environmental microbiology

Environmental Microbiology (2012) doi:10.1111/j.1462-2920.2012.02698.x

Cyanobacterial construction of hot spring siliceous stromatolites in Yellowstone National Park

Charles Pepe-Ranney,[1] William M. Berelson,[2] Frank A. Corsetti,[2] Merika Treants[3] and John R. Spear[1]*

[1]*Department of Civil and Environmental Engineering, Colorado School of Mines, Golden, CO 80401, USA.*
[2]*Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089, USA.*
[3]*Department of Microbiology, University of Washington, Seattle, WA 98105, USA.*

Summary

Living stromatolites growing in a hot spring in Yellowstone National Park are composed of silica-encrusted cyanobacterial mats. Two cyanobacterial mat types grow on the stromatolite surfaces and are preserved as two distinct lithofacies. One mat is present when the stromatolites are submerged or at the water–atmosphere interface and the other when stromatolites protrude from the hot spring. The lithofacies created by the encrustation of submerged mats constitutes the bulk of the stromatolites, is comprised of silica-encrusted filaments, and is distinctly laminated. To better understand the cyanobacterial membership and community structure differences between the mats, we collected mat samples from each type. Molecular methods revealed that submerged mat cyanobacteria were predominantly one novel phylotype while the exposed mats were predominantly heterocystous phylotypes (*Chlorogloeopsis* HTF and *Fischerella*). The cyanobacterium dominating the submerged mat type does not belong in any of the subphylum groups of cyanobacteria recognized by the Ribosomal Database Project and has also been found in association with travertine stromatolites in a Southwest Japan hot spring. Cyanobacterial membership profiles indicate that the heterocystous phylotypes are 'rare biosphere' members of the submerged mats. The heterocystous phylotypes likely emerge when the water level of the hot spring drops. Environmental pressures tied to water level such as sulfide exposure and possibly oxygen tension may inhibit the heterocystous types in submerged mats. These living stromatolites are finely laminated and therefore, in texture, may better represent similarly laminated ancient forms compared with more coarsely laminated living marine examples.

Received 2 March, 2011; revised 6 December, 2011; accepted 6 January, 2012. *For correspondence. E-mail jspear@mines.edu; Tel. (+1) 303 273 3497; Fax (+1) 303 273 3413.

Introduction

Stromatolites are laminated, accretionary structures commonly interpreted as a lithified manifestation of microbial life (e.g. Awramik *et al*., 2005) and are observed in rocks as old as 3.4 billion years (Allwood *et al*., 2006). Stromatolites were very common in shallow marine environments of Precambrian age, but became much more rare after the diversification of metazoan life, approximately 542 million years ago (Awramik and Sprinkle, 1999). 'Living' marine stromatolites found in Shark Bay, Western Australia (e.g. Hoffman, 1976; Burns *et al*., 2004; Papineau *et al*., 2005) and Highborne Cay, Bahamas (e.g. Visscher *et al*., 1998; 2000; Reid *et al*., 2000; Paerl *et al*., 2001; Andres and Reid, 2006; Baumgartner *et al*., 2009; Foster *et al*., 2009) lend insight into the biological significance of ancient forms. The microstructure of the modern marine examples, however, is rather coarse compared with most ancient forms (Awramik and Riding, 1988; Grotzinger and Knoll, 1999). Berelson and colleagues (2011) reported finely laminated stromatolites growing in a hot spring in Yellowstone National Park (YNP), Wyoming, that constitute a somewhat better *textural* analogue to ancient stromatolites versus the modern marine examples. Berelson and colleagues (2011) focused on growth rates and a formal description of the stromatolites from a geological perspective. Here, we focus on the molecular biology, environmental context and evolutionary importance of the microbial communities building the finely laminated stromatolites from YNP.

Most ancient stromatolites are composed of calcium carbonate, whereas the YNP examples described here (Fig. 1) are siliceous. Early lithification is crucial for a mat to form a stromatolite over time, regardless of the geochemistry of the system. In YNP, early lithification occurs as silica precipitates in association with the mats, thus forming an excellent textural, if not geochemical, analogue for many finely laminated ancient stromatolites. Moreover, living (Golubic and Focke, 1978) and ancient (see Seong-Joo and Golubic, 1998 and Zhongying, 1986)

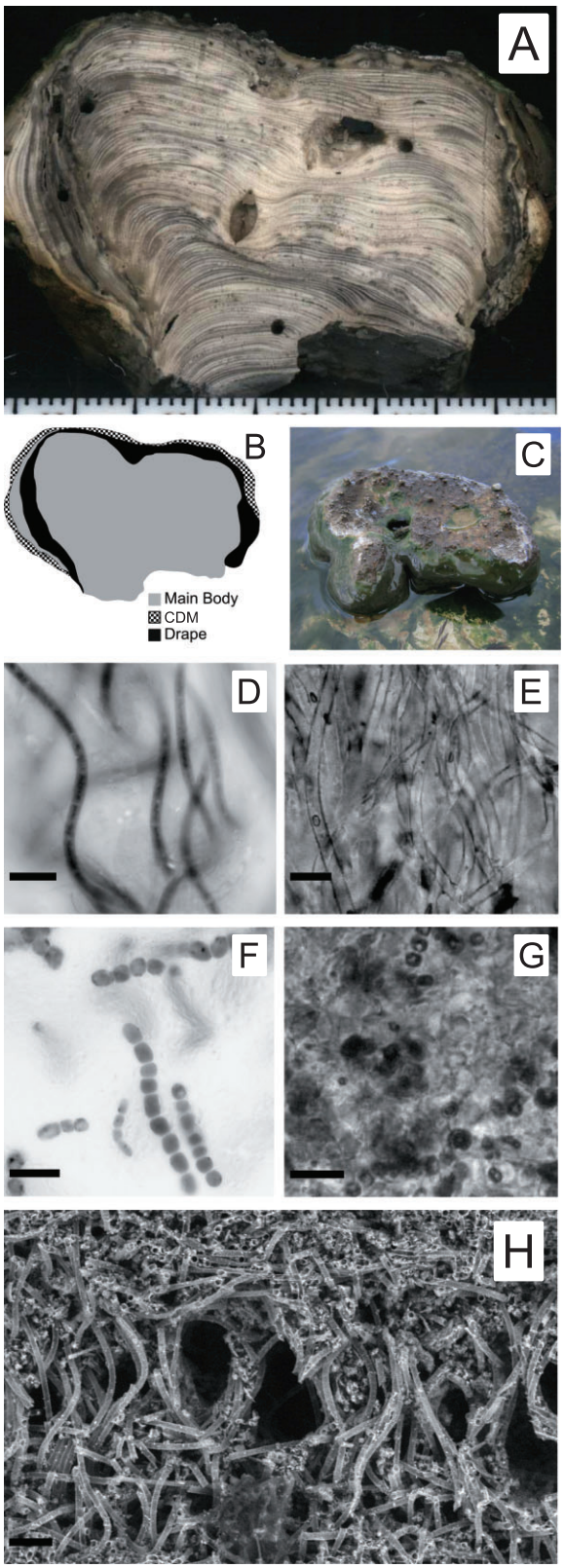

Fig. 1. Selected images of stromatolite facies. All scale bars are 10 μm unless otherwise indicated.
A. Cross-section of stromatolite from which CDM samples in August of 2006 were collected (reproduced with permission from Pepe-Ranney, C., Berelson, W.M., Corsetti, F.A., Treants, M., and Spear, J.R. submitted). Gradations in scale bar across bottom of image are millimetres.
B. Cartoon depicting the location of facies in cross-section 'A'.
C. Stromatolite from image 'A' *in situ*.
D. Autofluorescence image of predominant morphotype in SFM.
E. Thin section image of 'Main Body' lithofacies.
F. Autofluorescence image of predominant phylotype in CDM samples from August of 2006.
G. Thin section image of 'Drape' lithofacies.
H. Electron micrograph of silica tubes that comprise 'Main Body' lithofacies.

carbonate stromatolites appear to be formed by alternating layers of cyanobacterial filaments oriented vertically and prostrate, similar to the stromatolites discussed in this study. Thus, the mat behaviours are similar, even if the early lithification is provided by silica or carbonate.

The living, siliceous stromatolites of this study grow around the rim of a hot spring in shallow water centimetres in depth. The stromatolites are built by accretion of silica-encrusted cyanobacterial mats. The bulk of the structure is comprised of silica tubes that are the remains of filamentous cyanobacteria (Berelson *et al*., 2011) (see Fig. 1H). Fine, distinct laminations are the result of the uniform but alternating growth orientation of cyanobacterial sheaths, closely resembling a pattern described for another siliceous sinter at an unspecified YNP hot spring (Walter *et al*., 1972). Specifically, silica-encrusted filaments in a single layer are oriented either sub-normal or sub-parallel to the surface, and the textural differences between normal and parallel orientation give rise to light and dark laminations visible to the naked eye (Berelson *et al*., 2011) (Fig. 1A). The microstructure also resembles that described by Jones and colleagues (1998) insofar as it is comprised of lithified filaments in alternating orientation.

The objective of this study is to present the molecular phylogenetic make-up of the cyanobacteria in stromatolite-building mats. Molecular tools have been used to elucidate microbial diversity of other modern analogues. For instance, community characterizations of microbial populations in stromatolites by 16S rRNA gene sequence surveys have shown that microbial community membership correlates well with stromatolite shape in the Shark Bay system (Papineau *et al*., 2005), and that species richness increases with, and may link to, lithification in the Highborne Cay system (Baumgartner *et al*., 2009). In this study, information from a 16S rRNA gene sequence molecular survey of stromatolite mats was used to (i) investigate the cyanobacterial membership profiles of stromatolite-building mats and (ii) describe the phylogeny of stromatolite-associated cyanobacteria. We report on 'rare biosphere' membership relationships between

different surface mat types and a novel cyanobacterial phylotype that is constructing the predominant stromatolite lithofacies. Although microbial cell morphology is commonly conservative (the specific genetic component of Precambrian stromatolites will likely never be known), understanding the molecular taxonomic composition of a modern example that most closely resembles the texture of the ancient stromatolites represents an important step in understanding stromatolite morphogenesis.

Results

Site and sample description

Centimetre-scale living stromatolites with sub-millimetre lamination grow around the rim of a hot spring in upper Hayden Valley in YNP (Berelson *et al*., 2011). The water in the hot spring is ~ 56°C during the day and the pH is 5.7. pH does not fluctuate significantly; however, temperature varies with wind and weather conditions (Berelson *et al*., 2011). The general water chemistry of the hot spring has been reported by Spear and colleagues (2005) (see 'Obsidian Pool Prime'). We measured combined nitrogen on-site to assess whether the hot spring communities could rely on nitrogen from hot spring water or had to fix atmospheric nitrogen for growth. The NH_4^+, NO_2^- and NO_3^- concentrations in the hot spring were all below the detection limit (0.20, 0.080 and 0.10 p.p.m. all values as N respectively) when measured in October of 2010. The sulfide concentration of water near the perimeter of the north end of the hot spring was measured in August 2008 at 0.224 p.p.m. Most stromatolites appear to be growing from the rim of the pool inward, but isolated stromatolites away from the shore grow upward and outward in all directions. The stromatolites grow by accretion of silica-encrusted cyanobacterial mats (Berelson *et al*., 2011) and our study focuses on yet-to-be encrusted mats.

Two morphologically distinct cyanobacterial surface mats have been observed as part of the stromatolites; one possesses short coccoid chains and branching-filamentous cyanobacterial morphotypes in addition to silica diatom frustules and is found forming a layer of the stromatolite when the structures protrude from the water. The other has non-branching cyanobacterial filaments lacking heterocysts and is found when stromatolites are submerged or at the water surface. The mats with coccoid chains and branching morphotypes could be easily smeared off of stromatolites and were texturally less cohesive and fabric-like than the non-branching filamentous mats, which tightly adhered to surfaces as a cohesive fabric of uniform thickness. Here we call the exposed Mat type with Cocci/Diatom morphotypes 'CDM' and the Submerged Mat type with the non-heterocystous Filamentous morphotype 'SFM'.

To study the microbial communities of surface mats with molecular methods, we collected samples of each mat type. Specifically, two samples were of the CDM, each from an exposed but moist mat on a stromatolite elevated just above the water (see Fig. 1C); and two samples were collected on separate visits to the hot spring from the SFM on mostly submerged stromatolites. Samples of the SFM were collected in different seasons (winter and summer). Greater coverage of the cyanobacterial mats around the hot spring has been observed in all winter trips to the field site. It appears as though the lower light intensities in addition to the lower ambient temperatures and possibly reduced grazing during the winter are more suitable for the cyanobacteria. All samples were collected in the daytime and it is not known how community structure shifts over the short-term (i.e. throughout the day and night).

Microstructure observations by thin section and scanning electron microscopy (SEM) revealed two lithofacies inside the stromatolites (Berelson *et al*., 2011) that correspond to the surface mats. For clarity, this study discusses samples from the two types of surface mats on the stromatolites ('CDM' and 'SFM' above). Each type is eventually silica-encrusted and preserved as a distinct lithofacies. Morphotype similarities between each mat's cyanobacteria and each lithofacies' silica forms (see Fig. 1D–G) allowed us to discern the relationships between mats and lithofacies. The bulk of the stromatolites' interiors can be categorized into a lithofacies composed of non-branching silica filaments (called the 'Main Body Facies' by Berelson *et al*., 2011) (Fig. 1E), which corresponds to the SFM. The second lithofacies, referred to by Berelson and colleagues (2011) as the 'Drape Facies', corresponds to the CDM (Fig. 1G).

Bacterial small sub-unit (SSU) rRNA gene libraries of pyrosequences and nearly full-length Sanger sequences were produced for diversity and phylogeny investigations respectively. Pyrosequence library names and descriptions are summarized in Table 1 and *Experimental procedures*; all library names indicate the mat type of the library's sample (CDM or SFM).

Cyanobacteria comprise the most abundantly found phylum in each pyrosequence library (Fig. 2A) and the cyanobacterial 16S rRNA sequences from each surface sample are of predominantly one phylotype (see Fig. 2B). When mat samples are viewed under the microscope it appears that cyanobacteria comprise a greater fraction of biomass than representation in 16S rRNA gene sequence libraries indicate. The dominant phylotype of the SFM (Fig. 2B) corresponds to the dominant SFM morphotype by qualitative comparisons of phylotype and morphotype distributions and is likely preserved as the silica tubes, essentially building the stromatolites.

Table 1. Sample and corresponding library information.

Mat facies	Sequence information: No. of sequences, average length (SD)	Sequence library name	Sample description	Collection date
CDM	1800, 233 (17)	PCDM0806a	Surface mat above water surface with coccoidal and filamentous, heterocystous, true-branching cyanobacterial morphotypes	August 2006
CDM	2063, 233 (25)	PCDM0806b	Same as above	August 2006
SFM	2044, 225 (25)	PSFM0209	Surface mat at water's surface dominated by filamentous cyanobacterial morphotype without heterocysts	February 2009
SFM	1706, 217 (28)	PSFM0809	Same as above	August 2009

Predominant cyanobacterial phylotypes

CDM samples. Libraries PCDM0806a and PCDM0806b are dominated by one major and one minor phylotype (Fig. 2B). Figure 3 shows the distribution of non-singleton/doubleton cyanobacterial sequences in PCDM0806a and PCDM0806b. As shown in Fig. 2B, two lineages possess the majority of the cyanobacterial membership in each CDM pyrosequence library (99% and 98%). The most abundant phylotype in both samples (85% and 92% of total cyanobacterial sequences) is shown as the red cluster of Fig. 3; the average pairwise identity in this cluster is 99.22%. All sequences in the red clade have the same nearest neighbour by BLAST (97.86–99.57% identity) in the Silva SSURef104 database (Pruesse *et al.*, 2007), a database of near full-length, annotated and curated SSU rRNA gene sequences. The nearest-neighbour is a 16S sequence from a cultivar in the Pasteur Culture Collection (PCC) annotated as *Chlorogloeopsis* HTF sp. PCC 7518 (Accession X68780) (Wilmotte *et al.*, 1993). Cyanobacteria in the *Chlorogloeopsis* HTF are heterocystous and are arranged in short chains or aggregates of coccoidal cells (Castenholz, 2001). Morphotypes from CDM samples fitting this description have been observed by microscopy (Fig. 1).

The second most abundant phylotype in the CDM libraries (14% and 6% of total cyanobacterial sequences in the pyrosequence libraries) has 100% sequence identity to a 16S rRNA gene sequence from an unpublished study of microbial communities in Australia's Great Artesian Basin (Accession AF407696). The sequence also has 99.09% identity by BLAST to *Fischerella* (*Mastigocladus Laminosus*) cultivars isolated from Costa Rica hot springs (Accession DQ786171, Finsinger *et al.*, 2008). True-branching filaments characteristic of the *Fischerella* genus (Castenholz, 2001) have been observed by microscopy in our CDM samples. Figure S1 shows the phylogenetic placement of near full-length sequences that represent the two most abundant phylotypes of CDM pyrosequence libraries.

SFM samples. A total of 93% and 82%, respectively, of cyanobacterial sequences in the SFM libraries PSFM0809 and PSFM0209 fall into one cluster with 98.77% average intra-cluster identity – this cluster is coloured blue in Fig. 3. Each sequence in this dominant

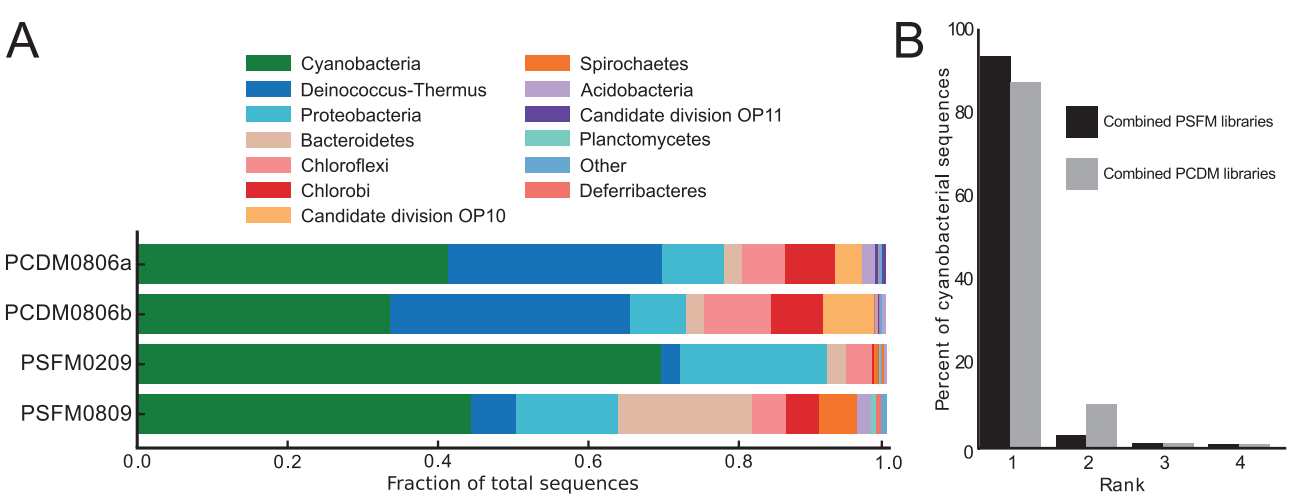

Fig. 2. A. The bar chart depicts the distribution of pyrosequencing reads into phyla for each pyrosequence library.
B. Rank abundance plot of cyanobacteria 97% identity OTUs in the combined PSFM and PCDM libraries respectively.

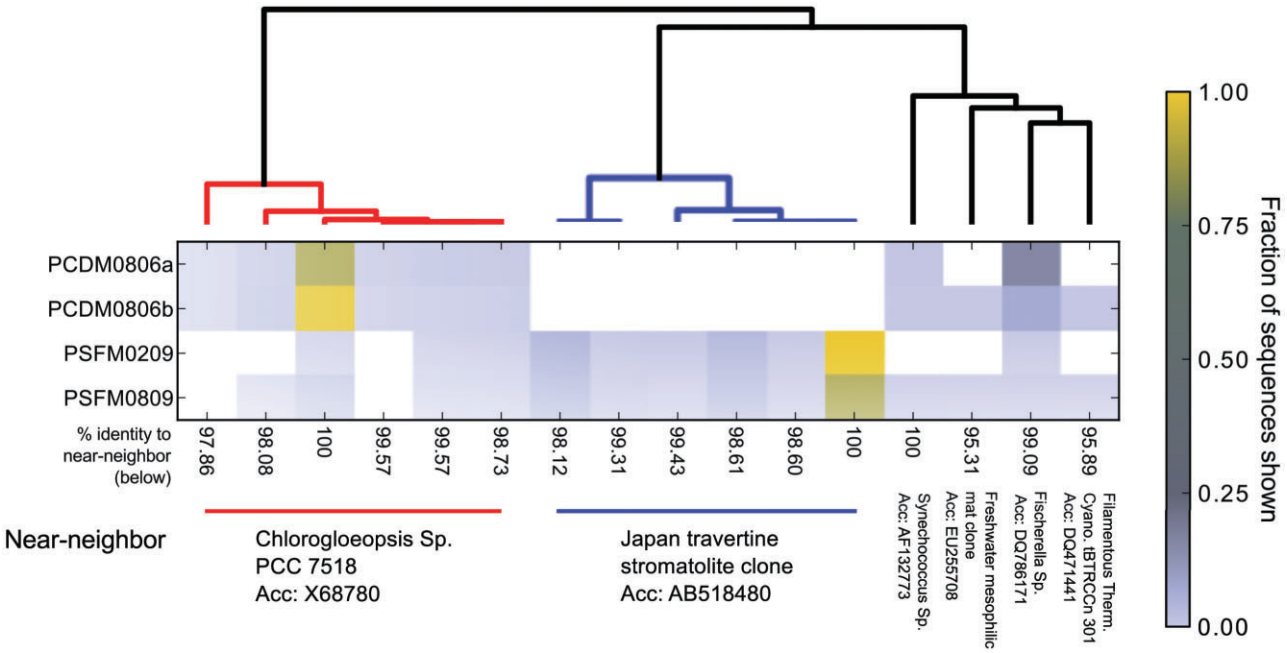

Fig. 3. The heat-map and dendrogram shown here depict the distribution and abundance of nonsingleton/doubleton cyanobacterial reads from the pyrosequence libraries in this study. The dendrogram is not a phylogeny but a furthest-neighbour clustering of all pairwise substitution per site comparisons of the sequences.

cluster has the same nearest neighbour (98.6% to 100% identity) by BLAST in Silva SSURef104. The nearest-neighbour is an unpublished sequence from a travertine hot spring in SW Japan (Accession AB518480) and is not classified beyond the phylum level by Greengenes (DeSantis *et al*., 2006), the Ribosomal Database Project (RDP) (Cole *et al*., 2005), or Silva. The SW Japan hot spring displays carbonate geobiological structures interpreted to be stromatolites, although the noted sequence is unpublished and it is unclear whether it is associated with a microstructure organization that is comparable to the YNP stromatolites.

Some rare sequences (< 3% of total sequences) in the SFM libraries are the predominant CDM phylotype. Conversely, no sequences from the CDM samples fall into the predominant cluster of the SFM samples (Fig. 3). To assess the phylogeny of the predominant phylotype of PSFM0809 and PSFM0209, nearly full-length 16S rRNA gene sequences were produced (Sanger sequenced) from the samples. One sequence, with optimal length (1364 bp) and quality (average QScore, 54.4), that shared 98.61% identity with the most abundant PSFM pyrosequence library sequence and up to 100% identity with other sequences in the predominant cluster (blue cluster Fig. 3) was selected to represent the predominant SFM phylotype. This near full-length representative sequence will hereafter be referred to as 'SFM-seq'. Sequences in GenBank (Benson *et al*., 2011) and SSUParc106, a curated database of all SSU rRNA gene sequences (short and long) from public repositories of biological sequences, that share high identity (97% or greater) to SFM-seq are summarized in Table 2 and include one cultivar 'Microcoleus' sp. PCC 8701 (Accession AY768403), that was isolated from a sulfur hot spring in Amelie Les Bains, Fance (PCC), in addition to several other travertine mat sequences from a Japan hot spring. The ribotype has also been reported once before in a study of YNP mesothermic microbial mats (see Accession FJ885932), but has not been published to date.

We determined the phylogenetic placement of SFM-seq and its SW Japan travertine near-neighbour by reconstructing their phylogeny in the context of a broad reference set of cyanobacterial 16S rRNA gene sequences. Although cyanobacteria have historically been classified into five subsections, the topology of cyanobacterial 16S phylogeny has shown more than five basal lineages (Honda *et al*., 1999; Turner *et al*., 1999; Wilmotte and Herdman, 2001). We chose to investigate the phylogeny of SFM-seq in the context of the cyanobacterial groups recognized by the RDP (Cole *et al*., 2009) that are based on the topology described by Wilmotte and Herdman (2001). The phylogenetic placement of SFM-seq and the SW Japan travertine sequence is outside all recognized subphylum cyanobacteria groups in the RDP taxonomy, suggesting the sequences constitute a novel basal lineage in the cyanobacteria (Fig. 4A).

Table 2. Sequences in SSUParc106 and/or GenBank with at least 97% identity to SFM-seq.

%ID to SFM-seq	Accession Number of top hit	Full name in SSUParc106	Study title in SSUParc106	Description in SSUParc106	PubMed ID
99.9[a]	FJ885932	Uncultured cyanobacterium	Variable community structures in microbial mats of the mesothermic border of hot spring pools in YNP	Uncultured cyanobacterium clone O1UDE03 16S ribosomal RNA gene, partial sequence.	
99.8[a]	AY768403	*Microcoleus* sp. PCC 8701	Cyanobacterial natural products genes: a source of novel genes for creating a metabolically engineered microbe	*Microcoleus* sp. PCC 8701 16S ribosomal RNA gene, partial sequence.	
99.0[b]	AB614537	N/A	Laminated microbial mat on Naganoyu travertine[c]	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: S1-101[d]	
98.6[b]	AB614557	N/A	Laminated microbial mat on Naganoyu travertine[c]	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: S2-52[d]	
98.5[a]	EF660480	Uncultured bacterium	A survey of the alkalithermophilic prokaryotic diversity from the hot spring waters in Coamo Puerto Rico	Uncultured bacterium clone F26 16S ribosomal RNA gene, partial sequence.	
98.5[a]	EF660494	Uncultured bacterium	A survey of the alkalithermophilic prokaryotic diversity from the hot spring waters in Coamo Puerto Rico	Uncultured bacterium clone F60 16S ribosomal RNA gene, partial sequence.	
98.3[b]	AB614561	N/A	Laminated microbial mat on Naganoyu travertine[c]	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: S2-60[d]	
98.2[a]	AB518480	Uncultured bacterium	Microbial process forming daily lamination in an aragonite travertine, Nagayu hot spring, SW Japan	Uncultured bacterium gene for 16S ribosomal RNA, partial sequence, clone: H2	
98.1[b]	AB614539	N/A	Laminated microbial mat on Naganoyu travertine[c]	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: S1-106[d]	
98.0[b]	AB614535	N/A	Laminated microbial mat on Naganoyu travertine[c]	Uncultured bacterium gene for 16S rRNA, partial sequence, clone: S1-43[d]	
97.9[a]	DQ146324	Uncultured *Nostoc* sp.	Watering, fertilization and slurry inoculation promote recovery of biological crust function in degraded soils	Uncultured *Nostoc* sp. isolate DGGE band 24 16S ribosomal RNA gene, partial sequence.	16710791
97.7[a]	DQ146332	uncultured *Lyngbya* sp.	Watering, fertilization and slurry inoculation promote recovery of biological crust function in degraded soils	Uncultured *Lyngbya* sp. isolate DGGE band 14 16S ribosomal RNA gene, partial sequence.	16710791
97.0[a]	DQ146331	Uncultured *Lyngbya* sp.	Watering, fertilization and slurry inoculation promote recovery of biological crust function in degraded soils	Uncultured *Lyngbya* sp. isolate DGGE band 13 16S ribosomal RNA gene, partial sequence.	16710791

a. %ID is the default identity definition in USearch version 4.1.93.
b. %ID is the identity definition of NCBI BLAST.
c. Study title in GenBank.
d. Sequence description in GenBank.

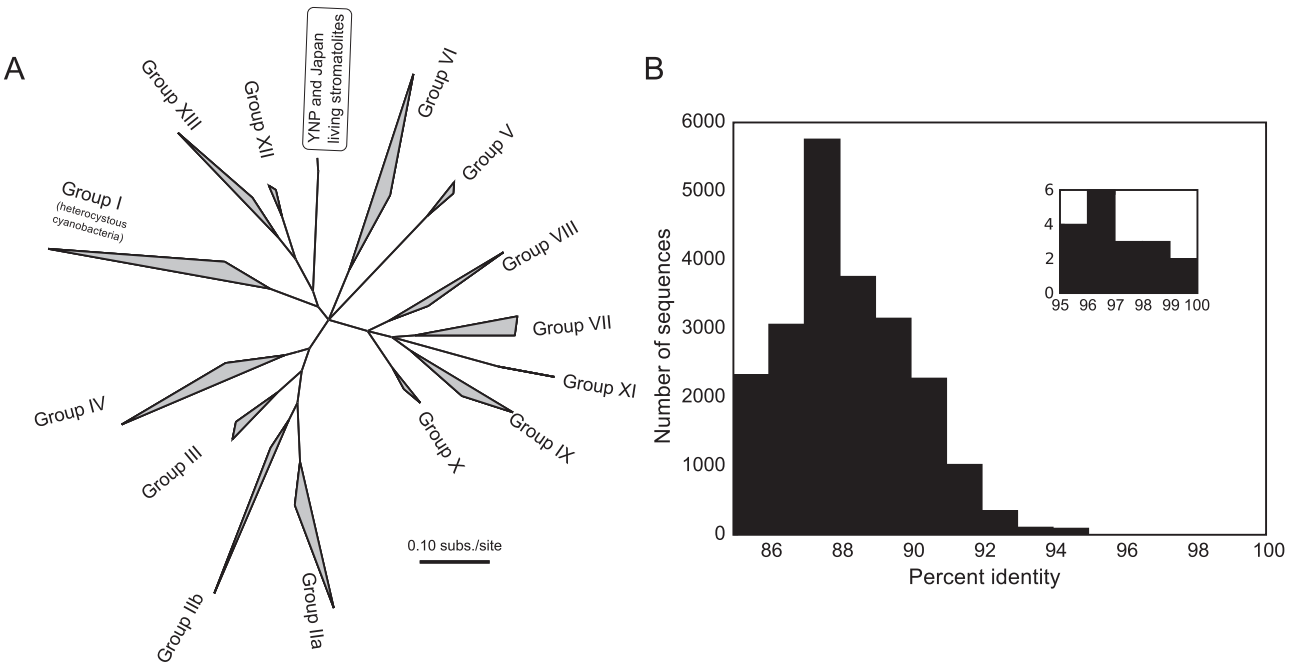

Fig. 4. A. Location of the stromatolite building phylotype in a broad, unrooted, maximum likelihood phylogeny of cyanobacteria. Groups are those named and recognized by RDP directed by study done by Wilmotte and Herdman (2001). Nodes with bootstrap support < 50% were collapsed and thus polytomies denote ambiguous branching order of descendant nodes.
B. Histogram of all identity values as determined by USearch (Edgar, 2010) between 'SFM-seq' and cyanobacterial sequences in SSUParc106.

An analysis of the diversity of cyanobacteria found in eight other 16S rRNA surveys of living stromatolites reviewed by Foster and Green 2011 (Burns *et al*., 2004; Elser *et al*., 2005; Baumgartner *et al*., 2009; Foster *et al*., 2009; Goh *et al*., 2008; Havemann and Foster, 2008; Santos *et al*., 2010) shows that a wide diversity of cyanobacteria are found to be associated with living stromatolites (Fig. S2). There does not appear to be any clear relationship between the major phylotypes found in this study and the cyanobacterial 16S rRNA sequences generated in surveys of other modern stromatolite analogues. There does appear to be a faint correlation between cyanobacterial membership with qualitative salinity (freshwater versus marine/hypersaline) in living stromatolite systems. Salinity has shown to have a high correlation with community membership on the global scale (Lozupone and Knight, 2007). No published 16S rRNA gene sequences from living stromatolite systems show significant identity to SFM-seq (see Fig. S2).

Major non-cyanobacterial phylotypes

As shown in Fig. 2, the pyrosequence libraries have overlapping membership at the phylum level in most of the abundantly found phyla. One clear difference is the low OP10 membership of 16S rRNA gene sequences in the SFM libraries. Only one OP10 16S rRNA gene sequence was recovered in each SFM library. OP10 is found in relatively higher abundance in the CDM libraries (7% and 3% of total sequences, Fig. S3). However, when examining the distribution of sequences in abundantly found phyla at higher resolution, there are several differences in the membership and structure of the CDM versus the SFM. Both the SFM and the CDM have members in the *Bacteroidetes*, OP10, *Chlorobi* and *Chlorflexi* groups, yet the most abundant operational taxonomic units (OTUs) with each of those phyla are not consistent between the SFM and the CDM (see Fig. S3). For instance, 82% and 92% of *Chlorobi* sequences in PCDM0806a and PCDM0806b are in the same 97% identity OTU but none of the SFM sequences fall into this OTU. Interestingly, *Chlorflexi* appears to be a major member of many cyanobacterial mats in YNP based on 16S rRNA gene surveys (see Miller *et al*., 2009a and Osburn *et al*., 2011), but only comprises a maximum of ~9% of sequences in any pyrosequence library generated in this study. The most abundant non-cyanobacterial phylum in the pyrosequence libraries is *Deinococcus–Thermus* (Fig. S3 and Table S1). The most abundant 97% identity OTU within *Deinococcus–Thermus* is closely related to *Meiothermus silvanus* an aerobic heterotroph isolated from hot springs in northern Portugal (Tenreiro *et al*., 1995).

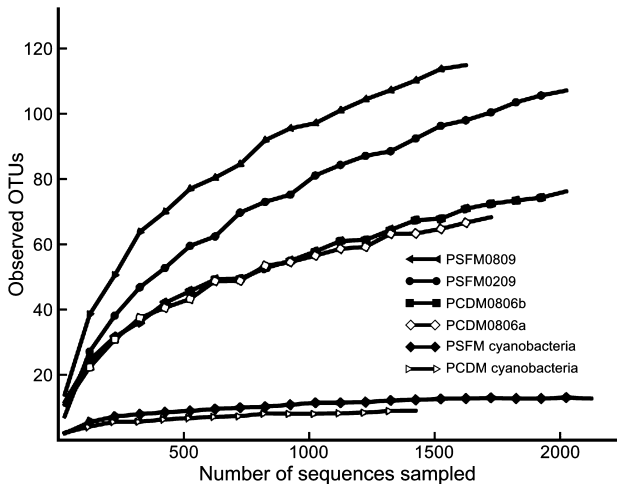

Fig. 5. Rarefaction curves for pyrosequence libraries and just cyanobacterial sequences for combined PCDM and PSFM libraries respectively. Sequences were clustered at 97% identity using UClust (Edgar, 2010). Points in curves represent the average observed OTUs for 10 re-samplings. Libraries were re-sampled at 100 sequence intervals.

Alpha diversity

Rarefaction curves show richness of the CDM is less than the SFM (Fig. 5). This trend holds for cyanobacterial specific richness as well. Table 3 summarizes the alpha diversity predictions of each sample as calculated by CatchAll (Bunge, 2011). The slopes of rarefaction curves for all pyrosequence libraries seem to be approaching asymptotes indicating the majority of the diversity in each sample has been recovered. Parametric richness estimates of entire pyrosequence libraries from each sample, however, indicate a significant amount of OTUs still have not been observed. In contrast to the total alpha diversity, coverage of just unique cyanobacterial sequences is greater. Only 13 unique cyanobacterial singletons are in the combined pyrosequence libraries. Extrapolating from the distribution of cyanobacterial reads combined for each mat type using CatchAll, only an additional 6 and 13 cyanobacterial sequences from the CDM and SFM, respectively, are predicted to be unobserved in the dataset (Table 3).

Discussion

Surface community structure

By 16S rRNA gene sequence composition, the Submerged, Filamentous morphotype Mat (SFM) possesses most of the cyanobacterial diversity of the Cocci, Diatom Mat (CDM) and identical sequences to the most abundant sequence of the CDM (Fig. 3). It has been postulated that rare members of microbial communities have the metabolic machinery to emerge with environmental changes (e.g. Sogin *et al*., 2006). In this stromatolite system, it appears that the *Chlorogloeopsis* HTF and *Fischerella* (*M. laminosus*) phylotypes emerge from the SFM following a shift in environmental conditions. If so, each progression from the 'Main Body' lithofacies to the 'Drape' lithofacies (Berelson *et al*., 2011) records an emergence event of *Chlorogloeopsis* HTF and *Fischerella* from rare to predominant members.

Chlorogloeopsis HTF and *Fischerella* are found in hot springs across the globe and both genera exploit nitrogen-limited waters where they can out-compete non-nitrogen-fixing cyanobacteria (Ward and Castenholz, 2000). *Fischerella* can also be found in high abundance in YNP where there is sufficient combined nitrogen to make nitrogen fixation unnecessary (Miller *et al*., 2006). *Synechococcus* mats are found in many hot spring outfalls in YNP, but *Fischerella* can predominate in waters below 58°C where the pH is low and/or sulfide levels are moderate (Castenholz, 1977). *Chlorogloeopsis* HTF (originally *Mastigocladus* HTF) was isolated by Castenholz (1969) from Icelandic hot springs where it was found to exhibit a higher upper-temperature limit than previously described *Mastigocladus* types. Similarly to *Fischerella*, in YNP *Chlorogloeopsis* HTF predominates in lower pH springs where *Synechococcus* is absent (Castenholz, 1978).

While the SFM preserves a limited reservoir of the predominant CDM phylotypes, the CDM does not appear to possess the predominant SFM cyanobacterium (Fig. 3). It is possible that the tube-building phylotype of the SFM would be observed in the CDM with deeper sequencing, although alpha diversity measurements indicate the sequencing effort has recovered most of expected unique

Table 3. Alpha diversity calculations for pyrosequence libraries.

OUT % identity	Phylotypes	Library name(s)	Total No. of observed OTUs	Estimated total OTUs	Estimated coverage (%)	Lower confidence bound[a]	Upper confidence bound[a]
100	Cyanobacteria	PCDM0806a and PCDM0806b	15	17.8	84	15.8	25.2
97	all	PCDM0806a	69	178.5	39	102.7	424.6
97	all	PCDM0806b	77	160.8	48	117.8	249
100	Cyanobacteria	PSFM0809 and PSFM0209	32	41.3	77	36.1	53.2
97	all	PSFM0809	117	198.3	59	161.7	265.2
97	all	PSFM0209	108	236	46	169.7	373.5

a. 95% confidence bound. For details see Bunge (2011).

cyanobacterial sequences (Fig. 5, Table 3). Therefore, we interpret progressions from 'Drape' to 'Main Body' lithofacies to include a non-conformity, or period of undetermined length for which there is no corresponding record of deposition, as it is unknown how long it takes the SFM to re-establish itself on the stromatolites. Measured growth rates on artificial growth substrates over the short term have thus far proven to be faster than rates determined by radiometric dating of stromatolites with multiple generations of Drape lithofacies (Berelson *et al*., 2011). There appear to be periods of unknown length where the stromatolites are not accreting and these no-growth periods may coincide with the time it takes to re-establish the SFM.

Berelson and colleagues (2011) describe ^{14}C signatures in the YNP stromatolites that suggest the fixed carbon of the CDM has a greater atmospheric to subsurface CO_2 input ratio than the SFM. Specifically, the 'Drape' lithofacies shows a generally younger carbon age than the 'Main Body' lithofacies, indicating the CDM autotrophs receive more carbon input from 'younger' atmospheric CO_2 as compared with SFM autotrophs that receive a greater carbon input from 'older', radiocarbon-dead subsurface CO_2 that vents from the hot spring. One explanation is that the CDM community exists on the stromatolites when water level in the hot spring drops and stromatolites protrude out of the pool. Indeed, the CDM has only been observed when stromatolites extend above the water's surface whereas the SFM is found on more submerged stromatolites. Nitrogen availability, temperature, sulfide and pH have all proven to be influential in selecting cyanobacterial phylotypes in hot spring communities (Ward and Castenholz, 2000) and could potentially change with water level in this system.

Chlorogloeopsis HTF and *Fischerella* are both heterocystous (Castenholz, 2001). The SFM cyanobacterium, contrastingly, is not phylogenetically related to heterocystous cyanobacteria (Fig. 4A) and heterocysts have proven to be a phylogenetically coherent trait (Hoffmann *et al*., 2005; Tomitani, 2006). Stal and colleagues (1993), in a study of the cyanobacterial mats in hypersaline ponds of Guererro Negro, Baja California Sur, reported a similar trend in relative abundance differences of heterocystous versus non-heterocystous cyanobacteria in relation to water level to that observed in this study. Nitrogen fixation and its relationship to oxygen tension could potentially explain the community structuring of submerged versus non-submerged mats provided combined nitrogen levels are low enough that the cyanobacteria must fix nitrogen. Presumably, lower diffusion rates in the submerged mat could allow for oxygen accumulation during the day beyond atmosphere saturation as observed in the submerged Baja mats (Stal *et al*., 1993). High oxygen tension would inhibit nitrogen fixation by heterocystous types that fix nitrogen during the day (Fay, 1992), giving the advantage to the tube-building phylotype which would presumably fix nitrogen at night once oxygen had been sufficiently depleted by respiration. Although the NH_4^+ and NO_3^- content of the stromatolite hot spring was below 0.20 and 0.10 p.p.m., respectively, cyanobacterial hot spring communities in YNP have been shown to survive without fixing nitrogen in waters with combined nitrogen levels of 0.15 p.p.m. (Miller *et al*., 2006). Therefore, it is unknown if either mat type is fixing nitrogen. Additionally, it has not been shown that the SFM accumulates oxygen in the light similarly to the marine Baja mats.

Defined temperature niches in *Synechococcus* species have been observed within Octopus Spring (YNP) cyanobacterial mats (Ward *et al*., 1998) and inter-genus niches have been observed along a temperature gradient in Hunter's Hot Springs, Oregon (Ward and Castenholz, 2000). Similar temperature adaptations may be selecting for the cyanobacterial phylotypes in the YNP stromatolites. The exposed mats would experience greater temperature fluctuations throughout a 24-hour cycle and with the changing seasons. Also, the CDM would experience lower temperatures on average than the SFM. *Fischerella* and *Chlorogloeopsis* HTF are considered to be high-temperature tolerant nitrogen fixers, however, and have been observed in waters up to 58°C and 64°C respectively (see Ward and Castenholz, 2000). *Fischerella* has been found to predominate at temperatures ~ 52°C but below 56°C in a Mammoth Hot-Springs outfall channel, YNP and White Creek, YNP (Miller *et al*., 2006). The proximity of the apparent upper temperature limit for *Fischerella* to this study's water temperature (56°C) may account for its lower abundance in submerged mats, though closely related cyanobacteria have been observed to occupy different temperature niches (Ward *et al*., 1998; Miller *et al*., 2009b), and therefore ribotype identity of a particular cyanobacterium probably is not indicative its temperature optimum. Nevertheless, the water temperature of this study's hot spring is within the range of temperatures where *Fischerella* and *Chlorogloeopsis* HTF communities have been previously observed. In fact, temperatures of 60°C have been used to enrich for *Chlorogloeopsis* HTF in the laboratory (Castenholz, 1978). It does not appear that a water temperature of 56°C would select against *Chlorogloeopsis* HTF but it may be inhibitory to *Fischerella*.

In contrast to temperature, the correlation of sulfide with the cyanobacterial restructuring is more consistent with field observations of *Fischerella* and *Chlorogloeopsis* HTF. The sulfide levels in the stromatolite hot spring are fairly high near the vent ~ 0.60 p.p.m. (Spear *et al*., 2005), and could inhibit some cyanobacteria (e.g. Ward and Castenholz, 2000; Miller and Bebout, 2004). *Chlorogloeopsis*

HTF and *Fischerella* appear to have a low and moderate tolerance to sulfide respectively (Ward and Castenholz, 2000), and we would expect the submerged mats to experience more consistent exposure to the sulfidic water. Sulfide concentrations of 0.15 p.p.m. were sufficient to inhibit photosystem II function by 50% in a *Fischerella* strain isolated from the Boiling River, YNP (Miller and Bebout, 2004); however, the strain was isolated from a low sulfide environment and sulfide tolerance has not been proven to be reliably extrapolated to a cyanobacterium based on the phenotype of its close relatives (Miller and Bebout, 2004). It should also be noted that the sulfide concentration at the north end of the stromatolite hot spring has been measured lower (0.224 p.p.m.) than near the vent. Thus far, studies of the YNP stromatolites have been limited to samples collected from the north end. Stromatolite structures do grow on the southern rim, closer to the vent; however, it is unknown at this time if the 'Drape' lithofacies is found in the south end stromatolites or if the sulfide levels of the south end are higher than the north end.

The pH of the stromatolite hot spring (~ 5.7) is too low for many cyanobacteria, but both *Chlorogloeopsis* HTF and *Fischerella* have been found at lower pH and seem to be acid tolerant (Ward and Castenholz, 2000). The pH of the CDM may be higher than the SFM because pH affects due to carbon fixation would be less mitigated by the surrounding water, although respiration would dampen that difference. Regardless, as with temperature, *Chlorogloeopsis* HTF and *Fischerella* would not appear to be inhibited by the pH of the surrounding spring based on published field observations and therefore presumably pH is not strongly influencing cyanobacterial community structure in this system.

Phylogeny of stromatolite builder

The nearest-neighbour in SSURef104 to the sequence 'SFM-seq' representing the predominant SFM cyanobacterium that is building the bulk of the stromatolites and is responsible for the fine laminations is not classified beyond the phylum level in Greengenes, RDP or the Silva databases. As shown in Fig. 4B, SFM-seq does not share genus-level identity, 95% (Ludwig *et al*., 1998), with any other sequence in Silva SSURef104. Moreover, the phylogeny of SFM-seq shows it does not belong in the cyanobacterial sub-phylum groups currently recognized by the RDP (Fig. 4A). SFM-seq and its SW Japan stromatolite neighbour constitute a heretofore undocumented lineage within the cyanobacteria showing molecular studies have not recovered the full diversity of cyanobacterial 16S rRNA genes and that there is captured cyanobacterial diversity in sequence databases that remains undescribed from a phylogenetic standpoint.

Comparisons to other living stromatolite studies

Other molecular studies of living stromatolites have elucidated the complexity of stromatolite microbial membership and show the correlation of geochemistry and stromatolite characteristics with diversity (Papineau *et al*., 2005; Baumgartner *et al*., 2009). To our knowledge, no other living stromatolite system has characteristics such that the growth of the stromatolites can be attributable to one or any specific phylotype(s). In contrast, the stromatolites presented here are predominantly built by one cyanobacterial phylotype. Furthermore, this cyanobacterial phylotype constitutes a novel and basal lineage in the cyanobacterial 16S rRNA gene phylogeny.

It should also be noted that molecular studies have not shown any strong commonalities in the microbial members of the geographically distinct living stromatolites in Shark Bay, Australia and Exuma Sound, Bahamas (Papineau *et al*., 2005). These two living systems exhibit different lamination patterns and local conditions, and therefore, common microbial membership would not be expected and would not necessarily reflect significance of any particular phylotype with respect to stromatolite growth. As shown in Fig. S2, a wide diversity of cyanobacterial 16S rRNA sequences have been recovered from living stromatolites.

In the YNP examples and some SW Japan stromatolites, lamination and growth can be attributed to cyanobacteria (Takashima and Kano, 2008). Yet the full significance of such a commonality is unknown without more detailed information on the origins and geographical extent of the Japan travertine cyanobacterium that displays such close relation to SFM-seq. Several studies have focused on carbonate stromatolites in SW Japan hot springs (Chizuru and Akihiro, 2005; Okumura *et al*., 2008; Takashima and Kano, 2008). A study by Takashima and Kano (2008) show the presence of filamentous cyanobacteria on SW Japan carbonate stromatolites in the Shionoha hot spring and attributed lamination to cyanobacterial metabolism effects. While the above studies all discuss travertine stromatolites that exhibit daily lamination, it appears that the lamination is distinct both in texture and in the influencing microbes depending if the travertine is aragonite or calcite (Okumura *et al*., 2008) and where the travertine deposit is in relation to the hot spring vent (Chizuru and Akihiro, 2005; Okumura *et al*., 2008). Therefore, it may be only happenstance that the two systems share the presence of a particular cyanobacterium.

Regardless, we describe and produce a nearly full-length 16S rRNA gene sequence from a cyanobacterium that is a stromatolite-building microbe. Additionally, particular SW Japan travertine stromatolites (Takashima and Kano, 2008) and the YNP examples show that cyanobac-

terial growth rhythms can create laminae in conditions suitable for rapid lithification (for instance, in silica supersaturated hot springs).

Additionally, cyanobacteria have been historically associated with stromatolites in the rock record. This is in part due to the presumed effect on pCO_2 by carbon fixation and also to the conspicuous and easily identifiable morphologies of cyanobacteria (Papineau *et al.*, 2005). In contrast, recent studies have highlighted the roles of non-cyanobacterial members in the interplay between Ca^{2+} and pCO_2 concentrations that lead to $CaCO_3$ precipitation or dissolution (e.g. Visscher *et al.*, 2000; Paerl *et al.*, 2001) and have shown that carbonate stromatolites are not formed under the influence of cyanobacteria alone. Likewise, molecular analyses by Papineau and colleagues (2005) found the fraction of cyanobacteria in clone libraries of 16S rRNA genes to be less than expected in living stromatolites found at Shark Bay, Australia. Similarly, Walter and colleagues (1972) proposed that *an*oxygenic phototrophs and not cyanobacteria were influencing the growth of some living stromatolites in YNP. Here, in contrast, we present living YNP stromatolite growth mainly attributable to cyanobacteria.

Major non-cyanobacterial phylotypes

In carbonate systems, microbial effects on the budget of carbonate-associated solutes drive precipitation. In contrast, a system where accretion is driven by cooling and evaporation of Si super-saturated water or by binding of detrital solids, the construction of stromatolites would be more associated with the surface chemistry of microbial cells and mats. It may be that while the non-cyanobacterial members play crucial roles in the health and function of this study's mat ecosystem(s), they cannot be directly associated with stromatolite construction. Clearly, the SFM and CDM possess different non-cyanobacterial microbial members (see Fig. S3), but whether these differences in membership are caused by the shift in the composition of the primary-producers in the mats or environmental effects or both is not known at this time.

Summary

Our molecular investigation of the cyanobacterial mats on stromatolites in YNP elucidated the biological significance of the two stromatolite lithofacies as described by Berelson and colleagues (2011). Specifically, transitions from the 'Main Body' lithofacies to the 'Drape' lithofacies mark past emergence events of *Chlorogloeopsis* HTF and *Fischerella* from rare members of the community to the large majority of surface biomass. The cyanobacterial community re-structuring is likely driven by physiological traits under selective environmental pressures that vary with water level, including sulfide exposure and potentially oxygen tension. Transition from the CDM back to the SFM seems to require the re-establishment of the predominant cyanobacterium that may coincide with a period of no accretion. The SFM is dominated by a non-heterocystous, filamentous, novel cyanobacterium that is preserved as silica tubes that are the constructive microstructure behind the stromatolites' lamination. This cyanobacterium is interestingly also found to be associated with travertine stromatolites in SW Japan hot springs. The YNP living stromatolites are new, compelling textural analogues for ancient stromatolites and display distinct microbiological characteristics in comparison to popularly studied living marine examples.

Experimental procedures

Sample collection

Table 1 summarizes the sample information. Surface samples were collected by scraping mats from stromatolites using sterile razor blades. Samples were placed in cryovials and frozen in liquid nitrogen for transport and kept at −80°C for long term storage. 16S rRNA gene sequence libraries from each sample were named as follows: the first letter of the library name denotes the sequencing technology used to create the library ('P' for pyrosequencing); the next three letters of each sample name denotes its mat type; and four numbers indicate the date of the sample (MMYY). For example, the pyrosequence library from a SFM mat sample or incipient SFM collected in August of 2009 is labelled PSFM0809. Lastly, samples of the same surface mat type collected on the same date are lettered.

NH_4^+, NO_2^- and NO_3^- content of the hot spring were measured on-site using colorimetric assays with CHEMetrics kit numbers K-1403, K-7004 and K-6913 (CHEMetrics, Calverton, VA, USA) for NH_4^+, NO_2^- and NO_3^- respectively. Sulfide was measured on-site using the methylene blue colorimetric assay (K-9510, CHEMetrics). A V-2000 Multi-analyte Photometer (CHEMetrics) was used in the field to read the sample ampoules.

Microscopy, SEM and thin-sectioning

A portion of each sample was thawed in the lab, fixed in 2% paraformaldehyde in 0.2 μm filtered hot spring water, homogenized and collected on a 0.2 μm black polycarbonate membrane filter (Millipore, Billerica, MA, USA). Cyanobacteria on the membrane filter were viewed by autofluorescence using a HQ : R fluorescence filter (Chroma Technology Corp, Rockingham, VA, USA) and a Leica DM RXA microscope. SEM micrographs were taken

with a Hitachi TM-1000 Tabletop Microscope. For thin-sectioning, stromatolite samples were dried, impregnated with petropoxy, and cut with a diamond saw. The cut surface was polished flat, cemented to a microscope slide with additional petropoxy, and ground down to a thickness of ~ 40 μm for petrographic analysis.

DNA extraction, PCR, cloning, capillary sequencing

DNA was extracted using the MoBio Powersoil DNA extraction kit (MoBio, Carlsbad, CA, USA). A one-minute bead-beating step was employed for lysis in place of the 10 min vortexing step outlined in the manufacturer's protocol. For Sanger capillary sequencing, bacterial primers 8F and 1492R (Lane, 1991) were used to amplify 16S rRNA genes from environmental DNA samples. PCR, cloning/transformation and sequencing were completed as described by Sahl and colleagues (2010).

Sanger sequence assembly and quality trimming

Bases were called from each chromatogram using Phred (Ewing and Green, 1998; Ewing *et al*., 1998) and reads for each clone were assembled using Phrap (Phil Green, http://www.phrap.org). Phred and Phrap were wrapped by a custom Python script that was also developed to trim the reads by quality score (details in Appendix S1).

PCR and pyrosequencing

For pyrosequencing, bacterial 16S rRNA genes were amplified from environmental DNA using 8F and 338R (Amann *et al*., 1995) primers. PCR was conducted with Promega PCR MasterMix (Promega). The PCR program was as follows: Initial denaturation for 2 min at 94°C followed by 30 cycles of annealing (52°C for 20 s), elongation (72°C for 20 s) and denaturation (94°C for 1 min). The amplicon region has proven to yield reliable results in phylogenetic studies even with short reads (Liu *et al*., 2007; 2008). Eight-base barcodes specific to each sample were attached to reverse primers by a 2-base linker to allow for post-sequencing binning of reads by sample (for general method overview see Hamady *et al*., 2008). Additionally, the primers included adapter sequences to be compatible with Roche's 454 GSFLX sequencing platform. Amplicons were normalized using the SeQualprep kit (Invitrogen), pooled, and gel purified (Montage DNA gel extraction kit, Millipore) prior to being sent to the sequencing facility (Anschutz Medical Campus, University of Colorado Denver).

Sequence analysis

Quality control of pyrosequences. Pyrosequences were binned by barcodes and initially quality filtered based on vital parameters identified by Huse and colleagues (2007) using the 'split_libraries.py' script in the QIIME software package (Caporaso *et al*., 2010). Specifically, sequences with ambiguous characters, errors in the barcode or primer, sequence length less than 120 nt or greater than 275 nt, average quality score below 27, or homopolymer runs greater than 6 nt were discarded. The remaining sequences were denoised using DeNoiser version 0.851 (Reeder and Knight, 2010). Chimeras were identified by ChimeraSlayer (Haas *et al*., 2011) and discarded.

Taxonomic classifications. The phylogenetic content of each sample was determined from the pyrosequence libraries. Taxonomic classifications for pyrosequences were made by recruiting pyrosequences to classified reference sequences in the Silva SSURef102_NR database using BLAST (Altschul *et al*., 1990) and extrapolating reference annotations of the top BLAST hit for each pyrosequence. Quality filtering of reference sequences is described in Appendix S1.

Alpha diversity measurements. Richness calculations for all sequences in a given sample were calculated from OTU counts after clustering reads at 97% identity. For richness estimates of just the cyanobacterial phylotypes, reads were clustered at 100% identity. All clustering of sequences for alpha diversity analyses was done using UClust (Edgar, 2010) via QIIME (Caporaso *et al*., 2010) with input order to UClust determined by read abundance. Rarefaction curves (Fig. 5) were generated using QIIME. Parametric estimates of richness were determined using CatchAll (Hong *et al*., 2006; Bunge, 2011). The richness estimates and corresponding 95% confidence bounds from the parametric model that fit the data best are reported in Table 3.

Sequence alignment. Multiple sequence alignments of nearly full-length, and short-length cyanobacterial 16S rRNA gene sequences were constructed using SSU-align (Cannone *et al*., 2002; Nawrocki *et al*., 2009) with a custom, cyanobacteria-specific covariance model constructed from sequences annotated as cyanobacteria in the Silva taxonomy field of the Silva SSURef102_NR database (Pruesse *et al*., 2007) (details in Appendix S1).

All alignments were filtered with SSU-align using posterior probabilities of aligned characters calculated by SSU-align. Specifically, positions where less than 95% of characters had posterior alignment probabilities of 95% or greater were masked for phylogenetic reconstruction. All alignments were visually inspected before further analysis.

Phylogenetic reconstructions. Only nearly full-length sequences were used to construct phylogenies. RAxML

(Stamatakis, 2006) version 7.2.7 was used to reconstruct phylogenies depicted in Figs S1 and S2. All RAxML trees were calculated utilizing the gamma distribution for rate heterogeneity and the GTR model of nucleotide substitution. The final topology of each phylogeny represents the best topology as determined by likelihood scores for 10 and 25 separate RAxML searches starting from random parsimony trees for Figs S1 and S2 respectively. The initial rearrangement setting for the best tree search and number of rate categories for the gamma distribution were selected by RAxML for each run. Bootstrap replicates were found using the rapid-bootstrapping algorithm of RAxML (Stamatakis *et al*., 2008), and the number of bootstrap replicates necessary for each phylogeny was determined using the RAxML frequency or the extended majority rule criteria (Pattengale *et al*., 2010). SH-like branch support values (Anisimova and Gascuel, 2006) were also calculated for backbone trees using RAxML. Pplacer (Matsen *et al*., 2010) was used to insert short and redundant sequences into backbone RAxML trees. Sequences to be inserted by pplacer were aligned to the same covariance model as the backbone sequences (see *Experimental procedures*: *Sequence analysis*: *Sequence alignment*) for each tree and were subject to the same column masking as the backbone sequences. Nodes below indicated support values were collapsed using a custom Python script that incorporated tree manipulation modules from PyCogent (Knight *et al*., 2007).

Selecting reference sequences. Two sets of cyanobacterial reference SSU rRNA sequences were selected to portray the phylogeny of environmental sequences. Each reference set is discussed at length in Appendix S1.

Accession numbers. Nearly full-length sequences generated in this study to represent 'SFM-seq' and the *Chlorogloeopsis* HTF and *Fischerella* phylotypes are deposited in GenBank under accession numbers JF303685, JF303684 and JF303683 respectively. We have made the presented pyrosequence information available via our lab website (http://inside.mines.edu/~jspear/resources.html).

Acknowledgements

We would like to acknowledge the Yellowstone Center for Resources and Christie Hendrix for assistance with a scientific research and collecting permit to J.R.S. Funding for this study was provided by a National Science Foundation Microbial Biology Postdoctoral Start-up Award to J.R.S., and by a US Air Force Office of Scientific Research award to J.R.S. We thank Brad Bebout for teaching us how temporal versus spatial separation of nitrogen fixation and oxygenic photosynthesis can shape community structure. Three anonymous reviewers provided helpful and detailed critiques of the manuscript. Thanks to the Pace Lab (CU-Boulder) for Sanger sequencing. We also thank Junko Munakata-Marr, Hallgerd Eydal, Jackson Lee, Jason Sahl, Shannon Ulrich and Chase Williamson for helpful edits and comments on this manuscript and Burnham Petrographics for producing the thin sections. Lastly, thanks go to Charles Robertson for helpful discussions on covariance models.

References

Allwood, A.C., Walter, M.R., Kamber, B.S., Marshall, C.P., and Burch, I.W. (2006) Stromatolite reef from the early archaean era of Australia. *Nature* **441:** 714–718.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403–410.

Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* **59:** 143–169.

Andres, M.S., and Reid, R.P. (2006) Growth morphologies of modern marine stromatolites: a case study from Highborne Cay, Bahamas. *Sediment Geol* **185:** 319–328.

Anisimova, M., and Gascuel, O. (2006) Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst Biol* **55:** 539–552.

Awramik, S.M., and Riding, R. (1988) Role of algal eukaryotes in subtidal columnar stromatolite formation. *Proc Natl Acad Sci USA* **85:** 1327 –1329.

Awramik, S.M., and Sprinkle, J. (1999) Proterozoic stromatolites: the first marine evolutionary biota. *Hist Biol* **13:** 241–253.

Awramik, S.M., Grey, K., Hoover, R.B., Levin, G.V., Rozanov, A.Y., and Gladstone, G.R. (2005) Stromatolites: biogenicity, biosignatures, and bioconfusion. In *Astrobiology and Planetary Missions*, Vol. 5906. Hoover, R.B., Levin, G.V., Rozanov, A.Y., and Gladstone, G.R. (eds). San Diego, CA, USA: SPIE, pp. 59060P–590609.

Baumgartner, L., Spear, J., Buckley, D., Pace, N., Reid, R., Dupraz, C., and Visscher, P. (2009) Microbial diversity in modern marine stromatolites, Highborne Cay, Bahamas. *Environ Microbiol* **11:** 2710–2719.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Sayers, E.W. (2011) GenBank. *Nucleic Acids Res* **39:** D32–D37.

Berelson, W.M., Corsetti, F.A., Pepe-Ranney, C., Hammond, D.E., Beaumont, W., and Spear, J.R. (2011) Hot spring siliceous stromatolites from Yellowstone National Park: assessing growth rate and laminae formation. *Geobiology* **9:** 411–424.

Bunge, J. (2011) Estimating the number of species with CatchAll. Pacific Symposium on Biocomputing Proceedings 2011. Fairmont Orchid, Hawaii.

Burns, B.P., Goh, F., Allen, M., and Neilan, B.A. (2004) Microbial diversity of extant stromatolites in the hypersaline marine environment of Shark Bay, Australia. *Environ Microbiol* **6:** 1096– 1101.

Cannone, J., Subramanian, S., Schnare, M., Collett, J., D'Souza, L., Du, Y., *et al*. (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs: correction. *BMC Bioinformatics* **3:** 15.

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

Castenholz, R. (1978) The biogeography of hot spring algae through enrichment cultures. *Mitt Int Ver Limnol* **21:** 296–315.

Castenholz, R.W. (1969) The thermophilic cyanophytes of Iceland and the upper temperature limit. *J Phycol* **5:** 360–368.

Castenholz, R.W. (1977) The effect of sulfide on the blue-green algae of hot springs II. Yellowstone National Park. *Microb Ecol* **3:** 79–105. Available at: [Accessed August 24, 2011].

Castenholz, R.W. (2001) Bergey's Manual of Systematic Bacteriology. Volume 1: Phylum BX. Cyanobacteria. Oxygenic photosynthetic bacteria. In *The Archaea and the Deeply Branching and Phototropic Bacteria*, Vol. 1, 2nd edn. Garrity, G.M., Boone, D.R., and Castenholz, R.W. (eds). New York, USA: Springer-Verlag, pp. 473–599.

Chizuru, T., and Akihiro, K. (2005) Depositional processes of travertine developed at Shionoha Hot Spring, Nara Prefecture, Japan. *J Geol Soc Jpn* **111:** 751–764.

Cole, J.R., Chai, B., Farris, R.J., Wang, Q., Kulam, S.A., McGarrell, D.M., *et al*. (2005) The ribosomal database project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res* **33:** D294–D296.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., *et al*. (2009) The ribosomal database project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* **37:** D141–D145.

DeSantis, T., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E., Keller, K., *et al*. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72:** 5069–5072.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:** 2460–2461.

Elser, J.J., Schampel, J.H., Garcia-Pichel, F., Wade, B.D., Souza, V., Eguiarte, L., *et al*. (2005) Effects of phosphorus enrichment and grazing snails on modern stromatolitic microbial communities. *Freshw Biol* **50:** 1808–1825.

Ewing, B., and Green, P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8:** 186–194.

Ewing, B., Hillier, L., Wendl, M.C., and Green, P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8:** 175–185.

Fay, P. (1992) Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev* **56:** 340– 373.

Finsinger, K., Scholz, I., Serrano, A., Morales, S., Uribe-Lorio, L., Mora, M., *et al*. (2008) Characterization of true-branching cyanobacteria from geothermal sites and hot springs of Costa Rica. *Environ Microbiol* **10:** 460–473.

Foster, J., and Green, S. (2011) Microbial diversity in modern stromatolites. In *Stromatolites: Interaction of Microbes with Sediments, Cellular Origin, Life in Extreme Habitats and Astrobiology*. Tewari, V., and Seckbach, J. (eds). Dordrecht, Netherlands: Springer, pp. 385–405.

Foster, J.S., Green, S.J., Ahrendt, S.R., Golubic, S., Reid, R.P., Hetherington, K.L., and Bebout, L. (2009) Molecular and morphological characterization of cyanobacterial diversity in the stromatolites of Highborne Cay, Bahamas. *ISME J* **3:** 573–587.

Goh, F., Allen, M.A., Leuko, S., Kawaguchi, T., Decho, A.W., Burns, B.P., and Neilan, B.A. (2008) Determining the specific microbial populations and their spatial distribution within the stromatolite ecosystem of Shark Bay. *ISME J* **3:** 383–396.

Golubic, S., and Focke, J.W. (1978) *Phormidium hendersonii* Howe; identity and significance of a modern stromatolite building microorganism. *J Sediment Res* **48:** 751–764.

Grotzinger, J.P., and Knoll, A.H. (1999) Stromatolites in precambrian carbonates: evolutionary mileposts or environmental dipsticks? *Annu Rev Earth Planet Sci* **27:** 313–358.

Haas, B.J., Gevers, D., Earl, A., Feldgarden, M., Ward, D.V., Giannokous, G., *et al*. (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* **21:** 494–504.

Hamady, M., Walker, J., Harris, J., Gold, N., and Knight, R. (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* **5:** 235–237.

Havemann, S.A., and Foster, J.S. (2008) Comparative characterization of the microbial diversities of an artificial microbialite model and a natural stromatolite. *Appl Environ Microbiol* **74:** 7410–7421.

Hoffman, P. (1976) Chapter 6.1: Stromatolite morphogenesis in Shark Bay, Western Australia. In *Stromatolites*, Vol. 20. Walter, M.R. (ed.). Amsterdam, The Netherlands: Elsevier, pp. 261–271.

Hoffmann, L., Komarek, J., and Kastovsky, J. (2005) System of cyanoprokaryotes (cyanobacteria) state in 2004. *Algol Stud* **117:** 95–115.

Honda, D., Yokota, A., and Sugiyama, J. (1999) Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. *J Mol Evol* **48:** 723–739.

Hong, S., Bunge, J., Jeon, S., and Epstein, S.S. (2006) Predicting microbial species richness. *Proc Natl Acad Sci USA* **103:** 117–122.

Huse, S., Huber, J., Morrison, H., Sogin, M., and Welch, D.M. (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* **8:** R143.

Jones, B., Renaut, R.W., and Rosen, M.R. (1998) Microbial biofacies in hot spring sinters; a model based on Ohaaki Pool, North Island, New Zealand. *J Sediment Res* **68:** 413–434.

Knight, R., Maxwell, P., Birmingham, A., Carnes, J., Caporaso, J.G., Easton, B., *et al*. (2007) PyCogent: a toolkit for making sense from sequence. *Genome Biol* **8:** R171.

Lane, D. (1991) 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacteria Systematics*. Stackebrandt, E., and Goodfellow, M. (eds). Chichester, UK: Wiley and Sons, p. 329.

Liu, Z., Lozupone, C., Hamady, M., Bushman, F., and Knight, R. (2007) Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res* **35:** e120.

Liu, Z., DeSantis, T.Z., Andersen, G.L., and Knight, R. (2008) Accurate taxonomy assignments from 16S rRNA

sequences produced by highly parallel pyrosequencers. *Nucleic Acids Res* **36:** e120.

Lozupone, C., and Knight, R. (2007) Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* **104:** 11436–11440.

Ludwig, W., Strunk, O., Klugbauer, S., Klugbauer, N., Weizenegger, M., Neumaier, J., *et al*. (1998) Bacterial phylogeny based on comparative sequence analysis (review). *Electrophoresis* **19:** 554–568.

Matsen, F., Kodner, R., and Armbrust, E.V. (2010) pplacer: linear time maximum-likelihood and bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **11:** 538.

Miller, S.R., and Bebout, B.M. (2004) Variation in sulfide tolerance of photosystem II in phylogenetically diverse cyanobacteria from sulfidic habitats. *Appl Environ Microbiol* **70:** 736–744.

Miller, S.R., Purugganan, M.D., and Curtis, S.E. (2006) Molecular population genetics and phenotypic diversification of two populations of the thermophilic *Cyanobacterium Mastigocladus laminosus*. *Appl Environ Microbiol* **72:** 2793–2800.

Miller, S.R., Strong, A.L., Jones, K.L., and Ungerer, M.C. (2009a) Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl Environ Microbiol* **75:** 4565–4572.

Miller, S.R., Williams, C., Strong, A.L., and Carvey, D. (2009b) Ecological specialization in a spatially structured population of the thermophilic *Cyanobacterium Mastigocladus laminosus*. *Appl Environ Microbiol* **75:** 729–734.

Nawrocki, E.P., Kolbe, D.L., and Eddy, S.R. (2009) Infernal 1.0: inference of RNA alignments. *Bioinformatics* **25:** 1335–1337.

Okumura, T., Shiraishi, F., Naganuma, T., Yukimura, K., Takashima, C., Nishida, S., *et al*. (2008) Microbial contribution to lamina formation in an aragonite travertine. In *Geobiology of Stromatolithes: International Kalkowsky-Symposium Gottingen, October 4-11.2008; Abstract Volume and Field Guide to Excursions*. Reitner, J., Queric, N., and Reich, M. (eds). Göttingen, Germany: Universitatsverlag Gottingen, pp. 105–106.

Osburn, M.R., Sessions, A.L., Pepe-Ranney, C., and Spear, J.R. (2011) Hydrogen-isotopic variability in fatty acids from Yellowstone National Park hot spring microbial communities. *Geochim Cosmochim Acta* **75:** 4830–4845.

Paerl, H.W., Steppe, T.F., and Reid, R.P. (2001) Bacterially mediated precipitation in marine stromatolites. *Environ Microbiol* **3:** 123–130.

Papineau, D., Walker, J.J., Mojzsis, S.J., and Pace, N.R. (2005) Composition and structure of microbial communities from stromatolites of Hamelin Pool in Shark Bay, Western Australia. *Appl Environ Microbiol* **71:** 4822–4832.

Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R.P., Moret, B.M.E., and Stamatakis, A. (2010) How many bootstrap replicates are necessary? *J Comput Biol* **17:** 337–354.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B., Ludwig, W., Peplies, J., and Glockner, F. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35:** 7188–7196.

Reeder, J., and Knight, R. (2010) Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nat Methods* **7:** 668–669.

Reid, R.P., Visscher, P.T., Decho, A.W., Stolz, J.F., Bebout, B.M., Dupraz, C., *et al*. (2000) The role of microbes in accretion, lamination and early lithification of modern marine stromatolites. *Nature* **406:** 989–992.

Sahl, J.W., Fairfield, N., Harris, J.K., Wettergreen, D., Stone, W.C., and Spear, J.R. (2010) Novel microbial diversity retrieved by autonomous robotic exploration of the world's deepest vertical phreatic sinkhole. *Astrobiology* **10:** 201–213.

Santos, F., Peña, A., Nogales, B., Soria-Soria, E., Del Cura, M.A.G., González-Martín, J.A., and Antón, J. (2010) Bacterial diversity in dry modern freshwater stromatolites from Ruidera Pools Natural Park. *Spain Syst Appl Microbiol* **33:** 209–221.

Seong-Joo, L., and Golubic, S. (1998) Multi-trichomous cyanobacterial microfossils from the Mesoproterozoic Gaoyuzhuang Formation, China: paleoecological and taxonomic implications. *Lethaia* **31:** 169–184.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., *et al*. (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* **103:** 12115–12120.

Spear, J.R., Walker, J.J., McCollom, T., and Pace, N.R. (2005) Hydrogen and Bioenergetics in the Yellowstone Geothermal Ecosystem. *Proc Natl Acad Sci USA* **102:** 2555–2560.

Stal, L.J., Paerl, H.W., Bebout, B., and Villbrandt, M. (1993) Heterocystous versus non- heterocystous cyanobacteria in microbial mats. In *Microbial Mats: Sturcture, Development and Environmental Significance, Vol. 35 of Ecological Sciences*. Stal, L.J., and Caumette, P. (eds). New York, USA: Springer-Verlag, pp. 403–414.

Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688–2690.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol* **57:** 758–771.

Takashima, C., and Kano, A. (2008) Microbial processes forming daily lamination in a stromatolitic travertine. *Sediment Geol* **208:** 114–119.

Tenreiro, S., Nobre, M.F., and Costa, M.S. (1995) *Thermus silvanus* sp. nov. and *Thermus chliarophilus* sp. nov., two new species related to thermus ruber but with lower growth temperatures. *Int J Syst Bacteriol* **45:** 633–639.

Tomitani, A. (2006) The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives. *Proc Natl Acad Sci USA* **103:** 5442–5447.

Turner, S., Pryer, K.M., Miao, V.P.W., and Palmer, J.D. (1999) Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J Eukaryot Microbiol* **46:** 327–338.

Visscher, P.T., Reid, R.P., Bebout, B.M., Hoeft, S.E., Macintyre, I.G., and Thompson, J.A. (1998) Formation of lithified micritic laminae in modern marine stromatolites

(Bahamas); the role of sulfur cycling. *Am Mineral* **83:** 1482–1493.
Visscher, P.T., Reid, R.P., and Bebout, B.M. (2000) Microscale observations of sulfate reduction: correlation of microbial activity with lithified micritic laminae in modern marine stromatolites. *Geology* **28:** 919 –922.
Walter, M.R., Bauld, J., and Brock, T.D. (1972) Siliceous algal and bacterial stromatolites in hot spring and geyser effluents of Yellowstone National Park. *Science* **178:** 402–405.
Ward, D.M., and Castenholz, D.W. (2000) Cyanobacterial mats and stromatolites. In *The Ecology of Cyanobacteria: Their Diversity in Time and Space*. Whitton, B.A., and Potts, M. (eds). Dordrecht, Netherlands: Kluwer Academic, pp. 61–120.
Ward, D.M., Ferris, M.J., Nold, S.C., and Bateson, M.M. (1998) A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol Mol Biol Rev* **62:** 1353–1370.
Wilmotte, A., and Herdman, M. (2001) Phylogenetic relationships among the cyanobacteria based on 16S rRNA sequences. In *Bergey's Manual of Systematic Bacteriology. Volume 1: The Archaea and the Deeply Branching and Phototropic Bacteria*, 2nd edn. Garrity, G.M., Boone, D.R., and Castenholz, R.W. (eds). New York, USA: Springer-Verlag, pp. 487–493.
Wilmotte, A., der Auwera, G.V., and Wachter, R.D. (1993) Structure of the 16S ribosomal RNA of the thermophilic cyanobacterium *Chlorogloeopsis* HTF ('*Mastigocladus laminosus* HTF') strain PCC7518, and phylogenetic analysis. *FEBS Lett* **317:** 96–100.
Zhongying, Z. (1986) Solar cyclicity in the precambrian microfossil record. *Palaeontology* **29:** 101–111.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Maximum likelihood tree depicting the phylogenetic placement of *Chlorogloeopsis* and *Fischerella* phylotypes found in this study. We are using Group IV cyanobacteria as the outgroup for this tree. *Gloeobacter violaceus* (PCC 7421), commonly used as an outgroup in cyanobacteria phylogenies, falls in Group IV according in the RDP taxonomic organization of cyanobacteria. Outlined sequences were generated in this study. The asterisk '*' denotes sequences that were placed into the tree by pplacer. Values at nodes are bootstrap support percentages. Nodes without support values were created by pplacer insertion. Polytomies denote undefined descendant branching order based on bootstrap values (i.e. nodes with bootstrap support < 50% were collapsed). Filled clades have cultured representation.
Fig. S2. Phylogeny of all cyanobacterial 16S SSU rRNA gene sequences from eight molecular studies of modern stromatolite studies as reviewed by Foster and Green (2011). Accession numbers and short descriptions for reference sequences used to build the backbone tree are shown. Arrow indicates position of 'SFM-seq' tree. The qualitative salinity for each stromatolite derived 16S rRNA gene sequence is also shown. Clades with bootstrap values below 50% were collapsed and therefore polytomies indicate ambiguous branching order. *Gloebacter violaceus* PCC 7421 is used as the outgroup.
Fig. S3. Distribution of 16S rRNA gene non-cyanobacterial pyrosequences into 97% identity OTUs. OTUs with less than or equal to five total 16S rRNA gene sequences summed across every pyrosequence library are omitted from the figure.
Table S1. Best BLAST hit in SSURef104 to major non-cyanobacterial OTUs.
Appendix S1. Supplementary methods.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.