



Viable cyanobacteria in the deep continental subsurface

Fernando Puente-Sánchez^{a,1,2}, Alejandro Arce-Rodríguez^{b,c,3}, Monike Oggerin^{d,4}, Miriam García-Villadangos^a, Mercedes Moreno-Paz^a, Yolanda Blanco^a, Nuria Rodríguez^d, Laurence Bird^{e,5}, Sara A. Lincoln^e, Fernando Tornos^f, Olga Prieto-Ballesteros^d, Katherine H. Freeman^e, Dietmar H. Pieper^c, Kenneth N. Timmis^b, Ricardo Amils^{d,g}, and Víctor Parro^a

^aDepartment of Molecular Evolution, Centro de Astrobiología, Instituto Nacional de Técnica Aeroespacial-Consejo Superior de Investigaciones Científicas (INTA-CSIC), 28850 Torrejón de Ardoz, Madrid, Spain; ^bInstitute of Microbiology, Technical University Braunschweig, D-38023 Braunschweig, Germany; ^cMicrobial Interactions and Processes Group, Helmholtz Zentrum für Infektionsforschung, 38124 Braunschweig, Germany; ^dDepartment of Planetology and Habitability, Centro de Astrobiología, INTA-CSIC, 28850 Torrejón de Ardoz, Madrid, Spain; ^eDepartment of Geosciences, The Pennsylvania State University, University Park, PA 16802; ^fInstituto de Geociencias, CSIC-Universidad Complutense de Madrid, 28040 Madrid, Spain; and ^gCentro de Biología Molecular Severo Ochoa, CSIC-Universidad Autónoma de Madrid, 28049 Madrid, Spain

Edited by David M. Karl, University of Hawaii, Honolulu, HI, and approved August 29, 2018 (received for review May 11, 2018)

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 extant, primitive, and putative extraterrestrial biospheres.

astrobiology | extreme environments | deep/dark biosphere | endolithic cyanobacteria | metagenomics

The deep terrestrial biosphere is receiving increasing interest as it harbors a significant fraction of the total microbial biomass of the planet (1–3), yet also imposes severe energy and nutrient limitations on its inhabitants (4). Subsurface life is dependent on buried organic matter and lithogenically sourced compounds such as molecular hydrogen (5). Despite its inhospitality, the deep subsurface is a stable and sheltered environment, which makes it a good candidate habitat for the development of early life on Earth, as well as on potential extraterrestrial scenarios (6). Despite the increasing interest in deep subsurface ecosystems, obtaining reliable data on their microbiology is severely hampered by the high cost and technical difficulty of retrieving pristine samples, its heterogeneity and extent, and the vanishingly small microbial loads (7). The majority of deep subsurface habitats thus remain significantly underexplored (6).

The Iberian Pyrite Belt (IPB) in southwestern Spain hosts one of the largest sulphide deposits in the world, as well as the Río Tinto Mars analog (8–10). In previous work, we characterized pyrite-rich drill core samples down to 166-m depth with several molecular ecology techniques, and revealed a subsurface ecosystem with active iron and sulfur cycles (9). To investigate the microbiota and their activities in a deeper, more pristine location, we carried out a second drilling at a site ~500 m from the previous one, where geophysical studies and tritium measurements showed the presence of an aquifer formed by ~60-y-old groundwater at a depth of around 400 m (10). A 613-m-deep borehole was drilled, implementing procedures for aseptic sampling and for tracing potential contamination events during drilling and subsequent processing of the retrieved rock cores. These procedures minimized contamination of samples with

extraneous microbes, enabled us to identify any samples that had been contaminated, and if so, to quantify any such contamination.

Results and Discussion

Biogeochemical Characterization of the Deep Subsurface of the Iberian Pyrite Belt. Rock cores from a 613-m-deep borehole were retrieved and processed following protocols for aseptic sampling and the tracing of potential contaminations (*SI Appendix, Fig. S1*). The geochemistry of the borehole was characterized with several complementary techniques (*SI Appendix, Fig. S2*), and microbial profiles were obtained via MiSeq 16S rRNA gene amplicon

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 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Data deposition: Raw sequence data and assemblies that support the findings of this study have been deposited in NCBI under the BioProject ID [PRJNA476489](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA476489). MG-RAST results are deposited under project ID [mgp83581](https://metagenomics.github.io/MG-RAST/).

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

²To whom correspondence should be addressed. Email: fpunte@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.

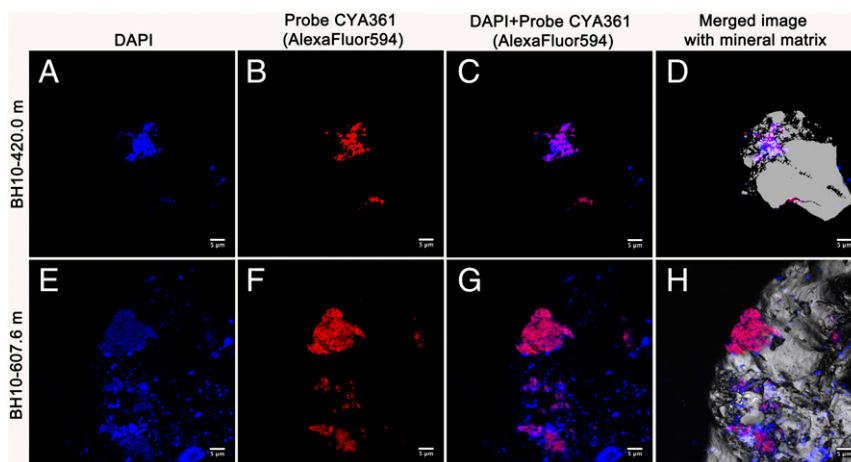


Fig. 2. Fluorescence micrographs (CARD-FISH) showing the presence of clusters of cyanobacterial cells attached to rock surfaces in deep subsurface samples. (A and E) Microbial DNA stained with DAPI (blue signal). (B and F) Hybridization signals with the specific cyanobacterial oligonucleotide probe CYA361 (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.)

Rivulareaceae and Xenococcaceae, as well as the noncyanobacterial families Sphingomonadaceae and Bradyrhizobiaceae (negative correlation, $P = 0.003$, $R^2 = 0.80$). These results, together with the presence of metabolic pathways for hydrogen utilization in the protein-coding metagenomic sequences affiliated with cyanobacteria (cyanobacterial pangenome), lead us to hypothesize 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 CO_2 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-intense conditions, where cytochrome *b₆f* 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).

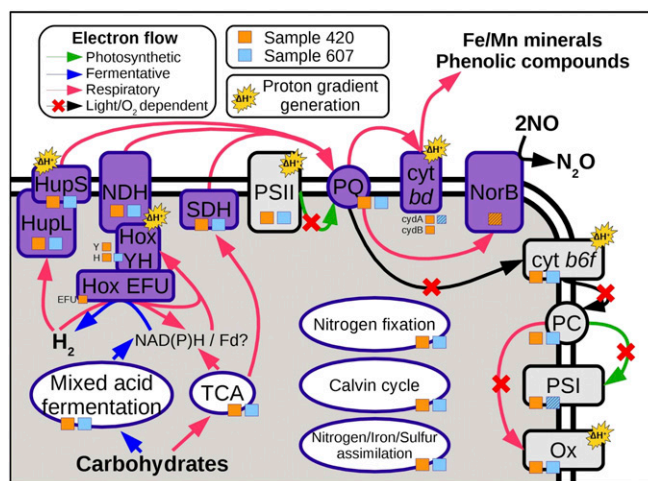


Fig. 3. Schematic representation of the photosynthetic, respiratory, and fermentative pathways detected in the cyanobacterial pangenomes of two deep subsurface metagenomes. Orange and blue squares indicate whether an enzyme was detected in sample 420 or sample 607, respectively. Enzymes detected in the 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 b₆f*, cytochrome *b₆f*; *cyt bd*, cytochrome *bd*; Fd, ferredoxin; Hox, bidirectional hydrogenase; Hup, uptake hydrogenase; NDH, NDH-1 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. Hoehler and Jørgensen (4), and more recently Starnawski 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). Further, they possess several defense mechanisms that, having likely 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 still a matter of ongoing research, their unequivocal presence in 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 Bort 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 return cuttings 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 lithoteque in Peñarroya. Selected cores were deposited in plastic bags, oxygen was displaced with N₂, sealed, and transported to a field laboratory located at the Museo Minero de Riotinto village. Upon arrival at the field laboratory, cores were placed in an anaerobic chamber (5% H₂, 95% N₂), logged, and photographed. The anaerobic chamber and the airlock were decontaminated daily with Virkon S (Antec International Limited), a mixture of surfactants, organic acids, and strong oxidizers with the ability to disrupt bacterial membranes and degrade their nucleic acids. Furthermore, the chamber and the airlock were cleaned with ethanol and a 50:50 bleach: water solution before the introduction of a new core sample. Once in the anaerobic chamber, aseptic subsamples were obtained by splitting cores with a hydraulic core splitter and drilling out the central untouched portion with a rotary hammer with sterile bits. Bit temperature was strictly controlled (maximum 40 °C) with an infrared thermometer. Subsamples intended for CARD-FISH were instead obtained by chipping away bits of the central portion of the split core with a sterile chisel and grounding them with a sterile mortar and pestle.

Physicochemical Characterization of Rock Core Samples. The concentrations of inorganic anions such as nitrite, 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.

Total Sugars, Proteins, and Total Organic Carbon Determination. Proteins, sugars, and total organic carbon were measured as described in ref. 11.

Lipid Extraction and Characterization. Rock powders were extracted using a modified Bligh–Dyer method (42). Samples were placed in a 250-mL Teflon bottle, submerged in a monophasic solution of 4:10:5 water:methanol:dichloromethane, and disrupted with a sonicator wand (Branson Ultrasonics) for 1 h while maintained on ice. Following sonication, the bottles were shaken at 200 rpm for 1 h and centrifuged at 1,500 rpm for 15 min. The supernatant was removed and the extraction repeated twice. A total of 10 mL of dichloromethane and 10 mL of water were added to the pooled supernatant to induce phase separation, and the organic phase was collected. The aqueous phase was extracted with 10 mL of dichloromethane two additional times, and the pooled organic phases were dried under N₂ and weighed. A total of 2.5 µg of pregnane diol was added to each sample as an internal standard. The extracts were acetylated by dissolving them in 50 µL of acetic anhydride and 50 µL of pyridine and heating them at 60 °C for an hour, after which they were dried and redissolved in 100 µL of dichloromethane. Samples were analyzed on a Trace 1310 GC coupled to an ISQ LT single quad mass spectrometer (Thermo Scientific). The programmable temperature vaporizing inlet was operated in constant temperature splitless mode at 300 °C. Separation was achieved on a Rxi-5HT fused silica column (Restek; 30 m, 0.25 mm inner diameter, 0.25 µm film) with a He flow rate of 1.5 mL/min using the following temperature program: 2 min hold at 40 °C; 25 °C/min to 120 °C; 6 °C/min to 320 °C; 30 min hold at 320 °C. The ISQ LT was operated in electron ionization mode with a 230 °C source temperature, scanning a mass range of 43–800 Da with a 0.2-s dwell time. Fatty acid methyl esters (formed from fatty acids during extraction in methanol-containing Bligh–Dyer solution) were quantified by comparing analyte peak areas to the area of the internal standard pregnane diacetate, assuming a 1:1 response factor. Finally, the calculated amount of fatty acid methyl esters in each sample was normalized to sample weight.

Sandwich Microarray Immunoassays with LDChip. Sandwich-type microarray immunoassays (SMIs) were performed as described previously (11). Briefly, printed microscope slides with LDChip300 antibody microarray were blocked with 0.5 M Tris-HCl in 5% BSA for 5 min and then in 0.5 M Tris-HCl with 2% BSA for 30 min. After washing with TBSTRR buffer (0.4 M Tris-HCl pH 8, 0.3 M NaCl, 0.1% Tween 20) and drying the chip by quick centrifugation, the slides were mounted on a portable multiarray analysis module (MAAM) cassette for nine samples. Approximately 0.5 g of ground core samples were resuspended in 2 mL of TBSTRR and sonicated [3 × 1 min cycle with a handheld Ultrasonic Processor, UP50H (Hielscher Ultrasonics)]. Coarse material 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 TBSTRR, 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 GenePix4100A 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 the field in 4% (vol/vol, final concentration) formaldehyde–1× PBS (137 mM NaCl; 2.7 mM KCl; 10 mM Na₂HPO₄; 1.8 mM KH₂PO₄) at 4 °C for 2 h. The samples were washed twice with 1× PBS and stored at –20 °C in 1:1 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 4', 6'-diamidino-2-phenylindole (DAPI). The oligonucleotide probe used in this study for targeting rRNA genes was CYA361 (5'-CCCATTGGCGAAATTC-3') (44) at 35% (vol/vol) formamide concentration. HRP-labeled probes

were synthesized by [Biomers.net](#) GmbH. Negative controls were performed with the control probe NON338 (5'-ACTCTACGGGAGGCAGC-3') (45). Additional controls were carried out by subjecting the sample to the whole hybridization process without adding the probe, to evaluate whether the complex fluorophore-tyramide could interact with some mineral giving rise to a false positive. No signal was obtained from these controls. Further controls were conducted to evaluate whether the CARD-FISH signal could have been overlapping with autofluorescence from the cyanobacterial cells. CYA361 probe with both Alexa 488-labeled tyramide and Alexa 594-labeled tyramide were tested separately, and fluorescence was measured at 488 nm (green), 594 nm (red), and 633 nm (far red). In none of these controls could we achieve any autofluorescence signal from the cyanobacterial cells. Samples were mounted onto ibidi μ -slides 8 Well (ibidi GmbH) embedded in Citifluor:Vectashield (4:1) and examined with a Nikon AiR+ Resonant Scanning Confocal System (Nikon) at the Confocal Microscopy service at the Centro de Biología Molecular Severo Ochoa, CSIC-UAM.

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 μ L of MoBio Ultra-Clean Soil DNA solution S1, the tubes were subjected to two FastPrep (MP Biomedicals) cycles (30 s each, power setting of 5.5 m/s) 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 Gene Linker UV Chamber (Bio-Rad Laboratories) or a Stratallinker 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 (Genetrix, 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 chamber and subjected to a total UV dose of 0.04 mJ/cm². The resulting MDA 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 16SF (5'-AGAGTTTGATCTGGCTCAG-3') and 16SR (5'-CACGAGCTGACGACAGCCG-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 μ L 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 m/s). Samples were then incubated at 80 °C in a water bath during 40 min and centrifuged at 2,500 \times 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 [SI Appendix, Table S1](#). DNA was additionally isolated from a drilling water sample, to trace potential contamination events occurring during the core retrieval process. Two 250-mL drill cooling water samples, collected with a 15-d difference, were pooled and filtered through a 0.22- μ m 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 Tube to account also for laboratory contaminations during extraction of the nucleic acids (DNA isolation control). DNA was finally eluted in 50 μ L of 10 mM Tris-HCl buffer and quantified using a NanoDrop 2000 spectrophotometer (Thermo Scientific). The V5-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-generated amplicon libraries were sent for 250-nt 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 μ g of amplified DNA were mixed with 1 \times fragmentase reaction buffer in a final volume of 18 μ L, vortexed thoroughly, and incubated on ice for 5 min. The fragmentation reaction was then started by mixing the samples with 2 μ L of NEBNext dsDNA fragmentase (New England Biolabs, Inc.) and carried out for 25 min at 37 °C. After incubation, the fragmentation was halted by the addition of 5 μ L 0.5 M EDTA. The ensuing DNA was purified with the QIAquick PCR Purification Kit (Qiagen) and eluted in a final volume of 35 μ L 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 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 preprocessed with *mothur* (51, 52) to remove chimeras and sequences with poor alignments to the SILVA reference database (53, 54). The preprocessed reads were analyzed with CFF (55) to achieve a subspecies resolution. The chimera-scanning step included in CFF's standard pipeline was omitted, as chimeras had already been removed with *mothur*'s distribution of the VSEARCH software (56). The resulting subspecies ESVs were classified with SINA (57), and also by homology search against the nt and 16S ribosomal RNA NCBI databases and the SILVA nr128 database distributed with *mothur*.

Correlations between the abundances of different bacterial taxa and hydrogen concentrations were tested by multiple linear regression. For each taxonomic level from phylum to family, multiple linear models were fitted to explain hydrogen concentrations as a function of the abundances of all of the possible combinations of one to six taxa, considering only the taxa whose global abundance was over 5%. Statistical analyses were performed in R (v. 3.4.4).

Metagenomic Analysis. Paired raw metagenomic reads were quality filtered with PRINSEQ (58) with the following parameters: -derep 14 -lc_method dust -lc_threshold 7 -ns_max_p 2 -trim_ns_left 2 -trim_ns_right 2 -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 5 -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 most abundant groups (Cyanobacteria, Ascomycota, Alphaproteobacteria, 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 SqueezeM pipeline (ref. 61; <https://github.com/jtamames/squeezeM>, commit 242644d1a1f6380b6df92b1c70924924cd1070c8) using the sequential mode. Briefly, the reads from each sample were assembled using *metaSPAdes* (62), and ORFs and rRNA coding sequences were predicted with *Prodigal* v2.6.2 (63) and *barrnap* v0.9-dev (<https://github.com/tseemann/barrnap>), respectively. ORFs were aligned against the Clusters of Orthologous Groups (64) and Kyoto Encyclopedia of Genes and Genomes (65) databases 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 [SI Appendix, Table S2](#).

ACKNOWLEDGMENTS. We thank Javier Tamames, Carlos Pedrós-Alió, Ramón Roselló-Mora, Barbara Sherwood-Lollar, and Chris McKay for useful 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

Museo Minero de Riotinto. This work was supported by the European Research Council Advanced Grant ERC250350-IPBSL, and the Spanish Ministerio de Economía y Competitividad/Fondo Europeo de Desarrollo Regional

Grants AYA2011-24803 and ESP2015-69540-R. F.P.-S. is currently funded by the Spanish Ministerio de Economía y Competitividad/Fondo Europeo de Desarrollo Regional Grant CTM2016-80095-C2-1-R.

- Whitman WB, Coleman DC, Wiebe WJ (1998) Prokaryotes: The unseen majority. *Proc Natl Acad Sci USA* 95:6578–6583.
- Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S (2012) Global distribution of microbial abundance and biomass in seafloor sediment. *Proc Natl Acad Sci USA* 109:16213–16216.
- McMahon S, Parnell J (2014) Weighing the deep continental biosphere. *FEMS Microbiol Ecol* 87:113–120.
- Hoehler TM, Jorgensen BB (2013) Microbial life under extreme energy limitation. *Nat Rev Microbiol* 11:83–94.
- Colman DR, Poudel S, Stamps BW, Boyd ES, Spear JR (2017) The deep, hot biosphere: Twenty-five years of retrogression. *Proc Natl Acad Sci USA* 114:6895–6903.
- Edwards KJ, Becker K, Colwell F (2012) The deep, dark energy biosphere: Intra-terrestrial life on earth. *Annu Rev Earth Planet Sci* 40:551–568.
- Wilkins MJ, et al. (2014) Trends and future challenges in sampling the deep terrestrial biosphere. *Front Microbiol* 5:481.
- Tornos F (2006) Environment of formation and styles of volcanogenic massive sulfides: The Iberian Pyrite Belt. *Ore Geol Rev* 28:259–307.
- Puente-Sánchez F, et al. (2014) Deep subsurface sulfate reduction and methanogenesis in the Iberian Pyrite Belt revealed through geochemistry and molecular biomarkers. *Geobiology* 12:34–47.
- Gómez-Ortiz D, et al. (2014) Identification of the subsurface sulfide bodies responsible for acidity in Rio Tinto source water, Spain. *Earth Planet Sci Lett* 391:36–41.
- Parro V, et al. (2011) A microbial oasis in the hypersaline Atacama subsurface discovered by a life detector chip: Implications for the search for life on Mars. *Astrobiology* 11:969–996.
- Escudero C, Vera M, Oggerin M, Amils R (2018) Active microbial biofilms in deep poor porous continental subsurface rocks. *Sci Rep* 8:1538.
- Schippers A, et al. (2005) Prokaryotic cells of the deep sub-seafloor biosphere identified as living bacteria. *Nature* 433:861–864.
- Lloyd KG, May MK, Kevorkian R, Steen AD (2013) Meta-analysis of quantification methods shows that archaea and bacteria have similar abundances in the seafloor. *Appl Environ Microbiol* 79:7790–7799.
- Karnahl U, Wasternack C (1992) Half-life of cytoplasmic rRNA and tRNA, of plastid rRNA and of uridine nucleotides in heterotrophically and photoorganotrophically grown cells of *Euglena gracilis* and its apoplast mutant W3BUL. *Int J Biochem* 24:493–497.
- Davis BD, Luger SM, Tai PC (1986) Role of ribosome degradation in the death of starved *Escherichia coli* cells. *J Bacteriol* 166:439–445.
- Cohn CA, et al. (2006) Pyrite-induced hydroxyl radical formation and its effect on nucleic acids. *Geochem Trans* 7:3.
- Franchi M, Gallori E (2005) A surface-mediated origin of the RNA world: Biogenic activities of clay-adsorbed RNA molecules. *Gene* 346:205–214.
- Harel Y, Ohad I, Kaplan A (2004) Activation of photosynthesis and resistance to photoinhibition in cyanobacteria within biological desert crust. *Plant Physiol* 136:3070–3079.
- Whitton BA (2012) Introduction to the cyanobacteria. *Ecology of Cyanobacteria II*, eds Whitton BA, Potts M (Springer, Dordrecht, The Netherlands), pp 1–11.
- Stal LJ (2012) Cyanobacterial mats and stromatolites. *Ecology of Cyanobacteria II*, eds Whitton BA, Potts M (Springer, Dordrecht, The Netherlands), pp 65–125.
- Philippot L, et al. (2010) The ecological coherence of high bacterial taxonomic ranks. *Nat Rev Microbiol* 8:523–529.
- Kormas KA, Smith DC, Edgcomb V, Teske A (2003) Molecular analysis of deep subsurface microbial communities in Nankai Trough sediments (ODP Leg 190, Site 1176). *FEMS Microbiol Ecol* 45:115–125.
- Rastogi G, et al. (2010) Microbial and mineralogical characterizations of soils collected from the deep biosphere of the former Homestake gold mine, South Dakota. *Microb Ecol* 60:539–550.
- Hubalek V, et al. (2016) Connectivity to the surface determines diversity patterns in subsurface aquifers of the Fennoscandian shield. *ISME J* 10:2447–2458.
- Bagnoud A, et al. (2016) Reconstructing a hydrogen-driven microbial metabolic network in Opalinus Clay rock. *Nat Commun* 7:12770.
- Mulkidjanian AY, et al. (2006) The cyanobacterial genome core and the origin of photosynthesis. *Proc Natl Acad Sci USA* 103:13126–13131.
- Tamagnini P, et al. (2002) Hydrogenases and hydrogen metabolism of cyanobacteria. *Microbiol Mol Biol Rev* 66:1–20.
- Zhang X, Sherman DM, Sherman LA (2014) The uptake hydrogenase in the unicellular diazotrophic cyanobacterium *Cyanothece* sp. strain PCC 7822 protects nitrogenase from oxygen toxicity. *J Bacteriol* 196:840–849.
- Dutta I, Vermaas WF (2016) The electron transfer pathway upon H₂ oxidation by the NiFe bidirectional hydrogenase of *Synechocystis* sp. PCC 6803 in the light shares components with the photosynthetic electron transfer chain in thylakoid membranes. *Int J Hydrogen Energy* 41:11949–11959.
- Schwarz C, Poss Z, Hoffmann D, Appel J (2010) Hydrogenases and hydrogen metabolism in photosynthetic prokaryotes. *Recent Advances in Phototrophic Prokaryotes*, ed Hallenbeck PC (Springer, New York), pp 305–348.
- De Rosa E, et al. (2015) [NiFe]-hydrogenase is essential for cyanobacterium *Synechocystis* sp. PCC 6803 aerobic growth in the dark. *Sci Rep* 5:12424.
- Pisciotta JM, Zou Y, Baskakov IV (2011) Role of the photosynthetic electron transfer chain in electrogenic activity of cyanobacteria. *Appl Microbiol Biotechnol* 91:377–385.
- Büsch A, Friedrich B, Cramm R (2002) Characterization of the norB gene, encoding nitric oxide reductase, in the nondenitrifying cyanobacterium *Synechocystis* sp. strain PCC6803. *Appl Environ Microbiol* 68:668–672.
- 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.
- Weber B, et al. (2015) Biological soil crusts accelerate the nitrogen cycle through large NO and HONO emissions in drylands. *Proc Natl Acad Sci USA* 112:15384–15389.
- Starnawski P, et al. (2017) Microbial community assembly and evolution in subsurface sediment. *Proc Natl Acad Sci USA* 114:2940–2945.
- Rajeev L, et al. (2013) Dynamic cyanobacterial response to hydration and dehydration in a desert biological soil crust. *ISME J* 7:2178–2191.
- Giordano M, Mobilili F, Pezzoni V, Hein MK, Davis JS (2000) Photosynthesis in the caves of Frasassi (Italy). *Phycologia* 39:384–389.
- Raanan H, et al. (2016) Towards clarifying what distinguishes cyanobacteria able to resurrect after desiccation from those that cannot: The photosynthetic aspect. *Biochim Biophys Acta* 1857:715–722.
- Fernández-Remolar DC, et al. (2008) Underground habitats in the Rio Tinto basin: A model for subsurface life habitats on Mars. *Astrobiology* 8:1023–1047.
- Talbot HM, Rohmer M, Farimond P (2007) Rapid structural elucidation of composite bacterial hopanoids by atmospheric pressure chemical ionisation liquid chromatography/ion trap mass spectrometry. *Rapid Commun Mass Spectrom* 21:880–892.
- Pernthaler A, Pernthaler J, Amann R (2004) Sensitive multi-color fluorescence in situ hybridization for the identification of environmental microorganisms. *Molecular Microbial Ecology Manual*, eds Kowalchuk GA, de Bruijn FJ, Head IM, Akkermans AD, van Elsas JD (Kluwer, Dordrecht, The Netherlands), pp 711–726.
- Schönhuber W, et al. (1999) In situ identification of cyanobacteria with horseradish peroxidase-labeled, rRNA-targeted oligonucleotide probes. *Appl Environ Microbiol* 65:1259–1267.
- Wallner G, Amann R, Beisker W (1993) Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* 14:136–143.
- Direito SO, Marees 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.
- Woyke T, et al. (2011) Decontamination of MDA reagents for single cell whole genome amplification. *PLoS One* 6:e26161.
- Bohorquez LC, et al. (2012) In-depth characterization via complementing culture-independent approaches of the microbial community in an acidic hot spring of the Colombian Andes. *Microb Ecol* 63:103–115, and erratum (2012) 63:238.
- Camarinha-Silva A, et al. (2014) Comparing the anterior rare bacterial community of two discrete human populations using Illumina amplicon sequencing. *Environ Microbiol* 16:2939–2952.
- Puente-Sánchez F, Aguirre J, Parro V (2016) A novel conceptual approach to read-filtering in high-throughput amplicon sequencing studies. *Nucleic Acids Res* 44:e40.
- 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.
- Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120.
- Quast C, et al. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596.
- 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* 35:7188–7196.
- Tikhonov M, Leach RW, Wingreen NS (2015) Interpreting 16S metagenomic data without clustering to achieve sub-OTU resolution. *ISME J* 9:68–80.
- Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: A versatile open source tool for metagenomics. *PeerJ* 4:e2584.
- Pruesse E, Peplies J, Glöckner FO (2012) SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829.
- Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864.
- Brown CT, Howe A, Zhang Q, Pyrkosz AB, Brom TH (2012). A reference-free algorithm for computational normalization of shotgun sequencing data. arXiv:1203.4802. Preprint, posted May 21, 2012.
- 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.
- Tamames J, Puente-Sánchez F (2018) SqueezeM, a fully automatic metagenomic analysis pipeline from reads to bins. bioRxiv:10.1101/347559. Preprint, posted June 14, 2018.
- Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: A new versatile metagenomic assembler. *Genome Res* 27:824–834.
- Hyatt D, et al. (2010) Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119.
- Tatusov RL, et al. (2003) The COG database: An updated version includes eukaryotes. *BMC Bioinformatics* 4:41.
- Kanehisa M (2002) The KEGG database. *Novartis Found Symp* 247:91–101, discussion 101–103, 119–128, 244–252.
- Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60.