

Ghost-rocks' microbiota: metagenomic insights into their influence on the biogeochemistry of karstic cave and groundwater

Guillaume Peugnet^{1,*}, Célie Pisapia¹, Bénédicte Ménéz¹, Maud Watkinson^{1,2}, Léna Lecourt¹, Nicolas Peugnet³, Julien Bouchez¹, Laurent Bruxelles^{4,5}, Emmanuelle Gérard¹

¹Université Paris Cité, Institut de physique du globe de Paris, CNRS, 75005 Paris, France

²BRGM, F-45060 Orléans, France

³Sorbonne Université, CNRS, LIP6, F-75005 Paris, France

⁴Université Jean Jaurès, TRACES, 31058 Toulouse, France

⁵Université Witwatersrand, Johannesburg, South Africa

*Corresponding author. Université Paris Cité, Institut de physique du globe de Paris, CNRS, Paris, France. Email: peugnet@ipgp.fr

Editor: Tillmann Lueders

Abstract

Microbial communities in the critical zone drive key geochemical processes, but many subsurface habitats remain poorly characterized. Ghost-rock karst systems in particular represent unexplored microbial niches. Here, we provide the first genome-resolved metagenomic comparison of ghost-rock and groundwater microbial communities from the Sterkfontein karst system (South Africa). Ghost-rock and groundwater communities host distinct taxonomic and metabolic assemblages. Groundwater communities are dominated by chemolithotrophs capable of oxidizing sulfur- and nitrogen-bearing compounds, and by heterotrophs degrading refractory, plant-derived organic matter. In contrast, primary producers in ghost-rocks likely rely on atmospheric chemosynthesis via trace gas oxidation, while glycogen metabolism and necromass recycling point to adaptations to oligotrophic and fluctuating hydrological conditions. Groundwater taxa with metal-interacting pathways may initiate bedrock colonization via metal oxidation, whereas ghost-rock communities include potential metal reducers that could drive iron and manganese oxide dissolution and influence trace element mobility. Together, these results underscore ghost-rocks as active microbial and geochemical hot spots within karst systems that may play a non-negligible role on biomineralization/bioweathering processes and on shaping (sub)terrestrial landscapes and global biogeochemical cycles.

Keywords cave microbiology; metagenomics; karst; porin-cytochrome complexes; atmospheric chemosynthesis

Introduction

Subsurface ecosystems host most of Earth's micro-organisms (Bar-On et al. 2018), whose activity can significantly influence critical zone (CZ) processes, including rock weathering (Akob and Küsel 2011, Wild et al. 2022). Subsurface ecosystems, such as fractured aquifers, are spatially heterogeneous and dynamic over time, generated as water circulates through fracture networks (Bochet et al. 2020). Significant differences exist between planktonic communities in groundwater and the biofilm communities attached to rocks (Flynn et al. 2013, Lazar et al. 2019), and these communities change in response to environmental cues (Yan et al. 2021). It is thus a major challenge to integrate these spatial and temporal heterogeneities when one seeks to understand processes that influence the dynamics of subsurface environments, particularly in the deeper part of the CZ. Consequently, the contribution of planktonic and biofilm communities to major biogeochemical processes in the deep CZ remain poorly understood (Smith et al. 2018).

In the shallow subsurface, carbonate karst systems are an important component of the CZ covering around 15% of the Earth's ice-free continents and forming complex cave networks (Covington et al. 2023). Caves can be classified according to their lithology and speleogenetic process (e.g. lava tubes caves, epigenic karstic carbonate caves, or hypogenic carbonate caves formed by sulfuric acid speleogenesis). From a microbiological perspective, non-karstic caves such as lava tubes may be comparable to epigenic karstic caves regarding organic carbon availability, and energy sources for microbial community.

Estimates suggest that only 0.02%–1.3% of micro-organisms in carbonate karstic cave are cultivable, depending on the methodology used (Bender et al. 2020). Microbial ecology studies of oligotrophic carbonate caves are limited in part due to very low biomass, with between 10^3 and 10^5 cells/ml in karst springs and cave lakes (Hershey and Barton 2018), and around 10^6 cells per gram of rocks (Barton et al. 2006). After the pioneer works in cave microbiology that offered the first insights in the cave mi-

crobiota composition and structure (Barton et al. 2004), metabarcoding and metagenomics have revealed a distinct cave microbiota, compared to surface soils (Ortiz et al. 2013). Sediments and biofilms on cave walls are dominated by Pseudomonadota, Actinobacteriota, and Acidobacteriota, whereas groundwater communities are dominated by Pseudomonadota, Verrucomicrobiota, and Patescibacteriota (Turrini et al. 2024). Shotgun metagenomics of speleothems (Ortiz et al. 2014), sediment (Wiseschart et al. 2019), or cave wall biofilms (Turrini et al. 2020) have revealed metabolic diversity and the importance of autotrophic organisms in these ecosystems. Microbial activity has been shown to influence greenhouse gas cycling (Martin-Pozas et al. 2022), and nutrient and elemental cycling, particularly of iron and manganese (Carmichael and Bräuer 2015).

A microbial origin has been suggested for ferromanganese minerals often encountered as crusts on cave walls (Spilde et al. 2005, Gázquez et al. 2011, Fonollá et al. 2020), stromatolite-like Mn patinas (Bernardini et al. 2021), and sediments or speleothems (Kotula et al. 2019). These oxides are thought to form via chemolithoautotrophic organisms that would mine reduced elements present in the bedrock and oxidize them to gain energy, therefore contributing to rock weathering. While iron-based chemolithoautotrophy has long been described (Winogradsky 1888), manganese-based chemolithoautotrophy is more recently documented (Yu and Leadbetter 2020). Alternative heterotrophic pathways for manganese oxides formation, involving reactive oxygen species (ROS) production, or multicopper oxidases (MCOs) have also been proposed (Carmichael and Bräuer 2015).

In this context, caves formed via ghost-rock karstification are of high interest for studying the interplay between microbial activity and mineral weathering. This new model of karstification is a two step process with (i) an isovolumetric chemical weathering of the bedrock that takes place under low hydrodynamic conditions: the bedrock, when submerged, undergoes dissolution of soluble ions, while leaving behind a residual alterite (the insoluble fraction), highly porous (up to 65% porosity), called ghost-rock; (ii) physical erosion triggered either by the lowering of the saturated zone, inducing an energetic water-flow capable of removing ghost-rock, or by the gravitational collapse of ghost-rock masses, thus creating caves and underground rivers characteristic of karstic systems (Dubois et al. 2014). Interestingly and following this multi-stage karstification process, some pockets of ghost-rocks may nevertheless remain in place without being completely emptied, representing considerable volumes of weathered rock, far greater than previously thought (e.g. up to 200 million m³ in the French Larzac system, Baral et al. 2024). They would thus constitute a largely unexplored component of karst systems, and their implications for the functioning of these karsts remain to be elucidated. In particular, due the high porosity of the residual alterite, they may act as hidden water reservoirs in the vadose zone of karsts (Champollion et al. 2018, Durand et al. 2025), contributing to global water cycle and more specifically to the drinking water resources of populations. However, the high porosity and water content of these ghost-rocks pockets make them ideal habitats for micro-organisms. Then, they might also be considered as putative niches for subsurface microbial ecosystems in which they may influence mineral weathering and/or biomineralization processes, either in bedrocks or in alterites, playing then a role on elemental (re)mobilization along the karst history. In particular, the identification of secondary metallic oxides (Fe, Mn) in these alterites, of which a microbial origin has

already been proposed (Dandurand et al. 2014, Dubois et al. 2014, Pisapia et al. 2021) raises the question of the impact of this putative microbial activity on rock weathering and metal cycling at depth, and on its broader influence on global metal biogeochemical cycles. However, to date, the microbial metabolic potential in ghost-rocks has never been studied.

The main objective of this study was then to investigate what kind and to which extent microbial ecosystems of a ghost-rock karst system may impact rock weathering/mineralization processes and influence the geochemistry of groundwaters, notably through metal cycling. For this purpose, we focused on the Sterkfontein karst system (South Africa) that presents one of the most impressive volume of ghost-rocks ever described (Bruxelles 2018). These alterites were preliminarily described to host large amounts of iron and manganese oxides jointly with the presence of micro-organisms whose metabolisms were still unknown (Pisapia et al. 2021). The objectives were then to achieve the first exhaustive description of the microbial ecology of this ghost-rock system, including alterites and groundwaters, and to highlight the main metabolic potentials at stake in these compartments of the cave system. Based on the working hypothesis that microbially-induced metal cycling might be central in these environments, either *via* the oxidation of trace metals from the dolomitic bedrocks as proposed by Spilde et al. (2005) or *via* the reduction of metal-oxides from the alterites as electron acceptors, we then specifically focused on the description of microbial metabolic potentials linked with iron and manganese cycling. To answer these questions, we used genome-resolved shotgun metagenomics, alongside with geochemical analyses of the studied caves. Notably, we used a global approach at the level of reads, contigs and MAGs, with a particular focus on metal-interacting pathways, including metallophores biosynthetic gene clusters (BGCs), multicopper oxidases (MCOs), and extracellular electron transfer (EET) systems.

Methods

Study site

The ghost-rock karst system of Sterkfontein in South Africa is located 35 km northwest of Johannesburg, and is part of the Cradle of Humankind UNESCO World heritage site (Fig. 1A). It develops in metric to pluri-metric dolomitic beds from the Malmani Subgroup (2.64–2.50 Gy), intercalated with centimetric chert beds. The system is highly labyrinthine, extending for more than 5 km, mostly along bedrock fractures, and harbors more than 400 caves (Bruxelles 2018). A tectonic uplift during the Oligocene led to incision by the Blaubank river, which, despite a low incision rate (down to 3 m per My throughout the Neogene and Quaternary), progressively removed ghost-rock, and opened the present cave system (Bruxelles 2018). This karst system presents one of the most important ghost-rock example ever discovered, with large volumes of ghost-rock pockets, in place in the caves (Fig. 1B) including at the contact with the water table, and easily accessible. It was then an ideal system where performing a first metagenomic study of these peculiar environments.

In that perspective, three caves from this wide karstic system were investigated: Sterkfontein Cave, Fault Cave and Glover's Cave (Fig. 1A). Sterkfontein and Fault Caves belong to the same

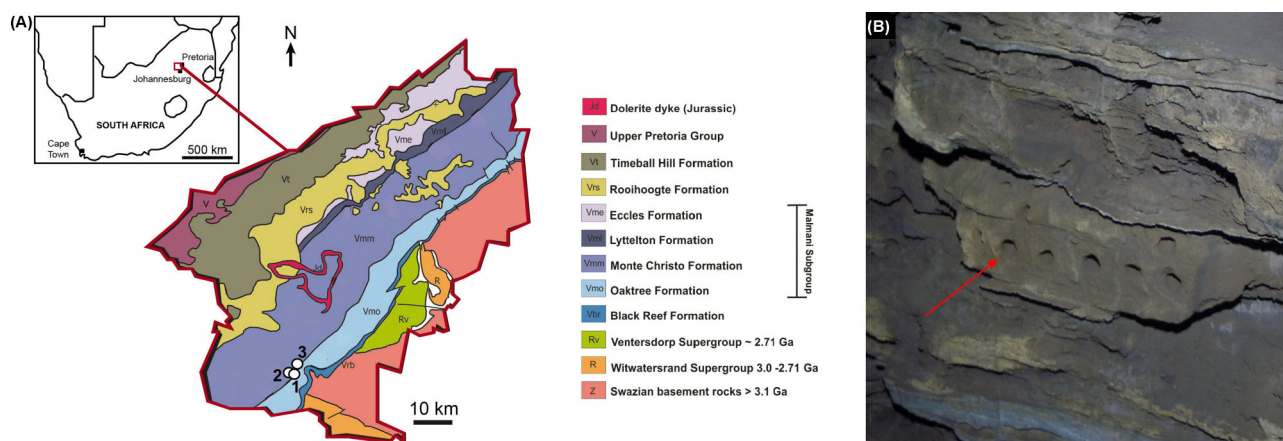


Figure 1 Sampling site localization and description. (A) Simplified geological map of the Cradle of Humankind, NW of Johannesburg, developed in the Transvaal Supergroup (modified from Makhubela et al. 2017). The location of the three studied caves, Sterkfontein Cave (1), Fault Cave (2), and Glover's Cave (3) is provided. (B) Photograph of a typical outcrop of porous ghost-rock in Fault Cave, sampled using 50 ml Falcon™ tubes (visible as holes, red arrow). The outcrop exhibits ghost-rock developed through alteration of dolomitic beds, with residual chert layers forming more resistant horizons. Thin, millimetric, siliceous partitions are visible within ghost-rock pockets, contributing to the structuring and stability of the outcrop.

network, the first one being currently open to tourism, while Fault cave, not accessible to the public, can be considered as a more pristine environment. Glover's Cave, located a few kilometers away in the Kromdraai system, but within the same geological context, was also investigated for representativity (Fig. 1A).

Sampling

All samples analyzed in this study (Supplementary Table S1) were collected during a field campaign in 2019. To investigate the microbial ecology of the karst system, both groundwater and ghost-rocks were sampled for metagenomic analyses in the targeted caves.

For each cave, 15 l of groundwater were collected at the water table using 5 l plastic bottles rinsed three times beforehand with *in situ* cave groundwater in order to limit external contamination. Groundwater samples were then filtered on the day of collection, using vacuum-pump filtration systems. All samples were first filtered through 3 µm Whatman® filters to remove large particles and aggregates followed by 0.2 µm Whatman® filters to recover microbial cells. Both sets of filters were then stored at –20°C until return to the laboratory and subsequent DNA extraction. During filtration, 500 ml of filtrate from each cave were collected in 100 ml Nalgene™ bottles (that were previously cleaned with three successive nitric acid wash) and acidified using a drop of 7 N nitric acid (HNO₃) for triplicated elemental analyses.

Ghost-rock samples were collected from each cave walls using 50 ml Falcon™ tubes, easily inserted due to the loose nature of the ghost-rock (Fig. 1B). The tubes were directly stored at –20°C until return to the laboratory for subsequent DNA extraction, and selected ghost-rock samples from Fault Cave were also used for geochemical analyses. Unaltered embedded bedrocks, either dolomites (two samples at Sterkfontein Cave) and cherts (one sample at Sterkfontein Cave and three at Fault Cave) collected during previous field campaigns, were also selected as primary, intact reference lithologies for geochemical analyses.

Geochemical analyses

Filtered groundwater samples from each cave were diluted using 1 M HNO₃ to obtain three final aliquots per cave at 0.5 M HNO₃, similarly to blanks and standards used for quantitative-inductively coupled plasma mass spectrometry (Q-ICP-MS) analyses.

Selected samples of embedded pristine dolomites and cherts, and ghost-rocks were finely crushed using a clean agate mortar. As most samples contained both carbonate and silicate fractions in variable proportions, 200 mg of each sample were used for sequential leaching. The carbonate phase was first digested in clean Teflon® vials using 0.5 M hydrochloric acid (HCl). After 10 min of sonication and overnight equilibration, solutions were evaporated to dryness during 24 h at 85°C. Solid residues were rinsed twice with Milli-Q® water, and the supernatants were recovered after centrifugation to obtain the carbonate leach for each sample. The remaining solid residues were evaporated to dryness overnight at 75°C. These residues, along with chert samples, were then digested in Teflon® vials using 1 ml of 27 M hydrofluoric acid (HF), 0.6 ml of 16 M HNO₃ and 300 µl of Milli-Q® water. Samples were left 24 h at 110°C, then evaporated to dryness overnight at 110°C. To neutralize fluorides and oxides, sample residues were treated with either aqua regia (i.e. concentrated HNO₃ and HCl mixture) or 12 M HCl during 24 h to 48 h at 125°C or 110°C, respectively. After complete digestion, samples were evaporated to dryness, then diluted with 16 M HNO₃ and heated at 120°C during 24–48 h.

All groundwater and solid samples were analyzed using Q-ICP-MS on an Agilent 7900 at the IPGP PARI platform (Paris, France) using standard bracketing method with SLRS-6 standard (NRC of Canada) for groundwater and USGS BCR-2 reference material for solid samples. Reproducibility on standards was commonly better than 10% for groundwater and 3% for solids. Blank analyses (HNO₃ 0.5 M) were performed after every eight samples. Results are detailed in Supplementary Table S2.

Total organic carbon (TOC) analyses were conducted at the Service d'analyse des roches et des minéraux (SARM, CRPG, Nancy,

France) using 250 mg of each solid sample that were decarbonated using water and 2% HCl, then filtered to remove chlorides. TOC was measured with a HORIBA EMIA-320V2 carbon/sulfur analyzer via infrared light absorption during combustion in an oxygen flow.

DNA extraction and sequencing

DNA extractions for both groundwater and ghost-rock samples were performed at IPGP (Paris, France). For groundwater samples, DNA was extracted from both the 3 μm and 0.2 μm filters collected at each of the three sampled caves (for a total of six samples; [Supplementary Table S1](#)) using the DNeasy® PowerWater® Kit (QIAGEN), following the manufacturer's protocol. For ghost-rock samples (two samples for Sterkfontein Cave, one sample for Fault Cave and one sample for Glover's cave, with four extraction replicates for each sample) DNA was extracted from 10 g sample aliquots using the DNeasy® PowerMax® Soil Kit (QIAGEN), following the manufacturer's protocol.

Extracted DNA concentrations were quantified using a Qubit™ 2.0 fluorometer with the Qubit™ dsDNA High Sensitivity Assay Kit (Thermo Fisher Scientific), and purity was assessed with a NanoDrop™ spectrophotometer (Thermo Fisher Scientific). All extraction blanks were below detection limit (<0.005 ng/ μl). Sample's extractions below detection limits were discarded. Only the samples with measurable concentrations were kept for subsequent pooling and concentration using a SpeedVac™ system (Thermo Fisher Scientific) to obtain sufficient DNA for shotgun metagenomic sequencing. DNA concentrations before pooling and concentration ranged from 0.018 to 0.140 ng/ μl for ghost-rock samples and from 0.538 ng/ μl to 8.200 ng/ μl for groundwater. The Glover Cave ghost-rock sample and the 3 μm filter from Fault Cave yielded insufficient DNA and were thus excluded from downstream analysis. The remaining samples were sent to Eurofins Genomics™ for sequencing on an Illumina NovaSeq platform (2 $\times$ 150 base pair, bp). Read number per sample are listed in [Supplementary Table S1](#). Nonpareil3 was used for the diversity and coverage estimates (Rodriguez-R et al. 2018) and results are listed in the [Supplementary Table S1](#). Average coverage for all samples was of $55\pm 9\%$, based on a sequencing effort of 7.3 ± 1.2 Gbp, providing a solid basis for downstream analyses. However, increasing the sequencing effort to approximately 20 Gbp, as suggested by Nonpareil3, would likely have improved the coverage. All raw sequencing reads are deposited in the NCBI Sequence Read Archive under the Bioproject PRJNA1298313.

Metagenomic analysis

This study combined metagenomic analyses at the read, contig and MAG levels to strengthen biological interpretations. Due to DNA extraction challenges and the limited number of samples, analyses at multiple levels provided complementary perspectives. The MAG level allows an integrated view of the metabolic pathways, but captures only a fraction of the total sequences. In contrast, read and contig level analyses, enabling the detection of specific genes, offer greater representativity. Thus, even though genome-scale pathway reconstruction is not possible, they can help confirm pathway prevalence inferred from MAGs by finding differentially abundant genes.

Raw sequences were quality-trimmed using Cutadapt (Martin 2011), with the following parameters: -m 50 -max-n 4 -nextseq-trim 20. Functional profiling was performed with Humann 3.0 against the UniRef90 database, yielding 12%–47% mapped reads across samples. Pathways were regrouped according to MetaCyc definitions (Caspi et al. 2020). The total coverage depth (in reads per kilobase) was normalized to count per millions (cpm), as recommended in the Humann documentation (Beghini et al. 2021). Two co-assemblies were generated, one for groundwater samples, and one for ghost-rock samples, using MEGAHIT (Li et al. 2015) with the following parameters: -min-contig-len 1000 -presets meta-large. Taxonomic annotation of the contigs was done using the “taxonomy” tool of MMseqs2, based on the GTDB database, with default parameters (Mirdita et al. 2021). About 96.4% and 96.9% of the contigs got a taxonomic assignment at the superkingdom level at least for the groundwaters and rocks co-assemblies, respectively. Community analysis was then based on the percentage of total contig coverage. Functional annotation of the contigs was done using Prokka v1.14.5 (Seemann 2014) and KofamScan (Aramaki et al. 2020). Significant differentially abundant KEGG Orthologs (KOs) between ghost-rock and groundwater samples were identified using LEfSe (Segata et al. 2011) with default parameters. It performs a Kruskal–Wallis sum-rank test followed by linear discriminant analysis (LDA) to evaluate effect size of KEGG annotations.

Contigs coverage was calculated using Bowtie2 (Langmead and Salzberg 2012), and binning was done with MetaBat2 (Kang et al. 2019) with a minimum contig length of 1500 bp. Bins were manually refined using Anvi'o (Eren et al. 2015), dereplicated with dRep (Olm et al. 2017), quality-controlled with CheckM (Parks et al. 2015), and taxonomically classified using GTDB-Tk (Chaumeil et al. 2020). This process yielded 273 MAGs of good quality ($> 50\%$ completeness, $< 5\%$ contamination), recruiting between 18% and 62% of reads from groundwater samples, and between 29% and 45% of reads from ghost-rock samples. Among these, the 94 high-quality MAGs ($> 90\%$ completeness, $< 5\%$ contamination) were deposited in NCBI under the Bioproject PRJNA1298313. The taxonomy, completeness, contamination, NCBI accession numbers and N50 values of the 273 MAGs are listed in [Supplementary Table S3](#). The prefix of the MAG names (South Africa Rocks, SAR or South Africa Water, SAW) differentiates those generated using the ghost-rocks co-assembly from those generated using the groundwaters co-assembly.

To estimate MAG coverage per sample types, mean contig coverage values provided by Anvi'o were normalized for each sample by its global mean contig coverage. Then, normalized MAG coverage values were averaged across the five groundwater samples, and the three ghost-rock samples.

Concerning functional gene analyses, genes related to autotrophy were identified with hmmsearch (Eddy, 2011) using HMM profiles from Zhou et al. (2022b). Carbon fixation through the hydroxypropionate/hydroxybutyrate cycle was investigated by Basic Local Alignment Search Tool (BLAST) using the sequences from *Nitrosopumilus maritimus* (Könneke et al. 2014, Otte et al. 2015). The classification of hydrogenases was conducted using the HydDB web server (Søndergaard et al. 2016). AntiSMASH v7.0.1 (Blin et al. 2023) was used to search for secondary metabolite biosynthetic gene clusters, with the -hmm-detection-strictness parameter defined as “strict”. Carbohydrate-active enzymes (CAZymes) were identified with dbCAN3 (Zheng et al. 2023) with default parameters.

ters. The CAZymes belonging to the GH13, GH57, GH77, GT5 and GT35 families were analyzed specifically, as they are characteristic for α -glucan degradation (Colpaert et al. 2020). A MAG is defined as a glycogen synthesizer if it possesses one of the following CAZyme combinations: [GT5 or GH77 or (GH13_6 + GH13_3)] + (GH13_8 or GH13_9 or GH57), based on Colpaert et al. (2020). This CAZyme-based approach was compared to a KofamScan-based approach, where glycogen synthetizers/degraders are detected by complete Kegg modules M00854-M00855. The preponderance of glycogen degraders/synthetizers MAGs was highlighted by their percentage of the total MAGs coverage.

MAG phylogenomic analysis was done using 40 marker genes extracted with fetchMG v1.2 (<https://github.com/motu-tool/fetchMGs>), aligned with Multiple Alignment using Fast Fourier Transform (MAFFT) using default parameters (Kato et al. 2002), and concatenated with catfasta2phyml.pl (<https://github.com/nylander/catfasta2phyml>). Phylogenetic reconstruction was performed with IQ-TREE (Minh et al. 2020) with ModelFinder (Kalyaanamoorthy et al. 2017) to select the best-fit model (LG+F+R10), and statistical support was assessed with 1000 ultrafast bootstrap (Hoang et al. 2018) and 1000 SH-like approximate likelihood ratio test replicates (Guindon et al. 2010). The alignment contained 16 969 parsimony-informative sites. The resulting tree was colored according to GTDB-Tk taxonomic assignments.

Search for metal-cycling metabolic pathways

Considering the putative importance of metal redox cycling in the Sterkfontein karst system, several classes of metal-interacting pathways were systematically screened in reconstructed MAGs.

First, micro-organisms are known to synthesize metallophores, which as specialized chelators for nutrient acquisition through metal chelation and mobilization (Kraemer et al. 2015). A metallophore BGC is a clustered group of genes that together encode a biosynthetic pathway for the production of a metallophore. They were searched with antiSMASH v7, which incorporates updated detection rules for additional metallophore clusters (Blin et al. 2023).

Second, MCOs are a type of oxidase that use Cu cofactors for the oxidation of a wide range of substrates, transferring the electrons to O₂. An Mn-oxidizing activity has been described in a few MCOs (Zhou and Fu 2020). They were searched using Prokka functional annotations.

Third, EET systems allow micro-organisms to use metals as terminal electron donors or acceptors (Shi et al. 2016). Most of the known EET systems typically involve multiple c-type cytochromes, in particular multiheme cytochromes (MHCs) (Shi et al. 2016). MHCs are often found in porin-cytochrome complexes (PCCs) (Hartshorne et al. 2009, Richardson et al. 2012, Shi et al. 2014, He et al. 2017), which have been described in numerous metal oxidizing or reducing micro-organisms (Szeinbaum et al. 2020, Eddie et al. 2022, Yu et al. 2022), including electroactive organisms from sediments and groundwaters (Arbour et al. 2020). Furthermore, most PCCs in both metal-oxidizers and -reducers are centered on at least one MHC with ≥ 10 heme-binding motifs (Shi et al. 2016). The presence of such MHCs has been proposed either to predict microbial iron reduction capacity (Hernsdorf et al. 2017) or to help distinguish iron-oxidizing bacteria from nitrite-oxidizers

among the Gallionellaceae (Hoover et al. 2023). Based on these insights, the search for PCCs in MAGs was done using three criteria, (i) the presence of one MHC with at least 10 heme-binding sites, (ii) the presence of at least one additional c-type cytochrome in close genomic proximity and (iii) the presence of one porin predicted to contain at least 10 β -strand transmembrane domains.

C-type cytochromes were detected using the heme_counter.pl script (https://github.com/seanmcallister/heme_counter), modified to find any CX₂₋₄CH motif (since variants to the canonical CXXCH motif are known; Sharma et al. 2010, Edwards et al. 2020). PRED-TMBB2 (using the MSA-version and HMM option) was used to search for putative porins (Tsirigos et al. 2016). Cellular localization of each PCC components was inferred with PSORTb v3.0.3 was used (Yu et al. 2010). All identified MHCs were clustered using MM-seqs2 (Steinegger and Söding 2017) with default parameters, corresponding to a minimum sequence identity of 90%. Genomic context and gene cluster illustrations were generated using the Python package pyGenomeViz (<https://github.com/moshi4/pyGenomeViz>) (Gilchrist and Chooi 2021).

As Cyc2 is a widely used protein for iron oxidation (McAllister et al. 2020, Zhou et al. 2022a), FeGenie (Garber et al. 2020) was used with default parameters for their characterization in the MAGs.

Results

Geochemical analyses

The geochemical composition of groundwater samples (Supplementary Table S2) reflected the dominant lithology of the dolomitic bedrock, with the predominance of Ca and Mg (53.5–93.7 mg/l and 32.8–55.1 mg/l for Ca and Mg respectively). Notably, Fault Cave and Sterkfontein Cave samples presented very similar concentrations, consistent with the hydrological connection between the two caves.

Major and trace elemental concentrations of rock samples, either ghost-rocks or embedding rocks (Supplementary Table S2) were further analyzed with principal component analysis (PCA; Fig. 2), with the first two PCA dimensions accounting for 85.3% of the total variance. These dimensions clearly differentiate dolomite, chert and ghost-rock samples, consistent with their dominant mineralogy.

PCA dimension 2 explained 12.8% of the variance, primarily controlled by Ca and Mg, with contributions from Sr, which clearly separated dolomite samples from ghost-rocks and cherts. Dolomite samples exhibited particularly high Mg and Ca concentrations (11.2 wt% and 19.6 $\pm$ 0.7 wt%, respectively) compared to cherts [(1.3 $\pm$ 1.6) $\times 10^4$ μ g/g and (2.2 $\pm$ 2.8) $\times 10^4$ μ g/g, respectively] and ghost-rocks [(7.9 $\pm$ 2.7) $\times 10^3$ μ g/g and (8.9 $\pm$ 4.7) $\times 10^3$ μ g/g, respectively]. The lower Ca and Mg concentrations in ghost-rocks and chert confirmed that secondary calcite or dolomite were not dominant in these samples, in contrast to secondary metal oxides (Pisapia et al. 2021). This was supported by data dispersion on dimension 1 (which explains 72.5% of the variance) that reflected variations in trace metals, particularly iron and manganese, clearly distinguishing ghost-rocks from unaltered lithologies. Ghost-rocks were enriched in Fe and Mn [(2.0 $\pm$ 0.6) $\times 10^4$ μ g/g and (2.6 $\pm$ 0.9) $\times 10^4$ μ g/g, respectively] compared to dolomites (2030 $\pm$ 21 μ g/g and 2460 $\pm$ 99 μ g/g, respectively) and cherts (782 $\pm$ 942 μ g/g and 406 $\pm$ 484 μ g/g, respectively). This en-

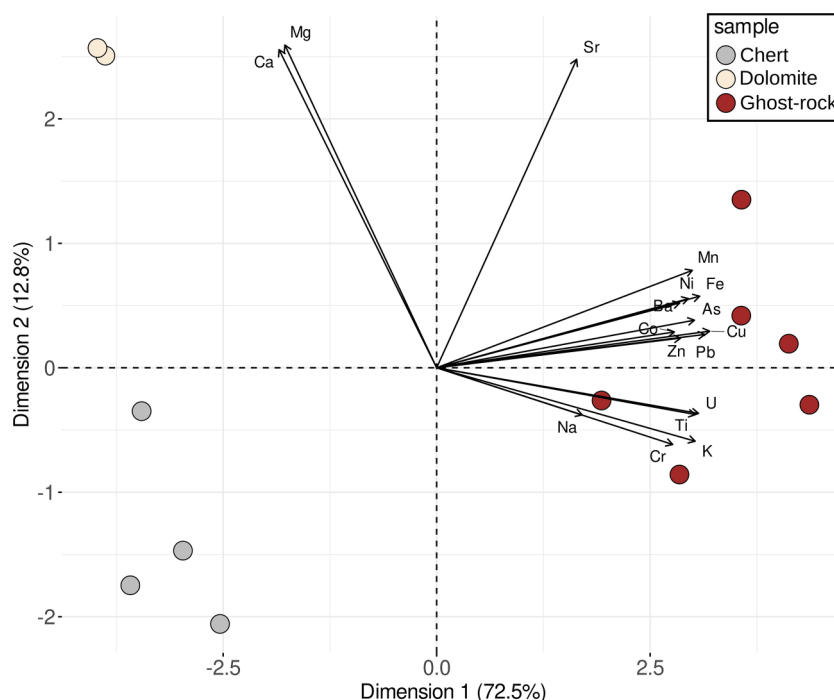


Figure 2 Principal component analysis (PCA) of embedded chert, dolomite, and ghost-rock samples based on major and trace element concentrations. The samples of embedded chert and dolomite are in grey and beige respectively, and the ghost-rock samples are in brown.

richment was consistent with the presence of Fe and Mn oxides in ghost-rocks (Pisapia et al. 2021). Ghost-rock samples were also enriched in heavy metals such as As, Pb, U, Ni, Zn, Ti, and Cu compared to dolomite and cherts (Supplementary Table S2). Overall, these results indicated that ghost-rocks were enriched in redox sensitive elements, such as iron and manganese, which may be bioavailable for chemolithotrophs.

Even though organic matter has previously been observed in ghost-rocks (Pisapia et al. 2021), and some DNA extractions yielded sufficient material for sequencing, the TOC content in all dolomite, chert and ghost-rocks samples were below the detection limit of 0.15 wt%, which confirmed that karstic environments, including ghost-rocks, can be considered as highly oligotrophic environments.

Taxonomic and functional description of the microbial communities at the contig level

In groundwater and ghost-rock samples, the most abundant phyla were Pseudomonadota ($38.1 \pm 12.5\%$ and $34.5 \pm 2.0\%$, respectively), Acidobacteriota ($14.2 \pm 7.6\%$ and $14.7 \pm 0.9\%$, respectively), Actinobacteriota ($10.0 \pm 5.2\%$ and $12.4 \pm 4.2\%$, respectively), Gemmatimonadota ($3.5 \pm 1.9\%$ and $8.3 \pm 1.3\%$, respectively), Thermoproteota ($6.0 \pm 5.0\%$ and $2.0 \pm 0.9\%$, respectively), Chloroflexota ($3.5 \pm 2.0\%$ and $5.3 \pm 0.2\%$, respectively), Methyloirabitolata ($2.9 \pm 1.5\%$ and $5.0 \pm 2.2\%$, respectively), Planctomycetota ($4.4 \pm 2.5\%$ and $3.5 \pm 0.6\%$, respectively), and Nitrospirota ($4.0 \pm 1.3\%$ and $3.2 \pm 1.7\%$, respectively) (Fig. 3A). The orders that were relatively more abundant in groundwaters compared to ghost-rocks were the Burkholderiales ($36 \pm 30\%$), the Nitrososphaerales ($5.4 \pm 5.3\%$), the Nitrospirales ($3.6 \pm 2.2\%$)

(Fig. 3B). The orders that were relatively more abundant in ghost-rocks compared to the groundwaters were the Gemmatimonadales ($7.8 \pm 1.8\%$), the Vicinamibacterales ($6.5 \pm 2.0\%$), the Steroidobacterales ($5.7 \pm 1.0\%$), the Acidimicrobiales ($4.8 \pm 2.6\%$), the Rokubacterales ($4.13 \pm 1.6\%$), the Rhizobiales ($3.8 \pm 0.9\%$), and the Pyrinomonadales ($3.2 \pm 0.2\%$) (Fig. 3B).

In order to highlight the functional differences between ghost-rock and groundwater samples, a LefSe analysis at the contig level was used to identify differentially abundant KOs and to pinpoint those that best discriminate groundwater from ghost-rock samples. In total, 403 annotations had a LDA score greater than 2, with 192 in groundwaters and 211 in ghost-rocks. They were grouped by KEGG pathways; 89 KOs were not assigned any KEGG pathway. All KEGG pathways with significant KOs are listed in Table 1. The complete list of significant KO annotations is provided in Supplementary Table S4.

Most significantly enriched KOs fell into broad functional categories essential for general cellular processes (notably transporters with 57 KOs and transcription factors with 32 KOs), as well as other generic functional groups (shown in black in Table 1). Such categories were expected to include KOs enriched in both groundwater and ghost-rocks. In contrast, several more specific functional categories contained KOs enriched exclusively in groundwater or in ghost-rocks.

Among functions enriched only in ghost-rocks, 17 KOs were involved in terpenoid and polyketides biosynthesis and 4 KOs related to starch and sucrose metabolism. Terpenoid and polyketide KOs were mostly involved in type II polyketide biosynthesis (K05551, K05552, K05554, K12420, K13327, K14631, K15926, K15928) and in carotenoid biosynthesis (K06443, K10211, K10212, K15746). The starch/sucrose-related KOs include key CAZymes involved in α -glucan catabolism (K01176: α -amylase [EC:3.2.1.1], K01187: α -glucosidase [EC:3.2.1.20], K01194: α , α -

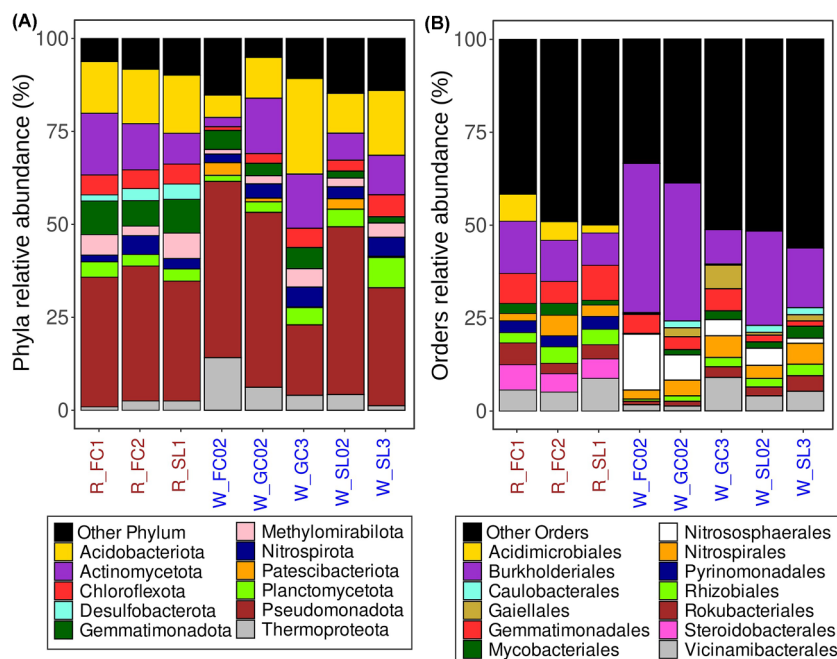


Figure 3 Microbial community at the contig level. (A) Phyla abundance based on the percentage of the total contig coverage. (B) Orders abundance based on the percentage of the total contig coverage.

Table 1 LEfSe analysis of KEGG annotations from all the contigs.

Kegg pathway	Ghost-rocks	Groundwater
00 362 Benzoate degradation	0	14
00 440 Phosphonate and phosphinate metabolism	0	5
00 907 Pinene, camphor and geraniol degradation	0	4
02 035 Bacterial motility proteins	0	7
03 036 Chromosome and associated proteins	0	4
03 400 DNA repair and recombination proteins	0	4
00 500 Starch and sucrose metabolism	4	0
09 109 Metabolism of terpenoids and polyketides	17	0
00 260 Glycine, serine and threonine metabolism	4	3
00 330 Arginine and proline metabolism	5	2
01 002 Peptidases and inhibitors	7	4
02 000 Transporters	31	26
02 020 Two-component system	4	11
02 044 Secretion system	2	8
03 000 Transcription factors	12	20
03 009 Ribosome biogenesis	2	4
09 108 Metabolism of cofactors and vitamins	7	2
Other categories (<4 KOs)	60	41

A total of 403 significant KEGG annotations was obtained and grouped according to KEGG pathway definitions. Categories containing less than four significant KEGG annotations in both groundwater and ghost-rocks are labeled as “Other categories” and are not discussed further. Categories enriched in groundwater and ghost-rocks are shown at the top and in the middle, respectively. Categories with KOs enriched in both groundwater and ghost-rock are at the bottom.

trehalase [EC:3.2.1.28], K02438: glycogen debranching enzyme [EC:3.2.1.196]). Two enriched transporter KOs (K10232: α -glucoside transport system substrate-binding protein, K21572: starch-binding outer membrane protein, SusD/RagB family) further suggest enhanced α -glucan uptake in ghost-rock communities.

Genes involved in aromatic compound degradation were strongly enriched in groundwaters. They included 14 KOs involved

in benzoate degradation, such as K00446 encoding catechol 2,3-dioxygenase [EC:1.13.11.2], and K16242-K16249, which correspond to the phenol/toluene 2-monooxygenase (NADH) operon [EC:1.14.13.244;1.14.13.243]. Two KOs related to aminobenzoate degradation were also significantly enriched (K03862 and K15064). They correspond to the vanillate monooxygenase [EC:1.14.13.82] and the syringate O-demethylase [EC:2.1.1.1-], respectively. Additional groundwater-specific KOs belonged

to monoterpene (pinene, camphor, and geraniol) degradation (K11731, K13774, K13777, K13778), phosphonate degradation (phnGHJ), DNA repair and chromosomal proteins, as well as motility/chemotaxis (K03415, K03776), and multiple flagellar proteins and regulators (K02423, K06603, K02398, K02402, K02403).

Overall, the contig-level analysis highlighted a clear enrichment of α -glucan degradation functions in ghost-rock and aromatic compound degradation in groundwater samples.

Metabolic specificities of groundwater and ghost-rock communities at the MAG level

Groundwater samples

To confirm and expand the results obtained at the contig level, the 273 reconstructed MAGs were analyzed. In groundwater samples, the community was dominated by members of the Gammaproteobacteria, Nitrospirata, and Thermoproteota phyla, together accounting for 54% of total MAG coverage, which was consistent with the results at the contig level. This dominance, however, was strongly influenced by a single gammaproteobacterial MAG (SAR200) assigned to *Limnobacter* sp. This MAG harbored pathways involved in benzene degradation (with the phenol/toluene 2-monooxygenase; K16242-K16246, K16249) and the catechol meta-cleavage pathway (with the catechol 2,3 dioxygenase; K00446), as well as the complete *soxABCDXYZ* complex for thiosulfate oxidation (K17222-K17227, K22622). Because no carbon fixation pathway was detected, SAR200 was considered as a chemolithoheterotroph.

Overall, chemolithoheterotrophic and chemolithoautotrophic MAGs represented almost half of total MAG coverage in groundwater samples (47.6%). Ten of the 17 MAGs belonging to the Nitrososphaerales order in the Thermoproteota phylum possessed the genes needed for carbon fixation through the 3-hydroxypropionate/4-hydroxybutyrate cycle. The most abundant Thermoproteota MAG (SAW134) contained homologues for all the enzymes demonstrated or proposed for *Nitrosopumilus maritimus* in the pathway with BLASTp identities between 77% and 91%. SAW134 also possessed genes for an ammonia transporter (K03320), the *ureABCDEFG* operon (K01428-K01430, K03187-K03190), and the ammonia monooxygenase complex *amoABC*, (K10944-K10946). Twelve MAGs within the Nitrospirata phylum possessed at least one subunit of the ATP-citrate lyase (K15230-K15231), a key enzyme for the reverse tricarboxylic acid (TCA) carbon fixation pathway. Among them, four MAGs also had genes for the nitrite oxidoreductase complex (K00370-K00371) needed for nitrite oxidation.

Other abundant autotrophic MAGs in groundwater possessed the genes encoding ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO; K01601-K01602) and the Calvin-Benson-Bassham (CBB) cycle for carbon fixation, and belonged to diverse Gammaproteobacteria genera, including *Methylibium* (SAW204, SAW232, SAW19), *Panacagrimonas* (SAW210, SAW216), *Aquabacterium* (SAW306, SAW104), and *Polaromonas* (SAW79, SAR163). Three of these had the genes encoding the complete *soxABCDXYZ* complex for thiosulfate oxidation (SAW79, SAR163, and SAW210), while the two *Aquabacterium* MAGs encoded a sulfide: quinone oxidoreductase (*sqr*; K17218) for sulfide oxidation. In addition, three MAGs (SAW104, SAW216 and SAW210) may be

metal oxidizers as they also possessed genes encoding a PCC (see next Section).

Among chemoorganoheterotrophic taxa, the SAR211 MAG, assigned to *Acidovorax delafieldii*, was the most abundant in groundwater samples. It possessed a complete glycolysis pathway and TCA cycle, as well as genes for dissimilatory nitrate reduction to ammonia (*nirBD*; K00362-K00363; *narGHI*; K00370-K00371, K00374) and pathways for benzoate, benzene and salicylate degradation.

As the analyses at the contig level showed the prevalence of pathways involved in aromatics and carboxylate degradation in groundwaters, their distribution was further verified at the MAG level. Only 15 MAGs possessed at least one complete KEGG module for aromatic compound degradation (SAR163, SAR200, SAR211, SAW152, SAW160, SAW19, SAW204, SAW210, SAW216, SAW232, SAW267, SAW277, SAW309, SAW315, SAW77). They represented 35.7% of total MAG coverage in groundwater but only 0.4% in ghost-rock. At the read level, pathways involved in aromatics degradation were also significantly relatively more abundant in groundwater samples (mean = 27.4 cpm) than in ghost-rock samples (mean = 4.0 cpm; $p = 0.03571$), especially those targeting catechol and protocatechuate degradation (Supplementary Fig. S1).

Carboxylates offers another source of organic matter for groundwater heterotrophs, as highlighted by read level-analyses, where pathways involved in carboxylate degradation were significantly relatively more abundant in groundwater samples (mean = 81.8 cpm) than in ghost-rock samples (mean = 5.7 cpm; $p = 0.03571$; Wilcoxon exact test) with D-galactarate and D-glucarate degradation being particularly enriched (Supplementary Fig. S3). Accordingly, 11 MAGs had genes for D-glucarate and/or D-galactarate dehydratase (Supplementary Fig. S2), including 7 MAGs belonging to the Burkholderiales order (*Polaromonas* SAW152, SAR163, SAW79; *Acidovorax* SAR211; *Rubrivivax* SAW179; *Methylibium* SAW204; and SAW317, unassigned Burkholderiales). They represented 8.3% and 1.9% of total MAGs coverage in groundwater and ghost-rock samples, respectively, confirming the enrichment of these pathways in groundwaters communities. The genomic context of these genes was very similar to the recently described operon needed for hexuronate catabolism in Pseudomonadota (Bouvier et al. 2018), including the presence of genes encoding the uronate dehydrogenase Udh and the uronate lactonase Uxul that convert galacturonate and glucuronate into galactarate and glucarate, respectively (Supplementary Fig. S2).

Ghost-rock samples

The predominant MAGs in ghost-rock samples belonged to the Actinobacteriota, Gemmatimonadota, Acidobacteriota, and Alphaproteobacteria, accounting for 51% of total MAG coverage (Fig. 4), which was consistent with the results at the contig level.

Chemolithoautotrophic MAGs account for 16.2% of total MAG coverage. Among them, six actinobacterial MAGs (SAR125, Acidimicrobiales; SAR66, Pseudonocardiaceae; SAR301, *Actinokineospora* sp.; SAR248, UBA4738; SAW70, Acidimicrobiales; SAW28, *Pseudonocardia*) encoded the RuBisCO and a complete CBB cycle. Moreover, all six possessed genes for high-affinity [NiFe]-hydrogenase 1 h, and four possess the *coxMLS* operon involved in CO oxidation.

Two additional autotrophic gammaproteobacterial MAGs (SAR13; Steroidobacteraceae and SAR127; genus CAIKBL01),

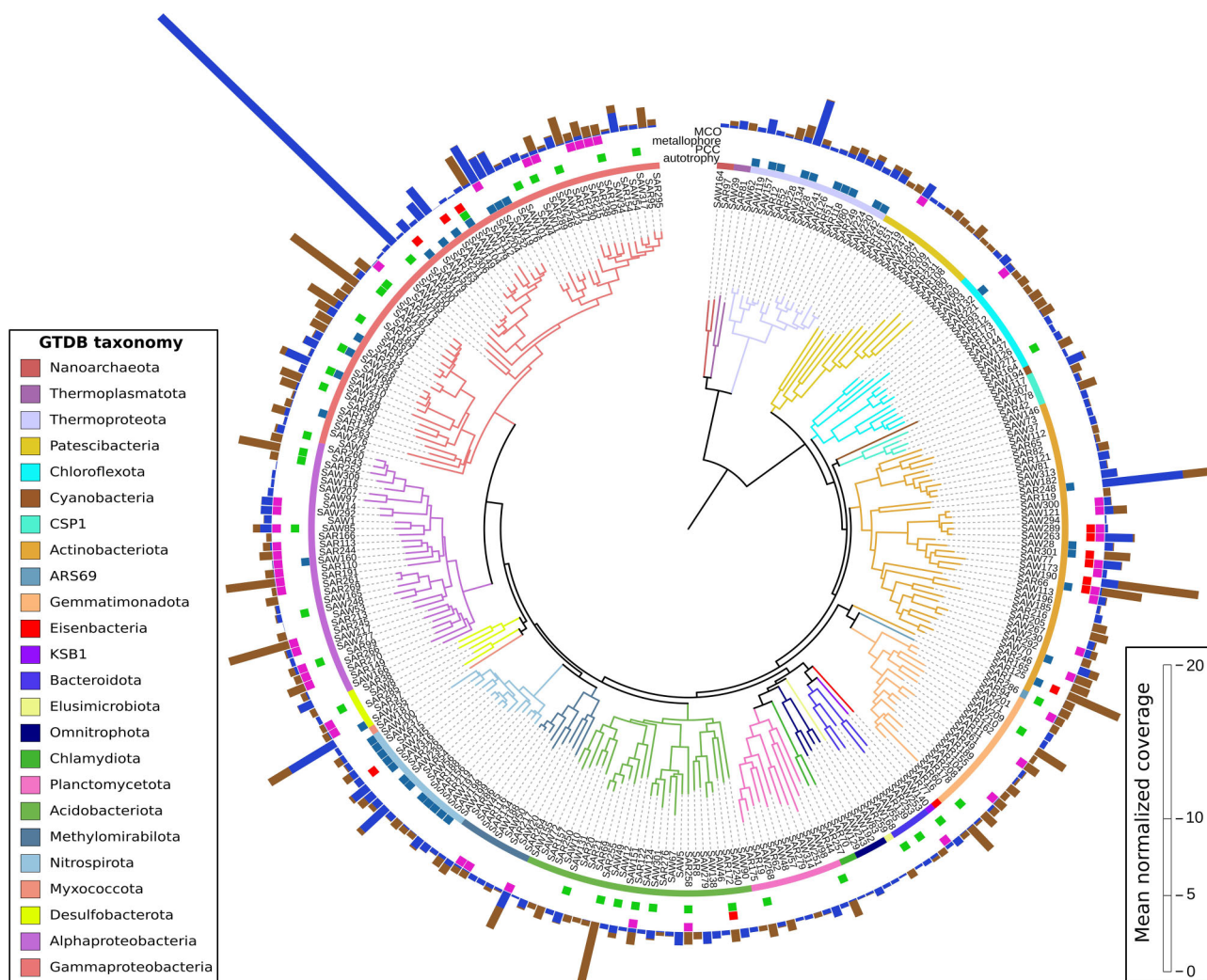


Figure 4 Phylogeny of the 273 MAGs obtained on the groundwater and ghost-rock samples. The branches are colored according to the taxonomy assigned by GTDB-Tk. The outer ring shows a histogram of mean normalized coverage of the MAG in groundwater samples (blue) and ghost-rock samples (brown). The four inner rings indicate MAGs harboring at least one of the following feature: multicopper oxidases (MCOs, pink square), metallophore biosynthesis pathway (red square), porin-cytochrome complexes (PCCs, green square) or carbon fixation pathway (blue square).

also possessed genes encoding RubisCO and the CBB cycle. SAR13 possessed, on one contig, a complete set of genes for the initial step of methanol oxidation (methanol dehydrogenase *xoxF*; K23995, glutathione-dependent formaldehyde-activating enzyme K03396; S-(hydroxymethyl)glutathione dehydrogenase K00121, and formate dehydrogenase K15022), and on another contig, genes encoding S-formylglutathione hydrolase (K01070) necessary to complete the glutathione-dependent pathway, enabling full oxidation of methanol to CO₂. SAR127 possessed several *coxMLS* genes necessary for CO oxidation, and *soxACXYZ* genes involved in thiosulfate oxidation (*soxB* is missing).

As the marker gene analysis at the contig level indicated a greater relative abundance of α -glucans (glycogen and starch) degradation pathway in ghost-rock samples, the presence of genes encoding CAZymes involved in this process within the MAGs was investigated. Based on the KofamScan annotations, the MAGs with a complete glycogen degradation or synthesis pathway represented 10.8 $\pm$ 2.4% and 31.2 $\pm$ 1.7% of the ghost-rock total MAGs

coverage respectively, whereas in groundwaters they represented 2.0 $\pm$ 1.6% and 15.8 $\pm$ 9.6% of the total MAGs coverage respectively (Fig. 5). Based on the dbCAN3 annotations, the glycogen synthetases were also relatively more abundant in ghost-rock samples (64.7 $\pm$ 7.4% compared to 39.0 $\pm$ 19.4% in groundwaters, Fig. 5). Overall, the glycogen metabolism represented a higher proportion of the MAGs relative abundance in ghost-rock samples than in groundwater samples. Concerning the read level, no significant difference was detected between groundwater and ghost-rock samples in the abundance of polysaccharide degradation (mean = 24.8 and 37.0 respectively; $p = 0.1429$), mostly due to the lack of differential abundance of chitin degradation pathways ($p = 0.37$ and $p = 1$, for chitin deacetylation and chitin degradation II, respectively). When restricting the analysis to α -glucan degradation pathways (i.e. pathways involved in glycogen and starch degradation), a significant difference emerged, with a higher relative abundance of these pathways in ghost-rock samples (mean = 28.4 cpm) compared to groundwater samples (mean = 14.4 cpm; $p = 0.03571$) (Supplementary Fig. S1C).

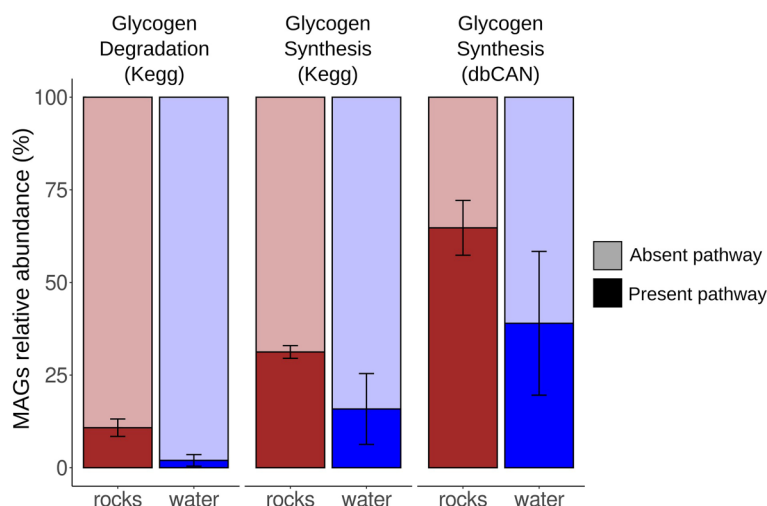


Figure 5 Relative abundance of the glycogen metabolism in the reconstructed MAGs. From left to right, abundance of the glycogen degraders and synthetizers based on KofamScan search, and glycogen synthetizers based on the dbCAN search. The MAG abundance is calculated as the percentage of total MAGs coverage. Ghost-rock samples are in brown and groundwater samples are in blue.

As the LEfSe analysis also showed that the production of certain secondary metabolites was characteristic of ghost-rock communities, the metabolic potential for secondary metabolites biosynthesis was further inspected across the reconstructed 273 MAGs. In total, 715 BGCs have been found, with up to 35 BGCs found in a single MAG. Most of these clusters showed no or low similarity to any entry in the minimum information about a biosynthetic gene cluster (MIBiG) database (Supplementary Fig. S3), suggesting the presence of novel or uncharacterized secondary metabolites. The majority of BGCs were involved in terpene (222 clusters), polyketide (96 clusters), and non-ribosomal peptide (41 clusters) biosynthetic pathways.

No significant difference was observed in the mean number of BGC per MAG between groundwater (mean = 1.64) and ghost-rock (mean = 2.07) samples ($p = 0.95$). Nevertheless, a slight but significant positive correlation existed between the mean coverage of MAGs in ghost-rock samples and their number of BGCs (Spearman's $r = 0.15$, $p = 0.012$). It was mainly due to the high abundance of three Actinobacteriota MAGs in ghost-rock samples: SAR301 (assigned to the genus *Actinokineospora*), SAR66 (*Pseudonocardia*ceae) and SAW190 (*Pseudonocardia*ceae, genus GCA-003244245), which contained 35, 30 and 19 BGCs, respectively (Supplementary Fig. S4).

127 MAGs possessed a putative O_2 reducing terminal oxidase (based on the presence of a complete Kegg module M00153-M00156 or M00416-M00417), which was expected as both groundwaters and sampled ghost-rocks were in contact with the atmosphere and oxygenated.

Analysis of the metal-interacting pathways

As ghost-rock samples are enriched in redox-sensitive elements, notably Fe and Mn (Fig. 2), reflecting active metal cycling and potential (bio)alteration/(bio)mineralization processes, a specific attention has been paid to genes and pathways involved in metal interactions, including metallophores, MCOs and PCC synthesis (Fig. 4).

Metallophore-producing BGCs were found in 12 MAGs, and 45 MAGs encoded at least one MCO, primarily among Alphaproteobacteria and Actinobacteriota, in both groundwater and ghost-rock samples. Concerning EET systems, 25 MAGs possessed at least one Cyc2 encoding gene, but all these genes were assigned to the cluster 3, implying that their role in iron oxidation should be considered uncertain (Keffer et al. 2021). Thirty six MAGs harbored at least one PCC, of which 29 have at least 10 multiheme cytochromes (Supplementary Table S2) and 27 also possessed *ccm* genes required for cytochrome c-type biogenesis. Considering the fragmented nature of the MAGs, we estimate that at least 75% of the MAGs possessing a PCC also encoded a functional cytochrome c biogenesis pathway. PCCs were not evenly distributed among phyla: they were all in groups of bacteria with an outer membrane, except for SAW271, assigned to the Chloroflexota phylum, and SAR286 and SAW140, which belonged to uncultivated candidate phyla (ARS69 and KSB1, respectively). However, the presence of PCCs in KSB1 MAGs had previously been reported (Suarez et al. 2023), along with Planctomycetota MAGs that also encoded PCCs. Among the 36 PCC-containing MAGs, 21 belonged to the Gammaproteobacteria class or Acidobacteriota phylum, and two to the Planctomycetota phylum. No clear difference was observed in PCC abundance between groundwater and ghost-rock samples. PCC containing-MAGs represented 28% of total MAG coverage in groundwater, mainly due to SAR200, which dominated this community, and only 17% in ghost-rocks. Only three PCC-containing MAGs also possessed a carbon fixation pathway (Fig. 4), suggesting that most MAGs corresponded to heterotrophic bacteria that may use their PCC systems to export electrons extracellularly.

The clustering of the multiheme cytochromes revealed several groups of PCCs sharing conserved genomic and structural features (Fig. 6). The first group comprising 11 PCCs was characterized by the presence of a large decaheme cytochrome (between 4 and 5 kbp) located adjacent to an outer membrane porin with around 20 transmembrane β -strands. In seven of these eleven PCCs, a ferredoxin-NADP⁺ reductase gene was also found in close genomic proximity, suggesting potential functional coupling. The second group, which included 8 PCCs had smaller decaheme cy-

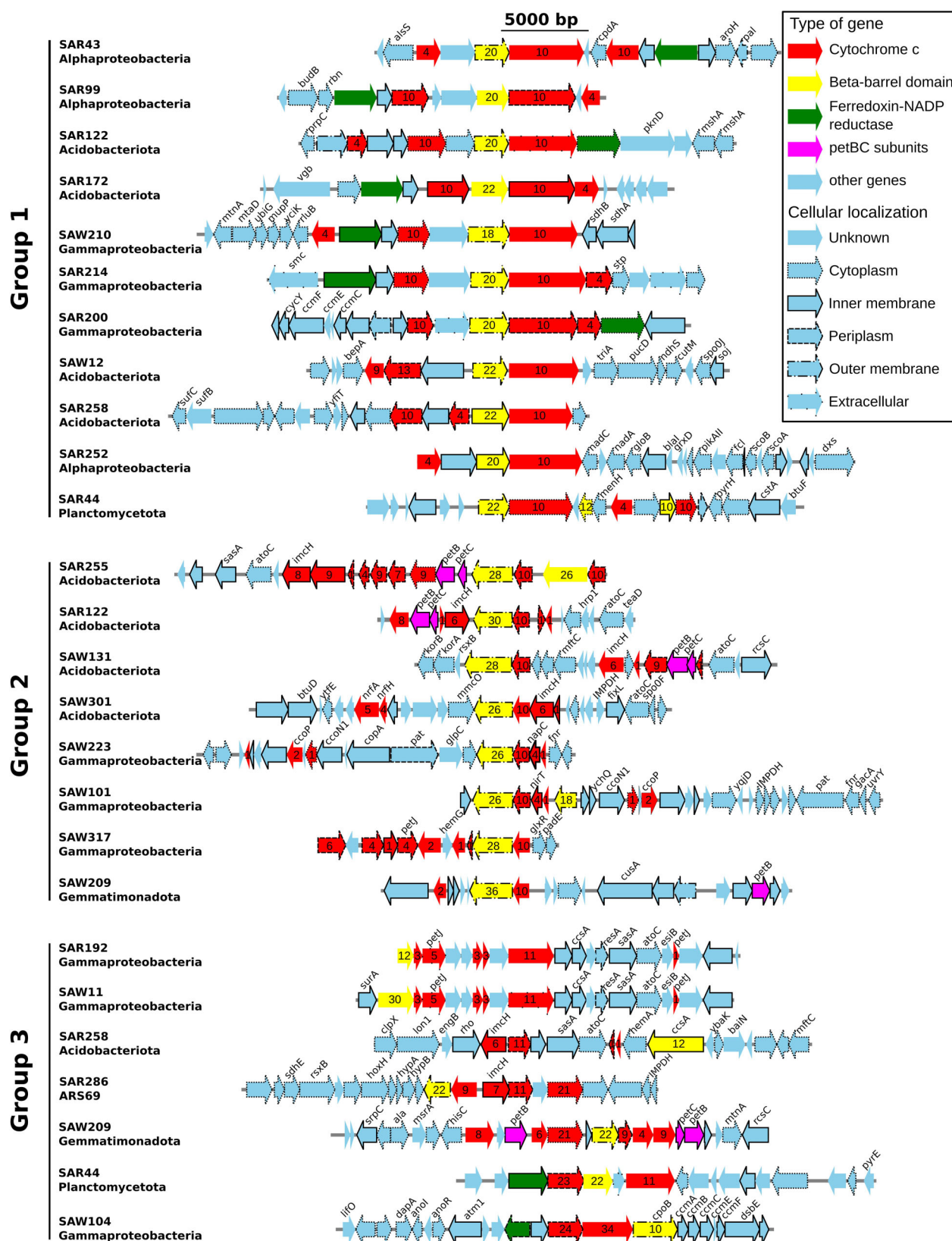


Figure 6 Genomic organization of extracellular electron transfer (EET) complexes detected in the MAGs. The phylum or class of the MAGs possessing the porin-cytochrome complexes (PCC) is indicated on the left. Arrows length is proportional to the length of the genes. Color indicates gene types: c-type cytochromes (red), porins (yellow), ferredoxin-NADP reductases (green), cytochrome bc₁ subunits (pink) or other genes (light blue). For c-type cytochromes and porins, the number inside the arrow corresponds to the number of heme-binding motifs and transmembrane β -barrel domains, respectively. Arrow borders denote predicted protein cellular localization (PSORTb), and functional annotations from Prokka are indicated above each arrow.

tochromes located next to an outer membrane porin containing between 25 and 36 transmembrane β -strands. Six of them also encoded inner-membrane cytochromes similar to ImcH or NapC/NirT that were necessary for electron transfer in iron oxidizers or reducers as they are known to interact with the quinone pool (Jain et al. 2022, Pimenta et al. 2023). The third group comprised 7 PCCs centered around undecaheme cytochromes or multiheme cytochromes with more than 20 CX₂₋₄CH heme-binding motifs. Undecaheme cytochromes have been described in iron-oxidizing Tenderiales (Eddie et al. 2022) and in the manganese-reducing bacterium *Candidatus Dechloromonas occulta* (Szeinbaum et al. 2020). Finally, several PPCs contained the *petB* and *petC* genes, encoding the cytochrome b and Fe-S subunits of the cytochrome bc₁ complex, respectively (SAR255, SAR122, SAW131, SAW209 in group 2, and SAW209 in group 3). The co-occurrence of these redox-active components was also described in “PCC4” systems reported by He et al. (2017), suggesting that similar electron transfer architectures may be conserved across phylogenetically diverse bacteria.

Discussion

Structure of the groundwater and ghost-rock microbial communities

Results based on both contig and MAG level revealed distinct communities inhabiting ghost-rocks (dominated by Actinobacteriota, Gemmatimonadota, Acidobacteriota, and Alphaproteobacteria) and groundwaters (enriched in Gammaproteobacteria, Nitrospirota, and Thermoproteota). This result was consistent with other karstic cave metabarcoding studies (Turrini et al. 2024).

While this is the first characterization of the metabolic potential of ghost-rock ecosystem, karstic groundwater communities have been more extensively studied. The groundwaters assemblage described here may represent the “core microbiome” of undisturbed groundwater, as described in Geesink et al. (2022). This community is described in this article as sustained by autotrophic and mixotrophic taxa that primarily use surface-derived inorganic nitrogen and rock-derived inorganic sulfur compounds, as well as heterotrophs capable of degrading refractory surface-derived plant biomass.

Concerning carbon sources, comparative analysis at the read, contig and MAG levels showed a preponderance in groundwater samples of aromatic compounds' catabolic pathways, involving protocatechuate, catechol, syringic acid, vanillic acid that may derive from lignin depolymerization, and funneling pathways (Brink et al. 2019). At both the read and MAG level (but not at the contig level), D-galactarate and D-glucarate degradation pathways, both linked to the breakdown of pectin and hemicellulose (Bouvier et al. 2018), were also more abundant in groundwaters. Overall, these results indicate that refractory, surface-derived plant compounds were an important carbon source for groundwater community. In contrast, at the read, contig and MAG levels, ghost-rock communities displayed a preponderance of pathways involved in metabolism of α -glucans, notably glycogen. Ghost-rocks sampled just above the water table experience intermittent flooding and desiccation along irregular supply of both allochthonous and autochthonous organic compounds, which would induce both oligotrophic and hydric stress. Glycogen and more generally α -

glucan function as major and durable carbon and energy storage molecules that enable prokaryotes to overcome environmental stresses such as desiccation and nutrient limitation (Colpaert et al. 2020, Wang et al. 2020). The ability to synthesize glycogen-like molecules likely provide a competitive advantage supporting survival during nutrient scarcity for the micro-organisms inhabiting the ghost-rocks. Glycogen may also serve as a major carbon source for heterotrophic necromass degraders in ghost-rock community. A process of α -glucan necromass turnover has previously been described in marine ecosystems, where it participates in intra-population energy conservation mechanisms (Beidler et al. 2024).

Concerning energy sources for chemolithotrophs, electron donors mobilized by the groundwater community were found to be thiosulfate (*Limnobacter*, *Polaromonas*, *Panacagrimonas*), nitrate/nitrite (Nitrospirota, Thermoproteota), or metals (*Aquabacterium*, *Panacagrimonas*). All cultivated *Limnobacter* species, isolated from marine, freshwater or soil environments, are thiosulfate oxidizers (Kim et al. 2024), and this study is the first report of groundwater community dominated by this genus. Conversely, thiosulfate-oxidizing *Polaromonas* (Taubert et al. 2022) and ammonia-oxidizing bacteria and archaea coexisting with nitrite-oxidizing bacteria (Krüger et al. 2023) were previously described in karstic groundwaters. The putative iron-oxidizing activity of *Aquabacterium* and *Panacagrimonas* is less documented but is consistent with the isolation of an autotrophic iron-oxidizing *Aquabacterium* strain from freshwater sediments (Straub et al. 2004). In contrast, the ghost-rock primary producers were mainly found to be Actinobacteriota putatively relying on H₂ and CO oxidation, and Gammaproteobacteria capable of methanol oxidation. Autotrophic H₂- and CO-oxidizing Actinobacteriota have already been described in diverse environments, including soils and marine sediments (Groster and Alvarez-Cohen 2013). This is also the case for the gammaproteobacterial Pseudomonadota that are known autotrophic methylotrophs combining methanol oxidation to CO₂ using the glutathione-linked formaldehyde oxidation pathway, and the CBB cycle, as described for *Paracoccus denitrificans* (Chistoserdova 2011). Volatile H₂, CO and methanol can be found as trace gases in the atmosphere (Stacheter et al. 2013, Bay et al. 2021) and their oxidation fuels a primary production process known as “atmospheric chemosynthesis”, which prevails in oligotrophic, low-light terrestrial environments such as Antarctic soils and cold deserts (Ji et al. 2017, Ray et al. 2022), and may in fact be more widespread globally (Bay et al. 2021). Ghost-rocks, sampled up to 20 cm into highly porous cave walls, are in direct contact with cave air and situated in the unsaturated karst zone, which offers surface-like geochemical conditions (Covington et al. 2023). Thus, it is hypothesized that primary production in ghost-rocks may be similar to that in low-light oligotrophic soils, relying on atmospheric trace gas oxidation rather than photosynthetic inputs (Ji et al. 2017, Ray et al. 2022).

Even though these results provided the first insights into ghost-rock microbial community composition and metabolic potential, several questions remain. The ghost-rock community is not merely a filtered subset of the groundwater microbiome. It is a distinct community that should be adapted to the specific conditions of the ghost-rock environment. However, its biogeographic origin remains uncertain. Given the strong connectivity between caves and surface systems, endemism, as suggested for some deep subsurface lineages (Colman et al. 2017), seems unlikely. The community

likely originates from a combination of sources, including groundwater, soil, and atmospheric inputs.

The dynamics of the ghost-rock community are also poorly understood. Ghost-rocks may act as water reservoirs (Champollion et al. 2018) with microbial community composition shifting in response to changing hydrologic conditions, notably during recharge periods. Hydrological boundaries are known to structure the community composition (Smith et al. 2018) and deeper, permanently flooded ghost-rock likely harbor different microbial assemblages less reliant on atmospheric chemosynthesis. Overall, addressing the spatio-temporal variations of ghost-rock microbial communities will need expanded sampling and long-term monitoring.

At geological timescales, microbial activity may have influenced both bedrock weathering, and ghost-rock formation. Ghost-rocks result from the dissolution of soluble bedrock components, and accumulation of insoluble residues, including metal oxides. The presence of metal-interacting pathways in these microbial communities suggest potential interactions between microorganisms and rock weathering processes.

Interaction with metals and weathering through karst history

The detection of metal-interacting pathways was partly limited by conservative antiSMASH parameters, and the fragmented nature of the MAGs, which reduce biosynthetic cluster recovery. Similarly, the criteria used to detect potential for EET, requiring a PCC centered on a multiheme cytochrome with at least 10 CX₂-₄CH motifs, may underestimate other EET systems. In particular, nanowires (such as OmcS and OmcE in *Geobacter*; Wang et al. 2019) and flavin-based electron transfer in some Gram positive bacteria (Light et al. 2018) were not investigated. Moreover, the PCC found could be related to non-metal metabolism, such as the DMSO reduction that uses similar system (Gralnick et al. 2006). Thus, cultivation-based approaches to obtain metal reducing/oxidizing isolates could be highly informative in order to complement and validate the putative metal interacting functions that were predicted through bioinformatics. Nevertheless, a significant number of PCC-containing MAGs was recovered, comparable to those from groundwater and sediment enrichments of electroactive micro-organisms (Arbour et al. 2020). Similar cytochrome-based EET systems have been reported in iron-rich caves (Parker et al. 2022), granitic diorite enrichment (Napieralski and Roden 2020) or a marine subseafloor tunnel (Suarez et al. 2023), although methodological differences complicate direct comparison. Overall, these findings suggest that metal-interacting pathways may influence the rates of rock weathering (Wild et al. 2022). Yet, it remains difficult to predict which micro-organisms initially colonize pristine rocks, whether they possess metal-interacting pathways, and how and to which extent these interactions drive rock weathering processes. Incubation experiments in bioreactors using groundwater from a German carbonate karst have shown that *Aquabacterium* spp. may have played a key role in biofilm formation on carbonate rocks, particularly iron-coated rocks (Sharma et al. 2024). In our groundwater dataset, three *Aquabacterium* MAGs were obtained, with the highest quality MAG (SAW104; 92.13% completeness; 3.08% contamination; contig size N50 = 97 849 bp) showing potential for metal oxidation. Such

taxa may thus contribute to initial colonization and bioweathering, promoting the formation of ghost-rock, and subsequently being replaced by the distinct heterotrophic ghost-rocks community described here. Indeed, diverse heterotrophic MAGs were observed within ghost-rocks, mainly Acidobacteriota and Pseudomonadota, that may act as metal reducers benefiting from the presence of iron and manganese oxides that constitute the ghost-rocks (Fig. 2; Pisapia et al. 2021). This reduction could solubilize adsorbed elements as substantial trace elements adsorption to iron and manganese oxides is well documented (Tebo et al. 2004, Shi et al. 2021). At the scale of karstic systems, ferromanganese deposits can immobilize trace elements (Friedrich and Catalano 2012), mitigating heavy metals contamination (Newsome et al. 2021). Conversely, dissimilatory reduction of iron oxides can release toxic metal such as adsorbed arsenic from aquifer sediments (Islam et al. 2004). Ghost-rock samples are enriched in redox-sensitive and potentially toxic heavy metals compared to embedded bedrock samples (Fig. 2) and may thus act as dynamic geochemical reservoirs. Microbial metal reduction could occur in anoxic micro-niches, or even under oxic conditions (Ceriotti et al. 2025), releasing adsorbed compounds into groundwater. Given the putative long water residence time of ghost-rocks (Champollion et al. 2018), the presence of metal reducers in ghost-rocks would favor partial dissolution of iron and manganese oxides, and significantly influence trace element mobility within karst aquifer. However, the precise nature of microbial impact on bedrock and ghost-rock weathering remains poorly understood. To elucidate these processes, *in situ* experiments combining geochemical monitoring and metatranscriptomics are needed to directly link microbial activity for microbial activity with mineral transformation.

Conclusion

The metagenomic analysis of the Sterkfontein karst system reveals highly distinct microbial communities between groundwaters and ghost-rocks, both taxonomically and metabolically. Groundwater communities are dominated by mixotrophic and autotrophic MAGs, primarily belonging to the Gammaproteobacteria, Nitrospirota and Thermoproteota, capable of utilizing ammonium, nitrite or thiosulfate as electron donors, and degrading refractory plant-derived organic matter. By contrast, ghost-rocks host predominantly necromass-degrading heterotrophic MAGs, some of them being putative metal reducers, from the Acidobacteriota and Pseudomonadota phyla, alongside autotrophic Actinobacteriota potentially relying on atmospheric chemosynthesis. The importance of α -glucan metabolism related to glycogen synthesis suggests carbon and energy storage strategies adapted to the oligotrophic and fluctuating hydrological conditions of ghost-rocks. Both ecosystems harbor MAGs with the potential to modify the redox state of metals, suggesting complementary roles in rock colonization and weathering. *Aquabacterium* species in groundwaters may initiate metal oxidation while putative heterotrophic metal reducers in ghost rocks may contribute to the dissolution of the metal oxides. Overall, the study highlights the ecological and geochemical significance of metal-interacting micro-organisms in influencing element cycling, mineral transformation, and the long-term evolution of karstic systems such as Sterkfontein.

Acknowledgements

Authors want to thank the South African Heritage Resource Agency, the French Ministry of Foreign Affairs, the French Embassy (Johannesburg), the University of the Witwatersrand (Johannesburg). Numerical computations were partly performed on the S-CAPAD/DANTE platform, IGP, France. We also want to thank the FIRE Doctoral school.

Author contributions

Guillaume Peugnet (Formal Analysis, Investigation, Validation, Visualization, Writing – original draft), Céline Pisapia (Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Visualization, Writing – review & editing), Bénédicte Ménez (Conceptualization, Supervision, Writing – review & editing), Maud Watkinson (Formal Analysis, Writing – review & editing), Léna Lecourt (Methodology), Nicolas Peugnet (Software), Julien Bouchez (Validation, Writing – review & editing), Laurent Bruxelles (Conceptualization, Writing – review & editing), and Emmanuelle Gérard (Conceptualization, Investigation, Supervision, Validation, Writing – review & editing)

Supplementary material

Supplementary material is available at *FEMSEC Journal* online.

Conflicts of interest

None declared.

Funding

This work was supported by the IGP programs “PPF Potamologie 2016” and “Soutien à idées nouvelles ou à risques 2019”; by the French CNRS Mission pour les Initiatives Transverses et Interdisciplinaires (MITI) Défi Origines 2020 program (n°227615) and the French National Research Agency ANR through the miKroAVEN project (ANR-24-CE01-6539-01).

Data availability

All raw sequencing reads and high quality MAGs were deposited on NCBI under the Bioproject PRJNA1298313. The datasets generated and analyzed in the present study are available from the corresponding author.

References

- Akob DM, Küsel K. Where microorganisms meet rocks in the Earth's Critical Zone. *Biogeosciences* 2011;**8**:3531–43. <https://doi.org/10.5194/bg-8-3531-2011>.
- Aramaki T, Blanc-Mathieu R, Endo H *et al.* KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 2020;**36**:2251–2. <https://doi.org/10.1093/bioinformatics/btz859>.
- Arbour TJ, Gilbert B, Banfield JF. Diverse microorganisms in sediment and groundwater are implicated in extracellular redox processes based on genomic analysis of bioanode communities. *Front Microbiol* 2020;**11**:1694. <https://doi.org/10.3389/fmicb.2020.01694>.
- Baral C, Séranne M, Camus H *et al.* Impact of alteration corridors on karst reservoir organisation and evolution of groundwater flow path: an example from the southern border of the Larzac Causse, southern France. *Bulletin de La Société Géologique de France* 2024;**195**:4. <https://doi.org/10.1051/bsgf/2023017>.
- Bar-On YM, Phillips R, Milo R. The biomass distribution on Earth. *Proc Natl Acad Sci USA* 2018;**115**:6506–11. <https://doi.org/10.1073/pnas.1711842115>.
- Barton HA, Taylor MR, Pace NR. Molecular phylogenetic analysis of a bacterial community in an oligotrophic cave environment. *Geomicrobiol J* 2004;**21**:11–20. <https://doi.org/10.1080/01490450490253428>.
- Barton HA, Taylor NM, Lubbers BR *et al.* DNA extraction from low-biomass carbonate rock: an improved method with reduced contamination and the low-biomass contaminant database. *J Microbiol Methods* 2006;**66**:21–31. <https://doi.org/10.1016/j.mimet.2005.10.005>.
- Bay SK, Dong X, Bradley JA *et al.* Trace gas oxidizers are widespread and active members of soil microbial communities. *Nat Microbiol* 2021;**6**:246–56. <https://doi.org/10.1038/s41564-020-00811-w>.
- Beghini F, McIver LJ, Blanco-Míguez A *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* 2021;**10**:e65088. <https://doi.org/10.7554/eLife.65088>.
- Beidler I, Steinke N, Schulze T *et al.* Alpha-glucans from bacterial necromass indicate an intra-population loop within the marine carbon cycle. *Nat Commun* 2024;**15**:4048. <https://doi.org/10.1038/s41467-024-48301-5>.
- Bender K, Glover K, Archey A *et al.* The impact of sample processing and media chemistry on the culturable diversity of bacteria isolated from a cave. *Int J Speleol* 2020;**49**:209–20. <https://doi.org/10.5038/1827-806X.49.3.2337>.
- Bernardini S, Bellatreccia F, Columbu A *et al.* Morpho-mineralogical and bio-geochemical description of cave manganese stromatolite-like patinas (Grotta del Cervo, Central Italy) and hints on their paleohydrological-driven genesis. *Front Earth Sci* 2021;**9**:642667. <https://doi.org/10.3389/feart.2021.642667>.
- Blin K, Shaw S, Augustijn HE *et al.* antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* 2023;**51**:W46–50. <https://doi.org/10.1093/nar/gkad344>.
- Bochet O, Bethencourt L, Dufresne A *et al.* Iron-oxidizer hotspots formed by intermittent oxic–anoxic fluid mixing in fractured rocks. *Nat Geosci* 2020;**13**:149–55. <https://doi.org/10.1038/s41561-019-0509-1>.
- Bouvier JT, Sernova NV, Ghasempour S *et al.* Novel metabolic pathways and regulons for hexuronate utilization in proteobacteria. *J Bacteriol* 2018;**201**:e00431–18. <https://doi.org/10.1128/JB.00431-18>.
- Brink DP, Ravi K, Lidén G *et al.* Mapping the diversity of microbial lignin catabolism: experiences from the eLignin database. *Appl*

- Microbiol Biotechnol* 2019;**103**:3979–4002. <https://doi.org/10.1007/s00253-019-09692-4>.
- Bruxelles L. Des fantômes et des hommes. Le rôle de la fantômi-sation dans la formation des karsts à hominins d'Afrique du Sud. *Karstologia* 2018;**69**:1–8. <https://hal.science/hal-03093010v1> (1 Nov. 2021, date last accessed)
- Carmichael S, Bräuer S. Microbial diversity and manganese cycling: a review of manganese-oxidizing microbial cave communities. In: Krüger M, van Berk W, Appel E (eds.), et al. (eds.), *Geomicrobiology and Biogeochemistry*. Berlin, Boston: De Gruyter, 2015;137–60. <https://doi.org/10.1515/9783110339888-009>.
- Caspi R, Billington R, Keseler IM et al. The MetaCyc database of metabolic pathways and enzymes—a 2019 update. *Nucleic Acids Res* 2020;**48**:D445–53. <https://doi.org/10.1093/nar/gkz862>.
- Cerriotti G, Bosco-Santos A, Borisov SM et al. Decoupling microbial iron reduction from anoxic microsite formation in oxic sediments: a microscale investigation through microfluidic models. *Front Microbiol* 2025;**16**:1504111. <https://doi.org/10.3389/fmicb.2025.1504111>.
- Champollion C, Deville S, Chéry J et al. Estimating epikarst water storage by time-lapse surface-to-depth gravity measurements. *Hydrol Earth Syst Sci* 2018;**22**:3825–39. <https://doi.org/10.5194/hess-22-3825-2018>.
- Chaumeil PA, Mussig AJ, Hugenholtz P et al. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 2020;**36**:1925–7. <https://doi.org/10.1093/bioinformatics/btz848>.
- Chistoserdova L. Modularity of methylotrophy, revisited. *Environ Microbiol* 2011;**13**:2603–22. <https://doi.org/10.1111/j.1462-2920.2011.02464.x>.
- Colman DR, Poudel S, Stamps BW et al. The deep, hot biosphere: twenty-five years of retrospection. *Proc Natl Acad Sci USA* 2017;**114**:6895–9003. <https://doi.org/10.1073/pnas.1701266114>.
- Colpaert M, Chabi M, Cenci U et al. Storage polysaccharides in prokaryotes: glycogen, granulose, and starch-like granules. In: Jendrossek D (ed.), *Bacterial Organelles and Organelle-like Inclusions*. Cham: Springer International Publishing, 2020. p.177–210. https://doi.org/10.1007/978-3-030-60173-7_8.
- Covington MD, Martin JB, Toran JL et al. Carbonates in the critical zone. *Earth's Future* 2023;**11**:e2022EF002765. <https://doi.org/10.1029/2022EF002765>.
- Dandurand G, Dubois C, Maire R et al. The Charente karst basin of the Touvre: alteration of the Jurassic series and speleogenesis by ghost-rock process. *Geol Belg* 2014; 1. Jan. <https://popups.uiliege.be/1374-8505/index.php?id=4131> (19 Apr. 2026, date last accessed).
- Dubois C, Quinif Y, Baele JM et al. The process of ghost-rock karstification and its role in the formation of cave systems. *Earth-Sci Rev* 2014;**131**:116–48. <https://doi.org/10.1016/j.earscirev.2014.01.006>.
- Durand L, Champollion C, Aubineau J et al. Geophysical imaging in carbonate critical zone highlights the significant of deep weathering. *Geomorphology* 2025;**488**:109953. <https://doi.org/10.1016/j.geomorph.2025.109953>.
- Eddie BJ, Bird LJ, Pelikan C et al. Conservation of energetic pathways for electroautotrophy in the uncultivated candidate order Tenderiales. *mSphere* 2022;**7**:e00223–22. <https://doi.org/10.128/msphere.00223-22>.
- Eddy SR Accelerated profile HMM searches. *PLoS Comput Biol* 2011;**7**:e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
- Edwards MJ, Richardson DJ, Paquette CM et al. Role of multiheme cytochromes involved in extracellular anaerobic respiration in bacteria. *Protein Sci* 2020;**29**:830–42. <https://doi.org/10.1002/pro.3787>.
- Eren AM, Esen ÖC, Quince C et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 2015;**3**:e1319. <https://doi.org/10.7717/peerj.1319>.
- Flynn TM, Sanford RA, Ryu H et al. Functional microbial diversity explains groundwater chemistry in a pristine aquifer. *BMC Microbiol* 2013;**13**:146. <https://doi.org/10.1186/1471-2180-13-146>.
- Fonollá C, Sanz E, Menéndez-Pidal I. Lateral ferruginous groundwater transfer as the origin of the iron crusts in caves: a case study. *J Cave Karst Stud* 2020;**82**:183–97. <https://doi.org/10.4311/2019ES0143>.
- Friedrich AJ, Catalano JG. Distribution and speciation of trace elements in iron and manganese oxide cave deposits. *Geochim Cosmochim Acta* 2012;**91**:240–53. <https://doi.org/10.1016/j.gca.2012.05.032>.
- Garber AI, Neilson KH, Okamoto A et al. FeGenie: a comprehensive tool for the identification of iron genes and iron gene neighborhoods in genome and metagenome assemblies. *Front Microbiol* 2020;**11**. <https://www.frontiersin.org/article/10.3389/fmicb.2020.00037> (18 Feb. 2022, date last accessed).
- Gázquez F, Calaforra JM, Forti P. Black Mn-Fe crusts as markers of abrupt palaeoenvironmental changes in El Soplao cave (Cantabria, Spain). *Int J Speleol* 2011;**40**:163–79. <https://doi.org/10.5038/1827-806X.40.2.8>.
- Geesink P, Taubert M, Jehmlich N et al. Bacterial necromass is rapidly metabolized by heterotrophic bacteria and supports multiple trophic levels of the groundwater microbiome. *Microbiol Spectr* 2022;**10**:e00437–22. <https://doi.org/10.1128/spectrum.00437-22>.
- Gilchrist CLM, Chooi Y-H. clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics* 2021;**37**:2473–5. <https://doi.org/10.1093/bioinformatics/btab007>.
- Gralnick JA, Vali H, Lies DP et al. Extracellular respiration of dimethyl sulfoxide by *Shewanella oneidensis* strain MR-1. *Proc Natl Acad Sci* 2006;**103**:4669–74. <https://doi.org/10.1073/pnas.0505959103>.
- Grosten A, Alvarez-Cohen L. RubisCO-based CO₂ fixation and C1 metabolism in the actinobacterium *Pseudonocardia dioxanivorans* CB1190. *Environ Microbiol* 2013;**15**:3040–53. <https://doi.org/10.1111/1462-2920.12144>.
- Guindon S, Dufayard J-F, Lefort V et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;**59**:307–21. <https://doi.org/10.1093/sysbio/syq010>.
- Hartshorne RS, Reardon CL, Ross D et al. Characterization of an electron conduit between bacteria and the extracellular environment. *Proc Natl Acad Sci USA* 2009;**106**:22169–74. <https://doi.org/10.1073/pnas.0900086106>.
- He S, Barco RA, Emerson D et al. Comparative genomic analysis of neutrophilic iron(II) oxidizer genomes for candidate genes in

- extracellular electron transfer. *Front Microbiol* 2017;**8**:1584. <https://doi.org/10.3389/fmicb.2017.01584>.
- Hernsdorf AW, Amano Y, Miyakawa S *et al.* Potential for microbial H₂ and metal transformations associated with novel bacteria and archaea in deep terrestrial subsurface sediments. *ISME J* 2017;**11**:1915–29. <https://doi.org/10.1038/ismej.2017.39>.
- Hershey OS, Barton HA. The microbial diversity of caves. In: Moldovan OT, Kováč Ľ, Halse S, (eds.), *Cave Ecology*. Cham: Springer International Publishing, 2018;69–90. https://doi.org/10.1007/978-3-319-98852-8_5.
- Hoang DT, Chernomor O, von Haeseler A *et al.* UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 2018;**35**:518–22. <https://doi.org/10.1093/molbev/msx281>.
- Hoover RL, Keffer JL, Polson SW *et al.* Gallionellaceae pangenomic analysis reveals insight into phylogeny, metabolic flexibility, and iron oxidation mechanisms. *mSystems* 2023;**8**:e00038–23. <https://doi.org/10.1128/msystems.00038-23>.
- Islam FS, Gault AG, Boothman C *et al.* Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature* 2004;**430**:68–71. <https://doi.org/10.1038/nature02638>.
- Jain A, Coelho A, Madjarov J *et al.* Evidence for quinol oxidation activity of ImoA, a novel NapC/NirT family protein from the neutrophilic Fe(II)-oxidizing bacterium *Sideroxydans lithotrophicus* ES-1. *mBio* 2022;**13**:e02150–22. <https://doi.org/10.1128/mbio.02150-22>.
- Ji M, Greening C, Vanwonterghem I *et al.* Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature* 2017;**552**:400–3. <https://doi.org/10.1038/nature25014>.
- Kalyaanamoorthy S, Minh BQ, Wong TKF *et al.* ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;**14**:587–9. <https://doi.org/10.1038/nmeth.4285>.
- Kang DD, Li F, Kirton E *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 2019;**7**:e7359. <https://doi.org/10.7717/peerj.7359>.
- Katoh K, Misawa K, Kuma K *et al.* MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res* 2002;**30**:3059–66. <https://doi.org/10.1093/nar/gkf436>.
- Keffer JL, McAllister SM, Garber AI *et al.* Iron oxidation by a fused cytochrome-porin common to diverse Iron-oxidizing bacteria. *mBio* 2021;**12**. <https://doi.org/10.1128/mbio.01074-21>. <https://doi.org/10.1128/mbio.01074-21>.
- Kim M, Song J, Sin SY *et al.* Cultivation of deep-sea bacteria from the Northwest Pacific Ocean and characterization of *limnobacter profundus* sp. nov., a phenol-degrading bacterium. *Front Mar Sci* 2024;**11**:1449548. <https://doi.org/10.3389/fmars.2024.1449548>.
- Könneke M, Schubert DM, Brown PC *et al.* Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc Natl Acad Sci USA* 2014;**111**:8239–44. <https://doi.org/10.1073/pnas.1402028111>.
- Kotula P, Andreychouk V, Pawlyta J *et al.* Genesis of iron and manganese sediments in Zoloushka Cave (Ukraine/Moldova) as revealed by $\delta^{13}\text{C}$ organic carbon. *Int J Speleol* 2019;**48**:222–35. <https://doi.org/10.5038/1827-806X.48.3.2255>.
- Kraemer SM, Duckworth OW, Harrington JM *et al.* Metallophores and trace metal biogeochemistry. *Aquat Geochem* 2015;**21**:159–95. <https://doi.org/10.1007/s10498-014-9246-7>.
- Krüger M, Chaudhari N, Thamdrup B *et al.* Differential contribution of nitrifying prokaryotes to groundwater nitrification. *ISME J* 2023;**17**:1601–11. <https://doi.org/10.1038/s41396-023-01471-4>.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–9. <https://doi.org/10.1038/nmeth.1923>.
- Lazar CS, Lehmann R, Stoll W *et al.* The endolithic bacterial diversity of shallow bedrock ecosystems. *Sci Total Environ* 2019;**679**:35–44. <https://doi.org/10.1016/j.scitotenv.2019.04.281>.
- Li D, Liu C-M, Luo R *et al.* MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 2015;**31**:1674–6. <https://doi.org/10.1093/bioinformatics/btv033>.
- Light SH, Su L, Rivera-Lugo R *et al.* A flavin-based extracellular electron transfer mechanism in diverse gram-positive bacteria. *Nature* 2018;**562**:140–4. <https://doi.org/10.1038/s41586-018-0498-z>.
- Makhubela TV, Kramers JD, Belyanin GA *et al.* Proterozoic ⁴⁰Ar/³⁹Ar ages from cave deposits of the Malapa, Sterkfontein and Dinaledi fossil sites, Cradle of Humankind. *South Africa Afr J Geol* 2017;**120**:21–44. <https://doi.org/10.25131/gssajg.120.1.21>.
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 2011;**17**:10–2. <https://doi.org/10.14806/ej.17.1.200>.
- Martin-Pozas T, Cuezva S, Fernández-Cortés Á *et al.* Role of subterranean microbiota in the carbon cycle and greenhouse gas dynamics. *Sci Total Environ* 2022;**831**:154921. <https://doi.org/10.1016/j.scitotenv.2022.154921>.
- McAllister SM, Polson SW, Butterfield DA *et al.* Validating the Cyc2 neutrophilic iron oxidation pathway using meta-omics of zeta-proteobacteria iron mats at marine hydrothermal vents. *mSystems* 2020;**5**. <https://doi.org/10.1128/msystems.00553-19>. <https://doi.org/10.1128/msystems.00553-19>.
- Minh BQ, Schmidt HA, Chernomor O *et al.* IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020;**37**:1530–4. <https://doi.org/10.1093/molbev/msaa015>.
- Mirdita M, Steinegger M, Breitwieser F *et al.* Fast and sensitive taxonomic assignment to metagenomic contigs. *Bioinformatics* 2021;**37**:3029–31. <https://doi.org/10.1093/bioinformatics/btab184>.
- Napieralski SA, Roden EE. The weathering microbiome of an outcropping granodiorite. *Front Microbiol* 2020;**11**:601907. <https://doi.org/10.3389/fmicb.2020.601907>.
- Newsome L, Farquharson N, Dowd T *et al.* Natural attenuation of lead by microbial manganese oxides in a karst aquifer. *Sci Total Environ* 2021;**754**:142312. <https://doi.org/10.1016/j.scitotenv.2020.142312>.
- Olm MR, Brown CT, Brooks B *et al.* dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 2017;**11**:2864–8. <https://doi.org/10.1038/ismej.2017.126>.
- Ortiz M, Legatzki A, Neilson J *et al.* Making a living while starving in the dark: metagenomic insights into the energy dynamics of a carbonate cave. *ISME J* 2014;**8**:478–91. <https://doi.org/10.1038/ismej.2013.159>.
- Ortiz M, Neilson JW, Nelson WM *et al.* Profiling bacterial diversity and taxonomic composition on speleothem surfaces in Kartch-

- ner caverns. *AZ Microb Ecol* 2013;**65**:371–83. <https://doi.org/10.1007/s00248-012-0143-6>.
- Otte J, Mall A, Schubert DM *et al.* Malonic semialdehyde reductase from the archaeon *Nitrosopumilus maritimus* is involved in the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle. *Appl Environ Microb* 2015;**81**:1700–7. <https://doi.org/10.1128/AEM.03390-14>.
- Parker CW, Senko JM, JAs A *et al.* Enhanced terrestrial Fe(II) mobilization identified through a novel mechanism of microbially driven cave formation in Fe(III)-rich rocks. *Sci Rep* 2022;**12**:17062. <https://doi.org/10.1038/s41598-022-21365-3>.
- Parks DH, Imelfort M, Skennerton CT *et al.* CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;**25**:1043–55. <https://doi.org/10.1101/gr.186072.114>.
- Pimenta AI, Paquete CM, Morgado TL *et al.* Characterization of the inner membrane cytochrome ImcH from *Geobacter* reveals its importance for extracellular electron transfer and energy conservation. *Protein Sci* 2023;**32**:e4796. <https://doi.org/10.1002/pro.4796>.
- Pisapia C, Ménez B, Bruxelles L *et al.* Impact of intraterrestrial microbial life on rock weathering at depth: case studies from cave environments. *Goldschmidt Conf Abstr* 2021;**2021**:7489. <https://doi.org/10.7185/gold2021.7489>.
- Ray AE, Zaugg J, Benaud N *et al.* Atmospheric chemosynthesis is phylogenetically and geographically widespread and contributes significantly to carbon fixation throughout cold deserts. *ISME J* 2022;**16**:2547–60. <https://doi.org/10.1038/s41396-022-01298-5>.
- Richardson DJ, Butt JN, Fredrickson JK *et al.* The ‘porin-cytochrome’ model for microbe-to-mineral electron transfer. *Mol Microbiol* 2012;**85**:201–12. <https://doi.org/10.1111/j.1365-2958.2012.08088.x>.
- Rodriguez-R LM, Gunturu S, Tiedje JM *et al.* Nonpareil 3: fast estimation of metagenomic coverage and sequence diversity. *MSystems* 2018;**3**. <https://doi.org/10.1128/msystems.00039-18>.
- Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;**30**:2068–9. <https://doi.org/10.1093/bioinformatics/btu153>.
- Segata N, Izard J, Waldron L *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;**12**:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.
- Sharma A, Taubert M, Pérez-Carrascal OM *et al.* Iron coatings on carbonate rocks shape the attached bacterial aquifer community. *Sci Total Environ* 2024;**917**:170384. <https://doi.org/10.1016/j.scitotenv.2024.170384>.
- Sharma S, Cavallaro G, Rosato A. A systematic investigation of multiheme c-type cytochromes in prokaryotes. *J Biol Inorg Chem* 2010;**15**:559–71. <https://doi.org/10.1007/s00775-010-0623-4>.
- Shi L, Dong H, Reguera G *et al.* Extracellular electron transfer mechanisms between microorganisms and minerals. *Nat Rev Micro* 2016;**14**:651–62. <https://doi.org/10.1038/nrmicro.2016.93>.
- Shi L, Fredrickson JK, Zachara JM. Genomic analyses of bacterial porin-cytochrome gene clusters. *Front Microbiol* 2014;**5**:657. <https://doi.org/10.3389/fmicb.2014.00657>.
- Shi M, Min X, Ke Y *et al.* Recent progress in understanding the mechanism of heavy metals retention by iron (oxyhydr)oxides. *Sci Total Environ* 2021;**752**:141930. <https://doi.org/10.1016/j.scitotenv.2020.141930>.
- Smith H, Zelaya AJ, De León KB *et al.* Impact of hydrologic boundaries on microbial planktonic and biofilm communities in shallow terrestrial subsurface environments. *FEMS Microbiol Ecol* 2018;**94**:fiy191. <https://doi.org/10.1093/femsec/fiy191>.
- Søndergaard D, Pedersen CNS, Greening C. HydDB: a web tool for hydrogenase classification and analysis. *Sci Rep* 2016;**6**:34212. <https://doi.org/10.1038/srep34212>.
- Spilde MN, Northup DE, Boston PJ *et al.* Geomicrobiology of cave ferromanganese deposits: a field and laboratory investigation. *Geomicrobiol J* 2005;**22**:99–116. <https://doi.org/10.1080/01490450590945889>.
- Stacheter A, Noll M, Lee C *et al.* Methanol oxidation by temperate soils and environmental determinants of associated methylotrophs. *ISME J* 2013;**7**:1051–64. <https://doi.org/10.1038/ismej.2012.167>.
- Steinberger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 2017;**35**:1026–8. <https://doi.org/10.1038/nbt.3988>.
- Straub KL, Schönhuber WA, Buchholz-Cleven BEE *et al.* Diversity of ferrous iron-oxidizing, nitrate-reducing bacteria and their involvement in oxygen-independent iron cycling. *Geomicrobiol J* 2004;**21**:371–8. <https://doi.org/10.1080/01490450490485854>.
- Suarez C, Hackl T, Wilen B-M *et al.* Novel and unusual genes for nitrogen and metal cycling in *planctomycetota*- and KSB1-affiliated metagenome-assembled genomes reconstructed from a marine subsea tunnel. *FEMS Microbiol Lett* 2023;**370**:fnad049. <https://doi.org/10.1093/femsle/fnad049>.
- Szeinbaum N, Nunn BL, Cavazos AR *et al.* Novel insights into the taxonomic diversity and molecular mechanisms of bacterial Mn(III) reduction. *Environ Microbiol Rep* 2020;**12**:583–93. <https://doi.org/10.1111/1758-2229.12867>.
- Taubert M, Overholt WA, Heinze BM *et al.* Bolstering fitness via CO₂ fixation and organic carbon uptake: mixotrophs in modern groundwater. *ISME J* 2022;**16**:1153–62. <https://doi.org/10.1038/s41396-021-01163-x>.
- Tebbo BM, Bargar JR, Clement B *et al.* Biogenic manganese oxides: properties and mechanisms of formation. *Annu Rev Earth Planet Sci* 2004;**32**:287–328. <https://doi.org/10.1146/annurev.earth.32.101802.120213>.
- Tsirigis KD, Elofsson A, Bagos PG. PRED-TMBB2: improved topology prediction and detection of beta-barrel outer membrane proteins. *Bioinformatics* 2016;**32**:i665–71. <https://doi.org/10.1093/bioinformatics/btw444>.
- Turrini P, Chebbi A, Riggio FP *et al.* The geomicrobiology of limestone, sulfuric acid speleogenetic, and volcanic caves: basic concepts and future perspectives. *Front Microbiol* 2024;**15**:1370520. <https://doi.org/10.3389/fmicb.2024.1370520>.
- Turrini P, Tescari M, Visaggio D *et al.* The microbial community of a biofilm lining the wall of a pristine cave in Western New Guinea. *Microbiol Res* 2020;**241**:126584. <https://doi.org/10.1016/j.micres.2020.126584>.
- Wang F, Gu Y, O’Brien JP *et al.* Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers. *Cell* 2019;**177**:361–9. e10. <https://doi.org/10.1016/j.cell.2019.03.029>.

- Wang L, Wang M, Wise MJ *et al.* Recent progress in the structure of glycogen serving as a durable energy reserve in bacteria. *World J Microbiol Biotechnol* 2020;**36**:14. <https://doi.org/10.1007/s11274-019-2795-6>.
- Wild B, Gerrits R, Bonneville S. The contribution of living organisms to rock weathering in the critical zone. *npj Mater Degrad* 2022;**6**:98. <https://doi.org/10.1038/s41529-022-00312-7>.
- Winogradsky S. Über eisenbakterien. *Botanische Zeitung* 1888;**46**:262–3.
- Wiseshart A, Mhuantong W, Tangphatsornruang S *et al.* Shotgun metagenomic sequencing from Manao-Pee cave, Thailand, reveals insight into the microbial community structure and its metabolic potential. *BMC Microbiol* 2019;**19**:144. <https://doi.org/10.1186/s12866-019-1521-8>.
- Yan L, Hermans SM, Totsche KU *et al.* Groundwater bacterial communities evolve over time in response to recharge. *Water Res* 2021;**201**:117290. <https://doi.org/10.1016/j.watres.2021.117290>.
- Yu H, Chadwick GL, Lingappa UF *et al.* Comparative genomics on cultivated and uncultivated freshwater and marine ‘*Candidatus Manganitrophaceae*’ species implies their worldwide reach in manganese chemolithoautotrophy. *mBio* 2022;**13**:e03421–21. <https://doi.org/10.1128/mbio.03421-21>.
- Yu H, Leadbetter JR. Bacterial chemolithoautotrophy via manganese oxidation. *Nature* 2020;**583**:453–8. <https://doi.org/10.1038/s41586-020-2468-5>.
- Yu NY, Wagner JR, Laird MR *et al.* PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 2010;**26**:1608–15. <https://doi.org/10.1093/bioinformatics/btq249>.
- Zheng J, Ge Q, Yan Y *et al.* dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res* 2023;**51**:W115–21. <https://doi.org/10.1093/nar/gkad328>.
- Zhou H, Fu C. Manganese-oxidizing microbes and biogenic manganese oxides: characterization, Mn(II) oxidation mechanism and environmental relevance. *Rev Environ Sci Biotechnol* 2020;**19**:489–507. <https://doi.org/10.1007/s11157-020-09541-1>.
- Zhou N, Keffer JL, Polson SW *et al.* Unraveling Fe(II)-oxidizing mechanisms in a facultative Fe(II) oxidizer, sideroxydans lithotrophicus strain ES-1, via culturing, transcriptomics, and Reverse Transcription-quantitative PCR. *Appl Environ Microbiol* 2022;**88**:e01595–21. <https://doi.org/10.1128/AEM.01595-21>.