Genomic and morphologic characterization of a planktonic *Thiovulum* (*Campylobacterota*) dominating the surface waters of the sulfidic Movile Cave, Romania

Mina Bizic[1,2,*#], Traian Brad[3,4,*#], Lucian Barbu-Tudoran[5], Joost Aerts[6], Danny Ionescu[1,2], Radu Popa[7, 8], Jessica Ody[9], Jean-François Flot[9, 10], Scott Tighe[11], Daniel Vellone[11], and Serban M. Sarbu[8, 12, 13]

[1] Leibniz Institute for Freshwater Ecology and Inland Fisheries, IGB, Dep 3, Experimental Limnology, Alte Fischerhütte 2, OT Neuglobsow, 16775 Stechlin, Germany
[2] Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Berlin, Germany
[3] "Emil Racoviţă" Institute of Speleology, Clinicilor 5-7, 400006 Cluj-Napoca Romania,
[4] Institutul Român de Știință și Tehnologie, Str. Virgil Fulicea nr. 3, 400022 Cluj-Napoca, Romania
[5] Center for Electron Microscopy, "Babeș-Bolyai" University, Clinicilor 5, 400006 Cluj-Napoca, Romania
[6] Department of Molecular Cell Physiology, Faculty of Earth and Life sciences, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands
[7] River Road Research, 62 Leslie St. Buffalo, NY 14211
[8] Emil G. Racoviţă Institute, Babes-Bolyai University, Cluj-Napoca 400006, Romania
[9] Evolutionary Biology and Ecology, Université libre de Bruxelles (ULB), C.P. 160/12, Avenue F.D. Roosevelt 50, 1050 Brussels, Belgium
[10] Interuniversity Institute of Bioinformatics in Brussels – (IB)², Brussels, Belgium
[11] Vermont Integrative Genomics Lab, University of Vermont Cancer Center, Health Science Research Facility, Burlington, Vermont, 05405
[12] Emil Racoviţă Institute of Speleology, str. Frumoasă 31, 010986 Bucureşti, Romania
[13] Department of Biological Sciences, California State University, Chico, USA

[*] Authors have contributed equally to the work

Corresponding authors

Mina Bizic: mbizic@igb-berlin.de

Traian Brad: traian.brad@iser.ro

ORIGINALITY SIGNIFICANCE STATEMENT

We identify *Thiovulum* sp. as a dominant bacterium in subsurface floating microbial agglomerations in a hypoxic cave system. We show for the first time that *Thiovulum* can be found in high abundance in planktonic environments dissociated from solid surfaces. We provide the complete genome of this organism and suggest that it is capable of dissimilatory nitrate reduction to ammonia using sulfur as the electron donor, thus contributing to both the cave's nitrogen and sulfur cycles. Based on sequence similarity, this clade of *Thiovulum* spp. may be common and important in other sulfidic caves. Last, we propose that the short peritrichous flagella-like structures of *Thiovulum* are type IV pili rather than actual flagella.

ABSTRACT

Life in Movile Cave (Romania) relies entirely on primary carbon fixation by bacteria oxidizing sulfide, methane and ammonia with oxygen, nitrate, sulfate, and ferric iron. There, large spherical-ovoid bacteria (12-16 µm diameter), rich in intracellular sulfur globules, dominate the stable microbial community in the surface water of a hypoxic Air Bell. These were identified as *Thiovulum* sp. (*Campylobacterota*). We obtained a closed genome of this *Thiovulum* and compared it to that of *Thiovulum* ES. The genes for oxidizing sulfide to sulfate are absent, therefore, *Thiovulum* likely avoids constant accumulation of elemental sulfur either by oxidizing sulfide to sulfite which is then excreted, or via dissimilatory nitrate reduction to ammonia using the formate-dependent nitrite reductase or hydroxylamine oxidoreductase. Thus, *Thiovulum, found also in other caves,* is likely important to both S and N cycles in subterranean aquatic ecosystems. Additionally, using electron microscopy, we suggest that in absence of motor-like structures along the membrane, the peritrichous flagella-like structures are type IV pili, for which genes were found in both *Thiovulum* genomes. These pili may play a role in veil formation, connecting adjacent cells. The force exerted by coordinated movement of such pili may partly explain the exceptionally fast swimming of these bacteria.

Key words: *Thiovulum*, sulfur, DNRA, Movile Cave, sulfide-oxidation, stable isotope analysis

INTRODUCTION

Cave ecosystems, typically characterized by stable conditions, provide a window into subsurface microbiology (Engel, 2015). In the absence of natural light, these ecosystems are typically fueled by chemolithoautotrophy via the oxidation of reduced compounds such as H_2S, Fe^{2+}, Mn^{2+}, NH_3, CH_4, and H^+. In sulfidic ecosystems, H_2S oxidation may be performed by animal-microbial symbioses such as the *Niphargus-Thiotrix* association (Bauermeister *et al.*, 2012; Flot *et al.*, 2014).

Discovered at the bottom of an 18-m deep geological survey shaft in 1986, Movile Cave is located near the town of Mangalia, SE Romania (43°49'32"N, 28°33'38"E), 2.2 km from the Black Sea shore. It consists of a 200 m long upper dry passage that ends in a small lake allowing access to a 40 m long, partially submerged lower cave level (Fig. 1). Thick and impermeable layers of clays and loess cover the limestone in which the cave is developed, preventing input of water and nutrients from the surface (Lascu *et al.*, 1994).

Sulfidic groundwater flows constantly at the bottom of Movile Cave's lower passages. This flow occurs down to over 1 m below the water surface at flow rates of about 5 l s^{-1} (Sarbu and Lascu, 1997), which equates to a flow velocity of approximately 5 cm s^{-1} in the Lake Room area. Yet, because of the morphology of the lower cave passages (Fig. 1) and slight difference in water temperatures, the water near the surface is practically stagnant. Earlier work by Riess *et al.* (1999) analyzed the change in the concentration of O_2 in the water column and found traces of oxygen up to a maximum depth of 0.8 mm, below which the water was anoxic.

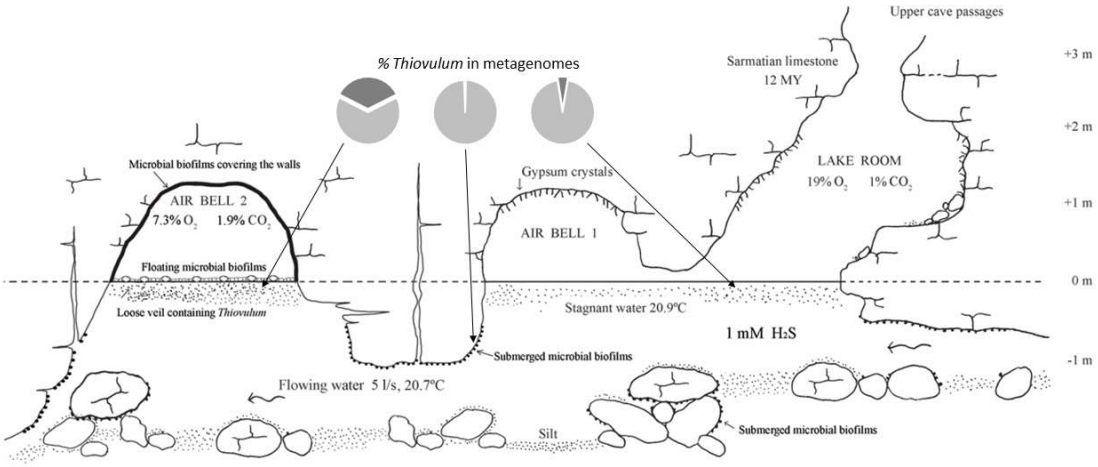


Figure 1. Longitudinal profile of the sampling area in Movile Cave (modified after Sarbu and Popa, 1992). The microbial community containing *Thiovulum* cells (depicted here as dots present beneath the water surface) was sampled in the Lake Room and in air bells 1 and 2. *Thiovulum* 16S rRNA made up 5 %, 0.9 % and 35 % of the 16S rRNA genes retrieved from metagenomic samples (dark gray in pies) from these cave sections, respectively.

Chemoautotrophic microorganisms living at the water surface oxidize reduced chemical compounds such as H_2S, CH_4 and NH_4^+ from the thermo-mineral groundwater (Sarbu and Kane, 1995; Sarbu, 2000). Most of the microbiological studies performed in Movile Cave (summarized in Kumaresan *et al.*, 2014) are based on samples of microbial biofilms floating on the water surface or covering rock surfaces in the cave's air bells (Fig. 1), where the atmosphere is poor in O_2 (7-10 %) and enriched in CO_2 (2.5 %) and methane (1-2 %) (Sarbu, 2000). *Thiobacillus*, *Thiothrix*, *Thioploca*, *Thiomonas* and *Sulfurospirillum* oxidize H_2S using O_2 or NO_3^- as electron acceptors (Rohwerder *et al.*, 2003; Chen *et al.*, 2009; Flot *et al.*, 2014). The methanotrophs *Methylomonas*, *Methylococcus* and *Methylocystis* (Hutchens *et al.*, 2004), *Methanobacterium* (Schirmack *et al.*, 2014) and *Methanosarcina* (Ganzert *et al.*, 2014) are also found in the cave, alongside other methylothrophs such as *Methylotenera*, *Methylophilus* and *Methylovorus* (Rohwerder *et al.*, 2003; Chen *et al.*, 2009). Chen *et al.*, (2009) further identified in the cave ammonia and nitrite oxidizers from the genera *Nitrospira* and *Nitrotoga*.

Directly below the water surface in the lower level of Movile Cave, we observed a loose floating veil resembling a slow-moving white cloud (Fig. 2 and Supplementary video 1). When the water was not disturbed, the veil formation was evenly dispersed, yet not deeper than 2-3 mm below the water surface. Occasionally, this veil was disturbed by feeding invertebrates including dendrocoelid flatworms (*Dendrocoelum obstinatum*), cyclopoid copepods (*Eucyclops subterraneus scythicus*), and niphargid amphipods (*Niphargus dancaui* and *Niphargus racovitzai*) (Brad *et al.*, 2015; Stocchino *et al.*, 2017; Sarbu *et al.*, 2019).

Using genetic and microscopic analysis, we concluded that this underwater agglomeration of bacteria is dominated by a species of the genus *Thiovulum*. To the best of our knowledge, this is the first description of planktonic *Thiovulum* swarms/veils at considerable distance from any solid surface. Here we provide further morphological and genomic information on this bacterium, offering new insights into its metabolic properties and raising novel questions.

MATERIALS AND METHODS

Replicate samples of water were collected into sterile containers from the surface of the small sulfidic lake and from the air-bells in the lower section of Movile Cave (Fig. 1). Unpreserved 50 ml water samples were immediately brought to the laboratory and inspected by optical microscopy, while other 50 ml water samples preserved with ethanol to a final concentration of 50 % and later used for DNA extraction. Additionally, 15-ml water samples were preserved with formaldehyde to a final concentration of 4 % and were later used to estimate the density of bacterial cells under an optical microscope by direct counts on a squared slide, using a defined sample volume (i.e. 5 ml).

Electron microscopy and elemental analysis

Two 15-ml samples collected from the water surface in Movile cave were centrifuged, and the microorganisms were resuspended and fixed for 2 h with 2.7 % glutaraldehyde in phosphate buffered saline (1× PBS). The cells were then rinsed three times in 1× PBS, and finally fixed for 1 h with 2 % osmic acid in 1× PBS. The cells were harvested again by centrifugation, dehydrated in graded acetone-distilled water dilutions, and embedded in epoxy resin. Sections of about 100 nm thickness were produced with a diamond knife (Diatome, Hatfield) using a Leica UC6 ultramicrotome (Leica Microsystems, Wetzlar, Germany) and were stained with lead citrate and uranyl acetate (Hayat, 2001). The grids were examined with a Jeol JEM transmission electron microscope. Samples for scanning electron microscopy (SEM) were fixed with 2.7 % glutaraldehyde in 1× PBS, air dried on 0.22 µm mesh-sized Millipore filters, sputter coated with 10 nm gold and examined on a JEOL JSM 5510 LV microscope (Jeol, Japan). Energy-dispersive X-ray spectroscopy (EDX) analysis was performed with an EDX analyzer (Oxford Instruments, Abingdon, UK) and with the INCA 300 software.

Stable isotope analysis

Surface water from the Lake Room (Fig. 1) was collected in a sterile 1 l Nalgene bottle and passed through a plankton net to remove any aquatic fauna and particles of floating microbial mats. The water was then filtered through ashed fiberglass filters (that had previously been heated at 1000 $^{\circ}$C for an hour to remove carbonates and organic molecules). The filters were then dried at 60 $^{\circ}$C for 24 h and sent to the Water Research Center of the University of Alaska in Fairbanks for carbon and nitrogen stable isotope analysis. Additional samples of water from the lake were collected in sterile 1 l Nalgene bottles and the H_2S was precipitated with Cd-acetate and then filtered through ashed fiberglass filters that retained the precipitate. Gypsum crystals were also collected from the cave walls above the water in the Lake Room to determine the stable isotope ratios of sulfur. After being dried, the filters were sent along with samples of gypsum to the Stable Isotope Laboratory of the University of Indiana for sulfur stable isotope analysis.

gDNA Extraction

Bacteria from water samples were concentrated using a vacuum pump and Nalgene's single-use analytical filter funnels (Thermo Fisher Scientific, MA, USA), with the included filter replaced with a 0.2 µm isopore membrane filter (Millipore Sigma, MA, USA). Prior to filtration, the glass assembly components were autoclaved at 121.5 °C for 30 min wet and 20 min dry at 1.4 bar (20 psi). Filters were placed into 50 µl conical tubes with 300 µl 1× PBS, 1.5 µl 2 % azide, and a sterile scalpel blade. Samples were minced using an OMNI bead ruptor elite (OMNI International, GA, USA) on a 4 m s^{-1} 30 sec program. The resulting sample was centrifuged to collect bacteria separated from the filters and extracted for gDNA using a modified version of the Omega BioTek Universal Metagenomics kit protocol (OMEGA Bio-tek, GA, USA). Fifteen µl of MetaPolyzyme (Millipore Sigma, MA, USA) was added to each sample and incubated at

35 °C for 13 h followed by three cycles of freeze and thaw alternating between 80 °C for 2 min and a -80 °C freezer for 10 min. Further digestion was subsequently performed by adding 35 µl Proteinase K (Omega Bio-tek, GA, USA) and incubated at 55 °C for 1 hour. Following complete enzymatic digestion, the sample was extracted using the manufacturer's protocol (Omega BioTek Universal metagenomics kit). Briefly, 500 µl ML1 buffer (CTAB) was added the digested sample and incubated at 55 °C for 15 min. One volume of Tris-stabilized (pH >7.5) phenol-chloroform-isoamyl alcohol mix (25:24:1) was used for purification and the resulting upper aqueous phase was removed and combined with RBB (Guanidinium) and 100 % EtOH and applied to a DNA silica column supplied with the kit. Final DNA was eluted in 35 µl of elution buffer. DNA was quantified using a Qubit spectrofluorometer and Nanodrop ND-1000 (ThermoFisher Waltham MA USA).

16S rRNA gene amplicon sequencing and processing

PCR reactions were performed in triplicate. Each 25 µl reaction consisted of 0.5 µl of Phusion Green Hot Start II high-fidelity DNA polymerase (Thermo Fisher Scientific, Sweden), 5.0 µl of 5× Phusion Green HF buffer, 4 µl DNase- and RNase-free water, 5.0 µl of 10 µM primer mix (1:1), 0.5 µl of 10 mM nucleotide mix and 10 µl of DNA extract (0.5 ng/µl). The PCR protocol was 98 ºC for 30 sec, 33 cycles of 98 ºC for 10 sec, 55 ºC for 30 sec, 72 ºC for 30 sec and a final 5-min extension at 72 ºC. The amplicon target was the V3-V4 region of the 16S rRNA gene, using the V3 forward primer S-D-Bact-0341-b-S-17, 5'-CCTACGGGNGGCWGCAG-3 (Herlemann *et al.*, 2011), and the V4 reverse primer S-D-Bact-0785-a-A-21, 5'-GACTACHVGGGTATCTAATCC-3 (Muyzer *et al.*, 1993), resulting in fragments of ~430 bp. The primers were dual barcoded in a way compatible with Illumina sequencing platforms (as described in Caporaso *et al.*, (2011). Product size and successful amplification was tested by running internal positive and negative controls from the triplicate plates on a 1.5 % (w/v) agarose gel. Triplicate PCR products were combined, and each triplicate sample was purified using SPRI beads (Agencourt® AMPure® XP, Beckman Coulter, CA, USA). DNA concentration of purified samples was determined with a Quant-iT high-sensitivity DNA assay kit and a Qubit® fluorometer (Invitrogen, Carlsbad, USA). All samples were diluted to similar concentrations prior to pooling diluted PCR products together in equimolar volumes (50 µl) in one composite sample (including positive and negative controls).

Composite samples were paired-end sequenced at the Vrije Universiteit Amsterdam Medical Center (Amsterdam, The Netherlands) on an Illumina MiSeq Sequencer with a 600-cycle MiSeq Reagent Kit v3 (Illumina, San Diego, Ca, USA) according to the manufacturer instructions. Sequences were quality trimmed using Trimmomatic (v 0.39) and paired using Bbmerge (Bushnell *et al.*, 2017). The paired sequences were dereplicated using the dedupe tool of the BBTools package (sourceforge.net/projects/bbmap/) aligned and annotated using the SINA aligner (Pruesse *et al.*, 2012) against the SILVA SSU database (v 138) (Quast *et al.*, 2012).

A maximum-likelihood tree was calculated using FastTree 2 (Price *et al.*, 2010). For the sake of legibility, the 908 *Thiovulum* sequence variants obtained were clustered at 97 % similarity using CD-HIT-EST (Huang *et al.*, 2010), resulting in 50 clusters. Three 16S rRNA sequences obtained from the genome assembly (see below) were included in the tree as well. All *Thiovulum* sequences available in the SILVA database (71 in total) were clustered separately in a similar manner, resulting in 22 clusters used in the final tree. Non-*Thiovulum* sequences belonging to the *Sulfurimonadaceae* family were used as an outgroup.

To obtain information on relative *Thiovulum* abundance, the raw short-read libraries (amplicon and metagenomic) were analyzed with phyloFlash (V 3.3; (Gruber-Vodicka *et al.*, 2020)).

Shotgun sequencing (Illumina and Oxford Nanopore)

Shotgun sequencing was accomplished using both Illumina and Oxford Nanopore sequencing technologies. For Illumina sequencing, 1 ng of genomic DNA from each sample was converted to whole-genome sequencing libraries using the Nextera XT sequencing reagents according to the manufacturer's instructions (Illumina, San Deigo CA). Final libraries were checked for library insert size using the Agilent Bioanalyser 2100 (Agilent Technologies Santa Clara, CA) and quantified using Qubit spectrofluorometry. The final sample was sequenced using paired end 2x150 sequencing on an Illumina MiniSeq system.

A first pass of Oxford Nanopore sequences was obtained using the SQK LSK109 ligation library synthesis reagents on a Rev 9.4 nanopore flow cell with the GridION X5 MK1 sequencing platform, resulting in a total of 131.8 Mbp of reads with a N50 of 1.3 kbp.

Additionally, sequencing was performed on several cellular aggregates that were confirmed microscopically to contain *Thiovulum* cells. For this, the aggregates were picked from environmental samples fixed in RNAlater (4 M $(NH_4)_2SO_4$; 15 mM EDTA (from 0.5 M, pH 8.0 stock); 18.75 mM Na-citrate (from 1M stock)), and washed several times in sterile 1× PBS buffer (137 mM NaCl; 2.7 mM KCl; 10 mM Na_2HPO_4; 1.8 mM KH_2PO_4; pH 7.4). The cell aggregates were lysed by freeze thawing and further, following the manufacturer's instructions, as part of the DNA amplification process using the Repli-G single cell amplification kit (Qiagene, Hilden, Germany). Libraries for Nanopore sequencing were prepared using the LSK-108 kit following the manufacturer's protocol but skipping the size selection step. The prepared libraries were loaded on MIN106 R9 flow cells, generating a total of 5.7 Gbp of reads with a length N50 of about 3.7 kbp. Basecalling for all Oxford Nanopore reads were done using Guppy 4.0.11.

All sequencing data generated in this study were deposited in NCBI Sequence Read Archive under accession number PRJNA673084.

Metagenomic data analysis

Nanopore reads were assembled using Flye 2.8.1-b1676 (Kolmogorov *et al.*, 2019) with default parameters. The resulting GFA was examined using Bandage (Wick *et al.*, 2015), allowing to delineate a set of high-coverage (>300X) contigs against a background of low-coverage (<100X) contigs. To verify that these high-coverage contigs corresponded to *Thiovulum*, the published proteome of *Thiovulum* ES (Marshall *et al.*, 2012). was aligned on the GFA using tblastn (Gertz *et al.*, 2006) within Bandage (parameters: minimum identity 70 %, minimum coverage 70 %), revealing that nearly all tblastn hits were concentrated on the high-coverage contigs and vice-versa. The GFA was therefore pruned to retain only the high-coverage contigs, which were all interconnected. The remaining 31 contigs were exported as FASTA then scaffolded using SLR (Luo *et al.*, 2019); the nine resulting scaffolds were mapped back to the GFA to resolve most repeats, and the remaining repeats were resolved manually until obtaining a circular genome. A final polishing step was performed with unicycler-polish from Unicycler v0.4.9b (Wick *et al.*, 2017) using the complete set of Illumina reads (for a total depth of coverage of 12X of the genome) and the subset of Nanopore reads longer than 5 kb (*ca*. 50X). Polishing consisted of two cycles of pilon 1.23 (Walker *et al.*, 2014), one cycle of racon 0.5.0 (Vaser *et al.*, 2017) followed by FreeBase (Garrison and Marth, 2012), then 30 additional cycles of short-read polishing using pilon 1.23, after which the assembly reached its best ALE score (Clark *et al.*, 2013).

The completeness of the *Thiovulum* genome obtained was assessed using CheckM (Parks *et al.*, 2015) and its continuity using the unicycler-check module in Unicycler v0.4.9b. Annotation

was performed using Prokka (Seemann, 2014), KEGG (Kanehisa *et al.*, 2016), the EggNOG 5.0 online tool (Huerta-Cepas *et al.*, 2019), PATRIC (Brettin *et al.*, 2015; Davis *et al.*, 2020) and RAST (Aziz *et al.*, 2008; Overbeek *et al.*, 2014) annotation servers. A COG (Tatusov *et al.*, 2000) analysis was done using the ANVIO tool (Eren *et al.*, 2015). To further elucidate the function of genes annotated as hypothetical proteins, a structural annotation was done using the Superfamily and SCOP databases using the Superfam online interface (Gough *et al.*, 2001).

RESULTS

Field observations

A pale-white loose veil, with a vertical thickness of 2 to 3 mm, was observed below and adjacent to the water surface in Movile Cave (Fig. 2). This resembled microbial veils described earlier as mucilaginous structures often produced by sulfur-oxidizing bacteria (Garcia-Pichel, 1989; Fenchel, 1994) under specific O_2 and H_2S conditions. Nevertheless, in Movile Cave, the dense agglomeration of cells does not form slime or strongly cohesive veils, but a loose cloud-like white veil close to the water surface.

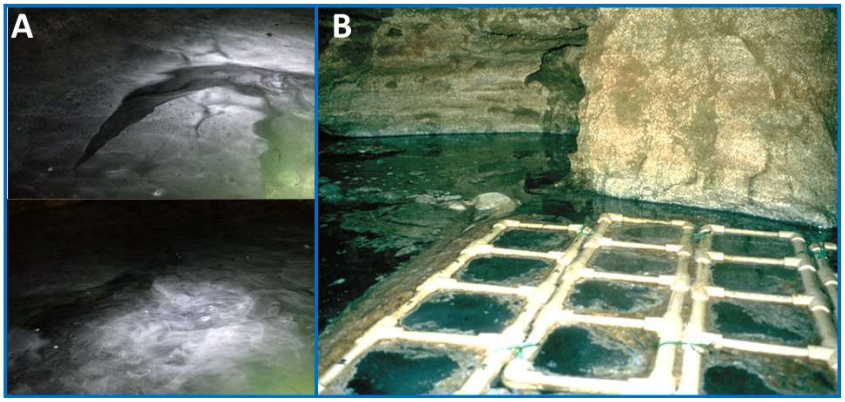


Figure 2. Images of subsurface cloud (veil) from the Lake Room (A) and Air Bell 2 (B) in Movile Cave. See also Supplementary Video 1.

Microscopy

Veils similar to those seen in the Lake Room (Fig. 2A) were also observed in air bell 1 and even more so in air bell 2 (Fig 2B) where they reached the highest densities. In air bell 2 these veils consist of large, spherical to ovoid, bacterial cells (Fig. 3A-B) identified as belonging to the genus *Thiovulum*. These cells have a diameter of 12-16 µm and occur in densities of approximately 5.5×10^3 cells/ml, resembling a dense bacterial culture (Fig. 2, Supplementary Video 1). Each cell contains up to 20-30 sulfur globules of various shapes and sizes that are responsible for the white appearance of the veil (Fig. 3). Transmission Electron Microscope (TEM) observations showed that these large cells are Gram-negative (Fig. 3C-D), like all *Campylobacterota* (formerly *Epsilonproteobacteria; Campylobacterales*), and confirmed the existence of irregularly shaped sulfur inclusions within the cells. Light microscopy suggested *Thiovulum* cells divide along the long cell axis (Fig. 3B). This was confirmed by the TEM images showing two dividing cells sharing the same cytoplasmatic spaces prior to the complete closing of the cell membrane (Fig. 3C). Short peritrichous flagella fragments (Fig. 3D) observed on the surface of the cells resemble those noticed earlier in *Thiovulum* species (Wirsen and Jannasch, 1978). Scanning electron microscopy (SEM) depicts the ball-like structure of the

sulfur inclusions in a series of connected *Thiovulum* cells (Fig. 3E). These cells are connected one to the other via multiples threads (Fig. 3F-G). Energy-dispersive X-ray (EDX) analysis (Fig. 3H-I) revealed that the intracellular globules are dominated by sulfur (20.9 - 26.1 %), along with elements common in organic matter such as carbon (49 - 49.2 %) and oxygen (21.1 - 24.6 %), and a few other elements in low abundance such as sodium (2.4 - 3.4 %) and phosphorus (1.2 - 2.2 %).

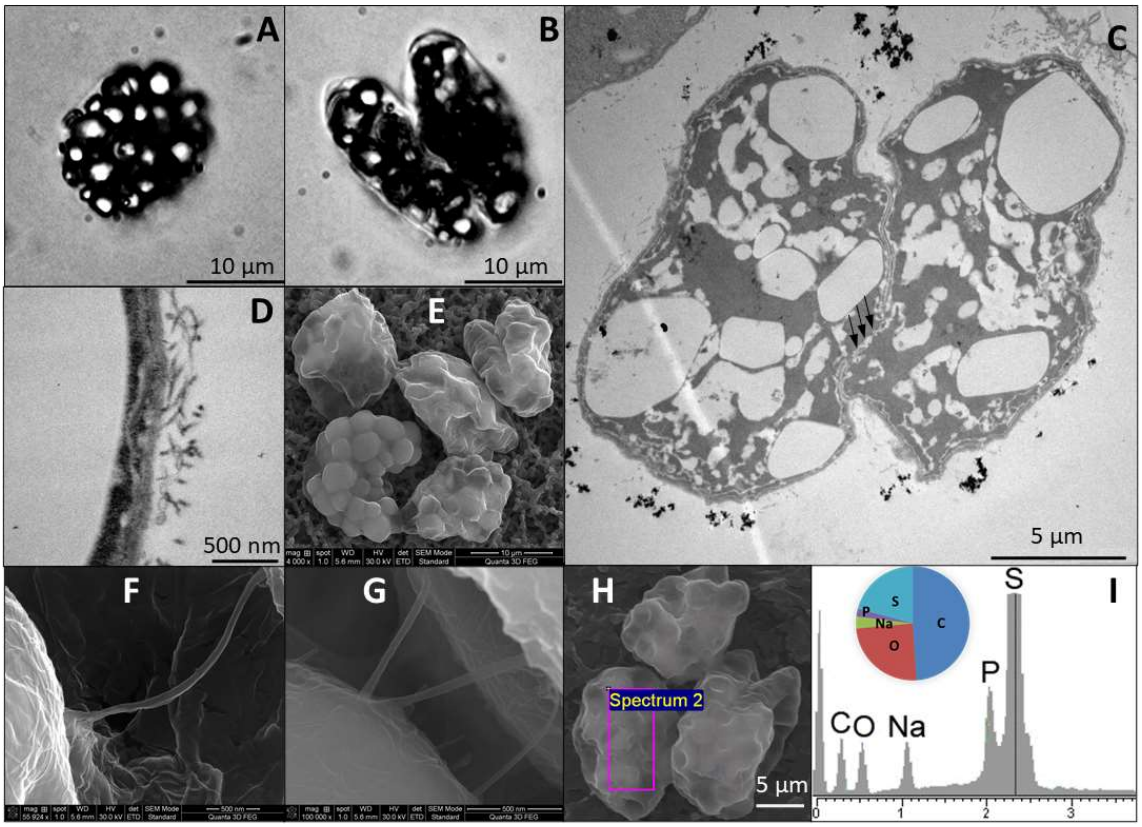


Figure 3. Optical images of giant globular cells colonizing the subsurface cloud/veil from Movile Cave (A-B) including a cell undergoing division (B). Each cell carries up to 30 sulfur inclusions (large bright spots in panels A, B). TEM images of *Thiovulum* show the cellular localization of sulfur inclusions of various shapes and sizes (panel C, white spots). Ovoid cells divide along their long axis (B, C). The region where the cell membrane is not fully closed between dividing cells is marked by with three black arrows (C). The cell wall (D) is covered in pili or short flagella. *Thiovulum* cells (E) are often connected one to another through thread-like structures (F-G). EDX analysis on *Thiovulum* cells (H) inspected under SEM show the typical elemental composition of the cells (I) and confirm the high sulfur content of the internal globules. Note that the height of the peaks in the EDX spectra do not correlate with the element's ratio, but with the X-ray signal intensity.

Laboratory observations

Water samples were collected from the cave and brought to the lab within one hour of sampling for optical microscopy observations. Upon exposure to the microscope light, the cells became very active and moved fast in all directions, a behavior similar to that observed in *Thiovulum majus* (Fenchel, 1994). After a few hours of swimming (Supplementary Video 2), the cells

slowed down and gathered in certain areas on the glass slide, while some cells turned around and swam in the reverse direction.

Within 24 h after being removed from the cave, the cells became inactive and most of them lysed, releasing their sulfur globules. Attempts to obtain the *Thiovulum* cells in culture were unsuccessful; however, we were able to devise a laboratory setup (Fig. S1) that resulted in a *Thiovulum*-enriched culture by mimicking as much as possible the cave conditions (i.e. constant temperature, darkness, high relative humidity, and sulfidic water from the cave).

Stable isotope analysis

Both the carbon and the nitrogen stable isotope ratios of the *Thiovulum* cells collected from the water surface in Movile Cave (Table 1) ($\delta^{13}C$ = -44.53 ‰ and $\delta^{15}N$ = -10.29 ‰) were among the lightest values measured in Movile Cave (Sarbu *et al.*, 1996).

Table 1. Stable isotope ratios of carbon and nitrogen of microbial samples collected in Movile Cave. $\delta^{13}C$ and $\delta^{15}N$ are expressed as ‰.

	$\delta^{13}C$ values (‰)	**$\delta^{15}N$ values (‰)**
Thiovulum microorganisms	-44.53 ‰	-10.29 ‰
Biofilms floating in the cave's air-bells	-44.54 to -45.82 ‰	- 10.09 to -10.84 ‰
Submerged microbial biofilms	-43.23 ‰	-9.98 ‰

Table 2. Stable isotope ratios of sulfur from HS^- from the cave water, S^o from sulfur granules within Thiovulum microorganisms swimming in the cave lake, and SO_4^- from the gypsum crystals covering the cave walls above the lake. $\delta^{34}S$ are expressed as ‰.

	$\delta^{34}S$ values (‰)
HS^- from the cave water	+5.5 and +8.3 ‰
S^o from sulfur granules within *Thiovulum* microorganisms	+4.2 ‰
SO_4^{2-} from the gypsum crystals covering the cave walls above the lake.	+1.1 to + 7.3 ‰

The sulfur stable isotope ratios of the H_2S from the groundwater in Movile Cave ranged between $\delta^{34}S$ = +5.5 and +8.3 ‰, and were close to the values measured for the oxidized sulfur in the gypsum crystals covering the cave walls above the lake: $\delta^{34}S$ = +1.1 to + 7.3 ‰ (Table 2). The sulfur granules from within *Thiovulum* in the cave lake had a $\delta^{34}S$ value of = +5.5 ‰, which is similar to the $\delta^{34}S$ measured for the sulfide in the water and the sulfate (gypsum) in the cave walls. No significant isotopic fractionation appears therefore to take place during the oxidation of HS^- to S^0.

Phylogenetic identification

Following quality control and trimming of the 16S rRNA amplicon library, *ca*. 4700 read pairs could be merged into 400 bp long sequences. Of these, 35 % were annotated as *Thiovulum*. The latter, together with other *Thiovulum* sequences, formed a cluster distinct from the phylogenetically close genus *Sulfurimonas* (Fig. 4A). All Movile Cave *Thiovulum* sequences form one cluster separated into three main subclusters, one of which harbor also the three genomic 16S rRNA sequences. Only one of the three clusters harbors previously known sequences of *Thiovulum*, some of which (DQ295692, FJ165502, KM410698, EF467575) were obtained from sulfidic caves. The 16S rRNA gene of the *Thiovulum* ES (accession number AH003133.2) was removed from the analysis as the sequence contained long stretches of Ns. However, when used, this sequence does not fall within the Movile Cave clusters.

The circular, closed contig obtained from the metagenomic assembly of samples from Movile Cave was taxonomically classified with the GTDB-Tk software against the GTDB-Tk database (Chaumeil *et al.*, 2019). The genome was confirmed as *Thiovulum* sp. by having 76.34 % amino acid identity to the single other available *Thiovulum* genome obtained from Elkhorn Slough (ES), California (*Thiovulum* ES) (Marshall *et al.*, 2012). Accordingly, a phylogenetic tree constructed from a multilocus alignment of single-copy marker genes from, *Thiovulum* ES (Marshall *et al.*, 2012), the *Thiovulum* genome from Movile Cave and all available *Sulfurimonas* genomes resulted as well in the two *Thiovulum* genomes clustering together (Fig. 4B).

To assess the abundance of *Thiovulum* in the different environments of the cave, phyloFlash (Gruber-Vodicka *et al.*, 2020) was used to identify and annotate 16S rRNA reads from the raw metagenomic data. *Thiovulum* was found in highest abundance (sequence frequency) in Air Bell 2 (35 %), followed by the Lake Room (5 %) and submerged microbial mats (0.9 %) (see pie charts in Fig. 1).

Genome analysis

The assembly of metagenomic data from Movile Cave resulted in a closed circular genomic sequence classified as *Thiovulum* sp. The genome length is 1.72 Mbp with a GC content of 28.4 %. Genome completeness was estimated using CheckM (Parks *et al.*, 2015) at 93 %. In absence of sufficient reference genomes for this genus, this value likely represents the full set of marker genes for *Thiovulum*. CheckM estimated a contamination of 0 % and a strain heterogeneity of 0: therefore, the *Thiovulum* genome assembly does not seem to contain any sequences from additional distant or closely related organisms. According to the annotation on the PATRIC server (Brettin *et al.*, 2015; Davis *et al.*, 2020), this genome contains 1849 coding sequences, of which 974 encode proteins of known function, 875 are hypothetical proteins, 35 are tRNAs and 4 are CRISPR array comprising a total of 87 repeats. Subsystem hierarchical annotation shows that among the genes attributed to a Subsystem the largest subsystems were cofactors and vitamins, protein metabolism, and amino acids and derivatives (Fig. 5A). The same annotation conducted de-novo on the *Thiovulum* ES genomes suggests that the Movile and ES strains share a core of 493 Subsystem functions and have 36 and 48 additional unique Subsystem functions, respectively (Fig. 5B and Supplementary Data 1). Despite this large functional similarity between the two genomes, the two genomes turned out to be highly divergent according to their level of sequence similarity with clusters of orthologous genes (COGs; Fig. 5B and Supplementary Data 1).

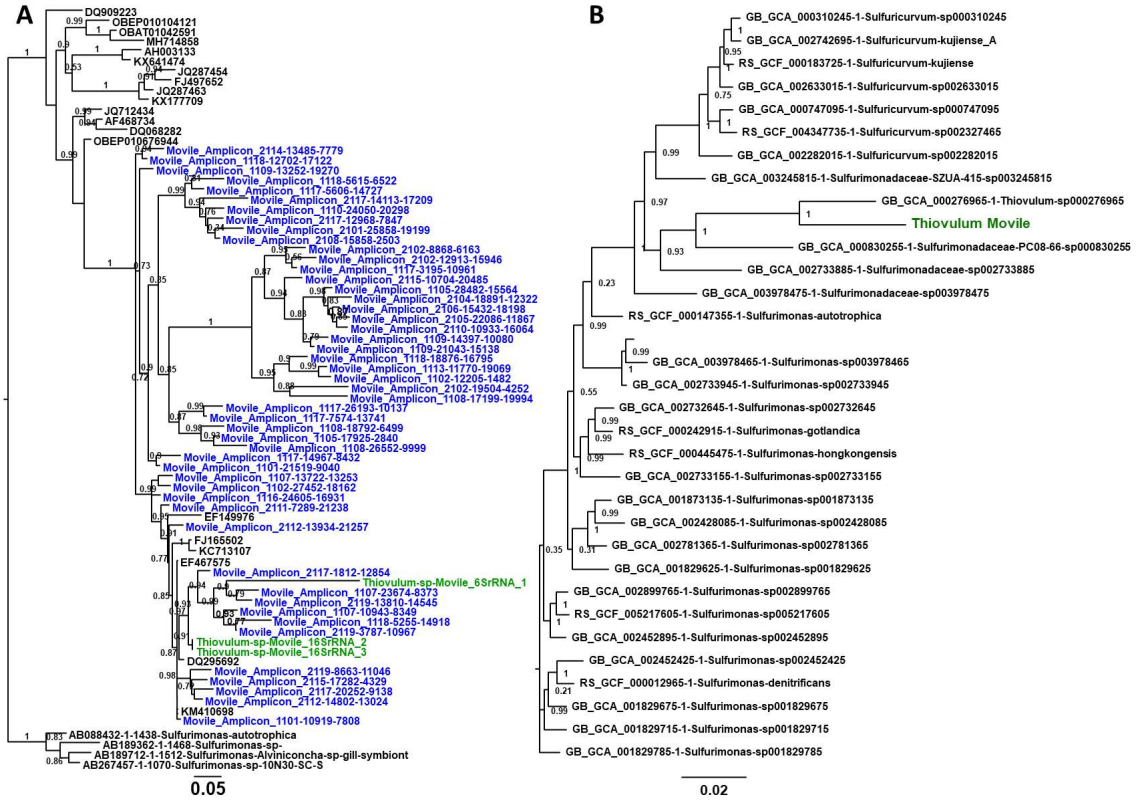


Figure 4. Phylogenetic placement of the 16S rRNA gene of *Thiovulum* from Movile Cave (A) and its genome (B). The Shimodaira-Hasegawa local support values (ranging from 0 to1) are shown next to each node.

Carbon fixation likely takes place via the reductive TCA cycle, similarly to *Thiovulum* ES. For both strains not all necessary genes for the rTCA cycle could be recognized by any single annotation method including EggNOG 5.0 (Huerta-Cepas *et al.*, 2019), KEGG (Kanehisa *et al.*, 2016), COG (Tatusov *et al.*, 2000) or Subsystems.

Both the KEGG and Subsystem annotations revealed only few genes involved in sulfur metabolism, with sulfide:quinone oxidoreductase being responsible for the formation of sulfur globules. Additionally, we identified in both *Thiovulum* strains the tetrathionate reductase subunit A and thiosulfate/3-mercaptopyruvate sulfurtransferase genes potentially involved in the formation of sulfite from tetrathionate via thiosulfate. COG annotation identified in both Movile and ES strains the presence of the bacterial heterodisulfide reductase (*dsrC*) (Venceslau *et al.*, 2014) alongside rhodanase sulfurtransferases, involved in the conversion of sulfide to sulfite. Subsequently, the presence of the *tauE* sulfite exporter gene was identified as well in both strains.

In addition to the membrane-bound nitrate reductases (Nar) found in *Thiovulum* ES, the Movile-Cave *Thiovulum* possesses also periplasmatic nitrate reductases encoded by the *nap* genes. Some of the genes involved in the synthesis of the formate-dependent nitrite reductase were identified in *Thiovulum* ES. These are confirmed in the current genome and are supplemented by precursor genes to the *nrfA* encoded cytochrome c-552 which reduced nitrite to ammonia. Additionally, the gene for hydroxylamine dehydrogenase, which is often encountered in genomes from *Campylobacterota* (Haase *et al.*, 2017), formerly referred to as *Epsilonproteobacteria* (Waite *et al.*, 2019), was also identified.

As *Thiovulum* sp. is a highly motile bacterium, we inspected motility and chemotaxis genes. All genes necessary for flagellar assembly were found, similarly to *Thiovulum* ES. The chemotaxis genes *cheA cheY*, *cheW*, *cheV* and *cheD* were identified as well alongside the *cetA* and *cetB* energy taxis genes, the parallel to the *Escherichia coli* aerotaxis (*aer*) gene. The *cheX* gene, which was not found in the genome of *Thiovulum* ES, was identified in the Movile Cave *Thiovulum*. Gene *cheB* was reported missing in *Thiovulum* ES, was identified in the Movile strain but also in *Thiovulum* ES upon COG reannotation. Additionally, 11 methyl-accepting chemotaxis proteins were identified.

In addition to flagella genes, the *pilA*, *pilE*, *pilT*, *pilN, pulO*, *fimV* genes responsible for the formation and retraction of type IV pili.

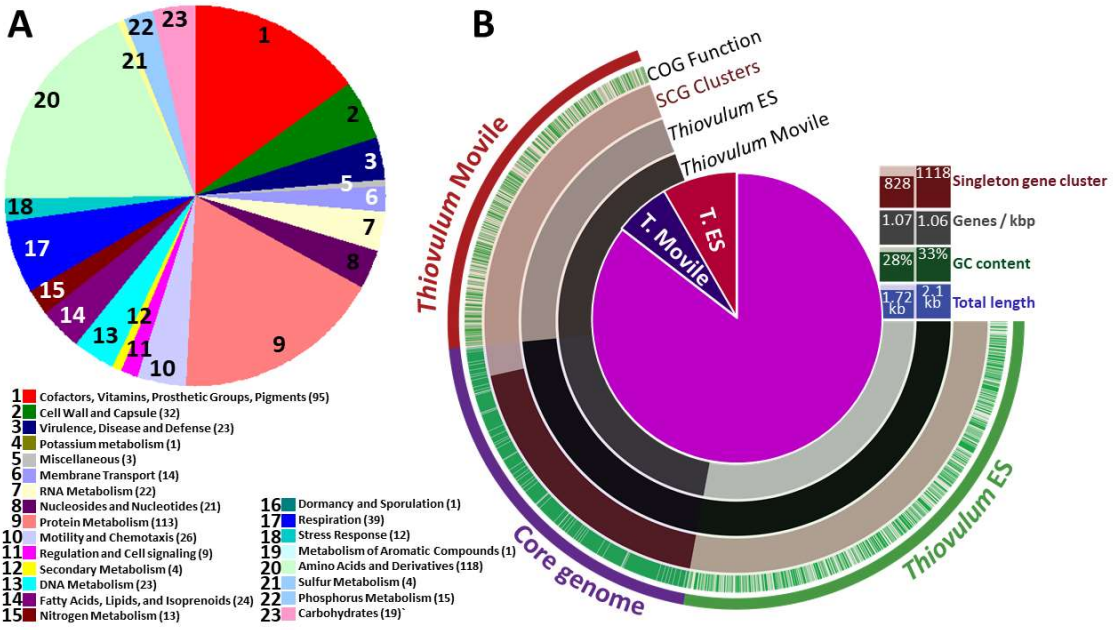


Figure 5. Distribution of genes of the Movile Cave *Thiovulum* associated with different Subsystems (A) and functional comparison to the *Thiovulum* ES genome using Subsystems and COG functional assignments (B) marking in both cases the core genome and the unique COGS and Subsystems.

DISCUSSION

In Movile Cave, the oxidation of reduced compounds such as H_2S, CH_4, and NH_4^+ is the only primary energy source, There, *Thiovulum* sp., a large sulfur oxidizer, often found in close proximity to sediments, microbial mats or surfaces, is part of the chemoautotrophic microbial community involved in *in situ* carbon fixation that represents the food base for the cave's abundant and diverse invertebrate community. These *Thiovulum* cells, exceeding 15 µm in diameter, are larger than most known sulfur-oxidizers: 5-10 µm (Fenchel, 1994) and 5 µm (Thar and Fenchel, 2005) and, together with the larger *Thiomargarita namibiensis* (Schulz *et al.*, 1999) and *Achromatium oxaliferum* (Babenzien, 1991), belong to the group of giant sulfur bacteria (Ionescu and Bizic, 2019). Here we investigated the morphological and genomic aspects of what appears to be a fully planktonic *Thiovulum* sp. The new genome was compared to the sole other existing genome of *Thiovulum*, strain ES. The latter, originating from a phototrophic marine mat, was reannotated for the purpose of this comparison to account for new available information, 8 years after its original publication.

Sulfide oxidation

Thiovulum (*Campylobacterota; Campylobacteria; Campylobacterales; Sulfurimonadaceae*) has been recognized to have several interesting features, including being among the fastest bacterial swimmers and its ability to form large veils consisting of interconnected cells. Many sulfur-oxidizing microorganisms including species of *Thiovulum* form veils by means of what appear to be mucous threads. These threads are used (Thar and Fenchel, 2001) by the cells to attach to solid surfaces (Fauré-Fremiet and Rouiller, 1958; De Boer *et al.*, 1961; Wirsen and Jannasch, 1978; Fenchel, 1994; Robertson *et al.*, 2006). In marine settings, such veils keep cells above sediments (Karavaiko *et al.*, 2006) at the oxic-anoxic interface where the optimal concentration of O_2 and H_2S can be found (Pruesse *et al.*, 2012).

In Movile Cave, the dense agglomeration of cells does not form slime or strongly cohesive veils, but a loose cloud-like white veil close to the water surface. Nevertheless, SEM analyses shows that the cells are at least partially interconnected. It has been hypothesized that the coordinated movement of *Thiovulum* cells generates currents directing H_2S or O_2 to the cells (Petroff and Libchaber, 2014). The fully planktonic localization of the cells in air bell 2 means that *Thiovulum* here cannot use surfaces to localize itself at the oxic anoxic interphase. Therefore, as sulfide concentration is high (245 µM) (Flot *et al.*, 2014), the cells likely need to coordinate their swimming to stay close to the surface where the O_2 necessary to their metabolism is present.

Thiovulum is a sulfide oxidizer as evidenced by its generation and accumulation of sulfur inclusions. Inclusions such as amorphous sulfur, polysulfide granules or orthorhombic sulfur inclusions have been described from other environments (Jorgensen and Revsbech, 1983; Robertson *et al.*, 2006). The amount and type of sulfur inclusions in cells is influenced by the concentrations of H_2S and O_2 in the environment. Typically, cells store elemental sulfur when H_2S is abundant in the environment, and later use the intracellular reserves of sulfur when the sulfide source in the environment is depleted (De Boer *et al.*, 1961). Sulfur inclusions were also shown to form when the supply of O_2 is limited and as a result the sulfur cannot be entirely oxidized to soluble sulfite, thiosulfate, or sulfate. Complete depletion of sulfur inclusions from cells is not likely in Movile Cave where abundant H_2S is available continuously and where O_2 is scarce in most habitats (Sarbu *et al.*, 1996). Furthermore, while genomic analysis of the Movile Cave and ES strains identified the SQR gene involved in this process, the genes required for further oxidizing the elemental sulfur to sulfate are not present.

Marshall *et al.*, (2012) proposed the *Thiovulum* needs to undergo frequent (daily) oxic/anoxic cycles to prevent continuous accumulation of elemental sulfur in the cell. We advance two additional options by which the Movile Cave *Thiovulum* and likely the entire genus may evade sulfur globule intoxication. First, the identification of a the C subunit of dissimilatory sulfite reductase (*dsrC*) (Venceslau *et al.*, 2014) and rhodanese genes, known to be involved in S^0 conversion to sulfite (Poser *et al.*, 2014), and a sulfite exporter (*tauE*), suggest *Thiovulum* may be able oxidize elemental sulfur to sulfite and transport the latter out of the cell. Second, we propose that *Thiovulum* is capable of dissimilatory nitrate reduction to ammonia (DNRA) using elemental sulfur (Slobodkina *et al.*, 2017), a process already shown in *Campylobacterota* (e.g. *Sulfurospirillum deleyianum*) (Eisenmann *et al.*, 1995). The Movile Cave *Thiovulum* contains not only the *nar* (*narGH*) genes for nitrate reduction but also the periplasmatic *nap* genes known for their higher affinity and ability to function in low nitrate concentrations (Pandey *et al.*, 2020). Additionally, we were able to identify the membrane component of the formate-dependent nitrite reductase, (*nrfD*) gene in both Movile and ES strains, as well as the formate-dependent periplasmic cytochrome C nitrite reductase, (subunit C-552) and the ferredoxin subunit of nitrite reductase in the Movile Cave strain. Last, the gene for hydroxylamine dehydrogenase (*hao*) was also identified in the Movile Cave *Thiovulum*. Hydroxylamine dehydrogenase is

known from other *Campylobacterota* (e.g. *Campylobacter fetus* or *Nautilia profundicola*) and was shown to mediate the respiratory reduction of nitrite to ammonia. In line with the findings of Marshall *et al.*, (2012), the *hao* gene was not found in the genome of *Thiovulum* ES upon re-annotation, suggesting that the *hao* gene may not be part of the core *Thiovulum* genome. Interestingly, normally *Campylobacterota* that utilize hydroxylamine dehydrogenase do not have the formate-dependent nitrite reductase. Similarly, *Campylobacterota*, typically use the periplasmic nitrate reductase (Nap) and do not have the membrane bound NarGHI system (Kern and Simon, 2009; Meyer and Huber, 2014). Interestingly, *Thiovulum* ES has only the Nar systems while the Movile Cave strain has both types. Nevertheless, while genomic information is suggestive of the presence or absence of specific enzymes and pathways, experiments or gene expression data are required to determine which of the genes are utilized and under which environmental conditions.

Carbon fixation

Marshall *et al.,* (2012) found all the genes necessary for the reverse TCA cycle to be present in *Thiovulum* ES, as is the case for the Movile Cave strain. However, while the isotopic fractionation of $^{34}S/^{32}S$ between the sulfide in the water and the sulfur granules of the cells matches the literature (*ca*. 2 ± 2 ‰, (Fry *et al.*, 1988), that of the ^{13}C is lighter than expected from an organism utilizing the rTCA cycle (van der Meer *et al.*, 2001). Such depletion in ^{13}C is rather typical of organisms utilizing the Calvin-Benson-Bassham (CBB) cycle (Pearson, 2010). Recently it was discovered that some *Campylobacterota* lost the genes needed for the TCA in favor of a gradual acquisition of those required for the CBB cycle (Assié *et al.*, 2020). Most of the genes involved in the CBB cycle are common in heterotrophs as well being part of alternative reactions. However, two enzymes are unique to the CBB cycle: phosphoribulokinase and ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) (Hügler and Sievert, 2011). Thus, given that $\delta^{13}C$ of cave dissolved inorganic carbon has been documented as low as -25 ‰ (Knierim *et al.*, 2015), and some bacteria using rTCA have a fractionation of -15 to -35 ‰, relative to the dissolved inorganic carbon (Preuß *et al.*, 1989), it is likely that *Thiovulum* has a stronger preference towards the lighter carbon isotopes than other rTCA utilizers.

Cell motility and veil formation

The loose veil formation from Movile Cave appears to move slowly in various directions, most probably driven by slow movement of the individual cells toward zones of optimal chemical concentrations of H_2S and O_2. Similarly, Thar and Fenchel (2001 and 2005) showed that sulfur oxidizers move in nutrient and oxygen gradients and the dynamics and aggregation of cells is due to chemotactic properties toward a particular concentration of oxygen. Accordingly, genes involved in chemotaxis are abundant in the *Thiovulum* genome. Unlike the ES strain the Movile strain contains the CheX phosphatase rather than CheC, highlighting the high variability and diversity among bacterial chemotaxis systems (Wuichet and Zhulin, 2010; Bardy *et al.*, 2017). Methyl-accepting chemotaxis proteins (MCPs) are key elements in the bacterial chemotactic response to environmental signal (Salah Ud-Din and Roujeinikova, 2017). Motile bacteria carry between 3-30 (average 13) MCPs with *Campylobacterota* having *ca*. 10 (Lacal *et al.*, 2010). We identified 11 MCPs in the Movile Cave *Thiovulum* two of which matched known sequences while the other 9 were identified using secondary structure, adding to the already large known variability of these proteins.

Thiovulum sp. often forms large veils of interconnected cells. The threads connecting the cells are thought to be secreted by the antapical organelle, located at the posterior side of the cell (De

Boer *et al.*, 1961; Robertson *et al.*, 2015). Short peritrichous flagella-like fragments (Fig. 3D) observed on the surface of the cells resemble those noticed earlier in *Thiovulum* species (Wirsen and Jannasch, 1978). All genes necessary for type IV pili and flagella assembly were found in the Movile Cave and ES *Thiovulum* sp. (Marshall *et al.*, 2012). Accordingly, together with evaluating available electron microscopy we suggest that these ideas need be revisited.

First, our own SEM images (as well as previous ones of connected *Thiovulum* cells) show connecting threads that are not exclusively polar and are much thinner than the stalk like structure shown by de Boer *et al*., (1961). We propose that these structures are rather type IV pili, which are known, among other functions, to connect cells to surfaces or other cells (Craig *et al.*, 2019). Alternatively, Bhattacharya *et al.*, (2019) have shown the formation of cell-connecting nanotubes constructed using the same enzymatic machinery used for flagella assembly. However, it is not possible to determine this in absence of TEM images of connected cells showing the presence of the reduced flagellar base Bhattacharya *et al*., (2019).

Second, we question the flagellar nature of the peritrichous structures around the cells. Inspecting the high-resolution electron microscopy images taken by Faure-Fremie and Rouiller (1958), there is no single structure visible resembling a flagellar motor. Additionally, the length of these structures based on de Boer *et al*., (1961) and a similar image in Robertson *et al*., (2015) suggest that these <3 µm structures are too short for typical flagella (> 10 µm in length (e.g. (Renault *et al.*, 2017)) and are closer to the 1-2 µm lengths known type for pili. Interestingly, *Caulobacter crescentus* swims at speeds of up to 100 µm s^{-1} using a single flagellum aided by multiple pili (Gao *et al.*, 2014). In light of this hypothesis, we inspected the images of the fibrillar organelle at the antapical pole of the cells. The high-resolution images presented in Fauré-Fremiet and Rouiller (1958) and in de Boer (1961) show an area of densely packed fibrillar structures. Considering our current knowledge in flagellar motor size (*ca*. 20 nm) it is highly unlikely that each of these fibers is an individual flagellum, thus, potentially representing a new flagellar organization. Petroff *et al*., (2015) investigated the physics behind the 2-dimensional plane assembly of *Thiovulum* veils and suggested it to be a direct result from the rotational movement which attracts cells to each other. Nevertheless, as seen, SEM images show cells are physically attached one to the other suggesting several mechanisms and steps may be involved. Interestingly, type IV pili retraction can generate forces up to 150 pN which are known to be involved in twitching motility in bacteria (Craig *et al.*, 2019). If coordinated, these may be part of the answer to the unexplained swimming velocity of *Thiovulum* (Garcia-Pichel, 1989), which is, at *ca*. 615 µm s^{1}, 5 to 10 times higher than that of other flagellated bacteria. Thus, genomic information and re-evaluation of electron microscopy data raise new questions concerning the nature of the extracellular structures on the surface of *Thiovulum* sp. and call for new targeted investigations into this topic.

CONCLUSIONS

Movile Cave is a system entirely depending on chemosynthesis. We showed that submerged near-surface, planktonic microbial accumulations are dominated by *Thiovulum*, a giant bacterium typically associated with photosynthetic microbial mats. The genome of this planktonic *Thiovulum* strain suggests that it can perform dissimilatory reduction of nitrate to ammonium. As this may be a main mechanism in preventing continuous accumulation of sulfur globules in the cells, *Thiovulum* likely plays a major role in the nitrogen cycle of the cave, providing readily available ammonia to the surrounding microorganisms. Kumaresan *et al.,* (2014), based as well on genetic evidence, suggest nitrogen fixation and deamination of methylamines as the source of ammonium in the cave. The coupling of DNRA to sulfide oxidation provides a direct and more productive source of ammonium. Our results show that

the Movile Cave strain is similar to other cave *Thiovulum* sp., suggesting that in those subterranean ecosystems *Thiovulum* may as well play a role in both N and S cycles.

This investigation of the genome, coupled with observations of current and previous microscopy images, question the number of flagella the cells has, bringing forth the possibility that the cell may make use of type IV pili for its rapid movement and cell-to-cell interactions.

The collective behavior of *Thiovulum* is still a puzzle and there may be more than one mechanism keeping the cells connected in clusters or in veils. Our SEM images suggest the cells are connected by thread-like structures. Petroff *et al*., (2015) show, on another strain, there is no physical connection between the cells. They suggest that the swimming behavior of individual cells keeps the cells together. More research is therefore needed to understand if these different mechanisms are driven by strain variability, or by different environmental conditions.

ACKNOWLEDGEMENTS

The authors thank GESS team for logistics with sampling in the cave and Viorel Atudorei for his valuable help with the interpretation of the stable isotope data. We would also like to thank Pheobe Laaguiby at the University of Vermont for performing the outstanding Oxford Nanopore sequencing and Bo Barker Jørgensen, Emily Fleming, Carl Wirsen, and Tom Fenchel for their valuable suggestions that led to the improvement of the quality of this manuscript. We thank Luca Zoccarato for the help with genomic graphical representation. We would also like to thank the extreme microbiome project for providing the DNA extraction reagents and methods as well as Laura Gray and Mehdi Keddache at Illumina Corp for providing partial sequencing reagents through its partnership with extreme microbiome project. T. Brad was supported by a grant of Ministry of Research and Innovation, project number PN-III-P4-ID-PCCF-2016-0016 (DARKFOOD), and by EEA Grants 2014-2021, under Project contract no. 4/2019 (GROUNDWATERISK). J.W. Aerts acknowledges the support from a grant from the User Support Programme Space Research (grant ALW-GO/13-09) of the Netherlands Organization for Scientific Research (NWO). M. Bizic was additionally funded through the DFG Eigene Stelle project (BI 1987/2-1). The computational resources for the assembly of the *Thiovulum* genome were provided to J.-F. Flot by the Consortium des Équipements de Calcul Intensif (CÉCI) funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11.

REFERENCES

Assié, A., Leisch, N., Meier, D. V, Gruber-Vodicka, H., Tegetmeyer, H.E., Meyerdierks, A., *et al*. (2020) Horizontal acquisition of a patchwork Calvin cycle by symbiotic and free-living Campylobacterota (formerly Epsilonproteobacteria). *ISME J* 14: 104–122.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., *et al*. (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9: 75.

Babenzien, H.-D. (1991) *Achromatium oxaliferum* and its ecological niche. *Zentralbl Mikrobiol* 146: 41–49.

Bardy, S.L., Briegel, A., Rainville, S., and Krell, T. (2017) Recent advances and future prospects in bacterial and archaeal locomotion and signal transduction. *J Bacteriol* 199: e00203-17.

Bauermeister, J., Ramette, A., and Dattagupta, S. (2012) Repeatedly evolved host-specific ectosymbioses between sulfur-oxidizing bacteria and amphipods living in a cave ecosystem. *PLoS One* 7: e50254.

Bhattacharya, S., Baidya, A.K., Pal, R.R., Mamou, G., Gatt, Y.E., Margalit, H., *et al*. (2019) A ubiquitous platform for bacterial nanotube biogenesis. *Cell Rep* 27: 334-342.e10.

De Boer, W.E., La Rivière, J.W.M., and Houwink, A.L. (1961) Observations on the morphology of *Thiovulum majus* Hinze. *Antonie Van Leeuwenhoek* 27: 447–456.

Brad, T., Fišer, C., Flot, J.F., and Sarbu, S.M. (2015) *Niphargus dancaui* sp. nov. (Amphipoda, Niphargidae) - A new species thriving in sulfidic groundwaters in southeastern Romania. *Eur J Taxon* 2015: (164).

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., *et al*. (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5: 8365.

Bushnell, B., Rood, J., and Singer, E. (2017) BBMerge – Accurate paired shotgun read merging via overlap. *PLoS One* 12: e0185056.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., *et al*. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* 108: 4516 LP – 4522.

Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019) GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics.

Chen, Y., Wu, L., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H., *et al*. (2009) Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J* 3: 1093–1104.

Clark, S.C., Egan, R., Frazier, P.I., and Wang, Z. (2013) ALE: a generic assembly likelihood evaluation framework for assessing the accuracy of genome and metagenome assemblies. *Bioinformatics* 29: 435–443.

Craig, L., Forest, K.T., and Maier, B. (2019) Type IV pili: dynamics, biophysics and functional consequences. *Nat Rev Microbiol* 17: 429–440.

Davis, J.J., Wattam, A.R., Aziz, R.K., Brettin, T., Butler, R., Butler, R.M., *et al*. (2020) The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res* 48: D606–D612.

Eisenmann, E., Beuerle, J., Sulger, K., Kroneck, P.M.H., and Schumacher, W. (1995) Lithotrophic growth of *Sulfurospirillum deleyianum* with sulfide as electron donor coupled to respiratory reduction of nitrate to ammonia. *Arch Microbiol* 164: 180–185.

Engel, A.S. (2015) Bringing microbes into focus for speleology: an introduction. In Engel, A.S (ed), Microbial Life of Cave Systems. Berlin, München, Boston: De Gruyter, pp. 1–22.

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., and Delmont, T.O. (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3: e1319.

Fauré-Fremiet, E. and Rouiller, C. (1958) Étude au microscope électronique d'une bactérie sulfureuse, *Thiovulum majus* Hinze. *Exp Cell Res* 14: 29–46.

672 Fenchel, T. (1994) Motility and chemosensory behaviour of the sulphur bacterium
673 *Thiovulum majus*. *Microbiology* 140: 3109–3116.

674 Flot, J.-F., Bauermeister, J., Brad, T., Hillebrand-Voiculescu, A., Sarbu, S.M., and
675 Dattagupta, S. (2014) *Niphargus–Thiothrix* associations may be widespread in sulphidic
676 groundwater ecosystems: evidence from southeastern Romania. *Mol Ecol* 23: 1405–1417.

677 Fry, B., Gest, H., and Hayes, J.M. (1988) $^{34}S/^{32}S$ fractionation in sulfur cycles catalyzed
678 by anaerobic bacteria. *Appl Environ Microbiol* 54: 250–256.

679 Ganzert, L., Schirmack, J., Alawi, M., Mangelsdorf, K., Sand, W., Hillebrand-
680 Voiculescu, A., and Wagner, D. (2014) *Methanosarcina spelaei* sp. nov., a methanogenic
681 archaeon isolated from a floating biofilm of a subsurface sulphurous lake. *Int J Syst Evol*
682 *Microbiol* 64: 3478–3484.

683 Gao, Y., Neubauer, M., Yang, A., Johnson, N., Morse, M., Li, G., and Tang, J.X. (2014)
684 Altered motility of *Caulobacter crescentus* in viscous and viscoelastic media. *BMC Microbiol*
685 14: 322.

686 Garcia-Pichel, F. (1989) Rapid bacterial swimming measured in swarming cells of
687 *Thiovulum majus*. *J Bacteriol* 171: 3560–3563.

688 Garrison, E. and Marth, G. (2012) Haplotype-based variant detection from short-read
689 sequencing. *arXiv* 1207.3907.

690 Gertz, E.M., Yu, Y.-K., Agarwala, R., Schäffer, A.A., and Altschul, S.F. (2006)
691 Composition-based statistics and translated nucleotide searches: improving the TBLASTN
692 module of BLAST. *BMC Biol* 4: 41.

693 Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001) Assignment of homology
694 to genome sequences using a library of hidden Markov models that represent all proteins of
695 known structure. *J Mol Biol* 313: 903–919.

696 Gruber-Vodicka, H.R., Seah, B.K.B., and Pruesse, E. (2020) phyloFlash: Rapid small-
697 subunit rRNA profiling and targeted assembly from metagenomes. *mSystems* 5: 521922.

698 Haase, D., Hermann, B., Einsle, O., and Simon, J. (2017) Epsilonproteobacterial
699 hydroxylamine oxidoreductase (εHao): characterization of a "missing link" in the multihaem
700 cytochrome *c* family. *Mol Microbiol* 105: 127–138.

701 Hayat, M. (2001) Principles and techniques of electron microscopy: biological
702 applications, 4th ed. London: Cambridge University Press.

703 Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., and Andersson,
704 A.F. (2011) Transitions in bacterial communities along the 2000 km salinity gradient of the
705 Baltic Sea. *ISME J* 5: 1571–1579.

706 Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010) CD-HIT Suite: a web server for
707 clustering and comparing biological sequences. *Bioinformatics* 26: 680–2.

708 Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K.,
709 Cook, H., *et al*. (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated
710 orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 47: D309–
711 D314.

712 Hügler, M. and Sievert, S.M. (2011) Beyond the Calvin cycle: Autotrophic carbon
713 fixation in the ocean. *Ann Rev Mar Sci* 3: 261–289.

Hutchens, E., Radajewski, S., Dumont, M.G., McDonald, I.R., and Murrell, J.C. (2004) Analysis of methanotrophic bacteria in Movile Cave by stable isotope probing. *Environ Microbiol* 6: 111–120.

Ionescu, D. and Bizic, M. (2019) Giant bacteria. In *eLS*. Chichester, UK: John Wiley & Sons, Ltd, pp. 1–10.

Jorgensen, B.B. and Revsbech, N.P. (1983) Colorless sulfur bacteria, *Beggiatoa* spp. and *Thiovulum* spp., in O_2 and H_2S microgradients. *Appl Environ Microbiol* 45: 1261–1270.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016) KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44: D457–D462.

Karavaiko, G.I., Dubinina, G., and Kondrat'eva, T. (2006) Lithotrophic microorganisms of the oxidative cycles of sulfur and Iron. *Mikrobiologiia* 75: 593–629.

Kern, M. and Simon, J. (2009) Electron transport chains and bioenergetics of respiratory nitrogen metabolism in *Wolinella succinogenes* and other Epsilonproteobacteria. *Biochim Biophys Acta - Bioenerg* 1787: 646–656.

Knierim, K.J., Pollock, E.D., Hays, P.D., and Khojasteh, J. (2015) Using stable isotopes of carbon to investigate the seasonal variation of carbon transfer in a northwestern Arkansas cave. *J Cave Karst Stud* 77: 12–27.

Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019) Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37: 540–546.

Kumaresan, D., Wischer, D., Stephenson, J., Hillebrand-Voiculescu, A., and Murrell, J.C. (2014) Microbiology of Movile Cave — a chemolithoautotrophic ecosystem. *Geomicrobiol J* 31: 186–193.

Lacal, J., García-Fontana, C., Muñoz-Martínez, F., Ramos, J.-L., and Krell, T. (2010) Sensing of environmental signals: classification of chemoreceptors according to the size of their ligand binding regions. *Environ Microbiol* 12: 2873–2884.

Lascu, C., Popa, R., and Sarbu, S.M. (1994) Le karst de Movile (Dobrogea de Sud). *Rev Roum Géogr* 38: 85–94.

Luo, J., Lyu, M., Chen, R., Zhang, X., Luo, H., and Yan, C. (2019) SLR: a scaffolding algorithm based on long reads and contig classification. *BMC Bioinformatics* 20: 539.

Marshall, I.P.G., Blainey, P.C., Spormann, A.M., and Quake, S.R. (2012) A single-cell genome for *Thiovulum* sp. *Appl Environ Microbiol* 78: 8555–8563.

van der Meer, M.T.J., Schouten, S., Rijpstra, W.I.C., Fuchs, G., and Sinninghe Damsté, J.S. (2001) Stable carbon isotope fractionations of the hyperthermophilic crenarchaeon *Metallosphaera sedula*. *FEMS Microbiol Lett* 196: 67–70.

Meyer, J.L. and Huber, J.A. (2014) Strain-level genomic variation in natural populations of *Lebetimonas* from an erupting deep-sea volcano. *ISME J* 8: 867–880.

Muyzer, G., Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* 59: 695–700.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., *et al*. (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42: D206–D214.

Pandey, C.B., Kumar, U., Kaviraj, M., Minick, K.J., Mishra, A.K., and Singh, J.S. (2020) DNRA: A short-circuit in biological N-cycling to conserve nitrogen in terrestrial ecosystems. *Sci Total Environ* 738: 139710.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25: 1043–1055.

Pearson, A. (2010) Pathways of carbon assimilation and their impact on organic matter values $\delta^{13}C$. In Timmis, K.N., *Handbook of Hydrocarbon and Lipid Microbiology*., pp. 143–156.

Petroff, A. and Libchaber, A. (2014) Hydrodynamics and collective behavior of the tethered bacterium *Thiovulum majus*. *Proc Natl Acad Sci U S A* 111: E537 LP-E545.

Petroff, A.P., Wu, X.-L., and Libchaber, A. (2015) Fast-moving bacteria self-organize into active two-dimensional crystals of rotating cells. *Phys Rev Lett* 114: 158102.

Poser, A., Vogt, C., Knöller, K., Ahlheim, J., Weiss, H., Kleinsteuber, S., and Richnow, H.H. (2014) Stable sulfur and oxygen isotope fractionation of anoxic sulfide oxidation by two different enzymatic pathways. *Environ Sci Technol* 48: 9094–9102.

Preuß, A., Schauder, R., Fuchs, G., and Stichler, W. (1989) Carbon isotope fractionation by autotrophic bacteria with three different CO_2 fixation pathways. *Zeitschrift fur Naturforsch - Sect C J Biosci* 44: 397–402.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2 – Approximately maximum-likelihood trees for large alignments. *PLoS One* 5: e9490.

Pruesse, E., Peplies, J., and Glöckner, F.O. (2012) SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28: 1823–1829.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al*. (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41: D590–D596.

Renault, T.T., Abraham, A.O., Bergmiller, T., Paradis, G., Rainville, S., Charpentier, E., *et al*. (2017) Bacterial flagella grow through an injection-diffusion mechanism. *Elife* 6: e23136.

Riess, W., Giere, O., Kohls, O., and Sarbu, S. (1999) Anoxic thermomineral cave waters and bacterial mats as habitat for freshwater nematodes. *Aquat Microb Ecol* 18: 157–164.

Robertson, L.A., Gijs Kuenen, J., Paster, B.J., Dewhirst, F.E., and Vandamme, P. (2015) *Thiovulum*. In Trujillo, M.E., Dedysh, S., DeVos, P., Hedlund, B., Kämpfer, P., Rainey, F.A., and Whitman, W.B. (eds), *Bergey's Manual of Systematics of Archaea and Bacteria.* Wiley, pp. 1–4.

Robertson, L.A., Kuenen, J.G., Paster, B.J., Dewhirst, F.E., and Vandamme, P. (2006) *Thiovulum* Hinze 1913, 195AL. In Brenner, D.J., Krieg, N.R., Staley, J.T., and Garrity, G.M. (eds), *Bergey's Manual® of Systematic Bacteriology.* New York: Springer, pp. 1189–1191.

Rohwerder, T., Sand, W., and Lascu, C. (2003) Preliminary evidence for a sulphur cycle in Movile Cave, Romania. *Acta Biotechnol* 23: 101–107.

Salah Ud-Din, A.I.M. and Roujeinikova, A. (2017) Methyl-accepting chemotaxis proteins: a core sensing element in prokaryotes and archaea. *Cell Mol Life* Sci 74: 3293–3303.

Sarbu, S.M. (2000) Movile Cave: A chemoautotrophically based groundwater ecosystem. In Wilken, H., Culver, D.C., and Humphreys, W.F. (eds), *Subterranean Ecosystems. Amsterdam*: Elsevier, pp. 319–343.

Sarbu, S.M. and Kane, T.C. (1995) A subterranean chemoautotrophically based ecosystem. *NSS Bull* 57: 91–98.

Sarbu, S.M., Kane, T.C., and Kinkle, B.K. (1996) A chemoautotrophically based cave ecosystem. *Science* 272: 1953–1955.

Sarbu, S.M. and Lascu, C. (1997) Condensation corrosion in Movile cave, Romania. *J Cave Karst Stud* 59: 99–102.

Sarbu, S.M., Lascu, C., and Brad, T. (2019) Dobrogea: Movile Cave. In Ponta, G.M,. and Onac, B.P. (eds), *Cave and karst systems of Romania*. Springer International Publishing, pp. 429–436.

Sarbu, S.M. and Popa, R. (1992) A unique chemoautotrophically based cave ecosystem. In Camacho, A.I. (ed), *The natural history of biospeleology*. Madrid: Mus. Nat. de Hist. Naturales, pp. 637–666.

Schirmack, J., Mangelsdorf, K., Ganzert, L., Sand, W., Hillebrand-Voiculescu, A., and Wagner, D. (2014) *Methanobacterium movilense* sp. nov., a hydrogenotrophic, secondary-alcohol-utilizing methanogen from the anoxic sediment of a subsurface lake. Int J Syst Evol Microbiol 64: 522–527.

Schulz, H.N., Brinkhoff, T., Ferdelman, T.G., Hernández Mariné, M., Teske, A., and Jørgensen, B.B. (1999) Dense populations of a giant sulfur bacterium in Namibian shelf sediments. *Science* 284: 493–495.

Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30: 2068–2069.

Slobodkina, G.B., Mardanov, A. V, Ravin, N. V, Frolova, A.A., Chernyh, N.A., Bonch-Osmolovskaya, E.A., and Slobodkin, A.I. (2017) Respiratory ammonification of nitrate coupled to anaerobic oxidation of elemental sulfur in deep-sea autotrophic thermophilic bacteria. *Front Microbiol* 8: 87.

Stocchino, G.A., Sluys, R., Kawakatsu, M., Sarbu, S.M., and Manconi, R. (2017) A new species of freshwater flatworm (Platyhelminthes, Tricladida, Dendrocoelidae) inhabiting a chemoautotrophic groundwater ecosystem in Romania. *Eur J Taxon* 2017: 1–21.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E. V (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28: 33–36.

Thar, R. and Fenchel, T. (2005) Survey of motile microaerophilic bacterial morphotypes in the oxygen gradient above a marine sulfidic sediment. *Appl Environ Microbiol* 71: 3682–3691.

Thar, R. and Fenchel, T. (2001) True chemotaxis in oxygen gradients of the sulfur-oxidizing bacterium *Thiovulum majus*. *Appl Environ Microbiol* 67: 3299–3303.

Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017) Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27: 737–746.

Venceslau, S.S., Stockdreher, Y., Dahl, C., and Pereira, I.A.C. (2014) The "bacterial heterodisulfide" DsrC is a key protein in dissimilatory sulfur metabolism. *Biochim Biophys Acta - Bioenerg* 1837: 1148–1164.

Waite, D.W., Chuvochina, M.S., and Hugenholtz, P. (2019) Road map of the phylum *Campylobacterota*. In Trujillo, M.E., Dedysh, S., DeVos, P., Hedlund, B., Kämpfer, P., Rainey, F.A., and Whitman, W.B. (eds) *Bergey's Manual of Systematics of Archaea and Bacteria, Major Reference Works*. pp. 1–11.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., *et al*. (2014) Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9: e112963.

Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13: e1005595.

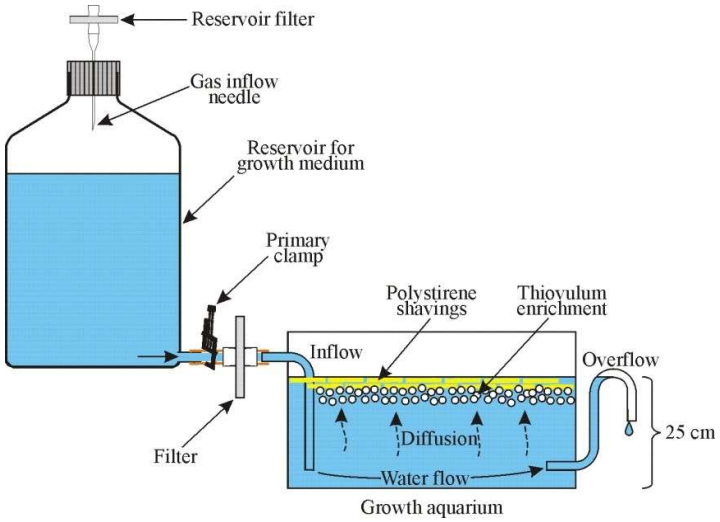
Wick, R.R., Schultz, M.B., Zobel, J., and Holt, K.E. (2015) Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31: 3350–3352.

Wirsen, C.O. and Jannasch, H.W. (1978) Physiological and morphological observations on *Thiovulum* sp. *J Bacteriol* 136: 765–774.

Wuichet, K. and Zhulin, I.B. (2010) Origins and diversification of a complex signal transduction system in prokaryotes. *Sci Signal* 3: ra50 LP-ra50.

860 Supplementary material

861



862

863

864 Fig. S1. Laboratory setup that produced an enrichment culture with high abundance of
865 *Thiovulum* cells from Movile Cave. The water source contained a mineral growth solution and
866 0.25 mM H_2S at pH 7. A floating layer approximately 5-10 mm thick of polystyrene (each 1-3
867 mm thin and 2-4 cm long) shavings was added to the water surface to maintain an anchoring
868 substrate for the cells near the water surface and also to limit the diffusion of O_2 in the water.
869 The setup was installed in a ventilated hood in the lab at 20 °C and the culture was established
870 within two weeks. This setup produced abundant *Thiovulum* growth, but because of its design,
871 it could not be sterilized and therefore did not allow producing a pure *Thiovulum* culture.