Metabolic specializations within a bacterial community to create living rocks

Running Title: Functional guilds in living rocks

Samantha C. Waterworth[a], Eric W. Isemonger[b], Evan R. Rees[a], Rosemary A. Dorrington[b] and Jason C. Kwan[a]#

[a] Division of Pharmaceutical Sciences, University of Wisconsin, 777 Highland Ave., Madison, Wisconsin 53705, USA

[b] Department of Biochemistry and Microbiology, Rhodes University, Grahamstown, South Africa

#Corresponding author. Email: jason.kwan@wisc.edu; Phone: 608-262-3829

Competing interests: The authors declare no competing financial interests

ABSTRACT

Stromatolites are complex microbial mats that form lithified layers and ancient forms are the oldest evidence of life on earth, dating back over 3.4 billion years. Their emergence aligns with the oxygenation of the Earth's atmosphere and insight into these ancient structures would shed light on the earliest days of Earth. Modern stromatolites are relatively rare but may provide clues about the function and evolution of their ancient counterparts. Previous studies have assessed microbial diversity and overall functional potential but not at a genome-resolved level. In this study, we focus on peritidal stromatolites occurring at Cape Recife and Schoenmakerskop on the southeastern South African coastline. We identify functional gene sets in bacterial species conserved across two geographically distinct stromatolite formations and show that these bacteria may promote carbonate precipitation through the reduction of sulfur and nitrogenous compounds and produce calcium ions that are predicted to play an important role in promoting lithified mats. We propose that abundance of extracellular alkaline phosphatases, in combination with the absence of transport regulatory enzymes, may lead to the precipitation of phosphatic deposits within these stromatolites. We conclude that the cumulative effect of several conserved bacterial species drives accretion in these two stromatolite formations.

INTRODUCTION

Stromatolites are organo-sedimentary structures that date back more than 3.4 billion years, forming the oldest fossils of living organisms on Earth [1]. The emergence of Cyanobacteria in stromatolites approximately 2.3 billion years ago initiated the Great

Oxygenation Event that fundamentally altered the Earth's redox potential and resulted in an explosion of oxygen-based and multicellular biological diversity [2]. Study of ancient stromatolites could provide insight into how microorganisms shaped early eukaryotic evolution but unfortunately these ancient microbial mats are not sufficiently preserved for identification of these microbes and individual bacteria cannot be classified further than Cyanobacteria due to morphological conservatism [1, 3]. The study of extant stromatolite analogues will therefore help to elucidate the biological mechanisms that led to the formation and evolution of their ancient ancestors. Modern stromatolites are formed through a complex combination of both biotic and abiotic processes. The core process revolves around the carbon cycle where photosynthesizing bacteria transform inorganic carbon into bioavailable organic carbon for respiration. Bacterial respiration in turn results in the release of inorganic carbon, which, under alkaline conditions, will bind cations and precipitate primarily as calcium carbonate [1]. This carbonate precipitate, along with sediment grains, can then become trapped within cyanobacterial biofilms forming the characteristic lithified layers.

Alteration of the pH and subsequently, the solubility index (SI), through microbial cycling of redox sensitive compounds such as phosphate, nitrogen, sulfur and other nutrients within the biofilm may promote mineralization or dissolution of carbonate minerals. This in turn regulates the rate of carbonate accretion and stromatolite growth. Particularly, photosynthesis and sulfate reduction have been demonstrated to increase alkalinity thereby promoting carbonate accretion, resulting in the gradual formation of lithified mineral layers [1, 4]. In some stromatolite formations such as those of Shark Bay in

Australia, there is abundant genetic potential for both dissimilatory oxidation of sulfur (which may promote dissolution under oxic conditions and precipitation under anoxic conditions) and dissimilatory reduction of sulfate (which promotes precipitation) [5–7].

The biogenicity of stromatolites has been studied extensively in stable environments, such as the hypersaline and marine stromatolites of Shark Bay, Australia and Exuma Cay, Bahamas, respectively [8, 9]. Overall, Cyanobacteria, Proteobacteria and Bacteroidetes appear to be abundant in both marine and hypersaline systems and the Cyanobacteria are proposed to be particularly vital to these formations through the combined effect of biofilm formation, carbon fixation, nitrogen fixation and tunneling (endolithic) activity [6, 8–10]. It is further hypothesized that Proteobacteria and Bacteroidetes influence the solubility index of the systems through sulfur cycling, anoxygenic phototrophy and fermentation of organic matter [11, 12].

Peritidal tufa stromatolite systems are found along the southeastern coastline of South Africa (SA). They are geographically isolated, occurring at coastal dune seeps separated by stretches of coastline [13]. In these SA systems, stromatolite formations extend from freshwater to intertidal zones and are dominated by Cyanobacteria, Bacteroidetes and Proteobacteria [14]. A key difference between SA stromatolites and hypersaline/marine stromatolites is chemical stability: The SA stromatolites are impacted by a far more chemically unstable environment due to the frequent mixing of fresh and tidal waters [15]. Thus the SA stromatolite formations face numerous environmental pressures such as desiccation, limited inorganic phosphorus and periodic

inundation by seawater, which affect the nutrient concentrations, temperature and chemistry of the system [15]. These formations are characterized by their proximity to the ocean, where stromatolites in the upper formations receive freshwater from the inflow seeps, middle formation stromatolites withstand a mix of freshwater seepage and marine over-topping and lower formations are in closest contact with the ocean [14]. Microbial communities within these levels therefore likely experience distinct environmental pressures, including fluctuations in salinity and dissolved oxygen [15]. While carbon predominantly enters these systems through cyanobacterial carbon fixation, it is unclear how other members of the stromatolite-associated bacterial consortia influence mineral stratification resulting from the cycling of essential nutrients such as nitrogen, phosphorus and sulfur. Since the SA peritidal stromatolite systems are in constant nutritional and chemical flux with varying influence from the fresh and marine water sources, they present an almost ideal *in situ* testing ground for investigating how stromatolite-associated microbial consortia interact with their environment. Identification of conserved bacterial species across both time and space and across varied environmental pressures would suggest that these bacteria are not only robust but likely play important roles within the peritidal stromatolite consortia.

Using a metagenomic approach, we sought to gain insight into the foundational metabolic processes that result in stromatolite formation. We assembled and annotated 183 putative bacterial genomes from peritidal stromatolites of two geographically isolated sites near Port Elizabeth, South Africa. We identified several temporally and

spatially conserved bacterial species and functional gene sets, which may play a central role in establishing and maintaining peritidal stromatolite microbial communities.

METHODS

Sample collection and DNA isolation. Samples for metagenomic analysis were collected from Schoenmakerskop (34°02'28.2"S 25°32'18.6"E) and Cape Recife (34°02'42.1"S 25°34'07.5"E) in January and April 2018, while samples used for 16SrRNA analysis were collected in July 2019. The samples were stored in RNAlater at -20 °C. DNA was extracted from ~1g of sample using Zymo quick DNA Fecal/Soil Microbe Miniprep Kit (Zymo Research, Cat No. D6010) according to the manufacturer's instructions.

Amplicon sequence analysis. Kapa HiFi Hotstart DNA polymerase (Roche, Cat No. KK2500) was used to generate amplicon libraries of the V4-V5 region of the 16s rRNA ribosomal subunit gene with the primer pair E517F (5'-GTAAGGTTCYTCGCGT-3') and E969-984 (5'-CAGCAGCCGCGGTAA-3') [16] using the following cycling parameters: Initial denaturation at 98 °C; 5 cycles (98 °C for 45 seconds, 45 °C for 45 seconds and 72 °C for 1 minute); 18 cycles (98 °C for 45 seconds, 50 °C for 30 seconds and 72 °C for 1 minute; final elongation step at 72 °C for 5 minutes. PCR products were purified using the Bioline Isolate II PCR and Gel kit (Bioline, Cat. No. BIO-52060). Samples were sequenced using the Illumina Miseq platform. Amplicon library datasets were processed and curated using the Mothur software platform [17]. Sequences shorter than 200 nucleotides or containing ambiguous bases or homopolymeric runs greater than 7 were discarded. Sequences were classified using Naive-Bayesian classifier against the

Silva database (v132) and Vsearch software [18] was used to remove chimeras. Reads that were 97% similar were combined into operational taxonomic units (OTUs) using the Opticlust method [19]. Abundance data was then standardized by total, transformed by square root and statistically analyzed using the Primer-e (V7) software package [20].

Metagenomic binning. Shotgun DNA libraries were prepared and sequenced using the IonTorrent Ion P1.1.17 Chip technology (Central Analytical Facilities, Stellenbosch University, South Africa). Adapters were trimmed and trailing bases (15 nts) were iteratively removed if the average quality score for the last 30 nts was lower than 16, which resulted in approximately 30–45 million reads per sample. Resultant metagenomic datasets were assembled into contiguous sequences (contigs) with SPAdes version 3.12.0 [21] using the –iontorrent and --only-assembler options with kmer values of 21,33,55,77,99,127. Contigs were then clustered in putative genomic bins using Autometa (Master branch, commit: a344c28) [22]. Bins were manually curated using a tool developed within the Kwan group (table_and_cluster_col branch, commit: 6e99c2f).

Identification of conserved taxa. Conservation of bacterial taxa was calculated using average nucleotide identities (ANIs) of all genomic bins, which were calculated in a pairwise manner using FastANI [23]. All genomic pairs sharing > 97% ANI were subset and considered conserved taxa. Temporal conservation of taxa was assessed between samples from upper formations of Schoenmakerskop and Cape Recife (Fig.1), collected in January and April 2018 respectively. Spatial conservation between sites was analyzed in a similar pairwise manner between upper, middle and lower regions of Schoenmakerskop and Cape Recife.

Genome taxonomic classification. Genomes were classified using the standalone GTDB-Tk tool (version 0.3.2) using the classify workflow and the Genome Taxonomy Database version 89 [24]. The tool appears unable to classify genomes less than 10% complete, as indicated in the respective tables.

Genome taxonomic clustering. Phylogeny of bacterial genomes carrying genes from functional groups investigated here (nitrogen, sulfur, calcium and phosphate metabolism) was inferred using JolyTree [25] with a sketch size of 5 000. JolyTree infers phylogeny through computation of dissimilarity of kmer sketches, which is then transformed for the estimation of substitution events of the genomes' evolution [25].

Genome annotation. Manually curated putative genomic bins were annotated using Prokka version 1.13 [26], with GenBank compliance enabled. Protein-coding amino-acid sequences from genomic bins were annotated against the KEGG database using kofamscan [27] with output in mapper format. KEGG orthologs were manually collected into functional groups and annotations were extracted from kofamscan output: Briefly, a dictionary of genomes was made with all KO numbers supplied in a KEGG annotation query list (i.e. functional group). Kofamscan output per genome was parsed and each time an entry matching that within the query list was found, a count for that annotation was increased by one, resulting in a count table of functional group KO annotations per genome. Calcium binding proteins and calcium transporting ATPases were counted from PROKKA annotations by searching "Leukotoxin" (Comparison of genes annotated as leukotoxins, against the NCBI database were identified as calcium binding rather than toxins) and "Calcium transporting ATPase" respectively.

Comparative analysis with Shark Bay. A total of 96 MAGs representative of the microbiota associated with Shark Bay stromatolites were downloaded from MG-RAST (https://www.mg-rast.org/linkin.cgi?project=mgp81948) [7]. These genomes were annotated using Prokka and kofamscan and classified using GTDB-Tk as described for the genomes of Cape Recife and Schoenmakerskop stromatolites.

Data availability. Raw 16S rRNA gene amplicon sequence files were uploaded to the NCBI sequence read archive database in BioProject PRJNA574289. Raw reads and binned genomes will be deposited in GenBank and respective accession numbers will be included in the accepted version of this manuscript. Assembled metagenomes were also uploaded to MG-RAST (Keegan et al., 2016) with the identifiers: mgm4784267, mgm4790047, mgm4797596, mgm4797628, mgm4797629, mgm4797630, mgm4802824, mgm4802825 and mgm4802891.

RESULTS AND DISCUSSION

Peritidal tufa stromatolites occur at several sites along the southeastern coast of South Africa. Two stromatolite sites, Cape Recife and Schoenmakerskop, which are approximately 3 km apart and have been well characterized with respect to their chemical environment [14,15] were chosen for this study. Stromatolite formations at both sites begin at a freshwater inflow and end before the subtidal zone (Fig. 1A-B) and are exposed to different levels of tidal disturbance. Samples were collected from the upper stromatolites at Cape Recife and Schoenmakerskop in January and April 2018 for comparisons over time and geographic space. Additional samples were collected from

middle and lower formations for extended comparison across the two sites (Fig. 1). Throughout this investigation, sample prefixes correspond to the site and time at which they were collected (Table S1).

Phylogenetic distribution of microbial communities

We assessed the diversity and structure of the bacterial communities in the stromatolites at Cape Recife and Schoenmakerskop using 16S rRNA gene amplicon sequence analysis. All communities were dominated by Cyanobacteria, Bacteroidetes, Alphaproteobacteria, Gammaproteobacteria and other unclassified bacteria (Fig. 2A), which was in agreement with a previous study at Schoenmakerskop [14] and other studies in the Dorrington laboratory (C. Damarajanan, E.W.Isemonger and R.A. Dorrington, unpublished data). Bacterial communities associated with the upper formations were distinct from those in middle and lower formations ($R>0.6$; $P=0.01$) (Fig. 2B) consistent with the different environmental conditions across the system.

Binning and phylogenetic classification of putative genomes

To characterize the metabolic potential of individual bacteria, we sequenced a shotgun metagenomic library from the 8 samples representative of the upper, middle and lower formations at both sites well as an additional two samples from the upper formations of each site, collected 4 months previously in January 2018. Binning of assembled metagenomic libraries using Autometa [22], resulted in a total of 183 bacterial genome bins from the 10 different metagenomic datasets (Table S1). These bins were manually curated for optimal purity and completion. Fundamental characteristics of each genome

bin were calculated, and the taxonomic identity was assigned using GTDB-Tk [28] (Table S2). The GTDB-Tk tool uses the Genome Taxonomy Database as a reference for classification, which is based on phylogeny inferred from concatenated protein alignments. This approach enabled the removal of polyphyletic groups and assignment of taxonomy from evolutionary divergence. The resulting taxonomy incorporates substantial changes in comparison to the NCBI taxonomy [28]. Equivalent NCBI taxonomic classifications have been provided for clarity in Table S2. Using relative coverage per sample as a proxy for abundance we found that genomes classified within the Cyanobacteriia class were consistently dominant in all collection points, while Alphaproteobacteria, Gammaproteobacteria and Bacteroidia were less abundant but notable bacterial classes (Fig. 3). This distribution appears to be approximately congruent with abundances observed in the 16S rRNA gene amplicon analyses (Fig. 2A).

Temporal and spatial conservation of bacterial species

We calculated pairwise average nucleotide identity (ANI) between all binned genomes and defined conserved species as genomes sharing more than 97% ANI [23] (Table S3). Species classified within the family Phormidiaceae were conserved in upper formations in both January and April samples, as well as within the middle formations (Table 1). Species within the genera *Acaryochloris* and *Hydrococcus,* and family Absconditabacterales were conserved across upper formations. Seven species were conserved across the middle formations, including species classified within genus *Rivularia* and Phormidesmiaceae and Spirulinaceae families (Table 1). Species

classified within family Elainellaceae but were distinctly different species wherein species A within family Elainellaceae was detected only in the upper pools of Cape Recife, whilst the Elainellaceae species B was conserved across the middle formations of both sites (Table 1). Bacterial species within genus *Microcoleus* and family Leptolyngbyaceae were spatially conserved, and present in the January samples of Schoenmakerskop but were not detected in metagenomes from the January sample of Cape Recife (Table 1).

Also of interest was the presence of conserved bacterial species within the order Absconditabacterales, which are classified under the Patescibacteria phylum. Patescibacteria are unusually small bacteria found in groundwater and produce large surface proteins hypothesized to help them attach or associate with other microorganisms that perform nitrogen, sulfur and iron cycling [29]. The presence of these conserved bacteria suggests that the inflow water seeps originate from groundwater. There was a lack of conserved bacterial species across the lower formations and may be due to increased variability at the subtidal interface and potentially limited sample size. Furthermore, some bacterial genomes defined here may represent transient bacteria that may not play any role in the stromatolite system.

Metabolic potential of binned genomes

Oxygenic photosynthesis by Cyanobacteria results in rapid fixation of carbon dioxide and an increase alkalinity [30]. Carbonate ions bind cations such as calcium and are precipitated under alkaline conditions, promoting the growth of stromatolite structures

[1]. Given their numerical dominance in Cape Recife and Schoenmakerskop stromatolites, and their predicted role in other stromatolites [1], the role of carbon cycling is likely performed by Cyanobacteria (Cyanobacteriia class). The identity of the bacteria that cycle redox sensitive sulfur, phosphate, nitrogen and calcium, and subsequently affect the alkalinity and solubility index enabling carbonate precipitation in these stromatolites remain unknown. We inspected PROKKA and KEGG annotations within stromatolite-associated bacterial genome bins to identify potential metabolic pathways that may drive the "alkalinity engine" and promote mineral deposition and accretion [1]. An overview of the results presented here are summarized in Fig. 4 and Table 1.

It has been shown in experimental models that the uptake of hydrogen ions during reduction increases the pH and results in the release of carbonate ions in stromatolites [31]. The increased concentration of carbonate drives an increase in the saturation index of calcite, resulting in precipitation [31]. Therefore, bacteria that reduce sulfate will promote the growth of stromatolites, whilst oxidizers will likely drive dissolution of the calcite precipitate. Amongst the Cape Recife and Schoenmakerskop stromatolites-associated bacteria, the capacity for sulfate reduction was confined to only a few genomes (Fig. S1 and S2). The complete set of genes required for assimilatory sulfate reduction (*sat*/*met3*, *cysC*, *cysH* and *sir* genes) [32] were recovered in four genomes, three of which were conserved *Acaryochloris* genomes (Fig. S3) and the complete set of genes for uptake and desulfonation of alkanesulfonates (*ssuABCDE*) [33] were detected exclusively in conserved *Hydrococus* species (Fig. S3). Alkanesulfonate

metabolism results in the release of sulfite and an aldehyde, the former of which can be reduced by sulfite reductase (*sir* gene) [34]. Both *Acaryochloris* and *Hydrococcus* species were detected exclusively in upper stromatolite formations and suggests freshwater inflow may carry sulfur compounds, however, in the absence of experimental data this remains purely speculative. Reduction of sulfate has previously been shown to promote the precipitation of carbonates in the form of micritic crusts in Bahamian and Australian stromatolites [7, 35] and it has been suggested that microbial cycling of sulfur played an important role in ancient Australian stromatolites, even prior to the emergence of Cyanobacteria [36, 37]. The cumulative removal of hydrogen by these reduction processes would suggest conserved *Hydrococcus* and *Acaryochloris* species may drive an alkaline pH within the system and potentially aid in calcite accretion. For a comparative analysis of these stromatolites with other, well-studied systems, a collection of 96 putative genomes from hypersaline Shark Bay (Australia) stromatolites [7] was downloaded, annotated and classified as performed for the genomes from the Cape Recife and Schoenmakerskop stromatolites (Table S4). This dataset was chosen for comparative analysis as it is the only other set of putative genomes isolated from stromatolite shotgun metagenomic data, however detailed functional potential analysis of individual genomes was not carried out previously [7]. Our analysis of the functional potential of binned genomes from hypersaline Australian stromatolites [7] showed a greater abundance of genomes capable of both dissimilatory and assimilatory sulfate reduction (Fig. S4) and the majority of these genomes were classified within the Desulfobacterota phylum (Table S4). This would suggest that reduction of sulfur compounds is important to stromatolite growth but the differing environmental pressures

(i.e. fresh vs hypersaline water) have resulted in different bacterial species performing this conserved function. Sulphur reduction in stromatolites has been previously attributed to strictly non-cyanobacterial, sulfate-reducing bacteria [38–40] but our analysis would suggest that cyanobacterial *Hydrococcus* and *Acaryochloris* species may be the drivers of sulfur reduction in the Cape Recife and Schoenmakerskop stromatolites.

The reduction of nitrogen, nitrates and nitrites has a similar effect to sulfide, in that the uptake of protons during the formation of ammonia, results in increased alkalinity that could lead to calcite precipitation [7, 41–44]. The release of ammonia also aids in calcification as the ammonia absorbs carbon dioxide, resulting in an increase in carbonate ions [43]. Metabolism of nitrogenous compounds has likewise been proposed to have emerged at approximately the same time as sulfur metabolism, as both processes share similar redox states [45]. Microbial reduction of nitrogen has been dated back 3.4 billion years [46] and ammonium availability during this time may have sustained developing microbial life [47]. Therefore, bacteria that can fix nitrogen or reduce nitrates/nitrites could potentially promote the growth of stromatolites and may have added to the formation of ancient analogues. We found that several bacteria were capable of ferredoxin-dependent assimilatory nitrate reduction (ANR) (*nirA-narB* genes) [48], nitrogen fixation (*nifDHK* genes) and dissimilatory nitrite reduction (DNR) (*nirBD/nrfAH* genes), all of which result in the production of ammonia [49] (Fig. S5 and S6). The potential for assimilatory nitrate reduction was detected in several genomes, primarily from the middle formations, including conserved *Rivularia* sp.,

Phormidesmiaceae, Phormidiaceae and Elainellaceae bacterial species (Fig. S7). The potential for dissimilatory reduction of nitrites, via either cytoplasmic *nirBD* or membrane-bound *nrfAH* nitrite reductases, was detected in several genomes across both the upper and middle formations, including conserved *Hydrococcus* bacterial species (Fig. S7). Nitrogen fixation during the night has previously been shown to be a driver of carbonate precipitation in stromatolites [1]. Genes associated with nitrogen fixation were identified in several bacteria associated with both upper and middle formations, including conserved *Hydrococcus* species across the upper formations and conserved Phormidiaceae, Spirulinaceae and Chloroflexaceae species across the middle formations, as well as non-conserved species within the *Blastochloris* genus (Fig. S5 - S7). These species may contribute to the growth of the stromatolites through the formation and release of ammonia. The capacity for nitrogen fixation being retained in Schoenmakerskop and Cape Recife is unexpected given the high seasonal concentrations of dissolved inorganic nitrogen ranging from 95-450 μM at the two sites [15]. Review of the genomes from Australian stromatolites revealed that 14 genomes were capable of nitrate reduction and 7 were capable of nitrogen fixation. The genomes were taxonomically diverse, but species within phyla Planctomycetota and Desulfobacterota were most prominent. As with the reduction of sulfur compounds, it would appear that the function remains conserved but that different bacteria perform these functions across diverse conditions of these stromatolites.

Endolithic bacteria bore into carbonate substrate through dissolution facilitated by uptake, transport and eventual release of calcium ions at the distal end of multicellular

cyanobacterial trichomes (Fig. 3) [50, 51]. Increased concentration of free calcium ions at the surface of the stromatolite may promote calcium carbonate precipitation [31]. Certain genes are upregulated during this boring activity, including key genes encoding calcium ATPases and calcium-binding proteins [51, 52]. We identified an incredible 1 182 gene copies encoding calcium-binding proteins and 315 calcium ATPases in the putative genomes of bacteria from Cape Recife and Schoenmakerskop stromatolites. The average gene copy number of calcium binding proteins and calcium ATPases was 6.5 and 1.7 respectively. The greatest abundance of gene copies was found in conserved *Microcoleus*, Phormidiaceae and Elainellaceae genomes (Table 1, Fig. S8), carrying 20–65 copies of the calcium binding protein gene and 1–7 copies of the calcium ATPase gene. The potential action of these bacteria on the formation of stromatolites is two-fold: first, the boring activity results in sand grains “welding” together and increases the structural stability of the stromatolite [10], and secondly, the release of concentrated calcium ions at the surface of the stromatolite could lead to rapid mineralization on contact with either carbonate ions [1, 53] or phosphate ions [54]. Endolithic bacteria have been identified in Bahamian stromatolites (*Solentia* sp.) [10] and may function in a similar manner to the conserved bacterial species identified here. Among the Australian stromatolites two genomes (S_098 and S305) (Fig. S4), classified within Elainellaceae and Phormidesmiales (Table S4), exhibited the genetic potential for endolithic activity. These genomes carry the 22–26 gene copies of the calcium binding protein and 3 gene copies (each) of the calcium transport ATPases. The abundances in the Australian stromatolites were not as high as that observed in bacteria from Cape Recife and Schoenmakerskop but the potential endolithic function appears to be

conserved within Elainellaceae species, which were conserved in the middle formations in Cape Recife and Schoenmakerskop. However, the shared ANI of the Australian and South African Elainellaceae species was between 75–77%, indicating that these genomes are only distant phylogenetic relatives.

Phosphatic structures recently observed within Cape Recife stromatolites [55] are particularly intriguing as they closely resemble ancient phosphatic stromatolites from the Paleoproterozoic and Neoproterozoic eras in which several major oxygenation events occurred [55]. Hydroxyapatite ($Ca_5(PO_4)_3(OH)$) is more easily precipitated than calcite ($CaCO_3$) and the release of inorganic phosphate into the biofilm by alkaline phosphatase activity could create a nucleation point for the initiation of apatite mineralization, the rate of which may be increased in the presence of an alkaline environment [56]. We have identified several conserved bacterial species here that can drive the alkalinity of the stromatolite system, as well as conserved species that concentrate calcium ions at the stromatolite surface (Fig. 3). Apatite could readily be precipitated if a concentrated source of phosphate ions was provided. We detected a large number of genes encoding alkaline phosphatases among the stromatolite-associated bacteria, with genes encoding extracellular PhoX being most abundant (190 copies across all genomes), with fewer copies of genes encoding intracellularly-localized PhoD (118 copies) and PhoA/B (85 copies) detected [57] (Fig. S9 and S10). Gene copies of *phoD* and *phoX* have been found to be variable in soil microbes but the majority of bacteria carry only one copy of *phoD* and one copy of *phoX* [58, 59]. Conserved bacterial species within *Acaryochloris*, *Rivularia* and Elainellaceae groups

were particularly notable (Table 1, Fig. S11), carrying between 2–4 copies of *phoX*. Several genomes from both sites carried genes encoding more than one type of alkaline phosphatase, which may indicate that these bacteria can utilize different enzymes, perhaps in response to a flux in available co-factors (i.e. calcium vs. zinc availability) [60].

Genes encoding the phosphate-sensing regulatory enzyme (*phoR*) of the canonical phosphate transporter (*pstABCS*) was absent in most stromatolite-associated bacteria (Fig. S9 -11) from the Cape Recife and Schoenmakerskop sites. The loss of *phoR* could potentially lead to toxic, uncontrolled P_i uptake [61] when external Pi levels are high. However, a recent experiment showed that under limited phosphate conditions, the growth of *Xanthomonas oryzae* Δ*phoR* deletion mutants was not significantly different from wild type strains [62] corroborating earlier studies in which *E. coli* Δ*phoR* deletion mutants had a greater biomass concentration than wild type strains under phosphate limited conditions [63]. There may be several other genes, or potentially divergent orthologs of *phoR* in these genomes [64, 65], which would account for the apparent absence of *phoR*, but the widespread absence of this gene across taxonomically diverse bacteria at Cape Recife and Schoenmakerskop, suggests that this seemingly important regulatory protein is not advantageous to this system. It is hypothesized that stromatolites of Cape Recife and Schoenmakerskop largely receive phosphorous from the oceans and have consistently been shown to experience limited inorganic phosphate availability [15, 66–68] and this consistent struggle for phosphate may have resulted in the loss of *phoR* as its role may have become redundant. Loss of PhoR has

also resulted in increased alkaline phosphatase activity [69] and consequently, it is possible that the combined effect of extracellular alkaline phosphatase abundance and loss of *phoR*, could result in a relative excess of local bioavailable phosphate in the extracellular stromatolite environment. Cyanobacteria in phosphorus-poor rivers have been hypothesized to generate bioavailable phosphate from trapped sediments in biofilms generating bioavailable phosphate concentrations 320-fold higher than the surrounding water [70]. An increase in both phosphate and calcium ions could result in the rapid precipitation of apatite, as observed in marine phosphorites, freshwater lakes and in some soil bacteria [54, 71, 72] and could account for the phosphatic deposits observed within the SA stromatolites [55].

Finally, conserved species classified within the genus *Rivularia* and several other non-conserved bacterial species harbored the 11 genes required for the transport (*phnCDEF*) and lysis (*phnGHIJKLM*) of phosphonate compounds (Fig. S9, S10 and S12) [73]. Phosphonates are characterized by the presence of a carbon-phosphorous bond and are biosynthesized by several organisms [74]. The C-P lyase (*phnGHIJLKM*) breaks the C-P bond within a wide array of phosphonate substrates resulting in a hydrocarbon and inorganic phosphate [74]. The inorganic phosphate may be used within the bacterial cell or released aiding in accretion through increased ion concentration [75]. However, phosphonates within an environment can prevent precipitation of calcite by binding to crystal growth sites [76] and the degradation of these compounds within the stromatolite formations may prevent chemical inhibition of stromatolite growth. The potential for phosphonate degradation was also detected in 8

genomes from the Australian stromatolites (Fig. S4), including the same species potentially capable of endolithic activity (S_098: Elainellaceae and S_305: Phormidesmiales). The conservation of phosphonate metabolism and potential endoliths as common features across stromatolites would indicate that these bacterial processes are vital within the stromatolite system.

It would appear that several bacterial groups are key players in the formation of stromatolites in Cape Recife and Schoenmakerskop, with conserved bacterial species acting as major contributors (Table 1). We propose the redox potential and solubility index in SA stromatolites is potentially controlled by conserved bacteria, including conserved *Hydrococcus*, *Acaryochloris* and Phormidiaceae bacterial species in upper formations, and Elainellaceae, Phormidiaceae, Phormdesmiaceae, *Rivularia* and Spirulinaceae bacterial species in middle formations (Table 1), through the generation and concentration of sulfide, ammonia and calcium ions for the rapid precipitation of carbonate compounds and subsequent growth of stromatolite structures. We further propose an abundance of alkaline phosphatases in combination with a suspected loss of transport regulatory protein PhoR, may result in increased local inorganic phosphate concentrations (Fig. 4). This free inorganic phosphate may bind with the calcium released by boring endolithic bacteria to form the phosphatic crusts previously observed in these stromatolites. Conservation of these functions, and in some cases taxonomic groups, across these and Australian stromatolites would suggest these processes may be especially essential functions within the system. The identification of conserved bacteria performing potentially vital roles within stromatolites of Cape Recife and

Schoenmakerskop may provide insight into what cyanobacterial species may have played a key role in the formation of ancient phosphatic stromatolites.

ACKNOWLEDGEMENTS

The authors acknowledge Caro Damarjanan who conducted a pilot study that led to this study. We also wish to thank Karthik Anantharaman (University of Wisconsin, Madison) for his helpful critique on this manuscript.

The authors acknowledge funding from the Gordon and Betty Moore Foundation (Grant number 6920) (awarded to R.A.D and J.K.), grants awarded to R.A.D by the South African National Research Foundation (UID: 87583 and 109680) and scholarship awarded to E.R.R from American National Science Foundation (DBI-1845890). This research was performed in part using the computer resources and assistance of the UW-Madison Center for High Throughput Computing (CHTC) in the Department of Computer Sciences. The CHTC is supported by UW-Madison, the Advanced Computing Initiative, the Wisconsin Alumni Research Foundation, and Wisconsin Institutes for Discovery, and the National Science Foundation and is an active member of the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science. The authors also acknowledge the Centre for High Performance Computing (CHPC, South Africa) for providing computing facilities for bioinformatics data analysis. Opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to any of the above-mentioned funding agencies.

CONFLICT OF INTERESTS

The authors declare no financial competing interests.

Supplementary information is available at *ISME Journal*'s website

REFERENCES

1. Dupraz C, Reid RP, Braissant O, Decho AW, Norman RS, Visscher PT. Processes of carbonate precipitation in modern microbial mats. *Earth-Sci Rev* 2009; **96**: 141–162.

2. Soo RM, Hemp J, Parks DH, Fischer WW, Hugenholtz P. On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. *Science* 2017; **355**: 1436–1440.

3. Awramik SM. The History and Significance of Stromatolites. In: Schidlowski M, Golubic S, Kimberley MM, McKirdy DM, Trudinger PA (eds). Early organic evolution: Implications for mineral and energy resources. 1992. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 435–449.

4. K. L. Gallagher S. Daniels C. Norris M. E. Cantino D.A. Knecht N. Stork A. Fowler C. Dupraz P. T. Visscher. Growth and mineralization of a biofilm of sulfate-reducing bacteria. 2010. Astrobiology Science Conference 2010: Evolution and Life:

Surviving Catastrophes and Extremes on Earth and Beyond, held April 26-20, 2010 in League City, Texas. LPI Contribution No. 1538, p.5498

5. Gallagher KL, Kading TJ, Braissant O, Dupraz C, Visscher PT. Inside the alkalinity engine: the role of electron donors in the organomineralization potential of sulfate-reducing bacteria. *Geobiology* 2012; **10**: 518–530.

6. Casaburi G, Duscher AA, Pamela Reid R, Foster JS. Characterization of the stromatolite microbiome from Little Darby Island, The Bahamas using predictive and whole shotgun metagenomic analysis. *Environ Microbiol* 2016; **18**: 1452–1469.

7. Wong HL, White RA, Visscher PT, Charlesworth JC, Vázquez-Campos X, Burns BP. Disentangling the drivers of functional complexity at the metagenomic level in Shark Bay microbial mat microbiomes. *ISME J* 2018; **12**: 2619–2639.

8. Khodadad CLM, Foster JS. Metagenomic and metabolic profiling of nonlithifying and lithifying stromatolitic mats of Highborne Cay, The Bahamas. *PLoS One* 2012; **7**: e38229.

9. Babilonia J, Conesa A, Casaburi G, Pereira C, Louyakis AS, Reid RP, et al. Comparative Metagenomics provides insight into the ecosystem functioning of the Shark Bay stromatolites, Western Australia. *Front Microbiol* 2018; **9**: 1359.

10. Macintyre IG, Prufert-Bebout L, Reid RP. The role of endolithic cyanobacteria in the formation of lithified laminae in Bahamian stromatolites. *Sedimentology* 2000; **47**: 915–921.

11. Dupraz C, Visscher PT. Microbial lithification in marine stromatolites and hypersaline mats. *Trends Microbiol* 2005; **13**: 429–438.

12. Visscher PT, Pamela Reid R, Bebout BM, Hoeft SE, Macintyre IG, Thompson JA JR. Formation of lithified micritic laminae in modern marine stromatolites (Bahamas): The role of sulfur cycling. *American Mineralogist* 1998; **83**: 1482–1493.

13. Smith A, Cooper A, Misra S, Bharuth V, Guastella L, Botes R. The extant shore platform stromatolite (SPS) facies association: A glimpse into the Archean? *Biogeosciences* 2018; **15**: 2189–2203.

14. Perissinotto R, Bornman TG, Steyn P-P, Miranda NAF, Dorrington RA, Matcher GF, et al. Tufa stromatolite ecosystems on the South African south coast. *S Afr J Sci* 2014; **110**: 01–08.

15. Rishworth GM, van Elden S, Perissinotto R, Miranda NAF, Steyn P-P, Bornman TG. Environmental influences on living marine stromatolites: Insights from benthic microalgal communities. *Environ Microbiol* 2016; **18**: 503–513.

16. Matcher GF, Dorrington RA, Henninger TO, Froneman PW. Insights into the bacterial diversity in a freshwater-deprived permanently open Eastern Cape estuary, using 16S rRNA pyrosequencing analysis. *Water SA* 2011; **37**: 381-390

17. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009; **75**: 7537–7541.

18. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* 2016; **4**: e2584.

19. Westcott SL, Schloss PD. OptiClust, an Improved Method for Assigning Amplicon-Based Sequence Data to Operational Taxonomic Units. *mSphere* 2017; **2**: e00073-17

20. Gorley RN, Clarke KR. PRIMER v7: User Manual/Tutorial. 2015. Massey University, Albany Campus.

21. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–477.

22. Miller IJ, Rees ER, Ross J, Miller I, Baxa J, Lopera J, et al. Autometa: Automated extraction of microbial genomes from individual shotgun metagenomes. *Nucleic Acids Res* 2019; **47**: e57.

23. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018; **9**: 5114

24. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. Selection of representative genomes for 24,706 bacterial and archaeal species clusters provide a complete genome-based taxonomy. *bioRxiv* 2019; 771964

25. Criscuolo A. A fast alignment-free bioinformatics procedure to infer accurate distance-based phylogenetic trees from genome assemblies. *Riogrande Odontol* 2019; **5**: e36178.

26. Seemann T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–2069.

27. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *bioRxiv* 2019.

28. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 2018; **36**: 996–1004.

29. Herrmann M, Wegner C-E, Taubert M, Geesink P, Lehmann K, Yan L, et al. Predominance of *Cand.* Patescibacteria in groundwater is caused by their preferential mobilization from soils and flourishing under oligotrophic conditions. *Front Microbiol* 2019; **10**: 1407.

30. Pace A, Bourillot R, Bouton A, Vennin E, Braissant O, Dupraz C, et al. Formation of stromatolite lamina at the interface of oxygenic-anoxygenic photosynthesis. *Geobiology* 2018; **16**: 378–398.

31. Bosak T, Greene SE, Newman DK. A likely role for anoxygenic photosynthetic microbes in the formation of ancient stromatolites. *Geobiology* 2007; **5**: 119–126.

32. Santos AA, Venceslau SS, Grein F, Leavitt WD, Dahl C, Johnston DT, et al. A protein trisulfide couples dissimilatory sulfate reduction to energy conservation. *Science* 2015; **350**: 1541–1545.

33. Ellis HR. Mechanism for sulfur acquisition by the alkanesulfonate monooxygenase system. *Bioorg Chem* 2011; **39**: 178–184.

34. Aguilar-Barajas E, Díaz-Pérez C, Ramírez-Díaz MI, Riveros-Rosas H, Cervantes C. Bacterial transport of sulfate, molybdate, and related oxyanions. *Biometals* 2011; **24**: 687–707.

35. Reid RP, Visscher PT, Decho AW, Stolz JF, Bebout BM, Dupraz C, et al. The role of microbes in accretion, lamination and early lithification of modern marine stromatolites. *Nature* 2000; **406**: 989–992.

36. Bontognali TRR, Sessions AL, Allwood AC, Fischer WW, Grotzinger JP, Summons RE, et al. Sulfur isotopes of organic matter preserved in 3.45-billion-year-old stromatolites reveal microbial metabolism. *Proc Natl Acad Sci U S A* 2012; **109**: 15146–15151.

37. Allen JF. A Proposal for Formation of Archaean Stromatolites before the Advent of Oxygenic Photosynthesis. *Front Microbiol* 2016; **7**: 1784.

38. Baumgartner LK, Reid RP, Dupraz C, Decho AW, Buckley DH, Spear JR, et al. Sulfate reducing bacteria in microbial mats: Changing paradigms, new discoveries. *Sediment Geol* 2006; **185**: 131–145.

39. Petrisor AI, Szyjka S, Kawaguchi T, Visscher PT, Norman RS, Decho AW. Changing microspatial patterns of sulfate-reducing microorganisms (SRM) during cycling of marine stromatolite mats. *Int J Mol Sci* 2014; **15**: 850–877.

40. Spring S, Sorokin DY, Verbarg S, Rohde M, Woyke T, Kyrpides NC. Sulfate-reducing bacteria that produce exopolymers thrive in the calcifying zone of a hypersaline cyanobacterial mat. *Front Microbiol* 2019; **10**: 862.

41. Rodriguez-Navarro C, Rodriguez-Gallego M, Ben Chekroun K, Gonzalez-Muñoz MT. Conservation of ornamental stone by *Myxococcus xanthus*-induced carbonate biomineralization. *Appl Environ Microbiol* 2003; **69**: 2182–2193.

42. Wei S, Cui H, Jiang Z, Liu H, He H, Fang N. Biomineralization processes of calcite induced by bacteria isolated from marine sediments. *Braz J Microbiol* 2015; **46**: 455–464.

43. Konopacka-Łyskawa D, Kościelska B, Karczewski J, Gołąbiewska A. The influence of ammonia and selected amines on the characteristics of calcium carbonate precipitated from calcium chloride solutions via carbonation. *Mater Chem Phys* 2017; **193**: 13–18.

44. Lee YS, Park W. Enhanced calcium carbonate-biofilm complex formation by alkali-generating *Lysinibacillus boronitolerans* YS11 and alkaliphilic *Bacillus* sp. AK13. *AMB Express* 2019; **9**: 49.

45. Thomazo C, Ader M, Philippot P. Extreme 15N-enrichments in 2.72-Gyr-old sediments: Evidence for a turning point in the nitrogen cycle. *Geobiology* 2011; **9**: 107–120.

46. Stüeken EE. Nitrogen in Ancient Mud: A Biosignature? *Astrobiology* 2016; **16**: 730–735.

47. Yang J, Junium CK, Grassineau NV, Nisbet EG, Izon G, Mettam C, et al. Ammonium availability in the Late Archaean nitrogen cycle. *Nat Geosci* 2019; **12**: 553–557.

48. Moreno-Vivián C, Cabello P, Martínez-Luque M, Blasco R, Castillo F. Prokaryotic nitrate reduction: Molecular properties and functional distinction among bacterial nitrate reductases. *J Bacteriol* 1999; **181**: 6573–6584.

49. Griffith JC. Insights into the soil microbial communities in New Zealand indigenous tussock grasslands. 2016. PhD, University of Otago.

50. Garcia-Pichel F, Ramírez-Reinat E, Gao Q. Microbial excavation of solid carbonates powered by P-type ATPase-mediated transcellular Ca^{2+} transport. *Proc Natl Acad Sci U S A* 2010; **107**: 21749–21754.

51. Guida BS, Garcia-Pichel F. Extreme cellular adaptations and cell differentiation required by a cyanobacterium for carbonate excavation. *Proc Natl Acad Sci U S A* 2016; **113**: 5712–5717.

52. Guida BS. Unique cellular, physiological, and metabolic adaptations to the euendolithic lifestyle in a boring cyanobacterium. 2016. PhD, Arizona State University.

53. Kamennaya NA, Ajo-Franklin CM, Northen T, Jansson C. Cyanobacteria as biocatalysts for carbonate mineralization. *Minerals* 2012; **2**: 338–364.

54. Cosmidis J, Benzerara K, Guyot F, Skouri-Panet F, Duprat E, Férard C, et al. Calcium-phosphate biomineralization induced by alkaline phosphatase activity in *Escherichia coli*: Localization, kinetics, and potential signatures in the fossil record. *Front Earth Sci Chin* 2015; **3**: 84

55. Buttner SH, Isaacs M, Isemonger EW, van Niekerk D, Sipler RE, Dorrington RA. Living phosphatic stromatolites - a link to Proterozoic microbial communities. (Under Review)

56. Gallagher KL, Braissant O, Kading TJ, Dupraz C, Visscher PT. Phosphate-related artifacts in carbonate mineralization experiments. *J Sediment Res* 2013; **83**: 37–49.

57. Luo H, Benner R, Long RA, Hu J. Subcellular localization of marine bacterial alkaline phosphatases. *Proc Natl Acad Sci U S A* 2009; **106**: 21219–21223.

58. Ragot SA, Kertesz MA, Mészáros É, Frossard E, Bünemann EK. Soil *phoD* and *phoX* alkaline phosphatase gene diversity responds to multiple environmental factors. *FEMS Microbiol Ecol* 2017; **93**: fiw212

59. Ragot SA, Kertesz MA, Bünemann EK. *phoD* Alkaline phosphatase gene diversity in soil. *Appl Environ Microbiol* 2015; **81**: 7281–7289.

60. Sebastian M, Ammerman JW. The alkaline phosphatase PhoX is more widely distributed in marine bacteria than the classical PhoA. ISME J 2009; **3**: 563–572.

61. Santos-Beneit F. The Pho regulon: A huge regulatory network in bacteria. *Front Microbiol* 2015; **6**: 402.

62. Zheng D, Xue B, Shao Y, Yu H, Yao X, Ruan L. Activation of PhoBR under phosphate-rich conditions reduces the virulence of *Xanthomonas oryzae* pv. oryzae. *Mol Plant Pathol* 2018; **19**: 2066–2076.

63. Marzan L, Shimizu K. Metabolic regulation of *Escherichia coli* and its *phoB* and *phoR* genes knockout mutants under phosphate and nitrogen limitations as well as at acidic condition. *Microb Cell Fact* 2011; **10**: 39.

64. Birkey SM, Liu W, Zhang X, Duggan MF, Marion Hulett F. Pho signal transduction network reveals direct transcriptional regulation of one two-component system by another two-component regulator: *Bacillus subtilis* PhoP directly regulates production of ResD. *Mol Microbiol* 1998; **30**: 943–953.

65. Kočan M, Schaffer S, Ishige T, Sorger-Herrmann U, Wendisch VF, Bott M. Two-component systems of *Corynebacterium glutamicum*: Deletion analysis and involvement of the PhoS-PhoR system in the phosphate starvation response. *J Bacteriol* 2006; **188**: 724–732.

66. Rishworth GM, Perissinotto R, Miranda NAF, Bornman TG, Steyn P-P. Phytoplankton community dynamics within peritidal pools associated with living stromatolites at the freshwater–marine interface. *Aquat Sci* 2017; **79**: 357–370.

67. Rishworth GM, Perissinotto R, Bird MS, Strydom NA, Peer N, Miranda NAF, et al. Non-reliance of metazoans on stromatolite-forming microbial mats as a food resource. *Sci Rep* 2017; **7**: 42614.

68. Rishworth GM, Perissinotto R, Bird MS, Pelletier N. Grazer responses to variable macroalgal resource conditions facilitate habitat structuring. *R Soc Open Sci* 2018; **5**: 171428.

69. Carmany DO, Hollingsworth K, McCleary WR. Genetic and biochemical studies of phosphatase activity of PhoR. *J Bacteriol* 2003; **185**: 1112–1115.

70. Wood SA, Depree C, Brown L, McAllister T, Hawes I. Entrapped sediments as a source of phosphorus in epilithic cyanobacterial proliferations in low nutrient rivers. *PLoS One* 2015; **10**: e0141063.

71. Danen-Louwerse HJ, Lijklema L, Coenraats M. Coprecipitation of phosphate with calcium carbonate in Lake Veluwe. *Water Res* 1995; **29**: 1781–1785.

72. Guang-Can TAO, Shu-Jun T, Miao-Ying CAI, Guang-Hui XIE. Phosphate-solubilizing and-mineralizing abilities of bacteria isolated from soils. *Pedosphere* 2008; **18**: 515–523.

73. Metcalf WW, Wanner BL. Evidence for a fourteen-gene, *phnC* to *phnP* locus for phosphonate metabolism in *Escherichia coli*. *Gene* 1993; **129**: 27–32.

74. White AK, Metcalf WW. Microbial metabolism of reduced phosphorus compounds. *Annu Rev Microbiol* 2007; **61**: 379–400.

75. Rott E, Steinmetz H, Metzger JW. Organophosphonates: A review on environmental relevance, biodegradability and removal in wastewater treatment plants. *Sci Total Environ* 2018; **615**: 1176–1191.

76. Kan AT, Fu G, Tomson MB. Adsorption and precipitation of an aminoalkylphosphonate onto calcite. *J Colloid Interface Sci* 2005; **281**: 275–284.

8.

Table 1. Conservation of bacterial species (ANI > 97%) and their functional potential across upper, middle and lower stromatolite formations at Cape Recife and Schoenmakerskop.

Conserved bacteria taxonomy (ANI >97%)	Upper (Jan)		Upper (April)		Middle (April)		Lower (April)		Notable metabolisms								
	CR	SCH	CR	SCH	CR	SCH	CR	SCH	Assim S reduction	Alkanesulfonate metabolism	Dissim N reduction	Assim N reduction	N fixation	Endolithic activity	Alkaline phosphatases	Phosphonate degradation
Phormidiaceae	X	X	X	X	X	X						X	X	X		
Absconditabacterales	X	X	X	X												
Acaryochloris	X	X	X	X					X						X	
Hydrococcus	X	X	X	X						X	X		X			
Leptolyngbyaceae		X	X	X												
Microcoleus		X	X	X		X								X		
Pleurocapsa	X	X		X												
Elainellaceae sp. A	X		X													
Scytonema	X		X													
Brachymonas	X		X													
Synechococcales	X		X													
Chloroflexaceae						X		X					X			
Elainellaceae sp. B					X	X		X				X		X	X	
JA-3-3Ab					X	X										
Bacteroidales					X	X										
Phormidesmiaceae					X	X		X				X				
Rivularia		X			X	X		X				X			X	X
Spirulinaceae					X	X							X			
UBA4181					X	X										
Calothrix				X		X										

Figure 1. Stromatolites were collected from four different points at two different sites along the SA coastline. (A) Aerial view of sampling locations within the Schoenmakerskop site and (B) sampling locations in the Cape Recife site. Samples were collected from upper (CSU: Schoenmakerskop Inflow Jan, SU: Schoenmakerskop Inflow April, CRU: Cape Recife Inflow Jan, RU: Cape Recife Inflow April), middle (SM1: Schoenmakerskop Middle 1 April, SM2: Schoenmakerskop Middle 2 April, RM1: Cape Recife Middle 1 April, RM2: Cape Recife Middle 1 April) and lower formations (SL: Schoenmakerskop Lower April, RL: Cape Recife Lower April).

Figure 2. Distribution and abundance of bacterial taxa in stromatolite formations. (A) Phylogenetic classification and average relative abundance (n=3) of dominant phyla in different sample sites indicated that all stromatolite samples are dominated by Cyanobacteria, Bacteroidetes, Alpha- and Gammaproteobacteria. (B) OTU abundance was used to cluster stromatolite biological replicates using Bray-Curtis non-dimensional scaling. Samples were isolated from upper (green), middle (red) and lower (blue) stromatolites formations in Cape Recife (triangles) and Schoenmakerskop (circles).

Figure 3. Taxonomic classification of putative genome bins in stromatolites collected from upper/inflow, middle and lower/marine formations of Schoenmakerskop and Cape Recife. Coverage per genome has been used as a proxy for abundance and summed for genomes within each respective taxonomic class. A key for the predicted taxonomic classification has been provided. Abrreviations are as follows: CSU: Schoenmakerskop Inflow Jan, SU: Schoenmakerskop Inflow April, CRU: Cape Recife Inflow Jan, RU: Cape

Recife Inflow April, SM1: Schoenmakerskop Middle 1 April, SM2: Schoenmakerskop Middle 2 April, RM1: Cape Recife Middle 1 April, RM2: Cape Recife Middle 1 April, SL: Schoenmakerskop Lower April, RL: Cape Recife Lower April.

Figure 4. A summarized overview of the bacterial metabolisms predicted in stromatolites from Cape Recife and Schoenmakerskop that may influence mineral precipitation and subsequent growth of stromatolite formations.

Figure S1. Abundance of sulfate metabolism genes in bacterial putative genomes from bacteria associated with stromatolites in Schoenmakerskop. Gene abundance per genome was counted from KEGG annotations detected using kofamscan [27] and are indicated with the scale provided. Genes are identified by both their assigned KO number and accepted abbreviation.

Figure S2. Abundance of sulfate metabolism genes in bacterial putative genomes from bacteria associated with stromatolites in Cape Recife. Gene abundance per genome was counted from KEGG annotations detected using kofamscan [27] and are indicated with the scale provided. Genes are identified by both their assigned KO number and accepted abbreviation.

Figure S3. Summary of genes encoding sulfate-reducing enzymes in stromatolite-associated bacterial genomes. Gene abundance is indicated per the scale provided. Genomes that do not carry any of the genes investigated here were not shown.

Phylogeny was inferred using JolyTree [24] with a sketch size of 5 000. Conserved genomes are indicated with an asterisk within highlighted taxonomic groups. Genome abbreviations indicate site, formation proximity to ocean and collection; CSU: Schoenmakerskop Inflow Jan, SU: Schoenmakerskop Inflow April, CRU: Cape Recife Inflow Jan, RU: Cape Recife Inflow April, SM1: Schoenmakerskop Middle 1, SM2:Schoenmakerskop Middle 2, RM1: Cape Recife Middle 1, RM2: Cape Recife Middle 1, SL: Schoenmakerskop Lower, RL: Cape Recife Lower.

Figure S4. Summarized nutrient metabolisms in genomes bins from stromatolites sampled from Shark Bay, Australia [7]. Gene abundance per functional group per genome was counted from KEGG annotations detected using kofamscan [27] and are indicated with the scale provided. Genes are identified by both their assigned KO number and accepted abbreviation.

Figure S5. Abundance of nitrogen metabolism genes in bacterial putative genomes from bacteria associated with stromatolites in Schoenmakerskop. Gene abundance per genome was counted from KEGG annotations detected using kofamscan [27] and are indicated with the scale provided. Genes are identified by both their assigned KO number and accepted abbreviation.

Figure S6. Abundance of nitrogen metabolism genes in bacterial putative genomes from bacteria associated with stromatolites in Cape Recife. Gene abundance per genome was counted from KEGG annotations detected using kofamscan [27] and are

indicated with the scale provided. Genes are identified by both their assigned KO number and accepted abbreviation.

Figure S7. Summary of genes encoding nitrate/nitrite-reducing and nitrogen fixation enzymes in stromatolite-associated bacterial genomes. Gene abundance is indicated per the scale provided. Genomes that do not carry any of the genes investigated here were not shown. Phylogeny was inferred using JolyTree [24] with a sketch size of 5 000. Conserved genomes are indicated with an asterisk within highlighted taxonomic groups. Genome abbreviations indicate site, formation proximity to ocean and collection; CSU: Schoenmakerskop Inflow Jan, SU: Schoenmakerskop Inflow April, CRU: Cape Recife Inflow Jan, RU: Cape Recife Inflow April, SM1: Schoenmakerskop Middle 1, SM2:Schoenmakerskop Middle 2, RM1: Cape Recife Middle 1, RM2: Cape Recife Middle 1, SL: Schoenmakerskop Lower, RL: Cape Recife Lower.

Figure S8. Summary of genes encoding calcium binding and transport enzymes in stromatolite-associated bacterial genomes. Gene abundance is indicated per the scale provided. Genomes carrying fewer than 6 genes (cumulative) in these two categories were collapsed. Phylogeny was inferred using JolyTree [24] with a sketch size of 5 000. Conserved genomes are indicated with an asterisk within highlighted taxonomic groups. Genome abbreviations indicate site, formation proximity to ocean and collection; CSU: Schoenmakerskop Inflow Jan, SU: Schoenmakerskop Inflow April, CRU: Cape Recife Inflow Jan, RU: Cape Recife Inflow April, SM1: Schoenmakerskop Middle 1,

SM2:Schoenmakerskop Middle 2, RM1: Cape Recife Middle 1, RM2: Cape Recife Middle 1, SL: Schoenmakerskop Lower, RL: Cape Recife Lower.

Figure S9. Abundance of phosphate metabolism genes in bacterial putative genomes from bacteria associated with stromatolites in Schoenmakerskop. Gene abundance per genome was counted from KEGG annotations detected using kofamscan [27] and are indicated with the scale provided. Genes are identified by both their assigned KO number and accepted abbreviation.

Figure S10. Abundance of phosphate metabolism genes in bacterial putative genomes from bacteria associated with stromatolites in Cape Recife. Gene abundance per genome was counted from KEGG annotations detected using kofamscan [27] and are indicated with the scale provided. Genes are identified by both their assigned KO number and accepted abbreviation.

Figure S11. Summary of genes encoding alkaline phosphatases and phosphate transport regulatory proteins in stromatolite-associated bacterial genomes. Gene abundance is indicated per the scale provided. Genomes that do not carry any of the genes investigated here were not shown. Phylogeny was inferred using JolyTree [24] with a sketch size of 5 000. Conserved genomes are indicated with an asterisk within highlighted taxonomic groups. Genome abbreviations indicate site, formation proximity to ocean and collection; CSU: Schoenmakerskop Inflow Jan, SU: Schoenmakerskop Inflow April, CRU: Cape Recife Inflow Jan, RU: Cape Recife Inflow April, SM1:

Schoenmakerskop Middle 1, SM2:Schoenmakerskop Middle 2, RM1: Cape Recife Middle 1, RM2: Cape Recife Middle 1, SL: Schoenmakerskop Lower, RL: Cape Recife Lower.

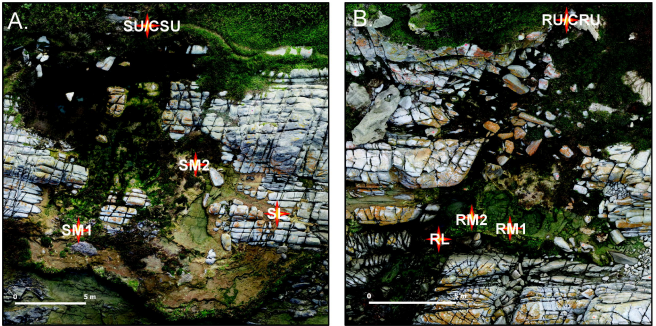
Figure S12. Summary of genes encoding phosphonate degradation enzymes in stromatolite-associated bacterial genomes. Gene abundance is indicated per the scale provided. Genomes that do not carry any of the genes investigated here were not shown. Phylogeny was inferred using JolyTree [24] with a sketch size of 5 000. Conserved genomes are indicated with an asterisk within highlighted taxonomic groups. Genome abbreviations indicate site, formation proximity to ocean and collection; CSU: Schoenmakerskop Inflow Jan, SU: Schoenmakerskop Inflow April, CRU: Cape Recife Inflow Jan, RU: Cape Recife Inflow April, SM1: Schoenmakerskop Middle 1, SM2:Schoenmakerskop Middle 2, RM1: Cape Recife Middle 1, RM2: Cape Recife Middle 1, SL: Schoenmakerskop Lower, RL: Cape Recife Lower.

Table S1. Summary of genomes binned per collected stromatolite sample

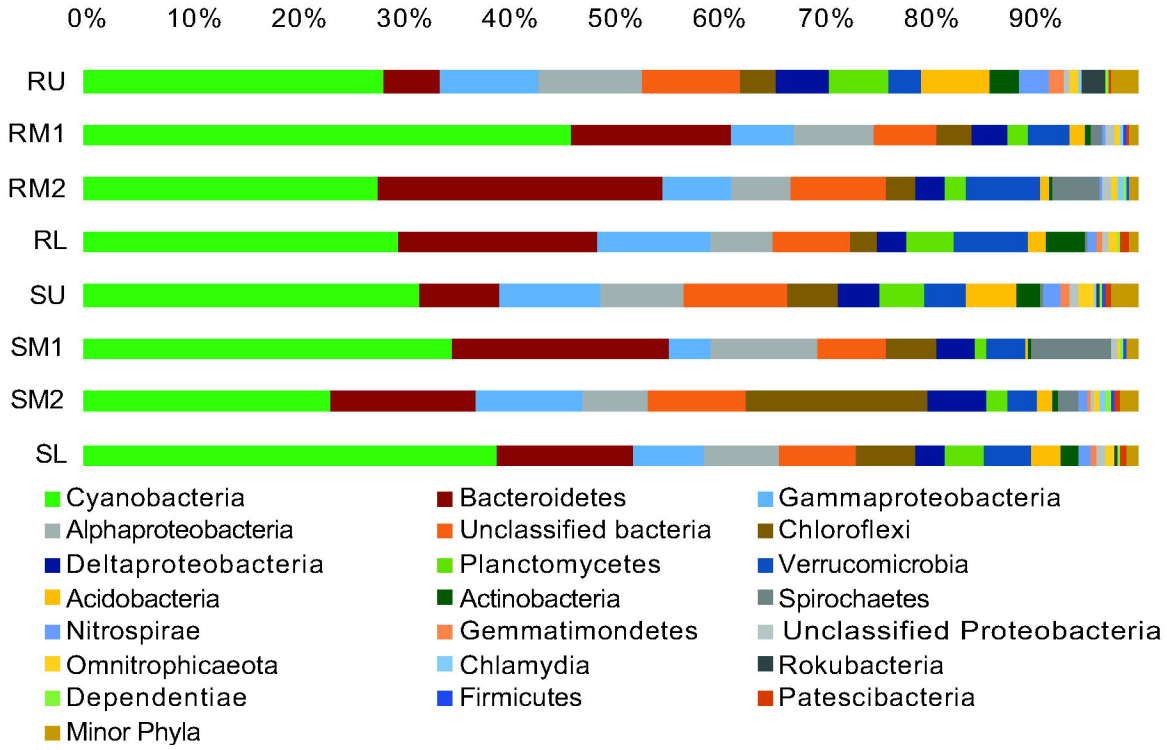
Table S2. Summary of characteristics of genomes binned from shotgun metagenomic data from sampled upper, middle and lower stromatolite formations from Cape Recife and Schoenmakerskop.

Table S3. Conserved bacterial species defined by shared ANI > 97% in stromatolite formations from Cape Recife and Schoenmakerskop.

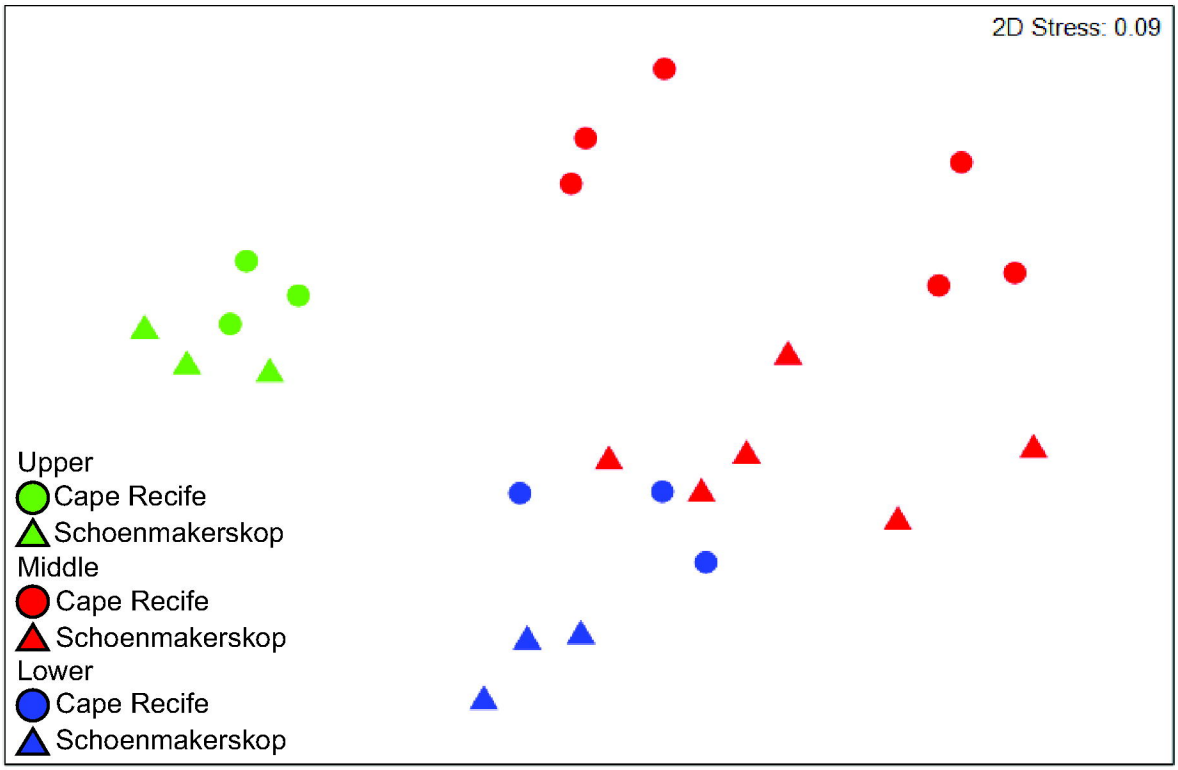
Table S4. Taxonomic classification of genomes from Shark Bay stromatolites

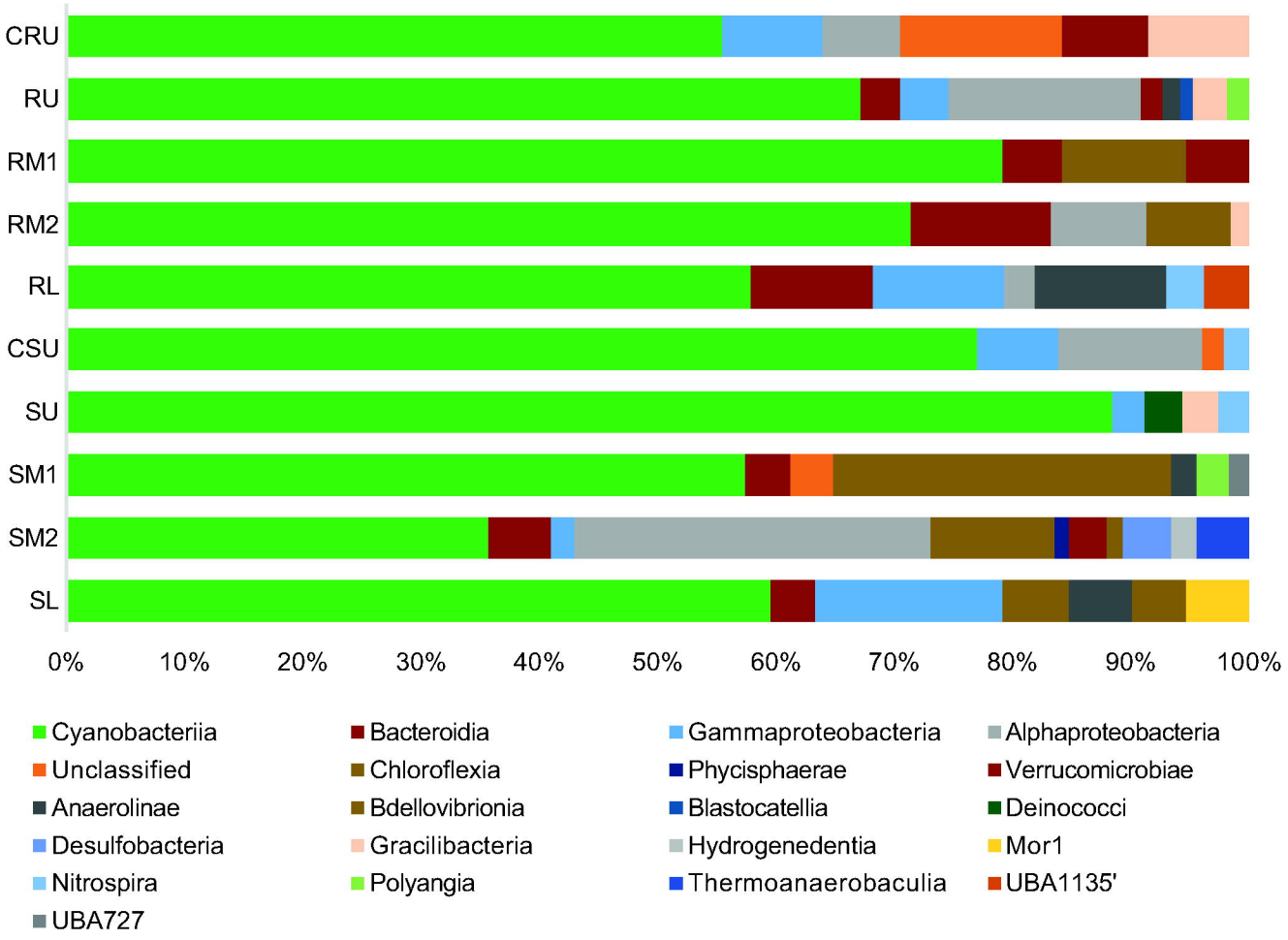

A.
SU
CSU
SM2
SL
SM1
0
5 m
B
RU
CRU
RM2
RM1
RL
0
5 m

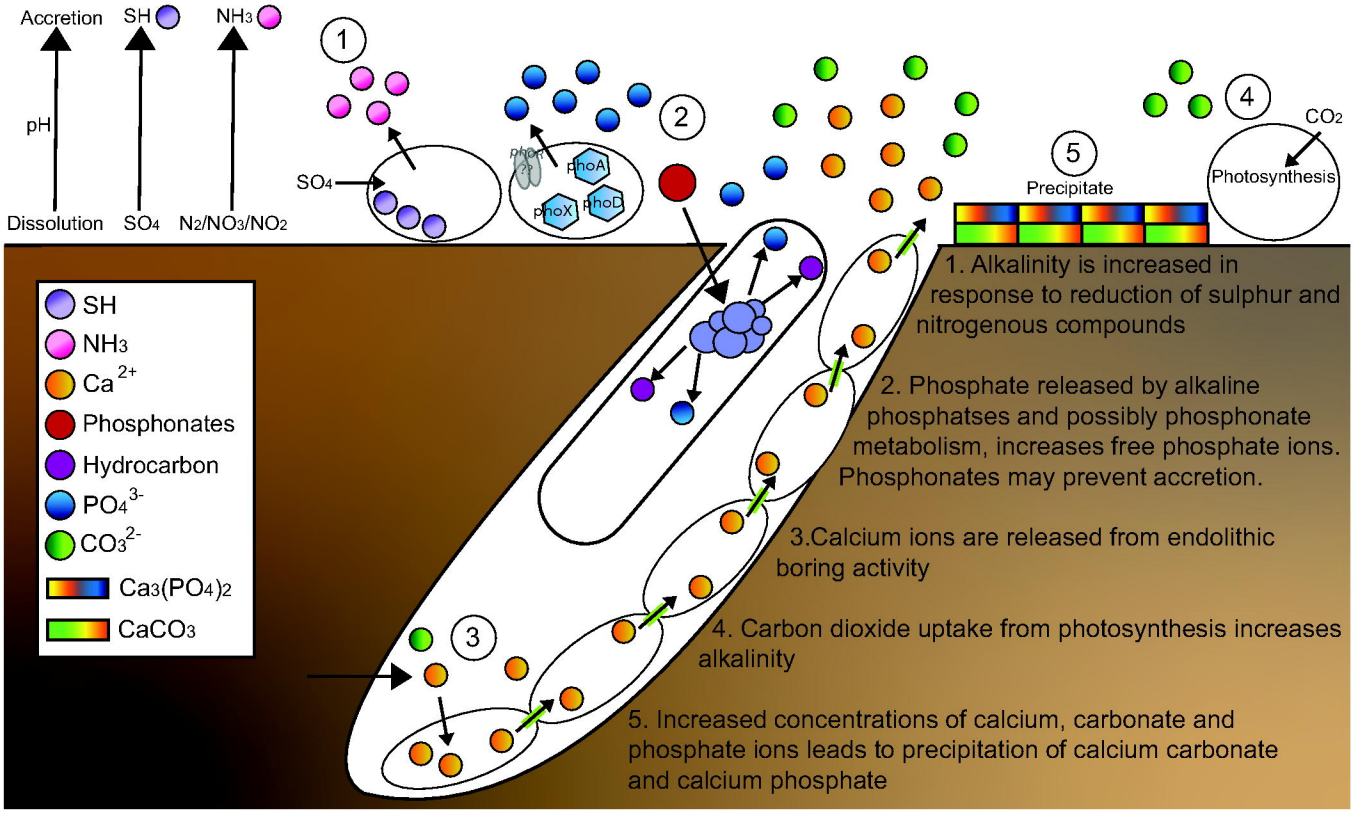
A.


0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
RU
RM1
RM2
RL
SU
SM1
SM2
SL
Cyanobacteria
Bacteroidetes
Gammaproteobacteria
Alphaproteobacteria
Unclassified bacteria
Chloroflexi
Deltaproteobacteria
Planctomycetes
Verrucomicrobia
Acidobacteria
Actinobacteria
Spirochaetes
Nitrospirae
Gemmatimondetes
Unclassified Proteobacteria
Omnitrophicaeota
Chlamydia
Rokubacteria
Dependentiae
Firmicutes
Patescibacteria
Minor Phyla

B.


2D Stress: 0.09
Upper
Cape Recife
Schoenmakerskop
Middle
Cape Recife
Schoenmakerskop
Lower
Cape Recife
Schoenmakerskop


CRU
RU
RM1
RM2
RL
CSU
SU
SM1
SM2
SL
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Cyanobacteriia
Bacteroidia
Gammaproteobacteria
Alphaproteobacteria
Unclassified
Chloroflexia
Phycisphaerae
Verrucomicrobiae
Anaerolinae
Bdellovibrionia
Blastocatellia
Deinococci
Desulfobacteria
Gracilibacteria
Hydrogenedentia
Mor1
Nitrospira
Polyangia
Thermoanaerobaculia
UBA1135'
UBA727


Accretion
SH
NH3
pH
Dissolution
SO4
N2/NO3/NO2
1
SO4
phoR
phoA
phoX
phoD
2
5
Precipitate
4
CO2
Photosynthesis
SH
NH3
Ca 2+
Phosphonates
Hydrocarbon
PO4 3-
CO3 2-
Ca3(PO4)2
CaCO3
3
1. Alkalinity is increased in response to reduction of sulphur and nitrogenous compounds
2. Phosphate released by alkaline phosphatses and possibly phosphonate metabolism, increases free phosphate ions. Phosphonates may prevent accretion.
3.Calcium ions are released from endolithic boring activity
4. Carbon dioxide uptake from photosynthesis increases alkalinity
5. Increased concentrations of calcium, carbonate and phosphate ions leads to precipitation of calcium carbonate and calcium phosphate