Comparative Metagenomics of Two Microbial Mats at Cuatro Ciénegas Basin II: Community Structure and Composition in Oligotrophic Environments

Germán Bonilla-Rosso,1 Mariana Peimbert,2 Luis David Alcaraz,3 Ismael Hernández,4 Luis E. Eguiarte,1 Gabriela Olmedo-Alvarex,4 and Valeria Souza1

Abstract

Microbial mats are self-sustained, functionally complex ecosystems that make good models for the understanding of past and present microbial ecosystems as well as putative extraterrestrial ecosystems. Ecological theory suggests that the composition of these communities might be affected by nutrient availability and disturbance frequency. We characterized two microbial mats from two contrasting environments in the oligotrophic Cuatro Ciénegas Basin: a permanent green pool and a red desiccation pond. We analyzed their taxonomic structure and composition by means of 16S rRNA clone libraries and metagenomics and inferred their metabolic role by the analysis of functional traits in the most abundant organisms. Both mats showed a high diversity with metabolically diverse members and strongly differed in structure and composition. The green mat had a higher species richness and evenness than the red mat, which was dominated by a lineage of *Pseudomonas*. Autotrophs were abundant in the green mat, and heterotrophs were abundant in the red mat. When comparing with other mats and stromatolites, we found that taxonomic composition was not shared at species level but at order level, which suggests environmental filtering for phylogenetically conserved functional traits with random selection of particular organisms. The highest diversity and composition similarity was observed among systems from stable environments, which suggests that disturbance regimes might affect diversity more strongly than nutrient availability, since oligotrophy does not appear to prevent the establishment of complex and diverse microbial mat communities. These results are discussed in light of the search for extraterrestrial life. Key Words: Cuatro Ciénegas—Metagenomics—Microbial mats—Oligotrophic—Phosphorus limitation—Stromatolites. Astrobiology 12, 659–673.

1. Introduction

Although modern microbial mats are geographically widespread (Gerdes, 2010), they are limited in their environmental occurrence to a selected few aquatic environments, both freshwater and marine (Krumbein et al., 1977; Bebout et al., 2002). This contrasts with their former distribution on Earth, as revealed by the abundance of biogenic, laminated reeflike structures in the fossil record from Precambrian shallow marine environments (Awramik, 1984) as early as 3.4 Ga (Allwood et al., 2006). It is believed that both an exclusion of grazing eukaryotes and a lack of competition for light with fast-growing algae are required for a microbial mat to develop, which restricts their current distribution mostly to extreme environments in modern Earth (Cohen, 1989; Bebout et al., 2002). Many current habitats can be considered environmental analogues to those found in...
Precambrian shallow waters (Foster and Mobberley, 2010). As such, microbial mats are the oldest ecosystems known to date and have proven to be successful ecological assemblages, given that they have survived billions of years of environmental change due to their stable, but adaptable, structural properties (Awramik, 1976; Green and Jahnke, 2010).

Since microbial mats are self-sustained heterogeneous ecosystems, they are suitable experimental models that can be used to test how ecosystems respond to rapid environmental disturbances, which cannot be tested with any other ecosystem (Paerl et al., 2003; Yannarell et al., 2007). They have also allowed for the reconstruction of putative metabolisms and the characterization of biosignatures of early life on Earth (Kasting, 2001; Bebout et al., 2004; Foster and Mobberley, 2010). Without doubt, the study of Earth’s most common, simple, and pervasive ecosystems at the molecular level will help us better understand the evolution of life on Earth and, as a consequence, will facilitate the search for life elsewhere (Foster and Mobberley, 2010).

Microbial mats can be found in extreme conditions, including cold (0.4–3.4°C; Bottos et al., 2008; Varin et al., 2010) and high concentrations of iron (Emerson and Revsbech, 1994), sulfur (Elshahed et al., 2003), and hydrocarbons (Mills et al., 2005). The vast majority of the literature, however, has focused on halophilic coastal mats (like those in Guerrero Negro, Mexico, or in Shark Bay, Australia) or hyperthermophilic mats (like those in Yellowstone) (Ward et al., 1998; Spear et al., 2003). Mats from coastal hypersaline environments can harbor diverse and complex ecosystems (Ley et al., 2006; Kunin et al., 2008). In contrast, mats that grow at high temperatures (64–82°C) are characterized by a low species diversity (Inskeep et al., 2010; Meyer-Dombard et al., 2011), most likely because temperature sets a limit for photosynthesis (Ward et al., 1998). This suggests that nutrient availability and primary production will determine the complexity of a microbial mat community, as has been previously proposed for plant communities (Grime, 1973; Tilman, 1990). The ecological characterization of communities from low-nutrient (oligotrophic) environments can aid the search of life on other planets by defining a nutrient availability range for the development of life, so that astrobiological targets can be more precisely defined.

Another ecological factor to be considered in the search for life beyond the confines of Earth is environmental disturbance, since communities will not develop where disturbances are too frequent or too intense, while an intermediate disturbance frequency in many cases will maximize complexity by creating several microniches (Grime, 1973; Connell, 1978; Rainey and Travisano, 1998). It is therefore relevant to study how communities are affected by disturbance regimes to be able to predict their possible existence in other planets.

The Cuatro Ciénegas Basin (CCB) is a naturally isolated valley in the Chihuahuan Desert (Coahuila, Mexico) that encompasses hundreds of permanent ponds, marshes, and desiccation pools (Minckley and Cole, 1968) (Fig. 1), which despite their low phosphorus content (Elser et al., 2005; Breitbart et al., 2009, Peimbert et al., 2012 in this issue) harbor varied and diverse microbial communities (Souza et al., 2006; Escalante et al., 2008). The CCB has received the attention of astrobiological research because it is considered a modern
analogue of the Precambrian oceans (Elser et al., 2006), which have been characterized as having had low levels of circulating phosphorus (Bjerrum and Canfield, 2002; Planavsky et al., 2010) and a reduced deposition of phosphorite rocks (Papineau, 2010). Moreover, complex microbial communities are known to have developed in Precambrian waters, since abundant phosphate-absorptive calcareous stromatolites have been preserved in the geological record (Grotzinger and Knoll, 1999; Gerdes, 2010). Finally, the CCB has also been identified as an analogue of the martian Olympia Undae gypsum sand dunes (Szynkiewicz et al., 2010) and the dynamic fluvial system of the Gale Crater of early Mars, one of the four landing sites targeted for the Mars Science Laboratory mission later this year (Golombek et al., 2011).

All this stresses the need to characterize the microbial communities that inhabit the CCB, while its oligotrophic waters and diversity of aquatic environments offer the opportunity to study the limiting effects of phosphorus availability on microbial community development. In the present study, we taxonomically characterized two aquatic, nonthermophilic, nonhalophilic oligotrophic microbial mats from two environments with contrasting disturbance regimes: a permanent green pool and a red desiccation pond. We analyzed their taxonomic structure and composition by using 16S rRNA gene clone libraries and metagenomics. A low diversity was expected in these oligotrophic environments, with a greater diversity in the mat from the desiccation pond due to intermediate disturbances. We also compared our data with two previously published metagenomes of microbial communities from the CCB (the stromatolites of Río Mesquites and Pozza Azul; Breitbart et al., 2009) and with the only other microbial mat metagenome available, the well-studied coastal halophilic microbial mat from the salines of Guerrero Negro (Baja California, Mexico; Spear et al., 2003; Kunin et al., 2008). We expected to find large differences in taxonomic composition, since Guerrero Negro and the CCB are geographically separated by a linear distance of 1200 km.

2. Material and Methods

Two 20×20 cm microbial mat cores were collected in July 2008 from a shallow, seasonal red desiccation pond in the Los Hundidos region (red mat, at 26°52’17”N, 102°01’11.3”W) and a permanent green pool in Pozas Azules Ranch (green mat, at 26°49’24.4”N, 102°00’53.2”W) (Fig. 1). Due to the nature of the desiccation ponds, the temperature disturbance frequency was 5 times higher for the red mat than for the green mat (Table 1). Samples were frozen in liquid nitrogen and transported to the laboratory. DNA extraction was performed by Freeze/Thaw, CTAB, phenolchloroform extraction as described previously (Zhou et al., 1996; Breitbart et al., 2009). The samples were further purified by electrodialysis as described by Rodriguez-Meja et al. (2008). Total DNA was amplified with GenomiPhi polymerase (GE Healthcare, Piscataway, NJ, USA) according to the manufacturer’s instructions. Ten independent reactions were carried out and later pooled before sequencing to reduce amplification bias. From the total DNA, 400 μL of one red mat (551 ng/μL) and 400 μL of the green mat (890 ng/μL) were independently pyrosequenced with 454 FLX (Roche Diagnostics, IN, USA) at CINVESTAV-LANGEBIO, Irupuato, Mexico. Metagenomic reads were noise-corrected with CD-HIT 454 (Li and Godzik, 2006). RNA sequences were identified and masked with rRNA-HMM and rRNA-Scan (Huang et al., 2009). Open reading frames calling was performed with GeneMark (Besemer and Borodovsky, 2005) and annotated with RAMMACAP (Li, 2009). Data sets were uploaded to the online server MG-RAST and subjected to its standard quality control pipeline (Meyer et al., 2008). Data is available via the MG-RAST portal (green mat ID is 4441363.3, and red mat ID is 4442466.3).

2.1. 16S rRNA gene clone libraries analyses

A 16S rRNA gene clone library was constructed for each of the two mats and one additional mat from an adjacent red desiccation pond in Los Hundidos (red mat 2). The PCR recombination mixture for amplification of 16S rRNA genes contain 1.5 mM MgCl2, 250 mM of each nucleotide, 4 mM of each primer, 1 U of Taq DNA polymerase (Roche, Mannheim, Germany), and 50 ng of isolated DNA. The bacterial 16S rRNA gene was amplified with the universal primers 27F (5-AGA GTT TGA TCC TGG CTC AG-3) and 1492R (5-GGT TAC CTT GTT ACG ACT T-3). Amplification was performed as follows: a “hot start” (95°C for 5 min) was followed by 25 cycles at 94°C for 40 s, 55°C for 40 s, and 72°C for 90 s with a 10 min extension at 72°C. Amplified products were purified with a Roche PCR purification Kit (Roche, Mannheim, Germany), and PCR products were cloned with a TOPO cloning kit (Invitrogen, Karlsruhe, Germany) following the instructions of the manufacturer. Plasmids were sequenced on one end by using dye terminator chemistry on an automated DNA sequencer (ABI3700, Applied Biosystems).

Raw sequences were chimera-checked with Mallard (Ashelford et al., 2006), trimmed of vector sequences, and checked for errors and low quality with BioEdit (Hall, 1999). Sequences were aligned and analyzed with ARB (Ludwig et al., 2004), assigned to a taxonomic category with the GreenGenes Classifier (DeSantis et al., 2006), and assigned to a phylogenetic position in reference to the SILVA database in ARB. Clustering of operational taxonomic units (OTUs) at 97% identity, alpha-diversity indexes, and rarefaction analyses were performed with MOTHUR (Schloss et al., 2009). Phylogenetic trees were constructed with PhyML software as implemented in ARB (GTR+I model, 1000 bootstraps), and edited with iTOL (Letunic and Bork, 2011).

---

**Table 1. Physiochemical Properties of the Aquatic Environments from Where the Samples Analyzed in This Study Were Taken**

<table>
<thead>
<tr>
<th>Environment</th>
<th>Tm (°C)</th>
<th>T°* (°C)</th>
<th>pH*</th>
<th>Conductivity (μS/cm)*</th>
<th>Stoichiometric N to P ratio**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>25 [15–24]</td>
<td>30.27</td>
<td>6.0</td>
<td>2.57</td>
<td>2:1</td>
</tr>
<tr>
<td>Red</td>
<td>30 [10–60]</td>
<td>48.26</td>
<td>5.5</td>
<td>117.6</td>
<td>157:1</td>
</tr>
<tr>
<td>Red 2</td>
<td>NA</td>
<td>47.58</td>
<td>7.3</td>
<td>83.5</td>
<td>NA</td>
</tr>
</tbody>
</table>

*T°* = Yearly average temperature mean; minimum and maximum temperatures are shown in brackets.

NA, Data not available.

---

**Notes:**

- *At time of sampling.
- **Numbers of nitrogen atoms per phosphorus atom. For reference, the classic “Redfield” ratio when nutrients are not limiting is 16:1.
- Tm = Yearly average temperature mean; minimum and maximum temperatures are shown in brackets.
- NA, Data not available.
2.2. Metagenomic protein gene phylogenetic marker analyses

The entire metagenomic data set of stromatolites from Poza Azul (ID 4440067.3) and Rio Mesquites (ID 4440060.4) were downloaded from MG-RAST (Meyer et al., 2008) and subjected to the exact same quality control and annotation analyses as described above. The 31 universally conserved, single-copy phylogenetic molecular markers identified with AMPHORA (Wu and Eisen, 2008) were searched for and assigned to taxonomic categories with MEGAN (Huson et al., 2007). Analyses were performed at genus level. Taxa-abundance matrices were built to calculate ecological distance matrices by using the Bray-Curtis distance and were subjected to a principal component analysis. Mantel’s test was performed to evaluate correlation between geographical and ecological distance. Community structure for the four data sets was evaluated with Renyi’s entropy profiles (Bent and Forney, 2008). All ecological analyses were performed with R (R Development Core Team, 2006). A list of the diversity metrics can be found in Supplementary Material S4b (Supplementary Data are available online at www.liebertonline.com/ast).

2.3. Metagenomic all-reads complement analyses

The resulting metagenomic data sets of the green and red microbial mats were uploaded and annotated with the MG-RAST automatic pipeline, where each single read in a metagenome can be assigned to a taxonomic category by comparing it with the available annotated databases in the SEED platform (Meyer et al., 2008). Hits with an e-value lower than $1 \times 10^{-5}$ were used to build taxonomic profiles for the metagenomes. The same analyses were performed with the pyrosequenced metagenomes of the Poza Azul and Rio Mesquites stromatolites (Breitbart et al., 2009) and the shotgun capillarily sequenced metagenomes from Guerrero Negro (Kunin et al., 2008), which were all downloaded from the MG-RAST portal (identification numbers 4440060.4, 4440067.3, and 4440963.3 to 4440972.3). The profiles of the 10 data sets from Guerrero Negro were pooled together since the referred study analyzed each layer of the microbial mat separately. These profiles allowed us to compare a larger number of reads from each metagenome by assigning them to different taxonomic categories at different taxonomic levels. We considered abundant taxa as those comprising 75% of the total annotated reads in each metagenome and compared these across metagenomes. The fragment recruitment analyses were calculated in MG-RAST with a maximum e-value cutoff of 1e-05. The presence of specific genes in the genomes that recruited the largest number of reads was analyzed in the Integrated Microbial Genomes portal (Markowitz et al., 2009).

3. Results

We evaluated diversity of two microbial mats, using three different approaches: 16S rRNA gene clone library construction, protein-coding phylogenetic marker genes from metagenomes, and all-reads analysis from metagenomes. These approaches are complementary, as each one analyzed different aspects of community structure and composition. For example, the abundance of minor phyla (Armatimonadaetes, Caldsericida, Nitrospora, OD1, Lentisphaera, and Deferribacteres) is only apparent with the clone library approach, while the protein marker approach gives a more precise structure estimation and allows the comparison between previously published metagenomes for which no clone libraries are available. The large proportion of cyanobacterial genomes is revealed by comparing the two previous approaches with the all-reads taxonomic composition.

3.1. 16S rRNA gene clone libraries analyses

Highly structured microbial mats were sampled from two different kinds of oligotrophic pools in the CCB: a green mat from a stable permanent pool and two red mats from two desiccation ponds. A total of 516 unique phylotypes were recovered from 724 sequences. Clustering at 97% identity produced 32 OTUs for the red mat and 260 for the green mat, comprising 23 phyla. The red mat had low evenness and was dominated by a few OTUs from the genus Pseudomonas (Shannon = 2.003, Simpson = 0.216), while the green mat showed a very high evenness with no dominant OTU (Shannon = 5.389, Simpson = 0.003), though the phyla Cyanobacteria, Clostridia, Gammaproteobacteria, Epsilonproteobacteria, and Deltaproteobacteria were abundant (Fig. 2). The Pseudomonas sequences belonged to the $P. \text{fluorescens}$ ($P. \text{fluorescens}$ and $P. \text{corrugata}$/P. brassicacearum) subgroups and the $P. \text{psychrerythrae}$ lineages (Mulet et al., 2010). The non-parametric richness estimator Chao estimated 1221 OTUs (C.I. = 904–1704) for the green mat and 80 OTUs (C.I. = 42–221) for the red mat. A total regional richness of 1198 OTUs (C.I. = 902–1643) is expected in microbial mats of the CCB, calculated from a composite pool of the individual libraries. The two microbial mat communities were phylogenetically and statistically different, both in terms of composition (Sorensen = 0.095, UniFrac U = 0.78) and structure (Morisita-Horn = 0.042, Weighted UniFrac W = 0.91). Both communities shared only 8 OTUs, with a total of 21 OTUs expected to be shared by the Chao richness estimator.

The 16S rRNA gene clone library of a second red mat (red mat 2) from a second desiccation pool was produced for comparison. Forty-two operational taxonomic units were recovered at 97% identity from 218 sequences. This red mat 2 was also dominated by Pseudomonas, although it was slightly more diverse than the first red mat (Shannon = 3.546, Simpson = 0.022, Chao = 122 [73–246]). The red mat 2 was more similar to the first red mat (Sorensen = 0.253, Morisita-Horn = 0.193) and shared 15 OTUs, with a total of 77 shared OTUs expected (Fig. 3).

3.2. Metagenomic protein gene phylogenetic marker analyses

The metagenomes of both mats were sequenced, which generated a total of 150,381,320 bp and 709,799 reads. The green mat metagenome consisted of 427,366 reads with an average length of 202.54 bp and a maximum length of 390 bp. The average GC content was 39.7%. The red mat metagenome consisted of 282,433 reads with an average length of 225.98 bp and a maximum length of 366 bp. The average GC content was 52.8%.

A total of 1066 sequences from 31 protein marker genes were retrieved from the data sets with AMPHORA (Wu and Eisen, 2008): 831 from the green mat and 235 from the
red mat. In agreement with the clone library, the red mat was less diverse (Shannon = 1.315, Simpson = 0.464, Chao = 23.5 ± 31.108 SE) and dominated by *Pseudomonas*. Genera *Sphingopixis* (Sphingomonadales), *Chitinophaga* (Sphingobacteria), and *Microcoleus* (Cyanobacteria) were also abundant. In contrast, the green mat was highly diverse (Shannon = 4.108, Simpson = 0.019, Chao = 148 – 33.618), with no dominant genus. Among the abundant genera were the phototrophic *Synechococcus* (Cyanobacteria) and *Chloroherpeton* (Chlorobi); the heterotrophic *Legionella* (Gamma-proteobacteria), *Algoriphagus* and *Bacteroides* (Bacteroidetes), and the deltaproteobacterial *Desulfococcus* and *Synthrophobacter* (Fig. 4).

The microbial mat metagenomes were compared with two other metagenomes from the CCB: a stromatolite from the permanent pool Pozas Azules and a riverine oncolite from Rio Mesquites (Breitbart et al., 2009). The analysis of the same 31 marker genes retrieved 785 sequences from Rio Mesquites and 238 from Pozas Azules. The most abundant genera in Pozas Azules all belonged to the superphylum Verrucomicrobia–Planctomycetes (*Chthoniobacter, Verrucomicrobia, Rhodopirellula, Akkermansia, Gemmata, Planctomyces*) and exhibited a high diversity (Shannon = 3.230, Simpson = 0.053, Chao = 145.2 ± 127.552). Rio Mesquites showed far less diversity (Shannon = 0.614, Simpson = 0.824, Chao = 132 ± 147.344), with most sequences belonging to the oligotrophic genus *Pelagibacter* (formerly known as the SAR11 cluster, Giovannoni et al., 2005). These findings are in agreement with recent clone libraries from DNA obtained from similar structures in the same system (Nitti et al., 2012).

The principal component analysis (Fig. 4) showed a larger similarity between the green mat and the Pozas Azules stromatolite because these samples aggregated together, though independently from the red mat and Rio Mesquites stromatolite. Ecological indexes also showed a larger similarity between the green mat and the Pozas Azules stromatolite (Sorensen = 0.194, Morisita-Horn = 0.2693) than between any other pair of samples (< 0.05).

Renyi’s community profiles revealed that the green mat was the most diverse sample, while the red mat was the least rich and Rio Mesquites showed the largest dominance. Although the two stromatolite samples showed a similar level of richness, they differed in dominance (Fig. 5). Mantel’s test was not significant and showed no correlation between ecological and geographical distances (r = 0.2821, p = 0.704). A guide to the interpretation of Renyi’s profiles can be found in Supplementary Material S4c.

### 3.3. Metagenomic all-reads complement analyses

The green mat and the Pozas Azules stromatolite samples displayed the highest evenness: 28 different orders composed 75% of the classified reads in the green mat and 22 orders in Pozas Azules. Among the most abundant taxa in the green mat were photosynthetic taxa (Cyanobacteria, Chloroflexales, Chlorobiales, Chromatiales, Rhodobacterales) and known heterotrophic taxa (Clostridiales, Bacillales, Burkholderiales). In Pozas Azules, the most abundant taxa were the Planctomycetes/Verrucomicrobia complex (Planctomycetes, Verrucomicrobia, Spartobacteria) and the Cyanobacteria (Chroococcales, Nostocales, Oscillatoriales). The red mat was dominated by reads from heterotrophic orders, with *Pseudomonadales, Burkholderiales, and Bacillales* comprising 50% of the total reads. Photosynthetic orders represented 15% of total reads. The Rio Mesquites stromatolite was dominated by reads from Cyanobacteria, with Nostocales, Chroococcales, and Oscillatoriales comprising 78% of the total reads.
To include another nonthermophilic microbial mat in the study, we analyzed the pooled taxonomic sequence profile from the coastal hypersaline mat metagenome from Guerrero Negro (Kunin et al., 2008). It also displayed a high evenness, with 75% of the reads distributed among 27 orders. Abundant taxa were Rhodobacterales, Rhizobiales, Clostridiales, Chroococcales, Actinomycetales, and Bacteroidales. We found a remarkable conservation in the relative proportion of the analyzed taxa across the green mat, Pozas Azules, and Guerrero Negro (Fig. 6).

Fragment recruitment revealed that reference genomes from members of the Nostocales recruited most reads from the green mat (1140 recruited by *Anabaena variabilis* and 820 from *Nostoc punctiforme*), but they did so with low similarity values (mode > 1e-10). The genomes of unicellular diazotrophic Cyanothece species (Chroococcales: Aphanotecoideae) recruited a large amount of reads from the green mat data set (1017 reads), while reference genomes from different strains of *Pseudomonas fluorescens* recruited most reads from the red mat (3996–5626 reads) with high similarity values (mode < 1e-30). The cyanobacteria *Microcoleus chthonoplastes* was the next reference genome that recruited most reads from the red mat (4847 reads). Given the cosmopolitan distribution and high similarity between *M. chthonoplastes* populations (Garcia-Pichel et al., 1996) and the fact that the recruited fragments from the red mat had very high similarity values (mode < 1e-30), the organisms present in the CCB red mat are very likely non-heterocystous filamentous cyanobacteria from *M. chthonoplastes*. Fragment recruitment plots can be found in Supplementary Material S3.

4. Discussion

4.1. On the dominant organisms in microbial mats

Cyanobacteria are important and often dominant components of microbial mats (Ward et al., 1998), and are responsible for the formation of the mat tissue with interlaced filaments. This study is no exception, since they were abundant in both mats, and order Chroococcales was the most abundant and diverse cyanobacterial order in both mats. Although no single OTU was dominant within Chroococcales, several species from genus *Synechococcus* were found. Organisms from this order would have had no role in the formation of the mat tissue, since it comprises only unicellular species. Order Oscillatoriales contains mat-forming species, and although it was among the most abundant groups in the green mat, no
sequences from the well-known mat-forming cosmopolitan Microcoleus chthonoplastes (García-Pichel et al., 1996) were found in GP. Order Nostocales was the most abundant group that contained known mat-forming Cyanobacteria. Abundant filamentous Nostocales from family Rivulariaceae (Calothrix and Tolypothrix) have been observed in other CCB microbial mats and stromatolites (García-Pichel et al., 2002; Domínguez-Escobar et al., 2011), and this group has also been reported to grow better under low-phosphorus conditions (Berrendero et al., 2008). Unfortunately, family Rivulariaceae cannot be accurately determined from protein sequences, since no reference genomes are available yet, which makes cultured members of this group good targets for future sequencing projects. In consequence, the mat-forming filamentous cyanobacteria in the green mat could be uncharacterized species from order Oscillatoriales, Nostocales, or both.

In contrast, most sequences from Oscillatoriales in the red mat show a remarkable similarity to M. chthonoplastes (BLAST hits below 1e-20), despite the fact that they comprise only 4% of the total metagenome. Nevertheless, only one 16S rRNA sequence from the clone library and only two protein-coding markers from the red mat displayed affinity to this organism. This is what would be expected from organisms with large genomes (M. chthonoplastes has a genome of 7.36 Mb), and this discrepancy has also been noted in other microbial mats (Sørensen et al., 2005; Dillon et al., 2009). All approaches agreed in that the red mat was dominated by Pseudomonas. Although unexpected, this is not surprising in that the existence of mats dominated by noncyanobacterial organisms like Thioploca (Oschmann, 2000), Beggiatoa (Mussmann et al., 2007), and even fungi mycelia (Verrecchia, 2000) have also been reported. Moreover, Pseudomonas species are recognized for their metabolic versatility, and many are acknowledged to form biofilms (Silby et al., 2009). More remarkable is the finding of several closely related but distinct lineages with affinity to P. fluorescens. Since the metagenomes described here represent only a snapshot of the red mat community evolution in time, we can infer that the composition observed in the red mats is actually a radiative burst of Pseudomonas and other heterotrophs after a desiccation disturbance. Even though it has already been shown that environmental heterogeneity promotes diversification in
P. fluorescens populations (Rainey and Travisano, 1998), the evolutionary consequences of these continuous disturbances are unknown. Our results suggest that this could lead to an adaptive radiation burst over the long term, but the observed pattern could also be caused simply by the metabolic plasticity of Pseudomonas. The similar composition of a second desiccation pond (red mat 2) suggests that the blooming of Pseudomonas could be an effect of the seasonal cycles common to all desiccation ponds in this area.

4.2. Taxonomic composition and functional trait analysis of the most abundant organisms

Although most organisms in these mats are uncharacterized, we can rely on the abundant literature on related species and genome projects to infer the general metabolic traits of the organisms that are abundant, according to the three methods, and have known and well-characterized relatives. In the green mat, despite the large amount of organisms represented, the main primary producers are likely to be Cyanobacteria, more specifically the abundant picooxygenic Synechococcus (Chroococcales). Moreover, the sequences found reveal high affinity to species OS-A and OS-B’ found in thermophilic mats in Yellowstone (Ward et al., 1998). Interestingly, these are the only Synechococcus genomes to carry nitrogen-fixing genes, as well as other genes that allow them to grow in oligotrophic environments such as genes for urea cycle, sulfolipid production, nitrate, phosphate and phosphonate utilization, and the biosynthesis and degradation of cyanophycin, a reserve polymer rich in nitrogen (Bhaya et al., 2007). This suggests that the most abundant phototrophs in the mat are not filamentous bacteria but efficient unicellular oligotrophic cyanobacteria and potentially important contributors to primary
productivity and inorganic nutrient incorporation. Nevertheless, large filamentous cyanobacteria might still be the dominant organisms when considering biomass contribution (S.J. Green et al., 2008).

The organic compounds synthesized by these primary producers in the green mat are likely recycled by the abundant and efficient oxygenic heterotrophs like *Legionella*, which are common in oligotrophic water bodies, even though they are auxotrophic for seven amino acids, including L-cysteine, and require iron salts for growth (Declerck, 2010). They satisfy their nutrient requirements from living organisms (such as amoeboid hosts and microbial biofilms) or decaying organic matter (Declerck, 2010). Therefore, the green microbial mat appears to provide physicochemical protection from the environment, while it concentrates large amounts of nutrients. *Algoriphagus* species are aerobic heterotrophs characterized by the degradation of high-molecular-weight polysaccharides, with *Algoriphagus* sp. PR1 genome coding for 145 polysaccharide degradation enzymes (Alegado et al., 2011). In the anoxic part of microbial mats, *Desulfococcus* species are able to completely degrade short- and long-chain fatty acids and aromatic compounds coupled to the reduction of sulfate to hydrogen sulfide (Muyzer and Stams, 2008) and can fix CO$_2$ via the Wood-Ljungdahl pathway (Platen et al., 1990). In contrast, the green-sulfur bacteria *Chloroherpethon* utilizes sulfide ions as electron donors for carbon assimilation (Bryant and Frigaard, 2006), while it contributes with carbon and nitrogen fixation to primary productivity and the production of phosphate granules via polyphosphate kinases. Even though Chlorobium are not properly classified as oligotrophs, they specialize in low light use and have been shown to dominate in environments with low organic matter input (Gonzalez et al., 2011).

The red mat, in contrast, is dominated by *Pseudomonas*, a metabolically versatile, ubiquitous heterotroph with broad catabolic and transport capabilities (Moore et al., 2006). We corroborated the *Pseudomonas* dominance with another clone library built from a second red mat, located in a similar desiccation pond that was adjacent to the first red mat pond (Fig. 3). *Pseudomonas fluorescens* species are aerobes; but some strains can use nitrate as electron acceptor instead of oxygen, and their genomes code for a large number of high-specificity nutrient transporters, as well as a large array of efflux systems for metal, organic solvent, and antibiotic detoxification (Silby et al., 2009). Their large genomes, high-affinity transporters, and broad sensing capabilities position *Pseudomonas* as copiotrophs (organisms adapted to grow in nutrient-rich environments) (Lauro et al., 2009). Nevertheless, *P. fluorescens* can use several carbon sources at very low concentrations, with a marked preference for amino acids (van der Kooij et al., 1982), and can form biofilms under starvation conditions to maximize exposure to diluted nutrients (Kroukamp et al., 2010), which suggests that they are at least tolerant to nutrient depletion. More relevant to the red mat environment, several *P. fluorescens* strains can efficiently solubilize mineralized inorganic phosphates (Fankem et al., 2008; Woo et al., 2010). Acidification increases solubilization by the production of carboxylic acids that have a high affinity for phosphate-bound ions (Khan et al., 2009) and reaches an optimum between a pH of 4.5 and 5.5 (Fankem et al., 2008). Phosphate solubilization is maximized with citrate and malate (Fankem et al., 2008), which are two of the most abundant carboxylic acids produced by *P. fluorescens*.
The aerobic heterotroph *Chitinophaga* is also abundant in the red mat. It is also a copiotroph, with a large genome of 9.1 Mb that contains a large and diverse collection of enzymes for the degradation of sugars (169 glycosyl hydrolases; Del Rio et al., 2010), most notably the degradation of chitin and casein (Sangkhobol and Skerman, 1981). In contrast, other abundant members in the red mat include the oligotrophic ultramicrobia (cell volume <0.1 μm³) *Sphingopyxis alaskensis* (Lauro et al., 2009) and *Janthinobacterium* sp. Marseille (*Minibacterium massiliensis*, Audic et al., 2007), whose common traits include very small cells and genomes, preference for amino acids over sugars, active iron scavenging, and a broad array of high-affinity but low-specificity transport systems. In addition, *S. alaskensis* exhibits degradation of high-energy yielding fatty acids and has a large number of genes for secondary metabolite catabolism and detoxification (Lauro et al., 2009). Surprisingly, the phosphorus metabolism in *S. alaskensis* is reduced to the non-specific alkaline phosphatases and a single ATP-dependent transporter, which is the only ABC transporter in the genome (Williams et al., 2009). In contrast, the genome of *Janthinobacterium* sp. contains ABC transporters for sulfate, sulfonate, thiosulfate, nitrate, most amino acids, phosphate, and phosphonates (Audic et al., 2007).

While *Microcoleus chthonoplastes*, which is also represented in the red mat, has a large genome (7.36 Mb) and is capable of performing both oxygenic and anoxygenic photosynthesis in the presence of sulfide (Stal, 1991), it cannot utilize nitrate or fix nitrogen (Zimmermann, 1989).

Mats dominated by *M. chthonoplastes* show very low nitrogen fixation (Camacho and de Wit, 2003), and most strains analyzed to date lack the nitrogenase gene (Bolhuis et al., 2009). Moreover, anoxygenic photosynthesis and nitrogen fixation in *M. chthonoplastes* mats are strongly dependent on phosphate abundance (Zimmermann, 1989; Camacho and de Wit, 2003), so it is not surprising that its genome contains genes for low- and high-affinity transporters (*pitA* and *pstSBCA*) and for the synthesis of the reservoir compound polyphosphate (*ppk*).

### 4.3. Food or stability? Oligotrophy and disturbance as limiting factors for the development of complex microbial mat communities

The functional trait analysis showed that, in the green mat, oligotrophic organisms are abundant primary producers that can fix nitrogen and optimize phosphorus utilization. Moreover, at least three of the five most abundant organisms are capable of CO₂ fixation. The fixed organic matter would then be recycled by non-oligotrophic heterotrophs. In contrast, oligotrophic organisms in the red mat are heterotrophs, while the dominant organisms are versatile copiotrophs. This reveals two very different strategies to cope with an oligotrophic environment, one mainly based on autotrophic primary production and the other on very efficient heterotrophic recycling. The results suggest that oligotrophy is not a limiting factor for the development of complex and functionally diverse microbial communities.

The analysis of diversity revealed that the green mat was by far the most diverse community and that communities in the more stable environments (the green mat and Pozas Azules) harbor higher richness and evenness than those in the more variable environments (the red mat and Río Mesquites). This suggests that disturbance frequency is a determinant factor of community structure in microbial mats, which is consistent with the exclusion of microbial mats by disturbing eukaryotes (Cohen, 1989; Bebout et al., 2002). In the red mat, the disturbance regime is apparently too frequent and may drive the reduction in abundance of several species, most notably autotrophs. These systems, as those growing in hyperthermophilic environments, demonstrate that microbial mats can exist with a simplified diversity, as long as a chemical gradient exists where primary producers are present and their nutrients are recycled by heterotrophs.

### 4.4. A common high-rank taxonomic composition as a result of trait conservation in microbial mats

Experimental studies on Guerrero Negro microbial mats have reported that, although salinity seems to affect community structure, cyanobacterial communities are only modestly affected by changes in sulfate and salinity concentrations (S.J. Green et al., 2008). Mathematical models of nutrient and population dynamics in microbial mats where complexity is reduced to a few functional groups show a large resemblance to natural mats (Decker et al., 2005). Moreover, functionally similar mats develop under varying oxygen concentrations because oxygenic phototrophs create a similar oxygen gradient in the upper layers (Herman and Kump, 2005). The role of oxygenic phototrophs is stressed not only as a source of nutrient incorporation but also as the generators of this gradient themselves. The potential metabolic differences between the two mats are mainly a result of differences between the proportion of photoautotrophs in the mats, which suggests that primary productivity is determinant to the community structure complexity. Since an unexpected diversity has been found in systems that produce a similar layered pattern with steep biogeochemical gradient (Jorgensen et al., 1986; Stolz, 2000; Kunin et al., 2008), it would appear that microbial mats are completely independent of the taxonomic composition of its conforming species and that similar systems may arise in any place where these gradient conditions are met. Hence, we compared our microbial mat communities with similar structures developed under different environmental conditions, the stromatolites from the CCB, and the hypersaline mat from Guerrero Negro. The recently published literature on microbialite community diversity comprising microbial mats (Ley et al., 2006; Abed et al., 2007; Allen et al., 2009), stromatolites (Burns et al., 2004; Papineau et al., 2005; Baumgartner et al., 2009), and endoevaporites (Sahl et al., 2008) reveals that microbial mat communities have a far larger expected richness than stromatolites and endoevaporites (Table S2b in Supplementary Material). As expected, we found very few shared organisms between the two CCB mats, and no species were shared between the four systems from the CCB. Moreover, only two genomes (*M. chthonoplastes* and *Desulfococcus oleovorans*) showed high recruitment in both Guerrero Negro and CCB metagenomes. However, a larger similarity at higher taxonomic levels was observed between the Guerrero Negro...
microbial mat, the green mat, and the Pozas Azules stromatolite even though these two mats were geographically separated by a linear distance of ca. 1200 km. The pattern appears unexpected, since other clone library–based investigations in which both stromatolites and microbial mats from the same sampling site were analyzed produced very different community compositions (for example Burns et al., 2004; Allen et al., 2009).

The finding of a common phylogenetic pattern at high taxonomic ranks suggests that layered microbial communities at the water-sediment interface assemble in biogeochemical gradients that fill defined ecological niches (photosynthesis, sulfate reduction, heterotrophy) according to their functional traits and independently of their phylogeny. It also supports the theory of the assembly of bacterial communities by functional genes rather than species (Burke et al., 2011) and the existence of general functional traits shared by organisms at deep phylogenetic nodes (J.L. Green et al., 2008; Philippot et al., 2010). Under this model, phylogenetically unrelated species are able to colonize the same niche in an ecosystem as long as they are ecologically equivalent (same trophic level or metabolic function), just as similar ecosystems will have communities with common functional “guilds,” but species within these guilds will be selected at random (Burke et al., 2011). This explains the success of microbial matlike communities because they benefit from a simple set of environmental requirements, namely, a water-sediment interface and a steep physicochemical gradient. Hence, complex communities are likely to be found wherever these conditions are met, with metabolically diverse functional guilds benefiting from the generation of diverse microniches and from the environmental buffering the mat structure provides. This would suggest that the prevalence of remarkably similar, but taxonomically distinct, mat structures in stable environments is a natural consequence of the undisturbed association of metabolically diverse organisms that exploit a locally microdiverse niche. The identification of a conserved taxonomic group core also provides a base system with which to test the effects of different kinds of disturbances on complex microbial ecosystems, while it also sets a start-up for the design of synthetic artificial model microbial communities in mesocosms.

Microbial mats appear to be the most ancient and pervasive ecosystems because they conform biogeochemical structures that are likely to be found wherever a gradient occurs across a sediment-water interface, and they comprise physicochemical structures that protect organisms from environmental extremes (S.J. Green et al., 2008) while concentrating nutrients. This could explain their success even under the low-phosphorus environments of the Precambrian oceans (Papineau, 2010). The fact that their structure depends more on the ecological traits than the taxonomic component of their members suggests that analogues of these kinds of ecosystems are also the most likely structures to be found beyond the confines of Earth. Their organosedimentary nature makes them ideal targets for the detection of evidence of former life in the geological record by means of the identification of stable isotopic signatures and molecules of biological origin.

Currently, a wide array of biosignatures has been proposed with which to narrow the number of astrobiological bodies that could host life, such as photosynthetic pigments (Seager et al., 2005), sulfur gases (Domagal-Goldman et al., 2011), methane (Hoehler et al., 2001), and ammonia (Des Marais et al., 2002). Theoretically, it would seem more likely to detect a complex community of unknown organisms than a simple one, since a complex community would contain and exude a larger variety of metabolic products, which would increase the probability of detecting one of these as biosignatures. Our results suggest that complex communities can be found in environments where nutrient concentrations are very low, which underscores the importance of considering microhabitats in otherwise nonviable environments. However, an environment with a frequent disturbance regime might significantly reduce diversity, while environments with intense disturbances might not host life at all. This suggests that more complex and, hence, more detectable communities are likely to be found in more stable environments, so that efforts in the search for life should incorporate disturbance intensity and frequency when narrowing the list of planets where search efforts are to be directed.

5. Conclusion

The CCB displays astounding microbial diversity despite being a mixed oligotrophic environment, which makes its ponds desirable astrobiological experimental models. The microbial mats from this study revealed two extremes of diversity, with the green mat showing a large diversity above phylum level and the red mat exhibiting diversity below genus level. Our results widen the ranges within which life may be found, since they show that complex communities can develop in environments with nutrient concentrations below detection levels, even though phosphate is readily mineralized and made biologically unavailable. However, our results also stress the need to incorporate disturbance intensity and frequency models into a more precise definition of astrobiological targets, since high disturbance regimes can lower diversity (with a concomitant lowering of measurable biosignatures) and even prevent the development of life.

Acknowledgments

We thank Rodrigo Gonzalez Chauvet for technical logistics and field assistance; E. Lopez, A. Islas, V. Lopez, F. Reverchon, E. Rebollar, M. Avitia, and A. Gutierrez for assistance in sample collection and DNA isolation; and Laura Espinosa from IE/UNAM for laboratory and technical assistance. We thank three anonymous reviewers, whose insightful comments greatly improved this manuscript. We also thank O. Rodriguez and C. Bixler for their inertiatic propulsion. This work was supported by grants CONACyT 057507, SEMARNAT 2006-C01-23459, and WWF-Alianza Carlos Slim L039 to V.S., and CINVESTAV-multidisciplinario to G.O.A. The manuscript was written while V.S. and L.E.E. were on sabattical at UCI with support from DGAPA to V.S. and UC-Mexus to L.E.E. G.B.R. was supported with Ph.D. scholarship CONACYT 196814 and Programa de Posgrado en Ciencias Biomédicas UNAM.

Abbreviations

CCB, Cuatro Ciénegas Basin; OTUs, operational taxonomic units.
References


DIVERSITY IN MICROBIAL MAT METAGENOMES


