Viable cyanobacteria in the deep continental subsurface

Fernando Puente-Sánchez,1,2 Alejandro Arce-Rodriguez,3,4 Monike Oggerin,4,5 Miriam García-Villadangos,5 Mercedes Moreno-Paz,6,7 Yolanda Blanco,8 Nuria Rodríguez,9 Laurence Bird,10,11 Sara A. Lincoln,12 Fernando Tornos,13 Olga Prieto-Ballesteros,14 Katherine H. Freeman,6 Dietmar H. Pieper,15 Kenneth N. Timmis,16 Ricardo Amils,14,15 and Victor Parro9

*Department of Molecular Evolution, Centro de Astrobiología, Instituto Nacional de Técnica Aeroespacial-Consejo Superior de Investigaciones Científicas (INTA-CSIC), 28850 Torrejón de Ardoz, Madrid, Spain; 1Institute of Microbiology, Technical University Braunschweig, D-38023 Braunschweig, Germany; 2Microbial Interactions and Processes Group, Helmholtz Zentrum für Infektionsforschung, 38124 Braunschweig, Germany; 3Department of Planetology and Habitability, Centro de Astrobiología, INTA-CSIC, 28850 Torrejón de Ardoz, Madrid, Spain; 4Department of Geosciences, The Pennsylvania State University, University Park, PA 16802; 5Instituto de Geociencias, CSIC-Universidad Complutense de Madrid, 28040 Madrid, Spain; and 6Centro de Biología Molecular Severo Ochoa, CSIC-Universidad Autónoma de Madrid, 28049 Madrid, Spain

Cyanobacteria are ecologically versatile microorganisms inhabiting most environments, ranging from marine systems to arid deserts. Although they possess several pathways for light-independent energy generation, until now their ecological range appeared to be restricted to environments with at least occasional exposure to sunlight. Here we present molecular, microscopic, and metagenomic evidence that cyanobacteria predominate in deep subsurface rock samples from the Iberian Pyrite Belt Mars analog (southwestern Spain). Metagenomics showed the potential for a hydrogen-based lithoautotrophic cyanobacterial metabolism. Collectively, our results suggest that they may play an important role as primary producers within the deep-Earth biosphere. Our description of this previously unknown ecological niche for cyanobacteria paves the way for models on their origin and evolution, as well as on their potential presence in current or primitive biospheres in other planetary bodies, and on the extent, primitive, and putative extraterrestrial biospheres.

Results and Discussion

Biotic evidence from the evidence for oxygenic photosynthesis and, have since come to colonize almost every environment on Earth. Here we show that their ecological range is not limited by the presence of sunlight, but also extends down to the deep terrestrial biosphere. We report the presence of microbial communities dominated by cyanobacteria in the continental subsurface using microscopy, metagenomics, and antibody microarrays. These cyanobacteria were related to surface rock-dwelling lineages known for their high tolerance to environmental and nutritional stress. We discuss how these adaptations allow cyanobacteria to thrive in the dark underground, a lifestyle that might trace back to their nonphotosynthetic ancestors.

Significance

Cyanobacteria were responsible for the origin of oxygenic photosynthesis, and have since come to colonize almost every environment on Earth. Here we show that their ecological range is not limited by the presence of sunlight, but also extends down to the deep terrestrial biosphere. We report the presence of microbial communities dominated by cyanobacteria in the continental subsurface using microscopy, metagenomics, and antibody microarrays. These cyanobacteria were related to surface rock-dwelling lineages known for their high tolerance to environmental and nutritional stress. We discuss how these adaptations allow cyanobacteria to thrive in the dark underground, a lifestyle that might trace back to their nonphotosynthetic ancestors.

Author contributions: F.P.-S., A.A.-R., K.N.T., R.A., and V.P. designed research; F.P.-S., A.A.-R., M.O., M.G.V., M.M.P., Y.B., N.R., L.B., S.A.L., F.T., O.P.-B., K.H.F., and D.H.P. performed research; F.P.-S., A.A.-R., M.O., Y.B., F.T., O.P.-B., and V.P. analyzed data; and F.P.-S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: Raw sequence data and assemblies that support the findings of this study have been deposited in NCBI under the BioProject ID PRJNA476489. MG-RAST results are deposited under project ID mgp83581.

1Present address: Systems Biology Program, Centro Nacional de Biotecnología, CSIC, 28049 Madrid, Spain.

2To whom correspondence should be addressed. Email: fpuente@cnb.csic.es.

3Present address: Molecular Bacteriology Group, Helmholtz Zentrum für Infektionsforschung, 38124 Braunschweig, Germany.

4Present address: Molecular Ecology Group, Max Planck Institut für Marine Mikrobiologie, 28359 Bremen, Germany.

5Present address: Isotope Biogeochemistry Laboratory, California Institute of Technology, Pasadena, CA 91125.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1808176115/-/DCSupplemental.

Published online October 1, 2018.
sequencing (SI Appendix, Fig. S3) and an in situ antibody microarray immunoassay (11) (SI Appendix, Fig. S4).

Detection of Deep Subsurface Cyanobacteria. Strikingly, the on-site immunoassay detected, inter alia, cyanobacterial markers in samples from several depths, and 16S rRNA gene sequencing showed a predominance of cyanobacteria, whose exact sequence variants (ESVs) were related to endolithic and hypolithic representatives of the genera *Calothrix*, *Chroococcidiopsis*, and *Microcoleus* (SI Appendix, Fig. S4 and Dataset S1). The presence of cyanobacteria was associated with local decreases in hydrogen concentrations (Fig. 1A). The cyanobacterial ESVs present in the subsurface samples were absent from the drilling fluid and the internal laboratory controls (SI Appendix, Fig. S3), which confirms that their detection is not a consequence of contamination during sample retrieval and processing, and that they are indigenous to the retrieved cores.

We further focused on the samples from 420 and 607 m of depth (from now on referred to as samples 420 and 607, respectively), as they showed higher amounts of fatty acids (SI Appendix, Fig. S2). The rocks from 420 m below the surface floor (mbsf) are dominantly made up of quartz, with minor proportions of pyrite, carbonates (ankerite), and white mica, and show a conspicuous layering between some millimeters and 1 cm. Permeability is mainly controlled by widespread unoriented fracturing, with very variable openings between 0.01 and 0.1 mm. The rocks from 607 mbsf, on the other hand, consist of alternating dark shale (with abundant centimeter-sized nodules of pyrite) and sandstone. Fractures appear in the abundant contact zones between both minerals (SI Appendix, Fig. S2). Overall, the rocks at both depths have low porosity, but the presence of fractures provides space for microbial colonization and allows for a limited input of water and nutrients.

Cyanobacteria were the most abundant organisms in the metagenomes of both samples, followed by the Ascomycota, Alphaproteobacteria, and Bacteroidetes groups (Fig. 1B and Dataset S2). These organisms may form a microbial consortium similar to those found in cyanobacterial crusts (SI Appendix, Supplementary Text). In previous work, we described the presence of biofilms in the same borehole using fluorescence microscopy (12), but only universal probes were used. In this work, we confirmed the presence of viable cyanobacteria by catalyzed reporter deposition fluorescent in situ hybridization (CARD-FISH) with specific probes against the cyanobacterial 16S ribosomal RNA. CARD-FISH is the best-practice method to search for viable cells—as defined by the presence of ribosomes—in deep subsurface settings (7, 13, 14). Ribosomal RNA has a half-life of days (15) and readily degrades upon cell starvation (16). Further, our samples contain pyrite, which is known to mediate the degradation of RNA via hydroxyl radicals under oxic and anoxic conditions (17). Sorption on certain mineral surfaces, such as clays, can increase the extracellular stability of ribosomal RNA, but complete degradation still occurs after a short time (18). Therefore, under the conditions of this study, positive CARD-FISH signals are a strong proof of extant viability.

CARD-FISH with specific probes revealed clusters of cyanobacterial cells tightly attached to the mineral matrix and associated with other microorganisms (Fig. 2). These cells did not show photosystem II-related autofluorescence, indicating that they lacked active photosynthetic pigments. The inactivation of the photosynthetic apparatus when under environmental stress is a known trait of desert-dwelling cyanobacteria, such as *Microcoleus* sp (19), which helps them cope with both desiccation and photoinhibition.

Cyanobacteria have long been known to be ecologically versatile microorganisms (20) capable of light-independent energy generation (21), but until now, their ecological range appeared to be restricted to environments with at least occasional or prior exposure to sunlight (22). A few studies have reported the presence of cyanobacteria in deep subsurface environments (23–25), but to the best of our knowledge, only in ref. 25 have the authors attempted to discuss their origin. They proposed that a bloom of aquatic cyanobacteria had been trapped thousands of years ago into a groundwater aquifer with no further connection with the surface. That scenario strongly differs from the one described in this study: we analyze rock samples instead of groundwater, the IPB subsurface aquifer has recent connection to the surface (10), and the cyanobacterial lineages detected in this work are endolithic rather than aquatic. We thus believe that our results correspond to modern cyanobacteria with the ability to colonize deep subsurface environments.

Hydrogen as an Electron Donor for Cyanobacteria. We found an apparent inverse correlation between the cyanobacteria predominance and hydrogen concentration in our samples (Fig. 1A). Hydrogen can be produced in the subsurface by several abiotic mechanisms, and its concentration in deep continental settings has recently been found to be controlled by biological sinks (26). To identify putative hydrogen consumers, we tested whether hydrogen concentration was dependent on taxa abundances using multiple linear regression. We considered the phylum, class, order, and family levels and tested models including all possible combinations of one to six taxa. Cyanobacteria was the only taxon that significantly explained hydrogen abundances when considered alone (negative correlation, \(P = 0.03, R^2 = 0.33\)). The addition of more taxa to the model helped explain residual variance. The best model included the cyanobacterial families

![Fig. 1. Distribution of cyanobacteria and other microbial taxa in core samples from the Iberian Pyrite Belt deep subsurface. (A) Relative abundance of different cyanobacterial clades based on 16S rRNA amplicons (bars) and hydrogen concentrations (purple area) across the borehole samples. Hydrogen concentrations are shown relative to the sample with the highest hydrogen concentration. (B) Taxonomic composition of the metagenomic reads from samples 420 and 607.](image-url)
In several cyanobacterial genera, the overreduction of plastoquinone (PQ) is observed, which can be attributed to the presence of respiratory enzymes such as quinol oxidase (33). This has a marked effect on the redox state of the cell. It is hypothesized that they obtain their energy by coupling the oxidation of hydrogen to the reduction of different electron acceptors (Fig. 3; see discussion below).

Hydrogenases are widespread in cyanobacteria, which are believed to have originated from hydrogenotrophic ancestors (27). We detected both uptake and bidirectional hydrogenases in the two retrieved cyanobacterial pangenomes (Dataset S2). The uptake hydrogenase (Hup) transfers electrons from hydrogen to an unknown acceptor from the electron transport chain, most likely plastoquinone (28). Its main function is to minimize energy losses during nitrogen fixation and protect nitrogenase from oxygen toxicity by transferring electrons from hydrogen to oxygen via the electron transport chain (29). Cyanobacteria also have a bidirectional hydrogenase (Hox), which is hypothesized to function as an electron valve, providing a rapid way to balance the redox state of the cell. Hox can transfer electrons from/to either NAD(P) or plastoquinone via the NDH-I complex, contributing to both hydrogen uptake (providing reducing power for CO₂ fixation) and hydrogen production (in dark-to-light transitions, and also coupled to fermentation) (21, 28, 30). Crucially, the cyanobacterial Hox has also been shown to be induced under dark anaerobic conditions (31), to participate in respiratory electron flow under prolonged darkness, and to be essential for growth when the photosynthetic and respiratory electron transport chains are overreduced (32).

**Cyanobacterial Electron Transport Chains in Deep Subsurface Environments.** In several cyanobacterial genera, the overreduction of plastoquinone triggers the transfer of electrons to extracellular acceptors, via a cytochrome bd quinol oxidase (33). This has a protective effect in light-intensive conditions, where cytochrome bd is unable to accept electrons from plastoquinone at a sufficient rate. We note that growth under the dark, anoxic conditions of the deep subsurface would also lead to an overreduction of the plastoquinone pool, potentially triggering electron transfer from plastoquinone to cytochrome bd quinol oxidase. We thus propose that this protection mechanism would also provide the means for the anaerobic oxidation of hydrogen or other compounds using extracellular electron acceptors such as iron and manganese oxides, or phenolic compounds derived from the degradation of recalcitrant organic matter by other members of the microbial community (SI Appendix, Supplementary Text).

An additional potential electron acceptor could be nitric oxide, as we found a quinol-dependent nitric oxide reductase in the cyanobacterial pangenome from sample 420 (Dataset S2). Cyanobacterial nitric oxide reductases connected to the electron transport chain have been proposed to participate in nitric oxide detoxification and energy conversion (34). Interestingly, incomplete denitrification by noncyanobacterial partners is predominant in cyanobacteria-dominated biological crusts, leading to the emission of nitric and nitrous oxides (35, 36). Thus, cyanobacteria might profit from nitric oxide reductases by using them to exploit the nitric oxide produced by other members of the consortium as an alternative electron acceptor (34).

**Fig. 2.** Fluorescence micrographs (CARD-FISH) showing the presence of clusters of cyanobacterial cells attached to rock surfaces in deep subsurface samples. 
(A and E) Microbial DNA stained with DAPI (blue signal). (B and F) Hybridization signals with the specific cyanobacterial oligonucleotide probe CYA361 (red signal). (C and G) Merged image of DAPI and probe hybridization signals (blue and red, respectively). (D and H) Merged image of DAPI and probe hybridization together with the mineral matrix. The gray and white signal shows the host mineral. (Scale bars, 5 μm in all cases.)

**Fig. 3.** Schematic representation of the photosynthetic, respiratory, and fermentative pathways detected in the cyanobacterial pangenomes of two deep subsurface metagenomes. Orange and blue squares indicate whether an enzyme was detected in sample 420 or sample 607, respectively. Enzymes detected in the metagenomic reads but not in the assemblies have their square marked with a diagonal hatching. Reactions dependent on light or oxygen, and thus unlikely to be active in anoxic deep subsurface environments, are marked with a red cross. Abbreviations: cyt bd, cytochrome bd; cyt bd, cytochrome bd; Fd, ferredoxin; Hox, bidirectional hydrogenase; Hup, uptake hydrogenase; NDH, NDH-I complex; NorB, quinol-dependent nitric oxide reductase; Ox, cytochrome c oxidase; PC, plastocyanin; PQ, plastoquinone; PSI, photosystem I; PSII, photosystem II; SDH, succinate dehydrogenase; TCA, tricarboxylic acid cycle.
Evolution of Light-Independent Electron Transport Chains in Cyanobacteria. 

Hochler and Jørgensen (4), and more recently Stanislawski et al. (37), have argued that the slow rates of biomass turnover in deep subsurface environments provide minimal opportunities for the introduction and propagation of beneficial mutations. Survivability will thus be determined by traits gained in other ecosystems showing similar (to some extent) restrictions, but with higher energy fluxes. In this context, endolithic cyanobacteria are perfect candidates for inhabiting the deep subsurface, as they are already adapted to living inside rocks and are able to withstand severe nutritional and environmental stresses and experience periodic anoxia during the diel cycle (38). Further, some cave-dwelling cyanobacteria survive for long periods in the near-total absence of light, where photosynthesis is no longer possible (39). For these reasons, they may have evolved defense mechanisms that, having evolved to cope with light stress and desiccation in their original habitats (40), could also be triggered under the reducing conditions found in the deep subsurface and result in functional electron transport chains. This proposed mechanism relies on traits that are conserved across cyanobacterial lineages, and might thus reflect the lifestyle of the nonphotosynthetic ancestor of cyanobacteria (27). Under this second hypothesis, part of the energy transduction machinery of such an ancestor would have been coopted to serve as stress defense mechanisms in cyanobacteria, while still retaining its original capabilities in the absence of light.

Conclusions

We report the existence of cyanobacteria-dominated microbial communities in the deep continental subsurface, and discuss their potential metabolism based on geochemical and metagenomic data. Our proposal of cyanobacterial hydrogenotrophy is consistent with a large body of literature, as well as several parallel lines of evidence presented in this work. While the dark metabolism of cyanobacteria is recognized, previous work on micro- and macro-algal samples from this and other studies calls for a reevaluation of their potential roles in deep subsurface ecosystems and increases their relevance in early life and astrobiological scenarios.

Materials and Methods

Drilling and Sampling. Boreholes were continuously cored by rotary diamond-bit drilling using a Boart Longyear HQ wireline system producing 3 m of 60-mm-diameter cores. Well water was used as a drilling fluid to lubricate the bit and displace air to the surface. Fluids were recirculated. To detect potential contamination of the samples, sodium bromide (200 ppm) was added to the drilling fluid as a marker. Upon retrieval from the drilling rig, cores were divided into 60-cm-length pieces, inspected for signs of alteration, and stored in boxes for permanent storage and curation in the Instituto Geológico Minero de España lithotube in Perlarroja. Selected cores were deposited in place with a core splitter and refrigerated with ethanol and a 5% formaldehyde solution. Aseptic handling of small fragments of rock. Alexa Fluor 594-labeled tyramide was added to each sample and was incubated for 30 min at room temperature. Hybridization was performed following the instructions. Buffer was used as a blank control sample in parallel analyses.

Physicochemical Characterization of Rock Core Samples. The concentrations of inorganic anions such as nitrate, nitrate, and sulfate, and small molecular weight organic acids such as acetate, formate, and oxalate, were estimated by ion chromatography as described elsewhere (11). pH was measured as described in ref. 41. The amounts of occluded hydrogen, methane, and CO2 in rock pores were measured as follows: 10 g of rock shards were placed into 100-ml vials under sterile and anoxic conditions. The vials were in turn sealed with a gas-tight rubber septum and an aluminum cap, and their headspace was flushed with nitrogen gas. After a year of incubation at room temperature, it was assumed that the gases originally present in the rock pores had reached equilibrium with the headspace. The concentrations of hydrogen, methane, and CO2 were then measured in a Bruker 450GC gas chromatographer using a Hayesep 80/100 column (Valco Instruments). The presence of Fe3+, Fe2+, and NH4+ was assessed by using the Reflectoquant system (Merck Millipore), in accordance with the manufacturer’s instructions.

Sandwich Microarray Immunoassays with LDChip. Sandwich-type microarray immunoassays (SMIs) were performed as described previously (11). Briefly, printed microsphere slides with LDChip300 antibody microarray were blocked with 5% (vol/vol) BSA for 30 min and then incubated with Alexa Fluor 594-labeled tyramide (prepared as described in ref. 41) for 30 min. After washing with TBS buffer (0.15 M NaCl, 10 mM NaPO4·2H2O, 0.1% Tween 20) and blocking with 1% (vol/vol) BSA in 0.01 M PBS for 30 min). After washing with TBS buffer (0.4 M Tris-HCl 0.1 M NaCl, 0.1% Tween 20) and drying the chip by vacuum centrifugation, the slides were mounted on a portable multiarray analysis module (MAAM) cassette for drying. Approximately 0.5 g of ground core samples were resuspended in 2 mL of TBS and sonicated (5 s x 1 min cycle with a handheld Sonifier Ultrasonic Processor, UPROR-100, Heat Systems-Ultrasonics) and transferred with the help of a 20-gauge needle. After centrifugation, the supernatant was removed by filtering through 10-μm nylon filter and 50 mL of the extracts were injected into each MAAM chamber and incubated for 1 h with the LDChip300 at ambient temperature. After a wash with TBS buffer, the chips were incubated with a fluorescently labeled antibody mixture for 1 h. The slides were then washed, dried, and scanned for fluorescence at 635 nm in a GeNePix100A scanner. Buffer was used as a blank control sample in parallel immunoassays. The scanned images were analyzed in the field with the GenePix Pro software (Genomic Solutions). The final fluorescence intensity was quantified as previously reported (11).

CARD-FISH. For CARD-FISH analysis, rock powders and small chips were fixed in 4% (vol/vol) formaldehyde for 4 h, then washed twice with 1× PBS and stored at −20 °C in 1× ethanol:1× PBS. Approximately 150 mg of small chips were subsequently subjected to analysis. Samples were embedded in 0.2% (wt/vol) agarose. Endogenous peroxidases were inactivated in 0.1% H2O2 in methanol for 30 min at room temperature. Hybridization was performed following the method described in ref. 43 with some modifications to facilitate the handling of small fragments of rock. Alexa Fluor 594-labeled tyramide was used as a fluorochrome. Samples were counterstained with 1 μg/ml 4′,6-diamidino-2-phenylindole (DAPI). The oligonucleotide probe used in this study for targeting rRNA genes was CYA361 (5′-CCCATTGCGGAAAATTCC-3′) (44) at 35% (vol/vol) formamide concentration. HRF-labeled probes were quantified by comparing analyte peaks from this and other studies calls for a reevaluation of their potential roles in deep subsurface ecosystems and increases their relevance in early life and astrobiological scenarios.
were synthesized by Blomers.net GmbH. Negative controls were performed with the control probe NON338 (5′-CTCTCTAAGGGACACTG-3′). (45). Additional controls with an amplicon of 16S rRNA were analyzed with the SqueezeM amplification method (46). The SqueezeM method was additionally isolated from a drilling water sample, to trace potential contaminations. The extraction of DNA from different depths with either one or several of the above-described methods is summarized in Table S1. DNA Extraction, Amplification, and Sequencing. DNA extraction was performed in a UV- and ethanol-sterilized flow chamber according to ref. 46. Briefly, 0.5 g of powdered core sample were introduced into an Ultra-Clean Bead Tube (MoBio Laboratories), whose original buffer had been previously removed and substituted with 1 mL of phosphate buffer (1 M sodium phosphate, 15% ethanol). After adding 60 µL of MoBio Ultra-Clean Soil DNA solution, the tubes were processed in a bead-beating chamber and subjected to two FastPrep (MP Biomedicals) cycles (30 s each, power setting of 5.5 mJ) separated by 1 min of ice cooling. Subsequently, the tubes were incubated in a thermomixer at 80 °C for 40 min, while shaking at 300 rpm. The MoBio Ultra-Clean Soil DNA extraction protocol was then followed from the addition of solution S2, according to the manufacturer’s instructions. All materials and stock solutions were UV sterilized for 5 min in either a GS Linker UV Chamber (Bio-Rad Laboratories) or a Stratalinker 1800 UV crosslinker (Stratagene) to eliminate trace DNA contaminations. The isolated DNA was later subjected to multiple displacement amplification (MDA) using either the MagniPhi Phi29 polymerase (Genetex, formerly X-Pol Biotech) or the REPLi-g Single Cell Kit (Qiagen). The nonenzymatic MDA reagents and the random hexamers were decontaminated following ref. 47. Briefly, they were aliquoted into 0.2-mL PCR tubes, which were laid down horizontally on the UV crosslinker chambers and subjected to total UV dose of 6.0 mJ. The resulting amplification products were finally purified using a Microspin G-50 column (GE Healthcare). Successful amplification was confirmed by PCR of the 16S rRNA gene using primers 1565F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1685r (5′-CACGGGCTGACGCGCCG-3′). Negative controls were also run from MDA reagents without template DNA. Furthermore, in the cases where the MDA-amplified DNA gave no 16S PCR products, another nine individual rounds of DNA isolation were performed using the same extraction protocol (for a total of ~5 g of powdered core sample). The 10 DNA isolations from the same rock core sample were immediately pooled, cleaned with phenol:chloroform:isoamyl alcohol (25:24:1), and precipitated with 70% ethanol (also UV sterilized) as described elsewhere. The DNA pellet was finally eluted in 50 µL of 10 mM Tris pH 8.0, amplified by MDA, and purified using the Microspin G-50 columns. Additionally, an upscaled modification of the previous protocol was applied to a set of samples using the PowerMax Soil DNA Isolation Kit (MoBio Laboratories). A total of 15 mL of phosphate buffer was mixed with up to 10 g of powdered core sample into the provided PowerMax Bead Tube. After the addition of 5 mL of PowerMax Soil DNA solution C1, the tubes were vortexed vigorously for 30 s and subjected to two FastPrep cycles (40 s, power setting of 6.0 mJ). Samples were then incubated at 80 °C in a water bath during 40 min and centrifuged at 2,500 × g for 3 min at room temperature. The supernatant was recovered in a new collection tube and the protocol was followed from the addition of solution C2. Once again, the solutions employed were UV sterilized as described above, and the resulting DNA was also subjected to MDA amplification, purification, and PCR of the 16S rRNA gene. The extraction of DNA from different depths with either one or several of the above-described methods is summarized in Table S1. DNA was additionally isolated from a drilling water sample, to trace potential contamination events occurring during the core retrieval process. Two 250-µL drill cooling water samples, collected with a 15-d difference, were pooled and filtered through a 0.22-nm pore size filter (Millipore). The filter was introduced into an empty Ultra-Clean Bead Tube and the DNA was isolated using the first protocol described above. Additionally, during the preparation of the aliquots used for MiSeq sequencing (see below) DNA was also extracted from an empty PowerMax Bead Bead Tube to account also for laboratory contamination events during extraction of the nucleic acids with non-MDA control). DNA was finally eluted in 50 µL of 10 mM Tris HCl buffer and quantified using a NanoDrop 2000 spectrophotometer (Thermo Scientific). The VS-V6 hypervariable regions of the bacterial 16S rRNA gene were PCR amplified using primers 807F and 1050R (48). The barcoding of the DNA amplicons, as well as the addition of illumina adaptors, was carried out as described previously (49). The PCR mix for 16S rRNA amplicon was sent for 250-bp paired-end sequencing on an illumina MiSeq platform (Illumina) at the Genome Analytics Platform of the Helmholtz Centre for Infection Research. For the construction of metagenomic libraries, 0.8 µg of amplified DNA were mixed with 1x fragmentase reaction buffer in a final volume of 18 µL, vortexed thoroughly, and incubated on ice for 5 min. The fragmentation reaction was then started by mixing the samples with 2 µL of NEBNext dsDNA fragmentation (New England Biolabs, Inc.) and carried out for 25 min at 37 °C. After incubation, the fragmentation was halted by the addition of 5 µL of 0.5 M EDTA. The ensuing DNA was purified with the QIAquick PCR Purification Kit (Qiagen) and eluted in a final volume of 35 µL before quantification with a Nanodrop (Thermo Scientific). Metagenomic libraries were prepared with the NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs, Inc.) using ~200 ng of fragmented DNA as initial input. Size selection of 400-500 bp DNA library fragments was carried out using the Agencourt AMPure XP magnetic beads (Beckman Coulter, Inc.) according to the NEBNext Ultra DNA Library Prep Kit instructions. Each metagenomic DNA library was sequenced (100 nt-paired end sequencing) with the illumina HiSeq. 2500 platform using the TruSeq SR Cluster Kit, v3-cBot-HS (Illumina). 16S Community Profiling. Raw 16S MiSeq paired reads were assembled and quality filtered with moira (50) (v 1.3.2) with the -q posterior flag, and then MISA (51) was used to remove self-matches (52). The reads were subjected to multiple displacement amplification (MDA) using either the MagniPhi Phi29 polymerase (Genetex, formerly X-Pol Biotech) or the REPLi-g Single Cell Kit (Qiagen). The nonenzymatic MDA reagents and the random hexamers were decontaminated following ref. 47. Briefly, they were aliquoted into 0.2-mL PCR tubes, which were laid down horizontally on the UV crosslinker chambers and subjected to total UV dose of 6.0 mJ. The resulting amplification products were finally purified using a Microspin G-50 column (GE Healthcare). Successful amplification was confirmed by PCR of the 16S rRNA gene using primers 1565F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1685r (5′-CACGGGCTGACGCGCCG-3′). Negative controls were also run from MDA reagents without template DNA. Furthermore, in the cases where the MDA-amplified DNA gave no 16S PCR product, another nine individual rounds of DNA isolation were performed using the same extraction protocol (for a total of ~5 g of powdered core sample). The 10 DNA isolations from the same rock core sample were immediately pooled, cleaned with phenol:chloroform:isoamyl alcohol (25:24:1), and precipitated with 70% ethanol (also UV sterilized) as described elsewhere. The DNA pellet was finally eluted in 50 µL of 10 mM Tris pH 8.0, amplified by MDA, and purified using the Microspin G-50 columns. Metagenomic Analysis. Paired raw metagenomic reads were quality filtered with PRINSEQ (58) with the following parameters: -deep 14 -1 lc method -l threshold 7 -ns max 2 -trim_left 20 -trim_right 20 -trim_qual_left 20 -trim_qual_right 20 -min_qual_mean 25. Reads with more than one mismatch in the barcode sequence were removed. The filtered dataset was then subjected to digital normalization using khmer (59). Due to low coverage, a single-step normalization was performed as recommended in the author’s webpage (khmer.readthedocs.org/en/v1.1/guide.html). To achieve this, the normalize-by-median.py script included with khmer was run with the following parameters: -c -k -20 -N -4 -x 2.5e8--paired. The filtered and normalized reads were annotated in the Metagenomic Rapid Annotations using Subsystems Technology (MG-RAST) server (60). Reads were classified based on the taxonomic classification reported by MG-RAST, and the four main phyla groups (Bacteria, Firmicutes, Actinobacteria, and Bacteroidetes) were analyzed independently to reconstruct their putative metabolisms, using the MG-RAST subsystem-based classification; if this led to missing enzymes for a pathway of interest, we manually searched for known synonyms in the annotation using the MG-RAST metagenome overview page. Additionally, paired raw reads were analyzed with the SqueezeEM pipeline (ref. 61; https://github.com/tamames/squeezeEM, commit 24264641d1f613808bd6f9b1c70924924dc10708c) using the sequential mode. Briefly, reads from each sample were assembled using Bfree (62) and barcodes and rRNA coding sequences were predicted with Prodigal v2.6.2 (63) and barnarr v0.9-dev (https://github.com/steemann/barnarr), respectively. ORFs were aligned against the Clusters of Orthologous Groups (COG) database and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database using diamond (66), and functionally annotated. Taxonomic annotation of ORFs was performed with a last common ancestor algorithm, based on the best hits against the NCBI nr database (downloaded on June 13, 2018). Binning was not attempted, as MDA introduces amplification biases which would affect contig coabundance patterns. Sequencing and assembly stats are summarized in Table S2. ACKNOWLEDGMENTS. We thank Javier Tamames, Carlos Pedros-Albino, Antoni Pujol-Roca, and O. Mora-Balsa,Ⓖons from the Biomining-Lobos (CSIC-CS) for their feedback and discussions of the draft manuscript and all the members of the Iberian Pyrite Belt Subsurface Life (IPBSL) team and the personnel of the
31. Schwarz C, Poss Z, Hoffmann D, Appel J (2010) Hydrogenases and hydrogen metabolism in cyanobacteria. J Bacteriol 192:2886–2892.
32. Philippot L, et al. (2010) The ecological coherence of high bacterial taxonomic ranks. Proc Natl Acad Sci USA 107:12596–12601.
33. Pisciotta JM, Zou Y, Baskakov IV (2011) Role of the photosynthetic electron transfer components with the photosynthetic electron transfer chain in thylakoid membranes. Microbiol Mol Biol Rev 75:719–762.
34. Bösch A, Friedrich B, Cramm R (2002) Characterization of the norA gene, encoding nitric oxide reductase, in the nondenitrfying cyanobacterium Synechocystis sp. strain PCC6803. Appl Environ Microbiol 68:668–672.
35. Abed RM, Lam P, de Beer D, Stief P (2013) High rates of denitrification and nitrous oxide emission in arid biological soil crusts from the Sultanate of Oman. ISME J 7:1862–1875.
36. Weber B, et al. (2013) Biological soil crusts accelerate the nitrogen cycle through large NO and HONO emissions in drylands. Proc Natl Acad Sci USA 112:15384–15389.
37. Starnawski P, et al. (2017) Microbial community assembly and evolution in subseafloric sediment. Proc Natl Acad Sci USA 114:2290–2295.
38. Rajeev L, et al. (2013) Dynamic cyanobacterial response to hydration and dehydration in a desert biological soil crust. ISME J 7:2178–2191.
39. Giordano M, Mobiili F, Pezzoni V, Heim MK, Davis JS (2000) Phototaxis in the caves of Frasassi (Italy). Physiolgia 59:25–34.
40. Raanen H, et al. (2016) Towards clarifying what distinguishes cyanobacteria able to resurface after desiccation from those that cannot. The phototrophic aspect. Biochim Biophys Acta 1857:715–722.
41. Fernández-Romolar DC, et al. (2008) Underground habitats in the Rio Tinto basin: A model for subsurface life habitats on Mars. Astrobiology 8:1023–1047.
42. Talbot HM, Rohmer M, Farrow M (2007) Rapid structural elucidation of composite bacterial haptophores by atmospheric pressure chemical ionisation liquid chromatography/mass spectrometry. Rapid Commun Mass Spectrom 21:880–892.
43. Pei YX, Portner H, Spindler F (2004) Sensitive multi-color fluorescence detection of site hybridization for the identification of environmental microorganisms. Molecular Microbial Ecology Manual, eds Kowalchuk GA, de Brujin FJ, Head IM, Akkermans AD, van Elsas JD (Kluwer, Dordrecht, The Netherlands), pp 71–76.
44. Schleifer KH, et al. (1999) Identification of cyanobacteria with horseradish peroxidase-labeled, RNA-targeted oligonucleotide probes. Appl Environ Microbiol 65:1259–1267.
45. Wallner G, Amann R, Beiser W (1993) Optimizing fluorescent in situ hybridization with mR-targeted oligonucleotide probes for flow cytometric identification of microorganisms. Cytometry 14:136–143.
46. Direito SO, Maeres A, Röling WF (2012) Sensitive life detection strategies for low-biomass environments: Optimizing extraction of nucleic acids adsorbing to terrestrial and Mars analogue minerals. FEMS Microbiol Ecol 81:111–123.
47. Woyte T, et al. (2011) Decontamination of MDA reagents for single cell whole genome amplification. PLoS One 6:e26611.
48. Bohorquez LC, et al. (2012) In-depth characterization via complementing culture-independent approaches of the microbial community in an acidic hot spring of the Colanandu Andes. Microb Ecol 63:103–115, and erratum (2012) 63:238.
49. Camarinha-Silva A, et al. (2014) Comparing the anterior nare bacterial community of two discrete human populations using Illumina amplicon sequencing. Environ Microbiol 16:2939–2952.
50. Puente-Sánchez F, Aguina J, Parro V (2016) A novel conceptual approach to read-filtering in high-throughput amplicon sequencing studies. Nucleic Acids Res 44:e40.
51. Schloss PD, et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541.
52. Kojicji JI, Westcott SL, Baster NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curion pipeline for analyzing amplicon data sets on the MiSeq Illumina platform. Appl Environ Microbiol 79:5112–5120.
53. Quast C, et al. (2013) The SILVA ribosomal RNA gene database project: Improvement of the database and web-based tools. Nucleic Acids Res 41:D590–D596.
54. Pruesse E, et al. (2007) SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:352–356.
55. Tikhonov M, Leach RW, Wingreen NS (2015) Interpreting 15 metagenomic data without clustering to achieve sub-OTU resolution. ISME J 9:686–80.
56. Rogens T, Flouri T, Nicholls B, Quince C, Mahé F (2016) VSEARCH: A versatile open source tool for metagenomics. PeerJ 4:e2584.
57. Pruesse E, Pielies J, Glückner FO (2012) SILVA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28:1823–1829.
58. Schneider R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863–864.
59. Brown CT, Howe A, Zhang Q, Pyrocos AB, Brom TH (2012). A reference-free algorithm for computational normalization of shotgun sequencing data. arXiv:1203.4802. Preprint, posted May 21, 2012.
60. Meyer F, et al. (2008) The metagenomics RAST server–A public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9:386.
61. Tamames J, Puente-Sánchez F (2008) SqueezeM, a fully automatic metagenomic analysis pipeline from reads to bins. bioRxiv:10.1101/347559. Preprint, posted June 14, 2018.
62. Meyer F, et al. (2017) The COG database: An updated version includes eukaryotes. BMC Bioinformatics 4:41.
63. Kanaispa M (2002) The KEGG database. Novartis Factsyn Sump 247:91–101, discussion 101–103, 119–128, 244–252.
64. Bunkii B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60.

Puente-Sánchez et al.