

# **Alma Mater Studiorum – Università di Bologna**

---

DOTTORATO DI RICERCA IN  
Biologia Cellulare e Molecolare

Ciclo XXXII

Settore Concorsuale: 05/12  
Settore Scientifico Disciplinare: BIO/19 Microbiologia Generale

## **Microbial diversity and metabolic potential in caves**

Presentata da:  
Daniele Ghezzi

Coordinatore di Dottorato:  
Prof. Giovanni Capranico

Supervisore:  
Prof. Davide Zannoni

Co-supervisori:  
Dr. Martina Cappelletti  
Dr. Francesco Sauro

Esame finale – Anno 2020

## **Abstract**

Caves are dark and oligotrophic habitats where chemotrophic microbial communities interact with the inorganic mineral rocks and typically organize themselves in complex biological formations, which are visible as biofilms, biodeposits or biospeleothems. In these environments, microorganisms contribute to the turnover of the matter and activate peculiar enzymatic reactions leading to the modification of the mineral rocks and to the production of metabolites with industrial and pharmaceutical potentials. In this PhD thesis, various molecular and geomicrobiological approaches were used to investigate the microbial diversity and potential activities in different cave systems, i.e. the orthoquartzite cave Imawarì Yeuta, the sulfidic cave Fetida and the ice cave Cenote Abyss. This study is aimed at gathering indications on the possible interactions that support microbial growth and its impact in cave environments. As a result, microbial taxa and functions associated to light-independent chemolithotrophic and heterotrophic activities were identified in the three caves, indicating the involvement of microorganisms in i) silica mobilization and amorphization processes and the formation of a novel type of silica-based stromatolite in Imawarì Yeuta Cave, ii) the formation of three types of biofilm/biodeposit involved in the sulphur cycle and in the speleogenesis of Fetida Cave, iii) the development of biofilms and their maintenance under psychrophilic conditions in samples collected from ice in Cenote Abyss. Additionally, the metabolic potential of more than one hundred isolates derived from these cave systems were evaluated in terms of antimicrobial activity. The results pointed out that unexplored and oligotrophic caves are promising environments for novel bioactive molecules discovery.

# Index

1. General Introduction	4
2. General Materials and Methods	8
2.1. Sampling	8
2.2. Total DNA extraction and Illumina sequencing of V4 and V5 hypervariable regions of 16S rRNA	8
2.3. Processing of 16S rRNA Illumina sequencing data and statistical analysis	9
2.4. Shotgun metagenomic sequencing and analysis	9
2.5. X-Ray fluorescence spectrometry and X-ray diffraction	9
2.6. Geochemical analyses	10
2.7. Scanning electron microscope (SEM)	11
3. Geomicrobiology of the orthoquartzite cave Imawarì Yeuta	12
3.1. Introduction	12
3.2. Materials and Methods	15
3.2.1. Samples list, DNA extraction and sequencing	15
3.2.2. Full-length 16S rRNA reconstruction and analysis through EMIRGE algorithm	16
3.2.3. q-PCR	17
3.2.4. Clone libraries of full-length 16S rRNA, <i>coxL</i> and <i>hypD</i> genes and screening through Restriction Fragment Length Polymorphism (RFLP)	17
3.2.5. Phylogenetic and statistical analyses	18
3.2.6. Geochemical analyses	19
3.3 Results and Discussion	19
3.3.1. Diversity and statistical analyses of all the microbial samples collected from Imawarì Yeuta	19
3.3.2. Characterization of the microbial communities associated to different silica amorphization levels	24
<i>Samples description</i>	24
<i>Microbial diversity featuring each type of speleothem</i>	25
<i>Microbial role in silica mobilisation</i>	30
<i>Comparison between OTU- and SV-based analyses for microbial diversity characterization</i>	34
3.3.3. Characterization of microbial communities associated to the inception of a novel type of siliceous stromatolite discovered in Imawarì Yeuta cave	35
<i>Samples description</i>	35
<i>Microbial diversity of the three sequential steps of stromatolite formation</i>	36
<i>The microbial community shifting during stromatolite formation</i>	44
<i>Clone libraries of <i>coxL</i> and <i>hypD</i> genes for Qz sample</i>	45
3.3.4. Carbon and nitrogen fixation potential in the metagenome of wall-related microbial communities	49
4. Geomicrobiology of the seawater-influenced sulfuric acid cave Fetida	53
4.1. Introduction	53
4.2. Materials and Methods	54
4.2.1. Samples list, DNA extraction and sequencing	54

---

4.2.2. Geochemical analyses	55
4.3 Results and Discussion	55
4.3.1. Samples description	55
4.3.2. Geochemistry, distribution and morphology of Fetida Cave biofilms	56
4.3.3. Microbial diversity in Fetida Cave biofilms	59
<i>Water filaments</i>	62
<i>Vermiculations</i>	66
<i>Gypsum moonmilk deposits</i>	70
4.3.4. Metagenomic insights into vermiculations	73
5. Geomicrobiology of the high-altitude ice cave Cenote Abyss	79
5.1. Introduction	79
5.2. Materials and Methods	80
5.2.1. Samples list, DNA extraction and sequencing	80
5.3. Results and Discussion	80
5.3.1. Samples description	80
5.3.2. Results on the diversity and taxonomy of the microbial communities in the three Cenote Abyss samples	82
5.3.3. Discussion on the role of microbial communities in the Cenote Abyss cave	86
6. Metabolic and phenotypic analyses of microbial strains isolated from different cave systems	89
6.1. Introduction	89
6.2. Materials and Methods	90
6.2.1. Bacterial strain isolation, purification and top-agar bioassay	90
6.2.2. Extraction of secondary metabolites and top-agar bioassay	91
6.2.3. Sample preparation and HPLC-MS analysis	91
6.2.4. Characterization of bacterial 16S rRNA and fungal ITS from pure cultures	91
6.2.5. <i>In vitro</i> silica solubilization test	92
6.2.6. Phenotype Microarray	92
6.3. Results and Discussion	93
6.3.1. Antimicrobial activity of cave isolates	93
<i>Library of isolates and inhibitory activity against pathogens</i>	93
<i>Taxonomy characterization of isolates</i>	94
<i>Evaluation of the production of bioactive molecules by ten active isolates</i>	97
<i>Phenotype Microarray of the Parabrurkholderia strains and the Sphingomonas strains</i>	102
6.3.2. <i>In vitro</i> evaluation of silica solubilising microbes	105
7. Conclusions	108
8. Bibliography	112
9. Acknowledgments	124

# 1. General Introduction

Microbial life has dominated most of the evolutionary history of our planet, occupying all available environmental niches on Earth's surface and subsurface<sup>1</sup>. Since more than three billion years ago, microbes have been keeping evolving themselves to exploit their metabolic capabilities in any kind of environment. In this respect, the subsurface ecosystem represents a unique container of microorganisms that are essential for the turnover of the matter and for the biogeochemical cycling of chemical elements in the environment<sup>2</sup>. The life forms inhabiting underground environments – such as caves – shift from the entrance zone to the deepest locations. In the entrance and twilight zones, carbon and energy sources are available in different forms, as plants and organic matter are present. Life forms can gain advantage by exploiting photosynthesis and chemosynthesis reactions, and the ecosystem is rich in autotrophic and heterotrophic microbial communities that are very similar to those detected on the surface<sup>3</sup>. In the deepest zones of a cave, the physical and chemical conditions change, as temperature and humidity tend to be constant<sup>3</sup>. Darkness and absence of plants make deepest locations of caves nutrient-limited (oligotrophic) environments. In these habitats, primary production is generally entrusted to autotrophic and chemolithotrophic bacteria that oxidize atmospheric gases (such as nitrogen, carbon monoxide and hydrogen) and inorganic substrates from the minerals rocks (such as iron and manganese) to gain energy for microbial growth<sup>4</sup>. These microbes are considered pioneers of the cave environment, in which they allow the colonization of complex microbial communities that possess unique structures from both a taxonomic and a metabolic point of view. Additionally, microbial cave life can also depend on small inputs of organic carbon transported into the underground through percolating waters, air circulation and fauna<sup>3</sup>. To survive in these nutrient-poor environments, microorganisms typically organize themselves in collective structures, offering cooperative and mutualistic relationships and producing, as results of their interaction, biosignatures that can be observed within caves<sup>5</sup>. Considering the extremely slow geological time required for the formation of caves, the mechanisms leading to the development of microbial communities in the subsurface are completely different from those observed on the Earth's surface<sup>6</sup>.

The microbial ecology of deep caves is strictly related to the interactions between microorganisms and the minerals composing the rocky substrates. In this context, the combination of all the biotic processes involved in the formation of minerals is called biomineralization. These processes can be either metabolism-dependent or –independent<sup>7</sup>. In the first case, the interaction between microbes and minerals is driven by enzymatic activities that alter the state of oxidation of chemical elements. Intra- and extra-cellular waste metabolic products can interact with mineral compounds and form

precipitates. In the second case, the cell wall of microorganisms can act as nucleation sites for the formation and precipitation of mineral compounds by absorbing metals and other chemical elements. Thus, the cell-mineral interaction is a key factor in caves as biomineralization processes positively influence both the biofilm growth and the microbial activity<sup>7</sup>.

Bioprecipitation and biomineralization processes, frequently in combination with abiotic factors, lead to the modification of the rocks and the formation of speleothems, which consist of secondary mineral deposits<sup>8</sup>. In nutrient-limited and low-energy caves featured by stable physical and chemical parameters, microbes are supposed to play a central role for the processes involved in the speleothems formation<sup>9</sup>. Many of these slow-growing microbes are still uncultured, and little is known about their metabolic potential. Depending on the nature of the mineral rock, the interaction between microbes and the rocky substrates provides the formation of secondary mineral deposits, which can be composed of sulphur, iron, manganese, nitrate, silicon and carbonate compounds. The properties of the cave rocks influence the taxonomy and the metabolic activities of the cave microbial communities, which in turn differ as respect to the typology of secondary mineral deposits<sup>10</sup>.

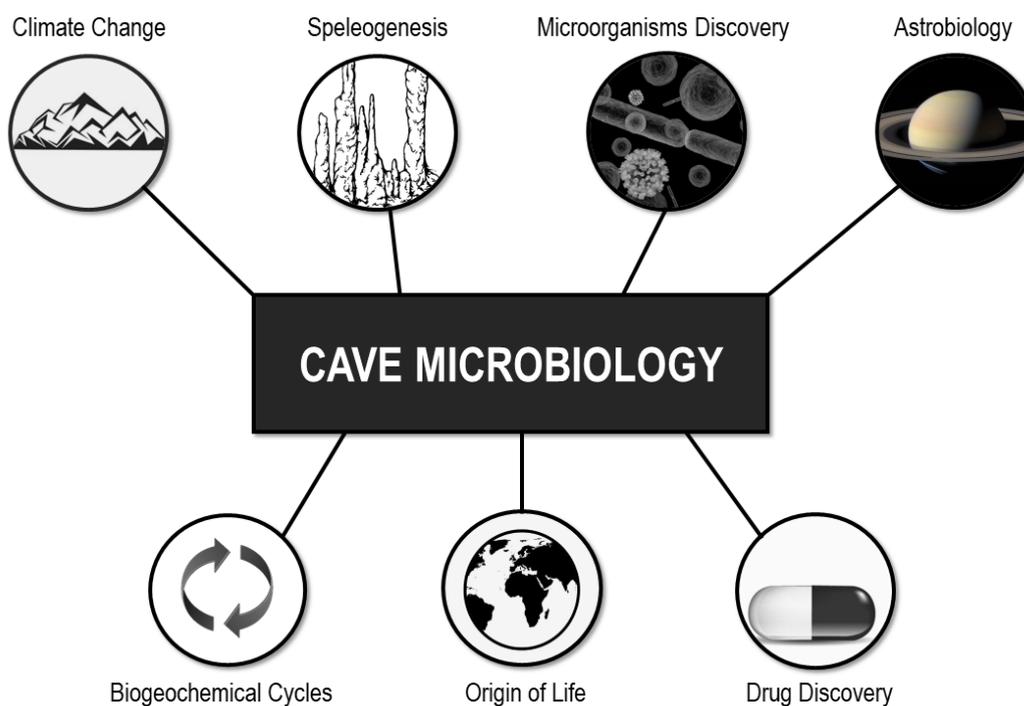
The interest on cave microorganisms is also associated to their ability to produce secondary metabolites and bioactive molecules with possible pharmaceutical and industrial applications, such as extremoenzymes, exopolysaccharides, biosurfactants, antitumorals, antibiotics<sup>11,12,13</sup>. Considering that the life forms inhabiting unexplored caves completely isolated from the exterior have had minimal or no contacts with human pathogens, the probability to find bacterial molecules active against current multidrug resistant bacteria is high. Up to date, the *Actinobacteria* phylum includes most of the microbial species that produce antimicrobial compounds and several microbiologists succeeded in isolating active actinobacterial species from different caves, including Azorean and Canadian volcanic caves, Siberian limestone caves<sup>14,15</sup>. Other studies described new bioactive compounds produced by cave bacteria belonging to additional microbial phyla, i.e. *Bacteroidetes*, *Firmicutes*, *Proteobacteria* and *Cyanobacteria*<sup>16</sup>.

Taken together, cave microbiology has great potentials towards the understanding of microbial impacts on the geochemical and mineralogical setting of the subsurface environments, on the ecosystem nutrient fluxes and on the processes of adaptation and evolution of bacterial and archaeal species in dark conditions and for drug discovery<sup>17</sup>.

The majority of caves formations is due to the dissolution of karstic rocks, including carbonate (limestone, conglomerate, dolostone), evaporates (halite, gypsum) and quartzites. According to the dissolutive chemistry leading to cave speleogenesis, the different caves comprise hydrothermal, sulfuric acid, mixing-corrosion, orthoquartzitic and evaporate-dissolution systems. Additionally, some caves are formed in low temperature locations, can contain ice or can be formed inside the ice

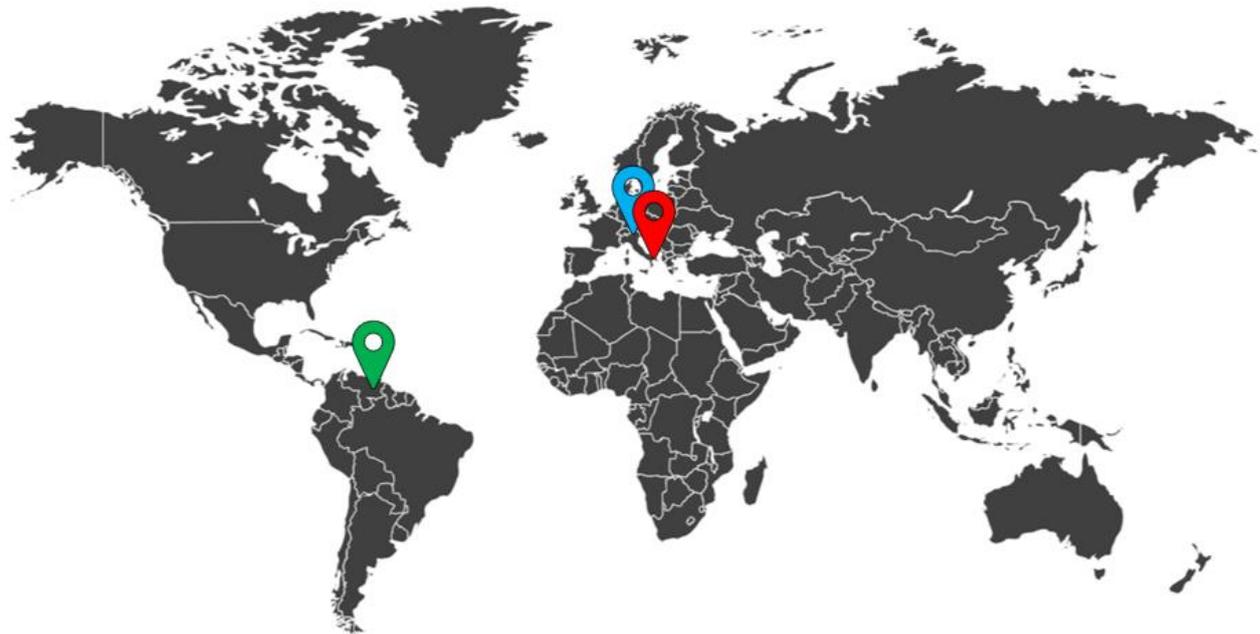
(glacier caves)<sup>18</sup>. Beyond the natural formation of caves due to the natural slow erosion process, other caves can also have an anthropogenic origin<sup>19</sup>. Each typology of cave system hosts selected microbial communities that develop in various macroscopic forms in accordance to the environmental conditions and the nature of the rocks<sup>20</sup>.

Lastly, since planetary unexplored caves share common features with the extraterrestrial environments providing, for example, protection from radiation and thermal fluctuations, they are also attractive targets for astrobiological studies<sup>21</sup> (Fig 1).



**Fig 1.** Main scientific topics of cave microbiology.

In this Doctoral thesis, the microbial communities inhabiting three different cave systems have been investigated from a microbiological point of view. Imawarì Yeuta is an orthoquartzitic cave located in the Venezuelan Tepui mountains, Fetida is a sulfuric acid cave located in the Italian Atlantic coast, Cenote Abyss is an Alpine ice cave (Fig 2). The three caves are very different in terms of chemical-physical parameters and host a wide variety of deposits and mats of biological origin. Further, the analysis of the metabolic potential of 168 microbial isolates from different cave environments has been performed for evaluating their abilities in producing bioactive molecules and in modifying mineral substrates.



-  **Imawari Yeuta**, Auyan-Tepui, Bolivar, Venezuela -  $5^{\circ}50'47.0''\text{N}$   $62^{\circ}22'30.0''\text{W}$
-  **Fetida**, Santa Cesarea Terme, Apulia, Italy -  $40^{\circ}02'04.3''\text{N}$   $18^{\circ}27'20.6''\text{E}$
-  **Cenote Abyss**, Cima Cunturines, Trentino-Alto Adige/Südtirol, Italy -  $46^{\circ}34'33.43''\text{N}$   $11^{\circ}58'40.52''\text{E}$

**Fig 2.** Geographical localization of the three caves.

## 2. General Materials and Methods

### 2.1. Sampling

Samples from cave rocky surfaces were collected after scraping with sterilized tools and stored in Eppendorf tubes. Water samples were collected from cave ponds with sterilized spoons and stored in falcon tubes. All samples were filled with LifeGuard and carried out in a portable fridge to the lab. Samples planned for DNA analysis were stored at -80 °C, whereas those designed for the isolation of pure cultures were stored at 4 °C.

### 2.2. Total DNA extraction and Illumina sequencing of V4 and V5 hypervariable regions of 16S rRNA

The cave samples were extracted for their total DNA using the PowerSoil DNA Isolation Kit (Qiagen), with slight modifications as previously described<sup>4</sup>. Two different primer sets were used to provide amplicons for Illumina MiSeq sequencing of V4 and V4-V5 hypervariable regions. V4 and V5 are the most reliable regions for representing the full-length 16S rRNA sequences in the phylogenetic analysis<sup>22</sup>. Table 1 reports the primer sets used in relation to the samples collected from specific caves.

Hypervariable region	Cave	Primer set	Sequence (5'-3')	Reference
V4-V5	Imawari Yeuta	515F	GTGYCAGCMGCCGCGGTA	23
		907R	CCCGYCAATTCMTTTRAGT	23
V4	Fetida, Cenote Abyss	515F	GTGYCAGCMGCCGCGGTA	23
		806R	GGACTACHVGGGTWTCTAAT	23

One µL of total DNA was added to a 50 µL (final volume) PCR reaction mixture containing 25 µL of Premix F (Epicentre Biotechnologies, WI, USA), 200 mM (each) forward and reverse primers, and 0.5 U of Ex Taq DNA polymerase (Takara Bio, Japan). Amplification reactions were carried out under the following thermocycling conditions: 95°C for 3 min, 30 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, with a final extension at 72°C for 5 min.

PCR amplicons were confirmed by electrophoresis with a 1% (w/v) agarose gel and then purified by AMPure XP beads (Beckman Coulter) prior to the index PCR. Nextera XT Index was incorporated into each of the individual samples during PCR. The thermal cycling program included a first denaturation step at 95°C for 3 min, followed by 8 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s, elongation at 72°C for 30 s, with a final extension at 72°C for 5 min.

Purified amplicons obtained from V4 region amplification were submitted to the Illumina MiSeq next-generation sequencing platform for indexing and pair-end sequencing (2x250 bp; reagent kit, v2) at the Medical University of Graz (Austria).

Purified amplicons obtained from V4-V5 region amplification were submitted to KAUST Genomic Core Lab for unidirectional sequencing reads on an Illumina MiSeq platform.

### **2.3. Processing of 16S rRNA Illumina sequencing data and statistical analysis**

Raw sequence reads were analysed by using one of the two different approaches, based on either Operational Taxonomic Units (OTU) picking through the Quantitative Insights Into Microbial Ecology (QIIME) platform version 1.9.1 or Sequence Variants (SV) analysis through QIIME2<sup>24</sup> version 2018.4 by using DADA2 package version 1.5.0 (as described by D'Angeli et al.<sup>20</sup>). Both the approaches were used because of the release of QIIME2 during the PhD Thesis. Indeed, Sequence Variants are the DADA2 outputs and are described as real amplicon denoised sequences allowing, for marker gene data analysis, higher specificity and resolution as compared to operational taxonomy units (OTUs)<sup>20</sup>. RDP Classifier and SILVA SSU 128 reference database<sup>25</sup> were used for taxonomical assignments of the 16S rRNA gene sequences.

After OTU picking or DADA2 analysis, the datasets were computed for the calculation of alpha and beta diversity indexes with Primer-E version 7 and with CALYPSO. Alignment of sequences with the GenBank best hits was performed through the ClustalW algorithm and the phylogenetic trees were constructed using the Juke-Cantor genetic distance model with 1000 bootstrap replicates.

### **2.4. Shotgun metagenomic sequencing and analysis**

Around 500 ng of total DNA was sequenced using Illumina HiSeq sequencing system following the manufacturer's protocol at the Medical University of Graz (Austria). The quality of sequencing datasets was initially assessed using FastQC software. Sequences with low Phred quality score (< 20) and shorter than 100 base pairs in length were filtered out. The remaining Illumina paired-end reads were assembled using MEGAHIT v1.1.233, using default parameters. Taxonomic and functional classifications of contigs were performed using MG-RAST (v4.0.3) against RefSeq23 (BLAT parameters of 60% similarity, 15 bp and E value of  $10^{-5}$ ) and KEGG databases, respectively.

### **2.5. X-Ray fluorescence spectrometry and X-ray diffraction**

Bulk chemical analyses were conducted by using a wave dispersive X-ray fluorescence spectrometer (WD-XRF) operating at BIGEA Department, University of Bologna (Italy). Ultra-fine powdered samples were mounted on rounded boric acid casts (~5 cm diameter, ~0.5 cm height), which were

prepared according with the matrix correction method. Thirty-five international reference materials were used for calibrating the raw results, allowing an accuracy better than 5% for elements >10 ppm, and between 10% and 15% for elements <10 ppm. Thermogravimetric TG–DTG–DTA measurements were performed by using a Setaram Labsys double-furnace apparatus and calcined Al<sub>2</sub>O<sub>3</sub> as reference substance, in order to calculate the volatile content. Powdered 0.5 g samples were placed in platinum crucibles and introduced into the furnace at 800±1 °C for ~24 hours drying before the final weighing.

Mineral phases were investigated by a Philips PW3710 X-Ray diffractometer (current: 20 mA, voltage: 40 kV, range 2θ: 5–80°, step size: 0.02° 2θ, time per step: 2 sec) at the University of Genova (Italy), which mounted a Co-anode. Acquisition and processing of data were carried out using the Philips High Score software package.

## 2.6. Geochemical analyses

Water temperature (T) and pH were measured by handheld field instruments (Hanna Instruments) after calibration on site. Accuracy was 0.1 °C and 0.01. pH was measured with pH stripes with range 2 to 9 and 0.5 pH unit increments. Dissolved silica concentration (DSi) was measured by using a field colorimetric test kit (Aquaquant 14410 Silicon - Merck). Inductively coupled plasma-mass spectrometry (ICP-MS) (method EPA 6020A) was applied for determination of multi-elemental sub µg/L concentrations (Al, Sb, As, Ba, Cd, Ca, Fe, Mg, Pb, K, Na, Zn) where the recovery of the Laboratory Control Sample (LCS) resulted between 85 and 115%, as expected by the method lines. Anion Chromatography (method EPA 9056A) was used to determine chloride, fluoride and nitrate, in the solution. NH<sub>4</sub> concentration was measured on the untreated sample with the method APAT CNR IRSA 4030 A2 MAN 29 2003. The concentration of S<sup>2-</sup> dissolved in the water was analyzed *in situ* using the spectrophotometer Hach DR/2010 (Loveland, USA), whilst pH, T, TDS (Total Dissolved Solids) were monitored using Hanna HI991001 instrumentation (Padova, Italy). Na<sup>+</sup>, K<sup>+</sup>, Mg<sup>2+</sup> and Ca<sup>2+</sup> in water samples were measured by Atomic Absorption Spectrophotometry (AA-6800 Shimadzu, Kyoto, Japan). Trace elements have been investigated using ICP-MS X SERIES 2 Thermo Scientific (Waltham, USA).

Gases (O<sub>2</sub>, CH<sub>4</sub>, SO<sub>2</sub>, and H<sub>2</sub>S) were analyzed using a MSA Altair4x multigas detector (Pittsburgh, USA). The range of values and their resolution were 0-30 ± 0.1 vol% for O<sub>2</sub>, 0-100 ± 1% LEL for CH<sub>4</sub>, 0-20 ± 0.1 ppm for SO<sub>2</sub>, and 0-200 ± 1 ppm for H<sub>2</sub>S. The cave air temperature was measured with the silicon band-gap sensor loggers Niphargus (Natural History Museum Brussels) and Hobo (ONSET).

Calcite, dolomite, and gypsum saturation indices (SI) were calculated using the ratio between ion activity product (KIAP) and solubility products (K<sub>sp</sub>). Each KIAP has been calculated using the Debye-Hückel equation to determine the ion activity coefficient. SI values close to 1 are indicative of saturated solution at equilibrium, whereas SI < 1 are indicative of undersaturation (i.e., corrosive-dissolutive conditions).

## **2.7. Scanning electron microscope**

For scanning microscope analyses, samples were first covered with a thin evaporated gold layer by sputtering, then introduced into a Vega3 Tescan scanning electron microscope (SEM) and a Zeiss Supra 40 VP field emission scanning electron microscopy (FESEM), operating at the DISTAV Department, University of Genova (Italy). The first operated at 20kV and was equipped with an EDAX-Apollo-X DPP3 energy-dispersive (EDS) X-Ray spectrometer, which was applied for major elements spectrometric measurements. Manganese Resolution of  $K\alpha = 126$  eV allowed the detection of chemical elements heavier than Boron (atomic number greater than 5). Acquisition and elaboration of data were performed by the TEAM Enhanced Version V4.2.2 EDS software. For FESEM images, we used accelerating voltages from 10 Å to 20kV.

### **3. Geomicrobiology of the orthoquartzite cave Imawarì Yeuta**

*Part of this chapter is included in: Sauro F, Cappelletti M, Ghezzi D, Columbu A, Hong P, Zowawi H, et al. (2018) Microbial diversity and biosignatures of amorphous silica deposits in orthoquartzite caves. Scientific Reports 8: 17569.*

#### **3.1. Introduction**

Silicon is the seventh most abundant element in the universe and the second most abundant element on Earth, after oxygen<sup>26</sup>. It can be broadly found in various forms including silicon dioxide (e.g., quartz, amorphous silica), which is the most abundant and important oxide in the Earth's crust. Considering the low solubility of quartz, cave systems formed in the quartzite are not very common<sup>26</sup>. Before the discovery of Venezuelan Tepui cave systems<sup>27</sup>, the quartzite caves described mainly consisted in small tectonic fractures. Therefore, the recent finding of extensive horizontal quartzitic subterranean systems in the Tepui has opened new questions on the processes leading the speleogenesis of these environments. In this respect, the number of microbiological works concerning quartzitic caves are still poor<sup>10,28</sup>.

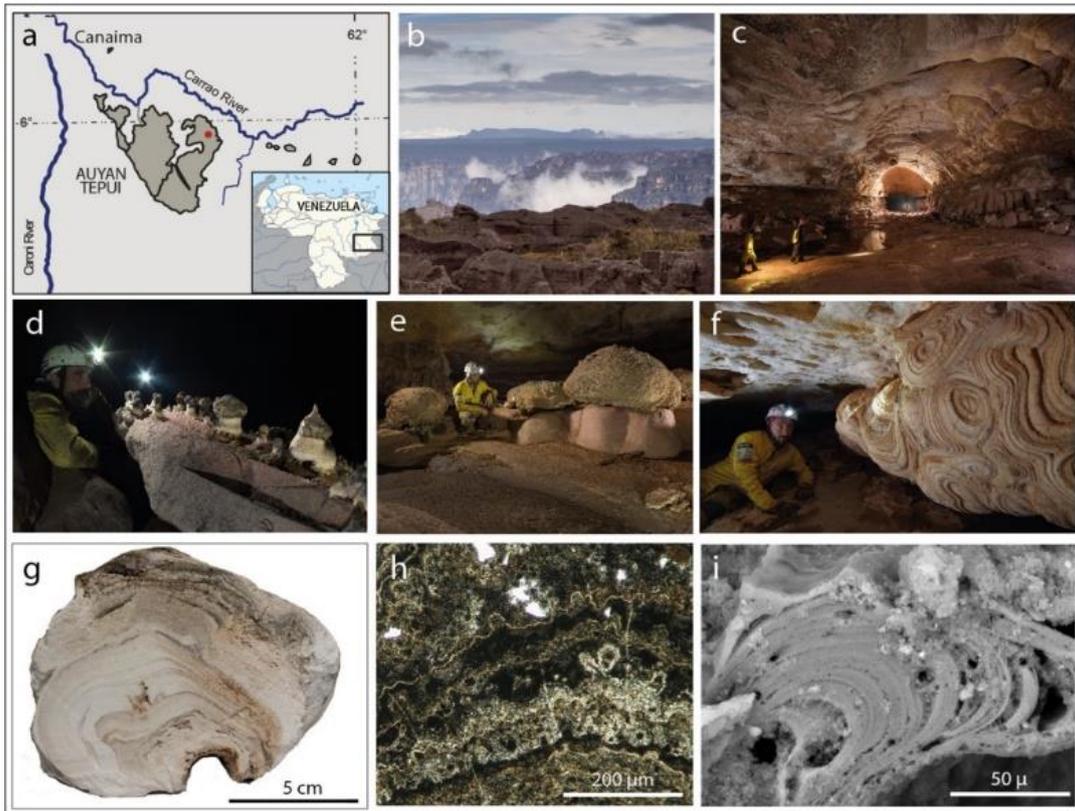
Silicification of bacteria and algae is common in hot springs where microorganisms play only a passive role, considering that silica precipitation is dominated by degassing, rapid cooling and pH changes at the discharge points<sup>29</sup>. In contrast to hydrothermal silica sinters, highly stable geochemical and environmental conditions are present in the Tepui caves, with constant temperatures (~ 15 °C, corresponding to the average yearly surface temperature) and water pH usually ranging from moderately acidic to neutral. In these conditions, other processes different from those occurring in hot springs are required to explain the mobilization and re-precipitation of important amounts of silica. In this context, the formation of silicified structures is highly attributable to a long-time persistent biological activity exerted by peculiar microorganisms<sup>10</sup>. Understanding the functional role of microorganisms in the formation of amorphous silica deposits and silica-stromatolites remains one of the most intriguing challenges in geomicrobiology. Indeed, the metabolic mechanisms ruling the sustainment of the microbial life in silica-rich and nutrient limited environments are still unknown.

Silicified biological structures are of high interest also for the comprehension of ancient natural environments, as those based on the interactions between amorphous silica and microbes, as well as those of Archean and Precambrian silica stromatolites<sup>30</sup>. Moreover, amorphous silica structures are good candidates as biosignatures for the investigation of life forms on other planets, because they represent potential analogues with the quartz-based formations detected on the Red Planet. Indeed, the oldest evidences for life on Earth are fossilized in amorphous and microcrystalline silica deposits (e.g., chert)<sup>31</sup>. This suggests that the orthoquartzite subsurface is one of the most important locations

in the search for extinct and extant extraterrestrial life<sup>32</sup>. Moreover, it was observed that environmental conditions of the Titan planet possibly enhance the reactivity of silicon compounds, providing clues on a putative silicon-based life<sup>33</sup>.

In the last twenty years, microbial diversity was studied in silica-rich environments including high-altitude tundra location, hyper-arid deserts, Antarctic soils and orthoquartzitic caves<sup>18,21</sup>. Among these, the orthoquartzitic Imawarì Yeuta cave represents a non-thermal and mild natural location which hosts a wide variety of unique amorphous silica deposits growing in several types of niches<sup>10</sup>. Caves carved in the Precambrian quartzitic table mountains (Gran Sabana, Venezuela), locally named Tepui, are among the less explored and most pristine places on Earth<sup>9</sup>. The 20-30 Ma old Imawarì Yeuta cave, discovered in the Auyan Tepui and first explored in 2013, is the longest and probably the oldest<sup>9</sup>. In addition to the absence of light, the isolation from the exterior atmosphere determines a general low nutrient availability and low organic carbon sources. For these reasons, Imawarì Yeuta represents a unique study model for understanding the biological activities of microorganisms leading to the development of silica-based secondary mineral deposits within oligotrophic, non-thermal and aphotic environments. Amorphous silica speleothems have been found in several caves of the Tepui area, as well as in lava tubes and granite caves, but never in such amount and diversity<sup>27,30</sup> (Fig 3). In Imawarì Yeuta, evidences of silica mobilization from the bedrock are ubiquitous, from the diagnostic dissolution features on quartz grains to the presence of impressive amounts of silica precipitated speleothems, which have been interpreted as biologically mediated or even as primary silica stromatolites<sup>10,27</sup>.

Beside the potential of Tepui caves in understanding the biological aspects in silica amorphization processes, the identification of complex microbial communities in Imawarì Yeuta opens interesting questions on their interaction with the mineral rocks, and the successful functional mechanisms developed to proliferate in oligotrophic conditions. In such locations, the composition of the microbial communities is connected to the variations of specific environmental parameters. Several cave microbiomes have been investigated in relation to pH values, temperatures changes, organic carbon and water availability, humidity, seasonal changes<sup>20</sup>. Consequentially, performing and implementing metagenomics and geochemical analyses from Tepui samples represents a key step for providing insights into the capabilities of complex microbial communities to adapt their metabolism in peculiar circumstances in order to exploit the little energy input of the environment they live in. Lastly, subsurface unexplored locations represents unique containers for the discovery of novel organisms, which have never interacted with the exterior atmosphere and with the life on the Earth's surface<sup>21</sup>. This aspect makes caves model ecosystems to study the evolution of life from its establishment on Earth.



**Fig 3.** Study area, cave system and silica deposits (Sauro et al. 2018<sup>10</sup>). Located in the southeastern corner of Venezuela (a), the Auyan Tepui table mountain (b) hosts the 23 km-long Imawari Yeuta cave system (c and red point in a). Examples of biologically mediated silica deposits in Imawari Yeuta cave: mushroom-like speleothems built by layered soft and highly porous amorphous silica in the hydrologically inactive areas of the cave (d); massive silica stromatolite-like columnar formations growing on pinkish orthoquartzite boulders (e); giant deposits of opaline silica with concentric growth bands completely covering the orthoquartzite walls of the cave (f). In cross-section most of the deposits are characterized by layered porous opaline silica (g) with typical wavy and crinkled lamina and thin opaque lamina under plane polarized light (h) and single micro-columnar features visible with SEM (i). Photos are provided by La Venta Archive (b, N. Russo; c, R. Shone; d and f, V. Crobu; e, R. De Luca).

The present chapter is divided into three four sections reporting the geomicrobiology studies on the orthoquartzitic Imawari Yeuta cave. The first section aims at giving an overview of the microbial ecology of the cave by investigating the taxonomy composition of microbial communities within several cave samples based on 16S rRNA gene. This analysis provides an overview of the microbiology of different habitats within Imawari Yeuta cave and the physicochemical variables correlated with the microbial diversity.

The second section aims at defining the microbial communities specifically associated to different silica-based speleothems featured by different silica amorphization phases (from the pristine quartz to amorphous silica) and collected from different cave locations (wall, pavement, water). On the basis of the shifting of specific taxonomy groups and the microscopy analysis, the role of specific taxa is suggested in the silica amorphization process.

The third section focuses on the geomicrobiology of three speleothems collected from the cave pavement representative of sequential silica amorphization steps associated with the formation of a novel (defined coralloid) silica stromatolite. A combination of different molecular and bioinformatics

approaches targeting 16S rRNA and functional genes encoding enzymes involved in atmospheric gas metabolism was used, with the attempt to improve the taxonomic resolution of the microbial diversity shifting during the stromatolite formation, to suggest a role of specific enriched taxa in silica amorphization processes on the cave pavement and to detect the metabolic potentials associated with oligotrophy.

The forth section describes a preliminary analysis of the metagenomes associated to three samples representative of consecutive silica amorphization steps collected from Imawarì Yeuta cave wall in order to perform a first screening of the putative metabolic pathways involved in the sustainment of the microbial communities and associated with the increase of silica amorphization.

### 3.2. Materials and methods

#### 3.2.1. Samples list, DNA extraction and sequencing

Samples were collected and stored as described in paragraph 2.1 before their processing in the Core laboratories of KAUST (Saudi Arabia) and Laboratory of Molecular and Applied Microbiology of the University of Bologna (Italy).

Table 2 summarizes all the samples analysed in this chapter.

Table 2. List of samples analysed in this chapter.						
Sample ID	Cave niche	Water activity	Silica content	Silica amorphization level	Methods of analysis	Section of this chapter
IY1	Floor	Dry	High	High	QIIME2 (SV)	3.3.1
IY2	Floor	Dry	High	High		
IY3	Floor	Dry	Low	Low		
IY4	Floor	Dry	Medium	Low		
IY5	Roof	Dry	High	High		
IY6	Roof	Dry	High	High		
IY7	Floor	Dry	Low	Low		
IY8	Roof	Dry	High	High		
IY9	Water	Wet	Saturated	High		
IY10	Water	Wet	Saturated	High		
IY11	Water	Wet	High	Low		
IY12	Water	Wet	Saturated	High		
IY13	Wall	Moist	High	High		
IY14	Wall	Moist	High	High		
IY15	Floor	Moist	High	Low		
IY16	Floor	Moist	High	Medium		
IY17	Floor	Moist	High	Medium		
IY18	Wall	Moist	High	Low		
IY19	Floor	Moist	High	Low		
Q	Wall	Moist	High	Low	FESEM, QIIME (OTU), QIIME2 (SV)	3.3.2
S	Floor	Moist	High	Medium		
WL	Wall	Moist	High	High		
F	Floor	Dry	High	High		
WB	Water	Water	High	High		
Qz	Floor	Moist	High	Low	FESEM, q-PCR, QIIME2 (SV), EMIRGE (e-OTU), RFLP (clone library)	3.3.3
AS	Floor	Dry	High	High		
CS	Floor	Dry	High	High		
Wall-mg1	Wall	Moist	High	Low	Shotgun metagenomics	3.3.4
Wall-mg2	Wall	Moist	High	Medium		
Wall-mg3	Wall	Moist	High	High		

Total DNA extraction, amplification of V4-V5 regions and Illumina sequencing were performed as described in paragraph 2.2. Considering the processing of 16S rRNA Illumina sequencing data and statistical analysis, both OTU picking and DADA2 analysis (SV) were performed in samples collected from Imawari Yeuta, as reported in Table 2 and described in paragraph 2.3.

### 3.2.2. Full-length 16S rRNA reconstruction and analysis through EMIRGE algorithm

Full-length 16S rRNA gene amplicons were analyzed through Illumina HiSeq and reconstructed by using EMIRGE<sup>34</sup>. In the literature, EMIRGE algorithm is usually used for reconstructing 16S rRNA gene from metagenomic raw data<sup>34</sup>. However, because of the low sequencing depth of the metagenome sequencing performed in this study (around high-quality 700000 reads resulted from the HiSeq of each sample), we assumed that the analysis of the 16S rRNA data from each metagenome was not enough to allow a deep taxonomy analysis of the microbial communities. Therefore, we decided to employ EMIRGE using 16S rRNA amplicons. The PCR reactions were performed in a final volume of 20 µl containing: 10 ng total DNA, primers (8F and 1100R, Table 3) 500 nM, Takara Ex Taq buffer with MgCl<sub>2</sub> (10 X), dNTP mix 200 µM, TAKARA Ex Taq Polymerase 0.5 U. The amplification reactions were carried out under the following thermocycling conditions: 95 °C for 2 min, 10 cycles of 96 °C for 30 s, 60 °C for 30 s, 72 °C for 1 min, followed by 94 °C for 25 s, 58 °C for 30 s, 72 °C for 1 min, with a final extension at 72 °C for 10 min. Libraries were prepared using NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs) that implied a first step of random enzymatic restriction of the amplicons. After Illumina MiSeq sequencing run at the Molecular Biology Core Lab of the Medical University of Graz (Austria), filtering and trimming of the sequencing data were performed using an in-house Galaxy set-up, which included the EMIRGE algorithm for the assembly of the high-quality reads (120 iterations). Reconstructed near-full length 16S rRNA sequences were clustered into OTUs (hereafter named e-OTUs) on the basis of a sequence identity threshold of 97%. The relative abundance of each e-OTU was calculated via the EMIRGE algorithm on the basis of “prior probabilities” of read coverage depth<sup>35</sup>. Chimeras were identified and removed with UCHIME. Taxonomic assignment of e-OTUs was performed by using SILVA SSU 132 reference database<sup>25</sup>.

Table 3. Primer set used for EMIRGE-based analysis of 16S rRNA.

Method	Target gene	Samples	Primer set	Sequence (5'-3')	Reference
EMIRGE	16S rRNA	Qz, AS, CS	8F	AGAGTTTGATCCTGGCTCAG	36
			1100R	GGGTTGCGCTCGTTG	36

### 3.2.3. q-PCR

Quantitative-PCR was conducted on the CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, USA) by using SYBR green-based reactions. The quantification of the bacterial and archaeal 16S rRNA genes was performed in triplicate for each sample using the corresponding primer pairs (Table 4) and the 20  $\mu$ L qPCR reaction mix containing: 10 ng total DNA, primers 300 nM each, 1x SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, Hercules, USA), water (Lichrosolv®; Merck, Darmstadt, Germany), using the thermocycling conditions: 95 °C for 15 min, 40 cycles of 94 °C for 15 s, 60 °C for 30 s, 72 °C for 40 s. Serial dilutions (across seven orders of magnitude (10<sup>1</sup>–10<sup>7</sup>)) of 16S rRNA gene PCR products from *Escherichia coli* and *Nitrososphaera viennensis* were used as standards for Bacteria and Archaea, as previously described<sup>37</sup>.

Method	Target gene	Samples	Primer set	Sequence (5'-3')	Reference
qPCR - Bacteria	16S rRNA	Qz, AS, CS	338F	GCTGCCTCCCGTAGGAGT	36
			517R	ATTACCGCGGCTGCTGG	36
qPCR - Archaea	16S rRNA	Qz, AS, CS	806F	GGACTACHVGGGTWTCTAAT	23
			945R	TAAACTYAAKGAATTGACGGG	23

### 3.2.4. Clone libraries of full-length 16S rRNA, *coxL* and *hypD* genes and screening through Restriction Fragment Length Polymorphism (RFLP)

Total genomic DNA was used as template for PCR amplification reactions targeting the bacterial 16S rRNA, *coxL* and *hypD* genes using the primers listed in Table 5. For the three target genes, one  $\mu$ L (~10 ng) of total DNA was added to a 50  $\mu$ L (final volume) PCR reaction mixture containing Takara Ex Taq buffer with MgCl<sub>2</sub> (10 X), dNTP mix 200  $\mu$ M, 1.25 U of Takara Ex Taq DNA polymerase (Takara Bio, Japan) and 500 nM (each) forward and reverse primers.

The amplification reaction of 16S rRNA gene was carried out under the following thermocycling conditions: 98 °C for 10 sec, 30 cycles of 98 °C for 10 sec, 55 °C for 30 sec, 72 °C for 90 sec, with a final extension at 72 °C for 20 min. The amplification reaction of *coxL* gene was carried out under the following thermocycling conditions: 98 °C for 10 sec, 30 cycles of 98 °C for 45 sec, 58 °C for 60 sec, 72 °C for 90 sec, with a final extension at 72 °C for 20 min. The amplification reaction of *hypD* gene was carried out under the following thermocycling conditions: 98 °C for 10 sec, 40 cycles of 98 °C for 30 sec, 60 °C for 40 sec, 72 °C for 45 sec, with a final extension at 72 °C for 20 min.

PCR amplification products were confirmed by electrophoresis with a 1% (w/v) agarose gel, purified with the Qiagen PCR purification kit (Qiagen, Hilden, Germany). PCR products were ligated into the pCRII vector supplied with the TOPO TA cloning kit (Invitrogen, San Diego, CA, USA), according

to the manufacturer's instructions, and cloned into *Escherichia coli* DH5 $\alpha$  for clone library construction.

Restriction Fragment Length Polymorphism (RFLP) analysis was performed for each library. Individual colonies containing inserts of the appropriate size were suspended in 20  $\mu$ L of TE pH 8 and boiled for 5 min. Cell debris were removed by centrifugation and 1- $\mu$ L portions of the supernatant were used as templates in PCR mixtures to re-amplify the gene inserts that were further used in restriction digestion with tetrameric restriction enzymes. The inserts were digested at 37 °C for 3 h with 5 U of *AluI/RsaI* enzymes for 16S rRNA and *MspI* enzyme for *coxL* and *hypD*. Restriction profiles were analyzed by 2% (w/v) high resolution agarose gel electrophoresis with high-resolution agarose (Metaphor, Tebu-bio). Clones were grouped manually based on restriction patterns. RFLP screenings were stopped when the rarefaction curves approached saturation (coverage > 80%).

For sequences identification, plasmids were purified from one representative clone of each group with a Qiagen plasmid purification kit (Qiagen). Sequencing was performed by the Eurofins Genomics Service (Germany) using T7 and T3 primers (Table 5). Sequences were checked for chimeras using the CHECK\_CHIMERA program at the Ribosomal Database Project (RDP) (<https://rdp.cme.msu.edu/>).

Method	Target gene	Sample	Primer set	Sequence	Reference
Clone library	16S rRNA	Qz	27F	AGAGTTTGATCMTGGCTCAG	38
			1492R	TACGGYTACCTTGTTACGACTT	38
Clone library	<i>coxL</i>	Qz	OmpF	GGCGGCTTYGGSAAASAAGGT	39
			O/Br	YTCGAYGATCATCGGRTTGA	37
Clone library	<i>hypD</i>	Qz	hypD-for2	GGNCCNGGCTGCCCGGTCTG	40
			hypD-rev	GGCGNNGTGGTTTCAAANCC	40
Clone library	16S rRNA, <i>coxL</i> , <i>hypD</i>	Qz	T7	TAATACGACTCACTATAGGG	Invitrogen
			T3	ATTAACCCCTCACTAAAGGGA	Invitrogen

### 3.2.5. Phylogenetic and statistical analyses of RFLP-based clone libraries

Phylogenetic trees of 16S rRNA nucleotide sequences and of CoxL and HypD amino acid sequences were constructed with the most closely related available sequences (Best Blast Hits) downloaded from GenBank database. All sequences resulting from each approach were aligned with ClustalW and used to construct a tree based on maximum-likelihood algorithm using MEGA7, with bootstrap values of 1000.

Diversity indexes and Bray-Curtis Distance Matrix were calculated through Primer-E v7. Clustering analyses were performed using both SV and e-OTUs on the basis of the presence and abundance of microbial genera.

### 3.2.6. Geochemical analyses

X-Ray fluorescence spectrometry and X-ray diffraction were performed as described in paragraph 2.5. Geochemical analyses were performed as described in 2.6. Scanning electron microscope analyses were performed as described in 2.7.

## 3.3 Results and discussion

### 3.3.1 Diversity and statistical analysis of all the microbial samples collected from Imawari Yeuta

Illumina MiSeq sequencing results of 19 samples collected from Imawari Yeuta cave are reported in Table 6. In total, 2968927 raw reads were generated with a minimum of 15394 reads per sample (IY19) and a maximum of 319010 (IY5). The read processing through DADA2 resulted in 5469 SVs, with a minimum of 84 SVs per sample (IY7) and a maximum of 751 (IY18).

Sample ID	Cave niche	Water activity	# reads	# SV
IY1	Floor	Dry	48391	174
IY2	Floor	Dry	202863	202
IY3	Floor	Dry	286335	193
IY4	Floor	Dry	195240	283
IY5	Roof	Dry	319010	119
IY6	Roof	Dry	65630	117
IY7	Floor	Dry	163365	84
IY8	Roof	Dry	193504	224
IY9	Water	Wet	99194	374
IY10	Water	Wet	213135	323
IY11	Water	Wet	115764	379
IY12	Water	Wet	253971	85
IY13	Wall	Moist	299378	530
IY14	Wall	Moist	66894	278
IY15	Floor	Moist	155558	582
IY16	Floor	Moist	16438	147
IY17	Floor	Moist	123576	674
IY18	Wall	Moist	135287	751
IY19	Floor	Moist	15394	252

Beta diversity analyses were performed in order to identify the physical and/or chemical features that better described the taxonomic composition of the different cave samples. As a result, nMDS and RDA plots at SV level showed that the 19 samples under analysis clustered mainly based on the water activity (Fig 4). On the other hand, other variables such as cave location, pH, silica amorphization and temperature correlated at lower extent with the microbial communities. Three main categories could be identified based on the water activity which clustered the samples that were “dry” (7 samples, IY1 to IY7), featured by completely absence of water, “moist”, featured by low water content (7 samples, IY8 to IY14), “wet”, collected from stagnant water ponds (4 samples, IY15 to IY19).

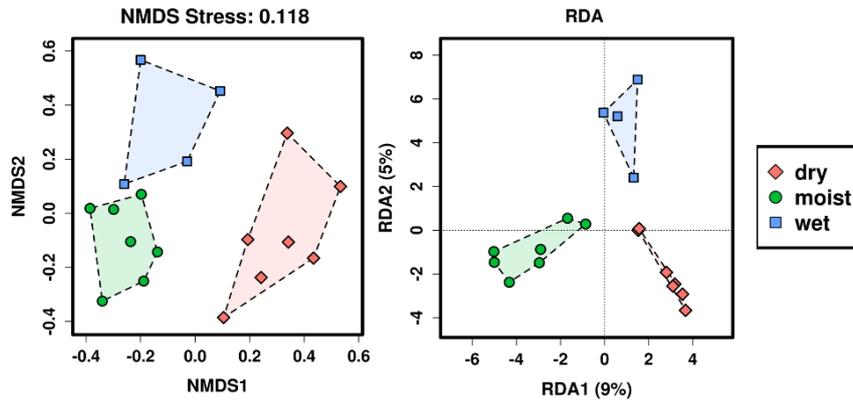


Fig 4. Beta diversity at SV level of the entire dataset.

Alpha diversity calculation was significantly different among the clustered samples for Shannon, Inverse Simpson and Chao1 indexes (Fig 5). Shannon ( $p = 0.03$ ) and Inverse Simpson ( $p = 0.05$ ) indexes were higher in moist samples compared to dry and wet ones. Chao1 index was lower in the dry samples ( $p = 0.016$ ). Although evenness calculation was not significantly different among the groups of samples, its value was lower in wet samples.

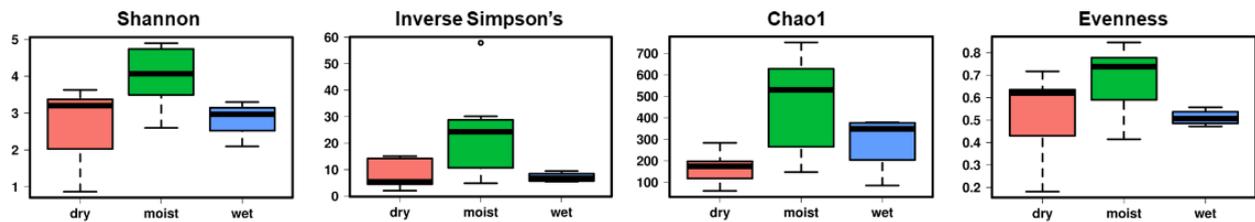
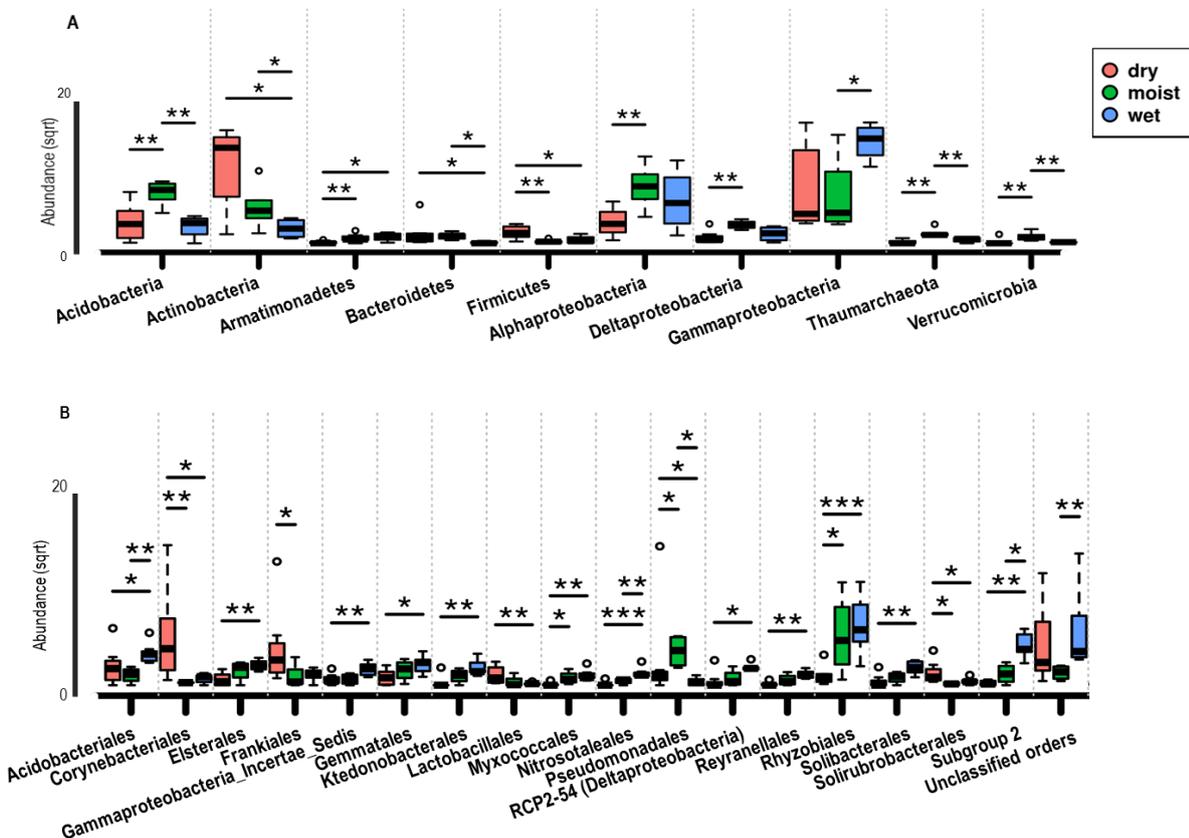


Fig 5. Alpha diversity indexes at SV level of the 19 samples grouped according to their water content.

These results indicate that low water activity is associated to low microbial diversity and richness in Imawari Yeuta. Despite this, the lower diversity of samples collected from stagnant water ponds (wet samples) as compared to the moist group of samples might be supported by the nature of the water. Indeed, moist samples are speleothems retrieved from rock surfaces close to small water streams or under percolating water sources, where the water is continuously in movement. This sort of dynamicity would allow the spread of nutrients and chemical elements that enriches the microbial communities and leads to the establishment of a broader diversity as respect to the microbial communities collected from stagnant waters. On the contrary, dry samples are featured by the microbial communities with lower diversity and dominated by few taxonomic groups that most probably have been selected for their capabilities of adaptation to desiccation and other environmental stress conditions. It can be therefore hypothesized that the restricted availability of nutrients and energy sources in the dark and oligotrophic Imawari Yeuta cave makes the microbial communities' development being strictly dependent on the little water that can enter from the outside by percolating

through the rocks and that moves into the cave. Previous studies indicated that in the absence of light, bacteria thrive exploiting the water runoff dripping into the cave through cracks in the overlying rock and produce energy from the compounds present in decaying organic matter in the soils above and from the minerals dissolved within the rock fissures<sup>41,42</sup>.

The core microbial communities of each group of samples were investigated at different taxonomic levels. *Bacteria* dominated the microbial communities of all 19 samples. *Archaea* were <1% in all samples except for one sample belonging to the moist group (IY16), in which *Thaumarchaeota* reached 3% of the total community (data not shown). Analysis of mean taxa proportion (at phylum and order levels) showed significant differences in the abundance of *Acidobacteria*, *Actinobacteria* and *Proteobacteria* among the clustered samples (Fig 6). In general, dry samples were dominated by *Actinobacteria*, whereas *Acidobacteria* were highly abundant in moist samples. *Proteobacteria* were commonly present in both moist and wet samples, with *Alphaproteobacteria* better characterizing moist samples and *Gammaproteobacteria* featuring samples collected from water ponds. Additionally, distinct samples are present within each group which show high abundance of other microbial phyla, such as *Chloroflexi*, *Verrucomicrobia*, *Firmicutes* and *Bacteroidetes*. At order levels, the abundance of Subgroup 2 (*Acidobacteria*), *Rhizobiales* (*Alphaproteobacteria*), *Corynebacteriales* (*Actinobacteria*), *Pseudomonadales* (*Gammaproteobacteria*), *Acidobacteriales* and *Solibacterales* (*Acidobacteria*) varied also significantly among the clusters (Fig 6).



**Fig 6.** Rank test at phylum/proteobacterial class (A) and order (B) level. For each level, the 30 most abundant taxonomic groups are considered. \* refers to p value < 0.05, \*\* refers to p < 0.01, \*\*\* refers to p < 0.001.

To explore the variation of the microbial community composition at lower taxonomic level among dry, moist and wet samples, we performed the linear discriminant analysis (LDA) effect size method (LEfSe). This test determines the taxa most likely to explain differences between sample groups by coupling standard tests for statistical significance with additional tests encoding biological consistency and effect relevance. At genus level, moist samples possessed a higher abundance and frequency of unclassified genera (mainly belonging to *Acidobacteria* and *Alphaproteobacteria*) as respect to dry and wet samples. *Candidatus Solibacter* (LDA score = 3.85,  $p = 0.00039$ ), *Bryobacter* (LDA score = 3.97,  $p = 0.0041$ ), *Reyranella* (LDA score = 3.93,  $p = 0.0002$ ) and several unclassified *Acidobacteria* (LDA score = 3.99,  $p < 0.001$ ) were significantly enriched in moist samples. The role of *Reyranella*, together with other members of *Rhodospirillales* and *Rhizobiales* orders, in the microbial community is probably linked to their capabilities to sustain the development of other microbial taxa by performing primary reactions like those concerning the nitrogen cycle in the cave<sup>29</sup>. *Acidobacteria* are ubiquitous in mineral-rich acidic soils, they are widely considered heterotrophs able to survive in low-nutrient conditions<sup>43</sup> with a direct involvement in mineral weathering processes. Initial studies associated *Acidobacteria* to acidic environments<sup>44</sup>. However, recent studies demonstrated their flexibility in colonizing both acidic and alkaline niches<sup>45</sup>. Specifically, *Acidobacteriales* (former Subgroup 1) and Subgroup 2 orders were associated to low pH values, whereas members belonging to Subgroup 4, 6, 7 and 16 mostly inhabit alkaline environments. The remaining orders, including *Solibacterales* (former Subgroup 3), were described as able to tolerate a broad range of pH values. Further, *Acidobacteriales* and *Solibacterales* orders, which correlate with moist samples, have been associated to quartz-based sandy soils<sup>46</sup>, assuming the capability to slow down their metabolism to persist in starving conditions and to survive on quartz. Among the few described genera, *Bryobacter* and *Candidatus Solibacter* have been both described as capable to survive in oligotrophy<sup>47</sup>.

*Stenotrophomonas* (LDA = 4.57,  $p = 0.027$ ) and *Pseudomonas* (LDA = 4.48,  $p < 0.0001$ ) were significantly associated to wet samples. The role of *Betaproteobacteriales* and *Pseudomonadales* in wet samples is probably linked to the presence silica-saturated waters, since members of these taxonomic groups have been described from other caves as capable to perform biomineralization processes<sup>48</sup> (Fig 7).

Among the *Actinobacteria*, *Mycobacterium* (LDA = 4.76,  $p = 0.024$ ) was the genus more characterizing the dry group of samples. Many actinobacterial strains, including some belonging to *Mycobacterium*, are able to persist in harsh conditions such as waterlessness<sup>49</sup>. In addition to the capacity to persist under harsh environmental conditions also in relation to their cell wall typically

containing long aliphatic hydrophobic chains named mycolic acids, *Mycobacterium* strains are capable to switch their metabolism according to the settings of the environment they live in<sup>49</sup>.

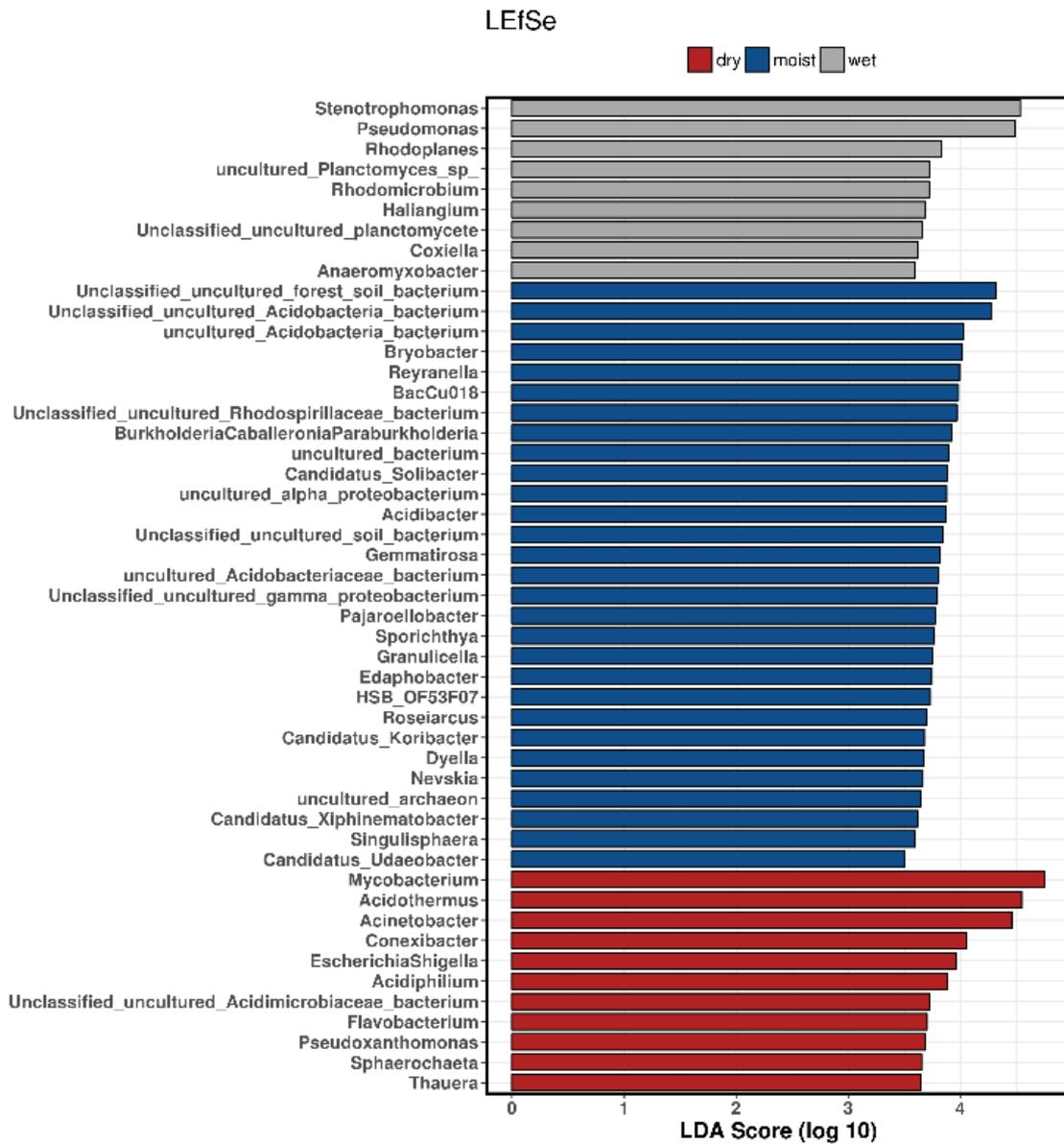
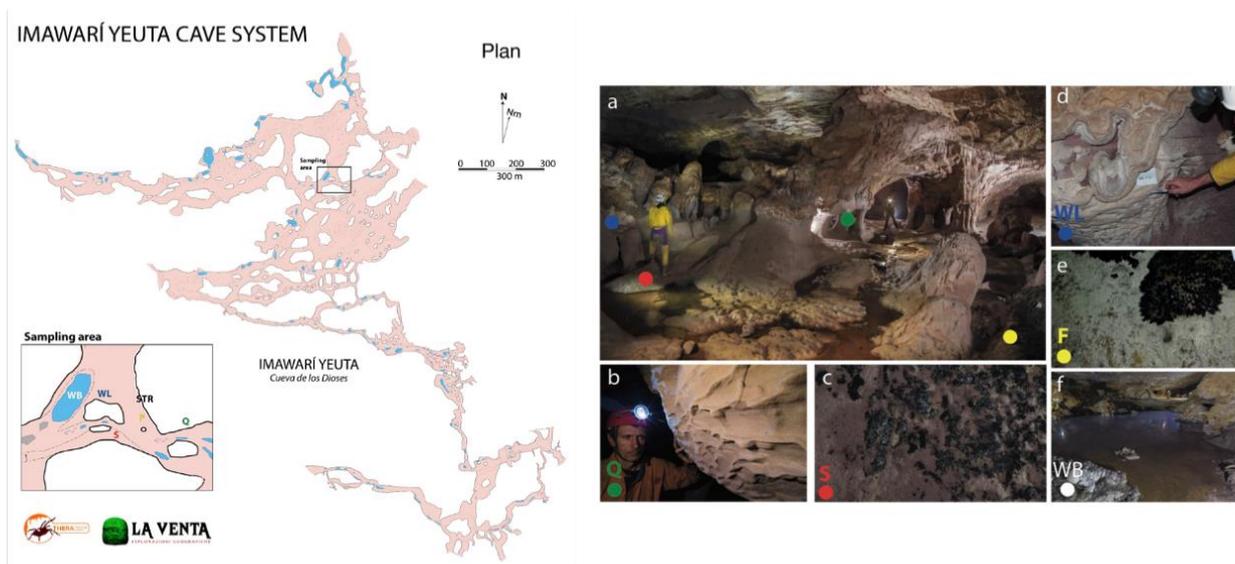


Fig 7. LDA score at genus level.

### 3.3.2. Characterization of the microbial communities associated to different silica amorphization levels

#### Sample description

Five samples were collected from different cave sub-environments (Fig 8) representing different peculiar biogeochemical niches from the pure bedrock (crystalline or non-amorphous silica) to the amorphous silica deposits and silica-saturated waters. Sample Q corresponds to a recently eroded quartzite wall, in which degradation produces loose quartz sand (sample S). Amorphization is absent in Q and minimal in S. Sample WL is a white soft paste of silica, where the amorphization is pervasive showing a transition from the rock surface to thick but soft laminated deposits. Sample F corresponds to a well-consolidated laminated amorphous silica speleothem on the cave floor. WB is representative of a standing water pool saturated with respect to silica, with evident iridescent violet patinas (Fig 8) floating on the surface and amorphous silica and sulphate deposits around the pool edges. SiO<sub>2</sub> dominates all sub-environments, but minor elements, such as iron and aluminium, slightly increase from Q to S and speleothems WL and F (Table 7). The pH of moisture wetting the different environments also increases from 4 in Q and S to 5 in the amorphous silica samples (Table 7). A similar trend is shown by cave water chemistry: in active stream waters (STR) silica content is low (0.1-1 mg L<sup>-1</sup>) and pH is acidic (3.5 to 4.5), while standing pool waters (WB) are saturated with respect to silica (>8 mg L<sup>-1</sup>), pH reaches 6 and minor components content like sulphates, chlorine and barium are significantly higher (Table 7).



**Fig 8.** Map of Imawari Yeuta cave system (left) and sampling points (right). Samples were obtained from orthoquartzite walls (b, green dots - Q), quartz sand lying on the cave floor (c, red dots - S), opaline silica growing on cave walls (d, blue dots - WL), opaline speleothem on the floor (e, yellow dots - F) and opaline silica precipitates and slime floating on cave ponds (f, white dots - WB). Modified from Sauro et al. (2018)<sup>10</sup>.

**Table 7. Distribution of major elements in waters (blu) and rock/speleothem samples (grey) from Imawari Yeuta cave.**

Sample name	SiO <sub>2</sub> mg/L	Al µg/L	Fe µg/L	K mg/L	Ba µg/L	Cl mg/L	SO <sub>4</sub> mg/L	NH <sub>4</sub> mg/L	Cu µg/L	Ca µg/L	Mg µg/L	pH	T °C
STR <sup>a</sup>	0.65	56	66	0.57	1.14	ND	ND	0.69	ND	ND	ND	3.9	14.5
WB	8.60	ND	ND	ND	2.82	1.15	0.63	ND	ND	ND	ND	6.1	14.8
Sample name	SiO <sub>2</sub> wt%	Al wt%	Fe wt%	K wt%	Ba ppm	Cl ppm	SO <sub>4</sub> ppm	Zn ppm	Cu ppm	Ca wt%	Mg wt%	pH	T °C
Q	98.0	1.6	0.1	0.0	0	11	0	0	16.2	0.00	0.02	4	14.9
S	98.48	0.93	0.15	0	0	ND	0	0	13.5	0.00	0.02	3.5	14.9
WL	93.9	1.6	0.7	0.05	37	52	295	75	88	0.01	0.1	5	14.9
F	86.9	0.8	0.4	0.01	75	150	170	77	23	0.03	0.04	5	14.9

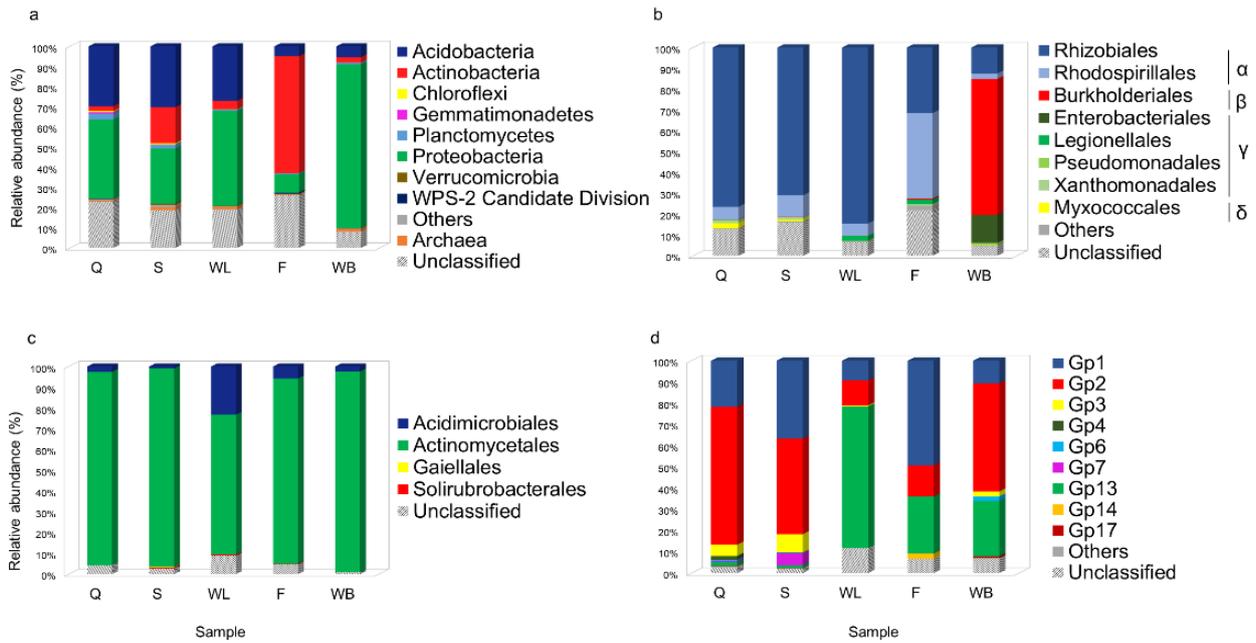
<sup>a</sup> STR stands for “stream water”, a sample representative of water flowing through the system (undersaturated in silica); ND = “not determined”.

**Microbial diversity featuring each type of speleothem**

At phylum level, *Proteobacteria*, *Actinobacteria* and *Acidobacteria* differently dominated the five microbial communities (Fig 9). Samples Q, S and WL showed higher bacterial diversity and more similar microbial community composition compared to F and WB (Table 8 and Fig 9), probably reflecting i) the changes of the community structure occurring during the first stages of silica speleothem formation compared to that of the mature amorphization stages and ii) the oligotrophic nutrient condition, air and water exposition featuring the cave wall compared to the floor environment. A strong relationship was shown among the wall-related samples when the phylum was considered as taxonomy level (Fig 9).

**Table 8. Summary of Illumina sequencing data analysis and alpha diversity indexes.**

Sample name	# reads	# OTU	Shannon	Simpson's
Q	19875	11822	8.26	0.9951
S	13576	8673	8.139	0.996
WL	9908	6642	7.854	0.9916
F	8695	6016	7.753	0.9901
WB	8437	4953	7.287	0.9877



**Fig 9.** Microbial community composition for the Imawari Yeuta cave samples representing progressive stages of silica precipitation in different niches. The category “Others” represents bacterial taxa that constitute < 0.5% of the microbial community in all samples. (a) Distribution of bacterial phyla and *Archaea* in cave samples. (b) Distribution of *Proteobacteria* at order and class levels. (c) Distribution of *Actinobacteria* at order level. (d) Distribution of *Acidobacteria* at order level. Modified from Sauro et al. (2018)<sup>10</sup>.

*Alphaproteobacteria* (mainly *Rhizobiales*) and *Acidobacteria* abounded in Q, WL, and S, while F and WB were characterized by *Actinobacteria* (mainly *Actinomycetales*) and *Betaproteobacteria*, respectively (Fig 9). On the other hand, the differences among wall-related samples included: i) a variety in *Acidobacteria* groups featuring each sample (Gp2 in Q, both Gp1 and Gp2 in S and Gp13 in WL.); ii) a higher abundance of *Actinobacteria* in S representing around 20% of the total microbial community, while representing <5% in Q and WL; iii) a decrease in *Deltaproteobacteria* going from Q to S and to WL (Fig 9); iv) a reduction of the number of bacterial phyla detected (Fig 9). Further, at OTU-taxonomy level, the wall-related samples shared only 2% of OTUs clustered with a 97% similarity cut-off (Fig 10). Therefore, the first stages of silica speleothem formation on the cave wall were featured by distinct microbial communities at species level even if they were spatially linked, while at phylum level, a variation in *Acidobacteria* groups and a decrease in *Deltaproteobacteria* population was observed. During the mature silica speleothem development, the change in the community composition regarded the shift from *Alphaproteobacteria* and *Acidobacteria* to *Actinobacteria* and *Betaproteobacteria*. These microbial diversity modifications might be related to the role that specific microbial groups or species can have in the different silica mobilization steps, to the selective pressure imposed on specific microbial groups or species by the environmental changes (pH, increase in major components), or, most probably, to a combination of the two phenomena.

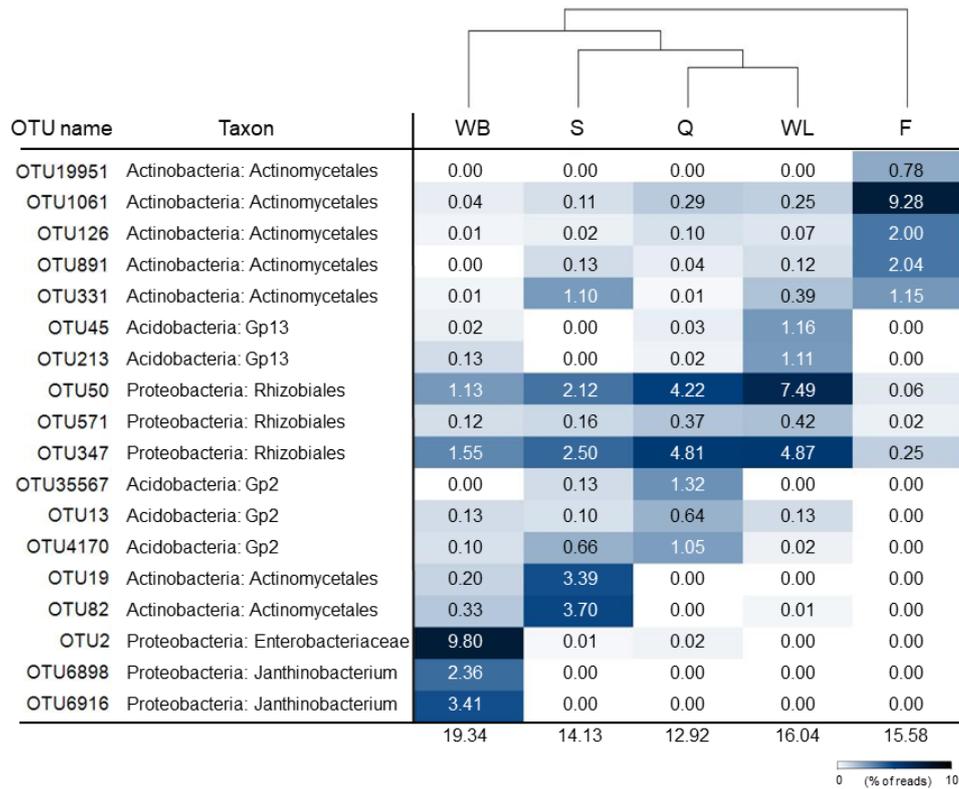


Fig 10. Relative abundance of the five most abundant OTUs retrieved from each sample (Sauro et al. 2018<sup>10</sup>).

The scarcity of nutrients and low buffering ability of quartz-sandstone support the high presence of *Rhizobiales* and *Acidobacteria* in the wall samples (Fig 9 and Fig 10). Particularly, abundant *Rhizobiales* OTUs in samples Q and WL were related to *Beijerinckiaceae* and *Methylocystaceae* families (Fig 10), which include genera able to fix nitrogen and metabolize C1-compounds<sup>50</sup>, representing fundamental functions of pioneer microbial communities. On the other hand, different *Acidobacteria* groups characterized the wall-related samples, suggesting a specific contribution to the diverse microscopic morphologies and/or a different response to the shift of pH and geochemical composition. *Acidobacteria* are ubiquitous in geochemically divergent subterranean environments, i.e. they proliferate in acidic to neutral and alkaline soils<sup>45</sup>, suggesting important contributions to biogeochemical cycles. Despite their possible fundamental role in several biogeochemical processes, insights into the actual physiology and environmental functions of members of this phylum remain limited. In addition to the cave wall origin, sample S was also influenced by the cave floor environment, which might have contributed to the higher abundance of *Actinobacteria* compared to Q and WL. In particular, both samples F and S had predominant OTUs affiliated to *Actinomycetales*, known to have an important role for heterotrophic interactions, carbon turnover and possible biomineralization process in caves<sup>51</sup>.

*Betaproteobacteria* *Janthinobacterium*-related OTUs predominated in WB (53%). OTUs belonging to *Janthinobacterium* also occurred in F but with much lower abundance (0.03%). This bacterial genus mediates cave mineralization processes, and is capable of producing the violacein pigment<sup>52</sup>, the possible reason why Imawarì Yeuta ponds appear violet (Fig 8). Moreover, the most abundant OTU in WB belonged to *Enterobacteriaceae* (Fig 10). Members of this family persist in cave environments, have different origins and possibly contribute to both biochemical and geochemical processes<sup>53</sup>.

High similarity was shared between these prevailing lineages in the different samples from Imawarì Yeuta cave and reference sequences detected in glacial/alpine sites, subterranean environments, lava caves, but also in other silica-dominated environments like Roraima Sur cave<sup>28</sup>, the only other quartzite cave microbiologically described in the Tepui area (Fig 11). Despite this matching, some differences were present in the microbial composition described for Imawarì Yeuta cave and for Roraima Sur cave. Specifically, samples collected from Roraima Sur cave were dominated by *Chloroflexi* and *Thaumarcheota*, that, conversely, were found in low abundance (<1%) in the Imawarì Yeuta cave samples under analysis.

Additionally, abundant species constituted only a small fraction of the microbial communities under analysis (data not shown), which were conversely dominated by rare species represented by very low abundant OTUs (with a frequency <0.01% of the total population). The high portion of rare biosphere contributes to the complexity of the microbial communities found in Imawarì Yeuta cave and might play important roles both in the silica mobilization/precipitation processes and in the evolution of microbial population during silica biospeleothem formation in the cave sub-environments.

The extremely low abundance of species shared among the samples, the strong diversity in microbial composition from Roraima Sur Cave, and the high portion of species falling within rare biosphere indicate the location and type of substrate rock (quartz rock) as only partial driving forces for the complex microbial community structures found in Imawarì Yeuta cave. Additional factors influencing the silica-based ecosystems might include water activity, cave location, silica-based speleothem evolution, nutrient availability in addition to the pH changes previously hypothesized<sup>54</sup>.

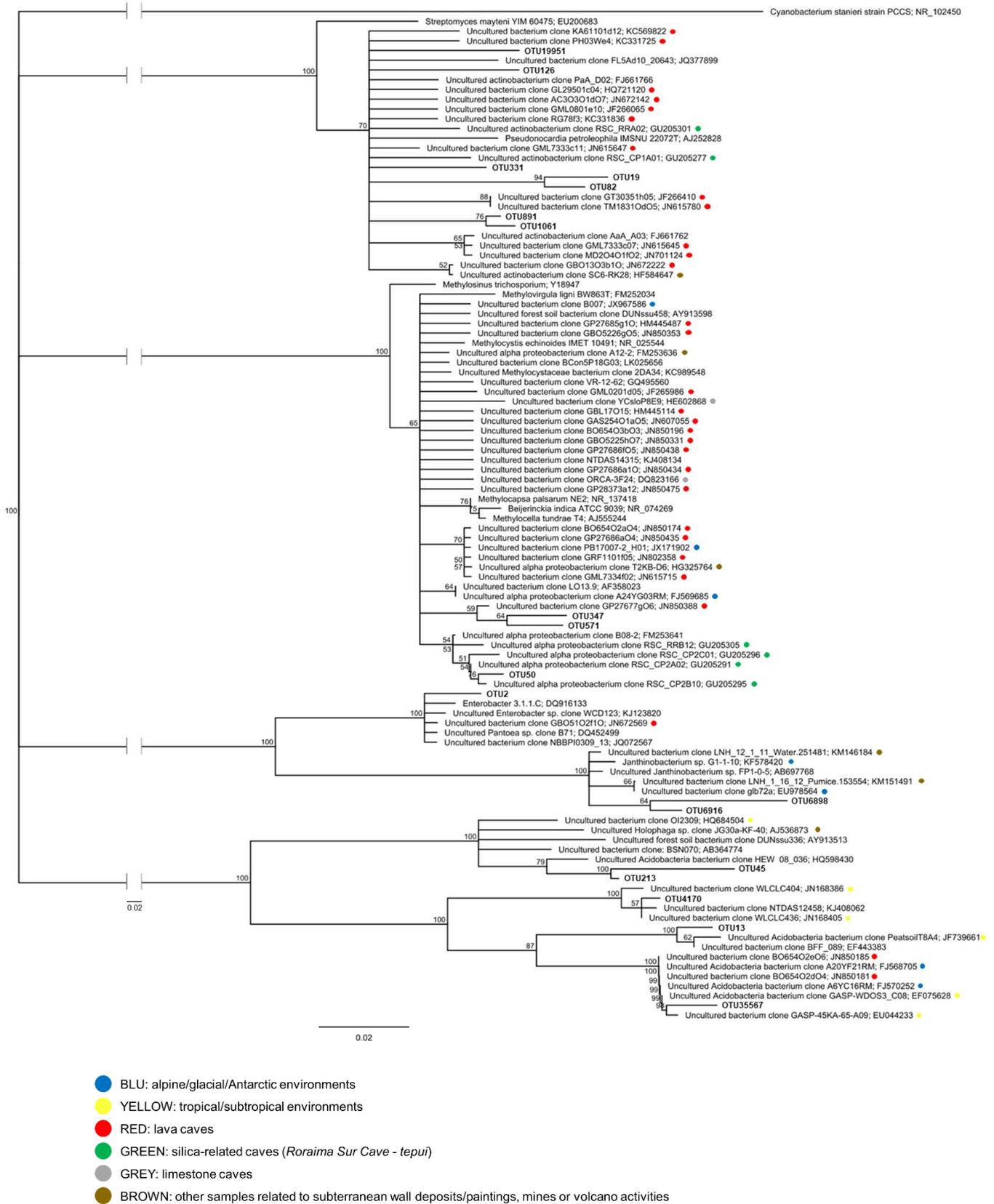


Fig 11. Phylogenetic tree of the most abundant OTUs retrieved from the samples under analysis (Sauro et al. 2018<sup>10</sup>).

### ***Microbial role in silica mobilisation***

Although the low T and pH environment is unsuitable for the mobilization of high amount of silica from the quartz-sandstone, as confirmed by analysis in STR (Table 7), amorphous silica abounds in Imawari Yeuta cave environments. Recent studies have shown that water condensation on cave walls drives dissolution processes on the quartz surfaces inside the rock, and therefore considerably increasing the rock porosity and mobilizing silica from the cave walls<sup>55</sup>. Bacterial communities are expected to take advantage of this process because diffusion involves not only silica but also other minor components of the rock like iron, zinc, barium and calcium necessary for microbial metabolism/growth. Although the direct role of microorganisms in silica mobilization and precipitation process remains strongly debated<sup>56</sup>, microbial metabolism might contribute to silica precipitation and silica-based speleothems formation in pristine environments, also in the light of the high microbial activity recently detected in other silica-dominated caves<sup>57</sup>. In particular, the silica extracted from the quartz-sandstone is believed to re-precipitate as amorphous silica in relation to the following biologically-mediated processes: i) the role of microbial cell surfaces with the related ultra-structures and extracellular polymeric substances (EPS) as nucleation site for silica precipitation<sup>29</sup>, ii) the increase of the pH due to microbial metabolic processes such as nitrogen fixation<sup>58</sup>, bacterial decomposition of proteins or amino acids (e.g. arginine) and urea degradation, CO<sub>2</sub> consumption, iii) the increase of the amount of inorganic cations and metal ions due to biological activity<sup>55</sup>. Biologically mediated silica precipitation in turn can lead to new silica mobilization from the rock by boosting further chemical diffusion.

The aggregates of filamentous structures and the tubular casts observed in the progression of silica speleothem formation in Imawari Yeuta cave (Fig 12) shared strong similarities with silica precipitates and silica-based peloids found in speleothems from other quartz-sandstone caves of the Tepuis<sup>54</sup>. In these other case studies, based only on micro-morphological analysis, silica precipitation and speleothem formation were ascribed to the activity of heterotrophic or autotrophic filamentous bacteria like cyanobacteria. In Imawari Yeuta cave, the sequencing analysis detected only traces (<0.05%) of cyanobacteria exclusively in Q and S (data not shown), therefore our data do not support their role in silica mobilization.



*Methylocella* genus and *Acidimicrobiales* order were described to have Fe-oxidizing activity<sup>59</sup>. In line with a possible role of rare or low abundant species in biogeochemical processes, the presence of OTUs related to Fe-oxidizing bacteria in all the analyzed samples makes this bacterial group a potential candidate for an involvement in silica speleothem formation.

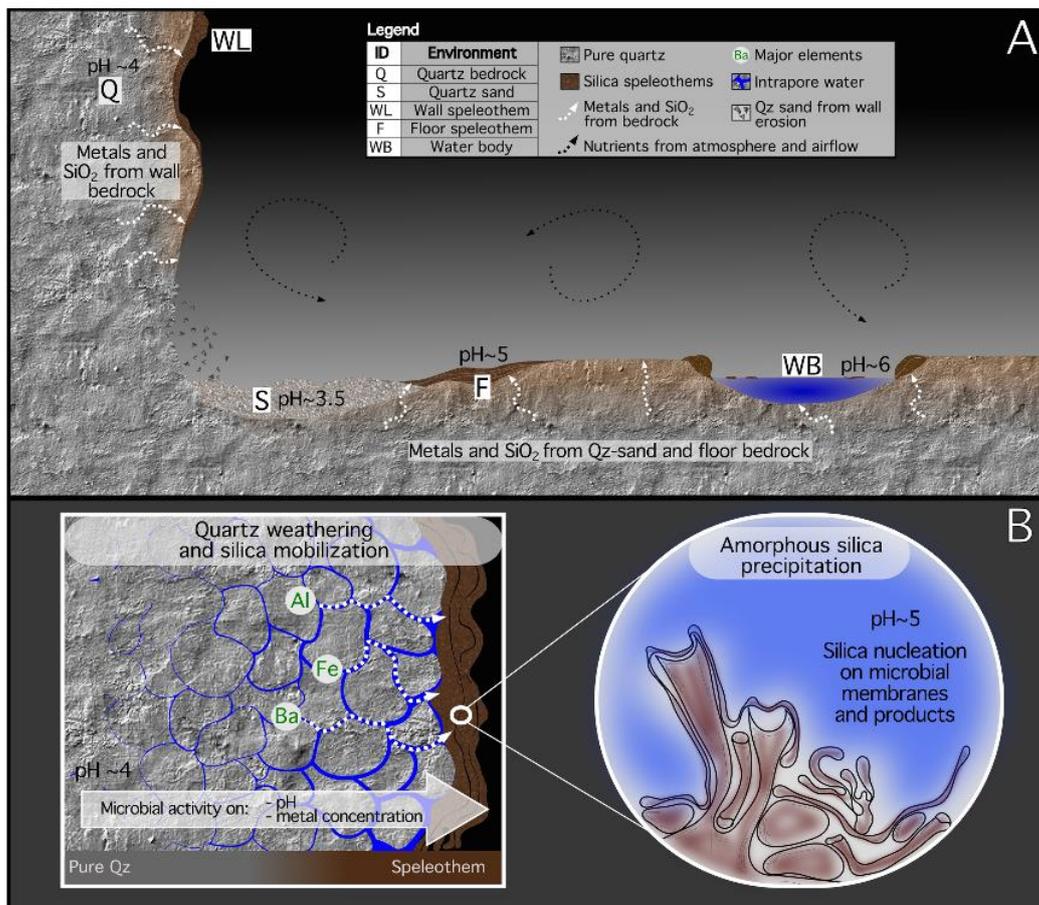
The increase of other metals and minor elements during the speleothem formation (Table 7) indicates additional biomineralization processes. The increase of Ba<sup>2+</sup> detected in WL and F speleothems and in the standing pool WB represents an interesting characteristic of Imawarì Yeuta cave. At near neutral pH conditions, even a limited concentration of barium in solution enhances the dissolution rate and solubility of quartz as much as forty times as compared to deionized water, having a strong influence on the overall silica mobilization potential<sup>60</sup>. Bacterial ability to mobilize, concentrate and precipitate barium compounds was demonstrated using bacterial isolates, and microbial biofilms were shown to play a role in the formation of barium-containing deposits found on volcanic rocks in catacombs<sup>60</sup>. In terrestrial environments, microbes may have a role in barium precipitation by i) oxidizing sulfur compounds to generate sulphate<sup>61</sup>, ii) providing biofilms thus favouring biomineralization<sup>62</sup>, or iii) bioaccumulating Ba in extracellular polymeric substances (EPS) or on cell walls functioning as nucleation sites<sup>60</sup>. Recently, biomineralization of barium was also found to occur intracellularly in filamentous bacteria symbiotic of marine sponges<sup>63</sup> showing an important role in silica mobilization. The tubular and filamentous structures in Imawarì Yeuta speleothems intriguingly show amorphous silica precipitation that are interpreted be due to biologically mediated concentration of different metals occurred in successive steps in the extracellular or intracellular environment.

Recently, a strain of *Betaproteobacteriales Janthinobacterium* was isolated at neutral pH from ferromanganese deposits in caves and was described to perform Mn oxidation<sup>64</sup>. Although little is known to date on its metal-oxidizing mechanisms, and only Mn oxidation was demonstrated, the role of *Janthinobacterium* in biomineralization processes might support its involvement in silica precipitation/mobilization processes occurring in sample WB of Imawarì Yeuta Cave.

Quartz dissolution increases under alkaline conditions, enhancing ionization of the orthosilicic acid. The pH increase detected in Imawarì Yeuta speleothems might be related to pH shifts occurring in the microenvironments surrounding microbial cells or mats. Local changes in pH and the production of metabolites that influence silica solubility (i.e. EPS, amino acids) are likely to result from bacterial metabolic processes, such as nitrogen fixation, bacterial decomposition of proteins or urea and CO<sub>2</sub> consumption. In the microbial population analysis of the silica-based speleothems, microbial groups able to fix nitrogen and perform other nitrogen cycle processes thrive. Specifically, members of *Beijerinckiaceae* and *Methylocystaceae* families proliferate in all the samples, including nitrogen-fixing genera such as *Methylocella* and *Methylocystis*. Both the ammonia oxidizer *Nitrosomonas* and

the nitrite-oxidizing *Nitrospirae* occur in Q and F, while only *Nitrospirae* appears in S (data not shown). These bacterial groups include members able to degrade urea into ammonia and CO<sub>2</sub> and indicate the occurrence of CO<sub>2</sub>-fixation-coupled ammonia oxidation processes in the studied cave ecosystem.

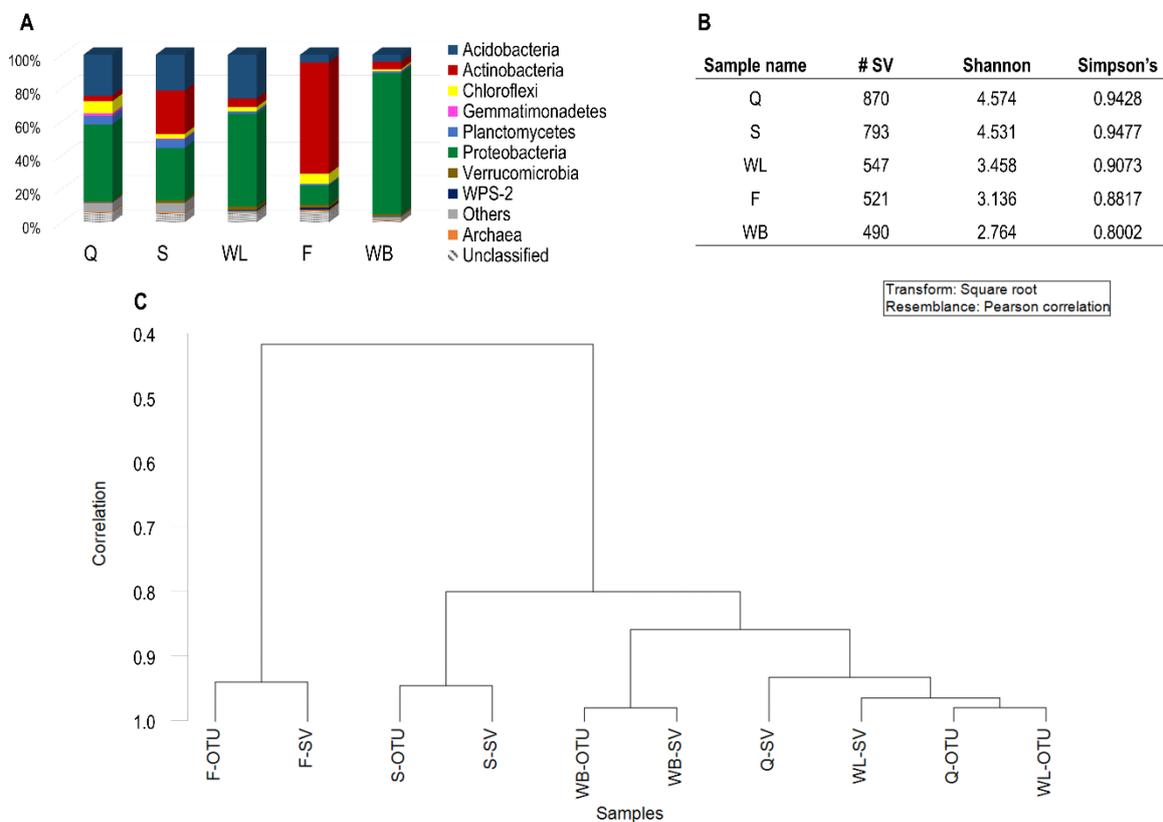
Taken together, our outcomes indicate that complex chemotrophic microbial communities colonize different niches in Imawari Yeuta and create the chemical conditions driving quartz dissolution through i) the increase of the amount of inorganic cations and metal ions in solution as a result of biomineralization processes; ii) the raise of pH mediated by microbial metabolisms (e.g. nitrogen fixation, decomposition of proteins or amino acids, urea degradation, CO<sub>2</sub> consumption). Silica solubilized from the rock can reprecipitate as amorphous species on microbial cell surfaces with their ultrastructures and extracellular polymeric substances (EPS) acting as nucleation sites as observed in Fig 12. Biologically mediated silica dissolution and reprecipitation in turn can lead to new silica mobilization from the rock by boosting further chemical diffusion. Fig 13 shows a working model of the mechanisms we propose are involved in the microbial-mediated silica solubilisation and precipitation in Imawari Yeuta cave.



**Fig 13.** Schematic representation of the Imawari Yeuta cave (a) and the processes of silica mobilization and precipitation (b) leading to the formation of biogenic silica deposits in *tepuì* caves (Sauro et al. 2018<sup>10</sup>).

**Comparison between OTU- and SV-based analyses for microbial diversity characterization**

The sections included in paragraph 3.3.2 describe the characterization of the microbial communities within Q, S, WL, F and WB samples, which was published in Scientific Reports journal by Sauro et al. (2018)<sup>10</sup>. This work was based on the Operational Taxonomic Unit (OTU) analysis using QIIME platform. For the sake of completeness, in this paragraph, the same raw data (fastq sequences) were analysed using the most updated version of QIIME, i.e. QIIME2 and DADA2 package. The DADA2 software provides Sequence Variants (SVs) data instead of OTUs. SVs are considered clustering units that correspond to real amplicon denoised sequences. Due to the development of improved algorithm like DADA2, which can resolve variation at even single nucleotide levels, the SV analysis allows higher resolution as compared to OTUs<sup>65</sup>. As a result of our QIIME2 analysis, the relative comparison among the five samples in terms of diversity and richness was similar using the two bioinformatic approaches (i.e. SV and OTU analyses) (Fig 14). Furthermore, the SV-based taxonomy composition of each sample highly correlated with the one obtained by using OTUs (Pearson correlation,  $\rho > 0.9$ ; Fig 14). A noteworthy difference consisted in the percentage of unclassified sequences, which resulted to be lower in SV analysis (0.5-2%) as compared to OTU analysis (10-20%) (Fig 14). This is likely due to the capacity of SVs to provide higher resolution in the taxonomic and phylogenetic analyses<sup>65</sup>.



**Fig 14.** A) Taxonomy composition of samples Q, S, WL, F and WB at phylum level based on Sequence Variants. B) Alpha diversity indexes of the five samples based on SVs. C) Pearson correlation of the five samples based on the taxonomy composition of the OTUs and SVs obtained through QIIME and QIIME2 platforms, respectively.

### 3.3.3. Characterization of microbial communities associated to the inception of a novel type of siliceous stromatolite discovered in Imawari Yeuta cave

#### *Sample description*

Three samples were collected on the floor of a cave gallery traversed by a small water stream in the inner part of Imawari Yeuta cave (with a discharge of 0.1 L-1 when the samples were collected, Fig 15). These samples were representative of sequential phases of a novel type of silica-based stromatolite that was firstly discovered in Imawari Yeuta Cave. In particular, Qz had the appearance of whitish and pinkish centimetric patches on the rock surface (Fig 15). Moving further and higher from sample Qz along the orthoquartzite floor there was a transition from the whitish patches described above (where Qz was sampled) to a whitish and soft paste of amorphous silica distributed in closely arranged dots. Some of these dots were collected and are representative of the sample AS. Blackish and harder branched formations - like cave coralloids - of amorphous silica have developed as compact masses above AS and they were collected and named CS (Fig 15). The film water on the floor at the sampling points had a pH of 5, which is higher than the one measured in the cave stream (between 3 to 4.5) but similar to that measured in standing pools of water with dissolved silica close to saturation in the same area of the cave (pH between 5 to 5.5)<sup>57</sup>. In term of mineralogy and composition, sample Qz is almost exclusively constituted by quartz grains, while AS and CS are composed of gel-like Opal-A (amorphous silica forms). Aside of silica, major elements homogeneously detected were Al and Fe, while trace elements and metals like SO<sub>4</sub>, Cu, Zn and Cl increased in abundance going from Qz to AS and CS. Barium also increased in the three samples, showing the higher content in CS sample (Table 9).

**Table 9. Distribution of major elements in the three floor samples.**

Sample	SiO <sub>2</sub> wt%	Al wt%	Fe wt%	K wt%	Ca wt%	Mg wt%	Ba ppm	Cu ppm	SO <sub>4</sub> ppm	Cl ppm	Zn ppm
Qz	98	1.6	0.11	0	0	0.02	0	16.2	0	11	0
AS	86.9	1.6	0.68	0.048	0.01	0.11	37	88.1	296	52	10
CS	93.9	0.8	0.35	0.012	0.03	0.04	75.7	23.7	171	151	471.4

Scanning electron microscopy showed the presence of microbial-like morphologies, which were different in the three samples (Fig 15). In particular, sample Qz showed sparse long filamentous structures directly on silica-based substrates. The microbial colonisation and the complexity of the biological structures increased in samples AS and CS. In particular, the first showed irregular distribution of cotton-shaped extracellular deposits and/or matrix mixed up with silicified tubular casts and peloids. The CS samples was composed by tiny coral bodies. In CS, tubular sheets, filaments and spore-like chains constituted a compact aggregate in which amorphous silica consolidated the

chains forming intertwined structures. Thin spider threads emerging from the speleothem were also visible (Fig 15).

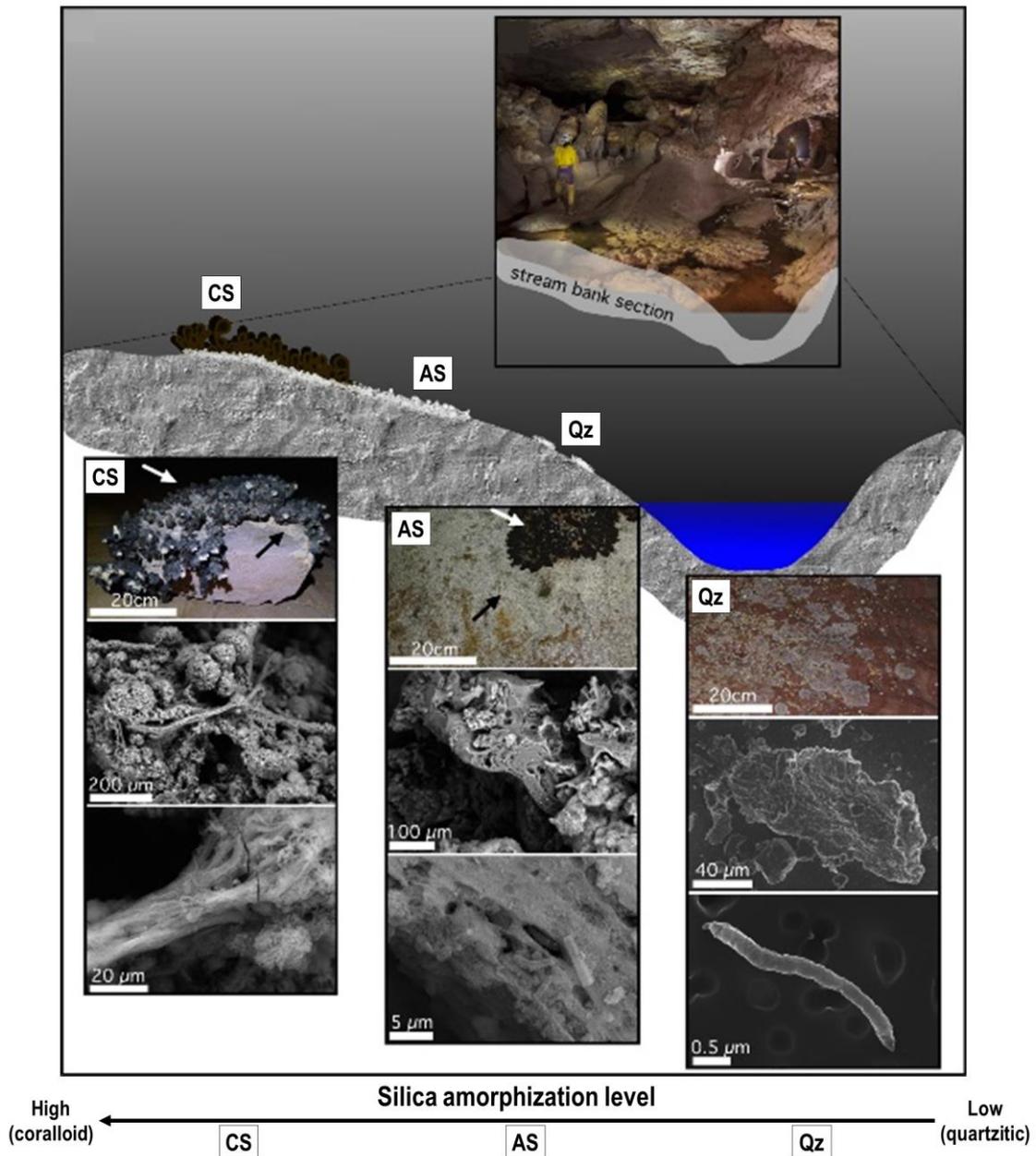


Fig 15. Graphical representation of the three samples collected from the floor and FESEM images.

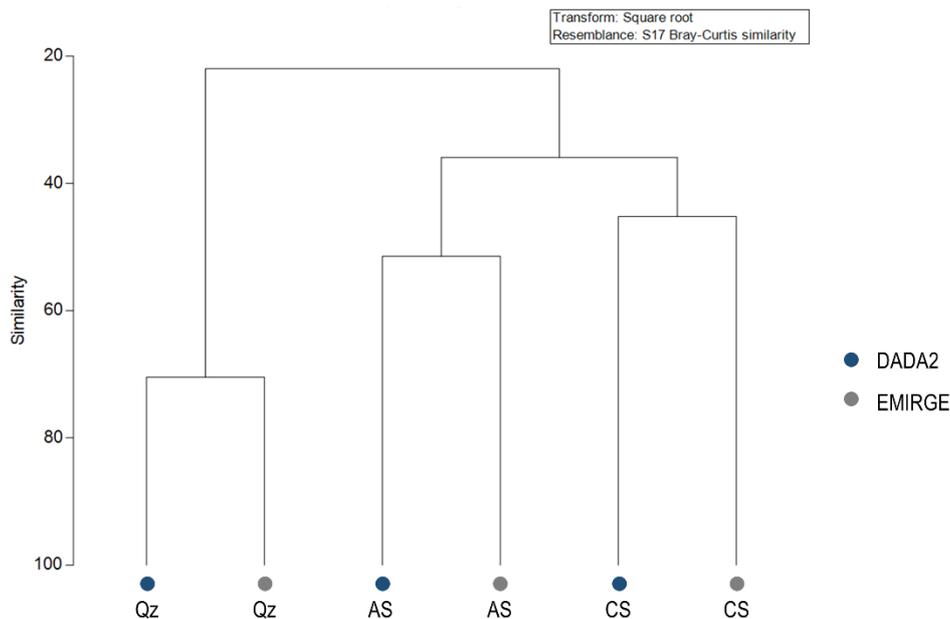
**Microbial diversity of the three sequential steps of stromatolite formation**

For all the three samples, the sampling depth and sequencing coverage resulted sufficient to describe the microbial diversity. Among the three samples, Qz showed the highest richness, while sample CS had the highest diversity in terms of Shannon and Simpson (Table 10). This suggested a negative correlation between microbial diversity and number of microbial species. The Pielou’s index also indicated an increase of evenness moving from Qz sample to AS and CS. This suggests a lower number of dominant species in the amorphous and coraloid silica samples as compared to the

quartzitic (non-amorphous silica) samples. Further, clustering analysis at genus level showed that AS and CS clustered together and separately from Qz, indicating a higher similarity in microbial community composition in AS and CS (Fig 16).

**Table 10. Summary of Illumina sequencing data analysis and alpha diversity indexes of the three samples collected from the pavement.**

DADA2					
Sample name	# reads	# SVs	Shannon	Simpson's	Evenness
Qz	177530	676	2.994	0.8278	0.4595
AS	56302	192	2.91	0.7565	0.5535
CS	185297	199	3.185	0.9159	0.6018
EMIRGE					
Sample name	# reads	# e-OTUs	Shannon	Simpson's	Evenness
Qz	703467	773	3.312	0.8135	0.4981
AS	780587	443	4.152	0.9358	0.6813
CS	656922	332	4.38	0.9665	0.7546



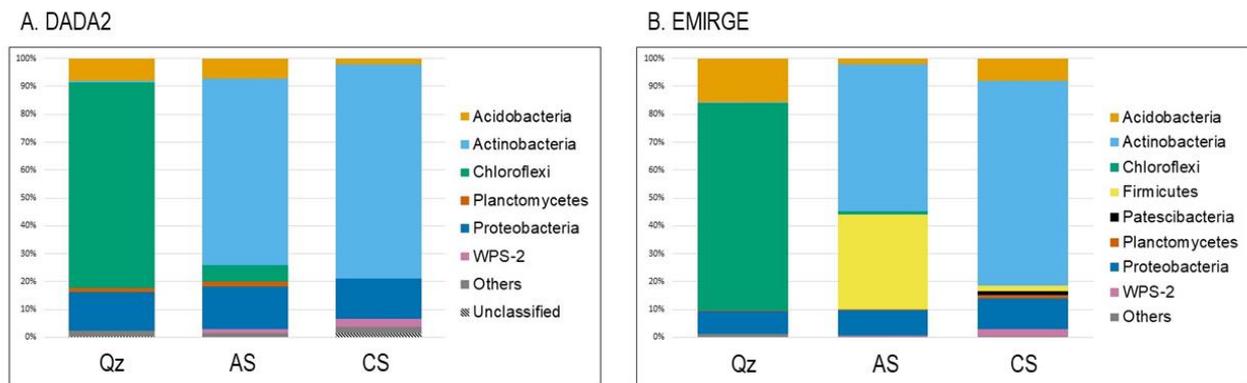
**Fig 16.** Clustering at genus level based on Bray-Curtis similarity of both DADA2 and EMIRGE datasets.

Bacterial and archaeal 16S rRNA genes were analysed with q-PCR in the three samples. Sample AS showed the highest number of bacterial 16S rRNA copies per gram of raw sample ( $3.92 \times 10^9$ ), followed by Qz ( $1.39 \times 10^8$ ) and CS ( $7.57 \times 10^5$ ). Archaeal 16S rRNA copies were three orders of magnitude lower in all samples ( $2.53 \times 10^4$  in Qz,  $9.5 \times 10^4$  in AS and  $8.48 \times 10^2$  in CS).

The taxonomy classification of the SVs was compared with the OTUs obtained through EMIRGE (hereafter named e-OTUs) for each sample under analysis. The only e-OTUs obtained using the bacterial primers were analysed because of the low abundance of SVs affiliated with archaeal taxa

(since the primers for SV analysis (515F and 907R, Table 1) targeted both *Bacteria* and *Archaea*) and because of the low number of archaeal 16S rRNA copies as respect to the number of bacterial 16S rRNA copies obtained through q-PCR. The Pearson correlation value between SV and E-OTUs was high enough in all samples ( $\rho > 0.7$ ) to indicate that the microbial characterization obtained through the analysis of the V4-V5 region was significantly represented by the near-full length 16S rRNA gene sequences obtained with the primer pairs targeting for bacteria (Fig 15).

Members of *Chloroflexi* dominated the microbial community in the pristine quartzitic surface (Fig 17). By using default parameters (minimum identity parameter of 0.95) for taxonomy assignment in SILVA database, Qz microbial community resulted to be dominated (>70% of the entire community) by unclassified bacteria. Only by reducing the minimum identity parameter to 0.8, these unclassified SVs and e-OTUs were affiliated to *Chloroflexi* phylum of *Ktedonobacteria* class and *Ktedonobacterales* order (Fig 17 and Table 11).



**Fig 17.** Microbial community composition at phylum level based on (A) DADA2 analysis of V4-V5 hypervariable region and (B) full-length 16S rRNAs reconstructed through EMIRGE.

In particular, in the SV dataset, almost the entire *Ktedonobacterales* community was constituted only by four SVs, which had a high similarity (>98%) among each other. These four SVs found correspondence (>97% similarity) with the six most abundant e-OTUs which were all affiliated to *Ktedonobacterales* order and were included in the phylogenetic analysis (Fig 18). In the phylogenetic analysis, these e-OTUs had limited sequence similarity (<90%) with other sequences present in the database. Among these, they showed a maximum similarity value of 90% with clone sequences retrieved from the orthoquartzitic Roraima Sur Cave<sup>28</sup>. These sequence identity values further decreased (71-83%) when clone sequences affiliated to *Ktedonobacterales* from other quartz-rich environments were considered. These environments corresponded to cold, arid and nutrient poor soils<sup>18,66</sup> (Fig 18).

**Table 11. Most abundant taxonomic groups in the three floor samples obtained through EMIRGE analysis. Darkening shades indicate higher abundance i.e. white=0-5%, light grey=5-20%, dark grey=20-40%, black: >40%.**

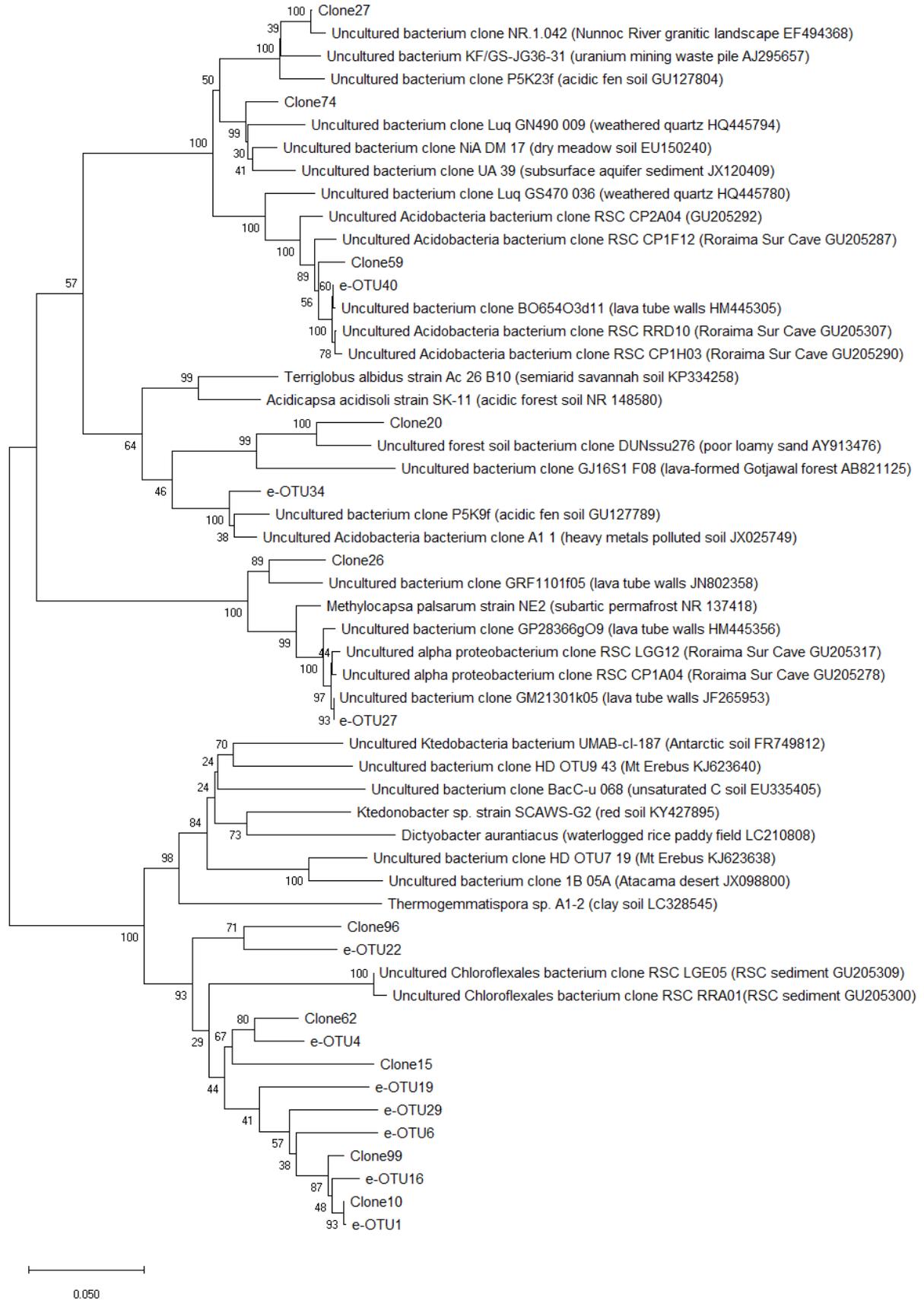
Phylum	Class	Order	Family	Genus	Qz	AS	CS
<i>Acidobacteria</i>	<i>Acidobacteriia</i>	<i>Acidobacteriales</i>	<i>Acidobacteriaceae</i>	Unclassified	0	1.14	5.94
<i>Acidobacteria</i>	<i>Acidobacteriia</i>	Subgroup 2	Unclassified	Unclassified	7.67	0.07	0.91
<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Corynebacteriales</i>	<i>Mycobacteriaceae</i>	<i>Mycobacterium</i>	0.02	2.02	32.77
<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Frankiales</i>	<i>Acidothermaceae</i>	<i>Acidothermus</i>	0.04	0.24	23.79
<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Pseudonocardiales</i>	<i>Pseudonocardiaceae</i>	<i>Crossiella</i>	0.08	40.42	7.43
<i>Actinobacteria</i>	<i>Actinobacteria</i>	<i>Pseudonocardiales</i>	<i>Pseudonocardiaceae</i>	Unclassified	0.02	7.87	0.48
<i>Chloroflexi</i>	<i>Ktedonobacteria</i>	<i>Ktedonobacteriales</i>	Unclassified	Unclassified	73.99	1.02	0
<i>Firmicutes</i>	<i>Bacilli</i>	<i>Bacillales</i>	<i>Bacillaceae</i>	<i>Bacillus</i>	0.01	18.70	1.34
<i>Firmicutes</i>	<i>Bacilli</i>	<i>Bacillales</i>	<i>Paenibacillaceae</i>	<i>Paenibacillus</i>	0	10.22	0
<i>Proteobacteria</i>	<i>Alphaproteobacteria</i>	Unclassified	Unclassified	Unclassified	0.22	0.14	6.07

In order to get deep into the analysis of the taxonomy of the sequences in Qz belonging to *Chloroflexi* phylum, we exploited RFLP-based clone library analyses to retrieve complete nucleotide sequences of the 16S rRNA gene. By analysing 100 RFLP profiles, 15 clusters were identified which showed a taxonomy affiliation that corresponded with the microbial community composition described through Illumina-based sequencing methods (Table 12). The most abundant cluster represented 67% of the whole microbial community in Qz and shared >99% of sequence similarity with *Ktedonobacteriales*-related e-OTU1. Other three clusters were classified as *Chloroflexi Ktedonobacteriales*, together representing 8% of the library (Table 12).

**Table 12. RFLP-based clone library analysis of Qz sample and comparison with the most similar e-OTUs.**

RFLP	% library	Taxonomy <sup>a</sup>	e-OTU	Query coverage (%)	Identity (%)
clone_10	67	<i>o_Ktedonobacteriales</i>	e-OTU_1	92	99.7
clone_15	1	<i>o_Rokubacteriales</i>	e-OTU_1	92	88.42
clone_20	2	<i>o_Acidobacteriales</i>	e-OTU_902	86	92.8
clone_24	1	<i>o_Acidobacteriales</i>	e-OTU_34	91	98.43
clone_26	5	<i>f_Beijerinckiaceae</i>	e-OTU_27	91	94.71
clone_27	2	<i>o_Subgroup 2</i>	e-OTU_137	92	99.64
clone_59	10	<i>o_Subgroup 2</i>	e-OTU_40	93	97.62
clone_62	4	<i>o_Ktedonobacteriales</i>	e-OTU_1	92	92.76
clone_63	1	<i>f_Chthoniobacteraceae</i>	e-OTU_1448	71	83.23
clone_66	1	<i>f_Phycisphaeraceae</i>	e-OTU_1409	84	80.96
clone_70	1	<i>o_Betaproteobacteriales</i>	e-OTU_1097	86	89.29
clone_73	1	<i>o_Subgroup 2</i>	e-OTU_85	92	96.06
clone_74	2	<i>o_Subgroup 2</i>	e-OTU_364	91	96.3
clone_96	1	<i>o_Ktedonobacteriales</i>	e-OTU_435	79	91.52
clone_99	1	<i>o_Ktedonobacteriales</i>	e-OTU_16	75	94.31

<sup>a</sup>“o\_” refers to bacterial orders, “f\_” refers to bacterial families.



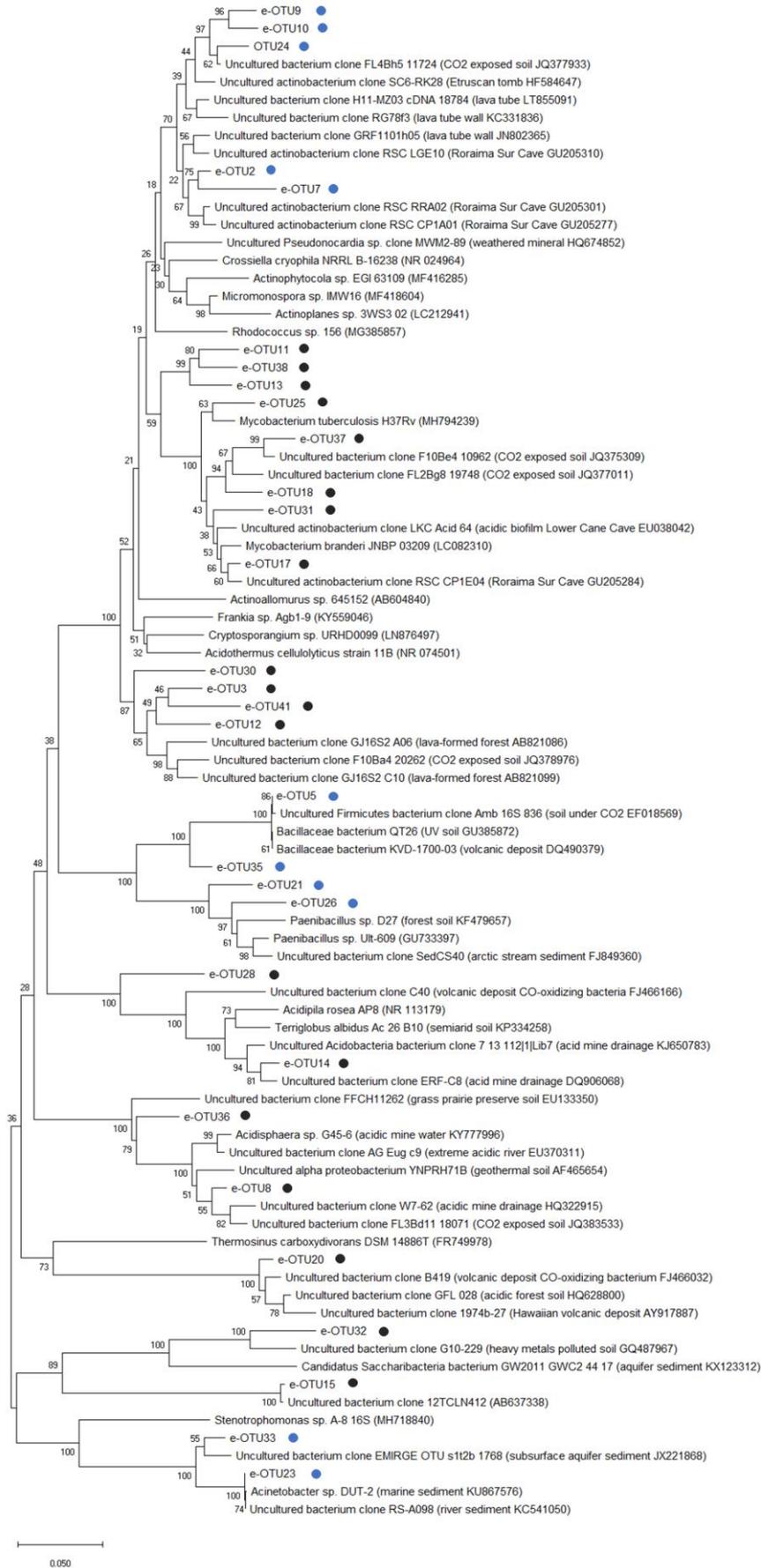
**Fig 18.** Phylogenetic analysis of the most abundant full-length 16S rRNA sequences of samples Qz. Sequences obtained through EMIRGE methods are named "e-OTU", whereas sequences obtained through clone library are named as "Clone".

In agreement with the taxonomy and phylogeny of the microbial community in Qz, we were able to state that a novel taxon of *Chloroflexi* phylum belonging to *Ktedonobacterales* order is dominant in the microbial community associated with the non-amorphous silica sample, representative of the very first colonization of the quartzitic cave surface. *Chloroflexi* represent a deep-branching lineage of the domain *Bacteria* and were first described as Gram-positive green non-sulfur bacteria that were isolated from a thermal environment<sup>67</sup>. They are mostly recognised as anoxygenic phototrophs, but many species are able to switch their metabolisms into heterotrophy in certain environmental conditions, such as aphotic environments<sup>68</sup>. Among these, the class *Ktedonobacteria* was introduced by Cavaletti in 2006<sup>69</sup> and up to date includes the two orders *Ktedonobacterales* and *Thermogemmatisporales*. *Ktedonobacteria* have been described as aerobic sporigen morphologically similar to mycelia-forming actinomycetes<sup>70</sup>, and are proposed to have both heterotrophic and chemolithotrophic metabolisms<sup>71</sup>. FESEM images of sample Qz from Imawari Yeuta showed filamentous structures that strongly resemble SEM morphologies of *Ktedonobacteria* strains obtained from geothermal soils and hot environments, most of them previously described to form spores in the latest stages of the maturation<sup>70</sup>. However, no spores were observed in sample Qz. This dissimilarity can be mostly associated to the absence of high temperatures in the cave. The high abundance of *Ktedonobacteria* members has been previously observed in microbial communities colonizing other silica-rich, nutrient-limited and acidic environments. Most of these environments were also featured by low temperatures, like the Atacama desert, the Antarctic Terra Nova Bay and the dark volcanic ice cave systems located in the Mt. Erebus, and by high CO<sub>2</sub> concentration, like a gas vent in Calatrava<sup>18,21</sup>. Members of undefined taxon of *Ktedonobacterales* order were also found to be dominant in two samples collected from the only other orthoquartzitic cave<sup>28</sup> (Roraima Sur Cave) described so far, which is also located on tepui mountains. Interestingly, the appearance and environmental conditions of Qz sample resemble that described for Roraima Sur Cave samples that were defined as markings on quartzitic ceiling and wall featured by acidic pH. In line with this, by comparing Qz 16S rRNA sequences from Illumina sequencing and clone library with the reference database, *Ktedonobacteria*-related sequences best matched those from Roraima Sur Cave. However, they shared <90% similarity, suggesting their belonging to distinct lower taxonomy levels (family and genus)<sup>72</sup>.

Moving towards late stages of the silica amorphization process, the microbial community composition shifted towards the dominance of *Actinobacteria* (50-70% in AS, 70-80% in CS, Fig 17) accompanied by a gradual decrease of *Chloroflexi* (<6% in AS and <0.1% in CS, Fig 17). *Actinobacteria* phylum was mainly composed by members of the genus *Crossiella* of *Pseudonocardiales* order in AS and by members of the genus *Acidothermus* of *Frankiales* order in

AS and CS, respectively. *Mycobacterium* genus of *Corynebacteriales* order was also detected in both the samples, although it was highly abundant in CS only. Members of these *Actinobacteria* groups were also present in Qz sample but they were found in traces (<0.5%). The progression towards amorphous and coralloid silica states resulted in the enrichment of taxa that were very low abundant in the quartzitic sample Qz. The only EMIRGE approach also revealed the high abundance (around 34%) of members of *Firmicutes* phylum within the AS microbial community, mainly represented by *Bacillus* and *Paenibacillus* genera (Fig 19). The discrepancy with the SV dataset might be due to the different primer sets used for the 16S rRNA gene amplification<sup>73</sup>.

In the phylogenetic tree, the dominant *Crossiella*-related e-OTUs clustered in two different clades (Fig 19). One clade was highly related (>98%) to clones retrieved from Roraima Sur Cave, while the other showed affiliation (>98%) with clones retrieved from soil exposed to high CO<sub>2</sub> concentration. Further, the *Bacillus*-related e-OTU5 abundant (9.1%) in AS community EMIRGE dataset, showed high similarity (99%) with sequences from deposits from volcano or soils exposed to radiation or CO<sub>2</sub>. On the other hand, in the only SV dataset, a *Ktedonobacterales*-related sequence was highly abundant (4.1%) in AS. This sequence showed low similarity with SVs from Qz (maximum sequence identity of 79%) and was related (similarity of 90-92%) with clone sequences retrieved from acidic sulfate soil and heavy metal-contaminated environments (data not shown). In CS microbial community, the dominant *Acidothermus*-related sequences formed one clade with a maximum of 96% similarity with reference sequences of acidophilic actinomycetes retrieved from acid environments (Fig 19).



**Fig 19.** Phylogenetic analysis of the most abundant full-length 16S rRNA sequences of sample AS and CS obtained through EMIRGE with abundances > 1%. Blue dots indicate e-OTUs characterizing sample AS; black dots indicate e-OTUs characterizing sample CS.

If on one hand the dominant microbial taxa conspicuously differed by moving through the silica amorphization steps, on the other hand the low abundant taxonomic groups were shared among the three floor samples and included members of *Acidobacteria*, *Alphaproteobacteria* and *Gammaproteobacteria* (Fig 17 and Table 11). Phylogenetic analyses of these sequences showed high similarities (>97% of sequence identity) with clone sequences of uncultured microorganisms retrieved from other cave systems, including Roraima Sur Cave, lava tube walls, and other volcanic caves featured by biological CO-oxidizing activities (Fig 18 and Fig 19).

### ***The microbial community shifting during silica-based stromatolite formation***

As a result, using both the sequencing approaches, during the development of the novel silica stromatolite found in Imawari Yeuta, the microbial diversity shifted from the dominance of *Chloroflexi Ktedonobacterales* to the dominance of *Actinobacteria*. In particular, each step of the silica speleothem development was featured by specific bacterial taxa (which accounted for more than half of the microbial population), i.e. *Ktedonobacterales (Chloroflexi)* dominated the crystalline quartz sample Qz, *Crossiella (Actinobacteria)*, *Bacillus* and *Paenibacillus (Firmicutes)* were predominant in amorphous silica sample AS, while *Mycobacterium* and *Acidothermus (Actinobacteria)* dominated the coralloid silica sample CS. The present speleothem progression study associates the increase of silica amorphization with the selection of Gram-positive acidophilic *Actinobacteria* and, at a lower extent, *Firmicutes*. This enrichment can be associated with the higher adhesion properties of Gram-positive bacteria able to promote stable connections with siliceous surfaces as compared to Gram-negative bacteria<sup>74</sup>. In this context, the high electronegativity of the cell wall of *Actinobacteria* and *Firmicutes* can facilitate the direct silicate binding, which is involved in siliceous mineral nucleation processes<sup>75</sup>. In addition to these cell wall properties, the intense polysaccharide production and the complex metabolic activities can also support the *Actinobacteria* role in silica solubilization and precipitation processes<sup>75</sup>. In relation to this, previous studies have indicated key activities of members of this phylum in rock destructive and constructive processes and in the precipitation of secondary minerals<sup>76</sup>. Considering the specific actinobacterial genera found in amorphous and coralloid silica samples under analysis, members of the predominant *Pseudonocardiaceae* family (to which *Crossiella* genus belongs) were previously found to populate mats in lava caves and weathered rocks in karst caves<sup>14</sup>, while *Acidothermus* strains were generally detected in acidophilic and hot environments and on cave walls of a thermal karst system<sup>77</sup>. We hypothesize that both the low pH and the nature of the quartzitic substrate has led to this selection. In a similar way, *Mycobacterium* strains have been described to thrive in silica-rich environments (e.g. human lungs in silicosis) and to colonize acidic environments. During the formation of the coralloid

silica in CS, pH slightly increases without reaching neutral values. Moreover, an increase of minerals and metals concentration was observed as respect to the first stages of the silica amorphization. In this regard, *Mycobacterium* has been reported to have an active role in iron mineralization supported by the activity of specific enzymes such as bacterioferritin A<sup>78</sup>. Among *Firmicutes* phylum, *Bacillus* spp. strains were widely studied for their biomineralization capacities associated with EPS production secretion and alkalisation<sup>79</sup>. In particular, *Bacillus* strains were demonstrated to have silica solubilising activities, which can support the role of these members in silica speleothem development<sup>46</sup>. Both the alkalisation and increase of metal amount during silica amorphization progression independently on the quartzitic cave location (wall, floor or water body)<sup>10</sup>. Similarly, in the present work, we revealed an alkalisation process and an increase of the amount of metals like barium and zinc during the formation of the coralloid silica speleothem. These results further supported the parallelism between silica amorphization, metal content and pH value previously reported<sup>10</sup>, which could both be the result and influence the enrichment of specific microbial taxa during the progression from the non-amorphous silica to the coralloid sample.

In addition to their biomineralization activities, both *Actinobacteria* and *Firmicutes* have heterotrophic metabolism associated to complex polymers breakdown and chemotrophic activities involved in N cycle, which may support and sustain microbial communities syntrophy relationships. Both *Actinobacteria* and *Firmicutes* members are able to survive under stressful and nutrient limiting conditions<sup>80</sup> with some species being particularly resistant thanks to the ability to form endospores and to be multidrug resistant (e.g. *Paenibacillus*)<sup>81</sup>. Further, some genera of *Actinobacteria* and *Firmicutes* were shown to be able to exploit mixotrophic metabolisms in case of necessity<sup>21</sup>. Together, these aspects sustain their dominance in oligotrophy and starving conditions represented by the samples under analysis.

### ***Clone libraries of coxL and hypD genes for Qz sample***

In order to get deeper knowledge of the *Ktedonobacteria*-dominating microbial community associated to the inception of a silica stromatolite, we performed additional clone library analyses on Qz sample targeting two functional genes, i.e. *coxL* gene encoding the large catalytic subunit of the carbon monoxide dehydrogenase (CODH)<sup>82</sup> and *hypD* gene encoding a protein involved in the maturation of [NiFe]-hydrogenases<sup>40</sup>. In this regard, the uptake and oxidation of CO and H<sub>2</sub> are traits that were suggested to be common among *Ktedonobacteria*<sup>83</sup> and have been used as markers of the capacity to utilize atmospheric trace gases<sup>41</sup>. The key role of CODH and [NiFe]-hydrogenases activities was recognised as crucial for life development under nutrient source limitation and was associated to possible pioneering colonisation capacities<sup>84</sup>. Indeed, *coxL* and *hypD* genes might be

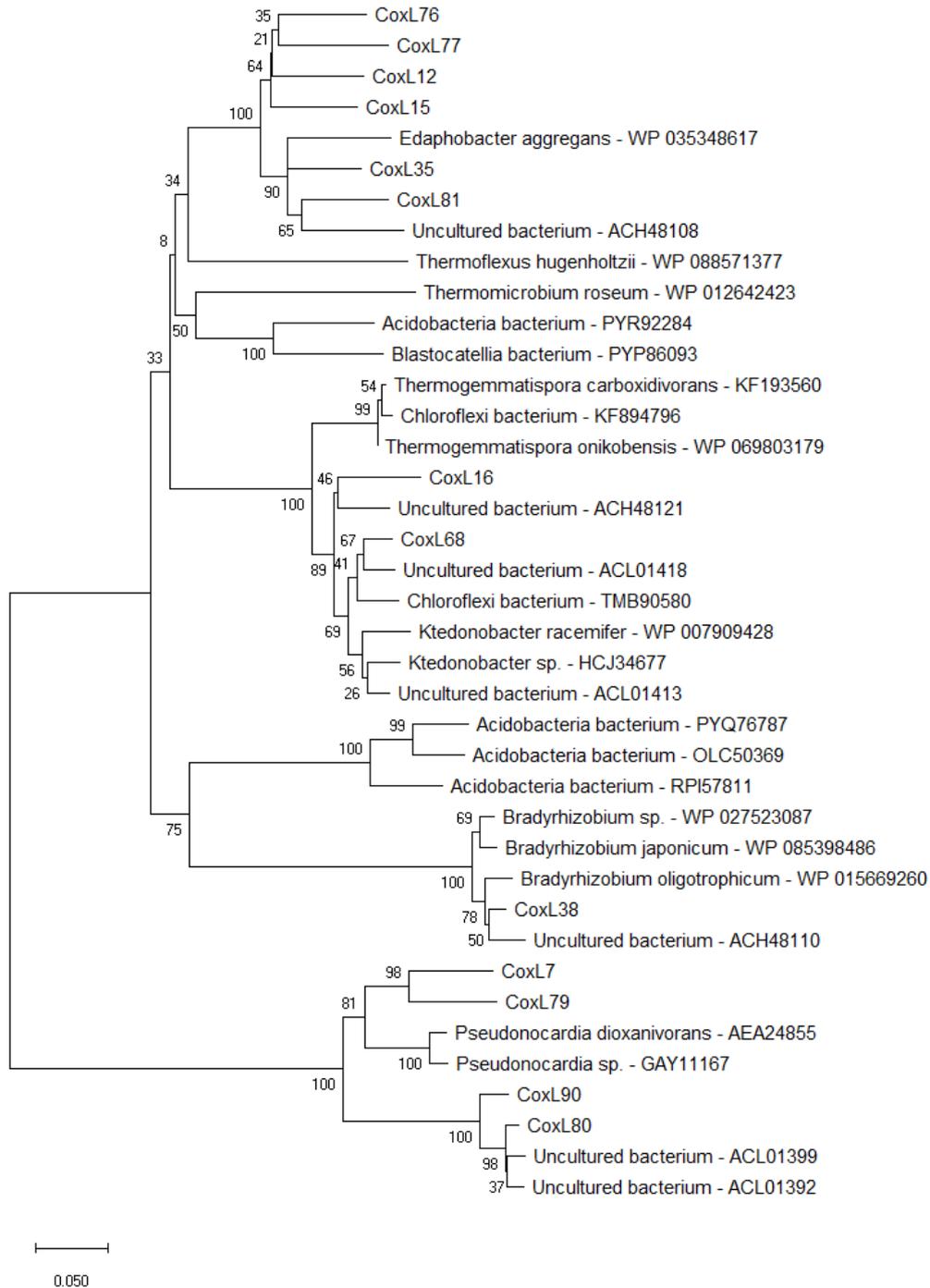
involved in the utilization of carbon monoxide and hydrogen as alternative energy sources in starving and dark environments where organic electron donor are scarce<sup>83</sup>. Therefore, in relation to this, the investigation of these two genes in *Ktedonobacterales* members dominating the pristine quartzitic Qz sample not only provided phylogenetic information but also gave hints on their capacity to act as first colonisers of the quartzitic sample.

The *coxL*-based library screening identified 13 clusters on the basis of the RFLP analysis (Table 13). Sequence analysis of representative clones from each group revealed that all sequences belonged to *coxL* form I, with the exception of one phylotype (represented by one clone) which belonged to *coxL* form II and was excluded from further analysis. The phylogenetic analysis of the amino acid sequences showed that they were mainly related to reference CoxL sequences related to *Acidobacteria Edaphobacter* and *Chloroflexi Ktedonobacter* (33% with 6 RFLP groups and 25% with 2 RFLP groups, respectively) mainly retrieved from agricultural soils and volcanic soils and deposits<sup>82</sup> (Fig 20).

**Table 13. RFLP-based clone library analysis of *coxL* genes in Qz sample.**

Clone <sup>a</sup>	% library	Best Blast classified Hit			
		Taxonomy	Phylum	Identity (%)	Accession number
CoxL7	3	<i>Pseudonocardia dioxanivorans</i>	Actinobacteria	86	WP_013674779.1
CoxL12	5	<i>Edaphobacter aggregans</i>	Acidobacteria	86	WP_035348617.1
CoxL15	3	<i>Edaphobacter aggregans</i>	Acidobacteria	85	WP_035348617.1
CoxL16	22	<i>Ktedonobacter sp.</i>	Chloroflexi	91	HAH00841.1
CoxL35	3	<i>Edaphobacter aggregans</i>	Acidobacteria	88	WP_035348617.1
CoxL38	22	<i>Bradyrhizobium oligotrophicum</i>	Proteobacteria	97	WP_015669260.1
CoxL68	3	<i>Ktedonobacter sp.</i>	Chloroflexi	94	HAH00841.1
CoxL76	11	<i>Edaphobacter aggregans</i>	Acidobacteria	84	WP_035348617.1
CoxL77	3	<i>Edaphobacter aggregans</i>	Acidobacteria	84	WP_035348617.1
CoxL79	3	<i>Pseudonocardia dioxanivorans</i>	Actinobacteria	85	WP_013674779.1
CoxL80	8	<i>Pseudonocardia dioxanivorans</i>	Actinobacteria	81	WP_013674779.1
CoxL81	8	<i>Edaphobacter aggregans</i>	Acidobacteria	86	WP_035348617.1
CoxL90	3	<i>Pseudonocardia dioxanivorans</i>	Actinobacteria	83	WP_013674779.1

<sup>a</sup>Representative clone of each RFLP-based cluster.



**Fig 20.** Phylogenetic analysis of CoxL sequences of Qz sample.

The *hypD*-based library analysis revealed 9 different phylotypes on the basis of the RFLP patterns (Table 14). Clones related to *Acidobacteria* sharing <89% of aminoacidic identity with database references represented five groups of the *hypD* library. Given the low identity percentage, they clustered apart forming a different clade in the tree. The other two groups were related to *Chloroflexi* and *Verrucomicrobia*, sharing 96% and 95% with the closest reference sequence respectively. Phylogenetic analysis of *Acidobacteria*- and *Verrucomicrobia*-related amino acid sequences revealed

similarities with clones retrieved from meadow soil. Phylogenetic analysis of *Ktedonobacteria*-related sequences revealed similarities with clones retrieved from clay soils in geothermal areas (Fig 21).

Clone <sup>a</sup>	% library	Best Blast classified Hit			
		Taxonomy	Phylum	Identity (%)	Accession number
HypD1	4	<i>Edaphobacter aggregans</i>	Acidobacteria	81	WP_035352780.1
HypD3	4	<i>Edaphobacter aggregans</i>	Acidobacteria	82	WP_035352780.1
HypD4	50	<i>Edaphobacter aggregans</i>	Acidobacteria	82	WP_035358998.1
HypD11	4	<i>Pedosphaera parvula</i>	Verrucomicrobia	93	WP_007412928
HypD16	11	<i>Chloroflexi bacterium</i>	Chloroflexi	94	TME01503.1
HypD28	10	<i>Edaphobacter aggregans</i>	Acidobacteria	81	WP_035358998.1
HypD29	1	<i>Acidobacteria bacterium</i>	Acidobacteria	82	PYX55871.1
HypD64	2	<i>Acidobacteria bacterium</i>	Acidobacteria	76	OFV79518.1
HypD72	14	<i>Edaphobacter aggregans</i>	Acidobacteria	82	WP_035358998.1

<sup>a</sup> Representative clone of each RFLP-based cluster.

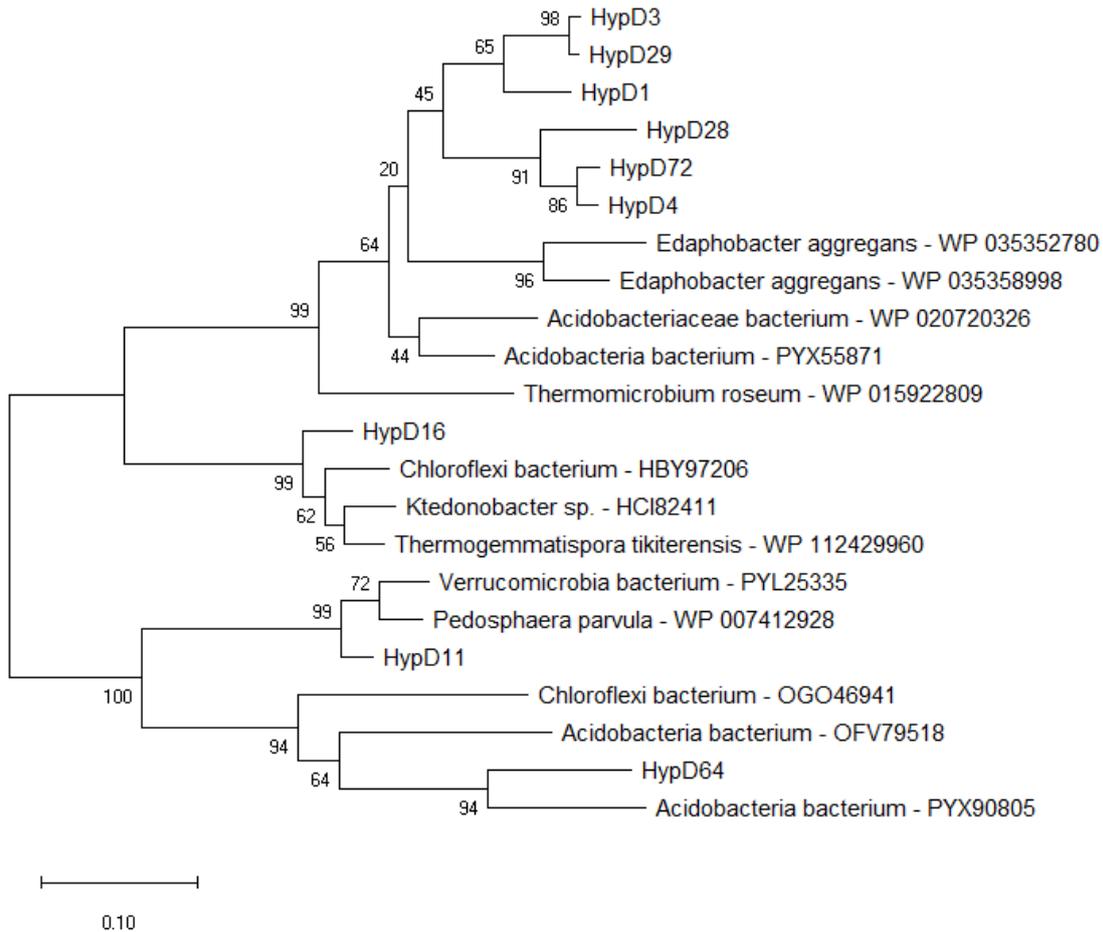


Fig 21. Phylogenetic analysis of HypD sequences of Qz sample.

Dominant CoxL and HypD sequences from Qz were related to the corresponding gene products of members of *Ktedonobacteriales* as well as *Edaphobacter*, which is an *Acidobacteria* genus. On the other hand, both phylogenetic trees did not show a clear clusterization of the amino acid sequences

in relation to the taxonomy information retrieved from their best matches with the database references. Indeed, the Qz dominant CoxL and HypD sequences, which affiliated to *Acidobacteria*, clustered separately from the reference *Acidobacteria* (Fig 20 and Fig 21). On the contrary, the CoxL and HypD sequences representing low abundant RFLP groups were classified with better certainty as they grouped closer to the corresponding reference sequences in the trees. The low similarity (<89%) between the dominant *Acidobacteria*-related CoxL and HypD sequences and the reference sequences might indicate their belonging to still undescribed bacterial taxa able to thrive at mesophilic temperature and could be associated with the possibly novel *Ktedonobacterales* family or genus described through 16S rRNA analysis. In this respect, some evolutionary events might have led to horizontal gene transfer (HGT) events. This is in line with recent evidences of HGT within the *Chloroflexi* phylum as a way to adapt in peculiar environmental conditions<sup>83</sup>, also considering the high number of transposases found in some genomes of *Chloroflexi* strains<sup>85</sup>.

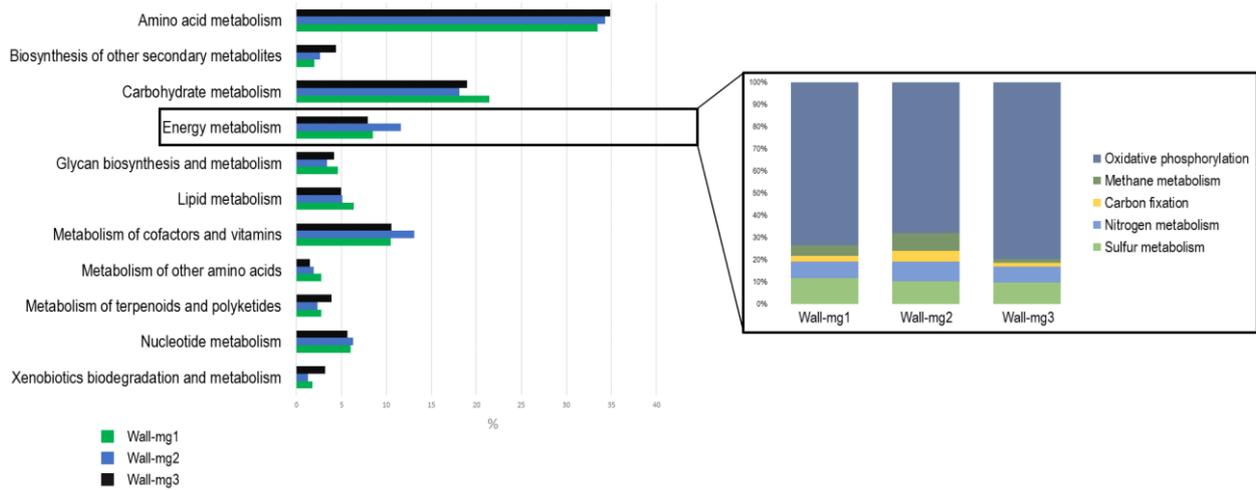
### 3.3.4. Carbon and nitrogen fixation potential in the metagenome of wall-related microbial communities

Metagenomic analyses of three samples collected from the wall of the cave were performed because of the higher level of oligotrophy expected on walls as compared to the cave pavement and roof. The three samples belong to the same cave wall and represent consecutive stages in the silica amorphization process. The study of the metabolic potential aims at providing functional insights into the capacity of wall-related microbial communities to thrive under oligotrophic conditions in silica-rich environments contributing to silica mobilization processes.

Metabolic potential analysis was performed by mapping contigs to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The results showed that the most represented genes belonged to Metabolism category (Table 15) accounting for 58-64% of the dataset, following by genes involved in Genetic information processing (16-20%), Environmental information processes (12-18%), Cellular processes (4-6%), Human diseases (1%) and Organismal systems (<1%).

**Table 15. Summary of metagenomes analysis from cave wall samples.**

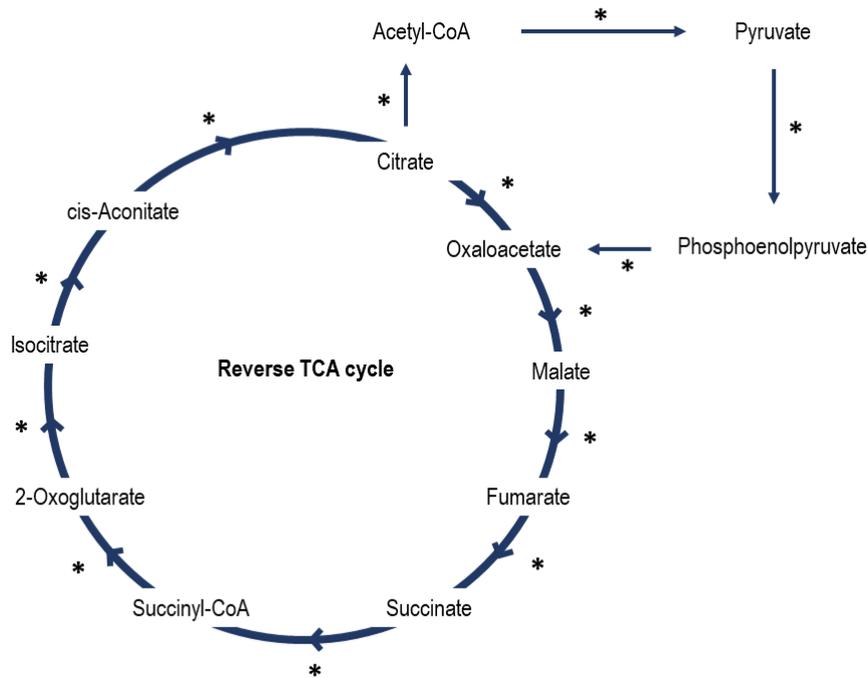
Sample	Reads	Contigs	N50 contig size	Predicted genes	Distribution (%) of sequences in major KEGG categories					
					Metabolism	Genetic information processing	Environmental information processing	Cellular processes	Human diseases	Organismal systems
Wall-mg1	496124	7869	806	8105	58.7	15.6	17.5	6.1	1.4	0.2
Wall-mg2	399960	10248	924	12050	58.1	20.6	15.4	4.7	1.0	0.3
Wall-mg3	616766	6855	733	6807	63.5	18.7	11.9	4.3	1.3	0.3



**Fig 22.** KEGG categories belonging to metabolism (left) and energy metabolism (right) retrieved from the three wall metagenomes.

Analysis of the sublevels in KEGG showed that genes involved in amino acid and carbohydrate metabolic pathways dominated the metabolism category (Table 15 and Fig 22). These two pathways are common in all life forms and therefore, in order to get deep into the analysis of metabolisms involved in microbial development under energetically unfavourable cave conditions, we focused on the functions included in the category of energy metabolism. Moreover, in Imawari Yeuta, the bacterial diversity of samples collected from the extremely oligotrophic wall habitat is higher compared to other cave samples (paragraph 3.3.1) opening fundamental questions on the way bacterial communities obtain energy and carbon sources to develop. Potential primary production strategies (associated to the capacity to utilize inorganic carbon and nitrogen) in Imawari Yeuta might be a crucial aspect considering the low organic carbon content within the cave. In environmental microbial communities, most of the energy is usually obtained through photosynthesis processes<sup>45</sup>. Conversely, in agreement with the deep origin of the sample, genes related to photosynthesis were absent. However, not only carbon fixation reactions can be carried by photosynthetic microorganisms (photoautotrophic), but also chemolithoautotrophic ones. KEGG analysis of the three wall samples identified putative genes from all six CO<sub>2</sub>-fixation pathways in the metagenomes. Interestingly, the metabolic pathways belonging to Calvin-Benson-Bassham (CBB) were poorly represented, whereas those belonging to reductive citric acid cycle Arnon-Buchanan (rTCA) was almost totally present. The CBB cycle is the most common autotrophic pathways, which generally characterizes phototrophic bacteria, but is also used by chemotrophs to gain energy from inorganic compounds by carbon fixation<sup>86</sup>. Within CBB, RuBisCO is described as the key enzyme present in microbial taxa thriving in oligotrophic environments including dark caves<sup>84</sup>. Nevertheless, in Imawari Yeuta, the RuBisCO-encoding *rbcL* genes were found only in sample Wall-mg2 and were affiliated to *Alphaproteobacteria*. On the other hand, all the three metagenomes possessed most of the genes

involved in the rTCA pathway. The rTCA pathway (Fig 23) produces carbon compounds from carbon dioxide and water, and it is probably one of the main strategies for bacteria aiming at producing energy. These results suggest that carbon fixation in Imawarì Yeuta relies mainly on rTCA. The presence of rTCA pathway instead of CBB may be due to an evolutionary aspect of this peculiar cave microbiome. Indeed, rTCA pathway is considered one of the pioneering metabolic processes leading to microbial life initial colonization in oligotrophic and dark habitats<sup>87</sup>.



**Fig 23.** Graphical representation of the rTCA cycle. Asterisks indicate pathways with genes present in the metagenomes.

Concerning the metabolism of nitrogen, any nitrogenase function were identified in the metagenomic data. This was surprising because nitrogen fixation is often a key process for the sustainment of complex microbial communities even in cave environment<sup>88</sup>. On the other hand, genes involved in nitrification and denitrification pathways, as well as in assimilatory and dissimilatory nitrate reduction, were identified (Fig 24). Nitrification processes potentially provide energy in oligotrophic environments<sup>45</sup>. In Imawarì Yeuta we found abundance of genes involved in nitrate reduction, such as *nirABSK* and *norAB* genes, which are involved in nitrite and nitric oxide reduction. *Proteobacteria* account as the main phylum ruling the nitrogen metabolism in the cave as most of the predicted genes were affiliated with the phylum members. Interestingly, in the first stage of silica amorphization (Wall-mg1 sample), these processes were mainly associated to *Betaproteobacterales*. By moving to the late stages (Wall-mg3 sample), *Rhizobiales* dominated.

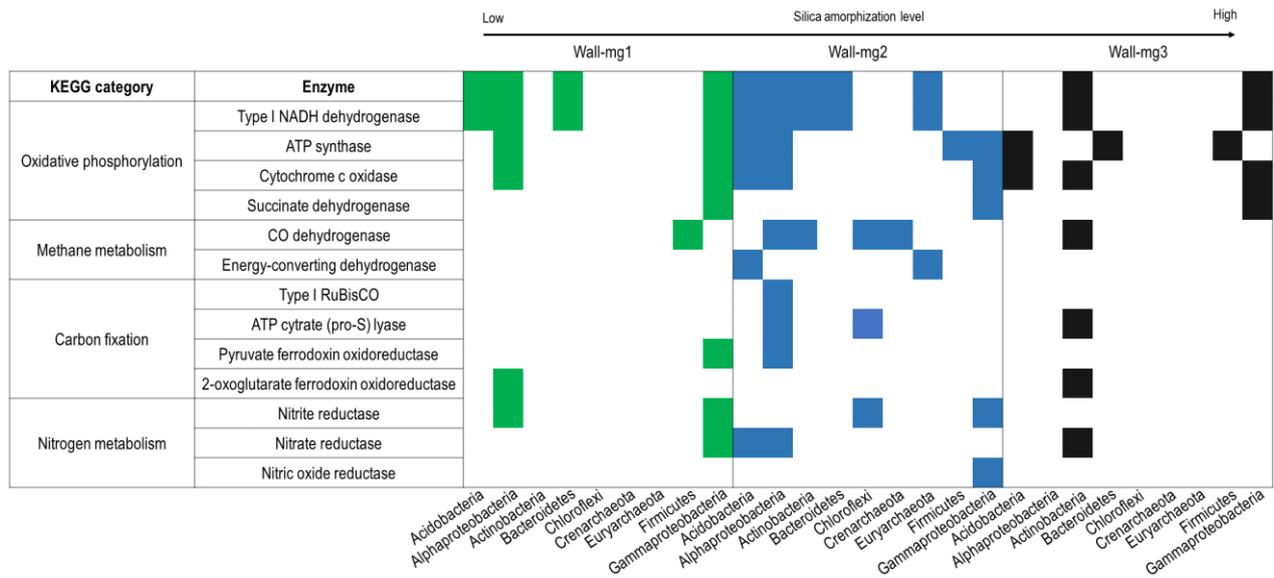


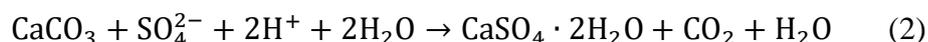
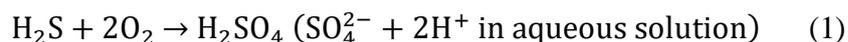
Fig 24. Summary of the key enzymes belonging to the energy metabolism category.

## 4. Geomicrobiology of the seawater-influenced sulfuric acid cave Fetida

Part of this chapter is included in: D'Angeli IM\*, Ghezzi D\*, Leuko S, Firrincieli A, Parise M, Fiorucci A, et al. (2019) *Geomicrobiology of the seawater-influenced sulfuric acid Fetida cave. PLoS One 14(8): e0220706. (\*Co-first authors)*

### 4.1. Introduction

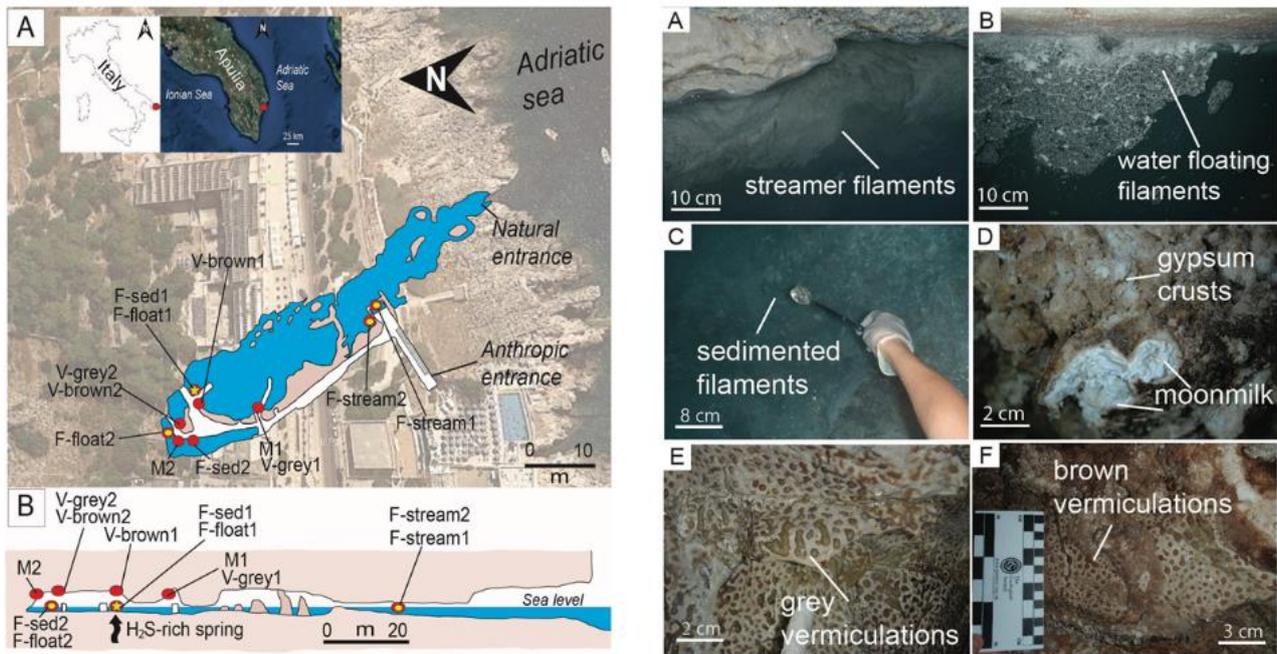
Sulfidic caves form by carbonic acid dissolution through the process of sulfuric acid speleogenesis (SAS), in which hydrogen sulfide is oxidized into sulfuric acid that is dissolved [see reaction (1)]. The origin of hydrogen sulfide can be abiotic, as a result of thermochemical sulfate reduction, or can be the product of microbial reduction and of the decomposition of sulfur-based organic matter. Sulphidic caves have been found all over the world being formed in carbonate rocks in different types of climates. Interestingly, Italy hosts around 25% of the known worldwide SAS systems<sup>19</sup>. In sulfidic caves, the dissolution of carbonates caused by sulfuric acid [see reaction (2)] is rapid, and immediately induces the replacement of the host rock by gypsum and the release of CO<sub>2</sub> into the environment.



The formed gypsum either solubilizes in the waters of the caves by inducing modifications of their structure or can deposit on the walls replacing the limestone. The oxidation of H<sub>2</sub>S provides an important energy source for sulfur oxidizing microorganisms, which are able to sustain the microbial ecosystem in aphotic sulfidic environments, acting as primary producers and supporting growth of other microorganisms<sup>89</sup>. Many different members of *Bacteria* and *Archaea* oxidize sulfur compounds by exploiting metabolic reactions that are much faster than the abiotic reactions, by using either nitrate or oxygen as electron acceptor<sup>75</sup>. When nitrate and oxygen are not available, some microbes are able to oxidize hydrogen sulfide producing elemental sulfur, which precipitates and provides energy for the cell. The biological oxidation of H<sub>2</sub>S also generates local acidity that cause carbonate dissolution and enhance the speleogenesis. Mineralized structures rich in sulfur compounds are also considered biosignatures and, together with the high temperatures and low oxygen concentration, can be used as analogues for the study of the development of life on Earth and on other planets<sup>90</sup>.

This chapter investigates the geomicrobiology of three types of biofilms/biodeposits (i.e. water filaments, vermiculations and moonmilk deposits) showing distinct distribution pattern in the sulfidic Fetida cave that opens at sea level (Fig 25). Fetida Cave represents a unique environment for the study of microbial biogeography and biodiversity in a sulfide-rich aphotic habitat influenced by seawater. In this work, the mineralogy, geochemistry and microbial diversity associated with the three different

biofilms/deposits from Fetida Cave are described in order to provide a combination of geochemical and biological insights into sulfur-rich environments influenced by seawater. From a microbiological point of view, the investigation of the role of the microbial communities aims at understanding the activity of specific microbes in the biogeochemical cycles, with particular focus on the metabolism of sulfur and nitrogen.



**Fig 25.** Map of Fetida cave with sampling points (left) and representative pictures of the different deposits collected from the cave (right) (D'Angeli et al. 2019<sup>20</sup>). This cave is located in Santa Cesarea Terme (Apulia, Italy). Geographic coordinates: 40°02'04.3"N 18°27'20.6"E) and is a 150 m long cave partially submerged. It consists of a main cave stream characterized by the mixture of seawater and H<sub>2</sub>S-rich upwelling fluids. Fluctuation of the seawater during the seasonal changes impact on the cave geology and microbiology, leading to modifications of the concentration of H<sub>2</sub>S in the cave atmosphere and water.

## 4.2. Materials and Methods

### 4.2.1. Samples list, DNA extraction and sequencing

Samples were collected and stored as described in paragraph 2.1 before their processing in the laboratory of the Medical University of Graz (Austria) and Laboratory of Molecular and Applied Microbiology of the University of Bologna (Italy).

Samples of each biofilm/deposit were collected from different sites along the cave during two sampling campaigns (October 2015 and December 2017). The sampling sites and corresponding sample name are displayed in Fig 25 and Table 16.

Table 16. List of samples collected from Fetida cave.					
Sample name	Type of deposit	Cave location	Colour of the deposit	Method of analysis	
F-stream-1	White filament	Water	White	FESEM, DADA2, geochemistry	
F-stream-2	White filament	Water	White		
F-float-1	White filament	Water	White		
F-float-2	White filament	Water	White		
F-sed-1	White filament	Water	White		
F-sed-2	White filament	Water	White		
V-brown-1	Vermiculation	Wall	Brown		
V-brown-2	Vermiculation	Wall	Brown		
V-grey-1	Vermiculation	Wall	Grey		
V-grey-2	Vermiculation	Wall	Grey		
M-1	Moonmilk	Roof	White		
M-2	Moonmilk	Roof	White		
Vg-mg1	Vermiculation	Wall	Grey		Shotgun metagenomics
Vg-mg2	Vermiculation	Wall	Grey		
Vb-mg1	Vermiculation	Wall	Brown		
Vb-mg2	Vermiculation	Wall	Brown		
Vv-mg1	Vermiculation	Wall	Violet		
Vv-mg2	Vermiculation	Wall	Violet		

Total DNA extraction, amplification of V4-V5 regions and Illumina sequencing were performed as described in paragraph 2.2. Processing of raw data with DADA2 package and statistical analyses were performed as described in paragraph 2.3.

#### 4.2.2. Geochemical analyses

Geochemical analyses were performed as described in 2.6. Scanning electron microscope analyses were performed as described in 2.7.

### 4.3. Results and Discussion

#### 4.3.1. Samples description

This work focuses on three types of biofilms/deposits which are related to microbial growth and have different morphologies and specific distribution, i.e. in the cave water, or on the walls and ceiling of Fetida Cave. In particular, the biofilms in the water can be categorized in white streamer filaments attached to the rock (named F-stream samples), white filaments floating on the water (named F-float samples) and biofilms sedimented on the stream bed (named F-sed samples) (Fig 25). Despite being more visible in the inner part of the cave (except for the F-stream samples), these filamentous biofilms were generally sparsely distributed along the cave water stream depending on the sea water movement and tidal conditions. During days of low tide and calm seawater, representative samples for each type of filament have been collected along the cave, at various distances from the sulfidic spring (Fig 25).

A total number of six filaments were collected along the cave length and were named F-float-1 and F-float-2, F-sed-1 and F-sed-2, F-stream-1 and F-stream-2 on the basis of their appearance (floating, sedimented or streamers). The latter two samples were collected at the cave entrance, as the streamer appearance was typically present on the two side walls of the water stream in this cave zone, while the sedimented and floating biofilms/filaments were generally more abundant inside the cave. F-sed-1 and F-sed-2 were collected at the bottom of the water stream (around 1 m below the water level). Only in the inner part of the cave, different biofilms/patinas covered the walls and ceiling showing two main different shapes, i.e. elongated spotted biofilms of different colors (mainly shades of grey and brown), which resemble the “biovermiculations” previously described<sup>91</sup>, and bright white deposits referable to moonmilk deposits. Moonmilk is a term to describe a soft, wet, plastic, fine grained speleothem (i.e. secondary cave deposit), which is typically found on the walls of caves<sup>92</sup>. Both vermiculations and moonmilk deposits copiously covered Fetida Cave walls and ceilings. While vermiculations were observed immediately above the water level (a.w.l), moonmilk deposits were solely found from 1 m a.w.l. upward, in areas where condensation was more abundant.

#### **4.3.2. Geochemistry, distribution and morphology of Fetida Cave biofilms**

Geochemical analyses indicate that the cave environment is mainly influenced by two factors, i.e. the seawater effect entering the cave from the natural cave entrance and the H<sub>2</sub>S-springs arising and mixing with the seawater inside the cave. The importance of the sulfuric acid processes strongly depends upon environmental (i.e. tides) and climatic conditions (i.e. wave action). In general, moving from the entrance towards the deep zone of the cave, the marine influence decreases and the effect of rising acidic H<sub>2</sub>S-rich waters increases. In particular, in correspondence of the in-cave sulfidic spring inlet, subaerial and submerged environments were featured by more constant physico-chemical parameters and higher concentration of H<sub>2</sub>S and temperatures (slightly thermal) compared to the cave entrance. Seawater plays a key role in buffering the H<sub>2</sub>S-rich spring water to circumneutral pH in Fetida Cave. This aspect differentiates Fetida Cave from the freshwater sulfidic caves (e.g. Frasassi Cave, Lower Kane Cave) previously studied, in which the carbonate dissolution was reported to be the main driver in lowering acidity<sup>93</sup>. The lower saturation indexes of the cave water as compared to seawater indicates that the host rock (limestone) dissolution is still in an active process in the inner portion of Fetida Cave. Sulfuric acid production arising from microbial oxidation by hydrogen sulphide (both in waters and cave walls) is supposed to influence this speleogenesis process.

In relation with this, abundant microbial biofilms/deposits were visible in the water pool and on the walls and ceilings of the inner zone of Fetida Cave (FC), suggesting that the development of resident microbial communities was closely related to the rising sulfidic fluids and H<sub>2</sub>S degassing. The water

biofilms visible in FC have a morphology similar to that described in other sulfidic cave water streams<sup>61</sup>. However, FC water biofilm pattern was not constant over the year and, generally, biofilms were less thick and dense as compared to the distribution of similar biofilms described in other sulfidic caves<sup>94</sup>. This is due to the fact that Fetida Cave is subject to seawater hydrodynamics which dilutes and occasionally washes the water biofilms away during exceptional tides or high waves. Unlike the biodeposits covering the cave walls and ceilings, water biofilms were also present at the entrance of Fetida. They had a prevalent morphology of thin filaments attached to the submerged rocks on the sides of the water stream (named in this chapter F-stream), whereas the water filamentous biofilms inside the cave were generally thicker, more abundant and floating on the stream or deposited at the bottom (named F-float or F-sed) (Fig 25). The development of F-stream was probably associated to specific environmental conditions occurring in the most external cave zone in association to i) the physical-chemical gradient created along the cave by the mixture of the warm sulfidic fluid with the marine water and ii) the water turbulence due to the higher exposure to the seawater currents. The microbial communities of F-stream showed important differences respect to F-sed and F-float and these are described in the next paragraph.

Abundant vermiculations covered the walls of the inner zone of FC which were featured by irregular morphology, mainly spotted shape, and different colors (i.e. light and dark grey, light and dark brown, red and black), being the grey and brown vermiculations the most abundant. The coloration was not apparently related to specific regions inside the cave, however, the vermiculations with the same aspect were grouped in clusters of different dimensions.

The moonmilk deposits were visible as bright white microcrystalline assemblages of gypsum. They are distinguished from gypsum crusts (that were also present in Fetida Cave and often surrounded the moonmilk deposits) because of their soft texture that resembles toothpaste. As they are intimately attached to the host rock, their development seems associated with carbonate dissolution and gypsum replacement. Indeed, their sampling required carving in the rock, whereas vermiculations were exposed and easy to sample.

Water filaments, vermiculations and moonmilk deposits were mainly distinguished by pH values. White filaments have neutral pH (pH around 7), grey and brown vermiculations are slightly acidic (pH of 5-5.5), whereas moonmilk deposits are extremely acidic with a pH of 0-1. The main difference between F-float and F-sed samples included the concentration of inorganic C that was more abundant in the sedimented biofilms as compared to the floating ones (Table 17). F-stream samples were dominated by Na and Mg, elements that are mainly associated to the seawater composition, while they were poor in N and C. Brown vermiculations resulted to be richer in total N, organic C, Ca, Mg, Fe, Mn, P, Co, and Cu, as compared to the grey vermiculations, which showed the highest

concentration in K (Table 18). Moonmilk was generally poor in all the elements except for Ca, which is mainly related to host rock dissolution. The elemental differences could be associated to the pigmentation of each biofilm, e.g. brown vermiculations contain hematite, while greyish ones are rich in muscovite and quartz, whereas moonmilk is exclusively made of gypsum.

**Table 17. Physico-chemical properties and composition of the Fetida cave biofilm and deposits.**

Type of biofilm	Main color	Mineralogy	T °C	pH	N tot <sup>a</sup>	C org <sup>a</sup>	C inorg <sup>a</sup>	C tot <sup>a</sup>	C:N
Water streamer filaments (F-stream)	White	ND	23.2 ±3.0	7.4	0.86 ±0.032	4.81 ±0.17	0.2 ±0.06	5.01 ±0.10	5.8
Filaments floating on the water (F-float)	White	ND	25.0 ±2.8	6.9	2.08 ±0.095	11.20 ±3.34	0.82 ±0.34	12.05 ±0.40	5.8
Sedimented water filaments (F-sed)	White	ND	24.2 ±2.1	6.9	0.51 ±0.025	3.75 ±0.40	5.57 ±0.14	9.32 ±0.31	18.3
Vermiculation (V-brown)	Brown	Quartz (SiO <sub>2</sub> ), Diopside (CaMgSi <sub>2</sub> O <sub>6</sub> ), Hematite (Fe <sub>2</sub> O <sub>3</sub> )	22.8 ±0.5	5-5.5	0.97 ±0.02	7.25 ±0.42	0.00	7.24 ±0.18	7.5
Vermiculation (V-grey)	Grey	Quartz (SiO <sub>2</sub> ), Calcite (CaCO <sub>3</sub> ), Muscovite [KAl <sub>2</sub> (Si <sub>3</sub> Al)O <sub>10</sub> (OH,F) <sub>2</sub> ], Gypsum [CaSO <sub>4</sub> ·2H <sub>2</sub> O]	22.8 ±0.5	5-5.5	0.31 ±0.04	2.34 ±0.19	0.00	2.32 ±0.13	7.5
Moonmilk (M)	White	Gypsum [CaSO <sub>4</sub> ·2H <sub>2</sub> O]	22.8 ±0.5	0-1	0.00	0.14 ±0.05	0.04 ±0.37	0.19 ±0.01	-

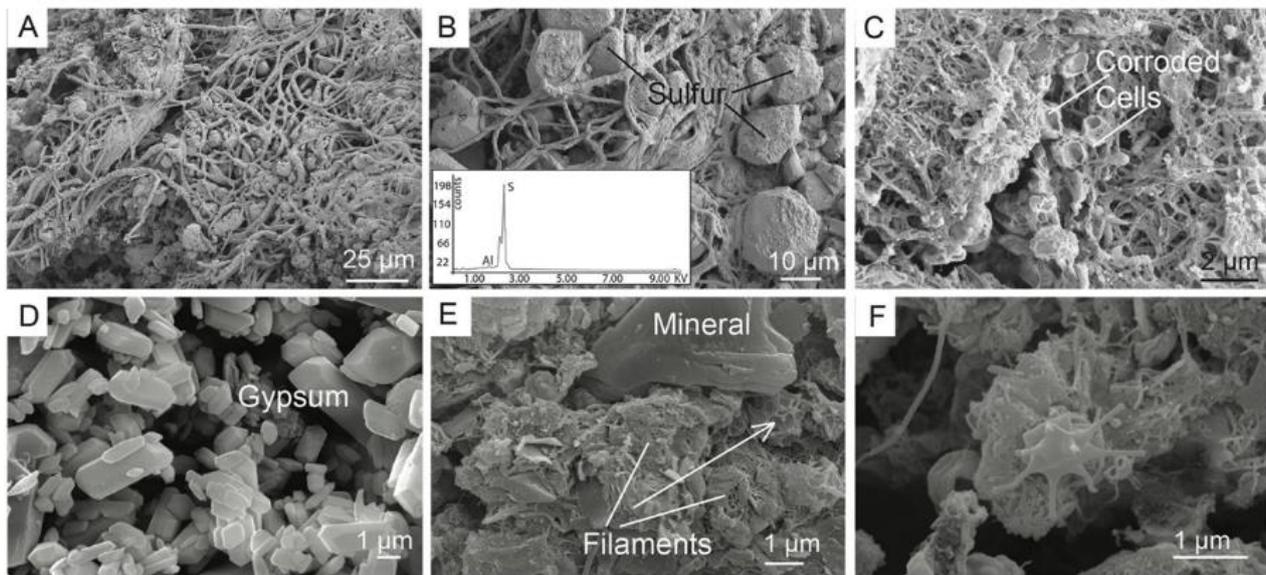
<sup>a</sup> Values are expressed in % dry weight.

**Table 18. Chemical elements in Fetida samples.**

Sample	Na	K	Ca	Mg	Fe	Mn	P	Co	Cu	Zn
F-stream-1	231.41±19.32	43.87±0.89	1.79±0.20	5.48±0.54	0.04±0.01	0.00	0.27±0.09	0.07±0.02	1.37±0.22	0.00
F-stream-2	192.31±10.87	37.25±3.95	2.91±0.38	5.63±0.41	0.29±0.05	0.01±0.00	0.86±0.08	0.26±0.01	18.99±1.41	0.33±0.03
F-float-1	93.15±7.57	41.40±3.44	8.40±0.28	4.31±0.12	3.44±0.22	0.11±0.01	3.31±0.38	2.51±0.22	26.74±2.60	0.04
F-float-2	101.37±5.16	28.63±1.55	6.59±0.43	2.26±0.67	2.88±0.17	0.10	2.52±0.12	1.93±0.10	36.55±3.06	0.03
F-sed-1	60.42±7.09	28.13±5.09	25.04±1.42	2.26±0.67	2.14±0.18	0.12±0.01	1.23±0.15	2.05±0.32	30.63±4.92	0.10±0.02
V-brown-1	2.13±0.47	14.85±1.64	41.28±1.11	15.71±0.55	36.97±2.16	1.37±0.12	8.73±1.08	19.61±1.71	58.89±8.45	0.14±0.01
V-brown-2	1.81±0.41	10.53±0.04	51.13±4.77	7.49±0.73	19.14±0.10	0.84±0.06	9.24±0.93	9.50±0.48	51.42±4.04	0.09±0.01
V-grey-2	2.80±0.03	52.28±9.72	6.97±1.13	4.00±0.60	9.72±1.90	0.19±0.03	1.74±0.38	6.87±1.33	25.72±2.92	0.07±0.01
M-2	0.05±0.04	0.08±0.06	1.56±0.05	0.03±0.00	0.01±0.01	0.00	0.04±0.04	0.00	0.63±0.26	0.00

FESEM was used to analyse the ultrastructural differences of the three types of biofilms (Fig 24). In F-stream, intact filamentous biological structures of different diameters were visible intercut with abundant elemental sulfur particles. Filamentous structures were also visible in F-sed although they appeared generally thinner, organized in an intricate net that entraps coccoid cell-like structures with a partially corroded appearance. This organization was similar to that observed in bacterial mats

collected nearby a sulfidic spring in Capo Palinuro<sup>95</sup>. The observation of damaged microbial-like structures can be associated to the constant exposure of F-sed to rising acidic sulfidic water. The moonmilk deposits are dominated by gypsum microcrystals, whereas biological structures are rare. Vermiculations (both brown and grey) showed an amorphous (possibly extracellular) matrix entwined with interlocking filamentous and net structures. Possible prosthecate bacteria were also visible in V-grey sample, the latter being previously described from oligotrophic cave environments<sup>91</sup>. Such complex arrangement in vermiculations is linked to phenomena of trapping and binding of particles which are dispersed in the surrounding environment or subaerially transported. This capacity of entrapping particles would also explain the extraordinary presence of diopside in vermiculations, a mineral generally absent in carbonate rocks, and possibly brought into Fetida Cave by wave action.



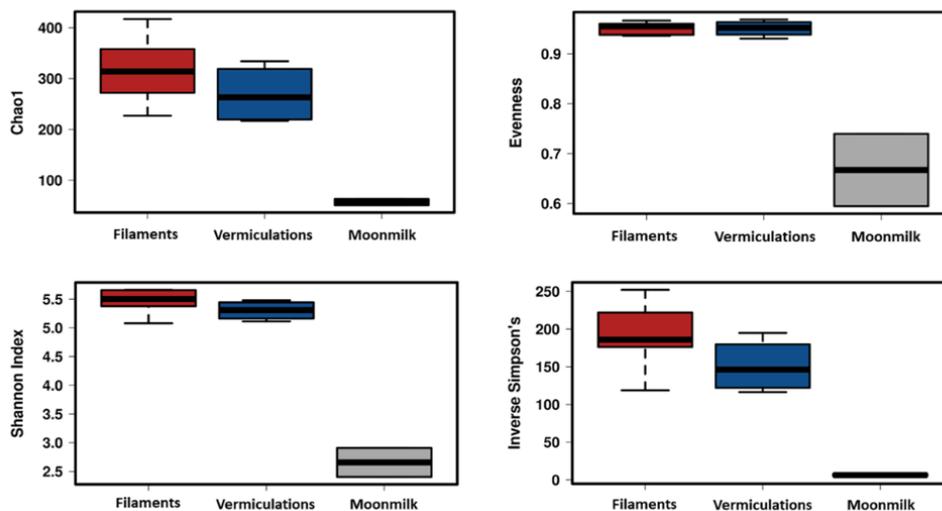
**Fig 26.** FESEM images of deposits collected from Fetida (D'Angeli et al. 2019<sup>20</sup>).

#### 4.3.3. Microbial diversity in Fetida Cave biofilms

As a result of the processing of the demultiplexed fastq files with DADA2 package, 154536 reads were obtained with an average length of 290 bp; they were clustered into a total of 2969 sequence variances (SVs) (Table 19).

Sample name	Type of deposit	# reads	# SV
F-stream-1	White filament	11451	272
F-stream-2	White filament	15970	338
F-float-1	White filament	9689	219
F-float-2	White filament	11426	274
F-sed-1	White filament	15649	347
F-sed-2	White filament	24789	398
V-brown-1	Vermiculation	19926	324
V-brown-2	Vermiculation	11467	218
V-grey-1	Vermiculation	9773	216
V-grey-2	Vermiculation	14090	298
M-1	Moonmilk	6577	57
M-2	Moonmilk	3729	51

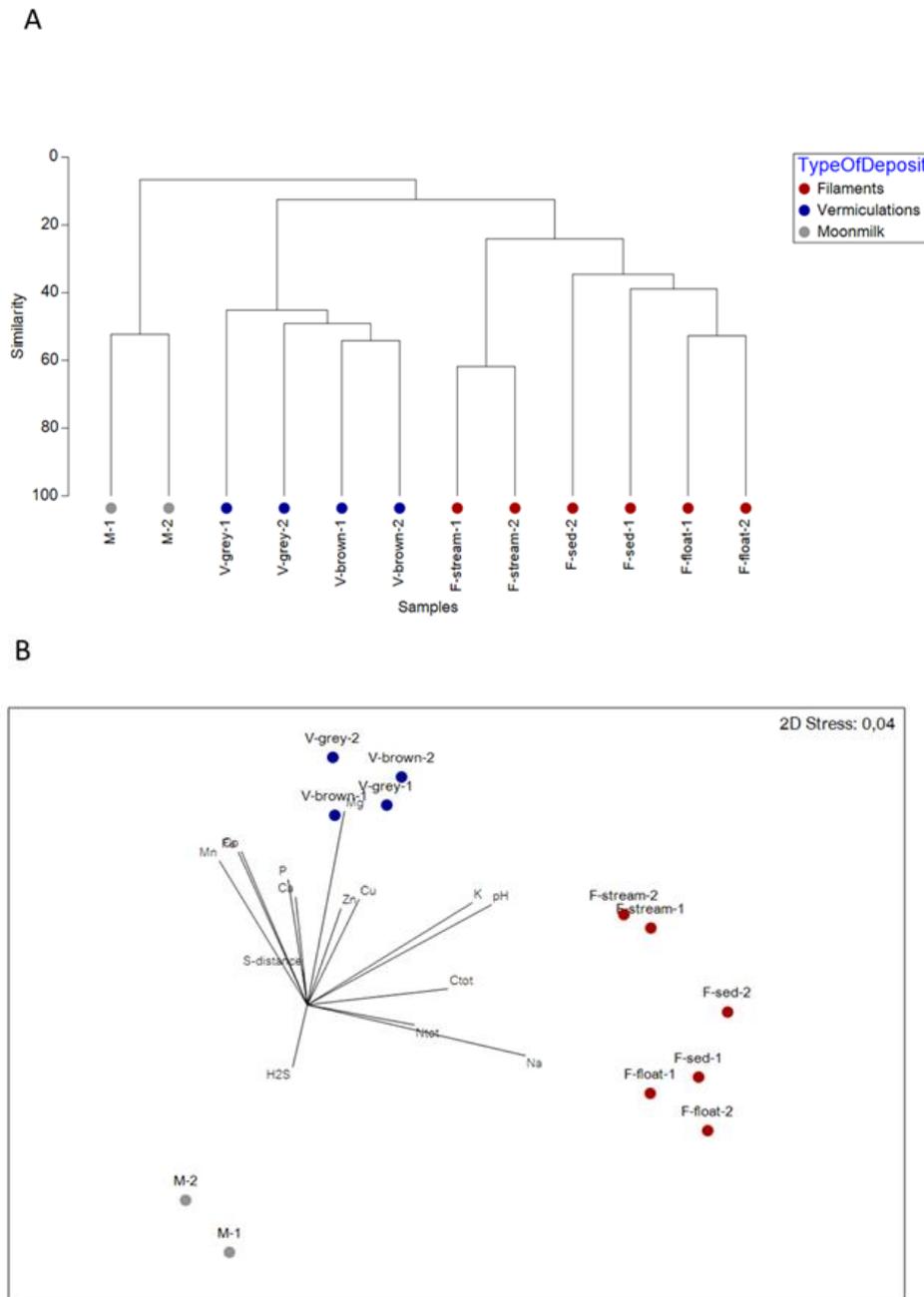
Alpha diversity indexes analysis showed that moonmilk has the lowest richness and diversity as compared to vermiculations and water filaments, with the latter displaying the highest values. In a previous work, the biodiversity associated to vermiculations from a different sulfidic cave (Frasassi cave) was higher than that identified in the white filaments collected from the water stream in the same cave system<sup>91</sup>. This is not true in FC, where the constant mixture of external seawater with the sulfidic fluid most probably leads to the development of complex microbial communities (in the water filaments) characterized by high biodiversity (Fig 27).



**Fig 27.** Alpha diversity indexes of samples grouped in the three different type of deposits (D'Angeli et al. 2019<sup>20</sup>). ANOVA test: Chao1,  $p = 0.002$ ; Evenness,  $p < 0.0001$ , Shannon,  $p < 0.0001$ , Inverse Simpson's:  $p = 0.001$ .

Bray-Curtis similarity and clustering analysis based on the taxonomy revealed that the samples clustered based on the substrate, i.e. waters (white filaments) and wall/ceiling (vermiculations and moonmilk). Within each group, F-stream samples clustered apart from F-sed and F-float samples, and moonmilk grouped apart from the vermiculations, being the latter further sub-clustered depending

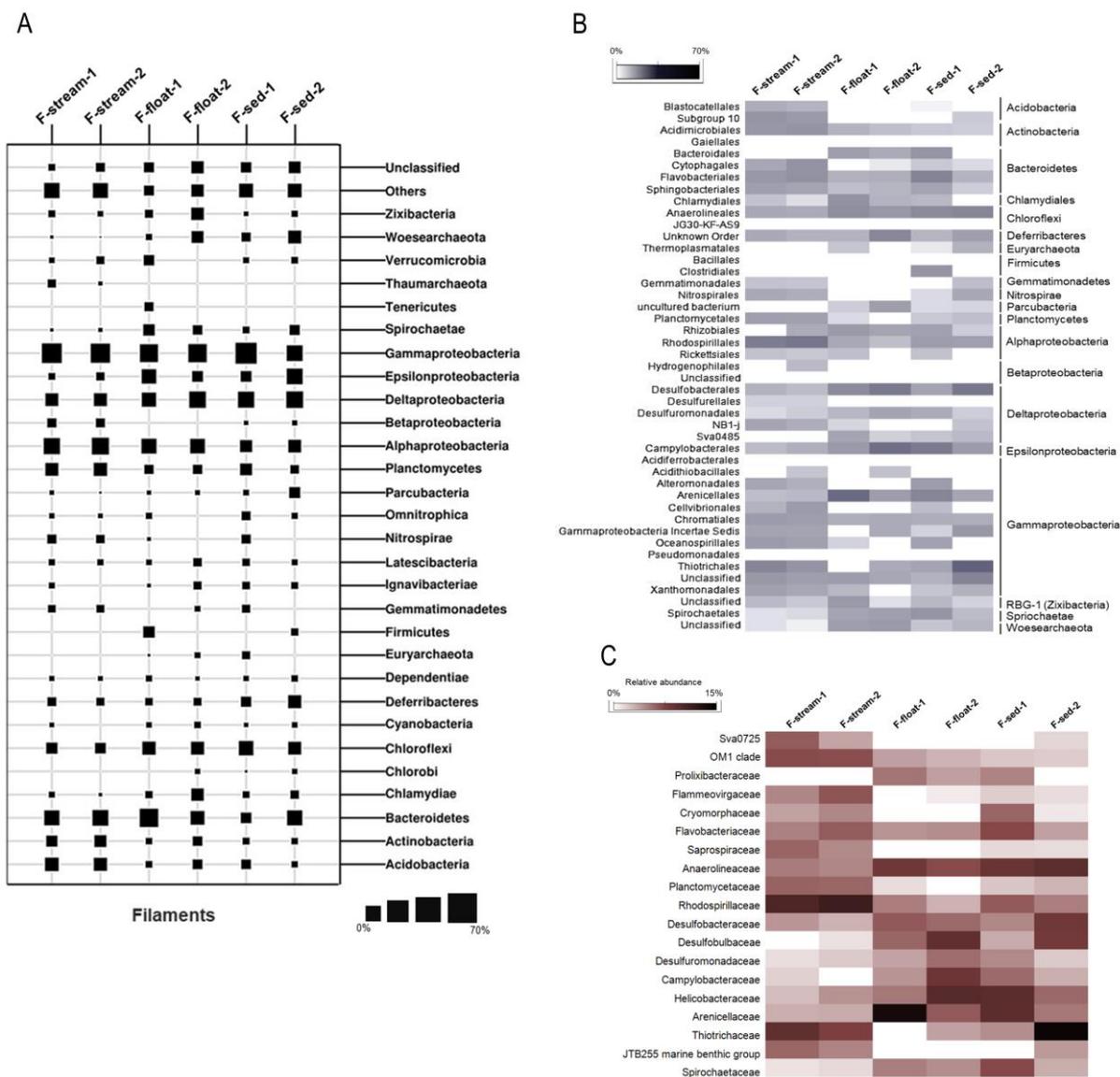
on the color (grey or brown). The non-metric multidimensional scaling (nMDS) plot confirmed that the clustering pattern of Fetida Cave samples is related to the type of deposit, morphology and location, as both F-stream groups were collected at the cave entrance (Fig 28). Pearson correlation of selected chemical elements and physical parameters indicated that the microbial communities of moonmilk were strongly associated with pH values, whereas vermiculations and white filaments were mainly affected by magnesium and by sodium and phosphate respectively. The concentration of Fe, Co, and Mn also affected the microbial diversity, albeit with a lower Pearson correlation ( $\rho > 0.4$ ).



**Fig 28.** Clustering of the different Fetida samples. (A) Bray-Curtis similarity analysis between samples based on SV taxonomy classification in SILVA. (B) Non-metric multidimensional scaling (nMDS) plot showing correlation between physical-chemical parameters and the microbial communities. Modified from D'Angeli et al. (2019)<sup>20</sup>.

### Water filaments

Despite the high variability of the cave water stream due to the fluctuation in geochemistry and hydrodynamics associated to the seawater and possible alternative organic carbon sources, a high presence of members related to sulfur metabolism and belonging to *Gammaproteobacteria*, *Deltaproteobacteria* and *Epsilonproteobacteria* were revealed in FC water biofilms (Fig 29). These microbial sequences were affiliated with those retrieved from deep-marine environments close to gas seeps (hydrothermal deep vents and cold seeps), and at a lower degree, from microbial mats and biofilms described in other sulfidic caves, these last characterized by more stable conditions and the presence of freshwater streams<sup>19</sup> (Fig 30).



**Fig 29.** Taxonomic characterization of microbial communities from white filaments at (A) phylum, (B) order and (C) family levels. Modified from D'Angeli et al. (2019)<sup>20</sup>.

*Gammaproteobacteria* of the *Thiotrichales* and *Arenicellales* orders dominated the FC filaments. The presence of these microbial groups, especially *Thiotrichales*, supports the filamentous and web structures of F-stream observed through microscopy, similar to that previously described for water biofilms collected from other cave environments<sup>89</sup> (Fig 30). *Arenicellales* order has been recently defined and mainly included marine bacteria, some of them isolated from deep-marine environments. While the present study on Fetida Cave is the first study describing members of *Arenicellales* order being associated to a marine sulfidic cave microbiology, filamentous sulfur-oxidizing members of *Thiotrichales* (of *Thiotrichaceae* family) were found to dominate microbial communities of water biofilms collected from the sulfidic Frasassi Cave<sup>91</sup>. Most of the *Arenicellales*- and *Thiotrichales*-related sequences were classified only up to family level (Fig 29). Interestingly, the recently described ‘*Candidatus* Thiopilula’ was dominant in the sedimented biofilms (F-sed-2). ‘*Ca.* Thiopilula’ was previously identified in oxygen minimum zone sediments and cold seeps through metagenomic approaches. The possible contribution of ‘*Ca.* Thiopilula’ in chemolithotrophic processes was supported by transcriptomic results indicating its sulfur oxidation and nitrogenous compounds reduction abilities in microbial mats collected from a deep cold seep<sup>75</sup>. Additional *Gammaproteobacteria* genera retrieved from the FC water filaments belonged to *Oceanospirillales* or *Chromatiales* orders including *Marinobacterium*, *Thiohalophilus*, *Granulosicoccus*, *Halothiobacillus* and *Sedimenticola* (Fig 30). Members of these genera are typically associated with marine and halophilic water habitats, featured by chemolithotrophic activities related to sulfur and sulfidic compounds oxidation, and nitrogen metabolism in anoxia and absence of light<sup>96</sup>. In association with chemolithotrophs, FC filaments host marine oligotrophic *Gammaproteobacteria* *Cellvibrionales* and *Alteromonadales* members, the latter being considered dominant colonizers of marine biofilms able to metabolize various hydrocarbon compounds, which are also possibly involved in nitrogen and sulfur metabolism in shallow-water hydrothermal vent ecosystems<sup>97</sup>.

*Epsilonproteobacteria* have been described to provide the main form of primary productivity in aphotic sulfur-driven microbial ecosystems, including cave sulfidic springs. Unlike filamentous microbial mats described in other sulfidic caves (e.g. Lower Kane Cave), *Epsilonproteobacteria* were not the dominant microbial group in Fetida Cave water biofilms, although their abundance increased by moving from the cave entrance towards the inner zone. The higher presence of *Epsilonproteobacteria* in the inner zone of the cave can be related to their enrichment occurring at low oxygen tension and high H<sub>2</sub>S concentration. We can hypothesize that, even though the general sulfide to oxygen ratio conditions and the seawater-related organic source negatively influence *Epsilonproteobacteria* growth, the environmental conditions in the inner zone of the cave can sustain their increase in the filament microbial communities. Indeed, in the cave inner zone the H<sub>2</sub>S arises

and the slowly flowing water limits oxygen diffusion and hosts deposited mats/filaments. Epsilonproteobacterial sequences in FC were exclusively related to *Helicobacteraceae* and *Campylobacteraceae* families, which were mostly represented by *Sulfurimonas*, *Arcobacter*, *Campylobacter* and *Sulfurovum* genera (Table 20), which are featured by sulfur- and sulfide-oxidizing activities associated to different freshwater and marine environments, including oil fields<sup>61</sup>.

**Table 20. Most abundant genera (and lowest taxonomy affiliation) in water filaments (with abundance >1% in at least one sample). Darkening shades indicate higher abundance i.e. white=0-0.9%, light grey= 1-2.4%, dark grey=2.5-5%, black: >5%.**

Phylum	Class	Order	Family	Genus	str-1	str-2	float-1	float-2	sed-1	sed-2
Acidobacteria	Blastocatellia	Blastocatellales	Blastocatellaceae	Blastocatella	1.39	1.02	0.00	0.00	0.00	0.00
Bacteroidetes	Bacteroidia	Bacteroidales	Prolixibacteraceae	Prolixibacter	0.00	0.00	2.07	1.00	1.66	0.00
Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Hoppeia	0.00	0.00	1.22	0.00	0.00	0.00
Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Maritimimonas	0.00	0.00	0.00	1.37	0.27	0.00
Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Ulvibacter	1.03	0.51	0.00	0.00	0.00	0.00
Chloroflexi	Anaerolineae	Anaerolineales	Anaerolineaceae	Thermomarinilinea	0.00	0.00	0.90	1.08	0.55	1.28
Deferribacteres	D. Incertae Sedis	Unknown Order	Unknown Family	Caldithrix	1.65	1.17	1.09	6.15	0.96	2.62
Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	1.79	1.41	0.00	0.00	0.00	1.22
Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Rhodopirellula	1.39	0.45	0.00	0.00	0.00	0.00
Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Pelagibacterium	0.00	0.00	1.20	0.55	0.57	0.00
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Defluviococcus	0.00	0.29	1.52	0.63	0.41	0.00
Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Pelagibius	2.45	3.28	0.00	0.00	1.14	0.27
Proteobacteria	Betaproteobacteria	Nitrosomonadales	Nitrosomonadaceae	Nitrosomonas	1.83	0.00	0.00	0.00	0.00	0.00
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobacteraceae	Desulfosarcina	0.00	0.00	1.61	0.65	0.33	0.73
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobacteraceae	Sva0081 group	1.20	0.63	0.58	0.00	0.78	0.99
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfobulbus	0.00	0.09	0.72	2.16	0.38	1.27
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	Desulfocapsa	0.00	0.00	1.02	1.37	0.00	0.00
Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobulbaceae	MSBL7	0.00	0.00	0.55	1.65	0.10	2.55
Proteobacteria	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	0.00	0.00	0.00	0.00	1.15	0.00
Proteobacteria	Deltaproteobacteria	Desulfuromonadales	Desulfuromonadaceae	Desulfuromusa	0.12	0.30	0.92	2.32	0.85	0.29
Proteobacteria	Deltaproteobacteria	Syntrophobacterales	Syntrophaceae	Desulfomonile	0.00	0.00	1.10	0.00	0.56	0.00
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Arcobacter	0.24	0.00	0.94	1.36	2.13	0.27
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter	0.00	0.00	0.00	2.16	0.00	0.05
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Sulfurospirillum	0.00	0.00	0.28	1.05	0.27	0.19
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	Sulfurimonas	0.00	0.00	1.09	5.70	4.82	1.18
Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	Sulfurovum	0.46	1.27	0.85	1.25	1.28	1.27
Proteobacteria	Gammaproteobacteria	Alteromonadales	Colwelliaceae	Colwellia	1.03	0.11	0.00	0.00	0.00	0.00
Proteobacteria	Gammaproteobacteria	Alteromonadales	Colwelliaceae	Thalassotalea	0.83	0.51	0.00	0.00	1.35	0.00
Proteobacteria	Gammaproteobacteria	Chromatiales	Granulosicoccaceae	Granulosicoccus	1.31	1.03	0.00	0.00	0.00	0.56
Proteobacteria	Gammaproteobacteria	Chromatiales	Halothiobacillaceae	Halothiobacillus	0.00	0.00	0.56	1.33	0.48	0.00
Proteobacteria	Gammaproteobacteria	G. Incertae Sedis	Unknown Family	Sedimenticola	0.50	0.31	0.00	1.60	0.00	1.41
Proteobacteria	Gammaproteobacteria	G. Incertae Sedis	Unknown Family	Thiohalophilus	1.33	1.43	0.00	0.00	0.00	0.67
Proteobacteria	Gammaproteobacteria	Oceanospirillales	Oceanospirillaceae	Marinobacterium	1.32	0.59	0.00	0.00	0.00	0.00
Proteobacteria	Gammaproteobacteria	Thiotrichales	Thiotrichaceae	CandidatusThiopila	0.00	0.00	0.00	0.00	0.00	12.45
Proteobacteria	Gammaproteobacteria	Thiotrichales	Thiotrichaceae	Cocleimonas	1.85	0.85	0.00	0.00	0.00	0.00
Proteobacteria	P. Incertae Sedis	Unknown Order	Unknown Family	Candidatus Thiobios	1.76	0.82	1.67	3.16	0.24	1.48
Spirochaetae	Spirochaetes	Spirochaetales	Spirochaetaceae	Spirochaeta 2	0.09	0.16	1.59	1.07	2.13	0.71

Besides the increase of sulfur-oxidizing *Epsilonproteobacteria* sequences, filaments inside the FC showed a higher concentration of the *Deltaproteobacteria*, *Chloroflexi* and *Deferribacteres* as compared to F-stream with a parallel decrease of the marine-associated taxa *Actinobacteria*, *Acidobacteria*, *Planctomycetes* and *Alphaproteobacteria*. This can be due to the selection imposed on the microbial diversity by the peculiar geo-physical-chemistry of the water inside the cave in relation to the higher concentration of H<sub>2</sub>S, the slower water flow and the absence of light. Among these, *Deltaproteobacteria* are known to include most of the sulfate reducers detected in sulfuric acid caves<sup>98</sup>. At genus level, members of *Desulfocapsa*, *Desulfobulbus*, *Desulfuromusa* and *Desulfovibrio* were found in Fetida Cave (Table 20), which are all known sulfur-reducers that are able to use the organic carbon released by sulfur-oxidizing bacteria and other primary producers, as the electron

donors<sup>89</sup>. The physical association between sulfur-oxidizing bacteria (of *Gamma*- and *Epsilonproteobacteria*) and *Deltaproteobacteria* has been frequently observed in microbial mats developing in marine and lacustrine sediments but also in other sulfuric acid caves, i.e. Frasassi and Acquisanta Terme, and it has been interpreted as a way to optimize microbial cooperation in the sulfur cycling<sup>89</sup>.

Other syntrophic cooperation involving *Deltaproteobacteria* in FC water filaments, might include members of *Anaerolineales* order and *Caldithrix* genus, which were previously found to be associated with this proteobacterial class in deep-marine sediments, probably contributing to chemoorganotrophic metabolisms under sulfate reducing conditions and metal reducing and oxidizing processes<sup>99</sup>.

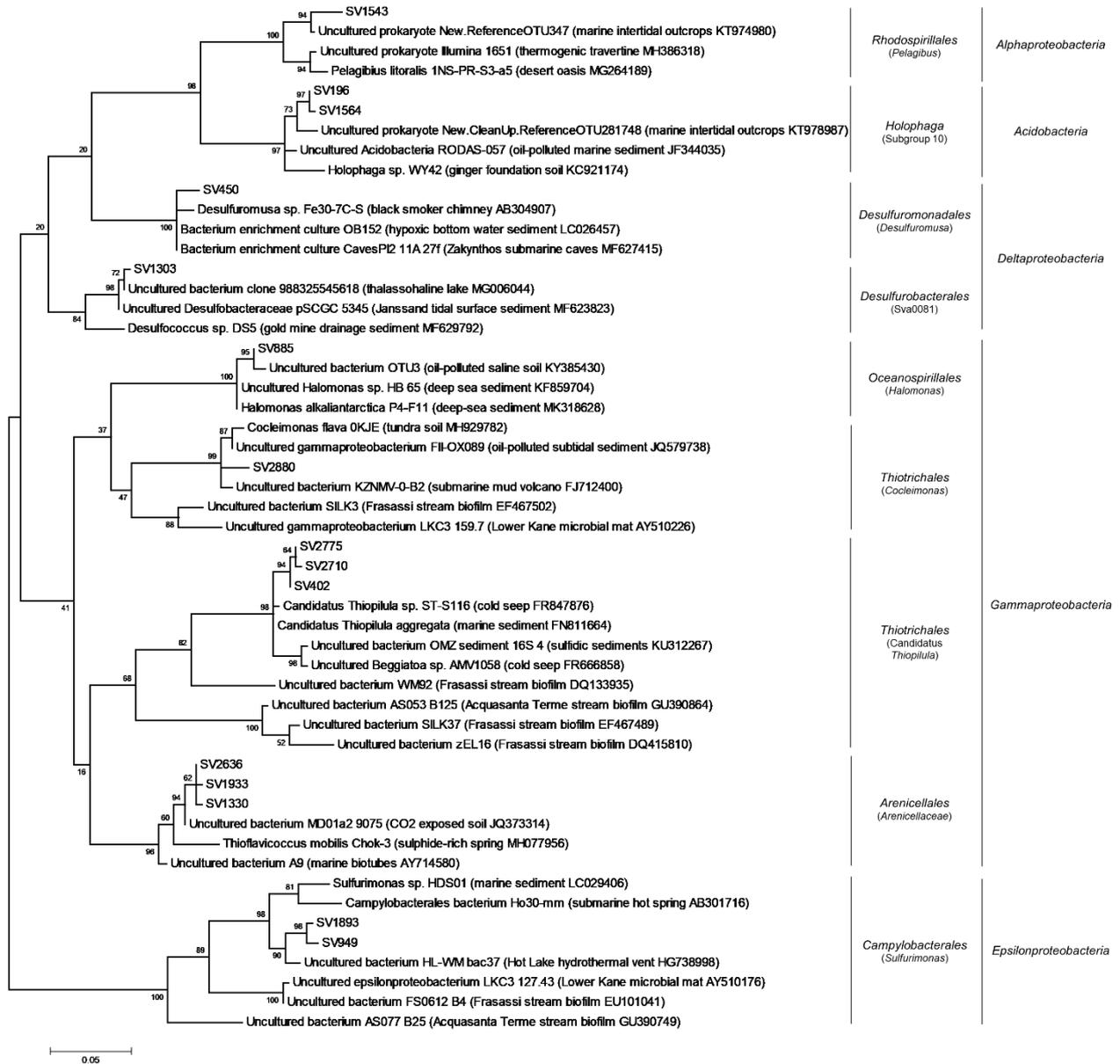
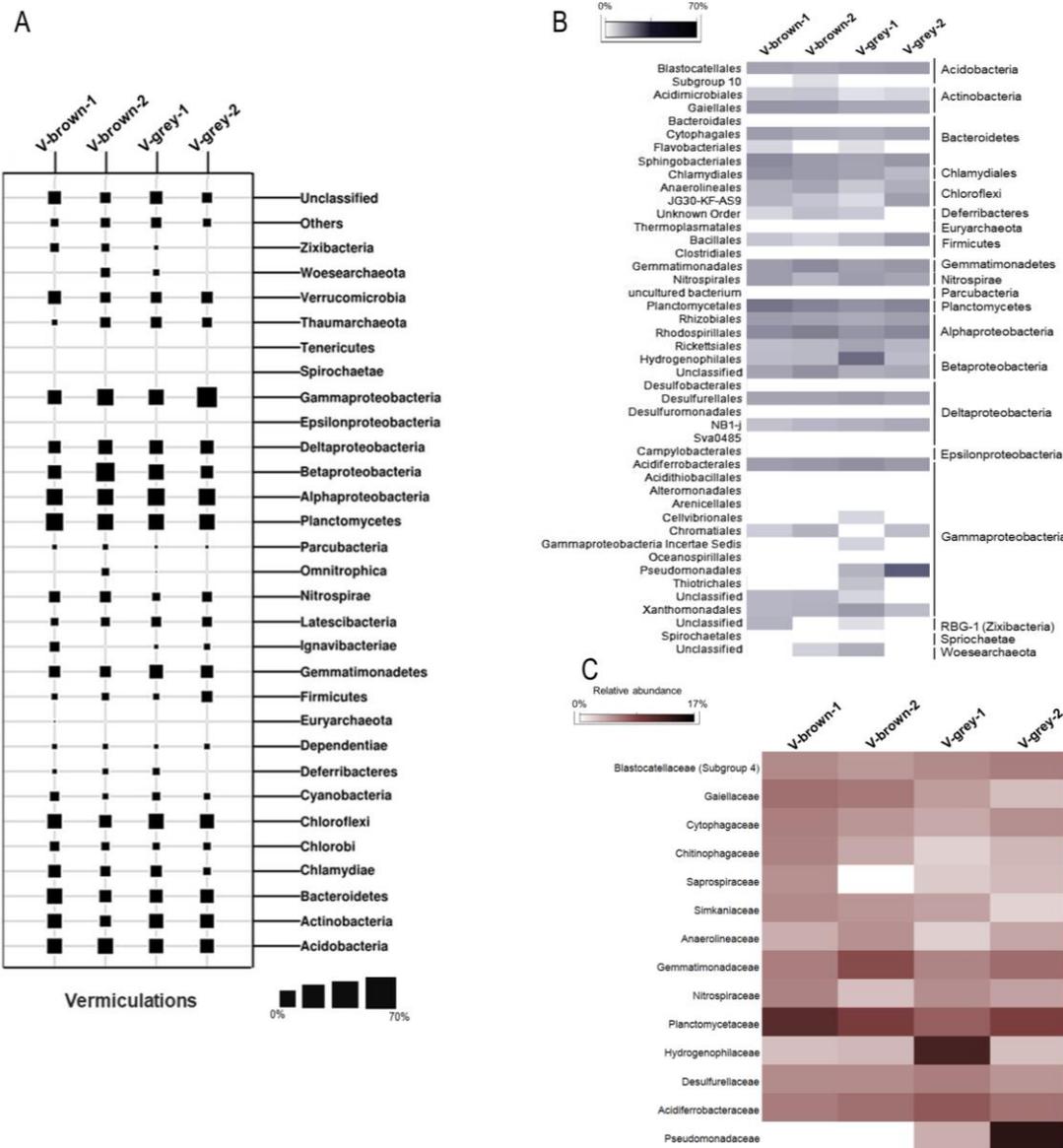


Fig 30. Phylogenetic analysis of the most abundant SVs retrieved from the white filaments (D'Angeli et al. 2019<sup>20</sup>).

### **Vermiculations**

Fetida Cave hosts a peculiar type of vermiculation that has been previously named “biovermiculation”<sup>91</sup>, because of their possible biological origin and the inclusion of highly diversified and active microbial populations. The biovermiculations typically develop in sulfidic caves and present complex and highly diversified geometric forms resembling carbonic-acid caves vermiculations, although they lack significant clay content<sup>91</sup>. Instead, the mineralogy of vermiculations from FC showed the abundance of quartz, in addition to either Mg and Fe-rich minerals, in brown vermiculations, or K- and Al-rich minerals in grey vermiculations (Table 18). The content of nitrogen and organic carbon within FC vermiculations were comparable to those reported by Jones et al. (2016)<sup>100</sup> for some biovermiculations, in which the biological origin of the included organic matter was demonstrated through isotopic analysis. In line with this, we have found that the microbial communities from the same type of deposits from Fetida Cave included chemolithotrophic microbial taxa previously associated with acidophilic and extreme metal-rich environments and mine tailings, wastewater habitats, activated sludges, marine environments and cave settings (Fig 32). In particular, the microbial communities in all the vermiculations under analysis showed microbial populations possibly involved in the nitrogen cycle (Fig 31), i.e. members of *Rhodobacterales*, *Rhodospirillales*, *Nitrospirales* orders and *Planctomycetaceae* family<sup>98</sup>, and in sulfur-reduction and -oxidation under acidophilic conditions and metal-rich environments, i.e. members of *Desulfurellaceae*, *Hydrogenophylaceae*, and *Acidiferrobacteraceae* families<sup>20</sup>. *Gemmatimonadetes* phylum, which is abundant in FC vermiculations, also includes potential sulfate reducing members as revealed by recent genomic analyses<sup>91</sup>.



**Fig 31.** Taxonomic characterization of microbial communities from vermiculations at (A) phylum, (B) order and (C) family levels. Modified from D'Angeli et al (2019)<sup>20</sup>.

In particular, chemosynthetic processes can be associated to *Nitrospira* of *Nitrospirales* and to *Sulfurifustis* of *Acidiferrobacterales* (Table 21), which are abundant genera in all FC vermiculation samples and are able to perform carbon fixation in association with ammonia- or sulfur-oxidation, respectively. *Nitrospira* members are able to catalyze the complete oxidation of ammonia via nitrite to nitrate and, unlike canonical ammonia-oxidizers, can grow under microaerophilic conditions, providing competitive advantage in nutrient-limited conditions and under biofilm growth, similar to vermiculation conditions<sup>101</sup>. Members of *Acidiferrobacterales* are able to gain energy from iron oxidation and to use not only O<sub>2</sub> but also Fe<sup>3+</sup>, NO<sub>3</sub><sup>-</sup> as electron acceptor, this being in line with the high concentration of Fe ions and hematite minerals detected in some biovermiculations, mainly in

the brownish (V-brown) samples (Table 18). Among these, *Sulfurifustis* strains were predicted to have a certain level of metabolic flexibility, due to the redundancy of genes involved in sulfur oxidation and inorganic carbon fixation (Table 21). *Sulfurifustis*-related sequences from Fetida Cave are phylogenetically related with sequences retrieved from Frasassi Cave vermiculations, classified as *Acidithiobacillus*, and water streamers, classified as *Sulfurovum*-like, which were associated to possible biomineralization processes<sup>91</sup>.

Lastly, in Fetida Cave vermiculations, chemoorganotrophs adapted to oligotrophic or contaminated environments were also detected such as members of *Gaiellaceae*, *Blastocatellaceae* and *Anaerolinaceae* families, in addition to copiotrophic bacteria (able to metabolize a wide array of carbon sources) belonging to *Sphingobacteriales*, *Cytophagales* and *Chlamydiales* orders, which were found in diverse terrestrial, aquatic and also underground habitats (Fig 32)<sup>102</sup>.

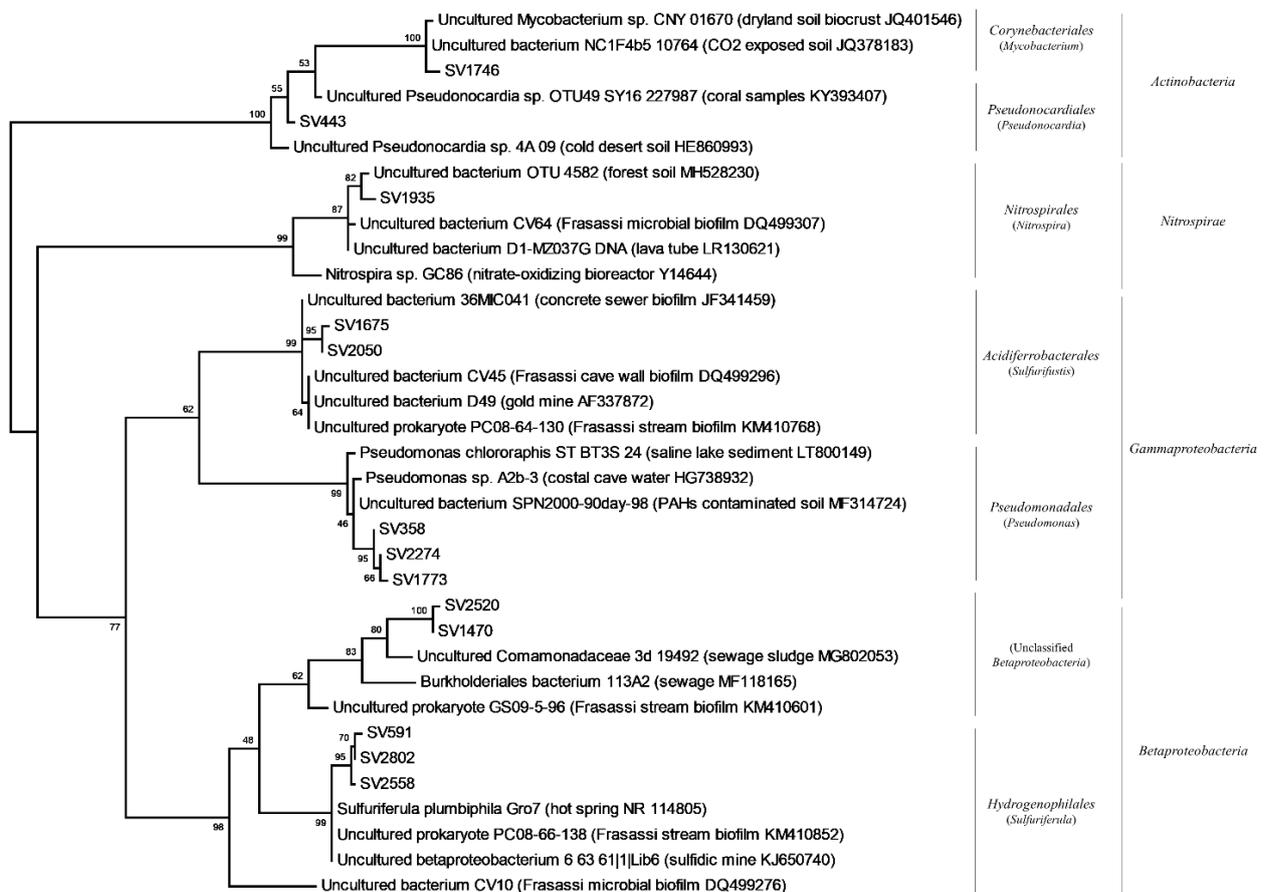
Members of the main chemolithotrophic and chemoorganotrophic taxa composing Fetida Cave (FC) vermiculations, i.e. *Betaproteobacteria*, *Gammaproteobacteria*, *Acidobacteria*, *Planctomyces* and *Nitrospirae*, were previously identified in biovermiculations from Frasassi Cave and Cueva de Villa Luz, indicating that this type of biofilm contains a core set of bacterial phyla which could have synergistic activities.

On the other hand, at lower taxonomy levels, specific bacterial groups distinguished vermiculation samples, even collected from the same cave (Table 21). In FC, members of sulfur-oxidizing autotroph *Sulfuriferula* genus (of *Hydrogenophylales* order), abundant in microbial consortia responsible for the weathering of sulfide minerals occurring under acidic conditions<sup>99</sup> and members of the highly adaptable chemoorganotroph *Pseudomonas* genus were present in the two V-grey samples, the first being predominant in V-grey-1 and in traces in V-grey-2, and the second being predominant in V-grey-2 and at 1% in V-grey-1. While *Hydrogenophilaceae* were present also in the brown vermiculations (although at <1%), *Pseudomonadaceae* were totally absent. The biovermiculation variability in terms of morphology, organic matter content and microbial composition might reflect various conditions of moisture, condensation exposure and geochemistry of the host rock featuring the different cave wall microniches that need to be further explored.

**Table 21. Most abundant genera (at lowest taxonomy affiliation) in vermiculations. Darkening shades indicate higher abundance i.e. white=0-0.9%, light grey= 1-2.4%, dark grey=2.5-5%, black: >5%.**

Phylum	Class	Order	Family	Genus	brown-1	brown-2	grey-1	grey-2
Acidobacteria	Blastocatellia	Blastocatellales	Blastocatellaceae	Blastocatella	0.00	0.00	0.00	1.17
Acidobacteria	Holophagae	Holophagales	Holophagaceae	Geothrix	1.46	0.00	0.00	0.00
Acidobacteria	Solibacteres	Solibacterales	Solibacteraceae	Elev-16S-1166	1.01	0.47	0.11	0.58
Actinobacteria	Actinobacteria	Corynebacteriales	Mycobacteriaceae	Mycobacterium	1.52	0.75	0.20	0.90
Actinobacteria	Thermoleophila	Gaiellales	Gaiellaceae	Gaiella	3.43	3.00	1.62	0.75
Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	0.00	0.00	0.00	1.53
Gemmatimonadetes	Gemmatimonadetes	Gemmatimonadales	Gemmatimonadaceae	Gemmatimonas	0.00	1.14	0.00	1.19
Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	2.62	0.67	2.16	1.48
Planctomycetes	Planctomycetacia	Planctomycetales	Planctomycetaceae	Gemmata	1.51	0.89	0.76	0.38
Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Brevundimonas	1.01	0.00	0.00	0.65
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	0.00	0.00	1.18	0.20
Proteobacteria	Betaproteobacteria	Hydrogenophilales	Hydrogenophilaceae	Sulfuriferula	0.00	0.00	12.08	0.30
Proteobacteria	Deltaproteobacteria	Desulfurellales	Desulfurellaceae	H16	1.77	2.06	2.57	1.87
Proteobacteria	Gammaproteobacteria	Acidiferrobacterales	Acidiferrobacteraceae	Sulfurifustis	2.87	3.47	4.60	2.92
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	0.00	0.00	1.10	17.24
Verrucomicrobia	Opitutae	Opitutaes	Opitutaceae	Opitutus	1.21	0.33	0.00	0.44

Despite lower than white filaments, FC vermiculations are featured by high biodiversity (having 18 phyla accounting for >1%), in consideration of the oligotrophic habitat provided by the cave wall. The corresponding SEM imaging showed intricate webs and filamentous microbial formation embedded in an irregular extracellular matrix (Fig 26). In this context, processes of *in situ* sediment particles entrapment and organic matter production can create a breeding ground for biovermiculation formation and the development of complex indigenous microbial communities.



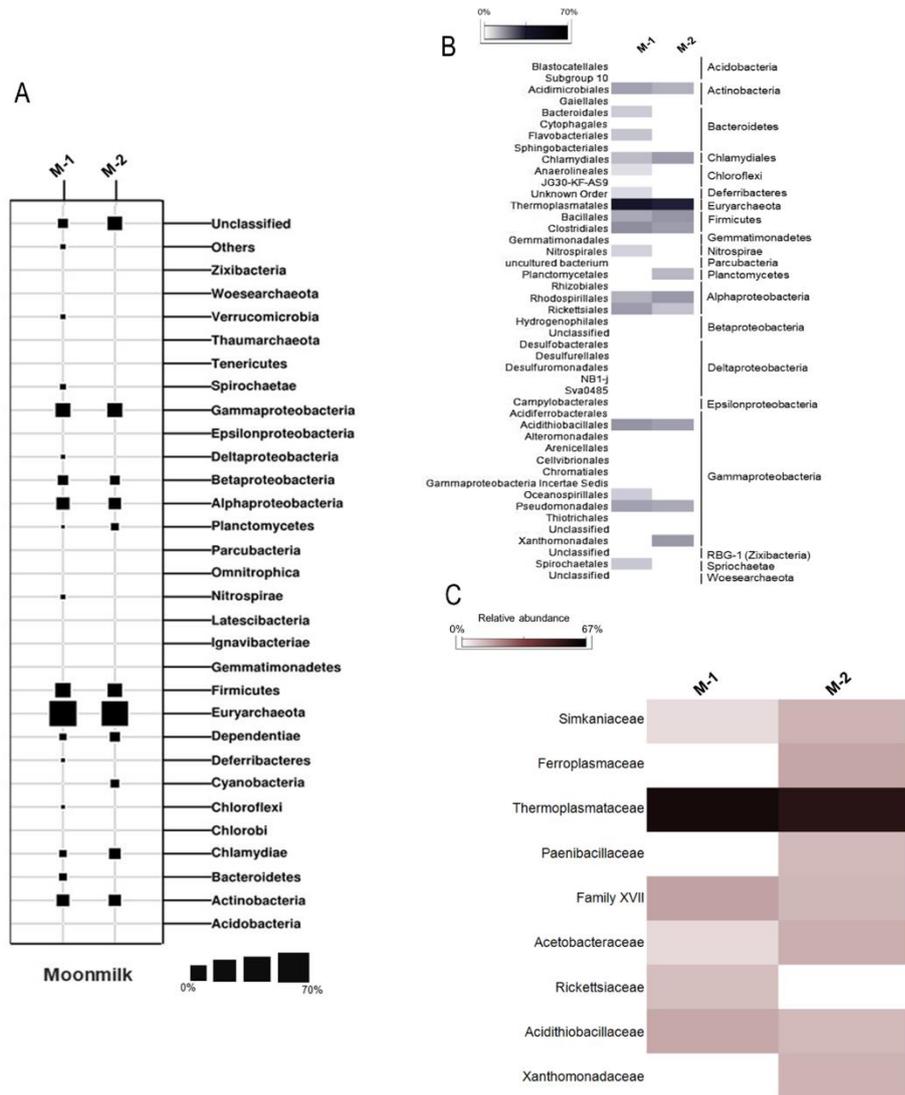
**Fig 32.** Phylogenetic analysis of the most abundant SVs retrieved from vermiculations (D'Angeli et al. 2019<sup>20</sup>).

***Gypsum moonmilk deposits***

Moonmilk is a generic term for a soft, wet, pasty texture material with white, grey or yellowish coloration, it generally consists of microcrystalline aggregates of carbonate precipitates with high water content and is present on the walls of many caves under diverse climatic conditions<sup>103</sup>. Fetida Cave (FC) hosts a peculiar type of moonmilk deposit made of gypsum, having bright white coloration, and, unlike the alkaline calcite moonmilk, with an extremely acidic pH, close to 0-1 (Table 17). This type of deposit has been previously observed in SAS systems, without being microbiologically characterized<sup>104</sup>.

FC's gypsum moonmilk presents low diversity microbial communities, which are strongly dominated by only one archaeal genus i.e. *Thermoplasma*, in some cases associated with *Ferroplasma*, the latter present at lower level (Fig 33). Both *Ferroplasma* and *Thermoplasma* are cell-wall lacking extremely acidophilic archaea with oligotrophic lifestyles and possible capacities to gain energy by sulfur respiration and iron oxidation. While only few isolates have been characterized, sequences related to these genera have been frequently retrieved from acidophilic and metal-rich environments<sup>105</sup>. Few other bacterial taxa, mostly facultative chemolithotrophic, were present in gypsum moonmilk microbial communities i.e. *Acidithiobacillus*, *Metallibacterium*, *Acidibacillus* and *Sulfobacillus*, which are known to be adapted to extremely acidic pH and/or metal-rich environments, some of them being possibly involved in the sulfur cycle and iron-oxidation processes (Table 22)<sup>59</sup>. The high proportion of oligotrophic archaeal populations in gypsum moonmilk with low total C and N values (Table 17) and lower numbers of prokaryotic primary producers (i.e. *Acidithiobacillus* spp.) is an interesting result that guides future metagenomic studies for the identification of key metabolic functions in the *Thermoplasma* population.

Table 22. Most abundant genera (at lowest taxonomic affiliation) in moonmilk. Darkening shades indicate higher abundance i.e. white=0-0.9%, light grey= 1-2.4%, dark grey=2.5-5%, black: >5%.						
Phylum	Class	Order	Family	Genus	M-1	M-2
Euryarchaeota	Thermoplasmata	Thermoplasmatales	Ferroplasmaceae	Ferroplasma	0.00	2.25
Euryarchaeota	Thermoplasmata	Thermoplasmatales	Thermoplasmataceae	Thermoplasma	66.92	57.09
Actinobacteria	Actinobacteria	Corynebacteriales	Mycobacteriaceae	Mycobacterium	1.90	1.45
Actinobacteria	Actinobacteria	Frankiales	Acidothermaceae	Acidothermus	0.00	1.18
Firmicutes	Bacilli	Bacillales	Alicyclobacillaceae	Acidibacillus	1.25	1.37
Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus	0.00	2.63
Firmicutes	Clostridia	Clostridiales	Family XVII	Sulfobacillus	4.03	1.69
Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Delftia	0.93	1.31
Proteobacteria	Gammaproteobacteria	Acidithiobacillales	Acidithiobacillaceae	Acidithiobacillus	4.20	2.57
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter	1.43	0.00
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	1.00	1.66
Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Metallibacterium	0.00	3.03



**Fig 33.** Taxonomic analysis of microbial communities from moonmilk at (A) phylum, (B) order and (C) family levels. Modified from D’Angeli et. al (2019)<sup>20</sup>.

The microbial community composition of gypsum moonmilk from FC strongly differs from that described in calcite moonmilk deposits, whose biodiversity is mainly characterized by members of aerobic chemoorganotrophic and facultative chemolithotrophs belonging to *Alpha*-, *Beta*- and *Gammaproteobacteria* and *Actinobacteria* involved in nitrogen and hydrogen oxidation and with optimal growth at circumneutral pH<sup>106</sup>. On the other hand, the microbial composition of Fetida’s moonmilk showed high similarities with the biodiversity described in acidic pendulous biofilms, named snottites collected from other sulfuric acid caves, i.e. Frasassi and Lower Kane caves. In particular, not only the dominant *Thermoplasma*-related sequences but also most of the abundant SVs within the FC gypsum moonmilk showed highly similarity with those retrieved from Frasassi Cave snottite (Fig 34). This is probably due to the extremely acidic pH and the similar sulfuric acid cave setting and mineralogy shared by the two types of biofilms. On the other hand, gypsum moonmilk

and snottites are featured by relevant differences in morphology and arrangement. In fact, gypsum moonmilk is a soft creamy white deposit composed of gypsum microcrystals, which develops on the overhanging exposed walls and ceiling of Fetida Cave from 1 m above the water table, whereas snottites are pendulous structures, developed on gypsum substrate, thriving from 0.5 to 4 m above the water stream. The different morphology and arrangement of these two deposits might be associated with the stability of geochemical and physical-chemical parameters and environmental conditions featuring the hosting caves (e.g. degassing H<sub>2</sub>S and O<sub>2</sub> content as a function of hydrodynamic conditions, condensation phenomena, possible meteoric water infiltration). These morphological differences might support the difference found in the dominance of specific bacterial taxa, which are affiliated to archaeal oligotrophic *Thermoplasma* genus in the Fetida's moonmilk, whereas they are affiliated to gammaproteobacterial chemolithoautotrophic *Acidithiobacillus* in Frasassi's snottite<sup>59</sup>. Probably, the extreme acidophilic community, which colonizes the moonmilk deposits of Fetida Cave, contributes to the precipitation of gypsum crystals.

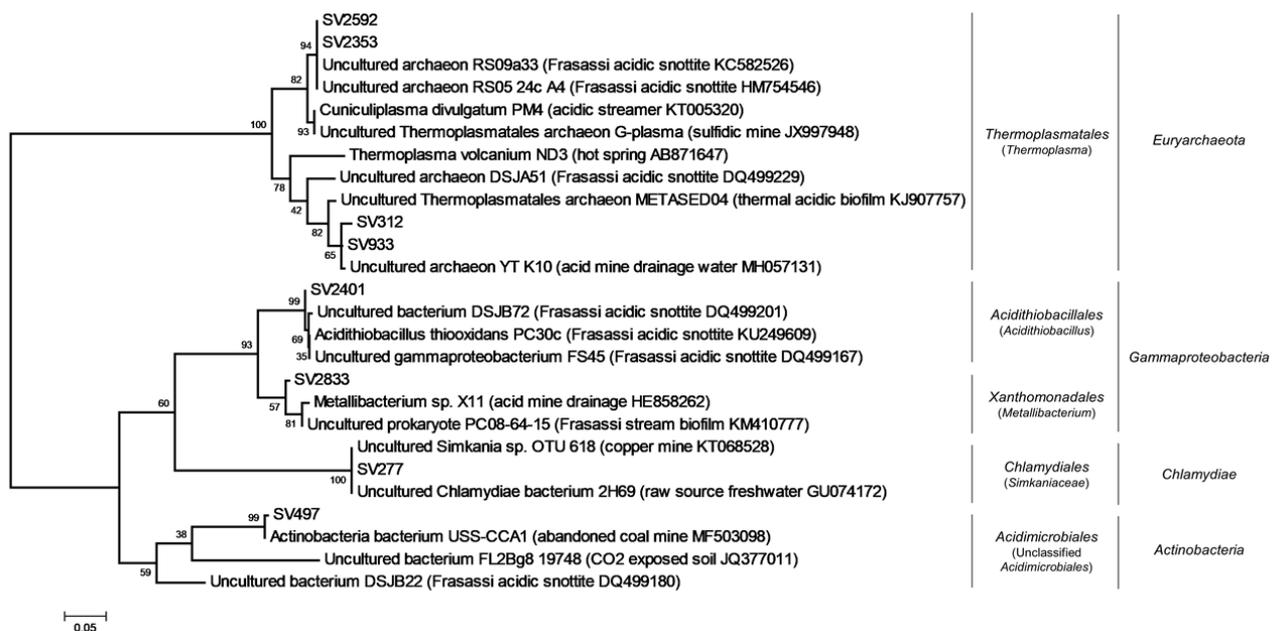


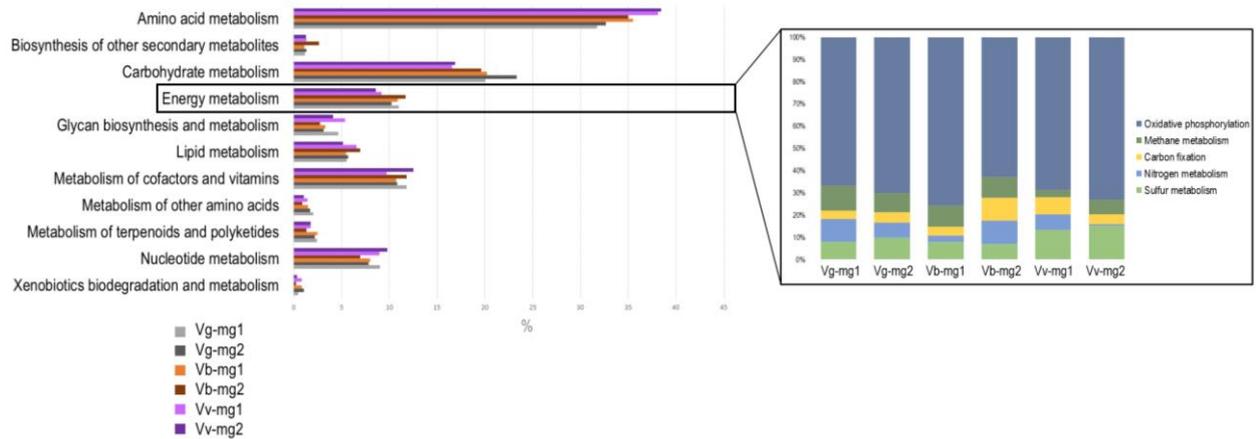
Fig 34. Phylogenetic analysis of the most abundant SVs retrieved from moonmilk samples (D'Angeli et al. 2019<sup>20</sup>).

#### 4.3.4. Metagenomic insights into vermiculations

**Table 23. Summary of metagenomes analysis from six vermiculations.**

Sample	Reads	Contigs	N50 contig size	Predicted genes	Distribution (%) of sequences in major KEGG categories					
					Metabolism	Genetic information processing	Environmental information processing	Cellular processes	Human diseases	Organismal systems
Vg-mg1	1274906	5296	747	5377	58.7	22.4	13.4	3.2	1.7	0.6
Vg-mg2	439114	8410	769	7717	56.6	19.3	16.6	5.6	1.2	0.7
Vb-mg1	464218	5515	710	5496	54.1	23.4	17.6	2.8	1.6	0.5
Vb-mg2	420618	3131	659	2851	59.7	18.1	16.8	3.3	1.5	0.6
Vv-mg1	405482	4347	812	4310	59.5	22.4	13.5	2.7	1.5	0.4
Vv-mg2	457282	4654	955	5249	59.6	23.8	12.0	2.9	1.3	0.5

Metagenomic analyses of the microbial community inhabiting six different vermiculations featured by three different colours collected from Fetida walls, which include grey (samples Vg-mg1 and Vg-mg2), brown (samples Vb-mg1 and Vb-mg2) and violet (samples Vv-mg1 and Vv-mg2) vermiculations. The only vermiculations were analysed during a first metagenomic study because of limited knowledge of the microbial functions in this type of deposits (while water filaments have more extensively described). Metabolic potential analysis was performed by mapping contigs to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Preliminary results showed that the most represented genes belonged to the category Metabolism, accounting for 54-59% of the dataset, followed by genes involved in Genetic information processing (18-24%), Environmental information processes (12-18%), Cellular processes (3-6%), Human diseases (1-2%) and Organismal systems (<1%) (Table 23). In particular, within the category of Metabolism, Energy metabolism included oxidative phosphorylation, methane metabolism, nitrogen metabolism and sulfur metabolism (Fig 35). Considering the nature of the cave and the high number of microbes putatively involved in sulfur and nitrogen cycles in vermiculations, we focused in an initial metagenomic study only on the investigation of genes involved in sulfur and nitrogen compounds transformations.

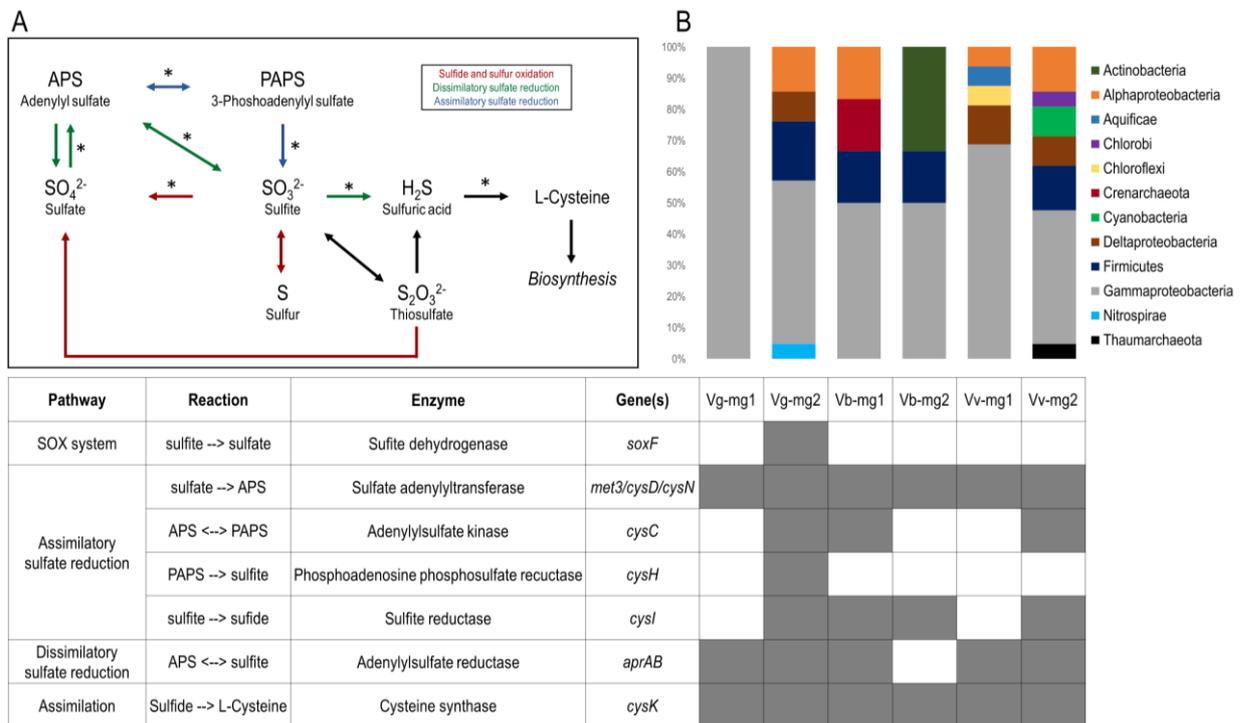


**Fig 35.** KEGG categories belonging to metabolism (left) and energy metabolism (right) retrieved from the six metagenomes.

The metagenomes of the six vermiculations revealed the dominance of genes involved in sulfate and sulphite transformation processes affiliated to *Gammaproteobacteria*, *Deltaproteobacteria* and *Firmicutes* (Fig 36). Especially Vg-mg2 (grey) vermiculation metagenome showed all the genes involved in the assimilatory and dissimilatory sulfur reduction pathway, while the other samples presented parts of them. In particular, in all vermiculations, metagenomic functions involved in sulfur reduction processes were encoded by the genes *met3*, *aprBA*, *cysCDHIN* genes. Among these, *met3* encodes the ATP sulfurylase and *cysD* and *cysN* encode the two subunits of sulfate adenylyltransferase; both these enzymes are involved in the activation of sulfate through the conjugation to AMP to form adenosine-5'-phosphosulfate sulfate (APS) in the sulfur assimilation pathway. Following this, the sequential steps catalyzed by CysC, CysH and CysI convert APS into H<sub>2</sub>S that can be utilized as sulfur source for amino acid synthesis<sup>107</sup>. The *aprAB* genes encode the APS reductase in the sulfur dissimilation pathway that catalyzes the conversion of APS to sulphite, which in turn can be converted into H<sub>2</sub>S for anabolic pathways. The genes detected in the vermiculation metagenomes and involved in both sulfur reduction pathways were mainly affiliated with the gammaproteobacterial *Thiobacillus*, *Thioalkavibrio*, *Xanthomonas* and the deltaproteobacterial *Desulfatibacillum*. This suggests a primary role of these bacterial genera in the sulfur metabolism in the vermiculation biofilm. Additionally, a high abundance of genes belonging to *Clostridia* class was observed in sample Vg-mg2. Most of these genes are involved in specific reactions within sulfur reduction (*cysCDIN*) and in the transport of nutrients such as sulfonate, nitrate, taurine and glutamine (*gltK*, *tauA*, *ssuA* genes). The ability of *Clostridium* strains to reduce sulfur was previously described by Sallam (2009)<sup>108</sup>. In all the samples under analysis, we detected low abundance of genes involved in sulfur oxidation, probably because of the low sequencing depth and/or the low abundance of sulfur-oxidizing bacteria, although *Hydrogenophylales*, which is highly present in some vermiculations, were previously described as sulfur oxidizer<sup>89</sup>. Only the metagenome belonging to Vg-mg2 (grey) vermiculation

possessed *soxF* gene affiliated with *Cupriavidus* of *Betaproteobacteriales*. SoxF constitutes the SOX system that catalyzes oxidation of thiosulfate, elemental sulfur (S<sup>0</sup>), sulfide, and sulfite to sulfate in sulfur-oxidizing bacteria<sup>100</sup>.

Together these results suggest that most of the sulfur available for the microbial communities inhabiting the vermiculations is present in oxidized forms (sulfate). We hypothesise that microbial communities in the vermiculations gain energy from the reduction of sulfate, which is released by the activity of sulfur-oxidizing bacteria (on the sulphide arising from the spring) present in the white filaments. Deeper metagenomic analyses of white filaments floating and sedimented in the cave waters will clarify this hypothesis.



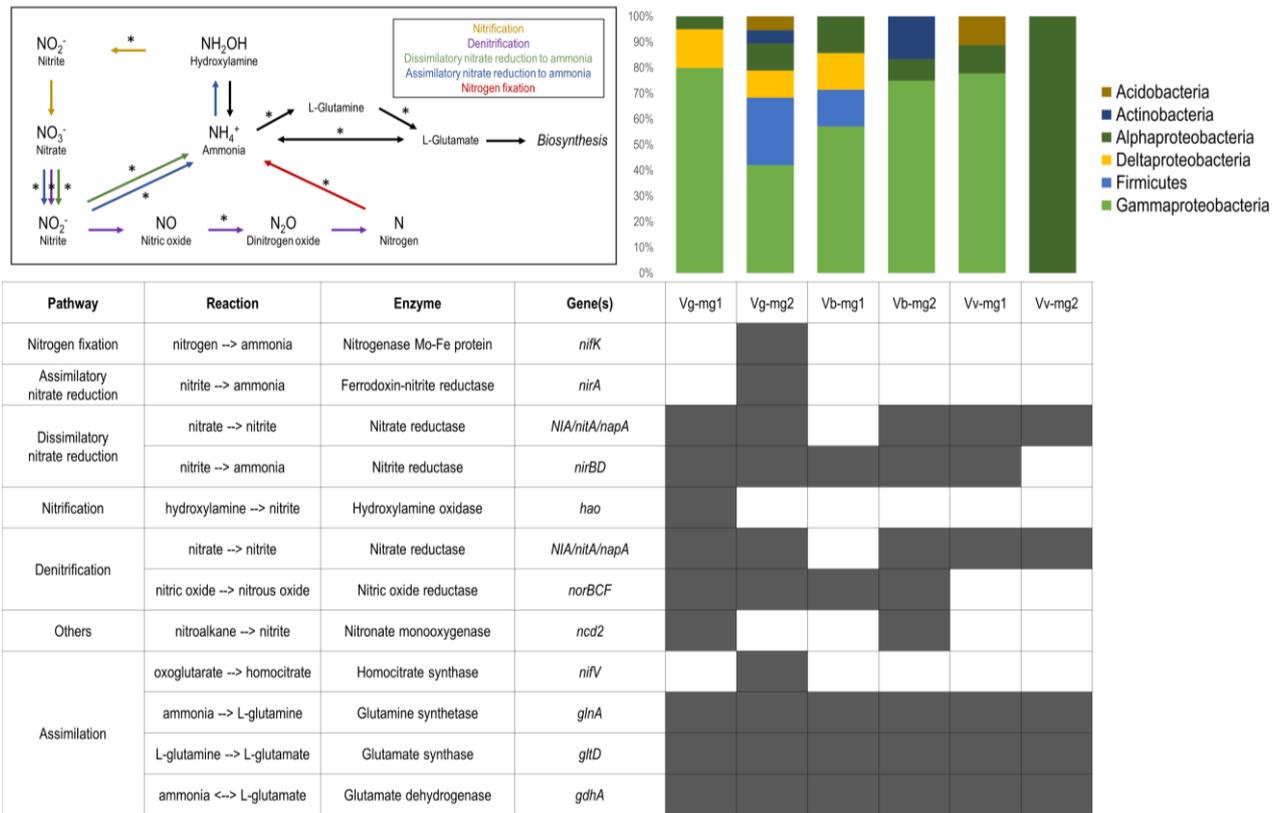
**Fig 36.** Metagenomic analysis of the sulfur metabolism (KEGG). (A) Graphical representation of the pathways involved in the sulfur metabolism. Asterisks indicate pathways with genes present in the metagenomes. (B) Taxonomic affiliation of the genes involved in the sulfur metabolism by using RefSeq database. The table summarizes the genes and enzymes involved in the sulfur cycle present in the six vermiculations.

Nitrogen metabolism is usually highly present in sulfuric caves<sup>109</sup>. More specifically, the transformation of nitrogen has been described as a key metabolism in cave vermiculations<sup>91</sup>. Metagenomic analyses showed that genes involved in several pathways of nitrogen metabolism were present in Fetida vermiculation microbial communities and these genes mainly belonged to *Gammaproteobacteria*, *Betaproteobacteria* and, at a lower extent, to *Alphaproteobacteria*, *Deltaproteobacteria*, *Firmicutes* and *Actinobacteria* (Fig 37). All the vermiculations under analysis possessed genes involved in denitrification processes and in the dissimilatory nitrate reduction to ammonium (DNRA) pathway. These genes included *napA*, *nirBD* and *norBCF*, which codify

enzymatic subunits involved in nitrate, nitrite and nitric oxide reductases respectively (Fig 37). Taxonomic classification of these genes revealed their affiliation to genera *Burkholderia*, *Cupriavidus* (*Betaproteobacterales*, formerly *Betaproteobacteria*), *Methylococcus*, *Pseudomonas* (*Gammaproteobacteria*) and *Chelativorans* (*Alphaproteobacteria*). On the other hand, enzymes involved in assimilatory nitrate reduction to ammonia (ANRA), nitrification and nitrogen fixation pathways were identified in the grey vermiculations only. In detail, *hao* genes, which are involved in the transformation of hydroxylamine to nitrate (nitrification), were detected in Vg-mg1 and were affiliated to *Desulfohalobium* genus of *Deltaproteobacteria*, whereas *nirA* and *nifK* genes, responsible of the ANRA and of the nitrogen fixation to ammonia respectively, were found in Vg-mg2 and belonged to *Clostridium* genus of *Firmicutes*. NirA and NifK activity in *Firmicutes* was previously documented in soil samples<sup>88</sup>, while the activity of these enzymes has never been associated to *Firmicutes* in biodeposits in sulfidic caves. Conversely, they were associated to members of *Gammaproteobacteria* and *Nitrospirae* which were pointed out as responsible of nitrate reduction and nitrogen fixation.

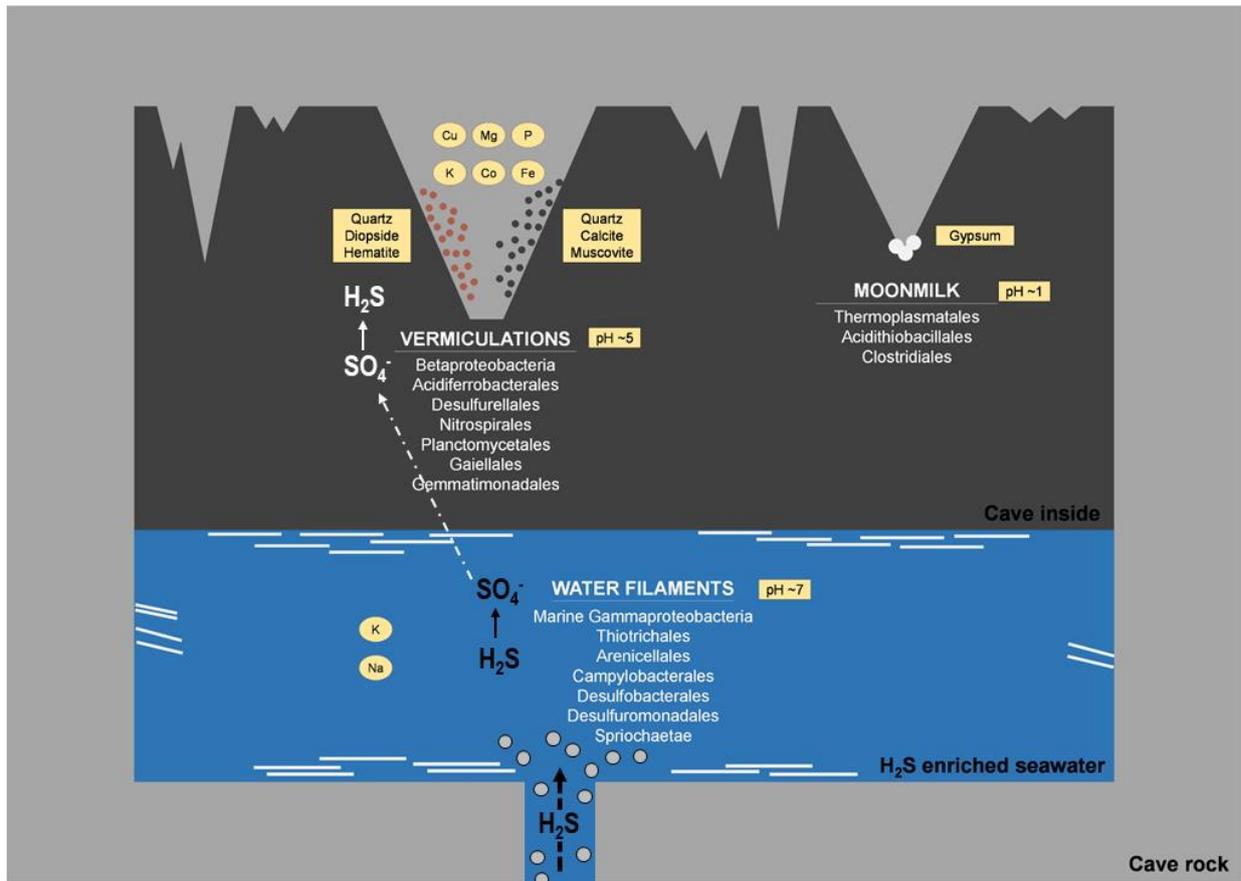
Unlike denitrification, which produces unreactive dinitrogen gas, DNRA and ANRA act to conserve bioavailable nitrogen in the environment producing soluble ammonium, which in turn can be used in further anabolic reactions by microorganisms for the production of complex molecules. Our metagenomic data indicated the presence of genes involved in denitrification and DNRA in all the vermiculations under analysis. However, denitrification contributes to the loss of nitrogen while DNRA is a major remineralization pathway, therefore the balance of these two processes is important to the nitrogen cycle of vermiculations in Fetida cave.

Although the microbial community analyses based on 16S rRNA of vermiculation suggested a role of *Plantomycetes* and *Nitrospirae* phyla in the nitrogen cycle, the metagenomic analysis did not show genes belonging to members of these phyla. This might be due to the low sequencing depth of the metagenomic data and to the low abundance of members of these two phyla, despite being present in all the vermiculations. Future investigations will be performed by co-assembling the metagenomes of all the vermiculations together and by re-sequencing selected samples with higher coverage.



**Fig 37.** Metagenomic analysis of the nitrogen metabolism (KEGG). (A) Graphical representation of the pathways involved in the nitrogen metabolism. Asterisks indicate pathways with genes present in the metagenomes. (B) Taxonomic affiliation of the genes involved in the nitrogen metabolism by using RefSeq database. The table summarizes the genes and enzymes involved in the nitrogen cycle present in the six vermiculations.

Figure 38 summarizes the microbial ecology of Fetida cave, showing the main microbial groups and the chemical elements characterizing each type of deposits, and the hypothesis of the principal reactions concerning the sulfur cycle of microbes inhabiting the different deposits.



**Fig 38.** Graphical representation of the distribution pattern of water filaments, vermiculations and moonmilk in Fetida cave along with indications on the metabolism of sulphide arising from the cave spring. Modified from D'Angeli et al (2019)<sup>20</sup>.

## **5. Geomicrobiology of the high-altitude ice cave Cenote Abyss**

### **5.1. Introduction**

Microbial life at the cryosphere represents a unique archive of peculiar and slow biogeochemical processes in ancient and chilly environmental conditions. The investigation of the geomicrobiology of ice layers of glacial locations allows the analysis of peculiar ecosystems, to provide clues for the past climate reconstruction and to understand the impact of climate change on ice<sup>110</sup>. The consequences of ice melting are associated to the exposition of the inner depth ice core leading to the development of new top layer of soil<sup>111</sup>. This phenomenon constitutes a crucial modification of the microbial ecology of iced places, as well as an alteration of the biodiversity of the flora and fauna. To date, the highly preserved cave ice accumulations represent understudied but potentially dynamic and robust archives for investigating the impact of climate and anthropogenic pollution on the diversity and viability of the ice entrapped microbiome<sup>110</sup>.

In glacial locations, frozen carbon stores are unavailable to microorganisms. With the warming of the climate and the increase of rainfalls, these stores become always more accessible and lead to additional microbial production of greenhouse gases like carbon dioxide, methane and nitrous oxide<sup>112</sup>. Nitrous oxide is an intermediate in the denitrification process of the nitrogen cycle. Its influence in the environment has been deeply studied in agricultural soils<sup>113</sup>. However, investigating the role of nitrous oxide in polar places and glaciers is fundamental for foreseeing future climate variations<sup>114</sup>.

Recent findings revealed the presence of complex microbial communities, comprising bacteria, archaea, and fungi in peculiar deposits located in ice caves, which include carbonate and volcanic caves located in the Alps and in the Antarctic mountains<sup>18,110,115</sup>. The study of the ice cave microbiome also suggested an involvement of psychrophilic microorganisms in the formation of secondary mineral deposits, such as ice speleothems<sup>116</sup>. However, the chemical reactions of the biogeochemical cycles in ice caves are very slow both for the low nutrient availability and for the very low temperatures characterizing the habitat. Many recent studies have attempted to measure microbial activities in glacial environments and have reported noteworthy rates of nitrogen fixation, nitrification and denitrification processes, confirming that psychrophilic microbes have an important role in the biogeochemical cycles of the cold biosphere<sup>117</sup>.

This chapter describes the microbial communities in the Cenote Abyss<sup>118</sup>, which is an impressive carbonatic ice cave of 285 m deep and hosts one of the most voluminous cave glaciers of the Dolomites. This abyss was first explored in 1994 by the Speleological Club Proteo of Vicenza, after the emptying of a lake at 2940 m a.s.l. in the Regional Park of Fanes, Sennes and Braises. The Cenote

Abyss ice cave is a remote and oligotrophic environment that represents an excellent accessible model system for understanding fundamental microbe-mineral interactions contributing to the subsurface biosphere. Moreover, the huge ice deposits identified within this cave are of great interests for studying the paleoclimate and the climate evolution of the Dolomites during the last thousands of years, with insights into the recent melting of the cave glacier.

## 5.2. Materials and methods

### 5.2.1. Samples list, DNA extraction and sequencing

During three expeditions organised in October 2015, September 2016 and October 2018, microbiological samples were collected from both the outside and the inside of the cave, and a complete survey of the cave was performed using laser scanning equipment<sup>118</sup>. Samples were collected and stored as described in paragraph 2.1 before their processing in the laboratory of the Medical University of Graz (Austria) and Laboratory of Molecular and Applied Microbiology of the University of Bologna (Italy).

Table 24 reports all the samples analysed in this chapter.

Sample	Depth (from the entrance)	Location
C1	30 m	Outermost
C2	40 m	Middle
C3	70 m	Innermost

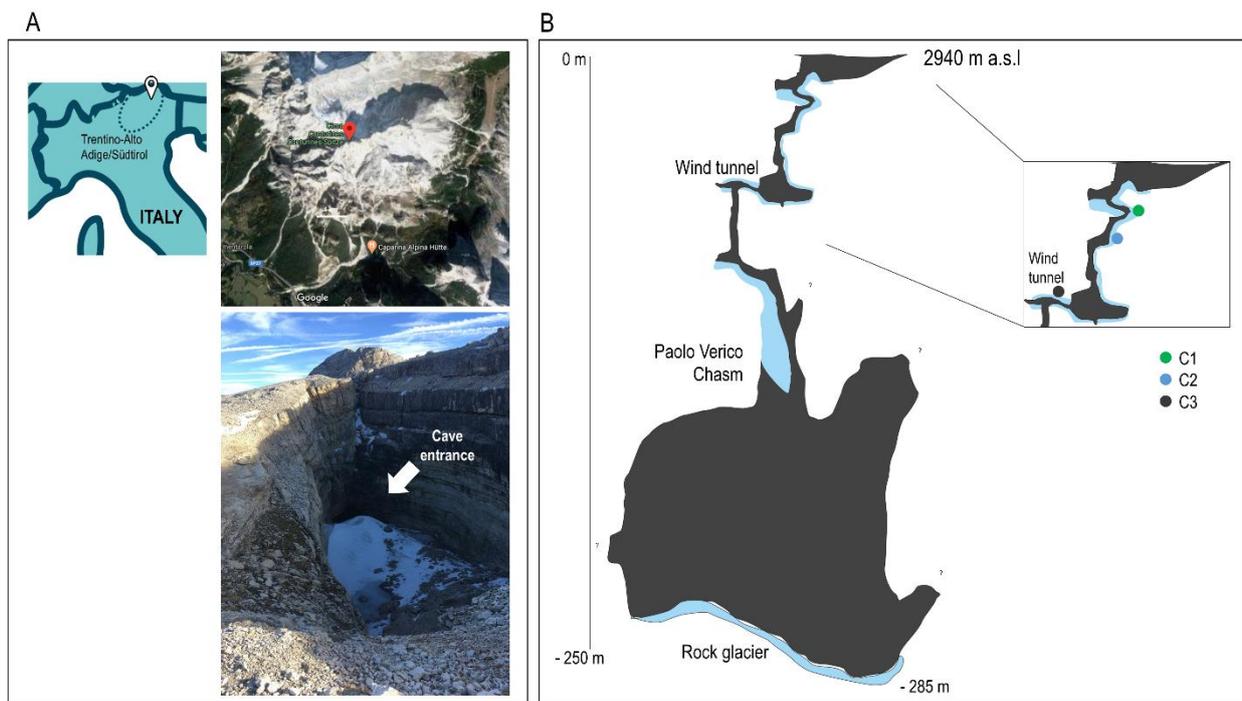
Total DNA extraction, amplification of V4-V5 regions and Illumina sequencing were performed as described in paragraph 2.2. Processing of raw data with DADA2 package and statistical analyses were performed as described in paragraph 2.3.

## 5.3. Results and discussion

### 5.3.1. Samples description

The area of the Cenote cave is characterised by a wide syncline structure comprising sedimentary rocks ranging from the Upper Triassic to the Lower Miocene and mainly composed of a succession of dolomite layers, with a total thickness of 1 km<sup>118</sup>. The cave starts with a karst depression (the original lake “delle due forcelle”), about 20 m deep and 50x30 m wide. The bottom of this doline is occupied by an ice and snow cone, characterised by elongated fractures parallel to the depression’s major axis. Entering the cave, it is possible to descend down a channel carved between the ice deposit and the rock wall. This conduit is characterised by ice walls sculptured with 1 m wide melting niches.

Samples C1 was collected from the ice walls of in this area, at 30 m of depths, and consists in an ice deposit with evidence of microbial colonisation in the form of black mud. From this area, it is possible to descend down a series of shafts developed at both the rock wall-ice contact and completely inside the ice mass, where an ice sample (C2) with brownish material was collected at 40 meters of depth. At a depth of 70 m the cave extends completely within the ice inside a tunnel carved by the airflow (the wind tunnel, Fig 39). Here, the innermost sample C3 was collected and appeared as an ancient ice layer with brownish mats. The cave then descends 30 m via another shaft opening in a long NW-SE-oriented passage that progressively opens into a large underground room. From this point onward the descent is along the ice deposit and the rock wall with the first ending after about 60 m with a huge ice tongue suspended above the final chamber. The free-hanging depth of this deep chasm is 165 m (Baratro Paolo Verico). The chamber at the bottom is 130 m long and about 40 m wide, progressively descending to the lowest point of the cave with a cone composed of a mixture of ice, sands and boulders (at 285 m of depth).



**Fig 39.** A) Geographical localisation of the Cenote Abyss and cave entrance. B) Schematic representation of the cave and sampling points.

**Table 25. Summary of Illumina sequencing data analysis and alpha diversity indexes.**

Sample	Location	# reads	# SV	Shannon	Inverse Simpson's	Evenness
C1	Outermost	3970	102	4.33	65.12	0.94
C2	Middle	3922	99	4.22	49.15	0.92
C3	Innermost	4001	157	4.46	54.26	0.88

A total of 11893 reads were obtained, with a maximum of 4001 sequences for C3 and a minimum of 3922 for C2 (Table 25). DADA2 denoising resulted with 102 SV for C1, 99 for C2 and 157 for C3. Evenness calculation was the highest in C1 and lowest in C3. Shannon and Inverse Simpson indexes showed that C2 had the lowest values in both cases. C1 and C3 had the highest value for Inverse Simpson and Shannon indexes calculation respectively. Bray-Curtis similarity at genus level showed that C1 and C2 clustered together in the dendrogram (Fig 40).

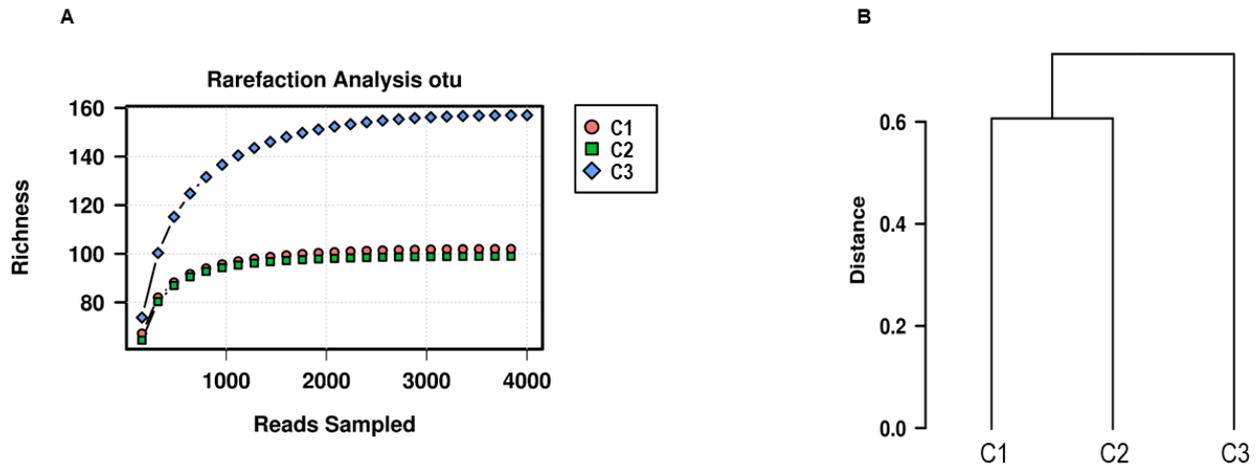


Fig 40. Rarefaction curves and Bray-Curtis clustering analysis at genus level of the three microbial communities.

### 5.3.2. Results on the diversity and taxonomy of the microbial communities in the three Cenote Abyss samples

From the taxonomy analysis, 22 bacterial and 3 archaeal phyla were detected in all three samples under analysis. Archaeal sequences were identified in C1 and C3 with abundances lower than 4%. *Thaumarchaeota* and *Euryarchaeota* were found in both samples, whereas the *Nanoarchaeota* phylum was present in C3 only (Fig 41). In all samples, the two most abundant phyla were *Proteobacteria* (29% in C1, 30% in C2 and 36% in C3) and *Bacteroidetes* (26% in C1, 31% in C2 and 15% in C3). *Proteobacteria* were mainly composed of *Gammaproteobacteria*, followed by *Alphaproteobacteria* and *Deltaproteobacteria*. In all samples, *Gammaproteobacteria* were rich in *Betaproteobacteriales* order that differed at family level. Indeed, C1 and C2 were rich in members of *Burkhoderiaceae* family (10% and 11%), whereas C3 were mainly composed of *Nitrosomonadaceae* (8%). The *Bacteroides* phylum was mainly represented by *Sphingobacteriaceae* of *Sphingobacteriales* order in C2 (16%) and C1 (8%), the latter also presenting high abundance of *Prolixibacteraceae* of *Bacteroidales* order (11%) (Fig 42). On the contrary, C3 was rich in *Chitinophagaceae* of *Chitinophagales* order (7%). When considering lower abundant phyla, the taxonomic profile resulted more diversified among the samples. *Actinobacteria* was the third most abundant phylum in C1 and C2 (19% and 25% respectively), whereas it was detected at low

abundance (1%) in C3. Within *Actinobacteria*, *Micrococcales* was the most represented order (8% in C1 and 13% in C2), followed by *Propionibacteriales* (4% in C1 and 2% in C2). On the other hand, C3 was highly rich in *Nitrospiraceae* of *Nitrospirae* phylum, which accounted for 8% of C3 microbial community and was totally absent in C1 and C2. *Acidobacteria*, mainly of *Blastocatellaceae* family, were highly present in C3 (8%) and found in traces in C1 and C2 (<1%). *Firmicutes* (mainly of *Clostridia* class) was detected at 12% in C1, 2% in C2 and <1% in C3. Verrucomicrobia, *Planctomycetes* and *Chloroflexi* were retrieved in all samples with abundances varying between 2% and 6%.

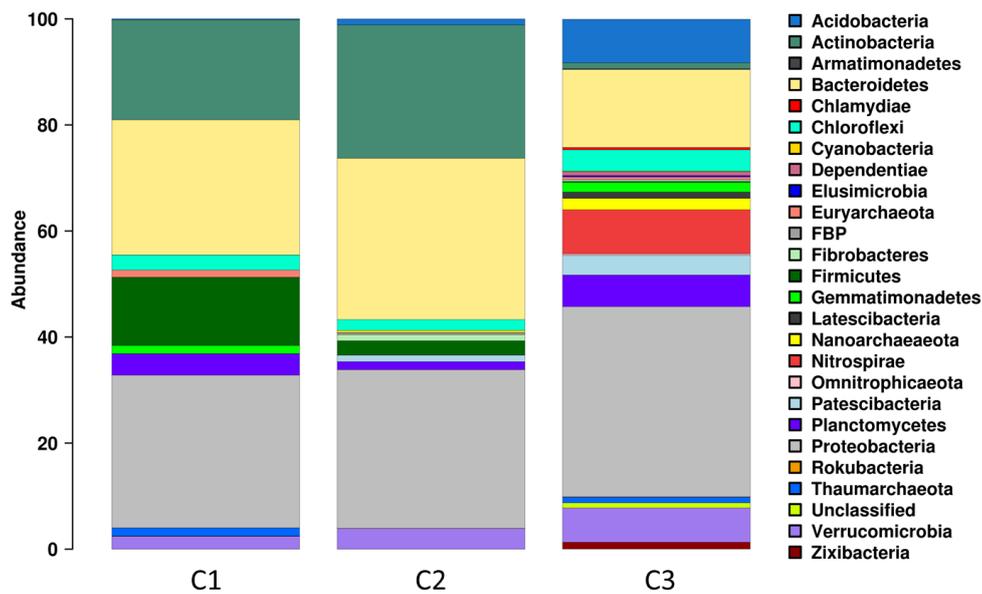


Fig 41. Taxonomic analysis of the three microbial communities at phylum level.

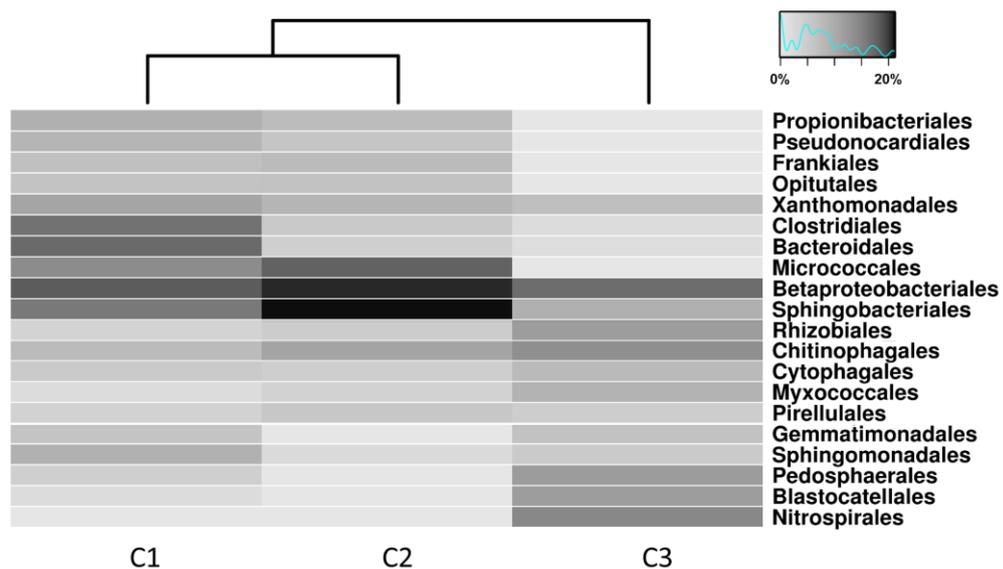
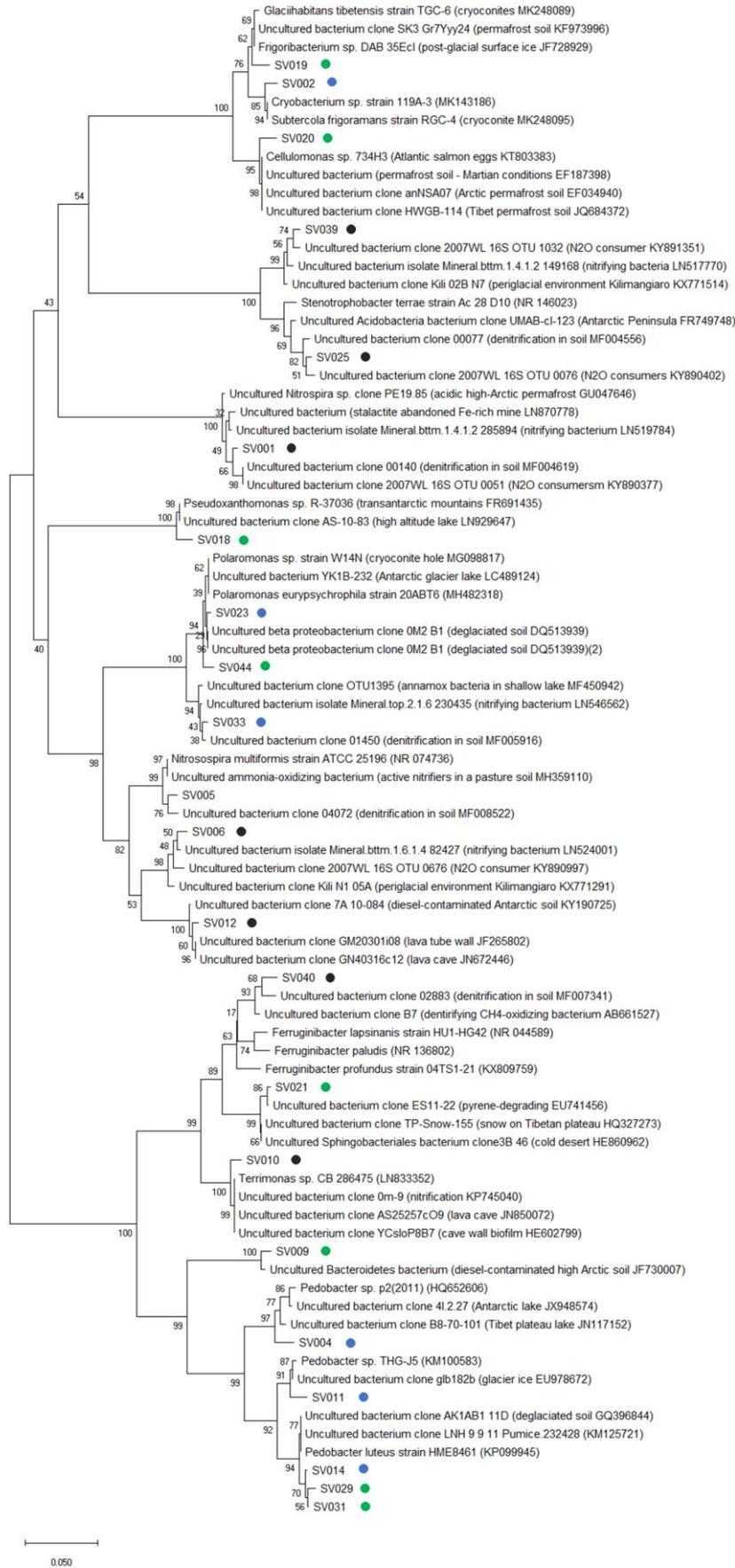


Fig 42. Heat map showing the 20 most abundant orders in the three microbial communities.

In line with the clustering analysis, at genus level the microbial community of C3 strongly differed from those belonging to C1 and C2 (Table 26). Sequences unclassified at genus level amounted at 31% in C1, 46% in C2 and 72% in C3. *Pedobacter* was the most representative genus within the *Bacteroides* phylum in C1 (8%) and in C2 (13%). Other abundant genera (2-6%) of *Bacteroides* were BSV13, WCHBI-32 (*Prolixibacteraceae*) and *Ferruginibacter* (*Chitinophagaceae*) in C1. Within *Gammaproteobacteria*, *Polaromonas* was present in both C1 (2.5%) and C2 (4.3%). *Pseudoxanthomonas* (3.6%) and *Rhizobacter* (2.5%) were found in C1 only, while *Nitrospira* (4.2%) was detected in C2. Actinobacterial *Micrococcales* were mostly represented by *Actinotalea* in C1 (3.6%) and *Parafrioglobacterium* in both C1 (2.4%) and C2 (6.2%). In general, no high abundant genera (>1%) belonging to C3 were shared with C1 and C2 except for *Ferruginibacter* (1.8% in C3). *Nitrospira* dominated C3 (8.3%). MND1 and IS-44 of *Nitrosomonadaceae* family were detected at 6%, whereas *Terrimonas* of *Chitinophagaceae* represented 3% of C3 microbial community.

Phylogenetic analyses of 16S rRNA sequences belonging to the most abundant taxa showed high similarities with sequences retrieved from cold environments and involved in the nitrogen cycle (Fig 43).

Phylum	Class	Order	Family	Genus	C1	C2	C3
Bacteroidetes	Bacteroidia	Sphingobacteriales	Sphingobacteriaceae	<i>Pedobacter</i>	8.01	12.67	0
Actinobacteria	Actinobacteria	Micrococcales	Microbacteriaceae	<i>Parafrioglobacterium</i>	2.39	6.2	0
Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	<i>Nitrospira</i>	0	0	8.32
Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	<i>Polaromonas</i>	2.47	4.28	0
Bacteroidetes	Bacteroidia	Bacteroidales	Prolixibacteraceae	BSV13	5.97	0.54	0
Bacteroidetes	Bacteroidia	Bacteroidales	Prolixibacteraceae	WCHBI-32	5.01	0.23	0
Actinobacteria	Actinobacteria	Pseudonocardiales	Pseudonocardaceae	<i>Pseudonocardia</i>	2.85	1.43	0
Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Nitrosomonadaceae	<i>Nitrospira</i>	0	4.23	0
Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	<i>Ferruginibacter</i>	2.32	0	1.83
Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Nitrosomonadaceae	MND1	0	0	3.97
Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Nitrosomonadaceae	IS-44	0	0	3.72
Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	<i>Pseudoxanthomonas</i>	3.58	0	0
Actinobacteria	Actinobacteria	Micrococcales	Cellulomonadaceae	<i>Actinotalea</i>	3.55	0	0
Actinobacteria	Actinobacteria	Propionibacteriales	Nocardioidaceae	<i>Aeromicrobium</i>	0.96	2.3	0
Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	<i>Sphingomonas</i>	2.6	0.26	0
Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	<i>Terrimonas</i>	0	0	2.85
Actinobacteria	Actinobacteria	Propionibacteriales	Nocardioidaceae	<i>Nocardioides</i>	2.57	0	0
Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	<i>Rhizobacter</i>	2.49	0	0
Firmicutes	Clostridia	Clostridiales	Gracilibacteraceae	<i>Lutispora</i>	1.39	0.92	0
Firmicutes	Clostridia	Clostridiales	Caldicoprobacteraceae	<i>Caldicoprobacter</i>	2.19	0	0



**Fig 43.** Phylogenetic analysis of the most abundant SVs retrieved from the three ice samples. Green dots indicate SVs characterizing sample C1; blue dots indicate SVs characterizing sample C2; black dots indicate SVs characterizing sample C3.

### 5.3.3. Discussion on the role of microbial communities in the Cenote Abyss cave

All life on Earth evolved from microorganisms and, billions of years later, the planet's smallest life forms are still fundamental to the biosphere<sup>119</sup>. Today, global warming is modifying some microbial cycles on a scale big enough to trigger damaging climate feedback loops. *Bacteria*, *Archaea* and *Fungi* have access to more organic material and produce extra carbon dioxide as the planet warms. In Earth's cold and glacial places, ice is melting. Glaciers are retreating fast due to rapidly warming temperatures. As the Earth warms, the melt line moves upwards so that the glacier melts faster at the bottom, shortening the glacier and reducing its mass<sup>111</sup>.

Cenote Abyss is a high-altitude oligotrophic ice cave that represents a study model both for describing the microbial life development in caves under psychrophilic conditions and for understanding the impact of climate change and ice melting on microbial diversity. Here we provide a first description of microbial communities inhabiting three samples collected from ice layers within the cave at different depth. Statistical analysis showed that the microbial communities associated to the outermost of the cave (samples C1 and C2) are more similar in terms of diversity and taxonomy as respect to the innermost sample C3 (Table 25 and Fig 38). Further, the abundances of the microbial community belonging to the innermost sample was featured by a higher domination of peculiar microbial taxa and, at the same time, was more diverse as compared to the C1 and C2. These differences might be a consequence of the recent disappearing of the lake that gave access to the cave entrance. This phenomenon led to the establishment for the first time of direct contacts between the cave ice layers and the surface atmosphere. Therefore, the outermost ice layers have begun to be exposed to the outside atmospheric gas composition, enhancing the metabolic activity of nitrogen-fixing bacteria. Diazotrophic (nitrogen fixing) bacteria are fundamental for providing nitrogen stocks in the form of ammonia to facilitate non-diazotrophic organisms' development in glacial oligotrophic environments<sup>114</sup>. Nitrogen fixing *Cyanobacteria* have been considered primary members involved in the formation of reduced nitrogen compounds. However, Cenote Abyss is practically devoid of *Cyanobacteria* (<0.3%), probably because of the lack of sunlight. On the other hand, members of the nitrogen-fixers *Polaromonas* and *Geobacter* genera<sup>120</sup> were detected in C1 and C2 samples, while they were absent in C3. This evidence suggests that atmospheric nitrogen fixation mostly occurs in the outermost parts of the cave. In the innermost and deepest parts of the cave, specific bacterial species of *Nitrospira*, *Nitrosomonadaceae* and *Blastocatellaceae* might have a role in enhancing processes involving the transformation of the metabolic products obtained from nitrogen fixation in the upper parts of the cave by performing other pathways in the N cycle (such as nitrification)<sup>114</sup>. The great depth of sample C3 together with the static cold conditions probably drove the microbial communities to select the most suitable bacterial species to face the oligotrophic and cold habitat. In

this respect, the metabolic activities of *Nitrospira* genus might represent a key factor for the sustainment of the entire microbial community in C3. Members belonging to *Nitrospira* have been previously found in Arctic soils and glaciers<sup>121</sup>. In the cave, *Nitrospira* is involved in nitrogen cycle through oxidation steps of reduced forms of nitrogen by exploiting metabolic pathways involving *amo* (ammonia monooxygenase), *hao* (hydroxylamine dehydrogenase) and *nxr* (nitrite oxidoreductase) genes<sup>121</sup>. Some *Nitrospira* species are also able to perform complete nitrification (Comammox bacteria)<sup>101</sup>. Further studies demonstrated the capability of *Nitrospira* and other ammonia-oxidizing bacteria to colonize mineral substrates and coatings<sup>122</sup>. With these findings, the capability of *Nitrospira* to perform complete nitrification can be exploited instead of partial in wastewater treatment, as complete nitrification leads to lower emissions of the greenhouse gases into the atmosphere (nitrous oxide and nitric oxide)<sup>123</sup>. Other taxonomic groups like *Terrimonas* and *Ferruginibacter* probably act as co-operators into the sustainment of the microbial community in C3 together with *Nitrospira*. *Terrimonas* and *Ferruginibacter* genera have been also detected in Arctic and glacial soils<sup>124</sup>. Their phylogenetic similarities with clones involved in nitrogen cycle confirm their role in nitrification and denitrification in the Cenote Abyss.

The two outermost samples shared many of the bacterial genera retrieved in cold environments and involved in the nitrogen cycle, such as *Polaromonas*, *Parafrioglobacterium* and *Pedobacter*. Members of these genera were previously described in microbial deposits inside ice caves<sup>18,110</sup>. In particular, *Pedobacter* members have been recently detected in association to ancient ice layers formation<sup>110</sup>. In this work, *Pedobacter* members were shown to possess genes involved in ammonia metabolism and transport, which are responsible for the utilization of fixed nitrogen. *Betaproteobacteriales* are also common in ice and glacier location<sup>125</sup>. Members of this order, which the psychrophilic *Polaromonas* species belong to, are characterized as r-strategists, i.e. capable of rapidly responding to environmental fluctuations and of utilizing several nutritional sources thanks to their extensive metabolic versatility<sup>126</sup>. *Polaromonas* strains are also able to enhance mineral weathering in an Alpine glacier forefield<sup>120</sup>, suggesting their role in glacial soils modification processes. *Micrococcales* represented the most abundant order of *Actinobacteria* in C1 and C2, with *Actinotalea* and *Parafrioglobacterium* as richest bacterial genera. Playing an important role in the soil development and biogeochemical cycling, *Actinobacteria* survive in extreme environments, being a key component of frozen habitats such as permafrost, glacier forefields, ice cores and cryoconites. In particular, members of *Actinotalea* genus were detected in perennial cave ice and associated to 900 years old ice<sup>115</sup>.

Figure 44 summarizes the microbial community structures in different depths of the cave along with the hypothesized role of microbes in nitrogen metabolism. Deeper studies on the cave microbiome metagenomes will aim at understanding the biogeochemical cycles governing the different layers of the cave.

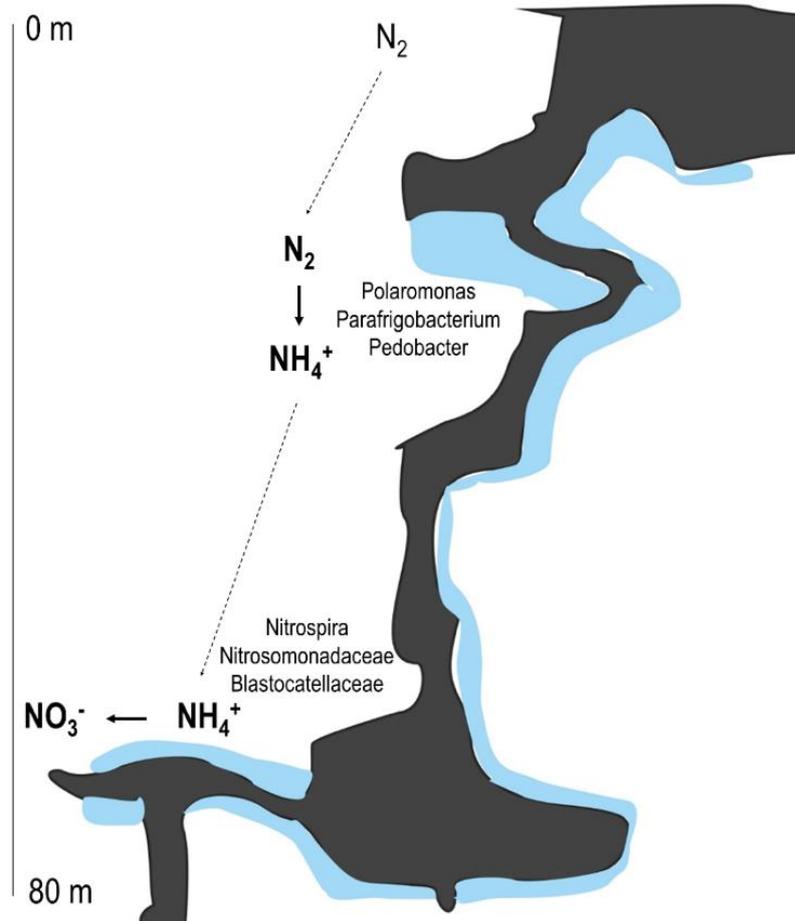


Fig. 44. Graphical representation of the hypothesized microbial activity in Cenote Abyss.

## **6. Metabolic and phenotypic analyses of microbial strains isolated from different cave systems**

### **6.1. Introduction**

The terrestrial subsurface microbiome has gained considerable amount of interest in the recent years because of its rich potential resource for biomining novel genes encoding enzymes and metabolites possessing antimicrobial activities. These genes make such organisms excellent candidates for new antibiotic and enzyme discovery for potential industrial development, including food industry, cleaning industry, including cleaning agents such as detergents and biocides, molecular biology studies, and in several features of biotechnology<sup>127</sup>. However, as opposed to the taxonomic and phylogenetic analyses, information on the metabolic potentials of cave microbial communities is still limited.

Among the useful molecules produced by cave bacteria, the discovery of new drugs represents one of the most investigated fields. Multi-drug resistant bacteria represent a severe problem for the human health that requires new active antimicrobials at wide spectrum<sup>128</sup>. This problem is getting more complex by the emergence of increasing number of Gram-positive and Gram-negative multi-drug resistant bacterial strains that are unresponsive to many antimicrobial treatments, including the use of amino glycosides, macrolides or fluoroquinolones and the first, second and third generation of penicillin and cephalosporin. In this contest, unexplored natural environments like caves represent unique containers of still undescribed microbial species potentially producers of a wide range of antimicrobials<sup>12</sup>. This potential is a consequence of the physiological and metabolic adaptability of the microbial communities during geological times, which led to a slow development of a peculiar microbial diversity in terms of taxonomy and chemicals production<sup>11</sup>. In deep caves, the lack of interactions of the microbial communities with the common human pathogens represents a promising starting point for the possible discovery of new antimicrobial molecules able to face the multidrug-resistance problem. Investigation on new chemicals has been recently focusing on caves and other extreme environments characterized by unfavourable life conditions, such as glaciers, deserts, volcanos, and geysers, where microbes activate peculiar metabolic pathways putatively leading to the production of unusual molecules<sup>79</sup>. Here, in stressful conditions, microorganisms establish specific synergistic cooperation to thrive, often organizing themselves in biofilm structures and biodeposits<sup>129</sup>. In some cases, microorganisms can also establish competition mechanisms against other microorganisms based on the production of bioactive molecules for trace metal scavenge and with antimicrobial activities.

*Streptomyces* represents the most active known genus in bioactive compounds production described so far. In relation with this, *Actinobacteria* is the phylum that is most studied from caves for the discovery of new antimicrobials. Recent studies on cave microbiomes have also demonstrated that other bacterial phyla are likely to produce novel antimicrobial agents, like *Bacteroidetes*, *Firmicutes*, *Proteobacteria* and *Cyanobacteria*<sup>12,16</sup>.

This chapter investigates the metabolic potential of bacterial isolates from different cave systems. In particular, we investigated their ability to produce antimicrobial molecules and to perform silica solubilisation. Four selected bacterial isolates of noteworthy interest have been further phenotypically characterized in relation to their capacity to metabolise specific carbon sources and to resist a selection of chemical compounds. Lastly, a preliminary characterization of putative antimicrobial molecules is reported.

## 6.2. Materials and methods

### 6.2.1. Bacterial strain isolation, purification and top-agar bioassay.

The various cave samples (biofilms and speleothems) used in this PhD thesis were collected from different cave systems, i.e. the Tepui orthoquartzitic cave system<sup>10</sup>, the ice cave Cenote Abyss<sup>118</sup> and the sulfuric caves Fetida<sup>20</sup> and Porretta Terme<sup>19</sup>.

The recovery of strains from raw cave samples was performed by using the enriched medium ISP2 agar (yeast extract 4 g/L, malt extract 10 g/L, dextrose 4 g/L, agar 20 g/L) in two different ways: i) direct cultivation from the cave raw sample on ISP2 medium with a sterilized cotton swap; ii) pre-treatment at 55 °C for 30 min in order to eliminate fast-growing bacteria and transfer on ISP2 plates. Plates were incubated for two weeks at 30 °C. Purifications were performed by re-cultivating single colonies on ISP2 agar plates.

For the top-agar bioassay test, each purified isolate was grown on an ISP2 plate for 7 days. Bioassays were performed by covering the grown isolates with a resuspension of the bacterial pathogen ( $2 \times 10^6$  CFU/mL) in 15 mL of LB (NaCl 10 g/L, tryptone 10 g/L, yeast extract 5 g/L) agar 0.7%. The cultures were incubated overnight at 37 °C. The inhibitory activity of the pathogen was observed by evaluating clarifying zones around the cave isolates.

The pathogenic bacterial strains used for the top-agar bioassay tests include the three Gram negative *Escherichia coli* NCTC 12923, *Klebsiella pneumoniae* ATCC 700603, *Pseudomonas aeruginosa* PAO1 and the Gram-positive *Staphylococcus aureus* ATCC 29213.

### **6.2.2. Extraction of secondary metabolites and top-agar bioassay.**

After the growth of the pure isolate, agar was fragmented into pieces with a sterile knife and put into falcon tubes with the addition of 20 mL of ethyl acetate, methanol or butanol. Extractions were performed in triplicate. The falcon tubes were incubated at 30 °C and 130 rpm for 18 hours. The liquid phases were centrifuged, and their supernatants were concentrated under vacuum, resuspended in 200 µL of 10% DMSO and kept frozen until their use for bioactivity screening and for LC-MS analysis.

For the bioassay, 15 mL of LB agar 0.7 % per plate was prepared inoculating  $2 \times 10^6$  CFU/mL of the bacterial pathogen before pouring. After solidification, 10 µL of the extract (resuspended in 10% DMSO) was spotted on the plate. 3 µL of kanamycin was spotted as negative control. The cultures were incubated overnight at 37 °C. The inhibitory activity on the pathogen was observed by evaluating clarifying zones.

### **6.2.3. Sample preparation and HPLC-MS analysis**

The HPLC-MS analyses were performed in NAICONS srl (<http://naicons.com>, Milano, Italy). Organic extracts for each isolate and medium blank were prepared in 10% DMSO. Liquid chromatography-mass spectrometry experiments were carried out in an HPLC LaChrom (VWR HITACHI-ELITE) coupled to ESI-IT mass spectrometer Amazon X (Bruker-Daltonics). The chromatographic analysis used a XTerra C-18 column (4.6 × 150 mm, 5µm) and a flow rate of 1.0 mL/min. The mobile phase consisted of 0.075% formic acid (A) and acetonitrile with 0.075% formic acid (B).

The solvent program started with 10% B for 3 min and a linear increase to 40% B in 17 min, followed by a linear increase to 100% B in 15 min, then followed by 10 min at 100% B. The total analysis time for each sample was 50 min. The sample (20 µL) was injected by an autosampler. The mass spectrometry was carried out in both positive and negative ionization modes with a spray voltage at 4 kV and capillary temperature at 250°C. The mass range was set from  $m/z$  70– 3000. Raw data files generated by HPLC-MS were first converted to netCDF files.

### **6.2.4. Characterization of bacterial 16S rRNA and fungal ITS from pure cultures**

Genomic DNA was extracted by performing Colony PCR. Bacterial 16S rRNA and fungal ITS were amplified by using bacterial universal primers 27F and 1492R and fungal primers ITS1 and ITS4 (Table 27). The 16S rRNA gene fragment was amplified under following PCR conditions: one cycle of 95 °C for 5 min followed by 28 cycles of 95 °C for 1 min, and annealing at 58 °C for 50 s with extension at 72 °C for 50 s, and a final extension step at 72 °C for 10 min. PCR products were purified

using PCR purification kit (Thermo Scientific, Waltham, USA), and sequenced commercially (Macrogen, South Korea). Sanger sequencing was performed with the external service Eurofins Genomics. Sequence alignment with references was performed with ClustalW. The neighbour-joining method with 1000 bootstrap in MEGA7 was used for construction of the phylogenetic tree.

Target sequence	Primer set	Sequence (5'-3')	Reference
Bacterial 16S rRNA	27F	AGAGTTTGATCMTGGCTCAG	36
	1492R	TACGGYTACCTTGTTACGACTT	36
Fungal ITS	ITS1	TCCGTAGGTGAACCTGCGG	130
	ITS4	TCCTCCGCTTATTGATATGC	130

### 6.2.5. *In vitro* silica solubilization test

Two approaches were used to identify the silica solubilizing bacteria among the cave isolates we obtained from different caves. The first approach consists in the cultivation of cave isolates on Glucose Agar medium (10 g/L Glucose, 2% agar) with the addition of 0.25% magnesium trisilicate before autoclaving. The ability in solubilizing silica was evaluated for the presence of halos around the cave isolate.

The second approach consists in investigating if the silica modification processes is due to active microbial metabolism or to indirect activity of the bacterial cells, by using both live and dead cells. At first, 100 mg of wall samples collected from Imawari Yeuta cave were resuspended in 1 mL of saline solution (0.85% NaCl). For the dead cells test, the resuspension was sterilized through autoclave (20 min at 121 °C). Further, for both dead and live cells tests, 100 µL of the resuspension were inoculated in falcon tubes in the presence of a home-made minimal medium, which mimics the mineral composition of the cave (50 mg/L BaCl, 10 mg/L ZnSO<sub>4</sub> x 7H<sub>2</sub>O, 13 mg/L CuCl x 2H<sub>2</sub>O, 200 mg/L MgSO<sub>4</sub> x 7H<sub>2</sub>O, 200 mg/L KCl, 50 mg/L AlCl<sub>3</sub>, 1 g/L NH<sub>4</sub>NO<sub>3</sub>, 0.5 g/L Soy Peptone), with the adding of 0.1% colloidal silica (<sup>C</sup>SiO<sub>2</sub>, Ludox-LS, Sigma Aldrich). Ludox-LS colloidal was chosen according to its relative purity, as described by Amores (2009). After 3 months of incubation in static and dark conditions at room temperature, 1 mL of each culture filtered by using 0.22 filters, which were subjected to scanning electron microscopy as described in paragraph 2.7.

### 6.2.6. Phenotype Microarray

The metabolic profiles of bacterial isolates were investigated by using GEN III MicroPlates™ (Biolog, Inc., Hayward, CA), which include 71 carbon sources and 23 antimicrobial chemicals. The GEN III plates were inoculated with inoculating fluid IF-A using protocol A in accordance with the OmniLog Data Collection Software Identification System User Guide. Data were collected every 15 min using the full data logger option for 72 h and analysed with DuctApe and with the OPM package in R.

## 6.3 Results and discussion

### 6.3.1 Antimicrobial activity of cave isolates

#### *Library of isolates and inhibitory activity against pathogens*

A library of 168 cave isolates was created using traditional isolation procedures (paragraph 6.2.1) starting from cave samples originated from orthoquartzitic Tepui cave systems [Imawari Yeuta (chapter 3)], Fetida cave (chapter 4), Cenote Abyss (chapter 5) and Porretta Terme (an anthropogenic carbonatic sulfuric thermal cave located in Emilia Romagna, Italy)<sup>19</sup>. In particular, 60 strains originated from Tepui caves, 89 from Porretta Terme, 10 from Fetida cave and 9 from Cenote Abyss. By using top-agar assay procedure, an initial screening was performed to detect possible antimicrobial activities by these strains. For this purpose, four important human pathogenic strains were chosen as target bacteria, i.e. *Escherichia coli* NCTC 12923, *Klebsiella pneumoniae* ATCC 700603, *Pseudomonas aeruginosa* PAO1 and the Gram-positive *Staphylococcus aureus* ATCC 29213 (paragraph 6.2.1).

As a result, ten cave isolates (6.6% of the total amount of screened strains) were active against at least one of the tested pathogens. The anti-microbial potential was associated with the appearance of a zone of growth inhibition around the tested cave isolate (Fig 45). Among these, seven originated from Tepui caves and three from Porretta Terme cave. The morphology of the colonies on ISP2 plates of these active isolates are shown in Fig 46. Interestingly, among the tested pathogens, *Pseudomonas aeruginosa* resulted to be the most resistant strain, whereas *Staphylococcus aureus* was the most frequently inhibited by cave isolates (Table 28). The *Staphylococcus aureus* strain that was tested in this study was an ATCC type-strain not showing resistance against known antibiotics. Future tests will evaluate the activity of the cave isolates against clinical *Staphylococcus aureus* strains that frequently are Methicillin Resistant strains possessing multiple antibiotic resistance<sup>131</sup>.

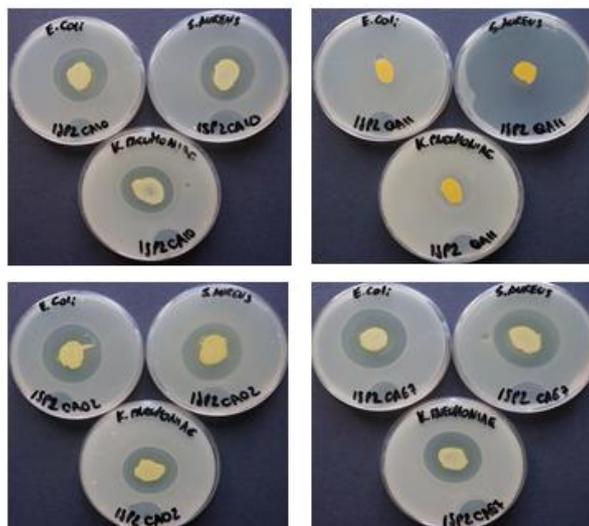


Fig 45. Clarifying zones representing the inhibitory activities of some cave isolates.

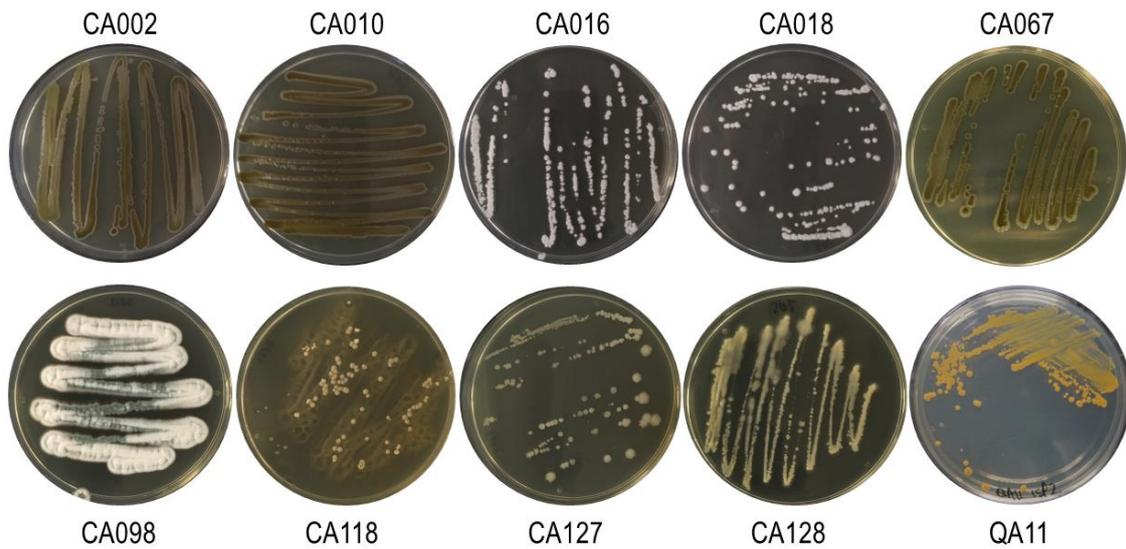


Fig 46. Pictures of the ten active cave isolates grown on ISP2 medium.

Table 28. Inhibitory activity of cave isolates against the target pathogens. Darkening shades indicate the diameter length of the clarifying zones: white=no halo, light grey= <2 cm, dark grey=2-4 cm; black: >4cm.				
Isolate	<i>E. coli</i>	<i>S. aureus</i>	<i>K. pneumoniae</i>	<i>P. aeruginosa</i>
CA002	Dark grey	Dark grey	Dark grey	White
CA010	Dark grey	Dark grey	Dark grey	White
CA016	Light grey	Light grey	Light grey	White
CA018	Light grey	Light grey	Light grey	White
CA067	Dark grey	Dark grey	Dark grey	White
CA098	White	Light grey	White	White
CA118	Light grey	Light grey	White	White
CA127	Light grey	Light grey	Light grey	Light grey
CA128	Light grey	Light grey	Light grey	Light grey
QA11	White	Black	White	White

### Taxonomic characterization of the microbial isolates

The ten active isolates were subjected to taxonomic identification by performing full-length 16S rRNA gene sequencing. The bacterial genera represented among the isolates were *Paraburkholderia* (*Betaproteobacteriales*, 3 isolates from Tepui caves – CA002, CA010, CA067), *Penicillium* (Fungi, 2 isolates from Tepui caves - CA016, CA018), *Rahnella* (*Gammaproteobacteria*, 1 isolate from Porretta Terme – CA118), *Serratia* (*Gammaproteobacteria*, 2 isolates from Porretta Terme – CA127, CA128), and *Sphingomonas* (*Alphaproteobacteria*, 1 isolate from Tepui caves – QA11) (Fig 46, Fig 47 and Table 29). For one Tepui isolate (CA098) we have not obtained phylogenetic information yet. Morphological analyses of the seven isolates from Imawarì Yeuta showed the presence of three micelia- and spore-forming fungi (CA016, CA018 and CA098) and four Gram-negative, yellow, non-

spore-forming, rod shaped bacteria (CA002, CA010, CA067 and QA11). Among these, QA11 produced an orange pigmentation both on solid and in liquid media that might indicate the presence of carotenoids<sup>132</sup>. Two isolates from Porretta Terme were also Gram-negative, non-spore-forming, straight short rods bacteria (CA127, CA128). One isolate was Gram-negative, spore-forming bacterium with an irregular growth shape visible on solid medium (CA118).

Isolate	Cave name	Country	Sample description	Cave niche	Morphology of the colony on ISP2 medium	Closer reference based on 16S rRNA gene
CA002	Imawari Yeuta	Venezuela	Violet slime on a lake surface	Water	Yellow-green	<i>Paraburkholderia briophyla</i> NR_042593 (99%)
CA010	Imawari Yeuta	Venezuela	Foam on a lake	Water	Yellow-green	<i>Paraburkholderia briophyla</i> MG770358 (99%)
CA016	Sima de la Lluvia	Venezuela	Red powder	Floor	White fungi	<i>Penicillium crysogenum</i> MK881028 (99%)
CA018	Sima de la Lluvia	Venezuela	White powder	Floor	White fungi	<i>Penicillium crysogenum</i> MK968256 (99%)
CA067	Cueva Pristina	Venezuela	White biofilm on weathered quartzite	Wall	Yellow-green	<i>Paraburkholderia briophyla</i> AM489500 (99%)
CA098	Imawari Yeuta	Venezuela	Silica stalactite	Roof	Green fungi	-
CA118	Porretta Terme	Italy	Vermiculation	Wall	White sporulating	<i>Rahnella</i> sp. AM403660 (99%)
CA127	Porretta Terme	Italy	White filament floating on water	Water	Whitish	<i>Serratia</i> sp. KM117223 (99%)
CA128	Porretta Terme	Italy	White filament sedimented in water	Water	Whitish	<i>Serratia</i> sp. AY689057 (99%)
QA11	Imawari Yeuta	Venezuela	Quartzitic rock	Wall	Orange	<i>Sphingomonas mali</i> MK302226 (98%)

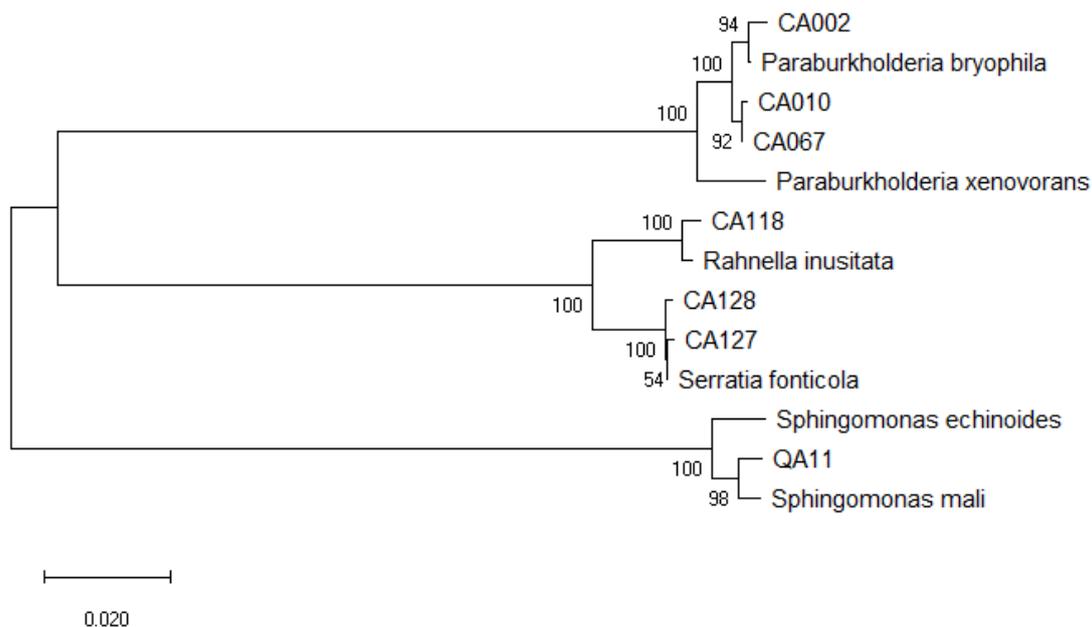


Fig 47. Phylogenetic analysis based on the bacterial 16S rRNA of cave isolates.

Filamentous fungi are well known for their production of substances with antimicrobial activities. Several studies focus on fungal diversity in cave systems, including the Domica cave in Slovakia and an Indian cave where *Penicillium* members (CA016 and CA018 isolates) have been identified<sup>133</sup>. The high rate of spore production of some fungi allows their penetration up to the deepest zones of a cave through air dispersion and dripping waters from the rock fractures. The Algerian Chaabe Cave evaluated the antimicrobial activity of 23 *Penicillium* species revealing their capability to inhibit, among others, strains of *E. coli*, *S. aureus* and *K. pneumoniae*, whilst none of the strains managed to inhibit the growth of *P. aeruginosa*<sup>134</sup>.

Members belonging to the *Serratia* genus (CA127 and CA128 isolates) have been found inhabiting cave niches and were described for their abilities to perform calcite biomineralization. The activity of metabolic extracts obtained from a strain of *Serratia marcescens* has been previously investigated<sup>135</sup>, revealing its antimicrobial potential against several bacterial pathogens. On the other hand, strains of *E. coli*, *S. aureus*, *K. pneumoniae* and *P. aeruginosa* resulted resistant to its activity. To the best of our knowledge, members of *Rahnella* (CA118 isolate) have never been described for their ability in producing bioactive molecules. Previous phylogenetic analyses allowed to detect members of *Rahnella* colonizing a Romanian perennial ice block of Scărișoara Ice Cave, however little is known about its role in cave systems<sup>116</sup>.

*Paraburkholderia* (CA002, CA010 and CA067 isolates) is a genus of the *Burkholderiaceae* family (*Gammaproteobacteria* class) that was first proposed in 2014 as a taxonomic group distinct from the *Burkholderia* species<sup>136</sup>. In the first study of Imawari Yeuta cave<sup>10</sup>, *Burkholderiaceae* members were found to dominate a water sample presenting a violet patina similar to that CA002 was isolated from. The taxonomical description of the *Paraburkholderia* genus derives because of the increase of the study of new isolates belonging to *Burkholderia*. *Burkholderia* genus contains pathogenic strains, whereas *Paraburkholderia* is related to non-pathogenic and environmental species. This distinction was attributed on account of the finding of robust conserved signature indels (CSIs) in protein sequences unique to the environmental and non-pathogenic species of (former) *Burkholderia* genus, which provide a category of molecular markers useful for understanding the phylogenetic relationships among taxonomically related microorganisms<sup>136</sup>. Genome analysis of peculiar *Burkholderia* strains revealed the presence of biosynthetic gene clusters involved in the production of bioactive molecules with antibacterial and antifungal capabilities<sup>137</sup>. On the other hand, few studies have been conducted so far on *Paraburkholderia* genus and little is known about the metabolic potential of its strains.

*Sphingomonas* species (QA11 isolate) contain glycosphingolipids (GSLs), specifically ceramide, instead of lipopolysaccharide (LPS) in their cell envelopes, and typically produce yellow-pigmented

colonies. The *Sphingomonas* genus includes species useful for the industrial production of gelling agents and other applications thanks to their ability to produce extracellular polymers (“sphingans”) and to biodegrade environmental contaminants<sup>138</sup>. Up to now, the *Sphingomonas* genus has never been described as putative producers of bioactive molecules with antimicrobial activity.

Analysis of biosynthetic gene clusters of the only two genomes present in database of *Paraburkholderia bryophila* and *Sphingomonas mali* through AntiSMASH revealed the presence of a scarce number of annotated gene clusters with low sequence homology with known clusters. Together these aspects make *Paraburkholderia* and *Sphingomonas* two potential genera for the discovery of novel antimicrobial compounds.

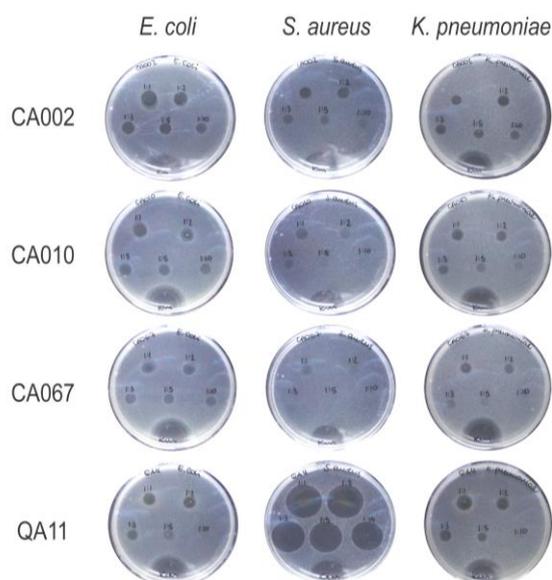
#### ***Evaluation of the production of bioactive molecules by the ten isolates previously screened through top-agar bioassay***

Organic extractions were performed both from solid and liquid growth cultures of each cave isolate selected for its ability to produce antimicrobials (Table 30). The total extractions (both from liquid and from solid media) were tested thorough soft agar overlay assay against the target pathogens. Our results revealed that the inhibitory activity was better retained when the extractions were performed starting from solid media bacterial growth. Many studies assessed that the metabolic pathways involved in secondary metabolites production are complex and regulated by numerous factors including the adhesion and formation of biofilms, which in some cases increase the production of secondary metabolites<sup>139</sup>. Further, considering that the source of isolation is often represented by cave rocky substrates, on which microbes develop and organize themselves in complex biofilm-like structures<sup>10,20</sup>, the growth on solid media probably denotes the best condition to enhance the production of bioactive molecules, representing the more analogous growth condition with the one in the cave.

Different organic solvents were tested in order to retain the antibacterial molecules, which included methanol, butanol and ethyl acetate (Table 30). The antimicrobial activity of metabolites was successfully retained using ethyl acetate for CA002, CA010, CA067, CA016 and CA018. Methanol resulted suitable for retaining the activity of QA11, CA118, CA127 and CA128. The activities CA010 and QA11 were also preserved when utilizing methanol and butanol respectively, though with a lower inhibitory effect on the pathogens growth. The need to utilize different organic solvents for metabolite extraction gave indications on the distinct chemical nature of the bioactive molecules produced by the different isolates.

Isolate		Ethyl acetate			Butanol			Methanol		
Code	Genus	<i>E. coli</i>	<i>S. aureus</i>	<i>K. pneum.</i>	<i>E. coli</i>	<i>S. aureus</i>	<i>K. pneum.</i>	<i>E. coli</i>	<i>S. aureus</i>	<i>K. pneum.</i>
CA002	<i>Paraburkholderia</i>									
CA010	<i>Paraburkholderia</i>									
CA016	<i>Penicillium</i>									
CA018	<i>Penicillium</i>									
CA067	<i>Paraburkholderia</i>									
CA098	TBD									
CA118	<i>Rahnella</i>									
CA127	<i>Serratia</i>									
CA128	<i>Serratia</i>									
QA11	<i>Sphingomonas</i>									

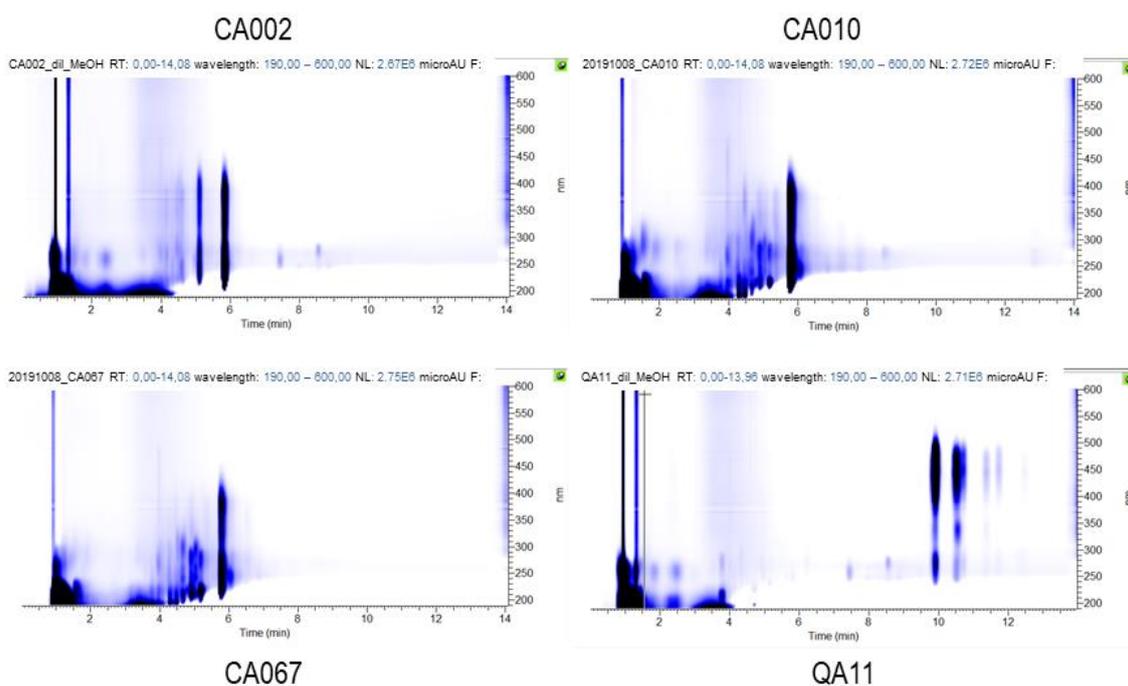
None of the three organic solvents resulted appropriate for retaining the activity of CA098. Therefore, forthcoming additional strategies will consist in using alternative solvents that have resulted active in previous studies, such as acidified ethanol<sup>140</sup> and chloroform<sup>141</sup>. The extracts of the *Paraburkholderia* strains (CA002, CA010 and CA067) resulted more active (on the basis of the growth inhibition zone extension) against *E. coli*, whereas *Sphingomonas* (QA11), the two *Serratia* (CA127 and CA128) and *Rahnella* (CA118) more extensively inhibited the growth of *S. aureus* (Fig 48). Further, by comparing these results with those obtained with the top-agar bioassay using live cells, we noticed an increase of the activity of the *Sphingomonas* isolate. Indeed, the live cells of QA11 resulted active against *S. aureus* only, while the secondary metabolites extracted were capable to inhibit also *E. coli* and *K. pneumoniae*, indicating that some active molecules of QA11 were not produced/secreted during the top-agar bioassay.



**Fig 48.** Inhibitory activity of metabolic extracts obtained from CA002, CA010, CA067 and QA11 against the target strains (*E. coli*, *S. aureus* and *K. pneumoniae*). For each extract, the total extraction and four dilutions (1:2, 1:3, 1:5 and 1:10, in 10% DMSO) are tested. Kanamycin (in the bottom of each plate) was used as negative control.

Considering the results obtained from the bioassay tests, we selected four strains showing the highest activity among those collected from Tepui cave systems, i.e. CA002, CA010, CA067 and QA11, to perform deep chemical analyses of the bioactive molecules.

With this purpose, LC-MS analyses were performed on the total extracts from the four selected isolates. In general, the UV spectra of the three *Paraburkholderia* considerably differed from the one belonging to *Sphingomonas* (Fig 49). Further, the UV spectra of CA010 and CA067 were more similar between each other, while the third *Paraburkholderia* strain showed a distinct metabolite profile respect to the other two. This result indicates that, in addition to a close phylogenetic relationship between CA010 and CA067 (Fig 47), the two strains also have similar metabolic profile. Moreover, an intense signal was present in the UV spectra of all the *Paraburkholderia* strains, which corresponded to a retention time of 5.50 minutes with a wavelength range between 225 nm and 435 nm. On the other hand, the *Sphingomonas* QA11 strain had the more intense signal at 10 minutes with a wavelength between 250 nm and 500 nm.



**Fig 49.** UV spectra of the metabolic extracts of CA002, CA010, CA067 and QA11 through LC-MS analysis.

The total metabolic extracts of the four strains were divided into 24 fractions using analytical HPLC. Each fraction was then tested against the target pathogens (Fig 50). As a result, among the three *Paraburkholderia* strains, the only active fractions were those obtained from CA002. This might be due to an insufficient amount of the extracts obtained from CA010 and CA067 for the fractionation or to an unsuitable type of separation column used for analytical HPLC (only a C18 column was used, which is a reversed-phase material that strongly retains nonpolar solutes).

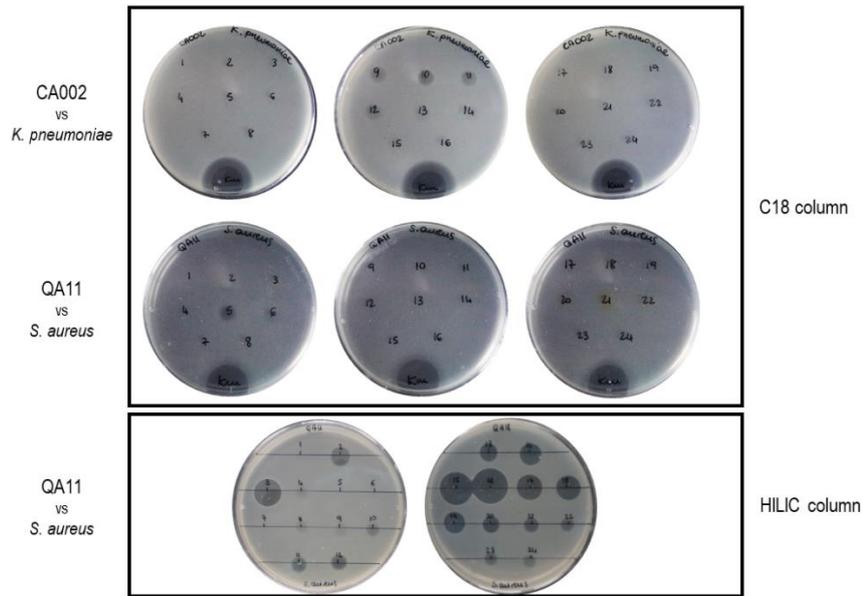


Fig 50. Inhibitory test of fractionated extracts.

The active fractions of CA002 included fractions from 9 to 13, with fraction number 10 showing the highest inhibitory activity (Fig 51). Low resolution LC-MS experiments performed on both active fractions (from 9 to 13) and non-active fractions (8 and 14) revealed the presence of one signal at 5.82 min that is specific of active fractions.

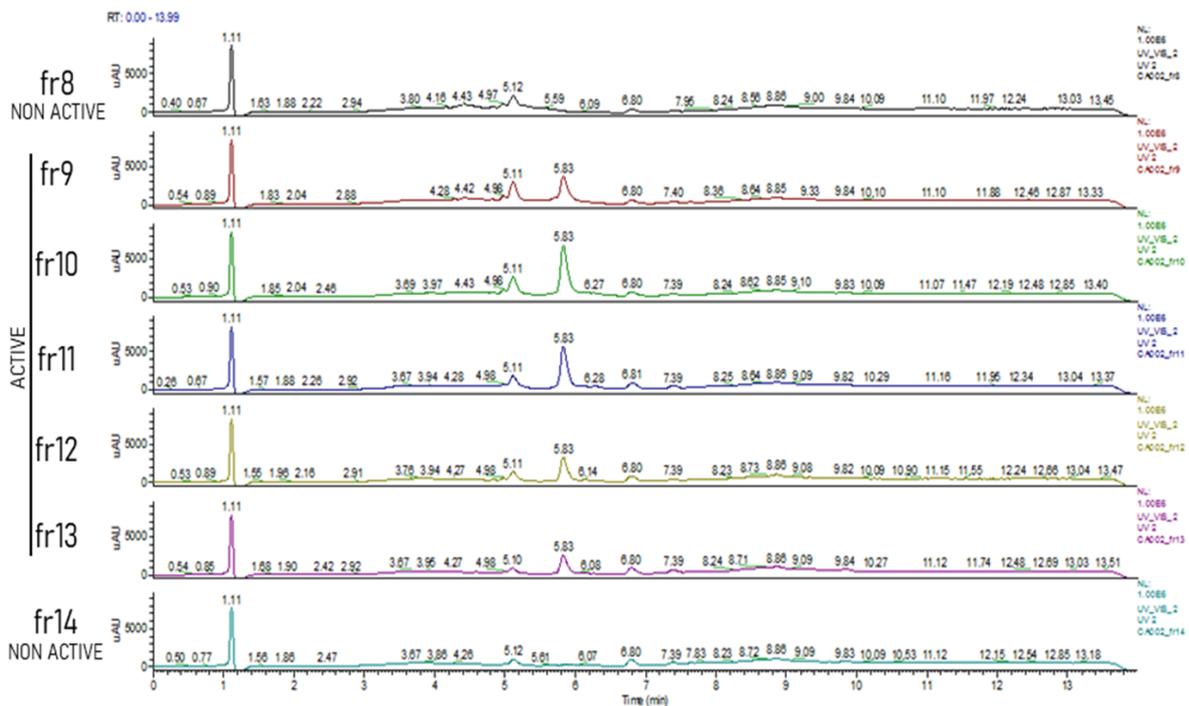


Fig 51. LC-MS low resolution analysis of CA002 fractions (fractions 8-14).

Among the fractions obtained with the C18 column of the *Sphingomonas* QA11 strain, only fraction 5 resulted to be active (Fig 50). Indeed, its intensity against the pathogen was limited and no peaks were detected in LC-MS in fractions 4 and 6 (data not shown). Considering the reduced inhibition of the fractionated extract as respect to the total extract, we hypothesized that the elution of the active molecule of QA11 was not appropriate when using C18 column (which is more suitable for separating hydrophobic molecules). Therefore, we performed a different fractionation process by using an HILIC column, which is specific for retaining polar hydrophilic compounds. This variation in analytical HPLC allowed to retain more active fractions, which showed their greatest inhibitory activity in fractions nos. 3, 16 and 18 (Fig 50). In particular, the inhibitory activity of fraction 3 was visible only against *S. aureus*, while the clarifying zone of fraction 16 and 18 was observed on *S. aureus* but also on *E. coli* and *K. pneumoniae*. These data suggest the presence of at least two antimicrobial compounds produced by QA11, one active only against *S. aureus* and one against all the tested pathogens. Low resolution LC-MS experiments performed on active fractions from 12 to 19 revealed the presence of one intense signal at 1.05 min that is specific of fraction 16 (Fig 52).

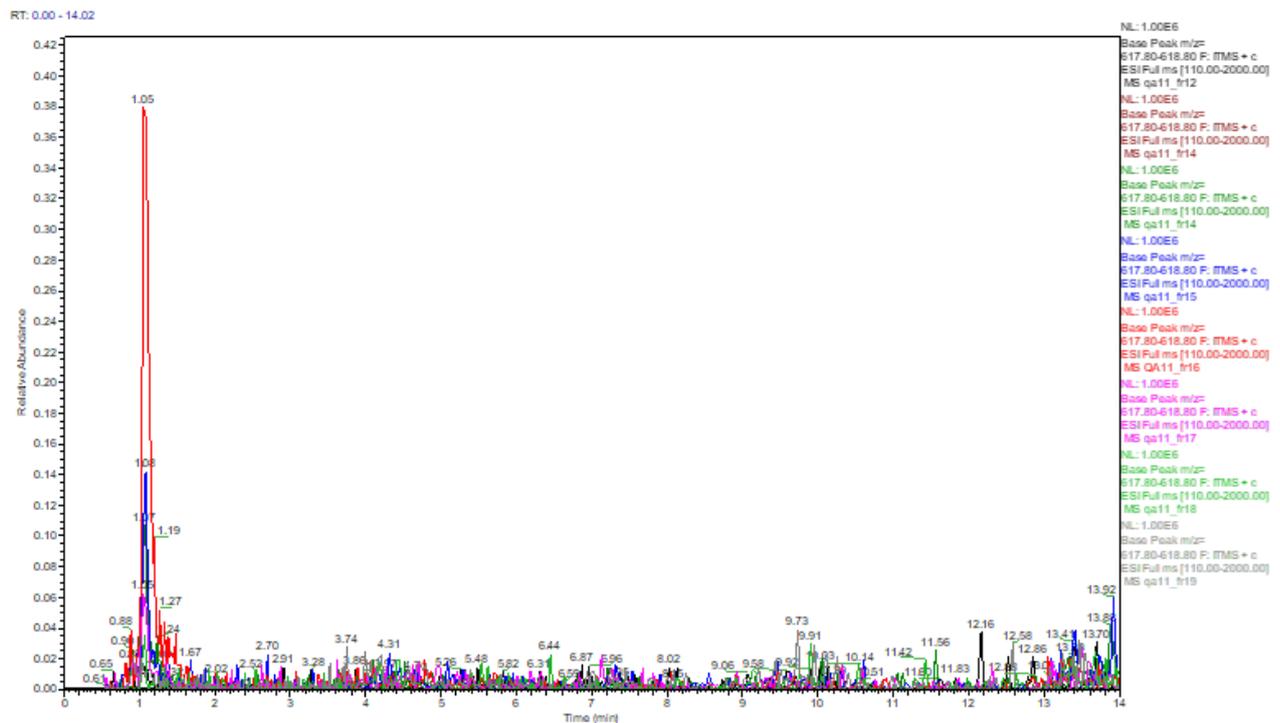
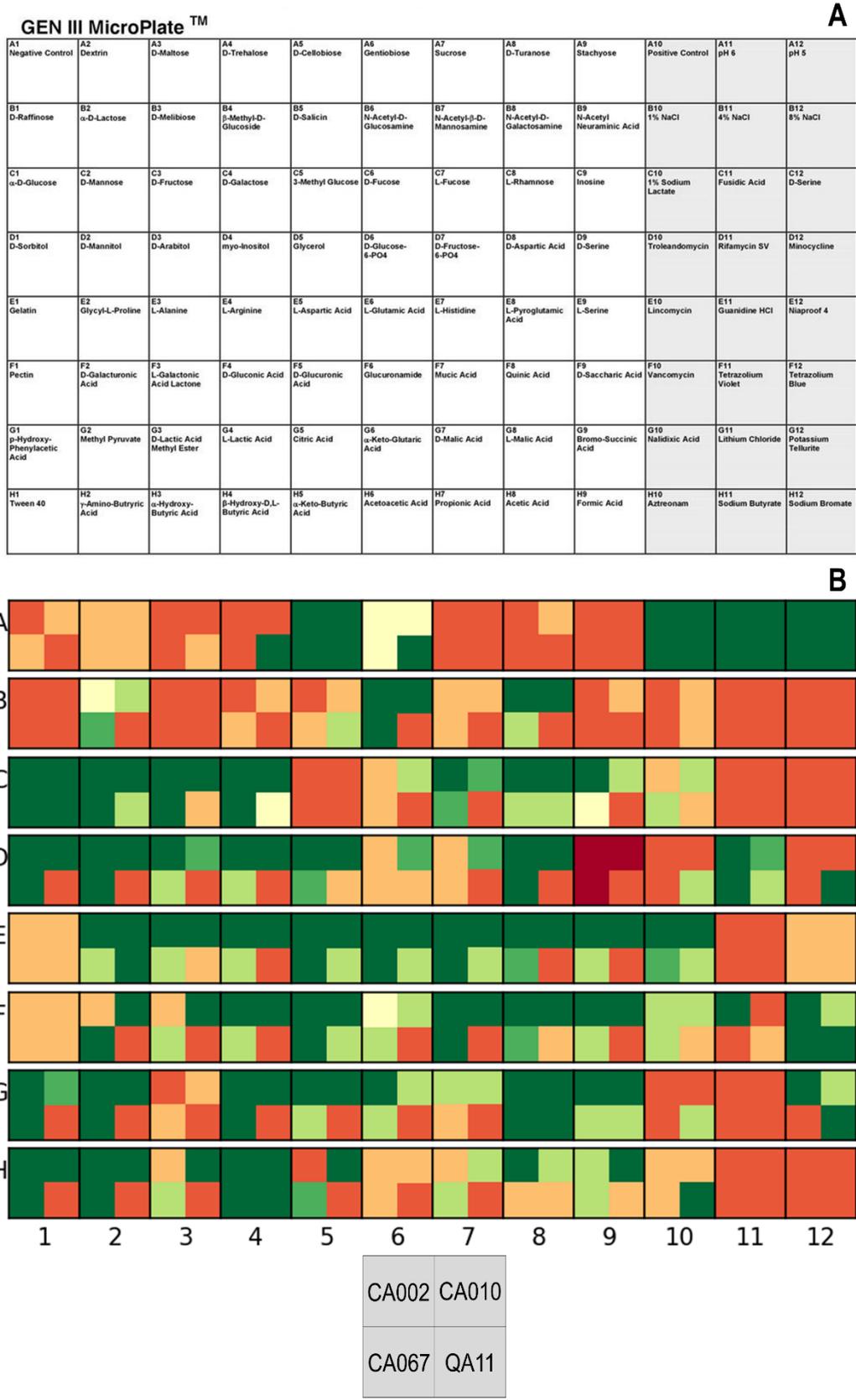


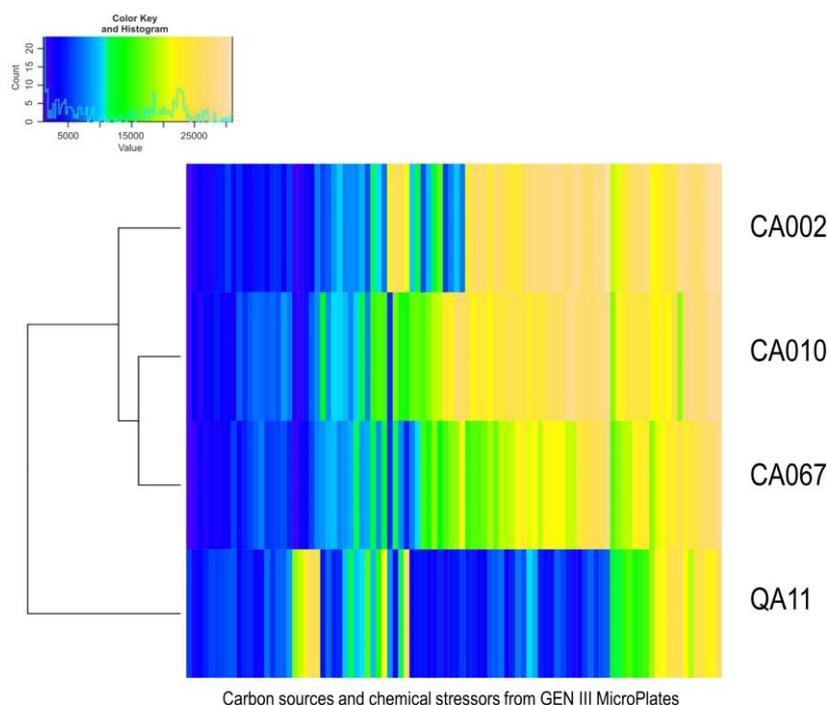
Fig 52. LC-MS low resolution analysis of QA11 fractions (fractions 12-19).

**Phenotype Microarray on the three *Paraburkholderia* strains and the *Sphingomonas* strains**

Metabolic profiles of the four strains (CA002, CA010, CA067 and QA11) isolated from Imawari Yeuta cave which showed the highest antimicrobial activity were investigated by using Biolog GENIII MicroPlates. The Biolog GEN III MicroPlate allows the parallel analysis of metabolic activities of a microorganism screening simultaneously 71 carbon sources and 23 chemical stressors (e.g. antibiotics, toxic metals). The test panel provides a “Phenotypic Fingerprint” that can be used to gather indications on the metabolic capacities of a microbial strains along with its capacity to resist/tolerate different stress conditions. As a result, the three *Paraburkholderia* strains utilized most of the 71 carbon sources present in GENIII plates, ranging from a minimum of 43 (CA002) to a maximum of 52 (CA010), whereas *Sphingomonas* QA11 exploited a selected number of carbon sources that accounted for 15 out of 71 tested only (Fig 53 and Fig 54). Twelve carbon sources were used by all the four tested strains. They included D-Cellobiose, D-Glucose, D-Mannose, L-Rhamnose, Gly-Pro, L-Aspartic Acid, L-Glutamic Acid, L-Histidine, D-Glucuronic Acid, L-Malic Acid, Bromo-Succinic Acid and  $\beta$ -Hydroxy-Butyric Acid. QA11 was the only strain able to utilize D-Threulose and D-Salicin, while D-Fucose and L-Fucose were used only by CA010. All four strains were able to grow at pH values of 5 and 6, and were resistant to antibiotics Lincomycin and Rifamycin SV, whereas they were sensible to Fusidic Acid, Lithium Chloride, Sodium Butyrate, D-Serina, Niaproof 4, Guanidine HCl, Sodium Bromate. Most of the tested strains were resistant to Potassium Tellurite except for CA067. Previous phenotypic studies using GEN III Microplates on five different *Burkholderia pseudomallei* strains revealed similar sensitivity fingerprint, although the carbon source utilization pattern was different<sup>142</sup>. *Sphingomonas* QA11 only was resistant to Troleandomycin, Minocycline, Nalidixic Acid and Aztreonam, whereas, unlike the other strains, it was sensible to Vancomycin.



**Fig 53.** Phenotype microarray of the four isolates CA002, CA010, CA067 and QA11. Panel A shows the scheme of the GEN III MicroPlates provided by Biolog™ with the indication of the carbon sources and chemical stressors added into each well. Panel B displays the metabolic activities (red – low, green – high) in the presence of the different carbon sources or chemical stressors (as indicated in Panel A). In particular, the activity in each well is divided into 4 squares, each one representing the activity of each strains, as represented by the scheme on the bottom of the image.



**Fig 54.** Heatmap of Phenotype Microarray. The colours indicate the rate of metabolic activity (blu – low, yellow – high) of each strain (CA002, CA010, CA067, QA11) in the presence of carbon source and chemical stressor.

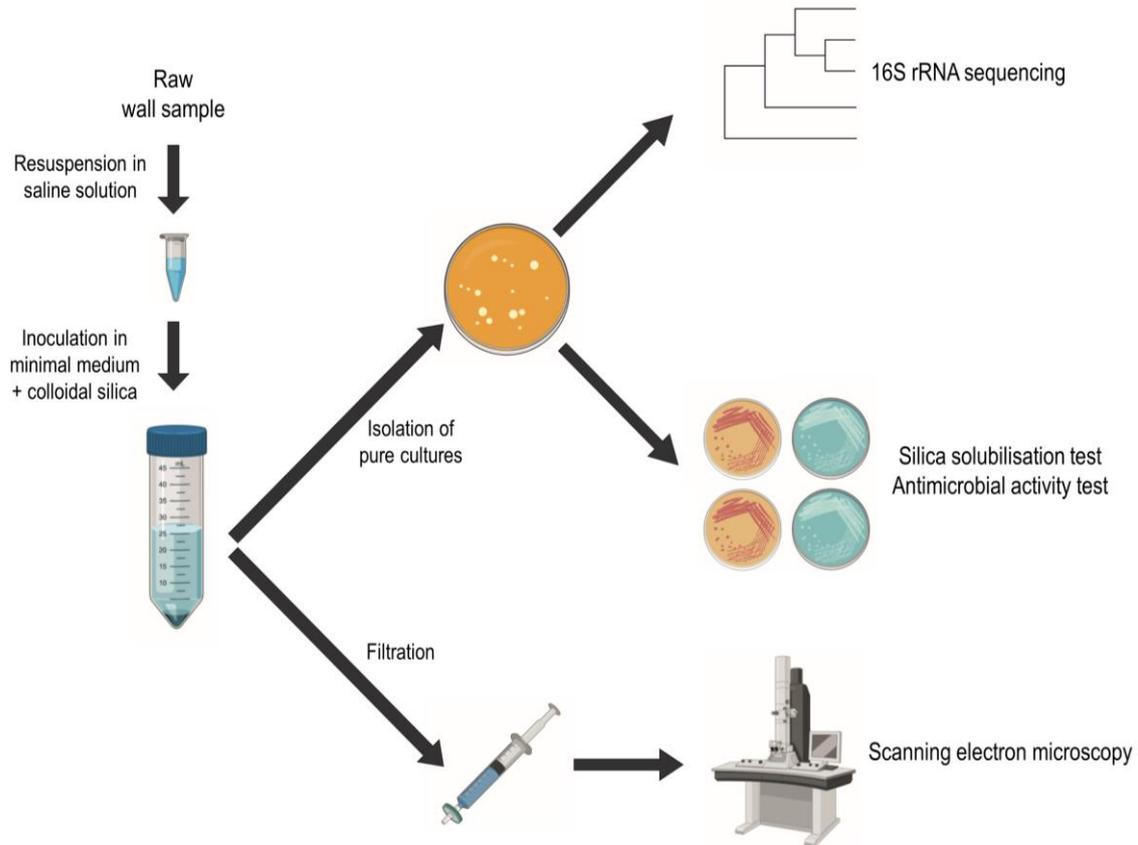
Statistical analysis of the similarities among the different metabolic profiles of the three *Paraburkholderia* strains showed that CA010 and CA067 clustered together as respect to CA002. This result was in accordance with the phylogenetic analysis based on 16S rRNA gene (Fig 47).

The antimicrobial activity of bacteria is regulated by several factors, including their metabolism of carbon compounds<sup>143</sup>. Our results showed that the metabolism of D-Glucose was present in all the four tested strains, and it has been commonly considered an enhancer of pathways leading the production of bioactive molecules<sup>144</sup>. Among the shared carbon sources used by the four strains, D-Cellobiose was highly metabolised (Fig 53). This metabolism might have a role in the bioactive molecules production. Indeed, it has been hypothesized that this disaccharide triggers antibiotic production in bacteria<sup>145</sup>, because of the observation that several aerobic cellulolytic bacteria (which metabolise cellobiose) are able to produce various antibiotics<sup>146</sup>. L-Rhamnose and polyhydroxybutyric acid (PHB) were also used by all the four strains. The biosynthesis of rhamnose lipids is known to be involved in the composition of bacterial biosurfactants, which have antibacterial activity<sup>147</sup>. Further, a possible role of PHB as a carbon reserve material used for antibiotic production was previously reported<sup>148</sup>.

In conclusion, the Phenotype Microarray provided indication on the capacity of the four anti-microbials producing strains to utilize different carbon sources and resist several stressors. These results give clues on the metabolic potentials of the three *Paraburkholderia* strains and the one *Sphingomonas* strains ruling the production of bioactive molecules.

### 6.3.2. *In vitro* evaluation of silica solubilising microbes

The solubilisation of silica was tested by cultivating 168 strains on two different culture media ISP2 and Glucose Agar (both media with the addition of an insoluble silicate compound, i.e. magnesium trisilicate). Figure 55 summarizes the work plan employed for the investigation of silica solubilizing microbes.



**Fig 55.** Graphical representation of the workplan for the investigation of silica solubilizing microbes.

As a result of the production of a clearing zone around the colony on plate indicative of the solubilization of magnesium trisilicate, 31% isolates of the library were able to perform silica solubilisation. Interestingly, the frequency of silica solubilization capacity among the strains obtained from Tepui was higher compared to that detected among the isolates from Porretta Terme. Indeed, around 44% of the Tepui isolates were active in silica solubilization, while only 22% of the strains from Porretta Terme cave had this capacity. This aspect can be related to the peculiar features of the two different cave environments where the strains were isolated. Tepui cave is a quartzitic and therefore silica-rich cave, while Porretta Terme is a carbonatic sulfidic cave.

Table 31. Isolates form orthoquartzitic Tepui caves able to solubilize silica and their collection site.		
Isolate	Cave niche	Location
CA001	water	Imawari Yeuta
CA002	water	Imawari Yeuta
CA010	water	Imawari Yeuta
CA011	water	Imawari Yeuta
CA012	water	Imawari Yeuta
CA013	roof	Imawari Yeuta
CA015	wall	Sima de la Lluvia
CA016	wall	Sima de la Lluvia
CA018	wall	Sima de la Lluvia
CA019	wall	Sima de la Lluvia
CA023	wall	Cueva Pristina
CA067	wall	Cueva Pristina
CA068	wall	Cueva Pristina
CA094	water	Imawari Yeuta
CA095	water	Imawari Yeuta
CA096	water	Imawari Yeuta
CA098	roof	Imawari Yeuta
CA104	floor	Sima Menor
CA105	floor	Sima Menor
CA110	wall	Imawari Yeuta
CA113	floor	Imawari Yeuta
CA114	floor	Imawari Yeuta
CA125	roof	Sima Menor
CA126	floor	Sima Menor
CA164	wall	Cueva Pristina

Interestingly, the solubilisation of magnesium trisilicate was mainly visible when the active isolates were grown on Glucose Agar plates, while we did not observe any clearing zone around bacterial and fungal colonies on ISP2 medium. This is probably linked to the lower number of nutrients and their defined nature (glucose is the only carbon source) present in Glucose Agar, while ISP2 medium is composed by complex carbon and nitrogen compounds. The high oligotrophy of the Glucose Agar might therefore be at the basis of the production of some metabolites involved in magnesium trisilicate, and more in general, silica solubilisation (Fig 56).

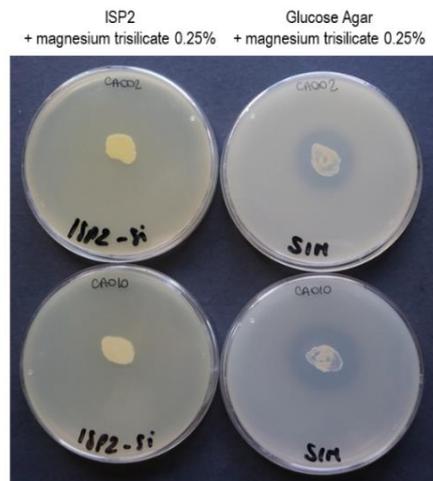
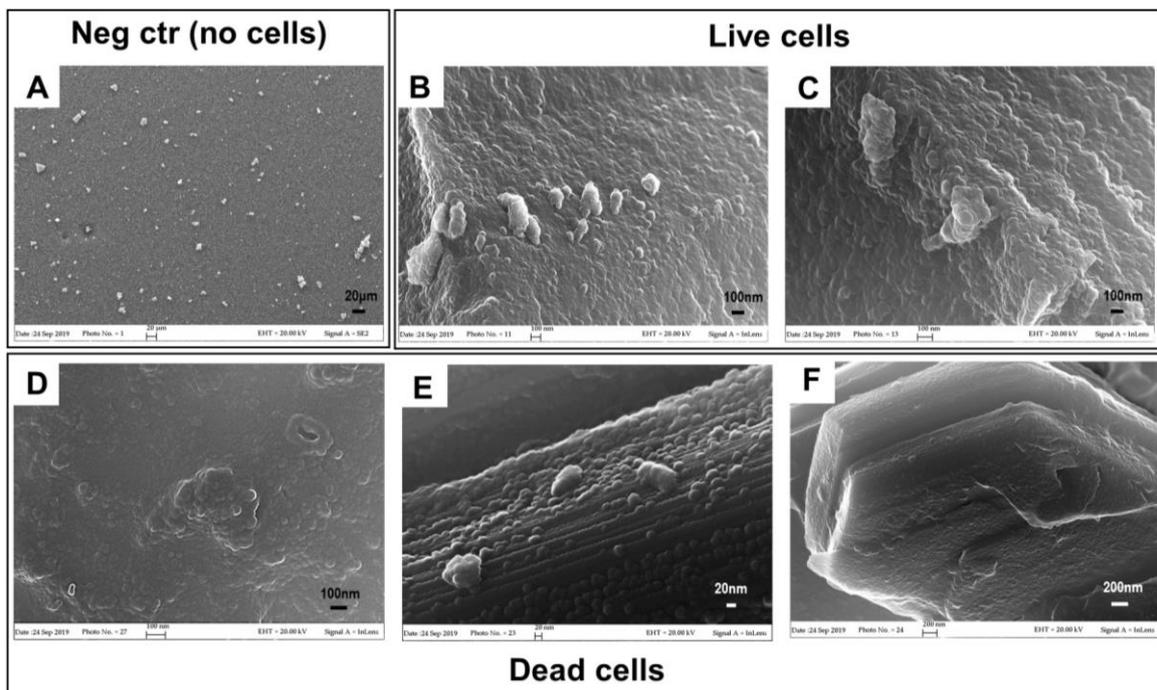


Fig 56. Example of activity of silica solubilizing bacteria (CA002 on the top and CA010 on the bottom of the picture) grown on ISP2 and Glucose Agar, with the addition of magnesium trisilicate (0.25%).

A high percentage of the silica solubilizing isolates from Tepui caves were collected from walls (on the basis of the plate assay), therefore we used wall samples to perform enrichment experiments. From the results shown in figure 57, differences between the inoculated cultures and the control ones (not inoculated with wall sample) are appreciable. Filters from the inoculated cultures with live cells resulted covered of globular silica structures (Fig 57 B and C) while the filter from control cultures shows silica in the form of small pieces of undefined shapes similar to glass splinters (Fig 57 A). The spheroid silica particles appear agglomerated towards coalescence and cover the microbial cells that are anchored to their surface (Fig 57 C). From the cultures with dead cells (autoclaved wall samples) (Fig 57 D-F), fingerprints of bacterial cell debris covered by silica can be observed (Fig 57 E). As respect to live cells treatment, the silica particles have crystalline forms presenting defined geometrical shapes (Fig 57 F).



**Fig 57.** Scanning electron microscopy images of the filters associated to the non-inoculated culture (A), the inoculated cultures with live cells (B and C) and the inoculated cultures with dead cells (D-F).

This preliminary experiment suggests that the presence of live cells influences silica weathering by forming highly organized cellular structures around globular shaped silica. The inoculation of dead cells also induced a certain level of silica amorphization although it seemed to remain at an earlier phase as compared to the live cells. Further studies will define the microbial role in silica mobilization and precipitation in continuous flow systems which are supposed to mimic better the dynamics of the cave environment.

## 7. Conclusions

Caves are hidden, mysterious and represent one of the most interesting formations on our planet, where microorganisms have adapted themselves according to cave conditions, including darkness and oligotrophy. Considering the complete isolation of deep unexplored caves from the Earth's surface, the investigation of the microbial diversity carries the potential for the discovery of rare microorganisms inhabiting these habitats. Microbial activities have been shown to contribute to the matter turnover in caves and to speleogenesis. In particular, effects of microorganisms on the mineral cave surfaces include degradation and corrosion events. Mechanical effects can occur due to secretion of exoenzymes, organic acids, and acidification of the extracellular environment due to the oxidation activities of chemical elements composing the mineral rocks. Microbial communities have also a role in the development of secondary mineral formations (speleothems) through the establishment of specific interactions with the mineral rocks, which induce mobilization and precipitation events. Besides the effects on cave formations by the microorganisms, the cave microbiology is important to understand how complex microbial communities evolved within a nutrient-limited environment. To survive in these nutrient-poor environments, microorganisms typically organize themselves in collective structures, offering cooperation and mutualistic relationships and producing, as results of their interaction, biosignatures that can be observed in caves. Within these syntrophic interactions, chemolithotrophic bacteria play a role as primary producers in the absence of light radiation and contribute to the maintenance of the biogeochemical cycles in the cave along with providing nutrients to heterotrophic microorganisms. Chemolithotrophic microbial activities typically involve the oxidation of methane, sulfide, inorganic hydrogen, nitrogen and minerals containing for instance manganese and iron. Additionally, microbial cave life can also depend on small inputs of organic carbon, transported into the underground through percolating waters, air circulation and fauna. However, the influx of organic carbon by these mechanisms is generally low and not a reliable source of energy, therefore most caves are considered oligotrophic. If on one hand cave microbes cooperate organizing themselves in complex biological structures, on the other hand, caves have a significant potential for exploring the microbial production of new antimicrobial substances and enzymes, in part due to the possible competition that can arise for the limited nutrient content.

The cave microbiology is still an understudied discipline and, during the last twenty years, has attracted research attention for evolutionary studies, biotechnological applications, and astrobiology.

In this PhD thesis the taxonomy and the metabolic potential of microbial communities inhabiting three different cave systems have been investigated. These caves are featured by different rock substrates, climatic conditions, type of nutrients and physicochemical parameters; therefore, they represent distinct and peculiar study models for the cave microbiology. The conclusions associated to each chapter can be summarized according to the results obtained with the different molecular, microbiological and geochemical approaches used in the study.

The conclusions of the study on the geomicrobiology of the orthoquartzite cave Imawari Yeuata are:

- The microbial community composition is mainly driven by the water activity. Other variables such as cave location, pH values and silica content are correlated at lower extent with the biodiversity detected in the different cave samples. On the basis of the water activity the samples could be grouped in three categories named “dry”, “moist” and “wet”. The moist samples showed the highest biodiversity and were featured by the presence of specific bacterial orders of *Alphaproteobacteria* (*Rhizobiales*) and *Acidobacteria* (*Solibacterales*, Subgroup 2). Dry and wet samples were featured by lower diversity and richness as compared to the moist group of samples. *Gammaproteobacteria* characterized the wet samples, whereas dry samples were significantly enriched in *Actinobacteria*. This can be associated to the specific capacity of *Actinobacteria* members to thrive under stress conditions (such as desiccation).
- Complex chemotrophic microbial communities colonize different cave niches (wall, water and floor) and they are most probably involved in creating the environmental conditions driving quartz dissolution through i) the increase of the amount of inorganic cations and metal ions in solution as a result of biomineralization processes; ii) the raise of pH mediated by microbial metabolisms such as nitrogen fixation, decomposition of proteins or amino acids, urea degradation, and CO<sub>2</sub> consumption. This study provides new insights into the relationship between microbial growth and silica mobilization and precipitation that is still unclear.
- The microbial communities associated to different silica amorphization steps in the formation of a new coralloid silica stromatolite are featured by peculiar microbial groups. In particular, starting from the most pristine sample (just eroded quartz) going to the most amorphous silica samples (coralloid silica) there is an enrichment of *Actinobacteria* members and the lowering of *Chloroflexi*. This can be associated to the possible activities that these microbial groups have in the silica mobilization and precipitation processes by increasing the dissolution of metal ions, increasing the pH, and providing nucleation sites. Moreover, sequencing data

- targeting 16S rRNA, *coxL* and *hypD* genes revealed that the most pristine sample was dominated by a novel taxon affiliated to the *Ktedonobacterales* order of *Chloroflexi* phylum.
- The analysis of metagenomes from cave wall revealed a high number of genes involved in the oxidative phosphorylation, indicating a key role of the aerobic metabolism for microbial growth on Imawarì Yeuta cave. Moreover, in relation to the carbon assimilation, a higher number of genes belonging to the reverse TCA cycle was detected as respect to the Calvin Benson Cycle. The nitrogen cycle was poorly represented in the wall samples genes. Although nitrogen fixation pathways seem to be absent, genes belonging to nitrification processes were detected.

The conclusions of the study on the geomicrobiology of the sulfidic Fetida cave influenced by seawater are:

- Each biofilm/deposit from Fetida Cave is characterized by specific microbial taxa, which are selectively enriched in each of the microbial communities growing as water filaments, vermiculation or in gypsum moonmilk deposits. This is related to the selection imposed by environmental factors linked to the cave environment in which each biofilm/deposit develops, which at various levels include the following aspects: the type of substrate (i.e. the cave wall/ceiling rock or cave water), the pH (extremely acidic in moonmilk, slightly acidic in vermiculations and neutral in water filaments), the amount of condensation on the wall/ceilings (i.e. high condensation for the moonmilk development), the amount and type of nutrient input (higher organic carbon is likely in the water stream which is constantly mixed with seawater entering from the coastline), the type of metal and mineral exposure (gypsum for moonmilk, Fe- or Al-containing minerals in vermiculations). In particular, the three deposits showed the presence of diverse chemolithotrophic bacterial and archaeal members which are able to utilize reduced and oxidized sulfur forms as energy sources.
- The investigation of metagenomes from the cave vermiculations revealed that *Gammaproteobacteria* and *Firmicutes* are the taxa mostly responsible of the sulfur and nitrogen cycles in Fetida. Fully represented dissimilatory and assimilatory sulfate reduction pathways characterize all the vermiculations, whereas the pathways of oxidation of reduced forms of sulfur are almost totally absent. Therefore, we suppose that most of the sulphide rising from the bottom of the cave waters is oxidized by microbial communities present in the white filaments floating and sedimented in the waters. The oxidized sulfur forms (volatile) are then, in turn, reduced by microbes composing the vermiculations covering the cave walls.

The conclusions of the study on the geomicrobiology of the high-altitude ice cave Cenote Abyss are:

- The 285 m deep Cenote Abyss is dominated by peculiar psychrophilic bacteria with a possible role in nitrogen cycle. In particular, the outermost location of the cave is inhabited by bacterial members that are involved in the production of reduced forms of nitrogen, including *Parafrigobacterium*, *Polaromonas* and *Pedobacter*. In the deepest locations of the cave, the microbial community shift towards the dominance of nitrifying bacteria including *Nitrospira*.

The conclusions of the study on the metabolic and phenotypic analysis of microbial isolated from cave systems are:

- The screening of a library of 168 microbial isolates derived from the different cave systems analysed in this study, revealed the potential of microorganisms growing in caves to produce antimicrobial molecules. The phenotypic characterization of the four most active isolates within the library allowed to hypothesize an association between the production of antimicrobial molecules and the capacity to metabolize specific carbon sources, including cellobiose and glucose. The chemical nature of the antimicrobials is under analysis.
- Around 33% of the library of cave isolates were able to solubilize silica as indicated by specific *in vitro* assays using solid and liquid cultures and SEM observations. This capacity was mainly observed under oligotrophic growth conditions, this being associated with possible specific metabolic pathways repressed or uninduced when complex and rich nutrients are present. These results provide the first evidence of the processes leading to the formation of silica-based speleothems in the orthoquartzite cave Imawari Yeuta under laboratory conditions.

## 8. Bibliography

1. Beale AD, Whitmore D, Moran D. Life in a dark biosphere: a review of circadian physiology in arrhythmic environments. *J Comp Physiol B Biochem Syst Environ Physiol*. 2016;186(8):947-968. doi:10.1007/s00360-016-1000-6
2. Northup DE, Barns SM, Yu LE, et al. Diverse microbial communities inhabiting ferromanganese deposits in Lechuguilla and Spider Caves. *Environ Microbiol*. 2003. doi:10.1046/j.1462-2920.2003.00500.x
3. Barton HA. Introduction to cave microbiology: A review for the non-specialist. *J Cave Karst Stud*. 2006;68(2):43-54.
4. Cappelletti M, Ghezzi D, Zannoni D, Capaccioni B, Fedi S. Diversity of Methane-Oxidizing Bacteria in Soils from “Hot Lands of Medolla” (Italy) Featured by Anomalous High-Temperatures and Biogenic CO<sub>2</sub> Emission. *Microbes Environ*. 2016;00(0):1-9. doi:10.1264/jsme2.ME16087
5. Léveillé RJ, Datta S. Lava tubes and basaltic caves as astrobiological targets on Earth and Mars: A review. *Planet Space Sci*. 2010. doi:10.1016/j.pss.2009.06.004
6. Westall F, Loizeau D, Foucher F, et al. Habitability on Mars from a microbial point of view. *Astrobiology*. 2013;13(9):887-897. doi:10.1089/ast.2013.1000
7. Gadd GM. Metals, minerals and microbes: Geomicrobiology and bioremediation. *Microbiology*. 2010;156(3):609-643. doi:10.1099/mic.0.037143-0
8. Tisato N, Torriani SFF, Monteux S, et al. Microbial mediation of complex subterranean mineral structures. *Sci Rep*. 2015;5:15525. doi:10.1038/srep15525
9. Sauro F, Lundberg J, De Waele J, Tisato N, Galli E. Speleogenesis and Speleothems of the Guacamaya Cave , Auyan Tepui , Venezuela. 16th International Congress of Speleology 2013.
10. Sauro F, Cappelletti M, Ghezzi D, et al. Microbial diversity and biosignatures of amorphous silica deposits in orthoquartzite caves. *SciRep*. 2018. doi:10.1038/s41598-018-35532-y
11. Riquelme C, Enes Dapkevicius M de L, Miller AZ, et al. Biotechnological potential of Actinobacteria from Canadian and Azorean volcanic caves. *Appl Microbiol Biotechnol*. 2017;101(2):843-857. doi:10.1007/s00253-016-7932-7
12. Ghosh S, Kuisiene N, Cheeptham N. The cave microbiome as a source for drug discovery: reality or pipe dream? *Biochem Pharmacol*. 2016. doi:10.1016/j.bcp.2016.11.018
13. Lavoie KH, Winter AS, Read KJH, Hughes EM, Spilde MN, Northup DE. Comparison of bacterial communities from lava cave microbial mats to overlying surface soils from Lava Beds National Monument, USA. *PLoS One*. 2017. doi:10.1371/journal.pone.0169339

14. Riquelme C, Rigal FO, Hathaway JJM, et al. Cave microbial community composition in oceanic islands: disentangling the effect of different colored mats in diversity patterns of Azorean lava caves. *FEMS Microbiol Ecol.* 2015;91(12). doi:10.1093/femsec/fiv141
15. Axenov-Gribanov DV, Voytsekhovskaya IV, Tokovenko BT, et al. Actinobacteria isolated from an underground lake and moonmilk speleothem from the biggest conglomeratic karstic cave in Siberia as sources of novel biologically active compounds. *PLoS One.* 2016;11(2). doi:10.1371/journal.pone.0149216
16. Tomova I, Lazarkevich I, Tomova A, Kambourova M, Vasileva-Tonkova E. Diversity and biosynthetic potential of culturable aerobic heterotrophic bacteria isolated from Magura Cave, Bulgaria. *Int J Speleol.* 2013. doi:10.5038/1827-806X.42.1.8
17. Engel AS. Microbial diversity of cave ecosystems. *Geomicrobiol Mol Environ Perspect.* 2010:219-238. doi:10.1007/978-90-481-9204-5\_10
18. Tebo BM, Davis RE, Anitori RP, Connell LB, Schiffman P, Staudigel H. Microbial communities in dark oligotrophic volcanic ice cave ecosystems of Mt. Erebus, Antarctica. *Front Microbiol.* 2015. doi:10.3389/fmicb.2015.00179
19. D'Angeli IM, Carbone C, Nagostinis C, et al. New insights on secondary minerals from Italian sulfuric acid caves. *Int J Speleol.* 2018;47(3):271-291. doi:10.5038/1827-806X.47.3.2175
20. D'Angeli IM, Ghezzi D, Leuko S, et al. Geomicrobiology of a seawater-influenced active sulfuric acid cave. *PLoS One.* 2019. doi:10.1371/journal.pone.0220706
21. Lynch RC, King AJ, Farías ME, Sowell P, Vitry C, Schmidt SK. The potential for microbial life in the highest-elevation (>6000 m.a.s.l.) mineral soils of the Atacama region. *J Geophys Res.* 2012;117. doi:10.1029/2012JG001961
22. Yang B, Wang Y, Qian P-Y. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. 2016. doi:10.1186/s12859-016-0992-y
23. Caporaso JG, Lauber CL, Walters WA, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA.* 2011;108:4516-4522. doi:10.1073/pnas.1000080107
24. Bolyen E, Ram Rideout J, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol.* 2019;37:852–857. <https://doi.org/10.1038/s41587-019-0209-9>
25. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013. doi:10.1093/nar/gks1219
26. Wray RA, Sauro F. An updated global review of solutional weathering processes and forms in

- quartz sandstones and quartzites. *Earth-Science Reviews*. 2017. doi:10.1016/j.earscirev.2017.06.008
27. Aubrecht R, Lánczos T, Šmída B, et al. Venezuelan sandstone caves: a new view on their genesis, hydrogeology and speleothems. *Geol Croat*. 2008;61(2-3):345-362. <http://hrcak.srce.hr/30674>.
28. Barton HA, Giarrizzo JG, Suarez P, et al. Microbial diversity in a Venezuelan orthoquartzite cave is dominated by the Chloroflexi (Class Ktedonobacterales) and Thaumarchaeota Group I.1c. *Front Microbiol*. 2014;5. doi:10.3389/fmicb.2014.00615
29. Handley KM, Turner SJ, Campbell KA, Mountain BW. Silicifying Biofilm Exopolymers on a Hot-Spring Microstromatolite: Templating Nanometer-Thick Laminae. *Astrobiology*. 2008. doi:10.1089/ast.2007.0172
30. Aubrecht R, Brewer-Carías C, Šmída B, Audy M, Kováčik E. Anatomy of biologically mediated opal speleothems in the World's largest sandstone cave: Cueva Charles Brewer, Chimantá Plateau, Venezuela. *Sediment Geol*. 2008;203(3-4):181-195. doi:10.1016/j.sedgeo.2007.10.005
31. Ruff SW, Farmer JD. Silica deposits on Mars with features resembling hot spring biosignatures at El Tatio in Chile. *Nat Commun*. 2016;7:13554. doi:10.1038/ncomms13554
32. Rice MS, Bell JF, Cloutis EA, et al. Silica-rich deposits and hydrated minerals at Gusev Crater, Mars: Vis-NIR spectral characterization and regional mapping. *Icarus*. 2010. doi:10.1016/j.icarus.2009.03.035
33. McKay CP, Smith HD. Possibilities for methanogenic life in liquid methane on the surface of Titan. *Icarus*. 2005;178(1):274-276. doi:10.1016/j.icarus.2005.05.018
34. Miller CS, Baker BJ, Thomas BC, Singer SW, Banfield JF. EMIRGE: Reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol*. 2011. doi:10.1186/gb-2011-12-5-r44
35. Miller CS, Handley KM, Wrighton KC, Frischkorn KR, Thomas BC, Banfield JF. Short-Read Assembly of Full-Length 16S Amplicons Reveals Bacterial Diversity in Subsurface Sediments. *PLoS One*. 2013;8(2). doi:10.1371/journal.pone.0056018
36. Turner S, Pryer KM, Miao VPW, Palmer JD. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J Eukaryot Microbiol*. 1999;46(4):327-338. doi:10.1111/j.1550-7408.1999.tb04612.x
37. Pausan MR, Csorba C, Singer G, et al. Measuring the archaeome: detection and quantification of archaea signatures in the human body. bioRxiv 334748. doi:10.1101/334748
38. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for

- phylogenetic study. *J Bacteriol.* 1991;173(2):697–703. doi:10.1128/jb.173.2.697-703
39. Weber CF, King GM. Distribution and diversity of carbon monoxide-oxidizing bacteria and bulk bacterial communities across a succession gradient on a Hawaiian volcanic deposit. *Environ Microbiol.* 2010. doi:10.1111/j.1462-2920.2010.02190.x
40. Beimgraben C, Gutekunst K, Opitz F, Appel J. HypD as a marker for [NiFe]-hydrogenases in microbial communities of surface waters. *Appl Environ Microbiol.* 2014;80(12):3776-3782. doi:10.1128/AEM.00690-14
41. Barton HA, Taylor MR, Pace NR. Molecular phylogenetic analysis of a bacterial community in an oligotrophic cave environment. *Geomicrobiol J.* 2004;21(1):11-20. doi:10.1080/01490450490253428
42. Barton HA, Jurado V. What's Up Down There? Microbial Diversity in Caves. *Microbe.* 2007;2:132-138
43. Lynch RC, Darcy JL, Kane NC, Nemergut DR, Schmidt SK. Metagenomic evidence for metabolism of trace atmospheric gases by high-elevation desert actinobacteria. *Front Microbiol.* 2014. doi:10.3389/fmicb.2014.00698
44. Quaiser A, Ochsenreiter T, Lanz C, et al. Acidobacteria form a coherent but highly diverse group within the bacterial domain: Evidence from environmental genomics. *Mol Microbiol.* 2003;50(2):563-575. doi:10.1046/j.1365-2958.2003.03707.x
45. Ortiz M, Neilson JW, Nelson WM, et al. Profiling Bacterial Diversity and Taxonomic Composition on Speleothem Surfaces in Kartchner Caverns, AZ. *Microb Ecol.* 2013. doi:10.1007/s00248-012-0143-6
46. Nishiyama M, Sugita R, Otsuka S, Senoo K. Community structure of bacteria on different types of mineral particles in a sandy soil. *Soil Sci Plant Nutr.* 2012. doi:10.1080/00380768.2012.729226
47. De Mandal S, Senthil Kumar N. Metagenomic analysis of bacterial community composition among the cave sediments of Indo-Burman biodiversity hotspot region. *PeerJ Preprints.* 2014. doi:10.7287/peerj.preprints.631v1
48. Velikonja BH, Tkavc R, Pašić L. Diversity of cultivable bacteria involved in the formation of macroscopic microbial colonies (cave silver) on the walls of a cave in Slovenia. *Int J Speleol.* 2014. doi:10.5038/1827-806X.43.1.5
49. Mohammadipanah F, Wink J. Actinobacteria from arid and desert habitats: Diversity and biological activity. *Front Microbiol.* 2016. doi:10.3389/fmicb.2015.01541
50. Diaz-Herraiz M, Jurado V, Cuezva S, et al. Deterioration of an Etruscan tomb by bacteria from the order Rhizobiales. *Sci Rep.* 2014;4:3610. doi:10.1038/srep03610

51. Miller AZ, Pereira MFC, Calaforra JM, Forti P, Dionísio A, Saiz-Jimenez C. Siliceous Speleothems and Associated Microbe-Mineral Interactions from Ana Heva Lava Tube in Easter Island (Chile). *Geomicrobiol J*. 2014. doi:10.1080/01490451.2013.827762
52. Valdes N, Soto P, Cottet L, et al. Draft genome sequence of *Janthinobacterium lividum* strain MTR reveals its mechanism of capnophilic behavior. *Stand Genomic Sci*. 2015;10:110. doi:10.1186/s40793-015-0104-z
53. Liu L, Liu Y, Zhang P, et al. Development of bacterial communities in biological soil crusts along a revegetation chronosequence in the Tengger Desert, northwest China. *Biogeosciences*. 2017. doi:10.5194/bg-14-3801-2017
54. Wong FKY, Lacap DC, Lau MCY, Aitchison JC, Cowan DA, Pointing SB. Hypolithic Microbial Community of Quartz Pavement in the High-Altitude Tundra of Central Tibet. *Microb Ecol*. 2010. doi:10.1007/s00248-010-9653-2
55. Bak EN, Larsen MG, Moeller R, et al. Silicates eroded under simulated martian conditions effectively kill bacteria-A challenge for life on mars. *Front Microbiol*. 2017. doi:10.3389/fmicb.2017.01709
56. Toporski JKW, Steele A, Westall F, Thomas-Keprta KL, McKay DS. The Simulated Silicification of Bacteria— New Clues to the Modes and Timing of Bacterial Preservation and Implications for the Search for Extraterrestrial Microfossils. *Astrobiology*. 2002. doi:10.1089/153110702753621312
57. Mecchia M, Sauro F, Piccini L, et al. Geochemistry of surface and subsurface waters in quartz-sandstones: Significance for the geomorphic evolution of tepui table mountains (Gran Sabana, Venezuela). *J Hydrol*. 2014;511:117-138. doi:10.1016/j.jhydrol.2014.01.029
58. De Mandal S, Chatterjee R, Kumar NS. Dominant bacterial phyla in caves and their predicted functional roles in C and N cycle. *BMC Microbiol*. 2017. doi:10.1186/s12866-017-1002-x
59. Macalady JL, Jones DS, Lyon EH. Extremely acidic, pendulous cave wall biofilms from the Frasassi cave system, Italy. *Environ Microbiol*. 2007. doi:10.1111/j.1462-2920.2007.01256.x
60. Teresa González-Muñoz M, Fernández-Luque B, Martínez-Ruiz F, et al. Precipitation of Barite by *Myxococcus xanthus*: Possible Implications for the Biogeochemical Cycle of Barium. *Appl Environ Microbiol*. 2003;69(9):5722-5725. doi:10.1128/AEM.69.9.5722-5725.2003
61. Kodama Y, Watanabe K. Isolation and Characterization of a Sulfur-Oxidizing Chemolithotroph Growing on Crude Oil under Anaerobic Conditions. *Appl Environ Microbiol*. 2003;69(1):107-112. doi:10.1128/AEM.69.1.107-112.2003
62. Northup DE, Kathleen H. Lavoie D. Geomicrobiology of Caves: A Review. *Geomicrobiol J*.

- 2001;18(3):199-222. doi:10.1080/01490450152467750
63. Keren R, Mayzel B, Lavy A, et al. Sponge-associated bacteria mineralize arsenic and barium on intracellular vesicles. *Nat Commun.* 2017;8. doi:10.1038/ncomms14393
64. Carmichael MJ, Carmichael SK, Santelli CM, Strom A, Bräuer SL. Mn(II)-oxidizing bacteria are abundant and environmentally relevant members of ferromanganese deposits in caves of the Upper Tennessee River Basin. *Geomicrobiol J.* 2013;30(9):779-800. doi:10.1080/01490451.2013.769651
65. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 2017. doi:10.1038/ismej.2017.119
66. Schmidt SK, Gendron EMS, Vincent K, et al. Life at extreme elevations on Atacama volcanoes: the closest thing to Mars on Earth? *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology.* 2018.
67. Thiel V, Tank M, Bryant DA. Diversity of Chlorophototrophic Bacteria Revealed in the Omics Era. *Annu Rev Plant Biol.* 2018;58. doi:10.1146/annurev-arplant-042817
68. Hug LA, Castelle CJ, Wrighton KC, et al. Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome.* 2013;1(1):22. doi:10.1186/2049-2618-1-22
69. Cavaletti L, Monciardini P, Bamonte R, et al. New lineage of filamentous, spore-forming, gram-positive bacteria from soil. *Appl Environ Microbiol.* 2006. doi:10.1128/AEM.00132-06
70. Yabe S, Sakai Y, Abe K, Yokota A. Diversity of <i>Ktedonobacteria</i> with Actinomycetes-Like Morphology in Terrestrial Environments. *Microbes Environ.* 2017. doi:10.1264/jsme2.ME16144
71. Yokota A. Cultivation of Uncultured Bacteria of the Class Ktedonobacteria in the Phylum Chloroflexi. *Makara J Sci J Sci.* 2012;16(161):1-8. doi:10.7454/mss.v16i1.1273/Makara
72. Yarza P, Yilmaz P, Pruesse E, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol.* 2014;12(9):635-645. doi:10.1038/nrmicro3330
73. Peiffer JA, Spor A, Koren O, et al. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci U S A.* 2013;110(16):6548-6553. doi:10.1073/pnas.1302837110
74. Gordienko AS, Kurdish IK. Surface electrical properties of bacillus subtilis cells and the effect of interaction with silicon dioxide particles. *Biophysics (Oxf).* 2007;52(2):217-219. doi:10.1134/S0006350907020121
75. Jones AA, Bennett PC. Mineral Microniches Control the Diversity of Subsurface Microbial

- Populations. *Geomicrobiol J.* 2014. doi:10.1080/01490451.2013.809174
76. Cockell CS, Kelly LC, Marteinsson V. Actinobacteria-An Ancient Phylum Active in Volcanic Rock Weathering. *Geomicrobiol J.* 2013. doi:10.1080/01490451.2012.758196
77. Borsodi AK, Knáb M, Krett G, et al. Biofilm Bacterial Communities Inhabiting the Cave Walls of the Buda Thermal Karst System, Hungary. *Geomicrobiol J.* 2012;29(7):611-627. doi:10.1080/01490451.2011.602801
78. Mohanty A, Subhadarshane B, Barman P, Mahapatra C, Aishwarya B, Behera RK. Iron Mineralizing Bacterioferritin A from Mycobacterium tuberculosis Exhibits Unique Catalase-Dps-like Dual Activities. *Inorg Chem.* 2019;58(8):4741-4752. doi:10.1021/acs.inorgchem.8b02758
79. Cheeptham N, Sadoway T, Rule D, et al. Cure from the cave: Volcanic cave actinomycetes and their potential in drug discovery. *Int J Speleol.* 2013. doi:10.5038/1827-806X.42.1.5
80. Li J, Zhou H, Peng X, Wu Z, Chen S, Fang J. Microbial diversity and biomineralization in low-temperature hydrothermal iron-silica-rich precipitates of the Lau Basin hydrothermal field. *FEMS Microbiol Ecol.* 2012. doi:10.1111/j.1574-6941.2012.01367.x
81. Oliveira C, Gunderman L, Coles CA, et al. 16S rRNA gene-based metagenomic analysis of Ozark cave bacteria. *Diversity.* 2017. doi:10.3390/d9030031
82. King GM, Weber CF. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat Rev Microbiol.* 2007. doi:10.1038/nrmicro1595
83. Islam ZF, Cordero PRF, Feng J, et al. Two Chloroflexi classes independently evolved the ability to persist on atmospheric hydrogen and carbon monoxide. *ISME J.* 2019;13:1801-1813. doi:10.1038/s41396-019-0393-0
84. Ji M, Greening C, Vanwonterghem I, et al. Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature.* 2017. doi:10.1038/nature25014
85. Epelde L, Lanzén A, Blanco F, Urich T, Garbisu C. Adaptation of soil microbial community structure and function to chronic metal contamination at an abandoned Pb-Zn mine. *FEMS Microbiol Ecol.* 2015. doi:10.1093/femsec/fiu007
86. Minic Z, Thongbam PD. The Biological Deep Sea Hydrothermal Vent as a Model to Study Carbon Dioxide Capturing Enzymes. *Mar Drugs.* 2011. doi:10.3390/md9050719
87. Muchowska KB, Varma SJ, Chevallot-Beroux E, Lethuillier-Karl L, Li G, Moran J. Metals promote sequences of the reverse Krebs cycle. *Nat Ecol Evol.* 2017;1(11):1716-1721. doi:10.1038/s41559-017-0311-7
88. Ortiz M, Legatzki A, Neilson JW, et al. Making a living while starving in the dark: Metagenomic insights into the energy dynamics of a carbonate cave. *ISME J.* 2014.

- doi:10.1038/ismej.2013.159
89. Hamilton TL, Jones DS, Schaperdoth I, Macalady JL. Metagenomic insights into S(0) precipitation in a terrestrial subsurface lithoautotrophic ecosystem. *Front Microbiol.* 2014. doi:10.3389/fmicb.2014.00756
  90. Westall F, Foucher F, Bost N, et al. Biosignatures on Mars: What, Where, and How? Implications for the Search for Martian Life. *Astrobiology.* 2015;15(11):998-1029. doi:10.1089/ast.2015.1374
  91. Jones DS, Lyon EH, Macalady JL. Geomicrobiology of biovermiculations from the Frasassi cave system, Italy. *J Cave Karst Stud.* 2008;70(2):78-93.
  92. Northup DE, Dahm CN, Melim LA, et al. Evidence for geomicrobiological interactions in Guadalupe Caves. *J Cave Karst Stud.* 2000.
  93. Engel AS, Meisinger DB, Porter ML, et al. Linking phylogenetic and functional diversity to nutrient spiraling in microbial mats from Lower Kane Cave (USA). *ISME J.* 2010;4(1):98-110. doi:10.1038/ismej.2009.91
  94. Pojski N, Ozimec R, Mazija M, Cetkovi H, Luki L. Molecular Characterization of Aquatic Bacterial Communities in Dinaric Range Caves. *Water Environ Res.* 2016;88(7):617-630. doi:10.2175/106143016X14609975746488
  95. Airoidi L, Cinelli F. Variability of fluxes of particulate material in a submarine cave with chemolithoautotrophic inputs of organic carbon. *Mar Ecol Prog Ser.* 1996;139(1-3):205-217. doi:10.3354/meps139205
  96. Raddadi N, Giacomucci L, Totaro G, Fava F. *Marinobacter* sp. from marine sediments produce highly stable surface - active agents for combatting marine oil spills. *Microb Cell Fact.* 2017;1-13. doi:10.1186/s12934-017-0797-3
  97. Belkova NL, Tazaki K, Zakharova JR, Parfenova V V. Activity of bacteria in water of hot springs from Southern and Central Kamchatskaya geothermal provinces, Kamchatka Peninsula, Russia. *Microbiol Res.* 2007;162(2):99-107. doi:10.1016/j.micres.2006.01.006
  98. Pašić L, Kovče B, Sket B, Herzog-Velikonja B. Diversity of microbial communities colonizing the walls of a Karstic cave in Slovenia. *FEMS Microbiol Ecol.* 2010;71(1):50-60. doi:10.1111/j.1574-6941.2009.00789.x
  99. Emerson JB, Thomas BC, Alvarez W, Banfield JF. Metagenomic analysis of a high carbon dioxide subsurface microbial community populated by chemolithoautotrophs and bacteria and archaea from candidate phyla. *Environ Microbiol.* 2016. doi:10.1111/1462-2920.12817
  100. Jones DS, Schaperdoth I, Macalady JL. Biogeography of sulfur-oxidizing *Acidithiobacillus* populations in extremely acidic cave biofilms. *ISME J.* 2016;10(12):2879-2891.

doi:10.1038/ismej.2016.74

101. Koch H, van Kessel MAHJ, Lüscher S. Complete nitrification: insights into the ecophysiology of comammox *Nitrospira*. *Appl Microbiol Biotechnol*. 2019;103(1):177-189. doi:10.1007/s00253-018-9486-3
102. Tetu SG, Breakwell K, Elbourne LDH, Holmes AJ, Gillings MR, Paulsen IT. Life in the dark: metagenomic evidence that a microbial slime community is driven by inorganic nitrogen metabolism. *ISME J*. 2013;7(6):1227-1236. doi:10.1038/ismej.2013.14
103. Barton HA, Northup DE. Geomicrobiology in cave environments: Past, current and future perspectives. *J Cave Karst Stud*. 2007;69(1):163-178.
104. Galdenzi S, Maruoka T. Gypsum deposits in the Frasassi Caves, Central Italy. *J Cave Karst Stud*. 2003;65(2):111-125.
105. Golyshina OV, Tran H, Reva ON, et al. Metabolic and evolutionary patterns in the extremely acidophilic archaeon *Ferroplasma acidiphilum* YT. *Sci Rep*. 2017;7(1):1-12. doi:10.1038/s41598-017-03904-5
106. Jones DS. Geomicrobiology of highly acidic, pendulous biofilms ("snottites") from the Frasassi Caves, Italy. Senior Integrative Exercise. 2005.
107. Wasmund K, Mußmann M, Loy A. The life sulfuric: microbial ecology of sulfur cycling in marine sediments. *Environ Microbiol Rep*. 2017;9(4):323-344. doi:10.1111/1758-2229.12538
108. Sallam A, Steinbüchel A. *Clostridium sulfidigenes* sp. nov., a mesophilic, proteolytic, thiosulfate- and sulfur-reducing bacterium isolated from pond sediment. *Int J Syst Evol Microbiol*. 2009;59(7):1661-1665. doi:10.1099/ijs.0.004986-0
109. Desai MS, Assig K, Dattagupta S. Nitrogen fixation in distinct microbial niches within a chemoautotrophy-driven cave ecosystem. *ISME J*. 2013;7(12):2411-2423. doi:10.1038/ismej.2013.126
110. Paun VI, Icaza G, Lavin P, et al. Total and potentially active bacterial communities entrapped in a late glacial through holocene ice core from scarisoara ice Cave, Romania. *Front Microbiol*. 2019;10(MAY):1-19. doi:10.3389/fmicb.2019.01193
111. Garcia-Lopez E, Cid C. Glaciers and ice sheets as analog environments of potentially habitable icy worlds. *Front Microbiol*. 2017;8(JUL):1-13. doi:10.3389/fmicb.2017.01407
112. Voigt C, Marushchak ME, Lamprecht RE, et al. Increased nitrous oxide emissions from Arctic peatlands after permafrost thaw. *Proc Natl Acad Sci U S A*. 2017;114(24):6238-6243. doi:10.1073/pnas.1702902114
113. Hénault C, Grossel A, Mary B, Roussel M, LéOnard J. Nitrous Oxide Emission by Agricultural Soils: A Review of Spatial and Temporal Variability for Mitigation. *Pedosphere*.

- 2012;22(4):426-433. doi:10.1016/S1002-0160(12)60029-0
114. Margesin R, Collins T. Microbial ecology of the cryosphere (glacial and permafrost habitats): current knowledge. *Appl Microbiol Biotechnol*. 2019;103(6):2537-2549. doi:10.1007/s00253-019-09631-3
115. Itcus C, Pascu MD, Lavin P, Perşoiu A, Iancu L, Purcarea C. Bacterial and archaeal community structures in perennial cave ice. *Sci Rep*. 2018;8(1):1-14. doi:10.1038/s41598-018-34106-2
116. Hillebrand-voiculescu A, Itcus C, Ardelean I, et al. Searching for cold-adapted microorganisms in the underground glacier of Scarisoara Ice Cave, Romania. *Acta Carsologica*. 2014;43(2-3):319-329
117. Cook J, Edwards A, Takeuchi N, Irvine-Fynn T. Cryoconite: The dark biological secret of the cryosphere. *Prog Phys Geogr*. 2016. doi:10.1177/0309133315616574
118. Santagata T, Sauro F, Spötl C, et al. The Cenote Project: monitoring a high-altitude ice cave in the Dolomites, Italy. *Proc 17th Int Congr Speleol*. 2017:86-88.
119. Lynch MDJ, Neufeld JD. Ecology and exploration of the rare biosphere. *Nat Rev Microbiol*. 2015. doi:10.1038/nrmicro3400
120. Nash MV, Anesio AM, Barker G, et al. Metagenomic insights into diazotrophic communities across Arctic glacier forefields. *FEMS Microbiol Ecol*. 2018;94(9):1-12. doi:10.1093/femsec/fiy114
121. Wu X, Zhang W, Liu G, et al. Bacterial diversity in the foreland of the Tianshan No.1 glacier, China. *Environ Res Lett*. 2012;7(1). doi:10.1088/1748-9326/7/1/014038
122. Palomo A, Jane Fowler S, Gülay A, Rasmussen S, Sicheritz-Ponten T, Smets BF. Metagenomic analysis of rapid gravity sand filter microbial communities suggests novel physiology of *Nitrospira* spp. *ISME J*. 2016;10(11):2569-2581. doi:10.1038/ismej.2016.63
123. Rodriguez-Caballero A, Pijuan M. N<sub>2</sub>O and NO emissions from a partial nitrification sequencing batch reactor: Exploring dynamics, sources and minimization mechanisms. *Water Res*. 2013;47(9):3131-3140. doi:10.1016/j.watres.2013.03.019
124. Havig JR, Hamilton TL. Snow algae drive productivity and weathering at volcanic rock-hosted glaciers. *Geochim Cosmochim Acta*. 2019;247:220-242. doi:10.1016/j.gca.2018.12.024
125. Hell K, Edwards A, Zarsky J, et al. The dynamic bacterial communities of a melting High Arctic glacier snowpack. *ISME J*. 2013;7(9):1814-1826. doi:10.1038/ismej.2013.51
126. Mattes TE, Alexander AK, Richardson PM, et al. The genome of *Polaromonas* sp. strain JS666: Insights into the evolution of a hydrocarbon- and xenobiotic-degrading bacterium, and features of relevance to biotechnology. *Appl Environ Microbiol*. 2008;74(20):6405-6416. doi:10.1128/AEM.00197-08

127. Cavicchioli R, Charlton T, Ertan H, Omar SM, Siddiqui KS, Williams TJ. Biotechnological uses of enzymes from psychrophiles. *Microb Biotechnol.* 2011;4(4):449-460. doi:10.1111/j.1751-7915.2011.00258.x
128. Marcolefes E, Leung T, Okshevsky M, et al. Culture-Dependent Bioprospecting of Bacterial Isolates From the Canadian High Arctic Displaying Antibacterial Activity. *Front Microbiol.* 2019;10. doi:10.3389/fmicb.2019.01836
129. Ravindran S. Bacteria work together to survive Earth's depths. *Proc Natl Acad Sci U S A.* 2017;114(5):788-790. doi:10.1073/pnas.1621079114
130. White TJ, Bruns TD, Lee SB, Taylor JW. Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. PCR - Protocols and Applications - A Laboratory Manual. 1990.
131. Giordani B, Costantini PE, Fedi S, et al. Liposomes containing biosurfactants isolated from *Lactobacillus gasseri* exert antibiofilm activity against methicillin resistant *Staphylococcus aureus* strains. *Eur J Pharm Biopharm.* 2019;139:246-252. doi:10.1016/j.ejpb.2019.04.011
132. Asker D, Beppu T, Ueda K. Unique diversity of carotenoid-producing bacteria isolated from Misasa, a radioactive site in Japan. *Appl Microbiol Biotechnol.* 2007;77(2):383-392. doi:10.1007/s00253-007-1157-8
133. Nováková A. Microscopic fungi isolated from the Domica Cave system (Slovak Karst National Park, Slovakia). A review. *Int J Speleol.* 2009;38(1):71-82. doi:10.5038/1827-806X.38.1.8
134. Belyagoubi L, Belyagoubi-Benhammou N, Jurado V, et al. International Journal of Speleology Official Journal of Union Internationale de Spéléologie Antimicrobial activities of culturable microorganisms (actinomycetes and fungi) isolated from Chaabe Cave Antimicrobial activities of culturable microorganisms (actinomycetes and fungi) isolated from Chaabe Cave, Algeria. *Alger Int J Speleol.* 2018;47(2):189-199. doi:10.5038/1827-806X.47.2.2148
135. Uğraş S, Sezen K, Kati H, Demirbağ Z. Purification and characterization of an antibacterial substance produced by pest-originated *Serratia marcescens* Mm3. *Turkish J Biol.* 2014;38(2):177-184. doi:10.3906/biy-1305-46
136. Sawana A, Adeolu M, Gupta RS. Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: Proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. *Front Genet.* 2014;5:1-22. doi:10.3389/fgene.2014.00429
137. Depoorter E, Bull MJ, Peeters C, Coenye T, Vandamme P, Mahenthiralingam E. *Burkholderia*: an update on taxonomy and biotechnological potential as antibiotic producers. *Appl Microbiol Biotechnol.* 2016;100(12):5215-5229. doi:10.1007/s00253-016-7520-x

138. White DC, Suttont SD, Ringelberg DB. The genus *Sphingomonas*: Physiology and ecology. *Curr Opin Biotechnol*. 1996;7(3):301-306. doi:10.1016/S0958-1669(96)80034-6
139. Manteca Á, Yagüe P. *Streptomyces* differentiation in liquid cultures as a trigger of secondary metabolism. *Antibiotics*. 2018;7(2):1-13. doi:10.3390/antibiotics7020041
140. Synytsya A, Monkai J, Bleha R, et al. Antimicrobial activity of crude extracts prepared from fungal mycelia. *Asian Pac J Trop Biomed*. 2017;7(3):257-261. doi:10.1016/j.apjtb.2016.12.011
141. Gebreyohannes G, Nyerere A, Bii C, Sbhatu DB. Investigation of Antioxidant and Antimicrobial Activities of Different Extracts of *Auricularia* and *Termitomyces* Species of Mushrooms. *Sci World J*. 2019;2019:1-10. doi:10.1155/2019/7357048
142. Bernhards RC, Cote CK, Amemiya K, et al. Characterization of in vitro phenotypes of *Burkholderia pseudomallei* and *Burkholderia mallei* strains potentially associated with persistent infection in mice. *Arch Microbiol*. 2017;199(2):277-301. doi:10.1007/s00203-016-1303-8
143. Sánchez S, Chávez A, Forero A, et al. Carbon source regulation of antibiotic production. *J Antibiot (Tokyo)*. 2010;63(8):442-459. doi:10.1038/ja.2010.78
144. Singh LS, Sharma H, Talukdar NC. Production of potent antimicrobial agent by actinomycete, *Streptomyces sannanensis* strain SU118 isolated from phoomdi in Loktak Lake of Manipur, India. *BMC Microbiol*. 2014;14(1):1-13. doi:10.1186/s12866-014-0278-3
145. Trüper HG. Culture and isolation of phototrophic sulfur bacteria from the marine environment. *Helgoländer Wissenschaftliche Meeresuntersuchungen*. 1970. doi:10.1007/BF01609883
146. Lynd LR, Weimer PJ, van Zyl WH and Pretorius IS. Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiol Mol Biol Rev*. 2002;66(3):506. doi:10.1128/MMBR.66.3.506-577.2002
147. Lang S, Wullbrandt D. Rhamnose lipids - Biosynthesis, microbial production and application potential. *Appl Microbiol Biotechnol*. 1999;51(1):22-32. doi:10.1007/s002530051358
148. Verma S, Bhatia Y, Valappil SP, Roy I. A possible role of poly-3-hydroxybutyric acid in antibiotic production in *Streptomyces*. *Arch Microbiol*. 2002;179(1):66-69. doi:10.1007/s00203-002-0500-9

## **9. Acknowledgments**

My first acknowledgments go to my family that has always cheered every step of my education path with great enthusiasm. Special thanks to Dr. Martina Cappelletti for her constant and dedicated guidance she provided during my PhD program from my first experiment until the PhD thesis writing. I would like to express my thankfulness to Prof. Davide Zannoni for his useful suggestions and advices. I am grateful to Dr. Francesco Sauro for the essential geological indications and teaching, inspiring collaboration, and proactive attitude. Further, I would like to acknowledge him for his key role in the PhD project funding. Additionally, I would like to thank all the former and present MAM (Molecular and Applied Microbiology) lab members. A special mention to my friend and colleague Paolo Costantini for his precious experimental support and spontaneous collaborative spirit. My heartfelt thanks also go to Dr. Stefano Fedi, Dr. Federica Sandri, Dr. Roberto Borghese and Dr. Andrea Firrincieli. Special thanks to the Master's degree students Andrea Gozzi and Lisa Foschi, the Bachelor's degree students Alessandro Caprini, Lorenzo Di Pietro and Giulia Cattabriga for their precious help in the experimental work. Additionally, I would like to thank Dr. Ilenia Maria D'Angeli, Prof. Jo De Waele and Dr. Andrea Columbu for their help in the geological data interpretation and for the cave expeditions for samples collection.

I would like to thank the Rector Prof. Francesco Ubertini, the Vice-Rector for Research Prof. A. Rotolo and the Governing Academic Bodies of the University of Bologna (UNIBO) for the financial support of the PhD scholarship and research project. My gratitude also goes to Prof. Peiyong Hong for her suggestions and indications for Illumina MiSeq data interpretation during my research period in the Division of Biological and Environmental Sciences and Engineering at the King Abdullah University of Science and Technology (Thuwal, Saudi Arabia) (Oct-Dec 2016). I would also like to thank Prof. Rob Capon, Dr. Zeinab Khalil and Ahmed Elbanna for their help and support during my research period in the Institute for Molecular Bioscience at the University of Queensland (Brisbane, Australia) (Sep-Nov 2017), where I started the isolation and analysis of cave microbial strains reported in this thesis. Further, I would like to acknowledge the agencies/associations and people involved in granting the permit for the speleological expeditions and samples collection: Instituto Nacional de Parques and the patronage of the Government of Bolivar State from Venezuela, the Embassy of the Bolivarian Republic of Venezuela in Italy, the Italian Speleological Society, Dr. Hosam Zowawi, Freddy Vergara, Santa Cesarea Terme spa and the Office Parks of the Autonomous Province of Bolzano. Many thanks to La Venta speleological association and Miles Beyond exploring teams for the organization of the expeditions to Venezuelan tepui and to Cenote Abyss in Cima Cunturines, Prof. Enrico Dinelli for XRF analyses at UNIBO, Prof. Cristina Carbone and Laura

Negretti for the SEM analysis at UNIGE. Funding for part of the microbial analyses was provided by Europlanet 2020 15-EPN1-029, 17-EPN3-021, 18-EPN5-031, 18-EPN4-088 grants. Europlanet 2020 RI has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654208. Within the Europlanet projects, I am grateful to Prof. Christine Moissl-Eichinger, Dr. Kaisa Koskinen and Dr. Alexander Mahnert at the Medical University of Graz (Austria), for their help with sample preparation for Illumina sequencing and data analysis with QIIME2 software. I would like to thank Dr. Stefano Donadio, Dr. Marianna Iorio and Dr. Sonia Maffioli from NAICONs srl (Milano) for their support in the chemical analyses of bioactive molecules. Finally, my acknowledgements go to all the researchers and authors here not mentioned, who contributed to the published manuscripts including part of the results discussed in this PhD thesis: Sauro et al. (2018) *Sci Rep* vol.8(1):17569, D'Angeli et al. (2019) *PLoS ONE* vol.14(8):e0220706.