



## OPEN ACCESS

## EDITED BY

Virginia Helena Albarracín,  
CONICET Center for Electron Microscopy  
(CIME), Argentina

## REVIEWED BY

Stephen Brian Pointing,  
National University of Singapore, Singapore  
Pedro Antonio Galleguillos,  
Centro de Investigación Científico y  
Tecnológico para la Minería (CICITEM), Chile

## \*CORRESPONDENCE

Kateřina Němečková  
✉ katerina.nemeckova@natur.cuni.cz

RECEIVED 27 February 2023

ACCEPTED 15 May 2023

PUBLISHED 06 July 2023

## CITATION

Němečková K, Mareš J, Procházková L,  
Culka A, Košek F, Wierzchos J, Nedbalová L,  
Dudák J, Tymlová V, Žemlička J, Kust A, Zima J,  
Nováková E and Jehlička J (2023) Gypsum  
endolithic phototrophs under moderate  
climate (Southern Sicily): their diversity and  
pigment composition.

*Front. Microbiol.* 14:1175066.

doi: 10.3389/fmicb.2023.1175066

## COPYRIGHT

© 2023 Němečková, Mareš, Procházková,  
Culka, Košek, Wierzchos, Nedbalová, Dudák,  
Tymlová, Žemlička, Kust, Zima, Nováková and  
Jehlička. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Gypsum endolithic phototrophs under moderate climate (Southern Sicily): their diversity and pigment composition

Kateřina Němečková<sup>1\*</sup>, Jan Mareš<sup>2,3</sup>, Lenka Procházková<sup>4</sup>,  
Adam Culka<sup>1</sup>, Filip Košek<sup>1</sup>, Jacek Wierzchos<sup>5</sup>, Linda Nedbalová<sup>4</sup>,  
Jan Dudák<sup>6</sup>, Veronika Tymlová<sup>6</sup>, Jan Žemlička<sup>6</sup>, Andreja Kust<sup>7,8</sup>,  
Jan Zima<sup>9</sup>, Eva Nováková<sup>9</sup> and Jan Jehlička<sup>1</sup>

<sup>1</sup>Institute of Geochemistry, Mineralogy and Mineral Resources, Faculty of Science, Charles University, Prague, Czechia, <sup>2</sup>Institute of Hydrobiology, Biology Centre of the Czech Academy of Sciences, České Budějovice, Czechia, <sup>3</sup>Center Algatech, Institute of Microbiology, The Czech Academy of Sciences, Třeboň, Czechia, <sup>4</sup>Department of Ecology, Faculty of Science, Charles University, Prague, Czechia, <sup>5</sup>Department of Biochemistry and Microbial Ecology, Museo Nacional de Ciencias Naturales - Consejo Superior de Investigaciones Científicas, Madrid, Spain, <sup>6</sup>Institute of Experimental and Applied Physics, Czech Technical University in Prague, Prague, Czechia, <sup>7</sup>Department of Earth and Planetary Science, University of Berkeley, Berkeley, CA, United States, <sup>8</sup>Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA, United States, <sup>9</sup>Department of Parasitology, Faculty of Science, University of South Bohemia, České Budějovice, Czechia

In this study, we used microscopic, spectroscopic, and molecular analysis to characterize endolithic colonization in gypsum (selenites and white crystalline gypsum) from several sites in Sicily. Our results showed that the dominant microorganisms in these environments are cyanobacteria, including: *Chroococidiopsis* sp., *Gloeocapsopsis pleurocapsoides*, *Gloeocapsa compacta*, and *Nostoc* sp., as well as orange pigmented green microalgae from the *Stephanosphaeria* clade. Single cell and filament sequencing coupled with 16S rRNA amplicon metagenomic profiling provided new insights into the phylogenetic and taxonomic diversity of the endolithic cyanobacteria. These organisms form differently pigmented zones within the gypsum. Our metagenomic profiling also showed differences in the taxonomic composition of endoliths in different gypsum varieties. Raman spectroscopy revealed that carotenoids were the most common pigments present in the samples. Other pigments such as gloeocapsin and scytonemin were also detected in the near-surface areas, suggesting that they play a significant role in the biology of endoliths in this environment. These pigments can be used as biomarkers for basic taxonomic identification, especially in case of cyanobacteria. The findings of this study provide new insights into the diversity and distribution of phototrophic microorganisms and their pigments in gypsum in Southern Sicily. Furthermore, this study highlights the complex nature of endolithic ecosystems and the effects of gypsum varieties on these communities, providing additional information on the general bioreceptivity of these environments.

## KEYWORDS

endoliths, gypsum, phototrophs, cyanobacteria, green algae, metagenomics

## 1. Introduction

Rock substrates are used as habitats by various microorganisms all over the globe (Gorbushina, 2007). These organisms are known as lithobionts and can be divided into several groups—epiliths (inhabiting the surface of rocks), endoliths (inhabiting inner parts of rocks close to the rock's surface), hypoliths (inhabiting spaces below rocks) and hypendoliths (inhabiting inner part of the rocks, close to the rock's bottom). Endoliths can further be divided into chasmoendoliths (colonizing rock fissures and cracks) and cryptoendoliths (colonizing the pores within rocks). Lithobionts represent a diverse group of organisms—microalgae, lichens, fungi, cyanobacteria, heterotrophic bacteria, and archaea (Golubic et al., 1981). These rock-dwelling communities are significant in many research areas such as the ecology of extreme environments (Friedmann, 1982), astrobiology (Vitek et al., 2014a), cultural heritage (Sanmartín et al., 2018), or even radioactive waste disposal (Pitonzo et al., 1999).

Generally, endolithic environments provide protection against changing exterior conditions such as solar radiation, temperature, water availability, etc. (Golubic et al., 1981; Walker and Pace, 2007). Despite that, not all rock-substrates are suitable for lithic growth. The capacity to support microbial growth is determined by its bioreceptivity, i.e., the ability of substrates to be colonized. Bioreceptivity is influenced by both intrinsic (roughness, pore size distribution, mineral composition and transparency, among other properties) and extrinsic factors (climatic conditions, availability of nutrients, composition of lithobiontic communities, synergism etc.) (Walker and Pace, 2007). Sedimentary rocks (e.g., sandstones (Archer et al., 2017), limestones (Wong et al., 2010), halite (Wierzchos et al., 2006) and gypsum (Culka et al., 2014)) are among the most common lithic substrates. Nevertheless, pyroclastic, magmatic or metamorphic rocks are suitable lithic habitats as well (e.g., ignimbrite or granites) (Crits-Christoph et al., 2016; Vitek et al., 2017).

Cyanobacteria are frequently found in these lithic habitats (primarily dominated by *Chroococcidiopsis* spp.), and thus the rock-substrates must allow sufficient penetration of light for photosynthesis (Walker and Pace, 2007). In many cases, photosynthetic and UV-screening pigments (e.g., carotenoids and scytonemin) produced by microbial communities dominated by cyanobacteria manifest themselves as colorful spots or lines on rocks; such phenomenon being known as “Tintenstriche” (Whitton, 2012). These has been found in various regions around the Earth (e.g., in the Atacama desert, Crits-Christoph et al., 2016 and Dry Valleys of Antarctica, Archer et al., 2017).

Exact phylogenetic determinations of the cyanobacterial and algal species in endolithic communities can be problematic, as the morphological differences are often insignificant. Currently, the best option for identification is a polyphasic approach. This approach is based on molecular criteria (e.g., 16S rRNA/18S rDNA sequencing) complemented by other features (e.g., morphology, ecology, and ecophysiology) (Komárek, 2016; Nedbalová et al., 2017; Procházková et al., 2019). For microbial profiling of the environmental samples, culture-independent methods such as metagenomics are the most suitable ones. For example, Meslier et al. (2018) used 16S rDNA amplicon high-throughput sequencing to analyze endolithic communities, and showed variances of the relative abundances in different substrates (e.g., granite, and gypsum).

Our study area, the southern region of Sicily, is characterized by easily accessible gypsum outcrops, which were formed during the Messinian salinity crisis (Basilone, 2018). The Messinian salinity crisis, also known as the Messinian event, was a geological period approximately 5 million years ago when the Mediterranean Sea went through a period of severe drying and desiccation (Basilone, 2018). The crisis had a significant impact on the geology of the region, leading to the formation of extensive evaporite deposits, including gypsum, which are still found in Sicily today. The Messinian evaporitic succession is mainly found in the central and southern parts of Sicily.

Relevant to our study, gypsum occurs as crystals of different habitus, e.g., tabular, prismatic, saber-shaped crystals. Based on their morphology and physical properties, the different varieties of gypsum can be distinguished. Transparent gypsum crystals with perfect cleavage are known as selenites. Alabaster is a massive and fine-grained gypsum variety. In scanning electron microscopy (SEM) images, selenitic crystals are uniform in shape (of different sizes), while alabastrine crystals are irregular-subhedral (Al-Wakeel, 2002).

Our previous study on gypsum endoliths from the Torre Salsa seaside in Southern Sicily, already had revealed the presence of various cyanobacteria in investigated stony specimens (e.g., *Gloeocapsa compacta* and *Nostoc* sp.) (Jehlička et al., 2019). The analysis of different layers of colonized gypsum using Raman microspectrometry revealed presence of the different UV-screening pigments such as scytonemin and gloeocapsin in those colonizations closer to the surface; while colonizations below these layers contained only carotenoids. Other photosynthesizing pigments (e.g., chlorophylls or phycobiliproteins) were also detected (Němečková et al., 2020, 2022).

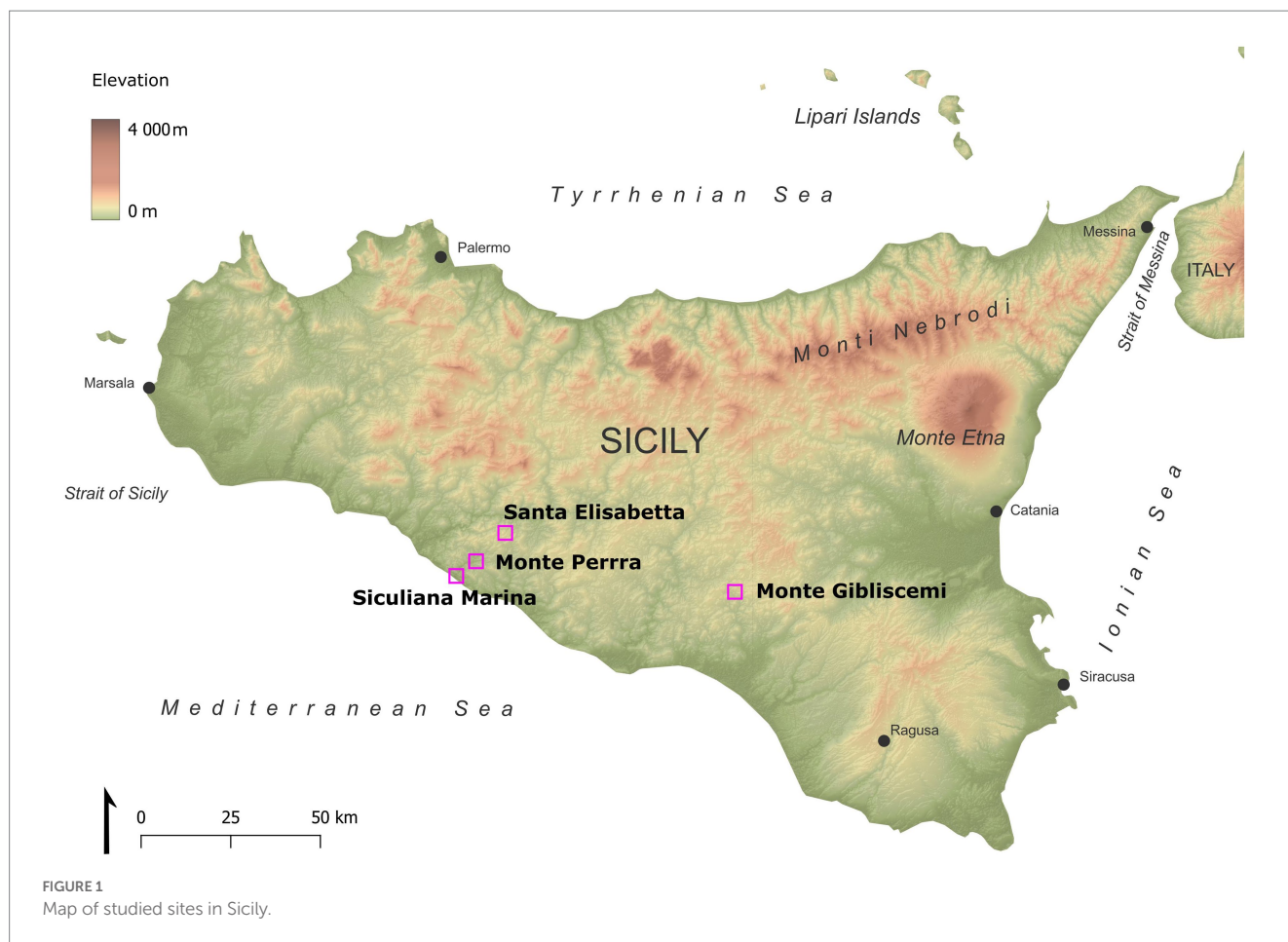
Here, we used molecular methods together with Raman microspectrometry and fluorescence microscopy in order to investigate the taxonomic composition of the endolithic communities (with a major focus on phototrophs) within the context of gypsum habitats at four sites in semi-arid Sicily: Siculiana Marina, Santa Elisabetta, Monte Perrera and Monte Gibliscemi. We also studied changes in the diversity within gypsum with different physical features. To our knowledge, this is the first study dealing with the microbial profiling of endoliths in varieties of the same substrate.

## 2. Methods

### 2.1. Sampling sites

The sites investigated were located in the Caltanissetta Basin in Sicily (Figure 1 and Table 1). Geologically, this area is formed by a sedimentary series, which include evaporites. The Caltanissetta Basin is composed of three stratigraphic units: salt deposits, a lower gypsum unit, and an upper gypsum unit (García-Veigas et al., 2018). The lower gypsum unit is further divided into primary and resedimented gypsum. The primary gypsum includes selenites, organic-rich shales, and dolomite rocks, while the resedimented component consists of halite, brecciated limestone, gypsum cumulate deposits, allochthonous selenites, and clastic gypsum. The upper gypsum unit is made up of gypsum interbedded with clastics (Caruso et al., 2015).

Different types of gypsum were investigated from the lower gypsum unit area, including transparent honey-colored selenitic crystals, white crystalline gypsum crystals, and crystalline gypsum aggregates. The samples were obtained from outcrops or fallen blocks (Figure 2).



Collected colonized gypsum samples (at least 5 pieces) were sealed in polyethylene bags and stored at room temperature in a dark and dry location until further analysis. The samples were collected in February of 2019. In southern part of Sicily, the long-term average temperature in February is  $\sim 13^{\circ}\text{C}$ , average daylight  $\sim 10\text{h}$ , rainfall ranging from 30 mm to 60 mm and the relative humidity is around 71%.

## 2.2. Micro-CT scan tomography

Six representative samples were chosen to describe the internal textures in detail using X-ray micro-computed tomography (micro-CT scan). The samples were chips or fragments and the dimensions did not exceed 2 cm in their largest dimension.

The micro-CT scan measurements were carried out using a custom-built cone-beam micro-CT scan system, operated in a laboratory of the Institute of Experimental and Applied Physics, Czech Technical University in Prague. The scanner used was equipped with a micro-focus X-ray tube (Hamamatsu L12531), and a CMOS-based flat-panel detector type (Dexela 1512). The tube is characterized by a tungsten transmission target, accelerating voltage up to 110 kVp, 16 W maximal power output, and  $2\ \mu\text{m}$  focal spot (considering a 2 W output power). The detector is equipped with a  $200\ \mu\text{m}$  thick micro-columnar CsI scintillation crystal and provides an array of  $1,944 \times 1,536$  pixels with a  $74.8\ \mu\text{m}$  pixel pitch. Further, the system provides a remote-control multi-axial positioning system, allowing one to accurately set the needed imaging geometry and a precise sample-rotation stage.

The exposition parameters varied based on the actual dimensions of each individual sample. The voxel size of all of the datasets was within the  $7.7\text{--}13\ \mu\text{m}$  range. The X-ray tube voltage was set to 80 or 90 kVp and was filtered by 2 mm of aluminum and 0.05 mm of copper. The CT scans were carried out with an angle step of  $0.15^{\circ}$ , resulting in 2400 projections per each dataset. The CT scan reconstruction was performed using filtered back-projection in the dedicated module of VG Studio MAX software (Volume Graphics, Heidelberg, Germany). Data analyses included denoising, ring artifact suppression, segmentation, and subsequent porosity calculation, and was carried out using Dragonfly software (Object Research System, Montreal, Canada). The histogram thresholding and deep learning extension (semantic segmentation models) were used to detect the boundaries of the gypsum fragments in 2D projections and to identify voxels representing empty volumes in the samples. The sum of empty-labeled voxels was used for the determination of the percentage of the apparent porous space in whole fragments or specific zones.

## 2.3. Raman spectroscopy

Selected cells (or colonies) were isolated from the gypsum crystals and put in a drop of water on a microscopic slide. In the case of cyanobacteria, the spectra were collected from these samples using a Renishaw inVia Reflex microspectrometer coupled with a Peltier-cooled CCD detector. Excitation was provided at 514.5 nm and 785 nm (power  $\sim 0.1\text{--}5\ \text{mW}$ ). Three spectra were recorded in

TABLE 1 Sampling sites and their basic characteristics.

Site characterization				Micro-CT scans characterization			
Site name		GPS	Characterization	Lithology	Dimensions (mm)	Weathered state	Void space (~porosity)
Siculiana Marina (SM)	1S	37.3388700 N, 13.3967000 E	Sites 1S and 1A are located near road to Siculiana Marina. Samples were collected from gypsum at western side of the outcrop.	Flat “honey-color” selenitic crystals (~5–10 cm in length) in aggregates	20 × 10 × 5	Almost uncorroded with secondary crust	0.8% (uncorroded part) 17.5% (crust)
	1A			White crystalline crystals (~10 cm in length) in aggregates	20 × 15 × 8	Almost uncorroded with secondary crust	1.2% (uncorroded part) 15.1% (crust)
Santa Elisabetta (SE)	2	37.4472489 N, 13.5474772 E	Gypsum outcrops at western side of Santa Elisabetta Mountain	Flat “honey-color” selenitic crystals (~15–20 cm in length) in aggregates	25 × 13 × 4	Uncorroded	1.8%
Monte Perrera (MP)	3	37.3807706 N, 13.4527739 E	Gypsum outcrops at south-east side of Monte Perrera	White hard coarse-grain aggregate gypsum crust with fine-grained matrix underneath, with presence of cracks	10 × 10 × 5	Partially corroded with secondary crust	14.8% (corroded part) 23.4% (crust)
					15 × 10 × 10	Deeply corroded with secondary crust	16.4%
Monte Gibliscemi (MG)	4	37.2094297 N, 14.2623708 E	Blocks of gypsum located at western side of Monte Gibliscemi	Coarse-grain aggregate	23 × 10 × 10	Uncorroded	8.1%

the 100–2,000  $\text{cm}^{-1}$  range with a spectral resolution of 1  $\text{cm}^{-1}$  full width at half maximum (CCD array detector/CCD chip with 576 × 400 pixels). A single Raman spectrum was collected using 5–10 scans each of 10 s exposure for an improved signal-to-noise ratio. The laser spot size was approx. 1.5  $\mu\text{m}$  in diameter when focused at the surface. By using a 50× objective (NA 0.75), it was possible to analyze single endolithic cells and colonies. A benzonitrile standard was used for spectral wavenumber calibration.

Endolithic coccoid algae from the Monte Gibliscemi site were analyzed using a Thermo Scientific DXR Raman microscope with an Olympus microscope. The spectrometer was calibrated by a software-controlled calibration procedure—using multiple neon emission lines (wavelengths calibration), and multiple polystyrene bands (laser frequency calibration). Spectra were obtained using 532 nm laser (power of 0.1–1 mW, 1,800 line/mm grating). At least 3 spectra were recorded using a high-resolution grating (range of 100–1,860  $\text{cm}^{-1}$ ) with a spectral resolution of 2  $\text{cm}^{-1}$ . To improve the signal-to-noise ratio, Raman spectra were collected using 64 scans and 2 s exposure. The laser spot size was ~2  $\mu\text{m}$  in diameter when focused on the surface. Endolithic colonies were analyzed using a 50× objective.

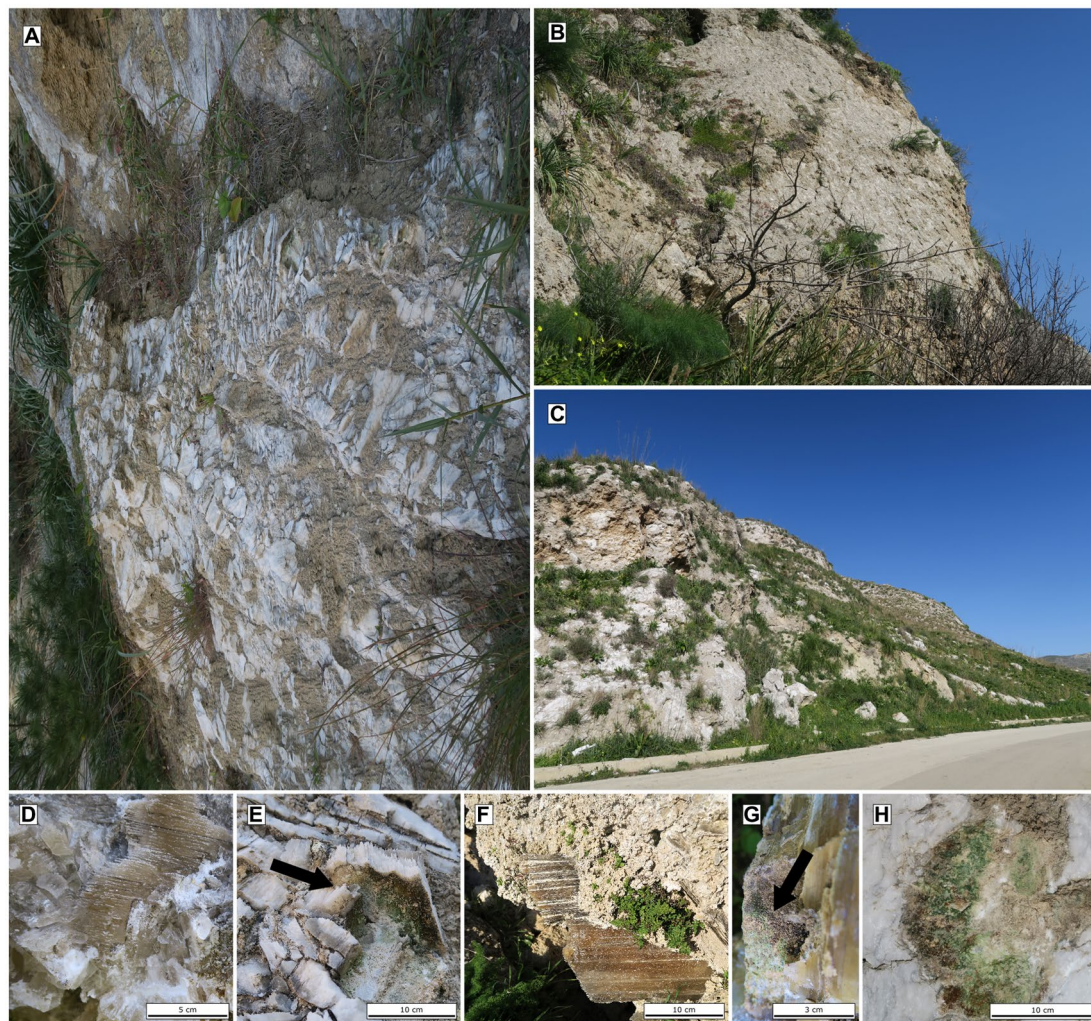
All Raman spectra obtained were exported into the Galactic \*.SPC format. The spectra were viewed and evaluated using the GRAMS AI spectroscopy software suite (v. 8.0, Thermo Electron Corp., Waltham, MA, United States). The wavenumber positions of the Raman bands were determined using the peak viewing settings and the positions corresponding to the maxima of the peaks.

## 2.4. Fluorescence microscopy

Small pieces of gypsum, colonized by pigmented endoliths were scraped and suspended in double-distilled water. The suspension was stained with SYBR Green I (SBI) (Molecular Probes), which is a fluorochrome specifically used for the staining of nucleic acids. Observations were made first in differential interference contrast (DIC) using a Zeiss AXIO Imager M2 fluorescence microscope (Carl Zeiss, Jena, Germany) plus an Apochrome x60, n = 1.4 Zeiss oil-immersion objective. A CCD AxioCam HRc Rev. 2 camera and AXIOVISION 4.7 software (Carl Zeiss, Oberkochen, Germany) were used to capture and record the DIC images. Images were acquired using a Multichannel Image Acquisition system, employing an eGFP filter set (Zeiss Filter Set 38; Ex/Em: 450–490/500–550 nm), a DAPI filter set (Zeiss Filter Set 49; Ex/Em: 365/420–470 nm), and a Rhodamine filter set (Zeiss Filter Set 20; Ex/Em: 540–552/567–647 nm).

## 2.5. Light microscopy

At least five pieces of colonized gypsum from each site were observed under the light microscope. The biomass (pigmented zones of gypsum) was removed from the stone surface/subsurface area using a sterile laboratory needle, and mechanically crushed on a microscopy slide in a drop of sterile water. If possible, the various pigmented layers were analyzed separately. Samples were observed using an Olympus BX 51 light microscope, equipped with DIC optics, and images of the



**FIGURE 2**  
 Images of gypsum outcrops at Siciliana Marina (A), Santa Elisabetta (B) and Monte Perrera (C). Detailed images of different gypsum habitats: selenites (D) and white crystalline-gypsum (E) at Siciliana Marina; selenites at Santa Elisabetta (F,G); gypsum crust at Monte Perrera (H).

microalgae were acquired with Olympus DP-72 digital camera using Olympus cell Sens Standard v. 2.1 image analysis software. Cyanobacterial taxa were identified according to Komárek and Anagnostidis (1999, 2005) and Komárek (2013). The relative abundance of individual taxa was estimated and expressed on a semiquantitative scale: dominant species ( $\geq 50\%$  of estimated biomass), frequent species (5–49% of the biomass), and scarce species ( $< 5\%$  of the biomass).

## 2.6. SSU rRNA gene barcoding of selected dominant phototrophs

Gypsum samples containing endoliths were selected based on species abundance and overall suitability for the single cell sequencing method (SCS) (e.g., no morphological overlap with other species in the sample) (Table 1).

### 2.6.1. Cyanobacteria

A modified isolation technique, originally developed for planktonic cyanobacteria, was used, applying a glass capillary under sterile conditions (Zapomělová et al., 2008; Mareš et al., 2015). Colonized areas

of the gypsum were scraped off and rehydrated on a microscopy slide in TE buffer (pH=8), and then mechanically crushed with another microscopy slide. The prepared specimen was observed under the Olympus IX71 inverted microscope and searched for single colonies of selected species. The single colony was separated by repeated washing through a series of droplets of TE buffer, and documented by microphotography. Subsequently, the isolated colony was either directly put into a sterile 0.2 mL Eppendorf tube for further processing (*Nostoc* sp., *Chroococidiopsis* sp.), or gently crushed using a clean cover slip, with five single cells of the target species collected per tube (*Gloeocapsa* spp., *Gloeocapsopsis pleurocapsoides*). After being kept frozen overnight at  $-20^{\circ}\text{C}$ , total genomic DNA from the isolated cells/colonies was amplified using the Multiple Displacement Amplification (MDA) method utilizing Repli-g Mini Kit (Qiagen). The MDA was performed following the manufacturer's instructions, which involved cell lysis, DNA amplification by Phi 29 polymerase (16 h at  $30^{\circ}\text{C}$ ), and a final denaturation step (5 min at  $65^{\circ}\text{C}$ ). The quality of MDA products was tested by gel electrophoresis, performed for 75 min at 80 V using 1.5% agarose gel. Successful MDA products were used as a template for PCR amplification of the 16S rRNA gene and the adjacent internal transcribed spacer (ITS) region using the following primer combinations (forwards/reverse): 16S378F/16S1494R,

and 16S27F/23S30R (Taton et al., 2003). The quality of the PCR products was tested by gel electrophoresis; performed for 30 min, at 90 V, using 1.5% agarose gel. PCR products were then purified using an Invisorb® Fragment CleanUp Kit (STRATEC Molecular GmbH, Germany) or a Monarch PCR and DNA Cleanup Kit (BioLabs® Inc., New England), and then sent for commercial Sanger sequencing at SeqMe (Dobříš, Czech Republic) using the same primers as for PCR, plus an additional internal sequencing primer (cyano6r) (Mareš et al., 2013).

### 2.6.1.1. Phylogenetic analysis

A set of SSU (16S) rRNA gene sequences for phylogenetic analysis was compiled from the data obtained in this current study, plus reference sequences from the GenBank database. The final data set included: (i) all sequences collected using the single cell/colony amplification protocol; (ii) the most abundant cyanobacterial amplicon sequence variants (ASVs) from the metagenomic analysis (ASVs reaching a relative abundance of at least 5% in at least one of the analyzed samples) (Benison and Karmanocky, 2014); frequent ASVs highly similar to the sequences of the dominant cyanobacterial morphotypes barcoded using single cell/colony sequencing (ASVs with pairwise sequence identity  $\geq 97\%$  to the single cell/colony sequences and a relative abundance  $\geq 1\%$  in at least one of the analysed samples); and (iv) closely related reference sequences identified by BLASTn against the NCBI non-redundant nucleotide database (using the single cell/colony sequences and abundant ASVs as queries). The resulting set of sequences was aligned using MAFFT v. 7 G-INS-I algorithm (Katoh and Standley, 2013), and subsequent manual removal of those regions with poor coverage. A maximum likelihood tree was reconstructed using RaxML v. 8 (Stamatakis, 2014) with 1,000 rapid bootstrap replications and applying the GTR + G + I substitution model.

### 2.6.1.2. Metagenomic profiling

For metagenomic analysis, altogether 26 samples from Sites 1 represented by selenitic (1S) and white crystalline-gypsum (1A) and Sites 2 represented by selenitic crystals were selected. Small parts of gypsum crystals exhibiting microbial pigmentation were firstly observed under the light microscope as described above and then scratched off with an ethanol-sterilized scalpel and needle into 2 mL Eppendorf tubes. Prior to DNA extraction, samples were homogenized in Eppendorf tubes using a Retsch MM 200 laboratory mill and 1 mm wolfram carbide beads shaken at  $30\text{ s}^{-1}$  frequency for 5 min. Lithobiontic DNA was extracted using the NucleoSpin® Soil Kit (Macherey-Nagel) with lysis buffer SL1 and Enhancer SX, in triplicates per sample and pooled before PCR.

A previously published in-house protocol (Brown et al., 2020) based on Earth Microbiome Project standards (EMP)<sup>1</sup> was employed for library construction and sequencing. Briefly, samples were amplified and barcoded using fused primers based on 515f and 926r, containing 12-bp Golay barcodes (Parada et al., 2016) and 5-bp barcodes, respectively, in a single PCR reaction (Marotz et al., 2019). The library was cleaned prior to the pooling of PCR products with AMPure XP (Beckman Coulter) magnetic beads; and finally purified using Pippin Prep (Sage Science) to remove non-target DNA, including the eukaryotic 18S rRNA gene PCR product. The library

was sequenced on Illumina MiSeq using v3 chemistry with  $2 \times 300$  nt output (Norwegian High Throughput Sequencing Centre and in-house MisSeq platform). The library contained two mock communities as the positive controls, i.e., one gDNA template with an equal composition of 10 bacterial species (ATCC® MSA-1000™), and one with a staggered composition of the same bacteria (ATCC® MSA-1001™). We have included three negative controls to account for contamination in the PCR library preparation (water template).

We obtained Illumina sequencing data with an average yield of 99654 pairs of reads per sample. Illumina 16S rRNA gene amplicons were initially processed as previously described (Brown et al., 2020) using scripts of USEARCH v. 9.2.64 (Edgar, 2013). The workflow included the initial quality-check (Andrews, 2010), demultiplexing, and primer trimming of the paired reads. The demultiplexed sequencing data with trimmed primers were loaded into R v. 4.1.0 (R Core Team, 2021), run under Rstudio version 1.4.1717 (Rstudio Team, 2021), and analysed following the DADA2 package pipeline v. 1.8 (Callahan et al., 2016) to infer a matrix of ASVs and their abundance in the individual samples. The DADA2 pipeline was run using the default parameters except for the filtering and trimming step, in which the parameters were adjusted to avoid unfavorable effects of any declining quality of the reverse reads by reducing their length [ $\text{truncLen} = c(210, 180)$ ] and a slight increase in their maximum expected error rate [ $\text{maxEE} = c(2, 5)$ ]. After the filtering procedures, removal of chimeras, and final merging of paired reads, 32.5% of the best quality reads (in average, 32380 pairs of reads per sample) were retained for further analysis. The hierarchical taxonomic classification of individual ASV was assigned based on SILVA database version 138.1 (Quast et al., 2012). The data set was subsequently imported into the *phyloseq* R package version 1.36.0 (McMurdie and Holmes, 2013), to export the taxonomy and the ASV abundance tables as well as a fasta file with the ASV sequences. The composition of the negative (blank) and positive control samples were checked manually. The control samples were then removed from further statistical analysis, and the data set was decontaminated by deleting minor amounts of contaminant bacteria found in the negative controls (using the *prune\_samples* and *prune\_taxa* functions in *phyloseq*). Chloroplast, mitochondrial, and eukaryotic sequences were also removed.

To visualize the relative abundance of higher taxa (bacterial phyla, cyanobacterial families) in individual samples using *phyloseq*, the data were normalized at median sequencing depth and rendered as plot bar graphs. To address the variability in sample composition data across defined categories (substrate, colored layer, locality), non-metric multidimensional scaling (NMDS) analysis was performed using R package *vegan* version 2.5.7 (Oksanen, 2009). The data was transformed into Bray-Curtis distances, and the statistical significance of differences between the categories was tested using permutational anova function *adonis* (999 permutations). Differences in alpha-diversity across the defined categories of samples were tested using pairwise t-tests (R package *stats* version) of diversity indices (Shannon, Simpson) calculated using the *diversity* function. The differences in the abundance of individual taxa across the defined categories of samples were tested using pairwise t-tests, and abundance data on individual taxa available as a *phyloseq* output.

## 2.6.2. Algae

Cocoid green algae were mostly present in orange subsurface zones of some gypsum samples. Endolithic algae were isolated from

1 <http://www.earthmicrobiome.org/protocols-and-standards/16s/>

crystalline gypsum rock from Monte Gibliscemi (Site 4). The crystalline gypsum from this site was selected due to the highly abundant and easily recognized orange-pigmented algal cells, without the presence of any other autotrophic species.

DNA was extracted from the orange colonization using the commercial kit DNeasy® PowerLyzer® PowerSoil® Kit (Qiagen) as described in the provided protocol. After being briefly vortexed again, the samples were centrifuged at 12,000 g for 2 min, the supernatant was used for testing DNA quality (isolate S14 Sicily) on NanoDrop®ND-1000 Spectrophotometer (NanoDrop Technologies, Inc.). The 18S rDNA region was then amplified by PCR using the forward primer FC (GGGAGGTAGTGACAAIAAATA), and the reverse primer RF (CCCGTGTGAGTCAAATTAAG) (Matsuzaki et al., 2015). Each 21.44 µL of PCR reaction for 18S rDNA amplification contained 2 µL of DNA isolates (diluted to a concentration of 5 ng µL<sup>-1</sup>), 4.32 µL 5x MyTaq Red Reaction Buffer (Bioline, Meridian Bioscience, United States), 0.43 µL of each 10 µM primer, 14.04 µL sterile Milli-Q water, and 0.22 µL of 5 U µL<sup>-1</sup> MyTaq HS Red DNA polymerase (Bioline, Meridian Bioscience, United States). Amplification reactions were performed using the following cycle parameters: for 3 min initial denaturation at 95°C, followed by 35 cycles (denaturation for 15 s at 95°C, annealing for 30 s at 56°C, extension for 40 s at 72°C), and final extension for 7 min at 72°C. The quality of the PCR products was tested by gel electrophoresis (45 min, 120 V, 1.5% agarose gel). PCR products were purified and sequenced using an Applied Biosystems automated sequencer (ABI 3730xl) at MacroGen Europe (Amsterdam, Netherlands). Chromatograms of the obtained sequences were examined using the FinchTV 1.4.0 (Geospiza, USA) program. Only sequences showing distinct single peaks in the electropherogram were used.

18S rDNA alignments were constructed for the phylogenetic analyses. The sequences were selected according to Gálvez et al. (2021) to encompass the closest relatives, to isolate S14 Sicily and other lineages in Chlamydomonadales. The 18S rDNA alignment contained 56 sequences (1,682 bp); the representatives of clade *Polytominia* and *Dunaliellinia* were selected as the outgroup. The best-fit nucleotide substitution model was estimated by jModeltest 2.0.1 (Posada, 2008). Based on the Akaike information criterion, the GTR + I + G model was selected for 18S rDNA. The phylogenetic trees were inferred by Bayesian inference (BI) and maximum likelihood (ML) according to Nedbalová et al. (2017), with a minor modification that the Markov chain Monte Carlo runs were carried out for 3 million generations in BI. Bootstrap analyses and Bayesian posterior probabilities were performed as described by Nedbalová et al. (2017).

All obtained sequences were compared to those in GenBank by BLAST and were submitted to the NIH genetic sequence database GenBank at NCBI during paper revision under accession numbers OQ509475 – OQ509403 and OQ520113.

### 3. Results

In this study, samples collected from four different sites in Sicily were initially analyzed by light microscopy. From these samples, the most abundant phototrophic endoliths were then sequenced and analysed using Raman spectroscopy. Based on the availability of samples with optimal biomass (i.e., at least 10–20 mg of dry mass) levels and different gypsum varieties present, two sites (Site 1 and 2)

have been selected for metagenomic analysis. This allowed us to gain insights into the microbial diversity and functional capabilities of the samples.

#### 3.1. Visual observation of colonized gypsum

In the field, gypsum rocks from 4 sites in Southern Sicily were evaluated for lithic colonization by visual examination of various coloration in fractured samples. The presence of colonization were also confirmed on the outcrops by the detection of biological pigments such as carotenoids by portable Raman spectrometer as previously described (Němečková et al., 2020).

At the **Siculiana Marina (Site 1)** two gypsum varieties were sampled—selenitic (1S) and white crystalline-gypsum (1A). Flat, thin, transparent selenitic crystal aggregates (each crystal was about 5–10 cm long and 3–5 cm wide) were found at Site 1 (Figure 2D). Aggregates were oriented in various directions on the outcrop. The inner parts of the crystals were slightly weathered, and after removing the crystals from the outcrop, a coarse-grained weathered matrix was revealed. Black and green pigmented endoliths were mainly observed in cleavage spaces of selenites and occasionally in coarse-grained matrix. The black-pigmented colonizations were always found in the uppermost areas closer to the sun-exposed zones while green colonizations were found below them. At the same outcrop, tabular white crystal aggregates were also observed (Figure 2E). Aggregates were around 10 cm long and wide, and around 5 cm thick. On the outcrop, aggregates were unordered and weathered in their inner parts, forming a fine-grained matrix beneath. The colonization had the same features as the colonization of selenitic crystals, i.e., the uppermost dark pigmented endoliths and inner green colonizations.

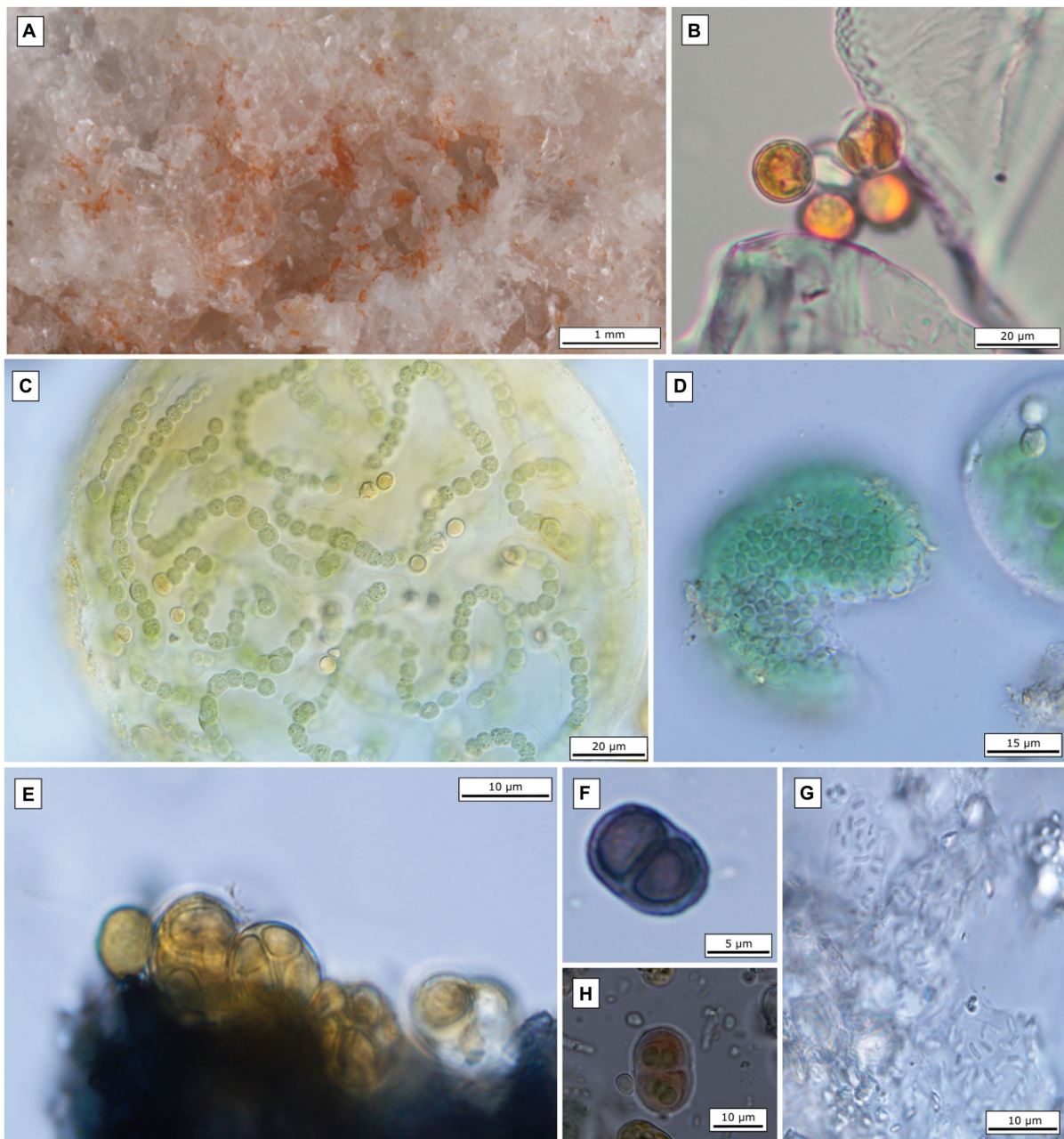
**Santa Elisabetta (Site 2)** contained flat, thin, transparent selenitic crystals (Figures 2F,G). Crystals were about 15–20 cm long and 5–10 cm wide. The character of the endolithic colonization was similar as in the case of 1S Site, but additional scarce orange zones were present between black and green colonizations.

**Monte Perrera (Site 3)** was characterized by a white and hard gypsum crust on south-east side of Monte Perrera Mountain (Figure 2H). In some areas, cracks colonized by dark endoliths were observed. Below the crust, endolithic colonization zones were in the fine-grained matrix underneath.

**Monte Gibliscemi (Site 4)** contained fallen blocks close to the western side of the mountain Monte Gibliscemi that were colonized by orange-pigmented endoliths found in the upper parts of fine-grained gypsum (Figure 3A).

#### 3.2. Micro-CT scans characterization of gypsum

To evaluate some of the textural features of selected specimens of gypsum, micro-CT scan analyses were conducted. Image analysis allowed us to investigate the three-dimensional inner architecture of gypsum fragments. A basic description of the observed characteristics and parameters is summarized in Table 2. Additional graphic materials of the mapped void space can be found in Supplementary material.



**FIGURE 3**

Microscopic images of the most abundant endolithic algae and cyanobacteria. (A) Gypsum crystals colonized by orange-pigmented algae at site 4; (B) microscopic image of coccoid algae isolated from the field gypsum sample (site 4); (C) *Nostoc* sp. (isolated from site Siculiana Marina); (D) *Chroococcidiopsis* sp. (isolated from site Siculiana Marina); (E) *Gloeocapsopsis pleurocapsoides* (isolated from site Monte Perrera); (F) *Gloeocapsa compacta* (isolated from site Monte Perrera); (G) *Gloeobacter violaceus* (isolated from site Monte Perrera); (H) *Gloeocapsa novacekii* (isolated from site Santa Elisabetta).

**Siculiana Marina (Site 1, 1S, 1A):** Gypsum from the Siculiana Marina site was composed of dense tabular crystals surrounded by secondary rims or crusts, which was observed for two gypsum samples (1A and 1S, Figures 4A,D). The crystals were predominantly broken by several cracks along the main cleavage plane, locally disturbed by cracks of different orientation (Figures 4B,C,E,F). These mostly linear features are generally open at the margins. Another system of small cracks and cavities that ignored the cleavage was also present, but these were predominantly isolated in the matrix, narrow,

and not easily accessible from the surface. The crystal surfaces were corroded, and shallow etch pits were visible mainly along the cleavage plane. However, in the case of the 1S sample, most of the corrosion marks are covered by secondary crusts. The calculated void space, including large cracks, was approximately 1.2 and 0.8% for the 1A and 1S samples. In contrast, the secondary crusts of the samples showed an entirely different, spongy-like texture that forms an open and highly interconnected porous space. The 1A sample was characteristic for the well-preserved secondary crust (Figures 4B,C),



TABLE 2 Relative abundance of cyanobacteria obtained by light microscopy.

Site	Layer	Overall relative abundance of observed phototrophs		
		< 5%	5–49%	> 50%
SM (1S)	Black	<i>Gloeobacter violaceus</i> , <i>Nostoc</i> sp., <i>Chroococcidiopsis</i> sp., <i>G. novacekii</i> , <i>Petalonema</i> sp., <i>Anathece</i> sp.	<i>Chroococcidiopsis</i> sp., <i>G. rupestris</i> , <i>Gloeocapsa</i> sp., <i>Gloeocapsopsis pleurocapsoides</i> *, <i>Chlorophyta</i>	<i>Nostoc</i> sp.*, <i>G. compacta</i> , <i>Petalonema</i> sp.
	Green	<i>Nostoc</i> sp., <i>Microcoleus</i> sp., <i>Anathece</i> sp., <i>Gloeocapsa novacekii</i> , <i>G. rupestris</i> , <i>G. compacta</i> , <i>Chroococcus</i> sp.	<i>Chroococcidiopsis</i> sp., <i>Chroococcus</i> sp., <i>Gloeocapsa</i> sp., <i>G. rupestris</i> , <i>G. novacekii</i> , <i>Chlorophyta</i>	<i>Chroococcidiopsis</i> sp.
SM (1A)	Black	<i>Petalonema</i> sp., <i>Chroococcidiopsis</i> sp., <i>Nostoc</i> sp., <i>G. novacekii</i> , <i>Gloeocapsa</i> sp.	<i>G. compacta</i> , <i>G. biformis</i> , <i>Gloeobacter violaceus</i> , <i>Gloeocapsa</i> sp., <i>G. rupestris</i> , <i>Chroococcidiopsis</i> sp., <i>Petalonema</i> sp., <i>G. pleurocapsoides</i>	
	Green	<i>G. compacta</i> , <i>Gloeocapsa</i> sp., <i>Nostoc</i> sp.	<i>Nostoc</i> sp.	<i>Chroococcidiopsis</i> sp.
SE (2)	Black	<i>Chroococcidiopsis</i> , <i>Chlorophyta</i>	<i>Gloeocapsa</i> sp., <i>Nostoc</i> sp., <i>G. biformis</i> , <i>G. rupestris</i> , <i>G. novacekii</i> *, <i>Chlorophyta</i>	<i>Gloeobacter violaceus</i> , <i>Chroococcidiopsis</i> sp.
	Green	<i>Petalonema</i> sp., <i>Chroococcidiopsis</i> sp.	<i>Chroococcidiopsis</i> sp., <i>Nostoc</i> sp., <i>Chlorophyta</i>	<i>Chroococcidiopsis</i> sp., <i>Gloeocapsa</i> sp., <i>Gloeobacter violaceus</i>
MP (3)	Black	<i>Nostoc</i> sp.	<i>G. pleurocapsoides</i>	<i>Gloeocapsa compacta</i> *
	Orange	<i>Gloeobacter violaceus</i>		
	Green	<i>G. pleurocapsoides</i>	<i>Chroococcidiopsis</i> sp.	
MG (4)	Orange	<i>Chlorophyta</i> *		

\*phototrophs selected for single cell sequencing.

where the pores occupied approximately 15.1% of the volume. The remaining secondary rim on the 1S sample was less than 1 mm thick but contained a similar porous space of 17.5%.

**Santa Elisabetta (Site 2, SE):** Micro-CT scan reconstruction of the SE sample revealed tabular well-developed crystals (Figure 4G) corresponding to the selenitic nature of gypsum found on this site. The investigated fragment was unweathered, but several linear fissures along the main cleavage plane were separating the individual flat crystals (Figures 4H,I). Occasionally, cracks of a different orientation were observed. This gypsum was predominantly dense, except for some domains where narrow cracks within the matrix could be observed. The void space occupied only 1.8% of the total volume of the fragment.

**Monte Perrera (Site 3, MP):** According to the micro-CT scan images, gypsum from the Monte Perrera site was represented by aggregates of oriented subhedral tabular-prismatic coarse crystals, which also were partially affected by corrosion, dissolution, and recrystallization. Micro-CT scan reconstruction of the MP1 sample showed that the gypsum grains are tightly packed (Supplementary Figure S1A), and three zones can be distinguished in the fragment (Supplementary Figures S1B,C). The first zone consisted of large gypsum grains with relatively narrow and tiny fissures, but with relatively large cavities. The second zone was represented by smaller elongated gypsum crystals, but, in contrast, they were frequently separated by more open fissures coupled with cavities, which were probably formed by dissolution of the gypsum. Together, the interstices formed approximately 14.8% of the void space in the crystalline part of the aggregate. The preserved remains of the secondary recrystallized gypsum crust represented the third zone. It was partially fractured with a network of small cracks, but it also contained large spheroidal cavities. These cavities are responsible for a relatively large volume of empty space in the crust (23.4%), although the cracks in the matrix are smaller

than in the crystalline portion. The CT scan images of the MP2 sample illustrated how the crystalline core was surrounded by the secondary gypsum crust (Supplementary Figure S1D). Individual subhedral crystals can still be observed in the core, but the shape of the grains and the interstices between them may suggest that the gypsum grains were deeply corroded. There was no distinct boundary, and the gypsum grains gradually transitioned into a fragmented and probably recrystallized mass, which also contained a system of small cracks and larger cavities (Figures 4E,F). The calculation performed for the part of the sample, where the crystalline core was intentionally excluded, showed that the crust matrix was relatively dense with approximately 16.4% of void space.

**Monte Gibliscemi (Site 4, MG):** The MG sample was composed of coarse subhedral tabular-prismatic crystals that were cemented by a crystalline matrix (Supplementary Figure S1G). The CT scans revealed fissures of different sizes between large gypsum crystals, and, in some cases, cracks along the cleavage plane could be found. A similar texture was also observed for the matrix, although on a different scale with smaller gypsum grains and subtle fissures (Supplementary Figures S1H,I). In general, this fragment is highly compacted, with interstices occupying 8.1% of the sample volume.

### 3.3. Bright field microscopy of endolithic phototrophs

**At Siculiana Marina (Site 1) and Santa Elisabetta (Site 2),** black pigmented colonies were composed predominantly of *Nostoc* sp. and/or *Gloeocapsa compacta* in those zones closer to the surface (distanced around 0.5–1 cm from the surface), while green colonization dominated by *Chroococcidiopsis* sp. was located underneath (Figure 3D). The black pigmented colonizations were more diverse,

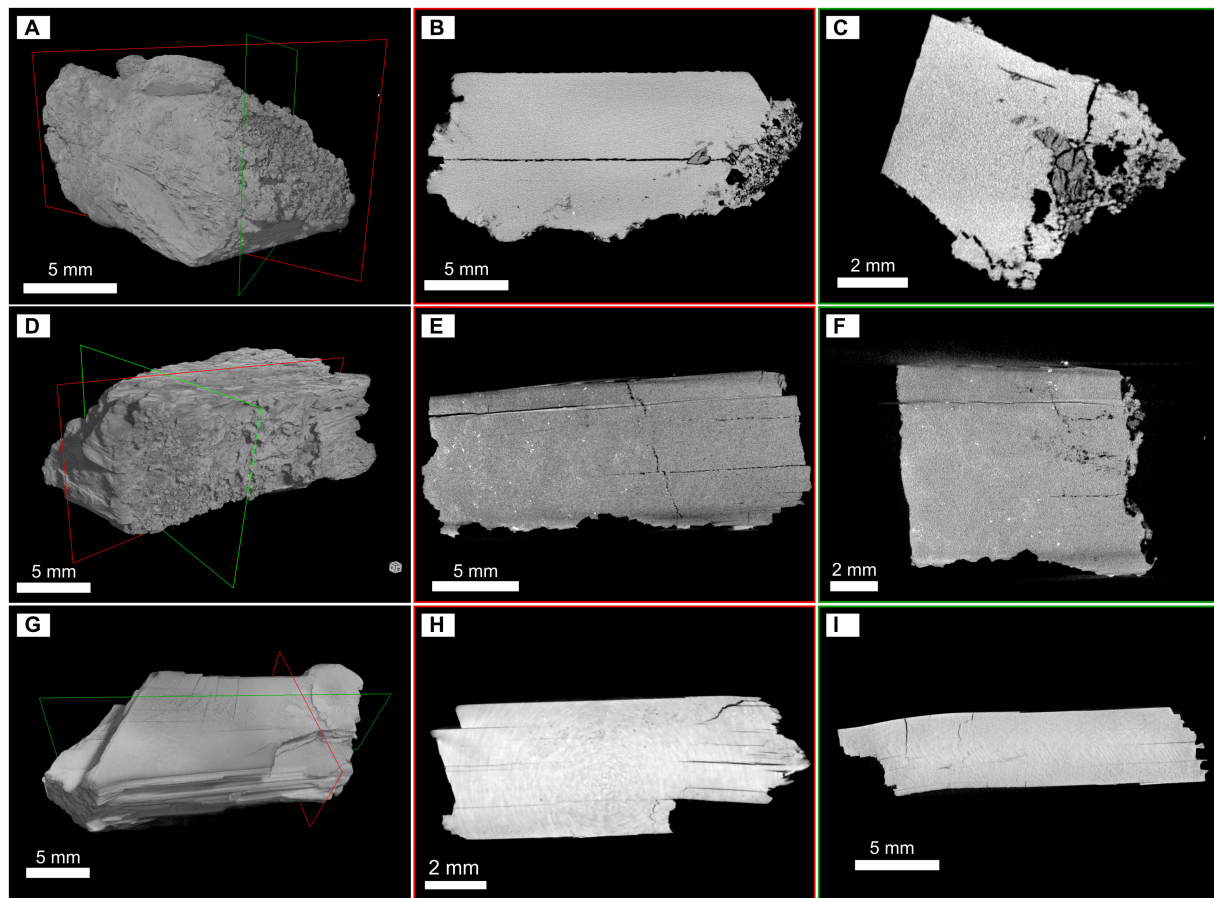


FIGURE 4

Micro-CT 3D reconstructions and selected slices illustrating the inner architecture and void space of gypsum fragments. Samples from Siculiana Marina—1A (A–C) and 1S (D–F); and from Santa Elisabetta (G–I). The red and green frames in the 3D images show the position of the displayed slices.

contained predominantly *G. compacta* and *Nostoc* sp. (but also other species—*G. biformis*, *G. novacekii*, *Chroococcidiopsis* sp. and *Gloeobacter violaceus*) (Figure 3). These black zones transformed diffusely into the inner green pigmented layer dominated by *Chroococcidiopsis* species.

**Monte Perrera site (Site 3)** showed a less diverse endolithic community. The uppermost black layer (around 1 cm from the surface) dominated by *G. compacta* was present. Below that, sporadic orange colonization of *G. violaceus* only, and inner green zones of predominated by *Chroococcidiopsis* sp. colonizations were recognised.

At **Monte Gibliscemi (Site 4)**, gypsum rocks were colonized by orange-pigmented coccoid cells of *Chlorophyta*, which were found in the upper parts of fine-grained gypsum without the visible presence of other endolithic organisms. The colonizations were around 0.5–4 cm from the surface (Figure 3B).

### 3.4. Raman microanalysis of endolithic pigments

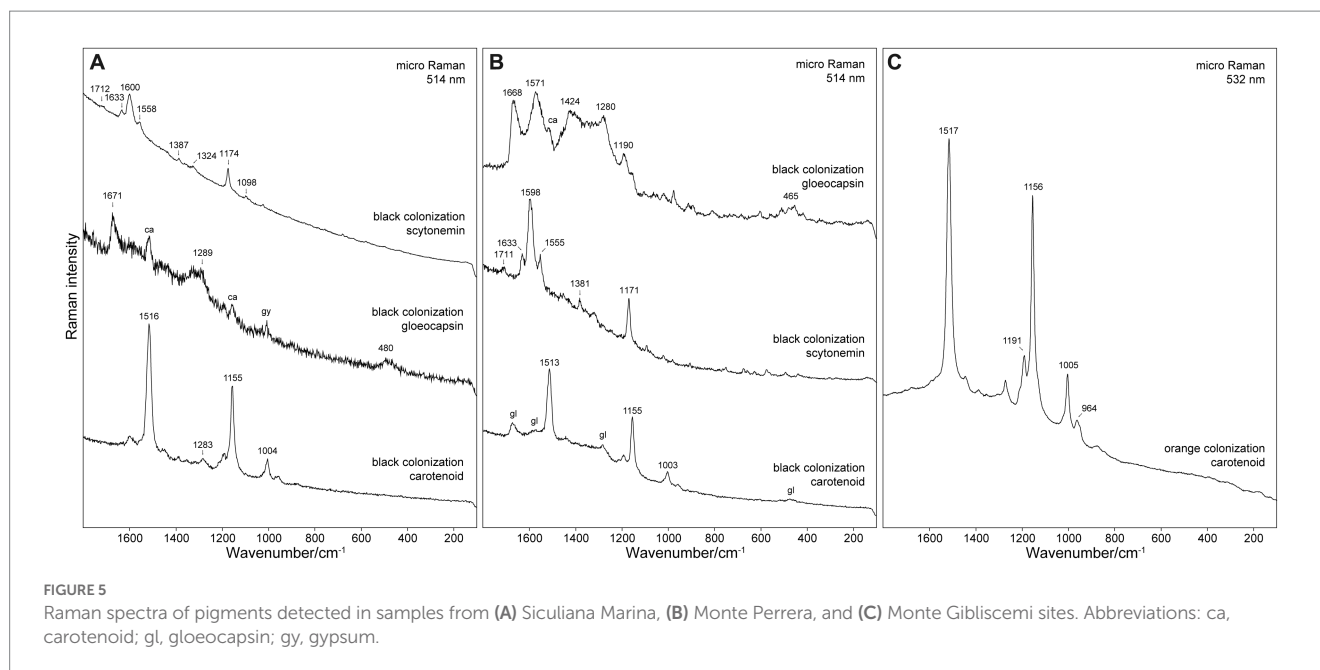
Raman spectroscopy allowed the detecting of the dominant pigments *in situ* in selected parts of the investigated specimens. In those samples in the uppermost dark pigmented layer, specialized pigments that protect against increased UV-radiation were detected.

As mentioned above, these upper layers were colonized mostly by *Nostoc* sp., *Gloeocapsa* sp. (especially *G. compacta*), and *Chroococcidiopsis* sp. Raman bands of detected pigments and their Raman spectra are shown in Figure 5, respectively.

Besides scytonemin (major pigment of *Nostoc* sp.) and gloeocapsin (major pigment of *Gloeocapsa* sp.), carotenoids were also detected in all analysed cell or colonies. As a matter of fact, it is very common that the Raman spectra of even the single cells contain the signal of carotenoids together with the signal of other pigments. This can be explained by the substantially increased sensitivity for the carotenoids due to resonance enhancement. The presented spectra of scytonemin and gloeocapsin in Figure 5 were selected as examples where the contribution of the Raman signal of carotenoids was minimal. Carotenoids were also detected in the green pigmented zones below. Carotenoids were the only protective pigments detected in the orange algae collected from Site 4. Raman bands of detected pigments can be found in Supplementary Table S1.

#### 3.4.1. Siculiana Marina (site 1)

At this site the major organism identified in the black layer was *Nostoc* sp., with the dominant pigment scytonemin. Detected Raman bands of scytonemin were 1,712 (weak intensity), 1,633 (medium), 1,600 (strong), 1,558 (m), 1,387 (w), 1,324 (w), 1,174 (ms), 1,098



(mw), 1,023 (w), 755 (w), and 681 (w)  $\text{cm}^{-1}$ . The most significant Raman band of scytonemin was detected at 1,600  $\text{cm}^{-1}$ , corresponding to the  $\nu(\text{CCH})$  aromatic ring quadrant stretching mode. The band at 1,558 can be assigned to the  $\nu(\text{CCH})$  p-disubstituted aromatic ring,  $\nu(\text{CCH})$  p-disubstituted aromatic ring, and the characteristic band at 1,174  $\text{cm}^{-1}$  of the  $\nu(\text{C}=\text{C}=\text{C})$  system (Edwards et al., 2000).

*Gloeocapsa compacta* was another major organism in the black layer with the dominant pigment gloeocapsin. Its Raman bands were detected at 1,671 (m), 1,289 (m), and 480 (mw)  $\text{cm}^{-1}$ . The analysis of *Gloeocapsopsis pleurocapsoides* cells revealed a strong carotenoid signal at 1,516 (vs), 1,449 (w), 1,388 (w), 1,283 (mw), 1,189 (mw), 1,155 (ms), 1,004 (m), and 957 (w)  $\text{cm}^{-1}$ . The band at 1,516  $\text{cm}^{-1}$  represents the C=C stretching mode of the conjugated chain in carotenoids. The band at 1,155  $\text{cm}^{-1}$  band represents the C-C stretching mode (coupled with C-H in-plane bending); and the band at 1,004  $\text{cm}^{-1}$  band is attributed to the C-CH<sub>3</sub> rocking mode.

### 3.4.2. Monte Perrera site (site 3)

*Gloeocapsa compacta* was the dominant species in the black layer, and the pigment gloeocapsin was detected as the major pigment of this cyanobacteria. The detected Raman bands were at 1,668 (s), 1,571 (s), 1,424 (ms), 1,340 (m), 1,280 (ms), 1,190 (m), and 465 (m)  $\text{cm}^{-1}$ . The bands around 1,668  $\text{cm}^{-1}$  can be assigned to  $\nu(\text{C}=\text{O})$  vibration, the bands around 1,571  $\text{cm}^{-1}$  to  $\nu(\text{C}=\text{C})$  aromatic ring vibration, and the bands around 1,276  $\text{cm}^{-1}$  to  $\delta(\text{CH}_2)$  vibration. The signal of carotenoid pigments was also detected in *G. compacta*, as corroborated by its Raman bands at: 1,513 (vs), 1,283 (w), 1,193 (m), 1,155 (s), 1,003 (m), and 958 (w). The pigment scytonemin was detected as a major pigment in *Nostoc* sp. cells: 1,711 (w), 1,633 (m), 1,598 (s), 1,555 (m), 1,381 (mw), 1,171 (ms), 573 (w), and 436 (w)  $\text{cm}^{-1}$  (see Figure 5B).

### 3.4.3. Monte Gibliscemi (site 4)

At this site the orange layer was colonized by orange-pigmented coccoid cells of *Chlorophyta* as corroborated by the strong signal of carotenoid pigment that was detected with bands at 1,517 (s), 1,447 (w), 1,389 (w), 1,274 (w), 1,191 (ms), 1,156 (s), 1,005 (m), and 964 (w)  $\text{cm}^{-1}$  (see Figure 5C).

## 3.5. Fluorescence microscopy of endolithic cyanobacteria

The presence of UV-screening pigment such as scytonemin within some cyanobacteria aggregates were indirectly detected by the visualization of autofluorescence of phycobiliproteins, and chlorophyll *a* signal attenuation observed using a DAPI filter but not using a Rhodamine filter. The excitation light of the DAPI filter (maximum at 365 nm) lay in UVA radiation range. Indeed, this UVA radiation can be absorbed by the UV-screening scytonemin, if present within the cyanobacteria aggregates. This situation is shown in Figure 6, where a large colony of *Nostoc* species in dark-pigmented part (DIC image) of the aggregate revealed a distinct attenuation of the autofluorescence signal when visualized with a DAPI filter. The same cyanobacteria aggregate had exhibited a uniform autofluorescence signal over the aggregate when observed with the Rhodamine filter, as scytonemin does not absorb light in the Rhodamine filter excitation range (450–490 nm). Similar results were also observed for other cyanobacteria taxa such as *G. pleurocapsoides* and *G. compacta*.

## 3.6. Molecular analysis

We used molecular methods based on rRNA for the phylogenetic barcoding of the most abundant species found—coccoid algae, *G. compacta*, *G. novacekii*, *G. pleurocapsoides*, *Nostoc* sp., and *Chroococcidiopsis* sp., as well as for the overall metagenomic studies of entire communities across two different varieties of gypsum (selenite vs. crystalline gypsum) and two differently macroscopically pigmented layers of colonizations (green vs. black) (Supplementary Table S2).

### 3.6.1. Metagenomic profiling of cyanobacteria

We obtained a total of 5,424 unique bacterial ASVs of the 16S rRNA gene, with 100–805 (354 on average) ASVs per sample. Taxonomic assignment of the ASVs to bacterial phyla consistently showed that the dominant group was Cyanobacteria (65.9% on average), followed by Proteobacteria (12.6%), Chloroflexi (4.5%),

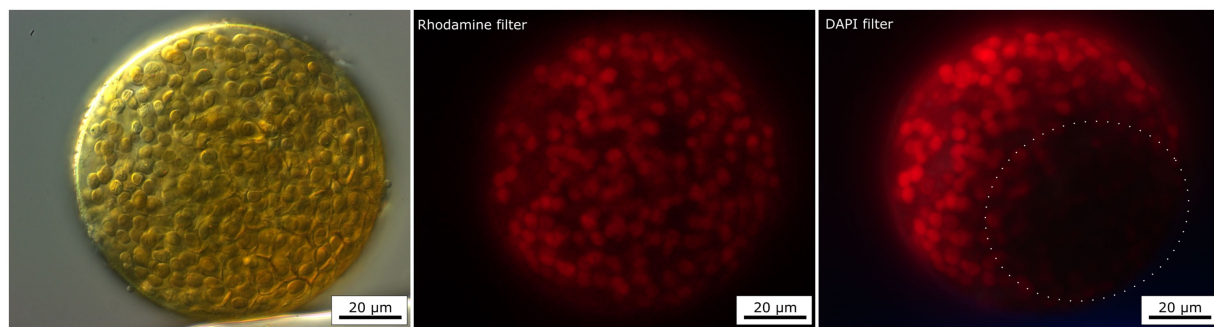


FIGURE 6

Fluorescence image of *Nostoc* colony using two different filters—Rhodamine and DAPI filter. The red autofluorescence of chlorophyll *a* and phycobiliproteins are partly screened by the presence of brown UV-screening pigment—scytonemin when excited by DAPI filter (dotted circle).

Bacteroidota (4.2%), Actinobacteriota (3.6%), Acidobacteriota (2.5%), Planctomycetota (2.3%), and Verrucomicrobiota (0.9%)—the latter showing a higher abundance in crystalline gypsum samples (Figure 7). Two families of coccoid cyanobacteria were the most abundant across the majority of the samples: Chroococcidiopsidaceae and Thermosynechococcaceae (Figure 7). Members of Leptolyngbyaceae, Nostocaceae, and Gloeobacteraceae were also generally present, and each of those groups was dominant in several individual samples.

Of the two localities for which sequencing data were obtained, Sicualiana Marina contained both types of the lithic substrate; whereas Santa Elisabetta contained only the selenite substrate. However, both localities included sets of green and black colored endolithic colonizations. Therefore, the data set was only partially representative with regard to the two types of gypsum; and the effect of the substrate type could only partly be separated from any unknown effects of the sampling locality such as, for example, the differences in the nutrient availability. Nevertheless, when we looked at the taxonomic composition between the two localities, the only striking difference was the dominance of Gloeobacteraceae in two (black and green) analyzed selenite samples collected from the Santa Elisabetta site. The two localities did not exhibit a significant difference in average biodiversity (Shannon index;  $p > 0.1$ ). The samples from Santa Elisabetta were more uniform (Supplementary Figure S1), which was not surprising given their lower number and occurrence of only selenite.

In terms of the type of substrate, selenites on average contained more Nostocaceae, Leptolyngbyaceae, and Gloeobacteraceae ( $p < 0.05$ ); but less Thermosynechococcaceae ( $p < 0.01$ ) when compared to crystalline gypsum (Figure 7). The composition of the crystalline gypsum samples was more uniform, and all of them were only dominated by Chroococcidiopsidaceae and Thermosynechococcaceae. However, in the crystalline gypsum samples, Gloeobacteraceae and Coleofasciculaceae sequences were always present, while missing in some samples of the selenites. In summary, these patterns could be expressed as a significantly higher average biodiversity (Shannon index;  $p < 0.01$ ), but a lower overall variability (Supplementary Figure S2) in crystalline gypsum versus selenite samples.

The dark-pigmented endolithic zones on average contained more Nostocaceae (filamentous cyanobacteria with colored sheaths) compared to the green-pigmented zones ( $p < 0.05$ ) (Figure 7). On the other hand, Thermosynechococcaceae (coccoid cyanobacteria with

colorless envelopes) were more abundant in the green areas ( $p < 0.01$ ). Chroococcidiopsidaceae (coccoid cyanobacteria with pigmented or non-pigmented envelopes) were present in high abundance within both of these differently colored zones. The biodiversity was not statistically different between the black-pigmented and green zones (Shannon index;  $p > 0.1$ ) however, the samples from green endolithic zones were more homogenous (Supplementary Figure S3).

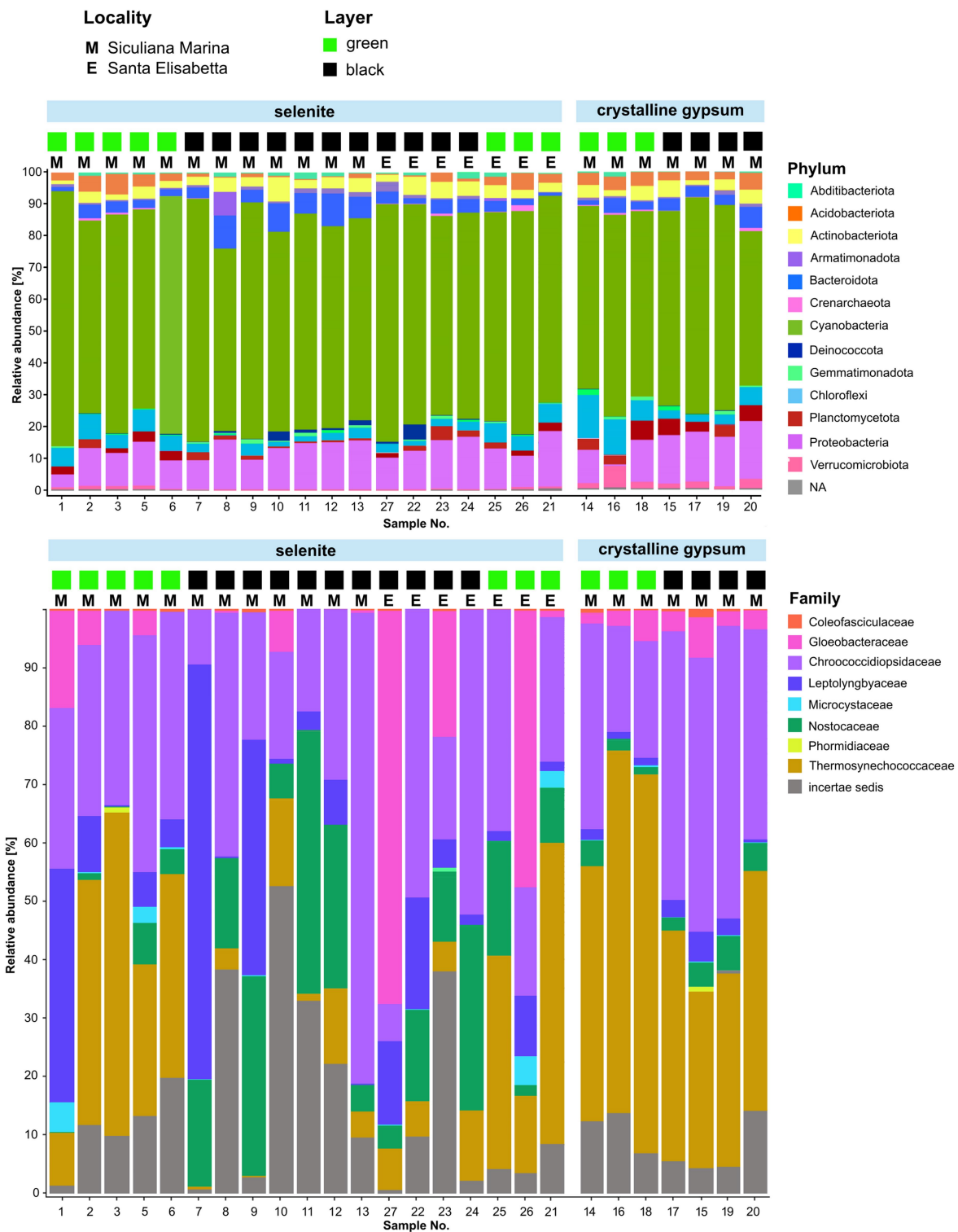
Significant differences in the overall ASV composition among the endolithic assemblages from selenite vs. white crystalline-gypsum, samples from green vs. dark layers, and samples from the two localities with sequencing data were further supported by NMDS analysis (Supplementary Figures S1–S3); although the variance between the localities was less pronounced. In permutational ANOVA tests, the type of substrate, the color of the endolithic layer, and the locality significantly explained 12% ( $p < 0.001$ ), 10% ( $p < 0.001$ ), and 9% ( $p < 0.01$ ) of the observed variability, respectively.

### 3.6.2. Phylogenetic analysis

Phylogenetic analysis allowed us to link the majority of the dominant ASVs discovered by metagenome profiling to cyanobacterial morphotypes observed by light microscopy.

The largest subgroup of dominant ASVs were classified to Thermosynechococcaceae (see Figure 8 for the list of ASVs), clustering in a deep-branching monophyletic lineage together with three of our single cell isolates of the “*Chroococcidiopsis*” morphotype. Another dominant group of ASVs, classified as Chroococcidiopsidaceae, exhibited considerable taxonomic diversity. It contained ASVs that could be assigned to several species within the lineage of *Gloeocapsa* spp. and *Gloeocapsopsis pleurocapsoides* (ASV9, 12, 13, 42, 55, 103, and 1,235), as established by our single cell isolates. Furthermore, it contained two abundant ASVs (8 and 23) clustering in the *Chroococcidiopsis sensu stricto* clade defined by the sequence of its type species, *C. thermalis*, two ASVs (36 and 81) in the *Sinocapsa* lineage, one very abundant “*Chroococcidiopsis* sp.” of the CC1 lineage (ASV1), and a single ASV (58) related to *Myxosarcina*. These results demonstrate extensive cryptic diversity of the coccoid endolithic cyanobacteria in the samples, especially regarding the *Chroococcidiopsis*-like morphotype.

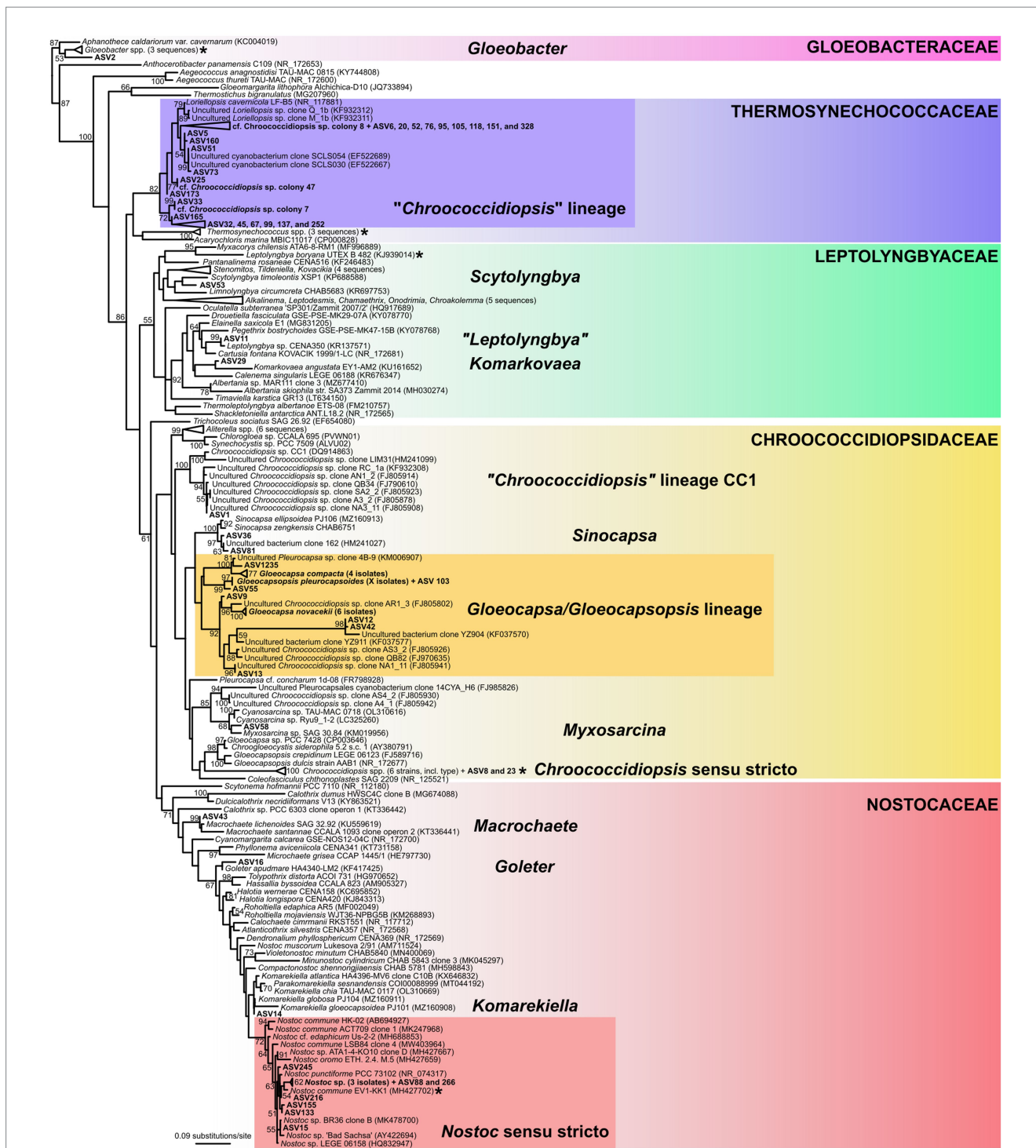
Another very abundant ASV (2), classified as Gloeobacteraceae, was resolved as a sister lineage to *Gloeobacter*



**FIGURE 7**  
Graphical representation of the relative abundance of prokaryotic phyla and cyanobacterial families in two locations (Siculiana Marina and Santa Elisabetta), in selenitic and crystalline gypsum and in differently pigmented endolithic zones (green versus black).

*violaceus*, in agreement with the observation of *G. violaceus* by light microscopy. The three most abundant ASVs of Leptolyngbyaceae were resolved in clades corresponding to *Scytolyngbya* (ASV53), *Komarkovaea* (ASV29), and “*Leptolyngbya*” sp. (ASV11); outside of the core *Leptolyngbya* cluster, showing the

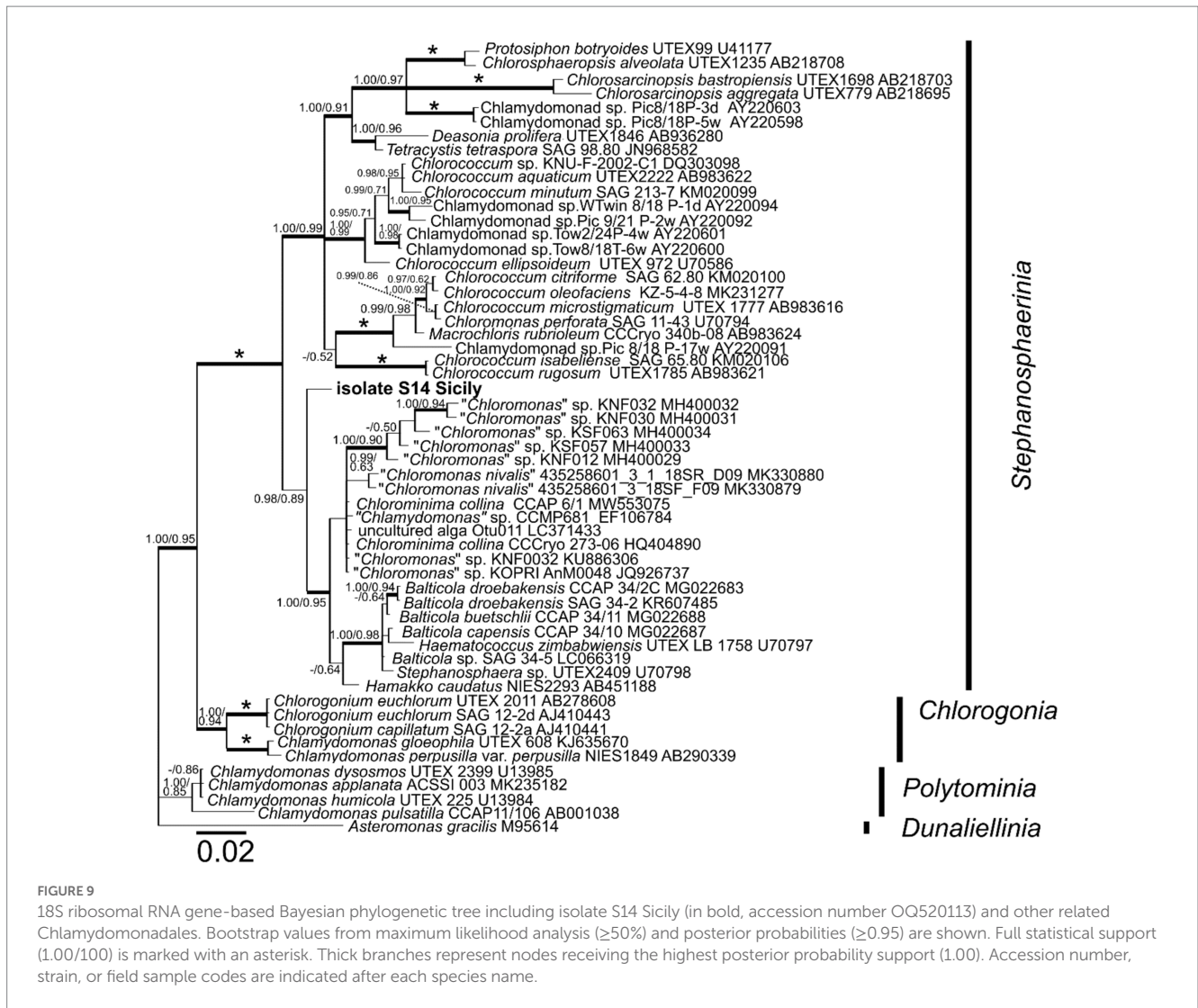
relatively broad diversity of simple filamentous cyanobacteria in the samples. These taxa were not captured by microscopic analysis (Table 2). A majority of ASVs classified as Nostocaceae were resolved in the *Nostoc sensu stricto* clade along with our sequences of *Nostoc* sp. single filaments identified by light microscopy. The



**FIGURE 8**  
 16S rRNA gene phylogenetic tree of cyanobacteria. The position of cyanobacterial morphotypes from the gypsum samples, barcoded by single cell sequencing, is shown in the tree along with the most abundant amplicon sequence variants (ASV) from metagenome analysis. Position of cyanobacterial families and genus-level lineages that were dominant in the metagenome analysis is highlighted. Position of the type species of the families is indicated by asterisks. The tree was inferred using Maximum Likelihood algorithm with 1,000 bootstrap pseudo-replicates, branch supports >50% are shown at nodes.

remaining three abundant ASVs of heterocytous cyanobacteria clustered into genera *Goleter* (ASV16) and *Komarekiella* (ASV14), which are morphologically cryptic to *Nostoc*, and to *Macrochaete* (ASV43), a heteropolar cyanobacterium which was not found by microscopy.

Several taxa identified in the samples using light microscopy, i.e., *Petalonema* sp., *Microcoleus* sp., and *Anathece* sp., were not found among the dominant ASVs; however, sequences classified as *Petalonema* and *Microcoleus* were found in the samples as low-abundance ASVs.



### 3.6.3. Sanger sequencing of orange-pigmented coccoid algae and their phylogenetic position

Using culture-independent sequencing, the acquired Sanger sequence (385 bp long) matched with 98.96% identity to several closely related strains or environmental sequences that were all assigned to the order Chlamydomonadales (Chlorophyta), clade *Stephanosphaerinia*: “*Chloromonas*” sp. KSF063 (MH400034.1), “*Chloromonas*” sp. KNF032 (MH400032.1), “*Chloromonas*” sp. KNF030 (MH400031.1), “*Chloromonas*” sp. 435258601\_3\_18SR\_D09 (MK330880.1), “*Chloromonas*” sp. 435258601\_3\_18SF\_F09 (MK330879.1), uncultured alga gene Otu011 (LC371433.1), “*Chloromonas*” sp. KNF0032 (KU886306.1), *Chlorominima collina* CCAP 6/1 (MW553075), and *Chlorominima collina* CCCryo 273–06 (HQ404890.1).

Phylogenetic analyses of the 18S rDNA marker placed isolate S14 Sicily in the clade *Stephanosphaerinia* (Chlorophyta) in a basal position in relationship to the clade containing *Balticola*, *Haematococcus* and *Stephanosphaera* species, and to the clade of *Chloromonas*-like and *Chlorominima* species (Figure 9).

## 4. Discussion

The majority of previous reports on the composition of endolithic communities have focused on extreme regions such as the Negev Desert (Friedmann et al., 1967; Archer et al., 2017), the Atacama Desert (Wierzchos et al., 2006), and the Dry Valleys of Antarctica (Friedmann et al., 1967; Archer et al., 2017).

In contrast, abundant endolithic communities in temperate regions have largely been neglected, with only a few exceptions, e.g., limestone in China (Horath et al., 2006; Tang et al., 2012) or lithic colonization of monumental stones; e.g., Jeronimos Monastery (Lisbon, Portugal) (Ascaso et al., 2002) or Romanesque churches of Segovia (Spain) (de Los Ríos et al., 2009). In this study, we used spectroscopic, microscopic, and molecular analysis to describe endolithic gypsum phototrophs from Sicily, a region typically with a dry temperate climate. By using multiple techniques, we were able to gain a more comprehensive understanding of the distribution, and diversity of these organisms in the region as well as their pigment composition.

It is known that gypsum can form many varieties differing only in their structure (Al-Wakeel, 2002). In this study, three different structures of gypsum were observed. CT-scan images showed that transparent selenites and white crystalline-gypsum from Siculiana Marina (Site 1) and Santa Elisabetta (Site 2) have a similar inner architecture and have a relatively smaller percentage of void space colonized by diverse endolithic communities.

At the remaining two other sites, Monte Perrera (Site 3) and Monte Gibliscemi (Site 4), only light microscopy was used to describe the cyanobacteria, due to the small number of collected samples and low biomass available for metagenomic analysis. At the Monte Perrera site, we identified fewer cyanobacterial species compared to other sites, even though the gypsum found here had the greatest void space, which would in theory provide more space for colonization by microorganisms. The pigmented zones at this site were dominated by a single species. The colonized zones found in the fine-grained gypsum at the Monte Gibliscemi site were only colonized by green algae, and cyanobacteria were virtually absent. These findings suggest that the microbial communities at these sites may be influenced by factors other than the availability of void space in the gypsum; e.g., light availability, pH, or the rate of erosion or weathering.

Differences in the relative abundance of endolithic microorganisms in gypsum of various physical features can be explained by combined effects of environmental factors promoting various ecological interactions within the microbial communities. For example, previous studies have shown that cracks and fissures in calcite allow direct access to sunlight, which then allows deeper colonizations compared to ignimbrite (Crits-Christoph et al., 2016). It was also observed that endolithic communities were less diverse in gypsum with large inner spaces, while more diverse in gypsum with high spatial heterogeneity. This relationship was described as the theory of pore connectivity stating that “low pore connectivity sustain higher microbial diversity because they prevent consumption of resources by more competitive organisms and provide spatial shelter for less competitive organisms” (Ertekin et al., 2020). In other words, the lack of interconnectedness of pores in the gypsum can enhance the diversity of the endolithic community.

In our samples, CT scan results showed that there was almost no difference in the inner microstructures of the two gypsum varieties analysed (Figure 4 and Supplementary Figure 4). However, in terms of light conditions, the transparent selenites should provide better conditions for the growth of photosynthetic microorganisms than do white crystalline gypsum. Indeed, metagenomic analysis has shown that samples of selenites hosted a greater biodiversity (Shannon index;  $p < 0.01$ ), suggesting that the transparency of the gypsum may play a more significant role than microstructure of the substrate in shaping the diversity of these particular endolithic communities.

#### 4.1. UV-screening pigments

UV-screening pigments are commonly distributed in endolithic settings, especially in extreme regions that are exposed to increased radiation (Wierzbos et al., 2015; Orellana et al., 2020; Vitek et al., 2020). Among these carotenoids usually dominate as they participated not only in the UV-protection but also in photosynthesis (Maguregui et al., 2012; Jehlička et al., 2016; Osterrothová et al., 2019; Němečková et al., 2020). In our study, carotenoids were detected using Raman

spectroscopy in all dominant species (*Nostoc* sp., *G. compacta*, and *G. pleurocapsoides*, orange coccoid algae of the clade *Stephanosphaerinia*). Three Raman bands at  $\sim 1,520$ , 1,155, and  $1,008\text{ cm}^{-1}$  are attributed to  $\nu_1$  (C=C),  $\nu_2$  (C-C) stretching mode, and  $\delta$  (C-CH<sub>3</sub>) bending mode (Jehlička et al., 2014b). The  $\nu_1$  band wavenumber in some cases can be used for more specific recognition of carotenoids. This band is correlated to the length of the polyene chain since the wavenumber of  $\nu_1$  (C=C) decreases as the number of double bonds in polyene chain increases (Kuzmany, 1980; Withnall et al., 2003; de Oliveira et al., 2010). Precise discrimination of carotenoids is important for taxonomic identification of microorganisms (Edwards et al., 2014; Jehlička et al., 2014a). However, the unambiguous carotenoid discrimination was not possible in the case of cyanobacteria as in biological systems the overall spectrum may also be influenced by other factors, e.g., molecular interactions or stages of life cycles (Collins et al., 2011; Osterrothová et al., 2019). These factors can manifest as shifts in the Raman spectra. In the case of the orange algal stages (*Stephanosphaerinia*) found at Monte Gibliscemi, obtained spectra were rather uniform. Three Raman bands of carotenoids were detected at 1,517, 1,156/55, and  $1,005\text{ cm}^{-1}$ . From morphologically-similar orange-pigmented coccoid algae isolated from gypsum found in the Atacama Desert, Raman spectra of carotenoids were also obtained at 1,521 and  $1,519\text{ cm}^{-1}$  (Wierzbos et al., 2015). These fall in the characteristic area of  $\beta$ -carotene, a common primary carotenoid (Jehlička et al., 2014b).

Besides carotenoids, two other protective pigments were detected—gloeocapsin and scytonemin. Gloeocapsin and scytonemin are dark UV screening pigments, unique for cyanobacteria. Using Raman spectroscopy, these have been previously detected in colonization of various lithic environments, e.g., in limestone (Belgium and Croatia) (Storme et al., 2015), gypsum (Sicily and Poland) (Němečková et al., 2021) or halite (the Atacama Desert, Chile) (Vitek et al., 2014b). In the current study, gloeocapsin was detected particularly in dark-pigmented *Gloeocapsa compacta* found in the upper zones of colonization, with the major Raman bands at 1,671 and  $1,571\text{ cm}^{-1}$ . Chemically, gloeocapsin is probably a derivative of parietin which has an anthraquinone skeleton (Storme et al., 2015), however, recently, it was also proposed that gloeocapsin could alternatively belong to another class of biomolecules capable of halochromism such as the coumarins, or flavonoids (Lara et al., 2022). Besides the *Gloeocapsa* genus, gloeocapsin was detected in epilithic *Solentia paulocellulare* (Storme et al., 2015) and *Phormidesmis nigrescens* ULC007 (Lara et al., 2022). Currently, the presence of gloeocapsin has only been confirmed in three distinct lineages of cyanobacteria (*Chroococcidiopsidales*, *Chroococcales*, *Leptolyngbyales*) (Lara et al., 2022; Strunecký et al., 2022), and more research is needed to determine the full extent of its occurrence.

On the other hand, scytonemin has been found in many cyanobacteria (García-Pichel and Castenholz, 1991; Dillon et al., 2002; Sorrels et al., 2009; Grant and Louda, 2013; Vitek et al., 2014b; Ekebergh et al., 2015; Gao, 2017). In this study, scytonemin was detected in two species—*Nostoc* sp. and *Gloeocapsopsis pleurocapsoides*, with the most prominent peak at around  $1,598\text{ cm}^{-1}$ , which corresponds to the  $\nu$ (CCH) aromatic ring quadrant stretching mode. Scytonemin has been detected by Raman spectroscopy in various lithic ecosystems (usually in extreme regions); e.g., in gypsum (Wierzbos et al., 2015) and halite (Vitek et al., 2012) from the Atacama Desert. In these settings, scytonemin is mainly produced by



cyanobacteria in the uppermost zones that are the most irradiated by sunlight (Balskus et al., 2011; Vitek et al., 2014b), which is in accordance with this study. However, Raman spectroscopy has also detected scytonemin in the hypodolitic colonies underneath the carotenoid-containing zones of gypsum from the Atacama Desert (Wierzbos et al., 2015). This trend may be due to the specific physical features of the colonized gypsum deposits, which allow for some transmission of ultraviolet radiation (Wierzbos et al., 2015).

Besides Raman spectroscopy, the presence of scytonemin was detected indirectly by fluorescence microscopy. Scytonemin absorbs from ~220 to 480 nm, with *in vivo*  $\lambda_{max}$  at 370 nm, which is close to the excitation by DAPI filter (excitation ~365 nm). Due to this, we could observe a lower signal of red autofluorescence of chlorophyll and phycobiliproteins in those zones where scytonemin had accumulated (see Figure 6). A similar trend was also observed in endoliths in halite (Vitek et al., 2014b). In this study, fluorescence analysis shows either a very weak or no red autofluorescence signal of chlorophyll and phycobiliproteins in the black aggregates using GFP (excitation ~450–490 nm). This decrease in red emission was interpreted as the result of scytonemin's screening effect of the incident green light together with the fact that photosynthetic pigments could be degraded, which would also lead to a decline in the autofluorescence signal (Vitek et al., 2014b). This indirect detection of scytonemin can complement other methods such as Raman spectroscopy. However, this only allows the visualization of the distribution of such compounds in cells, and not the direct detection of scytonemin (or gloeocapsin).

The detection of UV-screening pigments, particularly scytonemin and gloeocapsin, in temperate regions supports previous suggestions about the potential functions of these pigments. Scytonemin synthesis is known to be affected by various environmental stresses, such as temperature, oxidative stress, salinity, and desiccation (Dillon et al., 2002; Vitek et al., 2014b). In addition, Villar et al., have detected scytonemin in gypsum crust, but noted that no scytonemin was detected in rock substrates containing iron oxide, which itself protects against intense solar radiation (Villar et al., 2005). As for gloeocapsin, greater detailed information about its physiological role is still required, but it could be suggested that its biosynthesis may be triggered by various factors, as well.

## 4.2. Diversity of gypsum phototrophs

### 4.2.1. Cyanobacteria as the dominant group

Our metagenomic analysis of gypsum from the southern part of Sicily showed the dominance of *Cyanobacteria* over other phyla such as *Proteobacteria*, *Chloroflexi*, *Bacteroidota*, *Actinobacteriota*, *Planctomycetota*, and *Verrucomicrobiota*. Previous studies on metagenomics showed that cyanobacteria significantly dominate as well, e.g., in gypsum from the Atacama Desert (Crits-Christoph et al., 2016). The other detected phyla have also been described from previous studies in calcite and ignimbrite from the hyperarid Atacama Desert (Wierzbos et al., 2015; Crits-Christoph et al., 2016), sandstone, limestone, and granite from the semiarid regions of the Rocky Mountain region (US) (Walker and Pace, 2007), and interestingly, some of them were also detected in an endolithic community associated with the coral *Porites astreoides* (Wegley et al., 2007). Cyanobacteria, *Proteobacteria*, *Actinobacteria*, and *Acidobacteria*

were also found in gypsum from the Canadian Arctic (Ziolkowski et al., 2013). The similarities of the diversity on the level of phyla from various lithologies and geographical regions of the planet supports the hypothesis of a global metacommunity. This hypothesis says that endolithic ecosystems are comprised of organisms from a relatively limited reservoir of phylogenetic diversity and that these organisms are specifically adapted to their endolithic habitat (Walker and Pace, 2007). Nevertheless, their relative abundance may vary, for example, *Chloroflexi* and *Proteobacteria* were dominant over cyanobacteria in gypsum collected from the Lake St. Martin impact crater (Rhind et al., 2014). In halite from the Atacama Desert, *Euryarchaeota* were also more abundant compared to cyanobacteria (Wierzbos et al., 2018). As already mentioned, the variances of relative abundance can be related to the physical properties and inner architectures of the inhabited substrates. For instance, a higher taxonomic diversity was observed, in calcite than in ignimbrite (Crits-Christoph et al., 2016). This could be explained by the fact that calcite can form larger fissures and cracks compared to ignimbrite, which would lead to more competition for space and nutrient resources (Crits-Christoph et al., 2016). This also results in different metabolic pathways in these two substrates, genes for ABC transporters and osmoregulation are more diverse in the calcite habitat and genes for secondary metabolites are more abundant in the ignimbrite community (Crits-Christoph et al., 2016). The presence of some genes is also related to the nutrient availability in the environment. For example, no genes for diazotrophy have been detected in calcite and ignimbrite from the Atacama Desert where there are high nitrate accumulations from the atmosphere (Crits-Christoph et al., 2016) compare to. Besides that, the greater abundance of cyanobacteria could be related to the lower water availability, which means decreased photosynthesis and thus less CO<sub>2</sub> fixation, which is important for the development of heterotrophs (Wierzbos et al., 2018).

In general, in majority of endolithic ecosystems, cyanobacteria are the main primary producers, and the *Chroococcidiopsis* is frequently the major genus (Wierzbos et al., 2018). For example, in the dolomite in south China, nine cyanobacterial major phylotypes were found: *Chroococcidiopsis* sp., *Phormidium autumnale*, *Anabaena oscillarioides*, *Scytonema* sp., *Leptolyngbya* sp., *Calothrix* sp., two *Nostoc* sp., and *Brasilonema octagenarum* (Tang et al., 2012). In gypsum crust from Chott el Jerid, most cyanobacterial phylotypes were related to the genus *Leptolyngbya* within the *Oscillatoriales* and *Nostocales* (Stivaletta et al., 2010). The molecular studies show that the most abundant genomes found in the calcite and ignimbrite were also related to *Chroococcidiopsis* and *Gloeocapsa* (Crits-Christoph et al., 2016). In general, similar compositions have been observed all over the globe in many various substrates (e.g., gypsum crust, Stivaletta et al., 2010), limestone (Ferris and Lawson, 1997; Tang et al., 2012), which could also be explained by the hypothesis of global metacommunity. Also, it has been shown that the overall relative abundance tends to be greater in less extreme environments (Ziolkowski et al., 2013). Nevertheless, the authors acknowledge that it is possible that the higher diversity could be related to the limitations of the molecular methods used (Ziolkowski et al., 2013).

In the current study, we utilized a single cell/filament sequencing technique (Mareš et al., 2015; Kurmayer et al., 2018) to provide a direct link between cyanobacterial morphotypes observed using light microscopy and the results of metagenomic profiling. Such an approach is elaborate, and has only rarely been used in previous

studies (Bowers et al., 2022), however, it provides valuable data leading to a better understanding of the microbial communities composition. Another similar, but even more challenging, approach is to link the metagenomic data to barcoded laboratory strains isolated from the same samples (Águila et al., 2021).

The most interesting result was a partial dissection of the cryptic variability in coccoid cyanobacteria that dominated the samples. Surprisingly, the main *Chroococcidiopsis*-like morphotype (Figure 3D) analysed by single-cell sequencing was resolved as *Thermosynechococcaceae* (Figure 8), together with a number of major amplicon sequence variants (ASVs). Some of the remaining frequent ASVs could be linked to the *Chroococcidiopsis sensu stricto* lineage, and even to other lineages of cyanobacteria forming sarcinoid packets of cells (“*Chroococcidiopsis*” lineage CC1, *Sinocapsa*, “*Myxosarcina*”). The polyphyly of *Chroococcidiopsis*-like morphotypes have long been known (Cumbers and Rothschild, 2014); however, these taxa are still, for the most part, awaiting formal taxonomic treatment.

The pattern of *Chroococcidiopsis*-like phylotypes found in our study matches those previously reported worldwide from both hot and cold deserts (Bahl et al., 2011), except for the dominant clade in *Thermosynechococcaceae*, for which the morphological information had been missing up to now. Similar organisms were previously sequenced from endolithic limestone communities in the Rocky Mountains and the Mojave Desert, United States (Walker and Pace, 2007; Smith et al., 2014). Another environmental sequence in this clade was erroneously assigned to a true-branching heterocytous cyanobacterium *Loriellopsis cavernicola* (Lamprinou et al., 2011), probably due to laboratory contamination from the sample of calcite rock taken from a cave in Spain.

A complementary line of evidence was provided by single-cell sequencing of the abundant coccoid cyanobacteria containing dark UV-screening sheath pigments, *Gloeocapsa* spp. and *Gloeocapsopsis pleurocapsoides*. All of them clustered in the *Chroococcidiopsidaceae* clade together with a number of frequent amplicon sequence variants and previously published uncultured “*Chroococcidiopsis* sp.” sequences (Figure 8). Reliable sequences of these taxa have never been provided before; thus our results will help to distinguish these eco-physiologically different organisms in future metagenomic studies.

Another interesting finding was the relatively high proportion of *Gloeobacter* in the gypsum samples documented by both light microscopy and amplicon sequencing (Figures 3, 9). While it is known that this basal thylakoid-less cyanobacterium is more common than once thought (Mareš et al., 2013), its global occurrence is still understudied.

The remaining dominant cyanobacteria were classified to *Nostocales* (Nostocaceae) and *Leptolyngbyaceae*. Identity of the dominant *Nostoc* sp. morphotype (Figure 3C) was confirmed by single filament sequencing. However, also in this case, taxa morphologically cryptic to *Nostoc* (Goleter; Miscoe et al., 2016 and Komarekiella, Johansen et al., 2017) were discovered by the metagenomic analysis along with several other heterocytous cyanobacteria (Figure 8). Surprisingly, the thin filamentous cyanobacteria from *Leptolyngbyaceae* were completely overlooked by microscopy analysis. This is due to the fact that only limited proportion of the samples can be microscopically observed. On the other hand, some of the taxa seen in the microscope (*Microcoleus*, *Petalonema*, *Anathece*) were very scarce or missing in the amplicon sequencing data (e.g., due to better DNA

extraction of some species). These findings illustrate the limitations of both approaches.

Considering the community composition, all samples analysed by metagenomic profiling were generally quite similar, and differed only in the proportion of individual groups (Figure 7). As expected, an increased occurrence of cyanobacteria with dark pigmentation (*Nostocales*) was confirmed in the macroscopically dark endolithic layers compared to the green layers dominated by *Thermosynechococcaceae*, in which the sheath pigments are absent.

#### 4.2.2. Endolithic algae

The orange pigmented coccoid algae inhabiting fine-grained gypsum at Site 4 (isolate S14 Sicily) were found to be genetically related to green algae from the *Chlorophyceae* family originated from extreme habitats such as: Antarctic or Arctic freshwater (Jung et al., 2016; Cho et al., 2019), red snow on glacier and green snow fields at King George island (Gálvez et al., 2021), snowfields in Ryder Bay, Antarctic Peninsula (Davey et al., 2019), and Yukidori Valley, Eastern Antarctica (Segawa et al., 2018). Metagenomic sequencing of Arctic gypsum endoliths indicated the presence of few green algal sequences closely matching with *Trebouxia* sp. and *Trichosarcina* sp., known lichenizing symbionts (Ziolkowski et al., 2013). In other study, coccoid chlorophototrophs were responsible for gypsum layered orange to green cryptoendolithic colonization in Atacama, Northern Chile, but its sequences failed to be amplified by the environmental sequencing; likely due to the primer bias (Wierzychos et al., 2015). Another study on chasmoendoliths in Antarctica in granite and gypsum from the Atacama Desert showed the presence of *Trebouxiphyceae* (Yung et al., 2014; Ertekin et al., 2020). Further, known alga found in gypsum (spring associated) was the desmid *Oocardium stratum* (Zygnematophyceae, Streptophyta) (Rott et al., 2010). In our case study, direct DNA extraction isolated from gypsum generated a 18S rDNA sequence corresponding to those from the order of Chlamydomonadales, clade *Stephanosphaerinia*; mostly to genera of *Chloromonas*-like morphology and *Chlorominima* (Figure 9). However, the acquired sequence fragment was too short to allow any closer taxonomic assignment. Some of the related species to our Sicilian algal isolate were shown to form morphologically similar globular-shaped thick-walled cells with cytoplasmic lipid droplets in its life cycle that occupy most of the cell volume. The production of abundant lipid droplets is also correlated with the accumulation of carotenoids (Grünwald et al., 2001), and as mentioned above, carotenoids were also detected in our study by Raman spectroscopy.

## 5. Conclusion

Endolithic ecosystems, which are found in a variety of extreme environments, play a significant role as a model for astrobiology. In this study, we examined gypsum endoliths from Sicily, where gypsum outcrops are quite common. Our results from a metagenomic analysis supported the theory of a global metacommunity, indicating that endolithic communities of a similar general composition can be found across a wide range of environments. In addition to the classical amplicon sequencing approach, single cell sequencing of selected dominant phototrophs allowed us to dissect their phylogenetic relationships and link

abundant morphotypes to amplicon sequence variants and Raman spectroscopy data. Unlike in previous studies, the sites we studied were not exposed to extreme solar radiation; yet we still frequently detected UV-protective pigments, scytonemin and gloeocapsin. Overall, our study highlights the diverse and complex nature of endolithic ecosystems, focusing on the previously less studied gypsum substrate and variable colonization of its two forms.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found: NCBI – OQ509475 – OQ509403 + OQ520113.

## Author contributions

JJ, KN, and AC conceived the study. KN, JJ, and AC collected in-field samples. KN carried overall investigation, data interpretation, and writing of the original draft. JM conducted the molecular analysis of cyanobacteria and writing. LP and LN were responsible for the overall molecular analysis of algae and writing. AC and JJ contributed to the Raman analysis and to the writing. FK conducted the CT-scan analysis and writing. JW contributed to the fluorescence analysis and its interpretation. JD, VT, and JŽ assisted with the CT-scan analysis and methodology. AK performed the DNA extraction and validated the PCR protocol. JZ was involved in the 16S rRNA amplification and DNA library preparation. EN prepared the primers and protocols for amplicon analysis and metagenomics. KN, AC, and JJ edited the manuscript. JJ contributed to the conceptualization and provided supervision, resources, and project administration. All authors contributed to the article and approved the submitted version.

## References

- Águila, B., Alcántara-Hernández, R. J., Montejano, G., López-Martínez, R., Falcón, L. I., and Becerra-Absalón, I. (2021). Cyanobacteria in microbialites of Alchichica Crater Lake: a polyphasic characterization. *Eur. J. Phycol.* 56, 428–443. doi: 10.1080/09670262.2020.1853815
- Al-Wakeel, E.-S. I. (2002). Alabaster, and selenite gypsum: I-dehydration-rehydration comparison studies. *J. Mater. Res. Technol.* 18, 365–368.
- Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Archer, S. D., de los Ríos, A., Lee, K. C., Niederberger, T. S., Cary, S. C., Coyne, K. J., et al. (2017). Endolithic microbial diversity in sandstone and granite from the McMurdo Dry Valleys, Antarctica. *Polar Biol.* 40, 997–1006. doi: 10.1007/s00300-016-2024-9
- Ascaso, C., Wierzbos, J., Souza-Egipsy, V., Delosrios, A., and Rodrigues, J. (2002). *In situ* evaluation of the biodeteriorating action of microorganisms and the effects of biocides on carbonate rock of the Jeronimos Monastery (Lisbon). *Int. Biodeterior. Biodegrad.* 49, 1–12. doi: 10.1016/S0964-8305(01)00097-X
- Bahl, J., Lau, M. C., Smith, G. J., Vijaykrishna, D., Cary, S. C., Lacap, D. C., et al. (2011). Ancient origins determine global biogeography of hot and cold desert cyanobacteria. *Nat. Commun.* 2, 163–166. doi: 10.1038/ncomms1167
- Balskus, E. P., Case, R. J., and Walsh, C. T. (2011). The biosynthesis of cyanobacterial sunscreen scytonemin in intertidal microbial mat communities. *FEMS Microbiol. Ecol.* 77, 322–332. doi: 10.1111/j.1574-6941.2011.01113.x
- Basilone, L. (2018). *Lithostratigraphy of Sicily*. Cham: Springer.
- Benison, K. C., and Karmanocky III, F. J. (2014). Could microorganisms be preserved in Mars gypsum? Insights from terrestrial examples. *Geology* 42, 615–618. doi: 10.1130/G35542.1
- Bowers, R. M., Nayfach, S., Schulz, F., Jungbluth, S. P., Ruhl, I. A., Sheremet, A., et al. (2022). Dissecting the dominant hot spring microbial populations based on community-wide sampling at single-cell genomic resolution. *ISME J.* 16, 1337–1347. doi: 10.1038/s41396-021-01178-4
- Brown, J. J., Rodriguez-Ruano, S. M., Poosakkannu, A., Batani, G., Schmidt, J. O., Roachell, W., et al. (2020). Ontogeny, species identity, and environment dominate microbiome dynamics in wild populations of kissing bugs (Triatominae). *Microbiome* 8, 146–116. doi: 10.1186/s40168-020-00921-x
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869
- Caruso, A., Pierre, C., Blanc-Valleron, M.-M., and Rouchy, J.-M. (2015). Carbonate deposition and diagenesis in evaporitic environments: the evaporative and Sulphur-bearing limestones during the settlement of the Messinian salinity crisis in Sicily and Calabria. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 429, 136–162. doi: 10.1016/j.palaeo.2015.03.035
- Cho, S. M., Kim, S., Cho, H., Lee, H., Lee, J. H., Lee, H., et al. (2019). Type II ice-binding proteins isolated from an arctic microalga are similar to adhesion-like proteins and increase freezing tolerance in transgenic plants. *Plant Cell Physiol.* 60, 2744–2757. doi: 10.1093/pcp/pcz162
- Collins, A. M., Jones, H. D., Han, D., Hu, Q., Beechem, T. E., and Timlin, J. A. (2011). Carotenoid distribution in living cells of *Haematococcus pluvialis* (Chlorophyceae). *PLoS One* 6:e24302. doi: 10.1371/journal.pone.0024302
- Crits-Christoph, A., Robinson, C. K., Ma, B., Ravel, J., Wierzbos, J., Ascaso, C., et al. (2016). Phylogenetic and functional substrate specificity for endolithic microbial

## Funding

This project was supported by the Czech Science Foundation (Grant/Award No. 17-04270S and 21-03322S), Ministry of Education, Youth and Sports of the Czech Republic, National Programme of Sustainability I (Grant/Award No. LO1416), Charles University (Grant/Award Nos. UNCE/SCI/006 and UNCE 204069), ALGAMIC (Grant/Award No. CZ.1.05/2.1.00/19.0392). JM was supported by the Czech Science Foundation (GAČR) Project No. 22-06374S to accomplish phylogenetic and taxonomic analysis. JW was thankful for the financial support by the PGC2021-124362NB-I00 grant from MCI/AEI (Spain) and FEDER.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1175066/full#supplementary-material>

- communities in hyper-arid environments. *Front. Microbiol.* 7:301. doi: 10.3389/fmicb.2016.00301
- Culka, A., Osterrothová, K., Hutchinson, I., Ingley, R., McHugh, M., Oren, A., et al. (2014). Detection of pigments of halophilic endoliths from gypsum: Raman portable instrument and European Space Agency's prototype analysis. *Philos. Trans. A Math. Phys. Eng. Sci.* 372:20140203. doi: 10.1098/rsta.2014.0203
- Cumbers, J., and Rothschild, L. J. (2014). Salt tolerance and polyphyly in the cyanobacterium *Chroococcidiopsis* (Pleurocapsales). *J. Phycol.* 50, 472–482. doi: 10.1111/jpy.12169
- Davey, M. P., Norman, L., Sterk, P., Huete-Ortega, M., Bunbury, F., Loh, B. K. W., et al. (2019). Snow algae communities in Antarctica: metabolic and taxonomic composition. *New Phytol.* 222, 1242–1255. doi: 10.1111/nph.15701
- De los Ríos, A., Cámara, B., Del Cura, M. A. G., Rico, V. J., Galván, V., and Ascaso, C. (2009). Deteriorating effects of lichen and microbial colonization of carbonate building rocks in the Romanesque churches of Segovia (Spain). *Sci. Total Environ.* 407, 1123–1134. doi: 10.1016/j.scitotenv.2008.09.042
- de Oliveira, V. E., Castro, H. V., Edwards, H. G. M., and de Oliveira, L. F. C. (2010). Carotenes and carotenoids in natural biological samples: a Raman spectroscopic analysis. *J. Raman Spectrosc.* 41, 642–650. doi: 10.1002/jrs.2493
- Dillon, J. G., Tatsumi, C. M., Tandingan, P. G., and Castenholz, R. W. (2002). Effect of environmental factors on the synthesis of scytonemin, a UV-screening pigment, in a cyanobacterium (*Chroococcidiopsis* sp.). *Arch. Microbiol.* 177, 322–331. doi: 10.1007/s00203-001-0395-x
- Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604
- Edwards, H. G. M., Garcia-Pichel, F., Newton, E., and Wynn-Williams, D. (2000). Vibrational Raman spectroscopic study of scytonemin, the UV-protective cyanobacterial pigment. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 56, 193–200. doi: 10.1016/s1386-1425(99)00218-8
- Edwards, H. G. M., Hutchinson, I. B., Ingley, R., and Jehlička, J. (2014). Biomarkers and their Raman spectroscopic signatures: a spectral challenge for analytical astrobiology. *Philos. Trans. A Math. Phys. Eng. Sci.* 372:20140193. doi: 10.1098/rsta.2014.0193
- Ekebergh, A., Sandin, P., and Mårtensson, J. (2015). On the photostability of scytonemin, analogues thereof and their monomeric counterparts. *Photochem. Photobiol. Sci.* 14, 2179–2186. doi: 10.1039/C5PP00215J
- Ertekin, E., Meslier, V., Browning, A., Treadgold, J., and DiRuggiero, J. (2020). Rock structure drives the taxonomic and functional diversity of endolithic microbial communities in extreme environments. *Environ. Microbiol.* 23, 3937–3956. doi: 10.1111/1462-2920.15287
- Ferris, F., and Lowson, E. (1997). Ultrastructure and geochemistry of endolithic microorganisms in limestone of the Niagara escarpment. *Can. J. Microbiol.* 43, 211–219. doi: 10.1139/m97-029
- Friedmann, E. I. (1982). Endolithic microorganisms in the Antarctic cold desert. *Science* 215, 1045–1053. doi: 10.1126/science.215.4536.1045
- Friedmann, I., Lipkin, Y., and Ocampo-Paus, R. (1967). Desert algae of the Negev (Israel). *Phycologia* 6, 185–200. doi: 10.2216/i0031-8884-6-4-185.1
- Gálvez, F. E., Saldarriaga-Córdoba, M., Huovinen, P., Silva, A. X., and Gómez, I. (2021). Revealing the characteristics of the Antarctic snow alga *Chlorominima collina* gen. et sp. nov. through taxonomy, physiology, and transcriptomics. *Front. Plant Sci.* 12:662298. doi: 10.3389/fpls.2021.662298
- Gao, X. (2017). Scytonemin plays a potential role in stabilizing the exopolysaccharidic matrix in terrestrial cyanobacteria. *Microb. Ecol.* 73, 255–258. doi: 10.1007/s00248-016-0851-4
- García-Pichel, F., and Castenholz, R. W. (1991). Characterization and biological implications of scytonemin, a cyanobacterial sheath pigment. *J. Phycol.* 27, 395–409. doi: 10.1111/j.0022-3646.1991.00395.x
- García-Veigas, J., Cendón, D. I., Gibert, L., Lowenstein, T. K., and Artiaga, D. (2018). Geochemical indicators in Western Mediterranean Messinian evaporites: implications for the salinity crisis. *Mar. Geol.* 403, 197–214. doi: 10.1016/j.margeo.2018.06.005
- Golubic, S., Friedmann, E. I., and Schneider, J. (1981). The lithobiontic ecological niche, with special reference to microorganisms. *J. Sediment. Res.* 51, 475–478. doi: 10.1306/212F7CB6-2B24-11D7-8648000102C1865D
- Gorbushina, A. A. (2007). Life on the rocks. *Environ. Microbiol.* 9, 1613–1631. doi: 10.1111/j.1462-2920.2007.01301.x
- Grant, C. S., and Louda, J. W. (2013). Scytonemin-imine, a mahogany-colored UV/Vis sunscreen of cyanobacteria exposed to intense solar radiation. *Org. Geochem.* 65, 29–36. doi: 10.1016/j.orggeochem.2013.09.014
- Grünewald, K., Hirschberg, J., and Hagen, C. (2001). Ketocarotenoid biosynthesis outside of plastids in the unicellular green alga *Haematococcus pluvialis*. *J. Biol. Chem.* 276, 6023–6029. doi: 10.1074/jbc.M006400200
- Horath, T., Neu, T., and Bachofen, R. (2006). An endolithic microbial community in dolomite rock in Central Switzerland: characterization by reflection spectroscopy, pigment analyses, scanning electron microscopy, and laser scanning microscopy. *Microb. Ecol.* 51, 353–364. doi: 10.1007/s00248-006-9051-y
- Jehlička, J., Culka, A., and Mareš, J. (2019). Raman spectroscopic screening of cyanobacterial chasmoliths from crystalline gypsum—the Messinian crisis sediments from Southern Sicily. *J. Raman Spectrosc.* 51, 1802–1812. doi: 10.1002/jrs.5671
- Jehlička, J., Culka, A., and Nedbalová, L. (2016). Colonization of snow by microorganisms as revealed using miniature Raman spectrometers—possibilities for detecting carotenoids of psychrophiles on Mars? *Astrobiology* 16, 913–924. doi: 10.1089/ast.2016.1487
- Jehlička, J., Edwards, H. G. M., and Oren, A. (2014a). Raman spectroscopy of microbial pigments. *Appl. Environ. Microbiol.* 80, 3286–3295. doi: 10.1128/AEM.00699-14
- Jehlička, J., Edwards, H. G. M., Osterrothová, K., Novotná, J., Nedbalová, L., Kopecký, J., et al. (2014b). Potential and limits of Raman spectroscopy for carotenoid detection in microorganisms: implications for astrobiology. *Philos. Trans. A Math. Phys. Eng. Sci.* 372:20140199. doi: 10.1098/rsta.2014.0199
- Johansen, J. R., Hentschke, G. S., Pietrasiak, N., Rigonato, J., Fiore, M. F., and Sant'Anna, C. L. (2017). *Komarekiella atlantica* gen. et sp. nov. (Nostocaceae, Cyanobacteria): a new subaerial taxon from the Atlantic rainforest and Kauai, Hawaii. *Fottea* 17, 178–190. doi: 10.5507/fot.2017.002
- Jung, W., Campbell, R. L., Gwak, Y., Kim, J. I., Davies, P. L., and Jin, E. (2016). New cysteine-rich ice-binding protein secreted from Antarctic microalga, *Chloromonas* sp. *PLoS One* 11:e0154056. doi: 10.1371/journal.pone.0154056
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Komárek, J. (2013). “Cyanoprokaryota: 3rd part: heterocystous genera” in *Süßwasserflora von Mitteleuropa, Bd 19 (3)*, eds. B. Büdel, G. Gärtner, L. Krienitz and M. Schagerl (Berlin: Springer Spektrum), 1–1130.
- Komárek, J. (2016). A polyphasic approach for the taxonomy of cyanobacteria: principles and applications. *Eur. J. Phycol.* 51, 346–353. doi: 10.1080/09670262.2016.1163738
- Komárek, J., and Anagnostidis, K. (1999). “Cyanoprokaryota. I. Chroococcales” in *Süßwasserflora von Mitteleuropa, Begründet von A. PascherBd, 19/3 Cyanoprokaryota, 1. Teil Chroococcales*, eds. H. Ettl, G. Gärtner, H. Heynig and D. Mollenhauer (Heidelberg: Spektrum Akademischer Verlag), 1–548.
- Komárek, J., and Anagnostidis, K. (2005). “Cyanoprokaryota. 2. Teil: Oscillatoriales” in *Süßwasserflora von Mitteleuropa, Bd 19 (2)*, eds. B. Büdel, G. Gärtner, L. Krienitz and M. Schagerl (München: Elsevier GmbH), 1–759.
- Kurmayer, R., Christiansen, G., Holzinger, A., and Rott, E. (2018). Single colony genetic analysis of epilithic stream algae of the genus *Chamaesiphon* spp. *Hydrobiologia* 811, 61–75. doi: 10.1007/s10750-017-3295-z
- Kuzmany, H. (1980). Resonance Raman scattering from neutral and doped polyacetylene. *Phys. Stat. Solid. B* 97, 521–531. doi: 10.1002/psb.2220970217
- Lamprinou, V., Hernández-Mariné, M., Canals, T., Kormas, K., Economou-Amilli, A., and Pantazidou, A. (2011). Morphology and molecular evaluation of *Iphinoe spelaebios* gen. nov., sp. nov. and *Loriellopsis cavernicola* gen. nov., sp. nov., two stigonematalean cyanobacteria from Greek and Spanish caves. *Int. J. Syst. Evol. Microbiol.* 61, 2907–2915. doi: 10.1099/ijs.0.029223-0
- Lara, Y. J., McCann, A., Malherbe, C., François, C., Demoulin, C. F., Sforza, M. C., et al. (2022). Characterization of the halochromic Gloeocapsin pigment, a cyanobacterial biosignature for paleobiology and astrobiology. *Astrobiology* 22, 735–754. doi: 10.1089/ast.2021.0061
- Maguregui, M., Knuutinen, U., Trebolazabala, J., Morillas, H., Castro, K., Martínez-Arkarazo, I., et al. (2012). Use of *in situ* and confocal Raman spectroscopy to study the nature and distribution of carotenoids in brown patinas from a deteriorated wall painting in Marcus Lucretius house (Pompeii). *Anal. Bioanal. Chem.* 402, 1529–1539. doi: 10.1007/s00216-011-5276-9
- Mareš, J., Hrouzek, P., Kaňa, R., Ventura, S., Strunecký, O., and Komárek, J. (2013). The primitive thylakoid-less cyanobacterium *Gloeobacter* is a common rock-dwelling organism. *PLoS One* 8:e66323. doi: 10.1371/journal.pone.0066323
- Mareš, J., Lara, Y., Dadáková, I., Hauer, T., Uher, B., Wilmette, A., et al. (2015). Phylogenetic analysis of cultivation-resistant terrestrial cyanobacteria with massive sheaths (*Stigonema* spp. and *Petalonema alatam*, Nostocales, Cyanobacteria) using single-cell and filament sequencing of environmental samples. *J. Phycol.* 51, 288–297. doi: 10.1111/jpy.12273
- Marotz, C., Sharma, A., Humphrey, G., Gittel, N., Daum, C., Gilbert, J. A., et al. (2019). Triplet PCR reactions for 16S rRNA gene amplicon sequencing are unnecessary. *Biotechniques* 67, 29–32. doi: 10.2144/btn-2018-0192
- Matsuzaki, R., Kawai-Toyooka, H., Hara, Y., and Nozaki, H. (2015). Revisiting the taxonomic significance of aplanozygote morphologies of two cosmopolitan snow species of the genus *Chloromonas* (Volvocales, Chlorophyceae). *Phycologia* 54, 491–502. doi: 10.2216/15-33.1
- McMurdie, P. J., and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217
- Meslier, V., Casero, M. C., Dailey, M., Wierzos, J., Ascaso, C., Artieda, O., et al. (2018). Fundamental drivers for endolithic microbial community assemblies in the hyperarid Atacama Desert. *Environ. Microbiol.* 20, 1765–1781. doi: 10.1111/1462-2920.14106

- Miscoe, L. H., Johansen, J. R., Kocielek, J. P., Lowe, R. L., Vaccarino, M. A., Pietrasiak, N., et al. (2016). The diatom flora and cyanobacteria from caves on Kauai, Hawaii. *Acta Bot. Hung.* 58, 3–4.
- Nedbalová, L., Mihál, M., Kviderová, J., Procházková, L., Řezanka, T., and Elster, J. (2017). Identity, ecology and ecophysiology of planktic green algae dominating in ice-covered lakes on James Ross island (northeastern Antarctic peninsula). *Extremophiles* 21, 187–200. doi: 10.1007/s00792-016-0894-y
- Němečková, K., Culka, A., and Jehlička, J. (2022). Detecting pigments from gypsum endoliths using Raman spectroscopy: from field prospection to laboratory studies. *J. Raman Spectrosc.* 53, 630–644. doi: 10.1002/jrs.6144
- Němečková, K., Culka, A., Němec, I., Edwards, H. G. M., Mareš, J., and Jehlička, J. (2021). Raman spectroscopic search for scytonemin and gloeocapsin in endolithic colonizations in large gypsum crystals. *J. Raman Spectrosc.* 52, 2633–2647. doi: 10.1002/jrs.6186
- Němečková, K., Jehlička, J., and Culka, A. (2020). Fast screening of carotenoids of gypsum endoliths using portable Raman spectrometer (Messinian gypsum, Sicily). *J. Raman Spectrosc.* 51, 1127–1137. doi: 10.1002/jrs.5891
- Oksanen, J. (2009). *Vegan: Community ecology package*. R package version 1.15–14. Available at: <http://CRAN.R-project.org/package=vegan>.
- Orellana, G., Gómez-Silva, B., Urrutia, M., and Galetović, A. (2020). UV-A irradiation increases Scytonemin biosynthesis in Cyanobacteria inhabiting halites at Salar Grande, Atacama Desert. *Microorganisms* 8:1690. doi: 10.3390/microorganisms8111690
- Osterrothová, K., Culka, A., Němečková, K., Kaftan, D., Nedbalová, L., Procházková, L., et al. (2019). Analyzing carotenoids of snow algae by Raman microspectroscopy and high-performance liquid chromatography. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 212, 262–271. doi: 10.1016/j.saa.2019.01.013
- Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023
- Pitonzo, B. J., Amy, P. S., and Rudin, M. (1999). Effect of gamma radiation on native endolithic microorganisms from a radioactive waste deposit site. *Radiat. Res.* 152, 64–70. doi: 10.2307/3580050
- Posada, D. (2008). jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253–1256. doi: 10.1093/molbev/msn083
- Procházková, L., Leya, T., Křížková, H., and Nedbalová, L. (2019). *Sanguina nivaloides* and *Sanguina aurantia* gen. et spp. nov. (Chlorophyta): the taxonomy, phylogeny, biogeography and ecology of two newly recognised algae causing red and orange snow. *FEMS Microbiol. Ecol.* 95:fiz064. doi: 10.1093/femsec/fiz064
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219
- Rhind, T., Ronholm, J., Berg, B., Mann, P., Applin, D., Stromberg, J., et al. (2014). Gypsum-hosted endolithic communities of the Lake St. Martin impact structure, Manitoba, Canada: spectroscopic detectability and implications for Mars. *Int. J. Astrobiol.* 13, 366–377. doi: 10.1017/S1473550414000378
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria. Available at: <http://www.R-project.org/>
- Rstudio Team. (2021). *RStudio: Integrated Development Environment for R*. Boston, MA. Available at: <http://www.rstudio.com/>
- Rott, E., Holzinger, A., Gesierich, D., Kofler, W., and Sanders, D. (2010). Cell morphology, ultrastructure, and calcification pattern of *Oocardium stratum*, a peculiar lotic desmid. *Protoplasta* 243, 39–50. doi: 10.1007/s00709-009-0050-y
- Sanmartín, P., DeAraujo, A., and Vasanthakumar, A. (2018). Melding the old with the new: trends in methods used to identify, monitor, and control microorganisms on cultural heritage materials. *Microb. Ecol.* 76, 64–80. doi: 10.1007/s00248-016-0770-4
- Segawa, T., Matsuzaki, R., Takeuchi, N., Akiyoshi, A., Navarro, F., Sugiyama, S., et al. (2018). Bipolar dispersal of red-snow algae. *Nat. Commun.* 9:3094. doi: 10.1038/s41467-018-05521-w
- Smith, H. D., Baqué, M., Duncan, A. G., Lloyd, C. R., McKay, C. P., and Billi, D. (2014). Comparative analysis of cyanobacteria inhabiting rocks with different light transmittance in the Mojave Desert: a Mars terrestrial analogue. *Int. J. Astrobiol.* 13, 271–277. doi: 10.1017/S1473550414000056
- Sorrels, C. M., Proteau, P. J., and Gerwick, W. H. (2009). Organization, evolution, and expression analysis of the biosynthetic gene cluster for scytonemin, a cyanobacterial UV-absorbing pigment. *Appl. Environ. Microbiol.* 75, 4861–4869. doi: 10.1128/AEM.02508-08
- Stamatakis, A. (2014). RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stivaletta, N., Lopez-Garcia, P., Boihem, L., Millie, D. F., and Barbieri, R. (2010). Biomarkers of endolithic communities within gypsum crusts (southern Tunisia). *Geomicrobiol. J.* 27, 101–110. doi: 10.1080/01490450903410431
- Storme, J.-Y., Golubic, S., Wilmotte, A., Kleinteich, J., Velázquez, D., and Javaux, E. J. (2015). Raman characterization of the UV-protective pigment gloeocapsin and its role in the survival of cyanobacteria. *Astrobiology* 15, 843–857. doi: 10.1089/ast.2015.1292
- Strunecký, O., Ivanova, A. P., and Mareš, J. (2022). An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis. *J. Phycol.* 59, 12–51. doi: 10.1111/jpy.13304
- Tang, Y., Lian, B., Dong, H., Liu, D., and Hou, W. (2012). Endolithic bacterial communities in dolomite and limestone rocks from the Nanjiang canyon in Guizhou karst area (China). *Geomicrobiol. J.* 29, 213–225. doi: 10.1080/01490451.2011.558560
- Taton, A., Grubisic, S., Brambilla, E., De Wit, R., and Wilmotte, A. (2003). Cyanobacterial diversity in natural and artificial microbial mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a morphological and molecular approach. *Appl. Environ. Microbiol.* 69, 5157–5169. doi: 10.1128/AEM.69.9.5157-5169.2003
- Villar, S. E. J., Edwards, H. G. M., and Cockell, C. S. (2005). Raman spectroscopy of endoliths from Antarctic cold desert environments. *Analyst* 130, 156–162. doi: 10.1039/b410854j
- Vitek, P., Ascaso, C., Artieda, O., Casero, M. C., and Wierzchos, J. (2017). Discovery of carotenoid red-shift in endolithic cyanobacteria from the Atacama Desert. *Sci. Rep.* 7:11116. doi: 10.1038/s41598-017-11581-7
- Vitek, P., Ascaso, C., Artieda, O., Casero, M. C., and Wierzchos, J. (2020). Raman imaging of microbial colonization in rock—some analytical aspects. *Anal. Bioanal. Chem.* 412, 3717–3726. doi: 10.1007/s00216-020-02622-8
- Vitek, P., Jehlička, J., Ascaso, C., Mašek, V., Gómez-Silva, B., Olivares, H., et al. (2014b). Distribution of scytonemin in endolithic microbial communities from halite crusts in the hyperarid zone of the Atacama Desert, Chile. *FEMS Microbiol. Ecol.* 90, 351–366. doi: 10.1111/1574-6941.12387
- Vitek, P., Jehlička, J., Edwards, H. G. M., Hutchinson, I., Ascaso, C., and Wierzchos, J. (2012). The miniaturized Raman system and detection of traces of life in halite from the Atacama desert: some considerations for the search for life signatures on Mars. *Astrobiology* 12, 1095–1099. doi: 10.1089/ast.2012.0879
- Vitek, P., Jehlička, J., Edwards, H. G. M., Hutchinson, I., Ascaso, C., and Wierzchos, J. (2014a). Miniaturized Raman instrumentation detects carotenoids in Mars-analogue rocks from the Mojave and Atacama deserts. *Philos. Trans. A Math. Phys. Eng. Sci.* 372:20140196. doi: 10.1098/rsta.2014.0196
- Walker, J. J., and Pace, N. R. (2007). Endolithic microbial ecosystems. *Annu. Rev. Microbiol.* 61, 331–347. doi: 10.1146/annurev.micro.61.080706.093302
- Wegley, L., Edwards, R., Rodriguez-Brito, B., Liu, H., and Rohwer, F. (2007). Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*. *Environ. Microbiol.* 9, 2707–2719. doi: 10.1111/j.1462-2920.2007.01383.x
- Whitton, B. A. (2012). *Ecology of cyanobacteria II: their diversity in space and time*. Dordrecht: Springer Science and Business Media.
- Wierzchos, J., Ascaso, C., and McKay, C. P. (2006). Endolithic cyanobacteria in halite rocks from the hyperarid core of the Atacama Desert. *Astrobiology* 6, 415–422. doi: 10.1089/ast.2006.6.415
- Wierzchos, J., Casero, M. C., Artieda, O., and Ascaso, C. (2018). Endolithic microbial habitats as refuges for life in polyextreme environment of the Atacama Desert. *Curr. Opin. Microbiol.* 43, 124–131. doi: 10.1016/j.mib.2018.01.003
- Wierzchos, J., DiRuggiero, J., Vitek, P., Artieda, O., Souza-Egipsy, V., Škaloud, P., et al. (2015). Adaptation strategies of endolithic chlorophototrophs to survive the hyperarid and extreme solar radiation environment of the Atacama Desert. *Front. Microbiol.* 6:934. doi: 10.3389/fmicb.2015.00934
- Withnall, R., Chowdhry, B. Z., Silver, J., Edwards, H. G. M., and de Oliveira, L. F. (2003). Raman spectra of carotenoids in natural products. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 59, 2207–2212. doi: 10.1016/s1386-1425(03)00064-7
- Wong, F. K., Lau, M. C., Lacap, D. C., Aitchison, J. C., Cowan, D. A., and Pointing, S. B. (2010). Endolithic microbial colonization of limestone in a high-altitude arid environment. *Microb. Ecol.* 59, 689–699. doi: 10.1007/s00248-009-9607-8
- Yung, C. C., Chan, Y., Lacap, D. C., Pérez-Ortega, S., de los Rios-Murillo, A., Lee, C. K., et al. (2014). Characterization of chasmoendolithic community in miers valley, Mcmurdo dry valleys, Antarctica. *Microb. Ecol.* 68, 351–359. doi: 10.1007/s00248-014-0412-7
- Zapomelová, E., Hisem, D., Řeháková, K., Hrouzek, P., Jezberová, J., Komárková, J., et al. (2008). Experimental comparison of phenotypic plasticity and growth demands of two strains from the *Anabaena circinalis/A. crassa* complex (cyanobacteria). *J. Plankton Res.* 30, 1257–1269. doi: 10.1093/plankt/fbn081
- Ziolkowski, L., Mykytczuk, N., Omelon, C., Johnson, H., Whyte, L., and Slater, G. (2013). Arctic gypsum endoliths: a biogeochemical characterization of a viable and active microbial community. *Biogeosciences* 10, 7661–7675. doi: 10.5194/bg-10-7661-2013