

## Microbial residents of the Atlantis Massif's shallow serpentinite subsurface

Shahzad Motamedi,<sup>a,#</sup> Beth N. Orcutt,<sup>b</sup> Gretchen L. Früh-Green,<sup>c</sup> Katrina I. Twing,<sup>a\*</sup>  
H. Lizethe Pendleton<sup>a</sup>, and William J. Brazelton<sup>a</sup>

5

<sup>a</sup> School of Biological Sciences, University of Utah, Salt Lake City, Utah, USA

<sup>b</sup> Bigelow Laboratory for Ocean Sciences, East Boothbay, Maine, USA

<sup>c</sup> Institute of Geochemistry and Petrology, ETH Zürich, Zürich, Switzerland

10 Running Head: **Microbial residents of the Atlantis Massif**

#Address correspondence to Shahzad Motamedi, [shahzad.motamedi@utah.edu](mailto:shahzad.motamedi@utah.edu)

\*Present address: Department of Biology, Tennessee Technological University,  
Cookeville, Tennessee, USA

15

Word Count Abstract: 245

Word Count Importance: 106

Word Count Text: 5,055

20

## Abstract

The Atlantis Massif rises 4,000 m above the seafloor near the Mid-Atlantic Ridge and consists of rocks uplifted from Earth's upper mantle. Exposure of the mantle rocks to  
25 seawater leads to their alteration into serpentinites. These aqueous geochemical reactions, collectively known as the process of serpentinization, are exothermic and are associated with the release of hydrogen gas (H<sub>2</sub>), methane (CH<sub>4</sub>), and small organic molecules. The biological consequences of this flux of energy and organic compounds from the Atlantis Massif were explored by International Ocean Discovery  
30 Program (IODP) Expedition 357, which used seabed drills to collect continuous sequences of shallow (<16 meters below seafloor) marine serpentinites and mafic assemblages. Here, we report the first census of microbial diversity in samples of the drill cores, as measured by environmental 16S rRNA gene amplicon sequencing. The problem of contamination of subsurface samples was a primary concern during  
35 all stages of this project, starting from the initial study design, continuing to the collection of samples from the seafloor, handling the samples shipboard and in the lab, preparing the samples for DNA extraction, and analyzing the DNA sequence data. To distinguish endemic microbial taxa of serpentinite subsurface rocks from seawater residents and other potential contaminants, the distributions of individual  
40 16S rRNA gene sequences among all samples were evaluated, taking into consideration both presence/absence as well as relative abundances. Our results highlight a few candidate residents of the shallow serpentinite subsurface, including uncultured representatives of the Thermoplasmata, Acidobacteria, Acidimicrobiia, and Chloroflexi.

45

## Importance

International Ocean Discovery Program Expedition 357: “Serpentinization and Life” utilized seabed drills for the first time to collect rocks from the oceanic crust. The recovered rock cores represent the shallow serpentinite subsurface of the Atlantis Massif, where reactions between uplifted mantle rocks and water, collectively known as serpentinization, produce environmental conditions that can stimulate biological activity and are thought to be analogous to environments that were prevalent on the early Earth and perhaps other planets. The methodology and results of this project have implications for life detection experiments, including sample return missions, and provide the first window into the diversity of microbial communities inhabiting subseafloor serpentinites.

## Introduction

Subsurface environments may host the majority of microbial life on Earth (1, 2), but current estimates of subsurface biomass lack certainty due to a paucity of observations. Recent studies, including major ocean drilling expeditions, have greatly improved the quantification and characterization of microbial populations in deep marine sediments (3–6). Seafloor crustal rocks are also expected to be an important microbial habitat, in part because a significant fraction of the oceanic crust supports hydrothermal circulation of seawater that can bring nutrients and energy into subseafloor ecosystems (7–9). The mineral composition and geological history of the host rocks determine the local environmental conditions experienced by subseafloor microbes (10–13). Ultramafic habitats for life may have unique

attributes, as these rock types are unstable in the presence of seawater and undergo  
70 a series of serpentinization reactions that can support chemolithotrophy (14–16). A  
few microbial diversity surveys of mafic basaltic and gabbroic seafloor rocks have  
been conducted (17–19), but no studies have characterized the microbial diversity of  
ultramafic rocks.

75 International Ocean Discovery Program (IODP) Expedition 357 addressed this  
knowledge gap by targeting the Atlantis Massif, a submarine mountain located in the  
north Atlantic Ocean on the western flank of the Mid-Atlantic Ridge (20, 21). The  
massif is approximately 16 km across and rises 4,267 m from the seafloor. It is  
composed of variable amounts of ultramafic rocks uplifted from the upper mantle  
80 along a major fault and altered into serpentinites through the geochemical process of  
serpentinization (22, 23). Oxidation of iron minerals in the serpentinizing rocks  
results in the release of hydrogen gas and hydroxyl ions, contributing to the  
formation of extremely reducing, high pH fluids (24, 25). The high hydrogen  
concentrations that develop during serpentinization are conducive to the formation of  
85 methane and other organic compounds from the reduction of inorganic carbon (26–  
28). Circulation of organic-rich fluids through the serpentinites of the Atlantis Massif  
may support an active subseafloor ecosystem (29).

Most of our current knowledge of the biological implications of marine  
90 serpentinization comes from the exploration of the Lost City hydrothermal field, a  
collection of carbonate chimneys near the summit of the Atlantis Massif where high  
pH fluids exit the seafloor (22, 25, 30, 31). The chimneys are covered in  
mucilaginous biofilms formed by bacteria and archaea that are fueled by the high

concentrations of hydrogen, formate, and methane in the venting fluids (32–35). The  
95 high density of the microbial biofilm communities is likely enabled by the mixing of  
warm, anoxic hydrothermal fluids with oxygenated ambient seawater, enabling a  
wide range of metabolic strategies. In contrast, the much more voluminous, rocky  
subsurface habitats underlying the Lost City chimneys and throughout the Atlantis  
Massif probably experience much more diffuse hydrothermal circulation and perhaps  
100 limited exposure to ambient seawater. Our knowledge of this seafloor  
environment, however, is extremely limited (19).

The work reported here focused on two fundamental questions regarding  
serpentinization and life: How can we distinguish endemic microbial communities of  
105 low biomass seafloor rocks from other environments? Which microbes inhabit  
marine serpentinites? To address these questions, we conducted a cultivation-  
independent census of microbial diversity in the shallow (<16 mbsf) rocks collected  
during IODP Expedition 357 (21) and compared these to the communities found in  
seawater and potential sources of contamination.

110

Laboratory contamination of DNA sequencing datasets is a well-known problem (36–  
39), especially for low biomass subsurface samples (19, 40–42), that must be  
addressed by a combination of experimental and computational methods. No matter  
how cleanly DNA is prepared, contaminant sequences can appear for a myriad of  
115 reasons (37)(38). Sheik *et al.* (39) recently reviewed practices for identifying  
contaminants in DNA sequence datasets, highlighting the importance of collecting  
and sequencing control samples representing potential sources of environmental  
and laboratory contamination. These studies have highlighted the need to

complement careful laboratory protocols with computational methods that can  
120 distinguish true residents from multiple potential environmental and laboratory  
contaminants (43–45).

Here we report our methodological and computational strategies for minimizing and  
identifying contaminants in DNA sequencing datasets from extremely low-biomass  
125 seafloor serpentinites collected during IODP Expedition 357. Sequencing DNA from  
serpentinites required the development of a novel DNA extraction and purification  
protocol that recovers highly pure DNA without the use of any commercial kits or  
phenol. We then identified environmental and laboratory contaminants in the  
resulting DNA sequencing data with a computational workflow that considers the  
130 relative abundances of individual sequences among potential contamination sources.  
Finally, we describe the archaeal and bacterial taxa that are potential residents of  
seafloor serpentinites.

## Results

### 135 DNA Yields from Rocks

DNA was extracted from at least one rock core sample collected from all sites  
(**Figure 1, Text S1**) except M0073 (**Table 1, Data set S1**). DNA yields from lysates  
(post-lysis, pre-purification) ranged from ~100 to almost 500 ng of DNA per gram of  
rock sample. DNA yields from serpentinite and non-serpentinite rocks (including  
140 metagabbro, carbonate sand, basalt breccia and metadolerite) were similar (**Table 1  
and Data set S2**). These values may be inflated by the many impurities in the  
lysates, which had not been purified and were visibly cloudy. No trends in DNA yield  
with respect to mineralogy or locations of the boreholes were apparent. During initial

145 testing of the protocol, DNA quantification was conducted at each step of the  
extraction to verify that DNA levels never increased, which would indicate laboratory  
contamination. For the precious serpentinite samples, however, DNA was quantified  
only after lysis, washing, and the final purification. DNA was below the detection limit  
for all final, purified preparations, indicating a substantial loss of DNA from the  
lysates but also minimal laboratory contamination.

150

### DNA Sequencing Results

155 Despite the below-detection levels of purified DNA, a total of 1,125,191 sequencing  
read pairs of the 16S rRNA gene were obtained from all rock samples (50 samples  
including sequencing duplicates). Here we focus only on the DNA sequencing results  
from the 15 samples characterized as serpentinites, but the DNA sequencing data  
from all samples are included in our SRA BioProject PRJNA575221. At least 2,780  
read pairs were obtained from each sample, and >10,000 read pairs were obtained  
from five of the serpentinite rock samples (**Table 1**). We identified 2,063 amplicon  
sequence variants (ASVs) among all rock samples, 664 of which were found only in  
160 serpentinites. We obtained 27,524,256 read pairs from the water samples (76  
samples of surface, shallow, and deep including sequencing duplicates) (**Data set  
S3**), from which we identified 15,142 ASVs. Six samples of lab air yielded 66,846  
read pairs and 293 ASVs (**Data set S4**), despite below-detection levels of purified  
DNA.

165

The proportional abundances of all 17,081 unique ASVs among all rock, water, and  
air samples were used to calculate the community dissimilarity between each pair of  
samples. The results are visualized in the non-metric multidimensional scaling

(NMDS) plot in **Figure 2**, which shows a clear split between water and rock samples.

170 The bacterial communities of both serpentinite (orange points) and non-serpentinite (yellow points) rock samples are highly variable and are not consistently distinct from each other. The bacterial compositions of lab air samples (green points) overlap with those of the rock samples in this visualization.

### 175 Evaluation of Contamination Sources

To evaluate contamination in the serpentinites, we used SourceTracker2 (45) to estimate the proportion of each rock sample's DNA sequences that could be attributed to contamination from water or air. These results show that very few of the DNA sequences in serpentinites can be traced back to seawater (**Figure 3**). The number of sequences attributed to air was highly variable among the serpentinites, from nearly zero to >50% of total sequences. Most serpentinite samples were dominated by sequences that could not be attributed to a single source by SourceTracker2. These ASVs are designated as having an "unknown" source and may represent true inhabitants of the serpentinites or may be derived from another source that was not sampled in this study. Based on these results, serpentinite samples with >15% of sequences attributed to lab air were excluded from further analyses (0AMRd014A, 031C, 033A, 033C, 034A, 036A, 045B, 067C, 071A, and 073A).

190 SourceTracker2 computes probabilistic estimates of how the whole-community composition of samples relate to each other, but it is not designed to identify the source of each taxon. For example, multiple ASVs in sample 0AMRd057A occurred in both air and water samples, which some approaches would consider to be



evidence of contamination, even though SourceTracker2 estimated that <1% of  
195 sequences in this sample could be assigned to water or air sources (**Figure 3**).  
Therefore, we explored other methods for the identification of individual ASVs as  
contaminants.

#### Identification of Contaminants from Air and Seawater

200 Of the 293 ASVs detected in air samples, 187 also occurred in serpentinites and  
were therefore designated as contaminants (**Figure 4A**). Thus, all remaining ASVs  
were absent in air and were only detected in water and serpentinite samples. Next,  
two approaches, simple overlap (SO) and differential abundance (DA), were used in  
parallel to distinguish the serpentinite ASVs from the water ASVs. In the SO  
205 approach (**Figure 4B**), the ASVs shared between water and serpentinite samples  
were removed from the dataset regardless of their abundances in water or  
serpentinites (i.e. only considering presence/absence). Thus, the results from the SO  
approach include only those ASVs that are exclusively found in serpentinite samples  
(664 ASVs).

210

In the differential abundance (DA) approach (**Figure 4C**), the sequence counts (i.e.  
the number of merged paired reads) of each ASV shared between water and  
serpentinite samples were compared. ASVs that were significantly more abundant in  
water samples compared to serpentinites were identified and then removed from the  
215 dataset (**Table S3**). In this case, ASVs that are highly abundant in serpentinites and  
very rare (but still present) in water are not necessarily identified as contaminants, as  
they were in the SO approach. Three edgeR comparisons were performed between  
serpentinite samples and each depth category of water samples (serpentinites vs.

surface, serpentinites vs. shallow, and serpentinites vs. deep). The ASVs that were  
220 more abundant in each group of water samples (one example in **Figure 4C**) were  
designated as contaminants from water and excluded from further analyses. The  
ASVs identified as water contaminants and detected in serpentinites are listed in  
**Data set S5**. The DA approach results in a final list of 684 ASVs putatively  
considered to be true inhabitants of serpentinites (**Data set S6**), including 20  
225 additional ASVs compared to the SO approach. Seven of these 20 ASVs were  
present at >100 counts in serpentinite samples and less than a total of 53 counts in  
all water samples (**Data set S7**).

#### *Taxonomic Composition of Serpentinites*

230 To conservatively report a shorter list of likely serpentinite inhabitants, the following  
analyses and visualizations only include ASVs with at least 100 total counts among  
all samples. Taxonomic classifications of the putative serpentinite ASVs from the SO  
approach are summarized in **Figure 5**. For clarity, the taxonomic classifications of  
only the top 50 most abundant ASVs are shown, and more information about these  
235 50 ASVs is provided in **Data set S7**. The taxonomic compositions of serpentinite  
samples are variable, consistent with the spread of serpentinites shown in the NMDS  
plot (**Figure 2**). No significant differences with respect to sample handling  
(flamed/unflamed, shaved/unshaved, washed/unwashed) could be detected. Indeed,  
very few similarities between any two samples were apparent.

240 Sample 0AMRd057A (**Figure 6**) provided the deepest sequencing dataset (98,462  
merged paired reads among those ASVs with at least 100 counts). Most of the  
taxonomic diversity in this sample was represented by rare ASVs (i.e. those with

lower abundance than the top 50 shown in **Figure 5**). The most abundant ASV was  
245 classified as family Thermoplasmatales (phylum Euryarchaeota) and has 99%  
identity to a clone (NCBI accession number: AB825685, **Data set S7**) from a deep-  
sea sediment core in the Okinawa Trough (46). Two of the top ASVs in this sample  
were classified as two different classes of Chloroflexi (S085 and Anaerolineae), each  
of which was unique to sample 0AMRd057A and was most similar to clones from  
250 marine sediments (AM998115; HQ721355). Another top ASV unique to this sample  
was classified as Acidimicrobia clade OM1 (phylum Actinobacteria), which was  
nearly identical to a clone from a marine endolithic community (47). Two ASVs in  
0AMRd057A were classified as Acidobacteria Subgroup 21 and were most similar to  
sequences from marine sediments, including a clone (KY977768) from altered rocks  
255 of the Mariana subduction zone, where serpentinization also occurs (48).

The deepest sequencing datasets of the remaining samples were obtained from  
0AMRd045A, 0AMRd072A, 0AMRd075A, and 0AMRd076A, which were collected  
from the sites closest to the Lost City chimneys (boreholes M0072B and M0076B;  
260 **Figure 1**). Each of these samples were dominated by Clostridia and  
Alphaproteobacteria (**Figure 5**), and most of these ASVs match sequences  
associated with animal digestive tracts or soil. An exception is a *Rhodobacter* ASV  
unique to 0AMRd045A that matches several sequences from lakes (e.g. MF994002).  
An abundant Sphingomonadaceae ASV in 0AMRd075A matches a clone from  
265 groundwater near the Hanford Site, WA (KT431076). A *Dechloromonas* ASV unique  
to 0AMRd076A (Site M0076) is identical to a clone from an ocean drilling expedition  
to the South Chamorro Seamount (LC279322), which is also a site of seafloor

serpentinization (49). However, the genus *Dechloromonas* has also been detected on human skin (50).

270

Sample 0AMRd030A is dominated by two ASVs classified as genera *Acidovorax* (class Betaproteobacteria) and *Cobetia* (class Gammaproteobacteria), respectively. The *Acidovorax* ASV matches sequences from soil (e.g. CP021359) and an anoxic lake (KY515689)(51). *Cobetia marina* (previously *Halomonas marina*) is a ubiquitous marine bacterium (52) and also a model organism for the study of biofouling due to its ability to grow as a biofilm on a variety of surfaces (53).

275

Other abundant ASVs shown in **Figure 5** and listed in **Data set S7** include those classified as Acidobacteria, Actinobacteria, Betaproteobacteria, Gemmatimonadetes, and Ktedonobacteria, and these sequences were generally most similar to environmental sequences from soils or freshwater. One possible exception is an ASV that was abundant in sample 0AMRd031A and was classified as genus *Modestobacter*, a soil bacterium, but it was also identical to another species in the same family, *Klenkia marina*, which was isolated from marine sediment (NR\_156810).

280

285

## Discussion

### *Serpentinite samples have minimal environmental contamination*

Sequencing of environmental DNA from seafloor rocks is challenging due to their very low biomass and their abundance of minerals (e.g. phyllosilicates) that inhibit enzymatic reactions and bind DNA (54–56). The extremely low biomass of seafloor rocks also makes them susceptible to environmental and laboratory contamination.

290

Many precautions were taken during handling of the rock core samples to minimize, eliminate, and detect contamination, and our results agree with previous reports that  
295 these efforts were generally successful (20, 57, 58).

To assess environmental contamination of the cores, perfluoromethylcyclohexane (PFC) was injected into the seabed drilling fluids during coring. Orcutt *et al.* (57) reported that the interior zones of rocks whose outer surfaces had been shaved off  
300 with a sterile rock saw had greatly reduced levels of PFC. Only one sample in our study (0AMRd073), whose exterior had been shaved, contained detectable levels of PFC. Two additional samples (0AMRd030, 0AMRd036), which were too rubbly to be shaved with the rock saw and were instead washed with ultrapure water, contained detectable PFC. In general, however, washing rubbly samples with ultrapure water  
305 was largely effective in reducing the level of PFC, and presumably of any other surface contamination (57). The PFC tracer, however, was not designed to track any seawater contamination that might have occurred after drilling or any contamination that might have resulted from handling the rock cores shipboard or in the laboratory.

310 We assessed environmental contamination of the serpentinites by sequencing DNA from samples of seawater collected above the boreholes. The contribution of seawater DNA sequences into the serpentinites was minimal (**Figure 3**), further demonstrating that the sampling-handling precautions on the ship and at the Kochi Core Center were successful in eliminating environmental contamination that must  
315 have occurred at some level during core recovery on the seafloor. In addition, our results agree with those of Orcutt *et al.* (57) that there was no noticeable difference

in contamination levels between rocks whose outer surfaces had been subjected to flame sterilization compared to those that were not flamed.

320 *Serpentinite samples have minimal environmental DNA*

Our yields of ~200 ng of DNA per g of rock sample in the crude lysates (**Table 1**) were much higher than expected from the previously reported numbers of cells for these serpentinites (20), all of which were only slightly greater than the quantification limit of ~10 cells cm<sup>-3</sup>. Such a cell density would predict 20 fg of DNA cm<sup>-3</sup> (assuming  
325 2 fg of DNA per cell), which is approximately 7 orders of magnitude less DNA than observed. Therefore, the total environmental DNA must have been primarily extracellular. Nearly all of the DNA was lost during purification and fell below our detection limit, so much of it was probably highly fragmented. Indeed, neither the filtration units used for washing nor the electrophoretic purification method are  
330 expected to recover DNA below 600 bp.

Our lowest detection limit of purified DNA (which was concentrated from many parallel extractions of 30-40 grams of rock per sample) was 0.05 ng per gram of rock, which is still 10,000 times greater than that expected from the cell numbers.  
335 However, amplicon sequencing of 16S rRNA genes in the environmental DNA was moderately successful, most likely thanks to the elimination of PCR inhibitors by the electrophoretic SCODA purification. Because of the exceedingly low levels of input DNA, the probability of contamination during purification, PCR, and sequencing was very high. Indeed, laboratory contamination was evident in the serpentinites, as  
340 represented by their relatively high proportions of DNA sequences from samples of our laboratory's air (**Figure 3**).

### Identification of laboratory contaminants

We identified laboratory contaminants as amplicon sequence variants (ASVs) that  
345 occurred in samples of both air and serpentinites. This simple overlap (SO) approach  
is commonly used in environmental microbiology studies to remove all sequences (or  
taxonomic categories) that are shared between any sources of contamination and  
the sample of interest (39). This approach can be useful for removing ASVs from  
obvious sources of contamination. An important assumption of the SO approach is  
350 that the direction of contamination is from the suspected contamination source into  
the sample. However, this assumption is often untested. For example, if a low-DNA  
control sample is handled in close proximity (e.g. the same 96-well plate) to a high-  
DNA environmental sample, it may be more likely for "contamination" to occur from  
the environmental sample into the control sample. Subsequent use of the SO  
355 approach would erroneously identify an abundant member of the environmental  
sample as a laboratory contaminant.

We considered the possibility of "reverse contamination" from our core samples into  
our air samples to be unlikely for multiple reasons. First, most of the processing of  
360 the rock samples, including sawing, crushing, washing, and subsampling, occurred  
at the Kochi Core Center in Japan, and our samples of air were collected in our  
laboratory at the University of Utah. Second, our SourceTracker2 results implicated  
air as a major source of DNA into the serpentinites (**Figure 3**). Third, nearly all of the  
taxa detected in the air samples are typically associated with humans (e.g.  
365 *Pseudomonas*, *Staphylococcus*, *Acinetobacter*). Our list of laboratory air taxa (**Data  
set S4**) may be useful for identifying contaminants in future studies, as a

complement to previously reported lists of contaminants from laboratory reagents (37–39).

370 *Identification of environmental contaminants*

Identification of environmental contaminants from seawater into the serpentinites was more complex because of the greater possibility of mixing between seawater and rocks in both directions. The Lost City chimneys provide dramatic evidence of the flux of biological material from the rocky subsurface into ambient seawater (25),  
375 and flux from the boreholes in this study was also evident by the release of bubbles during drilling (20). Deep seawater near the Atlantis Massif cannot be assumed to be completely free of subsurface inhabitants, and samples of seawater at shallower depths or locations far removed from the Atlantis Massif would not be accurate representations of the environmental contamination that the serpentinite samples  
380 experienced during recovery on the seafloor. Therefore, the relative abundances of taxa in samples of deep seawater and serpentinites should be informative because true subsurface organisms that are mixed into ambient seawater should be immediately diluted and present in seawater only at very low abundances.

385 In such situations of environmental mixing, the SO approach can erroneously identify contaminants because it ignores abundance information. We identified at least one such example in our study by comparing the results of the SO approach with the differential abundance (DA) approach. An ASV classified as Acidobacteria Subgroup 21 occurred 264 times in a serpentinite sample (0AMRd057A; 0.3% of total  
390 sequence counts in that sample) from borehole M0075 and also occurred a total of 12 times among five deep seawater samples collected from above boreholes M0068,



M0072, and M0073 (**Data set S6**). The ASV was identical to several sequences from marine sediments as well as one from altered rocks of the Mariana subduction zone (KY977768) and another from seafloor lavas near the Loi'hi seamount (EU491098)(17).  
395

This was the most striking example in our study, but there were 19 additional ASVs flagged as contaminants by the SO approach but not by the DA approach. This relatively small difference (664 compared to 684 total ASVs in serpentinites) is probably a consequence of the success in minimizing and eliminating environmental contamination during handling of the rock cores, as discussed above and demonstrated by the minimal level of contamination from seawater shown in **Figure 3**. Differences between the SO and DA approaches are likely to be greater in studies where environmental contamination cannot be controlled so carefully.  
400

405

#### Identification of contaminants by taxonomy

After removal of DNA sequences identified as contaminants from seawater and air, the most abundant serpentinite ASVs (**Data set S6**) still contained many taxa that seemed unlikely to be true inhabitants of the subseafloor. Some are typically associated with animal digestive tracts (e.g. *Ruminiclostridium*), some are found in soil (e.g. various Acidobacteria), and others have been previously identified as reagent contaminants (e.g. *Micrococcus*)(37, 38). The sources of these likely contaminants could not be determined in this study, but the extraction and sequencing reagents are strong candidates, even though we took great care in preparation of our own reagents at every step and did not use any commercial extraction or purification kits. Also, we only sampled and sequenced DNA from the  
410  
415

air in our own laboratory at the University of Utah and did not sample the air inside the research vessel, the Kochi Core Center, or the DNA sequencing core facility at Michigan State University.

420

Identification of contaminants at the level of taxonomic category is risky, however, because organisms that are capable of persisting in the extreme environments of mostly sterile laboratory supplies, may also be capable of inhabiting the extreme environments of mostly sterile subsurface rocks (59). For example, a relatively abundant serpentinite ASV that was absent in all seawater and air samples is nevertheless a suspected contaminant because it was classified as *Acidovorax*, a ubiquitous soil bacterium and plant pathogen that has also been identified as a contaminant in laboratory reagents (38). The ability of *Acidovorax* to form biofilms and perform anaerobic iron oxidation (60), however, would be advantageous in the serpentinite subsurface. Because we are suspicious of these taxa but lack direct experimental evidence of a source other than serpentinites, we include these ASVs in our final tables but flag them as suspected contaminants (**Data set S7**).

430

#### *Putative inhabitants of the serpentinite subsurface*

435

Excluding all air and seawater ASVs and ignoring all suspicious taxa leaves us with a few putative inhabitants of subseafloor serpentinites (**Data set S7**). Many of these ASVs were unique to a single serpentinite sample shown in **Figure 6A**. This serpentinite was recovered from the most eastern site (borehole M0075A) at a depth of ~60 cm below the seafloor, and it had extensive talc-amphibole alteration, suggesting that it was eroded from the central dome during uplift of the massif. The most abundant taxa in this sample included Euryarchaeota, Chloroflexi,

440

Actinobacteria, and Acidobacteria, all of which matched DNA sequences previously found in marine sediments and rocks. The most abundant ASV was classified as Thermoplasmatales (an order within the Euryarchaeota) and was nearly identical to  
445 a clone from hydrothermal sediment collected during IODP Expedition 331 to the Mid-Okinawa Trough (46). Most, but not all, of the characterized members of the Thermoplasmatales are thermophilic, and they are typically involved in sulfur cycling (61, 62). The Thermoplasmatales are generally known as acidophiles (63), which is curious since serpentinization is associated with high pH.

450

The most abundant taxa in serpentinites recovered from the central dome of the massif (boreholes M0069, M0072, and M0076) were more difficult to interpret. At least two probable subsurface taxa (*Sphingomonadaceae* and *Dechloromonas*) were detected in these serpentinites, but many of the other sequences from these  
455 samples are identical to environmental sequences from studies of animal digestive tracts or soils. Therefore, serpentinites from the central dome, which are considered to be most representative of the massif's basement, appear to be more susceptible to undetected contamination, perhaps due to extremely low DNA yields from these samples.

460

Recently, Quéméneur *et al.* (58) reported enrichment cultures obtained from IODP Expedition 357 rock cores, including two serpentinite samples. One of these was a section of a serpentinite core (357-71C-5R-CC) that was adjacent to a sample included in the present study (0AMRd036A). Our parallel sample was heavily  
465 contaminated with air DNA (**Figure 3**) and did not contain any of the dominant taxa detected by Quéméneur *et al.* Many of the taxa detected in their cultivation

experiments were also present in our results, but only one genus reported by Quéméneur *et al.* (*Sphingomonas*) was included in our final list of candidate serpentinite taxa (**Data set S7**). Most of the successful enrichments reported by  
470 Quéméneur *et al.* were obtained from a carbonate-hosted basalt breccia, which was not included in our study.

In 2005, IODP expedition 304/305 recovered gabbroic rocks from the central dome of the Atlantis Massif, but no serpentinites (**Figure 1**). A few phylotypes of  
475 proteobacteria were detected in the gabbroic rock cores by sequencing of environmental 16S rRNA genes (19). None of these matched the candidate serpentinite taxa reported here. However, most of the functional genes detected by Mason *et al.* (19) with the GeoChip microarray, including genes associated with metal toxicity, carbon degradation, denitrification, and carbon fixation, are consistent  
480 with the taxa detected in our serpentinite samples.

## Conclusions

This study provides the first sequences of environmental DNA from subseafloor serpentinites, enabled by the high recovery of rock cores by IODP Expedition 357 to  
485 the Atlantis Massif (20). The extremely low biomass of the serpentinites presented multiple challenges, necessitating the development of a novel DNA extraction and purification protocol. We developed strategies to increase the yield of DNA appropriate for PCR amplification, while eliminating many potential sources of laboratory contamination. We hope that our methodology, as well as our  
490 identification of laboratory contaminants derived from dust particles, will prove to be useful to other researchers studying extremely low-biomass environments. We were

able to identify candidate residents of the serpentinite seafloor by employing a series of contamination-detection procedures, including evaluations of whole-sample compositions and of individual sequences. Our results highlight the importance of a  
495 multi-faceted approach to contamination control that features sampling of potential contamination sources as a complement to microbiologically clean sample-processing protocols. Multiple statistical and bioinformatic tools were used to identify contaminant DNA sequences in a careful process that did not unnecessarily ignore useful information, such as the relative abundances of individual DNA sequences in  
500 potential contamination sources. No computational tool can ever prove that an environmental DNA sequence is not a contaminant, and future studies will investigate candidate seafloor organisms with experimental studies.

## Materials and Methods

505 Detailed descriptions of samples and sample-processing protocols are available in the Supplemental Methods (**Text S1**). DNA was purified from rock samples with SCODA (synchronous coefficient of drag attenuation) technology implemented with the Aurora purification system (Boreal Genomics, Vancouver, British Columbia, Canada). 16S rRNA gene amplicon sequencing was conducted by the Michigan  
510 State University genomics core facility on all of the samples (rock, water, and lab air). All samples were sequenced twice (i.e. sequencing replicates). The V4 region of the bacterial 16S rRNA gene was amplified with dual-indexed Illumina fusion primers (515F/806R) as described elsewhere (64). 16S rRNA gene amplicon sequences were processed with cutadapt v. 1.15 (65) and DADA2 v. 1.10.1 (66). Taxonomic  
515 classification of all ASVs was performed with DADA2 using the SILVA reference alignment (SSURefv132) and taxonomy outline (67). Raw counts were converted to

proportions to normalize for variations in sequencing depth among samples. The non-metric multi-dimensional scaling (NMDS) plot was generated from a table of ASV proportional abundances across all sample categories (rocks, water, lab air) using the distance, ordinate, and plot\_ordination commands in the R package phyloseq v.1.26.1 (68). The major sources of contamination and the level of contamination in each sample were estimated with SourceTracker2, version 2.0.1 (45). Differential abundance was tested with the R package edgeR v. 3.24.3 (69) as recommended elsewhere (70). We used edgeR to contrast the total read counts of ASVs in serpentinite rock samples compared to three groups of water samples (surface, shallow, deep). ASVs that were absent in all serpentinites and ASVs with low variance ( $<1e-6$ ) were excluded from the comparisons. The output of the three edgeR tests (serpentinite samples compared to each of the three groups of water samples) was three lists of ASVs with significant differential abundances (false discovery rate (FDR)  $< 0.05$ ) in serpentinite samples or water samples. The final list of ASVs was created by deleting ASVs with greater abundances in any of the categories of water samples (as determined by the edgeR tests) from the original list of ASVs from which all air ASVs had already been removed. Finally, rare ASVs (those that did not have  $\geq 100$  counts in a single sample) were excluded from the final results, merely as a conservative abundance filter for reporting a final list of ASVs expected to be present in serpentinite rocks. No comparisons of diversity were attempted after removing rare ASVs.

### Data Availability

All unprocessed DNA sequence data are publicly available at the NCBI SRA under BioProject PRJNA575221. All supplementary data and protocols are available at

<https://github.com/Brazelton-Lab/Atlantis-Massif-2015>. All custom software and scripts are available at <https://github.com/Brazelton-Lab>.

## 545 **Acknowledgements**

This research used data and samples provided by the International Ocean Discovery Program (IODP) Expedition 357 supported by the European Consortium for Ocean Research Drilling (ECORD) and implemented by the ECORD Science Operator (ESO). The shipboard sample-processing protocols were executed by ESO  
550 Expedition Project Managers Carol Cotterill and Sophie Green, ESO Operations Superintendent David Smith, the crews of the R.R.S. *James Cook* and the MeBo and RD2 seabed drills, and shipboard scientists Susan Lang, Marvin Lilley, Yuki Morono, Marianne Quéméneur, and Matthew Schrenk. We are indebted to Fumio Inagaki for welcoming our use of sampling equipment at the JAMSTEC Kochi Core Center  
555 (KCC), to Nan Xiao at KCC for coordinating our sample processing and operation of core sectioning equipment, to Christopher Thornton, Katherine Hickock, and Susan Lang for manually processing the rock samples at KCC, and to ESO staff Ursula Röhl and David McInroy for helping to coordinate sample transfer and tracking. Christopher Thornton, Alex Hyer, Emily Dart, and Julia McGonigle provided  
560 laboratory and computational assistance. We also thank Jackie Goordial for comments on an earlier draft. The Expedition 357 core description team helpfully provided descriptions of the core lithologies. Funding for this work was provided by the NASA Astrobiology Institute Rock-Powered Life team, the U.S. National Science Foundation-funded US Science Support Program (subawards to WJB and BNO), the  
565 Deep Carbon Observatory funded by the Alfred P. Sloan Foundation (via subaward to BNO from the Marine Biological Laboratory), and by the Swiss National Science

Foundation (SNSF) project No. 200021\_163187 (to GFG) and SNSF contributions to the Swiss IODP.

570

## References

1. Whitman WB, Coleman DC, Wiebe WJ. 1998. Prokaryotes: The unseen majority. *Proc Natl Acad Sci* 95:6578–6583.
- 575 2. Bar-On YM, Phillips R, Milo R. 2018. The biomass distribution on Earth. *Proc Natl Acad Sci* 115:6506–6511.
3. Lipp JS, Morono Y, Inagaki F, Hinrichs K-U. 2008. Significant contribution of Archaea to extant biomass in marine subsurface sediments. *Nature* 454:991–994.
4. Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D’Hondt S. 2012. Global distribution of  
580 microbial abundance and biomass in subseafloor sediment. *Proc Natl Acad Sci* 109:16213–16216.
5. Orcutt BN, Sylvan JB, Knab NJ, Edwards KJ. 2011. Microbial Ecology of the Dark Ocean above, at, and below the Seafloor. *Microbiol Mol Biol Rev* 75:361–422.
6. Takai K, Mottl MJ, Nielsen SH. 2012. IODP Expedition 331: strong and expansive  
585 subseafloor hydrothermal activities in the Okinawa Trough. *Sci Drill* 13:19–27.
7. Schrenk MO, Huber JA, Edwards KJ. 2010. Microbial provinces in the subseafloor. *Annu Rev Mar Sci* 2:279–304.



8. Mottl MJ, Wheat CG. 1994. Hydrothermal circulation through mid-ocean ridge flanks: Fluxes of heat and magnesium. *Geochim Cosmochim Acta* 58:2225–2237.
- 590 9. Fisher AT, Becker K. 2000. Channelized fluid flow in oceanic crust reconciles heat-flow and permeability data. *Nature* 403:71–74.
10. Santelli CM, Edgcomb VP, Bach W, Edwards KJ. 2009. The diversity and abundance of bacteria inhabiting seafloor lavas positively correlate with rock alteration. *Environ Microbiol* 11:86–98.
- 595 11. Singer E, Chong LS, Heidelberg JF, Edwards KJ. 2015. Similar Microbial Communities Found on Two Distant Seafloor Basalts. *Front Microbiol* 6.
12. Edwards KJ, Wheat CG, Sylvan JB. 2011. Under the sea: microbial life in volcanic oceanic crust. *Nat Rev Microbiol* 9:703–712.
13. Smith AR, Fisk MR, Thurber AR, Flores GE, Mason OU, Popa R, Colwell FS. 2017. Deep  
600 Crustal Communities of the Juan de Fuca Ridge Are Governed by Mineralogy. *Geomicrobiol J* 34:147–156.
14. Amend JP, Rogers KL, Shock EL, Gurrieri S, Inguaggiato S. 2003. Energetics of chemolithoautotrophy in the hydrothermal system of Vulcano Island, southern Italy. *Geobiology* 1:37–58.
- 605 15. Lollar BS. 2004. Life’s Chemical Kitchen. *Science* 304:972–973.
16. Okland I, Huang S, Dahle H, Thorseth IH, Pedersen RB. 2012. Low temperature alteration of serpentinized ultramafic rock and implications for microbial life. *Chem Geol* 318–319:75–87.

17. Santelli CM, Orcutt BN, Banning E, Bach W, Moyer CL, Sogin ML, Staudigel H, Edwards  
610 KJ. 2008. Abundance and diversity of microbial life in ocean crust. *Nature* 453:653.
18. Mason OU, Meo-Savoie CAD, Nostrand JDV, Zhou J, Fisk MR, Giovannoni SJ. 2009.  
Prokaryotic diversity, distribution, and insights into their role in biogeochemical cycling  
in marine basalts. *ISME J* 3:231–242.
19. Mason OU, Nakagawa T, Rosner M, Nostrand JDV, Zhou J, Maruyama A, Fisk MR,  
615 Giovannoni SJ. 2010. First Investigation of the Microbiology of the Deepest Layer of  
Ocean Crust. *PLOS ONE* 5:e15399.
20. Früh-Green GL, Orcutt BN, Rouméjon S, Lilley MD, Morono Y, Cotterill C, Green S,  
Escartin J, John BE, McCaig AM, Cannat M, Ménez B, Schwarzenbach EM, Williams MJ,  
Morgan S, Lang SQ, Schrenk MO, Brazelton WJ, Akizawa N, Boschi C, Dunkel KG,  
620 Quéméneur M, Whattam SA, Mayhew L, Harris M, Bayrakci G, Behrmann J-H, Herrero-  
Bervera E, Hesse K, Liu H-Q, Ratnayake AS, Twing K, Weis D, Zhao R, Bilenker L. 2018.  
Magmatism, serpentinization and life: Insights through drilling the Atlantis Massif  
(IODP Expedition 357). *Lithos* 323:137–155.
21. Früh-Green GL, Orcutt BN, Green SL, Cotterill C, Morgan S, Akizawa N, Bayrakci G,  
625 Behrmann JH, Boschi C, Brazleton WJ. 2017. Expedition 357 Summary. *Proc Int Ocean  
Discov Program* 357.
22. Kelley DS, Karson JA, Blackman DK, Früh-Green GL, Butterfield DA, Lilley MD, Olson EJ,  
Schrenk MO, Roe KK, Lebon GT. 2001. An off-axis hydrothermal vent field near the  
Mid-Atlantic Ridge at 30 N. *Nature* 412:145.

- 630 23. Blackman DK, Karson JA, Kelley DS, Cann JR, Früh-Green GL, Gee JS, Hurst SD, John BE,  
Morgan J, Nooner SL. 2002. Geology of the Atlantis Massif (Mid-Atlantic Ridge, 30 N):  
Implications for the evolution of an ultramafic oceanic core complex. *Mar Geophys Res*  
23:443–469.
24. Allen DE, Seyfried WE. 2004. Serpentinization and heat generation: constraints from  
635 Lost City and Rainbow hydrothermal systems<sup>11</sup>Associate editor: J. C. Alt. *Geochim*  
*Cosmochim Acta* 68:1347–1354.
25. Kelley DS, Karson JA, Früh-Green GL, Yoerger DR, Shank TM, Butterfield DA, Hayes JM,  
Schrenk MO, Olson EJ, Proskurowski G, Jakuba M, Bradley A, Larson B, Ludwig K,  
Glickson D, Buckman K, Bradley AS, Brazelton WJ, Roe K, Elend MJ, Delacour A,  
640 Bernasconi SM, Lilley MD, Baross JA, Summons RE, Sylva SP. 2005. A Serpentinite-  
Hosted Ecosystem: The Lost City Hydrothermal Field. *Science* 307:1428–1434.
26. McCollom TM, Seewald JS. 2007. Abiotic synthesis of organic compounds in deep-sea  
hydrothermal environments. *Chem Rev* 107:382–401.
27. Holm NG, Charlou JL. 2001. Initial indications of abiotic formation of hydrocarbons in the  
645 Rainbow ultramafic hydrothermal system, Mid-Atlantic Ridge. *Earth Planet Sci Lett*  
191:1–8.
28. Konn C, Charlou J-L, Donval J-P, Holm NG, Dehairs F, Bouillon S. 2009. Hydrocarbons and  
oxidized organic compounds in hydrothermal fluids from Rainbow and Lost City  
ultramafic-hosted vents. *Chem Geol* 258:299–314.

- 650 29. Lang SQ, Früh-Green GL, Bernasconi SM, Lilley MD, Proskurowski G, Méhay S, Butterfield DA. 2012. Microbial utilization of abiogenic carbon and hydrogen in a serpentinite-hosted system. *Geochim Cosmochim Acta* 92:82–99.
30. Früh-Green GL, Kelley DS, Bernasconi SM, Karson JA, Ludwig KA, Butterfield DA, Boschi C, Proskurowski G. 2003. 30,000 years of hydrothermal activity at the Lost City vent  
655 field. *Science* 301:495–498.
31. Ludwig KA, Kelley DS, Butterfield DA, Nelson BK, Früh-Green G. 2006. Formation and evolution of carbonate chimneys at the Lost City Hydrothermal Field. *Geochim Cosmochim Acta* 70:3625–3645.
32. Schrenk MO, Kelley DS, Bolton SA, Baross JA. 2004. Low archaeal diversity linked to  
660 subsurface geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ Microbiol* 6:1086–1095.
33. Brazelton WJ, Schrenk MO, Kelley DS, Baross JA. 2006. Methane-and sulfur-metabolizing microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl Env Microbiol* 72:6257–6270.
- 665 34. Brazelton WJ, Mehta MP, Kelley DS, Baross JA. 2011. Physiological differentiation within a single-species biofilm fueled by serpentinization. *MBio* 2:e00127–11.
35. Lang SQ, Früh-Green GL, Bernasconi SM, Brazelton WJ, Schrenk MO, McGonigle JM. 2018. Deeply-sourced formate fuels sulfate reducers but not methanogens at Lost City hydrothermal field. *Sci Rep* 8:755.

- 670 36. Tanner MA, Goebel BM, Dojka MA, Pace NR. 1998. Specific Ribosomal DNA Sequences  
from Diverse Environmental Settings Correlate with Experimental Contaminants. *Appl  
Environ Microbiol* 64:3110–3113.
37. Barton HA, Taylor NM, Lubbers BR, Pemberton AC. 2006. DNA extraction from low-  
biomass carbonate rock: An improved method with reduced contamination and the  
675 low-biomass contaminant database. *J Microbiol Methods* 66:21–31.
38. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J,  
Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically  
impact sequence-based microbiome analyses. *BMC Biol* 12:87.
39. Sheik CS, Reese BK, Twing KI, Sylvan JB, Grim SL, Schrenk MO, Sogin ML, Colwell FS.  
680 2018. Identification and removal of contaminant sequences from ribosomal gene  
databases: lessons from the census of deep life. *Front Microbiol* 9:840.
40. Labonté JM, Lever MA, Edwards KJ, Orcutt BN. 2017. Influence of Igneous Basement on  
Deep Sediment Microbial Diversity on the Eastern Juan de Fuca Ridge Flank. *Front  
Microbiol* 8.
- 685 41. Santelli CM, Banerjee N, Bach W, Edwards KJ. 2010. Tapping the Subsurface Ocean Crust  
Biosphere: Low Biomass and Drilling-Related Contamination Calls for Improved Quality  
Controls. *Geomicrobiol J* 27:158–169.
42. Lever MA, Alperin M, Engelen B, Inagaki F, Nakagawa S, Steinsbu BO, Teske A. 2006.  
Trends in basalt and sediment core contamination during IODP Expedition 301.  
690 *Geomicrobiol J* 23:517–530.

43. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6:226.
44. Inagaki F, Hinrichs K-U, Kubo Y, Bowles MW, Heuer VB, Hong W-L, Hoshino T, Ijiri A, Imachi H, Ito M. 2015. Exploring deep microbial life in coal-bearing sediment down to ~2.5 km below the ocean floor. *Science* 349:420–424.
45. Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R, Kelley ST. 2011. Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* 8:761–763.
46. Yanagawa K, Nunoura T, McAllister S, Hirai M, Breuker A, Brandt L, House C, Moyer CL, Birrien J-L, Aoike K, Sunamura M, Urabe T, Mottl M, Takai K. 2013. The first microbiological contamination assessment by deep-sea drilling and coring by the D/V Chikyu at the Iheya North hydrothermal field in the Mid-Okinawa Trough (IODP Expedition 331). *Front Microbiol* 4.
47. Couradeau E, Roush D, Guida BS, Garcia-Pichel F. 2017. Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico). *Biogeosciences* 14:311–324.
48. Hyndman RD, Peacock SM. 2003. Serpentinization of the forearc mantle. *Earth Planet Sci Lett* 212:417–432.

- 710 49. Kawagucci S, Miyazaki J, Morono Y, Seewald JS, Wheat CG, Takai K. 2018. Cool, alkaline  
serpentinite formation fluid regime with scarce microbial habitability and possible  
abiotic synthesis beneath the South Chamorro Seamount. *Prog Earth Planet Sci* 5:74.
50. Grice EA, Kong HH, Renaud G, Young AC, Bouffard GG, Blakesley RW, Wolfsberg TG,  
Turner ML, Segre JA. 2008. A diversity profile of the human skin microbiota. *Genome*  
715 *Res* 18:1043–1050.
51. Schiff SL, Tsuji JM, Wu L, Venkiteswaran JJ, Molot LA, Elgood RJ, Paterson MJ, Neufeld  
JD. 2017. Millions of Boreal Shield Lakes can be used to Probe Archaean Ocean  
Biogeochemistry. *Sci Rep* 7:46708.
52. Arahal DR, Castillo AM, Ludwig W, Schleifer KH, Ventosa A. 2002. Proposal of *Cobetia*  
720 *marina* gen. nov., comb. nov., within the Family Halomonadaceae, to Include the  
Species *Halomonas marina*. *Syst Appl Microbiol* 25:207–211.
53. Mieszkin S, Martin-Tanchereau P, Callow ME, Callow JA. 2012. Effect of bacterial  
biofilms formed on fouling-release coatings from natural seawater and *Cobetia marina*,  
on the adhesion of two marine algae. *Biofouling* 28:953–968.
- 725 54. Direito SO, Marees A, Röling WF. 2012. Sensitive life detection strategies for low-  
biomass environments: optimizing extraction of nucleic acids adsorbing to terrestrial  
and Mars analogue minerals. *FEMS Microbiol Ecol* 81:111–123.
55. Saeki K, Sakai M. 2009. The influence of soil organic matter on DNA adsorptions on  
andosols. *Microbes Environ* 0904270086–0904270086.

- 730 56. Lever MA, Torti A, Eickenbusch P, Michaud AB, Šantl-Temkiv T, Jørgensen BB. 2015. A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. *Front Microbiol* 6:476.
57. Orcutt BN, Bergenthal M, Freudenthal T, Smith D, Lilley MD, Schnieders L, Green S, Früh-  
Green GL. 2017. Contamination tracer testing with seabed drills: IODP Expedition 357.  
735 *Sci Drill* 23:39–46.
58. Quéméneur M, Erauso G, Frouin E, Zeghal E, Vandecasteele C, Ollivier B, Tamburini C, Garel M, Ménez B, Postec A. 2019. Hydrostatic Pressure Helps to Cultivate an Original Anaerobic Bacterium From the Atlantis Massif Seafloor (IODP Expedition 357): *Petrocella atlantisensis* gen. nov. sp. nov. *Front Microbiol* 10.
- 740 59. Swanner ED, Nell RM, Templeton AS. 2011. *Ralstonia* species mediate Fe-oxidation in circumneutral, metal-rich subsurface fluids of Henderson mine, CO. *Chem Geol* 284:339–350.
60. Byrne-Bailey KG, Weber KA, Bose S, Knox T, Spanbauer TL, Chertkov O, Coates JD. 2010. Completed genome sequence of the anaerobic iron-oxidizing bacterium *Acidovorax*  
745 *ebreus* strain TPSY. *J Bacteriol* 192:1475–1476.
61. Arce-Rodríguez A, Puente-Sánchez F, Avendaño R, Martínez-Cruz M, de Moor JM, Pieper DH, Chavarría M. 2019. Thermoplasmatales and sulfur-oxidizing bacteria dominate the microbial community at the surface water of a CO<sub>2</sub>-rich hydrothermal spring located in Tenorio Volcano National Park, Costa Rica. *Extremophiles* 23:177–187.



- 750 62. Barton LL, Fardeau M-L, Fauque GD. 2014. Hydrogen sulfide: a toxic gas produced by  
dissimilatory sulfate and sulfur reduction and consumed by microbial oxidation, p.  
237–277. *In* The Metal-Driven Biogeochemistry of Gaseous Compounds in the  
Environment. Springer.
63. Yelton AP, Comolli LR, Justice NB, Castelle C, Deneff VJ, Thomas BC, Banfield JF. 2013.  
755 Comparative genomics in acid mine drainage biofilm communities reveals metabolic  
and structural differentiation of co-occurring archaea. *BMC Genomics* 14:485.
64. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a  
dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence  
data on the MiSeq Illumina sequencing platform. *Appl Env Microbiol* 79:5112–5120.
- 760 65. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput  
sequencing reads. *EMBnet.journal* 17:10–12.
66. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2:  
High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–  
583.
- 765 67. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: Accurate high-throughput multiple  
sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829.
68. McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive  
Analysis and Graphics of Microbiome Census Data. *PLOS ONE* 8:e61217.

69. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for  
770 differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–  
140.
70. McMurdie PJ, Holmes S. 2014. Waste Not, Want Not: Why Rarefying Microbiome Data Is  
Inadmissible. *PLOS Comput Biol* 10:e1003531.
71. Früh-Green GL, Orcutt BN, Green SL, Cotterill C, Morgan S, Akizawa N, Bayrakci G,  
775 Behrmann J-H, Boschi C, Brazleton WJ. 2017. Supplementary material for Volume 357  
expedition reports. *Proc Int Ocean Discov Program* 357.

## Figures and Tables

780

**Table 1:** DNA yield from each IODP Expedition 357 serpentinite sample after lysis, washing, and purification steps. The concentration of the final, purified DNA was below the stated detection limit for each sample. Numbers of 16S rRNA gene  
785 amplicon sequences are reported before and after quality filtering.

**Figure 1:** Multibeam bathymetry of the Atlantis Massif, with water depth in meters below sea level (mbsl) in color scale per legend. Drilled sites are shown as black circles. The yellow star is the location of the Lost City hydrothermal field. The  
790 distance between the most western site (M0073) and the most eastern site (M0075) is approximately 13 km (Adapted from Früh-Green, *et. al.* (20)).

**Figure 2:** A non-metric multidimensional scaling plot shows general patterns in the 16S rRNA gene microbial community compositions of all types of samples. Each data point represents the microbial community composition of a sample of seawater (deep, shallow, or surface), rock core (serpentinite versus non-serpentinite), or laboratory air. The distances between the data points represent the Morisita-Horn dissimilarities among the microbial communities.

**Figure 3:** Assessment of potential 16S rRNA gene library contamination in serpentinite samples with SourceTracker2. Each bar on this plot shows the percentage of ASVs from each source of contamination into each serpentinite sample. Green = lab air. Blue = seawater samples. Gray = undetermined.

**Figure 4:** Workflow for identifying DNA sequence contaminants. A) All ASVs detected in air samples were removed from the dataset. B) The Simple Overlap (SO) approach was used to remove all ASVs detected in water samples. C) The Differential Abundance (DA) approach was used to remove only those ASVs that were significantly more abundant in water samples compared to serpentinite samples. Differentially abundant ASVs are shown as red data points in the plot. The significance threshold was a false detection rate of 0.05. Only one of the comparisons is shown here (deep water samples vs. serpentinite samples). Shallow and surface water samples were compared in separate analyses not shown here.

**Figure 5:** Taxonomic summary of ASVs detected in serpentinite samples after contaminant removal. For clarity, only the top 50 ASVs among all samples (Data set S7) are shown here. N is the total ASV counts for each sample after excluding rare

ASVs. Samples 0AMRd076A, 075A, 072A, and 067A were collected from hole M0076B. Samples 0AMRd031A and 030A were collected from hole M0071B. No  
820 other samples in this figure were collected from the same hole.

**Figure 6:** A) Taxonomic summary of serpentinite sample 0AMRd057 (357-75A-1R-CC,0-4cm). The relative abundance of each taxonomic group is its proportion of 103,229 total sequence counts among the final 331 ASVs identified by the  
825 differential abundance (DA) approach. B) This sample was recovered from the core catcher section of the core from borehole M0075A. Scale is in millimeters (71).

### Supplemental files

830 **Data set S1:** Names and locations of all serpentinite and non-serpentinite rock samples.

**Data set S2:** DNA yield from non-serpentinite rock samples.

835 **Data set S3:** Names and locations of all seawater samples.

**Data set S4:** List of ASVs detected in samples of laboratory air.

**Data set S5:** List of ASVs detected in samples of seawater.

840

**Data set S6:** Final list of 684 ASVs in serpentinites determined by the DA method.

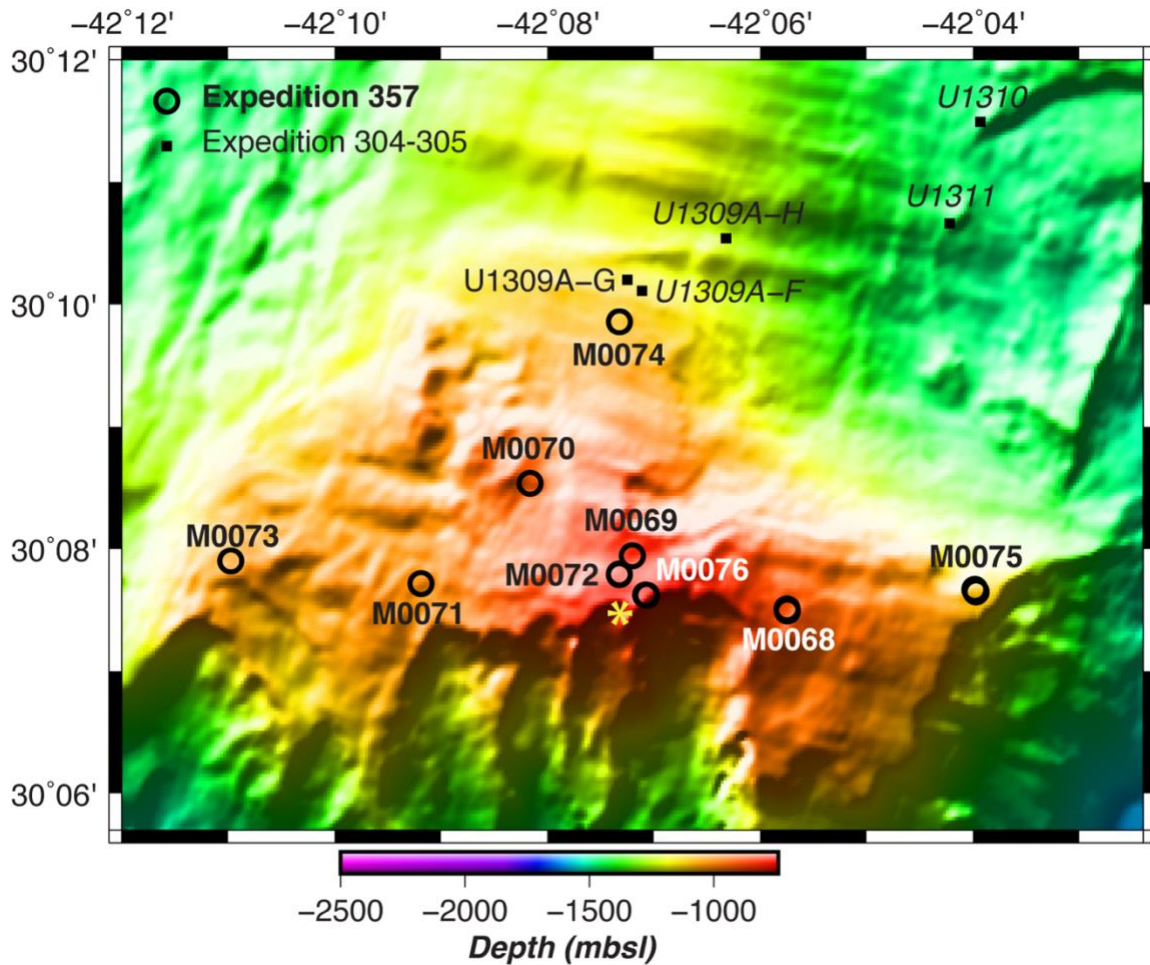
**Data set S7:** Top 50 most abundant ASVs in serpentinites (according to maximum proportional abundance per sample).

845

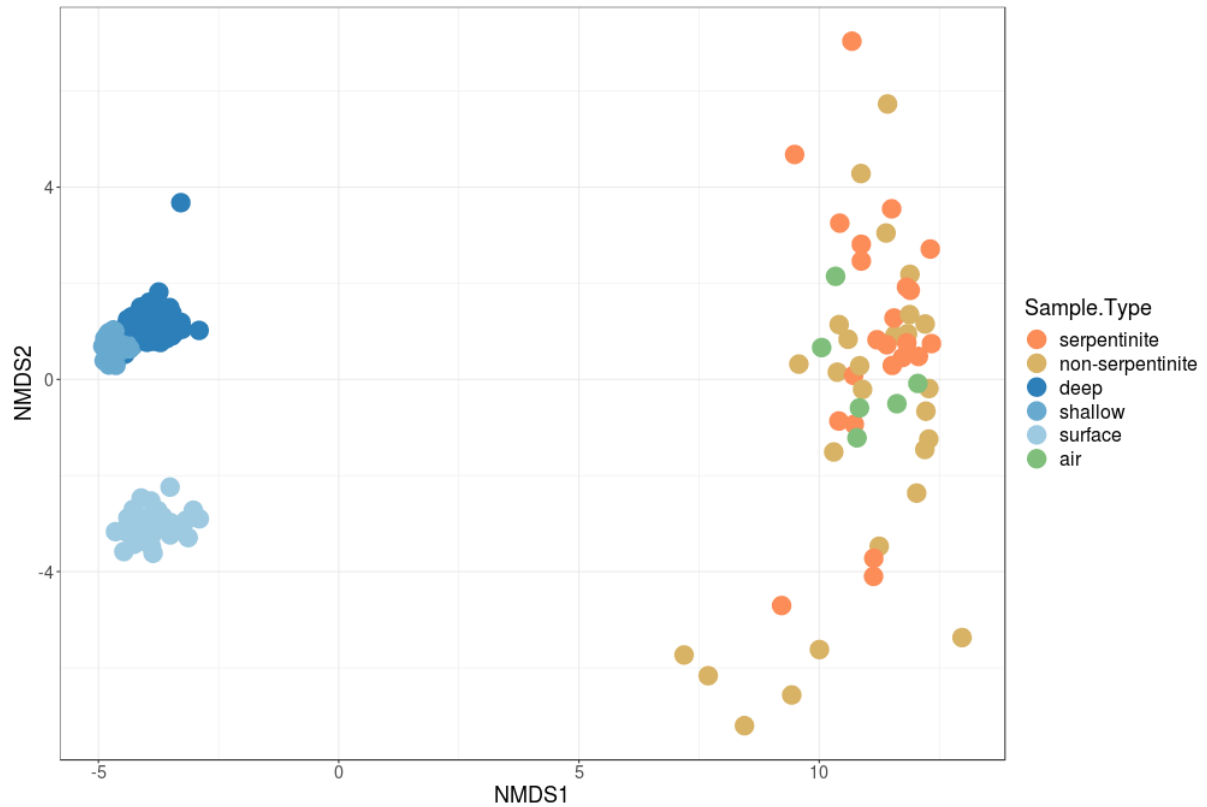
**Text S1:** Supplementary descriptions of samples and sample-processing methodology.

Sample ID	IODP ID	Extracted rock (g)	Lysate DNA (ng per g rock)	Post-washing DNA (ng per g of rock)	Purified DNA (ng per g rock)	Total sequences	Sequences after quality filtering
0AMRd005A	357-68B-8R-1,29-34cm	30	325	4.07	< 0.46	2985	1318
0AMRd014A	357-69A-7R-1,73-75cm	40.5	104	2.51	< 0.55	2928	1926
0AMRd030A	357-71B-2R-2,53-58cm	37	224	2.34	< 0.29	3893	2957
0AMRd031A	357-71B-2R-1,64-66cm	37.5	215	2.32	< 0.49	2780	2210
0AMRd033A	357-71C-2R-1,78-84cm	39	371	4.99	< 0.2	3920	3291
0AMRd034B	357-71C-2R-1,84-90cm	39	213	2.97	< 0.19	2901	2294
0AMRd036A	357-71C-5R-CC,5-10cm	38	248	2.03	< 0.21	4281	3442
0AMRd045A	357-72B-3R-1,34-36cm	38	473	15.2	< 0.2	10787	7225
0AMRd057A	357-75A-1R-CC,0-4cm	26	136	2	< 0.16	153,054	113,994
0AMRd067A	357-76B-5R-1,14-28cm	37	299	4.39	< 0.22	4699	2831
0AMRd071A	357-76B-7R-1,105-120cm	34.5	100	7.14	< 0.41	10,095	7742
0AMRd072A	357-76B-9R-1,34-41cm	37	110	2.64	< 0.49	122,361	82,246
0AMRd073A	357-76B-9R-1,43-52cm	47.4	149	4.71	< 0.82	5189	2079
0AMRd075A	357-76B-10R-1,91-111cm	38	300	1.49	< 0.52	12,082	9758
0AMRd076A	357-76B-10R-1,91-111cm	33	223	4.3	< 1.24	9399	7478

**Table 1.** DNA yield from each IODP Expedition 357 serpentinite sample after lysis, washing, and purification steps. The concentration of the final, purified DNA was below the stated detection limit for each sample. Numbers of 16S rRNA gene amplicon sequences are reported before and after quality filtering.

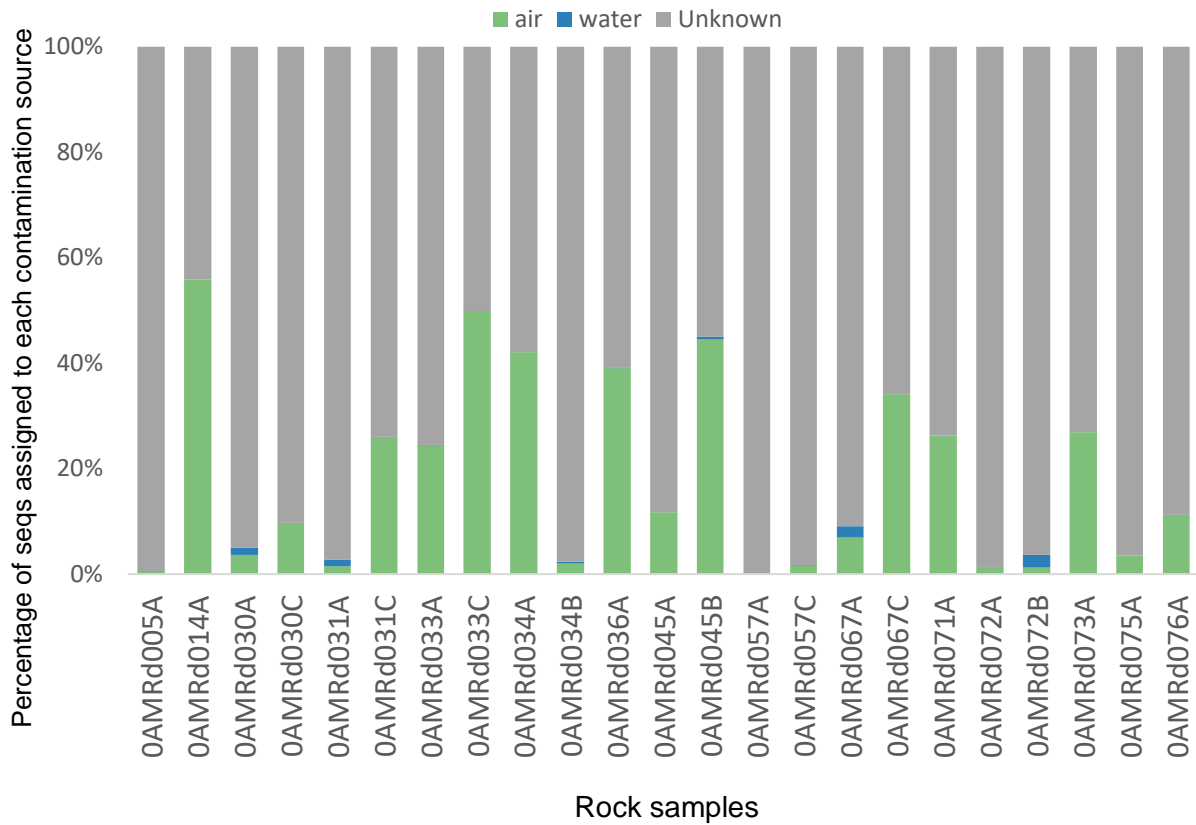


**Figure 1.** Multibeam bathymetry of the Atlantis Massif, with water depth in meters below sea level (mbsl) in color scale per legend. Drilled sites are shown as black circles. The yellow star is the location of the Lost City hydrothermal field. The distance between the most western site (M0073) and the most eastern site (M0075) is approximately 13 km (Adapted from Früh-Green, *et. al.* (20)).

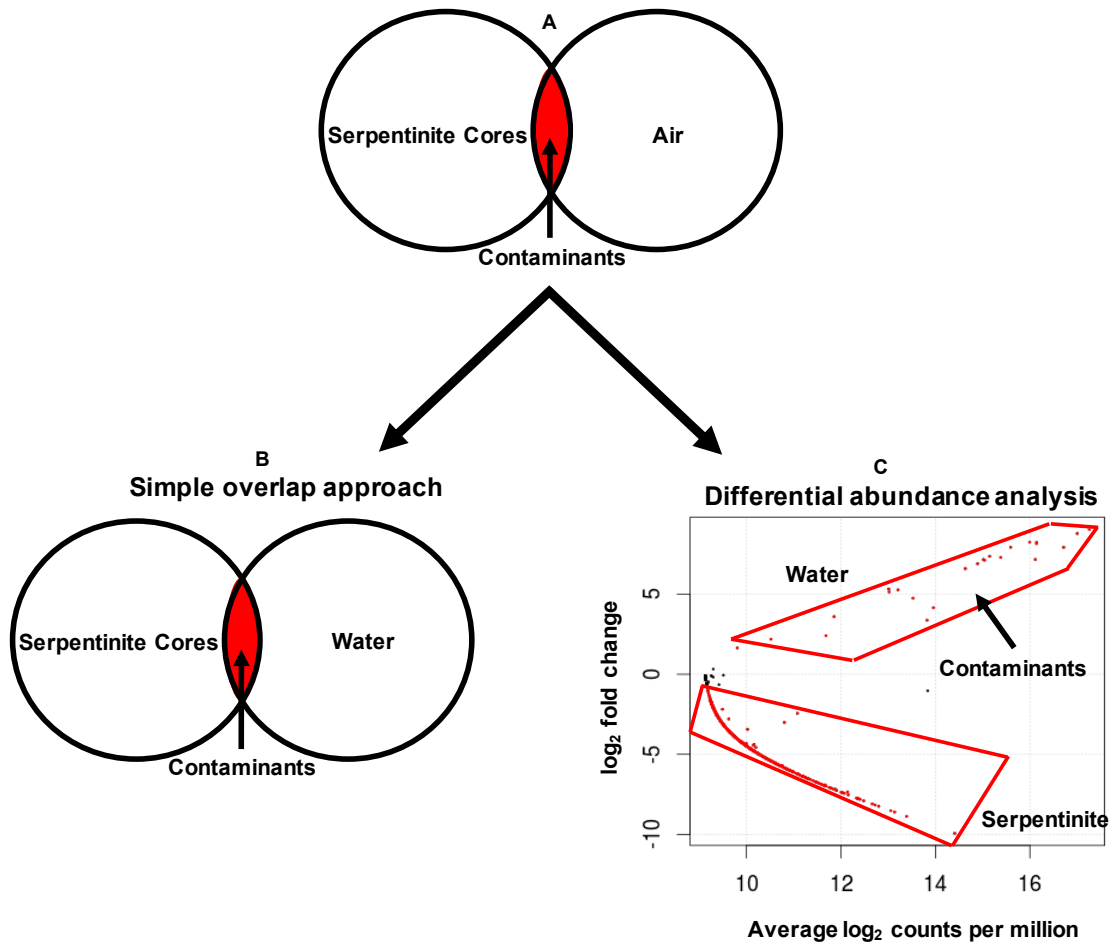


**Figure 2:** A non-metric multidimensional scaling plot shows general patterns in the 16S rRNA gene microbial community compositions of all types of samples. Each data point represents the microbial community composition of a sample of seawater (deep, shallow, or surface), rock core (serpentinite versus non-serpentinite), or laboratory air. The distances between the data points represent the Morisita-Horn dissimilarities among the microbial communities.

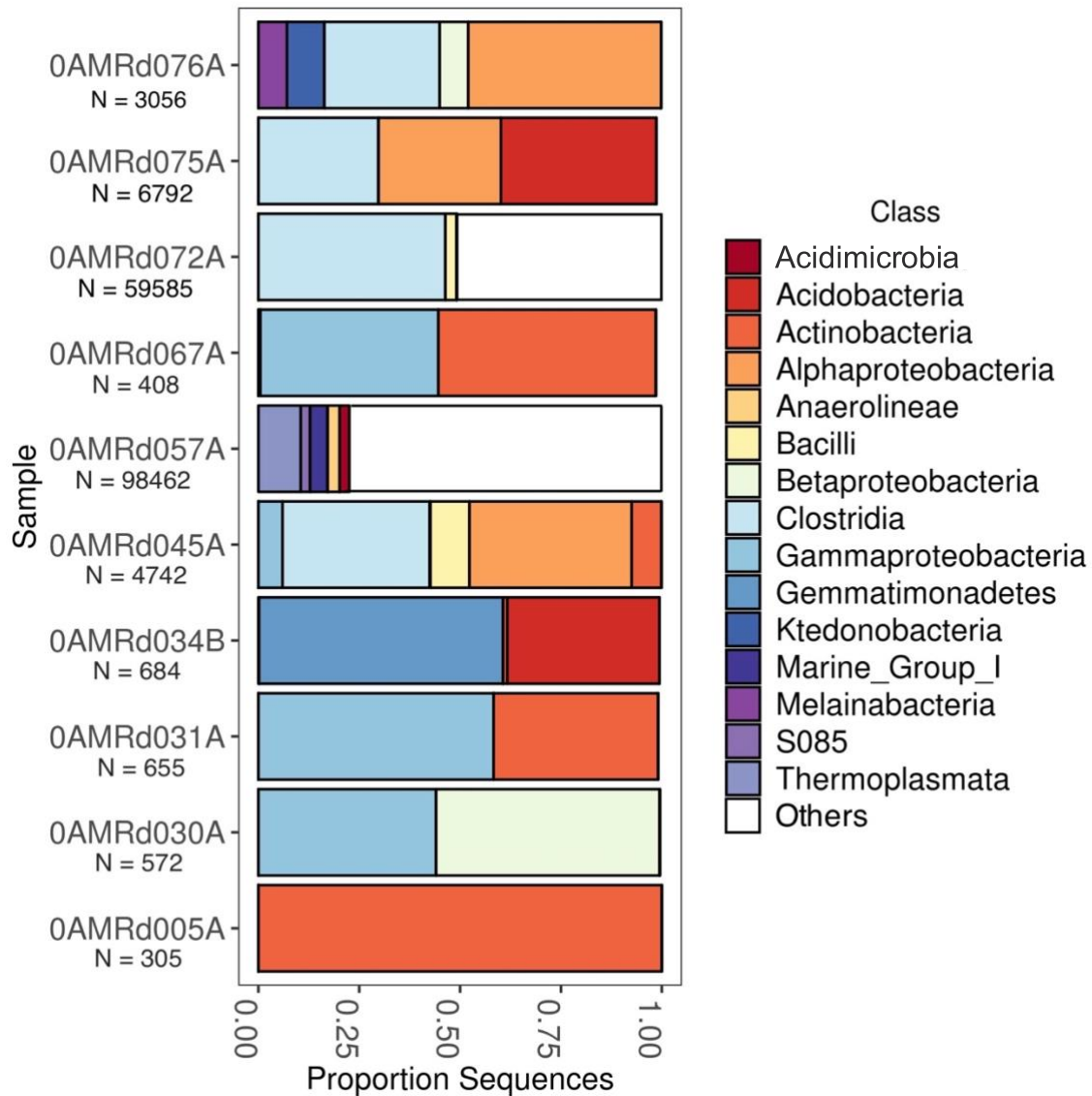




**Figure 3:** Assessment of potential 16S rRNA gene library contamination in serpentinite samples with SourceTracker2. Each bar on this plot shows the percentage of ASVs from each source of contamination into each serpentinite sample. Green = lab air. Blue = seawater samples. Gray = undetermined.



**Figure 4:** Workflow for identifying DNA sequence contaminants. A) All ASVs detected in air samples were removed from the dataset. B) The Simple Overlap (SO) approach was used to remove all ASVs detected in water samples. C) The Differential Abundance (DA) approach was used to remove only those ASVs that were significantly more abundant in water samples compared to serpentinite samples. Differentially abundant ASVs are shown as red data points in the plot. The significance threshold was a false detection rate of 0.05. Only one of the comparisons is shown here (deep water samples vs. serpentinite samples). Shallow and surface water samples were compared in separate analyses not shown here.



**Figure 5:** Taxonomic summary of ASVs detected in serpentinite samples after contaminant removal. For clarity, only the top 50 ASVs among all samples (Supplementary Table S8) are shown here. N is the total ASV counts for each sample after excluding rare ASVs. Samples 0AMRd076A, 075A, 072A, and 067A were collected from hole M0076B. Samples 0AMRd031A and 030A were collected from hole M0071B. No other samples in this figure were collected from the same hole.

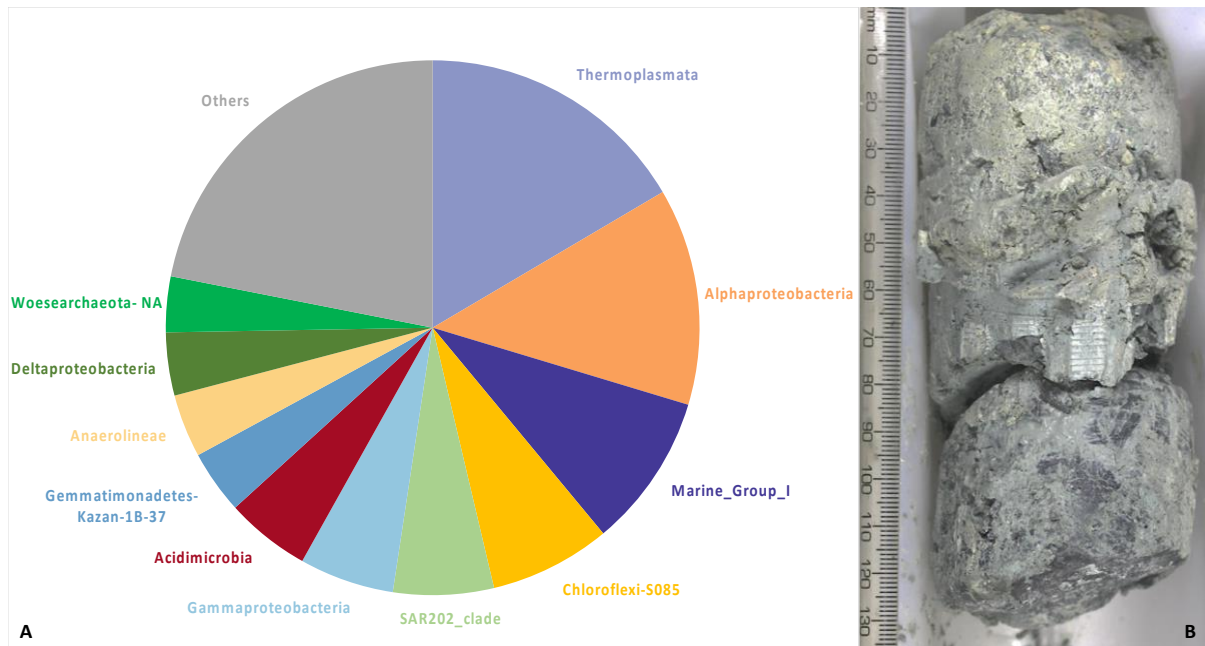


Figure 6: A) Taxonomic summary of serpentinite sample 0AMRd057 (357-75A-1R-CC,0-4cm). The relative abundance of each taxonomic group is its proportion of 103,229 total sequence counts among the final 331 ASVs identified by the differential abundance (DA) approach. B) This sample was recovered from the core catcher section of the core from borehole M0075A. Scale is in millimeters (71).