ISME

ARTICLE

Disentangling the drivers of functional complexity at the metagenomic level in Shark Bay microbial mat microbiomes

Hon Lun Wong[1,2] · Richard Allen White III[3] · Pieter T. Visscher[2,4] · James C. Charlesworth[1,2] ·
Xabier Vázquez-Campos [1] · Brendan P. Burns[1,2]

Abstract

The functional metagenomic potential of Shark Bay microbial mats was examined for the first time at a millimeter scale, employing shotgun sequencing of communities via the Illumina NextSeq 500 platform in conjunction with defined chemical analyses. A detailed functional metagenomic profile has elucidated key pathways and facilitated inference of critical microbial interactions. In addition, 87 medium-to-high-quality metagenome-assembled genomes (MAG) were assembled, including potentially novel bins under the deep-branching archaeal Asgard group (Thorarchaetoa and Lokiarchaeota). A range of pathways involved in carbon, nitrogen, sulfur, and phosphorus cycles were identified in mat metagenomes, with the Wood–Ljungdahl pathway over-represented and inferred as a major carbon fixation mode. The top five sets of genes were affiliated to sulfate assimilation (*cysNC cysNCD, sat*), methanogenesis (*hdrABC*), Wood–Ljungdahl pathways (*cooS, coxSML*), phosphate transport (*pstB*), and copper efflux (*copA*). Polyhydroxyalkanoate (PHA) synthase genes were over-represented at the surface, with PHA serving as a potential storage of fixed carbon. Sulfur metabolism genes were highly represented, in particular complete sets of genes responsible for both assimilatory and dissimilatory sulfate reduction. Pathways of environmental adaptation (UV, hypersalinity, oxidative stress, and heavy metal resistance) were also delineated, as well as putative viral defensive mechanisms (core genes of the CRISPR, BREX, and DISARM systems). This study provides new metagenome-based models of how biogeochemical cycles and adaptive responses may be partitioned in the microbial mats of Shark Bay.

Introduction

Microbial mats are organosedimentary, layered ecosystems, characterized by a diverse community that can

✉ Brendan P. Burns
  brendan.burns@unsw.edu.au

1 School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, NSW, Australia

2 Australian Centre for Astrobiology, University of New South Wales Sydney, Sydney, NSW, Australia

3 Institute of Biological Chemistry, Washington State University, Pullman, USA

4 Department of Marine Sciences, University of Connecticut, Storrs, CT, USA

facilitate mineral precipitation or dissolution mediated by both biotic and abiotic factors [1]. Molecular analyses have begun to uncover the complexity of mat communities from a range of environments, including hypersaline settings of Shark Bay [2–5], Guerrero Negro [6, 7] and Kiritimati Atoll [8], freshwater lakes [9, 10], acidic springs [11], alkaline lakes [12, 13], hot springs [14], and the arctic [15].

Shark Bay in Western Australia has some of the most extensive extant microbial mat systems [2, 16], and some of these mats are potential analogs to Precambrian stromatolites that can be dated back to 3.5 Ga [17]. Furthermore, Shark Bay microbial mats are subject to hypersalinity of more than 60 PSU [3, 5], desiccation stress, and high UV radiation [18, 19]. Daily tidal cycles also cause fluctuations in salinity and temperature at a local scale in the mat microenvironment [20, 21], adding another dimension to the extreme conditions present. The Shark Bay microbial mat systems are thus considered as an ecological model for understanding stress response and

resilience in microbial ecosystems, as well as a window to the past.

Early work in Shark Bay has focused primarily on understanding microbial diversity using targeted culturing and cloning/sequencing of 16S rDNA PCR products [2, 4]. These studies proposed the Shark Bay systems were dominated by Proteobacteria, Bacteroidetes, Cyanobacteria, and Haloarchaea [2, 4], and, more recently, deep iTag amplicon sequencing has revealed another dominant clade—Chloroflexi—as well as novel members, including OP8, GN04, and Parvarchaeota [3, 21, 22]. These latter studies were coupled with microelectrode measurements in Shark Bay and indicated a steep biogeochemical gradient with vertical depth in the microbial mats (O_2, H_2S, CH_4) [3, 21, 23]. It has been proposed that microbial interactions and biochemical cycling occur at a fine, millimeter scale [6], and steep nutrient gradients support niche differentiation, allowing different microorganisms to interact with each other and the environment in their specialized niche [24, 25]. It is also hypothesized that coupling of microbial metabolisms from different microbial taxa may help create distinct niches in the mats [3, 26, 27].

High-throughput sequence analysis of bacterial and archaeal populations (targeting 16S rDNA) in Shark Bay mats were recently undertaken at a millimeter scale, leading to the proposal of several putative niches [3, 21]. Phototrophic consortia at the surface mat layer were proposed as the main source of energy production to fuel the anoxic layers, along with a putative surface anoxic niche harboring a considerable abundance of anaerobic bacteria and archaea [3, 21]. Although sulfate-reducing bacteria outcompete hydrogenotrophic methanogens thermodynamically, both appear to co-exist near the mat surface, suggesting metabolic specialization that is impacted by daily tidal and salinity fluctuations [21]. Results suggested that tightly coupled cooperative niches are formed that are regulated by microbial interactions between different functional groups, including the Cyanobacteria, sulfate reducers/sulfate oxidizers, and hydrogenotrophic methanogens [21, 27, 28].

One caveat of 16S rDNA sequencing is that one cannot directly link function(s) to phylogeny, and thus metagenomic analyses are required as a next step for annotating microbial functional capacity in a given system. A recent study has identified metabolic pathways in Shark Bay microbial mats using 454-shotgun sequencing of bulk mat material [5]. Alternative modes of carbon fixation (3-hydroxypropionate/4-hydroxybutyrate pathway), heavy metal resistance genes, and osmoadaptive pathways were prevalent in the Shark Bay mat metagenomes [5]. However, advances in next-generation sequencing platforms have allowed a deeper sequencing depth to be obtained, and sequencing in discrete mat layers also offers a higher resolution to delineate the functional capacity of any putative niches discovered. Of particular interest, the lack of bacterial nitrifiers and corresponding nitrifying genes [3, 5], suggests incomplete or alternate modes of nitrogen cycling in Shark Bay, and thus metagenomic analyses at higher resolution are required to ascertain how communities overcome potential bottlenecks in these systems. Therefore, the aim of this study was to delineate for the first time the functional metagenomic potential of microbial mats in Shark Bay at the millimeter scale, via sequencing of total microbial mat DNA on the Illumina NextSeq platform. Genomic binning from metagenomes facilitated the recovery of novel, near-complete, and high-fidelity metagenome-assembled genomes (MAGs) for the first time and revealed potential microbial interactions crucial to ecosystem function in Shark Bay systems.

Materials and methods

Sample collection and measurement of key biogeochemical properties and parameters

Smooth mats were sampled on 16 June 2013 from Nilemah, Hamelin Pool, Shark Bay (26°27'336''S, 114°05'762''E), using previously reported methods [3]. At the time of sampling, water temperature was 24.8 °C, salinity 68 PSU, and pH 8.13. Sterilized scalpels were used to extract mat samples and subsequently place in sterile containers. At the point of sampling, mat samples were stored immersed in RNAlater and frozen to facilitate maintenance of nucleic acid integrity, and DNA extracted immediately upon sample return. Samples were sectioned upon return immediately prior to DNA extractions. Seawater was taken directly overlying the mats at the time of sampling and placed into sterile containers. In order to delineate key metabolic activities (e.g., photosynthesis, respiration, and sulfide oxidation), depth profiles of oxygen and sulfide were measured in the mats in situ using microelectrodes, and sulfate reduction was determined using the $^{35}SO_4^{2-}$-labeled silver foil technique as described previously [21, 23]. To semi-quantitatively determine elemental levels in seawater overlying mats, inductively coupled plasma optical emission spectrometry was employed. Based on a previous study in which specific elements were proposed to be important in mat community structure [3], the concentrations of As, Ca, Cu, Fe, K, Mg, Mn, Mo, P, and V, present in seawater samples were determined in triplicate. Dissolved organic carbon (DOC), total organic carbon (TC), and total nitrogen (TN) were determined through high-temperature catalytic oxidation [29], while the concentrations of nitrate and nitrite ions were measured through flow injection colorimetric analysis [30].

Nucleic acid extraction, library preparation, and shotgun sequencing

Mats were dissected into 2 mm layers using sterilized blades over a depth of 20 mm. Total community genomic DNA was extracted in duplicate from each microbial mat layer employing the MoBio PowerBiofilm DNA Isolation Kit (MO BIO Laboratories, Carlsbad, USA) as previously described [3]. The concentrations and purity of extracted DNA were determined spectrometrically on Xpose and the Qubit dsDNA HS assay kit, and the quality confirmed by PCR amplification [3–5]. Libraries were prepared using the Illumina Nextera DNA sample prep kit, with 40–50 ng of DNA input in each sample. The fragment sizes of the libraries were subsequently assessed by Perkin Elmer LabChip GXII and Agilent 2200 TapeStation. After the libraries were normalized and pooled, Illumina NextSeq 500 with a High Output v2 300-cycle kit was used to sequence the samples, with duplicate metagenomes obtained for each layer. The library pool was loaded at 1.7 pM with 1% PhiX spike-in.

Metagenomic data assembly and binning

Analysis of metagenomic data was done as described previously [31], using ATLAS (Automatic Tool for Local Assembly Structures). Raw paired-end Illumina reads (.fastq format) are extended for overlaps by using FLASH [32], after which φX174 is removed using Bowtie2 [33], the reads are trimmed with Trimmomatic [34], and then quality control is performed with FastQC. Overlapped paired-end reads (from FLASH) and unpaired reads were assembled using MEGAHIT [35]. The resulting contigs are sub-sampled for lengths >1 kbp with contig statistics including assembly size, number of contigs, contig length distribution, and $N_{50/90}$ values. Contigs are translated to protein-coding open reading frames (ORFs) using Prodigal [36] in meta-genome mode and annotated using DIAMOND blastp [37] for protein–protein searching. DIAMOND blastp high-scoring pairs are filtered to user-specified bitscore and e-value cut-offs (defaults >200 and $<1 \times 10^{-7}$, respectively). Functional annotation utilizes non-redundant RefSeq [38], EggNOG [39], dbCAN for CAZy families [40], ENZYME for enzyme commission number (EC) [41], COG [42], and PFAM/TIGRFAM [43] databases. ATLAS obtains KEGG [44] (i.e., KO number) annotations from EggNOG reference database. ATLAS uses RefSeq high-scoring pairs along with NCBI's taxonomy assignments reference tree via a modified majority voting method (MMVM) that utilizes lowest common ancestor (LCA) [45] to determine the LCA represented across all ORFs present within a single contig. Functional and taxonomic count data were obtained by mapping quality-controlled reads to assembled contig annotations using Bowtie2, then parsed using featureCounts of the Subread package [46]. While all layers of the Shark Bay smooth mat metagenomes were pooled and assembled together, functional and taxonomic count data were obtained by mapping individual-layer metagenomes against the pooled assembly. Binning was completed using MaxBin2 (version 2.2), in which differential coverages were used by mapping individual-layer metagenomes against the pooled assembly. Using ATLAS, the assembly of smooth mat metagenomes resulted in 945,688 contigs (>1 kbp), a maximum contig length of 470,099 bp, average contig length of 2841 bp, and N50 of 3695 bp. All meta-genome assembly statistics can be found in Supplementary Table 1.

Functional annotation of MAGs, phylogenetic tree construction, and data analyses

Quality control of MAGs was performed by CheckM to assign estimated completeness, contamination values, and statistics [47]. QUAST and tRNAscan-SE were used to assign statistics and number of amino acids encoded by MAGs, in order to meet the standards of Minimum Information about Metagenome-Assembled Genomes (MIMAG) [48–50]. MAG nucleotide contigs were translated to protein-coding ORFs with Prodigal [36]. The KEGG server (Ghostkoala) was used to annotate protein functions [44]. MAG taxonomic assignments were performed by CheckM, RAST, Phylosift, and JspeciesWS [47, 51, 52]. MAGs with incongruent taxonomic classifications had their ribosomal proteins extracted by Phylosift and were checked against the NCBI database (with threshold 40% amino-acid identity and cut-off value $1e^{-20}$). Identity of putative quorum sensing genes was confirmed with BLAST analyses of NCBI databases (with threshold 25% amino-acid identity and cut-off value $<1e^{-10}$) [53]. MAGs were annotated against PFAM/TIGRFAM to search for BREX and DISARM associated genes using InterProScan5 (with cut-off value $<1e^{-10}$; [54]). The *DESeq2* package in R (using the log2count function) was used to compare the abundance of genes across the ten mat layers. Network correlation analysis was conducted on genes involved in different nutrient cycles (C, N, P, and S) annotated against KEGG database with the *Hmisc* and *Igraph* library package showing only very strong correlations (Pearson's $p > 0.8$, $p < 0.001$) [55]. Correlation network graphs were then exported and visualized with Cytoscape version 3.5.1 [56]. For phylogenetic tree construction of the draft MAGs, ribosomal protein sequences generated from Phylosift were extracted and aligned using MAFFT [57]. Gaps in the alignment were removed using BMGE [58], and the ribosomal protein alignment concatenated as described previously
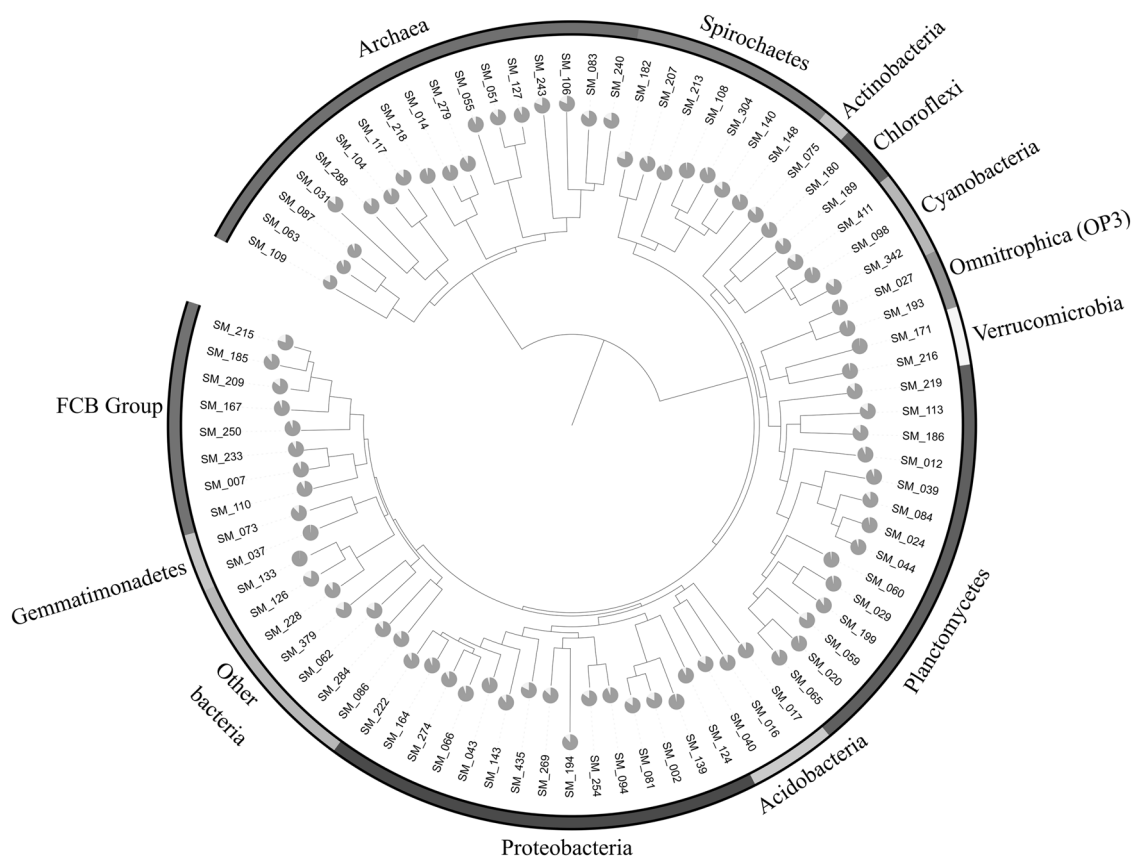
Fig. 1 Phylogenetic tree of 87 assembled Shark Bay smooth mat microbial metagenome-assembled genomes (MAGs). MAGs assigned to different phyla are highlighted in color, and the estimated completeness of each draft genome is represented by pie charts. SM_103, 152, 180, 194, 225, and 277 were excluded from the tree as coherent taxonomic classification could not be generated from these MAGs using Phylosift (color figure online)

[59]. IQ-TREE was used to construct the phylogenetic tree with a total of 1000 bootstrap replicates, with the output file uploaded to iTOL for visualization [60, 61].

Results

Elemental analysis

Elemental measurements of seawater overlying the Shark Bay microbial mats (Supplementary Table 2) indicate the mat environment harbors very high concentrations of magnesium (1960 mg/L) and copper (380 µg/L); however, there was extremely limited phosphorus (6.46 µM). DOC and TN were detected at 72.9 and 2.9 mg/L, respectively, and nitrate and nitrite ions were under detection limits (<0.1 and <0.02 mg/L, respectively).

MAG summary

For the first time in the Shark Bay mats, near-complete genome bins were retrieved from the metagenomic data, with 550 MAGs generated, of which 87 were of medium to high quality, according to the standards of MAGs (MIMAG) [48]. All MAGs in this study have >80% completeness and <10% contamination. Phylogenetic classification of these MAGs is depicted in Fig. 1, with the majority classified as Deltaproteobacteria, Planctomycetes, Spirochaetes, Chloroflexi, and Bacteroidetes. Potentially novel MAGs were assembled, comprising two Thorarchaetoa (SM_104, 117) and one Lokiarchaeota (SM_031)—two new archaeal phyla recently proposed [62] —as well as Micrarchaeota (SM_051, 127) and Woesearchaeota (SM_240) of the DPANN superphylum [63]. Statistics (completeness, contamination, bin size, strain heterogeneity, marker lineage, number of single-copy markers, GC content, N50, L50, longest contig, and number of tRNAs) of microbial assembled metagenomes (MAGs) from Shark Bay smooth mats can be found in Supplementary Table 3 (checkM). In addition, genes found in the MAGs for the pathways and functions discussed are listed in Supplementary Table 4, and identities of MAGs with incongruent taxonomic classification are listed in Supplementary Table 5.
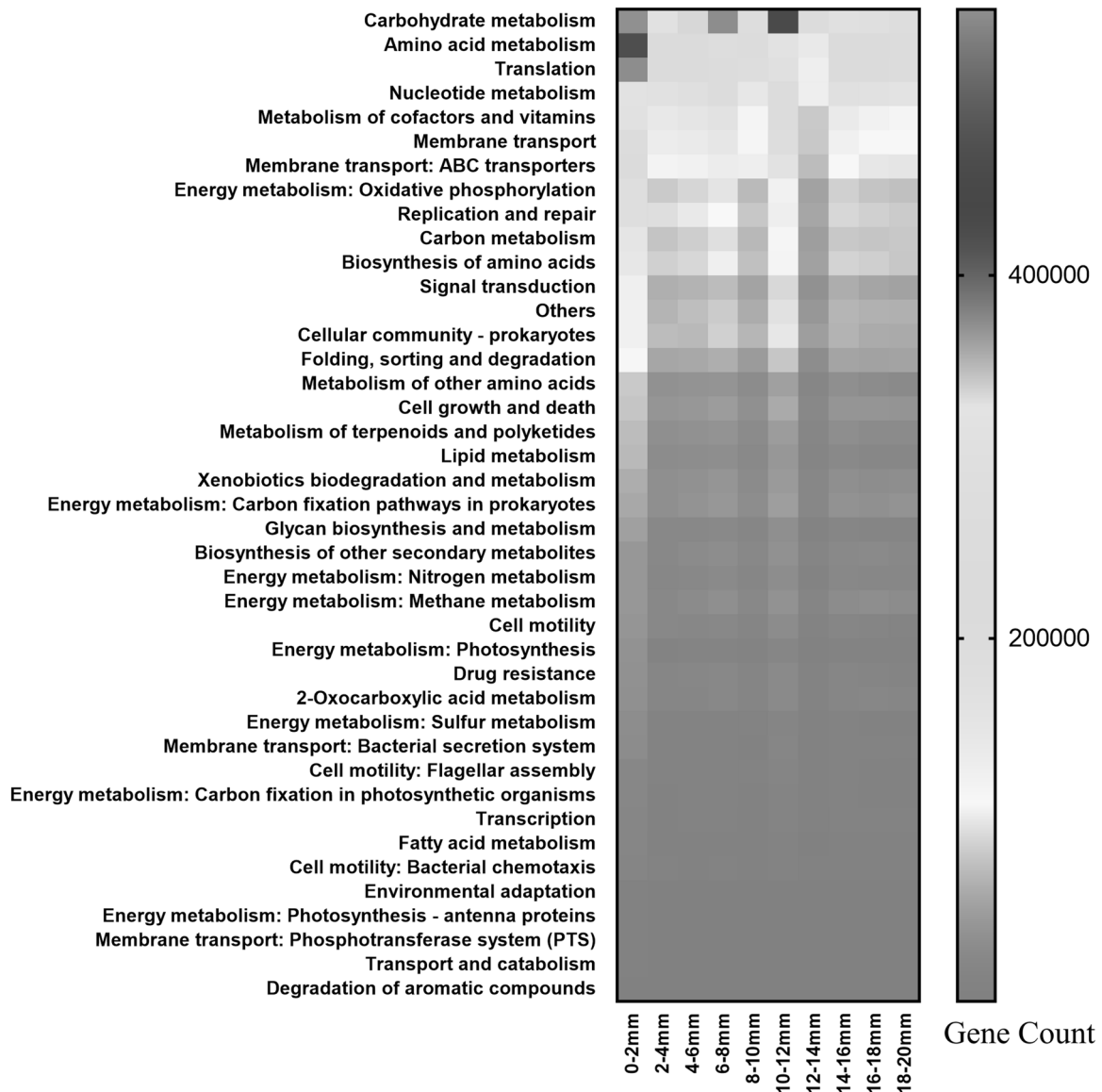
Fig. 2 Distribution of major functional pathways with depth in Shark Bay smooth mats. These are denoted at KEGG level 3 (defined as any kind of biological reaction and/or gene expression regulation), with the exception of the energy metabolism shown at KEGG level 2 (defined as molecular interactions/protein post-translational modifications).

X-axis indicates the different smooth mat layers in 2 mm increments, and y-axis indicates major functional genes/pathways. Scale bar represents gene count as the relative abundance of relevant genes associated with the displayed functions

Key functional gene distribution in Shark Bay mat metagenomes

Microorganisms align themselves in microbial mats along nutrient, biogeochemical, and light gradients, and this is also reflected in the spatial distribution of functional genes observed in the present study. The most abundant genes are affiliated to carbohydrate metabolism, followed by amino-acid metabolism and carbon, nitrogen, sulfur, methane, and phototrophic energy metabolisms (Fig. 2). Genes associated with energy metabolisms were most abundant at the surface layer/photo-oxic zone (0–2 mm), suggesting most of the

energy was generated at the photo-oxic zone [3]. This was also consistent with microelectrode data, where the mats displayed a relatively shallow oxycline observed even during peak photosynthesis (Fig. 3). Oxygen production and consumption was concentrated near the surface (0–3.5 mm), with anoxic conditions below this depth. This anoxia may be due to the lack of oxygenic photosynthesis, and/or because oxygen production and consumption are exactly in balance. In terms of nutrient cycling and metabolic pathways, most genes were found abundant at the surface, except for the 4-hydroxybutyrate (4HB) pathway (*abfD*, 4-hydroxybutyrate ligase), reverse TCA cycle (*ACLY, aclAB*), sulfate reduction
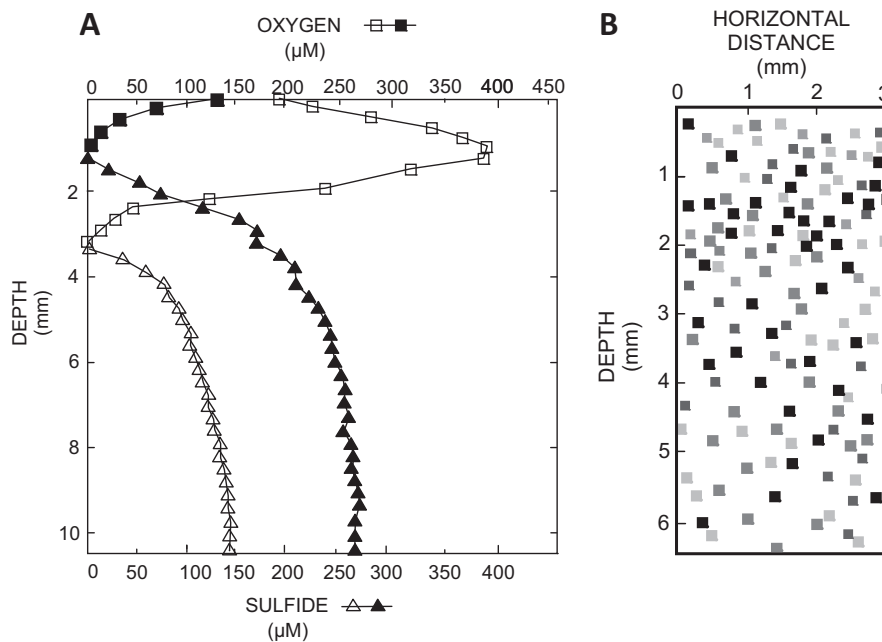
Fig. 3 Measurement of key biogeochemical properties in Shark Bay smooth mats. **a** In situ depth profiles of oxygen and sulfide concentrations. Oxygen and sulfide concentrations were measured during peak photosynthesis (1200–1400 h, open symbols) and the end of the night (0400–0500 h, closed symbols). Multiple profiles ($n = 3–7$) were measured and representative profiles are shown. Squares represent oxygen, triangles sulfide concentrations. **b** Two-dimensional distribution of sulfate reduction visualized using the ^{35}S-silver foil technique. Trace near the top of the panels indicates the surface of the mats. Pixels indicate hotspots of sulfate reduction; darker pixels represent higher rates

(cysNC, cysNCD, dsrAB), Wood–Ljungdahl pathway (cooS, coxSML, acsBCD), and C$_1$-associated metabolic pathways (hdrABC, mttB, mtrA) (Fig. 4). The wide range of functional genes identified suggests the smooth mat communities in Shark Bay possess versatile metabolic capabilities.

Phototrophy in Shark Bay mats

Light energy is one of the primary driving forces of various nutrient cycles in microbial mats [28, 64]. As expected, most of the phototrophy-related genes in the metagenomes were found in the present study at the mat surface, with a variety of psb, chl, and bch genes that encode photosystem, chlorophyll, and bacteriochlorophyll synthesis, respectively. Chlorophyll and bacteriochlorophyll synthesis genes were found more enriched at the photo-oxic zone (0–4 mm), but the latter has more varied depth distribution and only diminishes below 14 mm (Fig. 4). As expected, oxygenic photosynthesis genes and anoxygenic photosynthesis were mostly affiliated to Cyanobacteria and Alphaproteobacteria, respectively.
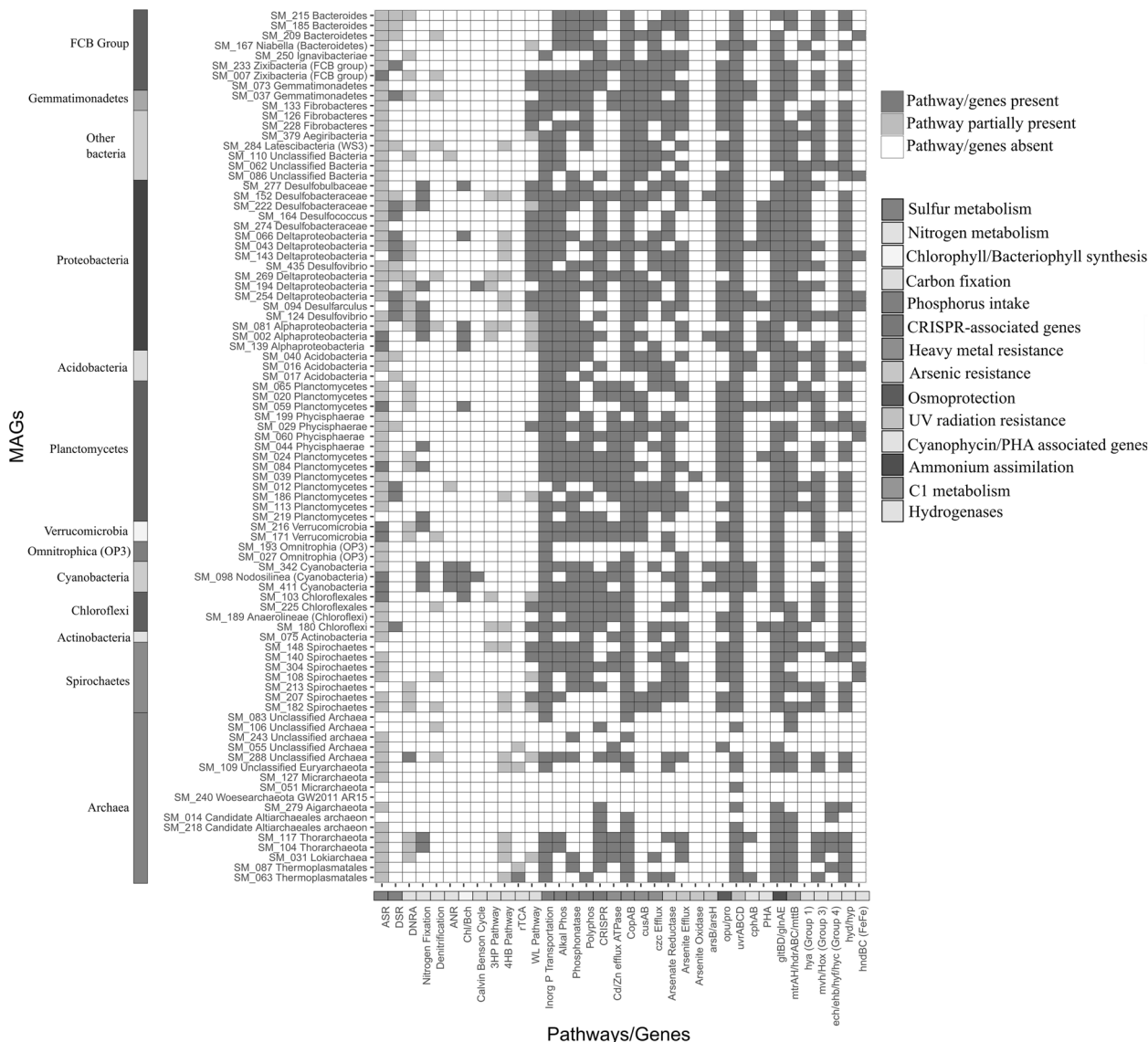
Carbon fixation

A range of carbon fixation pathways were identified in the smooth mat metagenomes, including the Calvin–Benson Cycle (CBB), acetogenesis via the Wood–Ljungdahl pathway (reductive acetyl-CoA), 3-hydroxypropinate cycle (3HP), 4-hydroxybutyrate pathway (4HB), as well as a low abundance of reverse TCA (rTCA) cycle genes (Fig. 3). Carbon monoxide dehydrogenase (cooS, coxSML) and acetyl-CoA synthase (acsB, cdhDE) are among the most abundant genes, indicating the potential significance of the Wood–Ljungdahl pathway in smooth mats (Fig. 4). Network correlation analysis can be used to describe putative direct or indirect interactions between microbial processes. Employing these analyses, reductive-acetate pathway genes (acsABCD) were found to be more abundant beneath the surface oxic zone, and these were negatively correlated to genes enriched at the surface (e.g., nifK, bchXY), while being positively correlated to genes involved in anaerobic pathways (e.g., hdrA, aclAB) (Supplementary Figure 1). Complete sets of genes encoding the Wood–Ljungdahl pathway of acetogenesis were found in a variety of MAGs, including Deltaproteobacteria, Spirochaetes, Planctomycetes, Chloroflexi, and Fibrobacteres MAGs (Fig. 5; Supplementary Table 4). The relative distribution of taxonomic groups involved in different nutrient cycles/metabolisms is shown in Supplementary Figures 2 to 5. The majority of the acetogens in these communities fall under the phylum Deltaproteobacteria. Genes encoding for the CBB were identified throughout the mat and were mostly affiliated to

Log 2 normalised counts
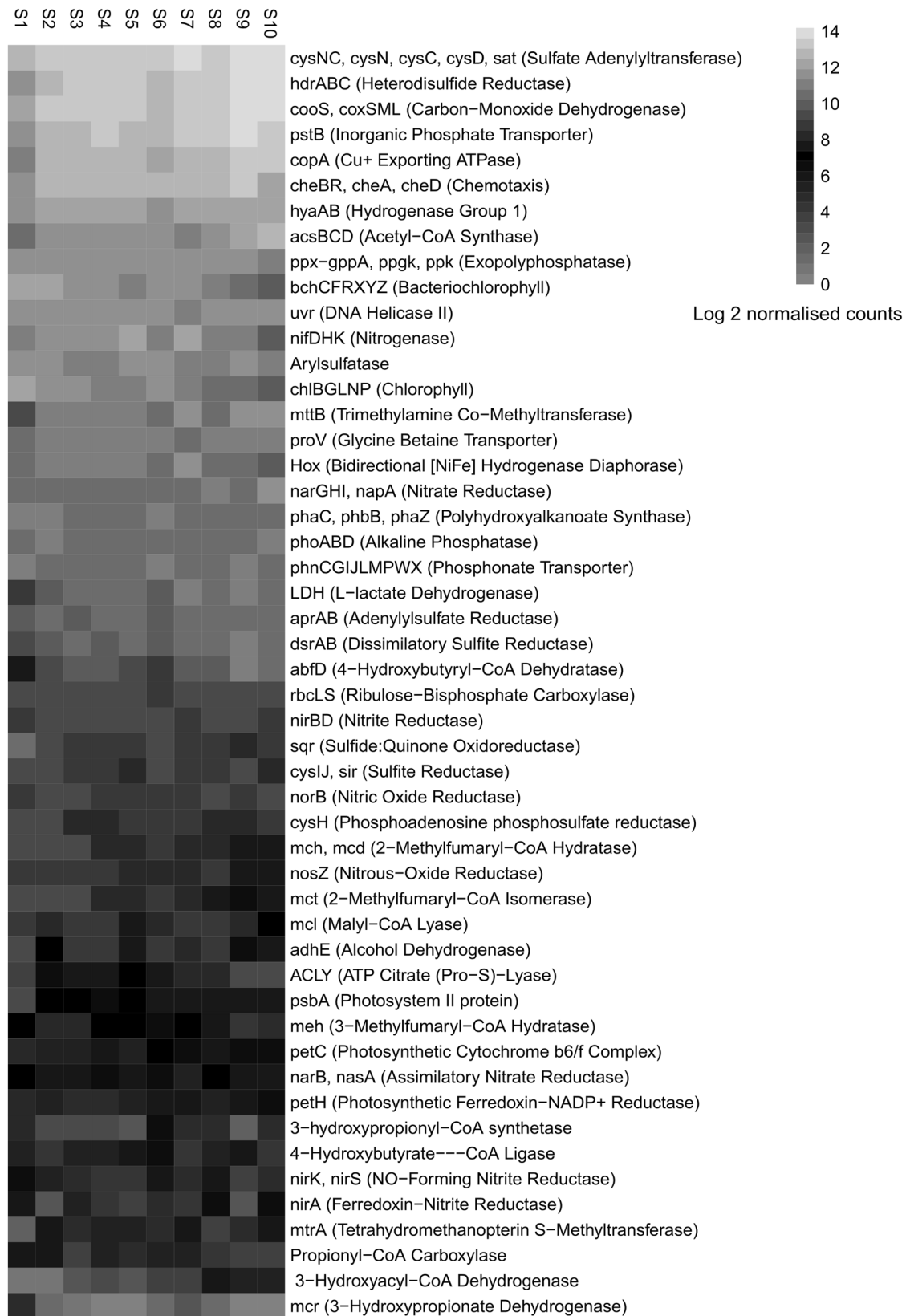
◀ **Fig. 4** Heatmap of distribution with depth of major functional genes in Shark Bay smooth mats. Differential analysis of count data (given as log2 normalized counts) was obtained using the *DESeq2* package in R and used to compare the abundance of genes with depth in the mats. A gradient from green to red indicates gene abundance across mat layers. Green represents genes that are more over-represented and red indicates genes that have the least relative abundance. X-axis indicates the different smooth layers in 2 mm increments (S1, 0–2 mm; S2, 2–4 mm; S3, 4–6 mm; S4, 6–8 mm; S5, 8–10 mm; S6, 10–12 mm; S7, 12–14 mm; S8, 14–16 mm; S9, 16–18 mm; S10, 18–20 mm), and the y-axis the major functional genes identified (color figure online)

Alphaproteobacteria and Cyanobacteria. RuBisCo genes (*rbcL*) were also identified in archaeal MAGs in Shark Bay (SM_031, 051, 127, 218, 243, 279, 288) (Fig. 5). However, they are likely non-photosynthetic archaeal RuBisCo (form II/ III) that only participate in the pentose bisphosphate pathway [65].

Genes encoding for proteins in the 3HP pathway (*mch, mct, mcl, meh*, 3-hydroxypropionyl-CoA synthetase) were more abundant in metagenomes at the mat surface,

Fig. 5 Color-coded table indicating major functional genes identified and their abundance in Shark Bay smooth mat draft MAGs. X-axis indicates specific genes likely involved in either nutrient cycling or environmental adaptation and y-axis the genome bin designations. Key: blue indicates the complete pathway identified in carbon, sulfur, and nitrogen nutrient cycles, genes present in heavy metal resistance, environmental adaptation, and hydrogenases; gray indicates the partial pathway identified in the carbon, sulfur, and nitrogen nutrient cycles; white indicates genes and pathways that are absent. On the left the color panel correspond to the phyla as shown in Fig. 1. On the right the colors represent different genes and pathways. Abbreviations, ASR assimilatory sulfate reduction, DSR dissimilatory sulfate reduction, ANR assimilatory nitrate reduction, Chl/Bch chlorophyll/bacterio-chlorophyll production, 3HP pathway 3-hydroxypropinate pathway, 4HB pathway 4-hydroxybutyrate pathway, rTCAd reverse Krebs cycle, WL pathway Wood–Ljungdahl pathway (color figure online)

while 4HB pathways (*abfD*, 3-hydroxyacyl-CoA dehydrogenase) were more enriched at the bottom (Fig. 4). 3HP genes were affiliated to Chloroflexales and Alphaproteobacteria. Two Alphaproteobacteria MAGs (SM_002, 081) encode for a near-complete 3HP pathway (both lacking 3-hydroxypropionyl-CoA synthetase and 3-hydroxypropionate dehydrogenase), while one Spirochaetes (SM_148), two Chloroflexi (SM_103, 180), and three Deltaproteobacteria MAGs (SM_152, 194, 269) encode for partial 3HP pathways (Supplementary Table 4). 3-Hydroxyacyl-CoA dehydrogenase (enoyl-CoA hydratase) and 4-hydroxybutyryl-CoA dehydratase were detected in a variety of bacterial MAGs, as well as the Asgard group Lokiarchaeota and Thorarchaeota, indicating that these MAGs possess genes for a putative partial 4HB pathway (Fig. 5). A low abundance of genes encoding the reverse TCA cycle were detected and were only observed in a Thermoplasmata MAG (SM_063), suggesting the rTCA cycle may not be a significant anaerobic carbon fixation mode in smooth mats in Shark Bay.

CaZy family and carbon metabolism

The CaZY enzymes facilitate synthesis and breaking down of polysaccharides and extracellular polymeric substance (EPS), providing a putative carbon source available for the heterotrophic community [40]. Top carbohydrate-active enzymes (CaZy) genes identified in mat metagenomes were glycoside hydrolase family 13 (GH13), carbohydrate-binding module family 48 (CBM48), and glycosyltransferase family 2 and 4 (GT4 and GT2). These genes were over-represented at the surface layer (0–2 mm) (Fig. 6). Complete Embden–Meyerhoff pathways were encoded in the smooth mat metagenomes, along with aldehyde dehydrogenase (*adhE*) and L-lactate dehydrogenase (*LDH*), demonstrating the potential sugar fermentation capability of smooth mat communities. Polyhydroxyalkanoate (PHA) synthase genes (*phaC/phbC*, *phaZ*) were also detected and over-represented at the photo-oxic zone (0–4 mm), and PHA serves as a potential storage of fixed carbon [66]. Alphaproteobacteria appear to be the main PHA producers (Supplementary Figure 2), and PHA synthase genes were only identified in Proteobacterial, Planctomycetes, and Chloroflexi MAGs (Fig. 5). Both Thorarchaeotal MAGs (SM_104, 117), five archaeal MAGs (SM_051, 055, 063, 087, 127), and one Latescibacteria (formerly WS3) MAG (SM_284) harbor fructose 1,6 bisphosphate aldolase (K01622). All cyanobacterial MAGs (SM_98, 342, 411) contain genes encoding cyanophycin (*cphA*), cyanophycinase (*cphB*), and carbon-concentrating mechanisms (*ccmKLMNO*). In addition to the Cyanobacteria, nine MAGs encode for cyanophycinase, including Gemmatimonadetes, Deltaproteobacteria, Bacteroidetes, Thermoplasmata, Thorarchaeota, Planctomycetes, and Acidobacteria (Fig. 5; Supplementary Table 4).

C$_1$-associated metabolisms

Heterodisulfide reductase, tetrahydromethanopterin S-methyltransferase, F420-non-reducing hydrogenase, and trimethylamine co-methyltransferase (*Hdr, mtr, mvh*, and *mtt*) were identified related to C$_1$-associated metabolisms. The key gene in methanogenesis, *mcrA*, encoding methyl-coenzyme M reductase, was not detected in any of the MAGs in the present study; however, *mcrA* was identified in the metagenomes when blasted against PFAM and TIGRFAM databases. Genes encoding a putative WL pathway using either tetrahydrofolate (THF) or tetrahydromethanopterin (THMPT) as a C$_1$ carrier were also identified, with a Thorarchaeota MAG (SM_104) encoding a complete THMPT–WL pathway (Supplementary Table 4). Of the 87 MAGs, 61 harbor *hdr, mtr, mvh,* or *mtt*, suggesting a high proportion of MAGs capable of C$_1$-associated metabolisms.

Sulfur cycling

Complete sets of genes responsible for both assimilatory and dissimilatory sulfate reduction (DSR) were identified in all depths in the mat metagenomes. Sulfate reduction genes were more enriched with depth (Fig. 4), and this complements the microelectrode data, which indicated that sulfide concentrations increased with depth, with a zone of 3–8 mm designated as a 'H$_2$S rich' zone (Fig. 3a). Interestingly however, ^{35}S-silver foil measurements indicate high sulfate-reducing activity also near the surface (Fig. 3b), supportive of either O$_2$-tolerant sulfate reducers [67] or a putative surface anoxic niche. Genes *dsrA* and *dsrB* (encoding dissimilatory sulfite reductase) were detected in the present study though not found in previous shotgun sequencing of the Shark Bay mats using the 454-sequencing platform [5], which may be due to higher coverage using the Illumina platform here. Assimilatory sulfate reduction genes were mostly affiliated to Alpha- and Deltaproteobacteria, Bacteroidetes, Chloroflexi, and Firmicutes; as depth increased, these genes were mostly recovered from Deltaproteobacteria and unclassified bacteria (Supplementary Figure 3). On the other hand, Deltaproteobacteria and Firmicutes harbor most of the genes involved in DSR. Partial assimilatory sulfate reduction pathways were present in the majority of MAGs (76 of 87), and ten MAGs encoded complete assimilatory sulfate reduction. Most Deltaproteobacterial MAGs harbored a complete set of genes for DSR, and surprisingly
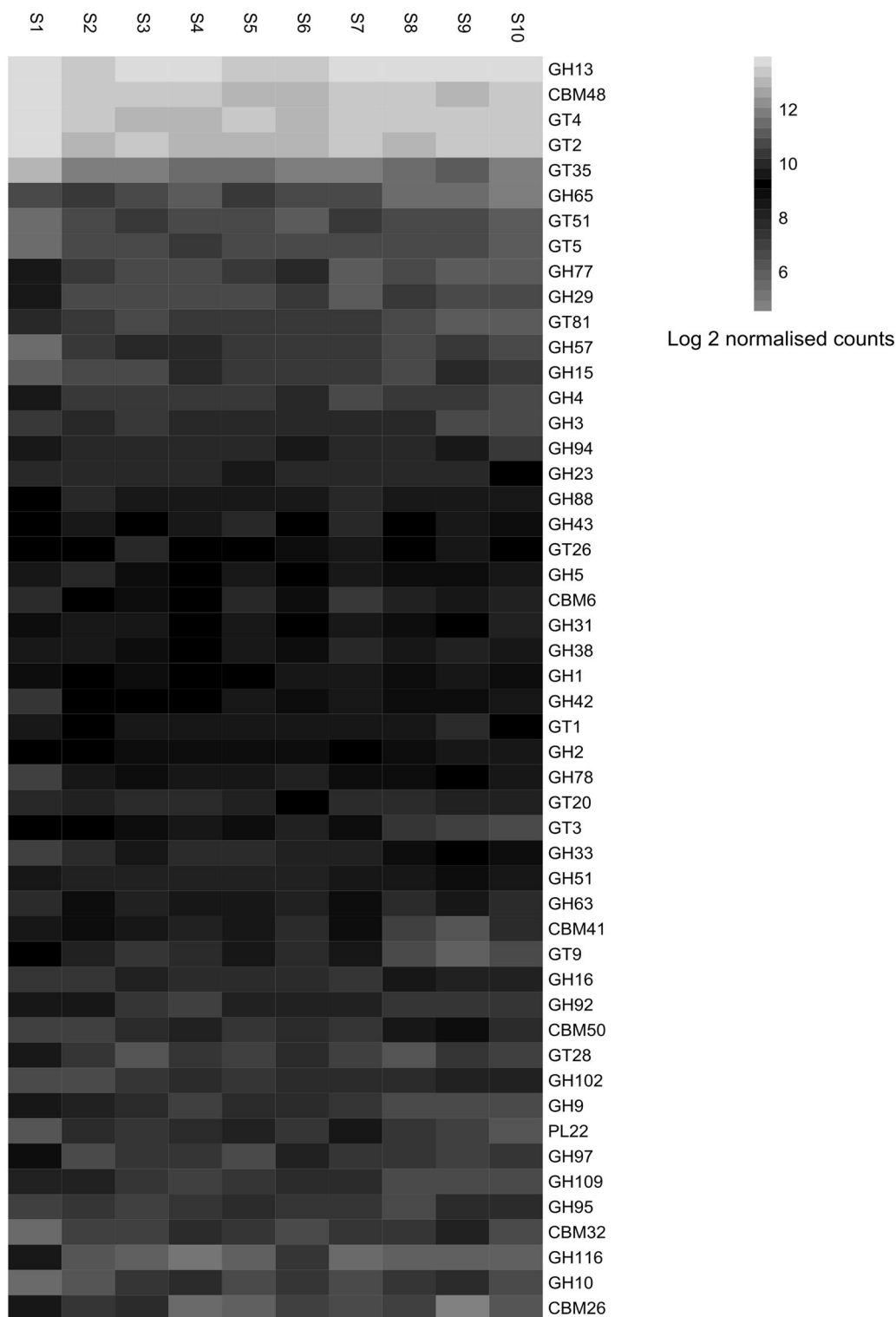
Fig. 6 Heatmap of distribution of genes encoding carbohydrate-active enzymes (CaZY) in Shark Bay smooth mats. Differential analysis of count data (given as log2 normalized counts) was obtained using the *DESeq2* package in R and used to compare the abundance of the top most 50 CaZY genes with depth in the mats. Green represents genes that are more over-represented and red indicates genes that have the least relative abundance. X-axis indicates the different smooth layers in 2 mm increments (S1, 0–2 mm; S2, 2–4 mm; S3, 4–6 mm; S4, 6–8 mm; S5, 8–10 mm; S6, 10–12 mm; S7, 12–14 mm; S8, 14–16 mm; S9, 16–18 mm; S10, 18–20 mm), and the y-axis the major functional genes identified. GH glycoside hydrolase, CBM carbohydrate-binding modules, GT glycosyltransferase, PL polysaccharide lyase

were also detected in a Chloroflexi MAG (SM_180), a Gemmatimonadetes MAG (SM_037), two Planctomycetes MAGs (SM_012, 186), and a candidate phylum Zixibacteria MAG (SM_233; Fig. 5; Supplementary Table 4). Three Bacteroidetes MAGs (SM_185, 209, 215) contain genes encoding an incomplete DSR pathway. The genetic basis for sulfide assimilation (*cysM, cysK*), organic sulfur scavenging (arylsulfatase), polysulfide metabolism (*sqr*), and sulfur oxidation (*sor*) were also identified in the mat metagenomes, though genes encoding thiosulfate oxidation (*sox*) were not found.

Nitrogen cycling

Complete pathways of dissimilatory nitrate reduction to ammonium (DNRA), assimilatory nitrate reduction, denitrification, and nitrogen fixation were identified in all layers of smooth mats. Nitrate reductase (*NarGHI*) was more enriched in deeper anoxic layers, while nitrite reductase (*nirKS*) and nitrous-oxide reductase (*nosZ*) were more abundant near the surface (Fig. 4). Nitrogen fixation genes (*nifDHK*) were more abundant beneath the surface, and were mostly affiliated to Deltaproteobacteria (Supplementary Figure 4). These data suggest that although cyanobacteria are likely the major players in nitrogen fixation in these systems, the Deltaproteobacteria may also play a role in this process in the Shark Bay mats. Analysis of metagenomes here suggests Bacteroidetes, Alpha- and Deltaproteobacteria appear to be responsible for DNRA and denitrification (Supplementary Figure 4). Assimilatory nitrate reduction genes were associated with Cyanobacteria at the surface, while in deeper layers they were associated with Gammaproteobacteria and Halobacteria (Supplementary Figure 4). Most of the MAGs have partial DNRA and denitrification pathways, with the exception of an unclassified archaeal MAG (SM_288) that encodes for a complete DNRA pathway. Nitrogen fixation genes spread widely among Planctomycetes, Verrucomicrobia, Alpha- and Deltaproteobacteria, and also in both Thorarchaeotal MAGs. Interestingly, it appears there is an incomplete nitrogen cycle in the Shark Bay mats, as genes from known nitrification pathways were not present in the mat metagenomes. In particular, genes encoding ammonia monooxygenase and hydrazine synthase (*amo, hzs*) were not detected. However, there may be genes classified as unknown or potentially novel proteins that could fulfill these roles in Shark Bay and future studies are needed to delineate this. Furthermore, genes encoding glutamate synthase (*gltBD*) and glutamine synthase (*glnAE*) were present in the metagenomes that are associated with nitrogen/ammonia assimilation.

Phosphorus uptake

Inorganic phosphate transport genes (*pstB*) were one of the most abundant found in the mat metagenomes (Fig. 4). Genes encoding alkaline phosphatase (*phoAB, phoD*), phosphonate metabolism (*phn*), polyphosphate kinase (*ppk*), and exopolyphosphatase (*ppx*) were also identified in all depths and among a variety of phyla, indicating likely widespread phosphorus-scavenging mechanisms in smooth mats (Supplementary Figure 5). Of the 87 MAGs, 80 possess phosphorus-salvaging genes, likely reflecting the importance of phosphorus uptake in the mats. Those lacking phosphorus-salvaging genes are all archaea (SM_014, 051, 055, 106, 218, 240, and 279).

Hydrogen metabolism

A wide variety of genes encoding [NiFe]-hydrogenases were identified in the metagenomes and MAGs, including *HoxEFHUY, hydAB, hyaABD, hypABCDEF MvhAG, ech*, and *ehb* hydrogenases, along with [Fe]-hydrogenases *hndBCD*, which could be involved in hydrogen production and consumption (Fig. 5). The extensive distribution of these hydrogenases among both aerobes and anaerobes suggests that H$_2$ is potentially a highly dynamic electron carrier produced and consumed by a wide range of microbial mat members. A high abundance of genes encoding oxygen-tolerant hydrogenases (*hyaAB* and *HoxEFU*) are evenly distributed throughout the mats (Fig. 4). The majority of the *HoxEFU* genes were affiliated to Cyanobacteria, unclassified bacteria, and Chloroflexi, while *hyaAB* were mostly related to Proteobacteria (Supplementary Figure 3). This indicates a potentially high level of hydrogen metabolism and may benefit the abundant hydrogenotrophs identified by 16S rDNA analyses [21]. Several genes encoding dehalogenases were also identified, albeit at low abundance: namely, 2-haloacid dehalogenase, haloacetate dehydrogenase, and haloalkane dehydrogenase.

Environmental adaptation

Unsurprisingly, given the high salinity fluctuations these microbial mats are exposed to [21], 51 of the MAGs were found to encode for osmoprotectant transport systems (*opuABCD*) and glycine betaine/choline transporter (*proV/opuABCD, betTS*). In addition, mat metagenomes contain a range of heavy metal resistance genes, including copper and silver (*copAB, cusAB*), cobalt–zinc–cadium (*czcABCD*), arsenic (arsenate reductase, arsenite transporter, arsenite oxidase, and arsenical resistance protein), and a range of other heavy metal resistance genes (Fig. 7). Copper efflux genes (*copAB, cusA*) and arsenite transporter genes (*ACR3*)
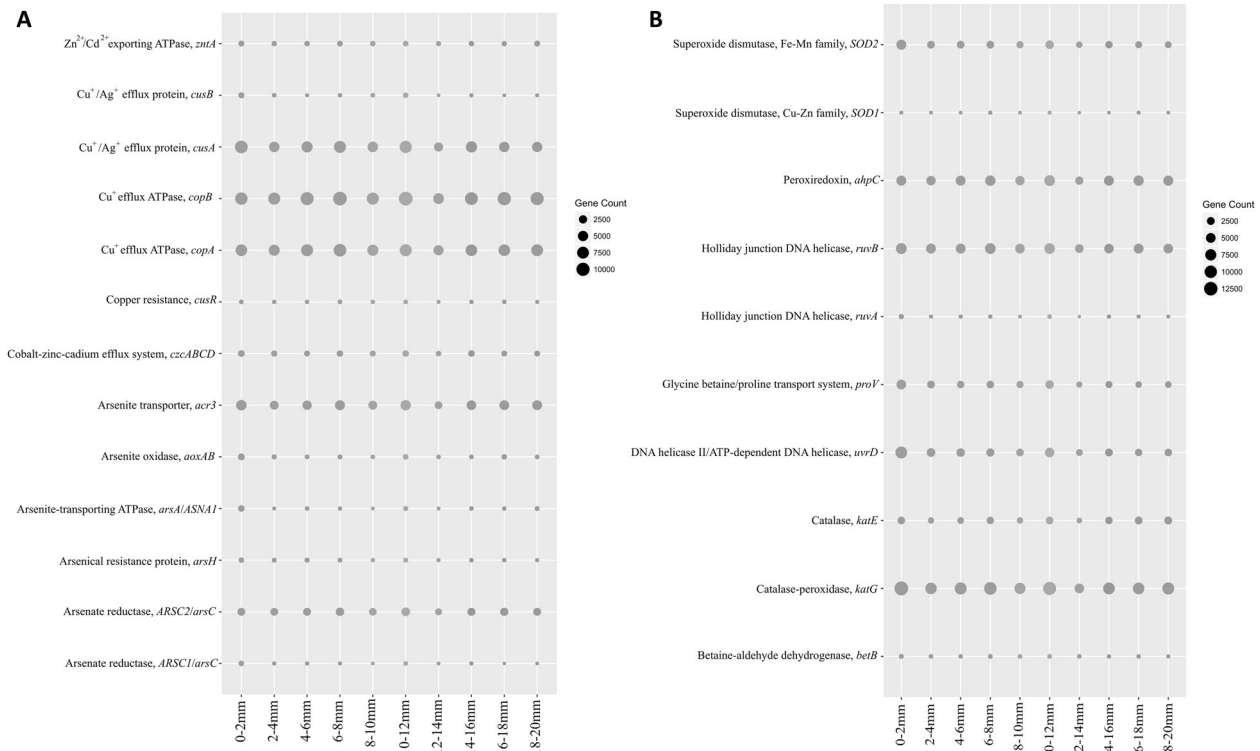
Fig. 7 Bubble plot illustrating traits of environmental adaptation at the metagenomic level in Shark Bay smooth mats delineated with depth. **a** Distribution of genes related to heavy metal resistance. **b** Distribution of genes related to hypersalinity, UV radiation, and oxidative stress

are most abundant. Of the 87 MAGs, 75 have arsenate resistance genes (*arc3, arsC, arsA, arsB*, and *arsH*) and one Planctomycetes MAG (SM_039) contains genes that encode for energetic arsenic respiration genes (*AoxAB*). In the Shark Bay metagenomes aoxAB genes were affiliated to Alphaproteobacteria, and these genes are responsible for respiratory arsenite oxidation and related to energetic metabolism in other mats [68].

Genes that encode for UV radiation resistance, excinuclease ABC (*uvrABC*) that facilitate nucleotide excision repair upon UV damage [69], were also identified and were distributed evenly at different depths. Holliday junction DNA helicase (*ruvAB*) were also identified that can alleviate ultraviolet lesions (Fig. 7; [70]). *UvrABC* genes were found in all MAGs except six archaeal MAGs (Fig. 5). High UV radiation usually leads to production of reactive oxygen species (ROS), including peroxide, superoxide, and hydroxyl radicals that are toxic to microbial communities, and this adaptive response capability was observed with high levels of catalase-peroxidase (*katG, katE*), superoxide dismutase (*SOD1, SOD2*), and peroxiredoxin (*ahpC*) in smooth mat metagenomes. Cell motility and chemotaxis-related genes (*cheBR, cheA, cheD*) were identified and are more abundant beneath the surface photo-oxic zone (Fig. 4). Chemotaxis genes are mostly affiliated to

adaptation. X-axis indicates the different smooth mat layers in 2 mm increments, and y-axis indicates major functional genes/pathways. Circular scale bar represents gene count as relative abundance of genes associated with the displayed functions

Deltaproteobacteria, Cyanobacteria, and Spirochaetes (Supplementary Figure 5).

Microbial communication involving signal molecules, often facilitated by the phenomenon of quorum sensing, is another process that can allow environmental adaptation and improve ecosystem function in complex microbial communities. Quorum sensing genes were identified in 14 MAGs in smooth mats. Two Alpha- and one Fibrobacteres MAG (SM_002, 139, 228) were confirmed to have both N-acyl homoserine lactones (AHL) synthesis as well as LuxR type response elements (Supplementary Table 4). LuxR type proteins were also detected in Alphaproteobacteria MAG (SM_081) although no synthesis proteins were detected. Several genes encoding for putative quorum quenching systems were also detected: AHL lactonases and AHL acylases. AHL lactonases were detected in two groups: Planctomycetes (SM_024, 039, 059) and Deltaproteobacteria (SM_254, 269). Putative AHL acylases were detected in Spirochaetes (SM_182), Gemmatimonadetes (SM_037), Alphaproteobacteria (SM_139), Deltaproteobacteria (SM_274), Acidobacteria (SM_040), and Chloroflexales (SM_225) (Supplementary Table 4).

Microbial defense systems against foreign DNA include clustered, regularly interspaced, short palindromic repeat (CRISPR) systems, which have been identified as an

adaptive microbial immune system that provides acquired immunity against viruses and plasmids [71]. CRISPR-related genes were predominant in the Shark Bay mat metagenomes, as well as the novel virus defense mechanism involving bacteriophage exclusion, termed BREX. Over half of the MAGs (49) encode for CRISPR-related genes, including CRISPR-associated endonuclease/helicase (*cas*) and CRISPR-associated proteins (*cmr, cst, csh, csb, csn,* and *csx*; Fig. 5). Core genes of the BREX systems were also identified, including the *pglZ* domain-associated protein and the *brxC/pglY* (DUF2791) domain associated protein. A newly discovered viral defense system, DISARM (defense island system associated with restriction-modification [72]), was also identified in mat metagenomes and draft MAGs. Twenty six of the MAGs were found containing a protein domain DUF1998 (pfam09369/*drmB*), which was located on the same contig with a pfam00271 domain (putative helicase domain/*drmA*; Supplementary Table 4). This indicates the smooth mat community may have the capacity to harbor the novel DISARM system for viral defence.

Discussion

The present investigation is the first to examine at a millimeter scale the functional genetic potential of Shark Bay microbial mats. Based on the complementary microelectrode and metagenomic data obtained, there appears to be a metabolic 'hotspot' at the surface of smooth mats, where most of the autotrophic genes were recovered, along with sulfur, carbon, nitrogen, and phosphorus cycling, as well as carbohydrate and amino-acid metabolisms. Although cyanobacteria are thought to be the primary producers in microbial mats [73], other major carbon fixation pathways were also identified in high abundance in smooth mats, in particular the Wood–Ljungdahl (WL) pathway, 3HP bi-cycling, and the 4HB pathway (Fig. 4). The top five abundant set of genes (excluding DNA polymerase and other housekeeping genes) in the Shark Bay mat metagenomes were affiliated to sulfate assimilation (*cysNC cysNCD, sat*), methanogenesis (*hdrABC*), Wood–Ljungdahl pathway (*cooS, coxSML*), phosphate transport (*pstB*), and copper efflux (*copA*), respectively. This is in concordance with the high sulfate reduction and methane production rates in these mats recorded previously [3, 21], as well as the extremely limited phosphorus and surprisingly high copper concentrations detected here. It is possible that the relative high abundance of genes encoding carbon monoxide dehydrogenases (*cooS, coxSML*) is due to the oligotrophic settings and high UV irradiation the mats in Shark Bay are exposed to [19, 74].

Vertical spatial arrangement of functional genes in the metagenomes corroborates with previous 16S rDNA data, in which cyanobacteria and Alphaproteobacteria are hypothesized to be in a surface phototrophic niche [3]. In the

metagenomes analyzed here, both these groups had high affiliation to phototrophic systems and the Calvin–Benson Cycle. The data suggest that nitrogen fixation activity is more prominent near the surface, as the genes for this process are negatively correlated to genes more enriched in deeper layers (*hdrB, acsAB*) (Supplementary Figure 1). It is proposed that the existence of surface anaerobic/suboxic microniches in these mats serves to protect the strictly anaerobic Deltaproteobacteria and archaea that were shown previously to be at a relatively high abundance in the surface oxic zone [3, 21]. Indeed, the presence of Deltaproteobacteria with high metabolic activity at the oxic surface has been reported for several other microbial mats [75–77].

Deep-branching putatively novel archaea

Asgard group archaeal MAGs were reconstructed from Shark Bay mat metagenomes sequenced in the present study, and this group has been suggested to be the origin of eukaryotes and exhibit eukaryotic features [62]. To our knowledge this is the first report of this group of archaea in microbial mats. Both Thorarchaeota MAGs identified here (SM_104, 117) encode for class I fructose-bisphosphate aldolase (K01622), which was suggested to be the ancestral gluconeogenic enzyme [78]. By reconstructing genomes from members of the Asgard group, the current investigation has significantly increased the amount of available genomic data on this enigmatic novel group of archaea, which may have critical functions in mat systems. Based on the MAG data, it is proposed that the Shark Bay Lokiarchaeota MAG (SM_031) is a hydrogen-dependent, anaerobic autotrophic archaeon, harboring a THMPT (H$_4$MPT)-dependent Wood–Ljungdahl pathway, characteristics also observed in a Lokiarchaeota genome in a previous study [53]. Members of the DPANN were also identified in here, which were also discovered in other saline habitats [79, 80]. Future studies may help unravel the role of 'microbial dark matter' [63, 81], such as the Asgard group, in Shark Bay microbial mats.

Carbon cycling

Genes encoding for the 3-hydroxypropionate/4-hydroxybutyrate pathway genes were identified in the metagenomes here, which is in agreement with previous work on these mats indicating a prevalence for this pathway [5]. Genes associated with the Wood–Ljungdahl pathway (*cooS, coxSML*) were over-represented, which may facilitate the oxidation of CO for energy or coupling to CO$_2$ fixation [82]. It has been proposed that heterotrophs utilize CO as an alternative carbon source for energy conservation in oligotrophic environments [83], and such an adaptive response may be prevalent in Shark Bay. It is possible that due to the oligotrophic settings and high UV irradiation the mats in

Shark Bay are exposed to [19, 74], photo-degradation of surface organic matter may occur, providing CO as an alternative carbon source for energy conservation [82, 83]. Moreover, the WL pathway is suggested to be one of the more ancient carbon fixation pathways and can be coupled to methanogenesis [84, 85]. As this pathway is proposed to be a major carbon fixation pathway in the modern day Shark Bay mats, it suggests these systems may be good analogues of ancient mats as previously proposed [21, 22].

Carbohydrate enzymes (CaZY) and EPS

EPS have been shown to be major components increasing the structural integrity of many microbial mats [86]. GT2 and GT4 enzymes are involved in formation of various polysaccharides [87, 88], which could potentially contribute to the construction of EPS in the smooth mats in Shark Bay. GH13 identified in metagenomes here (Fig. 6) represents the most abundant family of glycoside hydrolyases (GHase) as it targets a wide range of polysaccharides with alpha-1,4-glycosidic bonds. GH13 is also the most abundant GHase found in the Guerrero Negro hypersaline microbial mats [7], and mat-forming cyanobacteria have been shown to secrete alpha-1,4-oligosaccharides into EPS to reuse excess organic carbon [89]. In microbial mats, it has been shown previously that EPS can be reused as a result of cyanobacterial photorespiration [90]. Genes encoding all the described carbohydrate enzymes are over-represented at the surface of smooth mats (Fig. 6), reflecting a potential abundance of EPS and polysaccharides in this zone. Microbial mat surfaces usually have the highest productivity and organic carbon turnover rates [64], and it is proposed that smooth mats maintain an extra source of extracellular organic carbon through constant EPS construction and degradation. This could potentially provide a dynamic carbon source for the surface microorganisms identified here and in previous studies [3], such as Cyanobacteria, Proteobacteria, and Bacteroidetes.

Potential carbon storage

In addition to extracellular organic carbon sources, cyanobacteria are likely the main nitrogen assimilators at the surface, with the potential to store organic nitrogen and carbon in cyanophycin with genes encoding cyanophycin synthetase (*cphA*) (Fig. 5; Supplementary Table 4), as well as accessing nitrogen with cyanophycinase (*cphB*) [91, 92]. Genes encoding cyanophycinase (*cphB*) were detected in draft MAGs of Gemmatimonadetes, Deltaproteobacteria, Bacteroidetes, Thermoplasmata, Thorarchaeota, Planctomycetes, and Acidobacteria. Although it is not possible to ascertain whether active herbivory is taking place, at the least it does emphasize a wide range of

microorganisms possessing the capacity to potentially scavenge cyanophycin from dead cyanobacteria in the mats and access the C/N therein. The presence of *phaC/phaZ/phbB* genes suggests that glycogen or photosynthate fermented by the smooth mat community can potentially be stored as polyhydroxyalkanoates (PHA) by various Proteobacteria and Planctomycetes, proposing a scenario that links carbon flux and cross-feeding between cyanobacteria and heterotrophs in the mats. These add another dimension in smooth mat carbon cycling. Dynamics of PHA are closely related to carbon cycling, which tends to be accumulated by microorganisms under oligotrophic conditions [93]. Chloroflexi was the most abundant bacterial group identified in previous 16S rDNA analyses of these mats [3]; however, its role in microbial mat carbon cycling here is still unknown other than potentially fixing carbon through the 3HP pathway.

Methanogenesis/C$_1$ metabolism

Methanogenesis/C$_1$ metabolism-associated genes were identified in mat metagenomes (*hdrABC, mtrAH, mvhAG,* and *mttB*), and this is in concordance with recent work indicating a prevalence of methanogens in Shark Bay mats based on 16S rDNA analyses and methane production [21]. Previous work has suggested that hydrogenotrophic methanogenesis appears to be the main form of methane production in these mats [21]. Acetate produced from the Wood–Ljungdahl pathway could be putatively metabolized by hydrogenotrophic methanogens, through the utilization of the methyl branch for methane formation accompanied with energy production [94]. Interestingly, the primary gene for methane production, *mcrA*, encoding methyl-coenzyme M reductase, was absent in all of the MAGs in the present study and only detected in bulk metagenomes when blasted against PFAM and TIGRFAM databases. This may be due to incomplete coverage of the MAGs, or there may be another novel gene responsible for methanogenesis. The presence of genes in MAGs encoding both THMPT–WL and THF–WL pathways, along with *hdr/mtr/mtt*, suggest a potentially more generalized role of C$_1$ metabolism in these mats [7, 95, 96].

Putative partitioning of sulfur and nitrogen cycles

Dissimilatory sulfate reduction genes were affiliated to Chloroflexi, Deltaproteobacteria, and Firmicutes (Supplementary Figure 5). Deltaproteobacteria and Firmicutes possess genes that encode *aprAB* and *dsrAB*, which may then reduce activated sulfates to sulfides. In addition, Chloroflexi, Gemmatimonadetes, and candidate phylum Zixibacteria may potentially take part in DSR, as a Chloroflexi MAG (SM_180), a Gemmatimonadetes MAG (SM_037), and a candidate phylum Zixibacteria MAG

(SM_233) were found harboring all *aprAB* and *dsrAB* genes. This is corroborated by studies in marine and estuary sediments [97–99], suggesting that these phyla may play a previously unacknowledged role in sulfur cycling and possess a putative expanded metabolic function. Genes encoding sulfide:quinone reductase (*sqr*) were abundant particularly at the surface of the Shark Bay mats (Fig. 4), suggesting that sulfide can be putatively converted to polysulfide and subsequently assimilated, linking inorganic and organic sulfur cycles as polysulfides react promptly with organic matter [100, 101]. Sulfide assimilation genes (*cysM*, *cysK*) allow assimilation and conversion of sulfides to cysteine for further metabolism. The majority of the draft genomes (76 out of 87) encode for partial and complete assimilatory sulfate reduction pathways, indicating the potential importance of sulfur respiration in the Shark Bay ecosystems. Arylsulfatase is responsible for the desulfurization of sulfated polysaccharides, which could putatively have a role in EPS degradation and utilization of sulfur-abundant exopolysaccharides, such as those produced by sulfate-reducing activity and sulfur assimilation [10, 102]. Most of the arylsulfatases were affiliated to Alphaproteobacteria, Bacteroidetes, and Planctomycetes (Supplementary Figure 3), which reflects the putative importance of sulfur metabolism and the versatile tools of smooth mat communities in potentially utilizing alternative sulfur sources.

Interestingly, with the exception of one archaeal MAG (SM_288), there were no MAGs present in the Shark Bay mat metagenomes that encode for complete DNRA and denitrification, suggesting division of labor for both processes. As *NarGHI* that encodes nitrate reductase was more enriched in deeper layers while genes encoding *nirKS* (nitrite reductase) and *nosZ* (nitrous-oxide reductase) were more abundant near the surface, this also suggests partitioning of the nitrogen cycle with respect to depth. Furthermore, there is no evidence at the metagenomic level that the mat communities are capable of ammonia oxidation, as ammonia monooxygenase (*amo*) and hydrazine oxidoreductase genes (*hzs*) were not detected. This is also consistent with recent 16S rDNA analyses of Shark Bay microbial mats that suggest few bacterial nitrifiers present, leading to a potential incomplete nitrogen cycle in smooth mats [3, 5]. Any build-up of ammonium produced by DNRA microorganisms is hypothesized to be assimilated by other members of the mat community, as a high number of glutamine synthase (*glnA*, *glnE*) and glutamate synthase genes (*glnB, glnD*) were identified throughout the mat. A similar scenario has been observed in Socompa Lake stromatolites [68] and hypersaline mats from Hot Lake [103], where it was found that ammonium is likely assimilated by microbial community members instead of being oxidized.

Phosphorus scavenging

Phosphorus limitation has been reported in Shark Bay [104], and this is in agreement with analyses of water overlying the mats obtained in the present study. Inorganic phosphate transport genes (*pstB*) were in high abundance that is likely critical for inorganic phosphorus uptake. Genes encoding alkaline phosphatase (*phoAB*, *phoD*) were also found throughout the Shark Bay mat metagenomes, and these proteins are responsible for salvaging phosphorus from organic molecules in the community, providing a potential alternate mechanism to acquire phosphorus, as has been observed in microbial mats in Hot Lake and Cuatro Cienegas [10, 103]. Smooth mat metagenomes also contain a wide range of phosphonate-scavenging genes, including those encoding phosphonate transporters (*phnC*), phosphonate metabolism (*phnG, phnI, phnL, phnM, phnP*), and C-P bond lyases (*phnJ*), that can putatively breakdown phosphonate for assimilation. Other phosphatase genes detected include those that encode aminotransferase (*phnW*) and phosphonoacetaldehyde hydrolase (*phnX*), indicating the potential of Shark Bay mat communities to utilize organophosphorus compounds despite phosphorus limitation. Phosphonates are also suggested to be a prebiotic phosphorus carrier on the early Earth, and thus phosphatases may be a trait of ancient microbial mats [10]. Polyphosphate kinase genes were detected in the metagenomes here, and the proteins encoded can be used to breakdown polyphosphate, which can also serve as a phosphorus reservoir [105]. Although the eight archaeal MAGs identified in the mat metagenomes here lack any phosphorus scavenging genes, it has been suggested that archaea in hypersaline environments utilize genomic DNA as a phosphate storage polymer under phosphorus starvation [106].

Hydrogen metabolism

A wide range of hydrogenases were identified here, spanning from Group 1, Group 3, Group 4 [NiFe], and Group A [FeFe] hydrogenases in a range of bacteria and archaea, predicting active H_2 production and consumption in these mats [99, 107]. H_2-evolving bidirectional [NiFe] hydrogenases (*HoxEFU*) are found in a range of MAGs including Cyanobacteria, Chloroflexi, Proteobacteria (Fig. 5), and are capable of evolving H_2 via fermentation [66, 108]. H_2 production during N_2-fixation could be an important contributor to net H_2 production in microbial mats [109]. Bidirectional [NiFe] hydrogenase was mainly affiliated to Cyanobacteria at the surface in addition to unclassified bacteria beneath the surface, and it is proposed that H_2-production is most likely fermentative [66]. Membrane-bound H_2-uptake [NiFe]-hydrogenases (*hyaAB*) have been shown to contribute to electron input during microbial

aerobic and facultative anaerobic respiration [107]. Both types of hydrogenases could potentially contribute to the hydrogen that supports the hydrogenotrophic methanogens identified previously in these mats [21]. The hydrogenases identified here may provide a putative link between H_2 production and methanogenesis, and could potentially be an indicator of an ancient trait retained in the modern mats in Shark Bay. It is widely hypothesized that H_2 was the primordial electron donor, suggesting a potential evolutionary significance on early Earth [107]. Hydrogen-oxidizing methanogenic archaea and acetogenic bacteria were suggested to be crucial for conserving energy in anoxic environments in early Earth [110]. [NiFe]-hydrogenase group 3c (*MvhAG*) genes, found in one-third of the MAGs in the present study, is a critical energy-conserving step in methanogenesis and was proposed to be crucial for anoxic environments on early Earth [110]. Thus these mats in Shark Bay may putatively have retained processes similar to those found on early Earth.

Environmental adaptation

The high number of genes affiliated with copper resistance corroborates with the high copper levels measured in Shark Bay water in the present study. The most abundant heavy metal resistance genes were *copAB* and *cusAB*, which are efflux pumps that help maintain copper homeostasis [111]. A total of 23 MAGs analyzed here encode both an *acrB* extrusion pump and *arsC* reductase, indicating the potential for arsenic resistance [68]. Despite Shark Bay water having low arsenic levels (Supplementary Table 2), arsenic metabolism has been proposed to have a substantial role on primitive Earth, and may have been coupled to carbon and nitrogen cycles [112, 113]. AoxAB genes were identified in metagenomes in the present study, and it has been suggested these genes were already present in the early Archaean in which arsenite oxidation constitutes one of the oldest bioenergetic metabolisms of life on early Earth [114]. The microbial mats in Shark Bay are also subjected to considerable salinity stress [5, 21], and uptake of the osmoprotectant glycine betaine (GB) has been shown to be a primary osmoadaptive mechanism [5, 115, 116]. GB transporters facilitate GB uptake, which contributes to the osmotic stress tolerance in smooth mats [115, 116]. Genes encoding GB uptake systems were identified in the present study throughout the mat metagenomes, confirming this previous work.

Furthermore, smooth mat communities are equipped with genes encoding proteins that protect microorganisms from UV radiation (*uvrABC*) and oxidative stress (*katEG, SOD1/2, ahpC*), likely enabling survival and adaptation of the mat microbial community under extreme conditions. In particular, genes encoding catalase-peroxidase and

superoxide dismutase identified here could putatively help alleviate oxidative stress for anaerobes at the mat surface [10]. Other than the stress response genes identified, EPS is also suggested to offer protection against UV irradiation [117] and desiccation [118], and future work can delineate the abundance of EPS in these mats. In terms of protection against viral infection/lysis, the majority of MAGs encoded CRISPR-associated genes and indicate the potential immune mechanism of mat communities to defend against active viruses [71, 119]. Interestingly, two potentially novel virus defense mechanisms, BREX and DISARM [72, 120], were also detected in mat metagenomes, though their exact role in Shark Bay systems is unknown. It is also hypothesized that motility genes identified allow niche adaptation by facilitating movement of microorganism to optimal conditions, as steep chemical gradients are observed in the mats [3, 21]. As most of the energy generating metabolisms were found at the surface, it may be that the relative abundance of chemotaxis genes beneath the surface facilitates detection of— and movement to—sparse nutrients in specific mat zones.

Microbial communication in Shark Bay mats

Quorum sensing refers to intercellular communication allowing group-level coordination of phenotype responses. Previous studies have suggested quorum sensing may play an integral role in coordination of metabolisms and phenotypes such as biofilm formation in microbial mats [27, 121]. The present study represents the first evidence of the genetic potential for quorum sensing in the Shark Bay mat systems. Quorum sensing systems utilizing AHL signals are some of the most common, and, in the mat metagenomes here, genes encoding AHL synthases in two separate proteobacterial MAGs were identified (Supplementary Table 4). It is possible other synthases are present but not detected due to low conservation of these genes [122]. Both MAGs containing an AHL synthase also contained genes encoding a LuxR type response protein, with some MAGs containing multiple LuxRs potentially to bind different types of signal. There were also a number of MAGs containing LuxR proteins while not containing AHL synthases, and while this may be due to poor conservation of AHL synthases, it could also be a case of "cheaters" or organisms which can respond to AHLs without producing a signal [123]. Another common interaction with signals involves degradation enzymes that can disrupt quorum sensing via signal degradation. Genes encoding two key enzyme classes—lactonases and acylases [124, 125]— were found in multiple MAGs here (Supplementary Table Quorum sensing). The putative presence of multiple AHL degradation enzymes across diverse phyla indicates the potential importance of disrupting communication networks within
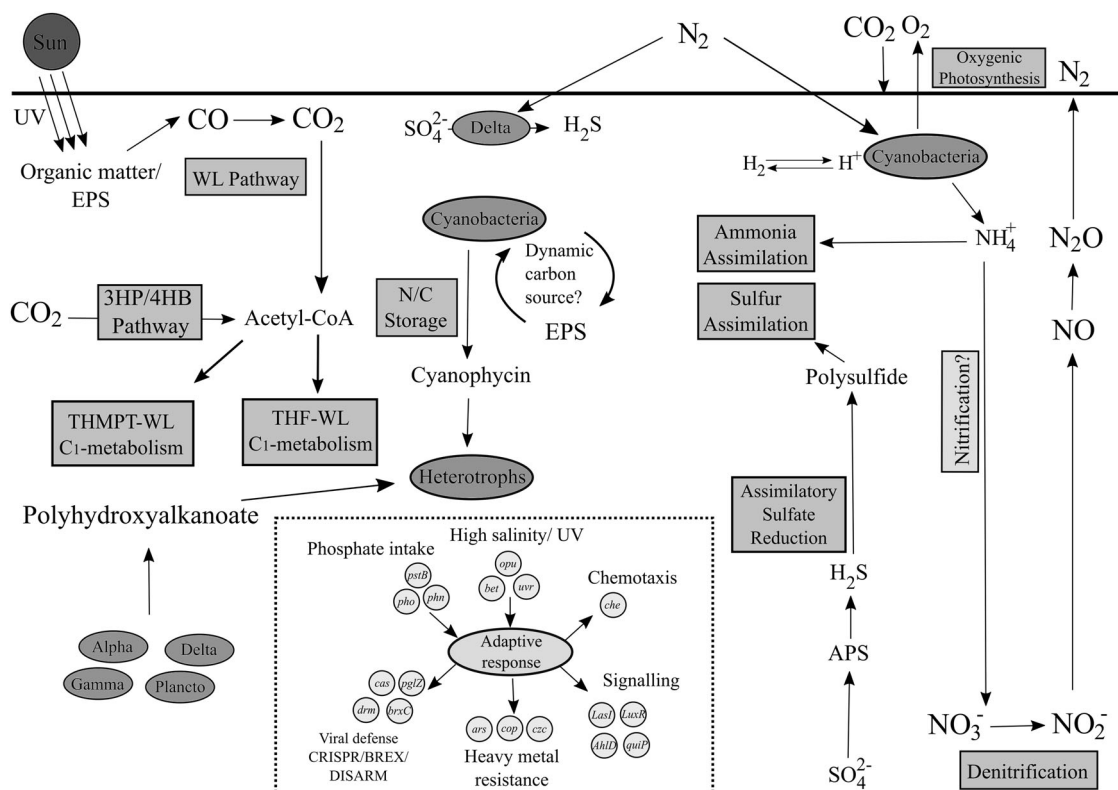
Fig. 8 Schematic illustrating major putative microbial functions and interactions in Shark Bay mats inferred from metagenomic data. Microbial processes that appear to be important in the cycling of major nutrients (carbon, nitrogen, sulfur, phosphorous) and for energy metabolism and adaptation are shown. Major pathways are indicated by rectangles, and organismal groups as ellipses. Putative major metabolisms/pathways are indicated in green, while those in yellow represent pathways where a major gene was not detected in the present study, and is potentially incomplete. Examples of genes identified here putatively associated with a given process are shown in blue circles.

Abbreviations: Alpha, Delta and Gamma refer to the corresponding Proteobacterial class; Plancto Planctomycetes, HM hydrogenotrophic methanogens, WL Wood–Ljungdahl pathway, *mvh, hdr, mtr, mtt* genes associated with methane metabolism, czc cobalt/zinc/cadmium; Cu (copper), As (arsenic) export mechanisms, *LasI/Luxr* genes associated with signal molecule synthesis/response, *AhlD/quiP* genes associated with signal molecule degradation, BREX bacteriophage exclusion systems, CRISPR clustered, regularly interspaced, short palindromic repeat systems (color figure online)

the Shark Bay systems. While the exact function of signal disruption here is not known, it is possible these quenching enzymes could target signals to interrupt other bacterial communication systems, or even modify/regulate the native signal molecule of the producing organism.

Conclusions

This study has delineated at a fine scale the metagenomic potential of Shark Bay mats, and future work will build on this platform by targeted global gene and protein expression. A schematic summarizing the major nutrient cycling pathways and putative microbial interactions inferred from the data is shown in Fig. 8. Mat metagenomes encode a variety of carbon fixation pathways, with the Wood–Ljungdahl pathway proposed as the main mechanism of fixing atmospheric carbon into organic carbon in Shark Bay mats. The relative abundance of anaerobic

pathways (e.g., sulfate reduction, Wood–Ljungdahl) even at the surface where oxygen levels were high— in conjunction with high rates of sulfate reduction measured in this zone— supports the possibility of putative surface suboxic microniches in Shark Bay mats with differing metabolisms tightly coupled (and potentially mediated) by quorum sensing [21]. The formation of anoxic/suboxic microniches has been observed in cyanobacterial sheaths [126], a scenario which could potentially be occurring in the Shark Bay mats. Mat microorganisms are likely to adapt to the extreme environment with genes encoding mechanisms to protect against hypersalinity, high UV irradiation, oxidative stress, and a potentially active viral community. High copper concentrations and extremely low phosphorus levels also likely shape the unique mat microbial community. Furthermore, smooth mat metagenomes have a high abundance of genes related to sulfur respiration, and carbon turnover potentially involving EPS, PHA, and cyanophycin scavenging. The presence of arsenic respiration genes, hydrogenases, and the

enrichment of the Wood-Ljungdahl pathway suggest smooth mats are potentially modern analogues of ancient stromatolites. To our knowledge, it is also the first time microbial mats have been shown to harbor the deep-branching Asgard archaeal groups, and it is likely other novel groups remain to be discovered in these systems that possess unique metabolic pathways.

Data availability

Metagenome sequences have been deposited in MG-RAST under project "Shark Bay smooth mat metagenome" and accession numbers 4761314.3 to 4761727.3. MAGs were deposited as project "Shark Bay smooth mat MAGs", under accession numbers 4762868.3 to 4762965.3. Details are listed in Supplementary Tables 1 and 3.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Dupraz C, Visscher PT. Microbial lithification in marine stromatolites and hypersaline mats. Trends Microbiol. 2005;13:429–38.
2. Allen MA, Goh F, Burns BP, Neilan BA. Bacterial, archaeal and eukaryotic diversity of smooth and pustular microbial mat communities in the hypersaline lagoon of Shark Bay. Geobiology. 2009;7:82–96.
3. Wong HL, Smith DL, Visscher PT, Burns BP. Niche differentiation of bacterial communities at a millimetre scale in Shark Bay microbial mats. Sci Rep. 2015;5:15607.
4. Burns BP, Goh F, Allen M, Neilan BA. Microbial diversity of extant stromatolites in the hypersaline marine environment of Shark Bay, Australia. Environ Microbiol. 2004;6:1096.
5. Ruvindy R, White RA III, Neilan BA, Burns BP. Unravelling core microbial metabolisms in the hypersaline microbial mats of Shark Bay using high-throughput metagenomics. ISME J. 2016;10:183–96.
6. Harris JK, Caporaso JG, Walker JJ, Spear JR, Gold NJ, Robertson CE. Phylogenetic stratigraphy in the Guerrero Negro hypersaline microbial mat. ISME J. 2013;7:50–60.
7. Kunin V, Raes J, Harris JK, Walker JJ, Ivanova N, von Mering C, et al. Millimetre-scale genetic gradients and community-level molecular convergence in a hypersaline microbial mat. Mol Syst Biol. 2008;4:198–203.
8. Schneider D, Arp G, Reimer A, Reitner J, Daniel R. Phylogenetic analysis of a microbialte-forming microbial mat from a hypersaline lake of the Kiritimati Atoll, Central Pacific. Plos One. 2013;8:1–14.
9. White RA III, Chan AM, Gavelis GS, Leander BS, Brady AL, Slater GF, et al. Metagenomic analysis suggests modern freshwater microbialites harbor a distinct core microbial community. Front Microbiol. 2015a;6:1531.
10. Breitbart M, Hoare A, Nitti A, Siefert J, Haynes M, Dinsdale E, et al. Metagenomic and stable isotopic analyses of modern freshwater microbialites in Cuatro Ciénegas, Mexico. Environ Microbiol. 2009;11:16–34.
11. Kozubal MA, Romine M, Jennings R, Jay ZJ, Tringe SG, Rusch DB, et al. Geoarchaeota: a new candidate phylum in the archaea from high-tempearture acidic iron mats in Yellowstone National Park. ISME J. 2013;7:622–34.
12. López-García P, Kazmierczak J, Benzerara K, Kempe S, Guyot F, Moreira D. Bacterial diversity and carbonate precipitation in the giant microbialites from the highly alkaline Lake Van, Turkey. Extremophiles. 2005;9:263–74.
13. Saghaï A, Zivanovic Y, Moreira D, Benzerara K, Bertolino P, Ragon M et al. Comparative metagenomics unveils functions and genome features of microbialite-associated communities along a depth gradient. Environ Microbiol. 2016. https://doi.org/10.1111/1462-2920.13456.
14. Thiel V, Wood JM, Olsen MT, Tank M, Klatt CG, Ward DM, et al. The dark side of the Mushroom Spring microbial mat: Life in the shadow of chlorophototrophs. Front Microbiol. 2016;7:919.
15. White RA III, Power IM, Dipple GM, Southam G, Suttle CA. Metagenomic analysis reveals that modern microbialites and polar microbial mats have similar taxonomic and functional potential. Front Microbiol. 2015b;6:966.
16. Jahnert RJ, Collins LB. Significance of subtidal microbial deposits in Shark Bay, Australia. Mar Geol. 2011;286:106–11.
17. van Kranendonk MJ, Philippot P, Lepot K, Bodrokos S, Piranjno F. Geological setting of Earth's oldest fossils in the ca. 3.5 GA Dresser formation, Pilbara Craton, Western Australia. Precambrian Res. 2008;167:93–124.
18. Bauld J. (1984). Microbial mats in marginal marine environments: Shark Bay, Western Australia, and Spencer Gulf, South Australia. In Cohen Y, Castenholz RW, Halvorson HO (Eds) *Microbial mats: Stromatolites*. Alan R Liss Inc, New York, pp. 39-58.
19. Palmisano AC, Summons RE, Cronin SE, Des Marais DJ. Lipophilic pigments from cyanobacterial (Blue-green algal) and diatom mats in Hamelin Pool, Shark Bay, Western Australia. J Phycol. 1989;25:633–61.
20. Visscher PT, Dupraz C, Braissant O, Gallagher KL, Glunk C, Casillas L, *et al*. Biogeochemistry of carbon cycling in hypersaline mats: linking the present to the past through biosignatures, In *Microbioal Mats*. Springer: Netherlands, 2010, pp 443-468.
21. Wong HL, Visscher PT, White RA, Smith D-L, Patterson M, Burns BP. Dynamics of archaea at fine spatial scales in Shark Bay mat microbiomes. Sci Rep. 2017;8:46160.
22. Wong HL, Ahmed-Cox A, Burns BP. Molecular ecology of hypersaline microbial mats: Current insights and new directions. Microorganisms. 2016;4:6.
23. Pagès A, Welsh DT, Teasdale PR, Grice K, Vacher M, Bennett WW, et al. Diel fluctuations in solute distributions and biogeochemical cycling in a hypersaline microbial mat from Shark Bay, WA. Mar Chem. 2014;167:102–12.
24. Grünke S, Felden J, Lichtschlag A, Girnth AC, de Beer D, Wenzhöfer F, et al. Niche differentiation among mat-forming, sulphide-oxidizing bacteria at cold seeps of the Nile Deep Sea Fan (Eastern Mediterranean Sea). Geobiology. 2011;9:330–48.
25. Macalady JL, Dattagupta S, Schaperdoth I, Jones DS, Druschel GK, Eastman D. Niche differentiation among sulfur-oxidizing bacterial populations in cave waters. ISME J. 2008;2:590–601.
26. Visscher PT, Stolz JF. Microbial mats as bioreactors: Populations, processes, and products. Palaeogeo Palaeoclim Palaeoecol. 2005;219:87–100.

27. Decho AW, Norman RS, Visscher PT. Quorum sensing in natural environments: emerging views from microbial mats. Trends Microbiol. 2010;18:73–80.

28. Gallagher KL, Kading TJ, Braissant O, Dupraz C, Visscher PT. Inside the alkalinity engine: The role of electron donors in the organomineralization potential of sulfate-reducing bacteria. Geobiology. 2012;10:518–30.

29. Sugimura Y, Suzuki Y. A high-temperature catalytic oxidation method for the determination of non-volatile dissolved organic carbon in seawater by direct injection of a liquid sample. Mar Chem. 1988;24:105–31.

30. Johnson KS, Petty RL. Determination of nitrate and nitrite in seawater by flow injection analysis. Limnol Ocean. 1983;28:1260–6.

31. White RA III, Brown J, Colby S, Overall CC, Lee JY, Zucker J, et al. ATLAS (Automatic Tool for Local Assembly Structures) – a comprehensive infrastructure for assembly, annotation, and genomic binning of metagenomic and metatranscriptomic data. PeerJ. 2017;5:e2843v1.

32. Magoc T, Salzberg S. Fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011; 27:2957–63.

33. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9:357–9.

34. Bolger A, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

35. Li D, et al. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31:1674–6.

36. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prodaryotic gene recognition and translation initiation site identification. BMC Bioinforma. 2010;11:119.

37. Buchfink B, et al. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015;12:59–60.

38. O'Leary NA, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44:D733–45.

39. Huerta-Cepas J, et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. 2016;44: D286–93.

40. Yin Y, et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012;40:445–51.

41. Bairoch A. The ENZYME database in 2000. Nucleic Acids Res. 2000;28:304–5.

42. Tatusov RL, et al. The COG database: an updated version includes eukaryotes. BMC Bioinforma. 2003;4:41.

43. Finn RD, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44:279–85.

44. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol. 2016;428:726–31.

45. Hanson NW, et al. LCA*: an entropy-based measure for taxonomic assignment within assembled metagenomes. Bioinformatics. 2016;32:3535–42.

46. Liao Y, et al. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–30.

47. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.

48. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol. 2017;35:725–31.

49. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–5.

50. Lowe TM, Chan PP. tRNAscan-SE On-line: Search and contextual analysis of transfer RNA genes. Nucleic Acids Res. 2016;44:W54–57.

51. Richter M, Rosselló-Móra R, Oliver Glöcker F, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 2015;32:929–31.

52. Darling AE, Jospin G, Lowe E, Matsen IV, Bik HM, Eisen JA. PhyloSift: phylogenetic analysis of genomes and metagenomes. Peer J. 2014;2:e243.

53. Sousa FL, Neukirchen S, Allen JF, Lane N, Martin WF. Lokiarchaeon is hydrogen dependent. Nat Microbiol. 2016;1:16034. Article no

54. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30:1236–40.

55. Csárdi G, Nepusz T. The Igraph software package for complex network research. Inter J Complex Syst. 2006;1695:1–9.

56. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, et al. A travel guide to Cytoscape plugins. Nat Methods. 2012;9:1069–76.

57. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30:3059–66.

58. Criscuolo A, Gribaldo S. BMGE (Block mapping and gathering with entrophy): a new software for selection of phylogeneticv informative regions from multiple sequence alignments. BMC Evol Biol. 2010;10:210.

59. Hug LA, Castelle CJ, Wrighton KC, Thomas BC, Sharon I, Frischkorn KR, et al. Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. Microbiome. 2013;1:22.

60. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the ultrafast bootstrap approximation. Mol Biol Evol. 2017;35:518–22.

61. Letunic I, Bork P. Interactive tree of life (iTOL) V3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44:W242–5.

62. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature. 2017;541:353–8.

63. Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, et al. Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. Curr Biol. 2015;25:690–701.

64. Bernstein HC, Brislawn C, Renslow RS et al. Trade-offs between microbiome diversity and productivity in a stratified microbial mat. ISME J. 2016. https://doi.org/10.1038/ismej. 2016.133.

65. Sato T, Atomi H, Imanaka T. Archaeal Type III RuBisCOs Function in a Pathway for AMP Metabolism. Science. 2007;315:1003–6.

66. Burow LC, Woebken D, Bebout BM, McMurdie PJ, Singer SW, Pett-Ridge J, et al. Hydrogen production in photosynthetic microbial mats in the Elkhorn Slough estuary, Monterey Bay. ISME J. 2012;6:863–74.

67. Baumgartner LK, Reid RP, Dupraz C, Decho AW, Buckley DH, Spear JR, et al. Sulfate-reducing bacteria in microbial mats: Changing paradigms, new discoveries. Sediment Geol. 2006;185:131–45.

68. Kurth D, Amadio A, Ordoñez OF, Albarracin VH, Gärtner W, Farias ME. Arsenic metabolism in high altitude modern stromatolites revealed by metagenomic analysis. Sci Rep. 2017;7:1024.

69. Leuko S, Neilan BA, Burns BP, Walter MR, Rothschild LJ. Molecular assessment of UVC radiation-induced damage repair in the stromatolitic halophilic archaeon, Halococcus hamelinensis. J Photoch Photobio 2011;B102:140–5.

70. Khan SR, Kuzminov A. Replication forks stalled at ultraviolet lesions are rescued via RecA and RuvABC protein-catalyzed disintegration in Escherichia coli. J Biol Chem. 2012;287: 6250–65.

71. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science. 2010;327:167–70.

72. Ofir G, Melamed S, Sberro H, Mukamel Z, Silverman S, Yaakov G, et al. DISARM is a widespread bacterial defence system with broad anti-phage activities. Nat Microbiol. 2017;3:90–8.

73. Sørensen K, Canfield D, Teske A, Oren A. Community Composition of a Hypersaline Endoevaporitic Microbial Mat. Appl Environ Microbiol. 2005;71:7352–65.

74. Burns BP, Anitori R, Butterworth P, Henneberger R, Goh F, Allen M, et al. Modern analogues and the early history of microbial life. Precambrian Res. 2009;173:10–8.

75. Visscher PT, Prin RA, van Gemerden H. Rates of sulfate reduction and thiosulfate consumption in a marine microbial mat. FEMS Microbiol Lett. 1992;86:283–8.

76. Canfield D, Des Marais D. Aerobic sulfate reduction in microbial mats. Science. 1991;251:1471–3.

77. Jørgensen B. Sulfate reduction and thiosulfate transformations in a cyanobacterial mat during a diel oxygen cycle. FEMS Microbiol Ecol. 1994;13:303–12.

78. Say RF, Fuchs G. Fructose 1, 6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. Nature. 2010;464:1077.

79. Casamayor EO, Triadó-Margarit X, Castañeda C. Microbial biodiversity in saline shallow lakes of the Monegros Desert, Spain. FEMS Microbiol Ecol. 2013;85:503–18.

80. Saghaï A, Gutiérrez-Preciado A, Deschamps P, Moreira D, Bertolino P, Ragon M, et al. Unveiling microbial interactions in stratified mat communities from a warm saline shallow pond. Environ Microbiol. 2017;19:2405–21.

81. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, et al. Insights into the phylogeny and coding potential of microbial dark matter. Nature. 2013;499:431–7.

82. Vavourakis CD, Ghai R, Rodriguez-Valera F, Sorokin DY, Tringe SG, Hugenholtz P, et al. Metagenomic insights into the uncultured diversity and physiology of microbes in four hypersaline soda lake brines. Front Microbiol. 2016;7:211.

83. King GM. Carbon monoxide as a metabolic energy source for extremely halophilic microbes: implications for microbial activity in Mars regolith. PNAS. 2015;112:4465–70.

84. Weiss M, Sousa F, Mrnjavac N, Neukirchen S, Nelson-Sathi S, et al. The physiology and habitat of the last universal common ancestor. Nat Microbiol. 2016;1:16116.

85. Fuchs G. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? Annu Rev Microbiol. 2011;65:631–58.

86. Braissant O, Decho AW, Przekop KM, Glunk C, Dupraz C, Visscher PT. Characteristics and turnover of exopolymeric substances (EPS) in a hypersaline microbial mat. FEMS Microbiol Ecol. 2009;67:293–307.

87. Kadnikov VV, Mardanov AV, Podosokorskaya OA, Gavrilov SN, Kublanov IV, Beletsky AV, et al. Genomic analysis of Melioribacter roseus, facultatively anaerobic organotrophic bacterium representing a novel deep lineage within Bacteroidetes/Chlorobi group. PLoS One. 2013;8:e53047.

88. Marvasi M, Visscher P, Casillas Martinez L. Exopolymeric substances from Bacillus subtilis: polymers and genes encoding their synthesis. FEMS Microbiol Lett. 2010;313:1–9.

89. Stuart RK, Mayali X, Lee JZ, Everroad RC, Bebout BM, et al. Cyanobacteria reuse of extracellular organic carbon in microbial mats. ISME J. 2016;10:1240–51.

90. Bateson MM, Castenholz R. Photoexcretion and fate of glycolate in a hotspring cyanobacterial mat. Appl Environ Microbiol. 1988;54:1738–43.

91. Füser G, Steinbüchel A. Analysis of genome sequences for genes of cyanophycin metabolism: Identifiying putative cyanophycin metabolizing prokaryotes. Macromol Biosci. 2007;7:278–96.

92. Liang B, Wu TD, Sun HJ, Vali H, Guerquin-Kern JL, Wang CH, et al. Cyanophycin mediates the accumulation and storage of fixed carbon in non-heterocystous filamentous Cyanobacteria from coniform mats. Plos One. 2014;9:e88142.

93. Villanueva L, del Campo J, Guerrero R. Diversity and physiology of polyhydroxyalkanoate-producing and-degrading strains in microbial mats. FEMS Microbiol Ecol. 2010;74:42–54.

94. Berg IA. Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. Appl Environ Microbiol. 2011;77:1925–36.

95. Sousa FL, and Martin WF. Biochemical fossils of the ancient transition from geoenergetics to bioenergetics in prokaryotic one carbon compound metabolism. BBA-Bioenergetics 2014;1837:964–81.

96. Liu Y, Zhou Z, Pan J, Baker BJ, Gu JK, Li M. Comparative genomic inference suggests mixotrophic lifestyle for Thorarchaeota. ISME 2018;J12:1021.

97. Wasmund K, Cooper M, Schreiber L, Lloyd KG, Baker BJ, Petersen DG, et al. Single-cell genome and group-specific dsrAB sequencing implicate marine members of the class Dehalococcoidia (Phylum Chloroflexi) in sulfur cycling. MBio. 2016;7: e00266–16.

98. Anantharaman K, Hausmann B, Jungbluth SP, Kantor RS, Lavy A, Warren LA, Rappé MS, Pester M, Loy A, Thomas BC, Banfield JF. Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. ISME J. 2018. https://doi.org/10.1038/s41396-018-0078-0.

99. Baker BJ, Lazar CS, Teske AP, Dick GJ. Genomic resolution of linkages in carbon, nitrogen, and sulfur among widespread estuary sediment bacteria. Microbiome. 2015;3:14.

100. Visscher PT, and H. van Gemerden. Sulfur cycling in laminated marine ecosystems. In: Biogeochemistry of Global Change: Radiatively Active Trace Gases. R.S. Oremland (ed). Chapman and Hall, New York, 1993, pp. 672-69

101. Sforna MC, Daye M, Philippot P, Somogyi A, Van Zuilen MA, Medjoubi K, et al. Patterns of metal distribution in hypersaline microbialites during early diagenesis: Implications for the fossil record. Geobiology. 2017. https://doi.org/10.1111/gbi.12218.

102. Braissant O, Decho A, Dupraz C, Glunk C, Przekop K, Visscher P. Exopolymeric substances of sulfate-reducing bacteria: Interactions with calcium at alkaline pH and implication for formation of carbonate minerals. Geobiology. 2007;5:401–11.

103. Mobberley JM, Lindemann SR, Bernstein HC, Moran JJ, Renslow RS, Babuta J, et al. Organismal and spatial partitioning of energy and macronutrient transformations within a hypersaline mat. FEMS Microbiol Ecol. 2017;93:10.1093.

104. Atkinson M. Low phosphorus sediments in a hypersaline marine bay. Estuar Coast Shelf S. 1987;24:335–47.

105. Kornberg A. Inorganic polyphosphate: toward making a forgotten polymer unforgettable. J Bacteriol. 1995;177:491–6.

106. Oren A. DNA as genetic material and as a nutrient in halophilic archaea. Front Microbiol. 2014;5:1–2.

107. Greening C, Biswas A, Carere CR, Jackson CJ, Taylor MC, Stott MB, et al. Genomic and metagenomic surveys of hydrogenase

distribution indicate H$_2$ is a widely utilised energy source for microbial growth and survival. ISME J. 2016;10:761–77.

108. Tamagnini P, Leitão E, Oliveira P, Ferreira D, Pinto F, Harris DJ, et al. Cyanobacterial hydrogenases: diversity, regulation and applications. FEMS Microbiol Rev. 2007;31:692–720.

109. Skyring GW, Lynch RM, Smith GD. Quantitative relationships between carbon, hydrogen and sulfur metabolism in cyanobacterial mats. In Cohen Y, Rosenberg E. (ed). *Microbial mats: Physiological ecology of benthic microbial communities* American Society for Microbiology, Washington, DC, 1989, pp 170-179.

110. Boyd ES, Schut GJ, Adams MW, Peters JW. Hydrogen metabolism and the evolution of biological respiration. Microbe. 2014;9:361–7.

111. Rensing C, Grass G. *Escherichia coli* mechanisms of copper homeostasis in a changing environment. FEMS Microbiol Rev. 2003;27:197–213.

112. Oremland R, Saltikov C, Wolfe-Simon F, Stolz J. Arsenic in the Evolution of Earth and Extraterrestrial Ecosystems. Geomicrobiol J. 2009;26:522–36.

113. Sforna MC, Philippot P, Somogyi A, van Zuilen MA, Medjoubi K, Schoepp-Cothenet B, et al. Evidence for arsenic metabolism and cycling by microorgansisms 2.7 billion years ago. Nat Geosci. 2014;7:811–5.

114. van Lis R, Nitschke W, Duval S, Schoepp-Cothenet B. Arsenics as bioenergetic substrates. Biochim Biophys Acta. 2013;1827:176–88.

115. Goh F, Barrow KD, Burns BP, Neilan BA. Identification and regulation of novel compatible solutes from hypersaline stromatolite-associated cyanobacteria. Arch Microbiol. 2010;192:1031–8.

116. Goh F, Jeon YJ, Barrow K, Neilan BA, Burns BP. Osmoadaptive strategies of the archaeon *Halococcus hamelinensis* isolated from a hypersaline stromatolie environment. Astrobiology. 2011;11:529–36.

117. Garcia-Pichel F, Castenholz RW. Occurrence of UV-absorbing, mycosporine-like compounds among cyanobacterial isolates and an estimate of their screening capacity. Appl Environ Microbiol. 1993;59:163–9.

118. Wright D, Smith S, Joardar V, Scherer S, Jervis J, Warren A, et al. UV Irradiation and Desiccation Modulate the Three-dimensional Extracellular Matrix of Nostoc commune (Cyanobacteria). J Biol Chem. 2005;280:40271–81.

119. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315:1709–12.

120. Goldfarb T, Sberro H, Weinstock E, Cohen O, Doron S, Charpak-Amikam Y, Shaked Afik, et al. BREX is a novel phage resistance system widespread in microbial genomes. EMBO J. 2015;34:169–83.

121. Decho AW, Visscher PT, Ferry J, Kawaguchi T, He L, Przekop KM, et al. Autoinducers extracted from microbial mats reveal a surprising diversity of N-acylhomoserine lactones (AHLs) and abundance changes that may relate to diel pH. Environ Microbiol. 2009;11:409–20.

122. Watson WT, Minogue TD, Val DL, von Bodman SB, Churchill MEA. Structural basis and specificity of acyl-homoserine lactone signal production in bacterial quorum sensing. Mol Cell. 2002;9:685–94.

123. Decho AW, Frey RL, Ferry JL. Chemical challenges to bacterial AHL signalling in the environment. Chem Rev. 2011;111:86–99.

124. Dong YH, Wang LH, Xu JL, Zhang HB, Zhang LH. Quenching quorum-sensing-dependent bacterial infection by an N-acyl homoserine lactonase. Nature. 2001;411:813–7.

125. Romero M, Diggle SP, Heeb S, Cámara M, Otero A. Quorum quenching activity in *Anabaena sp*. PCC7120: Identification of AiiC, a novel AHL-acylase. FEMS Microbiol Lett. 2008;280:73–80.

126. Paerl HW, Pinckney JL, Kucera SA. Clarification of the structural and functional roles of heterocysts and anoxic microzones in the control of pelagic nitrogen fixation. Limnol Oceanogr. 1995;40:634–8.