

Recycling and metabolic flexibility dictate life in the lower oceanic crust

<https://doi.org/10.1038/s41586-020-2075-5>

Received: 16 August 2019

Accepted: 10 January 2020

Published online: 11 March 2020

 Check for updates

Jiangtao Li^{1,2,9}, Paraskevi Mara^{2,9}, Florence Schubotz^{3,4}, Jason B. Sylvan⁵, Gaëtan Burgaud⁶, Frieder Klein⁷, David Beaudoin², Shu Ying Wee⁵, Henry J. B. Dick², Sarah Lott², Rebecca Cox², Lara A. E. Meyer^{3,4}, Maxence Quémener⁶, Donna K. Blackman⁸ & Virginia P. Edgcomb^{2,9}✉

The lithified lower oceanic crust is one of Earth's last biological frontiers as it is difficult to access. It is challenging for microbiota that live in marine subsurface sediments or igneous basement to obtain sufficient carbon resources and energy to support growth^{1–3} or to meet basal power requirements⁴ during periods of resource scarcity. Here we show how limited and unpredictable sources of carbon and energy dictate survival strategies used by low-biomass microbial communities that live 10–750 m below the seafloor at Atlantis Bank, Indian Ocean, where Earth's lower crust is exposed at the seafloor. Assays of enzyme activities, lipid biomarkers, marker genes and microscopy indicate heterogeneously distributed and viable biomass with ultralow cell densities (fewer than 2,000 cells per cm³). Expression of genes involved in unexpected heterotrophic processes includes those with a role in the degradation of polyaromatic hydrocarbons, use of polyhydroxyalkanoates as carbon-storage molecules and recycling of amino acids to produce compounds that can participate in redox reactions and energy production. Our study provides insights into how microorganisms in the plutonic crust are able to survive within fractures or porous substrates by coupling sources of energy to organic and inorganic carbon resources that are probably delivered through the circulation of subsurface fluids or seawater.

Diverse and abundant microorganisms are confirmed to be present in the basaltic upper crust^{5–7}, yet limited information exists about the distinctly different gabbroic lower crust that accounts for two-thirds of the volume of the crust^{8,9}. Expedition 360 of the International Ocean Discovery Program (IODP) drilled the Atlantis Bank oceanic core complex on the southwest Indian Ridge, Indian Ocean, where the lower oceanic crust is exhumed by detachment faulting to around 700 m below sea level, providing convenient access to an otherwise largely inaccessible realm (Fig. 1 and Supplementary Table 1).

Our approach combined cell counts, microscopy and shipboard quantification of ATP levels to assess the abundance and distribution of microbial communities, exoenzyme activity assays, carbon, hydrogen, nitrogen and sulfur (CHNS) and thin-section analyses of host rock samples, marker gene (iTAG) analyses of prokaryotic diversity, analyses of cultures to assess the viability and activities of selected taxa (including those that might be rare and may be missed by molecular approaches), and examination of expressed genes and lipid biomarkers. Rigorous efforts were made at all stages to control for contamination, including from drilling fluids (Supplementary Table 2) during shipboard sampling, and from laboratory and molecular kits (further details are provided in the Methods and Supplementary Information).

Samples along the 809-m depth of hole U1473A showed evidence of carbonate and/or clay-altered felsic veins (variable 0.01–6 mm

openings) within predominantly olivine gabbro, oxide gabbro and gabbro (Fig. 2) with around 0.1–5% porosity and 0.2–4 wt% water content. Borehole logging indicates temperatures of 15–18 °C. Total organic nitrogen (0.002–0.003 wt%) and total organic carbon (0.004–0.018 wt%) contents were extremely low (Supplementary Table 1).

Detection of biomass and biomarkers

Low-density (131–1,660 cells per cm³), heterogeneously distributed microbial communities were detected (Supplementary Table 3b), in the same range as recent studies of young, cool and oxic ridge-flank subsurface basalts in the Atlantic Ocean¹⁰ and samples from 0–15 m below seafloor (mbsf) obtained from the 3.5-million-year-old Atlantis Massif¹¹. Detection of archaeal intact polar lipids (IPLs) with a dominance of archaeol-based lipids over glycerol dialkyl glycerol tetraethers (GDGTs) indicates the existence of indigenous archaeal communities that are distinct from those that are observed in typical deep-biosphere sedimentary settings¹² (Fig. 3b, Extended Data Fig. 1 and Supplementary Table 3a). Bacterial diether glycerol (DEG) lipids were detected at all depths, which require less cellular maintenance than fatty-acid-containing lipids (Fig. 3c) and may indicate an adaptation to the stresses of living in the low-energy plutonic rocks of the deep biosphere¹³.

¹State Key Laboratory of Marine Geology, Tongji University, Shanghai, China. ²Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, MA, USA. ³MARUM, University of Bremen, Bremen, Germany. ⁴Department of Geosciences, University of Bremen, Bremen, Germany. ⁵Department of Oceanography, Texas A&M University, College Station, TX, USA. ⁶Université de Brest, EA 3882, Laboratoire Universitaire de Biodiversité et Ecologie Microbienne, ESIAB, Technopôle Brest-Iroise, Plouzané, France. ⁷Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA. ⁸Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. ⁹These authors contributed equally: Jiangtao Li, Paraskevi Mara, Virginia P. Edgcomb. ✉e-mail: vedgcomb@whoi.edu

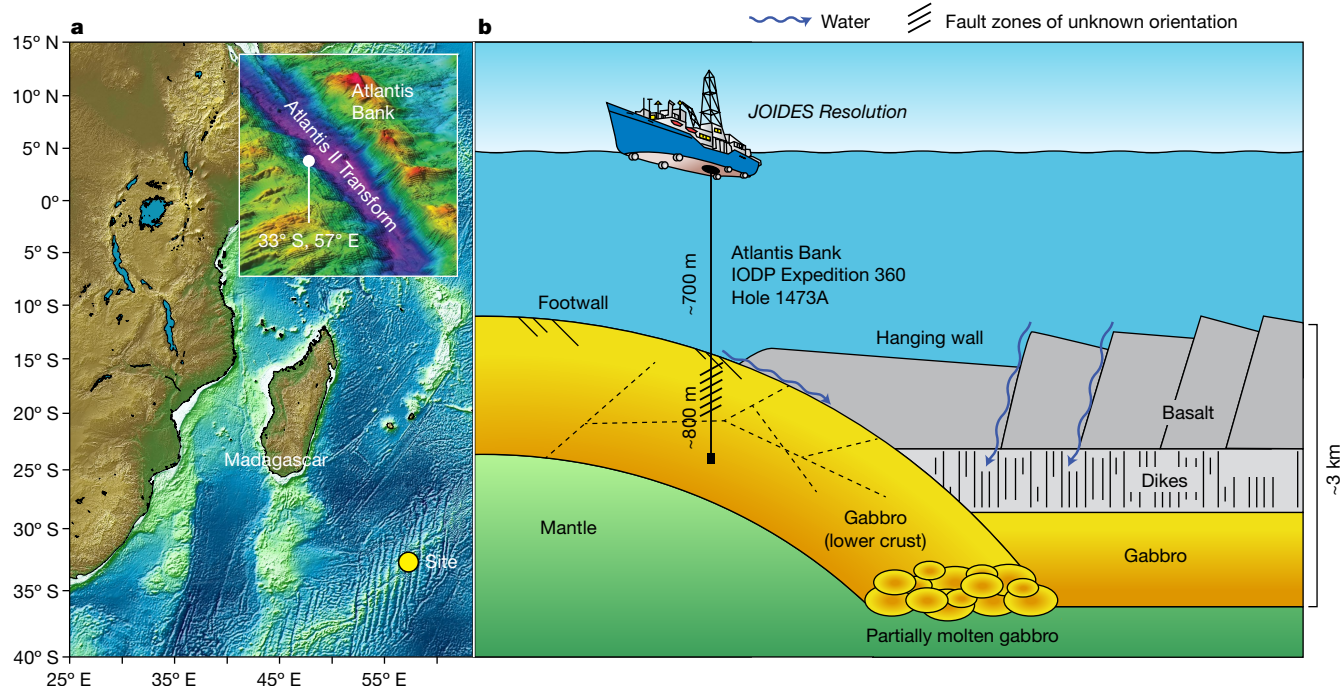


Fig. 1 | Study site of the exploration of the intrusive oceanic crust by IODP Expedition 360. a, Map of study site. **b**, Schematic of lower oceanic crust exhumed at the seafloor at Atlantis Bank. Dashed lines indicate the detected

presence of fractures along the depth of hole U1473A. The inset map is adapted from a previous study³¹; the background map was made using GeoMapApp (<http://www.geomapp.org/>).

The high IPL-to-core-lipid ratios support a likely in situ source of detected lipids. After cell death, the labile head groups of membrane lipids are enzymatically cleaved, leaving behind the core lipids that can be preserved in the rock record over millions of years¹⁴. Concentrations of core lipids were as low as for their IPL counterparts, ranging between 4 and 3,700 pg g⁻¹ of rock (Supplementary Table 3a). The resulting IPL-to-core-lipid ratios are high, on average more than 70% for glycosidic (G)-GDGTs/(G)-GDGTs + core GDGTs and 35% for glycosidic archaeols (G-ARs)/(G-ARs + core ARs), pointing to the minimal accumulation of fossil material. In comparison, IPLs typically comprise less than 10% of the sum of core lipids and IPLs in sedimentary settings¹².

Raman spectroscopy of a sample from 182 mbsf shows Raman bands that are consistent with amino acids and sugars. Scanning electron microscopy shows filamentous inclusions of organic compounds associated with iron and manganese oxides, surrounded by, and inter-grown with, calcite (CaCO₃), evidence of in situ origin (Fig. 3a). The viability of the cells is further supported by (1) the detection of methane production in long-term enrichments (average production from 9.62×10^{-4} – 5.66×10^{-2} nmol CH₄ g⁻¹ of rock per day) (Extended Data Fig. 2); (2) the cultivation of more than 100 unique fungal isolates from samples along the entire depth of the hole; (3) ATP concentrations above detection levels (1 pg cm⁻³) in 6 out of 11 rock samples (Fig. 2); and (4) detectable alkaline phosphatase exoenzyme activity rates of 0.04–2.3 pmol cm⁻³ h⁻¹ in all 5 samples for which measurements were attempted (Supplementary Table 3b).

Marker gene (iTAG) libraries reveal low diversity (Simpson scores (1–D) of 0.48–0.93) of putative heterotrophic and chemoautotrophic taxa, many of which have been described from deep-sea and polyextreme habitats (Extended Data Fig. 3). These include Nitrosopumilaceae, SAR202 Chloroflexi, Nitrospinaceae, SAR406 Marinimicrobia, SAR324 Deltaproteobacteria and Thioglobaceae, among other heterogeneously distributed taxa. Variation in the taxonomic composition among samples suggests that the availability of carbon and energy substrates is also heterogeneous. The SAR324 (marine group B) lineage of Deltaproteobacteria was detected in 8 out of 11 samples and is

known to have wide metabolic flexibility, including carbon fixation through Rubisco, and an ability to oxidize alkanes, methane and/or sulfur to generate energy^{15,16}—all potential metabolisms in the crustal basement. The detected DEG lipids could also be in part derived from these Deltaproteobacteria¹⁷. The putative sulfur and hydrogen oxidizers SAR406 Marinimicrobia were also detected in all samples¹⁸.

Non-metric multidimensional scaling (Extended Data Fig. 4) indicates that the measured ATP concentrations, fracturing intensity and graphic lithology all had the most-significant effect on observed depth variance in diversity for the 11 samples ($P \leq 0.05$). Locations of fracturing and alteration are where seawater from above, or low-temperature fluids from below, can most readily permeate the rock. Unfortunately, it was not possible to sample borehole fluids for important studies of in situ fluid chemistry during Expedition 360. Other studies of the seafloor lithosphere show that such fluids can carry the required substrates to support microbial life¹⁹.

mRNA reveals diverse strategies

Recovery of mRNA was lower for the shallowest samples from 11 and 168 mbsf, consistent with observations that the top of the drill hole was primarily dense unaltered gabbro with few fractures to support microbial populations (Supplementary Table 4; further details for all analyses are provided in the Supplementary Discussion). Detection of transcripts for various metabolic processes indicates that microorganisms have taken the prerequisite steps towards these processes. Additional measurements of activities, for example, through stable-isotope probing, in addition to those performed in this study are needed to demonstrate the occurrence of these processes. Expression of genes unique to peptidoglycan synthesis was detected in 5 out of 11 samples, indicating that cells are engaging in cell division and/or maintenance of cell integrity (Supplementary Table 4). Transcripts involved in biosynthesis of different amino acids were detected in 10 out of 11 samples as well as transcripts implicated in vitamin E and B12 synthesis (for example, *cbiC* and *cbiJ*) in 9 samples (Supplementary Table 4). The *cbiJ*

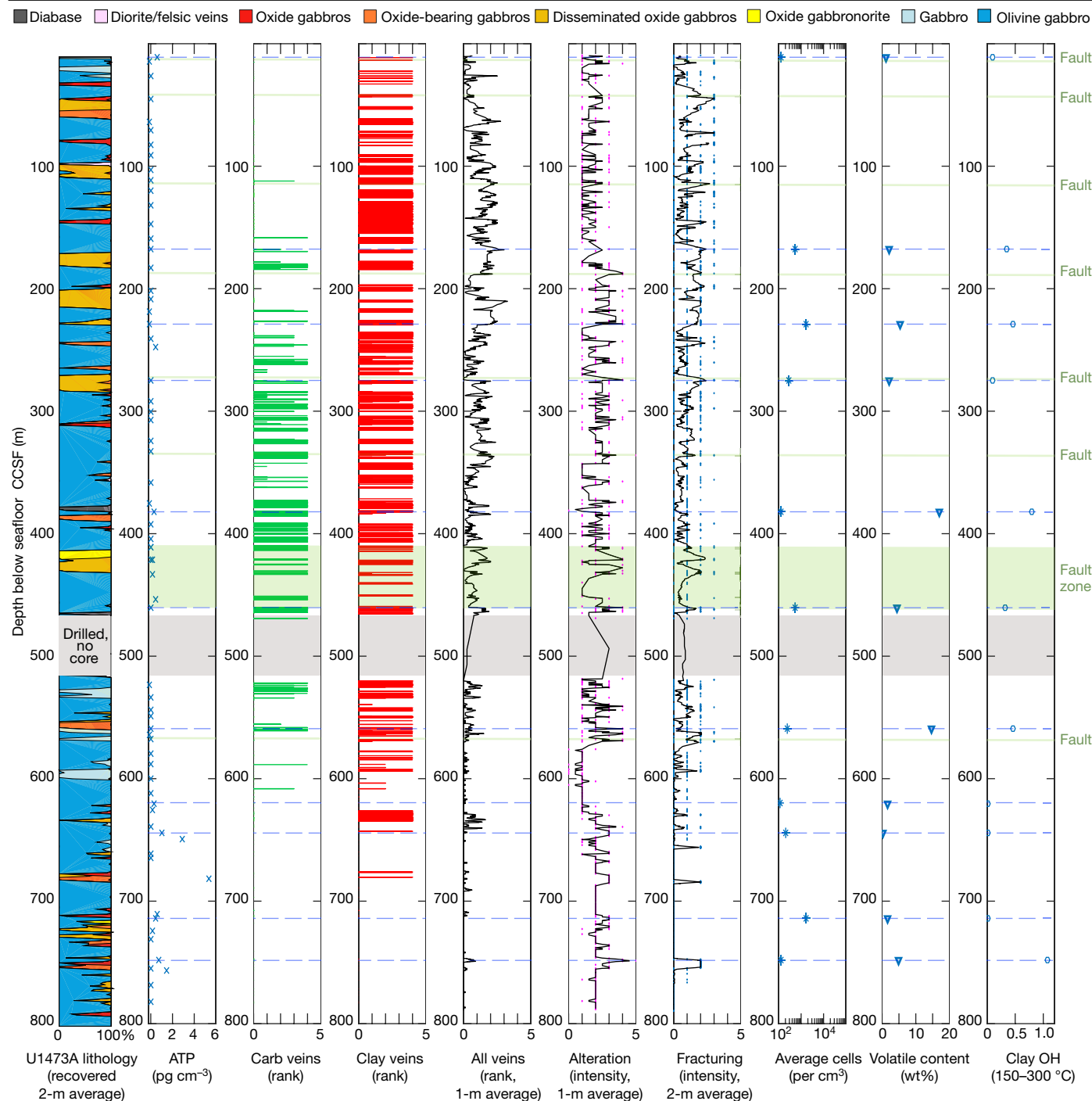


Fig. 2 | Summary of hole U1473A of IODP Expedition 360. Downhole properties of the gabbroic section determined by shipboard macroscopic descriptions and measured microbiological indicators. Labels indicate the property shown in each panel. Carb, carbonate. A key to lithology is included; rank and intensity range between 0 and 5. Faults are indicated by pale-green lines. Microbiological sample depths are highlighted by dashed blue lines. The

light-grey horizontal zone around 500 mbsf indicates a zone that was drilled but not cored. ATP was measured once per sample. CCSF, core composite depth below seafloor. Replicate cell-count data are provided in Supplementary Table 3b. Source data are available at [http://web.iodp.tamu.edu/DESCReport/\(file name, 360_U1473A_macroscopic.xlsx\)](http://web.iodp.tamu.edu/DESCReport/(file%20name,%20360_U1473A_macroscopic.xlsx)).

sequence was annotated to *Chroococcidiopsis* sp., a cyanobacterial taxon that may contribute to anaerobic production of cobalamin or pseudocobalamin in the lower crust using a hydrogen-based lithoautotrophic metabolism²⁰.

Although we observed the production of methane in long-term enrichments (Extended Data Fig. 2), methanogens appear to be rare in our samples, explaining the low recovery of transcripts involved in methane metabolism (Supplementary Table 4), and no ribosomal

RNA gene sequences of known methanogens were recovered (Extended Data Fig. 3). However, the abundance of SAR324 Deltaproteobacteria in our iTAG libraries indicates potential methanotrophy^{15,16} coupled with methane production in our long-term enrichments (Extended Data Fig. 2). Some methane metabolism is probably present, but its importance is unclear in these samples. Detection of glycosidic archaeol and acyclic GDGT (GDGT-0) may indicate the presence of methanogens; however, these lipids may also derive from other archaeal sources.

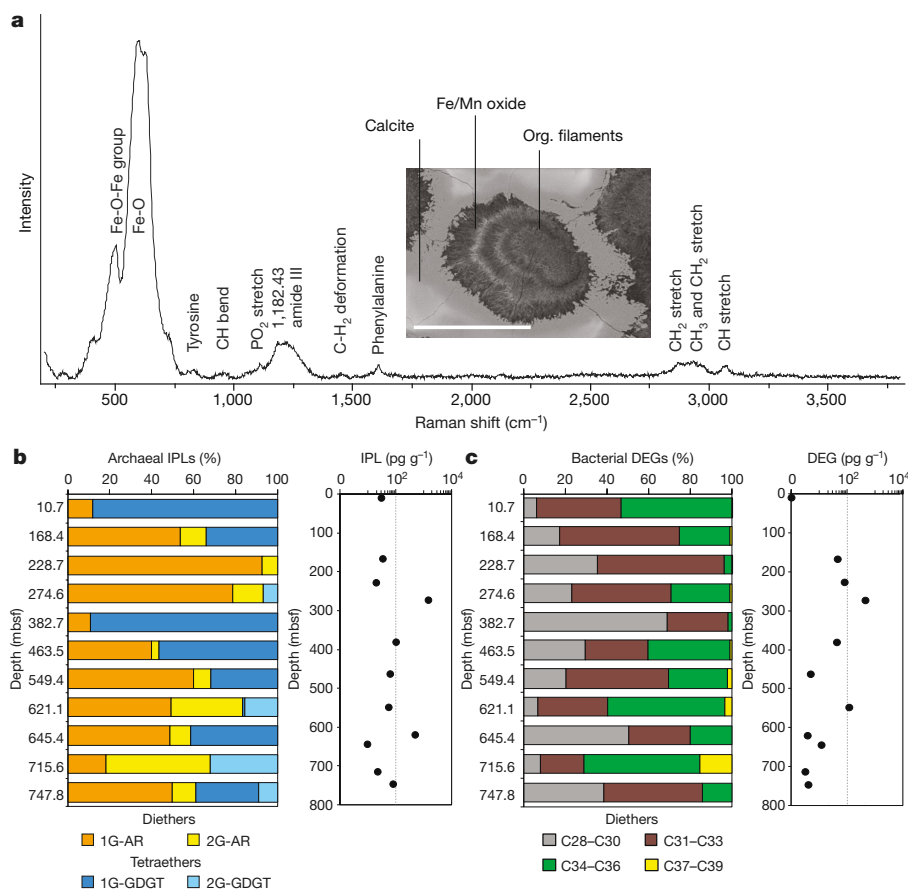


Fig. 3 | Biological signatures at hole U1473A revealed by Raman spectroscopy and membrane lipid analyses. **a**, Raman spectral features of organic inclusions in sample 21R2 from 182 mbsf. Inset, back-scattered electron image of an organic inclusion. Scale bar, 100 μm . Spectral measurements were performed at two spots on this particular feature, 10 independent times per spot, with similar results. **b**, Diversity (left) and concentration (right) of

archaeal IPLs within the 11 independent samples. **c**, Diversity (left) and concentration (right) of bacterial DEGs within the 11 independent samples. The summed carbon chain lengths from C28 to C39 are shown. 1G, monoglycosidic; 2G, diglycosidic; AR, archaeol. Lipid data are from single measurements owing to sample constraints. Source Data are available online.

Transcripts involved in nitrogen metabolism were detected in 9/11 samples, including genes associated with denitrification and dissimilatory nitrite reduction to ammonium, the latter of which was found in zones where fracturing and vein intensity were greatest (274.6, 460.4 and 619.6 mbsf). Transcripts involved in sulfur metabolism were detected in most samples (Supplementary Table 4).

Autotrophy might be expected in the gabbroic basement given the likelihood of only ephemeral sources of organic carbon via circulating fluids and the potential for seawater reactions with olivine-rich rocks that can produce molecular hydrogen (H_2), methane (CH_4) and other short-chain hydrocarbons that can be used as reducing agents for metabolic energy production¹⁹. Indeed, fluid inclusions present within gabbro from the Atlantis Bank represent a source of abiogenic CH_4 and H_2 , which can be released upon dissolution or fracturing of the mineral host²¹. Although the expression of genes that are potentially involved in 5 of the 6 known carbon-fixation pathways was observed (Supplementary Table 4), transcriptomes suggest that heterotrophy is also active, and that in situ microorganisms efficiently recycle and store the available organic compounds (Fig. 4). For example, the expression of transglycosylase genes associated with peptidoglycan degradation annotated to Desulfobacterales and *Thermus* was detected at 460.4 and 643.9 mbsf. Furthermore, polycyclic aromatic hydrocarbons that can be formed through natural processes (for example, pyrolysis in hotter, deeper layers²²) may serve as carbon sources in the lower oceanic crust. We detected the expression of genes involved in polyaromatic and aromatic hydrocarbon degradation in 10 out of 11 samples; these

genes were primarily annotated to *Pseudomonas* (Supplementary Table 4). Recycled proteins are another important source of carbon. Expression of oligo-, DD-endo-, amino- and carboxypeptidases was detected (Supplementary Table 4). Expression of anabolic and catabolic genes implicated in the synthesis and degradation of amino acids known to be involved in Stickland reactions were detected in 10 of the 11 samples, accounting for up to 24% of normalized reads, indicating that these reactions may have a role in the nutrient-poor lower ocean crust (Fig. 4). Stickland reactions involve the oxidative deamination of one amino acid with a higher oxidation state, and the reductive deamination of another amino acid with a lower oxidation state, leading to ATP synthesis by substrate-level phosphorylation²³. Heterotrophic processes such as the degradation of amino acids may be the dominant terminal steps of organic carbon mineralization to CO_2 in these samples; however, the incomplete coverage of expressed genes in the recovered metatranscriptomes means that we cannot exclude contributions from additional pathways, including sulfate reduction, that may be expressed at low levels.

Many biochemical pathways produce acetyl-coenzyme A (acetyl-CoA), propionyl-CoA or acetoacetyl-CoA as intermediates that may contribute to the production of polyhydroxyalkanoates (PHAs)—bacterial and archaeal biopolymers known to serve as compatible solutes in halophiles²⁴. PHAs also serve as carbon-storage molecules, reducing equivalents and energy sources under carbon, nutrient and oxygen limitation^{25,26}. The primary role of PHAs in our samples is likely to be as energy storage in an environment in which the abundance of nutrients

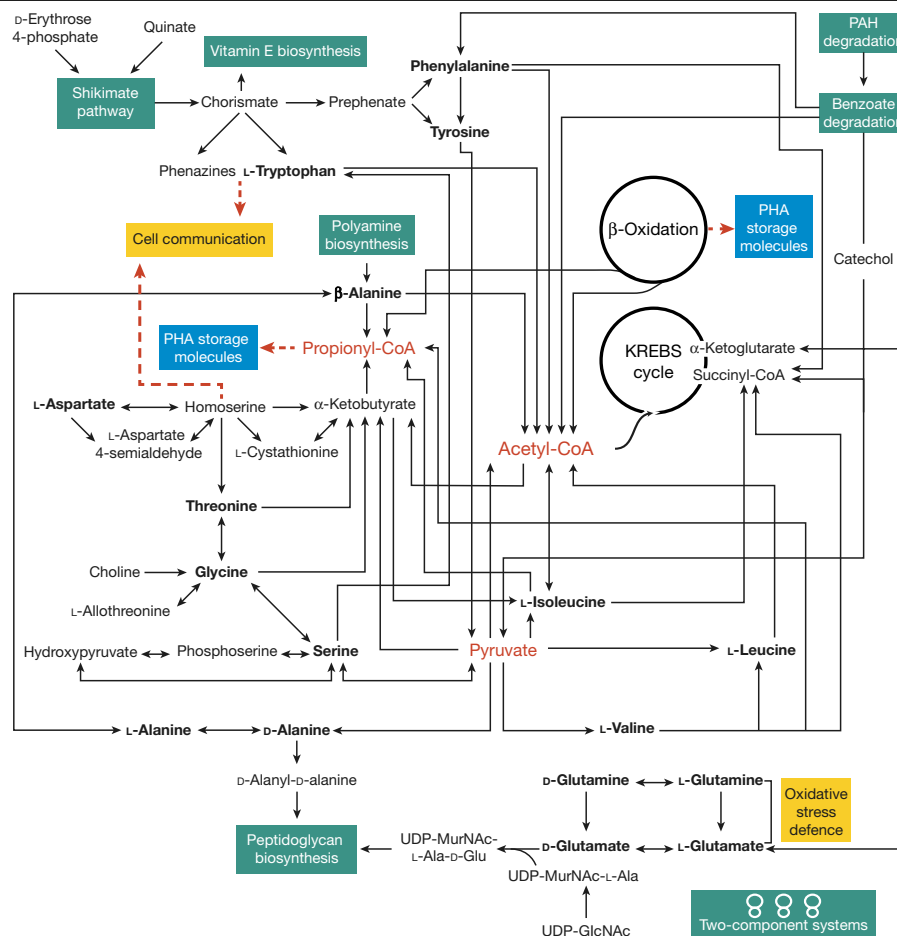


Fig. 4 | Schematic representation of metabolic processes inferred from observed transcripts in core samples from IODP Expedition 360. As the expression of all of the genes in the presented pathways was not detected in every sample, the schematic is presented as a working hypothesis only. PAH, polycyclic aromatic hydrocarbons; UDP-MurNAc, uridine diphosphate *N*-acetylmuramic acid; UDP-GlcNAc, uridine diphosphate *N*-acetylglucosamine. Amino acids are highlighted in bold. Blue boxes highlight PHA storage molecules, the production of which is supported by the detected transcripts.

Green boxes highlight biosynthetic processes, degradation pathways and two-component systems supported by transcripts. Yellow boxes highlight cellular activities supported by transcripts. Dashed red arrows represent expected (but unobserved) connections based on the data. The KREBS cycle (also known as the tricarboxylic acid cycle) generates energy via the oxidation of acetyl-CoA. β -Oxidation is the process of degradation of fatty acids to generate acetyl-CoA. The two processes have been circled.

can vary considerably over time. PHAs such as poly-3-hydroxybutyrate esters can be synthesized through an operon of three genes (*phaCAB*). Expression of *phaB* or *phaC* was detected in samples from 10.7 and 619.6 mbsf and both were detected in the sample from 714.9 mbsf; these genes were primarily annotated to *Pseudomonas*. Propionyl-CoA can be used for the biosynthesis of poly(3-hydroxybutyrate-co-3-hydroxyvalerate), a type of PHA that is produced by bacteria and archaea²⁷. Formation of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) requires the action of acetyl-CoA carboxylase and propionyl-CoA carboxylase. Expression of both genes was detected in the sample from 619.6 mbsf (Supplementary Table 4). β -Oxidation of fatty acids through 3-ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase and acyl-CoA dehydrogenase can produce propionyl-CoA, as can the catabolism of amino acids using 2-oxoisovalerate dehydrogenase and methylmalonyl-CoA mutase. One or more of these genes was detected in all samples, including those obtained from the two deepest samples (Fig. 4 and Supplementary Table 4). Collectively, our data suggest that PHAs have an important role in survival in the deep biosphere.

Inorganic phosphorus starvation is hypothesized to occur in the basaltic basement²⁸ and is a possible condition in the lower oceanic crust. The levels of inorganic phosphorus are efficiently regulated in bacteria by a two-component system²⁹ that involves ABC transporters.

Expression of phosphate transporters was detected in samples from 274.6, 460.4, 558.5, 619.6 and 747.7 mbsf. Alkaline phosphatase is part of the *pho* regulon that encodes extracellular enzymes capable of obtaining inorganic phosphorus. Genes of the *pho* regulon were expressed in samples from 619.6 and 643.9 mbsf, consistent with the detected alkaline phosphatase activity (Supplementary Table 3b). We hypothesize that recycled organic carbon may be used as a source of phosphorus for cells in this environment (Supplementary Tables 3, 4).

Our transcriptome data of life in the lower oceanic crust reveal heterotrophic activities that reflect the competition for limited and sporadically available resources, adaptations for withstanding long periods of resource scarcity, and efficient recycling of pools of organic matter in this challenging environment (further discussion of the results is provided in the Supplementary Discussion).

Conclusion

Circulation of fluids through fault zones and fractures in the lower oceanic crust can facilitate delivery of volatiles (for example, H_2 , CO_2 and CH_4), nutrients, and abiotic and biotic electron donors and acceptors. Fluid pathways may represent advantageous habitats for microorganisms, yet the distribution of life in the deep crust remains

poorly understood. We reveal a low-biomass, but viable and/or active community in the exhumed lower oceanic crust, expanding our view of the extent of Earth's biosphere. Cellular activities appear to be very low based on enzyme activity measurements and mRNA recovery. Unexpectedly, heterotrophic processes may dominate over more-familiar autotrophic processes found at seafloor hydrothermal vents and in shallow marine sediments. Microorganisms can adapt to life in this 'slow lane' at least in part by using available fermentable organic molecules³⁰. Given the global expanse of the lower oceanic crust within known temperature limits for life, even low-biomass and slow-growing communities may make non-trivial contributions to global nutrient cycling. Future exploration of deeper lower ocean crust that is not exposed to faulting is required to determine whether the diversity and activities of microbiota present at those locations are similar to those found below the Atlantis Bank.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2075-5>.

- Shah Walter, S. R. et al. Microbial decomposition of marine dissolved organic matter in cool oceanic crust. *Nat. Geosci.* **11**, 334–339 (2018).
- D'Hondt, S., Rutherford, S. & Spivack, A. J. Metabolic activity of subsurface life in deep-sea sediments. *Science* **295**, 2067–2070 (2002).
- Jørgensen, B. B. Deep subseafloor microbial cells on physiological standby. *Proc. Natl Acad. Sci. USA* **108**, 18193–18194 (2011).
- Hoehler, T. M. & Jørgensen, B. B. Microbial life under extreme energy limitation. *Nat. Rev. Microbiol.* **11**, 83–94 (2013).
- Tully, B. J., Wheat, C. G., Glazer, B. T. & Huber, J. A. A dynamic microbial community with high functional redundancy inhabits the cold, oxic subseafloor aquifer. *ISME J.* **12**, 1–16 (2018).
- Santelli, C. M., Edgcomb, V. P., Bach, W. & Edwards, K. J. The diversity and abundance of bacteria inhabiting seafloor lavas positively correlate with rock alteration. *Environ. Microbiol.* **11**, 86–98 (2009).
- Jungbluth, S. P., Bowers, R. M., Lin, H. T., Cowen, J. P. & Rappé, M. S. Novel microbial assemblages inhabiting crustal fluids within mid-ocean ridge flank subsurface basalt. *ISME J.* **10**, 2033–2047 (2016).
- Shrenk, M. O., Huber, J. A. & Edwards, K. J. Microbial provinces in the subseafloor. *Ann. Rev. Mar. Sci.* **2**, 279–304 (2010).
- Mason, O. U. et al. First investigation of the microbiology of the deepest layer of ocean crust. *PLoS ONE* **5**, e15399 (2010).
- Zhang, X., Feng, X. & Wang, F. Diversity and metabolic potentials of subsurface crustal microorganisms from the western flank of the Mid-Atlantic Ridge. *Front. Microbiol.* **7**, 363 (2016).
- Früh-Green, G. L. et al. Magmatism, serpentinization and life: insights through drilling the Atlantis Massif (IODP Expedition 357). *Lithos* **323**, 137–155 (2018).
- Lipp, J. S. & Hinrichs, K.-U. Structural diversity and fate of intact polar lipids in marine sediments. *Geochim. Cosmochim. Acta* **73**, 6816–6833 (2009).
- Valentine, D. L. Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* **5**, 316–323 (2007).
- Summons, R. E. & Lincoln, S. A. in *Fundamentals of Geobiology* (eds Knoll, A. H.) 269–296 (John Wiley and Sons, 2012).
- Swan, B. K. et al. Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* **333**, 1296–1300 (2011).
- Sheik, C. S., Jain, S. & Dick, G. J. Metabolic flexibility of enigmatic SAR324 revealed through metagenomics and metatranscriptomics. *Environ. Microbiol.* **16**, 304–317 (2014).
- Grossi, V. et al. Mono- and dialkyl glycerol ether lipids in anaerobic bacteria: biosynthetic insights from the mesophilic sulfate reducer *Desulfatibacillum alkenivorans* PF2803^T. *Appl. Environ. Microbiol.* **81**, 3157–3168 (2015).
- Hawley, A. K. et al. Diverse Marinimicrobia bacteria may mediate coupled biogeochemical cycles along eco-thermodynamic gradients. *Nat. Commun.* **8**, 1507 (2017).
- Kelley, D. S. et al. A serpentinite-hosted ecosystem: the Lost City hydrothermal field. *Science* **307**, 1428–1434 (2005).
- Puente-Sánchez, F. et al. Viable cyanobacteria in the deep continental subsurface. *Proc. Natl Acad. Sci. USA* **115**, 10702–10707 (2018).
- Klein, F., Grozeva, N. G. & Seewald, J. S. Abiotic methane synthesis and serpentinization in olivine-hosted fluid inclusions. *Proc. Natl Acad. Sci. USA* **116**, 17666–17672 (2019).
- Zolotov, M. & Shock, E. L. Abiotic synthesis of polycyclic aromatic hydrocarbons on Mars. *J. Geophys. Res. Planets* **104**, 14033–14049 (1999).
- Fonknechten, N. et al. *Clostridium sticklandii*, a specialist in amino acid degradation: revisiting its metabolism through its genome sequence. *BMC Genomics* **11**, 555 (2010).
- Cai, L. et al. Comparative genomics study of polyhydroxyalkanoates (PHA) and ectoine relevant genes from *Halomonas* sp. TD01 revealed extensive horizontal gene transfer events and co-evolutionary relationships. *Microb. Cell Fact.* **10**, 88 (2011).
- Jendrossek, D. & Handrick, R. Microbial degradation of polyhydroxyalkanoates. *Annu. Rev. Microbiol.* **56**, 403–432 (2002).
- Liu, G. et al. Enoyl-CoA hydratase mediates polyhydroxyalkanoate mobilization in *Haloflex mediterranei*. *Sci. Rep.* **6**, 24015 (2016).
- Han, J. et al. Complete genome sequence of the metabolically versatile halophilic archaeon *Haloflex mediterranei*, a poly(3-hydroxybutyrate-co-3-hydroxyvalerate) producer. *J. Bacteriol.* **194**, 4463–4464 (2012).
- Lin, H.-T. et al. Inorganic chemistry, gas compositions and dissolved organic carbon in fluids from sedimented young basaltic crust on the Juan de Fuca Ridge flanks. *Geochim. Cosmochim. Acta* **85**, 213–227 (2012).
- Santos-Beneit, F. The Pho regulon: a huge regulatory network in bacteria. *Front. Microbiol.* **6**, 402 (2015).
- Zinke, L. A. et al. Thriving or surviving? Evaluating active microbial guilds in Baltic Sea sediment. *Environ. Microbiol. Rep.* **9**, 528–536 (2017).
- Dick, H. J. B. et al. The Atlantis Bank gabbro massif, Southwest Indian Ridge. *Prog. Earth Planet. Sci.* **6**, 64 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Article

Methods

Data reporting

No power analyses were done to pre-estimate sample size. The sample size was determined by sample availability. The molecular control samples and actual samples were identified by different numbers and were treated blindly until the final stages during which these labels were replaced with the actual depth-associated identifiers.

Study site

The Atlantis Bank was uplifted from beneath sheeted dikes and pillow lavas in the local rift valley³² at the palaeo ridge-transform intersection. On the crest of the bank, there is a 25-km² wave-cut platform that is the location of an 809-m hole (U1473A) drilled during IODP Expedition 360 (32° 42.3402' S, 57° 16.6910' E) along a plate-spreading flow line at the centre of the gabbro massif. The eroded platform exposes massive foliated gabbro and oxide gabbro mylonites with enclaves of un-deformed olivine gabbro³¹. Oolitic sands and deep subaerial weathering occur on the flanks of the Bank down to 2,500 m depth, indicating the massif once rose to as much as 1.8-km above sea level, before subsiding to its present 700 m depth³¹, where it lies more than 3 km above the surrounding seafloor of the same age. Flexural uplift at the inside corner high can only account for around 1 km of this uplift, the rest is associated with later normal faulting accompanying isostatic uplift that occurred around 12 million years ago^{33,34}. A notable feature of the gabbro massif is that the boundary between crust and mantle rock is exposed along the transform valley wall for some 30 km. This structural geometry suggests that the overlying gabbro massif is in direct contact with partially serpentinized peridotite at a depth of around 4,500 m below site U1473A³³.

Sampling

Hole U1473A was drilled with a rotary coring bit, recovering cored rock in 10-m increments within core barrels that were brought on deck of the drill ship *JOIDES Resolution* and laid out on trays before splitting the cored rock samples into approximately 10–20-cm-long whole round core sections that were immediately selected and removed from the core liner for microbiological analysis by scientists who wore gloves and masks; each rock sample was placed into a sterile plastic bag for transit to the microbiology laboratory. Prioritized core samples exhibited evidence of low-temperature alterations by aqueous fluids, including carbonate veins and clay-altered felsic veins. We collected 68 samples for microbiological analysis; 11 of these—spanning the depth of hole U1473A—were selected for analyses from among those that showed the highest concentration of altered felsic and/or carbonate veins.

In the microbiology laboratory on-board the ship, each core section was rinsed three times with sterile distilled and deionized water, changing the bag each time, and then sprayed with 200-proof ethanol. At this point, after around 5 min, the core section was transferred to a custom laminar flow hood equipped with a HEPA filter and air supply that maintained positive air pressure. Within the hood, samples were photographed and the exterior approximately 1–2 cm of each whole round core was removed using a sterile chisel and rock hammer within a sterilized custom-made 0.3-m × 0.2-m × 0.1-m stainless-steel rock box. After removing core exteriors, core interiors were divided for different analyses. Three times during the cruise (during the first, second and third week of drilling), 1-litre samples of drilling fluid were filtered onto 45-mm 0.2-μm-pore-size Millipore Express Plus polycarbonate filters and frozen at –80 °C for DNA and RNA analysis. At the same time, 50-ml samples of drilling mud (Sepiolite) were also collected and frozen. Materials for DNA, RNA and lipid analyses were stored in sterile 50-ml Falcon tubes, whereas microscopy and thin-section samples were placed inside sterilized aluminium foil. These samples were carefully labelled, placed within plastic bags and immediately frozen at –80 °C. All equipment was rinsed and flame-sterilized between samples.

Contamination controls

Rotary coring contaminates the exteriors of core samples due to the circulation of drilling fluids (a mixture of Sepiolite and surface seawater) around the drilling bits. Extreme care was taken to remove or minimize this contamination and to not introduce new contaminations during sample handling and analysis. During IODP Expedition 360, a new less-volatile tracer, perfluoromethyldecalin, was successfully tested and calibrated and thus used to quantify the intrusion of drilling fluids into the interior of samples. As a further control for laboratory contamination, open Petri dishes containing microbiological medium used to culture fungi were placed inside the laminar flow hood during sample processing and these plates then were stored at room temperature for the duration of the cruise. In addition, extraction blanks, procedural controls and samples (of drilling fluids) were analysed for lipid and nucleic acid analyses (see the 'Lipid extraction and UHPLC–MS', 'DNA extraction and small subunit ribosomal-RNA marker-gene analysis' and 'RNA extraction and metatranscriptome analysis' sections).

Cell counts

Paraformaldehyde (4 ml; 4% solution in 100 mM phosphate-buffered saline) was added to autoclaved 7-ml plastic tubes. Then, 1 ml of powdered rock material was added to each tube, bringing the volume to 5 ml total. Two replicate tubes were prepared for cell counts for each sample and—where possible—vein material alone was aliquoted into one sample and whole-rock powder was aliquoted into the other. Preserved samples were stored at 4 °C for onshore analysis. For cell counts, 1 ml of the fixed sample slurry was used in the quantification procedure³⁵, which was modified to use 40 cycles of sonication instead of 20 to better release cells attached to the powdered rock. Cells were enumerated on filters stained with a 1:40 dilution of the stock SYBR Green I in TE buffer by counting either around 400 fields of view if fewer than a total of 40 cells were detected, or at least 40–50 cells in fewer fields when possible. The limit of quantification was defined as 3× the s.d. of the mean of the negative-control counts. One negative control was processed and analysed for every 11 experimental samples. All samples were counted in duplicate.

ATP and exoenzyme assays

ATP concentration was assessed for all samples on board the *JOIDES Resolution* using luminescence methods³⁶ with the ATP Bioluminescence Assay Kit (Sigma-Aldrich) and a Turner Designs BioSystems 20/20 luminometer (Promega). The presence of ATP is indicative of microbial biomass. A standard curve of 0 (sterile MilliQ water), 1 and 100 ng l^{–1} ATP standard was run with each analysed set of samples. ATP in core samples was measured by placing 500 μl of ATP Assay Mix into a clean, sterile microcentrifuge tube after which the tube was incubated at room temperature for 3 min to allow hydrolysis of endogenous ATP, thus decreasing the background signal. Then around 0.5 cm³ of powdered sample was added (all tubes were weighed) to the same tube and the solution was immediately transferred to a clean 1.9-ml screw-top glass vial (acid washed) with the cap off, using a 1-ml pipette tip with the tip cut off. The vial was placed into the luminometer and results were read immediately. From the beginning of the expedition until drilling reached 218 mbsf, samples for ATP assays used leftover powdered rock in the steel rock box after separation of sample material for other assays. After observing largely undetectable ATP concentration, a second approach was used because microbial cells—if present—are likely to be more concentrated along cracks, within veins and vugs. Thus, starting at 218 mbsf, material from these specific features were included for these analyses.

Alkaline phosphatase activity was measured using fluorogenic substrate 4-methylumbelliferyl phosphate (Sigma-Aldrich) and its reference standard, methylumbelliferone. Fluorescence was measured using black, flat-bottom, 96-well microplates in a Spark 10M

Multimode Microplate Reader (Tecan). Fluorescence of methylumbelliferone is greatest at pH 10, therefore 25 μ l of 0.4 M NaOH was added to the wells (final concentration 40 mM) to be read. Then, 25 μ l of 1 M EDTA was added to the wells (100 mM final concentration) to prevent precipitation of carbonate from sampled veins³⁷. Fluorescence was measured with an excitation wavelength of 380 nm and emission of 454 nm for all substrates and standards. We mixed 1 cm³ powdered rock with 5 cm³ of sterile artificial seawater (ASW) in a 8-ml serum vial with 90:5:5 N₂:CO₂:H₂ headspace. Then, 700 μ l of each slurry was transferred with a sterile syringe into a 1.5-ml Eppendorf tube after set-up but before sealing the vial; this sample served as T_0 , which comprised triplicate 200 μ l technical replicates. The 700- μ l samples were briefly centrifuged (60 s at 2,500 rpm) and the supernatant was used for the T_0 analyses. Two additional samples were taken using the same methods as for T_0 after at least 2 weeks and then again after 4–6 weeks to generate a slope of activity. Incubations were kept at 10 °C, the inferred *in situ* temperature, for the duration of each assay.

Autoclaved, powdered rock from each of the samples was tested to determine the amount of fluorophore adsorbance to rock powder. Adsorbance was found to behave in a systematic way, resulting in a straight line when comparing fluorescence standards in ASW alone with fluorescence standards plus rock powder in ASW, although this relationship was found to be different when measured after 4 h versus days later. Therefore, a correction factor for adsorbance was applied to the enzyme data for the initial measurement (T_0 , $y = 1.90x - 676$), which was taken less than 2 h after experiment initiation, versus the second and third measurements (T_1 and T_2 , $y = 4.64x - 303$), which were taken days to weeks later. Negative controls consisting of the same ASW used for the sample incubations plus substrate, but no sample, were consistently below the limit of detection. The limit of quantification for the alkaline phosphatase assay, defined as 3 $\times$ the s.d. of the blank, was 0.0242 pmol cm⁻³ of rock h⁻¹ based on the analysis of 8 blanks.

Carbon and nitrogen analysis

Powdered rock material from each sample (produced in the laminar flow hood on the *JOIDES Resolution* using a sterile mortar and pestle) was immediately transferred to sterile, muffled glass containers and stored in a desiccator until analysis of carbon and nitrogen according to established methods^{38,39}. In brief, samples were weighed into methanol-rinsed silver boats (4 mm $\times$ 6 mm, Costech). Then, 96-well glass plates (combusted for 4 h at 450 °C) holding these samples were placed in a vacuum desiccator that also contained an open dish with about 50 ml fresh, concentrated (12 N) HCl. An inverted crystallization dish was placed over the samples to protect them from water condensation. The desiccator was closed and pumped out with an air-driven aspirator, to a reading of about 0.5 atm and the desiccator was placed in an oven kept between 60 and 65 °C. Acidification was allowed to run for 60 to 72 h, as described previously⁴⁰. When acidification was complete, the samples were removed and set in the oven to dry (60–65 °C). Subsequently, the samples were placed in a vacuum desiccator charged with indicating silica gel (Fisher S162-500, activated by heating above 220 °C for several hours) and pumped down again and dried for about 24 h. Samples were then analysed on a Costech 4010 Elemental Analyzer connected via a Finningan-MAT ConFlo-II interface to a DeltaPlus isotope ratio mass spectrometer.

Thin-section preparation, scanning electron microscopy and Raman analyses

Thin-section billets were cut from dedicated subsamples of the 11 samples examined in this study. Thin sections were prepared by High Mesa Petrographics. Mosaic images were taken of all thin sections in transmitted and reflected light and were used to guide the scanning electron microscopy (SEM) and Raman analyses. Uncoated thin sections were screened with SEM in low-vacuum mode to search for and visualize carbon-rich inclusions of possible organic origin. SEM was performed

on a Hitachi TM3000 scanning electron microscope equipped with a Bruker energy-dispersive spectroscopy system for imaging and semi-quantitative element analysis. Promising samples with possible organic inclusions were analysed using a computer-controlled, high-resolution confocal Raman system (Horiba LabRam HR) equipped with three lasers (633 nm, 532 nm and 473 nm), a motorized stage and a SWIFT fast-mapping option. Confocal Raman spectroscopy enables the non-destructive analysis and recognition of living and fossil (once-living) microorganisms in altered igneous rocks. The achievable lateral and spectral resolution of this instrument is better than 1 μ m and 2 cm⁻¹, respectively. Spectra were analysed using the BioRad KnowItall software and spectral databases to identify organic compounds.

Lipid extraction and UHPLC–MS lipid biomarker analysis

Crushed core samples stored in Falcon tubes at –80 °C were first milled for 10 min to a fine powder and subsequently extracted with a modified Bligh and Dyer method according to a previous study⁴¹. Before milling and extraction of each sample, a procedure blank was performed. First, a milling blank was performed using combusted sea sand (fired at 450 °C for 5 h) to clean the mill and to limit cross-contamination of samples. Subsequently, this sea sand was transferred to geo-cleaned (rinsed three times with a mixture of methanol and dichloromethane) Teflon containers used for the extraction of the samples and solvent-extracted in the same manner as the samples. For this, 100 ng of an internal standard (C46 glycerol trialkyl glycerol tetraether) and around 50 ml of a solvent mixture of dichloromethane:methanol:buffer (2:1:0.8, v/v) was added to the sample in the Teflon container and ultrasonicated for 10 min using a geo-cleaned ultrasonic stick. After ultrasonication, the samples were centrifuged (1,750 rpm at 10 min) and the supernatant was transferred to a fired separatory funnel. The samples were extracted in four steps, for the first two steps a phosphate buffer (K₂HPO₄, 50 mM at pH 7.4) was used, in the third step the phosphate buffer was replaced by 5% trichloroacetic acid (50 g l⁻¹ at pH 2) and in the last step only dichloromethane:methanol (9:1, v/v) was used. Equal amounts of dichloromethane and deionized MilliQ water were added to the extract collected in the separatory funnel, the mixture was shaken, and the organic phase was collected as the total lipid extract and blown to dryness under a gentle stream of nitrogen.

An aliquot of the total lipid extract was analysed using ultrahigh-pressure liquid chromatography (UHPLC) coupled to mass spectrometry (MS) on a Dionex Ultimate 3000RS UHPLC connected to an ABSciEX QTRAP4500 Triple Quadrupole/Ion Trap MS (UHPLC-triple quad-MS) using a Turbolon electrospray ion (ESI) source. Separation of compounds was achieved on a Waters Acquity BEH C18 column (1.7 μ m, 2.1 mm $\times$ 150 mm) equipped with a guard column of the same material following a previously published protocol⁴². Compounds of interest were screened with multiple-reaction monitoring and selected-ion monitoring techniques as described previously⁴². Concentrations of lipids were determined relative to the internal C46 glycerol trialkyl glycerol tetraether standard and were corrected for individual response factors using commercially available standards (diC16-DEG, archaeol) and isolated standards from cultures (GDGT-0, 1G-AR, 2G-AR, 1G-GDGT-0 and 2G-GDGT-0). Sciex Analyst 1.6.3 and Sciex MultiQuant 3.0.3 (AB Sciex) were used for triple-quadrupole MS data acquisition and data processing.

The presence of crenarchaeol was confirmed by core GDGT analysis according to a previously published study⁴³. In brief, an aliquot of the total lipid extract was analysed using a Dionex Ultimate 3000RS UHPLC connected to a Bruker maXis ultrahigh-resolution quadrupole time-of-flight MS, equipped with an atmospheric pressure chemical ionization (APCI) II source. Compounds were separated using two acquity BEH HILIC amide columns (1.7 μ m, 2.1 mm $\times$ 300 mm) in tandem maintained at 50 °C, and *n*-hexane as eluent A and *n*-hexane:isopropanol (90:10, v/v) as eluent B (a detailed protocol has been published previously⁴³). Drilling mud and extraction blank contamination controls were also

run for lipid biomarker analyses (Extended Data Figs. 5, 6). Bruker Compass 1.9 and Bruker data analysis v.4.4 (Bruker Daltonics) were used for quadrupole time-of-flight data acquisition and processing.

DNA extraction and small subunit ribosomal-RNA marker-gene analysis

Rock material was crushed while still frozen in a Progressive Exploration Jaw Crusher (Model 150), the surfaces of which were sterilized with 70% ethanol and RNase AWAY (Thermo Fisher Scientific) inside a laminar flow hood. Powdered rock material was returned to the -80°C freezer until extraction. DNA was extracted from 20, 30 or 40 g of powdered rock material, depending on the quantity of rock available. A DNeasy PowerMax Soil Kit (Qiagen) was used according to the manufacturer's protocol modified to include three freeze–thaw treatments before the addition of Soil Kit solution C1. Each treatment consisted of 1 min in liquid nitrogen followed by 5 min at 65°C . DNA extracts were concentrated by isopropanol precipitation overnight at 4°C . The low biomass in our samples required whole-genome amplification before PCR amplification of marker genes. Genomic DNA was amplified by multiple displacement amplification using the REPLI-g Single Cell Kit (Qiagen) as described. Multiple displacement amplification bias was minimized by splitting each whole-genome amplification sample into triplicate 16 μl reactions after 1 h of amplification and then resuming amplification for the manufacturer-specified 7 h (8 h total). DNA was also recovered from samples of drilling mud and drilling fluid (surface water collected during the coring process) for negative controls, as well as two 'kit control' samples, in which no sample was added, to account for any contaminants originating from either the DNeasy PowerMax Soil Kit or the REPLI-g Single Cell Kit. Bacterial small subunit ribosomal RNA (rRNA) gene fragments were PCR amplified from multiple displacement amplification samples and sequenced at the Georgia Genomics and Bioinformatics Core (University of Georgia). The primers used were: Bac515-Y and Bac926R⁴⁴. Dual-indexed libraries were prepared with (HT) iTruS (Kappa Biosystems) chemistry and sequencing was performed on an Illumina MiSeq $2 \times 300\text{-bp}$ system with all samples combined equally on a single flow cell.

Raw sequence reads were processed using Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), FLASH (<http://ccb.jhu.edu/software/FLASH/>) and FASTX Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) for trimming and removal of low-quality and/or short reads. Quality filtering included requiring a minimum average quality of 25 and rejection of paired reads that were less than 250 nucleotides. Operational taxonomic unit (OTU) clusters were constructed at 99% similarity with the script `pick_otus.py` within the Quantitative Insights Into Microbial Ecology (QIIME) v.1.9.1 software⁴⁵ and `ucrust`. Any OTU that matched an OTU in one of our control samples (drilling fluids, drilling mud, extraction and whole-genome amplification controls) was removed (using `filter_otus_from_otu_table.py`) along with any sequences of land plants and human pathogens that may have survived the control filtering due to clustering at 99% (`filter_taxa_from_otu_table.py`). As an additional quality-control measure, genera that are commonly identified as PCR contaminants were removed^{46,47}.

Unclassified OTUs were queried using BLAST against the GenBank nr database and further information about these OTUs is provided in the Supplementary Discussion. OTUs that could not be assigned to Bacteria or Archaea were removed from further analysis. For downstream analyses, any OTUs not representing more than 0.01% of the relative abundance of sequences overall were removed as those are unlikely to contribute considerably to in situ communities. The OTU data table was transformed to a presence or absence table and the Jaccard method was used to generate a distance matrix using the `dist.binary()` function in the R package `ade4`. A hierarchical clustering dendrogram was created using `hclust()` and the stability of the clusters was evaluated using the `clusterboot()` function in the `fpc` package in R with 500 iterations.

RNA extraction and metatranscriptome analysis

Frozen rock material was crushed as above and then ground quickly into a fine powder using a precooled sterilized mortar and pestle, after which RNA extraction started immediately. The jaw crusher was cleaned and rinsed with 70% ethanol and RNaseZap RNase Decontamination Solution (Invitrogen) between samples. About 40 g of material was extracted for each sample using the RNeasy PowerSoil Total RNA Isolation Kit (Qiagen) according to the manufacturer's protocol with the following modifications. Each sample was evenly divided into 8 bead tubes (Qiagen) and then 2.5 ml of Bead solution was added into the bead tube followed by 0.25 ml of solution SR1 and 0.8 ml of solution SR2. Bead tubes were frozen in liquid nitrogen and then thawed at 65°C in a water bath three times. RNA was purified using the MEGAclear Transcription Clean-up Kit (Ambion) and concentrated with an overnight isopropanol precipitation at 4°C . Trace amounts of contaminating DNA were removed from the RNA extracts using TURBO DNA-free (Invitrogen, USA) as described by the manufacturer. To ensure that DNA was removed thoroughly, each RNA extract was treated twice with TURBO DNase (Invitrogen). A nested PCR reaction (2×35 cycles) using bacterial primers⁴⁸ was used to confirm the absence of DNA in our RNA solutions.

RNA was converted to cDNA using the Ovation RNA-Seq System V2 kit (NuGEN) according to the manufacturer's protocol to preferentially prime non-rRNA sequences. The cDNA was purified with the MinElute Reaction Cleanup Kit (Qiagen) and eluted into 20 μl elution buffer. Extracts were quantified using a Qubit Fluorometer (Life Technologies) and cDNAs were stored at -80°C until sequencing using 150-bp paired-end reads on an Illumina NextSeq 550.

To control for potential contaminants introduced during drilling, sample handling and laboratory kit reagents, we sequenced a number of control samples as described above. This included two samples that controlled for potential nucleic acid contamination, a 'method' control to monitor possible contamination from our laboratory extractions, which included around 40 g sterilized glass beads processed through the entire protocol in place of rock, and a 'kit' control to account for any signal coming from trace contaminants in kit reagents, which received no addition. In addition, three field controls were extracted: a sample of the drilling mud (Sepiolite), and two drilling seawater samples collected during the first and third weeks of drilling. cDNA obtained from these controls was sequenced together with the rock samples and co-assembled.

Trimmomatic (v.0.32)⁴⁹ was used to trim adaptor sequences (leading = 20, trailing = 20, sliding window = 04:24, minlen = 50). Paired reads were further quality checked and trimmed using FastQC (v.0.11.7) and FASTX toolkit (v.0.014). Downstream analyses used paired reads. After co-assembling reads with Trinity (v.2.4.0) from all controls (minimum length, 150 bp), Bowtie2 (v.2.3.4.1)⁵⁰ was used (with the parameter 'uncon') to align all sample reads to this co-assembly. Reads that mapped to our control co-assembly allowing 1 mismatch were removed from further analysis (23.5–68.5% of sequences remained in sample datasets; Supplementary Table 4). Trinity (v.2.4.0) was used for de novo assembly of the remaining reads in sample datasets (minimum length, 150 bp). Bowtie aligner was used to align reads to assembled contigs, RSEM⁵¹ was used to estimate the expression level of these reads, and TMM was used to perform cross-sample normalization and to generate a TMM-normalized expression matrix. Within the Trinotate suite, TransDecoder (v.3.0.1) was used to identify coding regions within contigs and functional and taxonomic annotation was made by BLASTx and BLASTp against UniProt, Swissprot (release 2018_02) and RefSeq non-redundant (nr) protein sequence databases (*e*-value threshold of 1×10^{-5}). BLASTp was used to look for sequence homologies with the same *e* values. HMMER (v.3.1b2) was used to identify conserved domains by searching against the Pfam (v 31.0) database⁵². SignalP (v.4.1)⁵³ and TMHMM (2.0c)⁵⁴ were used to predict signal peptides and transmembrane domains. RNAMMER (v.1.2)⁵⁵ was used to identify

rRNA homologies of archaea, bacteria and eukaryotes. Because the Swissprot database does not have extensive representation of protein sequences from environmental samples, particularly deep-sea and deep-biosphere samples, annotations of contigs used for analyses of selected processes that we report were manually cross checked by BLASTx against the GenBank nr database.

Aside from removing any reads that mapped well to our control co-assembly (1 mismatch), as an extra precaution, any sequence that exhibited $\geq 95\%$ sequence identity over $\geq 80\%$ of the sequence length to suspected contaminants (for example, human pathogens, plants or taxa known to be common molecular kit reagent contaminants, and not described from the marine environment) as described previously^{46,56} were removed. This conservative approach potentially removed environmentally relevant data that were annotated to suspected contaminants due to poor taxonomic representation from environmental taxa in public databases; however, it affords the highest possible confidence about any transcripts discussed. Additional functional annotations of contigs were obtained by BLAST against the KEGG, COG, SEED and MetaCyc databases using MetaPathways (v.2.0) to gain insights into particular cellular processes, and to provide overviews of metabolic functions across samples based on comparisons of fragments per kilobase per million (FPKM)-normalized data. All annotations were integrated into a SQLite database for further analysis.

Statistics

Ocean-drilling legs with diverse sampling requirements can provide only a limited volume of rock material for microbiology from each depth horizon. Although still providing tremendously valuable insights into the deep lithified biosphere, the absence of replicate samples limits the types of statistical analyses that we can perform. A clustering analysis of curated transcripts within the functional categories presented in this paper was performed using log + 1-transformed FPKM values and the Ward method, and the distance matrix was constructed using the Manhattan method and pvcust in R. Non-metric multidimensional scaling (NMDS) analysis was performed on the Jaccard distance matrix of prokaryotic OTU presence or absence data for 11 samples but not on transcript data because we report here only a carefully curated subset of total reads for each sample.

Taxonomic assignments of discussed contigs are presented in Supplementary Table 4, and were selected from among the top-10 BLASTx hits. In cases where different taxa were included within the top-10 hits for a conserved domain, either the top hit or the top consensus hit to a known marine group was selected.

Fungal culturing

Culturing efforts to specifically isolate fungi were performed using several culture conditions using not only classical culture-based approaches but also laser-nephelometry-based high-throughput culturing. Three different culture media were used (a modified Sabouraud (1 g l⁻¹ mycological peptone, 3% sea salts, 15 g l⁻¹ agar), a yeast nitrogen base (7 g l⁻¹ YNB powder (Difco), 3% sea salts, 15 g l⁻¹ agar), a minimal medium as defined previously⁵⁷ and formulated with or without crushed oceanic crustal rock (10 g l⁻¹), with variable carbon sources (galactose, celluloses or chitin at 15 g l⁻¹), with a mix of essential amino acids, and with or without a mix of antibiotics (chloramphenicol and penicillin G). Each experiment was processed at different temperatures ranging from 5 to 35 °C. After isolation, a dereplication step by mini- or microsatellite-primed (MSP)-PCR was processed to obtain unique isolates⁵⁸ that were identified based on sequencing of several genetic markers (ITS, 18S, 28S or 26S rRNA genes, RPB2, tubulin and actin genes).

Methane analysis

Roughly 5 cm³ of rock from depths of 247.71, 279.55, 621.09, 724.68 and 747.78 mbsf were added to sterile ASW to 27 ml total in 30-ml

serum vials. The control for the methane measurements was a vial of 27 ml of ASW with no rocks added. The vials were sealed with butyl stoppers and gassed with a mixture of 90% N₂, 5% H₂ and 5% CO₂, and incubated at 10 °C. Methane in the headspace of the vials was measured at 25 weeks and 60 weeks after incubation. The gas in the headspace of the vials was extracted and measured using gas chromatography–flame ionization detection (GC–FID), following protocols similar to previously described methods⁵⁹. Calibrations were performed for the measurements using methane tanks at 0 ppm, 1.0 ppm and 6.41 ppm, and the GC–FID loop was flushed for 1 min before each calibration measurement. The syringe used to inject the gas samples into the GC–FID loop was flushed three times with the 90% N₂, 5% H₂ and 5% CO₂ gas mixture before each sample extraction and injection.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

iTAG data for the 11 samples as well as 13 different negative controls (drilling muds and fluids, seawater and kit controls) are deposited in the NCBI BioProject under accession number PRJNA497074. Results are presented at approximately phylum level; however, taxonomic assignments at finer resolution are available from the corresponding author upon request. Raw reads for transcriptome data have also been deposited in BioProject under accession number PRJNA497074; iTAGs have been deposited in the NCBI SRA under accession numbers SRR8136794–SRR8136814 and transcript raw reads can be found in the SRA under accession numbers SRR8141073–SRR8141077. Assemblies for curated portions of the data presented are available upon request to the corresponding author. All relevant data are available from the corresponding author or are included with the manuscript as Supplementary Information, and the Source Data for Fig. 2 are available at <http://web.iodp.tamu.edu/DESCReport/> (file name, 360_U1473A_macroscopic.xlsx). Source Data for Fig. 3 and Extended Data Figs. 1, 2, 6 are provided with the paper.

32. Fox, P. J. & Gallo, D. G. A tectonic model for ridge-transform-ridge plate boundaries: implications for the structure of oceanic lithosphere. *Tectonophysics* **104**, 205–242 (1984).
33. Dick, H. J. B. et al. Dynamic accretion beneath a slow-spreading ridge segment: IODP hole 1473A and the Atlantis Bank oceanic core complex. *J. Geophys. Res. Solid Earth* **124**, 12631–12659 (2019).
34. Baines, A. G. et al. Mechanism for generating the anomalous uplift of oceanic core complexes: Atlantis Bank, southwest Indian Ridge. *Geology* **31**, 1105–1108 (2003).
35. Morono, Y., Terada, T., Kallmeyer, J. & Inagaki, F. An improved cell separation technique for marine subsurface sediments: applications for high-throughput analysis using flow cytometry and cell sorting. *Environ. Microbiol.* **15**, 2841–2849 (2013).
36. Lundin, A., Hasenson, M., Persson, J. & Pousette, A. Estimation of biomass in growing cell lines by adenosine triphosphate assay. *Methods Enzymol.* **133**, 27–42 (1986).
37. Coolen, M. J. & Overmann, J. Functional exoenzymes as indicators of metabolically active bacteria in 124,000-year-old sapropel layers of the eastern Mediterranean Sea. *Appl. Environ. Microbiol.* **66**, 2589–2598 (2000).
38. Pella, E. Elemental organic analysis. Part 1, historical developments. *Am. Lab.* **22**, 116–125 (1990).
39. Pella, E. Elemental organic analysis. Part 2: State of the art. *Am. Lab.* **22**, 28–32 (1990).
40. Whiteside, J. H. et al. Pangean great lake paleoecology on the cusp of the end-Triassic extinction. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **301**, 1–17 (2011).
41. Sturt, H. F., Summons, R. E., Smith, K., Elvert, M. & Hinrichs, K. U. Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry—new biomarkers for biogeochemistry and microbial ecology. *Rapid Commun. Mass Spectrom.* **18**, 617–628 (2004).
42. Klein, A. T. et al. Investigation of the chemical interface in the soybean–aphid and rice–bacteria interactions using MALDI-mass spectrometry imaging. *Anal. Chem.* **87**, 5294–5301 (2015).
43. Becker, K. W. et al. An improved method for the analysis of archaeal and bacterial ether core lipids. *Org. Geochem.* **61**, 34–44 (2013).
44. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).

45. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
46. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).
47. Sheik, C. S. et al. Identification and removal of contaminant sequences from ribosomal gene databases: lessons from the Census of Deep Life. *Front. Microbiol.* **9**, 840 (2018).
48. Cole, J. R. et al. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* **37**, D141–D145 (2009).
49. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
50. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
51. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
52. Finn, R. D. et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).
53. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
54. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
55. Lagesen, K. et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100–3108 (2007).
56. Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B. & Chiodini, R. J. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog.* **8**, 24 (2016).
57. Le Calvez, T., Burgaud, G., Mahé, S., Barbier, G. & Vandenkoornhuise, P. Fungal diversity in deep-sea hydrothermal ecosystems. *Appl. Environ. Microbiol.* **75**, 6415–6421 (2009).
58. Burgaud, G., Arzur, D., Durand, L., Cambon-Bonavita, M.-A. & Barbier, G. Marine culturable yeasts in deep-sea hydrothermal vents: species richness and association with fauna. *FEMS Microbiol. Ecol.* **73**, 121–133 (2010).
59. Valentine, D. L. et al. Propane respiration jump-starts microbial response to a deep oil spill. *Science* **330**, 208–211 (2010).

Acknowledgements We thank the captain, crew and all who sailed on *JOIDES Resolution* for IODP Expedition 360, whose support was essential; T. Sehein and M. Torres-Beltran for assistance with statistical analyses; Q. Ma, E. S. Taylor (WHOI Creative) and Andrew Newman Design for assistance with figures; S. Yvon-Lewis for providing materials and assistance with methane measurements; and M. Sogin, J. Huber, B. Orcutt, K. Lloyd, J. Biddle and S. D'Hondt for helpful discussions on the relative merits of different molecular data-handling options for contamination controls. This study was funded by National Science Foundation grants OCE-1658031 to V.P.E. and F.K., OCE-1658118 to J.B.S., and OCE-1450528 and OCE-1637130 to H.J.B.D. F.S. acknowledges funding from the DFG under Germany's Excellence Strategy (no. EXC-2077-390741603) and the Gottfried Wilhelm Leibniz Program (HI 616-14-1). Support to J.L. was provided by the National Science Foundation of China (no. 41772358) and the Ministry of Science and Technology of China (no. 2012CB417302).

Author contributions V.P.E., J.B.S., F.K., F.S. and J.L. acquired funding. V.P.E. and J.B.S. collected samples, performed shipboard assays and established enrichment cultures. J.L. extracted RNA. P.M., J.L. and V.P.E. performed mRNA analyses. D.B., S.L. and R.C. extracted DNA, and D.B. and S.Y.W. analysed iTAG data. F.S. and L.A.E.M. performed lipid biomarker analyses. G.B. and M.Q. performed fungal isolations. F.K. performed Raman spectroscopy experiments and J.L. carried out SEM with F.K. J.B.S. and S.Y.W. performed cell counts and exoenzyme assays, and analysed methane-generation experiments. H.J.B.D. and F.K. provided the geological context, and D.K.B. compiled downhole-core description data. P.M. conceived the working hypotheses in Fig. 4. V.P.E. wrote the manuscript draft and all authors contributed to review and editing.

Competing interests The authors declare no competing interests.

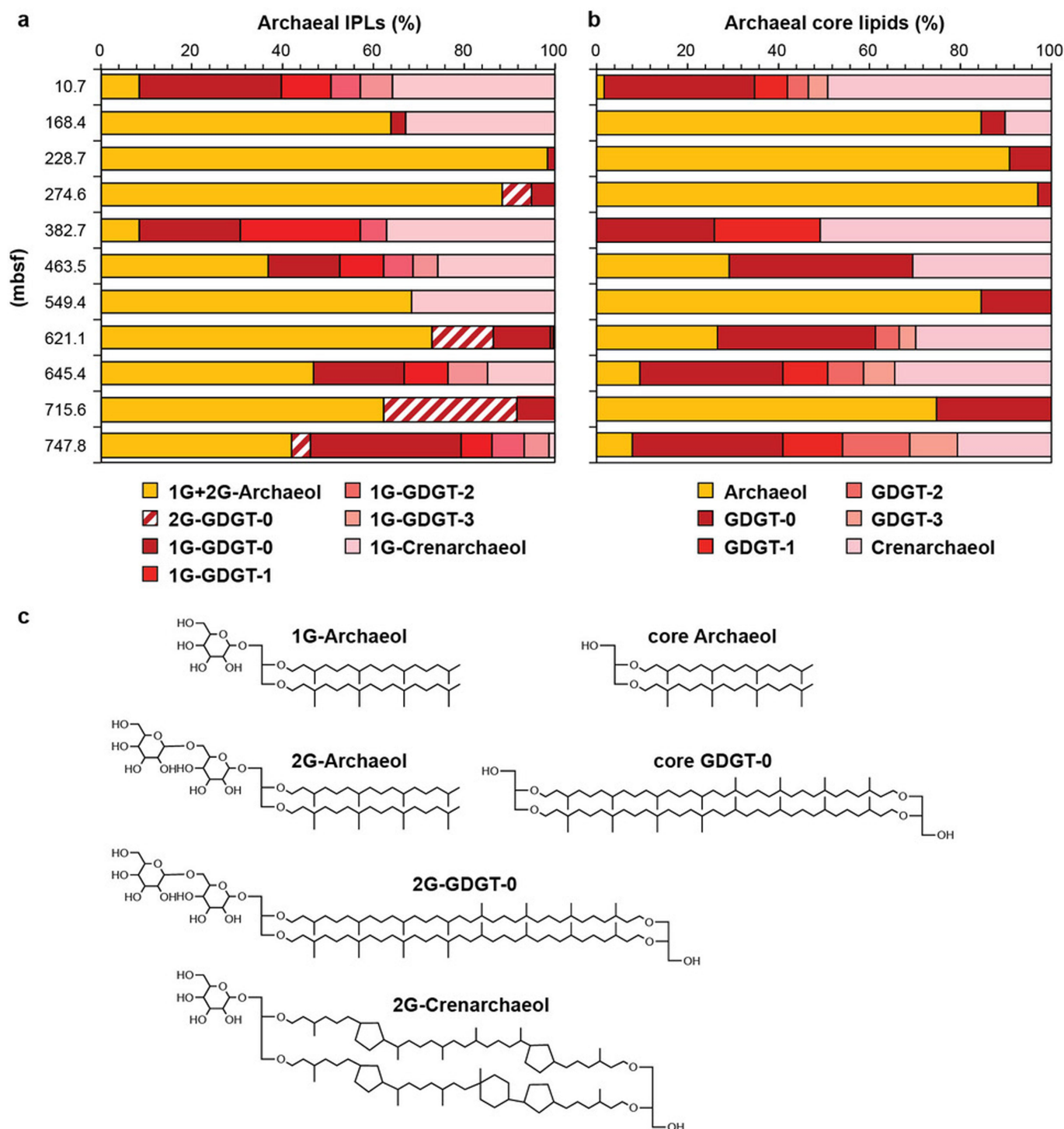
Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-2075-5>.

Correspondence and requests for materials should be addressed to V.P.E.

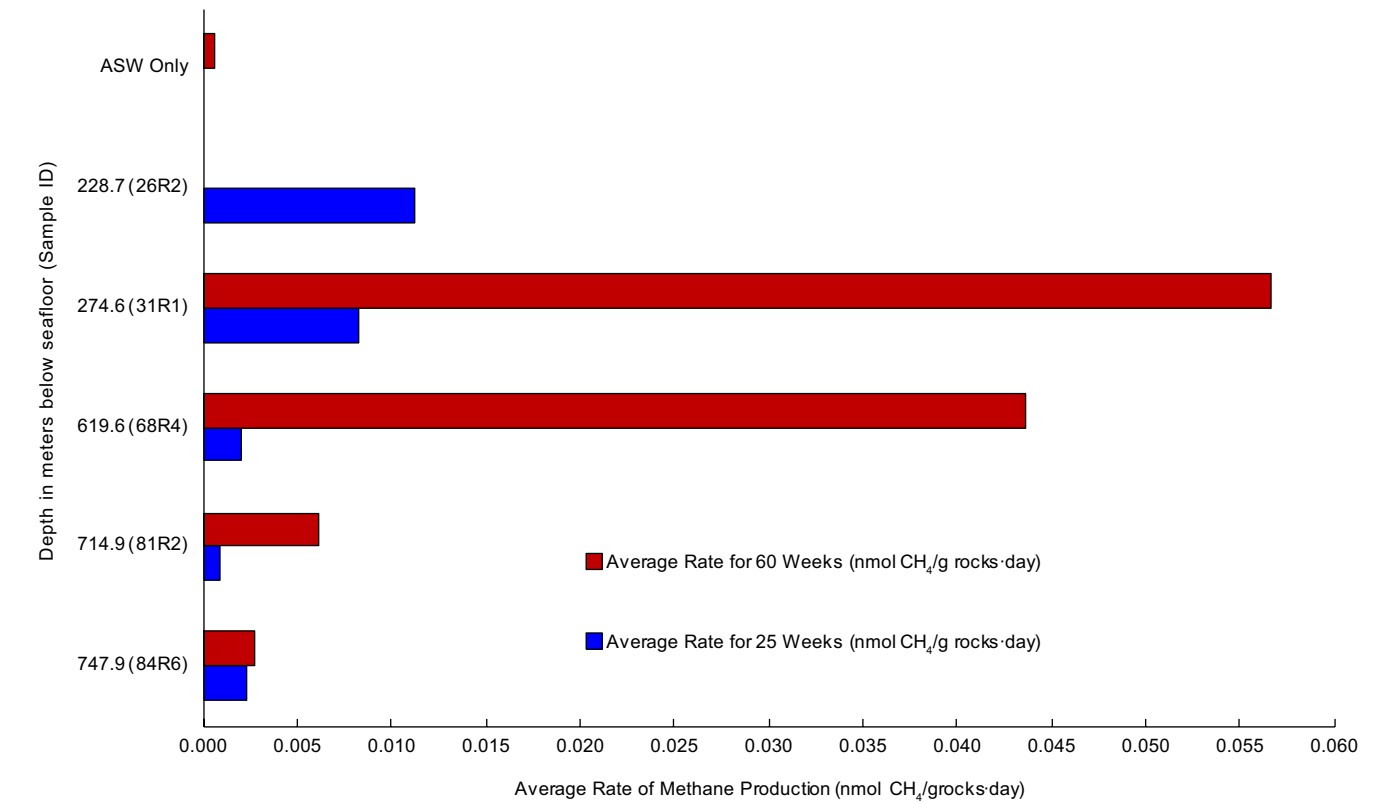
Peer review information *Nature* thanks Bo Barker Jørgensen, Jennifer Biddle and Steven D'Hondt for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



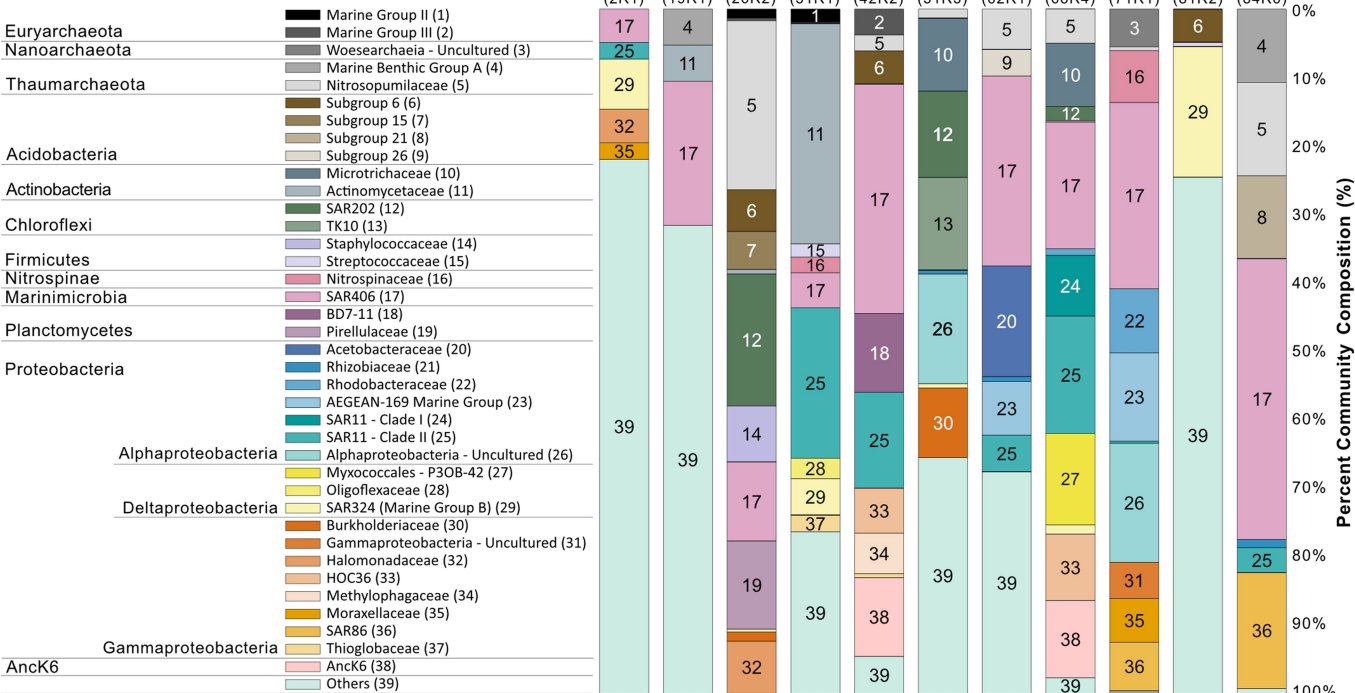
Extended Data Fig. 1 | Archaeal core lipid analyses. a, b, Downhole changes in core lipid composition of archaeal IPLs (**a**) and archaeal core lipids (**b**). The relative abundances of the diether lipid archaeol and tetraether lipids with zero, one, two and three rings (GDGT-0, GDGT-1, GDGT-2 and GDGT-3,

respectively) and crenarchaeol are shown. **c,** Structures of the most abundant archaeal lipids. 1G, monoglycosidic; 2G, diglycosidic; AR, archaeol; GDGT, glycerol dialkyl glycerol tetraether. Lipid data are from single measurements owing to sample constraints. Source Data are available online.

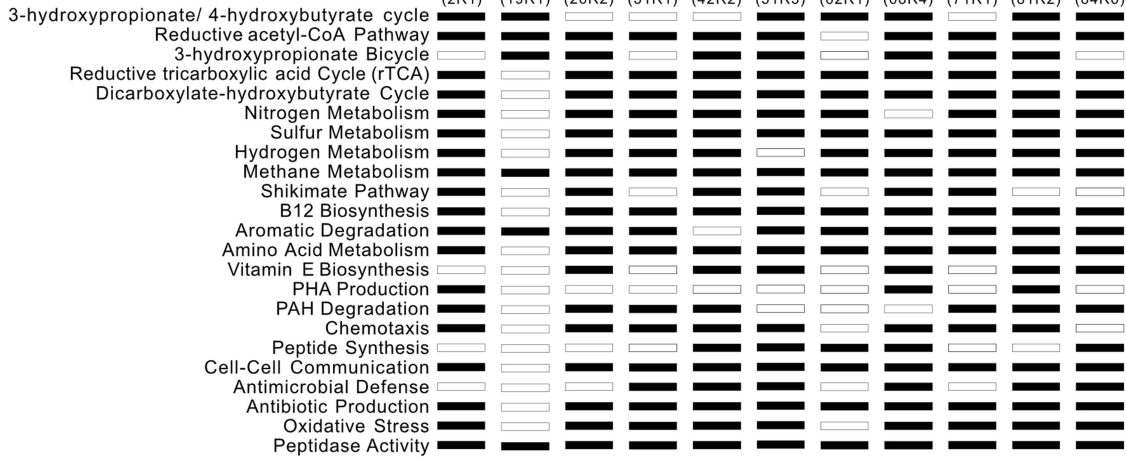


Extended Data Fig. 2 | Rate of methane production observed in long-term enrichment experiments. The rate at 25 weeks was derived from the value measured at 25 weeks minus the initial value (0) divided by the time elapsed in days. The rate at 60 weeks was derived from the difference between methane measured at 60 weeks and methane measured at 25 weeks divided by the time elapsed since the 25-week measurement. Data are from single measurements owing to limited sample availability, except for the ASW blank ($n = 3$), for which the mean is displayed. Source Data are available online.

Taxonomy

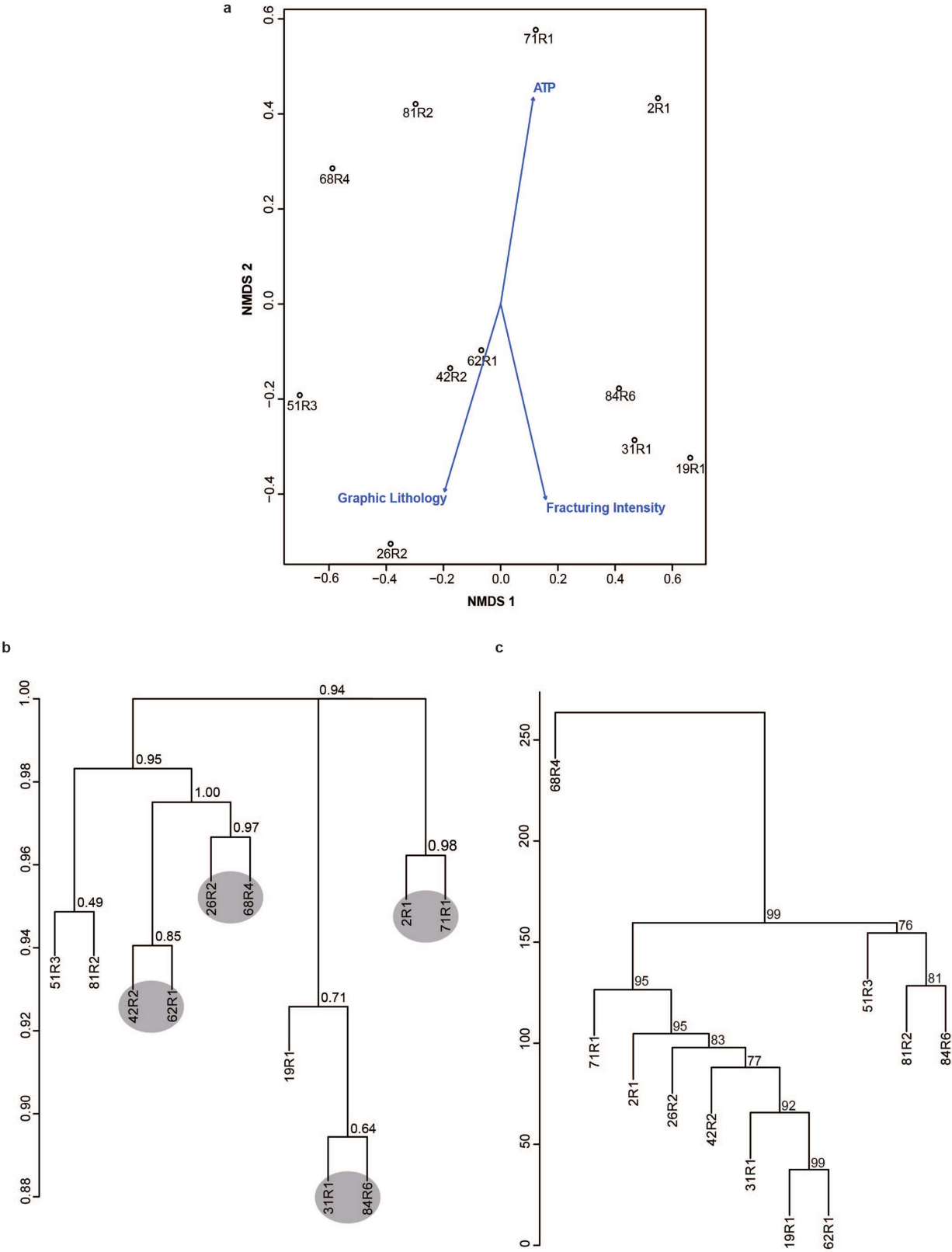


Metabolic Pathways



Extended Data Fig. 3 | 16S rRNA iTAG composition and metabolic processes detected in metatranscriptomes. Top, taxonomic composition at family level or deeper (where possible) of 16S rRNA iTAG sequences for taxa present as more than 1% of the total abundance in at least one sample. Taxa present at an abundance of less than 1% in every sample were grouped into the 'Others' category. Bottom, an overview of categories of expressed genes in each sample

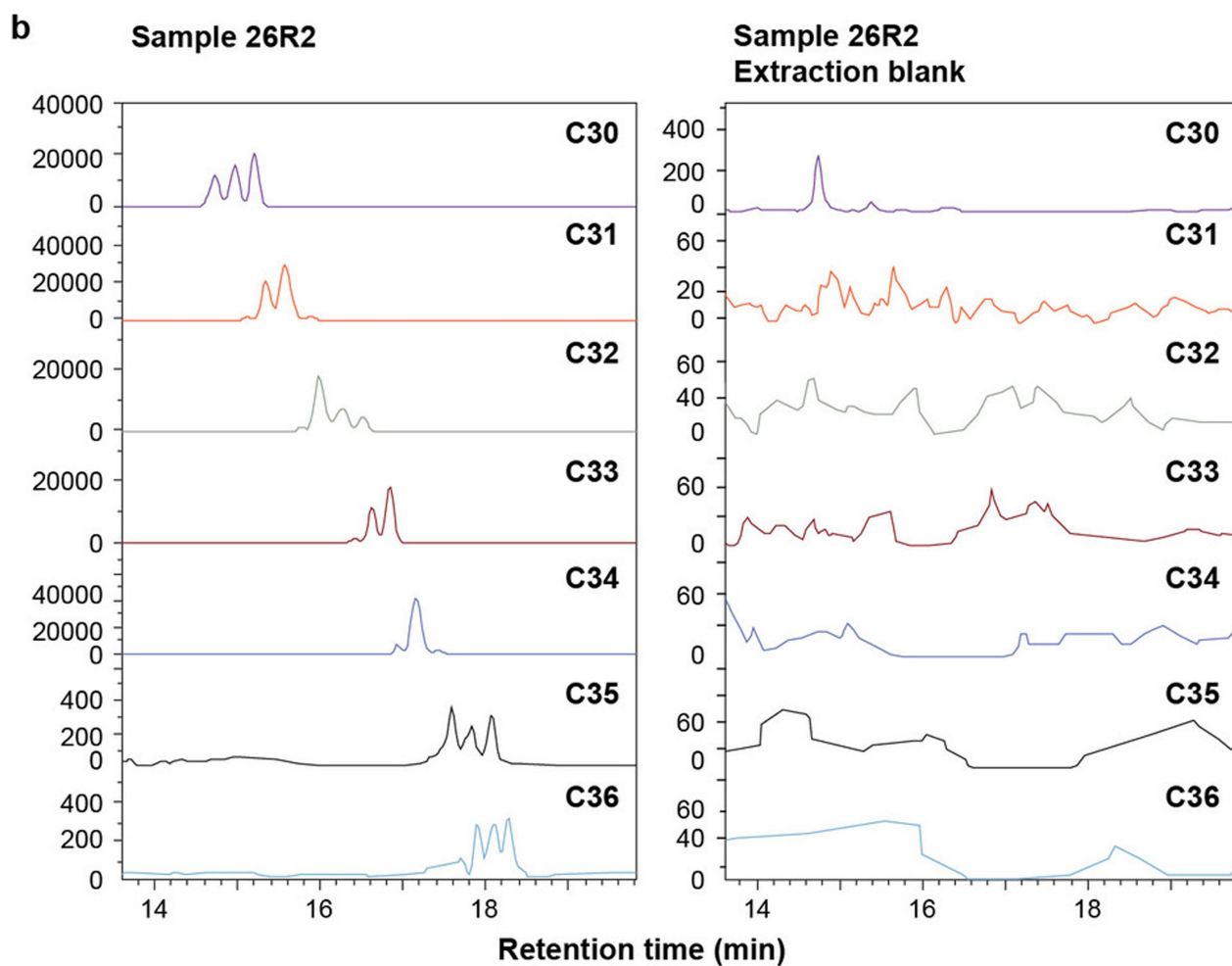
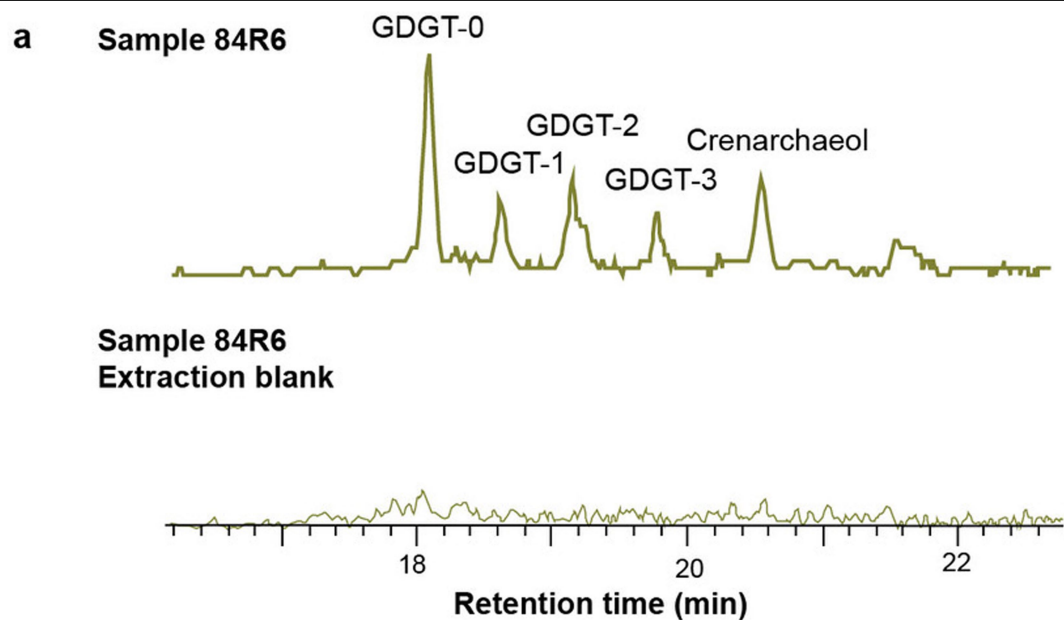
is given below the iTAG composition for each sample. The presence (black bars) or absence (white bars) indicates the detection or non-detection of genes associated with the processes and activities listed in that sample. A detailed discussion of metabolic pathways is provided in the Supplementary Information. Annotations at higher taxonomic resolution are available upon request from the corresponding author.



Extended Data Fig. 4 | See next page for caption.

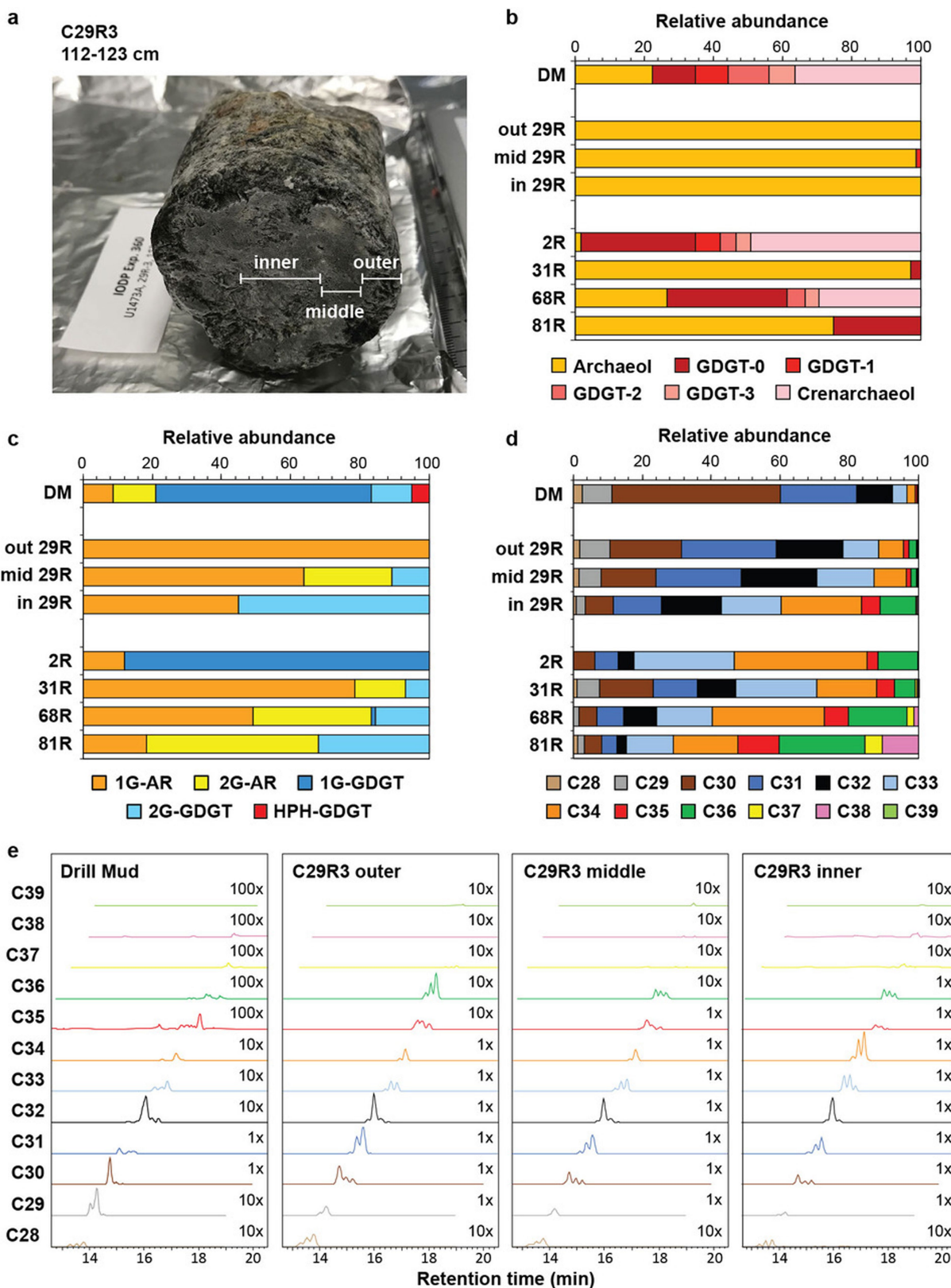
Extended Data Fig. 4 | Non-metric multidimensional scaling and clustering analyses of detected prokaryotic OTUs and transcripts. a, Non-metric multidimensional scaling analysis performed on the Jaccard distance matrix (Ward clustering using hclust; see Methods) of prokaryotic OTU presence or absence data for 11 biologically independent samples spanning the depth of hole U1473A showing only environmental vectors supported by $P < 0.05$. P values were generated as $P = N + 1/n + 1$ after a goodness-of-fit statistic, which is the squared correlation coefficient, and was calculated on 999 random permutations of the data using the vegan package of R. **b,** Clustering analysis of presence or absence prokaryotic OTU data for 11 samples spanning the depth of hole U1473A. Hierarchical clustering dendrogram based on a Jaccard

distance matrix of the presence or absence data for 99% OTUs from 11 IODP Expedition 360 samples. Jaccard similarity values of > 0.8 , calculated with the clusterboot function in R, suggest a stable cluster. Shading highlights several samples that share certain OTUs. From left to right (1–4): 1, samples share SAR11 clade II and SAR406; 2, samples share Nitrosopumilaceae; 3, samples share SAR11 clade II; and 4, samples both contain the lowest OTU counts. **c,** Clustering analysis of curated transcripts for 11 samples within the functional categories presented. Clustering analysis was based on log + 1-transformed FPKM values and the Ward method, and a distance matrix was constructed using the Manhattan method and the pvclust function in R.



Extended Data Fig. 5 | Lipid sample blanks. **a**, Representative UHPLC-APCI-MS chromatogram, showing the detection of core GDGTs in sample 84R6 and the sample extraction blank. **b**, Representative multiple-reaction-monitoring

HPLC-ESI-MS chromatograms for selected bacterial DEG lipids with summed chain lengths from C30 to C36 in sample 26R2 and the sample extraction blank. Lipid data are from single measurements owing to sample constraints.



Extended Data Fig. 6 | Drilling mud contamination control during lipid analysis. **a**, Depiction of outer, middle and inner sections subsampled from whole round core (WRC) sample 29R3 (259.03 mbsf) in order to test the influence of drilling mud (DM) on the lipid biomarker composition. **b–d**, Composition of archaeal core lipids (**b**), IPLs (**c**) and bacterial DEGs (**d**) in the drilling mud, subsampled WRC sample 29R3 and selected rock samples

analysed in this study. **e**, Extracted ion chromatograms of targeted DEG lipids from C28 to C39 summed carbon chain lengths in the drilling mud and subsampled WRC sample 29R3, showing the differences in relative abundance and isomer composition among these samples. Lipid data are from single measurements owing to sample constraints. 2R1, 10.7 mbsf; 31R1, 274.6 mbsf; 68R2, 619.6 mbsf; 81R2, 714.9 mbsf. Source Data are available online.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

For Raman analysis of thin section samples we used the LabSpec 6® Spectroscopy Suite Software (https://www.horiba.com/en_en/products/detail/action/show/Product/labspec-6-spectroscopy-suite-software-1843/) for data acquisition and background subtraction of acquired spectra. The processed spectra were then compared with the BIO-RAD KnowItAll® Raman Spectral Library databases for identification of organic substances and with the RRUFF (TM) Project Raman database for identification of minerals (Lafuente et al. (2015) (Lafuente B, Downs R T, Yang H, Stone N (2015) The power of databases: the RRUFF project. In: Highlights in Mineralogical Crystallography, T Armbruster and R M Danisi, eds. Berlin, Germany, W. De Gruyter, pp 1-30). For lipid analyses, Sciex Analyst 1.6.3 and Sciex MultiQuant 3.0.3 (AB Sciex LP, Concord, Canada) were used for triple quadrupole mass spectral data acquisition and data processing. Bruker Compass 1.9 and Bruker Data Analysis Version 4.4 (Bruker Daltonics, Bremen, Germany) were used for quadrupole time-of flight lipid data acquisition and processing.

Data analysis

The software R was used with the pvclust package for nMDS and clustering analysis of OTU data. Trimmomatic v. 0.32 was used to trim adapter sequences in transcript data, FastQC v. 0.11.7 and FASTX-toolkit v. 0.014 were used to quality check and trim paired reads in transcriptome libraries, Trinity v. 2.4.0 was used for assembly of transcriptomes, Bowtie v. 2.3.4.1 was used to align reads to our assembly and for de novo assembly of reads in our libraries, Bowtie aligner was used to align reads to assembled contigs, RSEM was used to estimate expression levels of reads in our libraries, TMM was used to perform cross-sample normalization and to generate a TMM-normalized expression matrix. The Trinotate suite (TransDecoder v. 3.0.1) was used to identify coding regions, BLASTx and BLASTp against UniProt, Swissprot (release 2018_02) and RefSeq nr were used to look for homologies between our reads and databases. HMMER v. 3.1b2 was used to identify conserved domains by searching against the Pfam v. 31.0 database. SignalP v. 4.1 and TMHMM 2.0c were used to predict signal peptides and transmembrane domains, RNAMMER v.1.2 was used to identify rRNA homologies. BLAST against the KEGG, COG, SEED and MetaCyc databases using MetaPathways v. 2.0 was used to interpret functions and pathways active in our transcriptome data. For our iTAG data Raw sequence reads were processed through Trim Galore [http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/], FLASH (ccb.jhu.edu/software/FLASH/) and FASTX Toolkit [http://hannonlab.cshl.edu/fastx_toolkit/] for trimming and removal of low quality/short reads. Quality filtering included requiring a minimum average quality of 25 and rejection of paired reads less

than 250 nucleotides. Operational Taxonomic Unit (OTU) clusters were constructed at 99% similarity with the script pick_otus.py within the Quantitative Insights Into Microbial Ecology (QIIME) v.1.9.1 software and 'uclust'. The OTU data table was transformed to a presence/absence table and the Jaccard method was used to generate a distance matrix using the dist.binary() function in the R package ade4. A hierarchical clustering dendrogram was created using hclust() and the stability of the clusters was evaluated using the clusterboot() function in the fpc package in R with 500 iterations.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

iTAG data for the 11 samples as well as 13 different negative controls (drilling muds and fluids, seawater, kit controls) are deposited in GenBank under BioProject PRJNA497074. Results are presented at ~phylum level, however taxonomic assignments at finer resolution are available from the authors upon request. Raw reads for transcriptome data have also been deposited in GenBank SRA under BioProject PRJNA497074. Assemblies for curated portions of the data presented are available upon request to the authors.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Samples available for microbiological studies were selected by the Co-Chief Scientists of Expedition 360. They were limited to ~10-20cm long whole round rock sections from recovered cores. Replication was not possible because there was a single hole drilled for all studies conducted during Expedition 360, and only the 10-20 cm sections of 10m cores were allocated to microbiologists. Each of the individual 10-20cm core sections was processed to remove exterior rock and remaining material was subdivided for all analyses presented here.
Data exclusions	Outside of subtraction of molecular data matching any contamination controls (as noted in methods), no data were excluded. All data collected from analyzed samples is presented.
Replication	Cell counts were replicated by processing two replicate 1ml tubes of material from each core sample. Alkaline phosphatase activity measurements were conducted on triplicate technical replicates of material for those analyses and results were consistent. Methane production measurements are single measurements due to limited sample availability, except for the artificial seawater blank (n=3), for which the mean value is presented. All attempts at replication were successful. Sufficient material was not available for replicating other analyses. Lipid biomarker measurements were single measurements due to limited sample material. Raman spectral analysis presented in Figure 3a was made at two spots on this particular feature of interest, 10 independent times per spot, with similar results. In total, we performed >2000 analyses (with shorter integration times) on similar features in the same thin section sample with similar results.
Randomization	Eleven rock samples were selected from our collection for analyses presented here, purely on visual observations of the rock samples, prioritizing samples from the full depth of the hole drilled, and samples with evident veins and signs of alteration.
Blinding	Blinding was not applicable to this study of rock samples and associated control samples beyond the fact that all samples were number coded and treated identically throughout preparation and analyses.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involvement in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging