Deciphering the community structure and the functional potential of a hypersaline marsh microbial mat community

Christina Pavloudi[1,2,*], Haris Zafeiropoulos[1,3,*]

[1] Institute of Marine Biology, Biotechnology, Aquaculture (IMBBC), Hellenic Centre for Marine Research (HCMR), Heraklion, Crete, Greece
[2] Department of Biological Sciences, The George Washington University, District of Columbia, USA
[3] Department of Biology, University of Crete, Voutes University Campus, Heraklion, Greece

Corresponding author: Christina Pavloudi (cpavloud@hcmr.gr, cpavloudi@gwu.edu)
*: shared co-authorship

Abstract

Microbial mats are vertically stratified communities of microorganisms characterised by pronounced physiochemical gradients allowing for high species diversity and a wide range of metabolic capabilities. High Throughput Sequencing has the potential to reveal the biodiversity and function of such ecosystems in the cycling of elements.

The present study combines 16S rRNA amplicon sequencing and shotgun metagenomics on a hypersaline marsh in Tristomo bay (Karpathos, Greece). Samples were collected in July 2018 and November 2019 from microbial mats, deeper sediment, aggregates observed in the water overlying the sediment, as well as sediment samples with no apparent layering.

Metagenomic samples' co-assembly and binning revealed 250 bacterial and 39 archaeal metagenome-assembled genomes, with completeness estimates higher than 70% and contamination less than 5%. All MAGs had KEGG Orthology terms related to osmoadaptation, with the "salt in" strategy ones being prominent. Halobacteria and Bacteroidetes were the most abundant taxa in the mats. Photosynthesis was most likely performed by purple sulphur and non-sulphur bacteria. All samples had the capacity for sulphate reduction, dissimilatory arsenic reduction and conversion of pyruvate to oxaloacetate. Overall, both sequencing methodologies resulted in similar taxonomic compositions and revealed that the formation of the microbial mat in this marsh exhibits seasonal variation.

Keywords: shotgun metagenomics, metagenome-assembled genomes (MAGs), 16S rRNA, extreme environments, sediment, salt crust

One-Sentence summary: This study investigates the taxa and the metabolic processes that potentially occur in a hypersaline marsh microbial mat showing the layered structure of microbial communities and the role of key-processes that support life in this extreme environment.

Introduction

Microbial mats are vertically stratified communities of functional groups of microorganisms embedded in an organic matrix, which may also contain minerals such as silicates and carbonates (Stal 2012; Bolhuis, Cretoiu and Stal 2014; Prieto-Barajas *et al.* 2018). They grow on a solid substrate (e.g., sand) and the vast majority of microbial mats utilise inorganic carbon as carbon source, hence they are autotrophic (Bolhuis, Cretoiu and Stal 2014). Microbial mats are highly productive ecosystems (Villagrasa *et al.* 2019) characterised by pronounced physiochemical gradients which allow for the presence of high species diversity, encompassing a wide range of metabolic capabilities; thus, mats are ideal models to study a whole ecosystem (Al-Thani *et al.* 2014) and are considered as natural laboratories (Villanueva *et al.* 2007). These physicochemical gradients provide microenvironments for various microbial functional groups, which exhibit a certain physiology with which they fulfil a specific function (van Gemerden 1993).

Microbial mats comprise an extensive diversity of microorganisms belonging to different species which are embedded in a matrix of extracellular polymers (EPS) and exchange signals and nutrients, thus enabling a flow of resources and energy for the survival of the overall community (Ruvindy *et al.* 2016; Prieto-Barajas, Valencia-Cantero and Santoyo 2018). The role of microbial mats has been vital throughout Earth's history since they produced and released reduced gases, e.g. O_2, H_2, CH_4, in the early earth's atmosphere (Hoehler, Bebout and Des Marais 2001) and they are regarded as modern analogs of early ecosystems (Krumbein, Paterson and Zavarzin 2003). In addition, they are considered to constitute the first ecosystems, along with stromatolites (Prieto-Barajas, Valencia-Cantero and Santoyo 2018), and probably are the oldest (van Gemerden 1993) and simplest of the self-organised structures that may have first appeared on earth (Guerrero, Piqueras and Berlanga 2002).

Regardless of the vertical structure, marine microbial mats are comprised of four main functional groups: i) oxygenic phototrophs (CYN) (primarily Cyanobacteria), ii) aerobic heterotrophic bacteria (HET), iii) sulphate-reducing bacteria (SRB) and iv) sulphide-oxidising bacteria (SOB) (Visscher and Stolz 2005). It is suggested that archaeal abundance and diversity can be limited in microbial mats (Hugenholtz *et al.* 1998; Sievert *et al.* 2000; Robertson *et al.* 2009; López-López *et al.* 2013), with halophilic and methanogenic taxa to be the most predominant among them (Bolhuis, Cretoiu and Stal 2014). However, archaeal taxa such as methanogens, DPANN and Asgard archaea might play a key role in some cases (Potter, Bebout and Kelley 2009; Kelley *et al.* 2012; Wong *et al.* 2020) while novel archaeal taxa are still discovered in such environments (Kozubal *et al.* 2013; Wong *et al.* 2020). Although microbial mats can have this very pronounced layered redox stratification during the day, they can become anoxic during nighttime (Des Marais 1995). Microbial mats function as a consortium where coupling of biogeochemical cycles and processes occurs (Paerl, Pinckney and Steppe 2000), allowing the products of the metabolism of one group to be available and used by another (Prieto-Barajas, Valencia-Cantero and Santoyo 2018). In addition, the metabolic rates of mat microorganisms are so high that the community production per unit mass competes with that of rainforests (Jørgensen 2001; Krumbein, Paterson and Zavarzin 2003).

Microbial mats can be distinguished in six categories (Bolhuis, Cretoiu and Stal 2014; Prieto-Barajas, Valencia-Cantero and Santoyo 2018): i) intertidal or coastal, ii) hypersaline, iii) hot spring, iv) mats in oligotrophic environments, v) psychrophile and vi) acid microbial mats. Intertidal mats are formed on beaches with low slopes and fine sandy sediments (Stal 2012) and they experience strong salinity fluctuations, large temperature changes (Bolhuis, Cretoiu and Stal 2014) and irregular floods (Prieto-Barajas, Valencia-Cantero and Santoyo 2018). On the other hand, hypersaline microbial mats are found in natural occurring salt lakes and man-made salterns (Bolhuis, Cretoiu and Stal 2014) and are exposed to salinities up to the crystallisation point of halite (Jørgensen 1994), high temperatures and high solar radiation (Bolhuis, Cretoiu and Stal 2014).

High Throughput Sequencing (HTS) technologies and methods have been widely used to study real-world microbial communities. They have enabled the study of ecosystems with no prior knowledge of the resident species, uncovering unknown and uncultivated strains (Hedlund *et al.* 2014). Metabarcoding studies are common, well-established and less computationally demanding than shotgun metagenomics (Bell *et al.* 2021). However, taxonomic biases may arise from differential

efficiency of PCR primer pairing in different species (van der Loos and Nijland 2021) while the short barcoding sequences may limit the resolution. On the other hand, shotgun metagenomic sequencing enables profiling up to the level of strains by obtaining information from random sampling of virtually all genomic regions (Clooney *et al.* 2016; Segata 2018; Dávila-Ramos *et al.* 2019). Therefore, microbiome metabolic functions and entire biochemical pathways that occur in a sample can be explored after processing the metagenomic information (Sharpton 2014). Over the recent years, HTS approaches have been used to study the taxonomic and the functional profiles of the microbial communities present in microbial mats (Chen *et al.* 2020; Wong *et al.* 2020; Kindler *et al.* 2022). Several novel high-level taxa have been discovered, e.g. Zixibacterial order GN15 (Wong *et al.* 2020), and a better understanding on both their adaptive responses in such environments has been established. On top of that, further insight on the mechanisms governing such assemblages has been gained, e.g. the role of photoheterotrophy (Kindler *et al.* 2022).

The present study was conducted in the Tristomo marsh in the island of Karpathos (Aegean Sea, Greece) (Figure 1A). The study area is included in the Natura 2000 network (site GR4210003) and also in the Greek catalogue of small island wetlands (Y421KAR001) (Figure 1B). It is a seasonal brackish water marsh formed at the edge of a small plain where a seasonal stream ends, characterised as an intertidal marsh (type H) according to the Ramsar convention. Freshwater enters the marsh from the precipitation and drainage basin, while the wetland interacts mainly with the sea through the waves but also underground (WWF Greece 2022). On the coastal front, the cobbled beach is full of litter accumulated by the waves. Due to the close proximity of the marsh with the sea, it occasionally receives saline water, therefore could be characterised as intertidal; however, evaporation creates hypersaline conditions and crystallised salt forms an upper layer above the actual microbial mat, something that is observed in hypersaline mats (Figure 1C, 1D).

The aim of the present study was to identify the microbial communities in samples from the hypersaline Tristomo marsh, as well as their functional and metabolic capabilities. Our main goals were to test whether the formation of the microbial mat is characterised by seasonality, and which are the mechanisms of osmoadaptation of the microbial communities in this hypersaline environment, as a response to the increased levels of salinity.

Methods

Sample collection

Samples were collected in July 2018 and November 2019 from the Tristomo marsh (Figure 1). Details on the sample collection are given in Table 1. Sediment samples were collected using cylindrical sampling corers (internal sampling surface 15.90 square centimetres) (Figure 1E). In the cases where microbial mat layers were clearly observed (July 2018), the top layers were collected separately from the bottom layer. In addition, microbial aggregates observed floating in the marsh were also collected. In the cases where microbial mat layers were not clearly formed (November 2019), there was no slicing during sample collection. In November, samples were collected from

three different locations in the marsh, distinguished by the colour of the sediment's upper layer (black, purple and orange).

Samples were placed in 50 ml falcon tubes (Sarstedt, Nümbrecht, Germany) and were stored at −20 °C, until further processing in the laboratory. Upon return to the laboratory, they were used for molecular analysis, i.e., DNA extractions, as well as for the measurement of the Particulate Organic Carbon (POC) and chloroplast pigments concentration (chlorophyll-a, phaeopigments and chloroplastic pigment equivalents (CPE)). For the latter, the samples were processed at the Environmental Chemistry Lab of the IMBBC (HCMR), based on standard techniques (Yentsch and Menzel 1963; Hedges and Stern 1984). Water temperature and dissolved oxygen concentration were measured in the water overlaying the sediments by means of a portable multi-parameter (WTW Multi 3420 SET G). Salinity was also measured with the portable multi-parameter but after dilution of samples with dH_2O since the initial measurement was out of limits (TetraCon® 925 sensor range: 0 - 70). Sampling was conducted under authorization from the relevant licensing authority (Directorate General for the Protection and Development of Forests and the Rural Environment, Directorate of Forest Management) of the Ministry of Environment and Energy. Additional authorization was also provided from the Management Agency of Dodecanese Protected Areas.

DNA extraction, PCR amplification and 16S rRNA sequencing

DNA was extracted as in (Henckel, Friedrich and Conrad 1999) and (Lueders, Manefield and Friedrich 2004). Approximately 0.7 g of wet sediment were added to a 2-ml screw-cap vial, prefilled with ~0.7 g of 0.1 mm (diameter) zirconia/silica beads (11079101z, BioSpec, USA). The vials were filled with 750 µl of 120 mM NaPO4 buffer (pH 8) and 250 µl TNS solution (500 mM Tris-HCl pH 8, 100 mM NaCl, 10 % SDS (w/v)) and placed horizontally in a vortex for 10 minutes at maximum speed. Immediately after that the vials were centrifuged for 10 min at 20,800 rcf and 4 °C and the supernatants were transferred to new 2-ml vials. One volume of phenol/chloroform/isoamylalcohol (P/C/I; 25:24:1; pH 8; Carl-Roth, Karlsruhe, Germany) was added to the aqueous supernatant. Vials were vigorously shaken for 20 s and centrifuged for 5 min at 20,800 rcf and 4 °C. Supernatants were transferred to new 2-ml vials, and one volume of chloroform/isoamylalcohol (C/I; 24:1; Carl-Roth) was added. Vials were again vigorously shaken for 20 s and then centrifuged for 5 min at 20,800 rcf and 4 °C. Supernatants were transferred to new 2-ml vials and C/I extraction was repeated to successfully remove all phenol remnants. Supernatants were transferred to new 2-ml vials and 1.5 ml of polyethylene glycol (30 % (w/v) polyethylene glycol 6000 in 1.6 M NaCl) was added to precipitate nucleic acids and the vials were centrifuged for 90 min at 20,800 rcf and 4 °C. Supernatants were discarded and the pellets were washed with 1 ml 70% ethanol (4 °C) and centrifuged for 30 min. Supernatants were again discarded, pellets were left for air drying (~5 min) to remove leftover ethanol and resuspended with 50 µl 10mM Tris.

PCR amplification, library preparation and MiSeq sequencing was performed as in (Pavloudi *et al.* 2017). Briefly, PCR amplification was performed following the two-Step PCR approach and targeting the V3–V4 region of the 16S rRNA gene with the primers 341F (5´-CCTACGGGNGGCWGCAG-3´) (Herlemann *et al.* 2011; Klindworth *et al.* 2013) and 805RB (5´-GACTACNVGGGTATCTAATCC-3´) revised for detection of SAR11 bacterioplankton (Apprill *et al.* 2015; Pavloudi *et al.* 2017).

The PCR negative control sample (blank) was also sequenced, so that possible contamination during the library preparation could be assessed. The raw sequence reads were processed with PEMA (version 2.1.4) (Zafeiropoulos *et al.* 2020) using VSEARCH for the creation of OTUs. Taxonomic assignment was performed with the SILVA database (version 132) (Quast *et al.* 2013). The detailed parameters of the PEMA processing are given in Supplementary File 1. The phyloseq (version 1.36) (McMurdie and Holmes 2013), vegan (version 2.5.7) (Oksanen *et al.* 2020) and ggplot2 (version 3.3.5) (Wickham, Chang and Wickham 2016) packages were used in R (version 4.1.1) (R Core Team 2021) for the creation of barcharts, for the nMDS and PERMANOVA, variation partitioning analysis, db-RDA and mantel test. The scripts of Steinberger (2020) were used for the simper and the Kruskal-Wallis tests.

Shotgun metagenomics sequencing

Six samples were selected for shotgun sequencing (Elos01, Elos02, Elos03, Elos07, Elos10 and Elos12), based on the results of the 16S rRNA amplicon sequencing, their potential for the identification of novel unknown lineages and the representation of the majority of collected sample types. Sample preparation was performed using the Nextera™ DNA Flex Tagmentation and sequencing was done at two lanes of a HiSeq 4000 (2x150bp) at the Norwegian Sequencing Centre (NSC). All the raw sequence files of this study (both 16S rRNA and shotgun metagenomes) were submitted to the European Nucleotide Archive (ENA) (Cummins *et al.* 2022) with the study accession number PRJEB46254 (available at http://www.ebi.ac.uk/ena/data/view/PRJEB46254). In addition, the analysis of both the amplicon as well as of the shotgun metagenomic data was performed by MGnify's pipeline version 5.0 and is available at https://www.ebi.ac.uk/metagenomics/studies/MGYS00006059 and https://www.ebi.ac.uk/metagenomics/studies/MGYS00006060 respectively (Mitchell *et al.* 2020).

Assembly and binning

Since the samples were sequenced in two lanes, the fastq files of each sample were concatenated before proceeding with the analyses. Metagenome raw reads were processed with the MetaWRAP workflow (version 1.3.2) (Uritskiy, DiRuggiero and Taylor 2018). Reads were trimmed and qualified using Trim Galore (version 0.5.0) (Krueger 2022), which is a wrapper around Cutadapt (version 1.18) (Martin 2011) and FastQC. The clean reads were concatenated, and their co-assembly was implemented through the corresponding metaWRAP module, using MEGAHIT v1.1.3; by the term "co-assembly", we denote the analysis of the community across the samples, meaning the reads from all metagenomic samples were handled as a single entity. The quality of the co-assembly was evaluated using QUAST (Gurevich *et al.* 2013). Binning was then performed using the clean reads and the co-assembly. The metaWRAP module for binning was performed using MetaBAT 2 (version 2.12.1) (Kang *et al.* 2019: 2) and MaxBin 2 (version 2.2.6) (Wu, Simmons and Singer 2016). CheckM (version 1.0.12) (Parks *et al.* 2015) was used by the metaWRAP module to assess the quality of the bins produced by MetaBAT 2 and MaxBin 2. Bins were then consolidated and refined using Binning_refiner (Song and Thomas 2017) as wrapped in the Bin_refinement module of metaWRAP.

The Bin_refinement module was invoked with the default values for minimum completion (70%) and maximum contamination (5%). The consolidated bins set was further improved using the reassemble_bins module of metaWRAP. To this end, bwa (version 0.7.17-r1188) (Li and Durbin 2009), spades (version v3.13.0) (Nurk *et al.* 2017) and CheckM were used. To estimate bins' abundances in each sample (in genome copies per million reads), the corresponding metaWRAP module was performed invoking Salmon (version 0.13.1) (Patro *et al.* 2017). The refined bin-set was also used for the blobology module of metaWRAP; taxonomic annotation of the co-assembled contigs was performed using megaBLAST and the nt database of NCBI.

The co-assembled contigs and the refined bins set were then used as input to Anvi'o (version 7.1) (Eren *et al.* 2015). Bowtie 2 (version 2.3.5) (Langmead and Salzberg 2012) was used to build BAM files and mapping and Prodigal (version 2.6.3) (Hyatt *et al.* 2010) for gene prediction. BAM files were also made from the clean reads of each sample. A contigs database was built (using the anvi-gen-contigs-database program) after converting the contigs name as Anvi'o suggests (see contigs-per-bin.sh script) and it was decorated with hits from HMM models (anvi-run-hmms). An anvi profile was then built for each of the samples' bam file (anvi-profile) and they were merged (anvi-merge) into a single profile. The refined bins along with their corresponding renamed contigs were imported as a *collection* in the merged profile database (anvi-import-collection). At this point, a first Anvi'o summary was recovered (anvi-summarize). Bins with a redundancy >10% were manually refined and a second summary of the bins set was created; "redundancy" here is referring to the anvi'o measure of "how many copies of each single-copy core gene is found within a genome". Afterwards, a simper analysis was performed to identify MAGs that could significantly differentiate the sample categories (i.e. mat, aggregates, sediment).

Taxonomic composition

Based on the returned co-assembly from METAWRAP and the clean reads, communities' taxonomic composition was assessed using Kraken2 (Wood, Lu and Langmead 2019) and the standard Kraken 2 database (NCBI: January 2022). GTDB-Tk (version 1.7.0) (Chaumeil *et al.* 2020) was used to classify genomes with the Genome Taxonomy Database (GTDB, version r202) (Parks *et al.* 2022). GTDB-Tk made use of pplacer (version 1.1.alpha19-0-g807f6f3) (Matsen, Kodner and Armbrust 2010) and FastANI (version 1.32) (Jain *et al.* 2018).

Functional annotation

Functions were predicted at two levels: both at the MAG level, as well as at the sample level. For the functional annotation at the MAG level, using the anvio' contigs database and the anvi-run-kegg-kofams program, the anvio' contigs database was annotated with HMM hits from KOfam, a database of KEGG Orthologs (KOs). Likewise, using the anvi-run-ncbi-cogs, NCBI's Clusters of Orthologous Groups (COGs) based annotations were added. The MAGs that correspond to the refined bins as they were retrieved after the metaWRAP and the anvio refinement steps, were annotated with KEGG modules; manually defined functional units of gene and reaction sets (Kanehisa *et al.* 2012). MAGs were "translated" to an anvio *collection* (i.e., a virtual construct storing bins of items in an anvi'o

profile database) and this collection was used along with the anvi-estimate-metabolism program to determine which enzymes are present in each MAG and compute the completeness of each metabolic module (scripts can be found under the anvio folder). An nMDS was constructed based on the presence/absence of modules in the MAGs using the jaccard similarity index.

For the functional annotation at the sample level, the clean reads as they were returned by the corresponding metaWRAP module and the DiTing tool (Xue *et al.* 2021) were used to estimate the contribution of each sample to the biogeochemical cycles incorporated in DiTing. DiTing used MEGAHIT (Li *et al.* 2015) to build the assembly of each sample separately (so, the co-assembly described in the *"Assembly and binning"* section was not used for this step) and Prodigal (Hyatt *et al.* 2010) to retrieve the Open Reading Frames (ORFs). KofamScan (Aramaki *et al.* 2020) was used for the annotation of the ORFs using KEGG ORTHOLOGY terms. The relative abundances of metabolic and biogeochemical functional pathways in each sample were then determined by DiTing.

Extracting molecular functions indicating osmoadaptation

Based on a review on strategies of adaptation of microorganisms to high salt concentrations (Gunde-Cimerman, Plemenitaš and Oren 2018), a list of compatible solutes in halotolerant/halophilic microorganisms, alkali metal–cation transporters and putative osmosensors was retrieved. Subsequently, a list of molecular function (KO) terms related to the processes responsible for all the above was created, available on Github, and their presence in the MAGs was investigated. In addition, based on a review on photosynthetic light harvesting (Croce and van Amerongen 2014), a list of light harvesting systems was created, and their related KO terms were retrieved. Afterwards, the samples where those systems were present were identified.

MAGs reference phylogenies

Bacteria_76 and Archaea_71 are sets of single copy genes (SCG) that are shared among all Bacteria and Archaea (Eren *et al.* 2015). Their intersection leads to a set of 25 SCG; i.e. the maximum number of SCG that 2 taxa potentially share. This set was used to build the phylogenetic tree of the reconstructed MAGs. The anvi-get-sequences-for-hmm-hits program of Anvi'o was used to extract and align the amino acid sequences of each of these genes from all the MAGs independently. This Anvi'o program makes use of MUSCLE (version 3.8.1551) (Edgar 2004) to return an alignment of the extracted sequences. Once all the amino acid sequence alignments were extracted, they were trimmed using Clipkit (version 1.1.5) (Steenwyk *et al.* 2020). A super matrix was then built using the single copy genes of the intersection. In cases where a MAG lacked a gene (due to sequencing limitations, assembly and/or binning related challenges etc.), alignment gaps were filled with dashes; the initial and the trimmed per gene alignments, the final super matrix alignment and a table showing exactly which SCGs were present in which MAG are available on the project's GitHub repository under the SCG folder. Using IQ-TREE2 (Hoang *et al.* 2018; Minh *et al.* 2020) the phylogeny of the reconstructed MAGs was built using 1,000 bootstrap replicates (-B 1,000) and 1,000 bootstrap replicates for Shimodaira–Hasegawa-like approximate likelihood ratio test (SH-aLRT) (-alrt 1000). The best-fit model (LG+R10) was retrieved using ModelFinder

(Kalyaanamoorthy *et al.* 2017). Using Barrnap (Seemann 2014) the 16S rRNA gene was extracted from the retrieved MAGs. The phylogeny of the MAGs and their relative abundances were integrated and visualised using GraPhlAn (Asnicar *et al.* 2015). All bioinformatics analyses were supported by the IMBBC High Performance Computing system (Zafeiropoulos *et al.* 2021).

Results

Taxonomic composition from 16S rRNA amplicon analysis

The DNA concentrations of the samples are shown in Table S1 and the results of the processing of the sequences are shown in Table S2. Sequencing of samples Elos08 and Elos11 was not successful and therefore, they were not included in the following analyses. The final number of OTUs, after removal of the OTUs that were also found on the blank sample, was 2,689 (see finalTable_noblanks.tsv). Overall, the most abundant phyla, as assessed by the relative abundance percentages of each replicate sample averaged per sampling station, were Bacteroidetes (~17%), Euryarchaeota (~16%), Proteobacteria (~15%) and Halanaerobiaeota (~10%, class Halanaerobiia) (Figure 2). Among the Bacteroidetes, the most abundant class was Bacteroidia (~14%), followed by Rhodothermia (~3%). Bacteroidetes had higher abundances in the aggregate samples (~25% on average in Elos03, Elos06 and Elos07) and lowest in the mat samples (~11% on average in Elos01 and Elos04). Euryarchaeota had very low abundances in samples Elos09, Elos10 and Elos13 and the higher abundances in the microbial mat samples (~28% on average in Elos01 and Elos04); along with the Halanaerobiaeota (~24% on average in Elos01 and Elos04) they were the dominant phyla of the mat samples. Halanaerobiaeota had low abundances in the orange and pink aggregates as well as in Elos09, Elos10 and Elos13. Among the Euryarchaeota, the most abundant class was Halobacteria (~14%) followed by Thermoplasmata (~2%).

Proteobacteria were almost equally distributed among all the three sample categories, i.e., mat, aggregates, sediments (~15%, ~21% and ~13% respectively), as well as among the classes Alphaproteobacteria (~7%), Gammaproteobacteria (~4%) and Deltaproteobacteria (~4%) (Figure S1). Although Patescibacteria had a low average abundance (~6%), they were dominant in Elos09 (~46%). In addition, Chloroflexi were almost absent from the top layers and the aggregates, but they were found in the bottom sediment layer and in the combined sediment samples. Cyanobacteria were about ~1% on average of all the samples.

The nMDS of the microbial OTUs (Figure S2) showed that their spatial pattern differs both by their type and the year of sampling, which was also confirmed by the PERMANOVA results (Type: $F.Model = 2.0396$, $p < 0.05$; Year: $F.Model = 2.3098$, $p < 0.05$).

Co-assembly, binning & taxonomic composition from shotgun metagenomics analysis

Shotgun metagenomic sequencing of the chosen six samples resulted in 744 million reads totalling 112.3 Gbp, with each sample ranging between 16.77 and 21.78 Gbp. Co-assembling of all the samples resulted in 1,5 million contigs totalling 5.04 Gbp (see assembly_report.html). The per-sample assemblies returned a total of 11.2 million contigs with a sum of 10.15 Gbp. Number of reads per sample, before and after the quality control, their length and the corresponding number of contigs are shown in Table S3.

Krona plots of the community profiles can be viewed through the kronagram.html. Based on the taxonomic profiles retrieved from Kraken2 (Figure S3), after removing sequences belonging to Viruses and sequences that could not be classified (~1%), Euryarchaeota (class Halobacteria) represent the majority of the total archaeal taxa (~30% on average); however, they were almost absent from sample Elos10 (abundance ~1%) while they were dominant in sample Elos01 (~59%). As far as bacterial taxa are concerned, the most abundant ones were Alphaproteobacteria (~19%), followed by Actinobacteria (~13%) and Gammaproteobacteria (~10%). Betaproteobacteria, delta/epsilon Proteobacteria subdivisions and Bacteroidetes/Chlorobi group had similar abundances (~5%, 4% and 5% respectively) with the latter having the highest abundance in sample Elos12 (~10%). Cyanobacteria were limited in all samples (~2% on average). Also, Kraken2 analysis did not identify any Nanoarchaeota; however, it identified in sample Elos01 the other archaeal taxa that are their hosts and namely a) *Ignicoccus hospitalis*, b) *Acidilobus* sp. 7A, c) *Vulcanisaeta* spp., d) *Pyrobaculum* spp., e) *Metallosphaera* spp., f) *Caldivirga* sp. and g) *Sulfolobus* sp.

Krona plots with the taxonomic profiles of each sample are available on GitHub. Prodigal predicted millions of genes per sample ranging from 2.1 (Elos03) to 4.3 million (Elos10). Their metabolic capacity/potential is further described in the “Biogeochemical cycles” section. Based on the blobology results, among the 1,513,505 co-assembled contigs a set of 102,250 were binned (see Figure S4); according to megaBlast and the nt database of NCBI, among the binned contigs 53,536 were bacterial and 2,230 archaeal while 60 contigs were assigned as viral and 739 as eukaryotic. The corresponding numbers for the case of the unbinned contigs were 1-2 orders of magnitude higher; thus, the number of unbinned contigs were 430,028 bacterial, 126,690 archaeal, 15,828 eukaryotic and 1,805 viral respectively (Figure S5).

MAGs phylogeny, functional annotation and distribution across samples

In line with the quality definitions described in (Bowers *et al.* 2017), metagenome binning generated a total of 289 MAGs; details are shown in the Supplementary File 2. According to the CheckM software 194 MAGs were reconstructed with a completeness higher than 90% and a contamination lower than 5% and all the rest had a completeness >70% and a contamination score <6%. According to the anvio summary (using the co-assembly as contigs database, the merged samples as profile and the reconstructed MAGs, i.e. the refined bins, as a collection) the redundancy of 10 (bin127, bin114, bin156, bin243, bin268, bin276, bin12, bin269, bin252, bin226) of the reconstructed MAGs was >10% (see binning_results.png). After the manual refinement of these 10 MAGs, a total of: (i)

178 bacterial *high* quality (completeness > 90%, contamination <5%), (ii) 70 bacterial and 39 archaeal *medium* quality (50% < completeness <90% and 5% <contamination <10%), and (iii) 2 bacterial MAGs of *low* quality (bin263 and bin182 with a completeness score <50%) were retrieved (see binning_reassembled.png). Combining the anvio summary results (see bin_by_bin folder in SECOND_SUMMARY) and the Barrnap outcome (see arc_rrnas and bac_rrnas on GitHub repo), the 16S rRNA gene was identified in 100 out of the total 250 bacterial MAGs. Likewise, from a total of 39 archaeal MAGs, 16S rRNA gene was found in 28 of them. Contigs included on those MAGs represented 1.03 Gbp of assembled reads. A set of 25 MAGs had a completeness of 100% and contamination less than 5% while 5 bacterial (MAG 143, MAG 66, MAG 129, MAG 189 and MAG 76) and 1 archaeal (MAG 232) MAGs among them had a contamination of 0%. Overall, bacterial MAGs had higher completeness scores.

Phylogenomic placement

The GTDB-Tk returned phylogenetic trees of the GTDB partition and the MAGs assigned to the corresponding domain for the cases of bacteria and archaea including the 2 low quality included (bin_182 : Proteobacteria and bin_263: Verrucomicrobiota). The phylogeny of the reconstructed MAGs (Figure 3) was built based on single-copy genes present on both Archaea and Bacteria, using the total number of MAGs even if some of the MAGs did not have all the 25 single-copy genes. Although not all these 25 single-copy genes were found in every MAG, still, the mean number of occurrences of a gene among the 289 MAGs was 266.68, ranging from 211 to 278. In general, the archaeal MAGs had the fewer single-copy genes, most probably due to their lower completeness. Using the total number of MAGs, the phylogeny of the reconstructed MAGs highlights the robustness of the method as even for those MAGs the phylogenetic signal was enough to place them among the representatives of their phylum. Thorough investigation of the tree pointed out that the only two discrepancies were that the sole MAG of the RBG-13-61-14 phylum (bin_124) that was placed among the Myxococcota representatives and that a representative of the Patescibacteria phylum (bin_61) was not placed close to the rest of Patescibacteria but as the closest relative of the representatives of the Chloroflexota phylum.

The novel candidate phylum (bin_202) was placed within the same clade with Eisenbacteria (bin_31). In general, bootstrap values were >90% with only exceptions a number of clades with representatives of the Nanoarchaeota phylum (of which the completeness and the number of single-copy genes present was relatively lower).

Distribution of MAGs across samples

Based on the classification of the retrieved MAGs (Figure S6; MAG abundances per sample (metaWRAP): link), the most abundant phylum was Bacteroidota (~28% on average), which almost dominated sample Elos03 (~57%) and Elos07 (~40%). The second most abundant phylum was Proteobacteria (~13% on average), with abundances ranging from ~2% in Elos02 to ~23% in Elos01 and ~22% in Elos07. Planctomycetota and Desulfobacterota were found at about ~8% and ~7% respectively, with the latter being absent from Elos03 and very rare in Elos07 (~2%). The only phylum

that was present only in the microbial aggregates, i.e., in Elos03 and Elos07, and was absent from all the other samples was Myxococcota.

The most abundant archaeal phylum was Nanoarchaeota (~5% on average), which was mostly found in Elos01 (~16%) and in much lower abundances in the other samples. Thermoplasmatota and Asgardarchaeota were found in similar abundances (~3% and ~2% on average respectively) and they were also absent from Elos03 and Elos07. Halobacteriota (~2% on average) were not found in Elos10 and Elos12 and were mostly present in Elos01 (~6%).

Elos10 was the sample with the highest number of bins (Figure S7) even if it was the one with the lowest number of reads. In addition, it seems to be closer to Elos02, the bottom layer sediment sample from July, and to sample Elos12. The microbial aggregates (Elos03 and Elos 07) form another cluster, distinct from the other samples, but closer to Elos01, the microbial mat sample.

The MAG 202 that represents a novel candidate phylum was present in samples Elos12, Elos10 and Elos02.

Based on the results of the simper analysis, there were no MAGs that could significantly contribute to the differentiation of the sample categories (i.e., mat, aggregates, sediment). The only comparison where certain MAGs were found to contribute, was the one between aggregates and combined sediment samples (Table S4).

Functional annotation of MAGs

The reconstructed MAGs were annotated with KofamScan with a range of KEGG ORTHOLOGY terms ranging from 354 to 2,879 terms (Figure S8), leading to 1 to 87 complete KEGG modules (Figure S9). The archaeal MAGs had, in general, lower completeness scores, lower number of KO terms assigned and less complete modules.

As it is shown in Figure S10, MAGs form distinct clusters both based on the taxonomy, i.e. if they are bacterial or archaeal (PERMANOVA: F.Model = 35.767, p<0.001), as well as based on their completeness (PERMANOVA: F.Model = 5.2156, p<0.001). The modules that contribute most to this clustering, as identified by the simper analysis, as well as the significance of any given module's contribution, are shown in Table S5. Examples of these modules were related to oxygenic photosynthesis and nitrogen, sulphur and carbon cycles. When examined separately, again they differ significantly by completeness (PERMANOVA: Bacteria: F.Model = 3.4053, p < 0.001; Archaea: F.Model = 2.4452, p < 0.001).

Comparison of taxonomies between amplicon and metagenomic analysis

As expected, the taxonomic composition of the microbial communities in our samples depends on the analytical procedure that was followed in each case. However, when the similarity matrices of the samples were compared, the pattern deriving from the relative abundances of the microbial OTUs as derived from the amplicon survey, was highly correlated with the one deriving from the classification of the shotgun metagenomic bins (Mantel statistic: r=0.83, p<0.001). The pattern deriving from the Kraken2 analysis was also correlated both with the amplicon survey, as well as

with the classification of the retrieved bins (Mantel statistic: r=0.45, p<0.05; r=0.52, p<0.05, respectively), but on a lesser degree.

Physicochemical analysis

The physicochemical variables are given in Table 2. According to the variation partitioning analysis (Table 3), the combination of oxygen and temperature explained 64% of the total variation in the Kraken2 community similarity matrix. For the other cases, i.e the amplicon data matrix and the classification of the retrieved MAGs, residuals were higher than 0.60 and therefore no significant models were retrieved.

Functional profiles at the sample level

A set of 783,693 unique proteins were predicted from a total of 3,532,725 hits with NCBI COGs. The MEGAHIT assembler as implemented in the framework of DiTing, returned the assembly of each sample ranging from 1,323,538 (Elos03) to 2,773,933 (Elos10) contigs.

The relative abundances of metabolic and biogeochemical functional pathways in each sample are available in the DiTing folder on the GitHub repository. As shown in Figure 4 and Figure S11, pathways belonging to the carbon cycle, central metabolism and other metabolism were the most abundant (~23%, ~19% and ~18% of the total pathways on average respectively). More specifically, the Reductive citrate cycle and the Dicarboxylate-hydroxybutyrate cycle dominate the carbon cycle (~8% and ~5%, ~18% respectively). In the central metabolism, pathways like the Embden-Meyerhof glycolysis pathway and tricarboxylic acid (TCA) cycle were most abundant (~7% each). In contrast, the Entner-Doudoroff pathway, i.e. an alternative glycolytic pathway was found in abundances an order of magnitude lower than the Embden-Meyerhof glycolysis pathway (Figure S12). Wood–Ljungdahl pathway that enables the use of hydrogen as an electron donor, was mostly found in bottom sediment layer sample (Elos02), but also in the combined sediments from the November 2019 sampling (Elos10 and Elos12) and to a lesser extent in the other samples. Bacterial chemotaxis, flagellar assembly and dissimilatory arsenic reduction are also among the most abundant pathways (~9% and~7%, ~3% respectively).

Regarding the methane cycle, methanogenesis pathways were found in the bottom sediment layer sample (Elos02), but also in the combined sediments from the November 2019 sampling (Elos10 and Elos12), as the Wood–Ljungdahl pathway; the most abundant methanogenesis pathway was the formation of methane from acetate. Interestingly, methane oxidation was almost absent in the samples.

Regarding the sulphur cycle, the most abundant pathways were the assimilatory and dissimilatory reduction of sulphate to sulphite (~1% each). As shown in Figure 5, thiosulphate oxidation as well as sulphite oxidation, but to a lesser extent, contribute to the sulphate pool. Sulphur disproportionation to sulphide and sulphite was absent in the aggregate samples, i.e., Elos03 and Elos07. In addition, although DMSO reduction was an abundant pathway in all the samples, DMS oxidation was very rare and mostly found in Elos03, i.e., the orange aggregate (Figure S13).

As shown in Figure 6, dissimilatory nitrate reduction to nitrite and nitrite to ammonia (DNRA) was prevalent in all samples, but it was mostly found in the combined sediment samples (Elos10 and Elos12). Denitrification (i.e., nitrite to nitric oxide and nitric oxide to nitrous oxide) was mostly found in sample Elos12. Nitrification, i.e., conversion of hydroxylamine to nitrite was almost absent from samples Elos01 (the microbial mat) and the microbial aggregates (Elos03 and Elos07); no *amoABC* genes were identified and *hao* genes were found only in bacterial MAGs. Nitrogen fixation was mostly abundant in the sample Elos07 (the pink aggregate).

Anaplerotic genes were also very abundant in our samples (~2%), and in particular the Pyruvate Carboxylase Pathway which produces oxaloacetate from pyruvate and replenishes the intermediates of the TCA cycle.

Finally, KO terms that were not added to one of the DiTing pathways regarding, K04641 (bacteriorhodopsin) and K04642 (halorhodopsin) were more abundant in the microbial mat sample, as well as in the microbial aggregates, and absent from Elos10. Sensory rhodopsin (K04643) was found in all samples, but in much higher abundances in Elos01, Elos03 and Elos07.

Osmoadaptation and light harvesting in the reconstructed MAGs

215 KOs related to osmoadaptation were identified; out of those, 61 were found in the MAGs and 59 were found in the samples (see GitHub repo). Each MAG had at least one of those KOs, with some MAGs having several osmoadaptation KOs (Figure S14). As shown in Figure S15, it seems that the preferred osmoadaptation strategy in all of the samples was related to the presence of alkali metal–cation transporters, while strategies related to uptake or production of compatible solutes were less abundant.

Regarding the light harvesting pigments/systems, we examined LH2 complex (“peripheral light-harvesting complex”), chlorosomes, phycobilisome and FMO protein (Figure S16). The K08930 term, corresponding to light-harvesting protein B-800-850 alpha chain, was present in all the samples. It was highly abundant in the mat and the aggregates (Elos01, Elos03 and Elos07) and much less in the sediment samples. Proteins related to phycobilisomes (K02094, K02096, K02097, K02290) were found in all of our samples and were most abundant in Elos 03 and Elos10.

KOs related to chlorosomes (various proteins of the chlorosome envelope), ranging from K08945 to K08954, were not detected in our samples. Similarly, the K08944 term, which corresponds to FMO protein, was absent from all the samples.

Discussion

Community composition and functional potential

Overall, there seems to be an agreement in all three ways that taxonomic composition has been derived for our samples. Even though six out of the eleven samples were sequenced using shotgun metagenomics, still we can compare them with the amplicon data and derive conclusions. In fact, the amplicon survey produced results that were highly correlated to the results of the classification of the retrieved bins; such congruence between the two methodologies has been reported in recent studies (Chan *et al.* 2015; Regalado *et al.* 2020) although there is also evidence for the opposite (Tessler *et al.* 2017). The results from Kraken2 were also correlated to the amplicon and the results of the classification of the retrieved MAGs, although the correlation was not as strong. However, Kraken2 is not restricted in the short amplicon length (Johnson *et al.* 2019) and therefore it was able to reach species level resolution (Lu and Salzberg 2020).

The top sediment layers, i.e. the microbial mat, was characterised by the presence of Halobacteria/Halobacteriota, Halanaerobiia and Nanoarchaeota which are all halophilic and were therefore able to withstand the salt crust that was on top of the mat (Norton, McGenity and Grant 1993; Casanueva *et al.* 2008; Çınar and Mutlu 2020; Akpolat *et al.* 2021). They were almost absent in the deeper sediment layers and they were absent in the winter, when there was no obvious microbial mat formed, despite the high salinity in sample Elos12. Therefore, the limiting factor for their presence could have been the lower temperature during the winter season since the optimum temperature for representatives of the Halobacteria is higher than 35°C (Grant 2015). Halobacteria, and their bacteriorhodopsin, were most likely responsible for the purple layer in the microbial mat, since they have been shown to cause striking red colours in salt flats (Stoeckenius, Lozier and Bogomolni 1979). Another organism responsible for the red colouration could have been the algae *Dunaliella salina* (Oren 2016; Naghoni *et al.* 2017); however, *Dunaliella* sp. was very rare in our shotgun metagenomic data as analysed by MGnify, therefore the red colouration cannot be attributed to algae. Nanoarchaeota are obligate symbionts and they have found in association with Crenarchaeota/Thermoproteota (Huber *et al.* 2002; Podar *et al.* 2013; Munson-McGee *et al.* 2015; Wurch *et al.* 2016; Merkel *et al.* 2017). Out of the known host-symbiont pairs and the putative hosts that have been proposed, we identified two of the known hosts (*Ignicoccus hospitalis* and *Acidilobus sp. 7A*) and five of the putative hosts (Jarett *et al.* 2018). However, the Nanoarchaeotal representatives were assigned at taxonomies higher than the species level. Thus, we can only presume that the symbionts *Nanoarchaeum equitans* and *Nanopusillus acidilobi* of the aforementioned hosts were present in our data.

It was hypothesised that the microbial aggregates would be more closely related to the microbial mat samples and this was confirmed by our results. However, the microbial aggregates were mostly characterised by the presence of Bacteroidetes/Bacteroidota and Myxococcota and the absence, or very low abundance, of Desulfobacterota, Thermoplasmatota and Asgardarchaeota.

In the photic zone of the microbial mats, phototrophic microorganisms transduce light into energy using either pigments (chlorophyll and/or bacteriochlorophylls) or retinal-based rhodopsins (Kurth *et*

al. 2021). In hypersaline microbial mats, phototrophy is possible because light can penetrate salt crusts and therefore the only limiting factor is the capability of the microbial communities to actually perform phototrophy under salt-saturated conditions (Meier *et al.* 2021). Out of the phyla that have been reported to perform (bacterio)chlorophyll-based phototrophy, i.e. Cyanobacteria, Proteobacteria, Chlorobi/Chlorobia, Chloroflexi/Chloroflexota, Firmicutes, Acidobacteria/Acidobacteriota, Eremiobacterota and Gemmatimonadetes/Gemmatimonadota (Zeng and Koblížek 2017; Zheng *et al.* 2022), only Proteobacteria and Chloroflexi/Chloroflexota were found in high abundances. Cyanobacteria were very rare in our samples, which can be attributed to the high salinity of the marsh (DiLoreto *et al.* 2019) which can lead to osmotic stress and inhibition of photosynthesis (Sudhir and Murthy 2004). So, Cyanobacteria were not the foundation of the microbial mats in our study, as has been shown in other examples of hypersaline microbial mats (Bolhuis, Cretoiu and Stal 2014; Wong, Ahmed-Cox and Burns 2016), which was expected as oxygenic photosynthesis is completely inhibited at saturation-level salinities (40%) (Meier *et al.* 2021). However, cyanobacterial genera that are characterised by strategies and survival mechanisms that allow them to grow in such high salinities (Oren 2015), such as *Euhalothece* and *Halothece*, were present in our samples. Chloroflexi/Chloroflexota were most present in the bottom sediment layer and in the combined sediment samples. Representatives that were identified from our samples are *Chloroflexus aurantiacus* (Pierson and Castenholz 1974) and *C. aggregans* (Hanada *et al.* 1995); they can grow photoautotrophically and photoheterotrophically under anaerobic conditions but also chemotrophically under aerobic conditions, using sulphide or hydrogen as an electron donor (Tang *et al.* 2011; Kawai *et al.* 2019, 2021). Given that it is unclear if they can harvest light deeper in the sediment and considering the low abundance of Cyanobacteria, it is most probable that Chloroflexi/Chloroflexota grow chemotrophically in our samples. According to our data, and since it is shown that light harvesting proteins are mainly in the mat and aggregate samples and less in the deeper sediment samples, Proteobacteria seem to be taking over the photosynthetic pathways. More specifically, some of the species that have been found to perform photosynthesis and are present in our samples are (from the Alphaproteobacteria class) *Citromicrobium* sp. (Jiao, Zhang and Zheng 2010), *Rhodopseudomonas palustris*, *Rhodobacter* sp., *Rhodospirillum rubrum*, *Roseobacter denitrificans*, *Bradyrhizobium* sp., *Roseobacter* sp. (Larimer *et al.* 2004; Bryant and Frigaard 2006) and (from the Gammaproteobacteria class) *Marichromatium purpuratum* (Shiung *et al.* 2018), *Congregibacter litoralis* (Fuchs *et al.* 2007; Spring *et al.* 2009), *Allochromatium* spp. (Kyndt and Meyer 2020). Overall, based on their taxonomic composition, the microbial mats under study seem to resemble microbial mats from an irregularly inundated tidal flat in Oman (Meier *et al.* 2021).

Regarding functional annotation of MAGs and their clustering according to taxonomy, it seems that it is driven by modules that are present in Archaea and absent in Bacteria, and vice versa. For example, regarding cysteine biosynthesis (M00021), this pathway is still unexplored in Archaea and although it has been found in certain species, e.g. *Methanosarcina barkeri*, it is suggested that there might be a different cysteine biosynthesis pathway in Archaea (Kitabatake *et al.* 2000). Likewise, there are no archaeal homologs for the bacterial pantothenate biosynthetic genes (Ronconi, Jonczyk and Genschel 2008), therefore pantothenate biosynthesis (M00913) is one of the pathways that

574 contributes to the differentiation between bacterial and archaeal MAGs. In addition, acetogen
575 (M00618) is only found in Bacteria.

576 Seasonality

577 Our second hypothesis was that samples from the deeper sediment layers collected in the summer
578 would be more similar to the combined sediment samples collected in the winter. This was also
579 confirmed by our results, as these samples were clustered together. However, despite the
580 similarities between the samples, there were also certain differences; as expected, there is a high
581 degree of spatial variation in the marsh under study (Dillon *et al.* 2009). Overall, in the deeper
582 sediment layers or where microbial mat was not formed, acetogens such as *Moorella thermoacetica*,
583 *Clostridium aceticum* and *Acetobacterium woodii* (Schuchmann and Müller 2016), might have been
584 utilising the Wood–Ljungdahl pathway, i.e. using H_2 hydrogen as an electron donor and CO_2 as an
585 electron acceptor. Acetogens are either competing directly with hydrogenotrophic methanogenic
586 archaea or interacting syntrophically with acetotrophic methanogens (Ragsdale and Pierce 2008).
587 Utilisation of acetate for methanogenesis is present in the genera *Methanosarcina* and *Methanothrix*
588 (Ferry 1992) which were both found in our samples. However, while methanogens were abundant
589 in our samples, they might have been contributing little to anaerobic mineralization, since in salinities
590 of 180‰ or less, they are inhibited by the increased activity of sulphate-reducing bacteria (Sørensen,
591 Canfield and Oren 2004). In addition, sulphate-reducing bacteria might have also been using the
592 Wood-Ljungdahl pathway in reverse (Ragsdale and Pierce 2008). It seems that other hypersaline
593 microbial mats, such as Guerrero Negro and Shark Bay, are characterized by a high number of
594 genes related to sulphur metabolism (Wong, Ahmed-Cox and Burns 2016) and hypersalinity in
595 coastal pans has been shown not to inhibit the activity of sulphate-reducing bacteria (Porter,
596 Roychoudhury and Cowan 2007). Therefore, the high salinity in our samples is not expected to
597 inhibit sulphur cycling. Sulphate-reducing microorganisms (SRM) from the Euryarchaeota lineage
598 (Muyzer and Stams 2008) were very abundant in the microbial mat and in the aggregates, while
599 SRM belonging to Deltaproteobacteria (Muyzer and Stams 2008) were present in the deeper and
600 the combined sediment samples. However, the occurrence of SRMs is not synonymous to the
601 occurrence of sulphate reduction in the given habitat (Muyzer and Stams 2008).
602 It has been proposed that arsenic and sulphur cycling can sustain high microbial metabolic rates in
603 permanently anoxic mats (Visscher *et al.* 2020). Bacteria capable of performing anoxygenic
604 photosynthesis using arsenite (As(III)) as an electron donor, such as *Ectothiorhodospira* sp. and
605 *Halorhodospira halophila* (Hoeft McCann *et al.* 2017), were present in our samples. It is suggested
606 that might be performing oxidation of As(III) to arsenate (As(V)), which is afterwards reduced back
607 to As(III) (Hoeft *et al.* 2004), which could explain the high occurrence of dissimilatory arsenic
608 reduction in our samples. Arsenate can be an important electron acceptor in the biogeochemical
609 cycling of carbon (Oremland *et al.* 2000); thus, arsenate reduction has a great potential to precipitate
610 carbonates and it is energetically better than sulphate reduction (Visscher *et al.* 2020), although
611 there is high potential for the latter to be also occurring our samples.

Thaumarchaeota, which are involved in nitrification in marine ecosystems (Veuger *et al.* 2013), were also present in our samples, according to the Kraken2 results, but in very low abundances, in contrast to other hypersaline microbial mats (Ruvindy *et al.* 2016). On the other hand, common nitrifying bacteria such as *Nitrobacter* where abundant in our samples; however, since the potential for nitrification was mostly present in the combined sediments and in the deeper layer, there seems to be a certain degree of hypersalinity limitation on the growth of nitrifying bacteria, as has been previously suggested (Jeffries *et al.* 2012). Interestingly, no *amoABC* genes were identified and, as expected, *hao* genes were found only in bacterial MAGs (Holmes, Dang and Smith 2019); it has been suggested that ammonia can be oxidised to hydroxylamine by nirS (Liu *et al.* 2018), which could potentially explain the lack of amoABC genes in our samples. Potential for denitrification was also present in our samples, although mostly in the combined sediment, and not limited by the increased salinity (Laverman *et al.* 2007).

Osmoadaptation

It has been debated if hypersaline environments are thermodynamic limiting the occurrence of self-sustaining microbial communities (Oren 2011) or if they are biologically permissive (Lee *et al.* 2018). Cells need to implement strategies to counteract the osmotic stress (Galinski 1995; Gunde-Cimerman, Plemenitaš and Oren 2018) but these strategies come with an energetic cost (Meier *et al.* 2021). It seems that the "salt in" strategy prevails over the "compatible solute" strategy in the samples of the present study; although they are both present, the former is much more abundant probably due to the high abundance of Halobacteria/Halobacteriota and Halanaerobiia (Mukhtar, Malik and Mehnaz 2020). This is in contrast to what has been observed in the hypersaline microbial mats of Shark Bay where choline and betaine uptake and betaine biosynthesis were the identified osmoadaptation strategies (Ruvindy *et al.* 2016).

It has been suggested that hypersaline microbial mats and, in particular, communities below salt crusts, cannot rely solely on primary production from anoxygenic phototrophy and mineralization from sulphate reduction (Meier *et al.* 2021). Instead, import of reduced substances and periods of reduced salinity are required, to allow the occurrence of oxygenic photosynthesis (Meier *et al.* 2021); in our study site this occurs in winter where evaporation is not as strong as in the summer, when combined with the increased precipitation, lowers the salinity of the marsh. During winter months, both the salt crust and the layering of the microbial mat disappears, as in Cardoso *et al.* (2019), and it seems that this temporal change and seasonal development of the microbial mats under study is the necessary element for the survival of the microorganisms. In addition, the potential for anaplerotic reactions, that were abundant in our samples, may play an important role in replenishing the intermediates of the TCA cycle, which is quite abundant in out samples, and thus allowing microbial growth with a carbohydrate as the sole carbon source (Tong 2013; Choi *et al.* 2016). Although it seems that hypersaline environments are "thermodynamically moderate", DNA based studies can only identify the members of a community and not their metabolic activities. Therefore, future studies on hypersaline microbial mats should focus on the combination of metabolomics, metatranscriptomics and metagenomics, in order to elucidate the functional repertoire of microbial

communities, their metabolic potential and their metabolic and ecological interactions. Metabolic modelling of the microbial assemblages can shed further light on the effects of the environmental challenges on the mat construction as well as on which processes are taking place within each mat layer and among its different layers.

Funding

This work was supported by the project RECONNECT [MIS 5017160] financed by the Transnational Cooperation Programme Interreg V-B "Balkan-Mediterranean 2014-2020" and co-funded by the European Union and national funds of the participating countries. There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We would like to thank the reviewers for their valuable comments and suggestions on our manuscript. We would also like to thank the Management Agency of the Dodecanese Protected Areas (MADPA) and especially Mr Dinos Protopapas and Mr Giorgos Prearis (captain of the R/V Saria) for providing assistance during our visit to the Tristomo bay. In addition, we would like to thank a) Mr Antonis Potirakis and Mr Stelios Ninidakis, the best HPC system administrators, for their help and support during cluster maintenance and installation of third-party software, b) Dr. Christos A. Christakis (ORCID iD: 0000-0002-7075-0996) for his suggestion to choose a few samples for shotgun metagenomic sequencing so that higher coverage could be achieved, c) Dr. Paschalis Natsidis (ORCID iD: 0000-0002-2446-3208) for his insight on the creation of the phylogenomic trees, d) Dr Jon Bent Kristoffersen (ORCID iD: 0000-0002-8804-1864) for performing the MiSeq sequencing at the IMBBC DNA Sequencing platform and e) Prof. Jens Carlsson (ORCID iD: 0000-0002-9262-5627) for allowing us to use the *bob* server of Area 52 lab in University College Dublin. This research was supported in part through computational resources provided by IMBBC (Institute of Marine Biology, Biotechnology and Aquaculture) of the HCMR (Hellenic Centre for Marine Research). Funding for establishing the IMBBC HPC has been received by the MARBIGEN (EU Regpot) project, LifeWatchGreece RI and the CMBR (Centre for the study and sustainable exploitation of Marine Biological Resources) RI.

References

Akpolat C, Fernández AB, Caglayan P *et al.* Prokaryotic communities in the thalassohaline Tuz Lake, Deep Zone, and Kayacik, Kaldirim and Yavsan Salterns (Turkey) assessed by 16S rRNA amplicon sequencing. *Microorganisms* 2021;**9**:1525.

Al-Thani R, Al-Najjar MAA, Al-Raei AM *et al.* Community structure and activity of a highly dynamic and nutrient-limited hypersaline microbial mat in Um Alhool Sabkha, Qatar. *PLOS ONE* 2014;**9**:e92405.

Apprill A, McNally S, Parsons R *et al.* Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat Microb Ecol* 2015;**75**:129–37.

Aramaki T, Blanc-Mathieu R, Endo H *et al.* KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 2020;**36**:2251–2.

Asnicar F, Weingart G, Tickle TL *et al.* Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* 2015;**3**:e1029.

Bell KL, Petit III RA, Cutler A *et al.* Comparing whole-genome shotgun sequencing and DNA metabarcoding approaches for species identification and quantification of pollen species mixtures. *Ecol Evol* 2021;**11**:16082–98.

Bolhuis H, Cretoiu MS, Stal LJ. Molecular ecology of microbial mats. *FEMS Microbiol Ecol* 2014;**90**:335–50.

Bowers RM, Kyrpides NC, Stepanauskas R *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.

Bryant DA, Frigaard N-U. Prokaryotic photosynthesis and phototrophy illuminated. *Trends Microbiol* 2006;**14**:488–96.

Cardoso DC, Cretoiu MS, Stal LJ *et al.* Seasonal development of a coastal microbial mat. *Sci Rep* 2019;**9**:9035.

Casanueva A, Galada N, Baker GC *et al.* Nanoarchaeal 16S rRNA gene sequences are widely dispersed in hyperthermophilic and mesophilic halophilic environments. *Extremophiles* 2008;**12**:651–6.

Chan CS, Chan K-G, Tay Y-L *et al.* Diversity of thermophiles in a Malaysian hot spring determined using 16S rRNA and shotgun metagenome sequencing. *Front Microbiol* 2015;**6**.

Chaumeil P-A, Mussig AJ, Hugenholtz P *et al.* GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 2020;**36**:1925–7.

Chen R, Wong HL, Kindler GS *et al.* Discovery of an abundance of biosynthetic gene clusters in Shark Bay microbial mats. *Front Microbiol* 2020;**11**.

Choi PH, Jo J, Lin Y-C *et al.* A distinct holoenzyme organization for two-subunit pyruvate carboxylase. *Nat Commun* 2016;**7**:12713.

Çınar S, Mutlu MB. Prokaryotic community compositions of the hypersaline sediments of Tuz Lake demonstrated by cloning and High-Throughput Sequencing. *Microbiology* 2020;**89**:756–68.

Clooney AG, Fouhy F, Sleator RD *et al.* Comparing apples and oranges?: next generation sequencing and its impact on microbiome analysis. *PLOS ONE* 2016;**11**:e0148028.

Croce R, van Amerongen H. Natural strategies for photosynthetic light harvesting. *Nat Chem Biol* 2014;**10**:492–501.

Cummins C, Ahamed A, Aslam R *et al.* The European Nucleotide Archive in 2021. *Nucleic Acids Res* 2022;**50**:D106–10.

Dávila-Ramos S, Castelán-Sánchez HG, Martínez-Ávila L *et al.* A review on viral metagenomics in extreme environments. *Front Microbiol* 2019;**10**.

Des Marais DJ. The biogeochemistry of hypersaline microbial mats. In: Jones JG (ed.). *Advances in Microbial*

Ecology. Boston, MA: Springer US, 1995, 251–74.
Dillon JG, Miller S, Bebout B *et al.* Spatial and temporal variability in a stratified hypersaline microbial mat community. *FEMS Microbiol Ecol* 2009;**68**:46–58.
DiLoreto ZA, Bontognali TRR, Al Disi ZA *et al.* Microbial community composition and dolomite formation in the hypersaline microbial mats of the Khor Al-Adaid sabkhas, Qatar. *Extremophiles* 2019;**23**:201–18.
Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7.
Eren AM, Esen ÖC, Quince C *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 2015;**3**:e1319.
Ferry JG. Methane from acetate. *J Bacteriol* 1992;**174**:5489–95.
Fuchs BM, Spring S, Teeling H *et al.* Characterization of a marine gammaproteobacterium capable of aerobic anoxygenic photosynthesis. *Proc Natl Acad Sci* 2007;**104**:2891–6.
Galinski EA. Osmoadaptation in Bacteria. In: Poole RK (ed.). *Advances in Microbial Physiology*. Vol 37. Academic Press, 1995, 273–328.
van Gemerden H. Microbial mats: A joint venture. *Mar Geol* 1993;**113**:3–25.
Grant WD. Halobacterium. *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons, Ltd, 2015, 1–11.
Guerrero R, Piqueras M, Berlanga M. Microbial mats and the search for minimal ecosystems. *Int Microbiol* 2002;**5**:177–88.
Gunde-Cimerman N, Plemenitaš A, Oren A. Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS Microbiol Rev* 2018;**42**:353–75.
Gurevich A, Saveliev V, Vyahhi N *et al.* QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;**29**:1072–5.
Hanada S, Hiraishi A, Shimada K *et al.* Chloroflexus aggregans sp. nov., a filamentous phototrophic bacterium which forms dense cell aggregates by active gliding movement. *Int J Syst Evol Microbiol* 1995;**45**:676–81.
Hedges JI, Stern JH. Carbon and nitrogen determinations of carbonate-containing solids1. *Limnol Oceanogr* 1984;**29**:657–63.
Hedlund BP, Dodsworth JA, Murugapiran SK *et al.* Impact of single-cell genomics and metagenomics on the emerging view of extremophile "microbial dark matter." *Extremophiles* 2014;**18**:865–75.
Henckel T, Friedrich M, Conrad R. Molecular analyses of the methane-oxidizing microbial community in rice field soil by targeting the genes of the 16S rRNA, particulate methane monooxygenase, and methanol dehydrogenase. *Appl Environ Microbiol* 1999;**65**:1980–90.
Herlemann DP, Labrenz M, Jürgens K *et al.* Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* 2011;**5**:1571–9.
Hoang DT, Chernomor O, von Haeseler A *et al.* UFBoot2: Improving the Ultrafast Bootstrap approximation. *Mol Biol Evol* 2018;**35**:518–22.
Hoeft McCann S, Boren A, Hernandez-Maldonado J *et al.* Arsenite as an electron donor for anoxygenic photosynthesis: description of three strains of Ectothiorhodospira from Mono Lake, California and Big Soda Lake, Nevada. *Life* 2017;**7**:1.
Hoeft SE, Kulp TR, Stolz JF *et al.* Dissimilatory arsenate reduction with sulfide as electron donor: experiments with Mono Lake Water and isolation of strain MLMS-1, a chemoautotrophic arsenate respirer. *Appl Environ Microbiol* 2004;**70**:2741–7.
Hoehler TM, Bebout BM, Des Marais DJ. The role of microbial mats in the production of reduced gases on the early Earth. *Nature* 2001;**412**:324–7.

Holmes DE, Dang Y, Smith JA. Nitrogen cycling during wastewater treatment. *Advances in Applied Microbiology*. Vol 106. Elsevier, 2019, 113–92.
Huber H, Hohn MJ, Rachel R *et al.* A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* 2002;**417**:63–7.
Hugenholtz P, Pitulle C, Hershberger KL *et al.* Novel Division Level Bacterial Diversity in a Yellowstone Hot Spring. *J Bacteriol* 1998;**180**:366–76.
Hyatt D, Chen G-L, LoCascio PF *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;**11**:119.
Jain C, Rodriguez-R LM, Phillippy AM *et al.* High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;**9**:5114.
Jarett JK, Nayfach S, Podar M *et al.* Single-cell genomics of co-sorted Nanoarchaeota suggests novel putative host associations and diversification of proteins involved in symbiosis. *Microbiome* 2018;**6**:161.
Jeffries TC, Seymour JR, Newton K *et al.* Increases in the abundance of microbial genes encoding halotolerance and photosynthesis along a sediment salinity gradient. *Biogeosciences* 2012;**9**:815–25.
Jiao N, Zhang R, Zheng Q. Coexistence of two different photosynthetic operons in Citromicrobium bathyomarinum JL354 as revealed by whole-genome sequencing. *J Bacteriol* 2010;**192**:1169–70.
Johnson JS, Spakowicz DJ, Hong B-Y *et al.* Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* 2019;**10**:5029.
Jørgensen BB. Diffusion processes and boundary layers in microbial mats. In: Stal LJ, Caumette P (eds.). *Microbial Mats*. Berlin, Heidelberg: Springer, 1994, 243–53.
Jørgensen BB. Space for hydrogen. *Nature* 2001;**412**:286–9.
Kalyaanamoorthy S, Minh BQ, Wong TKF *et al.* ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;**14**:587–9.
Kanehisa M, Goto S, Sato Y *et al.* KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 2012;**40**:D109–14.
Kang DD, Li F, Kirton E *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 2019;**7**:e7359.
Kawai S, Martinez JN, Lichtenberg M *et al.* In-situ metatranscriptomic analyses reveal the metabolic flexibility of the thermophilic anoxygenic photosynthetic bacterium Chloroflexus aggregans in a hot spring Cyanobacteria-dominated microbial mat. *Microorganisms* 2021;**9**:652.
Kawai S, Nishihara A, Matsuura K *et al.* Hydrogen-dependent autotrophic growth in phototrophic and chemolithotrophic cultures of thermophilic bacteria, Chloroflexus aggregans and Chloroflexus aurantiacus, isolated from Nakabusa hot springs. *FEMS Microbiol Lett* 2019;**366**:fnz122.
Kelley CA, Poole JA, Tazaz AM *et al.* Substrate Limitation for Methanogenesis in Hypersaline Environments. *Astrobiology* 2012;**12**:89–97.
Kindler GS, Wong HL, Larkum AWD *et al.* Genome-resolved metagenomics provides insights into the functional complexity of microbial mats in Blue Holes, Shark Bay. *FEMS Microbiol Ecol* 2022;**98**:fiab158.
Kitabatake M, So MW, Tumbula DL *et al.* Cysteine biosynthesis pathway in the archaeon Methanosarcina barkeri encoded by acquired bacterial genes? *J Bacteriol* 2000;**182**:143–5.
Klindworth A, Pruesse E, Schweer T *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 2013;**41**:e1.
Kozubal MA, Romine M, Jennings R deM *et al.* Geoarchaeota: a new candidate phylum in the Archaea from

high-temperature acidic iron mats in Yellowstone National Park. *ISME J* 2013;**7**:622–34.
Krueger F. Trim Galore. 2022.
Krumbein WE, Paterson DM, Zavarzin GA. Fossil and recent biofilms: a natural history of life on earth. *Foss Recent Biofilms Nat Hist Life Earth* 2003.
Kurth D, Elias D, Rasuk MC *et al.* Carbon fixation and rhodopsin systems in microbial mats from hypersaline lakes Brava and Tebenquiche, Salar de Atacama, Chile. *PLOS ONE* 2021;**16**:e0246656.
Kyndt JA, Meyer TE. Genome sequences of Allochromatium palmeri and Allochromatium humboldtianum expand the Allochromatium family tree of purple sulfur photosynthetic Bacteria within the Gammaproteobacteria and further refine the genus. *Microbiol Resour Announc* 2020;**9**:e00774-20.
Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–9.
Larimer FW, Chain P, Hauser L *et al.* Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris. *Nat Biotechnol* 2004;**22**:55–61.
Laverman AM, Canavan RW, Slomp CP *et al.* Potential nitrate removal in a coastal freshwater sediment (Haringvliet Lake, The Netherlands) and response to salinization. *Water Res* 2007;**41**:3061–8.
Lee CJD, McMullan PE, O'Kane CJ *et al.* NaCl-saturated brines are thermodynamically moderate, rather than extreme, microbial habitats. *FEMS Microbiol Rev* 2018;**42**:672–93.
Li D, Liu C-M, Luo R *et al.* MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinforma Oxf Engl* 2015;**31**:1674–6.
Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;**25**:1754–60.
Liu X, Shu Z, Sun D *et al.* Heterotrophic Nitrifiers Dominate Reactors Treating Incineration Leachate with High Free Ammonia Concentrations. *ACS Sustain Chem Eng* 2018;**6**:15040–9.
van der Loos LM, Nijland R. Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Mol Ecol* 2021;**30**:3270–88.
López-López A, Richter M, Peña A *et al.* New insights into the archaeal diversity of a hypersaline microbial mat obtained by a metagenomic approach. *Syst Appl Microbiol* 2013;**36**:205–14.
Lu J, Salzberg SL. Ultrafast and accurate 16S rRNA microbial community analysis using Kraken 2. *Microbiome* 2020;**8**:124.
Lueders T, Manefield M, Friedrich MW. Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environ Microbiol* 2004;**6**:73–8.
Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011;**17**:10–2.
Matsen FA, Kodner RB, Armbrust EV. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 2010;**11**:538.
McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* 2013;**8**:e61217.
Meier DV, Greve AJ, Chennu A *et al.* Limitation of microbial processes at saturation-level salinities in a microbial mat covering a coastal salt flat. *Appl Environ Microbiol* 2021;**87**:e00698-21.
Merkel AYu, Pimenov NV, Rusanov II *et al.* Microbial diversity and autotrophic activity in Kamchatka hot springs. *Extremophiles* 2017;**21**:307–17.
Minh BQ, Schmidt HA, Chernomor O *et al.* IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;**37**:1530–4.
Mitchell AL, Almeida A, Beracochea M *et al.* MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Res* 2020;**48**:D570–8.

Mukhtar S, Malik K, Mehnaz S. Osmoadaptation in halophilic bacteria and archaea. *Res J Biotechnol* 2020;**15**:154–61.
Munson-McGee JH, Field EK, Bateson M *et al.* Nanoarchaeota, their Sulfolobales host, and Nanoarchaeota virus distribution across Yellowstone National Park hot springs. *Appl Environ Microbiol* 2015;**81**:7860–8.
Muyzer G, Stams AJM. The ecology and biotechnology of sulphate-reducing bacteria. *Nat Rev Microbiol* 2008;**6**:441–54.
Naghoni A, Emtiazi G, Amoozegar MA *et al.* Microbial diversity in the hypersaline Lake Meyghan, Iran. *Sci Rep* 2017;**7**:11522.
Norton CF, McGenity TJ, Grant WDY 1993. Archaeal halophiles (halobacteria) from two British salt mines. *Microbiology* 1993;**139**:1077–81.
Nurk S, Meleshko D, Korobeynikov A *et al.* metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 2017;**27**:824–34.
Oksanen J, Blanchet FG, Friendly M *et al.* vegan: Community Ecology Package. 2020.
Oremland RS, Dowdle PR, Hoeft S *et al.* Bacterial dissimilatory reduction of arsenate and sulfate in meromictic Mono Lake, California. *Geochim Cosmochim Acta* 2000;**64**:3073–84.
Oren A. Thermodynamic limits to microbial life at high salt concentrations. *Environ Microbiol* 2011;**13**:1908–23.
Oren A. Cyanobacteria in hypersaline environments: biodiversity and physiological properties. *Biodivers Conserv* 2015;**24**:781–98.
Oren A. Life in Hypersaline Environments. In: Hurst CJ (ed.). *Their World: A Diversity of Microbial Environments*. Cham: Springer International Publishing, 2016, 301–39.
Paerl HW, Pinckney JL, Steppe TF. Cyanobacterial–bacterial mat consortia: examining the functional unit of microbial survival and growth in extreme environments. *Environ Microbiol* 2000;**2**:11–26.
Parks DH, Chuvochina M, Rinke C *et al.* GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 2022;**50**:D785–94.
Parks DH, Imelfort M, Skennerton CT *et al.* CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;**25**:1043–55.
Patro R, Duggal G, Love MI *et al.* Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 2017;**14**:417–9.
Pavloudi C, Kristoffersen JB, Oulas A *et al.* Sediment microbial taxonomic and functional diversity in a natural salinity gradient challenge Remane's "species minimum" concept. *PeerJ* 2017;**5**:e3687.
Pierson BK, Castenholz RW. A phototrophic gliding filamentous bacterium of hot springs, Chloroflexus aurantiacus, gen. and sp. nov. *Arch Microbiol* 1974;**100**:5–24.
Podar M, Makarova KS, Graham DE *et al.* Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park. *Biol Direct* 2013;**8**:9.
Porter D, Roychoudhury AN, Cowan D. Dissimilatory sulfate reduction in hypersaline coastal pans: Activity across a salinity gradient. *Geochim Cosmochim Acta* 2007;**71**:5102–16.
Potter EG, Bebout BM, Kelley CA. Isotopic Composition of Methane and Inferred Methanogenic Substrates Along a Salinity Gradient in a Hypersaline Microbial Mat System. *Astrobiology* 2009;**9**:383–90.
Prieto-Barajas CM, Alcaraz LD, Valencia-Cantero E *et al.* Life in hot spring microbial mats located in the trans-Mexican Volcanic Belt: A 16S/18S rRNA gene and metagenomic analysis. *Geomicrobiol J* 2018;**35**:704–12.

Prieto-Barajas CM, Valencia-Cantero E, Santoyo G. Microbial mat ecosystems: Structure types, functional diversity, and biotechnological application. *Electron J Biotechnol* 2018;**31**:48–56.
Quast C, Pruesse E, Yilmaz P *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:D590–6.
R Core Team. R: A language and environment for statistical computing. 2021.
Ragsdale SW, Pierce E. Acetogenesis and the Wood-Ljungdahl Pathway of CO2 fixation. *Biochim Biophys Acta* 2008;**1784**:1873–98.
Regalado J, Lundberg DS, Deusch O *et al.* Combining whole-genome shotgun sequencing and rRNA gene amplicon analyses to improve detection of microbe–microbe interaction networks in plant leaves. *ISME J* 2020;**14**:2116–30.
Robertson CE, Spear JR, Harris JK *et al.* Diversity and stratification of Archaea in a hypersaline microbial mat. *Appl Environ Microbiol* 2009;**75**:1801–10.
Ronconi S, Jonczyk R, Genschel U. A novel isoform of pantothenate synthetase in the Archaea. *FEBS J* 2008;**275**:2754–64.
Ruvindy R, White III RA, Neilan BA *et al.* Unravelling core microbial metabolisms in the hypersaline microbial mats of Shark Bay using high-throughput metagenomics. *ISME J* 2016;**10**:183–96.
Schuchmann K, Müller V. Energetics and application of heterotrophy in acetogenic Bacteria. *Appl Environ Microbiol* 2016;**82**:4056–69.
Seemann T. Barrnap: BAsic Rapid Ribosomal RNA Predictor. *Barrnap BAsic Rapid Ribosomal RNA Predict - GitHub Repos* 2014.
Segata N. On the road to strain-resolved comparative metagenomics. *mSystems* 2018;**3**:e00190-17.
Sharpton TJ. An introduction to the analysis of shotgun metagenomic data. *Front Plant Sci* 2014;**5**.
Shiung I-I, Chang M-J, Chang Y-T *et al.* Photosynthetic purple sulfur bacterium Marichromatium purpuratum RuA2 induces changes in water quality parameters, the occurrence of sulfonamide resistance gene and microbial community structure of marine aquaculture. *Aquaculture* 2018;**493**:68–78.
Sievert SM, Ziebis W, Kuever J *et al.* Relative abundance of Archaea and Bacteria along a thermal gradient of a shallow-water hydrothermal vent quantified by rRNA slot-blot hybridization. *Microbiology* 2000;**146**:1287–93.
Song W-Z, Thomas T. Binning_refiner: improving genome bins through the combination of different binning programs. *Bioinforma Oxf Engl* 2017;**33**:1873–5.
Sørensen KB, Canfield DE, Oren A. Salinity responses of benthic microbial communities in a solar saltern (Eilat, Israel). *Appl Environ Microbiol* 2004;**70**:1608–16.
Spring S, Lünsdorf H, Fuchs BM *et al.* The photosynthetic apparatus and Its regulation in the aerobic Gammaproteobacterium Congregibacter litoralis gen. nov., sp. nov. *PLOS ONE* 2009;**4**:e4866.
Stal LJ. Cyanobacterial mats and stromatolites. In: Whitton BA (ed.). *Ecology of Cyanobacteria II: Their Diversity in Space and Time*. Dordrecht: Springer Netherlands, 2012, 65–125.
Steenwyk JL, Iii TJB, Li Y *et al.* ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. *PLOS Biol* 2020;**18**:e3001007.
Steinberger A. asteinberger9/seq_scripts. 2020.
Stoeckenius W, Lozier RH, Bogomolni RA. Bacteriorhodopsin and the purple membrane of halobacteria. *Biochim Biophys Acta BBA - Rev Bioenerg* 1979;**505**:215–78.
Sudhir P, Murthy SDS. Effects of salt stress on basic processes of photosynthesis. *Photosynthetica* 2004;**42**:481–6.
Tang K-H, Barry K, Chertkov O *et al.* Complete genome sequence of the filamentous anoxygenic

phototrophic bacterium Chloroflexus aurantiacus. *BMC Genomics* 2011;**12**:334.
Tessler M, Neumann JS, Afshinnekoo E *et al.* Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Sci Rep* 2017;**7**:6589.
Tong L. Structure and function of biotin-dependent carboxylases. *Cell Mol Life Sci* 2013;**70**:863–91.
Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 2018;**6**:158.
Veuger B, Pitcher A, Schouten S *et al.* Nitrification and growth of autotrophic nitrifying bacteria and Thaumarchaeota in the coastal North Sea. *Biogeosciences* 2013;**10**:1775–85.
Villagrasa E, Ferrer-Miralles N, Millach L *et al.* Morphological responses to nitrogen stress deficiency of a new heterotrophic isolated strain of Ebro Delta microbial mats. *Protoplasma* 2019;**256**:105–16.
Villanueva L, Navarrete A, Urmeneta J *et al.* Analysis of diurnal and vertical microbial diversity of a hypersaline microbial mat. *Arch Microbiol* 2007;**188**:137–46.
Visscher PT, Gallagher KL, Bouton A *et al.* Modern arsenotrophic microbial mats provide an analogue for life in the anoxic Archean. *Commun Earth Environ* 2020;**1**:1–10.
Visscher PT, Stolz JF. Microbial mats as bioreactors: populations, processes, and products. In: Noffke N (ed.). *Geobiology: Objectives, Concepts, Perspectives*. Amsterdam: Elsevier, 2005, 87–100.
Wickham H, Chang W, Wickham MH. Package 'ggplot2'. Create elegant data visualisations using the grammar of graphics. 2016.
Wong HL, Ahmed-Cox A, Burns BP. Molecular ecology of hypersaline microbial mats: current insights and new directions. *Microorganisms* 2016;**4**:6.
Wong HL, MacLeod FI, White RA *et al.* Microbial dark matter filling the niche in hypersaline microbial mats. *Microbiome* 2020;**8**:135.
Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol* 2019;**20**:257.
Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 2016;**32**:605–7.
Wurch L, Giannone RJ, Belisle BS *et al.* Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from a terrestrial geothermal environment. *Nat Commun* 2016;**7**:12115.
WWF Greece. Inventory report: KAR001 - Tristomo marsh. *GrIsWet – Greek Isl Wetl Database* 2022.
Xue C-X, Lin H, Zhu X-Y *et al.* DiTing: a pipeline to infer and compare biogeochemical pathways from metagenomic and metatranscriptomic data. *Front Microbiol* 2021;**12**.
Yentsch CS, Menzel DW. A method for the determination of phytoplankton chlorophyll and phaeophytin by fluorescence. *Deep Sea Res Oceanogr Abstr* 1963;**10**:221–31.
Zafeiropoulos H, Gioti A, Ninidakis S *et al.* 0s and 1s in marine molecular research: a regional HPC perspective. *GigaScience* 2021;**10**:giab053.
Zafeiropoulos H, Viet HQ, Vasileiadou K *et al.* PEMA: a flexible Pipeline for Environmental DNA Metabarcoding Analysis of the 16S/18S ribosomal RNA, ITS, and COI marker genes. *GigaScience* 2020;**9**, DOI: 10.1093/gigascience/giaa022.
Zeng Y, Koblížek M. Phototrophic Gemmatimonadetes: a new "purple" branch on the bacterial tree of life. In: Hallenbeck PC (ed.). *Modern Topics in the Phototrophic Prokaryotes: Environmental and Applied Aspects*. Cham: Springer International Publishing, 2017, 163–92.
Zheng R, Cai R, Wang C *et al.* Characterization of the first cultured representative of "Candidatus Thermofonsia" Clade 2 within Chloroflexi reveals its phototrophic lifestyle. *mBio* 2022;**13**:e00287-22.