Published in final edited form as:
Environ Microbiol. 2016 May ; 18(5): 1452–1469. doi:10.1111/1462-2920.13094.

Characterization of the stromatolite microbiome from Little Darby Island, The Bahamas using predictive and whole shotgun metagenomic analysis

Giorgio Casaburi[1], **Alexandrea A. Duscher**[1], **R. Pamela Reid**[2], and **Jamie S. Foster**[1,*]

[1]Department of Microbiology and Cell Science, University of Florida, Space Life Science Lab, Merritt Island, FL, USA

[2]Rosenstiel School of Marine Sciences, University of Miami, Miami, FL, USA

Summary

Modern stromatolites represent ideal ecosystems to understand the biological processes required for the precipitation of carbonate due to their long evolutionary history and occurrence in a wide range of habitats. However, most of the prior molecular work on stromatolites has focused on understanding the taxonomic complexity and not fully elucidating the functional capabilities of these systems. Here, we begin to characterize the microbiome associated with stromatolites of Little Darby Island, The Bahamas using predictive metagenomics of the 16S rRNA gene coupled with direct whole shotgun sequencing. The metagenomic analysis of the Little Darby stromatolites revealed many shared taxa and core pathways associated with biologically induced carbonate precipitation, suggesting functional convergence within Bahamian stromatolites. A comparison of the Little Darby stromatolites with other lithifying microbial ecosystems also revealed that although factors, such as geographic location and salinity, do drive some differences within the population, there are extensive similarities within the microbial populations. These results suggest that for stromatolite formation, 'who' is in the community is not as critical as metabolic activities and environmental interactions. Together, these analyses help improve our understanding of the similarities among lithifying ecosystems and provide an important first step in characterizing the shared microbiome of modern stromatolites.

Introduction

Stromatolites are quintessential geomicrobiological ecosystems with a fossil record that extends back 3.5 billion years (Grotzinger and Knoll, 1999). These long-lived features are sedimentary structures formed as a result of the synergy between the metabolisms of microbial mats and the environment (Reid *et al.*, 2000). Ancient stromatolites formed massive carbonate reefs comparable in size to modern coral reefs, dominating life on our planet for 80% of Earth's history (Awramik, 1984). Modern stromatolites, in contrast, are relatively rare, but have been found in diverse habitats including freshwater, marine and

*For correspondence. jfoster@ufl.edu; Tel. 321-525-1047.

Supporting information
Additional Supporting Information may be found in the online version of this article at the publisher's website:

hypersaline environments (Playford and Cockbain, 1976; Reid *et al.*, 1995; Reid et al., 2003; Laval *et al.*, 2000; Breitbart *et al.*, 2009; Santos *et al.*, 2010; Farías *et al.*, 2013; Schneider *et al.*, 2013; Perissinotto *et al.*, 2014). One of the most well-studied stromatolite systems has been the island of Highborne Cay located in the Northern Exuma Cays, The Bahamas, which for more than 20 years has served as a model to understand the geological, biochemical and biological processes associated with stromatolite formation and accretion under normal sea water conditions (e.g. Reid *et al.*, 1995; Visscher *et al.*, 1998; Reid *et al.*, 2000; Baumgartner *et al.*, 2009a; Stolz *et al.*, 2009; Khodadad and Foster, 2012).

These previous studies have identified several key guilds of microbes that together, through their metabolic activities, generate steep geochemical gradients throughout the microbial mat community that can influence the precipitation potential of calcium carbonate (Paerl *et al.*, 2001; Dupraz and Visscher, 2005; Dupraz *et al.*, 2009). These functional groups have been identified through biogeochemical and 16S rRNA gene analyses and include: oxygenic and anoxygenic phototrophs, aerobic heterotrophs, sulfate reducers, sulfide oxidizers and fermenters (Visscher *et al.*, 1998; 2000; Steppe *et al.*, 2001; Baumgartner *et al.*, 2009a, b; Foster *et al.*, 2009; Khodadad and Foster, 2012). It is the net activity of these different functional groups that influence carbonate precipitation by altering the pH within the microbial mat. For example, some metabolisms, such as photosynthesis and some types of sulfate reduction, promote carbonate precipitation through an increase in pH, whereas other metabolisms, such as sulfide oxidation and aerobic respiration, can decrease the pH, thereby promoting dissolution of calcium carbonate (Dupraz and Visscher, 2005; Visscher and Stolz, 2005; Gallagher *et al.*, 2012).

Although considerable progress has been made in delineating key metabolisms and microbes associated with modern stromatolites of Highborne Cay, the underlying molecular mechanisms and cellular interactions that drive these geochemical gradients are not well characterized (Khodadad and Foster, 2012). Additionally, very little is known regarding underlying microbial and functional gene diversity of other stromatolites found throughout The Bahamas. Stromatolites have also been observed in the Southern Exuma Cays including Bock Cay, Iguana Cay, Lee Stocking Island and Little Darby Island (Dill *et al.*, 1986; Pinckney *et al.*, 1995b; Reid *et al.*, 1995; 2010; Feldmann and McKenzie, 1998) and along the eastern margin of Exuma Sound in Schooner Cays (Dravis, 1983); however, studies of these sites have been limited to biogeochemical and morphological analyses.

To more fully understand the mechanisms of stromatolite formation and accretion it is necessary to begin to characterize the communities beyond biogeochemistry and taxonomic diversity and assess the associated stromatolite microbiome, that is, the totality of the microbes, functional genes and environmental interactions that result in the deposition of calcium carbonate. Not all microbial mats are conducive to the formation of stromatolite structures (Dupraz *et al.*, 2009; Foster and Green, 2011), and characterizing the microbiome represents a holistic approach to understanding the connectivity and microbial interactions that lead to stromatolite formation. In this study, we begin to characterize the microbiome of stromatolites from Little Darby Island, located in the Southern Exuma Cays by assessing the microbial and functional gene diversity using two complementary approaches, 16S rRNA gene and shotgun metagenomic sequencing analysis. Much like the stromatolites of

Highborne Cay, the stromatolites of Little Darby Island are subtidal, forming under normal seawater conditions and experiencing frequent burial events. Additionally, these ecosystems accrete through the trapping and binding of sediments and carbonate precipitation by the surface microbial mat community (Fig. 1; Reid *et al.*, 2010). By expanding the analysis of stromatolite-forming communities beyond Highborne Cay, we can begin to assess those taxa and pathways shared in Bahamian stromatolites and determine whether these similarities extend beyond The Bahamas to other lithifying microbial habitats. Together, these results may help delineate the essential biological processes and pathways needed for modern stromatolite formation.

Results and discussion

Overview of stromatolite environment of Little Darby Island, The Bahamas

The stromatolites of Little Darby Island are located on the northeast side of the island and are distributed parallel to the shore in a narrow band approximately 300 m in length and 100 m wide. The stromatolites occur in the subtidal zone at depths ranging from 1 m to 3 m and vary in height from 2 cm to 1 m (Fig. 1A). The Little Darby stromatolites undergo extensive sand transport through wave action and experience partial and/or full burial events throughout the year. Microbial mats occur on the surfaces of the stromatolite build-ups and typically exceed 2 cm in thickness. The upper 2 mm of the microbial mats are enriched in exopolymeric substances (EPS) giving the surface mat a caramel colour (Fig. 1B). Microscopic analysis of the caramel surface layer revealed a layer of vertically orientated cyanobacterial filaments morphologically identified as primarily *Schizothrix* spp. and *Phormidium* spp. (Fig. 1C). Additional pronounced layers of cyanobacteria were visible in cross-sections of the stromatolite-forming mats at a depth of 5 mm (Fig. 1B).

Taxonomic diversity of Little Darby stromatolites

To assess the microbial communities associated with Little Darby stromatolites, a two-pronged approach was used that included targeted bacterial 16S rRNA gene amplicon and whole shotgun metagenomic sequencing. Recent studies have shown numerous limitations when relying on a single marker gene for assessing microbial diversity, including primer bias, lower taxonomic resolution and stability issues regarding operational taxonomic units (OTU) clustering (Yousef *et al.*, 2009; He *et al.*, 2015). Whole genome shotgun metagenomics overcomes the limitations of targeting a single gene and can provide a comprehensive assessment of community diversity (Gilbert *et al.*, 2010). However, in complex communities, such as lithifying microbial mats, genomic databases can be far less representative and sometimes taxonomically skewed compared with some well-curated 16S rRNA gene databases (Poretsky *et al.*, 2014). Therefore, we used a dual approach to assess the overall community complexity of the Little Darby stromatolites. For the amplicon library, recovered reads were quality filtered in QIIME resulting in 2 305 124 high quality filtered paired-end reads (see experimental procedures for additional detail). To help improve the diversity estimates, a cut-off filter of 0.005% as the minimum total observation count of an OTU was applied in QIIME, as previously described (Bokulich *et al.*, 2013). The stringent filtering resulted in 2157 remaining OTUs that represented 13 phyla within microbial mats (Fig. 2A). Of these phyla, the *Proteobacteria* (62.2%) and the *Cyanobacteria*

(31.6%) were dominant with the mats. *Bacteroidetes* and *Spirochetes* represented only 3.4% and 1.26% of the community, respectively, whereas the other phyla each comprised less than 1% of the community (Fig. 2A). Other phyla were observed within the 16S rRNA gene libraries but at levels below 0.005% abundance. These results strongly correlate with previous 16S rRNA gene sequencing of Highborne Cay stromatolites (Baumgartner *et al.*, 2009a; Foster and Green, 2011).

Within the *Proteobacteria*, the most represented classes were the *Alphaproteobacteria* (61.5%) and *Deltaproteobacteria* (3.2%) (Fig. 2A). The *Alphaproteobacteria* in the Little Darby stromatolites were enriched in taxa known to be anoxygenic phototrophs, specifically purple non-sulfur bacteria derived from the orders *Rhodobacterales* (15.3%), *Rhizobiales* (11.3%) and *Rhodospirillales* (7.8%), whereas the *Deltaproteobacteria* were primarily comprised of sulfate-reducing bacteria associated with the taxa *Desulfobacterales* (1.7%) and *Desulfovibrionales* (0.2%). The cyanobacterial population within the stromatolite-forming mats was dominated by the subclass *Oscillatoriophycideae* (39%) primarily with the orders *Chroococcales* (9%) and *Oscillatoriales* (0.6%). Other cyanobacterial orders were recovered but at much lower abundance, such as the *Nostocales* (0.2%) and *Pleurocapsales* (0.3%); however, it is important to note that most of the recovered cyanobacterial sequences were unable to be classified beyond the phylum level and that many of the cyanobacterial taxa previously identified in Little Darby Island stromatolites using morphological approaches, such as *Schizothrix* spp. and *Solentia* spp. (Reid *et al.*, 2010), have no representative genome available.

Whole genome shotgun metagenomic sequencing of DNA derived from the seven distinct stromatolite heads generated a total of 7 927 495 quality-filtered, paired-end reads. The high-quality reads were processed using both Metagenome Composition Vector (MetaCV), which targets primarily prokaryotes and the MG-RAST pipeline, which targets all three domains. The majority of the metagenomic reads were derived from *Bacteria* (>70%). *Archaea* constituted 10% of the total data set, whereas *Eukarya* were represented by less than 3% of the recovered sequences with 16% of the total sequences reported as unclassified in both analysis programs. Analysis of the metagenomic data revealed a much more complex bacterial community within the stromatolite-forming mats compared with the 16S rRNA gene libraries (Fig. 2B) and provided additional insight into the archaeal and eukaryotic community (Figs S1–2). Based on MetaCV analysis, there was a total of 2883 OTUs, a 33% increase compared with the 16S rRNA gene library, representing a total of 27 bacterial phyla in the shotgun metagenomic library. Similarly to the amplicon library, the *Proteobacteria* (50.0%) and *Cyanobacteria* (28.4%) were the dominant phyla within the community. The metagenomic sequencing analysis provided a much higher taxonomic resolution of the *Bacteria* with over 232 families observed in the data set (Fig. 2B). The results also revealed a much more complex *Alphaproteobacteria* population specifically in the *Rhizobiales* with recovered sequences sharing similarity to the facultative methyltrophs of the families *Xanthobacteraceae* and *Methylobacteriaceae*, which have the ability to grow on methylamine, fructose, fucose, xylose, arabinose and some organic acids (Green, 2006). Substrates, such as these, have been found in high abundance in the EPS that comprise stromatolite-forming mats (Kawaguchi and Decho, 2002). However, as in the 16S rRNA gene analysis, the dominant *Alphaproteobacteria* in the shotgun library were the

phototrophic purple non-sulfur bacteria *Rhodobacteraceae* (10.5%) and *Rhodospirallaceae* (6.5%), which correlates with the alphaproteobacterial populations found in other lithifying microbial mat systems such as the thrombolites of Highborne Cay, The Bahamas; hypersaline stromatolites of Hamelin Pool, Shark Bay; Kiritimati Atoll, Kiribati; and Socampa, Argentina (Foster and Green, 2011; Khodadad and Foster, 2012; Farías *et al.*, 2013; Schneider *et al.*, 2013).

Metagenomic sequencing analysis of the cyanobacterial population strongly correlated with the 16S rRNA gene analysis indicating that the microbial community was primarily comprised of *Cyanobacteria* derived from the subclass *Oscillatoriophycideae* and shared sequence similarity to genera such as *Cyanothece*, *Synechococcus* and *Acaryochloris* (Fig. S3). Shotgun metagenomic sequencing also indicated that *Nostocales* (5.2%) occur at a higher relative abundance compared with the 16S rRNA gene analysis. However, one major caveat to taxonomic identification of *Cyanobacteria* is the relatively low diversity of sequenced reference genomes for data comparison. This under-representation in the MetaCV database may account for the lack of *Pleurocapsales* and representation of the *Oscillatoriales* family by a single genus *Trichodesmium* (Fig. S3). These results reinforce the need for a greater diversity of cyanobacterial cultivars to be sequenced.

The shotgun metagenomic library also provided information on the archaeal and eukaryotic populations within the Little Darby stromatolites. Methanogens dominated the archaeal population and were diverse, containing representatives of all known orders, *Methanobacteriales*, *Methanococcales*, *Methanopyrale*s, *Methanosarcinales*, *Methanomicrobiales* and the recently recognized *Methanocellales* (Fig. S1). Of the different orders of methanogens, the metabolically versatile *Methanosarcinales* were the most abundant. Members of this order can grow and generate methane from a wide range of compounds including hydrogen, acetate, methanol and many other single carbon molecules (Anderson *et al.*, 2009; Sakai *et al.*, 2011). Genes associated with archaeal methane metabolism accounted for approximately 5% of the recovered archaeal functional genes (data not shown).

Another enriched archaeal taxa were the *Nitrosopumilales*, an ammonia-oxidizing *Thaumarchaeota*, which has been shown to be prominent in unlaminated, thrombolite-forming microbial mats from Highborne Cay, The Bahamas (Mobberley *et al.*, 2012; 2013; 2015). Additionally, the Little Darby stromatolite communities were enriched with *Halobacteria*, which are obligate heterotrophs. *Halobacteria* have been observed in a wide range of lithifying microbial mat systems, such as the hypersaline stromatolites of Shark Bay (Burns *et al.*, 2004; Goh *et al.*, 2006; Leuko *et al.*, 2007; Ruvindy *et al.*, 2015). Although their precise ecological role in stromatolite formation has not been delineated, these *Haloarchaea* are metabolically diverse, known to form biofilms (Fröls *et al.*, 2012), metabolize carbohydrates (e.g. pentoses, hexoses; Andrei *et al.*, 2012) and grow photoheterotrophically (Hartmann *et al.*, 1980); all of which are important elements in lithifying microbial mat ecosystems (Dupraz *et al.*, 2009).

The shotgun metagenomic sequencing revealed that diatoms dominated the eukaryotic population within the Little Darby stromatolite-forming mats, and were enriched in the

orders Bacillariophyceae, Fragilariophyceae, Coscinodiscophyceae and Mediophyceae (Fig. S2). Previous studies on the stromatolites of Highborne Cay have identified several dominant species of diatoms including the stalked diatoms *Striatella unipunctata* and *Licmorpha remulus* (class Fragilariophyceae) as well as several Naviculid (class Bacillariophyceae) tube diatoms (Stolz *et al.*, 2009). Diatoms have been shown to play an important role in the accretion of Bahamian stromatolites (Reid *et al.*, 2000; Stolz *et al.*, 2009; Bowlin *et al.*, 2012). Both the tube and stalked diatoms can form networks that trap sediment grains and become bound into the stromatolite framework by filamentous *Cyanobacteria* forming unlithified grain layers (Bowlin *et al.*, 2012). The high abundance of diatoms within the eukaryotic population likely reflects the collection of samples in August, where water temperatures were > 27°C. Previous work has shown a positive correlation of the relative abundance of stalked and tube diatoms under elevated temperatures (Bowlin *et al.*, 2012). The other dominant eukaryotic taxa within the Little Darby stromatolite ecosystems including Rhodophyta, Chlorophyta, Streptophyta and Heterokontophyta (Stramenopiles) have all been well documented in other lithifying microbial ecosystems including Highborne Cay, Cuatro Ciénegas, Mexico and Hamelin Pool, Shark Bay Western Australia (Breitbart *et al.*, 2009; Myshrall *et al.*, 2010; Mobberley *et al.*, 2012; 2013; Edgcomb *et al.*, 2014). Functional genes associated with eukaryotic taxa were primarily associated with genetic information processing and cellular processes, such as cell growth, transport and death (data not shown). However, there were numerous carbohydrate metabolisms genes associated with sugar metabolism (e.g. fructose, xylose, galactose), suggesting the eukaryotes may be contributing to the heterotrophic degradation of EPS material within the stromatolite forming mats.

An overlay of the 16S rRNA gene amplicon and metagenomic libraries is visualized in Fig. 3. The radar plot depicts the relative abundance of recovered genes associated with the 92 dominant shared bacterial families within the stromatolite-forming community and are expressed as a percentage on a log scale. There is a high correlation (r = 0.91, Pearson) between the two data sets regarding the relative abundance of the dominant bacterial families, suggesting that both approaches capture the overall diversity of the dominant taxa within the community. For example, in some phyla, such as *Cyanobacteria* and *Firmicutes*, there was a strong correlation between the relative abundances of families between the 16S rRNA gene and metagenomic data sets (Fig. 3). However, for some taxa there was extensive variability in the relative abundance observed between the data sets. In the *Proteobacteria*, there were several examples of families, such as the alphaproteobacterial families Brucellaceae and Bradyrhizobiaceae, which had a higher relative abundance in the shotgun metagenomic library. These differences between the data sets may reflect issues regarding many of the clustering programs associated with 16S rRNA gene analysis. Several recent studies have identified that OTU clustering of sequences can be impacted by the number of sequences being clustered (He *et al.*, 2015). Although there was a strong correlation between the two data sets, assigning a taxonomic classification to metagenomic reads provided a more comprehensive assessment of the complexity of the stromatolite forming mat community structure than with 16S rRNA gene analysis alone.

Functional gene complexity of stromatolite-forming microbial mats using both predictive and whole shotgun sequencing analyses

For decades ribosomal RNA gene analysis has provided a robust mechanism to survey the diversity of complex microbial communities (e.g. Woese *et al.*, 1990); however, it provided little direct evidence of the metabolic capabilities. With the recent, rapid expansion of sequenced reference genomes, it has now become possible to use this data to help predict the functional composition of communities metagenome (Zaneveld *et al.*, 2010; Collins and Higgs, 2012; Langille *et al.*, 2013). Using the program Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt), the gene content of the target ecosystem can be inferred using the reference genomes of one or more of its closest relatives as well as a reconstruction of the organisms' ancestral genome (Langille *et al.*, 2013). This program has been effectively used to predict accurately the functional complexity of the metagenomes of the human microbiome, soils and the hypersaline, non-lithifying microbial mats (Langille *et al.*, 2013). However, some environmental metagenomes, such as stromatolite ecosystems, have the potential problem that there may not be enough reference genomes available to make accurate predictions. To assess the effectiveness of using the 16S rRNA gene as a prognostic tool for metagenomic analyses in stromatolites, we generated a simulated prediction of the metabolic functions of the Little Darby stromatolites from 16S rRNA gene sequences with PICRUSt. This predictive metagenomic approach utilized the QIIME output for taxonomic profiling and directly compared it with the whole shotgun sequenced metagenome library.

For the 16S rRNA gene library, we identified a total of 6909 KEGG (Kyoto Encyclopedia of Genes and Genomes) functions, corresponding to 328 level 3 KEGG orthology (KO) entries, whereas the post-quality filtered whole metagenome shotgun reads analysed in MetaCV showed a total of only 2944 KEGG functions, corresponding to 257 level 3 KO entries. We filtered out those entries occurring at a low relative abundance (<1%) and compared the two data sets. In total, we identified 88 dominant level 3 KOs in both data sets, mostly related to metabolism (Fig. 4). There was a moderate to strong correlation between the 16S rRNA gene PICRUSt prediction and the whole metagenome sequencing ($r = 0.78$, Pearson) and the two data sets never varied more than 4% in each KO pathway (Fig. 4). Photosynthesis, chlorophyll metabolism and carbon fixation pathways in photosynthetic organisms were dominant within the metagenome with a higher representation of genes associated with these metabolisms in the whole shotgun library compared with the 16S rRNA gene predicted libraries. The underestimation of genes related to photosynthetic pathways in the 16S rRNA gene predictions further confirms the need for additional cyanobacterial reference genomes to be targeted for sequencing to fill phylogenetic gaps in the databases.

Other metabolic pathways that were enriched in the Little Darby stromatolite metagenome included genes associated with aerobic oxidative phosphorylation, nitrogen fixation, glycolysis, as well as methane metabolism. Together, the metagenome reflected all of the major functional groups, in varying relative abundances, typically associated with lithifying microbial mat systems (Dupraz and Visscher, 2005; Visscher and Stolz, 2005) and overlap with the dominant pathways in the stromatolites and thrombolites of Highborne Cay (Khodadad and Foster, 2012; Mobberley *et al.*, 2013; 2015).

In addition to the various microbial metabolisms, there were also a high abundance of genes associated with two-component signalling systems (Fig. 4), including genes associated with quorum sensing and its regulation (e.g. *luxQ*, *luxU*, *luxO*), as well as numerous genes associated with OmpR family that respond to a range of environmental stress conditions, such as high light, nutrient limitation (e.g. phosphate, nitrogen, Mg^{2+}, Mn^{2+}) and osmotic stress suggesting a strong ability of stromatolite-forming microorganisms to sense and consequently respond to changes in different environmental conditions. Environmental factors, such as sand burial events, temperature and photosynthetic active radiation have been shown to have a profound impact on the rates of metabolic activity within lithifying microbial ecosystems, as well as provide control of the cycling and development of the dominant microbial mat communities (Dupraz *et al.*, 2009; Bowlin *et al.*, 2012). Moreover, lithifying microbial mat ecosystems are usually oligotrophic and must scavenge nutrients from the surrounding water. A high number of sequences associated with sensing and responding to nutrient limitations, such as the phosphate regulon (*phoR-phoP*), were observed in the metagenome. The results corresponded with other microbialite ecosystems, such as the alkaline lake Alchichica and fresh water springs of Cuatro Ciénegas in Mexico (Breitbart *et al.*, 2009; Valdespino-Castillo *et al.*, 2014). Much like the Mexican microbialites, the phosphate regulon and assimilation genes from the Little Darby stromatolites were derived from a diverse group of organisms, primarily the *Proteobacteria* classes *Alphaproteobacteria*, *Gamma-proteobacteria* and *Deltaproteobacteria*, *Cyanobacteria* class *Oscillatoriophycideae* and *Actinobacteria* (Fig. S4).

Together, the PICRUSt prediction and shotgun library approaches suggest a strong correlation between phylogeny and function, further supporting that, even for environmental metagenomes, such as stromatolites, 16S rRNA gene libraries can be used to provide insight into the metabolic capabilities of complex microbial ecosystems. Although whole metagenome shotgun sequencing is becoming increasingly more affordable, providing the necessary sequencing and sampling depth for all environments may still be cost-prohibitive. Therefore, these results suggest that using phylogenic markers, such as the 16S rRNA gene, can provide an important first look into the functional complexity of lithifying mat ecosystems as well as the microbial diversity.

To delineate the taxa associated with most highly represented genes of the stromatolite shotgun metagenome library, genomic reads were matched with metabolic function in MetaCV and then grouped by phylogenetic relationships in MEGAN 4 (see *Experimental procedures*). A heat map correlating the relative abundance of each taxon with the specific gene is shown in Fig. 5. Of the different metabolic KO groups, the most abundant genes were associated with photosynthesis, carbohydrate synthesis and degradation, as well as nitrogen and sulfur metabolism. Within the photosynthesis KO, the most abundant reads shared sequence similarity to genes associated with photosystem I and II, chlorophyll and accessory pigment synthesis, and the electron acceptor ferredoxin (Fig. 5). Most of these genes were assigned to the cyanobacterial order *Chroococcales*, with a few genes associated with the filamentous order *Oscillatoriales*. High rates of photosynthesis have been previously shown to increase the localized alkalinity by depleting CO_2 resulting in disassociation of bicarbonate into CO_2 and OH^-(Dupraz *et al.*, 2009). The role of photosynthesis, as a major driver in the precipitation of carbonate through changes in the carbonate alkalinity, has been

well documented (Chafetz, 1986; Paerl *et al.*, 2001; Dupraz and Visscher, 2005; Dupraz *et al.*, 2009).

Another group of highly represented genes within the stromatolite metagenome were associated with the synthesis and degradation of sugars. Several genes associated with xylose fucose, mannose metabolism, as well as genes associated with the production of alginate (e.g. phosphomannomutase), were highly represented within the stromatolite metagenome. The taxa associated with these genes varied but were primarily derived from the *Alphaproteobacteria* orders *Rhodobacteriales*; however, some genes associated with fucose synthesis were associated with the cyanobacterial orders *Chroococcales* and *Nostocales*. Previous studies have characterized the EPS within Highborne Cay stromatolites and have shown that more than 50% of the material is comprised of carbohydrates such as galactose, xylose and fucose (Kawaguchi and Decho, 2000). Additionally, lithified stromatolite forming mats (i.e. Type 3 mats) of Highborne Cay, exhibited a high propensity to utilize hexose (e.g. galactose, mannose), deoxy sugars (fucose) and acidic sugars (e.g. ketoglutaric acid) (Khodadad and Foster, 2012). Together, these results suggest that Bahamian stromatolites have an increased metabolic capacity for the heterotrophic degradation and rearrangement of EPS material, which has the potential to alter the Ca^{2+} binding affinity of the EPS, thereby promoting carbonate precipitation (Dupraz *et al.*, 2009).

The predominant genes associated with nitrogen metabolism were nitrogenases primarily associated with the *Cyanobacteria*, including heterocystous forming *Nostocales* and the non-heterocystous *Chroococcales* with similarity to *Cyanothece* and *Synechococcus.* However, genes were also recovered from numerous other taxa including the *Firmicutes*, *Proteobacteria* and *Acidobacteria*. Similar results were observed in the metagenome and metatranscriptome of the thrombolites of Highborne Cay, where there was an enrichment of *nif* genes expressed at midday from a diversity group of microbes (Mobberley *et al.*, 2015) as well as from *nifH* surveys of Mexican microbialites in Cuatro Ciénegas, Lake Alchichica and the Bacalar costal lagoon (Falcón *et al.*, 2007; Beltrán *et al.*, 2012). However, in addition to nitrogen fixing genes, there was also a high representation of genes in the stromatolite metagenome associated with denitrification, including reductases for nitrate, nitrite, nitric oxide and nitrous oxide from a wide range of taxa. Although primarily associated with the alphaproteobacteria orders *Rhodobacteriales* and *Rhizobiales*, a high relative abundance of denitrification genes were also recovered from other *Proteobacteria* (*Beta-*, *Delta-* and *Gammaproteobacteria*) as well as the cyanobacterial order *Chroococcales*. The dissimilatory reduction of nitrate to N_2 has the potential to result in a net loss of carbonate thereby having a negative impact on carbonate precipitation within the stromatolites (Visscher and Stolz, 2005).

Although sulfur metabolism was not a dominant level 3 KO pathway in either of the shotgun or 16S rRNA gene PICRUSt predictive metagenome (Fig. 4), genes encoding the three key enzymes associated with dissimilatory sulfate reduction and sulfide oxidation were represented within the stromatolite metagenome. Specifically, genes encoding sulfate adenylytransferase (ATP sulfurylase), adenylysulfate reductase (APS reductase) and sulfite reductase were found primarily in the *Deltaproteobacteria*, *Gammaproteobacteria* and *Chromatiales* within the Little Darby stromatolites. Sulfate-reducing bacteria are a major

functional group in microbial mats systems and have been identified as having a major role in the precipitation of calcium carbonate (Visscher *et al.*, 1998; 2000; Dupraz *et al.*, 2009), although more recent studies have shown that the net precipitation potential of sulfate reduction is highly dependent on the type of electron donor (Gallagher *et al.*, 2012). Together, these results further confirm that lithifying microbial ecosystems, such as the Little Darby stromatolites, utilize a diverse network of microbes and functional genes to aid in the influx and cycling of nutrients within the stromatolites. Further, through the coordinated activities of these microbial metabolisms, the geochemical microenvironment within the stromatolites can be altered, thereby potentially generating conditions (e.g. changes in localized pH) that either promote the precipitation or dissolution of carbonate (Visscher and Stolz, 2005; Dupraz *et al.*, 2009).

Comparison of Bahamian stromatolites to the global microbialite population from varying environments

In addition to profiling the taxa and functional complexity of the Little Darby stromatolite microbiome, we compared this system with other carbonate-based microbialites throughout the globe. Publically available data sets that used next generation sequencing (454 and Illumina platforms) for 16S rRNA gene analysis within lithifying microbial mat ecosystems were mined for the comparison (Table 1). As stromatolites are highly novel communities, a subsampled open-reference OTU picking approach in QIIME to be inclusive of unique taxa was needed (Rideout *et al.*, 2014); therefore, we targeted only 16S rRNA gene libraries that had overlapping amplicons (V2-V4 regions). Unfortunately, this approach prohibited the inclusion of several prominent microbialite studies, such as the 16S rRNA gene analyses of microbialites collected throughout Mexico (Cuatros Ciénegas, Alchichica, Sian Ka'an and Bacalar; Centeno *et al.*, 2012), which targeted the V5–6 region. Additionally, several earlier studies that used clone libraries to survey stromatolites from across the globe (e.g. Papineau *et al.*, 2005; Allen *et al.*, 2009; Baumgartner et al., 2009b; Foster *et al.*, 2009; Goh *et al.*, 2009; Santos *et al.*, 2010; Foster and Green, 2011) were not included due to their limited sequencing depth.

A total of 10 different study sites were included in the comparison for a total of 171 samples (Table 1). The sites included a wide range of microbialite habitats including several freshwater (Pavilion Lake, British Columbia; South African estuaries), marine (Highborne Cay and Little Darby Island, The Bahamas), hypersaline (Hamelin Pool, Shark Bay, Western Australia; Socompa Lake, Argentina; Storr's Lake, San Salvador, The Bahamas) and ultrahypersaline (Kiritimati Atoll, Kiribati) locations (Lim *et al.*, 2009; Mobberley *et al.*, 2012; Dupraz *et al.*, 2013; Farías *et al.*, 2013; Schneider *et al.*, 2013; Perissinotto *et al.*, 2014; Russell *et al.*, 2014; Paul *et al.*, 2015). Two non-lithifying microbial mat sites were also included to assess whether there where major differences between mat communities that undergo lithification; the hydrothermal vent mats from Santorini Crater in the Mediterranean Sea and the hypersaline microbial mats of Guerrero Negro (Table 1). All the sequences from each data set were merged, normalized and analysed in QIIME (see *Experimental procedures*).

The alpha diversity of each community was assessed using rarefaction curves derived from both the Shannon Diversity and the Faith's Phylogenetic Diversity Index (Fig. S5). The results indicated that the highest levels of microbial diversity were within the hypersaline stromatolites of Shark Bay Australia, whereas the lowest levels of diversity were observed in the microbialites of the Swartkops estuaries and the non-lithifying hydrothermal vent communities of Santorini Caldera, which may reflect the lower sample number and sequencing depth of the two sites. A beta diversity analysis on the different communities was completed using Principal Coordinate Analysis (PCoA) of the unweighted UniFrac rarefied distance matrix (Fig. 6). Some distinctive clustering was apparent in the PCoA plot indicating that geography does plays a role in driving some of the microbial diversity within lithifying microbial mat ecosystems [Fig. 6A; $P = 0.001$; R = 0.93, analysis of similarity (ANOSIM)]. The most isolated system appeared to be the Hamelin Pool stromatolites located in Shark Bay Western Australia. Hamelin Pool is home to the most extensive and diverse communities of stromatolites in the world and is thought to have been isolated from the Shark Bay system approximately 6000 years ago (Logan *et al.*, 1974; Playford and Cockbain, 1976). The extended geographic isolation of these stromatolites may be a major factor in distinctive clustering of the Hamelin Pool community from the other targeted habitats.

Another environmental factor that appeared to drive differences in the microbial diversity between habitats was salinity (Fig. 6B). Salinity has been long known to be a contributor to driving differences in diversity within terrestrial and aquatic-based microbial communities (e.g. Pinckney *et al.*, 1995a; Lozupone and Knight, 2007; Dupont *et al.*, 2014). Although the correlation was not as strong as geography, differences between the populations did appear to be linked to the salinity of the surrounding environment ($P = 0.001$; R = 0.59 ANOSIM). For example, the two most distant microbialite communities were the ultrahypersaline Kiritimati Atoll (up to 170 ppt) and freshwater Pavilion Lake (<1 ppt) samples (Fig. 6B). Samples from the marine (Little Darby Island and Highborne Cay; 35–38 ppt) and moderately hypersaline (Guerrero Negro, Storr's Lake; Socampo Lake; 66–100 ppt) sites clustered together. Interestingly, sites such as the Swartkops Estuary in South Africa (5–35 ppt) and Storr's Lake, San Salvador, The Bahamas (26–> 90 ppt), which undergo extensive transitions in their salinity throughout the year clustered with the open marine Bahamian microbialites of Little Darby and Highborne Cay (Perissinotto *et al.*, 2014; Paul *et al.*, 2015). Although metagenomic analyses do not currently exist for most of these sites, including the Swartkops Estuary and Storr's Lake, the results do suggest that the communities may harbour key taxa or clades that may have a large variation in gene content and metabolic capabilities, thereby enabling a quick response to the regular freshening events that occur in these habitats. Similar results have been seen along salinity gradients in the Baltic Sea where the complexity of pangenome of some taxa may enable organisms to thrive under a wide range of salinities (Dupont *et al.*, 2014).

Although geography and salinity appear to be important factors driving differences between the communities, the overall variation was low (Fig. 6; PC1 17%; PC2 5%; PC3 4%), suggesting that there were extensive taxonomic similarities within the different lithifying and non-lithifying mat systems. For example, the non-lithifying microbial mats from Santorini Crater and Guerrero Negro clustered closely with several lithifying communities,

including all the Bahamian sites, Socompa Lake and the Swartkops Estuary stromatolites. These results suggest that the taxa themselves are not the distinguishing characteristic of why some microbial mat systems undergo lithification and others do not, suggesting that 'who' in the community is not as critical as their metabolic capabilities or interactions with their environment to promote microbialite formation.

Conclusions

Taken together, the results of this study expand our understanding of the stromatolite microbiome by characterizing not only the taxa associated with these systems but also the full range of metabolic capabilities. Metagenomic sequencing of the Little Darby stromatolites revealed the relative abundances and gene products associated with several key phototrophic (e.g. photosynthesis, EPS production) and heterotrophic (e.g. sulfate reduction, EPS degradation) metabolisms typically associated with carbonate precipitation. Carbonate precipitation in stromatolites relies on the coupling of both phototrophic and heterotrophic metabolic processes, and it is the balance between these processes that determines the net precipitation potential. Additionally, the direct comparison of whole shotgun sequencing with 16S rRNA gene-based analyses demonstrated the effectiveness of using predictive metagenomic analyses in environmental ecosystems, such as stromatolites, to help provide a first look at the functional capabilities of the community. However, these metagenomic approaches have also revealed pronounced gaps in the database for key taxa, such as *Cyanobacteria*, indicating the importance of continuing to refine and expand the catalogue of reference genomes from environmental systems. Lastly, a comparison of the Little Darby Island stromatolites with both lithifying and non-lithifying microbial ecosystems worldwide revealed that although geographic location and salinity do drive some differences within the microbial communities, the targeted ecosystems were highly similar. These results suggest that the taxonomic composition or 'who' is in the community is not as important as their functional capabilities and their interactions with their environment. As metagenomic and metatranscriptomic analyses of stromatolites become increasingly available, it will become critical to provide more comprehensive functional comparisons to begin to more fully characterize the shared microbiome of modern stromatolites and assess the mechanisms and interactions that lead to carbonate precipitation in microbial ecosystems.

Experimental procedures

Sample collection and DNA extraction

Samples of stromatolites were collected from waters around Little Darby Island located in the southern Exuma Cays, The Bahamas (76°49′ W, 115 24°43′N) in August 2012. The upper 1 cm of living stromatolite-forming microbial mat material was collected from seven different stromatolite heads and immediately placed in RNA later (Life Technologies, Grand Island, NY). The samples were transported to the Space Life Sciences lab and stored at −80°C until processed. Each stromatolite sample was vertically sectioned and genomic DNA was extracted using a modified xanthogenate method as previously described and used for either 16S rRNA gene amplification or metagenomic analysis (Foster *et al.*, 2009).

16S rRNA gene amplicon library construction

To generate the amplicon libraries, DNA was PCR amplified using a bacterial domain-specific primers that targeted the V1–2 region of the 16S rRNA gene (27F and 338R; Lane, 1991; Liu *et al.*, 2007). The PCR reactions were conducted in triplicate for each sample and contained the following final concentrations: 1 x *Pfu* reaction buffer (Stratagene, La Jolla, CA), 280 μM dNTPs, 2.5 μg of BSA, 600 nM each primer, 0.75 ng of genomic DNA, 1.25 U of *Pfu* DNA polymerase (Stratagene) and nuclease free water (Sigma, St Louis, MO) in a volume of 25 μl. The reactions were initially denatured at 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, 64°C for 1 min, 72°C for 1 min and a final extension at 72°C for 7 min. Negative controls were conducted for all reactions. PCR products were gel extracted and concentrated using an UltraClean Purification kit (MoBio, Carlsbad, CA), and replicate amplicons were pooled in equimolar ratios. A subset of genomic DNA consisting of a minimum of three extractions from each head was pooled and concentrated using a DNA Clean & Concentrator-25 kit (Zymo Research, Irvine, CA). The pooled stromatolite DNA was spiked into the amplicon libraries and sequenced using the Illumina MiSeq platform generating two libraries of paired-end reads (2 × 250 bp).

Bioinformatics analyses of 16S rRNA amplicon libraries

The 16S rRNA amplicon libraries were analysed using the Quantitative Insights Into Microbial Ecology tool (QIMME v. 1.8; Caporaso *et al.*, 2010). A pre-quality filtering step was conducted using the default parameters in QIMME. For a sequence to be retained the following criteria had to be met: (i) a minimum average quality Phred score of 25; (ii) a minimum and maximum sequence length (200–1000, 454 data set); (iii) no more than six ambiguous bases (N) or homopolymers.

In addition to the amplicon libraries generated in this study, the NCBI database was mined for all 16S rRNA gene sequences derived from microbialite (e.g. stromatolites and thrombolites) ecosystems generated using next-generation sequencing (Table 1). Two additional non-lithifying microbial mat systems were included as outgroups (Table 1). A total of 171 samples were recovered from short read archive (SRA) database and merged with the data sets from this study and analysed in QIMME using a subsampled open-reference OTU picking approach. The OTU picking procedure was computed with UCLUST (at 97% identity) using the 16S rRNA Greengenes database (v. 13_8) as a reference database (DeSantis *et al.*, 2006; Edgar, 2010). Sequences that did not match the Greengenes database were clustered as *de novo* as to not lose the overall novel diversity. Taxonomic assignment was also computed using UCLUST with a 0.9% similarity against a representative set of 16S rRNA gene sequences from the Greengenes database, obtaining a final Biological Observation Matrix (McDonald *et al.*, 2012). The observations are reported as OTUs and the matrix contained counts corresponding to the number of times each OTU was observed in each sample. For visualization purposes, the taxonomic composition data were transported in Krona (Ondov *et al.*, 2011) showing only the most abundant taxa (>0.1%).

The alpha and beta diversity analyses were computed in QIIME using the script core_diversity_analyses.py, at a rarefaction depth of 450 sequences/sample. Alpha diversity was computed for each rarefied OTU table, using the Shannon Diversity index and the

Faith's Phylogenetic Diversity Index. Beta diversity was performed using unweighted UniFrac distance matrices and plotted as principal coordinate analysis (PCoA) in Emperor (Vazquez-Baeza *et al.*, 2013) in QIIME. A nonparametric test and the ANOSIM (Clarke, 1993) method were used to determine the statistical significance related to the diversity analyses.

Metabolic profile prediction of Little Darby Island with 16S rRNA gene analysis

To predict the functional gene content of the 16S rRNA gene amplicon libraries, PICRUSt v. 1 .0 (Langille *et al.*, 2013) was used, and the results were then compared with actual metagenomic data as described below. An OTU table generated in QIIME was normalized by dividing each OTU by the predicted 16S rRNA gene copy number abundance using the script *normalize_by_copy_number.py*. The metagenome functional prediction was then computed by multiplying each normalized OTU abundance by each predicted functional trait abundance, producing a table of KO predictions using the script *predict_metagenomes.py*. The resulting table was collapsed at KO level 3 within the pathway hierarchy of KEGG using the script *categorize_by_function.py*.

Metagenome sequencing and analysis

Genomic DNA derived from the Little Darby stromatolites was sequenced using the Illumina MiSeq platform and the paired-end reads were quality filtered considering the Phred quality scores and read length, using SICKLE (v. 1.2; Joshi and Fass, 2011) with default parameters. High-quality filtered reads were assigned for functionality at different KEGG levels, using MetaCV v. 2.3.0 (Liu *et al.*, 2013) with default parameters. Metagenome Composition Vector classifies short metagenomic reads into specific taxonomic and functional groups using a composition and phylogeny-based algorithm. The reads were not assembled due to the variable relative abundance of different taxa within the stromatolite-forming mats, which may result in potential chimeric contigs (Liu *et al.*, 2013; Casaburi, 2015; Howe and Chain, 2015). The output was loaded into the statistical program R and initially filtered at a correlation score > 50 to retain only the top matched genes and then filtered for abundance and for different KEGG classification levels. Taxonomic information was extracted using MEtaGenome ANalyzer V.5.6.5 to highlight phylogenetic relationships within the stromatolite community (Huson *et al.*, 2007). To compare the predicted metabolic profile with the 16S rRNA gene sequences with the metagenome data set a Pearson Correlation was performed correlating the KEGG entries identified at level 3 of KO classification.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Christina Khodadad and Artemis Louyakis for their technical assistance on the project as well as the owners of Little Darby Island for access to the study site. Permits were provided by The Bahamas Ministry of Agriculture, Marine Resources and Local Government, and logistical support was provided by the Little Darby Island Research Station. The project was supported by the NASA Astrobiology: Exobiology and Evolutionary Biology Program Element (NNX14AK14G) and the Florida Space Grant Consortium.

References

Allen MA, Goh F, Burns BP, Neilan BA. Bacterial, archaeal and eukaryotic diversity of smooth and pustular microbial mat communities in the hypersaline lagoon of Shark Bay. Geobiology. 2009; 7:82–96. [PubMed: 19200148]

Anderson I, Ulrich LE, Lupa B, Susanti D, Porat I, Hooper SD, et al. Genomic characterization of methanomicrobiales reveals three classes of methanogens. PLoS ONE. 2009; 4:e5797. [PubMed: 19495416]

Andrei AS, Banciu HL, Oren A. Living with salt: metabolic and phylogenetic diversity of archaea inhabiting saline ecosystems. FEMS Microbiol Lett. 2012; 330:1–9. [PubMed: 22339687]

Awramik, SA. Ancient stromatolites and microbial mats. In: Cohen, Y.Castenholz, RW., Halvorson, HO., editors. Microbial Mats: Stromatolites. New York: Alan R. Liss; 1984. p. 1-22.

Baumgartner LK, Spear JR, Buckley DH, Pace NR, Reid RP, Visscher PT. Microbial diversity in modern marine stromatolites, Highborne Cay, Bahamas. Environ Microbiol. 2009a; 11:2710–2719. [PubMed: 19601956]

Baumgartner LK, Dupraz C, Buckley DH, Spear JR, Pace NR, Visscher PT. Microbial species richness and metabolic activities in hypersaline microbial mats: insight into biosignature formation through lithification. Astrobiology. 2009b; 9:861–874. [PubMed: 19968463]

Beltrán; Y, Centeno CM, Garcia-Oliva F, Legendre P, Falcón LI. N2 fixation rates and associated diversity (*nifH*) of microbialite and mat-forming consortia from different aquatic environments in Mexico. Aquat Microb Ecol. 2012; 67:15–24.

Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat Methods. 2013; 10:57–59. [PubMed: 23202435]

Bowlin EM, Klaus J, Foster JS, Andres M, Custals L, Reid RP. Environmental controls on microbial community cycling in modern marine stromatolites. Sediment Geol. 2012; 263–264:45–55.

Breitbart M, Hoare A, Nitti A, Siefert J, Haynes M, Dinsdale E, et al. Metagenomic and stable isotopic analyses of modern freshwater microbialites in Cuatro Ciénegas, Mexico. Environ Microbiol. 2009; 11:16–34. [PubMed: 18764874]

Burns BP, Goh F, Allen M, Neilan BA. Microbial diversity of extant stromatolites in the hypersaline marine environment of Shark Bay, Australia. Environ Microbiol. 2004; 6:1096–1101. [PubMed: 15344935]

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010; 7:335–336. [PubMed: 20383131]

Casaburi, G. Bioinformatics Approaches for the Study of the Human Gut Microbiome. Saarbrücken, Germany: Edizioni Accademiche Italiane; 2015.

Centeno CM, Legendre P, Beltran Y, Alcantara-Hernandez RJ, Lidstrom UE, Ashby MN, Falcon LI. Microbialite genetic diversity and composition relate to environmental variables. FEMS Microbiol Ecol. 2012; 82:724–735. [PubMed: 22775797]

Chafetz HS. Marine peloids – a product of bacterially induced precipitation of calcite. J Sediment Petrol. 1986; 56:812–817.

Clarke KR. Nonparamentric multivariate analyses of changes in community structure. Aust J Ecol. 1993; 18:117–143.

Collins RE, Higgs PG. Testing the infinitely many genes model for the evolution of the bacterial core genome and pangenome. Mol Biol Evol. 2012; 29:3413–3425. [PubMed: 22752048]

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006; 72:5069–5072. [PubMed: 16820507]

Dill RF, Shinn EA, Jones AT, Kelly K, Steinen RP. Giant subtidal stromatolites forming in normal salinity water. Nature. 1986; 324:55–58.

Dravis JJ. Hardened subtidal stromatolites, Bahamas. Science. 1983; 219:385–386. [PubMed: 17815316]

Dupont CL, Larsson J, Yooseph S, Ininbergs K, Goll J, Asplund-Samuelsson J, et al. Functional tradeoffs underpin salinity-driven divergence in microbial community composition. PLoS ONE. 2014; 9:e89549. [PubMed: 24586863]

Dupraz C, Visscher PT. Microbial lithification in marine stromatolites and hypersaline mats. Trends Microbiol. 2005; 13:429–438. [PubMed: 16087339]

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.

Dupraz C, Fowler A, Tobias C, Visscher PT. Stromatolitic knobs in Storr's Lake (San Salvador, Bahamas): a model system for formation and alteration of laminae. Geobiology. 2013; 11:527–548. [PubMed: 24118887]

Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010; 26:2460–2461. [PubMed: 20709691]

Edgcomb VP, Bernhard JM, Summons RE, Orsi W, Beaudoin D, Visscher PT. Active eukaryotes in microbialites from Highborne Cay, Bahamas, and Hamelin Pool (Shark Bay), Australia. ISME J. 2014; 8:418–429. [PubMed: 23924782]

Falcón LI, Cerritos R, Eguiarte LE, Souza V. Nitrogen fixation in microbial mat and stromatolite communities from Cuatro Cienegas, Mexico. Microb Ecol. 2007; 54:363–373. [PubMed: 17450393]

Farías ME, Rascovan N, Toneatti DM, Albarracin VH, Flores MR, Poire DG, et al. The discovery of stromatolites developing at 3570 m above sea level in a high-altitude volcanic lake Socompa, Argentinean Andes. PLoS ONE. 2013; 8:e53497. [PubMed: 23308236]

Feldmann M, McKenzie JA. Stromatolite-thrombolite associations in a modern environment, Lee Stocking Island, Bahamas. Palaios. 1998; 13:201–212.

Foster, JS., Green, SJ. Microbial diversity in modern stromatolites. In: Seckbach, J., Tewari, V., editors. Cellular Origin, Life in Extreme Habitats and Astrobiology: Interactions with Sediments. Netherlands: Springer; 2011. p. 385-405.

Foster JS, Green SJ, Ahrendt SR, Hetherington KL, Golubic S, Reid RP, Bebout L. Molecular and morphological characterization of cyanobacterial diversity in the marine stromatolites of Highborne Cay, Bahamas. ISME J. 2009; 3:573–587. [PubMed: 19148145]

Fröls S, Dyall-Smith M, Pfeifer F. Biofilm formation by haloarchaea. Environ Microbiol. 2012; 14:3159–3174. [PubMed: 23057712]

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. [PubMed: 22925453]

Gilbert JA, Field D, Swift P, Thomas S, Cummings D, Temperton B, et al. The taxonomic and functional diversity of microbes at a temperate coastal site: a 'multi-omic' study of seasonal and diel temporal variation. PLoS ONE. 2010; 5:e15545. [PubMed: 21124740]

Goh F, Leuko S, Allen MA, Bowman JP, Kamekura M, Neilan BA, Burns BP. Halococcus hamelinensis sp. nov., a novel halophilic archaeon isolated from stromatolites in Shark Bay, Australia. International Journal of Systematic and Evolutionary Microbiology. 2006; 56(6):1323–1329. [PubMed: 16738110]

Goh F, Allen MA, Leuko S, Kawaguchi T, Decho AW, Burns BP, Neilan BA. Determining the specific microbial populations and their spatial distribution within the stromatolite ecosystem of Shark Bay. ISME J. 2009; 3:383–396. [PubMed: 19092864]

Green, PN. The prokaryotes. Springer; New York: 2006. Methylobacterium; p. 257-265.

Grotzinger JP, Knoll AH. Stromatolites in Precambrian carbonates: evolutionary mileposts or environmental dipsticks? Annu Rev Earth Planet Sci. 1999; 27:313–358. [PubMed: 11543060]

Hartmann R, Sickinger HD, Oesterhelt D. Anaerobic growth of halobacteria. Proc Natl Acad Sci USA. 1980; 77:3821–3825. [PubMed: 6933439]

He Y, Caporaso JG, Jiang XT, Sheng HF, Huse SM, Rideout JR, et al. Stability of operational taxonomic units: an important but neglected property for analyzing microbial diversity. Microbiome. 2015; 3:20. [PubMed: 25995836]

Howe A, Chain PS. Challenges and opportunities in understanding microbial communities with metagenome assembly (accompanied by IPython Notebook tutorial). Front Microbiol. 2015; 6:678. [PubMed: 26217314]

Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. Genome Res. 2007; 17:377–386. [PubMed: 17255551]

Joshi, NA., Fass, JN. Software. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.21).

Kawaguchi T, Decho AW. Biochemical characterization of cyanobacterial extracellular polymers (EPS) from modern marine stromatolites (Bahamas). Prep Biochem Biotech. 2000; 30:321–330.

Kawaguchi T, Decho AW. Isolation and biochemical characterization of extracellular polymeric secretions (EPS) from modern soft marine stromatolites (Bahamas) and its inhibitory effect on $CaCO_3$ precipitation. Prep Biochem Biotech. 2002; 32:51–63.

Khodadad CL, Foster JS. Metagenomic and metabolic profiling of nonlithifying and lithifying stromatolitic mats of Highborne Cay, The Bahamas. PLoS ONE. 2012; 7:e38229. [PubMed: 22662280]

Lane, DJ. 16S/23S rRNA sequencing. In: Stackebrandt, E., Goodfellow, M., editors. Nucleic Acid Techniques in Bacterial Systematics. Chichester, England: Wiley; 1991. p. 115-175.

Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol. 2013; 31:814–821. [PubMed: 23975157]

Laval B, Cady SL, Pollack JC, McKay CP, Bird JS, Grotzinger JP, et al. Modern freshwater microbialite analogues for ancient dendritic reef structures. Nature. 2000; 407:626–629. [PubMed: 11034210]

Leuko S, Goh F, Allen MA, Burns BP, Walter MR, Neilan BA. Analysis of intergenic spacer region length polymorphisms to investigate the halophilic archaeal diversity of stromatolites and microbial mats. Extremophiles. 2007; 11:203–210. [PubMed: 17082971]

Lim DS, Laval BE, Slater G, Antoniades D, Forrest AL, Pike W, et al. Limnology of Pavilion Lake, BC, Canada – characterization of a microbialite forming environment. Fund Appl Limnol. 2009; 173:329–351.

Liu J, Wang H, Yang H, Zhang Y, Wang J, Zhao F, Qi J. Composition-based classification of short metagenomic sequences elucidates the landscapes of taxonomic and functional enrichment of microorganisms. Nucleic Acids Res. 2013; 41:e3. [PubMed: 22941634]

Liu Z, Lozupone C, Hamady M, Bushman FD, Knight R. Short pyrosequencing reads suffice for accurate microbial community analysis. Nucleic Acids Res. 2007; 35:e120. [PubMed: 17881377]

Logan, BW., Hoffman, P., Gebelein, CD. Alga mats, cryptalgal fabrics and structures, Hamelin Pool, Western Australia. In: Logan, BW., editor. Evolution and Diagenesis of Quaternary Carbonate Sequences, Shark Bay, Western Australia. Tulsa, Oklahoma: American Association of Petrology, Geology and Mineralogy; 1974. p. 140-194.

Lozupone CA, Knight R. Global patterns in bacterial diversity. Proc Natl Acad Sci USA. 2007; 104:11436–11440. [PubMed: 17592124]

McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D, et al. The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. GigaScience. 2012; 1:7. [PubMed: 23587224]

Mobberley JM, Ortega MC, Foster JS. Comparative microbial diversity analyses of modern marine thrombolitic mats by barcoded pyrosequencing. Environ Microbiol. 2012; 14:82–100. [PubMed: 21658172]

Mobberley JM, Khodadad CL, Foster JS. Metabolic potential of lithifying cyanobacteria-dominated thrombolitic mats. Photosynth Res. 2013; 118:125–140. [PubMed: 23868401]

Mobberley JM, Khodadad CL, Visscher PT, Reid RP, Hagan P, Foster JS. Inner workings of thrombolites: spatial gradients of metabolic activity as revealed by metatranscriptome profiling. Sci Rep. 2015; 5:12601. [PubMed: 26213359]

Myshrall K, Mobberley JM, Green SJ, Visscher PT, Havemann SA, Reid RP, Foster JS. Biogeochemical cycling and microbial diversity in the modern marine thrombolites of Highborne Cay, Bahamas. Geobiology. 2010; 8:337–354. [PubMed: 20491947]

Ondov BD, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a web browser. BMC Bioinform. 2011; 12:385.

Paerl HW, Steppe TF, Reid RP. Bacterially mediated precipitation in marine stromatolites. Environ Microbiol. 2001; 3:123–130. [PubMed: 11321542]

Papineau D, Walker JJ, Mojzsis SJ, Pace NR. Composition and structure of microbial communities from stromatolites of Hamelin Pool in Shark Bay, Western Australia. Appl Environ Microbiol. 2005; 71:4822–4832. [PubMed: 16085880]

Paul VG, Wronkiewicz DJ, Mormile MR, Foster JS. Mineralogy and microbial diversity of the microbialites in the hypersaline Storr's Lake, The Bahamas. Astrobiology. 2015 in review.

Perissinotto R, Bornman TG, Steyn RP, Miranda NF, Dorrington RA, Matcher GF, et al. Tufa stromatolite ecosystems on the South African south coast. S Afr J Sci. 2014; 110:1–8.

Pinckney J, Paerl HW, Bebout BM. Salinity control of benthic microbial mat community production in a Bahamian hypersaline lagoon. J Exp Mar Biol Ecol. 1995a; 187:223–237.

Pinckney J, Paerl HW, Reid RP, Bebout BM. Ecophysiology of stromatolitic microbial mats, Lee Stocking Island, Exuma Cays, Bahamas. Microb Ecol. 1995b; 29:19–37. [PubMed: 24186636]

Playford, PE., Cockbain, AE. Modern algal stromatolites at Hamelin Pool, a hypersaline barred basin in Shark Bay, Western Australia. In: Walter, MR., editor. Stromatolites – Developments in Sedimentology. Amsterdam, the Netherlands: Elsevier; 1976. p. 389-411.

Poretsky R, Rodriguez RL, Luo C, Tsementzi D, Konstantinidis KT. Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. PLoS ONE. 2014; 9:e93827. [PubMed: 24714158]

Reid, P., Foster, JS., Radtke, G., Golubic, S. Modern marine stromatolites of Little Darby Island, Exuma Archipelago, Bahamas: environmental setting, accretion mechanisms and role of euendoliths. In: Reitner, J.Thrauth, MH.Stüwe, K., Yuen, D., editors. Advances in Stromatolite Geology. Berlin, Germany: Springer; 2010. p. 77-90.

Reid RP, Macintyre IG, Browne KM, Steneck RS, Miller T. Modern marine stromatolites in the Exuma Cays, Bahamas – uncommonly common. FACIES. 1995; 33:1–17.

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. [PubMed: 10984051]

Reid RP, James NP, Macintyre IG, Dupraz CP. Shark Bay stromatolites: microfabrics and reinterpretation of origins. FACIES. 2003; 49:299–324.

Rideout JR, He Y, Navas-Molina JA, Walters WA, Ursell LK, Gibbons SM, et al. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. PeerJ. 2014; 2:e545. [PubMed: 25177538]

Russell JA, Brady AL, Cardman Z, Slater GF, Lim DS, Biddle JF. Prokaryote populations of extant microbialites along a depth gradient in Pavilion Lake, British Columbia, Canada. Geobiology. 2014; 12:250–264. [PubMed: 24636451]

Ruvindy R, White RA III, Neilan BA, Burns BP. Unravelling core microbial metabolisms in the hypersaline microbial mats of Shark Bay using high-throughput metagenomics. ISME J. 2015 May 30.

Sakai S, Takaki Y, Shimamura S, Sekine M, Tajima T, Kosugi H, et al. Genome sequence of a mesophilic hydrogenotrophic methanogen *Methanocella paludicola*, the first cultivated representative of the order *Methanocellales*. PLoS ONE. 2011; 6:e22898. [PubMed: 21829548]

Santos F, Pena A, Nogales B, Soria-Soria E, Del Cura MA, Gonzalez-Martin JA, Anton J. Bacterial diversity in dry modern freshwater stromatolites from Ruidera Pools Natural Park, Spain. Syst Appl Microbiol. 2010; 33:209–221. [PubMed: 20409657]

Schneider D, Arp G, Reimer A, Reitner J, Daniel R. Phylogenetic analysis of a microbialite-forming microbial mat from a hypersaline lake of the Kiritimati atoll, Central Pacific. PLoS ONE. 2013; 8:e66662. [PubMed: 23762495]

Steppe TF, Pinckney JL, Dyble J, Paerl HW. Diazotrophy in modern marine Bahamian stromatolites. Microb Ecol. 2001; 41:36–44. [PubMed: 11252162]

Stolz JF, Reid RP, Visscher PT, Decho AW, Norman RS, Aspden RJ, et al. The microbial communities of the modern marine stromatolites at Highborne Cay, Bahamas. Atoll Res Bull. 2009; 567:1–29.

Valdespino-Castillo PM, Alcantara-Hernandez RJ, Alcocer J, Merino-Ibarra M, Macek M, Falcon LI. Alkaline phosphatases in microbialites and bacterioplankton from Alchichica soda lake, Mexico. FEMS Microbiol Ecol. 2014; 90:504–519. [PubMed: 25112496]

Vazquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. EMPeror: a tool for visualizing high-throughput microbial community data. GigaScience. 2013; 2:16. [PubMed: 24280061]

Visscher PT, Stolz JF. Microbial mats as bioreactors: populations, processes and products. Palaeogeogr Palaeoclimatol Palaeoecol. 2005; 219:87–100.

Visscher PT, Reid RP, Bebout BM, Hoeft SE, Macintyre IG, Thompson JA. Formation of lithified micritic laminae in modern marine stromatolites (Bahamas): the role of sulfur cycling. Amer Mineral. 1998; 83:1482–1493.

Visscher PT, Reid RP, Bebout BM. Microscale observations of sulfate reduction: correlation of microbial activity with lithified micritic laminae in modern marine stromatolites. Geology. 2000; 28:919–922.

Woese CR, Winker S, Gutell RR. Architecture of ribosomal RNA: constraints on the sequence of ‘tetra-loops. Proc Natl Acad Sci USA. 1990; 87:8467–8471. [PubMed: 2236056]

Yousef M, Ketany M, Manevitz L, Showe LC, Showe MK. Classification and biomarker identification using gene network modules and support vector machines. BMC Bioinform. 2009; 10:337.

Zaneveld JR, Lozupone C, Gordon JI, Knight R. Ribosomal RNA diversity predicts genome diversity in gut bacteria and their relatives. Nucleic Acids Res. 2010; 38:3869–3879. [PubMed: 20197316]

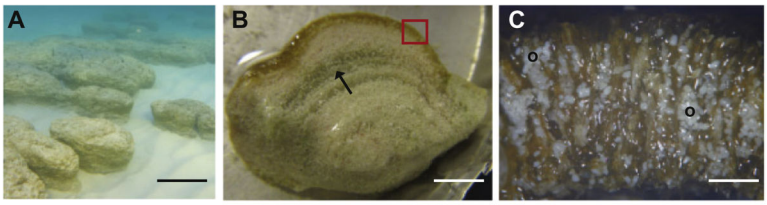

Fig. 1.

Overview of stromatolites of Little Darby Island, The Bahamas.

A. Stromatolites located in the shallow subtidal zone. Bar, 50 cm.

B. Cross-section of microbial mat from the surface of stromatolites. Red box represents area visualized in C. Arrow points to subsurface green layer of enriched cyanobacteria. Bar, 5 mm.

C. Cyanobacteria-rich area with extensive exopolymeric substances and trapped white oolitic sand grains (o). Bar, 500 μm.

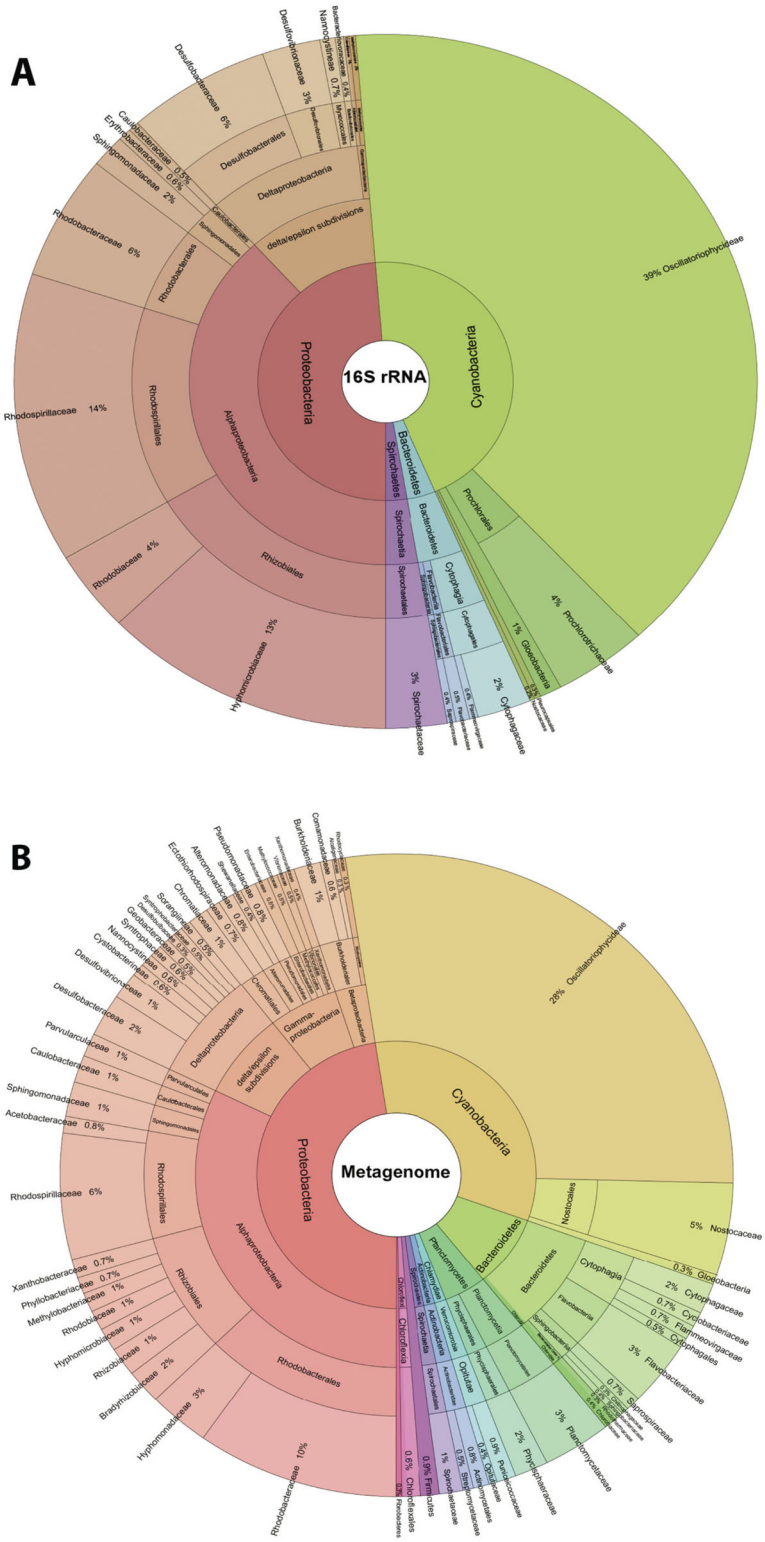

Fig. 2.

Taxonomic abundance and diversity of Little Darby stromatolites. Krona plot visualizing the taxonomic hierarchies of (A) the 16S rRNA gene analysis performed with QIIME and (B)

whole shotgun metagenome sequencing analysed with metaCV. The inner ring represents the different phyla associated with the mats, and as the plot progresses outwards, there is an increasing taxonomic resolution for each ring (i.e. class, order, family respectively). Sequences have been filtered and only those OTUs with relative abundance greater than 0.1% are reported.

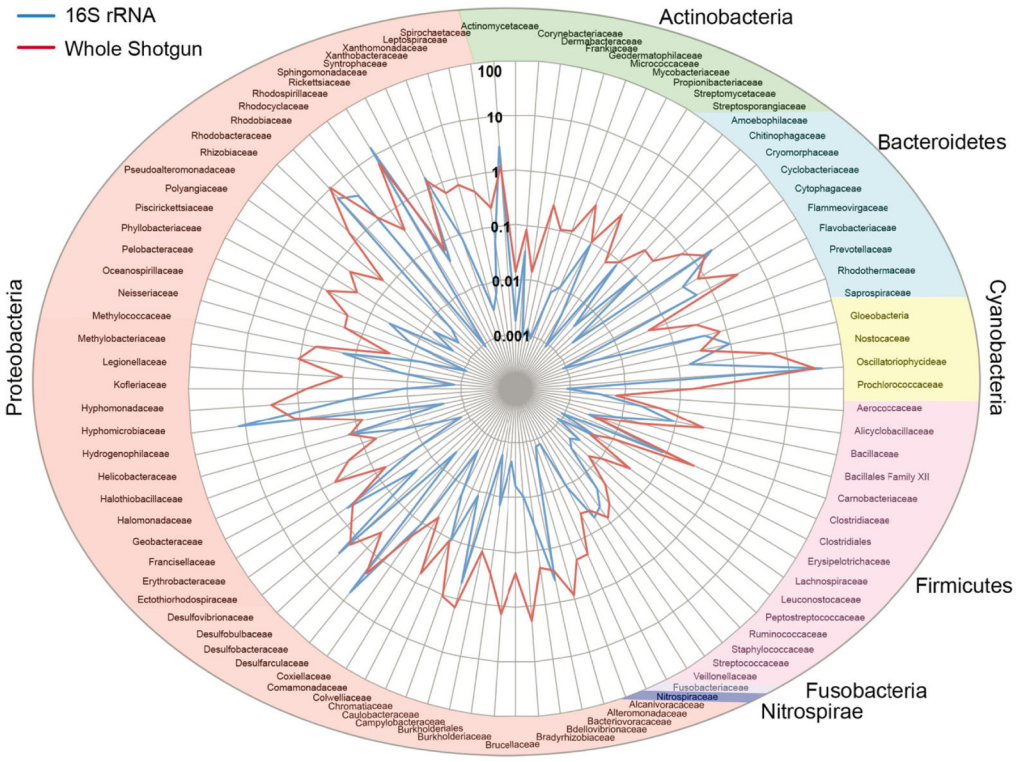

Fig. 3.

Radar plot depicting dominant bacterial families shared between 16S and metagenomic data sets. Relative abundance of bacterial families in 16S rRNA gene (blue line) and whole shotgun (red line) metagenome libraries expressed as a percentage on a log scale. Colours denote families within different phyla.

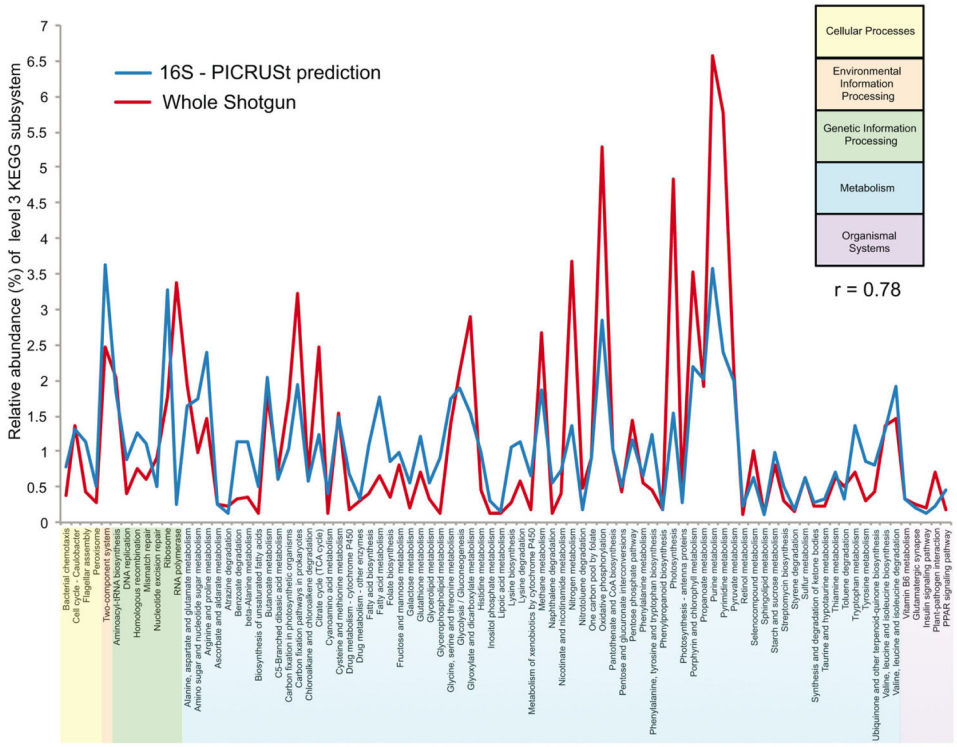

Fig. 4.

Functional comparison of Little Darby stromatolite metagenome from 16S rRNA metabolic prediction and whole shotgun sequencing. Comparison of relative abundances (%) of level 3 KEGG pathways from shotgun metagenomic sequences and 16S rRNA metabolic prediction (PICRUSt). Gene pathways within the same colour belong to the same level 1 of KEGG orthology classification. Pearson correlation value (r) is shown to estimate functional similarities derived from the two data sets.

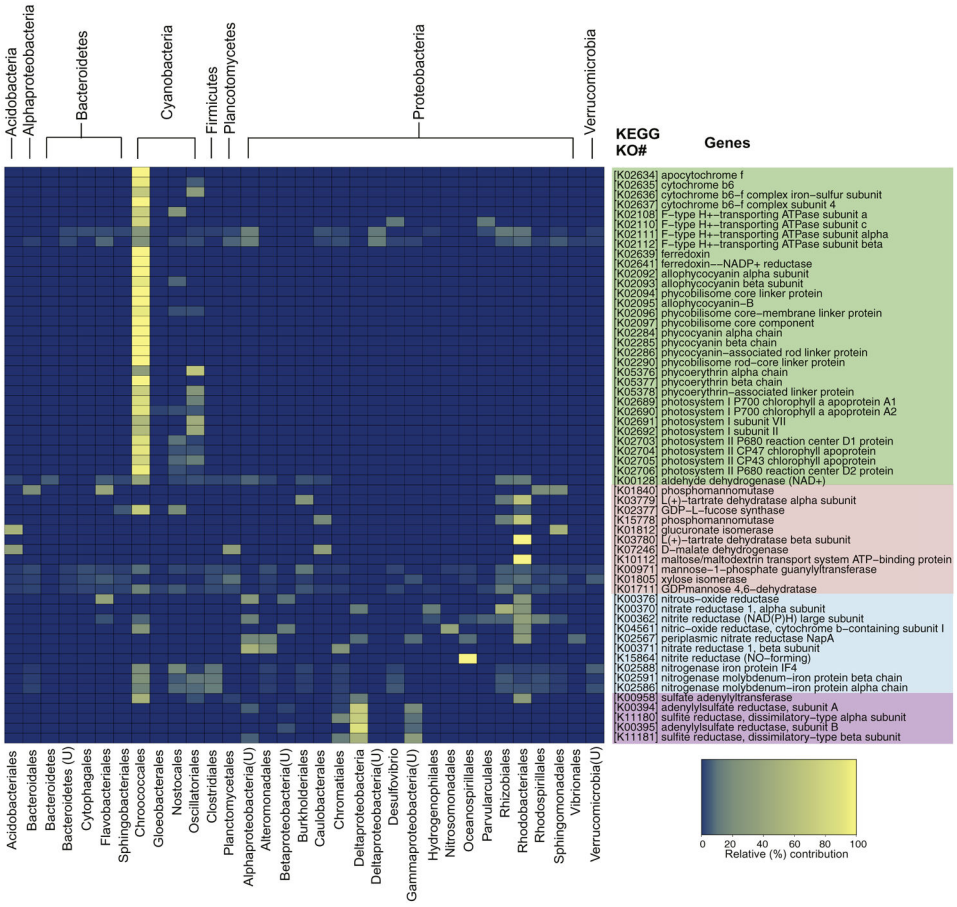

Fig. 5.

Relative contribution bacterial taxa to selected gene pathways. Heat map showing the relative contribution (%) of bacterial orders to selected metabolic pathways expressed as KO entries and associated products derived from metagenomic shotgun sequencing library. Different colours cluster genes belonging to the same pathway; photosynthesis (green), carbohydrate (pink), nitrogen (blue) and sulfur (purple) metabolisms. Unclassified bacterial orders (U) are reported as bacterial classes.

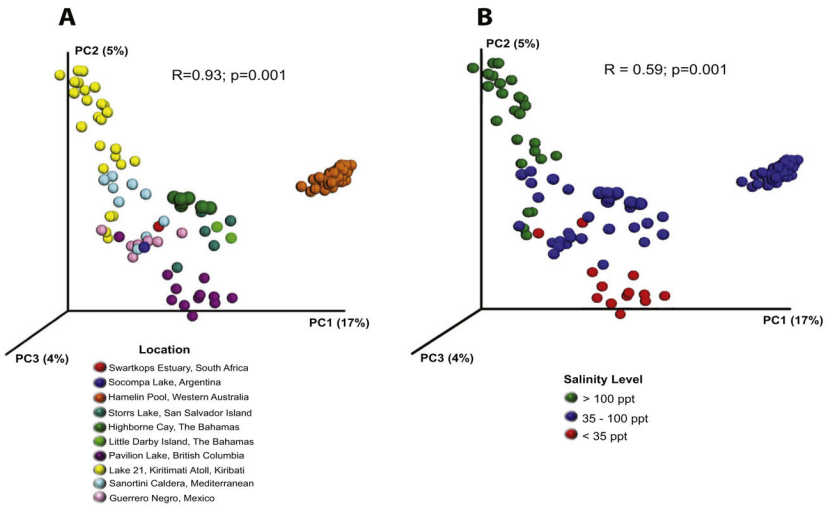

Fig. 6.

Comparative diversity analyses of lithifying and non-lithifying microbial mats across geographical locations. Principal coordinates analysis plotting the unweighted UniFrac distance matrix generated from rarefied taxon abundances clustering according to geographical area. Analysis of similarity indicated that (A) geography (R = 0.93) and (B) salinity (R = 0.59) played a role in significant driving differences between the microbial mat communities ($P = 0.001$). Salinity was divided into three major categories based on the habitat and included: > 100 ppt (Kiritimati Atoll); 35–100 ppt (Guerrero Negro, Hamelin Pool; Highborne Cay; Little Darby Island; Santorini Caldera; Socompa Lake; Storr's Lake); and <35 ppt (Swartkops Estuary; Pavilion Lake).

Table 1

Metadata associated with metagenomic sequence data collected from stromatolite ecosystems.

Sample location	Habitat	SRA #[a]	Sequencing Platform	Sample #[b]
Pavilion Lake, British Columbia	Freshwater	SRP035880	454 GS FLX Titanium	13
Swartkops Estuary, South Africa	Freshwater	SRP039055	454 GS Junior	1
Highborne Cay, Bahamas	Marine	SRP004035	454 GS FLX Titanium	8
Little Darby Island, Bahamas	Marine	SRS101923	Illumina MiSeq	2
Hamelin Pool, Western Australia	Hypersaline	SRP055055	Illumina GAIIx	75
Socompa Lake, Argentina	Hypersaline	SRP007748	454 GS FLX Titanium	1
Storrs Lake, San Salvador Island	Hypersaline	SRP031628	454 GS FLX Titanium	5
Lake 21, Kiritimati Atoll	Ultrahypersaline	SRP015407	454 GS FLX Titanium	50
Santorini Caldera, Mediterranean	Marine NL[c]	SRP019274	454 GS FLX Titanium	10
Guerrero Negro, Mexico	Hypersaline NL	SRP030038	454 GS FLX Titanium	8

[a] Accession number of the National Center for Biotechnology Information Short Read Archive.

[b] Sample number reflects the number of sequencing runs.

[c] Non-lithifying (NL) microbial mat systems.