Complex subsurface hydrothermal fluid mixing at a submarine arc volcano supports distinct and highly diverse microbial communities

Anna-Louise Reyensbacha,1,2, Emily St. Johna,2, Jennifer Meneghinb, Gilberto E. Floress, Mircea Podar, Nina Dombrowskia, Anja Spangd, Stephane L’Haridonf, Susan E. Humphrisg, Cornel E. J. de Rondeh, Fabio Caratori Tontini, Maurice Tiveyb, Valerie K. Stuckeri, Lucy C. Stewart, Alexander Diehl, and Wolfgang Bachi,2

aCenter for Life in Extreme Environments, Biology Department, Portland State University, Portland, OR 97201; bDepartment of Biology, California State University, Northridge, CA 91330; cBiosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831; dRoyal Netherlands Institute for Sea Research, Department of Marine Microbiology and Biogeochemistry, NL-1790 AB Den Burg, The Netherlands; eDepartment of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, SE-75123 Uppsala, Sweden; fCNRS, Institut Français de Recherche pour l’Exploitation de la Mer, Laboratoire de Microbiologie des Environments Extrêmes, Université de Bretagne Occidentale, F-29280 Plouzané, France; gDepartment of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; hDepartment of Earth Systems and Resources, GNS Science, Avalon, Lower Hut 5010, New Zealand; iLaboratories and Collections, GNS Science, Avalon, Lower Hut 5010, New Zealand; jToha Science, Wellington 6011, New Zealand; kFaculty of Geosciences, University of Bremen, 28359 Bremen, Germany; and lMARUM - Center for Marine Environmental Sciences, University of Bremen, 28359 Bremen, Germany

Edited by Edward F. DeLong, University of Hawaii at Manoa, Honolulu, HI, and approved November 3, 2020 (received for review September 10, 2020)

Hydrothermally active submarine volcanoes are mineral-rich biologic oases contributing significantly to chemical fluxes in the deep sea, yet little is known about the microbial communities inhabiting these systems. Here we investigate the diversity of microbial life in hydrothermal deposits and their metagenomic-inferred physiology in light of the geological history and resulting hydrothermal fluid paths in the subsurface of Brothers submarine volcano north of New Zealand on the southern Kermadec arc. From metagenome-assembled genomes we identified over 90 putative bacterial and archaeal genomic families and nearly 300 previously unknown genera, many potentially endemic to this submarine volcanic environment. While magmatically influenced hydrothermal systems on the volcanic resurgent cones of Brothers volcano harbor communities of thermoacidophiles and diverse members of the superphylum “DPANN,” two distinct communities are associated with the caldera wall, likely shaped by two different types of hydrothermal circulation. The communities whose phylogenetic diversity primarily aligns with that of the cone sites and magmatically influenced hydrothermal systems elsewhere are characterized predominately by anaerobic metabolisms. These populations are probably maintained by fluids with greater magmatic inputs that have interacted with different (deeper) previously altered mineral assemblages. However, proximal (a few meters distant) communities with gene-inferred aerobic, microaerophilic, and anaerobic metabolisms are likely supported by shallower seawater-dominated circulation. Furthermore, mixing of fluids from these two distinct hydrothermal circulation systems may have an underlying imprint on the high microbial phylogenetic diversity. Collectively our results highlight the importance of considering geologic evolution and history of subsurface processes in studying microbial colonization and community dynamics in volcanic environments.

Significance

Much of Earth’s volcanism occurs in the deep sea, yet little is known about the microbial communities inhabiting such extreme and dynamic systems. Using a multidisciplinary approach to study distinct hydrothermal systems at Brothers submarine arc volcano, we provide insights into how microbial community composition and function reflect subtly different fluid chemistries resulting from subsurface fluid interactions with distinct alteration mineral assemblages. These variations can be traced to the subsurface hydrogeologic history beneath Brothers volcano. Further, we show that these systems represent oases of phylogenetically diverse Archaea and Bacteria. Our findings highlight the importance of geologic legacy in understanding drivers of microbial diversity, assembly, and evolution and may have insights into processes that drove early diversification of life on Earth.

Author contributions: A.-L.R., S.E.H., C.E.J.d.R., and M.T. designed research; A.-L.R., E.S.J., S.L., V.K.S., and A.D. performed research; G.E.F., M.P., F.C.T., M.T., and L.C.S. contributed new reagents/analytic tools; A.-L.R., E.S.J., J.M., G.E.F., M.P., N.D., A.S., V.K.S., A.D., and W.B. analyzed data; and A.-L.R. and E.S.J. wrote the paper with contributions from all authors.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

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

S Submarine volcanoes account for ~75% of global volcanic activity (1), contributing significantly to oceanic biological productivity (2) and mineral resources (3). The Kermadec portion of the Kermadec–Tonga intraoceanic volcanic arc has more than 30 major volcanoes, of which 80% are hydrothermally active, making it the most active intraoceanic arc in the world (4). Hydrothermal activity associated with these arc volcanoes is commonly dominated by the discharge of magmatic volatiles, in contrast to midocean ridge vent systems, which are dominated by seawater circulation through oceanic crust. Brothers volcano on the Kermadec arc is unusual in that it hosts both types of hydrothermal systems. At the Upper Cone (UC) and Lower Cone (LC) sites inside the caldera (Fig. 1), relatively low-temperature (~120 °C), highly acidic (pH to 1.9) acid-sulfate fluids derived from disproportionation of magmatic sulfur gases (SO2 and HS) discharge from native sulfur mounds, and extensive Fe oxyhydroxide crusts are common. By contrast, only ~3 km away, high-temperature (~320 °C), less acidic (pH >2.8), but metal-rich fluids are expelled from ~20-m-tall chimneys composed of significant Cu-Zn-Ba ± Au mineralization at the Northwest Caldera Wall (NWC) and Upper Caldera Wall (UCW) sites (5–8) (Fig. 1).

metagenomics | deep-sea hydrothermal | thermophiles | Archaea | volcanics
Changes in the hydrothermal fluid regime at Brothers volcano have been linked to the evolution of the caldera (9). In the precaldera stage, the volcano hosted a hydrothermal system dominated by magmatic volatiles and metal-rich brines. The magmatic volatiles mixed with seawater as they ascended toward the seafloor to produce low-pH, vapor-rich, metal-poor fluids, very much like those discharging at the Cone sites today. The metal-rich brines were segregated and temporarily sequestered in the subsurface. As the volcano grew through the eruption of thick layers of volcaniclastics, caldera collapse eventually occurred, allowing seawater to infiltrate the volcano through fault-controlled permeability. This seawater interacted with wallrock and the segregated brines and transported the associated metals to the seafloor to discharge high-temperature, moderately acidic, metal-rich fluids, forming Cu-Zn-Au-rich chimneys on the NWC and UCW. The recent International Ocean Discovery Program Expedition 376 on the NWC (9, 10) showed that, while the shallower parts of the volcanic sequence are characterized by an alteration mineral assemblage indicative of wallrock reactions with seawater, deeper (underlying) parts of the stratigraphy record evidence of earlier alteration reactions with the magmatically influenced hydrothermal fluids (9). After caldera collapse, a larger resurgent cone (UC) developed followed by a more recent but smaller subsidiary cone (LC) that today both host modern magmatically influenced hydrothermal systems (6, 7).

While several studies have analyzed the microbial diversity of Brothers volcano (11, 12), the impact of both its complex geological evolution and the existence of two proximal but distinct hydrothermal systems on the microbial community diversity is unknown. Studies of terrestrial hydrothermal environments have shown that the geological history of a system impacts microbial community assembly (13) and can be the underlying cause for biogeographically distinct microbial populations (14). Likewise, at deep-sea hydrothermal vent systems, differences in subsurface hydrothermal processes produce fluids of contrasting chemistry that result in different microbial communities associated with the diffuse venting fluids (15, 16) and mineral deposits (17, 18) that develop as the hydrothermal fluids mix with cold seawater. Here, we use amplicon and metagenome sequencing to show that Brothers volcano supports at least three distinct and phylogenetically diverse microbial communities that reflect the complex subsurface hydro-geology of the system.

Results and Discussion

Microbial Community Diversity of Brothers Volcano Hydrothermal Deposits Reveals Two Distinct Communities on the Caldera Wall. To explore the patterns of microbial diversity associated with hydrothermal deposits from the geologically and geochemically different areas at Brothers volcano, we collected hydrothermal deposits from 16 actively discharging hydrothermal vents and two diffuse flow samples from the UCW, NWC, LC, and UC sites (Fig. 1, Dataset S1A, and SI Appendix, Fig. S1). Wherever possible, hydrothermal fluids from vent orifices were collected to examine possible geochemical drivers of differences in microbial community assemblies (Dataset S1B).

Similar to previous analyses of Brothers vent fluids (6, 7), fluids from the NWC and UCW feature chemistry indicative of seawater and rock interactions at high temperatures (~300 °C). These vent fluids reveal subcritical phase separation with chloride values above and below background seawater concentrations (540 mmol/kg), with most being greater than seawater. Magnesium and sulfate have been removed from the end-member fluids, and potassium, calcium, trace metals, and sulfide are all elevated (Dataset S1B). In contrast, samples from the cone sites do not show phase separation and indicate a greater degree of magmatic degassing (as compared to water–rock interactions) with large releases of CO₂, H₂, and shallow mixing of gases and seawater. While both cones have vent field systems primarily showing magmatic degassing, fluids sampled in our study at the UC were hotter at 200 °C than previously recorded (7), compared to the 45 to 60 °C fluids from the LC. The pH in the UC samples was pH 2.1 ± 0.1 and higher in the LC samples (pH 5.4 ± 0.3).

Initially we assessed the microbial community diversity and composition associated with the deposits (and one fluid sample) from the different Brothers volcano sites using 16S ribosomal RNA (rRNA) gene amplicon sequencing (Dataset S1C). The taxonomic composition associated with the magmatically influenced cone sites differed from that of the NWC and UCW sites. Furthermore, there also appeared to be two distinct communities associated with the different caldera wall sites, NWC-A, and...
NWC-B+UCW (Fig. 2A; NWC-A versus NWC-B+UCW, analysis of similarities [ANOSIM] $R = 0.896$, $P = 0.002$). Several of these different communities were almost adjacent sites, for example S145 and S146 (Fig. 1). Despite their close proximity, the microbial community diversity of the NWC-A samples clustered somewhat with samples from the cone sites, while those from NWC-B+UCW formed a unique cluster (Fig. 2A).

Based on similarity percentages breakdown (SIMPER) analysis (Dataset S1D), the Epsilonbacteraeota and Deltaproteobacteria* (in NWC-A) and the Gammaproteobacteria, Aquificae, Aigarchaeota/Geothermarchaeota group (SILVA nomenclature Aigarchaeales) and Hydrothermarchaeota (in NWC-B+UCW) are some of the taxa that contribute to the compositional differences seen between these distinct Caldera Wall communities. The shared Epsilonbacteraeota between the cone sites and NWC-A drive some of the similarity in composition between these communities. Thermoacidophiles, like members of Thermoplasmata, are also shared between the cone sites and NWC-A. In addition, members of the Candidate Phyla Radiation (CPR) and the Diapherotrites, Aenigmarchaeota, Nanoarchaeota, Nanohaloarchaeota (DPANN) Archaea (19) contribute to differences in community structure between the UC and LC and the other sites.

**Geochemical Energy Calculations Confirm Importance of Sulfide Oxidation at All Sites.** Fluid chemistry could be one potential factor driving the above observed community structure differences. The notable differences between the distribution of the Epsilonbacteraeota and the Gammaproteobacteria at the Brothers sites suggest that these taxa may be examples of ecological indicators of the environmental conditions driving the microbial composition shifts. In general, Epsilonbacteraeota occupy niches of lower oxygen and higher sulfide concentrations, while Gammaproteobacteria thrive in higher oxygen and lower sulfide levels (20). Unfortunately, due to insufficient geochemical data for the

---

*These have been recently reclassified, as Desulfobacterota, Waite et al., Int. J. Syst. Evol. Microbiol. DOI: 10.1099/ijs.0.004213.
fluids associated with the samples, especially from NWC-A (only one dilute sample for S142) and NWC-B+UCW (four samples; Dataset S1B), we could not determine a statistically relevant correlation between end-member fluid chemistry and the observed microbial community structure. Further, it is possible that the observed differences in microbial community composition on adjacent sites on the NWC are due to adaptation to subtly different chemistries (not detectable through bulk chemistry measurements) as a result of fluids following different pathways and intersecting different subsurface rock assemblages. Nonetheless, catabolic energy availability calculations for autotrophs of measured vent fluids upon mixing with seawater confirmed that sulfide oxidation was likely the most favorable energy-yielding metabolism at all sites examined, with iron oxidation likely representing an additional important potential process at the caldera wall sites (21). Furthermore, modeling the fluid geochemistry from NWC-B samples suggested that hydrogen oxidation may provide an additional source of energy for these communities.

**Functional Metagenomic Diversity Tracks Microbial Community Diversity.** These thermodynamic calculations provided initial information about the potential energy availability for autotrophy in the Brothers volcano communities. However, natural microbial communities are functionally more complex and usually represent a continuum of other processes in addition to autotrophy.
such as syntrophy, symbiosis, and heterotrophy, all contributing to the overall microbial community composition. To determine to what extent differences in community composition (NWC-A, NWC-B+UCW, UC, and LC) resemble differences in functional complexity, we sequenced the metagenomes of all 18 samples (Dataset S2A) and identified genes coding for proteins involved in energy conservation, mediating key steps of biogeochemical cycles, and indicative of oxygen sensitivity (Fig. 3A and B; Dataset S2R, and SI Appendix, Fig. S2A–F), including the functional processes explored through thermodynamic modeling.

The level of similarity based on the normalized abundance patterns of 123 functional genes/gene categories in the assemblies (Fig. 2B) followed a pattern similar to that of the amplicon diversity (Fig. 2A), with NWC-A affiliated more with the cone sites and NWC-B+UCW forming an independent cluster (NWC-A versus NWC-B+UCW, ANOSIM R = 0.872, P = 0.002). While some of the functional genes confirm the thermodynamic model predictions, namely the importance of sulfide oxidation (sulfide-quinone oxidoreductase, sqr; SI Appendix, Fig. S2A) at all sites, other genes provide insights into potential functional differences that may be driving the differences in community diversity between sites (Fig. 3A and B and SI Appendix, Fig. S2A–F). For example, genes associated with metabolic processes in low-oxygen or anoxic environments (22, 23) are all in higher relative abundance in the NWC-A samples, including the bd-type terminal oxidase (cydA), superoxide reductase (dfs), and the [NiFe] Group 1b hydrogenase, which couples hydrogen oxidation to sulfate, fumarate, or nitrate reduction in certain anaerobes (24) (SI Appendix, Fig. S2B and C). Additionally, NWC-A communities have a lower relative abundance of the microaerophile-associated [NiFe] Group 1d hydrogenase gene (24) and the putative aerobic-type xanthine/oxidase genes (coxA) and superoxide dismutase (SOD2) (Fig. 2B–D). By contrast, the NWC-B+UCW communities appear to be functionally more diverse, possessing a range of genes related to growth under anaerobic, microaerophilic, and aerobic conditions. This suggests that these communities occupy more complex geochemical gradients, or that geothermal fluid inputs are mixed or temporally variable. Furthermore, iron oxidation and iron reduction potential appear to be variable across all sites (SI Appendix, Fig. S2A–C). These functional differences were explored statistically, the analysis further confirmed that the MAGs obtained from NWC-A samples were significantly different from those found at NWC-B+UCW (NWC-A versus NWC-B+UCW, ANOSIM R = 0.931, P = 0.002; Fig. 2C). Putative anaerobes related to the genera Archaeoglobus, Thermococcus, Hippea, Nitratifactor, and Mesoaciditoga and the family Caldisericaceae were prevalent MAGs from the NWC-A samples, while putative microaerophilic Aquificae and Thermoprotei (Aeropyrum-related) and putative carboxydothermophilic Hydrothermarchaeota MAGs characterized the NWC-B+UCW samples (Dataset S3D). Furthermore, from the LC site samples in particular a diversity of previously undescribed members of the DPANN superphylum were assembled (Datasets S3E and S4 and SI Appendix, Figs. SSC and S6 and SI Text). One DPANN lineage common to several sites was the Woesearchaeota. We successfully enriched for a novel woesearchaeote with only Bacteria (Thermales and Aquificales) in the coculture (Dataset S5A and SI Appendix, SI Text). The reduced metabolic repertoire of the assembled metagenome of this woesearchaeote suggests that it forms a symbiotic association with bacterial hosts (Dataset S5B).

Several of the MAGs most likely contributed to the functional differences described above (data available at FigShare (21). For example, 143 (%76) of NWC-A MAGs, including members of the Epislonibacterota, Thermodesulfobacteria, Archaeoglobi, Thermoprotei, Thermococc, Thromoplasmata, and several DPANN, all encoded for superoxide reductase (dfs), which is often prevalent in anaerobes (22), while very few (n = 40, or 19%) MAGs encoded the superoxide dismutase (SOD2) typically found in aerobes. By contrast, over 200 (59%) of the MAGs from NWC-B+UCW encoded for SOD2. Likewise, the genes for terminal oxidases also pointed to the enriched patterns seen in the assemblies, where the high-oxygen-adapted, low-affinity aa3-type heme-copper oxidase (coxA) (23) was detected in 56% of the NWC-B+UCW MAGs but only in 24% of the NWC-A MAGs. Conversely, the low-oxygen-adapted, high-affinity bd-type terminal oxidase (cydA) (23) was identified in 43% of the NWC-A MAGs but only in 26% of the NWC-B+UCW MAGs. The coxA distribution in the NWC-B+UCW assembled genomes was associated with the Gammmaproteobacteria, Thermoprotei,
and Aquificae and to a lesser extent with Heimdallarchaeota, Aigarchaeota, Geothermarchaeota, and Hydrothermarchaeota. The cyaL genes from NWC-A were mostly associated with members of the Epsilonbacteraeta (such as Desulfurellales), Deltaproteobacteria/Thermodesulfobacteria and, to a lesser degree, with the Thermococci. Furthermore, the gene encoding for the putative large subunit of the aerobic-type xanthine/carbon monoxide dehydrogenase family (coxL) was only present in 4% of NWC-A MAGs, while it was detected in 18% of NWC-B+UCW MAGs (including MAGs affiliated with members of the Alphaproteobacteria, Chloroflexi, Calditrichaeota, Thermoprotei, and members of the candidate division KSB1). The differences between the [NiFe] Group 1b (oxygen-sensitive) and [NiFe] Group 1d hydrogenases (oxygen-tolerant) from the NWC-A and NWC-B+UCW tracked with anaerobic Epsilonbacteraeta in NWC-A ([NiFe] Group 1b hydrogenases, 7% of MAGs) and microaerophilic Aquificae and the candidate division KSB1 in NWC-B+UCW ([NiFe] Group 1d hydrogenases,
Further, the enrichment in CO₂ fixation pathways, namely the rTCA cycle and the CBB pathway, could be linked to Epsilonbacteraeota MAGs in NWC-A and the prevalence of Gammaproteobacteria MAGs in NWC-B+UCW, respectively.

Hydrothermal Magmatic Inputs Drive Similarities in Microbial Composition at Brothers Volcano and the Backarc Mariner Vent Field. Collectively, the amplicon, MAG, and functional data point to the NWC-A sites supporting mostly anaerobic communities, while those associated with the NWC-B+UCW sites are more mixed with respect to oxygen usage potential (anaerobic, microaerophilic, and aerobic). Given that the microbial communities at the cone sites track more closely with the NWC-A communities, we hypothesized that the NWC-A community and functional diversity may be influenced by hydrothermal fluids with mostly magmatic inputs, while those of NWC-B+UCW have a stronger oxygenated seawater (due to mixing) signature. Recent drilling on the wall of the caldera has documented the influence of both types of hydrothermal fluids through the presence of mineral assemblages indicative of rock reactions with acid-sulfate-type fluids intercalated and overprinted by assemblages resulting from reactions with seawater-dominated fluids (9).

Fig. 5. Maximum-likelihood phylogenetic tree of archaeal MAGs from Brother volcano, constructed in GTDB-Tk using 122 archaeal marker genes. SH branch support (0.8 to 1.0) is shown with dark circles. Major clades are collapsed into triangles, and triangle coloring is based on GTDB-Tk phylum-level taxonomic classifications. Individual clades are labeled using NCBI taxonomy with GTDB-Tk nomenclature shown in brackets, and novel clades are named using GTDB-Tk taxonomy. Taxonomic groups containing Brothers volcano MAGs are indicated by blue text, with the number of MAGs shown in parentheses. The scale bar shows expected substitutions per amino acid. The uncollapsed tree used to create this figure is available as SI Appendix, Fig. S4 and at iTOL (29) (https://itol.embl.de/shared/alrlab).
In order to explore this further, we tested whether NWC-A microbial communities would be more similar in structure to microbial communities from other magnetically influenced deep-sea vents, such as those associated with the Mariner vent field on the Eastern Lau Spreading Center–Valu Fa Ridge (ELSC-VFR) (17). Here, the deep-sea vent fields (ABE, Tui Maillia, Mariner, and Vai Lili) run north to south along a basalt to andesite gradient (31). Chemical species such as H₂S, Fe, and Mn decrease southward whereas pH and alkalinity increase, with the exception of the southern Mariner vent field, where fluids are significantly hotter, more acidic and metal-rich, and richer in dissolved gases such as H₂S and CO₂ relative to the other vent fields, consistent with the greater input of magmatic gases (32). Indeed, when comparing our Brothers volcano amplicon diversity data with that from ELSC-VFR, the NWC-A communities are more similar to those from Mariner (Mariner versus NWC-A, ANOSIM R = 0.414, P = 0.003), while the NWC-B+UCW communities are unique to Brothers volcano (NWC-A versus NWC-B+UCW, ANOSIM R = 0.896, P = 0.002; Fig. 2D). Thus, it is likely that the differences in the microbial communities on the caldera wall reflect subtle differences (sometimes undetectable by bulk geochemical measurements) in temporal fluxes of fluid mixing and fluid paths in the subsurface (Fig. 6), much like differences observed in the Yellowstone caldera where acid and alkaline hot springs can exit adjacent to each other (33, 34).

Subsurface Heat Flow Patterns Help Explain the Observed Microbial Differences between Sites. The collective phylogenetic, functional, and metagenomic differences of the microbial communities associated with hydrothermal deposits at Brothers volcano (Fig. 6) correspond to the complexity in the subsurface hydrogeology at Brothers volcano. This is supported by recent heat flow studies that show both caldera-scale (2 to 3 km) upflow and recharge and more local-scale (100 to 200 m) heat flow patterns (35). The geological evolution of Brothers volcano and the accompanying changes in the subsurface fluid regime (e.g., phase separation, water–rock reactions, mixing of differently sourced fluids, etc.) produce geochemical gradients that provide a range of geochemical disequilibria that support a diversity of Archaea and Bacteria. In particular, we posit that the NWC-A communities are supported by a mix of magnetically influenced and seawater-influenced hydrothermal fluids, while the NWC-B+UCW communities are supported by shallower seawater-derived hydrothermal fluids. Differences in these microbial communities may be due to adaptation to subtly different chemistries, generated as hydrothermal fluids follow different paths and interact with both the earlier (deeper) acid-sulfate conditions and the more recent (shallow) alkaline conditions. Sulfide was measured using a modified Cline (methylene blue) method and a HACH DR2800 spectrophotometer. Appropriate dilutions were made with efforts to minimize air exposure and loss of analyte. Onshore, major cations and metal concentrations were analyzed by inductively coupled plasma optical emission spectroscopy using acidified and filtered splits with appropriate dilutions and matrix-matched standards. Check standards were run every 10 samples and verified to be within 5% of expected values, and samples were run in triplicate. Major anion samples were filtered and refrigerated prior to analysis with a Dionex ICS3000 equipped with an AG/AS-18 column. Concentrations of dissolved H₂ and CH₄ in vent fluid samples were analyzed using on-board gas chromatography within 12 h after recovery of fluid samples. Volatiles were extracted by a syringe headspace extraction. A Thermo Scientific Trace GC Ultra gas chromatograph, equipped with a packed Molsieve 60/80 column (Sigma-Aldrich), was operated with N₂ as carrier gas at 50 °C. The concentration of CH₄ was quantified using a flame ionization detector. The device was calibrated on a daily basis with reference gases of 1.02 mol % H₂ or 0.987 mol % CH₄ in a N₂ matrix. Samples for determination of dissolved CO₂ were stored upside down in preweighed He-filled and subsequently evacuated glass serum vials to avoid atmospheric CO₂ contamination. CO₂ concentrations were determined onshore using a Thermo Scientific Trace GC Ultra equipped with a packed HaySep 80/100 column (Sigma-Aldrich) and operated with helium as carrier gas at 50 °C. The system was calibrated with reference gas of pure CO₂ with 99.995 mol %. Each sample was measured in triplicate and a control standard was measured prior to each sample block. The accuracy and precision of the measurements are within 10%. The method is described in Reeves et al. (41) and was adopted with an improved calculation procedure for headspace liquid–vapor partitioning. End-member fluid compositions were calculated for the NW Caldera and Upper Caldera based on Mg = 0 extrapolations from background seawater concentrations. The measured vent fluid compositions were used to calculate the Gibbs energies for catalytic reactions in mixtures of vent fluid and seawater following the procedure outlined in Amend et al. (42) and shown in Dataset S18.

Amplification and Sequencing of Environmental Samples. Microbial communities were characterized by sequencing the V4 region of the 16S rDNA gene following the Earth Microbiome Protocol (43) with primers 515F′ (5′-GTCTACGGGNGGCWGCAG-3′) (44) and 806R (5′-GACTACHVGGGTATCTAATAC-3′). Twelve-base-pair barcodes (generated from the primer) were added to the 5′-end of each PCR product. Twenty-five μL PCR reactions were prepared with 12.5 μL of GoTaq Master Mix (Promega Corp.), 0.5 μL of forward and reverse primers (0.2 μM final concentration), 10.5 μL of nuclease-free PCR water (Promega Corp.), and 1 μL of template DNA. Thermal cycling conditions began with an initial denaturation at 94 °C for 3 min, followed by 35 cycles of 94 °C for 45 s, 50 °C for 60 s, and 72 °C for 90 s. The final extension was at 72 °C for 10 min. After amplification, PCR products were pooled, visualized using gel electrophoresis, and then quantified using the Quant-iT Picogreen dsDNA kit (Invitrogen). Three hundred nanograms of each product were pooled, cleaned using the Invitrogen PureLink PCR purification kit (Carlsbad), and quantified using the Qubit dsDNA High Sensitivity (HS) kit (Invitrogen). Sequencing was performed on an Illumina MiSeq using the their microbial community more carefully, we may be able to use the microbial community diversity as indicators of potential subsurface magmatic and hydrothermal processes.

Materials and Methods

Sample Collection, DNA Extraction, and Hydrothermal Fluid Chemistry Analysis. Hydrothermal deposits were collected from the Brothers volcano (~34.8708° S, 179.0667°) E NW Caldera, UCW, and LC and UC sites during the research vessel (RV) Thomas Thompson March 2018 TN350 expedition using the remotely operated vehicle (ROV) Jason. Water samples were collected using isobaric gas-tight fluid samplers (38) and major water samplers (39). Additional deposits were collected from the ELSC-VFR during the RV Roger Revelle (RR1507) expedition in April/May 2015. Once shipped, deposits were processed as described previously (17). DNA was extracted from homogenized deposits using the DNeasy PowerSoil kit (Qiagen). For DNA from water, 1 L of water was filtered using a 0.2-μm Sterivex filter (Merck) and DNA was extracted as described by Anderson et al. (40).

Geochemical Measurements. pH was measured onboard at thermal equilibrium with a Metromeh 780 pH meter calibrated daily with standard buffers. A Hach digital titrator and prepared titrant solutions of HCl and NaOH (Fish-erbrand 0.1 M appropriately diluted and calibrated) were used for alkalinity/ acidity titrations. Sulfide was measured using a modified Cline (methylene blue) method and a HACH DR2800 spectrophotometer. Appropriate dilutions were made with efforts to minimize air exposure and loss of analyte. Onshore, major cations and metal concentrations were analyzed by inductively coupled plasma optical emission spectroscopy using acidified and filtered splits with appropriate dilutions and matrix-matched standards. Check standards were run every 10 samples and verified to be within 5% of expected values, and samples were run in triplicate. Major anion samples were filtered and refrigerated prior to analysis with a Dionex ICS3000 equipped with an AG/AS-18 column. Concentrations of dissolved H₂ and CH₄ in vent fluid samples were analyzed using on-board gas chromatography within 12 h after recovery of fluid samples. Volatiles were extracted by a syringe headspace extraction. A Thermo Scientific Trace GC Ultra gas chromatograph, equipped with a packed Molsieve 60/80 column (Sigma-Aldrich), was operated with N₂ as carrier gas at 50 °C. The concentration of CH₄ was quantified using a flame ionization detector. The device was calibrated on a daily basis with reference gases of 1.02 mol % H₂ or 0.987 mol % CH₄ in a N₂ matrix. Samples for determination of dissolved CO₂ were stored upside down in preweighed He-filled and subsequently evacuated glass serum vials to avoid atmospheric CO₂ contamination. CO₂ concentrations were determined onshore using a Thermo Scientific Trace GC Ultra equipped with a packed HaySep 80/100 column (Sigma-Aldrich) and operated with helium as carrier gas at 50 °C. The system was calibrated with reference gas of pure CO₂ with 99.995 mol %. Each sample was measured in triplicate and a control standard was measured prior to each sample block. The accuracy and precision of the measurements are within 10%. The method is described in Reeves et al. (41) and was adopted with an improved calculation procedure for headspace liquid–vapor partitioning. End-member fluid compositions were calculated for the NW Caldera and Upper Caldera based on Mg = 0 extrapolations from background seawater concentrations. The measured vent fluid compositions were used to calculate the Gibbs energies for catalytic reactions in mixtures of vent fluid and seawater following the procedure outlined in Amend et al. (42) and shown in Dataset S18.
Fig. 6. The proposed influence of subsurface hydrogeological processes on the microbial phylogenetic and functional diversity associated with Brothers volcano and its hydrothermal systems. Schematic cross-section adapted from de Ronde et al. (9) from southeast to northwest. Overall, the shallow subsurface is influenced by seawater-dominated hydrothermal circulation (blue arrows = recharge; red arrows = discharge). Recharge of seawater occurs both at a volcano scale through the caldera floor and faults along the caldera wall and more locally at the 100- to 200-m scale at the vent sites (see Caratori Tontini et al. (35)). Red arrows denote heated (modified) seawater after reaction with the rocks; purple arrows represent hot, magmatically influenced fluids; red-brown arrows represent magmatic volatiles. Cross-hatching represents zones of inferred magmatic salt [see de Ronde et al. (9)]. (A) 2D plots of the average relative abundance of key taxa (NCBI taxonomy) amplicons discussed in the text from each area are shown, with bubble size corresponding to percent average relative abundance (Dataset S1). (B) bubble plots of the average relative abundance of Archaea was calculated separately using only archaeal taxa, and bubble size represents the percent average relative abundance of key taxa (NCBI taxonomy) amplicons discussed in the text from each area are shown, with bubble size corresponding to percent average relative abundance (Dataset S1). (A) Bubble plots of the average relative abundance of key taxa (NCBI taxonomy) amplicons discussed in the text from each area are shown, with bubble size corresponding to percent average relative abundance (Dataset S1). (B) Bubble plots of the average relative abundance of key taxa (NCBI taxonomy) amplicons discussed in the text from each area are shown, with bubble size corresponding to percent average relative abundance (Dataset S1).
mean contig lengths in two) and thus MEGAHIT assemblies were chosen for further metagenomic analyses. To generate differential read coverage profiles, reads were mapped to the contigs with Bowtie2 v.2.2.9 (51) and paired using SAMtools v.1.3.1 (52).

Annotation of Metagenome Assemblies. To assess functional differences between samples, proteins on assembled contigs from each metagenome were predicted using Prodigal v.2.6.3 (53) and annotated with the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (54) using GhostKOALA (55). KEGG orthology (KO) hits corresponding to catalytic hydrogenase subunits (K00436, K00440, K00592, K02681, K12142, K14090, K14106, K14123, K14126, K18106, K18332, and K23549) were further characterized using HyDad v.24 (24), and only two hydrogenase hits, proteins of K02950, K02992, K0296, K02926, K02909, K02994, K02979, K02959, K02948, K02831, K02881, and K02986, as previously described (57). Normalized values were also averaged across samples from each area (UC, LC, NWC-A, and NWC-B+UCW), and a heat map was constructed using Morpheus (https://software.broadinstitute.org/morpheus).

Metagenome Binning. Contigs (>2,000 bp) were binned as draft MAGs with MetaBAT v.0.32.4 (58) based on tetranucleotide frequency and differential read coverage. CheckM (59) was used to estimate MAG completion and contamination, and Phylotyph (v.1.0.1) (60) was used to screen for 16 ribosomal RNA genes (rRNA). MetaBAT v.0.32.4 (58) and MetaProk v.0.6.0 (61) were used to bin contigs into MAGs. The Tax4Fun pipeline was used to predict functional proteins, and detailed nomenclature proposed by Waite et al. (72), which is largely congruent with GTDB data release 86 v.3 taxonomy at the class level and lower taxonomic ranks. Detailed taxonomy of the DPANN superphylum MAGs is also included in the phylogenetic tree depicted in SI Appendix, Fig. 56.

Curation of Representative Genomes. Representative MAGs belonging to taxonomic clades of interest were selected for additional curation using Emerging Self-Organizing Maps (62) with the tetramerFreq script package (https://github.com/tetramerFreq/Binning) (63). After calculating the tetranucleotide frequency of individual MAGs and multiple reference genomes, ESOM maps were trained using the K-batch option with the recommended number of rows and columns. MAGs were manually rebinned, and outlier contigs with ≤90% confidence were removed. Identified curated MAGs are denoted on the phylogenetic trees (discussed below) by an asterisk.

Phylogenetic Analyses. As a framework for this study, genome bins were used to construct maximum-likelihood phylogenetic trees for the Bacteria (n = 21,780 taxa and 5,035 amino acid positions) and Archaea (n = 1,371 taxa and 9,461 amino acid positions) with an ultrafast bootstrap approach (n = 400 taxa and 9,461 amino acid positions) with an ultrafast bootstrap approach and SH-like approximate likelihood tests, each run with 1,000 bootstrap replicates.

Taxonomic Classification and Novel Taxonomic Rank Designation. MAGs were classified using the GTDB-Tk v.0.2.2 (data release 86 v.3) classify workflow (27) and with the NCBI taxonomy browser. In order to explore novel family and genus level ranks for MAGs, we used both the GTDB-Tk workflow and average amino acid identity (AAI) matrices (using the Enveomics AAI Matrix Tool, enve-omics.ce.gatech.edu/g-matrix) to inform our decision. In general, we used the GTDB-Tk assignment first and confirmed the ranking with AAIs. In a few cases, compelling evidence based on AAI similarity superseded the GTDB-Tk ranking. For comparison, archaeal taxonomic ranks were also determined using AAI comparisons between MAGs and reference genomes. AAI matrices for the DPANN superphylum were generated separately using CompareM v.0.0.20.23 (https://github.com/idparks134/CompareM).

Nomenclature. We recognize the archaeal and bacterial nomenclature is in flux. Here we primarily use the NCBI-assigned taxonomies or the GTDB-recommended taxonomy in parentheses. Dataset S38 provides both nomenclatures for reference. We refer to the Epsilonproteobacteria (formerly Epsilonbacteraeta; GTDB Campylobacteraeta) throughout the text using detailed nomenclature proposed by Waite et al. (72), which is largely congruent with GTDB data release 86 v.3 taxonomy at the class level and lower taxonomic ranks. Detailed taxonomy of the DPANN superphylum MAGs is also included in the phylogenetic tree depicted in SI Appendix, Fig. 56.

Annnotation and Comparison of MAGs. Open reading frames (ORFs) within MAGs were predicted using Prodigal v.2.6.3 (53) as implemented in Prokka v.1.13 (73). For comparison with assembly annotation data, the MAGs were annotated as described above using GhostKoala (55), HyDad v.24, and FeGenie (56).

Statistical Analyses. To assess site differences at Brothers volcano, Bray–Curtis similarity matrices were generated in PRIMER v.6.1.13 (74) using the relative abundance of key functional genes/gene categories identified in the assemblies (n = 123 genes/gene categories) and the relative abundance of medium- to high-quality MAGs based on normalized read coverage and GTDB-Tk taxonomy (n = 251 taxonomic assignments). Bray–Curtis dissimilarity matrices comparing communities from Brothers volcano samples and from both Brothers volcano and ELSC-VFR samples were also generated in QIME2 as part of the “core-metrics-phylogenetic” workflow in the “diversity” plugin, at a sampling depth of 3,000 sequences per sample. Similarity and distance matrices were used to create nonmetric multidimensional scaling (NMDS) plots in PRIMER. One-way ANOSIM was performed on each matrix to identify statistically significant differences between sites, and one-way SIMPER analysis was used to identify specific contributors to the similarities and differences between sites where applicable.

Data Availability. Amplicon reads from Brothers volcano, Bray–Curtis similarity matrices were generated in PRIMER v.6.1.13 (74) using the relative abundance of key functional genes/gene categories identified in the assemblies (n = 123 genes/gene categories) and the relative abundance of medium- to high-quality MAGS based on normalized read coverage and GTDB-Tk taxonomy (n = 251 taxonomic assignments). Bray–Curtis dissimilarity matrices comparing communities from Brothers volcano samples and from both Brothers volcano and ELSC-VFR samples were also generated in QIME2 as part of the “core-metrics-phylogenetic” workflow in the “diversity” plugin, at a sampling depth of 3,000 sequences per sample. Similarity and distance matrices were used to create nonmetric multidimensional scaling (NMDS) plots in PRIMER. One-way ANOSIM was performed on each matrix to identify statistically significant differences between sites, and one-way SIMPER analysis was used to identify specific contributors to the similarities and differences between sites where applicable.

Acknowledgments. We thank the captain and crew of the RV Thompson and the engineers from Woods Hole Oceanographic Institution for the successful operation of ROV Jason. The project was funded by NSF grants OCE-1558356 (Principal Investigator S.E.H.) and OCE-1558795 (Principal Investigator A.D.T.). The project was funded by NSF grants OCE-1558356 (Principal Investigator S.E.H.) and OCE-1558795 (Principal Investigator A.D.T.). The project was funded by NSF grants OCE-1558356 (Principal Investigator S.E.H.) and OCE-1558795 (Principal Investigator A.D.T.). The project was funded by NSF grants OCE-1558356 (Principal Investigator S.E.H.) and OCE-1558795 (Principal Investigator A.D.T.). The project was funded by NSF grants OCE-1558356 (Principal Investigator S.E.H.) and OCE-1558795 (Principal Investigator A.D.T.). The project was funded by NSF grants OCE-1558356 (Principal Investigator S.E.H.) and OCE-1558795 (Principal Investigator A.D.T.).
29. A.-L. Reysenbach et al

27. P.-A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-Tk: A toolkit to delineate biogeographic provinces among Yellowstone thermophiles.

24. D. Søndergaard, C. N. S. Pedersen, C. Greening, HydDB: A web tool for hydrogenase enzymes in seven hydrothermal vent systems of the Kermadec intraoceanic arc. Chem. Geol. 528, 119298 (2020).

23. J. C. A. E. de Ronde et al, Submarine hydrothermal activity and gold-rich mineralization at Brothers volcano, Kermadec arc, New Zealand. Miner. Depos. 46, 541–584 (2011).

22. Y. Sheng et al, Variability in microbial communities in black smoker chimneys from volcanic vent systems of the Kermadec intraoceanic arc. Geophys. Res. Solid Earth 113, B08S06 (2008).

21. K. Taka et al, Ultra-high-throughput microbial community analysis on the Ilumina HiSeq and MiSeq platforms. ISME J. 6, 1621–1624 (2012).

20. G. E. Flores et al, Evolution of a submarine magmatic-hydrothermal system: Impacts of subseafloor phase-densification processes on the chemistry of hydrothermal fluids at a flow-through hydrothermal vent at Brothers volcano, Kermadec arc, New Zealand. J. Geophys. Res. Solid Earth 113, B08S06 (2008).

19. L. A. Hug et al, Additional files for Complex subsurface hydrothermal fluid mixing at a submarine arc volcano supports distinct and highly diverse microbial communities. mSystems. https://doi.org/10.6084/m9.figshare.c.5099348. Deposited 20 November 2020.

18. S. Nakagawa et al, Prodigal: Prokaryotic gene recognition and translation initiation site identification. Bioinformatics 22, 3674–3676 (2006).

17. G. E. Flores et al, Evolution of a submarine magmatic-hydrothermal system: Impacts of subseafloor phase-densification processes on the chemistry of hydrothermal fluids at a flow-through hydrothermal vent at Brothers volcano, Kermadec arc, New Zealand. J. Geophys. Res. Solid Earth 113, B08S06 (2008).

16. C. Kleint et al, Geochemical characterization of highly diverse hydrothermal fluids from seafloor massive sulfide deposits. Earth Planet. Sci. Lett. 212, 1–10 (2003).

15. A. V. Caetano et al, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Methods 21, 935–940 (2020).

14. C. Nakagawa et al, Variation in microbial community and venting chemistry in a�αiluroidea lava fuels phytoplankton bloom in the North Pacific Ocean. Sci. Adv. 5, eaau6785 (2019).

13. D. Payne et al, Geologic legacy spanning > 200 years explains unique Yellowstone hot spring geochemistry and biodiversity. Environ. Microbiol. 21, 4180–4195 (2019).

12. K. Takai et al, Prodigal: Prokaryotic gene recognition and translation initiation site identification. Bioinformatics 22, 3674–3676 (2006).

11. M. B. Stott et al, Culture-independent characterization of a novel microbial community at a hydrothermal vent at Brothers volcano, Kermadec arc, New Zealand. J. Geophys. Res. Solid Earth 113, B08S06 (2008).

10. J. S. Seewald et al, KOH: A new gas-tight isobaric sampler for hydrothermal fluids. Deep. Res. Part I Oceanogr. Res. Pap. 49, 189–196 (2002).

9. D. Søndergaard et al, Axial Seamount harbor distinct subseafloor microbial communities. Geobiology 11, 98–106 (2013).

8. P. A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-Tk: A toolkit to delineate biogeographic provinces among Yellowstone thermophiles. Nat. Microbiol. 4, 1613–1633 (2019).
63. G. J. Dick et al., Community-wide analysis of microbial genome sequence signatures. Genome Biol. 10, R85 (2009).
64. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797 (2004).
65. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313 (2014).
66. I. Letunic, P. Bork, Interactive tree of life (iTOl) v4: Recent updates and new developments. Nucleic Acids Res. 47, W256–W259 (2019).
67. K. Zaremba-Niedzwiedzka et al., Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature 541, 353–358 (2017).
68. K. S. Makarova, Y. I. Wolf, E. V. Koonin, Archaeal clusters of orthologous genes (archCOGs): An update and application for analysis of shared features between Thermo cocci, Methanococcales, and Methanobacteriales. Life 5, 818–840 (2015).
69. K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066 (2002).
70. A. Criscuolo, S. Gribaldo, BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. 10, 210 (2010).
71. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274 (2015).
72. D. W. Waite et al., Comparative genomic analysis of the class Epsilonproteobacteria and proposed reclassification to Epsilonbacteraeota (phyl. nov.). Front. Microbiol. 8, 682 (2017).
73. T. Seemann, Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069 (2014).
74. K. R. Clarke, R. N. Gorley, Primer v6: User Manual/Tutorial (PRIMER-E Ltd, 2006).