

RESEARCH

Open Access



# Competition-cooperation in the chemoautotrophic ecosystem of Movile Cave: first metagenomic approach on sediments

Iulia Chiciudean<sup>1\*</sup>, Giancarlo Russo<sup>2</sup>, Diana Felicia Bogdan<sup>1,3</sup>, Erika Andrea Levei<sup>4</sup>, Luchiana Faur<sup>5,6</sup>, Alexandra Hillebrand-Voiculescu<sup>6,7</sup>, Oana Teodora Moldovan<sup>6,8</sup> and Horia Leonard Banciu<sup>1,9\*</sup>

## Abstract

**Background:** Movile Cave (SE Romania) is a chemoautotrophically-based ecosystem fed by hydrogen sulfide-rich groundwater serving as a primary energy source analogous to the deep-sea hydrothermal ecosystems. Our current understanding of Movile Cave microbiology has been confined to the sulfidic water and its proximity, as most studies focused on the water-floating microbial mat and planktonic accumulations likely acting as the primary production powerhouse of this unique subterranean ecosystem. By employing comprehensive genomic-resolved metagenomics, we questioned the spatial variation, chemoautotrophic abilities, ecological interactions and trophic roles of Movile Cave's microbiome thriving beyond the sulfidic-rich water.

**Results:** A customized bioinformatics pipeline led to the recovery of 106 high-quality metagenome-assembled genomes from 7 cave sediment metagenomes. Assemblies' taxonomy spanned 19 bacterial and three archaeal phyla with *Acidobacteriota*, *Chloroflexota*, *Proteobacteria*, *Planctomycetota*, *Ca. Patescibacteria*, *Thermoproteota*, *Methylo-mirabilota*, and *Ca. Zixibacteria* as prevalent phyla. Functional gene analyses predicted the presence of CO<sub>2</sub> fixation, methanotrophy, sulfur and ammonia oxidation in the explored sediments. Species Metabolic Coupling Analysis of metagenome-scale metabolic models revealed the highest competition-cooperation interactions in the sediments collected away from the water. Simulated metabolic interactions indicated autotrophs and methanotrophs as major donors of metabolites in the sediment communities. Cross-feeding dependencies were assumed only towards 'currency' molecules and inorganic compounds (O<sub>2</sub>, PO<sub>4</sub><sup>3-</sup>, H<sup>+</sup>, Fe<sup>2+</sup>, Cu<sup>2+</sup>) in the water proximity sediment, whereas hydrogen sulfide and methanol were assumedly traded exclusively among distant gallery communities.

**Conclusions:** These findings suggest that the primary production potential of Movile Cave expands way beyond its hydrothermal waters, enhancing our understanding of the functioning and ecological interactions within chemo-lithoautotrophically-based subterranean ecosystems.

**Keywords:** Chemoautotrophic, Genome-scale metabolic models, Competition-cooperation interactions, Cross-feeding dependencies, Sulfidic cave, Romania

## Background

The most studied chemoautotrophically-based ecosystems are deep-sea hydrothermal vents and cold seeps [1–4]. But carbon fixation via chemosynthesis also plays a crucial role in reduced sulfur-rich water and sediments of circumneutral saline and soda lakes [5–7] and sulfidic caves [8–10]. In contrast to deep-sea chemoautotrophic

\*Correspondence: iulia.chiciudean@ubbcluj.ro; horia.banciu@ubbcluj.ro

<sup>1</sup> Department of Molecular Biology and Biotechnology, Babeş-Bolyai University, Cluj-Napoca, Romania  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

systems, their subterranean counterparts, namely chemoautotrophic caves, have been largely underexplored. To date, only a few cave ecosystems that rely totally or partially on chemosynthesis have been explored, including Ayyalon Cave (Israel), Movile Cave (Romania), Frasassi Cave (Italy), Lower Kane Cave and Cesspool Cave (USA), and Villa Luz Cave (Mexico) [9, 11–16]. In these caves, microorganisms produce organic matter in situ, starting from inorganic compounds as nutrient and energy sources.

The Movile Cave (SE Romania) was the first to be mentioned as defying the conventional view of subterranean ecosystems as supported by aboveground photosynthesis. Movile Cave is a small-surfaced, closed chemoautotrophic system [8] driven by in-house sulfur and methane oxidation and CO<sub>2</sub> fixation as primary production processes. The cave, located in south-eastern Romania (Dobrogea region), developed in oolitic and fossil-rich limestone of the Sarmatian age (Late Miocene) and was sealed off during the Quaternary by a thick and impermeable layer of clays and loess [17]. Movile Cave has a complex geological evolution with an ongoing speleogenesis driven by two main processes: the sulfuric acid corrosion in the partially submerged lower cave level; and the condensation-corrosion processes active in the upper level of the cave [18]. The upper gallery (approx. 200 m long) is dry, whereas the lower gallery (approx. 40 m long) is partially flooded by sulfidic hydrothermal waters (T ~ 20.9 °C). The two cave levels converge in the so-called Lake Room (Fig. 1). Air pockets (Air Bells) are present in the submerged gallery (Fig. 1). Here, an active redox interface is created on both the water's surface and cave walls, colonized by floating microbial mats and biofilms. Most of our knowledge of this ecosystem comes from studies of trophic chain components such as microbial mats associated with the surface of the hydrothermal water and its close proximity [19–22]. In addition, the rich endemic invertebrates of Movile Cave live in the proximity of the sulfidic water. The fauna is abundant in the Lake Room [15, 19], particularly at the water surface in the Air-Bells [23], where the redox potential is high.

Except for a snapshot investigation of the rock's surface collected at about 2 m away from the water [24]—which revealed the presence of sulfur oxidizers and methylotrophic bacteria—Movile Cave has not yet been explored for the microbial life associated with sediments distant from the sulfidic waters. Also, no metagenome-assembled genomes (MAGs) reconstruction studies have been carried out, and the genomic information available in databases about this unique environment is limited to a *Ca. Methylomonas* sp. LWB [19] and a *Thiovulum* sp. [22] genome assemblies from the microbial mat.

Considering the unique features of Movile Cave as a model chemoautotrophic subterranean ecosystem and its overlooked microbial communities away from the sulfidic water, we hypothesized that the scarcity of high-energy organic substrates beyond the primary energy source of sulfidic water would strengthen the ecological interactions among community members to optimize the flow of nutrients and ultimately, the energy gain. To address this hypothesis, genome-resolved metagenomics was employed to infer the diversity and abundance of spatially distinct sediment-associated microbiomes and to predict their metabolic abilities, ecological interactions and roles within the cave's ecosystem. Our genome resolved metagenomic study adds 106 higher-quality MAGs to the Movile Cave microbial genomes. The results broaden our understanding of sediment microbial communities' role within the complex food web of Movile Cave, which supports 52 endemic invertebrate species [15].

## Materials and methods

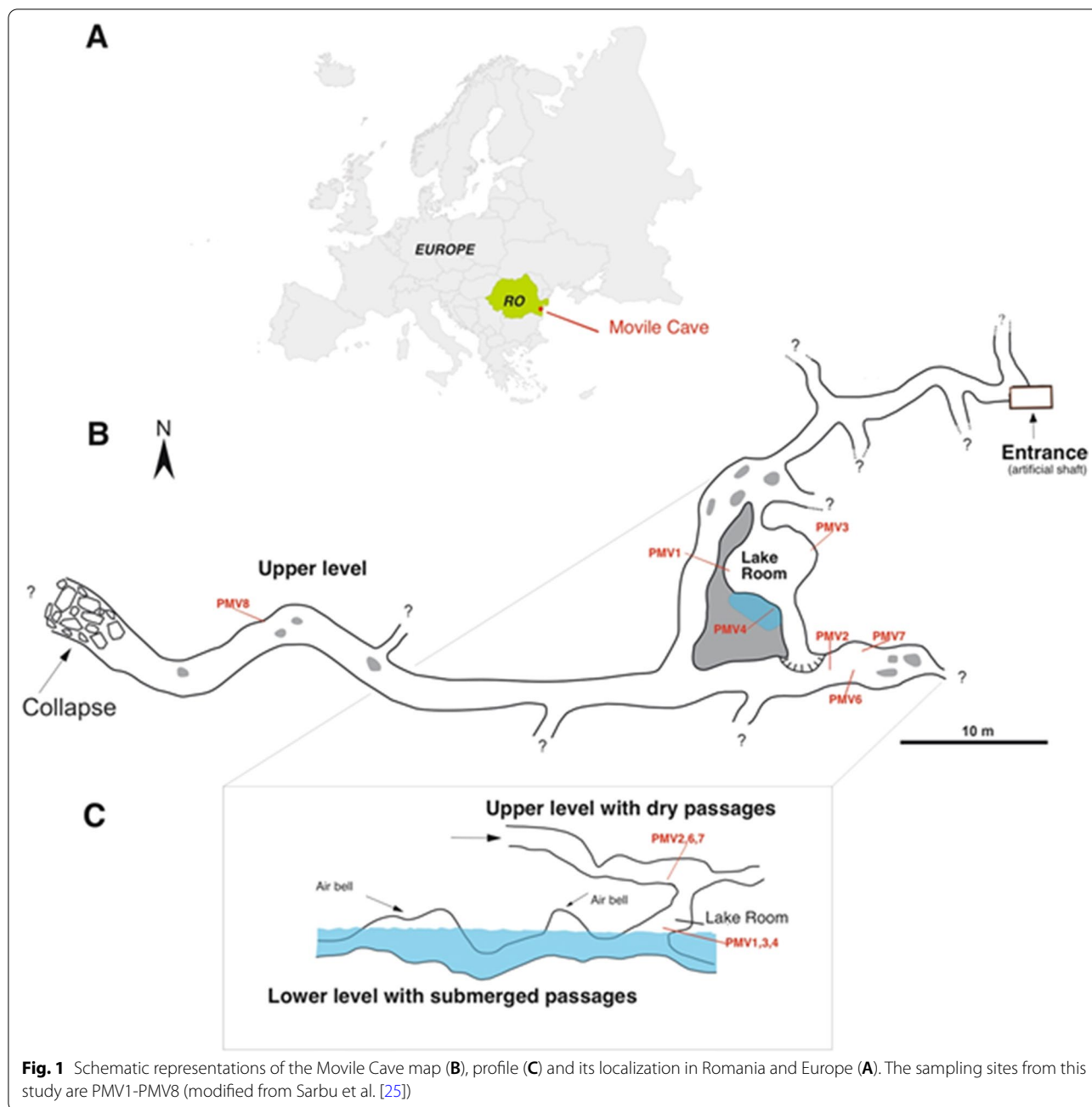
### Study site and sample collection

Sediment samples were collected from seven sites located in the cave's upper and lower levels (galleries) (Fig. 1) in December 2019. Sampling sites were selected based on their physical characteristics (detailed in Table 1 and Additional file 1: Fig. S1) and positions related to the Lake Room. Three sampling sites were near the water, in the Lake Room (PMV1, PMV3, PMV4), and the other four (PMV2, PMV6, PMV7, PMV8) in the upper, dry galleries. The PMV5 sample consisted of sulfidic water and was not the subject of this study. Stable environmental conditions typical to a subterranean habitat combined with hydrothermal activity drive a constant temperature of the water and air at ≈ 20.9 °C within Movile Cave. Aseptically collected top sediments (up to 50 g) were kept at 4 °C after sampling and while transporting, then at – 20 °C until processing.

### Mineralogy and geochemistry measurements

Powdered X-ray diffraction analyses were performed on sediments in order to establish their mineralogy. Samples were analyzed with a Rigaku Ultima IV diffractometer in parallel beam geometry equipped with CuK $\alpha$  radiation (wavelength 1.5406 Å). The XRD patterns were collected in 2 $\theta$  range between 5 to 80 with a speed of 2°/min and a step size of 0.02°. PDXL software from Rigaku, connected to the ICDD database was used for phase identification. The quantitative determination was made using the RIR (Reference Intensity Ratio) methodology.

The pH and electric conductivity were measured in 1:5 solid to water extract using a Seven Excellence (Mettler Toledo) multimeter. The N, C and H were measured on freeze-dried samples using a Flash 2000 (Thermo



Scientific, Thermo Fisher Scientific, Waltham, MA, United States) analyzer. For measurement of metals, samples were digested using a 1:3 mixture of 65% HNO<sub>3</sub> and 37% HCl at reflux conditions. Major elements (Na, K, Ca, Mg, Al, Fe, S, P) in the digested samples were determined by inductively coupled plasma optical emission spectrometry (ICP-OES) using a 5300 Optima DV spectrometer (Perkin Elmer, Shelton, CT, United States), while trace metals and metalloids by inductively coupled plasma mass spectrometry (ICP-MS) using an ELAN

DRC II spectrometer (Perkin Elmer, Waltham, MA, United States). Additional measurements were done for the aqueous fraction of the PMV4 sample. Total nitrogen (TN) was measured in unfiltered water, while dissolved nitrogen (DN) was measured in water samples filtered through 0.45 μm pore size PTFE syringe filters by catalytic combustion and chemiluminescence detection using a Multi N/C 2100 (Analytik Jena, Jena, Germany) analyzer. Total carbon (TC) and total inorganic carbon (TIC) were measured from unfiltered samples, while dissolved

**Table 1** Description of the sites sampled in Movile Cave during December 2019

Sample codes	Sampling sites characteristics	Sediment description	Microclimate
<b>PMV1</b>	<b>Lower gallery, Lake Room, inside a wall diverticulum</b>	<b>Fine grey silt</b>	<b>Air: Humidity=99%; Temp=20.8 °C; CO<sub>2</sub>=17 420 ppm</b>
<i>PMV2</i>	<i>Upper gallery; floor, 50–70 cm away from the wall</i>	<i>Yellowish silt/clay</i>	<i>Air: Humidity=98%; T=20.6 °C; CO<sub>2</sub>=17 010 ppm</i>
<b>PMV3</b>	<b>Lower gallery, Lake Room depressions in the wall</b>	<b>Fine grey silt</b>	<b>Air: Humidity=99%; Temp=0.8 °C; CO<sub>2</sub>= 7 420 ppm</b>
<b>PMV4</b>	<b>Lower gallery, Lake Room; water:sediment interface (water edge)</b>	<b>Fine black silt</b>	<b>Water: pH=7.26; O<sub>2</sub>=0.12 mg/L; EC=1510 µS; TDS=754 ppm; Temp=21.5 °C; Air: Humidity=99%; Temp=20.8 °C; CO<sub>2</sub>=17 420 ppm</b>
<i>PMV6</i>	<i>Upper gallery; floor, 50–70 cm away from the left wall</i>	<i>Coarse yellowish silt/clay</i>	<i>Air: Humidity=98%; T=20.6 °C; CO<sub>2</sub>=17 010 ppm</i>
<i>PMV7</i>	<i>Upper gallery; floor, 50–70 cm away from the right wall</i>	<i>Yellowish silt/clay</i>	<i>Air: Humidity=98%; T=20.6 °C; CO<sub>2</sub>=17 010 ppm</i>
<i>PMV8</i>	<i>Upper gallery; wall diverticulum</i>	<i>Coarse grey silt</i>	<i>Air: Humidity=97%; T=20.8 °C; CO<sub>2</sub>=14 800 ppm</i>

Stations in the lower level are marked in bold, while stations in the upper level are in italic

carbon (DC) and dissolved inorganic carbon (DIC) from water samples filtered through 0.45 µm pore size PTFE syringe filters by catalytic combustion and non-dispersive infrared detection, using a Multi N/C 2100 (Analytik Jena, Jena, Germany) analyzer. Total dissolved solids (TDS) were determined by gravimetric method. Anions (NO<sub>2</sub><sup>-</sup>, NO<sub>3</sub><sup>-</sup>, F<sup>-</sup>, Cl<sup>-</sup>, PO<sub>4</sub><sup>3-</sup>, SO<sub>4</sub><sup>2-</sup>) were measured by ion-chromatography, using a 761 IC compact ion chromatograph (Metrohm, Herisau, Switzerland), equipped with a Metrosep 5–100/4 column and a Metrosep A Supp 4/5 mm guard column after filtration of samples through 0.45 µm cellulose acetate membrane filters. Na, Mg, K and Ca were measured in the filtered samples after acidulation with 65% HNO<sub>3</sub> to a pH < 2 by ICP-OES.

#### DNA extraction and sequencing

Community DNA from sediments was extracted using the DNeasyPowerMax Soil Kit (Qiagen, Gaithersburg, MD, United States); precipitated with 1:10 (v:v) of 3 M Na-Acetate (pH 5.2) and 1:1 (v:v) of 100% ethanol; air-dried and finally resuspended in 100 µL of sterile TE buffer. DNA extracts were quantified with Qubit<sup>TM</sup>dsDNA BR Assay Kit and a Qubit 4.0 fluorometer (Invitrogen, Thermo Fisher Scientific, Waltham, MA, United States). Novogene Europe Co., Ltd (Cambridge, UK) performed the library preparation and shotgun metagenomics sequencing. For library preparation, Illumina protocol NEBNext<sup>®</sup> Ultra<sup>TM</sup>DNA Library Prep Kit (New England BioLabs, Ipswich, MA, United States) was used starting from 200 ng DNA. Sonication was used to randomly fragment the DNA to a size of around 350 bp. Sequencing indexes were added to each sample. The PCR amplification was conducted after the DNA fragments were end-polished, A-tailed, and ligated with the adapter. AMPure XP system was used for PCR product purification. Agilent 2100 Bioanalyzer was used to analyze the size distribution of the DNA library, and qPCR was

performed to quantify DNA according to conventional Illumina library quantification protocol. The samples were sequenced and paired-end reads were generated (PE150) on the Illumina NovaSeq 6000 platform.

#### Reads assembly and annotation

Metagenomic raw reads (fastq) were adapter-trimmed using Trimmomatic v0.39 [26] and, their quality was evaluated using FastQC v0.11.9 [27]. The resulting sequences were assembled into contigs using MEGAHIT v1.2.9 [28], with the following parameters: *no-mercy*, *k-min=31*, *k-max=101*, *k-step=10*. The assembly of reads was done separately, for each metagenome (sample). Assembled contigs (> 1 kb) were binned using two binners: MaxBin v2.0 [29] and MetaBAT2 v2.12.1 [30]. DAS Tool (v1.1.2) was applied to integrate the bins generated from the two methods [31]. Only bins with over 70% completion and less than 10% contamination based on the results of CheckM v1.1.3 [32] were analyzed. Protein-coding sequences (CDS) of selected metagenome-assembled genomes (MAGs) were predicted and annotated with Prodigal v2.6.1 [33] and DIAMOND v2.0.5 [34] against Uniprot databases. Furthermore, the functional annotation of CDS and the prediction of metabolic pathways related to sulfur, nitrogen, carbon, and methane metabolism was carried out using the KEGG's Ghost Koala tool [35]. For confirmation, CDS of interest were identified by blastp against NCBI nr database. In addition to the results obtained by functional annotation, the shotgun preprocessed reads were mapped to a set of reference sequences of nitrogen and carbon fixation hallmark genes with Minimap2 v2.21 with the short reads settings (*-sr* option).

#### Taxonomic assignment and abundance of MAGs

Taxonomic classification of MAGs was assessed using Genome Taxonomy Database Toolkit (GTDB-Tk v1.5.0,

reference data for GTDB R06-RS202) [36] based on 120 bacterial and 122 archaeal marker genes. For phylogenetic analysis, closely related MAGs, reference genomes, or type strains genomes were selected from GTDB (July 2021). The marker genes alignment was concatenated and then imported into R using the *readAA Multiple Alignment* function from the package *Biostrings*. Phylogenetic distances were calculated using the function *dist.alignment* from the package *seqinr*, and a phylogenetic tree was built using the neighbor-joining tree estimation as implemented in the function *nj* in the R package *ape*. The tree was visualized using the online iTOL tool [37].

Finally, for MAGs coverage, the filtered shotgun reads were mapped back to contigs belonging to each MAG with *Minimap2* v2.21 [38], and the coverage was summarized with the *coverage* command from *SAMtools* v1.11 [39]. The relative abundance of each MAG in each metagenome was given by the number of reads mapped per Kb of MAG divided by Gb of corresponding metagenome (RPKG) [40].

#### Genome-scale metabolic reconstruction and interaction models

Genome-scale metabolic models derived from MAGs (metaGEMs) were reconstructed using *CarveMe* v1.5.1 [41], with the default CPLEX solver (v12.8.0.0) (IBM, 2019). Two sets of metaGEMs were reconstructed: one set without gap-filling for any particular medium, based only on genetic evidence; and another set gap-filled for a custom minimal medium that would guarantee the models' growth in a nutrient-poor environment as in Movile Cave. MEMOTE tests suite [42] was applied to the gap-fill models initialized for the minimal medium to ensure that the metaGEMs could generate biomass and reproduce growth in a scarce environment. The custom minimal medium composition is detailed in Additional file 2: Table S1. Community simulations were performed with SMETANA v1.1.0 [43] using the reconstructed metaGEMs. For the simulations, the samples were divided into two groups based on their distance from the sulfidic water: samples from the lower gallery in the Lake Room (PMV1, PMV3, PMV4) and samples from the upper gallery (PMV2, PMV6, PMV7, PMV8). Metabolic interactions were assessed for each sample, and metabolic dependencies were calculated for each condition (lower vs. upper gallery). Global metabolic interactions algorithms MRO (*metabolic resource overlap*) and MIR (*metabolic interaction potential*) were applied for both sets of reconstructed metaGEMs assuming complete or minimal media. MRO/MIR scores were used as a measure of how much the species compete for the same metabolites (MRO) and how many metabolites they can share to decrease their dependency on external

resources (MIP). Only the gap-filled models were used for metabolic dependency, calculated assuming a complete medium. The complete medium composition relies on the BIGG Database [44] (<http://bigg.ucsd.edu/>). By doing so, we imposed the growth of community members on the minimal medium as a constraint. However, we relaxed the community simulation to all possible metabolic exchanges of a complete environment (to identify the metabolic exchanges essential for the community's survival). The obtained SMETANA scores estimate the strength of metabolic coupling (metabolites exchange) in a community. Only a SMETANA score of 1 was considered for analysis. The interactions characterized by a score of 1 are viewed as essential metabolic exchange (essentialities) for the community, and this metabolic coupling is considered a "dependency". SMETANA score was used to highlight the cross-feeding interactions and the keystone MAGs (phyla). Significantly different ( $p < 0.001$ ) compound exchanges across conditions were identified with Wilcoxon rank-sum test using the *compare\_means* function of *ggpubr* package. Model's reconstruction and community simulations parameters and the simulation codes are summarized and detailed in Additional file 2: Tables S2, S3.

The figures were generated using the R packages *ggplot2* [45], *phheatmap* [46], and *circlize* [47].

#### Statistical analysis

We used multidimensional scaling (MDS) to represent the sediment samples in a bi-dimensional space. For the MDS we used a proximity matrix of similarity between the objects (the chemical characteristics of the sediment samples) to the coordinates of these same objects in a 2-dimensional space so that the objects may be visualized easily. Stress is the goodness-of-fit statistic that MDS tries to minimize. Stress varies between 0 and 1, with values near 0 indicating better fit. The analysis was done in XLSTAT 2021.4.1 (Addinsoft, New York).

## Results

### Physicochemical and mineralogical characterization

The sediments had a slightly alkaline pH, in the range of 7.9–9.2, and low (PMV1, PMV7 and PMV8) or high (PMV2, PMV3, PMV4 and PMV6) electrical conductivity ranging from 60.6 to 1738.8  $\mu\text{S cm}^{-1}$  (Table 2). The physicochemical characteristics of the sediment samples are shown in Table 2. Total carbon (TOC) and total nitrogen (TN) content, measured for the aqueous fraction of the sample collected from the water edge (PMV4) were low (TOC = 14.9  $\text{mg L}^{-1}$ , TN = 3.7  $\text{mg L}^{-1}$ ) with a TOC/TN of  $\sim 4 \text{ mg L}^{-1}$  (Additional file 3: Table S4).

In the upper cave gallery, the identified minerals are calcite, dolomite, aragonite, quartz, muscovite, gypsum,

**Table 2** Physicochemical and mineralogical composition of the sediment samples in Movile Cave

Parameter (unit)/Sample	PMV1	PMV2	PMV3	PMV4	PMV6	PMV7	PMV8
pH	8.5	8.3	7.9	8.3	8.1	8.9	9.2
Electrical conductivity ( $\mu\text{S cm}^{-1}$ )	270	751	955	1 237	1 739	119	60.6
N (%SU)	0.21	<0.01	<0.01	1.37	<0.01	<0.01	<0.01
C (%SU)	9.1	4.8	3.9	8.8	8	6.6	4.1
H (%SU)	0.4	0.4	0.8	0.3	0.4	0.4	0.6
Ca ( $\text{mg kg}^{-1}$ )	172 900	125 000	119 100	255 100	159 500	136 900	92 300
Al ( $\text{mg kg}^{-1}$ )	21 450	25 950	19 520	12 635	11 630	29 540	38 490
Mg ( $\text{mg kg}^{-1}$ )	42 960	29 860	6 834	27 590	51 980	36 320	28 120
Fe ( $\text{mg kg}^{-1}$ )	10 160	10 540	21 470	10 030	5 545	10 140	12 780
K ( $\text{mg kg}^{-1}$ )	5 103	4 279	4 983	2 145	2 140	5 816	6 185
Na ( $\text{mg kg}^{-1}$ )	415	289	126	336	564	374	240
S ( $\text{mg kg}^{-1}$ )	902	3 003	13 270	5 730	30 480	1 610	136
P ( $\text{mg kg}^{-1}$ )	162	239	357	110	163	186	124
Mn ( $\text{mg kg}^{-1}$ )	278	216	138	38.5	99.1	184	343
Ba ( $\text{mg kg}^{-1}$ )	86.5	135	60.3	429	87.3	183	110
Ti ( $\text{mg kg}^{-1}$ )	202	246	127	230	149	196	87.8
Sr ( $\text{mg kg}^{-1}$ )	75.4	106	52.2	1060	270	91.6	48.2
V ( $\text{mg kg}^{-1}$ )	29.6	26.9	261	15.5	16.4	24.2	33.6
As ( $\text{mg kg}^{-1}$ )	39.1	24.7	322.5	20.6	24.8	26.4	19.5
Zn ( $\text{mg kg}^{-1}$ )	9.9	14.9	19.4	138	5.7	10.8	17.7
Cr ( $\text{mg kg}^{-1}$ )	14.9	20.4	101	8	4.7	11.9	11.1
Mineralogy (> 5%)	dolomite, quartz, calcite	quartz, calcite, mus- covite	calcite	calcite, quartz, dolomite, mus- covite	dolomite, gypsum, quartz	dolomite, calcite, quartz, muscovite	dolomite, muscovite, aragonite

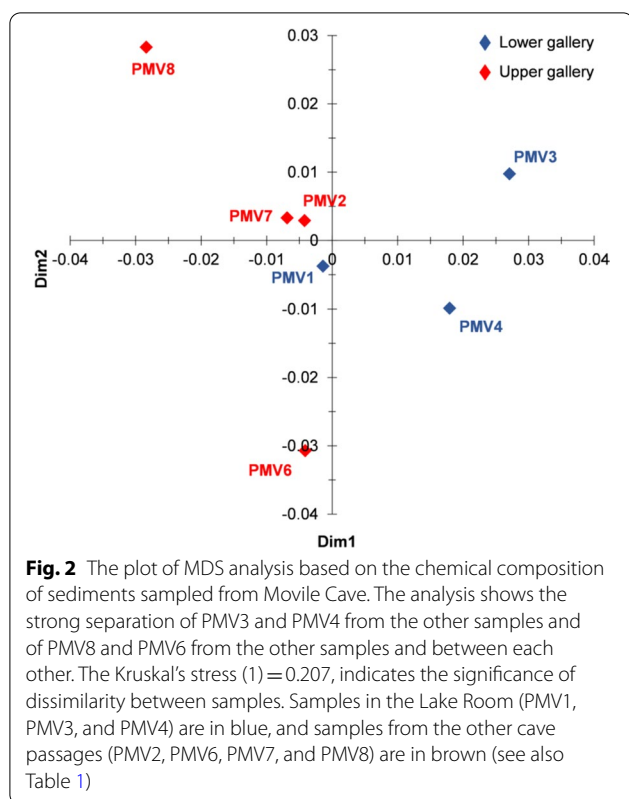
and in small amounts, goethite. Near the Lake, in the lower gallery, the association is dominated by dolomite and calcite, followed by quartz, gypsum and muscovite (Table 2). Carbonates are represented mainly by calcite derived from the limestone bedrock, dolomite and aragonite. Dolomite is present in substantial amounts in both levels of the cave. The bedrock mineralogy reveals the presence of magnesium-rich carbonate, which could be the explanation for the occurrence of aragonite, a signature of PMV8 mineralogy. Samples mineralogy is shown in Table 2 and Additional file 3: Table S5.

A multidimensional scaling (MDS; Fig. 2) analysis shows that PMV6 and PMV8 are more dissimilar from the other samples, PMV6 having the higher electrical conductivity and higher S, Na, Mg, Ca combined with low K, Al and Fe content. The presence of gypsum (sulfate mineral) in that sample (Table 2) explains the high content of S and Ca. In contrast, PMV8 had the lowest electrical conductivity and content in S, Ca (no gypsum in this sample), with a higher concentration of Al, Li, and Na. PMV3 and PMV4 were also separated, while PMV1, PMV2 and PMV7 were more similar along the MDS 1st Dimension. PMV3 had low Na and Mg, and high Fe (the only sample where goethite—iron(III)

oxide-hydroxide—is well represented), V, As, Cr, etc. PMV4 had the highest content of Ca, Ni, and Ba, which probably gives the black color to the sample. The other samples (PMV1, PMV2, and PMV7) were less dissimilar, with muscovite (hydrated phyllosilicate mineral of aluminum and potassium) and quartz ( $\text{SiO}_2$ ) in their mineralogical composition.

#### MAGs recovery and phylogeny

A total of 106 medium- to high-quality MAGs (>70% complete and <10% contamination) [48] were recovered from the sediment metagenomes of Movile Cave. The taxonomic assignment performed using GTDB-tk indicated that the recovered MAGs span 19 bacterial and three archaeal phyla. Out of those, 4 were candidate phyla *Ca. Patescibacteria*, *KSB1*, *Krumholzibacteriota* and *Zixibacteria*. The percentages of reads mapped to MAGs varied among samples from ca. 80% to 13% for MAG recovery efficiency (summarized in Additional file 4: Table S6). The low binning efficiency in some of the samples points toward a very diverse community (feature specific to sediment samples) where the low abundance species were unable to be assembled into MAGs as a consequence of low or partial coverage.



Based on the relative abundance of MAGs in the sample (expressed as RPKG), the most abundant phyla were *Acidobacteriota*, *Chloroflexota*, *Proteobacteria* and *Planctomycetota* (> 20 RPKGs in any of the samples) (Fig. 3A.). In the lower gallery, *Chloroflexota* was abundant in PMV1 and PMV3, while *Proteobacteria* (class *Gammaproteobacteria*) dominated PMV4. In the upper, dry gallery, *Acidobacteriota* was the most abundant phylum for PMV8 and PMV7. For the phyla with a lower relative abundance (< 20 RPKGs), *Ca. Patescibacteria* was the signature phylum for PMV2, *Thermoproteota* for PMV3, *Methylomirabilota* for PMV7 and *Ca. Zixibacteria* for PMV8 (Fig. 3B.).

Out of the 106 MAGs, 25 were classified as high-quality (> 90% complete and < 5% contamination) (Fig. 3C.), but only two were assigned to species level (PMV3.39 and PMV4.117 affiliated to *Chromohalobacter marismortui*) with available genomes in the NCBI database (gANI > 95%). One-third of the MAGs (32 of 106) was classified by Relative Evolutionary Divergence (RED) as new species of established GTDB taxa [49]. Given the taxonomic novelty of recovered MAGs, RefSeq type material genomes (GCF) and representatives MAGs from GeneBank (GCA) were incorporated into the phylogenetic tree construction for a better representation (Fig. 3C.)

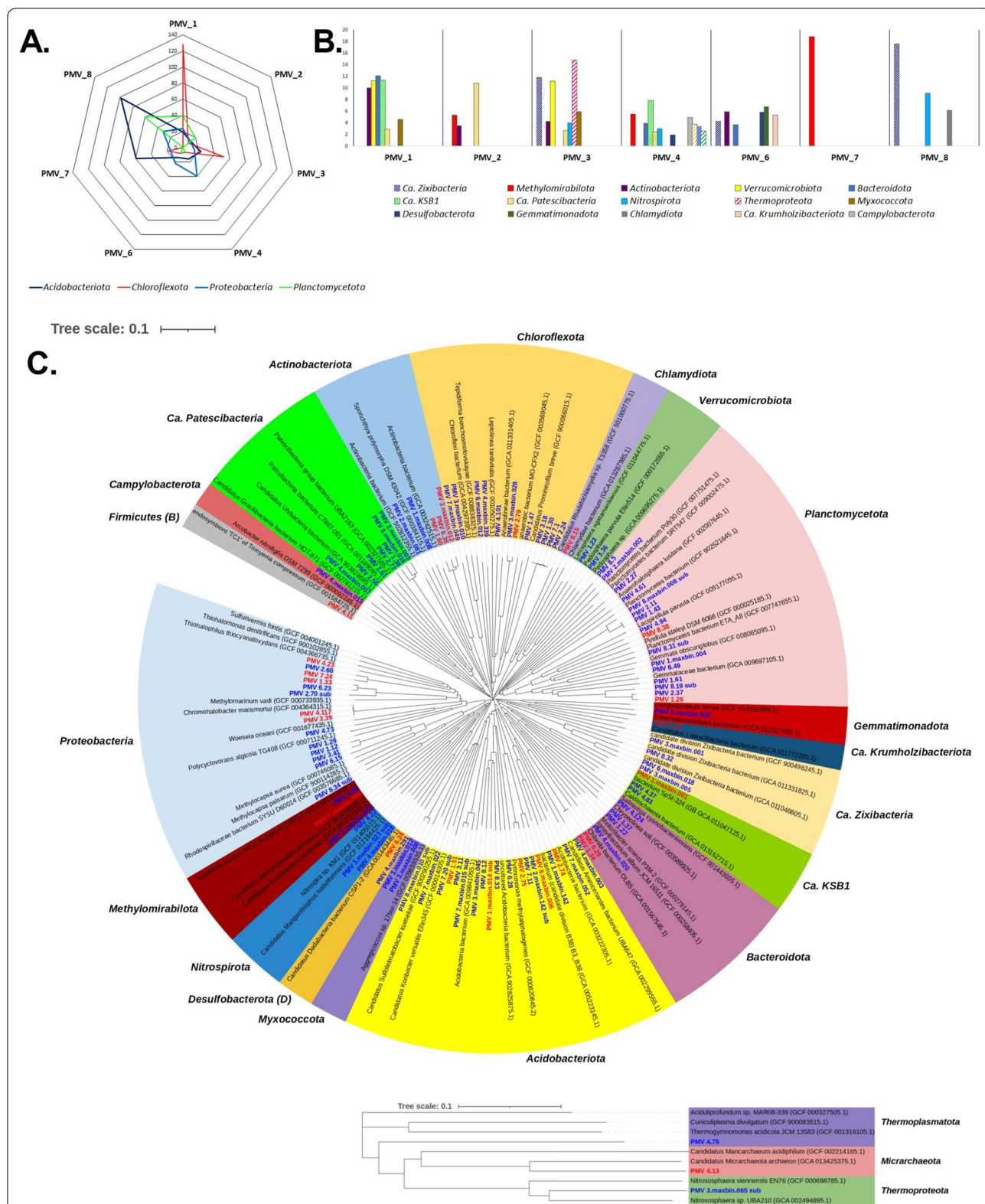
MAGs attributes, extended taxonomy and abundance are summarized in Additional file 4: Tables S7–S9.

### Potential for biogeochemical cycling of S and N, CH<sub>4</sub> oxidation and CO<sub>2</sub> fixation

#### Sulfur metabolism

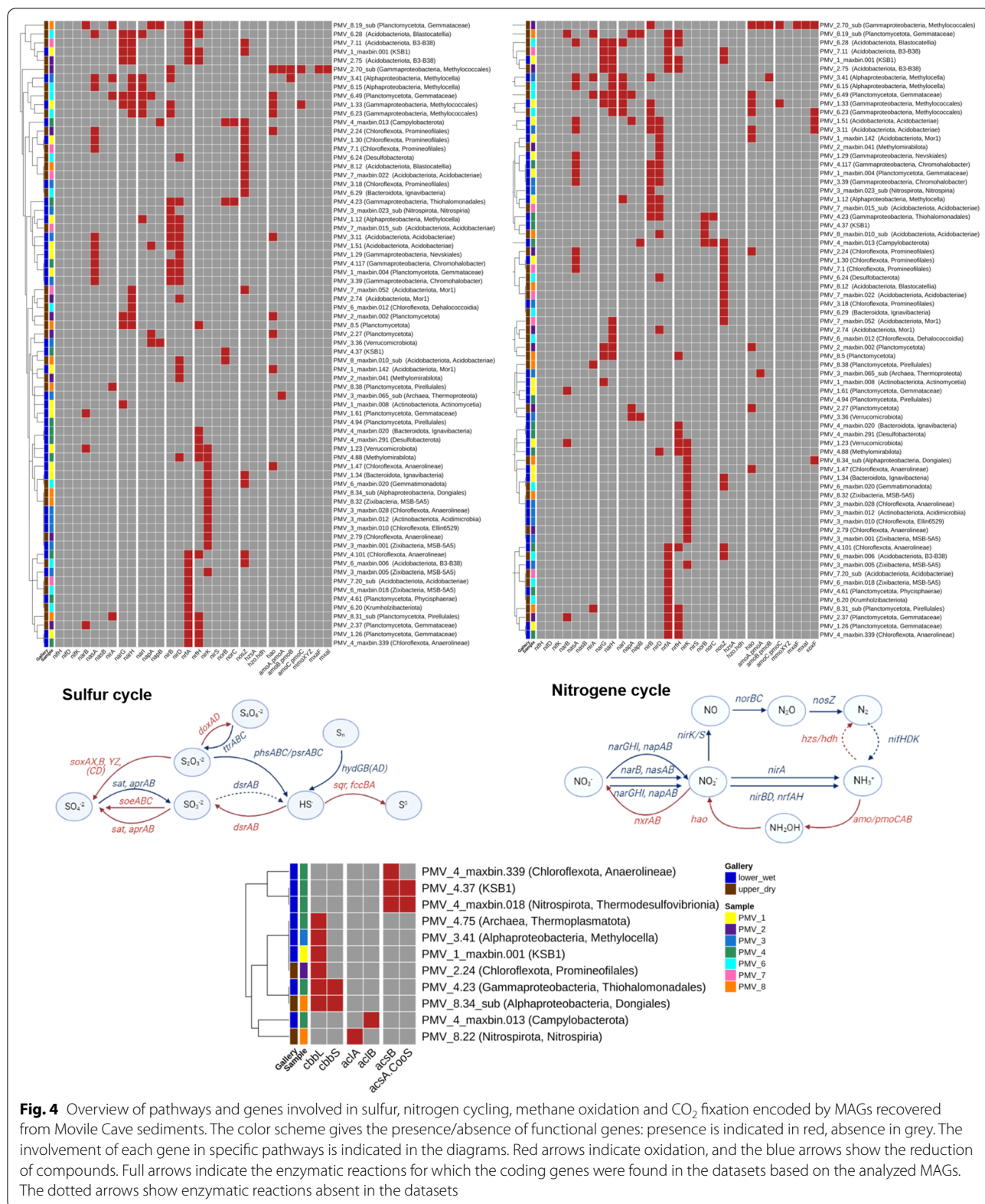
Microbial sulfur cycling has been proposed as a driving force for bacterial proliferation in microbial mats of Movile Cave [24] but not investigated in the cave's sediments. Here, we examined the presence of the marker genes for sulfur cycle in the metagenomes of Movile Cave's sediments. For **sulfur oxidation** (Fig. 4), the complete canonical pathway for thiosulfate ( $S_2O_3^{2-}$ ) to sulfate ( $SO_4^{2-}$ ) conversion (*soxAX*, *B,YZ*, *(CD)*) was annotated in the water edge dataset PMV4; and partially annotated in the upper gallery datasets PMV2 (*soxZ*), PMV6 (*sox(CD)*, *Y*) and PMV7 (*soxYZ*). In the PMV4 dataset, Sox-complex was encoded in MAGs affiliated to order *Thiohalomonadales* (class *Gammaproteobacteria*) (PMV4.23) and family *Arcobacteraceae* (class *Campylobacteria*) (PMV4\_maxbin.013). The *Thiohalomonadales*-affiliated MAG encoded an incomplete thiosulfate-oxidation pathway lacking Sox(CD) and carrying multiple copies of *soxB*. For the sulfide ( $HS^-$ ) oxidation, the MAG encoded both the pathway to elemental sulfur ( $S^0$ ) via flavocytochrome c sulfide dehydrogenase (*fccAB*), and to sulfate ( $SO_4^{2-}$ ) via the reverse-operating dissimilatory sulfite reductase pathway (*dsrAB* + *aprAB* + *sat*). The complete path for sulfite ( $SO_3^{2-}$ ) oxidation to sulfate ( $SO_4^{2-}$ ) (*soeABC*) was also annotated here. This *Thiohalomonadales* MAG also encoded (*hydBD*) the production of  $HS^-$  from polysulfides by sulphydrogenase complex (*hyd(G,B),(A,D)*).

In the rest of the datasets, Sox(YZ) subunits were encoded in *Proteobacteria* MAGs and Sox(CD) in a *Gemmatimonadota* MAG (PMV6\_maxbin.020). Thiosulfate oxidation with tetrathionate ( $S_4O_6^{2-}$ ) formation was partially encoded (*doxD*) in all datasets except for PMV8, and mostly in bins affiliated to *Chloroflexota*. Sulfide:quinone reductase (*sqr*) for the oxidation of  $HS^-$  with the zero-valent sulfur  $S^0$  formation was encoded in all datasets. The oxidation of sulfide via SQR was predicted as a widespread trait encoded in 27 MAGs. The pathway for sulfite ( $SO_3^{2-}$ ) oxidation to sulfate ( $SO_4^{2-}$ ) was also partially encoded (*soeB*) in PMV6 and PMV8 datasets. Sulfur respiration, dissimilatory **reduction of oxidized sulfur** compounds ( $SO_4^{2-}$ ,  $SO_3^{2-}$ ,  $S_2O_3^{2-}$ ,  $S^0$ ) coupled with sulfide ( $HS^-$ ) production, were also detected. Sulfate ( $SO_4^{2-}$ ) to sulfite ( $SO_3^{2-}$ ) reduction (*sat* + *aprAB*) was present in the PMV4 dataset, in a MAG affiliated to order *Thermodesulfobivibrionales* (phylum *Nitrospirota*) (PMV4\_maxbin.018). An NCBI blast (blastx against RefSeq Select database) showed the AprAB sequences similarities to



**Fig. 3** The abundance and phylogeny of MAGs recovered from sediments of Movile Cave. A. Most abundant phyla (relative abundance > 20 RPKGs). B. Less abundant phyla (relative abundance < 20 RPKGs). C. Phylogenetic tree of MAGs from Movile Cave sediments, including their closely related MAGs from GTDB (GCA) and NCBI type material genomes (GCF) (type strain and/or reference genomes). The neighbor-joining phylogenetic tree was constructed based on the GTDB marker genes. The MAGs detected in this study are shown in blue or red for medium- or high-quality MAGs, respectively





**Fig. 4** Overview of pathways and genes involved in sulfur, nitrogen cycling, methane oxidation and CO<sub>2</sub> fixation encoded by MAGs recovered from Mobile Cave sediments. The color scheme gives the presence/absence of functional genes: presence is indicated in red, absence in grey. The involvement of each gene in specific pathways is indicated in the diagrams. Red arrows indicate oxidation, and the blue arrows show the reduction of compounds. Full arrows indicate the enzymatic reactions for which the coding genes were found in the datasets based on the analyzed MAGs. The dotted arrows show enzymatic reactions absent in the datasets

*Deltaproteobacteria* ( $\approx 73\%$  to  $80\%$ ), placing them in the direct-operating dissimilatory pathway.

Thiosulfate/elemental sulfur ( $S_2O_3^{2-}/S^0$ ) disproportionation and tetrathionate respiration were partially encoded (*phsA*, *phsC*; *ttrA/B*) in different phylogenetic groups, but a *Planctomycetota*-affiliated MAG (PMV1.61) that might encode both as *ttrB* and *phsA* were present in the bin. The production of hydrogen sulfide from polysulfides by sulfhydrogenase complex (*hyd(G,B),(A,D)*) was partially encoded, especially by MAGs affiliated to *Acidobacteriota*, *Chloroflexota* and *Planctomycetota*.

### Nitrogen metabolism

Analogous to sulfur, we addressed the potential for biogeochemical N cycling across sediments of Movile Cave. Primary producers capable of obtaining energy for autotrophy by **nitrogen oxidation** (nitrification) were questioned. The first step of *nitrification*, the ammonia  $NH_3^+$  to nitrite  $NO_2^-$  oxidation, was suggested in Movile Cave sediments as ammonia monooxygenase AMO (*amoABC*) was encoded in PMV1 to PMV3 datasets. All three subunits of AMO's were encoded in an order *Methylococcales*-affiliated MAG (PMV2.70\_sub). However, the *amoABC* sequences were homologous ( $>65\%$  similarity) to the sequence encoding particulate methane monooxygenase pMMO in the genus *Methyloterricola*. The methanotrophs can oxidize both substrates ( $NH_3$  and  $CH_4$ ) but grow only on their characteristic substrate [50]. Interestingly, AmoA-like subunit was annotated in an archaeon MAG affiliated to *Nitrososphaera* genus (phylum *Thermoproteota*) (PMV3\_maxbin.65\_sub). For the second step of nitrification, the oxidation of  $NO_2^-$  to  $NO_3^-$ , none of the bins encoded the nitrite oxidoreductase (*nxrAB*) found in known nitrite-oxidizing genera. Also, no marker genes (*hzsA* and *hzo*) for the anaerobic ammonium oxidation (*anammox*) ( $NH_3$  to  $N_2H_4$  and then to  $N_2$ ) were annotated in the Movile datasets.

In the chemoautotrophic cave environment, the nitrogen demand can be met by converting inorganic nitrogen to a biologically useful form by **nitrogen reduction**. The conversion involves microbial *dinitrogen fixation* or *nitrate assimilation*. Surprisingly, the  $N_2$  fixation pathway was not detected in any sediment datasets, as no nitrogenase genes (*nif*) for the reduction of atmospheric molecular nitrogen ( $N_2$ ) to ammonia ( $NH_3^+$ ) was annotated in MAGs. This result was reinforced by mapping the metagenomics reads against a set of *nif* reference sequences (listed in Additional file 5: Table S10) as no *nifH/D* were found in the PMV datasets except PMV4 (Additional file 5: Fig. S2). The assimilatory process ( $NO_3^-$ ,  $NO_2^-$  conversion to  $NH_3^+$ ), catalyzed by nitrate and nitrite reductases (*nasA/B*, *narB*; *nirA*), was fully encoded in PMV3, PMV6, and PMV8 datasets. The bins

that encoded the nitrate-assimilatory enzymes were taxonomically diverse as the ability is however widely distributed among bacteria.

The first step in the *dissimilatory nitrate reduction to ammonium* (DNRA) and *denitrification*,  $NO_3^-$  to  $NO_2^-$  reduction, was fully encoded as the cytosolic nitrate reductases (*narGHI*) (PMV1, PMV6) or as the periplasmic nitrate reductases (*napAB*) (PMV3, PMV8). The  $NO_2^-$  to  $NH_4^+$  step of DNRA was encoded in all datasets as the cytoplasmic nitrite reductase (*nirBD*) or the periplasmic nitrite reductase complex (*nrfAH*). MAGs that encoded at least partially both steps of the DNRA pathway were affiliated to *Acidobacteriota* (PMV1.51 (*Nap-Nir*); PMV2.75 (*Nar-Nrf*); PMV7.11 (*Nar-Nrf*)), *Planctomycetota* (PMV8.19\_sub (*Nap-Nrf*); PMV8.5 (*Nar-Nrf*)) and *Gammaproteobacteria* (PMV6.23 (*Nar-Nir*)). Nitrite reductase (*nirK*), the hallmark enzyme of denitrification ( $NO_2^-$  to (NO) conversion) was encoded in all datasets except PMV7, in bins taxonomically unrelated with common denitrifiers, including *Ca. Zixibacteria* (PMV8.32) and *Ca. MethyloMirabilis* (PMV4.88). The conversion of NO to nitrous oxide ( $N_2O$ ) was fully encoded as nitric oxide reductase (*norBC*) in PMV4 in the sulfur oxidizing *Thiohalomonadales* MAG (PMV4.23). The last step of denitrification, conversion of  $N_2O$  to  $N_2$ , carried out by an oxygen-sensitive enzyme, the nitrous oxide reductase (*nosZ*), was annotated in all data sets and across taxonomically diverse bins.

### Methane metabolism

Via **methane oxidation**, methanotrophs can metabolize methane as their source of carbon and energy. In the oxidation pathway, methane ( $CH_4$ ) is oxidized to methanol ( $CH_3OH$ ) and then to formaldehyde ( $CH_2O$ ) which is incorporated into organic compounds via the serine or the ribulose monophosphate (RuMP) pathway. As previously mentioned, all subunits of the particulate methane monooxygenase pMMO (*pmoABC*) were predicted in the order *Methylococcales* MAG (PMV2.70\_sub). Separated subunits of pMMO were also annotated in MAGs affiliated to *Methylococcales* order (PMV1.33) and *Methylocella* genus (PMV3.41). No soluble methane monooxygenase (sMMO) was annotated across investigated metagenomes. The methanol oxidation to formaldehyde appeared encoded as  $Ln^{3+}$ -dependent methanol dehydrogenases (*xoxF*) and heterotetrameric methanol dehydrogenase (*mxoFI*) in the *Methylococcales*-affiliated MAG PMV2.70\_sub. Only XoxF methanol dehydrogenase was encoded in MAGs affiliated to the classes *Gammaproteobacteria* (PMV6.23), *Alphaproteobacteria* (order *Dongiales*) (PMV8.34\_sub) and *Acidobacteriae* (PMV1.51, PMV3.11).

**Methanogenesis** potential was absent in the Movile sediment datasets as suggested by the absence of methyl coenzyme-M reductase (*mcrA*) marker gene across investigated MAGs.

### CO<sub>2</sub> fixation

The CO<sub>2</sub> fixation is presumably critical in a chemoautotrophic ecosystem such as Movile Cave. To verify the autotrophic potential of the sediment communities, the presence of genes for the key enzymes of CO<sub>2</sub> fixation pathways, namely the Calvin–Benson–Bassham (CBB) cycle (key enzyme: ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO), genes *cbbL* and *cbbS*), the reverse tricarboxylic acid (rTCA) cycle (key enzyme: ATP citrate lyase, genes *aclA* and *aclB*) and the reductive acetyl-CoA, or Wood-Ljungdahl (WL) pathway (key enzyme: CO dehydrogenase/acetyl-CoA synthase (CODH/ACS complex), genes *acsA* and *acsB*) was investigated.

No carbon fixation key enzyme was encoded by MAGs from PMV6 and PMV7 datasets. RuBisCO large subunit (*cbbL*) (CBB cycle) was predicted in all other datasets in MAGs affiliated to *Ca. KSB1* (PMV1\_maxbin.001), *Chloroflexota* (PMV2.24), *Micrarchaeota* (PMV4.13) phyla and classes *Alphaproteobacteria* (genus *Methylocella* (PMV3.41) and order *Dongiales* (PMV8.34\_sub)) and *Gammaproteobacteria* (order *Thiohalomonadales*) (PMV4.23). The *Thiohalomonadales* (PMV4.23) and the *Dongiales* (PMV8.34\_sub) MAGs encoded both RuBisCO subunits (*cbbL/S*). Most of the enzymes of the rTCA cycle are shared with the TCA cycle, except for the ATP citrate lyase (ACL) used by autotrophic prokaryotes for the conversion of citrate to acetyl-CoA. The rTCA cycle hallmark enzyme subunits were annotated in MAGs affiliated with phyla *Campylobacterota* (PMV4\_maxbin.013) and *Nitrospirota* (PMV8.22).

For the anaerobic carbon fixation via WL pathway, the marker genes of CODH/ACS complex, *acsAB*, were found only in PMV4 dataset. The Western branch (carbonyl) was encoded by MAGs affiliated to phyla *Nitrospirota*, (order *Thermodesulfovibrionales*) (PMV4\_maxbin.018) (*AcsA* (β), *AcsB* (α), *AcsD* (δ), *AcsC* (γ)) and *Ca. KSB1* (PMV4.37) (*AcsA* (β), *AcsB* (α), *AcsC* (γ)). The Eastern branch (methyl), for CO<sub>2</sub> conversion to 5-methyltetrahydrofolate, was also encoded in the *Nitrospirota*-affiliated MAG. The *Ca. KSB1*-affiliated MAG encoded for methylenetetrahydrofolate dehydrogenase (*folD*) of WL pathway Eastern branch as well as the acetyl-CoA to acetate enzymes: acetate kinase (*ackA*), phosphate acetyltransferase (*pta*) and a putative phosphotransacetylase. This suggests that the *Ca. KSB1* MAG could be unable to gain ATP from acetyl-CoA degradation. Enzymes for the

Eastern branch (methyl) of WL pathway were encoded in all sediment datasets.

In addition to MAG analysis, mapping the metagenomics reads against a set of reference CO<sub>2</sub> fixation gene sequences (listed in Additional file 5: Table S10) highlighted the CBB cycle key genes as present in all datasets, including PMV6 and PMV7, and the WL pathway marker genes as hallmark genes for only PMV4 dataset (Additional file 5: Fig. S2).

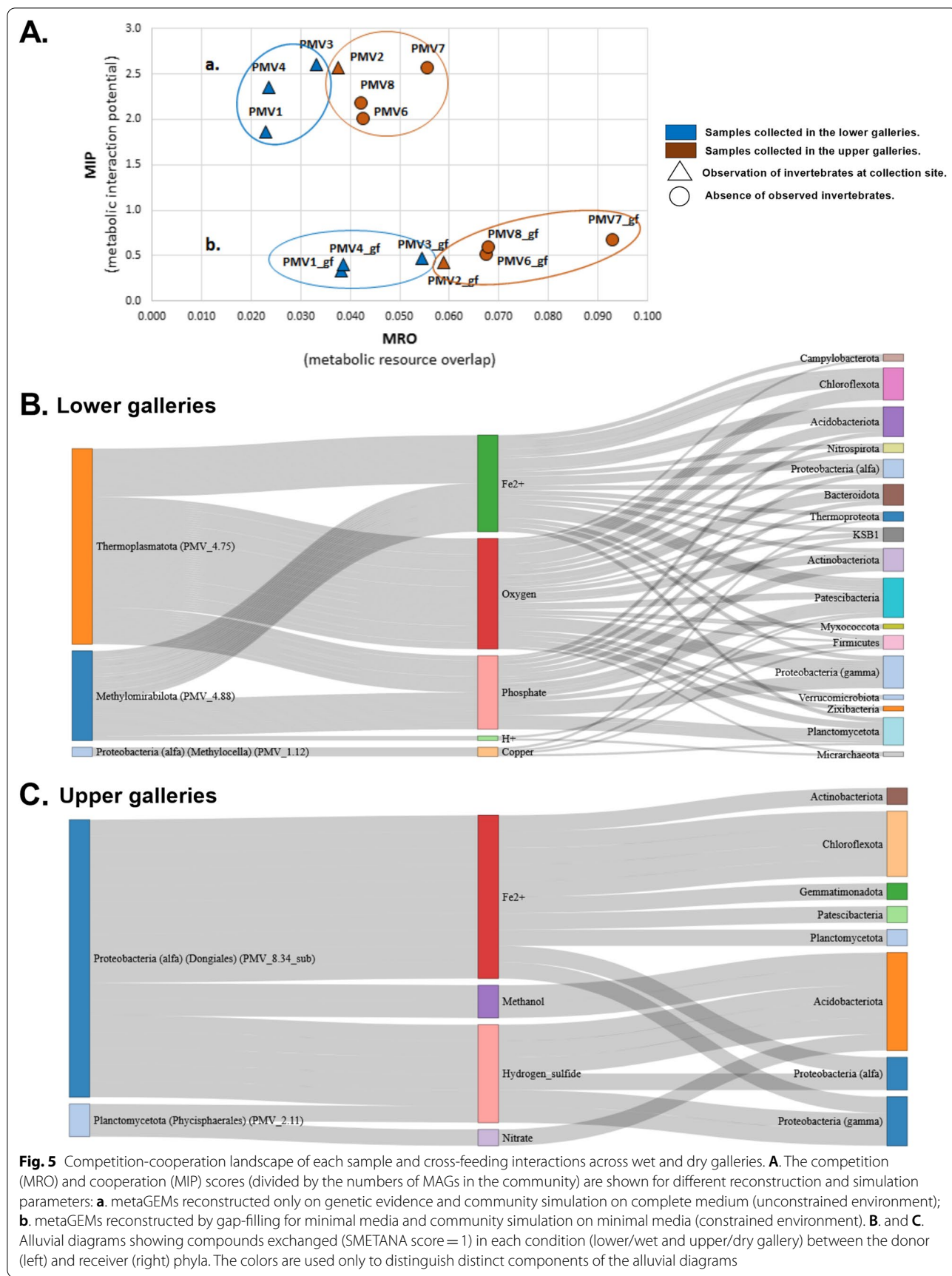
Interestingly, none of the genes of interest for biogeochemical cycling of S and N, CH<sub>4</sub> oxidation and CO<sub>2</sub> fixation were annotated in MAGs assigned to *Ca. Patescibacteria* or *Myxococcota* phyla.

All the functional genes implicated in the sulfur and nitrogen cycling, methane oxidation and CO<sub>2</sub> fixation annotated in MAGs from Movile Cave sediment are listed in Additional file 6: Tables S11–S14 and an overview of the genes, encoding MAGs and pathways are shown in Fig. 4.

### Potential metabolic interactions and dependencies within the communities

To gain a deeper insight into the microbial metabolic webs in Movile Cave sediments, we used the MAGs to construct metagenome-scale metabolic models (metaGEMs) for simulation and prediction of potential metabolic interactions and dependencies within the communities. Noteworthy, no metabolic models could be constructed for three MAGs belonging to *Ca. Patescibacteria* (PMV2.56, 2.61 and 2.72) thus were excluded from the analysis. The inability to construct metabolic models for MAGs belonging to *Ca. Patescibacteria* can be a consequence of the size of their reduced genome (usually < 1 Mb). The rest of the metaGEMs could generate biomass and reproduce growth in the simulation conditions, as evidenced by MEMOTE tests results included in the Additional file 7: Table S15. The competition-cooperation potential predicted by SMETANA (Fig. 5A.; Additional file 7: Table S16) was highest in the communities associated with the upper gallery (e.g. PMV7). This was evident, especially in the case of the necessary resources overlap (competition). The pattern persisted regardless of metaGEMs reconstruction or community simulation parameters. When community simulation assumed minimum nutrient availability, as presumed for the Movile Cave environment, the community members of the upper gallery (samples PMV6 to PMV8) exhibited the highest similarities of metabolic requirements (competition) and highest potential for community self-sufficiency (cooperation) (Fig. 5A.b.).

For an insightful analysis of the microbial metabolic webs in the Movile Cave sediments, we grouped the samples based on distance, namely samples from the lower,



'wet' gallery (PMV1, PMV3, PMV4) and samples from the upper, 'dry' gallery (PMV2, PMV6, PMV7, PMV8). We examined the cross-feeding interactions and keystone phyla across those two conditions ('wet' and 'dry'). Our simulation results (SMETANA score = 1) suggested that only 'currency' molecules and inorganic ions ( $O_2$ ,  $PO_4^{3-}$ ,  $H^+$ ,  $Fe^{2+}$ ,  $Cu^{2+}$ ) were exchanged in the lower gallery. Here, the donor MAGs belonged to archaea *Ca. Thermoplasmatota* (PMV4.75), bacteria candidate division NC10 (*Methyloirabilota*) (PMV4.88) and the bacteria genus *Methylocella* (PMV1.12) (Fig. 5B.). In the upper gallery, besides inorganic ions ( $Fe^{2+}$ ,  $NO_3^-$ ), hydrogen sulfide ( $H_2S$ ) and methanol ( $CH_4O$ ) were also exchanged. Here, the essential donors were affiliated to the *Alphaproteobacteria* within order *Dongiales* (PMV8.34\_sub) and *Planctomycetota* within order *Physcisphaerales* (PMV2.11) (Fig. 5C.). For both conditions, the taxonomic affiliation of receiver MAGs was diverse, represented by MAGs belonging to 8 and 17 phyla. The MAGs that interact in each simulation condition and their role as donors or recipients of exchanged metabolites are listed in Additional file 7: Tables S17 and S18.

Out of the predicted essential compounds (SMETANA score = 1) that are readily exchanged among MAGs,  $H_2S$ ,  $NO_3^-$ , methanol, and  $H^+$  exchanges were significantly different (Wilcoxon rank-sum test, BH adjusted p-values < 0.001) across upper and lower galleries (Additional file 7: Table S19).

## Discussions

Our current understanding of microbial life within the Movile Cave ecosystem was limited to the hydrothermal waters, as only microbial mats, water samples and lake sediments were previously investigated [18, 19]. Our study gives a first insight into the cave's chemoautotrophic and primary production potential beyond the hydrothermal waters, namely the cave's sediments.

Based on physicochemical and mineralogical characterization, sampled sediments showed a very different chemical composition, even if located just meters away from one another. PMV6 and PMV8 differed the most from the other samples, followed by PMV3 and PMV4. Mineralogy partially supports such chemical discrepancies, but some of the best-represented elements are of different, unknown origins, probably linked to the history of the region where the cave is located and the input of the underground sediments from the Miocene to the more recent climatic events.

The taxonomic novelty of recovered MAGs was high, with over 60% unaffiliated to a genus and 30% classified by RED as new species of established GTDB taxa. The community was diverse with 22 microbial phyla, mainly *Acidobacteriota*, *Chloroflexota*, *Proteobacteria*,

*Planctomycetota*, *Ca. Patescibacteria*, *Thermoproteota*, *Methyloirabilota*, and *Ca. Zixibacteria* based on the relative abundance of MAGs in samples. Despite the high novelty at lower taxonomic levels, the overall phyla diversity of sediment MAGs was typical of sulfidic and non-sulfidic cave microbiology [51–53]. Something worth noticing is the lack of *Betaproteobacteria* among recovered MAGs, since it was previously demonstrated for the Movile Cave ecosystem [16, 21].

Functional analyses suggested the presence of chemolithoautotrophic primary producers in the cave sediments from both galleries. As expected, sulfur oxidation, proposed as the ecosystem's driving force, was mainly present in the lakeside dataset PMV4. The complete oxidation pathways (Sox and Dsr pathways), RuBisCO and partial nitrate/nitrite respiration and denitrification were encoded by the *Thiohalomonadales*-affiliated MAG (class *Gammaproteobacteria*). As *Thiohalomonas* members are obligate chemolithoautotrophic facultative anaerobic sulfur-oxidizing bacteria [54], this MAG might play a key role as a primary producer within the ecosystem. It uses reduced sulfur compounds as energy sources, nitrate as an electron acceptor, and assimilating  $CO_2$  via the Calvin-Benson-Bassham cycle. Another primary producer inferred from the lake edge sample was an *Arcobacteraceae*-related MAG (phylum *Campylobacterota/Epsilonproteobacteria*) that encoded the capacity for sulfur oxidation (Sox) potentially coupled with nitrate-reduction (NapB) and  $CO_2$  fixation via the reverse TCA cycle (*aclAB*). Those are typical metabolic features of *Epsilonproteobacteria* chemolithotrophic primary producers from the deep-sea hydrothermal vent [55]. Similar metabolic traits were found for the recently characterized *Thiovulum* sp. that dominates Movile Cave's hypoxic Air Bells microbial mat [22]. Moreover, *Arcobacteraceae* family was found dominant in the white filaments from the thermal sulfidic spring of Fetida Cave, Italy [56]. Sulfur oxidation was not, however, limited to the lakeside sediments. Genes coding for Sox pathway (*soxCD*) were found in PMV2, 6 and 7 datasets of the upper, dry gallery across MAGs affiliated to *Alpha-*, *Gammaproteobacteria* and *Gemmatimonadota* phylum. The *Gemmatimonadota* role in the sulfur cycle is uncertain, but MAGs with similar attributes were identified in the Siberian soda lakes [6]. In the sediments from around the lake, sulfur respiration was also evident, which is less critical in self-sustaining ecosystems and a characteristic of heterotrophs. In addition to previous reports [16, 21] describing *Deltaproteobacteria* sulfate reducers in Movile Cave water and microbial mat samples, we assembled a hypothetical thermophilic sulfate reducer MAG affiliated to *Nitrospirota* (order *Thermodesulfobivibrionales*) possessing DSR pathway genes (*aprAB*). These results might indicate that

sulfate reduction is important in and around the sulfidic water but not away from it.

Considering the relatively high (0.2–0.3 mM) ammonium concentrations in the cave waters [57], it was formerly implied that nitrogen oxidation might also be a driving force for the ecosystem [21]. A few *Nitrospirota*-affiliated MAGs were identified in the upper and lower galleries, but none of the nitrite or ammonia-oxidation genes were annotated. No nitrifying bacteria were identified based on gene annotation, but an ammonia-oxidizing *Nitrososphaera* archaeon encoding for an AmoA-like subunit was assembled from the Lake Room PMV3 sediments. *Nitrososphaera* archaea are facultative chemolithoautotrophs with potential metabolic flexibility [58, 59]. Our metabolic annotation results suggest that the nitrogen cycle throughout the Movile Cave sediments is mostly driven by nitrate/nitrite respiration and denitrification. Those nitrogen reduction processes are linked to taxonomically diverse bacteria and spread among all investigated locations. Assimilatory and dissimilatory nitrate reduction to ammonium were predicted in all datasets, emphasizing the ability of the microbial communities to provide bioavailable N for the other trophic links in the Movile Cave ecosystem. Methanotrophs were also postulated as primary producers in the water and microbial mats [19–21, 24, 60] as 1–2% methane concentration was highlighted, especially in the Air-Bells [57]. In this study we were able to assemble methanotrophic MAGs from both lower and upper galleries. These were affiliated to the uncultured UBA1147 genus of *Methylococcales* (in PMV1, PMV2, PMV6) and *Methylocella* genus of *Rhizobiales* (in PMV1, PMV3, PMV6) and encoded for subunits of pMMO monooxygenase. Although no methane-oxidation gene could be annotated, *Methylomirabilales*-affiliated (candidate division NC10) MAGs were found in PMV2 (order *Rokubacteriales*), PMV4 and PMV7 (order *Methylomirabilales*) datasets. Interestingly *Methylomirabilales*-affiliated MAG was one of the most abundant MAGs in PMV7. The presence of *Methylococcales* (*Methylococcus*, *Methylomonas*, *Methylocaldum*) and *Rhizobiales* (*Methylocystis*/*Methylosinus*, *Methylocella*) methanotrophs was previously postulated in the microbial mats based on pMMOA-sequence clones from a CH<sub>4</sub>-enriched culture [19, 20]. The genus *Methylocella* comprises facultative methanotrophs that utilize multicarbon compounds (acetate, pyruvate, succinate, malate, and ethanol) [61]. While the order *Rokubacteriales* might comprise non-methanotrophic bacteria [62], *Methylomirabilales* are known autotrophic (CO<sub>2</sub> fixing via CBB cycle), denitrifying methanotrophs that can oxidize methane anaerobically. The molecular oxygen needed for methane oxidation is generated by reducing nitrate to dinitrogen gas and O<sub>2</sub> and bypassing the

nitrous oxide formation [63–65]. This is considered the fourth biological pathway known to produce oxygen besides photosynthesis, chlorate respiration, and the detoxification of reactive oxygen species [63].

The key enzymes for CO<sub>2</sub> fixation were found in 5 out of 7 sediment samples (except PMV6, PMV7), including PMV8 collected farthest from the lake. Noteworthy, in PMV8 MAGs both the CBB and rTCA enzymes were annotated. WL pathway of anaerobic CO<sub>2</sub> fixation was postulated only in the water edge dataset PMV4. A sulfur-respiring (*aprAB*) MAG affiliated to the order *Thermodesulfovibrionales* (phylum *Nitrospirota*) encoded for the WL pathway key enzymes (*acsAB*). However, the *Thermodesulfovibrionales* order includes members (genus *Thermodesulfovibrio*) that are chemoorganotroph, fermentative and dissimilatory sulfate-reducing bacteria [66]. Therefore, the MAG encodes the WL pathway in reverse to break down acetate to CO<sub>2</sub> and H<sub>2</sub>. Uncultured *Nitrospirota* MAG from PMV8 dataset encoded key genes for rTCA cycle, confirming early evidence revealed by stable-isotope probing (SIP) of water and microbial mat samples that *Nitrospirota* might be able of CO<sub>2</sub> fixation [21]. CBB cycle key enzymes were detected in most sediment samples, except for PMV6 and PMV7. The archaeal MAG encoding CBB capability found in the water's edge sample (PMV4) pertains to *Ca. Thermoplasmatota*'s class EX4484-6 that includes MAGs retrieved from marine hydrothermal vent sediments (BioProject PRJNA362212). Thermoplasmata members were also identified in the snottites (thick snot-like biofilm) in the Frasassi caves and in the gypsum moonmilk of Fetida Cave, both in Italy [51, 67, 68]. The presence of RuBisCO (*cbbL*) and sulfhydrogenase (*hydBG*) subunits in the genome may point toward an autotrophic organism that couples sulfur (S<sup>0</sup>, S<sub>n</sub>) reduction to hydrogen oxidation. Among the bacteria encoding RuBisCO, a MAG affiliated to the order *Dongiales* from the upper gallery PMV8 stands out. GTDB's *Dongiales* order comprises the members of the formerly known family *Rhodospirillaceae*. The RuBisCO enzyme presence and the lack of genes for the photosynthetic reaction center (*pufLM*) are intriguing for *Rhodospirillaceae*-like bacteria. *Rhodospirillaceae* are known as photoautotrophic, photoheterotrophic, and chemoheterotrophic bacteria [69], but not as chemoautotrophs. Since the sulfite dehydrogenase (*soeA*) subunit was also annotated, the *Dongiales*-affiliated MAG may be a chemolithoautotrophic bacterium that fix CO<sub>2</sub> and uses sulfur compounds as electron sources. Another interesting phylum encoding autotrophic features was *Ca. KSB1* phylum, which currently consists of a single class termed UBA2214. The two *Ca. KSB1*-related MAGs recovered in this study encode oxic- (CBB) and anoxic (WL) CO<sub>2</sub> fixation potential. Both MAGs also seem to carry nitrate

reduction by DNRA and denitrification, which can serve as a nitrogen retainer (DNRA) and a nitrogen remover (denitrification) in the environment.

N<sub>2</sub> and CO<sub>2</sub> fixation prediction obtained by MAGs functional analysis was mostly reinforced by metagenomic read mapping on the marker genes (Additional file 5: Figs. S2, S3). This reinforcement could also be considered a validation of the obtained MAGs-based metabolic prediction even if a low binning efficiency (percentages of reads mapped to MAGs) was the case for some samples.

The metabolic interactions were investigated to extend our understanding of the metabolic potential in sediment-associated microbial communities. Microbial communities in the lower, wet and upper dry galleries were analyzed using MAG-based metabolic models (metaGEM). The community metabolic modeling approach using metaGEM reconstruction and in silico simulation is only recently applied in different fields [70–74]. The community MRO and MIP distribution patterns support the expectation of lower nutrient availability in the dry gallery vs. the wet one. In scant nutrient environments, the microbes compete over the available nutrients (high MRO), and its members might need to have complementary biosynthetic capabilities to decrease their dependency on the scarce external resources (high MIP). In caves, it is known that selfish competition for resources can be replaced by cooperative and mutualistic associations, such as the ones seen in biofilms [75], maintaining bacterial communities with diverse metabolic pathways, interdependent and cooperative [76].

On the other hand, the MRO/MIP distribution pattern supports the supposition of higher nutrient availability in the areas where invertebrates were present. Those communities (PMV1, PMV2, PMV3, PMV4) had lower MRO, and MIP than the communities where no invertebrates were identified (PMV6, PMV7, PMV8). The simulated cross-feeding interactions highlighted the key donor MAGs for each condition as *Methylomirabilales* and *Methylocella* methanotrophs and *Thermoplasmatota* autotrophic archaeon for the lower, wet gallery; and the autotrophic *Dongiales* and a *Phycisphaerales*-related MAG for the dry gallery. Little could be deduced for the *Phycisphaerales*-affiliated donor MAG. None of the marker genes for sulfur, nitrogen, methane metabolism, or CO<sub>2</sub> fixation were assembled or annotated. This is not surprising since the phylum *Planctomycetes* has the highest values (35–65%) of protein sequences with unknown functions among bacterial phyla [77].

Moreover, *Phycisphaerales* MAG is the second of its genus (SLJD01). The first was assembled from the surface sediments of a hypersaline soda lake in Siberia (GCA\_007135295). As an observation, MAGs belonging to *Ca. Patescibacteria* could not be modeled or were on

the receivers' side of cross-feeding interactions. This is typical for *Ca. Patescibacteria* members with a reduced genomes size (usually <1.5 Mb) that lack essential biosynthetic capacities have metabolic dependence and may have a parasitic or symbiotic lifestyle [78–80].

In the absence of light, the chemolithoautotrophic ecosystems are fueled by the oxidation of reduced compounds such as H<sub>2</sub>S (HS<sup>-</sup>), CH<sub>4</sub>, NH<sub>3</sub> (NH<sub>4</sub><sup>+</sup>), Fe<sup>2+</sup> and H<sup>+</sup>. The metabolic modeling community simulations pointed to distinct metabolic dependencies of simple compounds in lower and upper galleries. We postulate that the compound accessibility influences the established microbial dependencies in the environment, as H<sub>2</sub>S, NO<sub>3</sub><sup>-</sup> (a result of ammonia oxidation), and CH<sub>4</sub>O (a result of methanotrophy) are more available in the lower gallery than in the upper one. Hence the organism's dependencies on those compounds (i.e., a likelihood of species A growth depending on metabolite X from species B) are established in the upper gallery, where those compounds are lacking in the environment. Similarly, in the case of O<sub>2</sub>, the dependencies appear in the lower gallery where O<sub>2</sub> concentration is reduced. Ferrous iron (Fe<sup>2+</sup>) dependencies are characteristics of both conditions, probably because under oxidizing conditions, iron is found mainly in Fe(III) (oxyhydr) oxide minerals (i.e. goethite).

## Conclusions

The present work addressed the diversity, biogeochemical potential and ecological interaction of sediment-associated microbiome located near and distant from the sulfidic hydrothermal waters feeding the chemolithoautotrophic-based Movile Cave. Metagenomic-based approaches have indicated that the diversity of the microbiomes detected in Movile Cave sediments spans a wide taxonomic range and is likely to have a high degree of novelty. This study pinpoints chemolithoautotrophy as an essential metabolic asset in an organic carbon-poor Movile Cave environment that is not confined to the high-redox potential hydrothermal water and nearby sediments but to as far as the most distant locations in the dry gallery. This assumption is supported by the recovery of autotrophic MAGs encoding CO<sub>2</sub> fixation ability via at least three different pathways (CBB, rTCA, WL). Sulfur oxidation was predicted for microbiomes detected nearby the sulfidic water, whereas ammonia-oxidation might be active in cave sediments in contrast with apparently absent nitrification.

Additionally, methanotrophy has been inferred across all sampled sediments. Therefore, it seems to play a key role in the primary production along the entire Movile Cave, not only in the water proximity. Despite simulation simplifications, our modeling approach postulates

that nutrient scarcity is the driving force of competition-cooperation patterns across Movile Cave. The metabolic annotation and simulations point towards the autotrophic and methanotrophic MAGs as key donors in the sediments. Cross-feeding interactions can reveal the limiting compounds in the environment and the notable differences between the lower and the upper galleries.

Our findings point to the potential ecological roles and interactions of the sediment-associated microbiome and add to the previous microbiological investigations focused on the sulfidic waters in Movile Cave, thus comprehensively expanding our understanding of the peculiar chemoautotrophically-based subterranean ecosystems. Nevertheless, prospective direct metabolic quantitative assessments adjoined by multi-omics and isolation and culturing efforts are needed to further unveil the full microbiological picture of this intriguing cave ecosystem.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40793-022-00438-w>.

**Additional file 1. Fig. S1** The display of sediment samples collected from Movile Cave. Photographic images of the sediments.

**Additional file 2. Table S1** Minimal media composition used in the reconstruction of metaGEMs and community simulations. Minimal compounds (primarily inorganic) that may be in the inorganic environment or can result from microbial metabolic traits (e.g., sulfur and nitrogen oxidation/reduction, CO<sub>2</sub> fixation, methanotrophy) previously shown in Movile Cave environment. The compound name and abbreviation are in accordance with the BIGG database (<http://bigg.ucsd.edu/>). **Table S2** Metagenome scale metabolic models (metaGEMs) reconstruction and community simulation parameters, results generated and used in this study. **Table S3** Codes used for CarveMe v.1.5.1 [41] metabolic model reconstructions (metaGEMs) and SMETANA v.1.1.0 [43] community simulations.

**Additional file 3. Table S4** The physicochemical composition of the aqueous fraction of PMV4 sample collected on the water edge in Movile Cave. **Table S5** Movile Cave sediment samples extended mineralogy.

**Additional file 4. Table S6** The binning efficiency (the percentages of reads mapped to the MAGs). **Table S7** Main features and statistics of Movile cave sediments-retrieved MAGs. **Table S8** Taxonomic assignment of MAGs by GTDB-tk v202. **Table S9** The binning efficiency and the relative abundance of MAGs in each metagenome.

**Additional file 5. Fig. S2** The number of reads from the 7 metagenomes mapped back to the reference sequences of the marker genes. **Fig. S3** The overall number of reads per metagenome mapped to the reference sequences of all considered marker genes (nitrogen fixation (nifH/D/K), carbon fixation (cbbL/S, acIA/B, acsA(cooS)/B)). **Table S10** Reference sequences for the marker genes used in predicting nitrogen and carbon fixation potential in sediments of Movile Cave based on the metagenomic read mapping analysis.

**Additional file 6. Table S11** The marker genes for sulfur-cycling and the MAGs that encoded them in each sample. **Table S12** The marker genes for nitrogen-cycle and the MAGs that encoded them in each sample. **Table S13** The methane oxidation marker genes and the MAGs that encoded them in each sample. **Table S14** CO<sub>2</sub> fixation marker genes and the MAGs that encoded them in each sample. Tables S9–S12 are the bases for Fig. 4 and contain the pathways, KEGG number assignment (KO), gene, sample number and the MAG's that possess genes from the addressed metabolic pathways.

**Additional file 7. Table S15** MEMOTE [42] test results regarding the potential of gap-fill models (metaGEMs) initialized for the minimal medium to produce biomass (growth). The biomass production potential is indicated by the biomass production values > 0. **Table S16** The competition (given by MRO score) and cooperation (given by MIP score) potential for the community in each sample, predicted by SMETANA [43]. **Table S17** The MAGs that interact in SMETANA simulation for the microbial community in the lower galleries. **Table S18** The MAGs that interact in SMETANA simulation for the microbial community in the upper galleries. **Table S19** Compounds exchanged that differed between upper and lower galleries according to Wilcoxon rank-sum test.

## Acknowledgements

We are grateful to Marius Kenezs and Cristian Sitar for providing some of the samples, Mihai Baciu for the access into the cave, Ionuț Cornel Mirea for the cave map editing. We acknowledge Francisco Zorrilla for his assistance in metabolic simulations and Wilcoxon rank-sum test statistical analysis.

## Author contributions

OTM, HLB, IC, GR conceived the study. OTM and LF performed sample collection. OTM, IC, AEL performed metadata analysis. LF performed the mineralogy and AEL performed the geochemistry measurements. IC and DFB performed DNA extraction. GR performed the bioinformatics analysis. IC and GR analyzed the metagenomic data. IC performed the metabolic modeling and community simulation analysis. IC wrote the manuscript and prepared the figures. OTM and HLB provided scientific input to the manuscript. AHV contributed to the final version of the manuscript. All authors read and approved the manuscript.

## Funding

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI-UEFISCDI, project number 16/2018 (DARK-FOOD), within PNCDI III, granted to OTM.

## Availability of data and materials

Raw metagenomics datasets and metagenome assembled genome derived from this work are publicly available under the NCBI Bioproject PRJNA777757. Metagenomes were deposited to NCBI-SRA with the accession numbers SAMN22875101-SAMN22875107. Metagenome-assembled genomes have been deposited in NCBI BioSample and are available under the accessions SAMN23768506-SAMN23768611.

## Declarations

### Ethical approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

### Consent for publication

All authors have read and commented on the manuscript and consent to the publication.

### Competing interests

The author(s) declare no competing interest.

### Author details

<sup>1</sup>Department of Molecular Biology and Biotechnology, Babeş-Bolyai University, Cluj-Napoca, Romania. <sup>2</sup>EMBL Partner Institute for Genome Editing, Life Sciences Center-Vilnius University, Vilnius, Lithuania. <sup>3</sup>Doctoral School of Integrative Biology, Babeş-Bolyai University, Cluj-Napoca, Romania. <sup>4</sup>National Institute for Research and Development for Optoelectronics, Research Institute for Analytical Instrumentation Subsidiary, Cluj-Napoca, Romania. <sup>5</sup>Emil Racovita Institute of Speleology, Geospeleology and Paleontology Department, Bucharest, Romania. <sup>6</sup>Romanian Institute of Science and Technology, Cluj-Napoca, Romania. <sup>7</sup>Biospeleology and Edaphobiology Department, Emil Racovita Institute of Speleology, Bucharest, Romania. <sup>8</sup>Cluj-Napoca Department, Emil Racovita Institute of Speleology, Cluj-Napoca, Romania. <sup>9</sup>Centre



for Systems Biology, Biodiversity and Bioresources, Babeş-Bolyai University, Cluj-Napoca, Romania.

Received: 27 May 2022 Accepted: 5 August 2022

Published online: 17 August 2022

## References

- McCollom TM, Shock EL. Geochemical constraints on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochim Cosmochim Acta*. 1997;61:4375–91.
- Pan J, Xu W, Zhou Z, Shao Z, Dong C, Liu L, et al. Genome-resolved evidence for functionally redundant communities and novel nitrogen fixers in the deyin-1 hydrothermal field. *Mid-Atlantic Ridge Microbiome*. 2022;10:8.
- Bell JB, Wouds C, van Oevelen D. Hydrothermal activity, functional diversity and chemoautotrophy are major drivers of seafloor carbon cycling. *Sci Rep*. 2017;7:12025.
- Jiang Q, Jing H, Jiang Q, Zhang Y. Insights into carbon-fixation pathways through metagenomics in the sediments of deep-sea cold seeps. *Mar Pollut Bull*. 2022;176: 113458.
- Lipsewars YA, Hopmans EC, Meysman FJR, Sinninghe Damsté JS, Villanueva L. Abundance and diversity of denitrifying and anammox bacteria in seasonally hypoxic and sulfidic sediments of the saline lake Grevelingen. *Front Microbiol*. 2016;7:1661.
- Vavourakis CD, Mehrshad M, Balkema C, van Hall R, Andrei A-Ş, Ghai R, et al. Metagenomes and metatranscriptomes shed new light on the microbial-mediated sulfur cycle in a Siberian soda lake. *BMC Biol*. 2019;17:69.
- Tourova TP, Kovaleva OL, Sorokin DYU, Muyzer G. Ribulose-1,5-bisphosphate carboxylase/oxygenase genes as a functional marker for chemolithoautotrophic halophilic sulfur-oxidizing bacteria in hypersaline habitats. *Microbiology*. 2010;156:2016–25.
- Sarbu SM, Kane TC, Kinkle BK. A Chemoautotrophically Based Cave ecosystem. *Science*. 1996;272:1953–5.
- Por FD. Ophel: a groundwater biome based on chemoautotrophic resources. The global significance of the Ayyalon cave finds, Israel. *Hydrobiologia*. 2007;592:1–10.
- Sarbu SM, Galdenzi S, Menichetti M, Gentile G. Geology and biology of the Frasassi caves in central Italy: an ecological multidisciplinary study of a hypogenic cave system. In: Wilkens H, Culver DC, Humphreys WF, editors. *Ecosystems of the World 30. Subterranean ecosystems*. Amsterdam: Elsevier. 2000. p. 359–78.
- Flot J-F, Wörheide G, Dattagupta S. Unsuspected diversity of *Niphargus* amphipods in the chemoautotrophic cave ecosystem of Frasassi, central Italy. *BMC Evol Biol*. 2010;10:171.
- Engel AS, Meisinger DB, Porter ML, Payn RA, Schmid M, Stern LA, et al. Linking phylogenetic and functional diversity to nutrient spiraling in microbial mats from Lower Kane Cave (USA). *ISME J*. 2010;4:98–110.
- Engel AS, Porter ML, Kinkle BK, Kane TC. Ecological assessment and geological significance of microbial communities from cesspool cave. *Virginia Geomicrobiol J*. 2001;18:257–74.
- Hose LD, Palmer AN, Palmer MV, Northup DE, Boston PJ, DuChene HR. Microbiology and geochemistry in a hydrogen-sulphide-rich karst environment. *Chem Geol*. 2000;169:399–423.
- Brad T, Iepure S, Sarbu SM. The chemoautotrophically based movile cave groundwater ecosystem, a hotspot of subterranean biodiversity. *Diversity*. 2021;13:128.
- Porter M, Engel AS, Kane T, Kinkle B. Productivity-diversity relationships from chemolithoautotrophically based sulfidic karst systems. *IJS*. 2009;38:27–40.
- Sarbu S, Lascu C. Condensation corrosion in Movile Cave, Romania. *J Cave Karst Stud*. 1997;59:99–102.
- Kumaresan D, Wischer D, Stephenson J, Hillebrand-Voiculescu A, Murrell JC. Microbiology of Movile Cave—a chemolithoautotrophic ecosystem. *Geomicrobiol J*. 2014;31:186–93.
- Kumaresan D, Stephenson J, Doxey AC, Bandukwala H, Brooks E, Hillebrand-Voiculescu A, et al. Aerobic proteobacterial methylotrophs in Movile Cave: genomic and metagenomic analyses. *Microbiome*. 2018;6:1.
- Hutchens E, Radajewski S, Dumont MG, McDonald IR, Murrell JC. Analysis of methanotrophic bacteria in Movile Cave by stable isotope probing. *Environ Microbiol*. 2004;6:111–20.
- Chen Y, Wu L, Boden R, Hillebrand A, Kumaresan D, Moussard H, et al. Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J*. 2009;3:1093–104.
- Bizic M, Brad T, Ionescu D, Barbu-Tudoran L, Aerts J, Popa R, et al. Cave *Thiovulaceae* differ metabolically and genomically from marine species. 2022;2020.11.04.367730.
- Sarbu SM, Vlasceanu L, Popa R, Sheridan P, Kinkle BK, Kane TC. Microbial mats in a thermomineral sulfurous cave. In: Stal LJ, Caumette P, editors. *Microbial Mats*. Berlin, Heidelberg: Springer; 1994. p. 45–50.
- Rohwerder T, Sand W, Lascu C. Preliminary evidence for a sulphur cycle in Movile Cave. *Romania Acta Biotechnol*. 2003;23:101–7.
- Sarbu SM, Lascu C, Brad T. Dobrogea: Movile Cave. In: Ponta GML, Onac BP, editors. *Cave and karst systems of Romania*. Cham: Springer International Publishing; 2019. p. 429–36.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
- Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 23 Feb 2022.
- Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 2015;31:1674–6.
- 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.
- Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7: e7359.
- Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol*. 2018;3:836–43.
- Parks DH, Imelfort M, Skennerton CT, Hugenholz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25:1043–55.
- Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010;11:119.
- Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2014;12:59–60.
- 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.
- Chaumeil P-A, Mussig AJ, Hugenholz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. 2019;36:1925–7.
- 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 Web Server issue:W242–5.
- Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34:3094–100.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–9.
- Cabello-Yeves PJ, Callieri C, Picazo A, Mehrshad M, Haro-Moreno JM, Roda-Garcia JJ, et al. The microbiome of the Black Sea water column analyzed by shotgun and genome centric metagenomics. *Environmental Microbiome*. 2021;16:5.
- Machado D, Andrejev S, Tramontano M, Patil KR. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res*. 2018;46:7542–53.
- Lieven C, Beber ME, Olivier BG, Bergmann FT, Ataman M, Babaei P, et al. MEMOTE for standardized genome-scale metabolic model testing. *Nat Biotechnol*. 2020;38:272–6.
- Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc Natl Acad Sci USA*. 2015;112:6449–54.

44. King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, et al. BiGG models: a platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* 2016;44:D515–22.
45. Wickham H. ggplot2. *WIREs Comput Stat.* 2011;3:180–5.
46. pheatmap: Pretty Heatmaps. <https://CRAN.R-project.org/package=pheatmap>. Accessed 24 Feb 2022.
47. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize implements and enhances circular visualization in R. *Bioinformatics.* 2014;30:2811–2.
48. Bowers RM, Kyrpidis 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. Parks DH, Chuvpochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol.* 2018;36:996–1004.
50. Stein LY, Roy R, Dunfield PF. Aerobic Methanotrophy and Nitrification: Processes and Connections. In: John Wiley & Sons, Ltd, editor. eLS. 1st edition. Wiley; 2012. p. 1–11.
51. D'Angeli IM, Ghezzi D, Leuko S, Fircincieli A, Parise M, Fiorucci A, et al. Geomicrobiology of a seawater-influenced active sulfuric acid cave. *PLoS ONE.* 2019;14: e0220706.
52. Adesso R, Gonzalez-Pimentel JL, D'Angeli IM, De Waele J, Saiz-Jimenez C, Jurado V, et al. Microbial community characterizing vermiculations from Karst Caves and its role in their formation. *Microb Ecol.* 2021;81:884–96.
53. Jones D, Lyon E, Macalady J. Geomicrobiology of biovermiculations from the Frasassi Cave system, Italy. *J Cave Karst.* 2008;70(2):78–93.
54. Sorokin DY, Merkel AY, Muyzer G, et al. Thiohalomonas. In: Trujillo ME, Dedysh S, DeVos P, Hedlund B, Kämpfer P, Rainey FA, et al., editors. *Bergey's manual of systematics of archaea and bacteria.* 1st ed. Wiley; 2020. p. 1–6.
55. Waite DW, Vanwonterghem I, Rinke C, Parks DH, Zhang Y, Takai K, et al. Comparative genomic analysis of the class *Epsilonproteobacteria* and proposed reclassification to *Epsilonbacteraeota* (phyl. nov.). *Front Microbiol.* 2017;0.
56. Jurado V, D'Angeli I, Martin-Pozas T, Cappelletti M, Ghezzi D, Gonzalez-Pimentel JL, et al. Dominance of *Arcobacter* in the white filaments from the thermal sulfidic spring of Fetida Cave (Apulia, southern Italy). *Sci Total Environ.* 2021;800: 149465.
57. Sarbu SM. Movable Cave: a chemoautotrophically based groundwater ecosystem. In: Wilken H, Culver DC, Humphreys WF, editors. *Subterranean Ecosystems.* Amsterdam: Elsevier; 2000. p. 319–43.
58. Tourna M, Stieglmeier M, Spang A, Konneke M, Schintlmeister A, Urlich T, et al. *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc Natl Acad Sci.* 2011;108:8420–5.
59. Spang A, Poehlein A, Offre P, Zumbärgel S, Haider S, Rychlik N, et al. The genome of the ammonia-oxidizing *Candidatus Nitrososphaera gargensis*: insights into metabolic versatility and environmental adaptations. *Environ Microbiol.* 2012;14:3122–45.
60. Wischer D, Kumaresan D, Johnston A, El Khawand M, Stephenson J, Hillebrand-Voiculescu AM, et al. Bacterial metabolism of methylated amines and identification of novel methylotrophs in Movable Cave. *ISME J.* 2015;9:195–206.
61. Dedysh SN, Dunfield PF. Chapter three - facultative and obligate methanotrophs: how to identify and differentiate them. In: Rosenzweig AC, Ragsdale SW, editors. *Methods in enzymology.* Academic Press; 2011. p. 31–44.
62. Ivanova AA, Oshkin IY, Danilova OV, Philippov DA, Ravin NV, Dedysh SN. Rokubacteria in Northern Peatlands: habitat preferences and diversity patterns. *Microorganisms.* 2022;10:11.
63. Ettwig KF, Butler MK, Le Paslier D, Pelletier E, Mangenot S, Kuypers MMM, et al. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature.* 2010;464:543–8.
64. Graf JS, Mayr MJ, Marchant HK, Tienken D, Hach PF, Brand A, et al. Bloom of a denitrifying methanotroph, '*Candidatus Methyloirabilis limnetica*', in a deep stratified lake. *Environ Microbiol.* 2018;20:2598–614.
65. Versantvoort W, Guerrero-Castillo S, Wessels HJCT, van Niftrik L, Jetten MSM, Brandt U, et al. Complexome analysis of the nitrite-dependent methanotroph *Methyloirabilis lanthanidiphila*. *Biochim Biophys Acta Bioenerg.* 2019;1860:734–44.
66. Maki JS, et al. *Thermodesulfobivrio*. In: Whitman WB, Rainey F, Kämpfer P, Trujillo M, Chun J, DeVos P, et al., editors. *Bergey's manual of systematics of archaea and bacteria.* 1st ed. Wiley; 2015. p. 1–9.
67. Macalady JL, Jones DS, Lyon EH. Extremely acidic, pendulous cave wall biofilms from the Frasassi cave system, Italy. *Environ Microbiol.* 2007;9:1402–14.
68. Jones DS, Albrecht HL, Dawson KS, Schaperdoth I, Freeman KH, Pi Y, et al. Community genomic analysis of an extremely acidophilic sulfur-oxidizing biofilm. *ISME J.* 2012;6:158–70.
69. Noviana Z, Vieira S, Pascual J, Fobofou SAT, Rohde M, Spröer C, et al. *Hypericibacter terrae* gen. nov., sp. nov. and *Hypericibacter adhaerens* sp. nov., two new members of the family Rhodospirillaceae isolated from the rhizosphere of *Hypericum perforatum*. *Int J Syst Evol Microbiol.* 2020;70:1850–60.
70. Basile A, Campanaro S, Kovalovszki A, Zampieri G, Rossi A, Angelidaki I, et al. Revealing metabolic mechanisms of interaction in the anaerobic digestion microbiome by flux balance analysis. *Metab Eng.* 2020;62:138–49.
71. Dal Bello M, Lee H, Goyal A, Gore J. Resource–diversity relationships in bacterial communities reflect the network structure of microbial metabolism. *Nat Ecol Evol.* 2021;5:1424–34.
72. Devika NT, Jangam AK, Katneni VK, Patil PK, Nathamuni S, Shekhar MS. In Silico Prediction of Novel Probiotic Species Limiting Pathogenic *Vibrio* Growth Using Constraint-Based Genome Scale Metabolic Modeling. *Front Cell Infect Microbiol.* 2021. <https://doi.org/10.3389/fcimb.2021.752477>.
73. Soto DF, Franzetti A, Gómez I, Huovinen P. Functional filtering and random processes affect the assembly of microbial communities of snow algae blooms at Maritime Antarctic. *Sci Total Environ.* 2022;805: 150305.
74. Zorrilla F, Patil KR, Zelezniak A. metaGEM: reconstruction of genome scale metabolic models directly from metagenomes. *Nucleic Acids Res.* 2020. <https://doi.org/10.1093/nar/gkab815>.
75. Barton H, Jurado V. What's up down there? microbial diversity in caves microorganisms in caves survive under nutrient-poor conditions and are metabolically versatile and unexpectedly diverse. *Microbe.* 2007;2:132–8.
76. Dong Y, Gao J, Wu Q, Ai Y, Huang Y, Wei W, et al. Co-occurrence pattern and function prediction of bacterial community in Karst cave. *BMC Microbiol.* 2020;20:137.
77. Overmann J, Abt B, Sikorski J. Present and future of culturing bacteria. *Annu Rev Microbiol.* 2017;71:711–30.
78. Lemos LN, Medeiros JD, Dini-Andreote F, Fernandes GR, Varani AM, Oliveira G, et al. Genomic signatures and co-occurrence patterns of the ultra-small *Saccharimonadia* (phylum CPR/Patescibacteria) suggest a symbiotic lifestyle. *Mol Ecol.* 2019;28:4259–71.
79. Lemos LN, Manoharan L, William Mendes L, Monteiro Venturini A, Satler Pylro V, Tsai SM. Metagenome assembled-genomes reveal similar functional profiles of CPR/Patescibacteria phyla in soils. *Environ Microbiol Rep.* 2020;12:651–5.
80. Chaudhari NM, Overholt WA, Figueroa-Gonzalez PA, Taubert M, Bornemann TLV, Probst AJ, et al. The economical lifestyle of CPR bacteria in groundwater allows little preference for environmental drivers. *Environ Microbiome.* 2021;16:24.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.