and funded field expeditions; SG and AM processed and sequenced samples; GU analyzed the data and wrote the manuscript; JT and JD edited the manuscript.

Acknowledgements

We thank Mike Sauria and Boris Brenerman for data analysis suggestions, Alexandra Galetovic for help with fieldwork, and Sarah Preheim, Michael Schatz and Vadim Uritsky for help in manuscript editing.

Bibliography

1. A. Shade et al., Fundamentals of microbial community resistance and resilience. Front Microbiol 3, 417 (2012).
2. M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, B. Walker, Catastrophic shifts in ecosystems. Nature 413, 591-596 (2001).
3. S. D. Allison, J. B. Martiny, Colloquium paper: resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci U S A 105 Suppl 1, 11512-11519 (2008).
4. A. Armstrong et al., Temporal dynamics of hot desert microbial communities reveal structural and functional responses to water input. Sci Rep 6, 34434 (2016).
5. C. P. McKay et al., Temperature and moisture conditions for life in the extreme arid region of the Atacama desert: four years of observations including the El Niño of 1997-1998. Astrobiology 3, 393-406 (2003).
6. D. Bozkurt, R. Rondanelli, R. Garreaud, A. Arriagada, Impact of Warmer Eastern Tropical Pacific SST on the March 2015 Atacama Floods. Monthly Weather Review 144, 4441-4460 (2016).
7. J. Wierzchos, M. C. Casero, O. Artieda, C. Ascaso, Endolithic microbial habitats as refuges for life in polyextreme environment of the Atacama Desert. Current Opinion in Microbiology 43, 124-131 (2018).
8. A. Crits-Christoph et al., Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community. Environ Microbiol 18, 2064-2077 (2016).
9. A. F. Davila et al., In situ metabolism in halite endolithic microbial communities of the hyperarid Atacama Desert. Front Microbiol 6, 1035 (2015).
10. K. M. Finstad et al., Microbial Community Structure and the Persistence of Cyanobacterial Populations in Salt Crusts of the Hyperarid Atacama Desert from Genome-Resolved Metagenomics. Front Microbiol 8, 1435 (2017).
11. A. Oren, Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. Front Microbiol 4, 315 (2013).
12. Dirección Meteorológica de Chile, Servicios Climáticos (2018 https://climatologia.meteochile.gob.cl/application/index/productos/RE2009).
13. N. Schulz, J. P. Boisier, P. Aceituno, Climate change along the arid coast of northern Chile. International Journal of Climatology 32, 1803-1814 (2012).
14. L. A. Ziolkowski, J. Wierzchos, A. F. Davila, G. F. Slater, Radiocarbon evidence of active endolithic microbial communities in the hyper-arid core of the Atacama Desert,. Astrobiology 13, 607-616 (2013).
15. S. Paul, S. K. Bag, S. Das, E. T. Harvill, C. Dutta, Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. Genome Biol 9, R70 (2008).
16. K. Finstad et al., Rates and geochemical processes of soil and salt crust formation in Salars of the Atacama Desert, Chile. Geoderma 284, 57-72 (2016).
17. G. V. Uritskiy, J. DiRuggiero, J. Taylor, MetaWRAP - a flexible pipeline for genome-resolved metagenomic data analysis. bioRxiv, (2018).
18. A. Shade et al., Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. Mbio 5, (2014).
19. C. A. Lozupone, J. I. Stombaugh, J. I. Gordon, J. K. Jansson, R. Knight, Diversity, stability and resilience of the human gut microbiota. Nature 489, 220-230 (2012).
20. A. Eng, E. Borenstein, Taxa-function robustness in microbial communities. Microbiome 6, 45 (2018).
21. Y. Nie et al., Species Divergence vs. Functional Convergence Characterizes Crude Oil Microbial Community Assembly. Front Microbiol 7, 1254 (2016).

22. J. E. Goldford et al., Emergent Simplicity in Microbial Community Assembly. bioRxiv, (2017).
23. S. Louca et al., High taxonomic variability despite stable functional structure across microbial communities. Nat Ecol Evol 1, 15 (2016).
24. S. Louca et al., Function and functional redundancy in microbial systems. Nature Ecology & Evolution 2, 936-943 (2018).
25. S. P. Hubbell, The Unified Neutral Theory of Biodiversity and Biogeography. . (Princeton Univ Press, Princeton: New Jersey, 2001).
26. L. Li, Z. S. Ma, Testing the Neutral Theory of Biodiversity with Human Microbiome Datasets. Sci Rep 6, 31448 (2016).
27. J. T. Lennon, S. E. Jones, Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol 9, 119-130 (2011).
28. S. R. Modi, J. J. Collins, D. A. Relman, Antibiotics and the gut microbiota. J Clin Invest 124, 4212-4218 (2014).
29. R. L. Rodriguez et al., Microbial community successional patterns in beach sands impacted by the Deepwater Horizon oil spill. ISME J 9, 1928-1940 (2015).
30. M. O. Sommer, G. Dantas, Antibiotics and the resistant microbiome. Curr Opin Microbiol 14, 556-563 (2011).
31. C. K. Robinson 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 17, 299-315 (2015).
32. D. M. Needham, J. A. Fuhrman, Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. Nat Microbiol 1, 16005 (2016).
33. J. G. Caporaso et al., QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7, 335-336 (2010).
34. C. Quast et al., The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41, D590-596 (2013).
35. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460-2461 (2010).
36. M. Waskom et al. (GitHub, 2017), pp. https://github.com/mwaskom/seaborn.
37. D. E. Wood, S. L. Salzberg, Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol 15, R46 (2014).
38. S. Nurk, D. Meleshko, A. Korobeynikov, P. A. Pevzner, metaSPAdes: a new versatile metagenomic assembler. Genome Res 27, 824-834 (2017).
39. S. Kumar, M. Jones, G. Koutsovoulos, M. Clarke, M. Blaxter, Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. Frontiers in Genetics 4, 237 (2013).
40. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14, 417-419 (2017).
41. I. A. Chen et al., IMG/M: integrated genome and metagenome comparative data analysis system. Nucleic Acids Res 45, D507-D516 (2017).
42. M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44, D457-462 (2016).
43. D. Hyatt et al., Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11, 119 (2010).

44. S. Wu, Y. Zhu, ProPAS: standalone software to analyze protein properties. Bioinformation 8, 167-169 (2012).
45. Y. Liu, D. N. Hayes, A. Nobel, J. S. Marron, Statistical Significance of Clustering for High-Dimension, Low-Sample Size Data. Journal of the American Statistical Association 103, 1281-1293 (2008).

Supplementary Figures

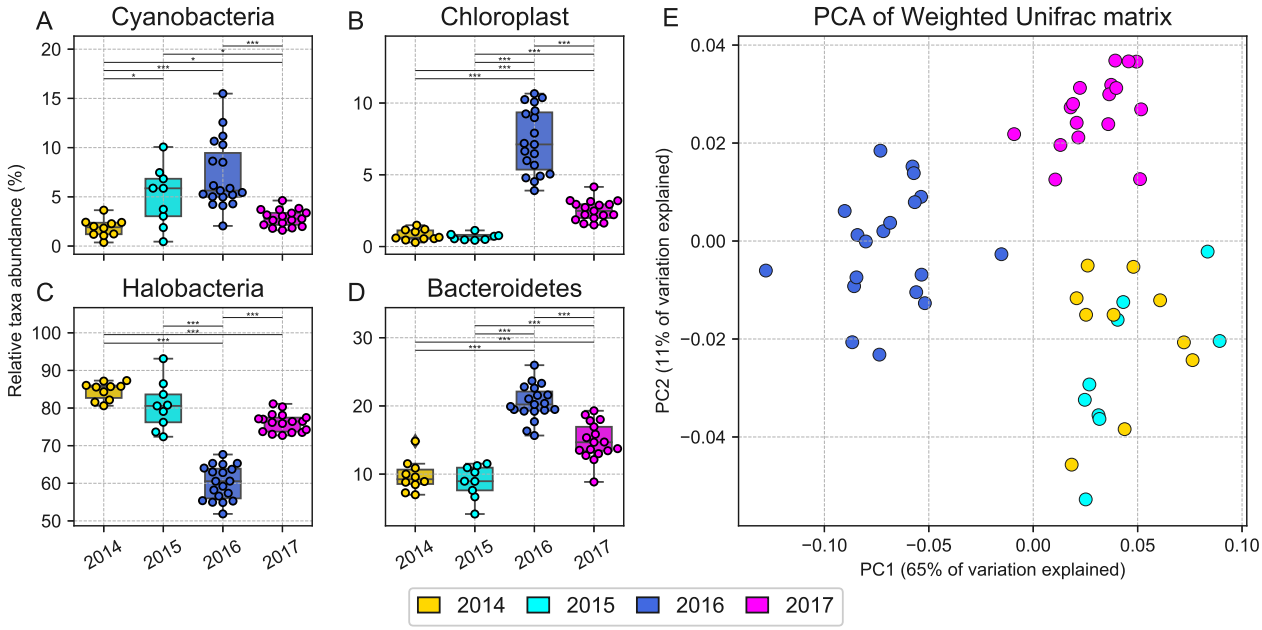

Figure S1: Taxonomic composition differences between halite samples harvested from Site 1 at different dates, inferred from 16S rDNA sequences clustered into OTUs at 97% identity and visualized through (A-D) relative abundance of major differentially abundant phyla and a (E) PCA plot of a Weighted Unifrac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

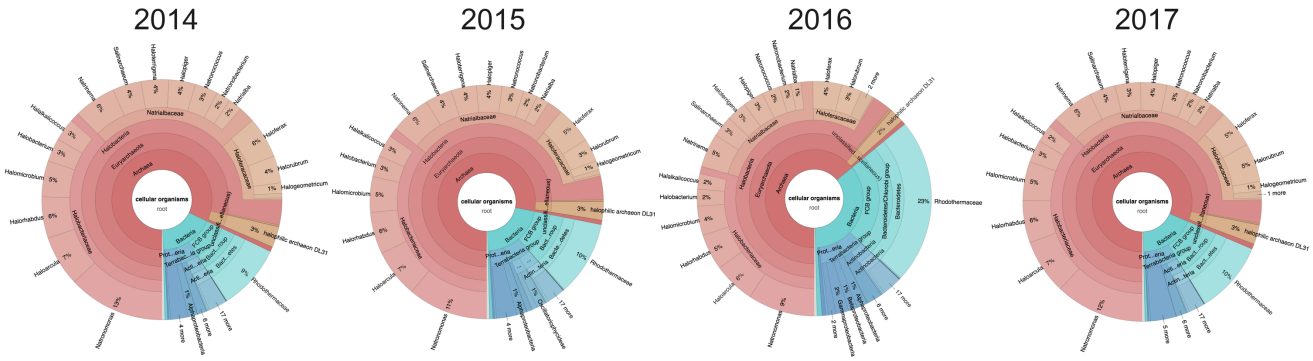

Figure S2: Average taxonomic composition of halite microbial communities from Site 1 sampled at different dates, estimated from WMG reads with KRAKEN and visualized with KronaTools.

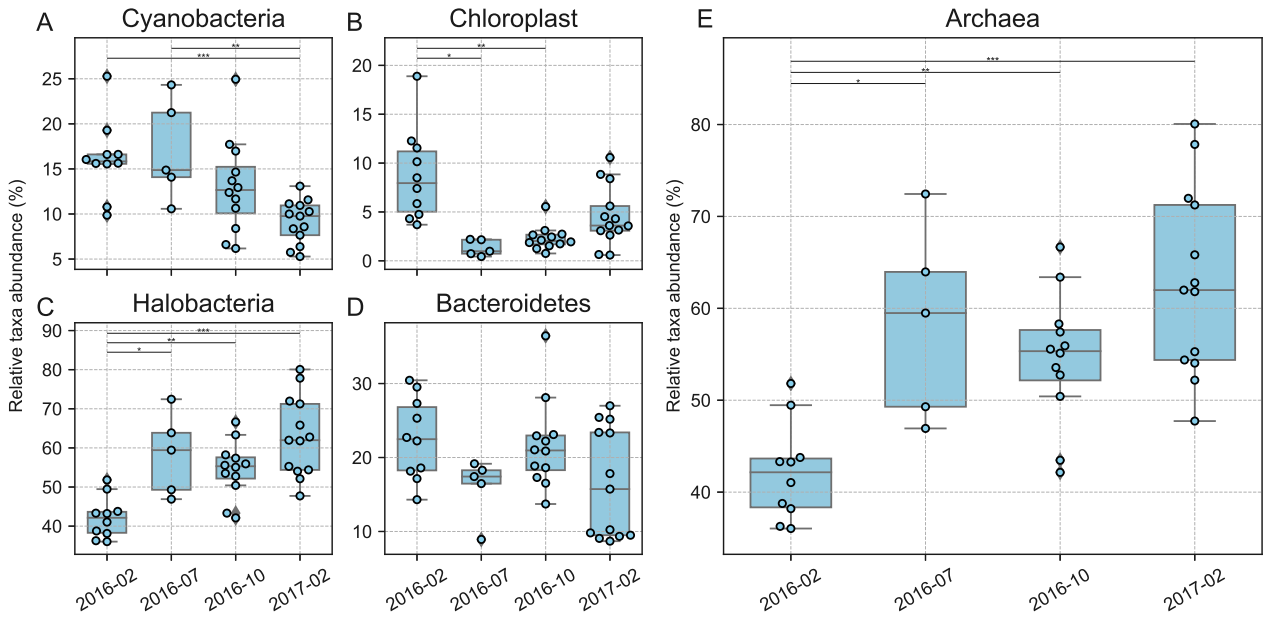

Figure S3: Taxonomic composition differences between halite samples harvested from Site 2 at different dates post-rain, infered from 16S rDNA sequences clustered into OTUs at 97% identity and visualized through (A-D) relative abundance of major differentially abundant phyla and (E) archaea abundance.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

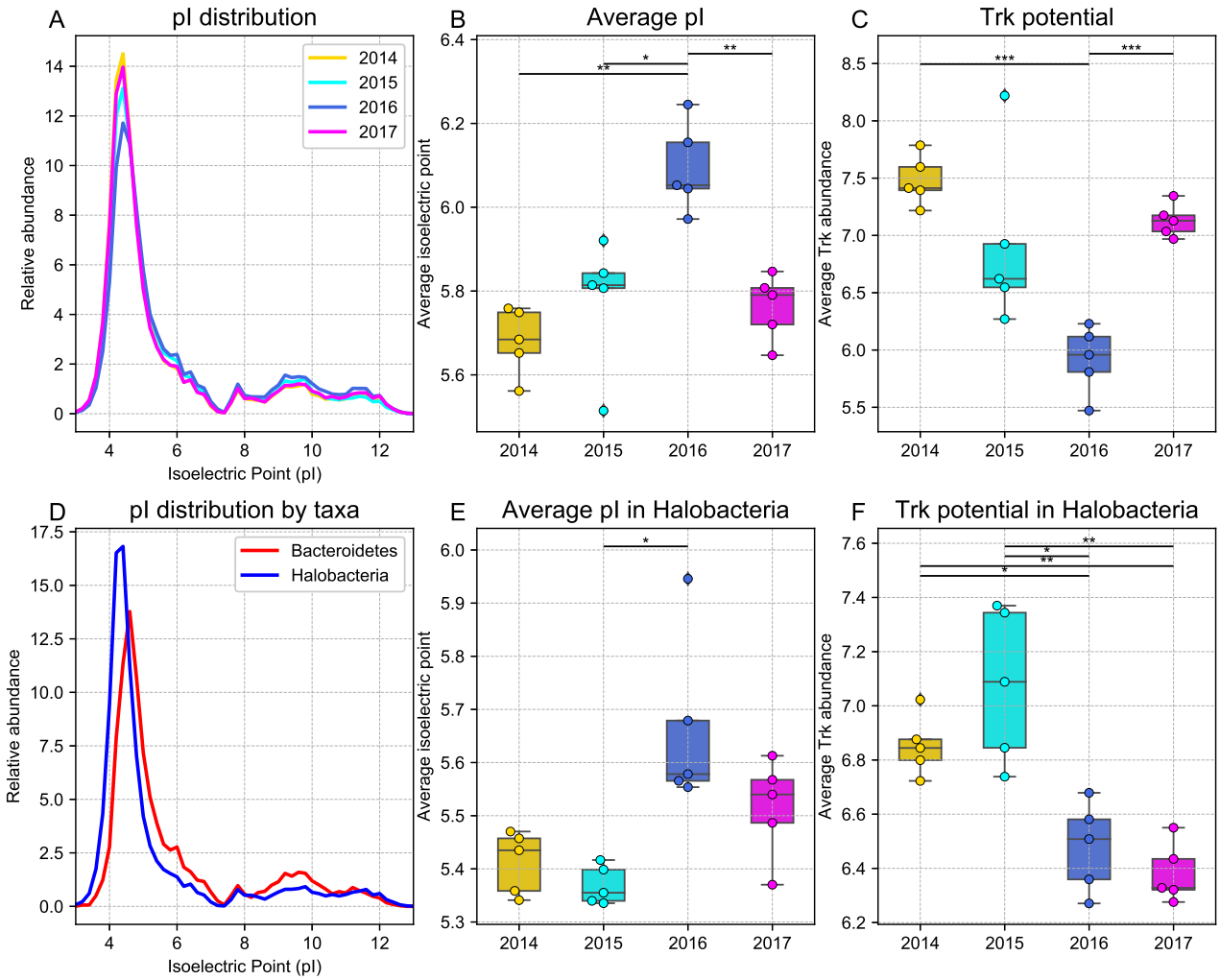

Figure S4: Analysis of the isoelectric points (pI) of proteins encoded in replicates of WMG assemblies from samples harvested at different dates, showing (A) the overall weighted distribution of the protein pIs, and the weighted average pI of proteins encoded in (B) all contigs and (E) only Halobacteria contigs. (D) pI distribution of proteins encoded in Bacteroidetes and Halobacteria contigs. Average potassium uptake potential across time point samples inferred from Trk gene relative abundance and quantified in (C) all contigs and (F) only Halobacteria contigs.

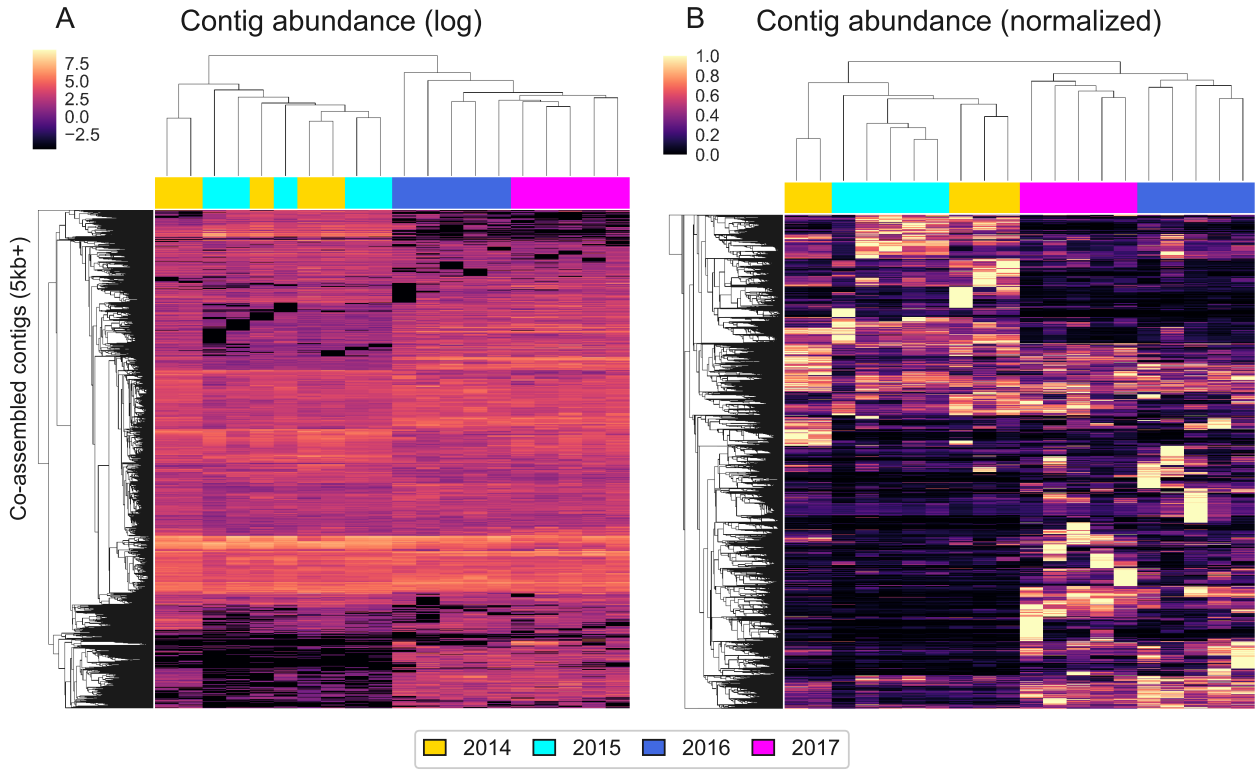

Figure S5: Hierarchical clustering (Euclidean metric) of relative abundances (fragments per million) of contigs > 5kbp in the WMG co-assembly, quantified with reads from samples harvested at different dates and displayed on (A) a log scale and (B) standardized to the maximum abundance of each contig.

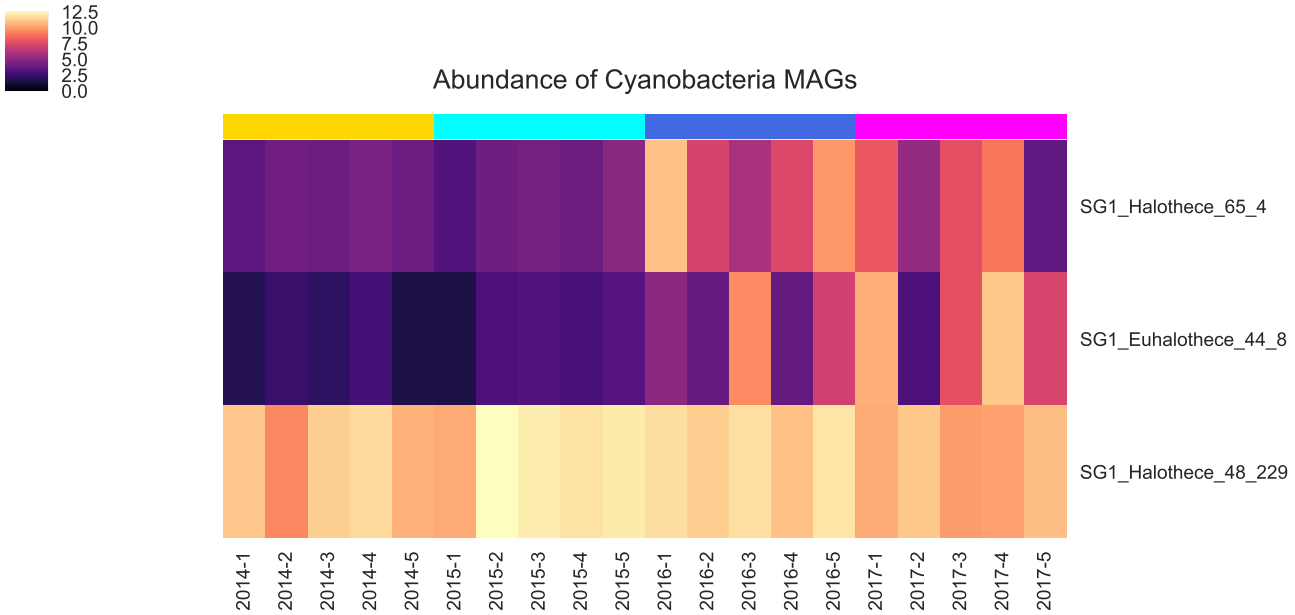

Figure S6: Hierarchical clustering (Euclidean metric) of photosynthetic MAG relative abundances (fragments per million), quantified with metaWRAP's quant_bins module, showing the emergence of two new Cyanobacteria MAGs after the rain.