frontiers
in Microbiology

Phylogenetic and Functional Substrate Specificity for Endolithic Microbial Communities in Hyper-Arid Environments

Alexander Crits-Christoph[1], Courtney K. Robinson[1], Bing Ma[2], Jacques Ravel[2], Jacek Wierzchos[3], Carmen Ascaso[3], Octavio Artieda[4], Virginia Souza-Egipsy[5], M. Cristina Casero[3] and Jocelyne DiRuggiero[1]*

[1] Biology Department, The Johns Hopkins University, Baltimore, MD, USA, [2] Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA, [3] Department of Biochemistry and Microbial Ecology, Museo Nacional de Ciencias Naturales – Consejo Superior de Investigaciones Científicas, Madrid, Spain, [4] Universidad de Extremadura, Plasencia, Spain, [5] Instituto de Ciencias Agrarias – Consejo Superior de Investigaciones Científicas, Madrid, Spain

Under extreme water deficit, endolithic (inside rock) microbial ecosystems are considered environmental refuges for life in cold and hot deserts, yet their diversity and functional adaptations remain vastly unexplored. The metagenomic analyses of the communities from two rock substrates, calcite and ignimbrite, revealed that they were dominated by *Cyanobacteria*, *Actinobacteria*, and *Chloroflexi.* The relative distribution of major phyla was significantly different between the two substrates and biodiversity estimates, from 16S rRNA gene sequences and from the metagenomic data, all pointed to a higher taxonomic diversity in the calcite community. While both endolithic communities showed adaptations to extreme aridity and to the rock habitat, their functional capabilities revealed significant differences. ABC transporters and pathways for osmoregulation were more diverse in the calcite chasmoendolithic community. In contrast, the ignimbrite cryptoendolithic community was enriched in pathways for secondary metabolites, such as non-ribosomal peptides (NRP) and polyketides (PK). Assemblies of the metagenome data produced population genomes for the major phyla found in both communities and revealed a greater diversity of *Cyanobacteria* population genomes for the calcite substrate. Draft genomes of the dominant *Cyanobacteria* in each community were constructed with more than 93% estimated completeness. The two annotated proteomes shared 64% amino acid identity and a significantly higher number of genes involved in iron update, and NRPS gene clusters, were found in the draft genomes from the ignimbrite. Both the community-wide and genome-specific differences may be related to higher water availability and the colonization of large fissures and cracks in the calcite in contrast to a harsh competition for colonization space and nutrient resources in the narrow pores of the ignimbrite. Together, these results indicated that the habitable architecture of both lithic substrates-chasmoendolithic versus cryptoendolithic – might be an essential element in determining the colonization and the diversity of the microbial communities in endolithic substrates at the dry limit for life.

Keywords: endoliths, cyanobacteria, extreme environment, Atacama Desert, hyper-arid environment, metagenomics

INTRODUCTION

The rate of desertification across our planet is accelerating as the result of human activity and climate change. Ecosystems of arid and hyper-arid environments are highly vulnerable, including the microbial communities supporting these biomes. Under extreme water deficit and high solar radiation, endolithic (inhabiting rock) habitats are considered environmental refuges for life (Friedmann, 1982; Cary et al., 2010; Pointing and Belnap, 2012; Wierzchos et al., 2012). In these microbial ecosystems, the rock substrate provides protection from incident UV, excessive solar radiation, and freeze–thaw events while providing physical stability, and enhanced moisture availability (Walker and Pace, 2007; Chan et al., 2012). Endolithic microbial communities are photosynthetic-based with primary producers supporting a diversity of heterotrophic microorganisms (Walker and Pace, 2007; Wierzchos et al., 2012).

Habitats as different as the underside of quartz rocks (Azúa-Bustos et al., 2011; Cowan et al., 2011; Chan et al., 2013; de los Ríos et al., 2014; Vikram et al., 2015), sandstones (Friedmann, 1982; de la Torre et al., 2003; Pointing et al., 2009), granite boulders (Wei et al., 2015), the inside of gypsum (Dong et al., 2007; Wierzchos et al., 2011, 2015) and halite evaporites (Wierzchos et al., 2006; Robinson et al., 2015), carbonaceous (DiRuggiero et al., 2013) and volcanic rocks (Wierzchos et al., 2013; Cámara et al., 2014) among others, have shown that life has found innovative ways to adapt to the extreme conditions of hyper-arid deserts.

The Atacama Desert is one of the driest deserts on Earth and its hyper-arid core has been described as "the most barren region imaginable" (McKay et al., 2003). In the Yungay area of the hyper-arid zone of the Atacama Desert, with decades between rainfall events and extremely low air relative humidity (RH; mean yr^{-1} values <35%), ancient halite crusts of evaporitic origin have been shown to provide sufficient moisture to sustain cryptoendolithic (within pores) microbial communities (Wierzchos et al., 2006; de los Ríos et al., 2010; Robinson et al., 2015). Another strategy for survival in this hyper-arid desert is the colonization of cracks and fissures of rock substrates, also called chasmoendolithic colonization (Dong et al., 2007; DiRuggiero et al., 2013). Our work in the hyper-arid core of the Atacama Desert revealed that such communities inhabit gypsum covered rhyolite and calcite rocks, where the complex network of cracks and fissures of the rocks promote water retention (DiRuggiero et al., 2013). Calcite is highly translucent, with a network of large connected fissures providing abundant space for microbial colonization within a few cm of the rock surface (DiRuggiero et al., 2013). We found that the diversity of the calcite community was significantly increased when the water budget of the ecosystem was potentially enhanced by dew formation on the rock in contrast to the rhyolite community, where the unique source of water were scarce rainfalls (DiRuggiero et al., 2013). More recently, we discovered communities colonizing ignimbrite, a pyroclastic rock composed of plagioclase and biotite embedded in a matrix of weakly welded glass shards (Wierzchos et al., 2013). In the ignimbrite the colonization is cryptoendolithic, with microorganisms colonizing small pores within the rock and only within a few mm of the rock surface (Wierzchos et al., 2013; Cámara et al., 2014).

From previous work on microbial communities in hyper-arid deserts, it is clear that the climate regime plays a major role in determining the habitability of a given substrate. However, the susceptibility of rocks to colonization also depends on the architecture of the lithic substrate (Wierzchos et al., 2015), including factors such as translucence, which allows transmission of photosynthetically active radiation (PAR), thermal conductivity, the presence of a network of pores and/or fissures connected to the rock surface, which is linked to the capacity at retaining water, and chemical composition. These factors are likely to impact greatly microbial colonization and diversity (Walker and Pace, 2007; Herrera et al., 2009; Cockell et al., 2011; Wierzchos et al., 2012).

While a number of studies have addressed hypolithic communities in hot and cold Deserts (reviewed in Makhalanyane et al., 2015), very little information is available about what rock properties are driving endolithic communities functioning and diversity. Here we used a metagenomic analysis of endolithic microbial communities in calcite (chasmoendolithic) and ignimbrite (cryptoendolithic) rocks collected in the extreme environment of the hyper-arid core of the Atacama Desert. The taxonomic and functional composition of the communities were characterized and interpreted within the context of the structural properties of each rock type.

MATERIALS AND METHODS

Sampling and Site Characterization

Calcite (composed mainly of calcium carbonate) rocks (VL) were collected in 2013 near Valle de la Luna (22°54′42″ S and 068°14′51″ W) and ignimbrite (pyroclastic volcanic glass) rocks (IG) were collected in the Lomas de Tilocalar area (23°57′25″ S and 068°10′13″ W) of the Atacama Desert in northern Chile. Sampling of the rocks with endolithic microbial communities was performed on April 22 and 23, 2013, respectively, with air temperature of 30°C and 10% of RH. Dry rock samples were stored in sterile Whirl-Pak plastic bags, at room temperature and under dark conditions, for future laboratory processing. DNA extractions and sample preparation for microscopy were carried out within two weeks of sample collection. Additional ignimbrite rocks were also collected from a nearby area 25 km to the North-East of Lomas to Tilocalar (PING) on April 24, 2013 under similar conditions. Microclimate data for the Lomas de Tilocalar area was collected in situ from April 2011 to 2013 using an Onset HOBO® Weather Station Data Logger (H21-001) as previously described (Wierzchos et al., 2013). Solar flux was measured using a PAR sensor for wavelengths of 400–700 nm (Wierzchos et al., 2013). Microclimate data for Valle de la Luna was extracted from historical records for the village of San Pedro de Atacama located 5 km east of our sampling site, as reported by DiRuggiero et al. (2013).

Microscopy Analyses

Colonized rock samples were processed for scanning electron microscopy in backscattered electron mode (SEM-BSE) observation and/or energy dispersive X-ray spectroscopy (EDS) microanalysis according to methods by Wierzchos and Ascaso (1994) and Wierzchos et al. (2011). SEM-BSE was used in combination with EDS to characterize the lithic substrates. Rock samples were observed using a scanning electron microscope (DSM960 Zeiss; Carl Zeiss) equipped with a solid-state, four diodes BSE detector plus an auxiliary X-ray EDS microanalytical system (Link ISIS Oxford, UK).

Fluorescence microscopy (FM) in structural illumination microscopy mode (SIM) using DAPI nucleic acids stain was performed on cell aggregates gently isolated from the chasmoendolithic habitat (Wierzchos et al., 2011). The samples were examined using a fluorescence microscope (AxioImager M2, Carl Zeiss, Germany) in SIM mode with a ApoTome (commercial SIM by Zeiss) system for 3 dimensional (3D) visualization of cell aggregates (Wierzchos et al., 2011).

DNA Extraction and Sequencing

For 16S rRNA gene sequencing, total genomic DNA was extracted from rock powder using the PowerSoil DNA Isolation kit (MoBio Laboratories Inc., Solana Beach, CA, USA) from four calcite and four ignimbrite rocks. DNA was amplified using the barcoded universal primers 338F and 806R spanning the V3–V4 hyper-variable region. Amplicons from at least three amplification reactions were pooled together, purified, and sequenced using the Illumina MiSeq platform by the Genomics Resource Center (GRC) at the Institute for Genome Sciences (IGS), University of Maryland School of Medicine. For metagenomes, the DNA was extracted from four rocks collected at each site and for each substrate from above (VL and IG), pooled, and sequencing libraries were prepared using the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA, USA), with an average insert size of 400 bp, and sequencing was performed on the Illumina HiSeq2500 platform by the GRC at IGS.

16S rRNA Gene Sequences Analysis

Illumina paired-end sequences were processed using the QIIME package (v1.6.0) (Caporaso et al., 2010) as previously described (DiRuggiero et al., 2013). Alpha and beta diversity metrics were calculated based on OTUs at the 0.03% cutoff (OTUs0.03) in QIIME using a maximum sequencing depth of 2400 sequences per sample. Statistical testing using Non-parametric Mann–Whitney tests were performed to indicate confidence in similarities/differences observed.

Metagenome Analysis and Assembly

Illumina sequencing of the metagenomic libraries produced 126,820,576 and 100,086,740 paired-end reads for the calcite and ignimbrite samples, respectively. Ribosomal RNA sequence reads were removed using Bowtie (v1.0) (Langmead et al., 2009) and by mapping to SILVA reference database (Pruesse et al., 2007). Both paired-end reads were filtered out if only one read was mapped to the rRNA database. Raw reads with low-quality bases as determined by base calling (phred quality score of 20 that corresponds to an error probability of 1%) were trimmed from the end of the sequence, and the sequence longer than 75% of the original read length were retained. The reads were assigned to a taxonomy using PhyloSift (Darling et al., 2014) and Kraken (Wood and Salzberg, 2014) and functional content using HUMAnN (Abubucker et al., 2012) and the Metagenomics analysis server MG-RAST (Meyer et al., 2008). An assembly was produced for the metagenome using the IDBA-UD assembler for metagenomic sequencing data (Peng et al., 2012). We use a k range of 20–100 with a pre-correction step before assembly. Assembled contigs were grouped into potential draft genomes using MaxBin (Wu et al., 2014), an expectation-maximization (EM) algorithm using tetranucleotide frequencies, abundance levels, and single-copy gene analysis. Taxonomy was assigned to each bin using both PhyloSift and Kraken (Darling et al., 2014; Wood and Salzberg, 2014).

Phylogenetic Analysis

The phylogenetic position of cyanobacterial metagenomic bins was determined using marker genes (via Phylosift) that were conserved across metagenomic bins with at least 50% estimated genome completeness from MaxBin (Wu et al., 2014). The marker genes were aligned individually to reference genes from the UniProt database using MUSCLE (Edgar, 2004) and the individual alignments were concatenated. Low-quality regions of the alignment were removed using Gblocks (Castresana, 2000) and a maximum-likelihood phylogenetic tree was created using FastTree (Price et al., 2010).

Population Genome Annotation

Cyanobacterial genomic bins Ca9 and Ig12 were selected for annotation as draft genomes based on both estimated genome completeness and n50 values of the binned contigs. RAST server (Aziz et al., 2008) was used to annotate and compare the draft genomes to known references. Putative phage contigs incorrectly binned were manually removed based on the annotation. The antibiotics and Secondary Metabolite Analysis Shell antiSMASH 2 (Blin et al., 2013) was used for the prediction of non-ribosomal peptides (NRPS) structures and synthetic gene clusters. The NORINE database[1] was used to determine chemical structures for putative products of NRPS and synthetic gene clusters. The draft genomes were compared using protein products predicted with Prodigal (Hyatt et al., 2010) and BLASTP (Gish and States, 1993) with a cutoff of 80% amino acid identity between matches. Results from the draft genomes comparisons were visualized using Circos (Krzywinski et al., 2009).

Sequence Data and Availability

All sequences were deposited at the National Center for Biotechnology Information Sequence Read Archive under Bioproject ID PRJNA285514 and accession numbers SAMN03754650 (ignimbrite) and SAMN03754649 (Calcite). The MG-RAST report for the data is available under ID

---

[1]http://bioinfo.lifl.fr/norine/

4600829.3 and 4600830.3[2,3]. Completed assemblies, annotation, and phylogenetic trees are available at http://figshare.com/articles/Phylogenetic_and_Functional_Substrate_Specificity_for_Endolithic_Microbial_Communities_from_the_Atacama_Desert/1606253.

RESULTS

We characterized at the molecular level the endolithic microbial communities from two rock types from the Atacama Desert, calcite from Valle de La Luna and ignimbrite from Lomas de Tilocalar. Both areas are extremely dry with less than ∼ 25 mm rainfall each year and average air RH between 18.3 and 40.5% (**Table 1**). In both locales, temperatures reached values above 32°C and maximum PAR values were about 2500 μmol s^{-1}m^{-2}.
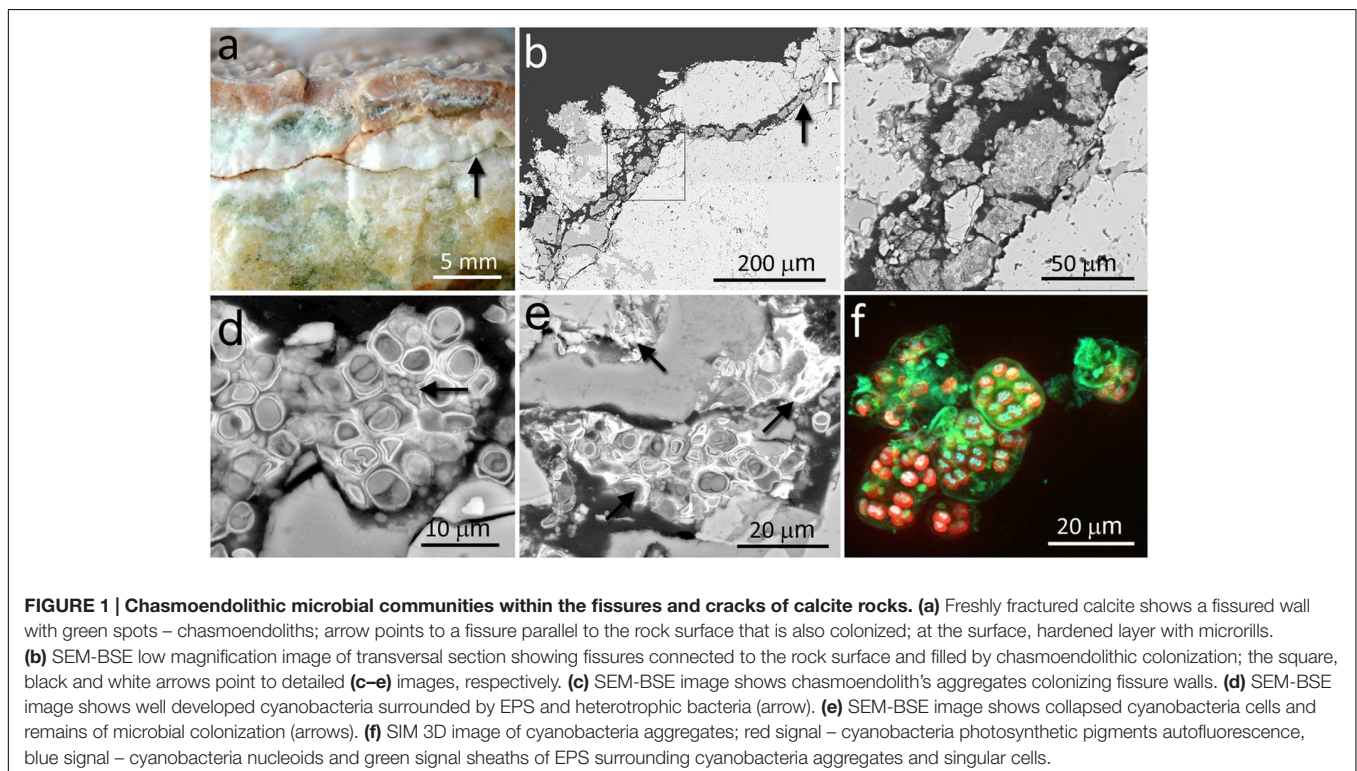
The calcite samples were composed of laminated calcite layers with a thickness of several centimeters. The calcite mineralogical composition and petrographic study were previously reported by DiRuggiero et al. (2013). The calcite rock surface was covered by a hardened semitransparent layer (up to 5 mm thick) with microrills forming small and short sinuous channels on the rock surface (**Figure 1a**). The most characteristic feature of the calcium carbonate rock was the presence of irregular narrow fissures and cracks that extended perpendicular and/or parallel to the rock surface (**Figures 1a,b**). A large number of these fissures and cracks, up to dozens of millimeters in length, were colonized by chasmoendolithic microorganisms detected upon fracturing the rock (**Figure 1a**). Additional fractures, parallel to the rock surface, also revealed similar chasmoendolithic colonization (black arrow on **Figure 1a**). Using SEM-BSE, we observed an almost continuous colonization of these fissures by chasmoendoliths (**Figures 1b–e**) and the connection of the fissures to the rock surface (**Figure 1**). We observed mostly well-developed cyanobacteria cells surrounded by one or two sheaths of capsular-like extracellular polysaccharidic substances

TABLE 1 | Microclimate data for the VL (calcite) and LdT (ignimbrite) sampling sites.

Sampling sites	PAR μmol m^{-2}s^{-1}		Air temperature (°C)			Air relative humidity (%)			Rainfall (mm)
	Average	Maximum	Mininum	Average	Maximum	Minimum	Average	Maximum	Average
Valle de la Luna(VL)[a]	–	2600	−1.8	13.1	32.2	16.7	40.5	80	27.8
Lomas to Tilocalar (LdT)[b]	442	2553	−4.6	15	44.8	1.2	18.3	100	25.6

Maximum, minimum, and yearly average values for photosynthetic active radiation (PAR), temperature, and relative humidity (RH); [a]*Data from DiRuggiero et al. (2013);* [b]*Data collected by HOBO weather station from April 2011–2013; "–" no data.*

FIGURE 1 | Chasmoendolithic microbial communities within the fissures and cracks of calcite rocks. (a) Freshly fractured calcite shows a fissured wall with green spots – chasmoendoliths; arrow points to a fissure parallel to the rock surface that is also colonized; at the surface, hardened layer with microrills. **(b)** SEM-BSE low magnification image of transversal section showing fissures connected to the rock surface and filled by chasmoendolithic colonization; the square, black and white arrows point to detailed **(c–e)** images, respectively. **(c)** SEM-BSE image shows chasmoendolith's aggregates colonizing fissure walls. **(d)** SEM-BSE image shows well developed cyanobacteria surrounded by EPS and heterotrophic bacteria (arrow). **(e)** SEM-BSE image shows collapsed cyanobacteria cells and remains of microbial colonization (arrows). **(f)** SIM 3D image of cyanobacteria aggregates; red signal – cyanobacteria photosynthetic pigments autofluorescence, blue signal – cyanobacteria nucleoids and green signal sheaths of EPS surrounding cyanobacteria aggregates and singular cells.

(EPSs) and heterotrophic bacteria present within some of the cyanobacteria aggregates (arrow in **Figure 1d**). Closer to the rock surface, cyanobacteria cells appeared deteriorated with collapsed cells and remains of microbial colonization (arrows in **Figure 1e**). Three-dimensional reconstruction of cyanobacteria showed characteristic microbial aggregates inside sack-like structures formed by EPSs (green signal in **Figure 1f**).

In contrast to the chasmoendolithic habitat found within the calcite fissures, the cryptoendolithic colonization within the volcanic ignimbrite presented a very different picture. The ignimbrite has a structure of vitrified foam identical to pumice rock with a large number of connected and non-connected bottle-shaped pores. The mineralogical, petrographic, and structural characteristics of the ignimbrite were previously reported by Wierzchos et al. (2013). The colonization zone of ignimbrite rocks was in the form of a narrow (1–3 mm thick) intensively green layer, parallel to the ignimbrite surface (**Figure 2a**), and within 2–3 mm of the rock surface. This brown-colored varnish rock covering the surface was composed of allochthonous clay minerals and filling some of the pores at the surface (white arrow in **Figure 2b**). Just below the ignimbrite surface cover, a cryptoendolithic colonization zone was observed (artificially cyan-colored spots in **Figure 2b**, black arrow) with many small pores. The microbial assemblage within the bottle-shaped pores was composed of cyanobacteria and heterotrophic bacterial cells, seemingly not associated with EPSs (**Figures 2c,d**). Some of the pores were not colonized whereas others were almost totally filled by cryptoendoliths (**Figure 2d**). Cyanobacteria formed aggregates and, in places, cells were associated with heterotrophic bacteria

(arrow in **Figure 2e**). In some pores, we also observed both damaged in healthy cyanobacteria together with heterotrophic bacteria, probably also damaged (**Figure 2f**).

Taxonomic Diversity of the Rock Communities

We obtained a total of 315,334 16S rRNA gene paired-end sequences, with an average read length of 408 bp, for four individual rocks from each of the two endolithic substrates, calcite, and ignimbrite. Taxonomic assignments of the sequence reads revealed that the dominant phyla, in both samples, were *Cyanobacteria*, *Actinobacteria*, and *Chloroflexi*, and that their relative distribution was significantly different between the two substrates (Supplementary Figure S1). Diversity metrics calculated for each community support the finding that the calcite substrate harbors a more diverse community in terms of richness and phylogenetic diversity (Supplementary Table S1 and Figure S2). Non-parametric Mann–Whitney tests reported that this difference is significant ($p < 0.05$) across all diversity metrics tested. Although the statistical power of this test is weak due to the small sample size of this comparison ($n = 8$), the ranges for each substrate of all diversity metrics are non-overlapping and distinct (Supplementary Table S1). Photosynthetic cyanobacteria dominated both communities and greater proportional abundances were found in all the ignimbrite samples than in the calcite samples; this was also supported by a non-parametric Mann–Whitney test ($p = 0.03$). Overall, 14 and 12 major OTUs of cyanobacteria (>1% abundance) were
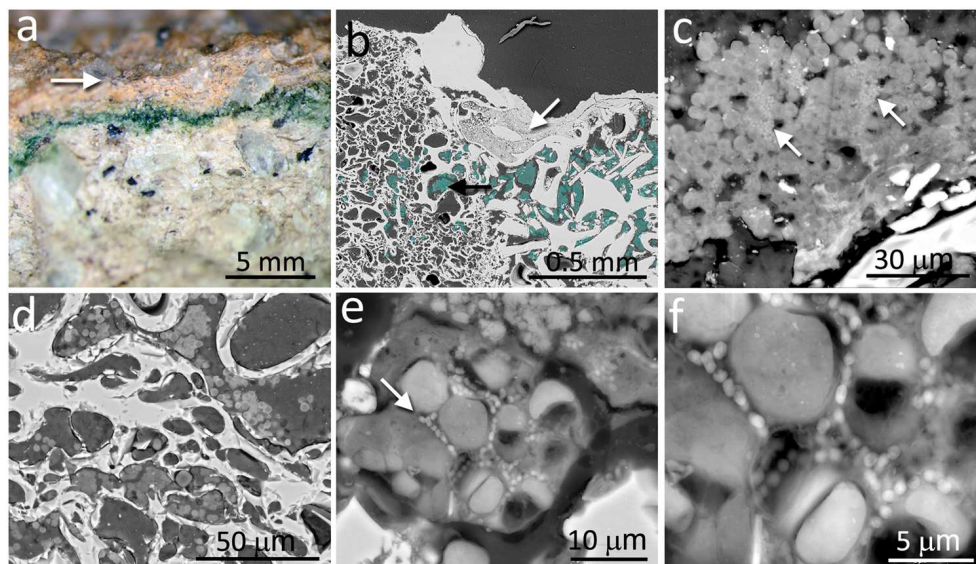

FIGURE 2 | Cryptoendolithic microbial communities within the pores of ignimbrite. (a) Freshly fractured ignimbrite shows a narrow green layer of cryptoendolithic microorganisms beneath the rock surface (arrow). **(b)** SEM-BSE image showing micromorphological features of rhyolitic ignimbrite composed of glass shards welded together forming bottle-shaped vesicles (left side of image) and irregular-shape pores (right side of image); white arrow points to surface pores filled by varnish rock mineral components; aggregates of cryptoendolithic microbial colonization were marked by cyan-color and black arrow points to aggregates shown in detail in **c**. **(c)** SEM-BSE image of cyanobacteria-composed aggregates containing heterotrophic bacteria (white arrows). **(d)** SEM-BSE image of empty and colonized bottle-shaped pores. **(e)** SEM-BSE image of a detailed view of cyanobacteria aggregates containing heterotrophic bacteria (white arrow). **(f)** Magnified central part of **(e)** showing heterotrophic bacteria in close position to collapsed cyanobacteria and their remains.

TABLE 2 | Number of sequence reads, Gbp of sequences, and contigs statistics for the calcite and ignimbrite metagenomes.

Sample	Number of paired-end reads[1]	Number of Gbp sequenced	Number of Gbp assembled	Number of contigs	Contig mean length (bp)	Contig maximum length (bp)
Ignimbrite (LdT)	100,086,740	10.1	0.22	166,509	1,351	101,619
Calcite (VL)	126,820,576	12.8	0.35	344,986	1,016	119,637

[1]*After quality control.*

identified in the calcite and ignimbrite communities, respectively, and none were shared across the two communities.
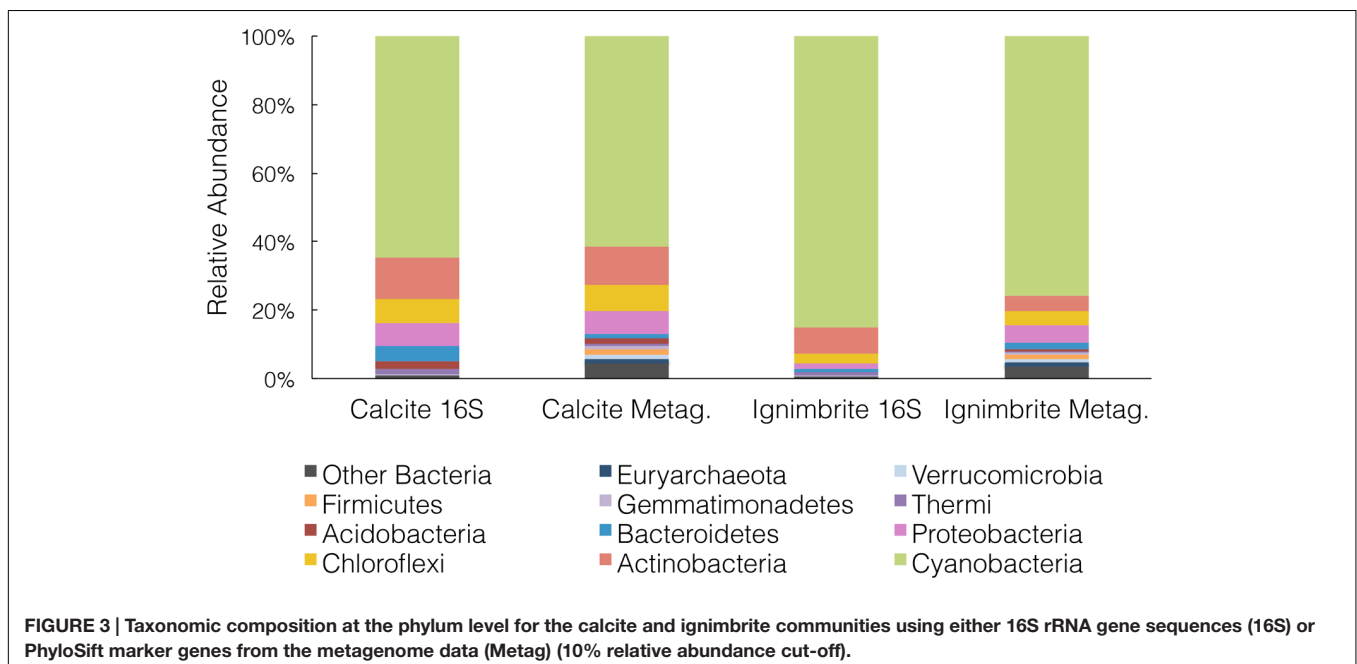
We generated over 22 Gbp of high quality, paired-end, metagenomic shotgun sequences for the calcite and ignimbrite substrates (**Table 2**). Taxonomic assignment of the metagenomic reads was performed with PhyloSift using all extracted marker genes (**Figure 3**). When compared to the 16S rRNA gene taxonomic distribution, averaged for each rock substrate, we found a general consensus between the two datasets (**Figure 3**), which supported a higher taxonomic diversity for the calcite community. In both substrates the metagenome analysis showed higher proportional abundance of rare taxa, probably as the result of variation in genome sizes across all taxa. Reconstruction of 16S rRNA gene sequences from the metagenomic dataset with EMIRGE (Miller et al., 2011) provided full-length, or nearly full-length, genes for all the major phyla observed in the 16S rRNA gene dataset (Supplementary Figure S3) that also indicated again a higher taxonomic diversity in the calcite community.
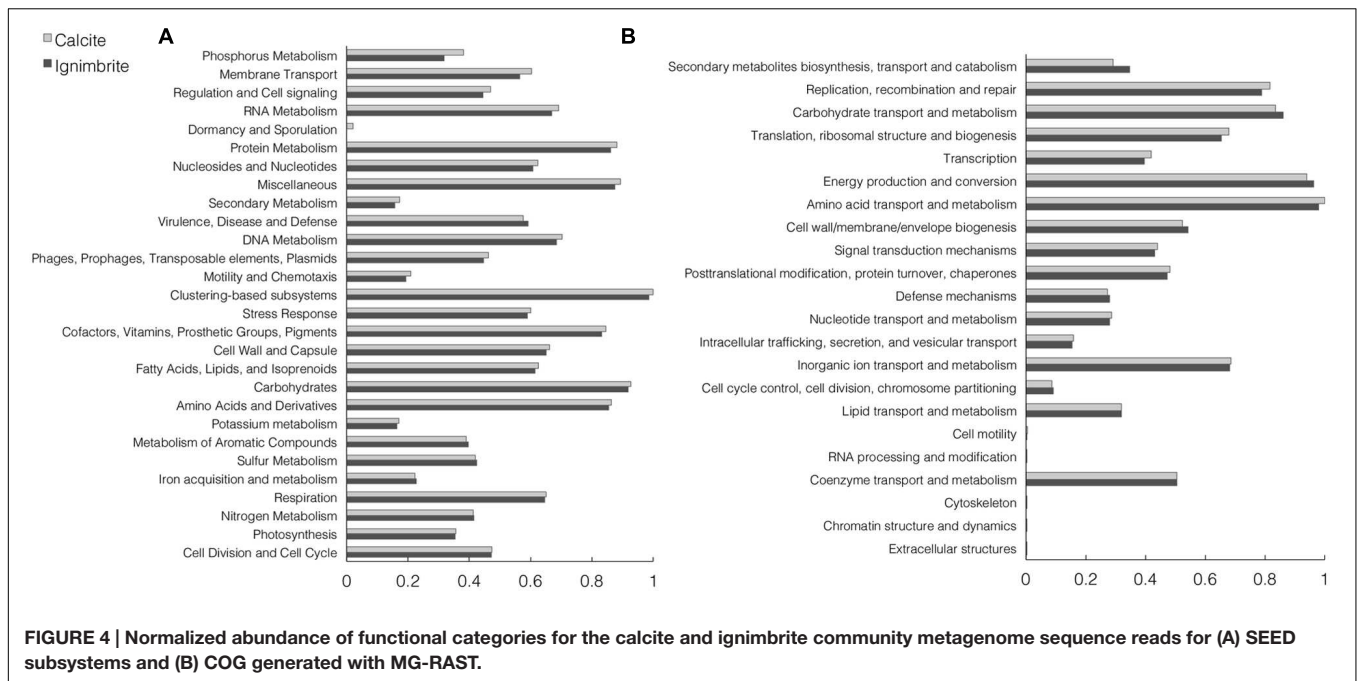
Functional Diversity of the Rock Communities

The functional composition of the calcite and ignimbrite metagenomes was analyzed with HUMAnN and MG-RAST using total sequence reads. The HUMAnN analysis showed that at the pathway level, the two communities maintained a strong degree of similarity in functional composition (Supplementary Figure S4). Functional assignment of sequence reads to the SEED or COG database by MG-RAST (**Figure 4**) also showed high similarity between the two communities. When using the SEED database to classify sequence reads into metabolic functions, categories with the most reads were for metabolisms related to "carbohydrates", "proteins", "amino acids and derivatives", and "cofactors, vitamins, prosthetic groups, and pigments" (**Figure 4**).

The metabolic potential for both the calcite and ignimbrite communities was further analyzed using the MG-RAST server. Genes encoding for pathways for photosynthesis and auxotrophic CO_2 fixation were found in both the calcite and ignimbrite metagenomes and included biosynthetic genes for photosystems I and II (PSI and PSII), phycobilisome (light harvesting antennae of PSII in cyanobacteria), the Calvin–Benson cycle, carboxyzomes, and photorespiration. Genes required for heterotrophy were found in both metagenomes, including genes for glycolysis and gluconeogenesis, central carbohydrate metabolism, trehalose biosynthesis, and lactate fermentation. Genes for nitrogen fixation were not detected, despite an intensive search using BLASTP and HMMER (Finn et al., 2011) on all predicted proteins from the sequence reads and from the assembled metagenomes. Using the SEED

FIGURE 3 | Taxonomic composition at the phylum level for the calcite and ignimbrite communities using either 16S rRNA gene sequences (16S) or PhyloSift marker genes from the metagenome data (Metag) (10% relative abundance cut-off).

FIGURE 4 | Normalized abundance of functional categories for the calcite and ignimbrite community metagenome sequence reads for (A) SEED subsystems and (B) COG generated with MG-RAST.
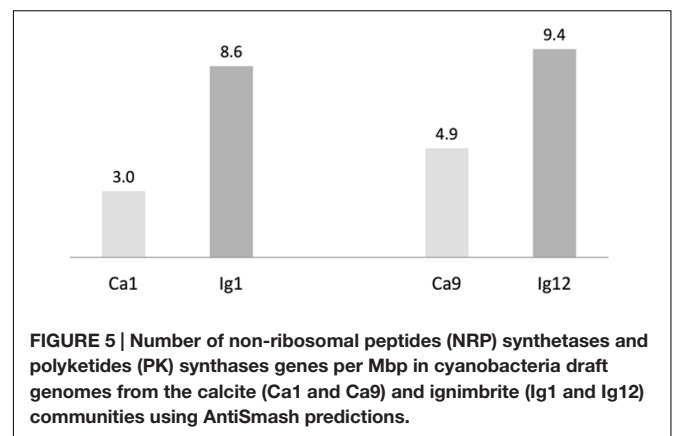
functional classification, we found that the major pathways for nitrogen acquisition in both communities were ammonia assimilation (direct uptake of ammonia and assimilation into glutamine and glutamate) and nitrate and nitrite ammonification (direct uptake of nitrate or nitrite and subsequent reduction to ammonia); minor pathways included allantoin utilization and cyanate hydrolysis (Supplementary Figure S5). Two major pathways for phosphorus acquisition were identified in both communities from the abundance of their protein encoding genes. One extensive suite of genes constituted a putative Pho regulon and another set of genes encoded for proteins involved in the transport and metabolism of phosphonates. With respect to iron transport, the calcite and ignimbrite metagenomes contained mostly siderophore/heme iron transporters. These ABC transporters are known to mediate the translocation of iron-siderophore complexes across the cytoplasmic membrane (Köster, 2001). We also found ABC transporters of the ferric iron type and, to a lesser extent, of the ferrous iron type and known to be induced by low pH (Köster, 2001). In cyanobacteria communities, photosynthesis results in an increase of pH as the result of CO_2 and HCO_3^- consumption and subsequent increase in CO_3^{-2} (Albertano et al., 2000), suggesting that slight variations in pH (above and below pH 7) might also be expected in the endolith colonization zone.

The functional analysis of the metagenomes from both communities also uncovered a diversity of stress response pathways, with most notably genes involved in carbon starvation, cold shock, oxidative stress, osmotic stress/desiccation tolerance, and secondary metabolites. In both the calcite and ignimbrite metagenomes, we found genes encoding for polyol transport (ABC transport systems), in particular glycerol, choline, betaine, and glycine betaine uptake systems (ABC transport systems), betaine and ectoine biosynthesis, and synthesis of glucans

(Supplementary Table S3). A gene encoding an aquaporin Z water channel, mediating rapid flux of water across the cellular membrane in response to abrupt changes in osmotic pressure (Calamita, 2000), was also found in high abundance. Also notable, the large number of genes dedicated to trehalose biosynthesis and uptake, which is an important compatible solute for osmoprotection in prokaryotes, fungi, and plants (Iturriaga et al., 2009). In particular, we found complete pathways for two ABC transport systems, a PTS transport system, a complete biosynthesis pathway, and several additional synthase genes. Secondary metabolites included a wide variety of pathways encoding for the synthesis of compounds such as non-ribosomal peptides (NRP), polyketides (PK), and alkaloids. The majority of genes encoding for NRP and PK represented a wide range of siderophores, including molecules with chemical similarity to pyoverdine, pyochelin, and yersiniabactin (Supplementary Table S2). Biosynthetic pathways for other secondary metabolites

FIGURE 5 | Number of non-ribosomal peptides (NRP) synthetases and polyketides (PK) synthases genes per Mbp in cyanobacteria draft genomes from the calcite (Ca1 and Ca9) and ignimbrite (Ig1 and Ig12) communities using AntiSmash predictions.

included those for alkaloids, phenylpropanoids, phenazines, and phytoalexins.

The two SEED subsystems of greatest difference between the two communities were phosphorous metabolism and membrane transport, with both functional categories showing higher gene abundance in the calcite community (**Figure 4A**). Functional diversity was also higher in the calcite community for genes involved in iron acquisition and transport, Mn transport, and the synthesis of compatible solutes such as betaine, ectoine, and glucans (Supplementary Table S3). One notable exception was a ferrichrome iron receptor, which was about three times more abundant in the ignimbrite than in the calcite community (Supplementary Table S3); the encoding genes were all annotated to cyanobacteria taxa. Overall, the pathways involved in osmoregulation were more diverse in the calcite than in the ignimbrite community (Supplementary Table S3). In addition, the functional categories most enriched in the calcite community were not from the most abundant taxa; for example, genes involved in the synthesis of osmoregulated periplasmic glucans, and most genes for the ectoine biosynthesis pathways, were from *Proteobacteria*, *Actinomycetales*, and *Firmicutes* rather than from *Cyanobacteria* or *Rubrobacter*. In contrast, when looking at gene abundance rather than diversity, we found considerably more NRPS and PKS related genes, especially siderophores biosynthetic patways, in the ignimbrite metagenome (Supplementary Table S2). The COG category for secondary metabolite biosynthesis/transport also shows similar higher levels of these genes in the ignimbrite metagenome (**Figure 4B**).

Sequence Assembly

The 12.8 Gbp of metagenome sequence data for the calcite community was assembled into a total of 350 Mbp of contigs (344,986) with an n50 value of 1,530 bp and a maximum contig length of 101,619 bp. The 10.1 Gbp metagenome sequence data for the ignimbrite community were assembled into 220 Mbp of contigs (166,509) with an n50 value of 2,631 bp and a maximum length of 119,637 bp (**Table 2**). Binning of contigs into population genomes using MaxBin generated 42 and 34 putative population genome bins, for the calcite and ignimbrite communities, respectively. All bins were individually assigned taxonomy using both Kraken and PhyloSift, and bins mostly composed of single taxa were further analyzed (**Table 3**). We found that the bins' G+C content closely matched that of the reference species, with G+C below 50% for *Cyanobacteria* and above 65% for *Rubrobacter* species. Interestingly, the relative abundance ratio of a bin did not necessarily match its estimated completeness. Rather, population genomes with the highest coverage produced lower quality assemblies, i.e., more contigs of smaller sizes (compare for example Ca01 with Ca09, and Ig01 with Ig12). This likely resulted from a higher representation of micro-heterogeneity at high coverage, for a given population genome, impairing the assembler functionality.

The most abundant population genomes found in the calcite and ignimbrite metagenomes were related to *Chroococcidiopsis* and *Gloeocapsa*, both members of the *Chroococcales*.

TABLE 3 | Most abundant species in the calcite and ignimbrite communities obtained by binning the metagenome sequence data.

Bin number	Taxa	%[a]	Estimated completenes (%)	Number of contigs	Size Mbp	G+C %	Bin number	Taxa	%[a]	Estimated completeness %	Number of contigs	Size Mbp	G+C %
	Calcite							**Ignimbrite**					
Ca0I	Gloeocapsa sp.	20.7	95.3	2329	7.7	48	Ig0I	Chroococcidiopsis sp.	32.8	73.8	2031	5.1	45
Ca02	Gloeocapsa sp.	9.2	46.7	2388	5.6	46	Ig02	Rubrobacter xylanophilus	15.1	38.3	1215	2.2	68
Ca03	Rubrobacter xylanophilus	4.2	43	413	1.2	65	Ig04	Rubrobacter xylanophilus	5.9	31.8	1900	6	67
Ca04	Rubrobacter xylanophilus	3.9	35.5	330	1.5	69	Ig0S	Frankineae sp.	5.1	74.8	460	4.5	75
Ca05	Acaryochloris sp.	3.6	68.2	1027	5.5	48	Ig06	Conexibacter woesei	2.7	87.9	357	3.9	75
Ca06	Conexibacter woesei	3.4	45.8	307	1.9	74	Ig07	Frankineae sp.	2.3	93.5	457	3.9	72
Ca07	Thermomicrobia sp.	3.2	83.2	1161	3.9	72	Ig0S	Thermobaculum sp.	2.0	86	661	3.5	69
Ca09	Gloeocapsa sp.	2.8	94.4	800	6.5	45	Ig0	Thermomicrobia sp.	2.0	89.7	548	5.1	63
Cal7	Conexibacter woesei	1.9	86	483	3.3	72	Igll	Conexibacter woesei	1.9	59.8	575	2.9	76
Cal9	Microlunatus sp.	1.6	91.6	538	4.3	71	Igl2	Chroococcidiopsis sp.	1.7	93.5	817	6.6	46

[a]*Relative abundance ratios; defined as actual coverage levels divided by summed coverage levels of all genomes.*

Cyanobacteria of the *Acaryochloris* genus were found only in the calcite community. The *Actinobacteria* were dominated by *Rubrobacter xylanophilus*, together with *Conexibacter woesei* and a number of other *Actinomycetales* species, the later differing between the two communities. *Chloroflexi* populations composed of *Thermomicrobia*, *Chloroflexus*, and *Thermobaculum* species were also found in the ignimbrite and calcite communities. The dominant *Proteobacteria* was *Sphingomonas*, and in the ignimbrite community, the low G+C% *Bacteroidetes* species *Segetibacter* was represented (**Table 3**). The phylogenetic analysis of the metagenomic bins indicated that there was significantly greater diversity within the calcite *Cyanobacteria* community, with bins distributed in two distinct clades, whereas only one *Cyanobacteria* clade was found within the ignimbrite dataset (Supplementary Figure S6). The ignimbrite *Cyanobacteria* population genomes were distinct at the phylogenetic level to that of the calcite, but one of the calcite metagenomic bins (Ca02) was consistently more closely related to the ignimbrite than the other calcite bins, further supporting the extent of the taxonomic diversity of the calcite community (Supplementary Figure S6).

A candidate draft genome for the dominant *Cyanobacteria* population genome in each community was selected based on estimated completeness, calculated from single gene copy analysis (MaxBin). The estimated completeness was 94.4% for *Gloeocapsa* sp. Ca09, from the calcite metagenome, and 93.5% for *Chroococcidiopsis* sp. Ig12, from the ignimbrite metagenome (**Table 3**). The genomic bins were annotated using the RAST server, and the two genomes were found to share similar hits for 64% of proteomes. Three contigs from Ca09 and five contigs from Ig12 were found to be entirely composed of phage-related and hypothetical proteins, making them likely to be cyanophage genomic fragments binned together with the host genome. This co-binning was probably due to the tendency of bacteriophages to share tetranucleotide frequencies, genes, and codon usage patterns with their host (Soueidan et al., 2014). These contigs were removed from the draft genomes for the analysis of annotation features and gene products. On average the two annotated proteomes had 69.4% amino acid identity, and only 27 and 24% of the gene products could be assigned a SEED category for Ca09 and Ig12, respectively. The draft genomes had multiple genes involved in choline and betaine biosynthesis subsystem and Aquaporin Z, both related to osmotic stress. A mycosporine synthesis gene cluster, *mysA*, *mysB*, and *mysC*, was found in the Ca09 but not in Ig12 draft genome; a BLAST search against the entire metagenomes found these genes to be absent from the ignimbrite metagenome. The Ca09 draft genome had six copies of the ferrichrome-iron receptor gene and the Ig12 draft genome had 12, while the closest reference genome in RAST, *Nostoc punctiforme* PCC 73102, only had one copy of this gene. Because these duplications have the same coverage, they belong to the same genomic bins, there are most likely true genome duplication rather than the result of genomic heterogeneity. Another iron related gene, the iron(III) dicitrate transport system *fecB* gene, had multiple duplications in both Ca09 and Ig12, but is absent in the *Nostoc punctiforme* RAST annotation.

We used antiSMASH to identify biosynthesis loci for secondary metabolism in the four major cyanobacteria population genomes from the calcite and ignimbrite metagenomes (**Table 3**). Similarly to the whole metagenome, we found an increased abundance of NRPS and PKS secondary metabolite biosynthetic genes and gene clusters in the Ig12 ignimbrite genome (**Figure 5**). Searches in the NORINE database[4] revealed that the closest chemical structures for the putative products of these gene clusters were often siderophores (**Figure 6**). A figure linking genes products with greater than 80% amino acid identity between the complement of contigs for the Ca9 and Ig12 draft genomes illustrates the close relatedness between these strains and also the greater abundance of NRPS/PKS gene clusters and ferrichrome-iron receptor genes in the draft genome from the ignimbrite community (**Figure 7**).
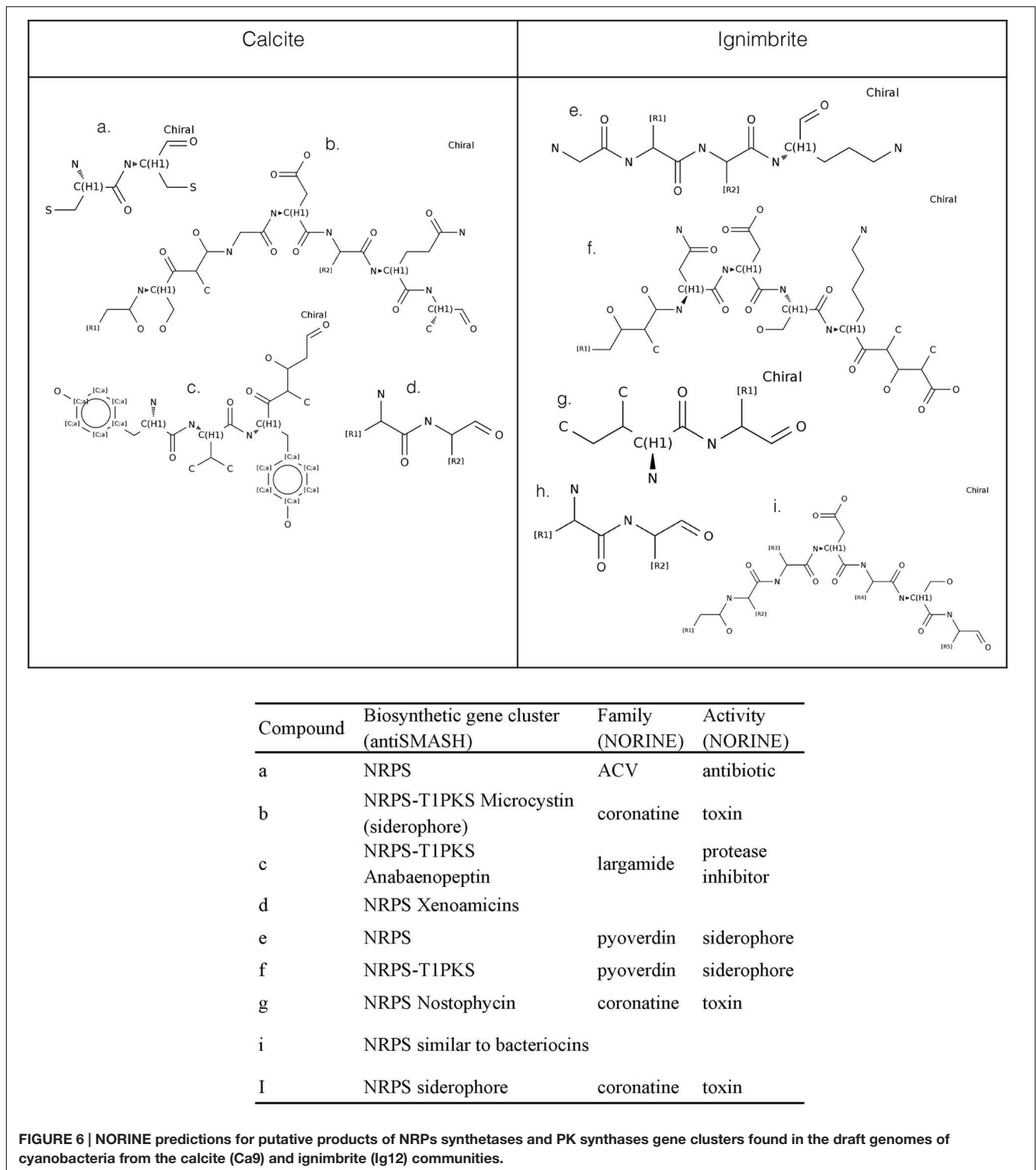
DISCUSSION

By linking taxonomy to function, our study of endolithic systems in the Atacama Desert provides an understanding of microbial adaptation strategies to extreme desiccation and solar irradiance. It also provides an integrated view of the biotic and abiotic factors shaping the functioning of these highly specialized ecosystems. Functional metagenomic reconstruction was carried out on two rock substrates, calcite and ignimbrite, and the context for the interpretation of the molecular data was provided by the examination of water availability, the mineralogical composition, and the type of colonization (chasmo vs. cryptoendolithic) for the two types of rock.

The Valle de la Luna (calcite rocks) and Lomas de Tilocalar (ignimbrite rocks) sampling sites are located in the hyper-arid zone of the Atacama Desert and, as such, both locations experience rare precipitations (DiRuggiero et al., 2013; Wierzchos et al., 2013). It is also the site of the highest solar irradiance recorded on Earth (Wierzchos et al., 2015). Moisture on the rock surface at Lomas de Tilocalar was only detected during and shortly after a rain event and the air RH is one of the lowest reported for the Atacama Desert (DiRuggiero et al., 2013; Wierzchos et al., 2013, 2015). In contrast, the yearly average of air RH at Valle de la Luna is about twice that in Lomas de Tilocalar and the presence of a hardened layer and microrills on the calcite rock surface was indicative of dew formation (DiRuggiero et al., 2013). This is important because dew formation, together with higher RH, can significantly increase the atmospheric water budget for the calcite microbial community (Kidron, 2000; Büdel et al., 2008; DiRuggiero et al., 2013). Higher maximum temperatures in Lomas de Tilocalar could have the opposite effect by potentially increasing the rate of evapotranspiration, limiting moisture availability within the ignimbrite (Houston and Hartley, 2003).
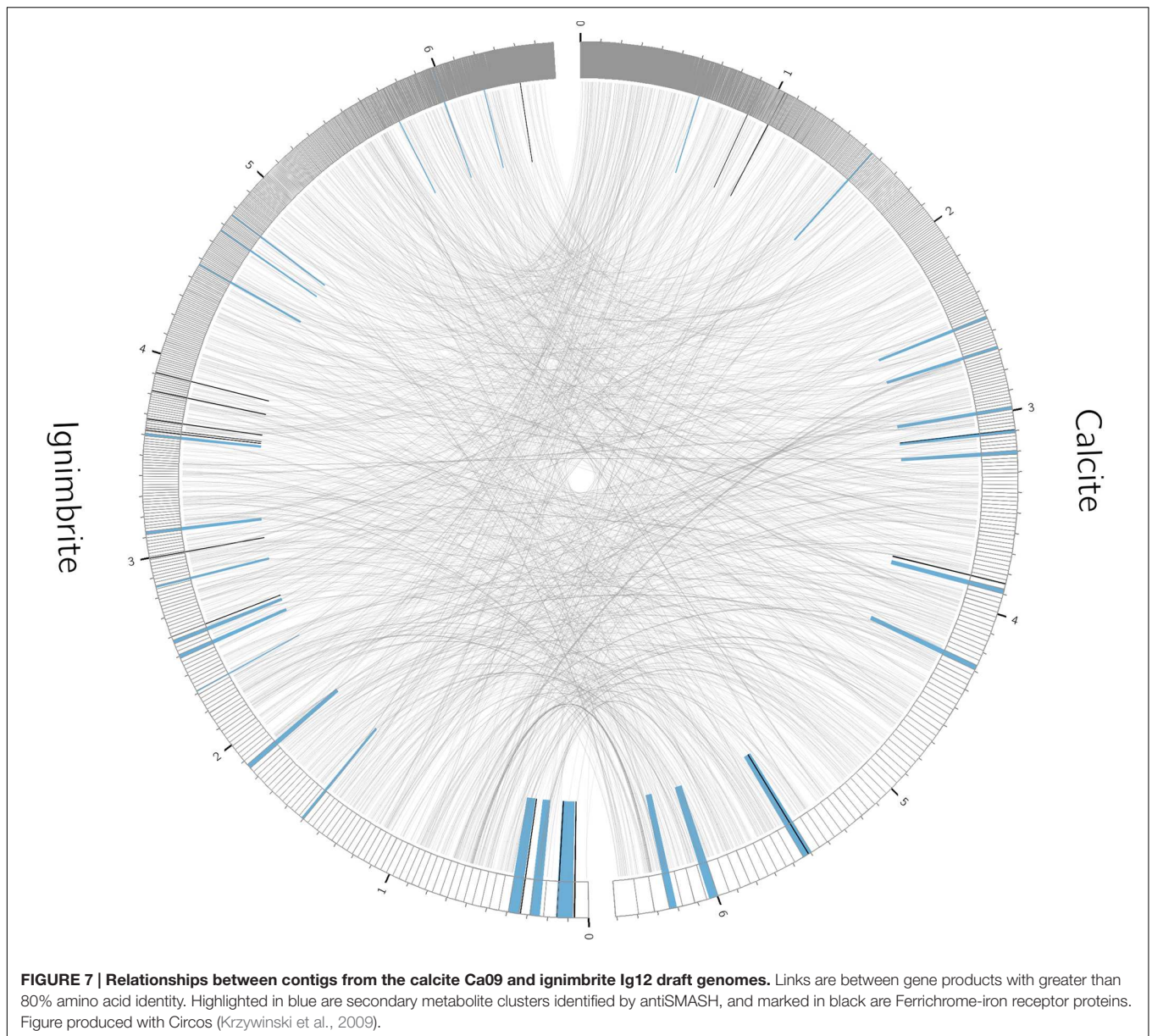
In addition to atmospheric water availability, the structure and geochemical composition of a rock substrate are also essential determinants of its bioreceptivity (Walker and Pace, 2007; Nienow, 2009). While the mineralogy between the two

---

[4]http://bioinfo.lifl.fr/norine/

Compound	Biosynthetic gene cluster (antiSMASH)	Family (NORINE)	Activity (NORINE)
a	NRPS	ACV	antibiotic
b	NRPS-T1PKS Microcystin (siderophore)	coronatine	toxin
c	NRPS-T1PKS Anabaenopeptin	largamide	protease inhibitor
d	NRPS Xenoamicins		
e	NRPS	pyoverdin	siderophore
f	NRPS-T1PKS	pyoverdin	siderophore
g	NRPS Nostophycin	coronatine	toxin
i	NRPS similar to bacteriocins		
I	NRPS siderophore	coronatine	toxin

FIGURE 6 | NORINE predictions for putative products of NRPs synthetases and PK synthases gene clusters found in the draft genomes of cyanobacteria from the calcite (Ca9) and ignimbrite (Ig12) communities.

substrates was significantly different, a laminated calcite layered rock and a pyroclastic rock composed of plagioclase and biotite embedded in a matrix of weakly welded glass shards, we argue here that the major differences lay within the physical properties of the rocks (DiRuggiero et al., 2013; Wierzchos et al., 2013;

Cámara et al., 2014). The calcite rocks displayed large colonized cracks and fissures, dozens of millimeters deep, all connected to the surface and providing direct access to surface water and increased light illumination. In contrast, the bottle-shaped pores found in the ignimbrite represented a limited colonization space.

FIGURE 7 | Relationships between contigs from the calcite Ca09 and ignimbrite Ig12 draft genomes. Links are between gene products with greater than 80% amino acid identity. Highlighted in blue are secondary metabolite clusters identified by antiSMASH, and marked in black are Ferrichrome-iron receptor proteins. Figure produced with Circos (Krzywinski et al., 2009).

Not all the pores were colonized, as observed by SEM-BSE, probably because they were not all connected to the surface. In addition, the dark varnish cover of the ignimbrite significantly decreases the incoming PAR radiation, which explains the very narrow colonization zone right underneath the rock's surface we observed.

Taken together, these observations point to a calcite environment far less extreme than that of the ignimbrite. This is substantiated by biodiversity estimates from 16S rRNA gene sequences, and from the metagenome data, showing a higher taxonomic diversity for the calcite than for the ignimbrite community. While both rock communities were dominated by members of the *Cyanobacteria*, we found a significantly higher proportion of chlorophototrophs (>20% difference) in the ignimbrite than the calcite community. In a previous

study of halite nodules, we reported a higher fraction of the cyanobacterium *Halothece* in samples collected in the driest and more extreme area of the desert than in Salars exposed to periodic fog events (Robinson et al., 2015). It was suggested that a higher proportion of photoautotrophs was required to support the associated heterotrophic populations because of lower primary productivity under water stress (Robinson et al., 2015). This might also be the case for the differences in chlorophototrophs abundance we observed between the calcite and ignimbrite rocks.

It is important to point out that the metagenome for each community was obtained with DNA extracted from four individual rocks each. The 16S rRNA gene data was obtained for each of the individual rocks and supported the taxonomic assignments of the metagenomes. Using the metagenomic data,

we identified the dominant cyanobacteria in the calcite substrate as *Gleocapsa* whereas previous work assigned them to the *Chroococcidiopsis* (DiRuggiero et al., 2013). This discrepancy arose because the *Gleocapsa* and *Chroococcidiopsis* taxa are closely related, we use metagenome data, and different taxonomic assignment algorithms produces different data. It is also possible that these belong to a new genus that is neither *Gleocapsa* nor *Chroococcidiopsis*. We did not find any archaeal sequences in our endolith metagenomes, which is consistent with low abundance or absence of archaea in previous studies of lithobiontic habitats (Wood et al., 2008; Pointing et al., 2009; Wong et al., 2010; Stomeo et al., 2012; DiRuggiero et al., 2013; Vikram et al., 2015)

In term of biogeochemical cycling genes, the identification of photosynthesis and carbon fixation pathways from the metagenomes confirmed that primary production was mainly carried out via photosynthesis. However, in contrast to Antarctic cryptoendolithic communities, we did not detect any genes for chemolithoautotrophy (Wei et al., 2015). Genes for diazotrophy were also absent from both metagenomes and the major sources of nitrogen for the communities were potentially ammonia and nitrates. These later nitrogen sources might not be limiting in the Atacama Desert where high nitrate concentration from atmospheric origin (Michalski et al., 2004) have been reported. This might explain the absence of nitrogen fixation in endoliths from this desert in contrast to endolithic niches in the McMurdo Dry Valleys (Chan et al., 2013; Wei et al., 2015).

The metabolic genes and pathways annotated from the calcite and ignimbrite metagenomes revealed that the rock substrate is a stressful environment with its own challenges. Nutrient limitation was suggested by a number of stress adaptation strategies, including an abundance of protein encoding genes for two major pathways for phosphate acquisition. In cyanobacteria, and a number of other bacteria, the Pho regulon, involved in transport and metabolism of phosphonates, is controlled by environmental inorganic phosphate levels and mediates an adaptive response to phosphate starvation (Monds et al., 2006; Adams et al., 2008). Additionally, phosphonates can also be used as an alternative source of phosphorous under limiting conditions (Dyhrman et al., 2006; Adams et al., 2008). The large number of genes involved in carbon starvation found in the metagenomes is also indicative of periods of low-nutrient stress in the rock environment. These genes play a role in long-term survival under extended periods of starvation (Kolter et al., 1993), provide resistance to additional stresses such as oxidizing agents and near-UV radiations (Watson et al., 1998), and might give a competitive advantage to the cells once growth resumes (Zambrano et al., 1993).

Adaptation to desiccation in the rock environment involved a number of strategies for desiccation avoidance and protection via the biosynthesis and accumulation of organic osmolytes. Production of reactive oxygen species in response to desiccating conditions substanciates the abundance of genes and pathways for oxidative stress that we found in the metagenomes (Fredrickson et al., 2008; Webb and DiRuggiero, 2012). The metagenome analyses revealed metabolic pathways for a broad spectrum of compatible solutes that have been found across

bacterial phyla (da Costa et al., 1998; Roberts, 2005). These solutes are thought to contribute to desiccation tolerance because in high concentrations they generate low water potentials in the cytoplasm without incurring metabolic damage (Yancey, 2005; Oren, 2007). Of note, is the high abundance of genes for polyol metabolism, which in algae, fungi, plants, and a few bacteria, is a key part of a biochemical protective strategy against water loss (da Costa et al., 1998; Holzinger and Karsten, 2013). Functional categories for osmoregulation were more diverse in the calcite community but were not from the most abundant taxa, suggesting that the calcite functional diversity was driven by the taxonomic diversity of the less abundant taxa in the community. We were not able to identified genes for EPS biosynthesis in the metagenomes but microscopic observations of the endolithic communities revealed abundant EPS in the calcite but not in the ignimbrite microbial assemblages. In the calcite environment with more liquid water available, and frequent wetting-drying cycles, EPS might play a significant role in water retention, as demonstrated in similarly dry environments (Hu et al., 2012). In contrast, in the ignimbrite, where the unique source of liquid water is very scarce rain events and complete dehydration periods lasting as long as 9 months, the environmental conditions might too harsh for EPS effectiveness and for the energy expenditure required for their biosynthesis.

Cyanobacteria are known to produce a wide range of secondary metabolites including PK and NRP (Leao et al., 2012; Puglisi et al., 2014). One of the major differences, at the functional level, between the calcite and ignimbrite metagenomes was in their potential to synthesize secondary metabolites. We found a considerably higher number of NRPS and PKS related genes in the ignimbrite metagenome, and the draft genomes of some of the most abundant cyanobacteria from the ignimbrite community, than in the calcite community. Biosynthetic pathways for other secondary metabolites included those for alkaloids, involved in host fitness and protection (Leflaive and Ten-Hage, 2007), phenylpropanoids, which have been shown to enhance antioxidant activity and tolerance to stress in cyanobacteria (Singh et al., 2014), phenazines known to contribute to behavior and ecological fitness (Pierson and Pierson, 2010), and phytoalexins with antimicrobial and often antioxidative activities (Patterson and Bolis, 1997). The synthesis of these secondary metabolites is often directed toward protection from radiation, allelopathy, resource competition, and signaling (Leao et al., 2012; Puglisi et al., 2014) and, in the case of the ignimbrite, might be an indication of the fierce competition for scarce resources and limited colonization space in the small pores of this substrate rock. Indeed the production of small antimicrobial compounds has been reported as a strategy for eliminating prior residents before colonization of a new space by a number of microorganisms (Hibbing et al., 2010). Additionally, SEM-BSE images of colonized pores from the ignimbrite rock revealed cell decaying cells in some of the pores, potentially reflecting the effects of antimicrobial compounds or resource competition. In contrast the chasmoendolithic environment of the calcite provided more physical space and therefore may promote interactions between members of the community for

metabolic cooperation. In addition to antimicrobials, database searches revealed that the closest chemical structures for some of the NRPS and PKS gene clusters found in both communities were for siderophores. Iron is an essential element for most organisms because of its essential role in redox enzymes, membrane-bound electron transport chains, but also in photosynthesis (Andrews et al., 2003). The abundance of siderophores in the rock communities, and the enrichments in genes for iron acquisition in the ignimbrite cyanobacteria draft genome when compared to the reference genome from *N. punctiforme*, strongly suggest iron starvation. This was further emphasized in the ignimbrite community with a large number of sequence reads assigned to gene clusters for siderophores and a ferrichrome-iron receptor. Furthermore, the mycosporine-like gene cluster found in the calcite cyanobacteria, but not the ignimbrite, may indicate differences in UV radiation stress experienced by the two communities (Garcia-Pichel et al., 1993; Oren, 1997).

Environmental stress is a major driver of microbial diversity and this is reflected in the functional annotation of metagenomes from Atacama Desert endoliths. Our analysis is reflective of adaptations to desiccation, osmotic stress, low nutrients, iron deficiency, and in the case of the ignimbrite, of a fierce competition for colonization space and resources among community members. The relative abundance of phototrophs in each community is likely indicative of the *in situ* level of primary production, a hypothesis that might be tested by careful field measurements of microbial activity.

REFERENCES

Abubucker, S., Segata, N., Goll, J., Schubert, A. M., Izard, J., Cantarel, B. L., et al. (2012). Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput. Biol.* 8:e1002358. doi: 10.1371/journal.pcbi.1002358

Adams, M. M., Gomez-Garcia, M. R., Grossman, A. R., and Bhaya, D. (2008). Phosphorus deprivation responses and phosphonate utilization in a thermophilic *Synechococcus* sp. from microbial mats. *J. Bacteriol.* 190, 8171–8184. doi: 10.1128/JB.01011-08

Albertano, P., Bruno, L., D'Ottavi, D., Moscone, D., and Palleschi, G. (2000). Effect of photosynthesis on pH variation in cyanobacterial biofilms from Roman catacombs. *J. Appl. Phycol.* 12, 379–384. doi: 10.1023/A:1008149529914

Andrews, S. C., Robinson, A. K., and Rodriguez-Quinones, F. (2003). Bacterial iron homeostasis. *FEMS Microbiol. Rev.* 27, 215–237. doi: 10.1016/S0168-6445(03)00055-X

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., and Edwards, R. A. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75

Azúa-Bustos, A., González-Silva, C., Mancilla, R. A., Salas, L., Gómez-Silva, B., McKay, C. P., et al. (2011). Hypolithic cyanobacteria supported mainly by fog in the coastal range of the Atacama Desert. *Microb. Ecol.* 61, 568–581. doi: 10.1007/s00248-010-9784-5

Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E., et al. (2013). antiSMASH 2.0–a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* 41, W204–W212. doi: 10.1093/nar/gkt449

Büdel, B., Bendix, J., Bicker, F. R., and Allan Green, T. G. (2008). Dewfall as a water source frequently activates the endolithic cyanobacterial communities in the granites of taylor valley, Antarctica. *J. Phycol.* 44, 1415–1424. doi: 10.1111/j.1529-8817.2008.00608.x

Calamita, G. (2000). The *Escherichia coli* aquaporin-Z water channel. *Mol. Microbiol.* 37, 254–262. doi: 10.1046/j.1365-2958.2000.02016.x

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.00301

Cámara, B., Suzuki, S., Nealson, K. H., Wierzchos, J., Ascaso, C., Artieda, O., et al. (2014). Ignimbrite textural properties as determinants of endolithic colonization patterns from hyper-arid Atacama Desert. *Int. Microbiol.* 17, 235–247. doi: 10.2436/20.1501.01.226

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Cary, S. C., McDonald, I. R., Barrett, J. E., and Cowan, D. A. (2010). On the rocks: the microbiology of Antarctic Dry Valley soils. *Nat. Rev. Microbiol.* 8, 129–138. doi: 10.1038/nrmicro2281

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334

Chan, Y., Lacap, D. C., Lau, M. C., Ha, K. Y., Warren-Rhodes, K. A., Cockell, C. S., et al. (2012). Hypolithic microbial communities: between a rock and a hard place. *Environ. Microbiol.* 14, 2272–2282. doi: 10.1111/j.1462-2920.2012.02821.x

Chan, Y., Van Nostrand, J. D., Zhou, J., Pointing, S. B., and Farrell, R. L. (2013). Functional ecology of an Antarctic Dry Valley. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8990–8995. doi: 10.1073/pnas.1300643110

Cockell, C. S., Pybus, D., Olsson-Francis, K., Kelly, L., Petley, D., Rosser, N., et al. (2011). Molecular characterization and geological microenvironment of a microbial community inhabiting weathered receding shale cliffs. *Microb. Ecol.* 61, 166–181. doi: 10.1007/s00248-010-9730-6

Cowan, D. A., Pointing, S. B., Stevens, M. I., Cary, S. C., Stomeo, F., and Tuffin, I. M. (2011). Distribution and abiotic influences on hypolithic microbial communities in an Antarctic Dry Valley. *Polar Biol.* 34, 307–311. doi: 10.1007/s00300-010-0872-2

da Costa, M. S., Santos, H., and Galinski, E. A. (1998). An overview of the role and diversity of compatible solutes in Bacteria and Archaea. *Adv. Biochem. Eng. Biotechnol.* 61, 117–153.

Darling, A. E., Jospin, G., Lowe, E., Matsen, F. A. T., Bik, H. M., and Eisen, J. A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ* 2:e243. doi: 10.7717/peerj.243

de la Torre, J. R., Goebel, B. M., Friedmann, E. I., and Pace, N. R. (2003). Microbial diversity of cryptoendolithic communities from the McMurdo Dry Valleys, Antarctica. *Appl. Environ. Microbiol.* 69, 3858–3867. doi: 10.1128/AEM.69.7.3858-3867.2003

de los Ríos, A., Cary, C. S., and Cowan, D. (2014). The spatial structures of hypolithic communities in the Dry Valleys of East Antarctica. *Polar Biol.* 37, 1823–1833. doi: 10.1007/s00300-014-1564-0

de los Ríos, A., Valea, S., Ascaso, C., Davila, A. F., Kastovsky, J., McKay, C. P., et al. (2010). Comparative analysis of the microbial communities inhabiting halite evaporites of the Atacama Desert. *Int. Microbiol.* 2, 79–89.

DiRuggiero, J., Wierzchos, J., Robinson, C. K., Souterre, T., Ravel, R., Artieda, O., et al. (2013). Microbial colonization of chasmoendolithic habitats in the hyper-arid zone of the Atacama Desert. *Biogeosciences* 10, 2439–2450. doi: 10.5194/bg-10-2439-2013

Dong, H., Rech, J. A., Jiang, H., Sun, H., and Buck, B. J. (2007). Endolithic cyanobacteria in soil gypsum: occurrences in Atacama (Chile), Mojave (USA), and Al-Jafr Basin (Jordan) Deserts. *J. Geophys. Res.* 112:G02030.

Dyhrman, S. T., Chappell, P. D., Haley, S. T., Moffett, J. W., Orchard, E. D., Waterbury, J. B., et al. (2006). Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* 439, 68–71. doi: 10.1038/nature04203

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367

Fredrickson, J. K., Li, S. M., Gaidamakova, E. K., Matrosova, V. Y., Zhai, M., Sulloway, H. M., et al. (2008). Protein oxidation: key to bacterial desiccation resistance? *ISME J.* 2, 393–403. doi: 10.1038/ismej.2007.116

Friedmann, E. I. (1982). Endolithic microorganisms in the Antarctic cold desert. *Science* 215, 1045–1053. doi: 10.1126/science.215.4536.1045

Garcia-Pichel, F., Wingard, C. E., and Castenholz, R. W. (1993). Evidence regarding the UV sunscreen role of a mycosporine-like compound in the cyanobacterium *Gloeocapsa* sp. *Appl. Environ. Microbiol.* 59, 170–176.

Gish, W., and States, D. J. (1993). Identification of protein coding regions by database similarity search. *Nat. Genet.* 3, 266–272. doi: 10.1038/ng0393-266

Herrera, A., Cockell, C. S., Self, S., Blaxter, M., Reitner, J., Thorsteinsson, T., et al. (2009). A cryptoendolithic community in volcanic glass. *Astrobiology* 9, 369–381. doi: 10.1089/ast.2008.0278

Hibbing, M. E., Fuqua, C., Parsek, M. R., and Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* 8, 15–25. doi: 10.1038/nrmicro2259

Holzinger, A., and Karsten, U. (2013). Desiccation stress and tolerance in green algae: consequences for ultrastructure, physiological and molecular mechanisms. *Front. Plant Sci.* 4:327. doi: 10.3389/fpls.2013.00327

Houston, J., and Hartley, A. (2003). The central Andean west-slope rainshadow and its potential contribution to the origin of hyperaridity in the Atacama Desert. *Int. J. Climatol.* 23, 1453–1464. doi: 10.1002/joc.938

Hu, C., Gao, K., and Whitton, B. A. (2012). "Semi-arid regions and deserts," in *Ecology of Cyanobacteria II: Their Diversity in Space and Time*, ed. B. A. Whitton (Amsterdam: Springer), 345–370.

Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119

Iturriaga, G., Suárez, R., and Nova-Franco, B. (2009). Trehalose metabolism: from osmoprotection to signaling. *Int. J. Mol. Sci.* 10, 3793–3810. doi: 10.3390/ijms10093793

Kidron, G. J. (2000). Analysis of dew precipitation in three habitats within a small arid drainage basin, Negev Highlands, Israel. *Atmosph. Res.* 55, 257–270. doi: 10.1016/S0169-8095(00)00063-6

Kolter, R., Siegele, D. A., and Tormo, A. (1993). The stationary phase of the bacterial life cycle. *Annu. Rev. Microbiol.* 47, 855–874. doi: 10.1146/annurev.mi.47.100193.004231

Köster, W. (2001). ABC transporter-mediated uptake of iron, siderophores, heme and vitamin B12. *Res. Microbiol.* 152, 291–301. doi: 10.1016/S0923-2508(01)01200-1

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25

Leao, P. N., Engene, N., Antunes, A., Gerwick, W. H., and Vasconcelos, V. (2012). The chemical ecology of cyanobacteria. *Nat. Prod. Rep.* 29, 372–391. doi: 10.1039/c2np00075j

Leflaive, J. P., and Ten-Hage, L. (2007). Algal and cyanobacterial secondary metabolites in freshwaters: a comparison of allelopathic compounds and toxins. *Freshw. Biol.* 52, 199–214. doi: 10.1111/j.1365-2427.2006.01689.x

Makhalanyane, T. P., Valverde, A., Gunnigle, E., Frossard, A., Ramond, J. B., and Cowan, D. A. (2015). Microbial ecology of hot desert edaphic systems. *FEMS Microbiol. Rev.* 39, 203–221. doi: 10.1093/femsre/fuu011

McKay, C. P., Friedmann, E. I., Gomez-Silva, B., Caceres-Villanueva, L., Andersen, D. T., and Landheim, R. (2003). 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. doi: 10.1089/153110703769016460

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., et al. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386. doi: 10.1186/1471-2105-9-386

Michalski, G., Böhlke, J. K., and Thiemens, M. (2004). Long term atmospheric deposition as the source of nitrate and other salts in the Atacama Desert, Chile: new evidence from mass independent oxygen isotopic compositions. *Geochim. Cosmochim. Acta* 68, 4023–4038. doi: 10.1016/j.gca.2004.04.009

Miller, C. S., Baker, B. J., Thomas, B. C., Singer, S. W., and Banfield, J. F. (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol.* 12:R44. doi: 10.1186/gb-2011-12-5-r44

Monds, R. D., Newell, P. D., Schwartzman, J. A., and O'Toole, G. A. (2006). Conservation of the Pho regulon in *Pseudomonas fluorescens* Pf0-1. *Appl. Environ. Microbiol.* 72, 1910–1924. doi: 10.1128/AEM.72.3.1910-1924.2006

Nienow, J. A. (2009). "Extremophiles: dry environments (including cryptoendoliths)," in *Encyclopedia of Microbiology*, 3rd Edn, ed. S. M. Schaechter (Oxford: Elsevier), 159–173.

Oren, A. (1997). Mycosporine-like amino acids as osmotic solutes in a community of halophilic cyanobacteria. *Geomicrobiol. J.* 14, 231–240. doi: 10.1080/01490459709378046

Oren, A. (2007). "Diversity of organic osmotic compounds and osmotic adaptation in cyanobacteria and algae," in *Algae and Cyanobacteria in Extreme Environments*, ed. J. Seckback (Berlin: Springer).

Patterson, G. M. L., and Bolis, C. M. (1997). Fungal cell-wall polysaccharides elicit an antifungal secondary metabolite (Phytoalexin) in the cyanobacterium Scytonema ocellatum. *J Phycol* 33, 54–60. doi: 10.1111/j.0022-3646.1997.00054.x

Peng, Y., Leung, H. C., Yiu, S. M., and Chin, F. Y. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi: 10.1093/bioinformatics/bts174

Pierson, L. S. III, and Pierson, E. A. (2010). Metabolism and function of phenazines in bacteria: impacts on the behavior of bacteria in the environment and biotechnological processes. *Appl. Microbiol. Biotechnol.* 86, 1659–1670. doi: 10.1007/s00253-010-2509-3

Pointing, S. B., and Belnap, J. (2012). Microbial colonization and controls in dryland systems. *Nat. Rev. Microbiol.* 10, 551–562. doi: 10.1038/nrmicro2831

Pointing, S. B., Chan, Y., Lacap, D. C., Lau, M. C., Jurgens, J. A., and Farrell, R. L. (2009). Highly specialized microbial diversity in hyper-arid polar desert. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19964–19969. doi: 10.1073/pnas.0908274106

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. doi: 10.1371/journal.pone.0009490

Pruesse, E., Quast, C., Knittel, K., Fuchs, B., Ludwig, W., Peplies, J., et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned

ribosomal RNA sequence data compatible with ARB. *Nuc Acids Res.* 35, 7188–7196. doi: 10.1093/nar/gkm864

Puglisi, M. P., Sneed, J. M., Sharp, K. H., Ritson-Williams, R., and Paul, V. J. (2014). Marine chemical ecology in benthic environments. *Nat. Prod. Rep.* 31, 1510–1553. doi: 10.1039/c4np00017j

Roberts, M. F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Syst.* 1:5. doi: 10.1186/1746-1448-1-5

Robinson, C. K., Wierzchos, J., Black, C., Crits-Christoph, A., Ma, B., Ravel, J., et al. (2015). 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. doi: 10.1111/1462-2920.12364

Singh, D. P., Prabha, R., Meena, K. K., Sharma, L., and Sharma, A. K. (2014). Induced accumulation of polyphenolics and favonoids in Cyanobacteria under salts tress protects organisms through enhanced antioxidant activity. *Am. J. Plant Sci.* 5, 726–735. doi: 10.4236/ajps.2014.55087

Soueidan, H., Schmitt, L. A., Candresse, T., and Nikolski, M. (2014). Finding and identifying the viral needle in the metagenomic haystack: trends and challenges. *Front. Microbiol.* 5:739. doi: 10.3389/fmicb.2014.00739

Stomeo, F., Makhalanyane, T. P., Valverde, A., Pointing, S. B., Stevens, M. I., Cary, C. S., et al. (2012). Abiotic factors influence microbial diversity in permanently cold soil horizons of a maritime-associated Antarctic Dry Valley. *FEMS Microbiol. Ecol.* 82, 326–340. doi: 10.1111/j.1574-6941.2012.01360.x

Vikram, S., Guerrero, L. D., Makhalanyane, T. P., Le, P. T., Seely, M., and Cowan, D. A. (2015). Metagenomic analysis provides insights into functional capacity in a hyperarid desert soil niche community. *Environ. Microbiol.* doi: 10.1111/1462-2920.13088 [Epub ahead of print].

Walker, J. J., and Pace, N. R. (2007). Endolithic microbial ecosystems. *Annu. Rev. Microbiol.* 61, 331–347. doi: 10.1146/annurev.micro.61.080706.093302

Watson, S. P., Clements, M. O., and Foster, S. J. (1998). Characterization of the starvation-survival response of *Staphylococcus aureus. J. Bacteriol.* 180, 1750–1758.

Webb, K. M., and DiRuggiero, J. (2012). "Radiation resistance in extremophiles: fending off multiple attacks," in *Polyextremophiles: Life Under Multiple forms of Stress*, eds J. Seckback, A. Oren, and H. Stan-Lotter (Amsterdam: Springer), 251–267.

Wei, S. T. S., Fernandez-Martinez, M., Chan, Y., Van Nostrand, J. D., de los Rios-Murillo, A., Chiu, J. M. Y., et al. (2015). Diverse metabolic and stress-tolerance pathways in chasmoendolithic and soil communities of Miers Valley, McMurdo Dry Valleys, Antarctica. *Polar Biol.* 38, 433–443. doi: 10.1007/s00300-014-1598-3

Wierzchos, J., and Ascaso, C. (1994). Application of backscattered electron imaging to the study of the lichen–rock interface. *J. Microsc.* 175, 54–59. doi: 10.1111/j.1365-2818.1994.tb04787.x

Wierzchos, J., Ascaso, C., and McKay, C. P. (2006). Endolithic cyanobacteria in halite rocks from the hyperarid core of the Atacama Desert. *Astrobiology* 6, 415–422. doi: 10.1089/ast.2006.6.415

Wierzchos, J., Camara, B., de Los Rios, A., Davila, A. F., Sanchez Almazo, I. M., Artieda, O., et al. (2011). Microbial colonization of Ca-sulfate crusts in the hyperarid core of the Atacama Desert: implications for the search for life on Mars. *Geobiology* 9, 44–60. doi: 10.1111/j.1472-4669.2010.00254.xGBI254

Wierzchos, J., Davila, A. F., Artieda, O., Cámara-Gallego, B., de los Ríos, A., Nealson, K. H., et al. (2013). Ignimbrite as a substrate for endolithic life in the hyper-arid Atacama Desert: implications for the search for life on Mars. *Icarus* 224, 334–346. doi: 10.1111/j.1472-4669.2010.00254.x

Wierzchos, J., de los Ríos, A., and Ascaso, C. (2012). Microorganisms in desert rocks: the edge of life on Earth. *Int. Microbiol.* 15, 173–183.

Wierzchos, J., DiRuggiero, J., Vítek, P., Artieda, O., Souza-Egipsy, V., Skaloud, P., et al. (2015). Adaptation strategies of endolithic chlorophototrophs to survive the hyperarid and extreme solar radiation environment of the Atacama Desert. *Front. Microbiol.* 6:934. doi: 10.3389/fmicb.2015.00934

Wong, F. K., Lau, M. C., Lacap, D. C., Aitchison, J. C., Cowan, D. A., and Pointing, S. B. (2010). Endolithic microbial colonization of limestone in a high-altitude arid environment. *Microb. Ecol.* 59, 689–699. doi: 10.1007/s00248-009-9607-8

Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46

Wood, S. A., Rueckert, A., Cowan, D. A., and Cary, S. C. (2008). Sources of edaphic cyanobacterial diversity in the Dry Valleys of Eastern Antarctica. *ISME J.* 2, 308–320. doi: 10.1038/ismej.2007.104

Wu, Y. W., Tang, Y. H., Tringe, S. G., Simmons, B. A., and Singer, S. W. (2014). MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* 2:26. doi: 10.1186/2049-2618-2-26

Yancey, P. H. (2005). Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *Exp. Biol.* 208, 2819–2830. doi: 10.1242/jeb.01730

Zambrano, M. M., Siegele, D. A., Almiron, M., Tormo, A., and Kolter, R. (1993). Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. *Science* 259, 1757–1760. doi: 10.1126/science.7681219