RESEARCH ARTICLE

Microbialite genetic diversity and composition relate to environmental variables

Carla M. Centeno[1], Pierre Legendre[2], Yislem Beltrán[1], Rocío J. Alcántara-Hernández[1], Ulrika E. Lidström[3], Matthew N. Ashby[3] & Luisa I. Falcón[1]

[1]Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico; [2]Département de Sciences Biologiques, Université de Montréal, Montréal, QC, Canada; and [3]Taxon Biosciences, Inc., Tiburon, CA, USA

Correspondence: Luisa I. Falcón, Instituto de Ecología, Universidad Nacional Autónoma de México, 3er Circuito Exterior sn, Ciudad Universitaria, Del. Coyoacán, Mexico D.F. 04510. Tel.: +52 55 5622 8222, ext. 46869; fax: 52 55 5622 8995; e-mail: falcon@ecologia.unam.mx

Received 30 March 2012; revised 28 June 2012; accepted 29 June 2012.
Final version published online 2 August 2012.

DOI: 10.1111/j.1574-6941.2012.01447.x

Editor: Gary King

Keywords
microbialites; microbial communities; diversity; composition; pyrosequencing.

Abstract

Microbialites have played an important role in the early history of life on Earth. Their fossilized forms represent the oldest evidence of life on our planet dating back to 3500 Ma. Extant microbialites have been suggested to be highly productive and diverse communities with an evident role in the cycling of major elements, and in contributing to carbonate precipitation. Although their ecological and evolutionary importance has been recognized, the study of their genetic diversity is yet scanty. The main goal of this study was to analyse microbial genetic diversity of microbialites living in different types of environments throughout Mexico, including desert ponds, coastal lagoons and a crater-lake. We followed a pyrosequencing approach of hypervariable regions of the 16S rRNA gene. Results showed that microbialite communities were very diverse ($H' = 6–7$) and showed geographic variation in composition, as well as an environmental effect related to pH and conductivity, which together explained 33% of the genetic variation. All microbialites had similar proportions of major bacterial and archaeal phyla.

Introduction

Microbialites are complex organo-sedimentary structures that trap, bind and precipitate detritus and minerals such as carbonates, silicates and sulphates (Dupraz *et al.*, 2009) through the activity of their biological constituents (Burne & Moore, 1987; Paerl *et al.*, 2001; Dupraz & Visscher, 2005; Desnues *et al.*, 2008). Hence, their characteristics and structural properties are the products of biological and ecological factors (Foster *et al.*, 2009; Last *et al.*, 2010). They can show various degrees of lithification (Sprachta *et al.*, 2001), differences in morphology, mode of growth and substrate (Pringault *et al.*, 2005). The grain size of the substrate, capillary attraction of water, penetration of light, sedimentation–erosion rates, wave action, grazing pressure (van Gemerden, 1993), inorganic and biologically mediated calcification (Burne & Moore, 1987) are fundamental for microbialite development (Dupraz & Visscher, 2005).

Extant microbialites are the nearest analogues to fossilized stromatolites, which consist of the oldest evidence of life on Earth, dating back to 3500 Ma (Allwood *et al.*, 2006; Schopf *et al.*, 2007). The microbial communities associated to these fabrics establish complex ecological associations that occur vertically, stratified in visible millimeter scale layers (DesMarais, 1995; Stal, 1995), forming small-scale ecosystems constituted by a variety of metabolically and phylogenetically distinct taxa (Burne & Moore, 1987; Paerl *et al.*, 2001). Microbialites have been argued to be the first organized ecosystems on the planet, where transformations of major constitutive elements are coupled to microbial metabolism (van Gemerden, 1993; Decho *et al.*, 2005; Leuko *et al.*, 2007; Paterson *et al.*, 2008). The actively growing portion of microbialites, which occurs in the surface layers, has been compared in terms of productivity (5000–6200 mg cm^{-2} day^{-1}) to tropical forests (6000 mg cm^{-2} day^{-1}), where a vast contribution of phototrophs, both oxygenic and anoxygenic, has been recognized (Stal, 2000). Microbialites are also known for their role in the N cycle harboring an important diversity of N_2-fixing phyla (Stal *et al.*, 2010).

Despite the ecological and evolutionary relevance of microbialites, the study of their genetic composition is

FEMS MICROBIOLOGY ECOLOGY

relatively recent (Burns *et al.*, 2004; Papineau *et al.*, 2005; Mobberley *et al.*, 2012). Historically, Cyanobacteria have been the most studied phylum, and their relevance as phototrophs and diazotrophs in the formation of the microbial tapestry and morphogenesis of microbialites has been recognized (Reid *et al.*, 2000; Neilan *et al.*, 2002; Dupraz & Visscher, 2005; Pringault *et al.*, 2005; Kromkamp *et al.*, 2007; Foster *et al.*, 2009; Gischler *et al.*, 2011). However, the identification of other bacterial and archaeal taxa associated to microbialites has remained limited (Breitbart *et al.*, 2009; Burns *et al.*, 2009; Bolhuis & Stal, 2011; Mobberley *et al.*, 2012). Although robust diversity estimations have been done using different molecular ecology techniques, to date no single study had been conducted in which both geographic and environmental gradients were included to identify potential abiotic variables that contribute to microbialite community composition. Microbialites are known to occur in several environments throughout Mexico (Gischler *et al.*, 2008, 2011; Breitbart *et al.*, 2009; Kazmierczak *et al.*, 2011; Beltrán *et al.*, 2012), a mega diverse country, in which species turnover increases due to strong environmental gradients. The aim of this study was to analyse whether microbial community composition, described using a pyrosequencing approach of hypervariable regions of the 16S rRNA gene, had a structuring effect related to environmental variables along a latitudinal gradient.

Materials and methods

Site description and sampling

Microbialites were collected in the summer of 2009 from 10 localities distributed throughout Mexico (Fig. 1, Table 1).

Northern Mexico

The Cuatro Cienegas Basin (CCB) consists of the largest oasis in the Chihuahuan Desert, where microbialites are conspicuous features in some of the more than 200 ponds, lakes and streams. The region is characterized by steeply folded Cretaceous sediments (Johannesson *et al.*, 2004), where 'Sierra de San Marcos' is a faulted anticline composed of cretaceous shallow water limestone, dolomite and gypsum (Garcia-Pichel *et al.*, 2004). Sampling sites included the Pozas Azules Ponds (PI and PII) that are found along the 'Sierra de San Marcos', where the vertical canyon slopes are rimmed by Quaternary karstified alluvial fans, giving way to a flat desert floor (Winsborough *et al.*, 1994) where we sampled the Rio Mesquites (RM) permanent stream.

Central Mexico

This region is characterized by the Trans Mexican Volcanic Belt, including several crater-lakes. Alchichica is the largest of these limnetic bodies and is a maars-alkaline and athalasohaline system with particular features including relatively high electric conductivity, subsaline water, as well as high contents of magnesium and calcium carbonates (Couradeau *et al.*, 2011). Caballero *et al.* (2003) mentioned that the main factors responsible for the chemistry of this lake are influx of water rich in sodium from volcanic materials, and bicarbonates from Cretaceous limestone. Previous work by Tavera & Komárek (1996) described columnar and spongy microbialites in Alchichica, following a microscopy-based identification strategy. Recent studies by Kazmierczak *et al.* (2011) identified that both microbialite morphologies comprised two generations and had differences in mineral composition, the spongy-types being composed mainly of hydromagnesite with an admixture of huntite and calcite, and dating to approximately 2800 years B.P., and the columnar-types being composed of aragonite and an admixture of Mg-calcite, dating to approximately 1100 years B.P. The different morphologies of Alchichica's microbialites are attributed to different environmental conditions, where the spongy forms would relate to a dry period and the columnar forms to a wet climate associated to large inflow of ground water. In this study we included both microbialite morphotypes (columnar AC, and spongy AS).

Southern Mexico-Yucatan Peninsula

The Yucatan Peninsula is a carbonate platform that began to emerge above sea level during the Oligocene and Miocene (Ferrusquia-Villafranca, 1993). The entire peninsula consists of karstic landscapes (Gischler *et al.*, 2008, 2011). In this study we sampled two coastal lagoons: Muyil, inside the Sian Kaʻan Biosphere Reserve, where we included microbialites (SS) and a microbial mat (SM); Laguna Bacalar, near the Mexico–Belize border, is the freshwater site with the largest occurrence of extant microbialites (Gischler *et al.*, 2008, 2011). Sampling included the Pirate Channel (BP), Los Rapidos (BR) and a soft, low degree of induration formation (BM). Bacalar microbialites have been radiometrically dated between 6790 and 9190 years B.P., corresponding to early Holocene (Gischler *et al.*, 2008), making them the oldest known extant microbialites.

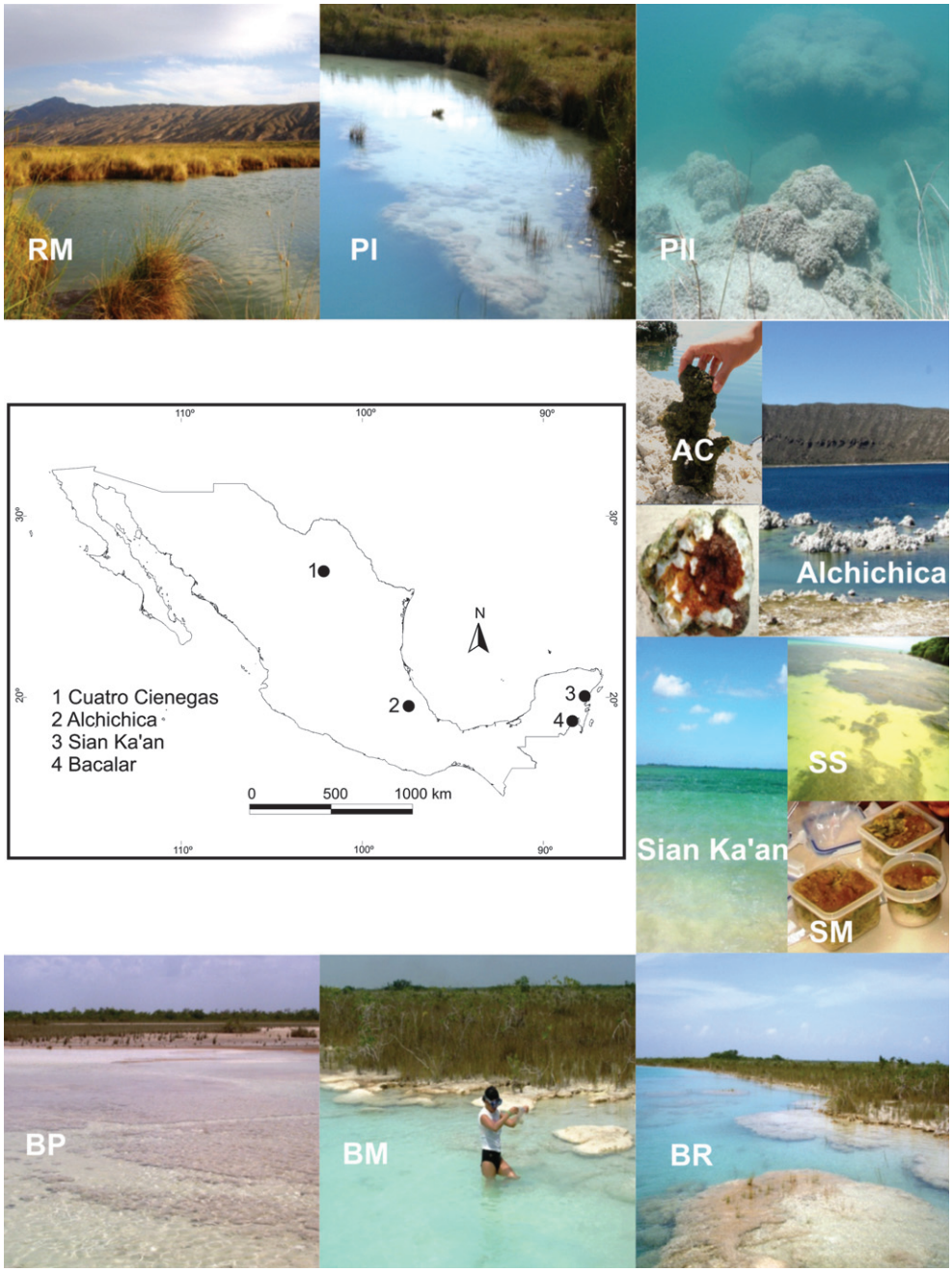

Fig. 1. Geographic location of the sampling sites. North Mexico, CCB: PI (Pozas Azules I), PII (Pozas Azules II) and RM (Rio Mesquites); Central Mexico, Alchichica crater-lake: (AC) columnar and (AS) spongy morphologies. Southern Mexico, Yucatan Peninsula, Sian Ka'an lagoon microbialite (SS) and microbial mat (SM); Bacalar lagoon, Pirate Channel (BP), soft microbialite (BM) and Los Rapidos (BR).

Samples from the surface layers (first 2 cm) of six microbialites (~ 10 g) were collected at each locality, placed into sterile and sealed containers, and frozen (−20 °C) until analysis. The total distribution of each microbialite reef formation was recorded with a GPS and the six sampling points were located along the distribution rage. Water samples (600 mL per site) were collected in sterile containers adjacent to microbialites and used as community composition controls.

Physical and biochemical analysis

Environmental variables (temperature, conductivity and pH) were determined *in situ* at each sampling point with

Table 1. Location and description of study sites

Locality	Site	Key	Geographic location	T (°C)	pH	Conductivity (mS cm^{-1})	Microbialite morphology
Yucatan Peninsula	Sian Ka'an microbialite	SS	20°02.59′N, 87°36.460′W	29.9	7.70	1.1	Continuous crust
	Sian Ka'an microbial mat	SM	20°04.40′N, 87°36.360′W	30.3	8.13	1.4	Microbial mat
	Bacalar Medio	BM	18°37.698′N, 88°29.543′W	32.0	7.80	2.2	Domes, low degree of lithification
	Bacalar Pirate Channel	BP	18°38.128′N, 88°23.087′W	31.8	7.90	2.4	Continuous crust
	Bacalar Rapidos	BR	18°35.347′N, 88°26.098′W	31.5	7.60	2.2	Coalesced giant domes, high degree of lithification
CCB	Rio Mesquites	RM	26°55.058′N, 102°06.148′W	29.5	7.60	2.9	Domes, high degree of lithification
	Pozas Azules I	PI	26°49.654′N, 102°01.460′W	28.8	7.40	2.7	Domes, low degree lithification
	Pozas Azules II	PII	26°49.607′N, 102°01.390′W	29.8	7.40	2.8	Domes, low degree lithification
Central Mexico	Alchichica Columnar	AC	19°24.361′N, 97°24.170′W	18.5	9.40	13.0	Domes, irregular shape, high degree lithification
	Alchichica Spongy	AS	19°24.998′N, 97°24.682′W	17.9	9.70	13.0	Columnar shape, high degree lithification

Sian Ka'an microbialite (SS) and mat (SM); BM, Bacalar lagoon; BM, soft microbialite; BP, Pirate Channel; BR, Los Rapidos; RM, CCB: Rio Mesquites; PI, Pozas Azules I; PII, Pozas Azules II (PII); Alchichica crater-lake: columnar (AC) and spongy (AS).

a Hydrolab Multisensor YSI (600 QS; OH) (Table 1). Total dissolved phosphates (PO_4^{3-}), nitrites (NO_2^-), nitrates (NO_3^-), and ammonium (NH_4^+) were quantified colorimetrically for the water column samples with a Bran Luebbe 3 autoanalyser (Norderstedt, Germany). Microbial carbon, nitrogen and phosphorous (Cmic, Nmic, Pmic), were measured for all microbialite samples by the fumigation-extraction protocols of Vance *et al.* (1987) and Joergensen & Mueller (1996). Cmic, Nmic and Pmic were expressed based on dry weight of samples, using at least three replicates per site (Table 2).

DNA extraction

Microbialites

Total community DNA from each sampling point (six sampling points, 10 locations = 60 samples) was extracted following Zhou *et al.* (1996) with modifications. To obtain a homogenized mixture, approximately 5 g of each microbialite was macerated using liquid nitrogen and extraction buffer (100 mM Tris-HCl, 20 mM NaCl, 100 mM EDTA, pH 8) with cetyltrimethylammonium bromide (CTAB 6%). A freeze-thaw cycle with continuous

Table 2. Biogeochemical variables from each locality

Key*	Cmic (μg g^{-1})	Nmic (μg g^{-1})	Pmic (μg g^{-1})	NH_4^+ (mg L^{-1})	NO_3^- (mg L^{-1})
SS	193.78	59.62	1.02	0.242	0.026
SM	632.41	62.81	1.19	0.352	bdv
BM	156.56	68.50	0.29	0.144	0.16
BP	411.65	21.83	0.41	0.122	0.274
BR	490.74	103.46	5.26	0.062	0.028
RM	209.45	ND	1.86	0.132	0.434
PI	381.94	30.70	1.22	0.029	0.425
PII	539.68	76.96	1.35	bdv	0.353
AC	704.02	87.95	0.83	0.028	bdv
AS	1008.79	148.17	15.73	0.018	bdv

Phosphates and nitrites in water column were below detection values (bdv) for all sites.
ND, not determined; mic, microbial.
*See Table 1 footnotes for key abbreviations.

maceration of samples was repeated three times. Samples were incubated for 30 min at 37 °C with lysozyme (30 mg mL^{-1}) (Sigma Aldrich, Carlsbad, CA). A volume of 10% sodium dodecyl sulphate (SDS) and proteinase K (10 mg mL^{-1}; Sigma Aldrich) was added and samples were incubated overnight at 55 °C. Samples were extracted twice with phenol : chloroform : isoamyl alcohol (25 : 24 : 1) and once with chloroform : isoamyl alcohol (24 : 1), recovering the overlayer after each centrifugation (2 800 ***g*** for 15 min). DNA was precipitated by the addition of 2 volumes of chilled absolute ethanol, 1/10 volume of 3 M sodium acetate and 2 μL of GlycoBlue (Ambion Inc., Austin, TX) and refrigerated at −20 °C overnight. The precipitate DNA was recovered by centrifugation (11 000 ***g*** for 15 min) and the pellet was washed with 70% ethanol and resuspended in molecular grade water (Sigma Aldrich). The DNA pellet was further purified of inhibitors with the DNeasy Blood & Tissue kit (Qiagen, Valencia, CA) according to manufacturer's instructions. After this second cleanup, samples were precipitated with chilled absolute ethanol and 3 M sodium acetate. DNA pellets were resuspended in 30 μL molecular grade water (Sigma Aldrich).

Water column

Samples from each locality were thawed and filtered through Durapore membranes (0.22 μm) (Millipore, Billerica, MA). DNA was extracted with extraction buffer (100 mM Tris-HCl, 20 mM NaCl, 100 mM EDTA, pH 8), 1% volume of 10% SDS and proteinase K (10 mg mL^{-1}). Next, samples were incubated at 55 °C for 30 min and further subjected to two freeze–thaw cycles (−20 and 50 °C, respectively). Nucleic acids were extracted twice with phenol followed by a chloroform extraction as described above. DNA was precipitated overnight at −20 °C with 2 volumes of 100% ethanol, 3 M sodium acetate and 2 μL of GlycoBlue. DNA was recovered by centrifugation (11 000 ***g*** for 15 min) and the pellet was washed with 70% ethanol, dried and resuspended in 30 μL of molecular grade water (Sigma Aldrich).

Prokaryotic 16S rRNA gene amplification and sequencing

Since the aim of this study was to compare microbialite community composition from different geographic localities, where distinct environmental gradients occur, total community DNA extracted from each sampling point was pooled at the same concentrations into one compound sample per locality, to represent genetic variation. A ~ 600-bp fragment including the V5 and V6 hypervariable regions of the 16S rRNA gene was PCR-amplified with five replicates per locality and the combined amplified product was agarose gel-purified (QIAquick; Qiagen). Amplification was done with universal primers for Bacteria and Archaea TX9 (5′-GGATTAGAWACCCBGGTAGTC-3′) and 1391R (5′-GACGGGCRGTGWGTRCA-3′) (Ashby *et al.*, 2007). Thermal cycling conditions included an initial denaturation step at 94 °C for 1 min, followed by 20–30 cycles (depending on the sample, see below) of denaturation at 94 °C for 30 s, primer annealing at 55 °C for 30 s, strand extension at 68 °C for 30 s and a final extension step at 68 °C for 2 min. To avoid saturation of amplicons, the optimal number of cycles per PCR reaction was identified when a band of 20 ng mL^{-1} was detected with gel electrophoresis. The PCR reaction for each locality consisted of a final volume of 50 μL and contained 2 μL of pooled DNA, 5 μL of 10× High Fidelity PCR Buffer, 2 mM $MgSO_4$, 0.4 μM each primer, 0.2 mM of each deoxynucleotide triphosphate, 1.5 U High Fidelity Platinum Taq DNA Polymerase (Invitrogen, Carlsbad, CA), 0.5 mg mL^{-1} of bovine serine albumin and 0.25% dimethyl sulphoxide.

In preparation for pyrosequencing on a Genome Sequencer FLX Titanium 454 Life Sciences System (Roche 454 Life Sciences, Branford, CT), a second round of PCR was performed using fusion primers that incorporated a variable length sample-specific nucleotide sequence for each locality ('barcodes'), and the 454 pyrosequencing adaptor A onto the 5′ end of the TX9 primer, and adaptor B onto the 5′ end of the 1391R primer. The barcodes enabled multiplexing several samples into a single 454 sequencing run, after PAGE purifying and quantitating the amplicons. The multiplexed amplicon library was sequenced using the standard 454 Life Sciences Lib-L emulsion PCR protocol and Titanium chemistry (Margulies *et al.*, 2005). Raw sequence reads were quality-filtered using the default 454 parameters as well as post instrument filters, which required all bases be Q20 or higher and the average of all bases in any read to be Q25 or greater. Sequences were trimmed to remove the TX9 primer from the 5′ end and at a conserved site distal to V6 on the 3′ (ca. position 1067, *Escherichia coli* numbering), resulting in an approximately 250-bp read covering the V5 and V6 regions. Raw data were first analysed with the standard Roche 454-signal processing with the GS FLX system software. Sequences that passed quality-filtering were then used to generate the dataset. Finally, to avoid diversity overestimation of the rare biosphere and chimeric sequences originated from pyrosequencing noise, only sequences occurring at least five times in the database were included in subsequent analysis. Pyrosequences were deposited in the Sequence Read Archive (SRA), NCBI, under BioProject SRA040061.

Diversity estimations

Diversity, richness and coverage estimations were calculated using MOTHUR v.1.14.0 software (Schloss *et al.*, 2009). Sequences were aligned using the Greengenes core set (May 2009). dnadist function was used to create distance matrices and operational taxonomic units (OTUs) clustered by pair-wise identity at 97% similarity using the furthest neighbour algorithm. To assess microbial genetic diversity of each locality, rarefaction curves and Goods *C* coverage were calculated, where: $C = 1 - (n/N)$, n = number of OTUs that have been sampled once, and N = total number of OTUs in sample. Further, two non-parametric richness estimators (ACE and Chao.1) plus the Shannon diversity index (H') (a measure of α-diversity that incorporates relative abundances) based on sequence frequencies, were obtained.

The phylum composition of the microbialite communities, based on the 16S rRNA gene sequences from each sampling locality, was determined with the Naïve Bayesian rRNA Classifier from the Ribosomal Database Project (RDP-II), with a confidence value of at least 80% (Wang *et al.*, 2007).

Statistical analysis

Prior to clustering and ordination, the response frequency data (10 sites, 10 899 reads) were transformed using the Hellinger transformation (Legendre & Gallagher, 2001), since Euclidean distance-based analysis is not suitable for presence–absence genetic datasets. After transformation, the genetic sequence data were well explained by a linear redundancy analysis (RDA) model of the environmental variables (adjusted $R^2 = 0.479$, permutational P-value = 0.014), that was not the case before transformation (adjusted $R^2 = 0.117$, permutational P-value = 0.526). The Hellinger-transformed data were used to compute a Hellinger distance matrix (Legendre & Legendre, 2012), which was subjected to Ward hierarchical clustering and principal coordinate analysis (PCoA). Although nowadays the use of FastUnifrac (Hamady *et al.*, 2010) is common for microbial community studies, our dataset was not suitable for analysis with this platform, since significance values could not be calculated for the vast number of locations here reported.

Environmental variables were correlated to the PCoA ordination axes and plotted on the principal coordinate diagram to form a biplot, as described in Legendre & Legendre (2012). Because these variables are intercorrelated, stepwise selection in RDA was used to identify the variables that best explained the genetic variation among sites.

The cluster analysis was obtained using the hclust() function of the STATS package in R (http://www.r-project.org/). Principal coordinate ordination was computed using the pcoa() function of the APE package (Paradis *et al.*, 2010). Stepwise selection of the environmental variables was computed using the ordistep() function of the VEGAN package (Oksanen *et al.*, 2011).

Results and discussion

Genetic diversity

Raw data consisted of 271 651 pyrosequence reads (average length 250 bp), of which 192 276 unique sequences were recovered after quality-filtering, where 10 899 sequences were seen at least five times of more. Site-specific diversity at a 97% cut-off (Fierer & Lennon, 2011) indicated that the microbial communities from coastal lagoons in the Yucatan Peninsula (BM, BP, BR, SM, SS) had the highest genetic diversity, followed by sites in the desert oasis of the CCB (PI, PII, RM); microbialites from Alchichica crater-lake (AC, AS) had the lowest diversity (Table 3). The coverage at this cut-off level was on average 65%, with high values of richness ($ACE_X = 4499$, $Chao.1_X = 3491$) and diversity ($H'_X = 6.85$), indicating that we had successfully represented the microbial communities, making these some of the most diverse communities known. Work by Bolhuis & Stal (2011) employed massive parallel 16S rRNA gene sequencing to analyse coastal bacterial and archaeal microbial mats along a tidal gradient, finding highly diverse communities ($H'_X = 5.9$). More recently, Mobberley *et al.* (2012) used pyrosequencing to analyse microbial communities of four thrombolytic types from Highborne Cay, Bahamas, revealing that each thrombolite was a distinctive community with high diversity values ($H'_X = 5.35$).

Previous work by Hamp *et al.* (2009) suggested that pyrosequencing reads should only be included when seen

Table 3. Estimated genetic richness (ACE and Chao.1 indexes), diversity (Shannon = H') and coverage for all study sites

Site*	No. sequence	OTUs	ACE	Chao.1	H'	Coverage (%)
SS	3686	1977	4749	4314	7.15	64.41
SM	2628	1501	5449	3648	6.90	59.93
BM	3108	1804	4373	4110	7.18	60.71
BP	3001	1527	3451	3156	6.88	67.48
BR	3762	2054	7322	4729	7.19	62.84
RM	2567	1517	5427	3681	7.00	59.06
PI	3227	1624	3700	3419	6.83	67.40
PII	3351	1688	3838	3561	6.96	67.62
AC	2169	1060	3510	2271	6.32	67.50
AS	2040	922	3172	2026	6.10	70.29

We report sequences seen in the total dataset 5× or more.
OTU, operational taxonomic unit at a 97% cut-off.
*See Table 1 footnotes for site abbreviations.

10× or more in the complete database, after carefully comparing annotated *E. coli* genomes manually curated, sequenced both by Sanger and 454 technologies. They concluded that pyrosequencing generates noise that is avoided with the 10× cut-off for unique sequences. In this study, we analysed the dataset for 5× and 10× cut-offs (Supporting Information, Table S1) and decided to work with the first, as richness estimators ACE and Chao.1, and coverage increased as compared with the 10× dataset. The Shannon diversity index did not vary significantly, suggesting that new, non-abundant OTUs were recovered in the 5× dataset that were not seen at greater cut-offs. As the percentage of unidentified OTUs did not increase between the 5× and 10× cut-offs, we are confident that the richness and diversity estimations at 5× for environmental and highly diverse communities are valid. Nevertheless, due to the characteristics of the pyrosequencing technique, we do not recommend working with datasets that have not been carefully curated, in order to avoid overestimations of the rare biosphere due to pyrosequencing noise.

Dendrogram and ordination analysis of the microbial composition (Fig. 2a and b) showed greater genetic similarity within sites in the same geographic region. The microbialites from desert ponds and a permanent stream in the CCB shared more phylotypes with those from coastal lagoons in the Yucatan Peninsula despite the fact that they are separated by more than 2500 km (Fig. 1). Yucatan Peninsula and CCB are karstic oligotrophic Quaternary systems, hence both present elevated calcium and carbonate concentrations in the water. Further similarities have been observed in the microbialite communities they harbour, such as similar thrombolitic morphologies, calcite as the principal mineralogy, and the presence of abundant grazers (bivalves and gastropods principally) (Gischler *et al.*, 2008). The analysis of 2169 sequences from the columnar microbialite and 2040 sequences from the spongy microbialite in Alchichica crater-lake indicate they have a different genetic composition, although sharing more phylotypes than with any other microbialite analysed here.

The analysis of spatial genetic variation following a stepwise selection procedure identified pH and conductivity as explanatory variables that together had an adjusted R^2 of 0.326 and a permutational P-value = 0.001 after 999 permutations, explaining 33% of the genetic variation among localities. Temperature was highly negatively correlated with pH and conductivity ($r = -0.91$ and -0.93, respectively). Microbial carbon, nitrogen and phosphorous (Cmic, Nmic, Pmic) were positively correlated with these two variables. Nitrates made a unique contribution to the explanation of the genetic variation separating the three CCB localities from the others (RDA P-value = 0.009) (Fig. 3). Elser *et al.* (2005) reported that CCB freshwater systems have high concentrations of dissolved inorganic nitrogen and low concentrations of P, suggesting a strong limitation of P on microbialite growth. In this study, the three localities sampled in the CCB had the highest levels of NO_3^-. Our results indicate that almost 70% of community structure could not be associated to the parameters analysed here, indicating the need to expand on temporal and spatial physicochemical characterization of the environments studied, as well as to focus on biological interactions, to fully understand microbial community composition.

This study suggests that microbialites that develop in similar environmental conditions have similar prokaryotic communities regardless of geographic distance, and that key environmental factors, identified in this study as pH and conductivity, affect microbialite communities diversity and structure. Fierer & Jackson (2006) pointed out that environmental factors such as pH influence the structuring of microbial communities which can be more important than geographic distance. Accordingly, Alchichica crater-lake had the highest values of conductivity and pH, and the communities with the lowest diversity, a pattern previously documented by others (Sabbe *et al.*, 2004; Zeglin *et al.*, 2011).

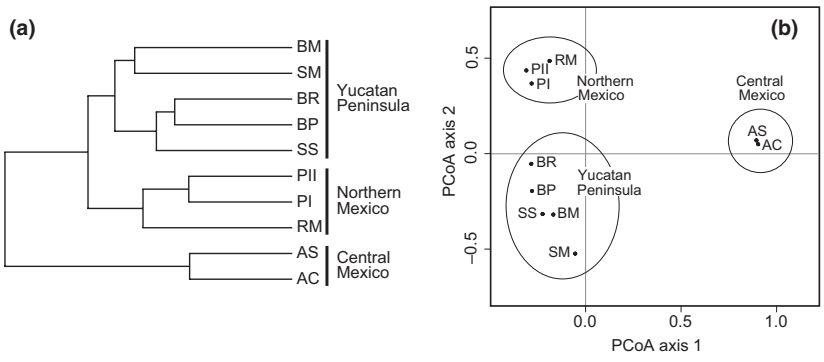

Fig. 2. (a) Hierarchical clustering dendrogram (Ward method) showing the genetic distance relationships among sites. (b) Principal coordinate ordination PCoA axis 1 accounts for 33% of the among-site variance, axis 2 accounts for 17%. The three study regions/environments are identified in both plots.

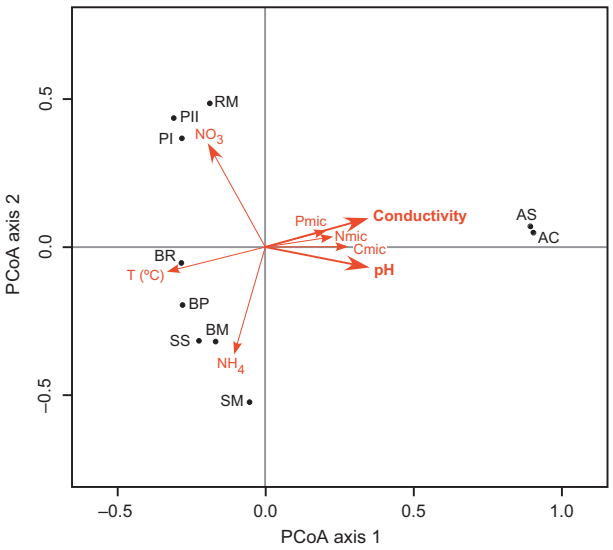

Fig. 3. Principal coordinate ordination biplot showing the study sites (points) and the explanatory environmental variables (arrows). The two most important variables, identified by stepwise selection in RDA, are in bold.

Community composition

Of the total number of sequences analysed, 98.7% were assigned to Bacteria and 1.3% to Archaea. A remarkable aspect of the community composition at the phyla level emerged from this analysis; all microbialites had the same dominant phyla. *Proteobacteria* were the most abundant phylum, ranging from 31% to 42%; *Alphaproteobacteria* had the highest frequencies, followed by *Betaproteobacteria*, *Gammaproteobacteria* and *Deltaproteobacteria*. *Planctomycetes* (5–11%), *Verrucomicrobia* (3–8%), *Cyanobacteria* (1–24%), *Bacteroidetes* (4–8%), *Acidobacteria* (1–2.4%), *Chloroflexi* (1–2.4%) and *Firmicutes* (1–3%) appeared less frequently (Fig. 4a). Phyla with low abundances included *Actinobacteria*, *Nitrospira*, *Chlamydiae*, *Spirochaetes*, *Chlorobi*, *Fusobacteria*, *Gemmatimonadetes* as well as candidate phyla (e.g. TM7, OP10, SR1), represented by few sequences (Fig. 4b). It is worth pointing out that unclassified bacteria represented 17–27% of total community composition (Fig. 4a). *Cyanobacteria* were more abundant in Alchichica crater-lake (22% and 24%, respectively; AC, AS) than in any other location, representing another difference with karstic, Yucatan Peninsula and CCB localities.

Archaeal sequences were present in low abundance (0.4–1.76%) in the 10 localities analysed, and were absent from one of the morphotypes of Alchichica crater-lake (AS) (Fig. 4c). Phylum *Euryarchaeota* was more abundant in Sian Ka′an lagoon (SS and SM). Phylum *Crenarchaeota* was found in all sites but was more abundant in Bacalar costal lagoon (BM). All sites had unclassified Archaea, which were most abundant in Sian Ka′an (SM and SS), Bacalar (BM) and Alchichica crater-lake (AC) (Fig. 4c).

The composition of control water column samples differed from that of microbialite communities below the phylum level. *Gematimonadetes*, *OD1*, *Fusobacteria*, *SRI*, *TM7* and *Spirochaetes* were not observed in water samples. Differences between water column and microbialite samples were evident for Archaea, which were absent from Sian Ka′an microbialite (SS), CCB Pozas Azules I pond (PAI) and both morphologies from Alchichica crater-lake (AS and AC) (Fig. S1 and Table S2).

Previous studies reported the bacterial diversity of giant microbialites from Lake Van, Eastern Anatolia, Turkey (López-García *et al.*, 2005), where the most abundant and diverse lineages were *Firmicutes*, followed by *Proteobacteria*, *Cyanobacteria* and *Actinobacteria*. In stromatolites from Hamelin Pool in Shark Bay, Australia, Papineau *et al.* (2005) reported that the community was composed of 90% bacteria and 10% archaea, the most abundant taxa being *Proteobacteria*, *Planctomycetes* and *Actinobacteria*. Ley *et al.* (2006) described the diversity of a microbial mat from Guerrero Negro, Baja California Sur, Mexico, and observed that *Chloroflexi* dominated clone libraries numerically, with *Proteobacteria* and *Bacteroidetes* the second most represented phyla, whereas *Cyanobacteria* constituted at most 10% of all sequences. Havemann & Foster (2008) examined the bacterial diversity of stromatolites from Highborne Cay, Bahamas, and found different phyla, *Proteobacteria* and *Cyanobacteria* being the most abundant. Using a metagenomic approach, Breitbart *et al.* (2009) determined the composition of microbial communities of a spherical oncolite from Rio Mesquites and a domal thrombolite from Pozas Azules II in CCB, Mexico. Those authors reported that sequences obtained from Pozas Azules II (PII) corresponded mainly to *Proteobacteria*, *Cyanobacteria* and *Planctomycetes*, whereas the Rio Mesquites community was largely represented by *Cyanobacteria*, followed by *Proteobacteria*. These results differ from those found in this study, especially for Rio Mesquites. This difference could be due to our inclusion of domal thrombolites from six locations in Rio Mesquites in a pooled sample, whereas Breitbart *et al.* (2009) studied an oncolite, a differentiated type of microbialite, clearly distinguishable from stromatolites and thrombolites (to name a few) by their concentric laminations (Garcia-Pichel *et al.*, 2004). Recently, Foster & Green (2011) estimated the diversity of microbialite-forming communities from different environments. These authors compiled 16S rRNA gene datasets derived from clone libraries. They found that *Alphaproteobacteria* and *Cyanobacteria* were the predominant phyla in all microbialites analysed. The patterns in community composition identified in this study for geographically distant sites are

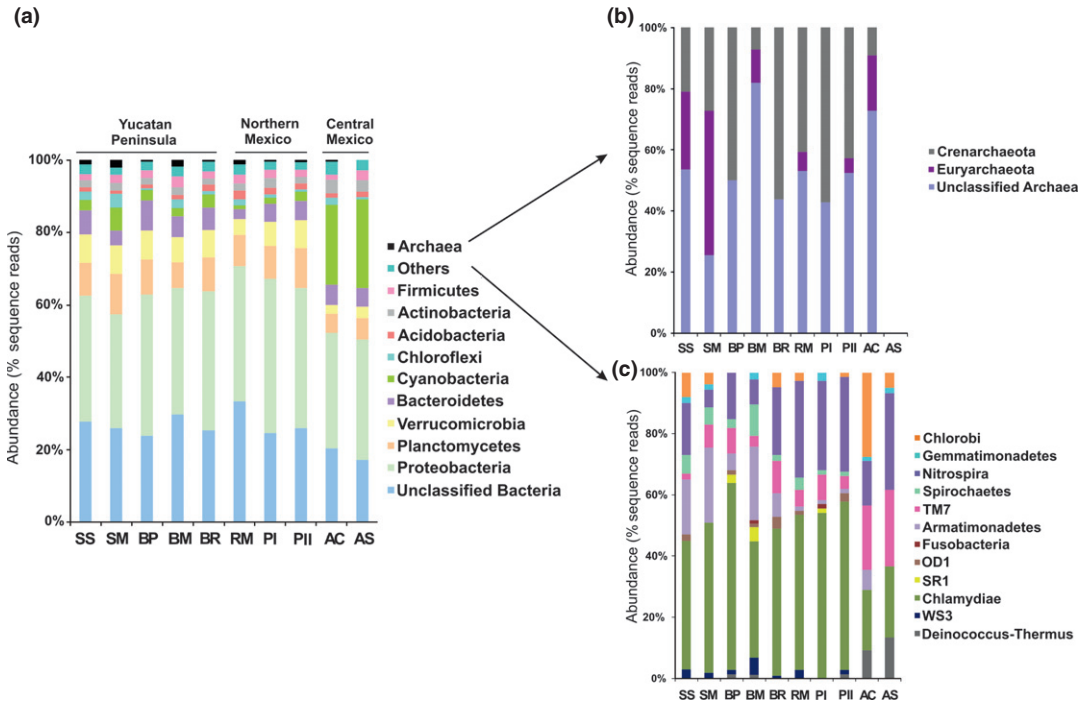

Fig. 4. Comparison of microbialite bacterial and archaeal community composition showing dominating phyla based on RDP-classifier. (a) Total sequences abundance per sampling site, (b) composition of rare bacterial phyla, (c) composition of archaeal phyla. Sian Ka'an microbialite (SS) and microbial mat (SM); Bacalar coastal lagoon, Piratas (BP), soft microbialite (BM), Los Rapidos (BR); CCB, Rio Mesquites (RM), Pozas Azules I (PAI), Pozas Azules II (PAII); Alchichica crater lake (AC) columnar and (AS) spongy) morphotypes.

not comparable with those found by Foster & Green (2011). Sequences in this study were obtained with the same sampling strategy, during the same season and year, and genomic extraction and amplification protocols were done in parallel for all. Thus, other studies analysed with different methodologies are not directly comparable.

In conclusion, the present study showed a geographical effect on the clustering of microbial communities and revealed that environmental variables such as pH and conductivity are partially correlated with the genetic composition of microbialite communities. Further, pyrosequencing of 16S rRNA gene hypervariable regions was successful at describing the diversity and composition of microbialites. It is evident that these communities are important depositories of carbonates, and their formation is associated with the carbonate pump, which will be affected under scenarios of global change (Petit *et al.*, 1999; Hoegh-Guldberg *et al.*, 2007). Increases in pH and changes in carbonate concentrations and conductivity will certainly affect these communities. We have yet to compute the contribution of microbialite-forming communities to the global carbon, phosphorous and nitrogen cycles. Further, we need to determine how species diversity is distributed along each locality harbouring microbialites. To date, we do not know whether these communities have homogeneous or heterogeneous distributions, which is fundamental for conservation strategies. We hope that these communities will not disappear, with all their associated diversity, from the face of our planet before we can fully study and understand them.

Acknowledgements

The authors would like to express their gratitude to C. Ezcurdia and O. Mejia for field assistance and R. Pérez-Ishiwara for technical assistance. Facilities for DNA extraction and amplification were provided by the Plant-Animal Interaction Laboratory (IE, UNAM). Graduate student fellowship from CONACyT-Mexico was awarded to C.M.C. This manuscript is required to obtain the PhD degree in the Posgrado de Ciencias Biológicas, UNAM. Field support in Cuatro Cienegas was provided by PRONATURA Noreste, A.C. inside the Pozas Azules Ranch. Field support in Sian Ka'an Biosphere reserve was provided by Community Tours SA de CV. Funding was provided by SEP-CONACyT No. 56045, No. 151796 (L.I.F.), PAPIIT No. IN225709-3, No. IT100212 (L.I.F.). All samples were collected under SEMARNAT collector permit DGOPA.09392.301109.4119.

References

Allwood AC, Walter MR, Kamber BS, Marshall CP & Burch IW (2006) Stromatolite reef from the Early Archaean era of Australia. *Nature* **441**: 714–718.

Ashby MN, Rine J, Mongodin EF, Nelson KE & Dimster-Denk D (2007) Serial analysis of rRNA genes and the unexpected dominance of rare members of microbial communities. *Appl Environ Microbiol* **73**: 4532–4542.

Beltrán Y, Centeno CM, García-Oliva F, Legendre P & Falcón LI (2012) N_2 fixation rates and associated diversity (*nifH*) of microbialite and mat-forming consortia from different aquatic environments in Mexico. *Aquat Microb Ecol* **67**: 15–24.

Bolhuis H & Stal LJ (2011) Analysis of bacterial and archaeal diversity in coastal microbial mats using massive parallel 16S rRNA gene tag sequencing. *ISME J* **5**: 1701–1712.

Breitbart M, Hoare A, Nitti A, Siefert J, Haynes M, Dinsdale E, Edwards R, Souza V, Rohwer F & Hollander D (2009) Metagenomic and stable isotopic analyses of modern freshwater microbialites in Cuatro Ciénegas, Mexico. *Environ Microbiol* **11**: 16–34.

Burne R & Moore L (1987) Microbialites: organosedimentary deposits of benthic microbial communities. *Palaios* **2**: 241–254.

Burns BP, Goh F, Allen M & Neilan BA (2004) Microbial diversity of extant stromatolites in the hypersaline marine environment of Shark Bay, Australia. *Environ Microbiol* **10**: 1096–1101.

Burns BP, Anitori R, Butterworth P, Henneberger R, Goh F, Allen MA, Ibañez-Peral R, Bergquist PL, Walter MR & Neilan BA (2009) Modern analogues and the early history of microbial life. *Precambrian Res* **173**: 10–18.

Caballero M, Vilaclara G, Rodríguez A & Juárez D (2003) Short-term climatic change in lake sediments from Lake Alchichica, Oriental, México. *Geofis Int* **42**: 529–537.

Couradeau E, Benzerara K, Moreira D, Gerard E, Kazmierczak J, Tavera R & López-García P (2011) Prokaryotic and eukaryotic community structure in field and cultured Microbialites from the alkaline lake Alchichica (Mexico). *PLoS One* **6**: e28767.

Decho AW, Visscher PT & Reid P (2005) Production and cycling of natural microbial exopolymers (EPS) within a marine stromatolite. *Palaeogeogr Palaeoclimatol Palaeoecol* **219**: 71–86.

DesMarais DJ (1995) The biogeochemistry of hypersaline microbial mats. *Adv Microb Ecol* **14**: 251–274.

Desnues C, Rodriguez-Brito B, Rayhawk S, Kelley S, Tran T, Haynes M, Liu H, Furlan M, Wegley L & Chau B (2008) Biodiversity and biogeography of phages in modern stromatolites and thrombolites. *Nature* **452**: 340–343.

Dupraz C & Visscher PT (2005) Microbial lithification in marine stromatolites and hypersaline mats. *Trends Microbiol* **13**: 429–438.

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

Elser JJ, Schampel JH, García-Pichel F, Wade BD, Souza V, Eguiarte L, Escalante A & Farmer JD (2005) Effects of phosphotus enrichment and grazing on modern stromatolitic microbial communities. *Freshw Biol* **50**: 1808–1825.

Ferrusquia-Villafranca I (1993) Geology of Mexico – a synopsis. *Biological Diversity of Mexico: Diversity and Distribution*, Chapter 1 (Ramammorthy TP, Bye R, Lot A & Fa J, eds), pp. 3–107. Oxford University Press, New York.

Fierer N & Jackson RB (2006) The diversity and biogeography of soil bacterial communities. *PNAS* **103**: 626–631.

Fierer N & Lennon JT (2011) The generation and maintenance of diversity in microbial communities. *Am J Bot* **98**: 439–448.

Foster JS & Green SJ (2011) Microbial diversity in modern stromatolites. *Stromatolites: Interaction of Microbes with Sediments. Cellular Origin, Life in Extreme Habitat and Astrobiology* **18**: 383–405.

Foster JS & Green SJ (2011) Microbial diversity in modern marine stromatolites. In: *Cellular Origin, Life in Extreme Habitats and Astrobiology: Stromatolites* (Tewari V. ed) pp. 385–405. Springer, Berlin.

Garcia-Pichel F, Al-Horani FA, Farmer JD, Ludwig R & Wade BD (2004) Balance between microbial calcification and metazoan bioerosion in modern stromatolitic oncolites. *Geobiology* **2**: 49–57.

Gischler E, Gibson M & Oschmann W (2008) Giant holocene freshwater microbialites, Laguna Bacalar, Quintana Roo, Mexico. *Sedimentology* **55**: 1293–1309.

Gischler E, Golubic S, Gibson MA, Oschmann W & Hudson JH (2011) Microbial mats and microbialites in the freshwater Laguna Bacalar, Yucatan Peninsula, Mexico. *Lect Notes Earth Sci* **131**: 187–205.

Hamady M, Lozupone C & Knight R (2010) Fast Unifrac: facilitating high-throughput phylogenetic analysis of microbial communities including analysis of pyrosequencing and Phylochip data. *ISME J* **4**: 17–27.

Hamp TJ, Jones WJ & Fodor A (2009) Effects of experimental choices and analysis noise on surveys of the 'rare biosphere'. *Appl Environ Microbiol* **75**: 3263–3270.

Havemann SA & Foster JS (2008) Comparative characterizacion of the microbial diversities of an artificial microbialite model and a natural stromatolite. *Appl Environ Microbiol* **74**: 7419–7421.

Hoegh-Guldberg O, Mumby PJ, Hooten AJ, *et al.* (2007) Coral reefs under rapid climate change and ocean acidification. *Science* **318**: 1737–1742.

Joergensen RG & Mueller T (1996) The fumigation-extraction method to estimate soil microbial biomass: calibration of the k(EN) value. *Soil Biol Biochem* **28**: 33–37.

Johannesson KH, Cortés A & Kilroy KC (2004) Reconnaissance isotopic and hydrochemical study of Cuatro Ciénegas groundwater, Coahuila, México. *J South Am Earth Sci* **17**: 171–180.

Kazmierczak J, Kempe S, Kremer B, López-García P, Moreira D & Tavera R (2011) Hydrochemistry and microbialites of

the alkaline crater lake Alchichica, Mexico. *Facies* **57**: 543–570.

Kromkamp JC, Perkins R, Dijkman N, Consalvey M, Andres M & Reid RP (2007) Resistance to burial of cyanobacteria in stromatolites. *Aquat Microb Ecol* **48**: 123–130.

Legendre P & Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**: 271–280.

Legendre P & Legendre L (2012) *Numerical Ecology*, 3rd English edn. Elsevier, Amsterdam, pp. 880.

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

Ley RE, Harris JK, Wilcox J, Spear JR, Miller SR, Bebout MB, Maresca JA, Bryant DA, Sogin ML & Pace NR (2006) Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. *Appl Environ Microbiol* **72**: 3685–3695.

López-García P, Kazmierczak J, Benzerara K, Kempe S, Guyot F & Moreira D (2005) Bacterial diversity and carbonate precipitation in the giant microbialites from the highly alkaline Lake Van, Turkey. *Extremophiles* **9**: 263–274.

Margulies M, Egholm M, William EA, Attiya S, Bader JS, Bembem LA, Berka J, Braverman MS, Chen YJ & Chen Z (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376–380.

Mobberley JM, Ortega M & Foster JS (2012) Comparative diversity analyses of modern marine thrombolites by barcoded pyrosequencing. *Environ Microbiol* **14**: 82–100.

Neilan BA, Burns BP, Relman DA & Lowe DR (2002) Molecular identification of cyanobacteria associated with stromatolites from distinct geographical locations. *Astrobiology* **2**: 271–280.

Oksanen J, Blanchet G & Kindt R *et al.* (2011). *Vegan: Community ecology package. R package version 1.17-9*. http://cran.rproject.org/package=vegan.

Paerl H, Steppe TF & Reid P (2001) Bacterially mediated precipitation in marine stromatolites. *Environ Microbiol* **3**: 123–130.

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

Paradis E, Bolker B, Claude J, Cuong HS, Desper R, Durand B *et al.* (2010) ape: Analyses of phylogenetics and evolution. R package version 2.6-2. http://cran.r-project.org/.

Paterson DM, Aspden RJ, Visscher PT, Consalvey M, Andres MS, Decho AW, Stolz J & Reid RP (2008) Light-dependant biostabilisation of sediments by stromatolite assemblages. *PLoS One* **3**: e3176.

Petit JR, Jouzel J, Raynaud D, Barkov NI, Barnola JM, Basile I, Bender M, Chappellaz J, Davis M & Delaygue G (1999) Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**: 429–436.

Pringault O, De Wit R & Camoin G (2005) Irradiance regulation of photosynthesis and respiration in modern marine microbialites built by benthic cyanobacteria in a tropical lagoon (New Caledonia). *Microb Ecol* **49**: 604–616.

Reid RP, Visscher PT, Decho AW, Stolz JF, Bebout BM, Dupraz C, Macintyre IG, Paerl HW, Pinckney JL & Prufert-Bebout L (2000) The role of microbes in accretion, lamination and early lithification of modern marine stromatolites. *Nature* **406**: 989–992.

Sabbe K, Hodgson DA, Verleyen E, Taton A, Wilmotte A, Vanhoutte K & Vyverman W (2004) Salinity, depth and the structure and composition of microbial mats in continental Antarctic lakes. *Freshw Biol* **49**: 296–319.

Schloss PD, Westcott S, Ryabin T, *et al.* (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Schopf JW, Kudryavtsevb BA, Czajac AD & Tripathic AB (2007) Evidence of Archean life: stromatolites and microfossils. *Precambrian Res* **158**: 141–155.

Sprachta S, Camoin G, Golubic S & Campion L (2001) Microbialites in a modern lagoonal environment: nature and distribution, Tikehau atoll (French Polynesia). *Palaeogeogr Palaeoclimatol Palaeoecol* **175**: 103–124.

Stal LJ (1995) Physiological ecology of cyanobacteria in microbial mats and other communities. *New Phytol* **131**: 1–32.

Stal LJ (2000) Cyanobacterial mats and stromatolites. *Ecology of Cyanobacteria: Their Diversity in Time and Space* (Whiton BA & Potts M, eds), pp. 61–120. Kluwer Academic Publications, The Netherlands.

Tavera R & Komárek J (1996) Cyanomicrobials in the volcanic lake of Alchichica, Puebla State, México. *Algol Stud* **83**: 511–538.

van Gemerden H (1993) Microbial mats: a joint venture. *Mar Geol* **113**: 3–25.

Vance ED, Brookes PC & Jenkinson DS (1987) An extraction method for measuring soil microbial biomass-C. *Soil Biol Biochem* **19**: 703–707.

Wang Q, Garrity GM, Tiedje JM & Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.

Winsborough BM, Seeler JS, Golubic S, Folk RL & Maguire B (1994) Recent fresh-water lacustrine stromatolites, stromatolitic mats and oncoids from northeastern Mexico. *Phanerozoic Stromatolites II* (Bertrand-Sarfati J & Monty C, eds), pp. 71–100. Kluwer, Amsterdam.

Zeglin LH, Dahm CN, Barrett JE, Gooseff MN, Fitpatrick SK & Takacs-Vesbach CD (2011) Bacterial community structure along moisture gradients in the parafluvial sediments of two ephemeral desert streams. *Microb Ecol* **61**: 543–556.

Zhou JZ, Bruns MA & Tiedje JM (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62**: 316–322.

Supporting Information

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

Fig. S1. Prokaryote community composition from water column controls (a), showing composition for less abundant bacterial phyla (b) and archaea (c), based on RDP-classifier.

Table S1. Estimated genetic richness diversity indices and coverage for hypervariable regions (V5–V6) of 16S rDNA gene for all study sites using a cut off at 97% similarity.

Table S2. Taxonomic assignation of water column controls based on RDP-II classifier.

Please note: Wiley-Blackwell is 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.