


Spatially distinct, temporally stable microbial populations mediate biogeochemical cycling at and below the seafloor in hydrothermal vent fluids

Caroline S. Fortunato,^{1,2*} Benjamin Larson,³
David A. Butterfield³ and Julie A. Huber ^{1,4}

¹Marine Biological Laboratory, Josephine Bay Paul Center, Woods Hole, MA, USA.

²Department of Biology, Wilkes University, Wilkes-Barre, PA, USA.

³Joint Institute for the Study of the Atmosphere and Ocean, University of Washington and NOAA Pacific Marine Environmental Lab, Seattle, WA, USA.

⁴Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA.

Summary

At deep-sea hydrothermal vents, microbial communities thrive across geochemical gradients above, at, and below the seafloor. In this study, we determined the gene content and transcription patterns of microbial communities and specific populations to understand the taxonomy and metabolism both spatially and temporally across geochemically different diffuse fluid hydrothermal vents. Vent fluids were examined via metagenomic, metatranscriptomic, genomic binning, and geochemical analyses from Axial Seamount, an active submarine volcano on the Juan de Fuca Ridge in the NE Pacific Ocean, from 2013 to 2015 at three different vents: Anemone, Marker 33, and Marker 113. Results showed that individual vent sites maintained microbial communities and specific populations over time, but with spatially distinct taxonomic, metabolic potential, and gene transcription profiles. The geochemistry and physical structure of each vent both played important roles in shaping the dominant organisms and metabolisms present at each site. Genomic binning identified key populations of SUP05, Aquificales and methanogenic archaea carrying out important transformations of

carbon, sulfur, hydrogen, and nitrogen, with groups that appear unique to individual sites. This work highlights the connection between microbial metabolic processes, fluid chemistry, and microbial population dynamics at and below the seafloor and increases understanding of the role of hydrothermal vent microbial communities in deep ocean biogeochemical cycles.

Introduction

At active hydrothermal systems, cold, oxidized seawater mixes with hot, chemically reduced hydrothermal fluids at and below the seafloor, forming low-temperature diffuse vent fluids that contain abundant energy sources to support taxonomically and metabolically diverse microbial communities (Jannasch and Mottl, 1985; Karl, 1995; Reysenbach and Shock, 2002; Huber *et al.*, 2007; Perner *et al.*, 2009; Amend *et al.*, 2011; Olins *et al.*, 2017). Examination of these mixed diffuse fluids allows for inferences about microbially mediated processes occurring beneath the seafloor, where various seafloor habitats are predicted to exist spanning oxic to anoxic conditions and temperatures of 5°C to greater than 100°C (Huber *et al.*, 2002; Butterfield *et al.*, 2004; Sievert and Vetriani, 2012; Fortunato and Huber, 2016).

Previous studies of diffuse fluid microbiology at hydrothermal vents have focused on the phylogenetic diversity within and between vent fields based on both the 16S rRNA gene as well as specific functional marker genes (e.g., Sievert *et al.*, 1999; Takai *et al.*, 2003; Huber *et al.*, 2007; Opatkiewicz *et al.*, 2009; Huber *et al.*, 2010; Akerman *et al.*, 2013; Perner *et al.*, 2013; Gonnella *et al.*, 2016; Meier *et al.*, 2016). Many studies have analyzed phylogenetic diversity across both space and time, over a range of scales. For example, a recent analysis of 63 million 16S rRNA gene sequences from ocean water samples and numerous hydrothermal vents along the Mid-Atlantic Ridge concluded that the global ocean acts as a seed bank for hydrothermal systems, with microbial community composition at any individual vent site reflecting both specific environmental parameters as well as random

Received 16 February, 2017; revised 20 October, 2017; accepted 22 November, 2017. *For correspondence. E-mail caroline.fortunato@wilkes.edu; Tel. (+570) 408 4766; Fax 570-408-7862.

dispersal (Gonnella *et al.*, 2016). In a spatial study of diffuse fluids from five hydrothermally active seamounts along the Mariana Arc, Huber *et al.* (2010) found that while evidence for large-scale geographic differences (> 1000 km) in bacterial community composition existed, at the local scale most vents hosted distinct communities of Epsilonbacteraeota regardless of seamount location. In a regional-scale study comparing the communities in diffuse fluids from two vent fields with distinct geology and chemistry located within 20 km of one another on the Mid-Cayman Rise, Reveillaud *et al.* (2016) showed that there were fine-scale phylogenetic differences in microbial community structure between the two vent fields, with fluid communities at each vent field more closely related to one another than communities found in fluids from the other vent field. In addition, smaller scale spatial studies of multiple diffuse vents within one vent field (Huber *et al.*, 2007; Opatkiewicz *et al.*, 2009; Perner *et al.*, 2010; Akerman *et al.*, 2013; Meier *et al.*, 2016; Olins *et al.*, 2017) show that the microbial community composition is shaped by both the geochemistry as well as the distance between individual diffuse vents. For example, Akerman *et al.* (2013) showed that patterns in abundance of Epsilonbacteraeota at Axial Seamount were correlated to vent methane concentrations as well as vent location, while at the Menez Gwen vent field, Meier *et al.* (2016) found a transition from chemolithoautotrophic communities close to sources of venting to heterotrophic communities with increasing distance from the venting source.

Layered onto these spatial studies are a number of temporal analyses, which have compared the microbial community composition of venting fluids and the geochemical properties of these fluids (Huber *et al.*, 2002; 2003; Opatkiewicz *et al.*, 2009; Perner *et al.*, 2009; 2013). Opatkiewicz *et al.* (2009) showed that over a 5-year time period at Axial Seamount, the microbial composition at five individual vents maintained a community structure that was distinct at each vent site despite changes in geochemistry over the same time period, suggesting stability in microbial communities at each site. Temporal studies across shorter time scales (less than one hour) have demonstrated that microbial community composition can change over short periods of time at a single vent, but temporal differences at a single vent site are less significant than spatial differences across separate vent sites (Perner *et al.*, 2009). These studies suggest that the availability of abiotic properties like oxygen, methane, and hydrogen sulfide in combination with alternations in flow and mixing regimes can play a role in determining the microbial communities present in venting fluids (Perner *et al.*, 2009; 2013).

Recently, a number of metagenomic and metatranscriptomic studies of microbial communities within and around hydrothermal vents have led to new insights about the

diversity of microbial metabolisms present and expressed under different geological and chemical conditions (e.g., Dick *et al.*, 2013; Urich *et al.*, 2014; Li *et al.*, 2015b; Anantharaman *et al.*, 2016; Fortunato and Huber, 2016; Meier *et al.*, 2016; 2017; Reveillaud *et al.*, 2016; Olins *et al.*, 2017). For example, hydrothermal vent plumes sampled above the seafloor were shown to vary chemically and be taxonomically and metabolically distinct from seafloor vent fluid and subseafloor communities (Dick *et al.*, 2013). Reveillaud *et al.* (2016) showed that subseafloor communities in diffuse fluids at two chemically and geologically distinct vent fields were broadly functionally similar, consistent with thermodynamic predictions. In addition, a study within a single vent field at Axial Seamount showed differences in gene expression patterns between both diffuse vents as well as within the intra-field waters between vent sites (Olins *et al.*, 2017). Finally, through the reconstruction of genomes from metagenomes (Tyson *et al.*, 2004; Dick *et al.*, 2009; Hug *et al.*, 2016), recent genomic binning studies at hydrothermal systems have revealed the presence of novel microbial and viral groups as well as their metabolic potential, including the discovery of auxiliary metabolic genes for sulfur oxidation within deep ocean viruses and the potential importance of hydrocarbon degradation distal to venting fluids, including within hydrothermal plumes (Anantharaman *et al.*, 2013; 2014; 2016; Li *et al.*, 2014; Sheik *et al.*, 2014; Meier *et al.*, 2016; 2017).

From these previous studies, it is clear that inter-vent field as well as intra-vent field microbial community diversity in diffuse fluids varies with both geochemistry and vent location, although it is difficult to tease apart significant patterns given low sample sizes and the dynamic nature of hydrothermal systems. In addition, the role of the heterogeneous subseafloor community must be considered in interpreting microbial community composition data from venting diffuse fluids, given that these mixed fluids entrain microbes living both at and below the seafloor. Subseafloor microbes may act as a seed bank to the surrounding vent ecosystem, yet our understanding of what controls the connectivity and stability of the subseafloor habitat is poorly constrained. Furthermore, the linkages between the identity of microbial lineages present at hydrothermal vents and their metabolic potential, gene expression patterns, and population dynamics is severely lacking in comparison to our understanding of microbial communities in global ocean waters.

Therefore, to address these gaps in our knowledge about microbial communities at deep-sea hydrothermal vents, we applied metagenomic, metatranscriptomic, and geochemical measurements to low-temperature venting fluids from three geologically distinct individual sites at Axial Seamount over a three year time period. We sampled the venting fluids from each site annually to

determine the geochemical signatures, taxonomic composition, metabolic potential, gene expression profiles, and population dynamics of microbial communities in venting fluids both at and below the seafloor. Results were compared to background seawater as well as the hydrothermal plume above the vents. Our study location is Axial Seamount, an active submarine volcano located on the Juan de Fuca Ridge, about 300 miles off the coast of Oregon, USA where there is a long history of geological, chemical, and biological studies (Chadwick *et al.*, 2010; Kelley *et al.*, 2014; Wilcock *et al.*, 2016). Building on the previous microbiological studies at Axial and elsewhere focused on microbial community composition in diffuse fluids, we hypothesized that individual vent sites maintain microbial communities and populations with distinct functional gene potential and gene expression profiles that reflect local subseafloor conditions. The results of this study increase our understanding of the key microbially mediated biogeochemical processes occurring across the dynamic mixing zones of deep-sea hydrothermal vents, and provide new insight into the complex environmental drivers that allow microbes to establish and thrive in deep-sea hydrothermal crustal environments.

Results

The diffuse vent sites Anemone, Marker 33, and Marker 113 were chosen for this study because they span three different vent fields within the Axial caldera and each vent has a distinct geological and chemical setting (Supporting Information Fig. S1, Table 1). Anemone vent is located on the southwest side of the caldera in the southern half of the ASHES vent field and sits on top of a low sulfide mound that forms as a result of mixing between hot source fluids and ambient seawater. Diffuse fluids sampled from Anemone for this study were less than 1 meter away from hot (nearly 300°C) hydrothermal fluid. In 2015, the hydrothermal plume sample was taken at 1500 m depth, approximately 42 m above Anemone vent. In contrast to Anemone, both Marker 33 and Marker 113 vent directly

from basalt. Marker 33 is located approximately 2.5 km east of Anemone along the eruption zone in the southeast side of the caldera, where fluids emanate directly from cracks and crevices in lobate lava erupted in 2011. Prior to the 2011 eruption, fluids at Marker 33 emanated from a long crack in broken sheet flow lava erupted in 1998 (Butterfield *et al.*, 2004). Finally, Marker 113 vents diffuse fluids from the top of a lava pillar covered in tube worms, near the edge of old lava flows, approximately 1 km south of Marker 33 and 2 km southeast of Anemone (Supporting Information Fig. S1, Butterfield *et al.*, 2004). There are no known high-temperature vents adjacent to either Marker 33 or Marker 113, therefore extrapolated zero-Mg end members are not valid for chemical species that are clearly reactive at low temperature. Statistical comparisons of fluid chemistry and microbiology were carried out using the actual measured fluid compositions that corresponded closely to the microbial material collected and analyzed.

While sampled vent temperatures at the three vent sites were of a similar range (18°C–40°C), other chemical parameters differed from vent to vent and to a lesser extent from year to year (Table 1). For example, dissolved hydrogen gas measurements were highest at Anemone, with an average concentration of 18 µM. Dissolved hydrogen gas concentrations at Marker 33 and Marker 113 were much lower compared to Anemone, with average concentrations of 1 µM and 0.9 µM respectively (Table 1). Nitrate concentrations showed a similar trend, with highest concentrations found at Anemone (average 35.5 µmol/kg), about half that concentration, on average, at Marker 33 (average 18 µmol/kg), and the lowest at Marker 113 (average 7.8 µmol/kg). The degree of mixing between seawater and hydrothermal fluid, shown as percent seawater in Table 1, was calculated using the conservative tracer magnesium. The percent seawater was highest on average at Marker 113 (average = 96.1%) and lowest at Marker 33 (average = 88.2%), with Anemone falling in the middle (average = 90%, Table 1). Cell counts at each vent were more variable from year to year with no clear pattern.

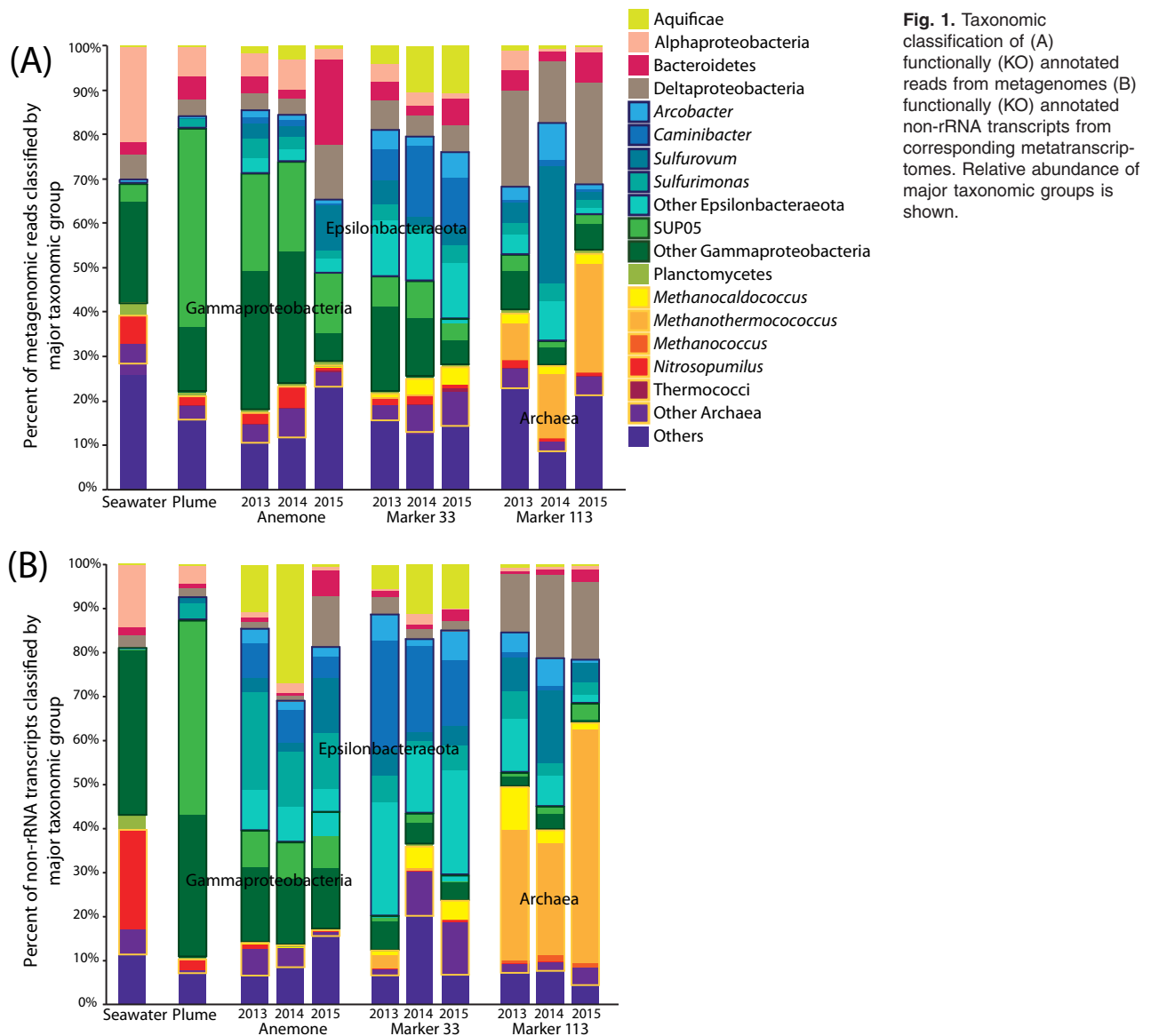
Table 1. Environmental measurements for Anemone, Marker 33 and Marker 113 diffuse vents for 2013–2015. Percent seawater was calculated via the conservative tracer magnesium (Mg): $[Mg]_{vent}/[Mg]_{seawater} * 100$. Seawater Mg concentration from 1500 m was 52.35 mmol kg⁻¹.

	Seawater 2015	Anemone 2013	Anemone 2014	Anemone 2015	Marker 33 2013	Marker 33 2014	Marker 33 2015	Marker 113 2013	Marker 113 2014	Marker 113 2015
Depth (m)	1520	1542	1539	1542	1516	1514	1516	1520	1518	1520
Temperature (°C)	2.0	28.2	20.3	17.9	27.3	18.5	40.6	24.5	24.3	25.4
pH	7.8	5.5	5.1	5.8	5.5	5.6	5.6	6.2	5.8	6.6
Percent Seawater	100	88.9	87.0	94.1	87.7	91.9	85.0	96.0	96.4	95.8
Mg (mmol/kg)	52.4	46.5	45.6	49.2	45.9	48.1	44.5	50.2	48.8	50.2
NO ₃ ⁻ (µmol/kg)	39.0	35.5	35.2	35.9	18.3	28.2	7.6	7.2	8.1	8.0
H ₂ (µM)	0.002	13.9	27.7	12.5	1.5	1.5	0.1	1.4	1.0	0.3
H ₂ S (µmol/kg)	0.0	1036.0	964.7	456.0	545.0	261.5	627.7	729.0	556.0	578.2
CH ₄ (µmol/kg)	0.002	14.8	31.7	13.7	19.0	6.4	27.6	16.9	38.0	22.3
Microbial cells per mL	2.5 × 10 ⁴	4.1 × 10 ⁵	1.7 × 10 ⁵	1.0 × 10 ⁶	4.2 × 10 ⁵	3.9 × 10 ⁵	7.2 × 10 ⁵	4.6 × 10 ⁵	6.8 × 10 ⁵	1.5 × 10 ⁶

The 2015 post-eruption chemistry was variable, but some trends did emerge. There was a 70%–90% decrease in the hydrogen concentrations at Marker 33 and Marker 113 from 2014 to 2015. At Anemone, hydrogen sulfide as well as methane was lowest in 2015 while at Marker 33 hydrogen sulfide and methane concentrations were highest in 2015. Additionally, post-eruption cell counts were elevated over previous years at each vent (Table 1; Topçuoğlu *et al.*, 2016).

Differences in the microbial community composition were observed across vents and in the hydrothermal plume above Anemone vent. Figure 1A shows the taxonomic composition of functionally (KO) annotated reads from each of the nine vent metagenomes as well as from background seawater taken at 1500 m depth and from the Anemone hydrothermal plume. Seawater was dominated by

Alphaproteobacteria and Gammaproteobacteria, which made up 21% and 27% of annotated reads respectively (Fig. 1A). The ammonia-oxidizing archaea *Nitrosopumilus* was also present, comprising 6% of annotated reads. In the plume metagenome, the relative abundance of the Gammaproteobacteria group SUP05 was much higher compared to background seawater, comprising over 45% of the total community. SUP05 are free-living Gammaproteobacteria that oxidize sulfur and hydrogen and are abundant in the ocean, especially in oxygen minimum zones and hydrothermal plumes (Anantharaman *et al.*, 2013; Shah *et al.*, 2016). Along with SUP05, other Gammaproteobacteria (15%), as well as a small percentage of Epsilonbacteraeota (2%), and archaea (5%) were also observed in the plume (Fig. 1A). Across all vents, Epsilonbacteraeota were more prevalent compared to seawater



and plume. Epsilonbacteraeota averaged 14%, 35% and 24% of the annotated reads at Anemone, Marker33 and Marker113 respectively (Fig. 1A). SUP05 and other Gammaproteobacteria annotated reads were relatively more abundant at Anemone than at the other vents. Annotated reads classified to thermophilic Epsilonbacteraeota like *Caminibacter* were more abundant at Marker 33. Methanogenic archaea were most abundant at Marker 113, comprising on average 18% of annotated reads (Fig. 1A). Taxonomic classification of 16S rRNA reads from the metagenomes show relatively similar composition pattern to the functionally annotated reads (Supporting Information Fig. S2). Although microbial community composition differed across the three vents, across years changes in community patterns were less pronounced. In 2015, there was an increase the abundance of annotated reads classified to Epsilonbacteraeota present at Anemone compared to the previous years and in 2014, annotated reads classified to *Sulfurovum* were more abundant at Marker 113 (Fig. 1A). Overall, however, temporal patterns within a vent were less distinct than spatial differences across vents.

The dominant groups identified in the metagenomes also had active gene transcription (Fig. 1B). In the plume, SUP05 comprised 44% of annotated transcripts and all Gammaproteobacteria totaled 77% of annotated transcripts. Transcripts identified as Gammaproteobacteria, *Sulfurimonas*, and the Aquificae were the most abundant at Anemone vent, consisting of, on average, 24%, 16%, and 13% of annotated transcripts. Temporally, transcripts from the Aquificae were more abundant in 2013 and 2014 compared to 2015 (Fig. 1B). The active microbial community at Marker 33 included the Epsilonbacteraeota *Caminibacter*, which totaled on average nearly 20% of annotated transcripts. Methanogens were active at Marker 33 and more so at Marker 113, where the methanogens ranged from 30% to 56% of annotated transcripts from 2013 to 2015 (Fig. 1B). In 2014, *Sulfurovum* was abundant and active at Marker 113, comprising 17% of annotated transcripts. This percentage was an increase compared to 2013 and 2015, where *Sulfurovum* transcripts only totaled 7% and 5% respectively (Fig. 1B). One obvious difference between the taxonomy of the metagenomes (Fig. 1A) and the taxonomy of the transcripts (Fig. 1B) was that annotated reads classified to Gammaproteobacteria were on average more abundant in the metagenomes compared to the metatranscriptomes. The opposite pattern was observed for the Epsilonbacteraeota which made up a smaller percentage of the annotated reads in the metagenomes but were more prominent in the transcript data (Fig. 1B).

Hierarchical clustering was used to understand the degree of similarity between samples based on the normalized abundance of all functional genes identified via the KO database. Hierarchical clustering of metagenomes showed that in general metagenomes separated by

sampling location, although there were exceptions (Fig. 2A). All Marker 113 metagenomes clustered together as well as all Marker 33 metagenomes. The 2015 Anemone metagenome separated from the 2013 and 2014 Anemone metagenomes, which clustered with the Anemone plume sample. The background seawater metagenome was the outlier. The metatranscriptomes also clustered by vent, although more variability was seen in comparison to the metagenomes, as observed through greater correlation distances (Fig. 2B). Temporal variability in the metatranscriptomes among the vent samples was also observed, with all three 2015 vent metatranscriptomes clustered separately from other years at each vent. The

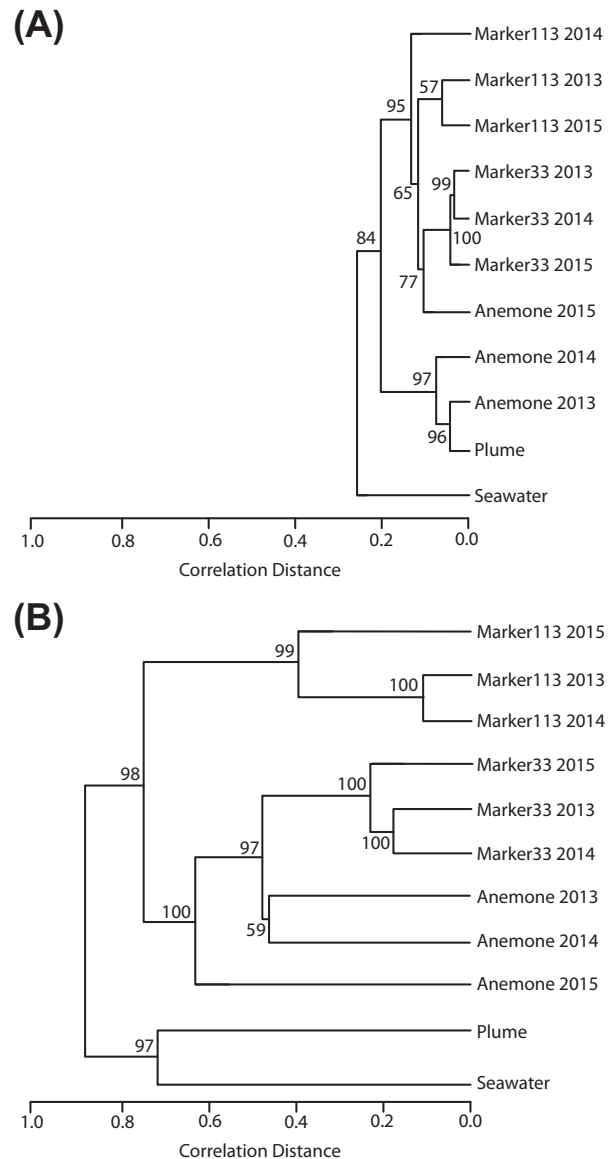


Fig. 2. Hierarchical clustering of (A) metagenomes and (B) metatranscriptomes based on the normalized abundance of genes or transcripts.

2015 samples were taken post-eruption and these differences in transcription profiles may indicate overall shifts in vent conditions. However, the 2015 samples still group more closely with the corresponding vent metatranscriptomes from previous years rather than grouping together, highlighting the spatial differences between the vents sites.

To better understand the relationship between microbial functional profiles and vent chemistry, we performed statistical analyses to compare the nine vent metagenomes and metatranscriptomes to environmental parameters. The seawater and plume samples were not included in this analysis. Environmental parameters used in analyses included temperature, pH, magnesium, dissolved hydrogen gas, hydrogen sulfide, methane, and nitrate. Magnesium, a conservative tracer, was included as an indicator of the degree of mixing between seawater and hydrothermal fluid at each vent. Functional diversity in metagenomes, as measured by normalized abundance of all genes annotated to the KO database, was most influenced by differences in pH, magnesium, hydrogen, nitrate, and methane ($\rho = 0.672$, Table 2). A Canonical Correspondence Analysis (CCA) showed these specific variables explained 76.1% of the functional patterns observed (Table 2). A CCA using all seven environmental variables showed that 90.8% of the functional variability could be explained (Table 2, Supporting Information Fig. S9). For the metatranscriptomes, pH was the most influential factor ($\rho = 0.721$). CCA analysis showed that 26.2% of the functional difference seen across the metatranscriptomes could be explained by pH, while a CCA including all environmental variables showed that 90.8% of functional differences could be explained (Table 2, Supporting Information Fig. S9).

Across vents and years, differences in the abundance and transcription of annotated genes for key metabolic pathways were apparent (Fig. 3). For nitrogen metabolism genes, the nitrogen-fixing gene *nifH* was most abundant at Marker 113, although transcription was limited (Fig. 3). The *nifH* genes at Marker 113 were mostly classified to thermophilic and hyperthermophilic methanogenic archaea (Supporting Information Table S2). However previous studies have indicated that few *nifH* genes are functional in these species (Mehta *et al.*, 2003; Ver Eecke *et al.*, 2013).

The abundance of annotated genes for denitrification was similar across vent samples, but transcription patterns differed. Annotated transcripts for the entire denitrification pathway were highest at Anemone and were almost entirely attributed to Epsilonbacteraeota (Supporting Information Table S3). In contrast, the only denitrification pathway gene that was highly expressed in background seawater was *nirK* and these annotated transcripts were classified to Thaumarchaeota and specifically to ammonia-oxidizing archaeal species (Supporting Information Table S3). In addition, annotated transcripts for the nitrate reductase gene, *narG* were also abundant in background seawater and were attributed to the Planctomycetes (Supporting Information Table S3). Transcripts affiliated with methanogenesis genes were most abundant at Marker 113. The abundance and transcription of Group 1 hydrogenase genes was similar across vents, with the exception of the 2014 Marker 113 sample where there was a large increase in the abundance of the Group 1 hydrogenases mostly classified to *Sulfurovum* (Figs 1 and 3, Supporting Information Tables S2 and S3). Sulfur metabolism genes and transcripts, specifically those associated with sulfur oxidation, were more abundant at Anemone and in the plume compared to the other two vents. These annotated sulfur oxidation genes and transcripts were mostly classified to Gammaproteobacteria (Supporting Information Tables S2 and S3). The most abundant carbon fixation pathway at each of the three vent sites was the reductive TCA cycle (rTCA) (Supporting Information Fig. S3), with most annotated rTCA genes mapping to Epsilonbacteraeota as well as to the phylum Aquificae. Annotated transcripts for genes associated with the Calvin cycle were abundant in the plume and were attributed to Gammaproteobacteria, specifically SUP05. Annotated genes and transcripts for the reductive Acetyl-CoA pathway, attributed predominantly to methanogenic archaea, were most abundant at Marker 113.

Genomic binning of each metagenome produced numerous metagenome assembled genomes (MAGs). Here, we only present analysis of MAGs classified to abundant taxonomic groups across multiple vent samples, including SUP05, Aquificales, *Sulfurovum*, and methanogenic archaea. These groups are also known to carry out

Table 2. BV-STEP and BIO-ENV Spearman rank coefficients (ρ) showing correlation between the functional profiles of diffuse vent communities (KO abundance) and environmental variables. Environmental variables were used in canonical correspondence analysis (CCA) to determine the percent of functional variability explained. CCAs were run using key variables determined via BV-STEP analysis as well as all environmental variables.

	BV-STEP factors	ρ	BIO-ENV factors	ρ	Variability explained BV-STEP factors	Variability explained All variables
Metagenomes	pH, Magnesium, Hydrogen Nitrate, Methane	0.672	pH Hydrogen	0.472 0.449	76.1%	90.8%
Metatranscriptomes	pH	0.721	pH Hydrogen	0.721 0.329	26.2%	90.8%

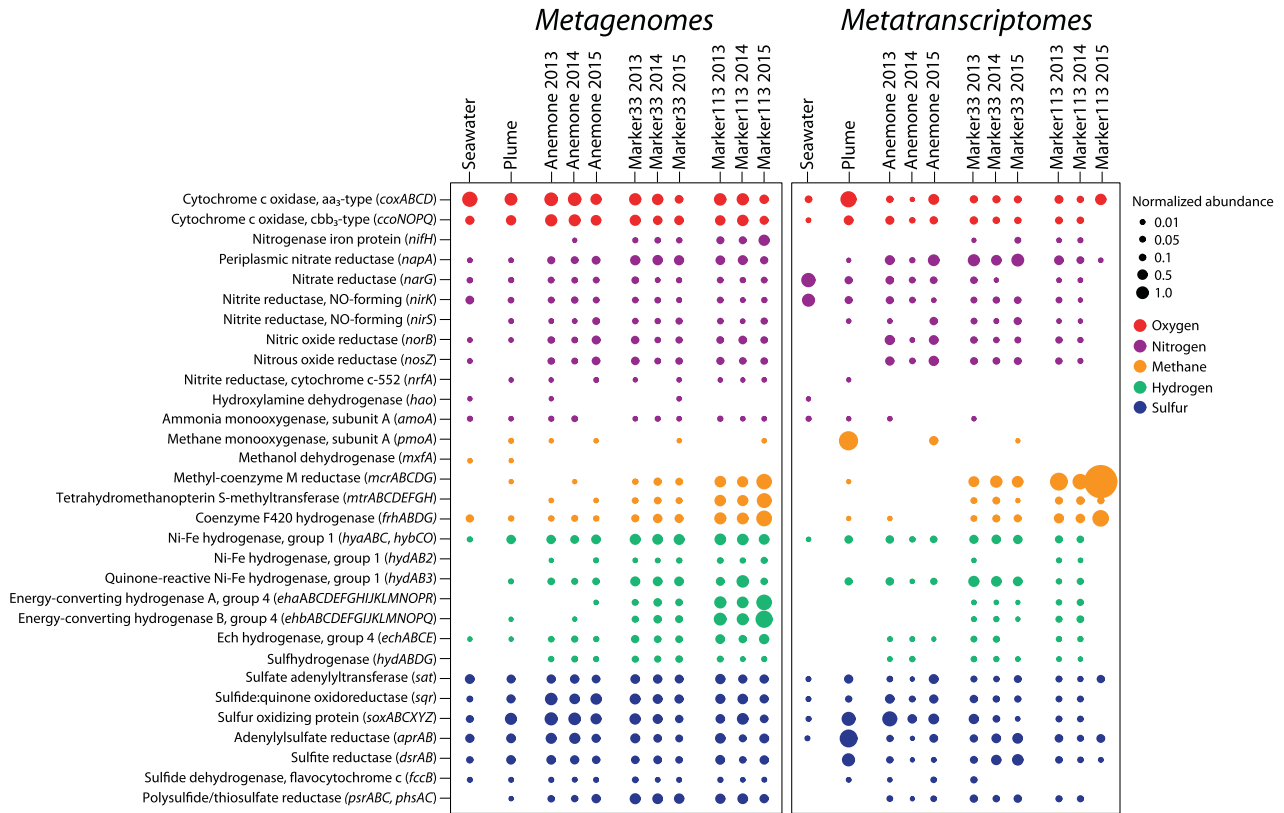


Fig. 3. Normalized abundance and expression of key genes for oxygen, nitrogen, methane, hydrogen and sulfur metabolisms within seawater, hydrothermal plume and at each vent site over three sampling years.

key metabolic processes such as sulfur and hydrogen oxidation and methanogenesis and are important players in the functional fingerprint at each vent. Heatmaps of the mean coverage for each MAG across metatranscriptomes show distinct patterns of gene transcription for each taxonomic group (Fig. 4). Heatmaps of mean coverage for each MAG across metagenomes show similar patterns (Supporting Information Fig. S4). For potential SUP05 MAGs, the heatmap shows that the MAGs cluster into two groups based on coverage (Fig. 4A). There is one group of SUP05 MAGs that is most transcribed at Anemone and also transcribed, to a lesser extent, at Marker 33 and in the plume, while a second group is only transcribed in high coverage in the plume. This separation in mean coverage between two potential SUP05 populations is also evident on a concatenated marker gene tree with other SUP05 related genomes (Supporting Information Fig. S5), with the first, more broadly distributed group closely related to SUP05 genomes from hydrothermal plumes and the second, plume-only group closely related to a MAG from a Lau basin hydrothermal plume as well as sequenced SUP05 genomes from oxygen minimum zones. Mean coverage of *Sulfurovum* MAGs was variable across vent metatranscriptomes, with little to no transcription in the

plume or seawater. *Sulfurovum* MAGs did not depend on vent location and only a few MAGs exhibited high gene transcription (Fig. 4B). The concatenated marker tree also showed no observable patterns across the MAGs (Supporting Information Fig. S6).

For the nine potential Aquificales MAGs, six MAGs exhibited gene transcription at Anemone while the other three were only transcribed at Marker 33 (Fig. 4C). The concatenated marker tree for these MAGs also revealed a phylogenetic pattern, with separation of MAGs based on family (Supporting Information Fig. S7). In the tree, two groups of Aquificales were observed, five MAGs clustering with genomes from the Desulfurobacteriaceae family and four MAGs clustering closely to sequenced genomes of the Aquificaceae family. The Desulfurobacteriaceae MAGs were highly transcribed both at Marker 33 and Anemone, while the second group of MAGs, possibly related to the Aquificaceae family, was only transcribed at Anemone (Fig. 4C, Supporting Information Fig. S7). Within the archaea, the methanogen MAGs identified clustered with sequenced genomes from mesophilic, thermophilic, and hyperthermophilic groups (Supporting Information Fig. S8). The heatmap of MAGs showed the most highly transcribed MAGs were observed at Marker 113.

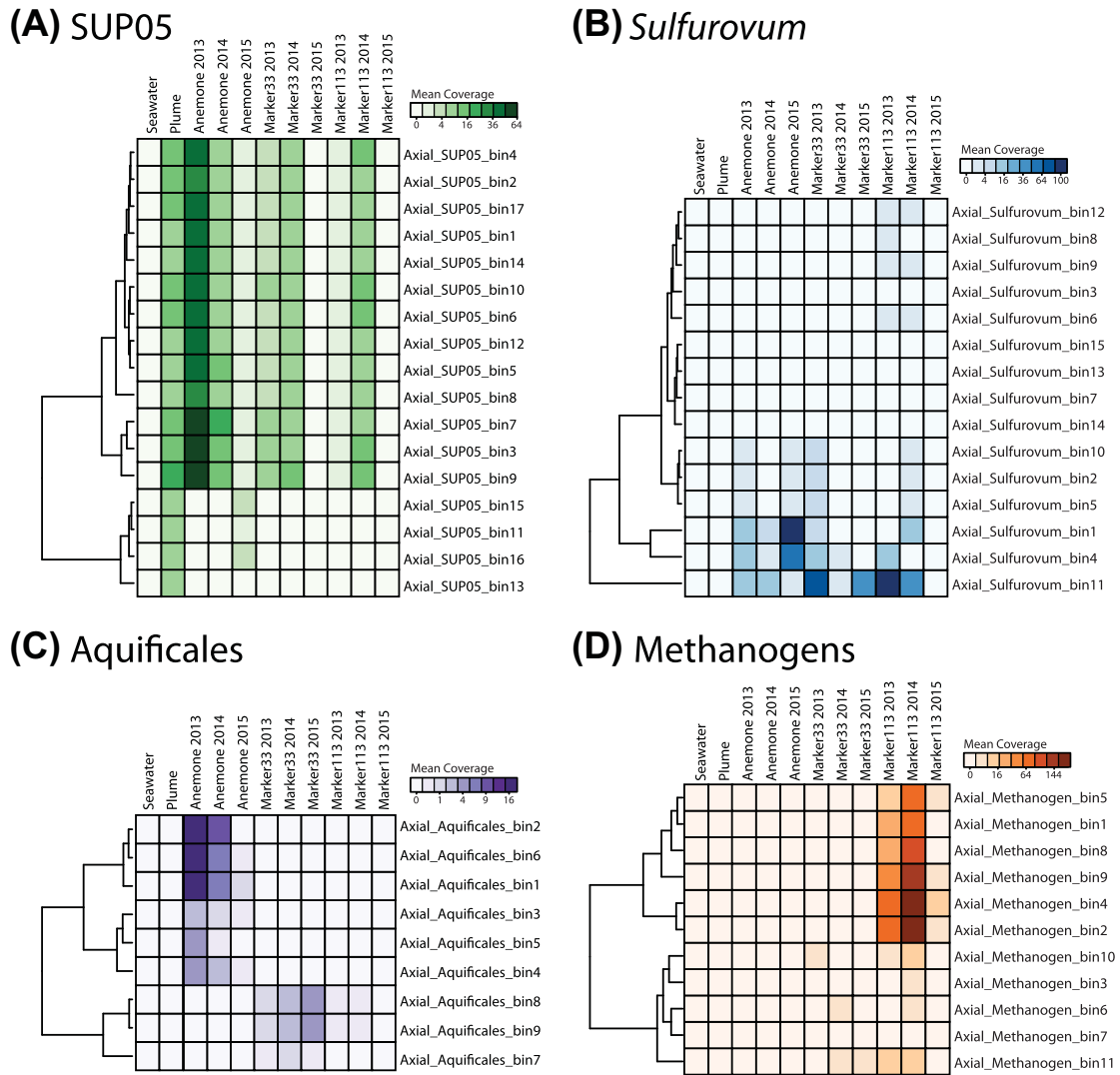


Fig. 4. Heatmaps of mean coverage across metatranscriptomes of metagenome assembled genomes (MAGs) identified as (A) SUP05, (B) *Sulfurovum*, (C) Aquificales and (D) methanogens. The dendrogram on each heatmap depicts the degree of similarity among MAGs based on mean coverage of metatranscriptomic reads across samples. Scales depict range of mean coverage across MAGs.

Methanogen MAGs were also observed at Marker 33, but were much less transcribed. The potential methanogen MAGs clearly separated by genus on a concatenated gene tree with sequenced genomes (Supporting Information Fig. S8). On the tree, the two *Methanocaldococcus* MAGs clustered with the sequenced *Methanocaldococcus bathoardescens* JH146 isolated from Marker 113 vent (Ver Eecke *et al.*, 2013; Stewart *et al.*, 2015).

Discussion

In this study, we expanded upon previous work at Axial Seamount focused on phylogenetic composition of microbial communities in venting fluids and used metagenomics and metatranscriptomic analyses to assess the spatial and temporal variability of microbial community composition,

metabolic potential, gene expression, and abundant populations across three low-temperature diffuse vents over three consecutive years. Our results show that over time, individual vent sites maintain microbial communities and populations with functional gene potential and expression profiles that are distinct to each site. The findings suggest that local seafloor geochemical and physical characteristics select for metabolic functions carried by different lineages, resulting in distinct microbial populations at each vent. This study is the first to assess both the taxonomic and metabolic fingerprints of microbial communities from diffuse fluids across multiple vents and years and lends further insight into the connection between microbial metabolisms, fluid chemistry, and microbial population dynamics at and below the seafloor.

The three diffuse fluid vents sampled in this study, Anemone, Marker 33, and Marker 113 all have distinct chemical signatures, and a statistical comparison of chemical variables to the functional gene profiles of each vent showed the influence of compounds like pH, dissolved hydrogen gas, and nitrate on the microbial metabolic potential and gene expression patterns across vents (Table 2). Both dissolved hydrogen and hydrogen sulfide are delivered by deep, hot source fluids ascending and mixing with nitrate- and oxygen-rich deep seawater in a mixing zone that constitutes the subseafloor microbial habitat. Both hydrogen and nitrate are highly reactive in the subseafloor mixing zone (Butterfield *et al.*, 2004; Bourbonnais *et al.*, 2012a,b) and their wide range of concentration across the three vent sites indicates variation in the extent of mixing and reaction among the sites. The concentrations of redox reactants hydrogen, hydrogen sulfide, nitrate, and oxygen influence the energy available for microbial metabolism in these environments (e.g., Amend *et al.*, 2011). Hydrogen sulfide is present at high concentrations in diffuse fluids at Axial and is not the limiting redox reactant even in mixtures that are 95%–98% seawater (Butterfield *et al.*, 2004, Table 1). This may explain the very weak correlation of hydrogen sulfide with the metagenomes and metatranscriptomes, $\rho = 0.195$ and $\rho = -0.147$ respectively. Dissolved hydrogen gas concentrations were low at Marker 33 (average 1 μM) and lower at Marker 113 (average 0.9 μM) (Table 1), and these two sites hosted a higher abundance and expression of various hydrogenases as well as a higher percentage of hydrogen-utilizing thermophilic Epsilonbacteraeota hyperthermophilic Aquificaceae, and methanogenic archaea compared to Anemone vent. These results indicate that the low dissolved hydrogen concentrations observed at Marker 113 and Marker 33 is likely due to a draw down of hydrogen for use in hydrogen oxidation and methanogenesis. Comparatively, genes for these two metabolic processes were not highly expressed at Anemone vent, where dissolved hydrogen gas concentrations were an order of magnitude higher (Table 1, Fig. 3). As with all transcript data, however, expression of hydrogen oxidation and methanogenesis genes does not quantitatively indicate activity of these proteins. Although methane concentrations were variable across years, methane was, on average, higher at Marker 113 (26 μM) than Anemone (21 μM), giving further indication of the role of methanogenesis in the drawdown of hydrogen. Similar to previous work on the importance of hydrogen in diverse hydrothermal vent ecosystems (Takai *et al.*, 2004; Amend *et al.*, 2011; Brazelton *et al.*, 2012; Ver Eecke *et al.*, 2012; Reveillaud *et al.*, 2016), our findings reinforce the critical role of dissolved hydrogen gas in structuring microbial communities at Axial Seamount, particularly in the warm, anoxic subseafloor habitat.

Although both our results and previous work at Axial indicate chemistry is clearly an important determinant of microbial community structure (Huber *et al.*, 2002; 2003; 2007; Opatkiewicz *et al.*, 2009; Akerman *et al.*, 2013; Anderson *et al.*, 2013), it is certain that diffuse flow chemistry and resulting energetics are not the only factor controlling the distribution and metabolism of microbes (Larson *et al.*, 2015). This is particularly important in the subseafloor, where fluid flow rates, fluid residence times, and the physical form of each vent likely play a role in structuring the subseafloor microbial habitat. Some of the large differences in chemistry and microbial communities between the isolated low-temperature, basalt-hosted vents of Marker 33 and Marker 113 and the low sulfide mound with high-temperature fluid of Anemone are very likely due to the physical extent of the mixing zone within the seafloor. In a study by Akerman *et al.* (2013) at Axial Seamount, 16S rRNA gene sequencing data of Epsilonbacteraeota showed that communities grouped by both individual vent sample and either basalt or sulfide-hosted diffuse venting. Our work expands these findings beyond single-gene approaches and shows that microbial metabolisms and populations vary depending on both the geologic setting and the chemistry of the host site.

Based on our meta-omics results and building on previous work at Axial, we constructed a conceptual model for each vent, as well as for the hydrothermal plume, depicting the dominant microbial metabolic processes occurring under the different geochemical conditions at each site (Fig. 5). The model depicts a simplified representation of the most abundant microbially-mediated processes occurring at and below the seafloor and lends insight into the numerous differences between the microbial metabolic fingerprints and population dynamics present at each site. At Anemone, which vents from a low sulfide mound, we saw large temperature fluctuations as we sampled fluids. This implies very hot fluids are mixing in the very shallow seafloor and are immediately exiting, resulting in a short residence time of the mixed fluid. The chemistry at Anemone supports this, as dissolved hydrogen gas is highest at this vent, corresponding to underutilization of this compound by microbes. Consequently, a more thermophilic anaerobic subsurface microbial community has little time within the seafloor to establish and thus the resident community at Anemone is mainly, but not entirely, mesophilic and microaerophilic, likely living at the subseafloor-seafloor interface (Fig. 5). Mesophilic and psychrophilic Gammaproteobacteria, namely SUP05, as well as Epsilonbacteraeota were the dominant organisms present and genes for metabolisms like sulfur oxidation, which can be carried out under aerobic and microaerobic conditions, were most abundant at Anemone (Figs 1 and 3). These results are supported by a study of diffuse vent sites within the ASHES vent field, which showed similar community composition

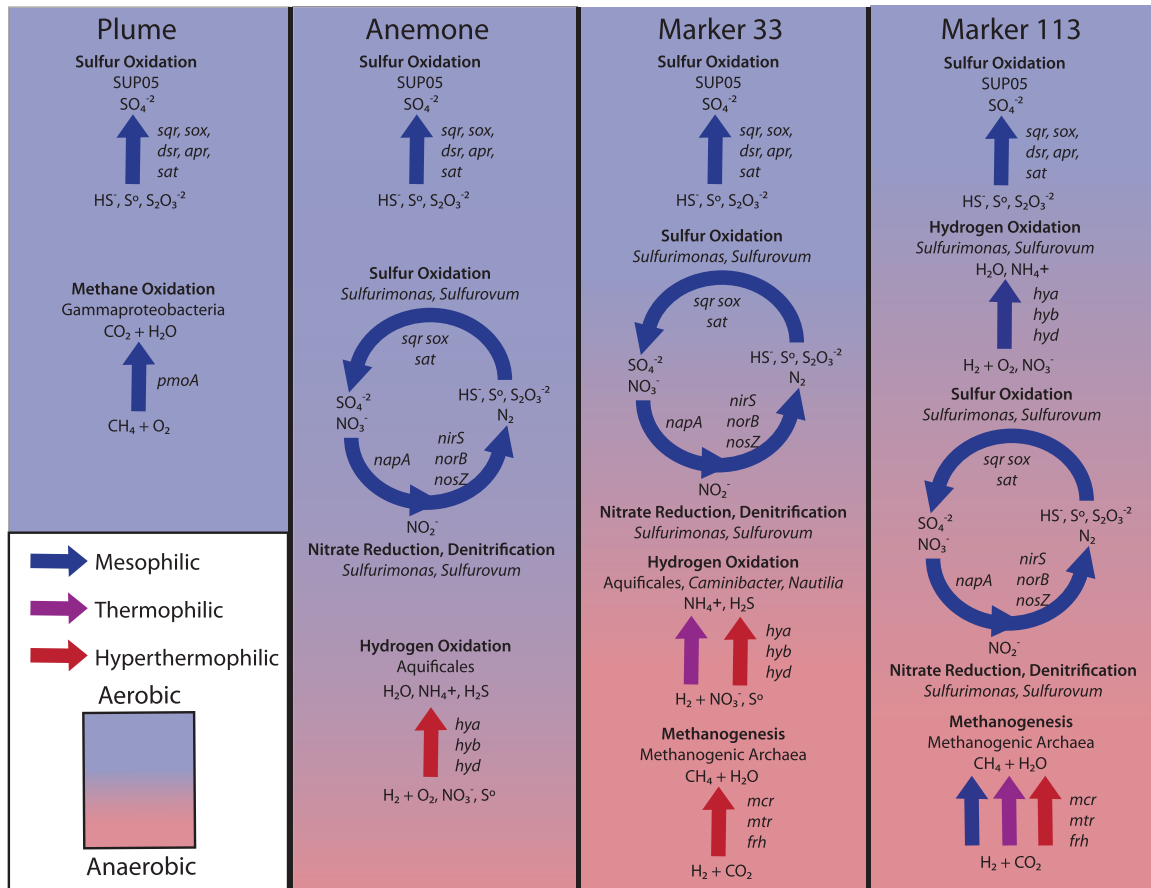


Fig. 5. Biogeochemical model of the dominant microbial metabolisms occurring at each vent, both above and below the seafloor, based on annotated gene and transcript abundance. A hydrothermal plume model is included for comparison.

and gene expression patterns as those observed at Anemone (Olins *et al.*, 2017). The genomic binning results also support this model, as SUP05 populations had higher gene expression levels at Anemone compared to other diffuse vents. SUP05 have been observed across many environments and determined to be important to biogeochemical cycles within oxygen minimum zones, hydrothermal plumes, and at diffuse vents (Canfield *et al.*, 2010; Anantharaman *et al.*, 2013; Anderson *et al.*, 2013; Dick *et al.*, 2013; Mattes *et al.*, 2013; Shah *et al.*, 2016; Meier *et al.*, 2017). Binning revealed two distinct groups of SUP05 populations, with the group most abundant at Anemone clustering with known SUP05 MAGs from the Guaymas Basin plume (Anantharaman *et al.*, 2013), while the second group, most abundant in the plume sample, clustered separately with a SUP05 MAG from Lau Basin (Anantharaman *et al.*, 2016) (Supporting Information Fig. S3). This separation indicates potentially location specific populations of SUP05 at Axial and implies possible temperature restrictions of SUP05 populations, with those abundant at vents able to tolerate the higher diffuse fluid temperature compared to cold background seawater.

At Marker 33, taxonomy of annotated reads in both the metagenomes and metatranscriptomes (Fig. 1) show the presence of higher temperature taxonomic groups such as thermophilic Epsilonbacteraeota, hyperthermophilic methanogens, and Aquificales at abundances higher than the other two vents. The presence of these groups, as well as genes for corresponding metabolisms like hydrogen oxidation, methanogenesis, and nitrate and sulfur reduction indicates the establishment of an active high temperature subsurface community compared to Anemone (Fig. 5). The percentage of seawater in Marker 33 samples (85%–92%) is also consistent with potentially hotter temperatures in the subsurface reservoir. Genomic binning resulted in the reconstruction of nine MAGs, at various levels of completeness, from the order Aquificales, a group of chemolithoautotrophic sulfur and hydrogen oxidizers that are thermophilic as well as hyperthermophilic, and can grow under anaerobic or microaerophilic conditions (Reysenbach *et al.*, 2001). The Aquificales MAGs were vent specific and only showed expression at Marker 33 and Anemone. Based on a tree of concatenated marker genes, two separate phylogenetic groups were observed, with four MAGs most

closely related to known genomes from the Aquificaceae family while the other five MAGs clustered with sequenced genomes from the family Desulfurobacteriaceae. A defining feature between the two families is that sequenced representatives from the Aquificaceae family are microaerophilic (Deckert *et al.*, 1998; Nunoura *et al.*, 2008) while those from the Desulfurobacteriaceae are strict anaerobes (L'Haridon *et al.*, 1998; Vetriani *et al.*, 2004; Stewart *et al.*, 2016). At Marker 33, all MAGs were closely related to the Desulfurobacteriaceae family, whereas the most abundant three MAGs from Anemone were related to the family Aquificaceae. The presence of microaerobic hydrogen oxidizing Aquificaceae at Anemone and not at Marker 33 is consistent with our conceptual model of seafloor conditions, with oxic conditions at Anemone favoring more aerobic metabolisms whereas Marker 33 has more reducing conditions, which favors the establishment of anaerobic hyperthermophiles like those from the Desulfurobacteriaceae family. Aquificales MAGs ranged in completeness from 15% to 69%, so broad functional differences could not be discerned, but it is clear from the gene expression patterns and placement on the marker tree that there are different populations present at Anemone and Marker 33 vents.

Marker 113 presents yet another possible setting for seafloor microbial communities. At Marker 113, anaerobic metabolisms like methanogenesis and sulfur reduction are dominant and methanogenesis is carried out by mesophilic and thermophilic methanogens, with few high temperature groups present. This suggests the establishment of both mesophilic and thermophilic anaerobic seafloor communities and utilization of dissolved hydrogen gas and generation of methane within the seafloor. Marker 113 had the lowest hydrogen and the highest methane concentrations, on average, of all three vents. The consistently higher percentage of seawater in Marker 113 fluids (> 96%, Table 1) compared to the other vents points to potentially cooler temperatures in the seafloor reservoir. Similar to Aquificales, binning of methanogenic archaea showed a vent specific pattern among MAGs, with mesophilic and thermophilic genera *Methanococcus* and *Methanothermococcus* at Marker 113 and the hyperthermophilic genus *Methanocaldococcus* at Marker 33, with few methanogens detected at Anemone. This separation in methanogenic genera between vents is again consistent with a more stable seafloor community developed at Marker 33 and 113, allowing methanogens to thrive, whereas the conditions at Anemone do not give enough time for these obligate anaerobes to grow. Binning results are consistent with taxonomic patterns observed in the 16S and functionally annotated reads. Our results are also consistent with previous studies at Marker 113, which used growth experiments, stable isotope probing, quantitative PCR, and meta-omic analyses to show diverse and active methanogens at

this site (Mehta and Baross, 2006; Ver Eecke *et al.*, 2012; Fortunato and Huber, 2016; Topçuoğlu *et al.*, 2016).

The combination of geochemical, metabolic potential, gene expression, and genomic binning data allowed us to identify key lineages carrying out important biogeochemical transformations in the diffuse hydrothermal vent environment, as well as depict potentially different habitats at and below the seafloor for each vent. Together, these results show that within one hydrothermally active seamount, individual vent sites can maintain microbial communities and populations that are both taxonomically and functionally distinct from one another and highlights the importance of local seafloor geochemical and physical characteristics for supporting these communities. For the seafloor, our results provide particularly important insights into the conditions under which microbes establish and thrive in the crustal environment. It is clear that the chemistry of the reduced hydrothermal fluid is important for microbes living beneath the seafloor and fluid dynamics and physical structuring are also essential for allowing microbes to exploit these energy sources and establish communities without being rapidly displaced from the seafloor habitat. Therefore, while predictions can be made about dominant microbial metabolisms based on host rock geological context (e.g., Amend *et al.*, 2011), fine-scale investigations reveal that the ability of an individual site to maintain an active seafloor population is intrinsically linked to the geochemistry of the system, the mixing conditions, and the physical constraints of an individual vent site. Such parameters are difficult to measure, but are key to understanding the establishment, diversification, stability, and connectivity of the seafloor habitat. While investigations spanning multiple sites will be necessary to put together a comprehensive picture of the role of microbial communities in venting fluids with respect to global biogeochemistry and the ocean seed bank, the data presented here provide the first glimpse of population dynamics for biogeochemically relevant microbial lineages in these diverse and underexplored ecosystems.

Experimental procedures

Sample collection

Diffuse hydrothermal vent fluids were collected at Axial Seamount in September 2013, August 2014, and August 2015 on board the *R/V Falkor* and *R/V Thompson* in 2013, *R/V Brown* in 2014, and *R/V Thompson* in 2015 using ROVs *ROPOS* and *JASON*. The 2015 sampling occurred approximately 5 months after Axial Seamount erupted. Diffuse fluids were sampled at three vents, Anemone, Marker 33 and Marker 113 during all three years of sampling (Supporting Information Fig. S1). Latitude and longitude for each vent are Anemone: N 45.9332, W 130.0137; Marker 33: N 45.9332, W 129.9822; Marker 113: N 45.9227, W 129.9882. For DNA and RNA collection, 3 L of diffuse fluid was pumped at a rate of 100–150 ml

min⁻¹ through a 0.22 µm, 47 mm GWSP filter (Millipore) using the Hydrothermal Fluid and Particle Sampler (HFPS) (Butterfield *et al.*, 2004). All filters were preserved *in situ* with RNALater (Ambion) as previously described in Akerman *et al.* (2013). In 2015, fluid from a hydrothermal plume approximately 42 m above Anemone vent (N 45.9335667, W 130.013667) and water from background seawater at 1500 m depth (N 46.27389, W 129.79548) were also collected using a Seabird SBE911 CTD and 10 L Niskin bottles (Supporting Information Fig. S1). From the plume and seawater samples, water was transferred from Niskin bottles into cubitainers and 3 L were filtered through 0.22 µm Sterivex filters (Millipore). The HFPS has an integrated temperature sensor at the point of intake to allow for constant temperature monitoring while drawing fluids. Fluid samples were also collected for chemistry and analyzed for magnesium, dissolved hydrogen gas, nitrate and methane concentrations following methods described in Butterfield *et al.* (2004). The pH was measured at each vent using a deep-sea glass pH electrode (AMT) plumbed to the HFPS.

Metagenomic and metatranscriptomic library preparation

Both DNA and RNA were extracted from each 0.22 µm, 47 mm or Sterivex filter and metagenomic and metatranscriptomic library preparation was carried out as described in detail in Fortunato and Huber (2016). Briefly, RNA was extracted using the mirVana miRNA isolation kit (Ambion). Ribosomal RNA removal, cDNA synthesis, and metatranscriptomic library preparation was carried out using the Ovation Complete Prokaryotic RNA-Seq DR multiplex system (Nugen) following manufacturer instructions. DNA was extracted using a phenol-chloroform method adapted from Crump *et al.* (2004) and Zhou *et al.* (1996) and metagenomic library construction was completed using the Ovation Ultralow Library DR multiplex system (Nugen) following manufacturer instructions. For 2013–2014 samples, sequencing was completed on an Illumina HiSeq 1000. For 2015 samples sequencing was completed on a NextSeq 500. Prior to library construction, DNA and cDNA was sheared using a Covaris S-series sonicator; 175 bp for 2013–2014, 275 bp for 2015. All libraries were paired-end, with a 30 bp overlap, resulting in an average merged read length of 160 bp for 2013–2014 libraries and 275 bp for 2015 libraries. Metagenomic and metatranscriptomic sequencing was carried out at the W.M. Keck sequencing facility at the Marine Biological Laboratory.

Metagenomic and metatranscriptomic analysis

For metagenome and metatranscriptome analysis, paired-end partially overlapping reads were merged and quality filtered using custom Illumina utility scripts (<https://github.com/meren/illumina-utils>). For each metagenomic library, merged reads were dereplicated then assembled using CLC Genomics Workbench (v 7.0) using default settings and a minimum contig length of 200 bp. Dereplicated libraries were only used for easing assembly, mapping was completed using all reads. Assembled contigs from each metagenomic library were submitted to the DOE Joint Genome Institute's Integrated Microbial Genome Metagenomic Expert Review (IMG/MER) annotation pipeline for Open Reading Frame (ORF) identification and functional and taxonomic annotation (Markowitz *et al.*,

2012). ORFs from each metagenomic library were annotated against the KEGG orthology (KO) database. Only annotations meeting the minimum requirements of an e-score of $1e^{-10}$, 30% amino acid identity and alignment length of 40 amino acids were included in functional analyses. To determine the number of reads per annotated ORF, reads from each metagenomic library were mapped to ORFs using Bowtie2 (v2.0.0-beta5, Langmead and Salzberg, 2012), with end to end alignment. For metatranscriptomes, non-rRNA transcripts were mapped to ORFs identified in the corresponding metagenome for each sample. After mapping, abundances for each ORF were normalized to gene or transcript length. Metagenomes were normalized to number of Reads per Kilobase per Genome (RPKG). Number of genomes per metagenome was estimated using hits to the single-copy gene, DNA-directed RNA polymerase beta subunit gene (*rpoB*). Metatranscriptomes were normalized to number of Transcripts Per Million reads (TPM) to allow for a more accurate comparison between samples. The taxonomy for each ORF was determined using the Phylogenetic Distribution tool in IMG, part of the IMG/MER annotation pipeline. This taxonomy was used to construct Fig. 1A and B. Normalized KO abundances were used for bubble plot construction and hierarchical clustering. Hierarchical clustering of samples (average-linkage method) was carried out using the statistical program *R* (v3.2.1 R-Development-Core-Team, 2011) and the *R* package *pvclust*. Sequencing and analysis statistics for all metagenomes and metatranscriptomes are listed in Supporting Information Table S1.

Ribosomal RNA identification and classification

For the metatranscriptomes, reads were mapped to SILVA SSU and LSU databases (release 123, Pruesse *et al.*, 2007) using Bowtie2 with a local alignment and default settings in order to identify all remaining rRNA reads. Identified rRNA reads were separated from each metatranscriptome using publicly available Perl scripts (<http://alrllab.research.pdx.edu/Aquificales/>). Ribosomal RNA from each metagenome was also identified but reads were not separated. After rRNA reads were identified in each metagenome, 16S rRNA reads were specifically targeted using the extensive Greengenes 16S rRNA taxonomic database (August 2013 release, McDonald *et al.*, 2012) using Bowtie2. 16S rRNA reads were then taxonomically identified via MOTHUR (v1.37, Schloss *et al.*, 2009), again using the Greengenes taxonomic database.

Comparison of functional profiles and environmental data

Chemical and environmental data were compiled for the nine diffuse vent samples collected from 2013 to 2015. The environmental variables included in analysis with functional profiles were: temperature, pH, magnesium (Mg), dissolved hydrogen gas (H₂), methane (CH₄), nitrate (NO₃⁻), and hydrogen sulfide (H₂S). Using Primer v6 (PRIMER-E), BV-STEP analysis was carried out to identify sets of variables that potentially influenced functional variability among the metagenomes and metatranscriptomes ($\rho > 0.95$, $\Delta\rho < 0.001$, 100 restarts) (Clarke and Ainsworth, 1993). BIO-ENV was used to rank each individual environmental variable by degree of association with functional variability. BV-STEP and BIO-ENV use the Spearman rank correlation coefficient (ρ) to determine the degree of association

between similarity matrices of normalized KO abundances (Bray-Curtis similarity) and environmental data (Euclidean distances). Using the environmental variables identified by BV-STEP, we performed a Canonical Correspondence Analysis (CCA) to determine the percent of functional variability explained by environmental variables (ter Braak, 1986). CCA analyses were run in *R* (v3.3.2, R-Development-Core-Team, 2011) using the *R* package *vegan*.

Genomic binning and analysis

Two different assembly approaches were used in order to reconstruct genomes from metagenomic data. We found that different assembly approaches yielded different, more complete metagenome assembled genomes (MAGs). In the first assembly approach, each individual metagenome, eleven total, was assembled using the assembler Ray (v2.3.1, Boisvert *et al.*, 2012) with default settings. In the second assembly approach, multiple metagenomes were assembled to create co-assemblies for each of the three vents, Anemone, Marker 33, and Marker 113. For example, all metagenomes from Anemone vent were assembled together in a co-assembly. This second method was carried out using the assembler Megahit and the *meta-large* preset (Li *et al.*, 2015a). Reads from each metagenome were then mapped to every individual or co-assembly using Bowtie2 (Langmead and Salzberg, 2012). Manual, supervised binning was then carried out using Anvi'o (v1.2.0, Eren *et al.*, 2015). MAGs were selected manually using patterns in tetranucleotide frequency within a sample and differential coverage between samples. A completion cutoff of 15% and contamination cutoff of under 10% of MAG completeness was used. Only one of the assembly approaches (individual or co-assembly) was used in MAG selection for each of the four taxonomic groups analyzed. Individual vent assemblies resulted in the most complete MAGs for SUP05 and Aquificales, while co-assemblies from each vent resulted in the selection of more complete *Sulfurovum* and methanogen MAGs. For comparison, MAG completeness and contamination were also computed using CheckM (Parks *et al.*, 2015). Taxonomic assignments for MAGs were determined using Phylosift (v1.0.1, Darling *et al.*, 2014) and confirmed using CheckM. Concatenated alignments of 37 reference marker genes were generated for each MAG using Phylosift. These alignments were then used to construct amino acid based maximum likelihood trees for key taxonomic groups using RAxML. Maximum likelihood trees were constructed using the protein model PROTGAMMAWAG with 100 bootstraps and an outgroup as identified on each tree (Aziz *et al.*, 2008). Mean coverage for each MAG within each metagenome was calculated using Anvi'o (v1.2.0, Eren *et al.*, 2015). Mean coverage for each MAG within each metatranscriptome was calculated in CLC Genomics Workbench (v7.0). Heatmaps of mean coverage were constructed in *R* (v3.3.2, R-Development-Core-Team, 2011) using the package *heatmap3*.

Data deposition

All raw metagenomic and metatranscriptomic read data are publically available through the European Nucleotide

Archive (ENA) under study accession numbers PRJEB7866, PRJEB12000 and PRJEB19456 for 2013, 2014 and 2015 respectively. Assembled contigs for each metagenomic library are publically available via IMG/MER under submission numbers 74596, 74599, 74602, 71132, 71134, 71135, 78375, 78377, 78401, 78368 and 78372. Contigs for each individual MAG are available through FigShare at DOI 10.6084/m9.figshare.5151547.

Acknowledgements

This work was funded by the Gordon and Betty Moore Foundation Grant GBMF3297, the NSF Center for Dark Energy Biosphere Investigations (C-DEBI) (OCE-0939564), contribution number 399, NOAA/PMEL, contribution number 4629, and JISAO under NOAA Cooperative Agreement NA15OAR4320063, contribution number 2017-0127. The data collected in this study includes work supported by the Schmidt Ocean Institute during cruise FK010–2013 aboard *R/V Falkor*. We thank the captains and crews of the *R/V Falkor*, *R/V Thompson* and *R/V Brown* as well as the ROV *ROPOS* and *JASON* groups. Chris Algar, Lisa Zeigler Allen, Rika Anderson, Jim Holden, Giora Proskurowski, Kevin Roe, Lucy Stewart, Begum Topçuoğlu, and Joe Vallino provided critical support in sample collection, experimental design and execution, and interpretation of results.

Conflict of Interest

All authors declare no conflicts of interest.

References

- Akerman, N.H., Butterfield, D.A., and Huber, J.A. (2013) Phylogenetic diversity and functional gene patterns of sulfur-oxidizing subseafloor Epsilonproteobacteria in diffuse hydrothermal vent fluids. *Front Microbiol* **4**: 185.
- Amend, J.P., McCollom, T.M., Hentscher, M., and Bach, W. (2011) Catabolic and anabolic energy for chemolithoautotrophs in deep-sea hydrothermal systems hosted in different rock types. *Geochim Cosmochim Acta* **75**: 5736–5748.
- Anantharaman, K., Breier, J.A., and Dick, G.J. (2016) Metagenomic resolution of microbial functions in deep-sea hydrothermal plumes across the Eastern Lau Spreading Center. *ISME J* **10**: 225–239.
- Anantharaman, K., Breier, J.A., Sheik, C.S., and Dick, G.J. (2013) Evidence for hydrogen oxidation and metabolic plasticity in widespread deep-sea sulfur-oxidizing bacteria. *Proc Natl Acad Sci U S A* **110**: 330–335.
- Anantharaman, K., Duhaime, M.B., Breier, J.A., Wendt, K.A., Toner, B.M., and Dick, G.J. (2014) Sulfur oxidation genes in diverse deep-sea viruses. *Science* **344**: 757–760.
- Anderson, R.E., Beltran, M.T., Hallam, S.J., and Baross, J.A. (2013) Microbial community structure across fluid gradients in the Juan de Fuca Ridge hydrothermal system. *FEMS Microbiol Ecol* **83**: 324–339.
- Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., *et al.* (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**: 1–15.

- Boisvert, S., Raymond, F., Godzaridis, É., Laviolette, F., and Corbeil, J. (2012) Ray Meta: scalable de novo metagenome assembly and profiling. *Genome Biol* **13**: 1–13.
- Bourbonnais, A., Lehmann, M.F., Butterfield, D.A., and Juniper, S.K. (2012a) Subseafloor nitrogen transformations in diffuse hydrothermal vent fluids of the Juan de Fuca Ridge evidenced by the isotopic composition of nitrate and ammonium. *Geochem Geophys Geosyst* **13**: n/a-n/a.
- Bourbonnais, A., Juniper, S.K., Butterfield, D.A., Devol, A.H., Kuypers, M.M.M., Lavik, G., *et al.* (2012b) Activity and abundance of denitrifying bacteria in the subsurface biosphere of diffuse hydrothermal vents of the Juan de Fuca Ridge. *Biogeosciences* **9**: 4661–4678.
- Brazelton, W.J., Nelson, B., and Schrenk, M.O. (2012) Metagenomic evidence for H₂ oxidation and H₂ production by serpentinite-hosted subsurface microbial communities. *Front Microbiol* **2**: 16.
- Butterfield, D.A., Roe, K.K., Lilley, M.D., Huber, J.A., Baross, J.A., Embley, R.W., and Massoth, G.J. (2004) Mixing, reaction and microbial activity in the sub-seafloor revealed by temporal and spatial variation in diffuse flow vents at axial volcano. In *The Subseafloor Biosphere at Mid-Ocean Ridges* Wilcock W.S.D., Delong, E.F., Kelley D.S., Baross J.A. and Cary S.C. (eds). Washington, DC, USA: American Geophysical Union, pp. 269–289.
- Canfield, D.E., Stewart, F.J., Thamdrup, B., De Brabandere, L., Dalsgaard, T., Delong, E.F., *et al.* (2010) A cryptic sulfur cycle in oxygen-minimum-zone waters off the Chilean coast. *Science* **330**: 1375–1378.
- Chadwick, W.W., Butterfield, D.A., Embley, R.W., Tunnicliffe, V., Huber, J.A., Noonan, S.L., and Clague, D.A. (2010) Axial seamount. *Oceanography* **23**: 38–39.
- Clarke, K.R., and Ainsworth, M. (1993) A method of linking multivariate community structure to environmental variables. *Marine Ecology-Progress Series* **92**: 205–219.
- Crump, B.C., Hopkinson, C.S., Sogin, M.L., and Hobbie, J.E. (2004) Microbial biogeography along an estuarine salinity gradient: Combined influences of bacterial growth and residence time. *Appl Environ Microbiol* **70**: 1494–1505.
- Darling, A.E., Jospin, G., Lowe, E., Matsen, F.A.I.V., Bik, H.M., and Eisen, J.A. (2014) PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ* **2**: e243.
- Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E., *et al.* (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**: 353–358.
- Dick, G.J., Andersson, A.F., Baker, B.J., Simmons, S.L., Thomas, B.C., Yelton, A.P., and Banfield, J.F. (2009) Community-wide analysis of microbial genome sequence signatures. *Genome Biol* **10**: 50.
- Dick, G.J., Anantharaman, K., Baker, B.J., Li, M., Reed, D.C., and Sheik, C.S. (2013) The microbiology of deep-sea hydrothermal vent plumes: ecological and biogeographic linkages to seafloor and water column habitats. *Front Microbiol* **4**: 124.
- Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., and Delmont, T.O. (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**: e1319.
- Fortunato, C.S., and Huber, J.A. (2016) Coupled RNA-SIP and metatranscriptomics of active chemolithoautotrophic communities at a deep-sea hydrothermal vent. *ISME J* **10**: 1925–1938.
- Gonnella, G., Böhnke, S., Indenbirken, D., Garbe-Schönberg, D., Seifert, R., Mertens, C., Kurtz, S., and Perner, M. (2016) Endemic hydrothermal vent species identified in the open ocean seed bank. *Nat Microbiol* **1**: 7.
- Huber, J.A., Butterfield, D.A., and Baross, J.A. (2002) Temporal changes in archaeal diversity and chemistry in a mid-ocean ridge subseafloor habitat. *Appl Environ Microbiol* **68**: 1585–1594.
- Huber, J.A., Butterfield, D.A., and Baross, J.A. (2003) Bacterial diversity in a subseafloor habitat following a deep-sea volcanic eruption. *FEMS Microbiol Ecol* **43**: 393–409.
- Huber, J.A., Cantin, H.V., Huse, S.M., Welch, D.B.M., Sogin, M.L., and Butterfield, D.A. (2010) Isolated communities of Epsilonproteobacteria in hydrothermal vent fluids of the Mariana Arc seamounts. *FEMS Microbiol Ecol* **73**: 538–549.
- Huber, J.A., Mark Welch, D., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield, D.A., and Sogin, M.L. (2007) Microbial population structures in the deep marine biosphere. *Science* **318**: 97–100.
- Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., *et al.* (2016) A new view of the tree of life. *Nat Microbiol* **1**: 16048.
- Jannasch, H.W., and Mottl, M.J. (1985) Geomicrobiology of deep-sea hydrothermal vents. *Science* **229**: 717–725.
- Karl, D.M. (1995) *The Microbiology of Deep-Sea Hydrothermal Vents*. Florida, USA: CRC Press.
- Kelley, D.S., Delaney, J.R., and Juniper, S.K. (2014) Establishing a new era of submarine volcanic observatories: Cabling Axial Seamount and the Endeavour Segment of the Juan de Fuca Ridge. *Mar Geol* **352**: 426–450.
- L'Haridon, S., Cilia, V., Messner, P., Raguénès, G., Gambacorta, A., Sleytr, U.B., *et al.* (1998) Desulfurobacterium thermolithotrophum gen. nov., sp. nov., a novel autotrophic, sulphur-reducing bacterium isolated from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* **48**: 701–711.
- Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**: 357–359.
- Larson, B.I., Houghton, J.L., Lowell, R.P., Farough, A., and Meile, C.D. (2015) Subsurface conditions in hydrothermal vents inferred from diffuse flow composition, and models of reaction and transport. *Earth Planet Sci Lett* **424**: 245–255.
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015a) MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676.
- Li, M., Jain, S., Baker, B.J., Taylor, C., and Dick, G.J. (2014) Novel hydrocarbon monooxygenase genes in the metatranscriptome of a natural deep-sea hydrocarbon plume. *Environ Microbiol* **16**: 60–71.
- Li, M., Baker, B.J., Anantharaman, K., Jain, S., Breier, J.A., and Dick, G.J. (2015b) Genomic and transcriptomic evidence for scavenging of diverse organic compounds by widespread deep-sea archaea. *Nat Commun* **6**: 8933.
- Markowitz, V.M., Chen, I.M.A., Chu, K., Szeto, E., Palaniappan, K., Grechkin, Y., *et al.* (2012) IMG/M: the integrated metagenome data management and comparative analysis system. *Nucleic Acids Res* **40**: D123–D129.

- Mattes, T.E., Nunn, B.L., Marshall, K.T., Proskurowski, G., Kelley, D.S., Kawka, O.E., *et al.* (2013) Sulfur oxidizers dominate carbon fixation at a biogeochemical hot spot in the dark ocean. *ISME J* **7**: 2349–2360.
- McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., *et al.* (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* **6**: 610–618.
- Mehta, M.P., and Baross, J.A. (2006) Nitrogen fixation at 92 degrees C by a hydrothermal vent archaeon. *Science* **314**: 1783–1786.
- Mehta, M.P., Butterfield, D.A., and Baross, J.A. (2003) Phylogenetic diversity of nitrogenase (nifH) genes in deep-sea and hydrothermal vent environments of the Juan de Fuca Ridge. *Appl Environ Microbiol* **69**: 960–970.
- Meier, D.V., Pjevac, P., Bach, W., Hourdez, S., Girguis, P.R., Vidoudez, C., *et al.* (2017) Niche partitioning of diverse sulfur-oxidizing bacteria at hydrothermal vents. *ISME J* **11**: 1545–1558.
- Meier, D.V., Bach, W., Girguis, P.R., Gruber-Vodicka, H.R., Reeves, E.P., Richter, M., *et al.* (2016) Heterotrophic Proteobacteria in the vicinity of diffuse hydrothermal venting. *Environ Microbiol* **18**: 4348–4368.
- Nunoura, T., Miyazaki, M., Suzuki, Y., Takai, K., and Horikoshi, K. (2008) Hydrogenivirga okinawensis sp. nov., a thermophilic sulfur-oxidizing chemolithoautotroph isolated from a deep-sea hydrothermal field, Southern Okinawa Trough. *Int J Syst Evol Microbiol* **58**: 676–681.
- Olins, H.C., Rogers, D.R., Preston, C., Ussler, W., Pargett, D., Jensen, S., *et al.* (2017) Co-registered geochemistry and metatranscriptomics reveal unexpected distributions of microbial activity within a hydrothermal vent field. *Front Microbiol* **8**: 1042.
- Opatkiewicz, A.D., Butterfield, D.A., and Baross, J.A. (2009) Individual hydrothermal vents at Axial Seamount harbor distinct seafloor microbial communities. *FEMS Microbiol Ecol* **70**: 413–424.
- Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.
- Perner, M., Petersen, J.M., Zielinski, F., Gennerich, H.-H., and Seifert, R. (2010) Geochemical constraints on the diversity and activity of H₂-oxidizing microorganisms in diffuse hydrothermal fluids from a basalt- and an ultramafic-hosted vent. *FEMS Microbiol Ecol* **74**: 55–71.
- Perner, M., Gonnella, G., Hourdez, S., Bohnke, S., Kurtz, S., and Girguis, P. (2013) In situ chemistry and microbial community compositions in five deep-sea hydrothermal fluid samples from Irina II in the Logatchev field. *Environ Microbiol* **15**: 1551–1560.
- Perner, M., Bach, W., Hentscher, M., Koschinsky, A., Garbe-Schonberg, D., Streit, W.R., and Strauss, H. (2009) Short-term microbial and physico-chemical variability in low-temperature hydrothermal fluids near 5 degrees S on the Mid-Atlantic Ridge. *Environ Microbiol* **11**: 2526–2541.
- Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W.G., Peplies, J., and Glockner, F.O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188–7196.
- R-Development-Core-Team (2011) *R: A Language and Environment for Statistical Computing*. USA: R Foundation for Statistical Computing.
- Reveillaud, J., Reddington, E., McDermott, J., Algar, C., Meyer, J.L., Sylva, S., *et al.* (2016) Subseafloor microbial communities in hydrogen-rich vent fluids from hydrothermal systems along the Mid-Cayman Rise. *Environ Microbiol* **18**: 1970–1987.
- Reysenbach, A.-L., Huber, R., Stetter, K.O., Ishii, M., Kawasumi, T., Igarashi, Y., *et al.* (2001) Phylum BI. Aquificae phy. nov. In *Bergey's Manual® of Systematic Bacteriology: Volume One: The Archaea and the Deeply Branching and Phototrophic Bacteria*. Boone, D.R., Castenholz, R.W., and Garrity, G.M. (eds). New York, NY: Springer New York, pp. 359–367.
- Reysenbach, A.L., and Shock, E. (2002) Merging genomes with geochemistry in hydrothermal ecosystems. *Science* **296**: 1077–1082.
- Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al.* (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.
- Shah, V., Chang, B.X., and Morris, R.M. (2016) Cultivation of a chemoautotroph from the SUP05 clade of marine bacteria that produces nitrite and consumes ammonium. *ISME J* **11**: 263–271.
- Sheik, C.S., Jain, S., and Dick, G.J. (2014) Metabolic flexibility of enigmatic SAR324 revealed through metagenomics and metatranscriptomics. *Environ Microbiol* **16**: 304–317.
- Sievert, S.M., and Vetrani, C. (2012) Chemoautotrophy at deep-sea vents past, present, and future. *Oceanography* **25**: 218–233.
- Sievert, S.M., Brinkhoff, T., Muyzer, G., Ziebis, V., and Kuever, J. (1999) Spatial heterogeneity of bacterial populations along an environmental gradient at a shallow submarine hydrothermal vent near Milos Island (Greece). *Appl Environ Microbiol* **65**: 3834–3842.
- Stewart, L.C., Llewellyn, J.G., Butterfield, D.A., Lilley, M.D., and Holden, J.F. (2016) Hydrogen and thiosulfate limits for growth of a thermophilic, autotrophic *Desulfurobacterium* species from a deep-sea hydrothermal vent. *Environ Microbiol Rep* **8**: 196–200.
- Stewart, L.C., Jung, J.-H., Kim, Y.-T., Kwon, S.-W., Park, C.-S., and Holden, J.F. (2015) *Methanocaldococcus bathoarsdens* sp. nov., a hyperthermophilic methanogen isolated from a volcanically active deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* **65**: 1280–1283.
- Takai, K., Gamo, T., Tsunogai, U., Nakayama, N., Hirayama, H., Nealson, K.H., and Horikoshi, K. (2004) Geochemical and microbiological evidence for a hydrogen-based, hyperthermophilic subsurface lithoautotrophic microbial ecosystem (HyperSLiME) beneath an active deep-sea hydrothermal field. *Extremophiles* **8**: 269–282.
- Takai, K., Inagaki, F., Nakagawa, S., Hirayama, H., Nunoura, T., Sako, Y., *et al.* (2003) Isolation and phylogenetic diversity of members of previously uncultivated epsilon-proteobacteria in deep-sea hydrothermal fields. *FEMS Microbiol Lett* **218**: 167–174.

- ter Braak, C.J.F. (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* **67**: 1167–1179.
- Topçuoğlu, B.D., Stewart, L.C., Morrison, H.G., Butterfield, D.A., Huber, J.A., and Holden, J.F. (2016) Hydrogen limitation and syntrophic growth among natural assemblages of thermophilic methanogens at deep-sea hydrothermal vents. *Front Microbiol* **7**: 1240.
- Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M., *et al.* (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**: 37–43.
- Urich, T., Lanzen, A., Stokke, R., Pedersen, R.B., Bayer, C., Thorseth, I.H., *et al.* (2014) Microbial community structure and functioning in marine sediments associated with diffuse hydrothermal venting assessed by integrated meta-omics. *Environ Microbiol* **16**: 2699–2710.
- Ver Eecke, H.C., Akerman, N.H., Huber, J.A., Butterfield, D.A., and Holden, J.F. (2013) Growth kinetics and energetics of a deep-sea hyperthermophilic methanogen under varying environmental conditions. *Environ Microbiol Rep* **5**: 665–671.
- Ver Eecke, H.C., Butterfield, D.A., Huber, J.A., Lilley, M.D., Olson, E.J., Roe, K.K., *et al.* (2012) Hydrogen-limited growth of hyperthermophilic methanogens at deep-sea hydrothermal vents. *Proc Natl Acad Sci U S A* **109**: 13674–13679.
- Vetriani, C., Speck, M.D., Ellor, S.V., Lutz, R.A., and Starovoytov, V. (2004) *Thermovibrio ammonificans* sp. nov., a thermophilic, chemolithotrophic, nitrate-ammonifying bacterium from deep-sea hydrothermal vents. *Int J Syst Evol Microbiol* **54**: 175–181.
- Wilcock, W.S.D., Tolstoy, M., Waldhauser, F., Garcia, C., Tan, Y.J., Bohnenstiehl, D.R., *et al.* (2016) Seismic constraints on caldera dynamics from the 2015 axial seamount eruption. *Science* **354**: 1395–1399.
- Zhou, J.Z., Bruns, M.A., and Tiedje, J.M. (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62**: 316–322.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Statistics for metagenomic and metatranscriptomic libraries. Reads were mapped to metagenomic Open Reading Frames (ORFs) using Bowtie2. Metagenomic annotations to the KEGG Ontology (KO) database were determined via IMG/MER. Percentages of total reads mapped to metagenomic ORFs and annotated to KO database are in parentheses.

Table S2. Taxonomy of hits for each gene represented in Fig. 3 for the eleven metagenomes. The percent contribution of the top two hits is listed.

Table S3. Taxonomy of hits for each gene represented in Fig. 3 for the eleven metatranscriptomes. The percent contribution of the top two hits is listed.

Table S4. Statistics for metagenome assembled genomes (MAGs). MAGs were selected manually using tetranucleotide frequency and differential coverage. Percent completeness and contamination was calculated in Anvi'o. For comparison these statistics were also calculated using CheckM.

Fig. S1. Map of Axial Seamount showing locations of the three diffuse vents sampled from 2013 to 2015. In 2015, both a hydrothermal plume sample (N 45.9335667, W 130.013667) and a background seawater sample (1500 m depth, N 46.27389, W 129.79548) were also taken.

Fig. S2. Taxonomic classification of 16S rRNA reads identified in each of the metagenomes. Relative abundance of major taxonomic groups is shown.

Fig. S3. Normalized abundance and expression of genes for five main carbon-fixation pathways identified within seawater, plume and at each vent site over three sampling years.

Fig. S4. Heatmaps of mean coverage across metagenomes of metagenome assembled genomes (MAGs) identified as (A) SUP05, (B) *Sulfurovum*, (C) Aquificales and (D) methanogens. The dendrogram on each heatmap depicts the degree of similarity among MAGs based on mean coverage of metagenomic reads across samples. Scales depict range of mean coverage across MAGs.

Fig. S5. Concatenated marker gene tree of SUP05 MAGs and sequenced SUP05 genomes. Amino acid alignments were constructed using the program Phylosift using a reference set of 37 marker genes. MAG completeness is denoted at the end of each MAG name. Trees were constructed using RAxML.

Fig. S6. Concatenated marker gene tree of *Sulfurovum* MAGs and sequenced *Sulfurovum* genomes. Amino acid alignments were constructed using the program Phylosift using a reference set of 37 marker genes. MAG completeness is denoted at the end of each MAG name. Trees were constructed using RAxML.

Fig. S7. Concatenated marker gene tree of Aquificales MAGs and sequenced Aquificales genomes. Amino acid alignments were constructed using the program Phylosift using a reference set of 37 marker genes. MAG completeness is denoted at the end of each MAG name. Trees were constructed using RAxML.

Fig. S8. Concatenated marker gene tree of methanogenic archaea MAGs and sequenced methanogenic archaea genomes. Legend denotes specific genera of methanogen. Amino acid alignments were constructed using the program Phylosift using a reference set of 37 marker genes. MAG completeness is denoted at the end of each MAG name. Trees were constructed using RAxML.

Fig. S9. Canonical Correspondence Analysis diagrams for (A) metagenomes and (B) metatranscriptomes. Analysis included all seven environmental variables.