# Conserved bacterial genomes from two geographically distinct peritidal stromatolite formations shed light on potential functional guilds

Running Title: Functional bacterial guilds in stromatolites

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

## ORIGINALITY-SIGNIFICANCE

Peritidal stromatolites are unique among stromatolite formations as they grow at the dynamic interface of calcium carbonate-rich groundwater and coastal marine waters. The peritidal space forms a relatively unstable environment and the factors that influence the growth of these peritidal structures is not well understood. To our knowledge, this is the first comparative study that assesses species conservation within the microbial communities of two geographically distinct peritidal stromatolite formations. We assessed the potential functional roles of these communities using genomic bins clustered from metagenomic sequencing data. We identified several conserved bacterial species across the two sites and hypothesize that their genetic

functional potential may be important in the formation of pertidal stromatolites. We contrasted these findings against a well-studied site in Shark Bay, Australia and show that, unlike these hypersaline formations, archaea do not play a major role in peritidal stromatolite formation. Furthermore, bacterial nitrogen and phosphate metabolisms of conserved species may be driving factors behind lithification in peritidal stromatolites.

## SUMMARY

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. Modern stromatolites are relatively rare but may provide clues about the function and evolution of their ancient counterparts. In this study, we focus on peritidal stromatolites occurring at Cape Recife and Schoenmakerskop on the southeastern South African coastline. Using assembled shotgun metagenomic data we obtained 183 genomic bins, of which the most dominant taxa were from the Cyanobacteriia class (Cyanobacteria phylum), with lower but notable abundances of bacteria classified as Alphaproteobacteria, Gammaproteobacteria and Bacteroidia. We identified functional gene sets in bacterial species conserved across two geographically distinct stromatolite formations, which may promote carbonate precipitation through the reduction of nitrogenous compounds and possible production of calcium ions. We propose that an abundance of extracellular alkaline phosphatases may lead to the formation 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 (Dupraz *et al.*, 2009). A more recent discovery of stromatolite-like structures in Greenland suggests that the structures may date as far back as 3.7 - 3.8 billion years (Nutman *et al.*, 2016), however, their biogenic origin remains under debate (Witze, 2016). 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 (Soo *et al.*, 2017). Ancient stromatolites could provide insight into how microorganisms shaped early eukaryotic evolution. Unfortunately ancient microbial mats are not sufficiently preserved for identification of these microbes and individual bacteria cannot be classified more specifically than phylum Cyanobacteria due to morphological conservatism (Awramik, 1992; Dupraz *et al.*, 2009). The study of extant stromatolite analogs may 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 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 (Dupraz *et al.*, 2009). This carbonate precipitate, along with sediment grains, can then become trapped within bacterial biofilms forming the characteristic lithified layers.

Alteration of the pH and subsequently, the solubility index (SI), may promote mineralization or dissolution of carbonate minerals through microbial cycling of redox sensitive compounds such as phosphate, nitrogen, sulfur and other nutrients within the biofilm. 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 (Dupraz *et al.*, 2009). In some stromatolite formations such as those of Shark Bay, 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) (Gallagher *et al.*, 2012; Casaburi *et al.*, 2016; Wong *et al.*, 2018).

The biogenicity of stromatolites has been studied extensively in the hypersaline and marine formations of Shark Bay, Australia and Exuma Cay, Bahamas, respectively (Khodadad and Foster, 2012; Babilonia *et al.*, 2018). The presence of Archaea has been noted in several microbial mat and stromatolite systems (Casaburi *et al.*, 2016; Balci *et al.*, 2018; Medina-Chávez *et al.*, 2019), particularly in the stromatolites of Shark Bay, where they are hypothesized to potentially fulfill the role of nitrifiers and hydrogenotrophic methanogens (Wong *et al.*, 2017). Although Cyanobacteria, Proteobacteria and Bacteroidetes appear to be abundant in both marine and hypersaline systems, Cyanobacteria are proposed to be particularly vital to these

93 formations through the combined effect of biofilm formation, carbon fixation, nitrogen
94 fixation and endolithic (boring) activity (Macintyre *et al.*, 2000; Khodadad and Foster,
95 2012; Casaburi *et al.*, 2016; Babilonia *et al.*, 2018).
96
97 Peritidal tufa stromatolite systems are found along the southeastern coastline of South
98 Africa (SA). They are geographically isolated, occurring at coastal dune seeps
99 separated by stretches of coastline (Smith *et al.*, 2018). In these systems, stromatolite
100 formations extend from freshwater to intertidal zones and are dominated by
101 Cyanobacteria, Bacteroidetes and Proteobacteria (Perissinotto *et al.*, 2014). The
102 stromatolites are impacted by fluctuating environmental pressures caused by periodic
103 inundation by seawater, which affects the nutrient concentrations, temperature and
104 chemistry of the system (Rishworth *et al.*, 2016). These formations are characterized by
105 their proximity to the ocean, where stromatolites in the upper formations receive
106 freshwater from the inflow seeps, middle formation stromatolites withstand a mix of
107 freshwater seepage and marine over-topping, and lower formations are in closest
108 contact with the ocean (Perissinotto *et al.*, 2014). The stromatolite formations at Cape
109 Recife and Schoenmakerskop are exposed to both fresh and marine water that has little
110 dissolved inorganic phosphate and decreasing levels of dissolved inorganic nitrogen
111 (Cape Recife: 82 - 9 µM, Schoenmakerskop: 424 - 14 µM) moving from freshwater to
112 marine influenced formations (Rishworth, Perissinotto, Bird, *et al.*, 2017). Microbial
113 communities within these levels therefore likely experience distinct environmental
114 pressures, including fluctuations in salinity and dissolved oxygen (Rishworth *et al.*,
115 2016). While carbon predominantly enters these systems through cyanobacterial carbon

116 fixation, it is unclear how other members of the stromatolite-associated bacterial
117 consortia influence mineral stratification resulting from the cycling of essential nutrients
118 such as nitrogen, phosphorus and sulfur. Since peritidal stromatolites exist in constant
119 nutritional and chemical flux with varying influence from the fresh and marine water
120 sources, they present an almost ideal *in situ* testing ground for investigating which
121 microbes are consistently present despite fluctuations in their environment. Identification
122 of conserved bacterial species across both time and space and across varied
123 environmental pressures would suggest that these bacteria are not only robust but likely
124 play important roles within the peritidal stromatolite consortia.
125
126 Previous studies of stromatolites collected from Lake Clifton, Pavilion Lake, Clinton
127 Creek, Highborne Cay and Pozas Azules have investigated the overall bacterial
128 composition using 16S rRNA gene analysis and/or overall potential gene function of
129 unbinned metagenomic datasets per bacterial taxa using tools such as MG-RAST
130 (Keegan *et al.*, 2016) to gain insight into the potential roles of the different bacterial
131 groups (Mobberley *et al.*, 2015; Centeno *et al.*, 2016; Gleeson *et al.*, 2016; Ruvindy *et*
132 *al.*, 2016; Warden *et al.*, 2016; White *et al.*, 2016). However, the presence of all genes
133 required for a complete functional pathway within a collection of bacteria does not
134 necessarily mean that the cycle can take place since they may not be present within a
135 single organism. Therefore, in order to assess the potential for functional roles,
136 individual bacterial genomes must be investigated. To date, only two studies have
137 successfully binned individual bacterial genomes from culture-independent,
138 metagenomic data originating from stromatolites from Shark Bay (Australia) (Wong *et*

*al.*, 2018) and Socompa Lake (Argentina) (Kurth *et al.*, 2017). The latter study obtained four high-quality bins (in accordance with MIMAG standards defined by (Bowers *et al.*, 2017)) and analyzed three which were believed to have carried genes from several closely-related genomes. Extracted 16S rRNA sequences were in conflict with whole-genome taxonomic classifications (Kurth *et al.*, 2017). The Shark Bay study obtained a total of 550 binned genomes, of which 87 (15.8%) were of medium to high quality (Bowers *et al.*, 2017; Wong *et al.*, 2018) and the data provided information on a number of potentially important processes that may contribute to the formation and maintenance of the hypersaline stromatolites. However, the study did not identify key microbial architects within these biogenic structures.

Using a metagenomic approach, we have sought to gain insight into the foundational bacteria responsible for metabolic processes that potentially result in formation of pertidal South African stromatolites. We obtained and annotated 183 putative bacterial metagenome-assembled genomes (MAGs), (of which 112 (61%) were of medium to high quality) from samples of two geographically isolated sites near Port Elizabeth, South Africa. We identified several temporally and spatially conserved bacterial species and functional gene sets, that are likely central in establishing and maintaining peritidal stromatolite microbial communities.

## RESULTS AND DISCUSSION

Two geographically isolated peritidal stromatolite sites, Cape Recife and Schoenmakerskop, which are 2.82 km apart, were chosen for this study. These two

sites have been extensively characterized with respect to their physical structure, nutrient and chemical environment (Perissinotto *et al.*, 2014; Rishworth *et al.*, 2016, 2019; Rishworth, Perissinotto, Bornman, *et al.*, 2017; Dodd *et al.*, 2018). The sites experience regular shifts in salinity due to tidal overtopping and groundwater seepage (Rishworth *et al.*, 2019). Comparison to a site with groundwater seepage but no stromatolite growth, has shown that the growth of peritidal stromatolites is promoted within this region by decreased levels of wave action, higher water alkalinity and decreased calcite and aragonite saturation (Dodd *et al.*, 2018). Furthermore, stromatolite growth is inhibited by increased levels of salinity, as lithified structures are not observed in both the marine waters and in deeper portions of formation pools, with higher salinity levels (Dodd *et al.*, 2018).

Stromatolite formations at both sites begin at a freshwater inflow and end before the subtidal zone (Fig. 1) and are exposed to different levels of tidal disturbance. Samples for this study 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 in April 2018 from middle and lower formations for extended comparison across the two sites (Fig. 1). For a detailed perspective of depth and the differentiation of the sampled zones a 3-dimensional rendering of the sample sites was constructed by a 3rd party (Caelum Technologies) using a combination of drone-based photogrammetry and geographic mapping using differential GPS (Fig. 1). Stromatolite formations were classified according to their tidal proximity as defined in previous studies (Perissinotto *et al.*, 2014; Rishworth *et al.*,

2016; Rishworth, Perissinotto, Bornman, *et al.*, 2017; Dodd *et al.*, 2018). Upper formations occur in the supratidal zone where they are constantly exposed to fresh water flowing from dune seeps. These formations will only be exposed to seawater during spring tides or extreme storm surges. The water from upper formations feeds into large pools, which in turn feed into the lower portion of the system. In the upper-middle intertidal zone, the formations form a slope into large pools where they will only receive seawater during peak high tide. Lower formations occur in areas where saline or brackish conditions are predominant. In Schoenmakerskop the lower zone is located in a semi stagnant pool which is frequently overtopped, whilst in the lower zone of Cape Recife, formations create a low flowing slope which ends at the subtidal zone. All sampling was conducted at low tide. Sample abbreviations and MAG prefixes used throughout this study correspond to the site and region from which they were sampled e.g. “SU” identifies the sample as originating from Schoenmakerskop, Upper formation in April (Table S1). Samples prefixed with a “C” identify samples collected in January e.g. “CSU” identifies the sample as originating from Schoenmakerskop, Upper formation in January (Table S1).

**Phylogenetic distribution of microbial communities**

We assessed the diversity and structure of the bacterial communities in triplicate samples taken from upper, middle and lower formations 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), in agreement with a

208 previous study at Schoenmakerskop (Perissinotto *et al.*, 2014). Bacterial communities at
209 each sampling site were statistically different from one another (Fig. 2B; ANOSIM: R =
210 0.976, p = 0.01 permutations = 999), whilst pairwise Kendall rank correlation tests
211 showed that replicates were not statistically different from one another (pairwise p-
212 values < $2.2 \times 10^{-16}$).
213
214 **Trends observed in metagenomes of stromatolite samples from Cape Recife and**
215 **Schoenmakerskop**
216 To characterize the metabolic potential within stromatolites, we generated shotgun
217 metagenomic libraries from 8 samples representative of the upper, middle and lower
218 formations at both sites in April 2018, as well as an additional two samples from the
219 upper formations of each site, collected 4 months previously in January 2018 (Table
220 S1). Following assembly of raw sequence reads, gene coverages were normalized to
221 the length-weighted average coverage per sample, revealing a high abundance of
222 genes encoding phosphate transport (*pstSCAB*), phosphate uptake regulation
223 (*phoURBP*) and alkaline phosphatases (*phoADX*) were observed across the board (Fig.
224 3). Additionally, genes involved in phosphonate metabolism (*phnCDEFGHIJKLM)* were
225 abundant in three of the four middle formations but were absent in upper formations.
226 The concentration of soluble phosphorus, although relatively low, is highest in water
227 surrounding the middle formations, as both fresh seep water and ocean overtopping
228 contribute to the total phosphate (Dodd *et al.*, 2018). It is thus unsurprising then, that the
229 greatest abundance of phosphate-metabolism genes are found in the middle
230 formations. The high abundance of genes encoding phosphate-metabolizing enzymes,

231 may also be indicative of how stromatolite communities cope with low dissolved
232 inorganic phosphorus in their environment. Genes encoding alkaline phosphatase
233 *phoX*, were the most abundant among the phosphatases observed in these stromatolite
234 samples, and represent a calcium-dependent enzyme that can function at low substrate
235 concentrations on a broad range of C-O-P substrates (Zaheer *et al.*, 2009). The
236 phosphonate transporter genes (*phnCDE*) are more prevalent than the genes encoding
237 the C-P lyase (*phnGHIJLM*) required for phosphonate degradations (Fig. 3). These
238 transporters have also been implicated in the transport of inorganic phosphate, in
239 addition to phosphonates (Stasi *et al.*, 2019), and the discrepancy between transporters
240 and metabolic genes may suggest an additional role of phosphate uptake in these
241 systems. Overall gene abundances indicated negligible presence of canonical
242 dissimilatory sulfate reduction/oxidation via *aprAB* and *dsrAB* encoded enzymes.
243 Similarly, there were low abundances of genes associated with sulfonate metabolism.
244 There was an abundance of genes associated with assimilatory sulfate reduction, but
245 the low abundance or absence of *cysC*, a gene encoding a key enzyme in this pathway,
246 indicated that this pathway may not be complete. Genes associated with assimilatory
247 nitrate reduction (*narB*, *nirA*) were the most prevalent markers of nitrogen metabolism.
248 All sites also contained a number of genes associated with dissimilatory nitrate
249 reduction where cytoplasmic NADH-dependent nitrate reductase *nirBD* appeared to be
250 favored over periplasmic cytochrome c nitrate reductase *nrfAH*. Interestingly nitrogen
251 fixation genes (*nifDHK*) appear to be more abundant in the middle stromatolite
252 formations of Schoenmakerskop (Fig. 3).
253

## Binning and phylogenetic classification of putative genomes

The 10 assembled metagenomes were binned using Autometa (Miller *et al.*, 2019), resulting in a total of 183 bacterial genome bins (Table S1). 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. 4A). This distribution appears to be approximately congruent with abundances observed in the 16S rRNA gene amplicon analyses (Fig. 2).

## 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 in two or more of the sampled regions. We identified 16 conserved taxa across the 10 sampled regions, with several species identified in 3 - 5 samples, across Schoenmakerskop and Cape Recife, however no single species was common to all sampled sites (Table S2, Fig. 4A). Conserved species were commonly the most abundant taxa present in each of the samples, accounting for approximately 30 - 80% of the species abundance (Table S2, Fig. 4A). Cyanobacterial species within the *Acaryochloris* and *Hydrococcus* genera and Absconditabacterales family were conserved across upper formations (Table 1, Fig. 4B, Table S3). Seven other species were conserved across the middle formations, including species classified within the *Rivularia* genus and Phormidesmiaceae and Spirulinaceae families (Table 1 and S3, Fig. 4B). Two distinct species within family Elainellaceaee were also conserved: species A was detected only in the upper pools of

277 Cape Recife, whilst species B was conserved across the middle formations of both sites
278 (Table 1 and S3, Fig. 4B).
279
280 Also, of interest was the presence of conserved bacterial species (order
281 Absconditabacterales), which are classified under the Patescibacteria phylum.
282 Patescibacteria are unusually small bacteria found in groundwater that produce large
283 surface proteins hypothesized to help them attach to, and exploit the ability of other
284 microorganisms performing nitrogen, sulfur and iron cycling (Herrmann *et al.*, 2019).
285 The presence of these conserved bacteria suggests that the inflow water seeps may
286 originate from groundwater.
287
288 Conserved *Rivularia*, Elainellaceae, Phormidesmiaceae and Chloroflexaceae species
289 were identified in the lower formation of Schoenmakerskop, but none of the genomes
290 identified in the lower formation of Cape Recife had greater than 97% shared ANI with
291 any other genome bin. The distribution of conserved taxa, wherein the upper and middle
292 formations appear to harbor distinct conserved taxa suggests that the differing nutrient
293 and physical characteristics between upper, middle and lower regions of the
294 stromatolite formation elicit specialization of the conserved bacterial community. The
295 lack of conserved bacterial species across the lower formation of Cape Recife may be
296 due to the choice of sampling site: The lower formation sample taken from Cape Recife
297 is in closer contact with the ocean than the lower formation sample taken from
298 Schoenmakerskop (Fig. 1 C - D). The proximity of the Schoenmakerskop lower
299 formation sample to the ocean was initially thought to be sufficient to delimit it as a

lower formation, but following this study it may be reclassified as a middle formation, as the bacterial community composition corresponds more closely with the other four middle formation samples (Fig. 4B).

**Metabolic potential of binned genomes**

Oxygenic photosynthesis by cyanobacteria results in rapid fixation of carbon dioxide and an increase in alkalinity (Pace *et al.*, 2018). Carbonate ions bind cations such as calcium and are precipitated under alkaline conditions, promoting the growth of stromatolite structures (Dupraz *et al.*, 2009). Given their numerical dominance in Cape Recife and Schoenmakerskop stromatolites, and their predicted role in other stromatolites (Dupraz *et al.*, 2009), cyanobacteria likely perform carbon sequestration. However, 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 remains unknown. We inspected PROKKA and KEGG annotations within all stromatolite-associated bacterial genome bins to identify potential metabolic pathways that may promote mineral deposition and accretion (Dupraz *et al.*, 2009). The results presented here are summarized in Table 1.

*Sulfur metabolism.* Reduction of sulfate has previously been shown to promote the precipitation of carbonates in the form of micritic crusts in Bahamian and Australian stromatolites (Reid *et al.*, 2000; Wong *et al.*, 2018) 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 (Bontognali *et al.*, 2012; Allen, 2016).

Amongst the Cape Recife and Schoenmakerskop stromatolite-associated bacteria, the potential capacity for sulfate reduction was confined to only a few genomes (Fig. 5 and Fig. S1). The complete set of genes required for assimilatory sulfate reduction (*sat*/*met3*, *cysC*, *cysH* and *sir* genes) (Santos *et al.*, 2015) were recovered in four genomes (Fig. S1), three of which were conserved *Acaryochloris* species (Fig. 5). The abundance of genes associated with assimilatory sulfate reduction in the conserved *Acaryochloris* genome bins accounted for 22 - 100% of gene counts observed in their respective metagenomes (Fig.5). This was calculated as abundance of geneX in binX, as a percentage of geneX abundance in the sample from which the bin was derived (Fig. 5). The greatest abundance genes for uptake and desulfonation of alkanesulfonates (*ssuABCDE*) (Aguilar-Barajas *et al.*, 2011; Ellis, 2011) were detected exclusively in conserved *Hydrococcus* species (Fig. 5), which account for 23 - 100% of gene abundance in the respective metagenomes. All genes required for a complete pathway found exclusively in bin RU2_2, but both SU_1_0 and CRU1_1 appear to be missing the *ssuE* gene. The *ssuE* gene is not required for growth using aliphatic sulfonate or methionine substrates, but is required for arylsulfonate metabolism (Kahnert *et al.*, 2000). This suggests that these *Hydrococcus* strains are all capable of some form of sulfonate metabolism, but not all can metabolize arylsulfonates (Bin CSU_1_8 is only 37% complete, and therefore of low quality and may be missing gene due to incompleteness). In both cases, these trends are in agreement with the patterns observed in the metabolic potential of the overall metagenome (Fig.3). *Hydrococcus* and *Acaryochloris* species are dominant in upper formations (Fig. 4 and Table S3) and the potential for cumulative removal of hydrogen by sulfate reduction by these species

346 may aid in the creation of an alkaline environment within the system. Seep waters
347 feeding both Cape Recife and Schoenmakerskop have relatively high levels of sulfate
348 (Dodd *et al.*, 2018), and it is possible that the *Hydrococcus* and *Acaryochloris* species
349 utilize this nutrient in the upper formations (closest to the seeps) as a selective
350 advantage, simultaneously increasing the alkalinity of the surrounding environment
351 through hydrogen assimilation. There was no evidence for the potential for dissimilatory
352 sulfate reduction in conserved genome bins (Fig. 5). Similarly, amongst all 183 genome
353 bins, only genomes representative of a *Thioiploca* sp. from the Schoenmakerskop lower
354 formation and a *Desulfobacula* sp. from the Schoenmakerskop middle formation
355 included all genes required for dissimilatory sulfate reduction (Fig. S1). The lack of
356 genes associated with conserved or dominant bacterial taxa in the middle and lower
357 pools suggest that either sulfate is not required by the consortia that inhabit the
358 middle/lower formations and is required only by bacteria exposed to ground water, or
359 that insufficient amounts of the sulfate continues into the lower pools where the middle
360 and lower formations are found. This suggests that calcite formation in these peritidal
361 stromatolites may be influenced by processes other than dissimilatory sulfate reduction,
362 in contrast to stromatolite formations in the Cayo Coco lagoonal network, Highborne
363 Cay and Eleuthera Island in the Bahamas (Visscher *et al.*, 2000; Dupraz *et al.*, 2004;
364 Pace *et al.*, 2018).
365
366 *Nitrogen metabolism.* The reduction of nitrogen, nitrates and nitrites can lead to calcite
367 precipitation (Rodriguez-Navarro *et al.*, 2003; Wei *et al.*, 2015; Konopacka-Łyskawa *et*
368 *al.*, 2017; Wong *et al.*, 2018; Lee and Park, 2019), and the released $NH_3$ can react with

$CO_2$ and $H_2O$, to form $2NH^{4+}$ + $CO_3^{2-}$ (Konopacka-Łyskawa *et al.*, 2017). Nitrogen metabolism is proposed to have emerged at approximately the same time as sulfur metabolism, dated to 3.4 billion years ago (Stüeken, 2016), as both processes share similar redox states (Thomazo *et al.*, 2011), and ammonium availability during this time may have sustained developing microbial life (Yang *et al.*, 2019). Therefore, bacteria that can fix nitrogen or produce ammonia from nitrates/nitrites could potentially promote the growth of stromatolites (Visscher and Stolz, 2005) and may have added to the formation of ancient analogs. Similarly, denitrification will likely lead to mineral dissolution and retardation of stromatolite growth (Visscher and Stolz, 2005). We found that several conserved bacteria carried the genes associated with ferredoxin-dependent assimilatory nitrate reduction (*nirA-narB* genes) (Moreno-Vivián *et al.*, 1999), nitrogen fixation (*nifDHK* genes) and dissimilatory nitrite reduction (*nirBD* genes) (Griffith, 2016), all of which result in the production of ammonia (Fig. 5).

The potential for assimilatory nitrate reduction was detected in several genomes, primarily from the middle formations, with particularly high gene abundances of *nirA-narB* genes conserved *Rivularia* sp. (Fig. 5), which are the dominant species in the middle formations (Fig. 4). 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 (Fig. S1). It appears that the gene abundances associated with assimilatory nitrate reduction in *Rivularia* sp. account for the majority (29 - 75%, Fig. 5) of *nirA-narB* genes observed in their respective metagenomes (Fig. 3). Among conserved bacteria, the potential for

392 dissimilatory nitrate reduction is relatively low, with *Hydrococcus* sp. carrying the
393 greatest abundance of *nirBD* genes (Fig. 5) which encode cytoplasmic nitrite
394 reductases. Non-conserved *Thioploca, Cyclobacteriaceae* and *Limnothrix* species
395 appear to account for the remainder of the *nirBD* genes observed in their respective
396 metagenomes (Fig. S1). Very few genomes included the *nrfAH* genes which encode
397 membrane-bound nitrite reductases.
398
399 Nighttime nitrogen fixation has previously been shown to be a driver of carbonate
400 precipitation in stromatolites (Dupraz *et al.*, 2009) and *nif* genes were identified in
401 several conserved bacteria associated with both upper and middle formations:
402 *Hydrococcus* species across both upper formations, *Microcoleus* in the
403 Schoenmakerskop upper formations and Phormidiaceae, Spirulinaceae and
404 Chloroflexaceae species the middle formations (Fig. 5). In addition to the conserved
405 bacteria carrying *nif* genes, non-conserved *Blastochloris* species appeared to account
406 for the majority (10% - 75%) of identified *nifDHK* genes (Fig. S1). The genetic capacity
407 for nitrogen fixation being retained in Schoenmakerskop and Cape Recife was
408 unexpected, given the high concentrations of dissolved inorganic nitrogen (DIN) ranging
409 from 95 - 450 mM at the two sites (Rishworth *et al.*, 2016).
410
411 **Phosphate and phosphonate metabolism**
412 Phosphatic structures that closely resemble fossilized phosphatic stromatolites have
413 recently been observed within Cape Recife stromatolites (Buttner *et al.*, 2019,
414 unpublished). Hydroxyapatite ($Ca_5(PO_4)_3(OH)$) is more easily precipitated than calcite

($CaCO_3$) and the release of $PO_4^{3-}$ into the biofilm by alkaline phosphatase activity could increase the potential for apatite mineralization, the rate of which may be increased in the presence of an alkaline environment (Gallagher *et al.*, 2013). Copies of alkaline phosphatase genes *phoD* and *phoX* are variable within bacterial genomes but the majority of bacteria carry only 1 copy of *phoD* and 1 copy of *phoX* (Ragot *et al.*, 2015, 2017). Conserved *Acaryochloris* and *Rivularia* species were particularly notable as they encoded between 2–4 copies of *phoX* (Fig. 5), and carried all genes required for phosphate transport (*pstSCAB*) (Fig. 5). Since the Cape Recife and Schoenmakerskop stromatolites experience limited inorganic phosphate availability (Rishworth *et al.*, 2016, 2018; Rishworth, Perissinotto, Bird, *et al.*, 2017; Dodd *et al.*, 2018), associated bacteria may generate bioavailable phosphate from trapped sediments in biofilms resulting in increased concentrations, similar to cyanobacteria living in phosphate-poor rivers (Wood *et al.*, 2015). An increase in both $PO_4^{3-}$ and $Ca^{2+}$ can result in rapid precipitation of apatite in marine phosphorites, freshwater lakes and in some soil bacteria (Danen-Louwerse *et al.*, 1995; Guang-Can *et al.*, 2008; Cosmidis *et al.*, 2015) and this could account for the phosphatic deposits observed within the SA stromatolites (Buttner *et al.*, 2019, unpublished).

Finally, conserved *Rivularia* species in the middle and lower formations and other non-conserved bacterial species harbored the 11 genes required for the transport (*phnCDE*) and lysis (*phnFGHIJKLM*) of phosphonate compounds (Fig. 5) (Metcalf and Wanner, 1993). Phosphonates are characterized by the presence of a carbon-phosphorus bond and are biosynthesized by many organisms (White and Metcalf, 2007). C-P lyase

(*phnGHIJLKM*) breaks the C-P bond within phosphonate substrates resulting in a hydrocarbon and inorganic phosphate (White and Metcalf, 2007). The inorganic phosphate may then be used within the bacterial cell or released aiding in accretion through increased ion concentration (Rott *et al.*, 2018). Conversely, phosphonates can prevent precipitation of calcite by binding to crystal growth sites (Kan *et al.*, 2005) and the degradation of these compounds within stromatolite formations may prevent chemical inhibition of stromatolite growth. Given the low availability of phosphate in these systems, phosphonates may also provide an auxiliary source of bioavailable phosphate. The potential for phosphonate degradation was also observed in 8 genomes from the Shark Bay stromatolites (Wong *et al.*, 2018). The conservation of phosphonate metabolism across stromatolites would indicate that these bacterial processes may be important within the generalized stromatolite system.

**Archaeal genomes in Cape Recife and Schoenmakerskop stromatolites**

Genome-resolved studies of hypersaline stromatolites in Shark Bay revealed that a large proportion of the microbial community consisted of archaeal species (Wong *et al.*, 2017, 2018). Scrutiny of the contigs from Cape Recife and Schoenmakerskop showed that none of the datasets comprised more than 1.5% archaeal genes. Assessment of contigs clustered into the Archaeal kingdom bins from the upper and middle formations showed little evidence for the presence of Archaea. However, two low-quality genomes were obtained from the two lower formations SL_Arch_1 and RL_Arch_2, which were classified as Woesearchaeales (order) and *Nitrosoarchaeum* (genus) respectively (Table S3). Alignment of the 16S rRNA gene obtained from SL_Arch_1 (Table S3)

against the nr database showed that it shared the greatest sequence homology with an uncultured archaeon clone GWA2 (87% identity). Similarly, analysis of the 16S rRNA gene sequence from RL_Arch_2 (Table S3) indicated that it shared greatest sequence homology with an uncultured archaeon clone 027 (99.79% identity). The coverage of these genomes is among the lowest in each of the lower formation samples (Table S3) and would suggest a low numerical abundance of these archaea. It is compelling that the only Archaea detected in this study are derived from the formations most influenced by saline waters, considering that the hypersaline stromatolites of Shark Bay comprise several archaeal species (Wong *et al.*, 2017, 2018), and it would be intriguing to identify whether salinity has an effect on the Archaeal composition of stromatolite-associated microbiota. However, as there are only two, low quality, non-conserved, relatively low-abundance genomes derived from Archaea, it is likely that this kingdom is less important in the stromatolites found at Cape Recife and Schoenmakerskop.

**Key bacteria for stromatolite formation at Cape Recife and Schoenmakerskop**

Several bacterial groups may be key players in the formation of stromatolites in Cape Recife and Schoenmakerskop, with conserved bacterial species potentially acting as major contributors (Table 1). There were distinct functional abilities between the bacterial communities associated between the upper and middle formations. The consortia in the upper formations were dominated by conserved *Hydrococcus* and *Acaryochloris* species, which appeared to be capable of sulfate and sulfonate metabolism, whilst the middle formations were dominated primarily by conserved *Rivularia* species, which carry the genetic capacity to metabolize various forms of

484 phosphate substrates and potentially reduce nitrate. The lack of all genes required for
485 phosphonate metabolism in *Rivularia* species in Schoenmakerskop Middle 2 (SM2) was
486 unexpected but may be explained by the lower abundance of *Rivularia* in this sample,
487 resulting in lower quality of genome (Table S3). Similarly, *Hydrococcus* bin CSU_1_8
488 was of low quality (Table S3) and as a result appeared to be missing 3 of the genes
489 required.
490
491 We propose the redox potential and solubility index in these stromatolites is potentially
492 influenced by conserved bacteria generating sulfide (*Acaryochloris* and *Hydrococcus*
493 species) and ammonia ions (*Rivularia* species) for the rapid precipitation of carbonate
494 compounds and subsequent growth of stromatolite structures. The presence of
495 *Microcoleus* species (Phormidiaceae family) and an unidentified genus of
496 Phormidiaceae in both upper and middle formations is notable, as previous studies
497 have found that lamellar precipitation of carbonate occurs on the filaments of cultured
498 Phormidiaceae bacteria isolated from freshwater tufa structures (Payandi-Rolland *et al.*,
499 2019). We further propose an abundance of alkaline phosphatases in several
500 conserved species, as well as the abundance of genes associated with phosphonate
501 degradation in *Rivularia* species, may result in increased local inorganic phosphate
502 concentrations. This free inorganic phosphate may be incorporated into the phosphatic
503 crusts observed in these stromatolites (Buttner *et al.*, 2019, unpublished).
504
505 The identification of conserved bacteria performing potentially important roles within
506 stromatolites of both Cape Recife and Schoenmakerskop may provide insight into what

cyanobacterial species may have played a key role in the formation of ancient phosphatic stromatolites that formed in shallow marine and peritidal environments (Misi and Kyle, 1994; Drummond *et al.*, 2015; Buttner *et al.*, 2019, unpublished; Shiraishi *et al.*, 2019). However, transcriptomics, nutrient uptake and additional hydrochemistry would be required to determine if the proposed role of the conserved bacteria is valid. Similarly, future studies will be needed to determine whether physical (e.g. flow rate) or chemical factors (e.g. nutrient availability) result in our observed difference in functional potential between the different formations.

## EXPERIMENTAL PROCEDURES

**Sample collection and DNA isolation.** Samples 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) at low tide, from the water surface level. Samples approximately 1cm deep were collected for 16S rRNA analysis in July 2019, from upper, middle and lower stromatolite formations (Fig. 1) in triplicate, approximately 1 cm apart. Sample cores for metagenomic shotgun sequencing were collected in April 2018, approximately 1cm deep, from upper, middle and lower stromatolite formations. Additional samples were collected in January 2018 from only the upper formations for metagenomic sequencing. All samples were stored in RNAlater and flash-frozen until delivery to the lab at which point, they were stored 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') (Matcher *et al.*, 2011) from triplicate samples 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 (Schloss *et al.*, 2009). 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 (Rognes *et al.*, 2016) was used to remove chimeras. Reads that were 97% similar were combined into operational taxonomic units (OTUs) using the Opticlust method (Westcott and Schloss, 2017). OTU abundance values were converted to relative values and the dataset was then transformed by square root and statistically analyzed using the Primer-e (V7) software package (Gorley and Clarke, 2015). Reads were not classified against the Greengenes database, as the creators of Mothur warn against this practice, citing poor alignment quality in the variable regions.

**Metagenomic binning.** As 16S rRNA sequence analysis indicated that there was no statistical difference between triplicate samples, and that sample regions are therefore homogenous, a single sample from sites CRU, RU, RM1, RM2, RL, CSU, SU, SM1,

552 SM2 and SL (Fig.1) collected in January and April 2018 was used to prepare shotgun
553 DNA libraries that were sequenced using IonTorrent Ion P1.1.17 Chip technology
554 (Central Analytical Facilities, Stellenbosch University, South Africa). Adapters were
555 trimmed and trailing bases (15 nts) were iteratively removed if the average quality score
556 for the last 30 nts was lower than 16, which resulted in approximately 30–45 million
557 reads per sample. Resultant metagenomic datasets were assembled into contiguous
558 sequences (contigs) with SPAdes version 3.12.0 (Bankevich *et al.*, 2012) using the --
559 iontorrent and --only-assembler options with kmer values of 21,33,55,77,99,127. The --
560 iontorrent option enables a read error correction step performed by IonHammer,
561 correcting homopolymeric runs inherent to IonTorrent sequencing chemistry (Ershov *et*
562 *al.*, 2019).

563 **Quantification of metabolism-associated genes**. All raw data, count tables and
564 scripts used to process the data can be accessed here:
565 https://github.com/samche42/Conserved_Stromatolite_bacteria_manuscript.git

566 Contigs in each of the 10 samples were clustered into kingdom bins using Autometa
567 (Miller *et al.*, 2019). Genes on all contigs within these bins were identified using Prodigal
568 v.2.6.3 (Hyatt *et al.*, 2010). Genes were then annotated against the KEGG database
569 using kofamscan (Aramaki *et al.*, 2019) with output in mapper format. Coverage
570 corrected KO annotation counts were collected using kegg_parser.py found in the
571 GitHub repo listed above. Briefly, a dictionary of collection sites was made with all KO
572 numbers supplied in a KEGG annotation query list (i.e. functional group). Kofamscan
573 output per collection site was parsed and each time an entry matching that within the
574 query list was found, a count for that annotation was increased by the coverage of the

575 contig on which it was located, resulting in a coverage-corrected count table of
576 functional group KO annotations per collection site. In order to assess differences
577 between collection sites, these values were transformed relative to average coverage
578 per sample (weighted by contig length) using weighted_contig_coverage_calculator.py
579 (see GitHub repo) according to Eq.1, where *i* is contig length and *j* is contig coverage.
580 The abundances were then transformed using log2 (Fig. 3).

581 **Equation 1.** Contig length weighted coverage = $\sum( (i / \sum(i)) \times j )$

582 Heatmaps for visualization of the data were created in R using the tidyr, ggplot2 and
583 viridis packages. Scripts can be found in the GitHub repo.

584 **Clustering of metagenomic data in genome bins.** Contigs were then clustered in
585 putative genomic bins using Autometa
586 (https://bitbucket.org/jason_c_kwan/autometa/src/master/, Master branch, commit:
587 a344c28) (Miller *et al.*, 2019). Bins were manually curated and resulted in 183 genomic
588 bins. CheckM (Parks *et al.*, 2015) was used to assess bin purity and completion using
589 default settings (Table S3). Approximately 37 genomes were of high quality, 75 were of
590 medium quality and 71 were of low quality, in accordance with MIMAG standards
591 defined by (Bowers *et al.*, 2017) (Table S3).

592 **Identification of conserved taxa.** Conservation of bacterial taxa was calculated using
593 average nucleotide identities (ANIs) of all genomic bins, which were calculated in a
594 pairwise manner using FastANI (Jain *et al.*, 2018). All genomic pairs sharing > 97% ANI
595 were subset and considered conserved taxa. Percentage of mapped regions used to
596 calculate ANI is reported in Table S2. Percentage of mapped regions is calculated as

597 the relative number of bidirectional fragments mapping to the total number of query
598 fragments.

599 **Genome taxonomic classification.** Genomes were classified using the standalone
600 GTDB-Tk tool (version 0.3.2) using the classify workflow and Genome Taxonomy
601 Database version 89 (Parks *et al.*, 2019). The tool is unable to classify genomes less
602 than 10% complete, as indicated in Table S3. The GTDB-Tk tool uses the Genome
603 Taxonomy Database as a reference for classification, which is based on phylogeny
604 inferred from concatenated protein alignments. This approach enabled the removal of
605 polyphyletic groups and assignment of taxonomy from evolutionary divergence. The
606 resulting taxonomy incorporates substantial changes in comparison to the NCBI
607 taxonomy (Parks *et al.*, 2018). Equivalent NCBI taxonomic classifications have been
608 provided for clarity in Table S3.

609 **Genome taxonomic clustering.** Phylogeny of bacterial genomes was inferred using
610 JolyTree (Criscuolo, 2019) with a sketch size of 10 000 (Fig. 5 and Fig. S1). JolyTree
611 infers phylogeny through computation of dissimilarity of kmer sketches, which is then
612 transformed for the estimation of substitution events of the genomes' evolution
613 (Criscuolo, 2019).

614 **Genome annotation.** Manually curated putative genomic bins were annotated using
615 Prokka version 1.13 (Seemann, 2014), with GenBank compliance enabled. Protein-
616 coding amino-acid sequences from genomic bins were annotated against the KEGG
617 database using kofamscan (Aramaki *et al.*, 2019) with output in mapper format. KEGG
618 orthologs were counted and processed as performed quantification of metabolism-

associated genes using kegg_parser.py. Relative percentage of genes in individual genomes was calculated by dividing the contig-coverage corrected gene abundance per gene by the total contig-coverage corrected gene abundance for the sample from which it was binned (Fig. 5 and Fig. S1). E.g. The contig-coverage corrected gene abundance of *phoU* in genome bin CRU1_1 was 3.52, and the total contig-coverage corrected gene abundance of *phoU* in the CRU metagenome sample was 29.90253 (See dataset Calculating_rel_perc.xlsx in GitHub repo). Therefore, the relative abundance of *phoU* in genome bin CRU1_1 was: 3.52/29.90 X 100 = 11.79% of total sample gene abundance.

**Archaeal genome binning and identification.** Contigs are classified within kingdom bins during Autometa binning. All contigs classified within the Archaea kingdom were given putative taxonomic assignments based on single-copy markers and clustered into bins. No bins could be acquired from the upper and middle formations from either collection site due to few or no contigs being of archaeal origin. Two genomes were obtained from the lower formation of each site respectively (Table S3). Genome quality was assessed using CheckM as described for bacterial bins. 16S rRNA and 23S rRNA sequences were extracted from each genome using barrnap 0.9 (https://github.com/tseemann/barrnap) using the archaeal databases (23S: SILVA-LSU-Arc, 16S: RF01959) as reference. These sequences were then aligned against the nr database using BLASTn (Johnson *et al.*, 2008) for putative identification.

**Data availability.** Raw 16S rRNA gene amplicon sequence files were uploaded to the NCBI sequence read archive (SRA) 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.

## ACKNOWLEDGEMENTS

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

The authors acknowledge funding from the Gordon and Betty Moore Foundation (Grant number 6920) (awarded to R.A.D and J.C.K.), and grants awarded to R.A.D by the South African National Research Foundation (UID: 87583 and 109680). Development of Autometa in J.C.K.'s laboratory and contributions by E.R.R. were supported by the U.S. 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, 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 Center for High Performance Computing (CHPC, South Africa) for providing computing facilities for bioinformatics data analysis.

The authors declare no conflict of interests.

## REFERENCES

665 Aguilar-Barajas, E., Díaz-Pérez, C., Ramírez-Díaz, M.I., Riveros-Rosas, H., and
666 Cervantes, C. (2011) Bacterial transport of sulfate, molybdate, and related oxyanions.
667 *Biometals* **24**: 687–707.
668 Allen, J.F. (2016) A Proposal for Formation of Archaean stromatolites before the advent
669 of oxygenic photosynthesis. *Front Microbiol* **7**: 1784.
670 Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., and
671 Ogata, H. (2019) KofamKOALA: KEGG ortholog assignment based on profile HMM and
672 adaptive score threshold. *bioRxiv*. DOI: https://doi.org/10.1101/602110
673 Awramik, S.M. (1992) The History and Significance of Stromatolites. In, Schidlowski,M.,
674 Golubic,S., Kimberley,M.M., McKirdy,D.M., and Trudinger,P.A. (eds), *Early Organic*
675 *Evolution: Implications for Mineral and Energy Resources*. Berlin, Heidelberg: Springer
676 Berlin Heidelberg, pp. 435–449.
677 Babilonia, J., Conesa, A., Casaburi, G., Pereira, C., Louyakis, A.S., Reid, R.P., and
678 Foster, J.S. (2018) Comparative metagenomics provides insight into the ecosystem
679 functioning of the Shark Bay Stromatolites, Western Australia. *Front Microbiol* **9**: 1359.
680 Balci, N., Gunes, Y., Sertcetin, E., and Kurt, M.A. (2018) Non-cyanobacterial community
681 shape the stromatolites of Lake Salda, SW, Turkey. American Geophysical Union, Fall
682 Meeting 2018, abstract #B43J-2972.
683 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., et al.
684 (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell
685 sequencing. *J Comput Biol* **19**: 455–477.
686 Bontognali, T.R.R., Sessions, A.L., Allwood, A.C., Fischer, W.W., Grotzinger, J.P.,
687 Summons, R.E., and Eiler, J.M. (2012) Sulfur isotopes of organic matter preserved in

688 3.45-billion-year-old stromatolites reveal microbial metabolism. *Proc Natl Acad Sci U S*
689 *A* **109**: 15146–15151.
690 Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy,
691 T.B.K., et al. (2017) Minimum information about a single amplified genome (MISAG)
692 and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat*
693 *Biotechnol* **35**: 725–731.
694 Casaburi, G., Duscher, A.A., Pamela Reid, R., and Foster, J.S. (2016) Characterization
695 of the stromatolite microbiome from Little Darby Island, The Bahamas using predictive
696 and whole shotgun metagenomic analysis. *Environ Microbiol* **18**: 1452–1469.
697 Centeno, C.M., Mejía, O., and Falcón, L.I. (2016) Habitat conditions drive phylogenetic
698 structure of dominant bacterial phyla of microbialite communities from different locations
699 in Mexico. *Rev Biol Trop* **64**: 1057–1065.
700 Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019) GTDB-Tk: A
701 toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*.
702 btz848.
703 Cosmidis, J., Benzerara, K., Guyot, F., Skouri-Panet, F., Duprat, E., Férard, C., et al.
704 (2015) Calcium-phosphate biomineralization induced by alkaline phosphatase activity in
705 *Escherichia coli*: Localization, kinetics, and potential signatures in the fossil record.
706 *Front Earth Sci* **3**: 84.
707 Criscuolo, A. (2019) A fast alignment-free bioinformatics procedure to infer accurate
708 distance-based phylogenetic trees from genome assemblies. *Riogrande Odontol* **5**:
709 e36178.
710 Danen-Louwerse, H.J., Lijklema, L., and Coenraats, M. (1995) Coprecipitation of

711 phosphate with calcium carbonate in Lake Veluwe. *Water Res* **29**: 1781–1785.

712 Dodd, C., Anderson, C.R., Perissinotto, R., du Plooy, S.J., and Rishworth, G.M. (2018)
713 Hydrochemistry of peritidal stromatolite pools and associated freshwater inlets along the
714 Eastern Cape coast, South Africa. *Sediment Geol* **373**: 163–179.

715 Drummond, J.B.R., Pufahl, P.K., Porto, C.G., and Carvalho, M. (2015) Neoproterozoic
716 peritidal phosphorite from the Sete Lagoas Formation (Brazil) and the Precambrian
717 phosphorus cycle. *Sedimentology* **62**: 1978–2008.

718 Dupraz, C., Reid, R.P., Braissant, O., Decho, A.W., Norman, R.S., and Visscher, P.T.
719 (2009) Processes of carbonate precipitation in modern microbial mats. *Earth-Sci Rev*
720 **96**: 141–162.

721 Dupraz, C., Visscher, P.T., Baumgartner, L.K., and Reid, R.P. (2004) Microbe-mineral
722 interactions: Early carbonate precipitation in a hypersaline lake (Eleuthera Island,
723 Bahamas). *Sedimentology* **51**: 745–765.

724 Ellis, H.R. (2011) Mechanism for sulfur acquisition by the alkanesulfonate
725 monooxygenase system. *Bioorg Chem* **39**: 178–184.

726 Ershov, V., Tarasov, A., Lapidus, A., and Korobeynikov, A. (2019) IonHammer:
727 Homopolymer-space hamming clustering for IonTorrent read error correction. *J Comput*
728 *Biol* **26**: 124–127.

729 Gallagher, K.L., Braissant, O., Kading, T.J., Dupraz, C., and Visscher, P.T. (2013)
730 Phosphate-related artifacts in carbonate mineralization experiments. *J Sediment Res*
731 **83**: 37–49.

732 Gallagher, K.L., Kading, T.J., Braissant, O., Dupraz, C., and Visscher, P.T. (2012)
733 Inside the alkalinity engine: The role of electron donors in the organomineralization

potential of sulfate-reducing bacteria. *Geobiology* **10**: 518–530.

Gleeson, D.B., Wacey, D., Waite, I., O'Donnell, A.G., and Kilburn, M.R. (2016) Biodiversity of living, non-marine, thrombolites of Lake Clifton, Western Australia. *Geomicrobiol J* **33**: 850–859.

Gorley, R.N. and Clarke, K.R. (2015) PRIMER v7: User Manual/Tutorial, Massey University, Albany Campus.

Griffith, J.C. (2016) Insights into the soil microbial communities in New Zealand indigenous tussock grasslands. (Thesis, Doctor of Philosophy). University of Otago. Retrieved from http://hdl.handle.net/10523/6906

Guang-Can, T.A.O., Shu-Jun, T., Miao-Ying, C.A.I., and Guang-Hui, X.I.E. (2008) Phosphate-solubilizing and-mineralizing abilities of bacteria isolated from soils. *Pedosphere* **18**: 515–523.

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

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.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., and Aluru, S. (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* **9**: 5114.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden,

T.L. (2008) NCBI BLAST: A better web interface. *Nucleic Acids Res* **36**: W5–9.

Kahnert, A., Vermeij, P., Wietek, C., James, P., Leisinger, T., and Kertesz, M.A. (2000) The ssu locus plays a key role in organosulfur metabolism in *Pseudomonas putida* S-313. *J Bacteriol* **182**: 2869–2878.

Kan, A.T., Fu, G., and Tomson, M.B. (2005) Adsorption and precipitation of an aminoalkylphosphonate onto calcite. *J Colloid Interface Sci* **281**: 275–284.

Keegan, K.P., Glass, E.M., and Meyer, F. (2016) MG-RAST, a metagenomics service for analysis of microbial community structure and function. *Microbial Environmental Genomics (MEG)* **1399:** 207–233.

Khodadad, C.L.M. and Foster, J.S. (2012) Metagenomic and metabolic profiling of nonlithifying and lithifying stromatolitic mats of Highborne Cay, the Bahamas. *PLoS One* **7**: e38229.

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

Kurth, D., Amadio, A., Ordoñez, O.F., Albarracín, V.H., Gärtner, W., and Farías, M.E. (2017) Arsenic metabolism in high altitude modern stromatolites revealed by metagenomic analysis. *Sci Rep* **7**: 1024.

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

Macintyre, I.G., Prufert-Bebout, L., and Reid, R.P. (2000) The role of endolithic

780 cyanobacteria in the formation of lithified laminae in Bahamian stromatolites.
781 *Sedimentology* **47**: 915–921.
782 Matcher, G.F., Dorrington, R.A., Henninger, T.O., and Froneman, P.W. (2011) Insights
783 into the bacterial diversity in a freshwater-deprived permanently open Eastern Cape
784 estuary, using 16S rRNA pyrosequencing analysis. *Water SA* **37.3**: 381 - 390.
785 Medina-Chávez, N.O., Viladomat-Jasso, M., Olmedo-Álvarez, G., Eguiarte, L.E., Souza,
786 V., and De la Torre-Zavala, S. (2019) Diversity of Archaea domain in Cuatro Cienegas
787 Basin: Archaean Domes. *bioRxiv* 766709. DOI: https://doi.org/10.1101/766709
788 Metcalf, W.W. and Wanner, B.L. (1993) Evidence for a fourteen-gene, *phnC* to *phnP*
789 locus for phosphonate metabolism in *Escherichia coli*. *Gene* **129**: 27–32.
790 Miller, I.J., Rees, E.R., Ross, J., Miller, I., Baxa, J., Lopera, J., et al. (2019) Autometa:
791 Automated extraction of microbial genomes from individual shotgun metagenomes.
792 *Nucleic Acids Res* **47**: e57.
793 Misi, A. and Richard Kyle, J. (1994) Upper Proterozoic carbonate stratigraphy,
794 diagenesis, and stromatolitic phosphorite formation, Irece Basin, Bahia, Brazil. *J*
795 *Sediment Res* **64**: 299–310.
796 Mobberley, J.M., Khodadad, C.L.M., Visscher, P.T., Reid, R.P., Hagan, P., and Foster,
797 J.S. (2015) Inner workings of thrombolites: Spatial gradients of metabolic activity as
798 revealed by metatranscriptome profiling. *Sci Rep* **5**: 12601.
799 Moreno-Vivián, C., Cabello, P., Martínez-Luque, M., Blasco, R., and Castillo, F. (1999)
800 Prokaryotic nitrate reduction: Molecular properties and functional distinction among
801 bacterial nitrate reductases. *J Bacteriol* **181**: 6573–6584.
802 Nutman, A.P., Bennett, V.C., Friend, C.R.L., Van Kranendonk, M.J., and Chivas, A.R.

(2016) Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. *Nature* **537**: 535–538.

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

Parks, D.H., Chuvochina, M., Chaumeil, P.-A., Rinke, C., Mussig, A.J., and Hugenholtz, P. (2019) Selection of representative genomes for 24,706 bacterial and archaeal species clusters provide a complete genome-based taxonomy. *bioRxiv* 771964. DOI: https://doi.org/10.1101/771964

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., and Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* **36**: 996–1004.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.

Payandi-Rolland, D., Roche, A., Vennin, E., Visscher, P.T., Amiotte-Suchet, P., Thomas, C., and Bundeleva, I.A. (2019) Carbonate precipitation in mixed cyanobacterial biofilms forming freshwater microbial tufa. *Minerals* **9**: 409.

Perissinotto, R., Bornman, T.G., Steyn, P.-P., Miranda, N.A.F., Dorrington, R.A., Matcher, G.F., et al. (2014) Tufa stromatolite ecosystems on the South African south coast. *S Afr J Sci* **110**: 01–08.

Ragot, S.A., Kertesz, M.A., and Bünemann, E.K. (2015) *phoD* Alkaline phosphatase gene diversity in soil. *Appl Environ Microbiol* **81**: 7281–7289.

826 Ragot, S.A., Kertesz, M.A., Mészáros, É., Frossard, E., and Bünemann, E.K. (2017) Soil
827 *phoD* and *phoX* alkaline phosphatase gene diversity responds to multiple environmental
828 factors. *FEMS Microbiol Ecol* **93**: fiw212
829 Reid, R.P., Visscher, P.T., Decho, A.W., Stolz, J.F., Bebout, B.M., Dupraz, C., et al.
830 (2000) The role of microbes in accretion, lamination and early lithification of modern
831 marine stromatolites. *Nature* **406**: 989–992.
832 Rishworth, G.M., Cawthra, H.C., Dodd, C., and Perissinotto, R. (2019) Peritidal
833 stromatolites as indicators of stepping-stone freshwater resources on the Palaeo-
834 Agulhas Plain landscape. *Quat Sci Rev* 105704.
835 Rishworth, G.M., van Elden, S., Perissinotto, R., Miranda, N.A.F., Steyn, P.-P., and
836 Bornman, T.G. (2016) Environmental influences on living marine stromatolites: Insights
837 from benthic microalgal communities. *Environ Microbiol* **18**: 503–513.
838 Rishworth, G.M., Perissinotto, R., Bird, M.S., and Pelletier, N. (2018) Grazer responses
839 to variable macroalgal resource conditions facilitate habitat structuring. *R Soc Open Sci*
840 **5**: 171428.
841 Rishworth, G.M., Perissinotto, R., Bird, M.S., Strydom, N.A., Peer, N., Miranda, N.A.F.,
842 and Raw, J.L. (2017) Non-reliance of metazoans on stromatolite-forming microbial mats
843 as a food resource. *Sci Rep* **7**: 42614.
844 Rishworth, G.M., Perissinotto, R., Bornman, T.G., and Lemley, D.A. (2017) Peritidal
845 stromatolites at the convergence of groundwater seepage and marine incursion:
846 Patterns of salinity, temperature and nutrient variability. *J Mar Syst* **167**: 68–77.
847 Rodriguez-Navarro, C., Rodriguez-Gallego, M., Ben Chekroun, K., and Gonzalez-
848 Muñoz, M.T. (2003) Conservation of ornamental stone by *Myxococcus xanthus*-induced

carbonate biomineralization. *Appl Environ Microbiol* **69**: 2182–2193.

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

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

Ruvindy, R., White, R.A., 3rd, Neilan, B.A., and Burns, B.P. (2016) Unravelling core microbial metabolisms in the hypersaline microbial mats of Shark Bay using high-throughput metagenomics. *ISME J* **10**: 183–196.

Santos, A.A., Venceslau, S.S., Grein, F., Leavitt, W.D., Dahl, C., Johnston, D.T., and Pereira, I.A.C. (2015) A protein trisulfide couples dissimilatory sulfate reduction to energy conservation. *Science* **350**: 1541–1545.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Seemann, T. (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068–2069.

Shiraishi, F., Ohnishi, S., Hayasaka, Y., Hanzawa, Y., Takashima, C., Okumura, T., and Kano, A. (2019) Potential photosynthetic impact on phosphate stromatolite formation after the Marinoan glaciation: Paleoceanographic implications. *Sediment Geol* **380**: 65–82.

Smith, A., Cooper, A., Misra, S., Bharuth, V., Guastella, L., and Botes, R. (2018) The

872 extant shore platform stromatolite (SPS) facies association: A glimpse into the
873 Archean? *Biogeosciences* **15**: 2189–2203.
874 Soo, R.M., Hemp, J., Parks, D.H., Fischer, W.W., and Hugenholtz, P. (2017) On the
875 origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. *Science*
876 **355**: 1436–1440.
877 Stasi, R., Neves, H.I., and Spira, B. (2019) Phosphate uptake by the phosphonate
878 transport system PhnCDE. *BMC Microbiol* **19**: 79.
879 Stüeken, E.E. (2016) Nitrogen in ancient mud: A biosignature? *Astrobiology* **16**: 730–
880 735.
881 Thomazo, C., Ader, M., and Philippot, P. (2011) Extreme 15N-enrichments in 2.72-Gyr-
882 old sediments: Evidence for a turning point in the nitrogen cycle. *Geobiology* **9**: 107–
883 120.
884 Visscher, P.T., Pamela Reid, R., and Bebout, B.M. (2000) Microscale observations of
885 sulfate reduction: Correlation of microbial activity with lithified micritic laminae in modern
886 marine stromatolites. *Geology* **28**: 919–922.
887 Visscher, P.T. and Stolz, J.F. (2005) Microbial mats as bioreactors: Populations,
888 processes, and products. In, Noffke,N. (ed), *Geobiology: Objectives, Concepts,*
889 *Perspectives*. Amsterdam: Elsevier, pp. 87–100.
890 Warden, J.G., Casaburi, G., Omelon, C.R., Bennett, P.C., Breecker, D.O., and Foster,
891 J.S. (2016) Characterization of microbial mat microbiomes in the modern thrombolite
892 ecosystem of Lake Clifton, Western Australia using shotgun metagenomics. *Front*
893 *Microbiol* **7**: 1064.
894 Wei, S., Cui, H., Jiang, Z., Liu, H., He, H., and Fang, N. (2015) Biomineralization

processes of calcite induced by bacteria isolated from marine sediments. *Braz J Microbiol* **46**: 455–464.

Westcott, S.L. and Schloss, P.D. (2017) OptiClust, an improved method for assigning amplicon-based sequence data to operational taxonomic units. *mSphere* **2**: e00073-17.

White, A.K. and Metcalf, W.W. (2007) Microbial metabolism of reduced phosphorus compounds. *Annu Rev Microbiol* **61**: 379–400.

White, R.A., Chan, A.M., Gavelis, G.S., Leander, B.S., Brady, A.L., Slater, G.F., et al. (2016) Metagenomic analysis suggests modern freshwater microbialites harbor a distinct core microbial community. *Front Microbiol* **6**: 1531.

Witze, A. (2016) Claims of Earth's oldest fossils tantalize researchers. *Nature News*. DOI:10.1038/nature.2016.20506

Wong, H.L., Visscher, P.T., White, R.A., III, Smith, D.-L., Patterson, M.M., and Burns, B.P. (2017) Dynamics of archaea at fine spatial scales in Shark Bay mat microbiomes. *Sci Rep* **7**: 46160.

Wong, H.L., White, R.A., Visscher, P.T., Charlesworth, J.C., Vázquez-Campos, X., and Burns, B.P. (2018) Disentangling the drivers of functional complexity at the metagenomic level in Shark Bay microbial mat microbiomes. *ISME J* **12**: 2619–2639.

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

Yang, J., Junium, C.K., Grassineau, N.V., Nisbet, E.G., Izon, G., Mettam, C., et al. (2019) Ammonium availability in the Late Archaean nitrogen cycle. *Nature Geoscience* **12**: 553–557.

Zaheer, R., Morton, R., Proudfoot, M., Yakunin, A., and Finan, T.M. (2009) Genetic and biochemical properties of alkaline phosphatases phoX family protein found in many bacteria. *Environ Microbiol* **11**: 1572–1587.

## TABLE AND FIGURE LEGENDS

**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) the Cape Recife site. Contour lines represent elevation at 5cm increments. Samples were collected from upper, middle and lower formations. The boundaries of which are indicated with a dotted blue line. (C - D) A 3-dimensional rendering of the sample sites at Schoenmakerskop and Cape Recife respectively constructed by a 3rd party (Caelum Technologies) using a combination of drone-based photogrammetry and geographic mapping using differential GPS. Samples were collected from upper, middle and lower formations; the boundaries of which are indicated with a dotted blue line. Abbreviations are as follows: CSU: Schoenmakerskop Upper Jan SU: Schoenmakerskop Upper April, CRU: Cape Recife Upper (Jan), RU: Cape Recife Upper (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 2.** Distribution and abundance of bacterial taxa in stromatolite formations based on 16S rRNA gene fragment amplicon libraries. (A) Phylogenetic classification and

941 average relative abundance (n=3) of dominant phyla in different sample sites indicated
942 that all stromatolite samples are dominated by Cyanobacteria, Bacteroidetes, Alpha-
943 and Gammaproteobacteria. (B) OTU abundance was used to cluster stromatolite
944 biological replicates using Bray-Curtis non-dimensional scaling and showed statistically
945 significant clustering of replicate samples from each of the sampled regions. Samples
946 were isolated from upper (green), middle (red) and lower (blue) stromatolites formations
947 in Cape Recife (triangles) and Schoenmakerskop (circles).
948
949 **Figure 3.** Summary of phosphate, nitrogen and sulfate transport and metabolism genes
950 in overall metagenomic data from stromatolites in Schoenmakerskop and Cape Recife
951 upper, middle and lower formations respectively. Gene abundance is expressed relative
952 to the length-weighted average contig coverage per sample and transformed using log2
953 scaling. Note: On the color scale, grey indicates genes which were not detected in the
954 respective sample. Abbreviations are as follows: CSU: Schoenmakerskop Upper Jan
955 SU: Schoenmakerskop Upper April, CRU: Cape Recife Upper (Jan), RU: Cape Recife
956 Upper (April), SM1: Schoenmakerskop Middle 1 (April), SM2: Schoenmakerskop Middle
957 2 (April), RM1: Cape Recife Middle 1 (April), RM2: Cape Recife Middle 1 (April), SL:
958 Schoenmakerskop Lower (April), RL: Cape Recife Lower (April).
959
960 **Figure 4.** Taxonomic classification of putative genome bins in stromatolites collected
961 from upper/inflow, middle and lower/marine formations of Schoenmakerskop and Cape
962 Recife. Coverage per genome has been used as a proxy for abundance and used to
963 scale the size of individual genome bars, expressed as a percentage of the sum of all

964 genome bin coverages. (A) The coverage of conserved (diagonal lines) and non-
965 conserved (solid colors) bacterial genomes. Taxonomic classification of each genome is
966 indicated by color. (B). Taxonomic classifications of conserved bacterial species in each
967 of the samples, generated with GTDB-Tk (Chaumeil *et al.*, 2019).
968
969 **Figure 5.** Summary of phosphate, nitrogen and sulfate transport and metabolism genes
970 in conserved genomes from bacteria associated with stromatolites in Schoenmakerskop
971 and Cape Recife upper, middle and lower formations respectively. Relative gene
972 abundance per sample is indicated using a gradient color scale. Relative gene
973 abundance was calculated by dividing coverage-corrected gene abundance in each
974 genome by the coverage-corrected gene abundance from the metagenomic sample
975 from which it came. CSU: Schoenmakerskop Upper Jan SU: Schoenmakerskop Upper
976 April, CRU: Cape Recife Upper (Jan), RU: Cape Recife Upper (April), SM1:
977 Schoenmakerskop Middle 1 (April), SM2: Schoenmakerskop Middle 2 (April), RM1:
978 Cape Recife Middle 1 (April), RM2: Cape Recife Middle 1 (April), SL: Schoenmakerskop
979 Lower (April), RL: Cape Recife Lower (April).
980
981 **Figure S1.** Relative abundance of phosphate, nitrogen and sulfate transport and
982 metabolism genes per genome relative to total gene abundance per sample from
983 Schoenmakerskop and Cape Recife upper, middle and lower formations. Relative gene
984 abundance per sample is indicated using a gradient color scale. Relative gene
985 abundance was calculated by dividing coverage-corrected gene abundance in each
986 genome by the coverage-corrected gene abundance from the metagenomic sample

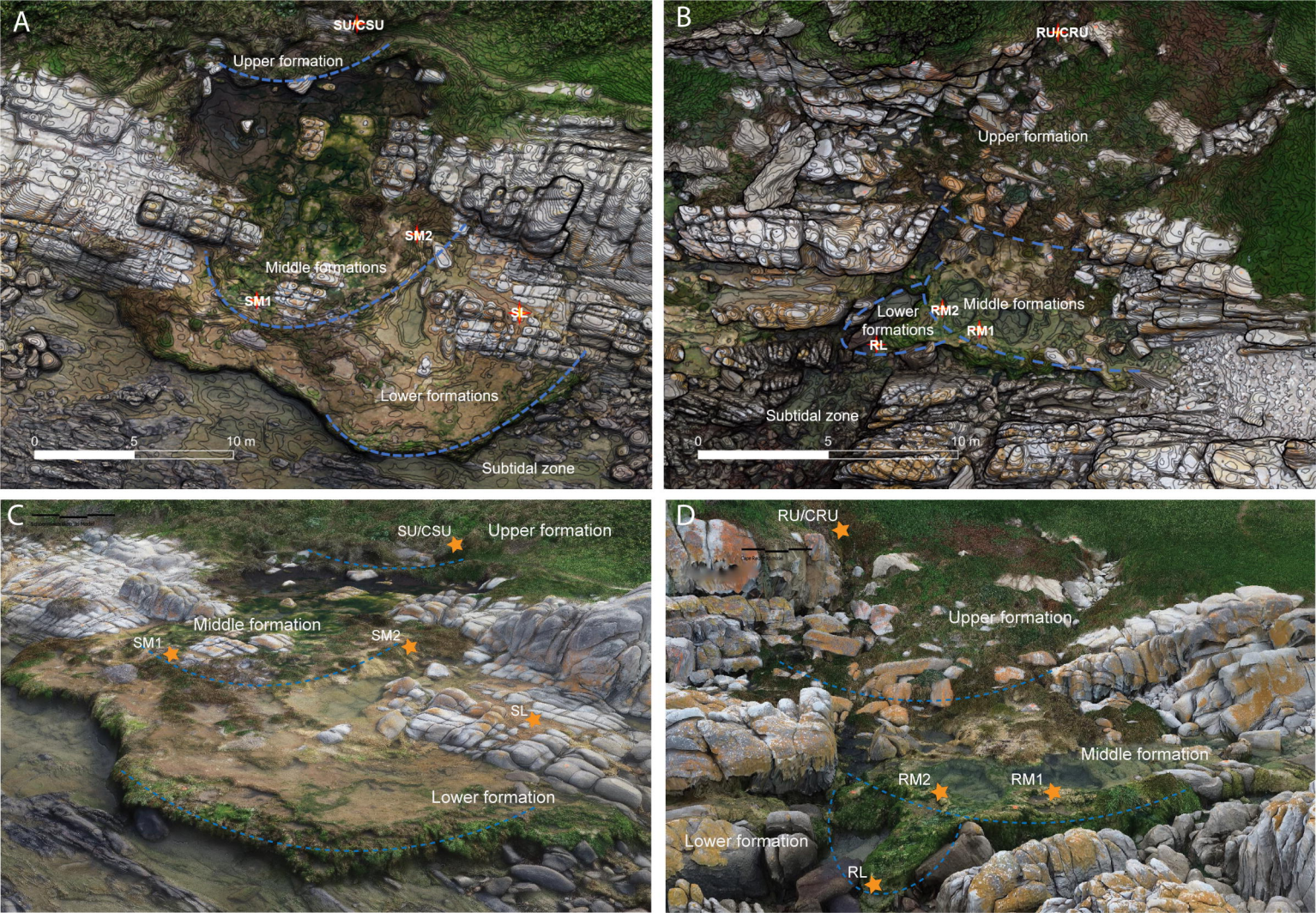
from which it came. CSU: Schoenmakerskop Upper Jan SU: Schoenmakerskop Upper April, CRU: Cape Recife Upper (Jan), RU: Cape Recife Upper (April), SM1: Schoenmakerskop Middle 1 (April), SM2: Schoenmakerskop Middle 2 (April), RM1: Cape Recife Middle 1 (April), RM2: Cape Recife Middle 1 (April), SL: Schoenmakerskop Lower (April), RL: Cape Recife Lower (April).

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

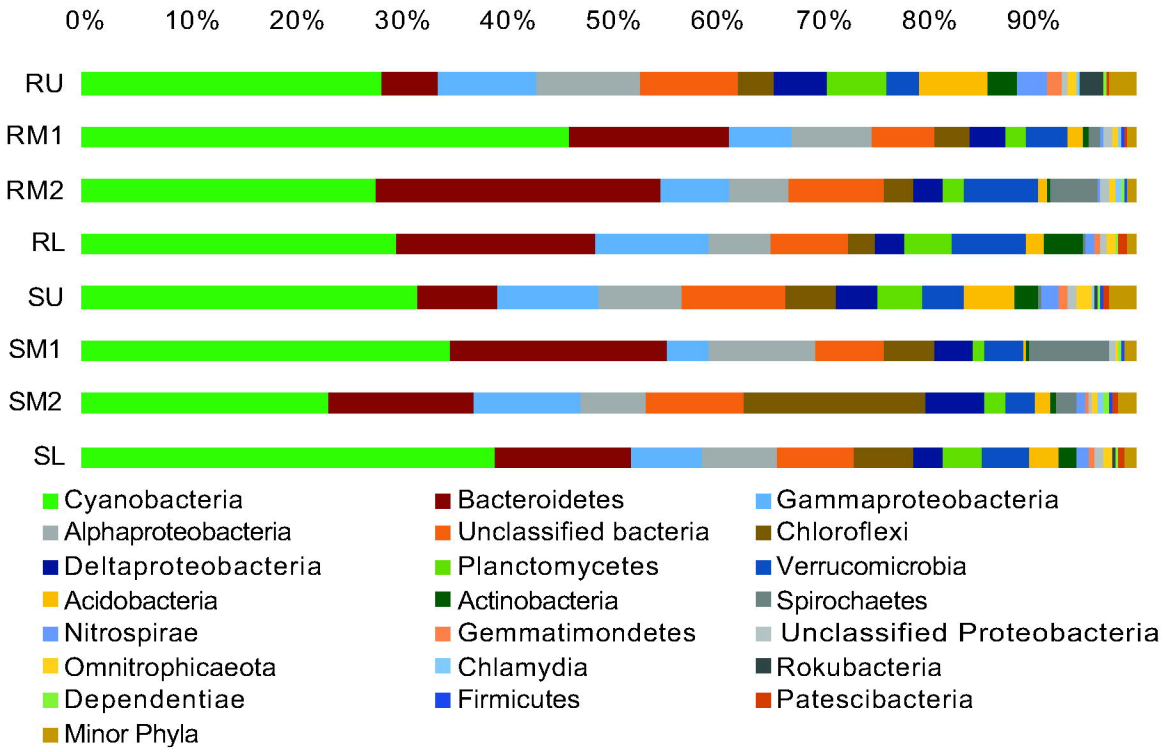
**Table S1.** Summary of genomes binned per collected stromatolite sample

**Table S2.** Conserved bacterial species defined by shared ANI greater than 97% in stromatolite formations from Cape Recife and Schoenmakerskop.

**Table S3.** Summary of characteristics and taxonomic classifications of genomes binned from shotgun metagenomic data from sampled upper, middle and lower stromatolite formations from Cape Recife and Schoenmakerskop. Bacterial bins are listed first, with the two archaeal bins listed at the bottom.


A
SU/CSU
Upper formation
SM2
Middle formations
SM1
SL
Lower formations
0
5
10 m
Subtidal zone
B
RU/CRU
Upper formation
Lower formations
RM2
Middle formations
RM1
RL
Subtidal zone
0
5
10 m
C
SU/CSU
Upper formation
Middle formation
SM1
SM2
SL
Lower formation
D
RU/CRU
Upper formation
Middle formation
RM2
RM1
Lower formation
RL

A.




B.