Statistical Analysis of Community RNA Transcripts between Organic Carbon and Geogas-Fed Continental Deep Biosphere Groundwaters

Margarita Lopez-Fernandez,[a*] Elias Broman,[a*] Domenico Simone,[a,b] Stefan Bertilsson,[c,d] Mark Dopson[a]

[a]Centre for Ecology and Evolution in Microbial Model Systems (EEMiS), Linnaeus University, Kalmar, Sweden
[b]SLU Bioinformatics Infrastructure, Swedish University of Agricultural Sciences, Uppsala, Sweden
[c]Department of Ecology and Genetics, Limnology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
[d]Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden

ABSTRACT Life in water-filled bedrock fissures in the continental deep biosphere is broadly constrained by energy and nutrient availability. Although these communities are alive, robust studies comparing active populations and metabolic processes across deep aquifers are lacking. This study analyzed three oligotrophic Fennoscandian Shield groundwaters, two "modern marine" waters that are replenished with organic carbon from the Baltic Sea and are likely less than 20 years old (171.3 and 415.4 m below sea level) and an extremely oligotrophic "thoroughly mixed" water (448.8 m below sea level) of unknown age that is composed of very old saline and marine waters. Cells were captured either using a sampling device that rapidly fixed RNA under *in situ* conditions or by filtering flowing groundwater over an extended period before fixation. Comparison of metatranscriptomes between the methods showed statistically similar transcript profiles for the respective water types, and they were analyzed as biological replicates. Study of the small subunit (SSU) rRNA confirmed active populations from all three domains of life, with many potentially novel unclassified populations present. Statistically supported differences between communities included heterotrophic sulfate-reducing bacteria in the modern marine water at 171.3 m below sea level that has a higher organic carbon content than do largely autotrophic populations in the H_2- and CO_2-fed thoroughly mixed water. While this modern marine water had signatures of methanogenesis, syntrophic populations were predominantly in the thoroughly mixed water. The study provides a first statistical evaluation of differences in the active microbial communities in groundwaters differentially fed by organic carbon or "geogases."

IMPORTANCE Despite being separated from the photosynthesis-driven surface by both distance and time, the deep biosphere is an important driver for the earth's carbon and energy cycles. However, due to the difficulties in gaining access and low cell numbers, robust statistical omics studies have not been carried out, and this limits the conclusions that can be drawn. This study benchmarks the use of two separate sampling systems and demonstrates that they provide statistically similar RNA transcript profiles, importantly validating several previously published studies. The generated data are analyzed to identify statistically valid differences in active microbial community members and metabolic processes. The results highlight contrasting taxa and growth strategies in the modern marine waters that are influenced by recent infiltration of Baltic Sea water versus the hydrogen- and carbon dioxide-fed, extremely oligotrophic, thoroughly mixed water.

Address correspondence to Margarita Lopez-Fernandez, margarita.lopezfernandez@lnu.se.

* Present address: Margarita Lopez-Fernandez, Institute of Resource Ecology, Helmholtz-Zentrum Dresden-Rossendorf, Dresden, Germany; Elias Broman, Department of Ecology, Environment and Plant Sciences, Stockholm University, Stockholm, Sweden.

The deep biosphere is the largest biome on earth, where the continental subsurface alone hosts up to 6×10^{29} cells from all three domains (1). Deep life has been demonstrated as active by, e.g., "viable/dead" PCR amplification (2), "omics" (3–5), and video evidence (6). A previous study at the Swedish Nuclear Fuel and Waste Management Company (SKB)-operated Äspö Hard Rock Laboratory (Äspö HRL) used a specially designed sampling device to fix cells under *in situ* conditions to ensure that RNA transcripts were unaffected by sampling procedures (3). In contrast, other studies used cell capture from flowing groundwater on filters over several days prior to fixation (see, e.g., reference 4). However, it is unknown if extended capture times alter the RNA transcript profile.

The extreme oligotrophy in the continental deep biosphere can limit cell numbers to 10^1 to 10^7 cells/ml (1), while Äspö HRL groundwaters contain 10^5 to 10^6 cells/ml (7). Due to the difficulty of obtaining deep biosphere samples and the large water volume needed to extract sufficient RNA for sequencing, no omics studies have provided sufficient replicates for valid statistics.

In this study, we combined RNA transcript data from the sampling device (3) and from cells captured over several days on filter holders to evaluate if the two methods are comparable (see File S1 in the supplemental material). Additionally, we statistically analyzed gene transcript counts pertaining to active microbial taxa and their metabolic processes between groundwaters of various ages and origins.

The studied groundwaters were two modern marine waters (MM-171.3 and MM-415.2) that are replenished from the Baltic Sea and have a residence time of <20 years and a "thoroughly mixed" water (TM-448.4) that is composed of different waters of multiple origins and unknown age (3, 7, 8). Cells were captured, and community RNA was extracted and sequenced according to File S1. The small subunit (SSU) rRNA sequences (File S2) were annotated against the SILVA database and normalized as relative abundances (File S3). The MM-415.2 filter holder metatranscriptomes only had two replicates and thus cannot be statistically compared to the others. However, this groundwater was clearly different in both its SSU and protein-coding RNA (pcRNA) transcripts (Fig. 1) and is discussed in File S4. Nonmetric multidimensional scaling (NMDS) of SSU rRNA transcript beta diversity suggested that the three water samples were statistically different in their microbial communities (permutational multivariate analysis of variance [PERMANOVA] 9,999 permutations, $P = 0.0011$; Fig. 1). Previous analysis of the sampling device (SD) TM-448.4-4 sample showed it was different from the SD TM-448.4-3 sample, as it had likely been recently exposed to an electron donor (3). Repetition of the NMDS without this outlier altered the significance between the three groundwaters ($P = 0.004$). Without TM-448.4-4, the grouping supports the notion that (i) the two methods give highly similar RNA transcript patterns and, therefore, sampling with filter holders over several days is valid, and (ii) in the absence of periodic availability of an electron donor (as for the SD TM-448.4-4 sample [3]), the deep biosphere communities were stable for a minimum of 2 years.

SSU rRNA-based phylogeny from all analyzed metatranscriptomes showed that a broad range of phyla from all three domains of life were viable and had protein-synthesizing potential (3) (Fig. 1). It also reinforced that the deep biosphere contains a large relative proportion of active candidate phyla from all three domains (e.g., *Patescibacteria*) along with many unclassified sequences. Statistically valid differences between the MM-171.3 and TM-448.4 groundwaters included sulfate-reducing bacteria (SRB) with *Desulfobulbaceae* in the MM-171.3 groundwater compared to *Desulfobacteraceae* and *Desulfurivibrio* in the TM-448.4 groundwater (File S5). This confirms that sulfur compound reduction is prevalent (see, e.g., references 9 and 10) with the predominantly organoheterotrophic SRB *Desulfobulbaceae* (11) in the MM-171.3 groundwater compared to autotrophic *Desulfurivibrio* spp. (12) in the ultraoligotrophic TM-448.4 water. In addition, increased 16S rRNA gene transcripts in the TM-448.4 groundwater that aligned within the *Syntrophus* genus demonstrated that syntrophy is likely to be an important survival strategy in these oligotrophic groundwaters (13).

Analysis of pcRNA transcripts identified 973 unique prokaryote genes (File S6). The NMDS analysis also showed that the community-level transcription profiles were
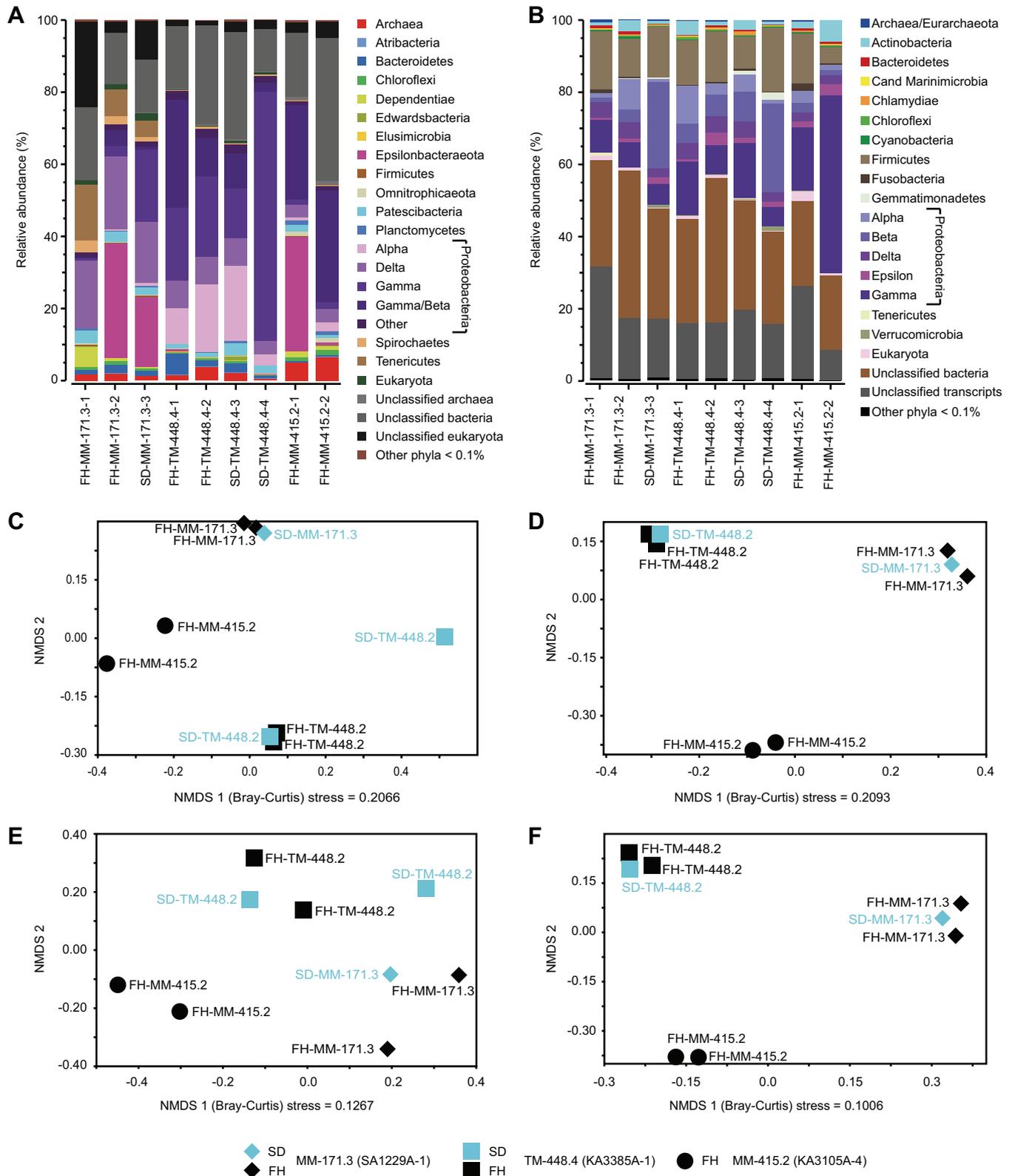
FIG 1 (A and B) Taxonomic annotation of the SSU rRNA (A) and protein-coding RNA (B) sequences showing stacked bars of the taxonomic phyla and *Proteobacteria* classes (*Betaproteobacteriales* shown separately) with a relative abundance of >0.1% for the modern marine (MM-171.3 and MM-415.2) and thoroughly mixed (TM-448.4) groundwaters. Rare taxa with a relative abundance of <0.1% are given as "other phyla." (C and D) NMDS Bray-Curtis dissimilarity (beta diversity) plots based on the SSU rRNA at the lowest taxonomic level that could be assigned to the SILVA database using the Ribosomal Database Project classifier (C) and a second NMDS without the SD TM-448.3-4 (KA3385A-1) outlier (D). (E and F) NMDS Bray-Curtis plots based on the annotated transcripts (i.e., UniProtKB identifiers) with an E value of <10 and TPM of >100 from the full data set (E) and without the SD TM-448.4-4 outlier (F). The sampling methods are filter holders (FH) and sampling device (SD). Cand, *Candidatus*.
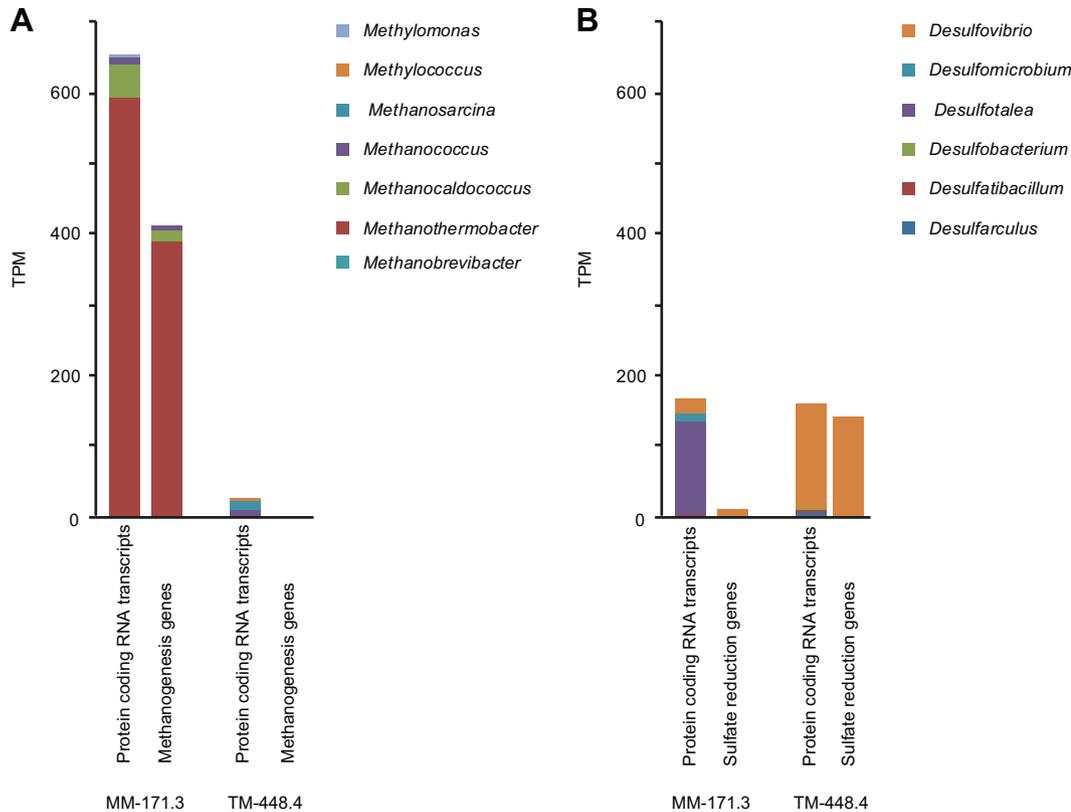
FIG 2 (A and B) Average of the significantly different (false-discovery rate [FDR] < 0.05; E value < 0.001) transcripts per million sequences (TPM) for the modern marine (MM-171.3) and thoroughly mixed (TM-448.4) groundwaters for protein-coding RNA transcripts assigned to methanogenic taxa (left) and genes attributed to methanogenesis from CO_2 (*fwdC*, *mtrACDEH*, *mcrABCG*; right) (A), as well as protein-coding RNA transcripts assigned to sulfate-reducing bacteria taxa (left) and genes attributed to sulfate reduction (cytochrome c_3, *rd2*, *aprA*, *dsrA*, and *dsvAB*; right) (B).

statistically different (P = 0.002; Fig. 1), and further removal of the SD TM-448.4-4 outlier gave a P value of 0.004. Altogether, 410 prokaryotic genes had significant differential expression between the MM-171.3 and TM-448.4 groundwaters (false-discovery rate [FDR] < 0.05; E value < 0.001). Transcripts encoding tricarboxylic acid (TCA) cycle (*mdh*, *fumC*, and *sucC*) and ATP synthase (*atpAG*) proteins had higher transcript counts in the MM-171.3 groundwater, while increased TM-448.4 transcripts encoded, e.g., ribosomal (e.g., *rpmB*, *rpsBK,* and *rplC*) and stress/repair (e.g., *dfx*, *recGN*, *cspAB*, *clpPX*, *dnaK*, and *hspC4*) proteins. Additionally, a qualitative comparison of the SD TM-448.4-4 outlier (3) with the other three replicates suggested that this outlier had more transcripts involved with, e.g., replication and metabolic processes. Overall, most overexpressed transcripts were seen in the MM-171.3 groundwater, robustly demonstrating that this community was actively growing while the TM-448.4 populations were in "metabolic standby" (3).

The metabolic process with the greatest number of statistically different MM-171.3 groundwater transcripts was methanogenesis from CO_2 (*fwdC*, *mtrACDEH*, and *mcrABCG* genes) attributed to *Methanothermobacter* spp. (14) within the *Euryarchaeota* (Fig. 2 and File S7 and S8). Sulfur oxidation coupled to nitrate reduction was also important with increased pcRNA transcripts attributed to *Sulfurimonas denitrificans* (15) and *Thiobacillus denitrificans* (16) in the MM-171.3 and TM-448.4 groundwaters, respectively (File S9). This difference was potentially because *S. denitrificans* can use, e.g., formate, while *T. denitrificans* is an obligate sulfur compound-oxidizing chemolithoautotroph that had statistically increased *cbbLS* transcripts encoding CO_2 fixation via the Calvin-Benson-Bassham cycle. Consistent with the SSU rRNA data, the pcRNA transcripts had significant differences in the SRB. These included transcripts from 37 genes attributed to the *Desulfobulbaceae* that were only present in the MM-171.3 water, while

dissimilatory sulfate-reducing genes *aprA*, *dsrA*, and *dsvAB* attributed to *Desulfovibrio* spp. were increased in the TM-448.4 groundwater (Fig. 2 and File S8). The importance of syntrophy was also further demonstrated by pcRNA transcripts in both the MM-171.3 and TM-448.4 waters attributed to *Syntrophus aciditrophicus* that grows alongside H_2 utilizers (17) predominantly present in the TM-448.4 groundwater (File S9). Finally, earlier observations of cyanobacteria in ancient deep terrestrial groundwaters (18, 19) were confirmed by increased *Synechocystis* pcRNA transcripts in the TM-448.4 water, also demonstrating their viability in these habitats.

This work presents for the first time a statistically robust omics study of deep subsurface crystalline rock groundwaters with different depths and geochemical characteristics. We conclude that cell capture over several days does not alter RNA transcript profiles compared to rapid *in situ* fixation in this extremely oligotrophic environment. Importantly, this analysis of the two methods validates published studies that have used capture times prior to RNA fixation over the several days needed to obtain sufficient biomass for biomolecule extraction from low-cell-density deep groundwaters. The similarity of the data obtained by the two methods was likely due to the long-term and stable oligotrophic conditions in the respective groundwaters. These novel findings also provide evidence on how the differences in active communities and metabolic processes are influenced by organic carbon versus geogas-fed modern marine and thoroughly mixed groundwaters, respectively. This benchmarking of deep biosphere metatranscriptome analyses paves the way for future and still-needed exploration of the living deep biosphere in a statistically sound way.

Data availability. The raw sequence data are available in the NCBI Sequence Read Archive BioProject numbers PRJNA400688 and PRJNA541524 for the sampling device and filter holders, respectively.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mBio .01470-19.

FILE S1, DOCX file, 0.1 MB.
FILE S2, XLSX file, 0.1 MB.
FILE S3, XLSX file, 0.2 MB.
FILE S4, DOCX file, 0.1 MB.
FILE S5, XLSX file, 0.1 MB.
FILE S6, XLSX file, 0.4 MB.
FILE S7, EPS file, 0.5 MB.
FILE S8, XLSX file, 0.1 MB.
FILE S9, EPS file, 1.9 MB.

REFERENCES

1. Magnabosco C, Lin LH, Dong H, Bomberg M, Ghiorse W, Stan-Lotter H, Pedersen K, Kieft TL, van Heerden E, Onstott TC. 2018. The biomass and biodiversity of the continental subsurface. Nat Geosci 11:707–717. https://doi.org/10.1038/s41561-018-0221-6.

2. Lopez-Fernandez M, Broman E, Wu X, Bertilsson S, Dopson M. 2018. Investigation of viable taxa in the deep terrestrial biosphere suggests high rates of nutrient recycling. FEMS Microbiol Ecol 94:fiy121. https://doi.org/10.1093/femsec/fiy121.

3. Lopez-Fernandez M, Simone D, Wu X, Soler L, Nilsson E, Holmfeldt K, Lantz H, Bertilsson S, Dopson M. 2018. Metatranscriptomes reveal all three domains of life are active, but are dominated by bacteria in the Fennoscandian crystalline granitic continental deep biosphere. mBio 9:e01792-18. https://doi.org/10.1128/mBio.01792-18.

4. Lau MCY, Kieft TL, Kuloyo O, Linage-Alvarez B, van Heerden E, Lindsay MR, Magnabosco C, Wang W, Wiggins JB, Guo L, Perlman DH, Kyin S, Shwe HH, Harris RL, Oh Y, Yi MJ, Purtschert R, Slater GF, Ono S, Wei S, Li L, Sherwood Lollar B, Onstott TC. 2016. An oligotrophic deep-subsurface community dependent on syntrophy is dominated by sulfur-driven autotrophic denitrifiers. Proc Natl Acad Sci U S A 113:E7927–E7936. https://doi.org/10.1073/pnas.1612244113.

5. Bagnoud A, Chourey K, Hettich RL, de Bruijn I, Andersson AF, Leupin OX, Schwyn B, Bernier-Latmani R. 2016. Reconstructing a hydrogen-driven microbial metabolic network in Opalinus clay rock. Nat Commun 7:12770. https://doi.org/10.1038/ncomms12770.

6. Borgonie G, Linage-Alvarez B, Ojo AO, Mundle SOC, Freese LB, Van Rooyen C, Kuloyo O, Albertyn J, Pohl C, Cason ED, Vermeulen J, Pienaar C, Litthauer D, Van Niekerk H, Van Eeden J, Lollar BS, Onstott TC, Van Heerden E. 2015. Eukaryotic opportunists dominate the deep-subsurface biosphere in South Africa. Nat Commun 6:8952. https://doi.org/10.1038/ncomms9952.

7. Lopez-Fernandez M, Åström M, Bertilsson S, Dopson M. 2018. Depth and dissolved organic carbon shape microbial communities in surface influenced but not ancient saline terrestrial aquifers. Front Microbiol 9:2880. https://doi.org/10.3389/fmicb.2018.02880.

8. Wu X, Holmfeldt K, Hubalek V, Lundin D, Åström M, Bertilsson S, Dopson M. 2016. Microbial metagenomes from three aquifers in the Fennoscandian shield terrestrial deep biosphere reveal metabolic partitioning among populations. ISME J 10:1192–1203. https://doi.org/10.1038/ismej.2015.185.

9. Itävaara M, Nyyssönen M, Kapanen A, Nousiainen A, Ahonen L, Kukkonen I. 2011. Characterization of bacterial diversity to a depth of 1500 m in the Outokumpu deep borehole, Fennoscandian Shield. FEMS Microbiol Ecol 77:295–309. https://doi.org/10.1111/j.1574-6941.2011.01111.x.

10. Purkamo L, Bomberg M, Nyyssönen M, Ahonen L, Kukkonen I, Itavaara M. 2017. Response of deep subsurface microbial community to different carbon sources and electron acceptors during similar to 2 months incubation in microcosms. Front Microbiol 8:232. https://doi.org/10.3389/fmicb.2017.00232.

11. Kuever J. 2014. The family Desulfobulbaceae, p 75–86. In Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), The prokaryotes: Deltaproteobacteria and Epsilonproteobacteria. Springer Berlin Heidelberg, Berlin, Germany.

12. Melton ED, Sorokin DY, Overmars L, Chertkov O, Clum A, Pillay M, Ivanova N, Shapiro N, Kyrpides NC, Woyke T, Lapidus AL, Muyzer G. 2016. Complete genome sequence of Desulfurivibrio alkaliphilus strain AHT2^T, a haloalkaliphilic sulfidogen from Egyptian hypersaline alkaline lakes. Stand Genomic Sci 11:67. https://doi.org/10.1186/s40793-016-0184-4.

13. Probst AJ, Ladd B, Jarett JK, Geller-McGrath DE, Sieber CMK, Emerson JB, Anantharaman K, Thomas BC, Malmstrom RR, Stieglmeier M, Klingl A, Woyke T, Ryan MC, Banfield JF. 2018. Differential depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the deep terrestrial subsurface. Nat Microbiol 3:328–336. https://doi.org/10.1038/s41564-017-0098-y.

14. Wasserfallen A, Nolling J, Pfister P, Reeve J, Conway de Macario E. 2000. Phylogenetic analysis of 18 thermophilic Methanobacterium isolates supports the proposals to create a new genus, Methanothermobacter gen. nov., and to reclassify several isolates in three species, Methanothermobacter thermautotrophicus comb. nov., Methanothermobacter wolfeii comb. nov., and Methanothermobacter marburgensis sp. nov. Int J Syst Evol Microbiol 50:43–53. https://doi.org/10.1099/00207713-50-1-43.

15. Han Y, Perner M. 2015. The globally widespread genus Sulfurimonas: versatile energy metabolisms and adaptations to redox clines. Front Microbiol 6:989. https://doi.org/10.3389/fmicb.2015.00989.

16. Beller HR, Chain PS, Letain TE, Chakicherla A, Larimer FW, Richardson PM, Coleman MA, Wood AP, Kelly DP. 2006. The genome sequence of the obligately chemolithoautotrophic, facultatively anaerobic bacterium Thiobacillus denitrificans. J Bacteriol 188:1473–1488. https://doi.org/10.1128/JB.188.4.1473-1488.2006.

17. Jackson BE, Bhupathiraju VK, Tanner RS, Woese CR, McInerney MJ. 1999. Syntrophus aciditrophicus sp. nov., a new anaerobic bacterium that degrades fatty acids and benzoate in syntrophic association with hydrogen-using microorganisms. Arch Microbiol 171:107–114. https://doi.org/10.1007/s002030050685.

18. Puente-Sánchez F, Arce-Rodríguez A, Oggerin M, García-Villadangos M, Moreno-Paz M, Blanco Y, Rodríguez N, Bird L, Lincoln SA, Tornos F, Prieto-Ballesteros O, Freeman KH, Pieper DH, Timmis KN, Amils R, Parro V. 2018. Viable cyanobacteria in the deep continental subsurface. Proc Natl Acad Sci U S A 115:10702–10707. https://doi.org/10.1073/pnas.1808176115.

19. Hubalek V, Wu X, Eiler A, Buck M, Heim C, Dopson M, Bertilsson S, Ionescu D. 2016. Connectivity to the surface determines diversity patterns in subsurface aquifers of the Fennoscandian shield. ISME J 10:2447–2458. https://doi.org/10.1038/ismej.2016.36.