Science Advances
AAAS

Manuscript Template

FRONT MATTER

Title:

Resilience and adaptation mechanisms of microbial community following catastrophic climate event

- Short title (<50 chars): Microbial community resilience to perturbation

Authors

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

Affiliations

[1]Department of Biology, Johns Hopkins University, Baltimore, MD, USA.

[2]Biomedical Department, CeBiB, Universidad de Antofagasta, Antofagasta, Chile.

[3]NASA Ames Research Center, Mountain View, CA, USA

[4]Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA.

*Correspondence to: james@taylorlab.org and jdiruggiero@jhu.edu.

Abstract (<150 words)

Microorganisms play a predominant role in the biosphere and, as such, understanding the mechanisms underlying their resistance and resilience to perturbations is essential to predict the impact of climate change on Earth's ecosystems. However, the temporal dynamics of microbial communities under natural environmental conditions remain relatively unexplored. The response of highly specialized desert extremophiles to a catastrophic rainfall provided the perfect opportunity to characterize and de-convolute the resilience mechanisms of communities following disturbance. We report the dynamics of their initial response, in which the community entered an unstable intermediate state after stochastic niche re-colonization, and that of their recovery, in which the community returned to its former functional potential but retained its newly-acquired strain composition. These response dynamics allowed for the inference of a general microbiome response model, which can be applied in future research to predict the taxonomic and functional flux in other ecosystems following environmental changes.

MAIN TEXT

INTRODUCTION

Microbial communities are essential to the functioning and evolution of our planet and their dynamics greatly affect ecosystems processing (*1*). Their taxonomic and functional diversity allow microbial communities to adapt to a wide range of environmental conditions and to respond rapidly to changes (*2, 3*). Mechanisms of resilience – the ability of communities to recover from perturbations – are of particular interest in the context of global climate change, as extreme weather events are becoming more frequent (*1*). Understanding adaptation strategies for microbial resilience is also critical to gain insights into microbial evolution and diversification and to better understand the dynamics of translationally relevant microbiomes following stress.

Previous studies have shown that acute disturbances can push a community's taxonomic structure toward alternative equilibrium states, while retaining the preexisting functional potential (*4*). Such changes have been observed in soil, aquatic, engineered, and human-associated ecosystems where experimental perturbations caused the community taxonomic composition to shift with relatively minor changes to the overall functioning of the community (*1, 2, 5, 6*). This functional stability is attributed to a redundancy of functions between multiple closely related taxa (*7*), ensuring that the functional potential of the community persists even after a major rearrangement of its taxonomic structure (*8-10*).

Transitions between alternative taxonomic states have been postulated to occur via an intermediate dis-equilibrium state, during which a perturbation produces drastically different environmental stressors, causing the community to radically reshape in composition (*1, 4*). This has been observed with antibiotic treatment that can lead to mass death events. The resulting re-structuring of the gut microbiome is major with long-lasting changes even after the former conditions are re-established. At this point, the community may re-enter an alternative equilibrium state with a community that is taxonomically different but functionally similar to the pre-treatment community (*6, 11*). However, little is known about the response dynamics to acute perturbations and in particularly the mechanisms that push a community's taxonomic and functional structure in and out of an intermediate state. Additionally, the response and recovery of natural communities following environmental disasters, rather than manipulative experiments, remain largely unexplored mechanistically because of the difficulty in avoiding multiple compounding environmental factors (*12, 13*). These gaps in the understanding of microbial community behavior limits our ability to effectively model and predict the responses of microbiomes to major perturbations, such as those resulting from climate change and natural or man-made ecological disasters.

To address this knowledge gap, and to build a conceptual model for modeling microbial community responses to extreme stress, we examined the temporal dynamics in response to a catastrophic climate perturbation of a unique microbial ecosystem found in the Atacama Desert, Chile. The hyper-arid core of the Atacama Desert is one of the harshest environments on Earth, with an average annual precipitation of less than 1mm and some of the highest ultraviolet (UV) and solar radiation on the planet (*14, 15*). Despite this, microbial communities have evolved strategies to survive and grow within various mineral substrates of the desert (*16*). One such community inhabits halite (salt rock) nodules found in evaporitic salt basins, including the Salar Grande basin (*17, 18*). In this community, the majority of the biomass is constituted of salt-in strategists *Halobacteria* (major archaea phylum) and *Bacteroidetes* (*18, 19*) – two taxonomically diverse groups of extreme halophiles that accumulate potassium ions to match the external osmotic pressure from sodium ions (*18, 20, 21*). This adaptation allows them to survive in extremely high-salt environment, but restricts their fitness to a narrow range of external salt concentration (*22, 23*). As such, these highly

specialized communities are more vulnerable to change compared to habitat generalists, particularly to sudden changes in external osmotic pressure.

Encased in salt rocks, halite communities have very limited nutrient input beyond atmospheric gasses, and obtain water almost exclusively from deliquescence, the ability of sodium chloride to produce concentrated brine when atmospheric relative humidity rises above 75% (*24*). Primary production is the major source of organic carbon in the community and is carried out by *Cyanobacteria* and, to a lesser extent, by a unique alga *(17)*. Each halite nodule represents a near-closed miniature ecosystem and thus can be treated as true independent biological replicates in longitudinal studies, allowing community changes to be tracked without external factors compounding the results. Combined with their sensitivity to changing osmotic conditions and slow growth rates, this makes halite microbiomes ideal for studying temporal dynamics of microbial communities and their ability to adapt to major environmental changes.

In August 2015, Northern Atacama received its first major rain in 13 years (*15, 25*). Such rain events have been observed to be devastating to the specialized hyper-arid microbiomes of the Atacama Desert (*26*). Our longitudinal study over 4 years not only captured the microbiome's short-term adaptations to this major natural environmental disaster, but also its recovery in the subsequent years, revealing two strikingly different community adaptation mechanisms.

RESULTS

Longitudinal sampling strategy and sequencing approach

To investigate the temporal dynamics of halite microbiomes, samples of halite nodules from two sites at Salar Grande were harvested at regular intervals from 2014 to 2017, capturing the rare rain events that occurred in 2015 throughout the desert (*15*). A nearby weather station (Diego Aracena airport, *ID85418*), located 40km North of the sampling site, recorded rainfalls in August 2015 (4.1mm). The previous notable precipitation in the area occurred in 2002 (4.1mm) (*25, 27*).

The main sampling site was revisited four times during the study – twice before the rain (Sep 2014, Jun 2015), and twice after the rain – 6 months (Feb 2016) and 18 months (Feb 2017) after (Table S1). For each time-point, 5 biological replicates were sequenced with whole-metagenomic (WMG) sequencing to investigate the functional potential and taxonomic structure of the communities over time, yielding a total of 70,689,467 paired-end reads (150bp paired-end, insert size 277±217bp). Additionally, 9-12 biological replicates were collected for ribosomal amplicons (16S rDNA) sequencing and were used for taxonomic profiling of the microbiomes; this yielded 535,233 paired-end reads (250bp paired-end, insert size 419±7bp). A nearby site was also sampled after the rain at a higher temporal resolution (Feb 2016, July 2016, Oct 2016, and Feb 2017), 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).

High-order taxonomic structure and functional potential were temporarily perturbed after the rain

The halite communities were found to be highly sensitive to the acute perturbation from the rain, as it induced a drastic change in their taxonomic structure. Weighted Unifrac analysis of the amplicon data, which compares the dissimilarity of communities based on weighted taxonomic composition, revealed that the halite communities were significantly different between time-points (PERMANOVA: $p<0.001$), with the taxonomic composition shifting following the rain (Fig. S2E). Samples from before the rain (2014 and 2015) and post-recovery communities (2017) clustered together, and away from post-rain (2016) communities (two-sided t-tests: $p<0.0001$; Fig. 1A, S2E). At the domain level, the halite community structure shifted from an *Archaea*-dominated community

before the rain (2014 and 2015) to a more balanced *Archaea-Bacteria* community 6-months after the rain (2016). The relative abundance of *Archaea* dropped significantly (two-sided t-tests: $p<0.0001$) in both 16S rDNA (Fig. 1B) and WMG sequencing (Fig. S1). 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* significantly decreased (Fig. S1, S2A-D, two-sided t-tests: $p<0.01$). A recovery to the pre-rain state of domain and phyla relative abundances was observed 18-months after the rain (2017) at Site 1 (Fig. S2) and also at a nearby supplementary Site 2 over an 18-month period, revealing a gradual recovery process (Fig. S3).

The functional potential of the community, determined by annotation of KEGG pathways in the WMG co-assembly, also significantly changed after the rain. Consistent with the taxonomy-based clustering, samples from before the rain (2014 and 2015) were distinctly separate from samples collected shortly after the rain (2016; Fig. 1C). The KEGG pathway abundances in 2014 samples were better correlated with that of 2015 and 2017 samples than 2016 samples (two-sided t-tests of Pearson correlations: $p<0.001$). 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 (Fig. 1D; ANOVA test, $p<0.01$, FDR<1%). Of these, the majority was significantly over- or under-represented in the samples collected shortly after the rain (2016-02; SigClust 2-group significance: $p<0.0001$).

Differences in salt adaptations likely drove the fate of the salt-in strategists

The most notable change in the functional composition of the community post-rain (2016) was an enrichment in proteins with a higher isoelectric point (*pI*), and a decrease in the potassium uptake potential (Trk genes), both of which are hallmarks of salt-in strategists. We found that the *pI* of proteins encoded in community gene pool shifted significantly after the rain, favoring higher *pI* composition (Fig. 3A; KS 2-sample test: $p<0.0001$). Because of the significantly different *pI* distributions in the proteomes of *Halobacteria* (*pI*=5.04) and *Bacteroidetes* (*pI*=5.80; Fig. 3D; KS2-sample test: $p<0.0001$), the shift in their relative abundances resulted in the average *pI* of the community to significantly increase after the rain (two-sided t-test: $p<0.01$; Fig. 3B). Consistent with salt-in adaptations, we also found that the average potassium uptake potential (estimated from *trk* gene abundances) significantly decreased after the rain (Fig. 3C). Interestingly, both the shift in the average protein pool *pI* and the change in potassium uptake potential were also observed within the highly heterogeneous *Halobacteria* phylum (Fig. 3E,F).

The rain permanently rearranged the fine-scale taxonomy of the halite community

Samples collected at different dates were significantly different in terms of presence or absence of microbial strains, as measured by the Unweighted Unifrac dissimilarity index (PERMANOVA: $p<0.001$), with samples harvested shortly after the rain (2016) being more distant from pre-rain samples than they were from each other (two-sided t-test: $p<0.0001$). Surprisingly, we found that the community did not return to its initial state after the perturbation, as the post-recovery samples (2017) clustered together with post-rain (2016) samples (Fig. 3A), and were less distant to 2016 samples than to the pre-rain samples (two-sided t-test: $p<0.0001$). The permanently altered OTU composition of the community, shown with Unweighted Unifrac clustering, strikingly contrasts with the successful recovery of the corse-grained taxonomic structure, as shown with Weighted Unifrac dissimilarity clustering (Fig. 1A).

The permanent rearrangement of the community strain membership was validated with WMG sequencing at the scale of individual contig abundances (Fig. S4). Based on contig read coverage across samples, we found that all post-rain samples clustered away from pre-rain samples (Fig. 3C;

SigClust 2-group significance: $p<0.01$). Additionally, pairwise Pearson correlation comparison confirmed that contig abundances of post-rain samples were better correlated with each other than with that of pre-rain samples (two-sided t-test: $p<0.0001$). These strain rearrangement dynamics were additionally investigated at the level of individually recovered metagenome-assembled genomes (MAGs). 91 high-quality MAGs (>70% completion, <5% contamination) were reconstructed with metaWRAP (*28*) and their abundances were tracked between samples. Pearson correlation comparison (two-sided t-test: $p<0.0001$) and group significance analysis (SigClust 2-group significance: $p<0.01$) confirmed the permanent shift in strain composition after the rain (Fig. 3B). While the strain composition did change during the post-rain recovery between 2016 and 2017, the resulting shift was more moderate when compared to the more drastic rearrangement immediately following the rain. Additionally, two conditionally rare taxa (*29*) of *Cyanobacteria* that were previously reported in only a small fraction of halite nodules (*19*), were found in high abundances in most of the samples after the rain (Fig. S5). 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 fine-grain taxonomy of the community has been permanently reshuffled.

The rain disrupted taxonomic membership of functional niches

To investigate the basis of the functional rearrangement of the halite community after the rain, we introduced a strain rearrangement index (*RI*), which quantifies the turnover of strains contributing to each community function. To compute the *RI*, genes from each KEGG Orthology identifier were catalogued and their abundances in each sample estimated from the read coverage of the contig that they were on. The standardized average net change in gene abundances that carry a given function between two samples represents the degree of taxonomic turnover within that functional category (see Methods). A relatively high *RI* for a given community function indicates that it is carried by different community members between two samples, but does not necessarily imply a high net change in its total abundance in the samples. Therefore, the distribution in *RIs* for all functions between two time-points quantifies changes in niche representation over that time (Fig. 3D). 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 return to its initial state. These findings indicate that functional redundancy of community members ensured a robust functional landscape in the halite microbial communities despite change in individual strains.

DISCUSSION

The response and recovery of the halite microbiome, a sensitive extremophile ecosystem, provided the perfect opportunity to characterize a natural community response dynamics to changing environmental conditions. A major rainfall proved to be devastating to the halophiles found within the salt nodules of Salar Grande, as was also found for other desert microbiomes that evolved to endure prolonged desiccation (*26*). The surviving community was comprised of organisms with higher average isoelectric points (*pI*) of their proteomes and average lower potassium uptake potential. Low proteome *pI* and high potassium uptake rates are features of salt-in strategists, as these adaptations allow them to balance high external salt concentrations (*22, 30*). Our observations suggest that the rain temporarily decreased the salt concentrations within the colonized pores (*24, 31*), rapidly changing the osmotic conditions within. We hypothesize that this

led to a mass death event of organisms poorly adapted to large osmotic changes immediately following the rain, while giving others an advantage.

The taxonomic rearrangement at the strain level after the rain was likely driven by neutral (i.e. random) processes (*32, 33*) similar to those governing the initial colonization of halite nodules. These rearrangements resulted in high inter-nodule taxonomic diversity (*19*) while the functional states remained. We suggest that each nodule was stochastically colonized by random draw, from the seed bank, of competitively equivalent organisms. A seed bank is a diverse genetic reservoir consisting of a large collection of low-abundance organisms (*1, 34*) and they are critical in microbiome functioning, particularly during prolonged unchanging environmental conditions, such as the past 13 years prior to the rain in northern Atacama. Seed banks conserve genetic and functional diversity, which in turn allows for rapid adaptation and restructuring of the microbial community following a drastic perturbation.

While the halite microbiome was able to recover from this catastrophic event, the effects of the perturbation lasted remarkably long (months), in contrast with studies in other desert systems where much quicker recoveries were documented (weeks) (*13*). This highlights the slow-growing nature of these extremophiles and suggests that the immediate effects of the rain on the halite community may have been even more dramatic than what we observed 6-months post-rain (*18, 35*). Eighteen months post-rain, the community was comprised of an entirely new set of organisms, but its functional potential recovered to a pre-rain state, suggesting that the community taxonomic structure entered an alternative equilibrium state during the recovery period (*4, 12*). The functional consistency of a community, disconnected from taxonomic variance, has previously been documented in a variety of microbiomes and stems from functional redundancy of closely related taxa (*6-8, 10*). In particular, isolated microbiomes such as miniature aquatic ecosystems found in bromeliad rosettes (similarly isolated as 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 (*9, 36*).

The pre-rain (2014) and recovered (2017) communities were very similar in terms of their functionally potential, while the intermediate state (2016) was very distinct (Fig. 1C, D). Therefore, the two compositional shifts that the halite microbiomes underwent following the rain – the initial response (2015-2016) and subsequent recovery (2016-2017) – resulted in a similar magnitude of change to the overall functional potential of the community. Taxonomically however, the two shifts were fundamentally distinct, as the individual taxa membership was drastically rearranged during the initial response to the rain but stayed unchanged during the recovery (Fig. 3B.C).

. The two different mechanisms by which the halite communities achieved almost identical net change in their functional potential as they entered and then exited their intermediate state (*12, 13*) offered a uniquely detailed view of microbial adaptation dynamics. These two types responses, or modes, allowed for inference of a general microbiome adaptation model, which can be potentially applied to explain and predict the taxonomic and functional flux in other ecosystems following major environmental changes (Fig. 4). The first mode (*Type I*; Fig. 4A) is a community rearrangement, resulting from adaptations to an acute major perturbation. In the halite nodules, the rain presented a major stress on the pre-existing communities by temporarily lowering external osmotic conditions and exerting a strong selective pressure on the salt-in strategists of the community. This produced gaps in existing functional niches and presented an opportunity for new organisms from the seed bank to come in through niche intrusion (*37*). The *Type I* shift is driven by neutral (random) processes characterized by changes in fine-scale (i.e. strains) taxonomic composition, which results in a high strain rearrangement index (*RI*).

The second mode (*Type II*; Fig. 4B) is an adjustment in existing community structure, and results from gradual changes in environmental conditions. After the rain passed and the osmotic conditions within the halite nodules returned to their initial levels, the halite community gradually returned to

its previous functional potential. However, because there were no major stress events to reset the strain composition of the communities, the newly dominant strains remained relatively unchanged during the recovery period. Instead, the communities achieved the desired functional potential through gradual changes in relative abundances of major taxa (Fig. 1, S2, S3), the strain composition of which remained unchanged. The taxonomic mechanisms behind the *Type II* response is relatively deterministic, as currently dominant taxa have the opportunity adjust their relative abundances based on fitness under the new selective pressures, preventing new organisms to take over. As a result, the strain composition of these major taxa remain largely unchanged, resulting in a low *RI*. In the halite microbiome, the *Type I* and a *Type II* shifts occurred in succession, leading the community first through an unstable intermediate state and then into an alternate equilibrium state (*4*). This intermediate dis-equilibrium intermediate has been reported in a number of communities after disaster events (*38*) or antibiotic administration (*37, 39*), but until now was difficult to investigate closely in natural ecosystems because of compounding complexity and fast microbial growth rates (*1, 4*). We postulate that *Type I* and *Type II* shifts observed in our model microbiome are integral to analogous structural rearrangement in other systems.

It is important to note that *Type I* and *Type II* functional shifts do not necessarily follow one another. If the initial environmental conditions are not re-established after a perturbation, such as after a permanent introduction of irrigation to desiccated soils, a *Type I* shift will most likely be the main mechanism for community adaptation, driven by the changes environmental conditions. Alternatively, in systems where environmental conditions shift gradually, such as aquatic microbiomes during seasonal changes, Type II shifts will likely drive the changes in the community's functional potential. We propose that *RI* measurements of such shifts may be useful in future studies to categorize such dynamics.

In conclusion, the tractable nature of our model microbiome allowed us to extrapolate general mechanisms of community response and resilience to acute shock. We demonstrated that a major disturbance can result in stochastic re-population of the community's functional niches, forcing a microbial community structure into an unstable intermediate. During the recovery period the newly dominant taxa adjust in abundance to reproduce the initial functional potential, allowing the community to enter an alternative equilibrium. Understanding the mechanisms behind the response and recovery components of microbial perturbation responses are vital to generally model and predict the taxonomic and functional flux of ecosystems following natural and man-made ecological disasters. Our proposed characterization and quantitation of two types of community shifts and our two-step model for community resilience can provide a framework for future work in predictive modeling of microbial communities.

MATERIALS AND 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 (*17*). All the sites were within 5 km of each other and, at each site, halite nodules were harvested within a 50m^2 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 (*17*) 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 (*17, 18*) 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 (*17*) by amplifying the hypervariable V3-V4 region with 515F and 926R primers (*40*). 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 (*41*). 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 (*42*) release as reference and USEARCH v6.1.554 (*43*). 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 (*44*) (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. 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 (*28*) 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 (*45*) (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) (*46*). 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 (*47*) module (--bins option specified), classified with the Classify_bins module (default parameters), and quantified by Salmon (*48*) with the Quant_bins module (default parameters). Contig read depth was estimated for each sample with the metaWRAP'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) (*49*) annotation service. Gene relative abundances were taken as the average read depth of the contigs carrying those genes (estimated with Salmon (*48*). KEGG KO identifiers were linked to their respective functions using the KEGG BRITE pathway classification (*50*). 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 (*51*) with the use of metaWRAP (*28*), and the *pI* of each ORF was calculate with ProPAS (*52*). 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 (Fig. 1C) and co-assembly contig abundances (Fig. 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 (*44*) (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) (*53*). 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.

H2: Supplementary Materials

Fig. S1. Average taxonomic composition of halite microbial communities from Site 1 sampled at different dates, estimated from WMG reads with KRAKEN and visualized with KronaTools.

Fig. S2. 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 plot of a Weighted Unifrac dissimilarity matrix comparing taxonomic composition. Error bars represent standard deviation; significance bars represent group signifficance based on a two tail t-test, and stars denote the p-value thresholds (*=0.01, **=0.001, ***=0.0001).

Fig. 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. Error bars represent standard deviation; significance bars represent group signifficance based on a two tail t-test, and stars denote the p-value thresholds (*=0.01, **=0.001, ***=0.0001).

Fig. S4. 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.

Fig. S5. 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.

Table S1. Description of sampling locations, dates, and replicate counts of biological samples collected for this study.

References and Notes

1. A. Shade *et al.*, Fundamentals of microbial community resistance and resilience. *Front Microbiol* **3**, 417 (2012).
2. F. Raymond, M. Deraspe, M. Boissinot, M. G. Bergeron, J. Corbeil, Partial recovery of microbiomes after antibiotic treatment. *Gut Microbes* **7**, 428-434 (2016).
3. L. A. David *et al.*, Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559-563 (2014).
4. M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, B. Walker, Catastrophic shifts in ecosystems. *Nature* **413**, 591-596 (2001).
5. S. D. Jurburg *et al.*, Legacy Effects on the Recovery of Soil Bacterial Communities from Extreme Temperature Perturbation. *Front Microbiol* **8**, 1832 (2017).
6. 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).
7. J. E. Goldford *et al.*, Emergent Simplicity in Microbial Community Assembly. *bioRxiv*, (2017).
8. A. Eng, E. Borenstein, Taxa-function robustness in microbial communities. *Microbiome* **6**, 45 (2018).
9. S. Louca *et al.*, High taxonomic variability despite stable functional structure across microbial communities. *Nat Ecol Evol* **1**, 15 (2016).
10. Y. Nie *et al.*, Species Divergence vs. Functional Convergence Characterizes Crude Oil Microbial Community Assembly. *Front Microbiol* **7**, 1254 (2016).
11. C. Jernberg, S. Lofmark, C. Edlund, J. K. Jansson, Long-term impacts of antibiotic exposure on the human intestinal microbiota. *Microbiology* **156**, 3216-3223 (2010).
12. 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).
13. A. Armstrong *et al.*, Temporal dynamics of hot desert microbial communities reveal structural and functional responses to water input. *Sci Rep* **6**, 34434 (2016).
14. 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).
15. 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).
16. 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).
17. 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).
18. A. Crits-Christoph *et al.*, Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community. *Environ Microbiol* **18**, 2064-2077 (2016).
19. 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).

20. E. F. Mongodin *et al.*, The genome of Salinibacter ruber: Convergence and gene exchange among hyperhalophilic bacteria and archaea. *PNAS*, 0509073102 (2005).
21. C. Monard, S. Gantner, S. Bertilsson, S. Hallin, J. Stenlid, Habitat generalists and specialists in microbial communities across a terrestrial-freshwater gradient. *Sci Rep* **6**, 37719 (2016).
22. A. Oren, Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. *Front Microbiol* **4**, 315 (2013).
23. R. S. Thombre, V. D. Shinde, R. S. Oke, S. K. Dhar, Y. S. Shouche, Biology and survival of extremely halophilic archaeon Haloarcula marismortui RR12 isolated from Mumbai salterns, India in response to salinity stress. *Sci Rep* **6**, 25642 (2016).
24. A. F. Davila *et al.*, In situ metabolism in halite endolithic microbial communities of the hyperarid Atacama Desert. *Front Microbiol* **6**, 1035 (2015).
25. Dirección Meteorológica de Chile, *Servicios Climáticos* (2018 https://climatologia.meteochile.gob.cl/application/index/productos/RE2009).
26. A. Azua-Bustos *et al.*, Unprecedented rains decimate surface microbial communities in the hyperarid core of the Atacama Desert. *Sci Rep* **8**, 16706 (2018).
27. N. Schulz, J. P. Boisier, P. Aceituno, Climate change along the arid coast of northern Chile. *International Journal of Climatology* **32**, 1803-1814 (2012).
28. G. V. Uritskiy, J. DiRuggiero, J. Taylor, MetaWRAP - a flexible pipeline for genome-resolved metagenomic data analysis. *bioRxiv*, (2018).
29. A. Shade *et al.*, Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. *Mbio* **5**, (2014).
30. 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).
31. 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).
32. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography.* . (Princeton Univ Press, Princeton: New Jersey, 2001).
33. L. Li, Z. S. Ma, Testing the Neutral Theory of Biodiversity with Human Microbiome Datasets. *Sci Rep* **6**, 31448 (2016).
34. J. T. Lennon, S. E. Jones, Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* **9**, 119-130 (2011).
35. 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).
36. S. Louca *et al.*, Function and functional redundancy in microbial systems. *Nature Ecology & Evolution* **2**, 936-943 (2018).
37. S. R. Modi, J. J. Collins, D. A. Relman, Antibiotics and the gut microbiota. *J Clin Invest* **124**, 4212-4218 (2014).
38. 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).
39. M. O. Sommer, G. Dantas, Antibiotics and the resistant microbiome. *Curr Opin Microbiol* **14**, 556-563 (2011).

40. D. M. Needham, J. A. Fuhrman, Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nat Microbiol* **1**, 16005 (2016).
41. J. G. Caporaso *et al.*, QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-336 (2010).
42. 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).
43. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461 (2010).
44. M. Waskom *et al.* (GitHub, 2017), pp. https://github.com/mwaskom/seaborn.
45. D. E. Wood, S. L. Salzberg, Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* **15**, R46 (2014).
46. S. Nurk, D. Meleshko, A. Korobeynikov, P. A. Pevzner, metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**, 824-834 (2017).
47. 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).
48. 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).
49. I. A. Chen *et al.*, IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res* **45**, D507-D516 (2017).
50. 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).
51. D. Hyatt *et al.*, Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
52. S. Wu, Y. Zhu, ProPAS: standalone software to analyze protein properties. *Bioinformation* **8**, 167-169 (2012).
53. 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).

Acknowledgments

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

Funding: 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 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.

Competing interests: The authors declare no competing interests.

Data and materials availability: 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.

Figures and Tables

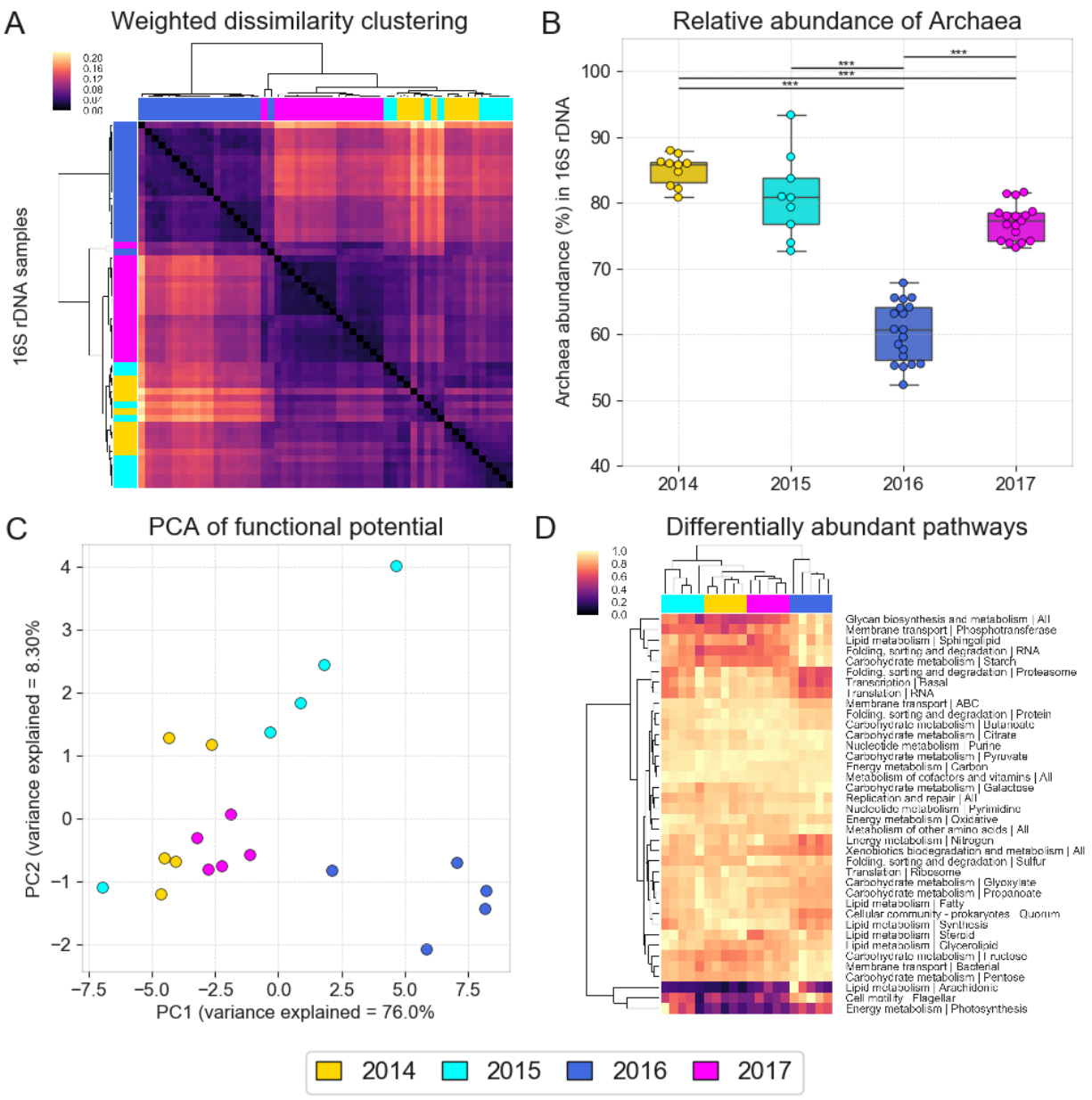

Fig. 1. Taxonomic and functional resilience after recovery period. Taxonomic composition of halite microbiomes over time, shown by (A) hierarchical clustering (correlation metric) of the Weighted Unifrac dissimilarity matrix and (B) the average relative abundance of archaeal sequences, based on 16S rDNA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

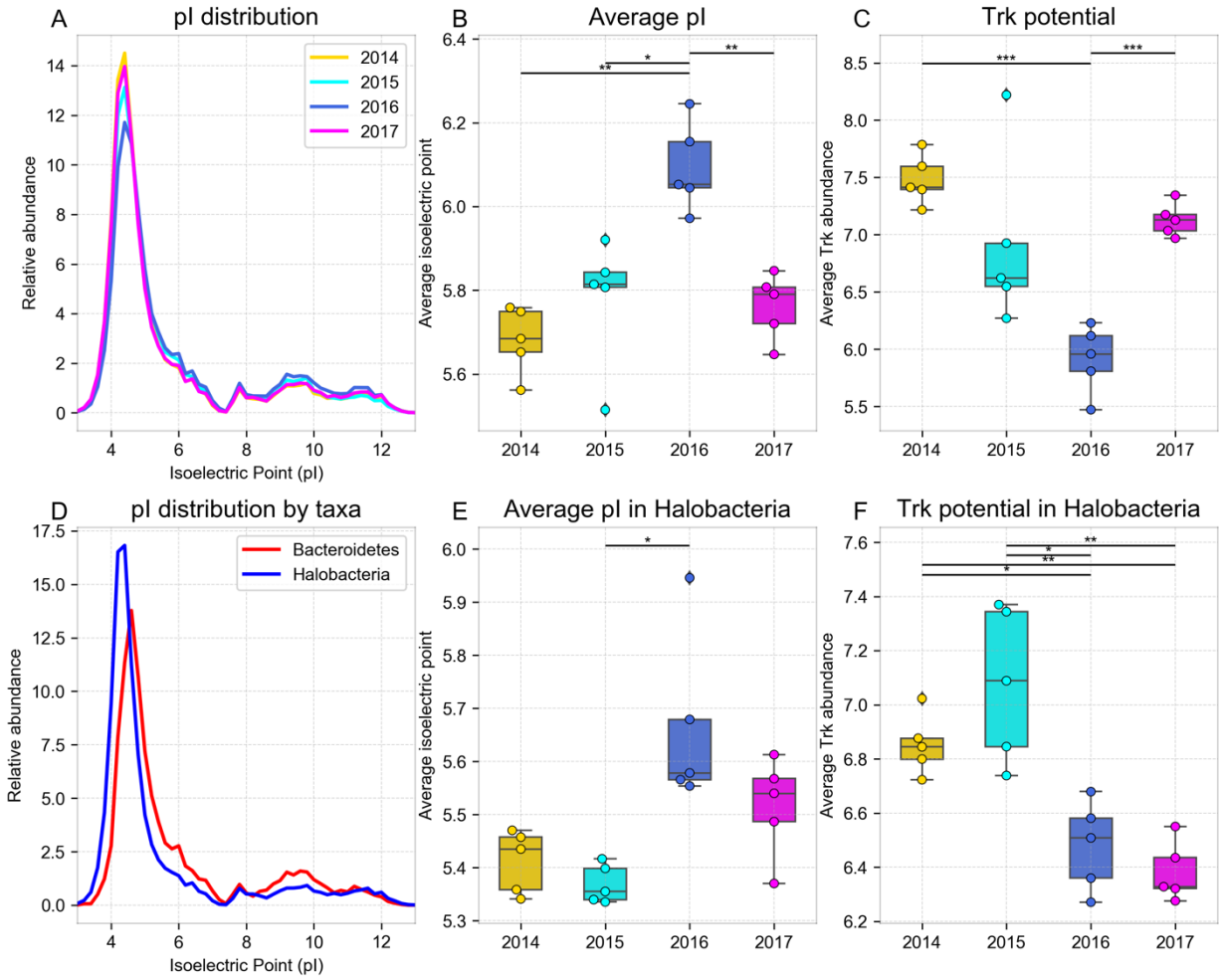

Fig. 2. Proteome adaptations across time. Analysis of the isoelectric points (*pI*) of proteins encoded in replicates of WMG assemblies from samples harvested at different dates, showing (A) the overall weighted distribution of the protein *pIs*, and the weighted average *pI* of proteins encoded in (B) 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. Error bars represent standard deviation; significance bars represent group signifficance based on a two tail t-test, and stars denote the p-value thresholds (*=0.01, **=0.001, ***=0.0001).

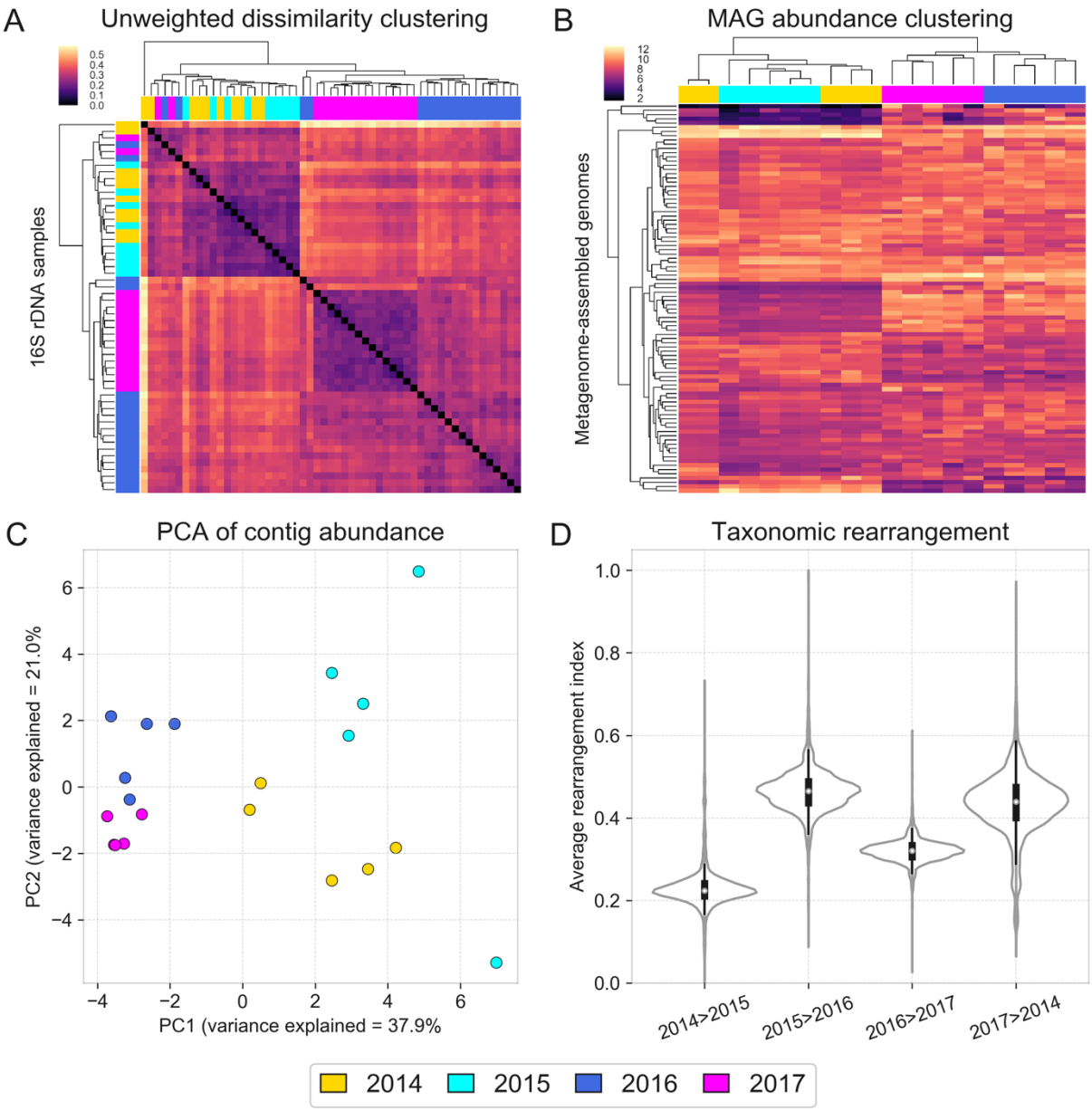

Fig. 3. Strain membership rearrangement. Changes in fine-scale composition of halite communities over time shown with (A) hierarchical clustering (correlation metric) of an Unweighted Unifrac dissimilarity matrix (based on 16S rDNA amplicon sequencing), (B) hierarchical clustering (Euclidean metric) 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.

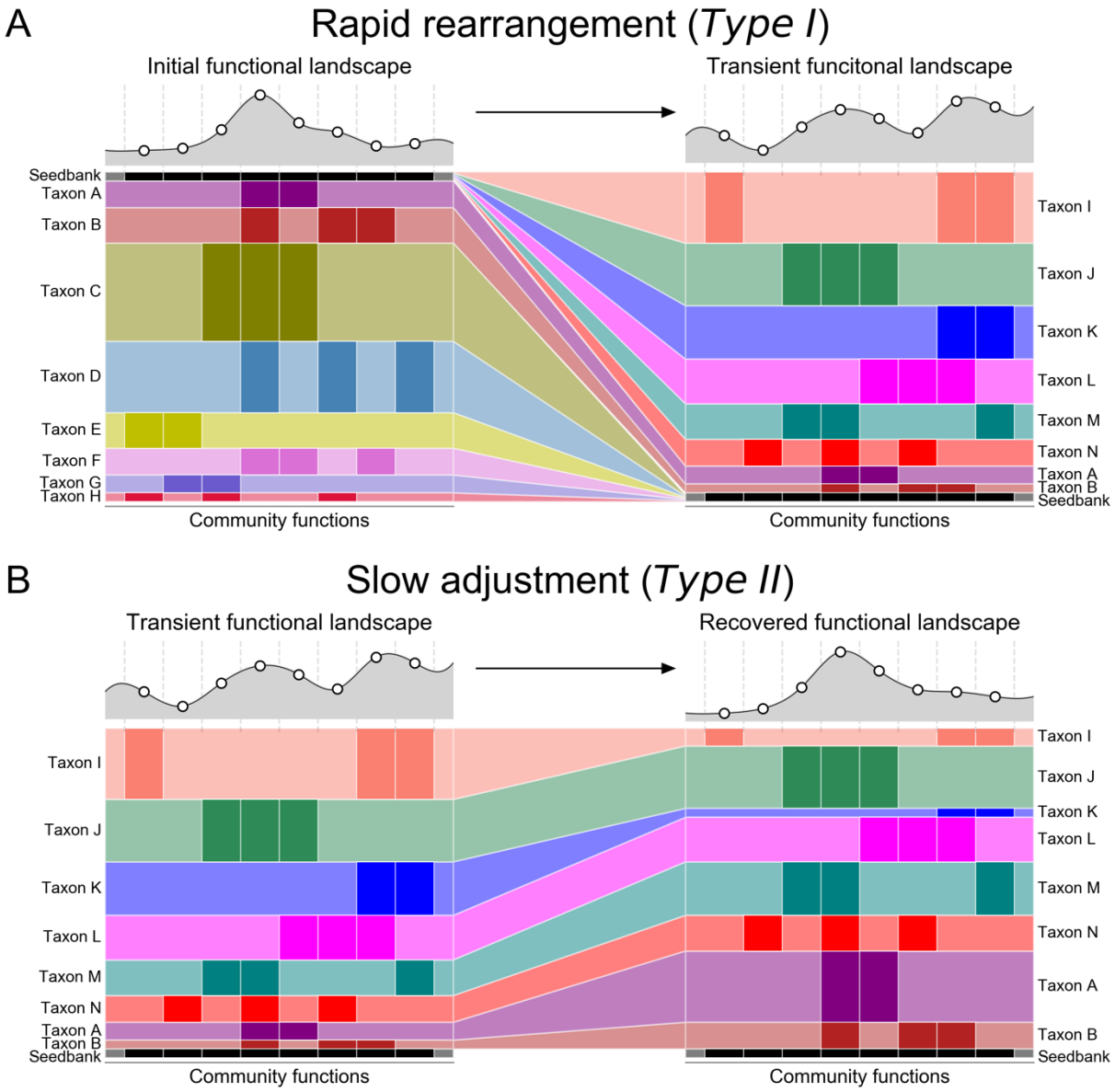

Fig. 4. Microbial community resilience model. Models of a microbiome adapting its functional potential in response to changing environmental conditions either with (A) a rapid rearrangement of the community's taxonomic structure resulting from new organisms from the seed bank displacing most previously dominant taxa through niche intrusion (as seen in the initial shock from the rainfall), or with (B) a gradual adjustment in relative abundance of major taxa (as seen in the halite community recovery).

Supplementary Materials

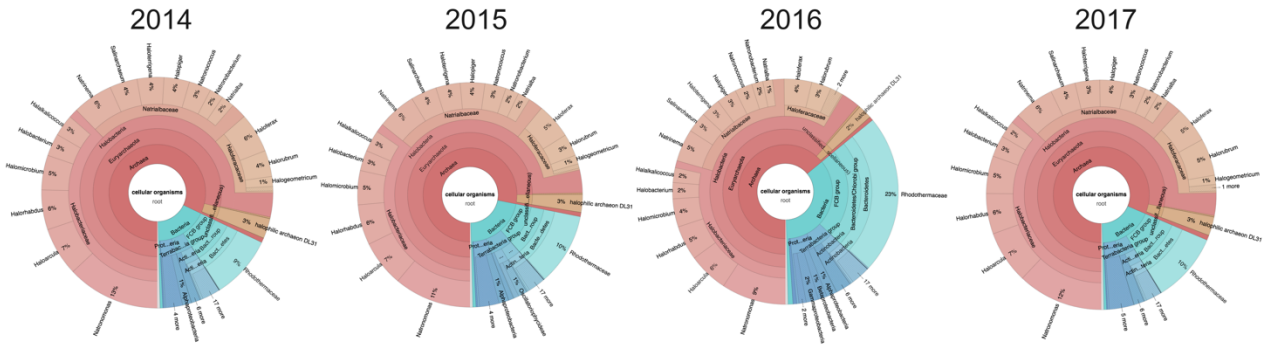

Fig. S1. Average taxonomic composition of halite microbial communities from Site 1 sampled at different dates, estimated from WMG reads with KRAKEN and visualized with KronaTools.

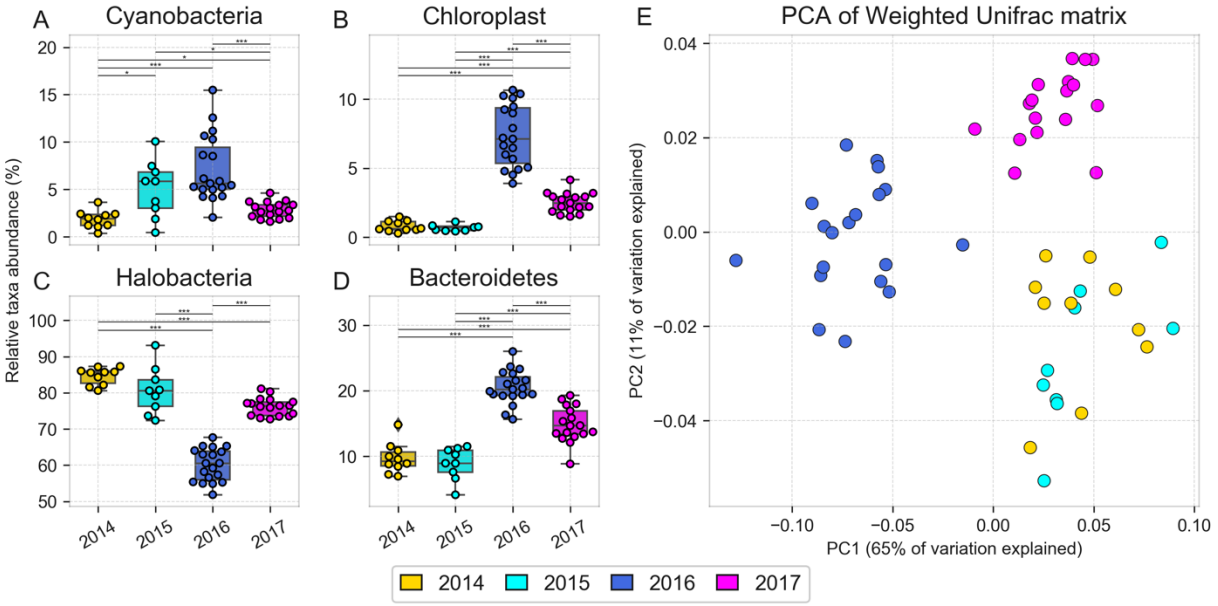

Fig. S2. Taxonomic composition differences between halite samples harvested from Site 1 at different dates, infered from 16S rDNA sequences clustered into OTUs at 97% identity and visualized through (A-D) relative abundance of major differentially abundant phyla and a (E) PCA plot of a Weighted Unifrac dissimilarity matrix comparing taxonomic composition. Error bars represent standard deviation; significance bars represent group signifficance based on a two tail t-test, and stars denote the p-value thresholds (*=0.01, **=0.001, ***=0.0001).

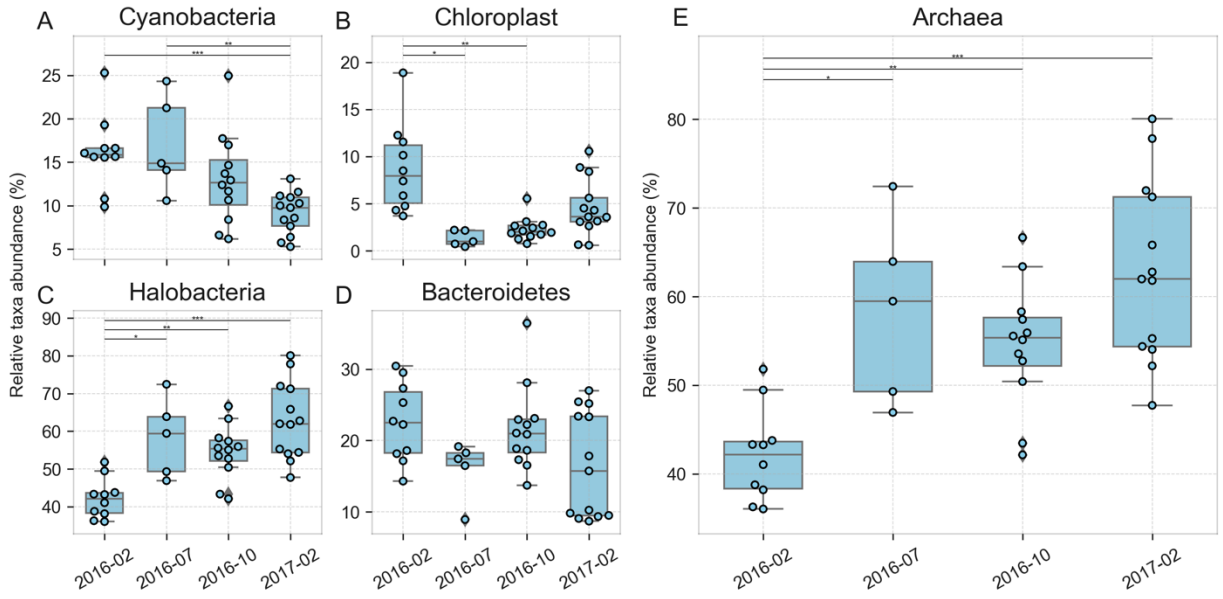

Fig. S3. Taxonomic composition differences between halite samples harvested from Site 2 at different dates post-rain, infered from 16S rDNA sequences clustered into OTUs at 97% identity and visualized through (A-D) relative abundance of major differentially abundant phyla and (E) archaea abundance. Error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

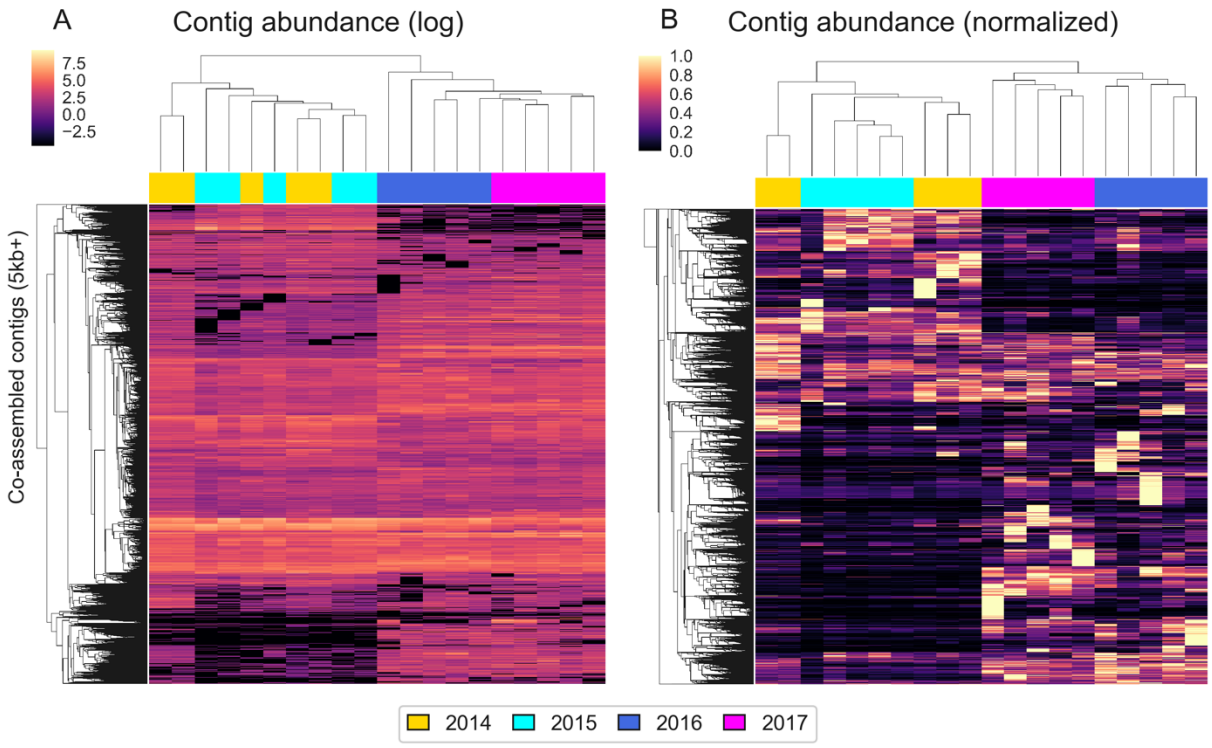

Fig. S4. Hierarchical clustering (Euclidean metric) of relative abundances (fragments per million) of contigs > 5kbp in the WMG co-assembly, quantified with reads from samples harvested at different dates and displayed on (A) a log scale and (B) standardized to the maximum abundance of each contig.

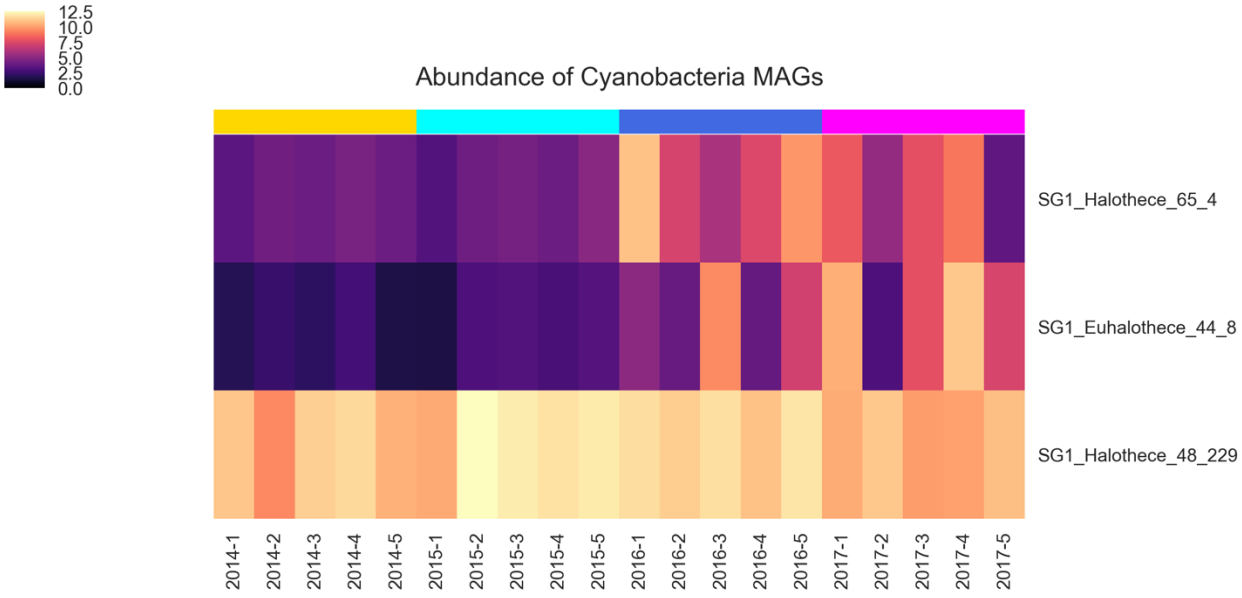

Fig. S5. Hierarchical clustering (Euclidean metric) of photosynthetic MAG relative abundances (fragments per million), quantified with metaWRAP's quant_bins module, showing the emergence of two new *Cyanobacteria* MAGs after the rain.

Site	Latitude	Longitude	Elevation (asl)	Collection dates	16S rDNA replication	WMG replication	Purpose
S1	20°57’ 12.006”S	70°1’ 10.5996”W	680m	Sep-14	10	5	Before-after rain comparison
				Jun-15	9	5	
				8-Feb-16	19	5	
				20-Feb-17	17	5	
S2	20°57’ 8.5212”S	70°1’ 1.2612”W	664m	8-Feb-16	12	NA	After rain recovery process
				11-Jul-16	5	NA	
				20-Oct-16	12	NA	
				20-Feb-17	13	NA	
S3	20°55’ 48.18”S	70°0’ 49.32”W	676m	Misc.	NA	15	Assembly and binning improvement

Table S1. Description of sampling locations, dates, and replicate counts of biological samples collected for this study.