

Bringing microbial diversity into focus: high-resolution analysis of iron mats from the Lō‘ihi Seamount

Jarrod J. Scott,^{1*} Brian T. Glazer² and David Emerson¹

¹*Ocean Microbiome & Blue Biotechnology Center, Bigelow Laboratory for Ocean Sciences, PO Box 380, East Boothbay, ME 04544, USA.*

²*Department of Oceanography, University of Hawai‘i at Mānoa, 1000 Pope Rd. Honolulu, HI 96822, USA.*

Summary

Thirty kilometers south of the island of Hawai‘i lies the Lō‘ihi Seamount, an active submarine volcano that hosts a network of low-temperature hydrothermal vents enriched in ferrous iron that supports extensive microbial mats. These mats, which can be a half a meter deep, are composed of ferric iron bound to organic polymers – the metabolic byproduct of iron-oxidizing Zetaproteobacteria. Though the role of Zetaproteobacteria in mat formation is well established, we have a limited understanding of how differences in diversity are related to mat morphology. We used Minimum Entropy Decomposition and ZetaOtu classification to demonstrate cryptic diversity between closely related Zetaproteobacteria while showing habitat and geographic specificity. Veiled mats, common structures at Lō‘ihi, exhibit distinct community composition and contain diversity not detected in other mat types, including specific Zetaproteobacteria and an unclassified Gammaproteobacteria. Our analyses also indicate that diversity can change dramatically across small spatial transects from points of active venting, yet we found comparatively few differences between major sampling sites. This study provides a better picture of the microbiome responsible for iron mat production at Lō‘ihi and has broad implications for our understanding of these globally distributed communities.

Introduction

Zetaproteobacteria were first described at Lō‘ihi Seamount (Moyer *et al.*, 1995), an active undersea volcano located at

the eastern edge of the Hawaiian archipelago, 30 kilometers south of the island of Hawai‘i (Staudigel *et al.*, 2010). The summit area of Lō‘ihi (at ~1000 m water depth, Fig. 1A) hosts numerous low-temperature (upper range 45–55°C) hydrothermal vent systems with fluids containing high concentrations of ferrous iron (Garcia *et al.*, 2006; Staudigel *et al.*, 2010). Since the initial discovery of Zetaproteobacteria (Emerson *et al.*, 2007), numerous studies from around the world have demonstrated their preponderance in various marine environments wherever iron is a likely energy source, including mid-ocean ridges (Toner *et al.*, 2009; Scott *et al.*, 2015), back-arc hydrothermal fields (Kato *et al.*, 2009; Edwards *et al.*, 2011), seamounts (Moyer *et al.*, 1995; Emerson and Moyer, 2002) and coast-al/coastal margin habitats (Dang *et al.*, 2011; Rubin-Blum *et al.*, 2014). A recent meta-analysis of 16S rRNA datasets indicated little evidence of Zetaproteobacteria outside of iron-rich marine environments (Scott *et al.*, 2015), and thus far, all pure culture isolates from this class are lithotrophic iron-oxidizers.

Iron is an important energy source due to its global abundance and role as an electron donor for chemolithoautotrophic growth via the oxidation of Fe(II) to Fe(III). Iron-oxidizing bacteria (FeOB) and their associated communities are widespread and prevalent in both terrestrial and marine systems wherever anoxic ferrous-rich subsurface waters mingle with oxygenated surface waters (Emerson *et al.*, 2010). In marine systems, the most well-characterized communities of FeOB are associated with ferrous-rich hydrothermal vents. Because vents are a major source of iron to the ocean (Fitzsimmons *et al.*, 2014; Resing *et al.*, 2015), and primary productivity in much of the sunlit surface ocean is iron-limited, understanding the dynamics and fate of biogenically processed iron has implications for the global carbon cycle (Emerson, 2016).

The requirements for growth on ferrous iron are unique due to relatively poor thermodynamics (approximately $-90 \text{ kJ mol}^{-1} \text{ Fe(II)}$), rapid chemical oxidation of Fe(II) at circumneutral pH, and precipitation of large amounts of insoluble iron oxides that threaten to encrust the cells (reviewed in Melton *et al.*, 2014; Emerson *et al.*, 2010). To prevent encrustation, several FeOB have developed morphological adaptations: they produce extracellular

Received 6 June, 2016; revised 27 October, 2016; accepted 16 November, 2016. *For correspondence. E-mail jscott@bigelow.org; Tel. (+1) 207 315 2567; Fax (+1) 207 315 2329.

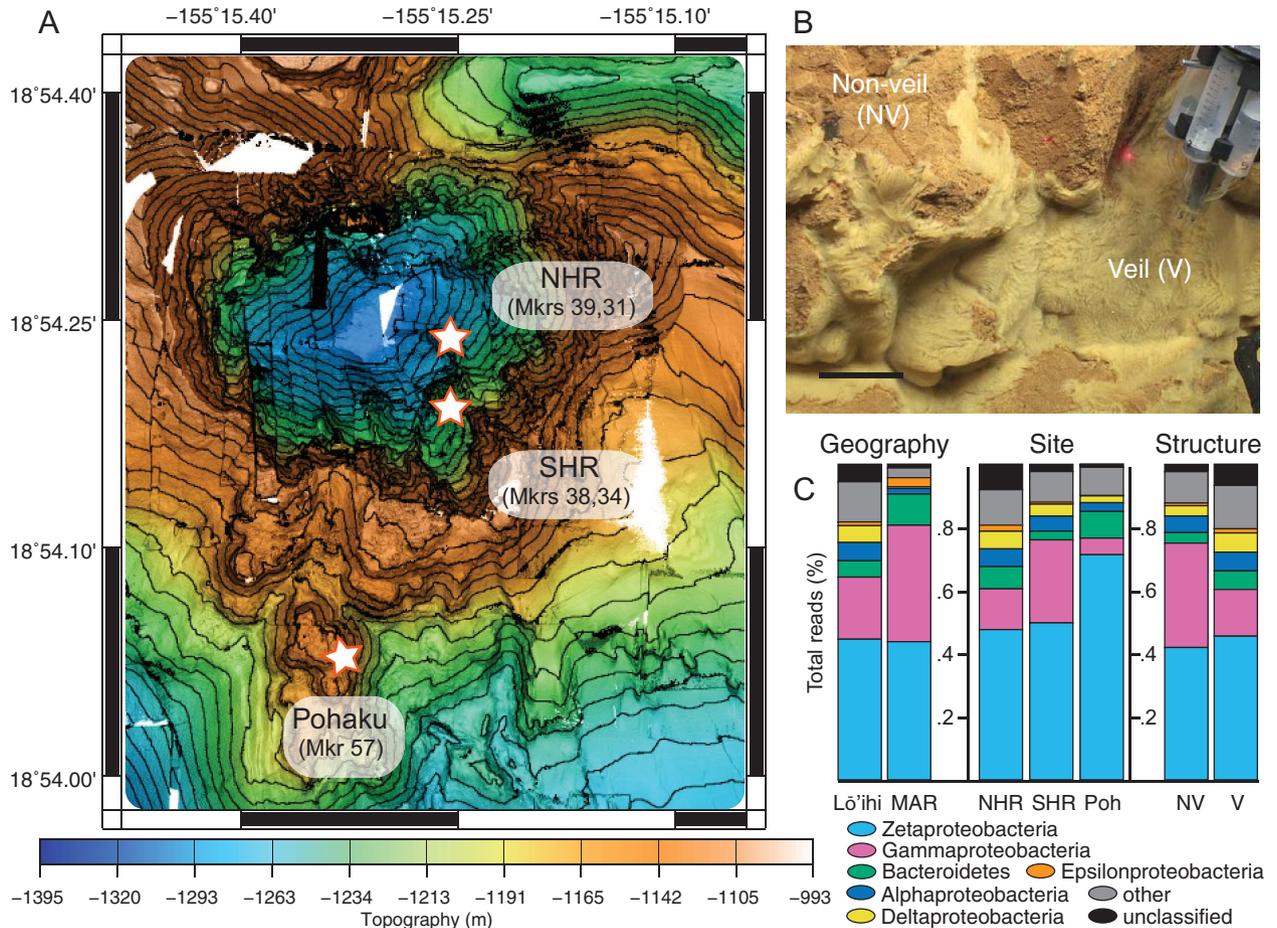


Fig. 1. Overview of study system.

A. Bathymetric map of Lō'ihi Seamount showing the relative location of major sampling sites [North Hiolo Ridge (NHR), South Hiolo Ridge (SHR) and Pohaku (Poh)] and associated markers.

B. Example of veil and non-veil mat structures found at Lō'ihi (scale bar, 10 cm).

C. Broad-scale taxonomic composition of iron mats encompassing geography (Lō'ihi and MAR), sites (NHR, SHR and Poh), and mat structure (non-veil and veil). Color scheme here and throughout adapted from (Wong, 2011).

For a summary of individual samples, including geographic location, sampling site, marker location and structure type, see Supporting Information Table S1.

mineral structures composed of poorly crystalline iron-oxyhydroxides and polysaccharides (Chan *et al.*, 2010).

At ferrous-rich hydrothermal vents, FeOB produce the most robust microbial mats of any chemolithoautotroph, in some cases reaching a meter or more in thickness (Edwards *et al.*, 2011). Though we continue to explore the intricacies of iron mat formation, it is clear that some filament-producing FeOB play a major role in mat architecture through the copious production of long (100's of μm) biogenic filaments of iron-oxyhydroxides, loosely arranged in a parallel orientation (Chan *et al.*, 2016). This results in a matrix with enough integrity to maintain itself while still being highly porous, thus influencing the flow and chemical composition of vent fluids as well as providing large surface areas that can be colonized by other microbes (Emerson and Weiss, 2004). These secondary colonizers,

including different FeOB and a variety of other taxa, may be more numerically abundant than the organisms responsible for production of the underlying mat matrix. The specificity of non-iron-oxidizing members of these iron mat communities is unknown. At present, we recognize two dominant mat-building morphotypes – sheath-formers and stalk-formers – yet it is unclear if either of these two mat morphotypes consists of multiple Zetaproteobacteria phylogenotypes. Sheath-formers produce thin and wispy veil-like mats while stalk-formers tend to make dense, more compact mats that may have a curd- or cauliflower-like appearance (Fig. 1B). What is apparent is that the most robust mat formations occur across chemical gradients largely driven by Fe(II) and O_2 , but physical gradients (e.g., temperature) must also play a role in structuring these microbiomes. In this sense, these systems are a useful

model for understanding when chemical and/or physical gradients form ecosystem boundaries that shape community structure, a fundamental question in microbial ecology (Denef *et al.*, 2010; Herlemann *et al.*, 2011; Meron, Buia, *et al.*, 2012; Meron, Rodolfo-Metalpa, *et al.*, 2012; Wilkins *et al.*, 2013).

Previous studies at Lō'ihi have shown the general predominance of Zetaproteobacteria in bulk mats (Rassa *et al.*, 2009; Jesser *et al.*, 2015) as well as other microbial taxa. These studies also demonstrated that the sheath-forming bacteria responsible for the production of veil-like mats are also Zetaproteobacteria and are not related to freshwater, sheath-forming iron-oxidizer, *Leptothrix ochracea*, as was originally thought (Fleming *et al.*, 2013). The work presented here expands on these studies by coupling amplicon sequencing and unsupervised oligotyping with the collection of high spatial resolution sampling to provide a more comprehensive picture of diversity in these structured microbial communities. Our working hypothesis was that community composition would be spatially stratified, both by distance from vent source and by depth in the mat, and there would be differences in communities related to major mat morphology (veil vs. non-veil mats, Fig. 1B).

Results

Sampling for this work focused on three main sites of diffuse venting at Lō'ihi (Fig. 1A). Two of these, North Hiolo Ridge and South Hiolo Ridge are part of the same complex located in the caldera of Pele's Pit that was formed by the 1996 eruption and the third site, Pohaku (Marker 57), is located outside the pit (Fig. 1A). For general site characteristics see Table 1. The hydrothermal fluid at Pohaku is notably cooler than Hiolo Ridge, but Fe(II) concentrations at all sites typically range between 400 and 600 μM . Once the vent fluid enters the overlying seawater there is a rapid decrease in Fe(II) concentration due to a combination of mixing and dilution with seawater, as well as oxidation of Fe(II) to iron-oxyhydroxides. This results in Fe(II) concentrations close to a vent orifice (1–5 cm) that ranges between 5.8 and 7 μM . A more detailed analysis of the vent fluid chemistry at these sites was published previously (Glazer and Rouxel, 2009).

Table 1. Summary of site characteristics.

Location	Site	Depth (m)	Temp ($^{\circ}\text{C}$)	Fe(II) (μM)	H ₂ S (μM) ^a	O ₂ vent (μM)	O ₂ ambient (μM)
N. Hiolo Ridge	Marker 39	1300.1	40.7	443	7.5	BD ^b	37
N. Hiolo Ridge	Texture Garden	1297.7	40.0	353	7.6	ND ^c	35
S. Hiolo Ridge	Marker 38	1271.4	46.3	566	BD	BD	38
S. Hiolo Ridge	Marker 34	1269.5	42.4	481	BD	ND	33
Pohaku	Marker 57	1178.7	25.9	400	3.0	ND	58

a. Value in vent orifice.

b. Below detection.

c. No data.

The summit of Lō'ihi is in an O₂ minima zone where ambient O₂ concentrations range from 30 to 60 μM .

Approximately 82 000 16S rRNA pyrotag reads were generated from 38 samples collected at sites within the Lō'ihi Seamount. After sequence processing and quality control, the dataset contained 65 484 high-quality reads (~270 bp average length) ranging from 423 to 6193 reads per sample (average, 1751). Taxonomic analysis revealed that roughly half of all reads classified as Zetaproteobacteria, 28% Gammaproteobacteria, 7% Bacteroidetes, 4% Alphaproteobacteria, 3% Deltaproteobacteria, 2% Unclassified and the remaining reads from numerous minor taxa. Combined samples from Lō'ihi contained comparable broad-level taxonomic profiles as MAR samples (Fig. 1C). Comparison between major Lō'ihi sampling sites (North Hiolo Ridge, South Hiolo Ridge and Pohaku) and structures (veil vs. non-veil) revealed similar patterns of broad-scale taxonomic composition (Fig. 1C). For a summary of individual samples, including total reads and diversity metrics, see Supporting Information Table S1.

Minimum Entropy Decomposition (MED)

MED analysis resulted in 305 final nodes after refinement (Supporting Information Fig. S1). Gammaproteobacteria and Zetaproteobacteria accounted for the majority of final nodes (29% and 24% respectively) followed by Deltaproteobacteria (7.2%), unclassified Bacteria (6.0%), Alphaproteobacteria (5.7%), Bacteroidetes (5.4%) and Epsilonproteobacteria (3.9%) (Supporting Information Fig. S1, Table S1). Non-metric multidimensional scaling (NMDS) analysis of Bray–Curtis dissimilarity matrix (from MED data) showed MAR samples clustering separately from Lō'ihi samples, however community similarity was not resolved between North and South Hiolo Ridge samples (Fig. 2). NMDS clustering showed that veil samples were distinct from non-veil samples (Fig. 2). Network analysis of MED nodes in Gephi (Bastian *et al.*, 2009) (v. 0.9.1, Force-Atlas2 layout using default settings) revealed similar clustering patterns of samples by mat structure and not site (Supporting Information Fig. S2).

Zetaproteobacteria diversity

Of the 80 MED nodes classified as Zetaproteobacteria (Supporting Information Figs S1 and S3A), 39 nodes were unique to Lō'ihi, 11 were unique to the MAR, and 30 were found in at least one sample from both geographic regions. We recovered 17 nodes that were each represented by 1000 or more reads accounting for 85% of all Zetaproteobacteria. ZetaHunter classification identified representatives from 17 different ZetaOtu. ZetaOtu1 and ZetaOtu2 had the highest number of nodes (11 and 15, respectively) accounting for 8% and 22% of total Zetaproteobacteria reads respectively. The other major groups were ZetaOtu4 (10 nodes, 9% reads), ZetaOtu9 (9 nodes, 24% reads), ZetaOtu14 (7 nodes, 10% reads), ZetaOtu6 (6 nodes, 7% reads) and ZetaOtu28 (4 nodes, 5% reads) (Fig. 3A). None of the 52 Zetaproteobacteria nodes were found in all samples across Lō'ihi and the MAR, however 4 were found in all Lō'ihi samples (veil and non-veil), specifically nodes from ZetaOtu1 (2110), ZetaOtu2 (2086, 2095) and Otu14 (2067) (Fig. 3A).

Comparative analyses

Lō'ihi Seamount & the Mid-Atlantic Ridge. LEfSe analysis detected 78 differentially abundant MED nodes across geographic regions (Lō'ihi vs. MAR) (Fig. 4). Roughly half (41) were enriched at Lō'ihi and the remainder at the MAR. MAR samples showed differential abundance in 20 Gammaproteobacteria MED nodes compared with 10 at Lō'ihi. Lō'ihi was also enriched in 7 Deltaproteobacteria nodes. Analysis of the 80 Zetaproteobacteria nodes showed differential abundance of 21 biomarkers between the two geographic regions (MAR, 10; Lō'ihi 11) – notably nodes from ZetaOtu9 (2155), ZetaOtu15 (1198), ZetaOtu17 (1833), ZetaOtu14 (1764) and ZetaOtu2 (1194) (enriched at the MAR) as well as ZetaOtu2 (2086, 2094), ZetaOtu1 (2110), ZetaOtu4 (2123, 2132), ZetaOtu6 (2076), ZetaOtu10 (1918) and ZetaOtu14 (2067) (enriched at Lō'ihi) (Figs 3 and 4). ZetaOtu17 was not detected in any Lō'ihi samples and only a small subset of samples from Lō'ihi had reads from either ZetaOtu9 or ZetaOtu15. Of the 8 nodes classified as ZetaOtu14, 2 were enriched at each geographic location. All samples contained nodes from ZetaOtu2 yet subtle variation was detected across geography. Nodes from ZetaOtu1, ZetaOtu4 and ZetaOtu6 were almost exclusively found at Lō'ihi.

Hiolo Ridge & Pohaku. Few differences were observed between North and South Hiolo Ridge. Only 5 of the 98 MED nodes tested were significantly different between the two sites (Fig. 4), none of which were Zetaproteobacteria. Comparison between Hiolo Ridge and Pohaku revealed 13 differences including 7 Zetaproteobacteria nodes. Pohaku was enriched for 4 nodes classified as ZetaOtu2 and 1

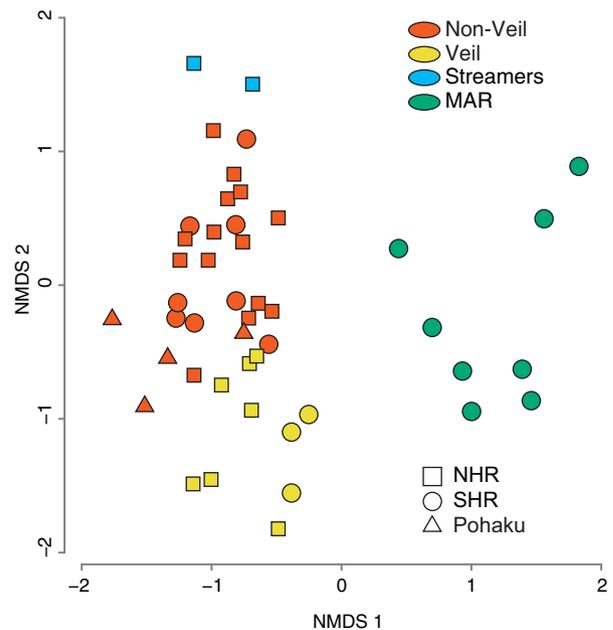
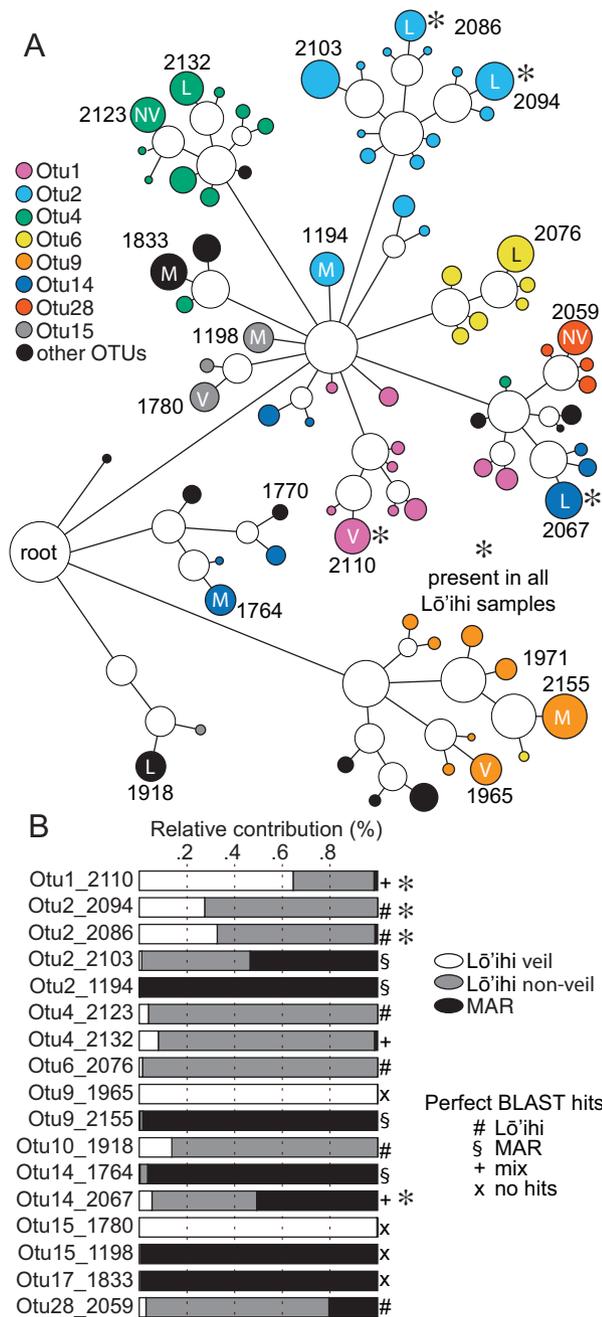


Fig. 2. Comparison of samples based on location and mat structure using NMDS analysis of Bray–Curtis dissimilarity matrix generated from MED analysis. Shape indicates samples by site and color indicates samples by structure. The green circles represent MAR samples.

ZetaOtu14 node (Fig. 4). Nodes 2110 (ZetaOtu1) and 2059 (ZetaOtu28) (enriched at Hiolo Ridge) were rare or undetected at Pohaku.

Mat structures. LEfSe analysis of veils and non-veils revealed 37 differentially abundant MED nodes (veil, 18; non-veil, 19) (Fig. 4), 12 of which were Zetaproteobacteria including an overrepresentation of nodes from ZetaOtu9 and ZetaOtu15 (both undetected in non-veils). Neither ZetaOtu9 nor ZetaOtu15 though were present in all veil samples. Veils were enriched for 7 different Gammaproteobacteria nodes including one (1668) that was undetected in non-veil samples. LEfSe analysis further showed that veils and non-veils were both enriched for representatives of ZetaOtu1 and ZetaOtu4. Non-veils were enriched for ZetaOtu28, a biomarker almost completely absent from veils. MED analysis did not return any nodes that were detected in all veil samples yet absent from non-veil samples, except for Gammaproteobacteria node 1668.

Within site assessments. We conducted several within-site assessments to further disentangle patterns of bacterial diversity from iron mats at the Lō'ihi Seamount. First, along North Hiolo Ridge (Fig. 5 and Fig. Supporting Information S3) we looked at samples from three sites containing veil and non-veil structures (Fig. 5A–C and Fig. Supporting Information S3) – the Texture Garden, Marker 39 B, and Marker 39 A. At the Texture Garden three veil samples were compared with six non-veil



samples all collected within 10 cm of one another (Fig. 5A). While most samples showed an abundance of Zetaproteobacteria reads, especially from ZetaOtu1 and ZetaOtu2 (Fig. 5A), only veil samples contained reads from the unclassified Gammaproteobacteria MED node 1668 while non-veils were abundant in MED node 1537 (Flavobacteriales) (Supporting Information Fig. S3). These patterns were similar for samples collected at Marker 39_B (Fig. 5B and Supporting Information Fig. S3) and Marker 39_A (Fig. 5C and Supporting Information Fig. S3).

Fig. 3. Network map of Zetaproteobacteria MED nodes highlighting sites of enrichment and ZetaHunter OTU classification.

A. Subset from the full MED network in Supporting Information Fig. S2 corresponding to all Zetaproteobacteria nodes. Final nodes are color-coded according to ZetaHunter OTU classification. MED IDs of all nodes represented by more than 1000 reads are included. Letters within nodes refer to differentially abundant features from the LEfSe analyses of MAR (M) vs. Lō'ihi (L) and veil (V) vs. non-veil (NV). The four nodes marked by asterisks indicate nodes found in all Lō'ihi samples. No nodes were found in all samples when the MAR was included. Nodes 1770 represents prototypical pure culture isolates from Lō'ihi (ZetaOtu11).
B. Chart showing the relative contribution (% total reads) of major nodes in (A) from Lō'ihi (veil and non-veil) and the MAR labeled according to ZetaOtu classification and MED node ID. Symbols indicate the origin of perfect BLAST hits specifically Lō'ihi (16S SAGs or environmental clones) or the MAR (16S SAGs). Mixed indicates that BLAST hits were from a variety of sites. Asterisks refer to the nodes found in all Lō'ihi samples.

Non-veil samples from the Texture Garden and Marker 39_B also showed an abundance of the Alphaproteobacteria order Kiloniellales (node 0953), while 20% of the reads from the non-veil sample at M39_A were associated with node 2001, in the Gammaproteobacteria order Marinicellales (Supporting Information Fig. 3).

We also looked at changes in community composition along small transects moving away from points of active venting. At North Hiolo Ridge (Markers 39/31 area) we sampled a vertical transect beginning at a vent (#5 in Supporting Information Fig. S4) moving in the direction of flow up the face of the formation. Within the vent orifice (39.6°C) we observed white, streamer-like structures similar to those produced by sulfur-oxidizing bacteria (Moyer *et al.*, 1995). This streamer sample had markedly different community composition compared with the other transect samples, specifically a low abundance of Zetaproteobacteria (<5%) coupled with higher abundance of Gammaproteobacteria and Epsilonproteobacteria. We also detected 2 unclassified taxa (0283 and 1422) that were rare in all other samples (Supporting Information Fig. S4, Table S2). Assessment of major Zetaproteobacteria nodes showed that the relative abundance of ZetaOtu6 decreased with distance from the vent, accounting for 15% of total community reads in sample #4 to less than 0.1% in sample #1 at the top of the face (Supporting Information Fig. S4). Similarly, ZetaOtu28 also decreased with distance from the vent opening. Both of ZetaOtu6 and ZetaOtu28 were exclusive to non-veil samples at Lō'ihi and only matched single amplified genome (SAG) 16S rRNA gene sequences retrieved previously from Lō'ihi (Field *et al.*, 2014).

At Marker 38 along the South Hiolo Ridge white streamers were not visible in the vent orifice. We found that the sample closest to the vent (#3 in Supporting Information Fig. S5) was predominantly Zetaproteobacteria, specifically ZetaOtu6

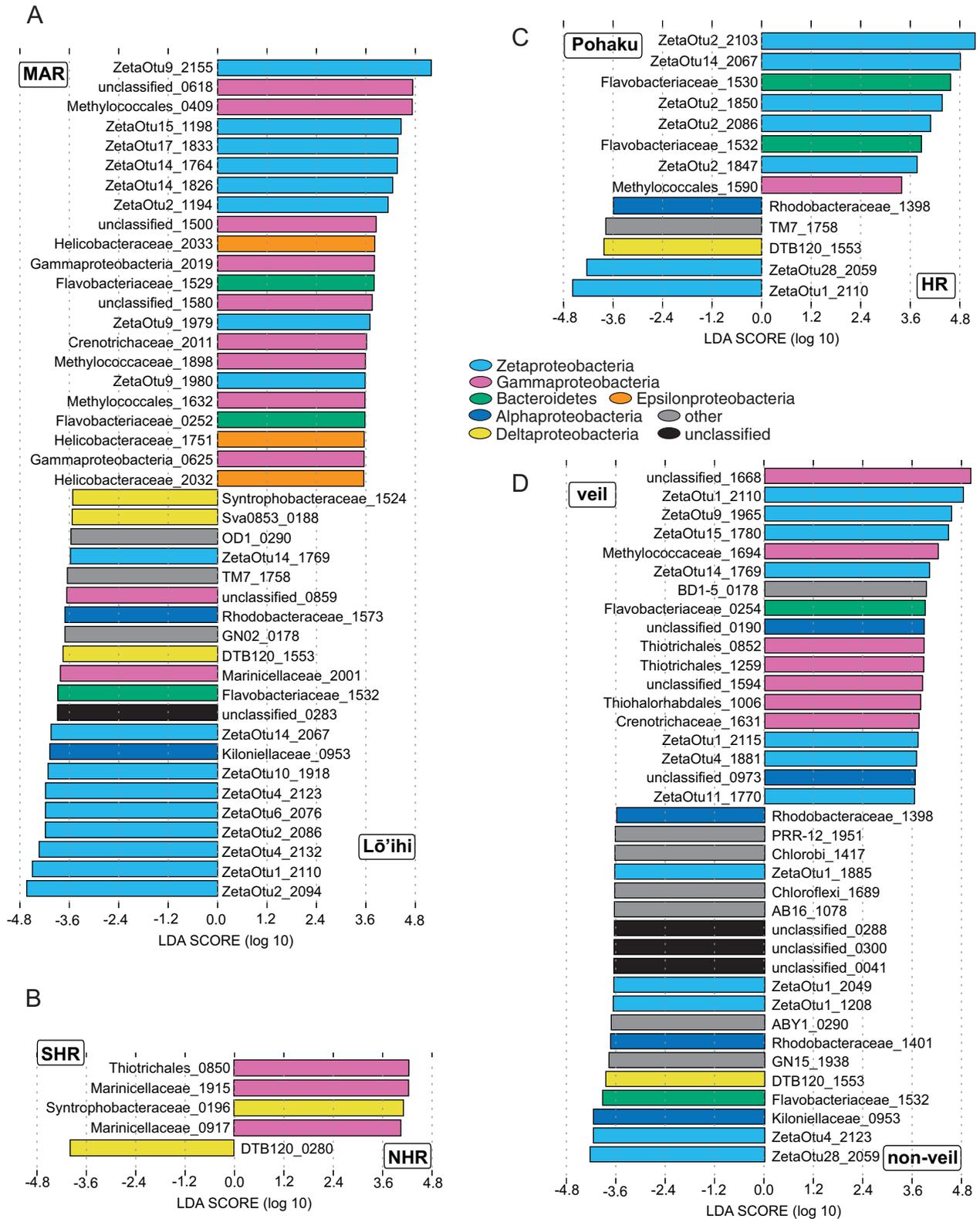


Fig. 4. Differentially abundant MED nodes detected by LEfSe for (A) geographic region, (B) and (C) Lō'ihī sites and (D) mat structure. Colors indicate broad-level taxonomic classification. For the comparison between Lō'ihī and the MAR (A) only 43 of 78 differentially abundant nodes are shown (LDA score ≥ 3.6). In the other cases, all differentially abundant nodes are presented. Biomarker names contain deeper taxonomic classification followed by MED node ID.

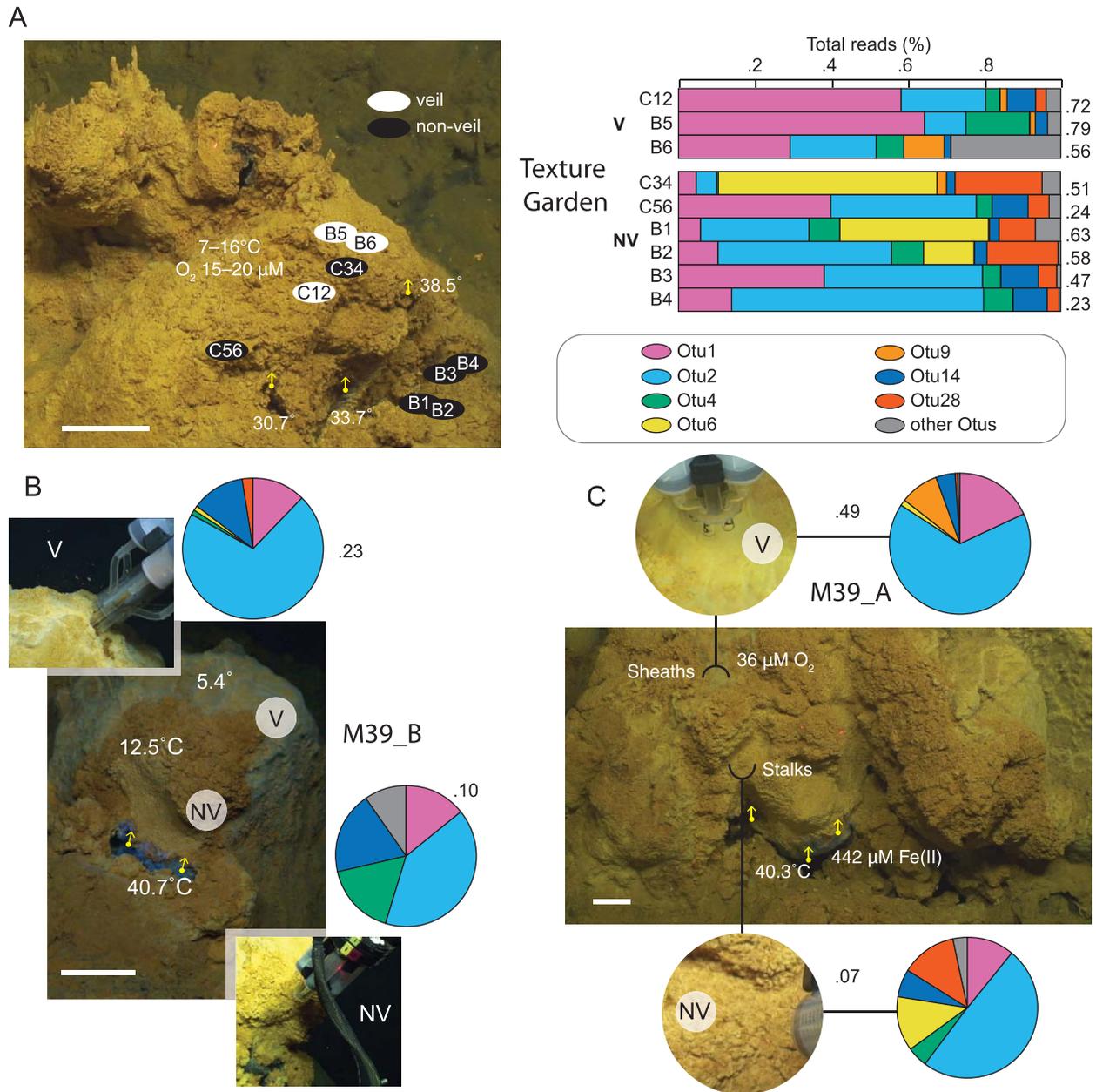


Fig. 5. Analysis of major ZetaOtu nodes at Marker 39. For an overview of the site see Supporting Information Fig. S3.

A. Comparison of ZetaOtu profiles between three veil and six non-veil samples collected around a series of diffuse flow vents at the Texture Garden site. Fe(II) concentrations were 18.3 μM near sample C56, 46.2 μM around the veil samples and 70.1 μM where the non-veil samples B1–B4 were collected.

B. Comparison between veil and non-veil samples from M39_B. Fe(II) concentrations were 36.1 μM in the veil and 465.5 μM in the non-veil sample.

C. Comparison between veil and non-veil samples from M39_A. Electrochemical measurements were not recorded from M39_A. Arrows indicate the direction of fluid flow with associated fluid temperature. Values indicate proportion of all Zetaproteobacteria reads to total community composition. Scale bars, 10 cm.

and ZetaOtu26. 33% and 98% of all reads from these ZetaOtu were found in this sample respectively. We observed that nodes from Kiloniellaceae, Flavobacteriaceae and Rhodobacteraceae increased with distance from the vent opening (Supporting Information Fig. S5). In addition, ZetaOtu2, comprising less than 0.1% of the

Zetaproteobacteria reads from the vent sample (#3), comprised over 50% of reads in the two samples further up the face (Supporting Information Fig. S5, Table S2). Fe²⁺ concentrations near the vent (#3 in Supporting Information Fig. S5) were 131.4 μM to and decreased to 30.4 μM further up the face while temperatures decreased from 21.3°C to 7.2°C.

Discussion

In this study, we used deep submergence assets and a precision sampling device to collect discrete, sub-centimeter scale, iron mat samples from several sites within the Lō'ihi Seamount. Our results confirm the overall abundance of Zetaproteobacteria at Lō'ihi iron mats and document the consistent presence of other bacterial lineages associated with these ferrous-rich ecosystems. Precision sampling, coupled with high-resolution Minimum Entropy Decomposition (MED) and the ZetaHunter OTU framework, allowed us to demonstrate cryptic diversity between closely related Zetaproteobacteria phylogenetic groups. We show that sheath-rich veiled mats, a common mat morphotype at Lō'ihi, contain diversity not evident in other mat types. We also present evidence that diversity can change significantly over small spatial scales along transects moving away from points of diffuse flow venting. Together these analyses provide a more in-depth picture of the microbiome responsible for the production of microbial iron mats fueled by Fe(II)-driven lithotrophy that are globally distributed at hydrothermal vents.

Zetaproteobacteria biogeography

Previous work has demonstrated biogeographic patterning for Zetaproteobacteria, including groups that exhibit cosmopolitan distribution and others that appear endemic to specific areas (McAllister *et al.*, 2011). Here we expand on these findings by coupling MED analysis with ZetaOtu classification to expose deeper biogeographic patterns within particular groups. McAllister *et al.* identified two ZetaOtus that exhibited cosmopolitan distribution throughout the Pacific Basin – ZetaOtu1 and ZetaOtu2 (McAllister *et al.*, 2011). Indeed, these ZetaOtus accounted for a large portion of total Zetaproteobacteria diversity in our dataset (26 nodes, 30% reads). ZetaOtu1 was abundant at Lō'ihi, especially node 2110, which was enriched in veil samples and widely distributed in non-veil samples (Figs 3 and 4). In contrast we found little evidence of ZetaOtu1 at the MAR suggesting that ZetaOtu1 may be restricted to the Pacific and not globally cosmopolitan. ZetaOtu2, the other putatively cosmopolitan group, contained 15 MED nodes and 23% of Zetaproteobacteria reads. Four of these nodes were present in samples from both Lō'ihi and the MAR. Node 2103 in particular was abundant in samples from both sites and phylogenetic inference showed that node 2103 was most similar to SAGs from both the MAR and Lō'ihi (Supporting Information Fig. S6). ZetaOtu2 nodes also showed evidence of geographic specificity including nodes 2086 and 2094, which were only found at Lō'ihi (Fig. 3). In contrast, node 1194 was almost completely restricted to the MAR (99.6% of reads) and most closely related to MAR SAGs (Supporting Information Fig. S6).

Phylogenetic analysis demonstrated that this clade was distinct from the remainder of ZetaOtu2 diversity. Field *et al.* also identified ZetaOtu2 as a dominant member of 16S rRNA gene sequences retrieved from SAG libraries (Field *et al.*, 2014). Together these results indicate that the ZetaOtu2 group has both cosmopolitan and endemic diversity. Given its abundance at both geographic sites, our findings also suggest that ZetaOtu2 may play an important role in general mat ecology.

We found additional ZetaOtus that were also abundant in samples from both sites including ZetaOtu9, ZetaOtu14, ZetaOtu15 and ZetaOtu28. Neither McAllister *et al.* nor Field *et al.* found representatives from ZetaOtu9 and ZetaOtu15 at Lō'ihi (McAllister *et al.*, 2011; Field *et al.*, 2014). This is perhaps not surprising as our data suggests that these ZetaOtus are restricted to Lō'ihi veils and select MAR samples (see discussion below) (Fig. 3). ZetaOtu28 – an abundant group found across numerous non-veil samples at (Fig. 3) – was also undetected at Lō'ihi in both studies. On the other hand we were unable to detect ZetaOtus described previously from Lō'ihi, specifically ZetaOtus 3, 12, 19 and 22 (McAllister *et al.*, 2011). We found several additional ZetaOtus that were scarce or absent at the MAR but abundant at Lō'ihi, notably ZetaOtu4, ZetaOtu6 and ZetaOtu10 (Fig. 3), all of which were previously described from Lō'ihi (McAllister *et al.*, 2011; Field *et al.*, 2014). ZetaOtu17 (2 nodes, 7% reads) – previously identified from the southern Pacific Ocean (Vailulu'u Seamount, Tonga Arc, East Lau Spreading Center, Kermadec Arc) (McAllister *et al.*, 2011) – was found only at the MAR and completely absent from our analysis of Lō'ihi samples. BLAST analysis of reads from ZetaOtu17 against the *nr* database returned no perfect hits (Fig. 3, Table 2). The closest hit to ZetaOtu17 (98% identity) was to a clone retrieved from the Southern Mariana Trough (Kato *et al.*, 2009) while the nearest 16S rRNA SAG hit was only 96%.

Communities & mat structures

At Lō'ihi, the veil structures were seen on the outer edges of iron mats, furthest away from areas of active venting. Geochemically, these areas have higher oxygen concentrations (typically from 20 to 30 μM), lower Fe(II) concentrations (from below detection to 40 μM), and cooler temperatures ($<10^\circ\text{C}$). They are also physically less exposed to the hydrodynamic forces of hydrothermal flow from the vent orifice. The most recognizable differences in community composition at Lō'ihi were between the veil and non-veil mats – reflecting the structural uniqueness of veiled mats (Fig. 1B). The underlying fabric of veiled mats is sheath-forming FeOB, while non-veil mats have various filamentous oxide morphotypes including helical stalks similar to those produced by *Mariprofundus ferrooxydans* (Chan *et al.*, 2010).

Table 2. BLAST survey of top Zetaproteobacteria MED nodes.

ZetaOtu	MED node	Enriched ^a	Perfect BLAST hits ^b	Lō'ihi 16S SAGs ^c	MAR 16S SAGs ^d	Env. clones ^e	Lō'ihi SAGs ^f	MAR SAGs ^g	Pure cultures ^h	AAI ⁱ
Otu1	2110	V	12	5	–	7	2	–	–	69.2
Otu2	2103	NE	20	1	18	1	–	3	–	99.4–99.6
Otu2	2094	L	20	20	–	–	3	–	–	91.1–96.9
Otu2	1194	M	8	–	7	1	–	1	–	–
Otu2	2086	L	20	20	–	–	–	–	–	–
Otu4	2132	L	14	5	–	9	–	–	–	–
Otu4	2123	NV	20	20	–	–	1	–	–	–
Otu6	2076	L	7	3	–	4	3	–	–	96.5–99.3
Otu9	2155	M	19	–	15	4	–	3	2	88.1–95.6
Otu9	1965	V	–	–	–	–	–	–	–	–
Otu10	1918	L	9	6	–	3	2	–	–	88.3
Otu14	2067	L	19	4	2	13	–	2	–	63.1
Otu14	1764	M	3	–	3	–	–	1	–	–
Otu15	1198	M	–	–	–	–	–	–	–	–
Otu15	1780	V	–	–	–	–	–	–	–	–
Otu17	1833	M	–	–	–	–	–	–	–	–
Otu28	2059	NV	3	2	–	1	–	–	–	–

The 17 Zetaproteobacteria MED nodes that were each represented by more than 1000 reads. Representative sequences from each node was BLAST-searched against NCBI's *nr* database.

a. Sites of node enrichment; Lō'ihi (L), MAR (M), veils (V), or non-veils (NE, not enriched). Values indicate total number of perfect BLAST matches (100% identity and 100% coverage) for a given node (out of 20).

b. Total number of BLAST hits.

c. Hits to 16S rRNA SAGs (single amplified genomes) from Lō'ihi (Field *et al.*, 2014).

d. Hits to 16S rRNA SAGs from the MAR (Scott *et al.*, 2015).

e. Hits to environmental clones from other studies.

f. Hits to Lō'ihi SAGs (Field *et al.*, 2014).

g. Hits to MAR SAGs (Scott unpublished data).

h. Hits to pure culture genomes.

i. AAI, Average Amino Acid Identity for SAGs within a node. No value (–) indicates that nodes returned no perfect BLAST match..

We did not detect any Zetaproteobacteria nodes that were exclusively found in all veil communities though 93% of ZetaOtu9 reads found at Lō'ihi were from veiled mats, especially node 1965 (Fig. 3). We also found that MAR samples were enriched in ZetaOtu9 (Fig. 4) but instead node 1965 the MAR contained node 2155 (Figs. 3 and 4). Node 2155 was most similar to MAR SAGs and a MAR pure culture isolate. However more data is needed to better understand the phylogenetic relationship of node 1965 to other ZetaOtu9 diversity (Supporting Information Fig. S6). It seems that ZetaOtu9 not only exhibits specific habitat preferences, but also cosmopolitan distribution. MAR communities and Lō'ihi veils were also enriched in ZetaOtu15 (Fig. 4), but again MED analysis resolved this diversity – the MAR contained all reads of node 1198 and veil samples contained all reads of node 1780 (Fig. 3), yet no reads from ZetaOtu15 were detected in non-veil samples from Lō'ihi. Phylogenetic inference confirmed the differences between these two nodes (Supporting Information Fig. S6). Analysis of major ZetaOtu9s indicated that veiled mats not only contained these unique Zetaproteobacteria groups but also common ZetaOtu9s (ZetaOtu1 and 2). It is interesting to note that McAllister *et al.* identified numerous clones of both ZetaOtu9 and ZetaOtu15

(McAllister *et al.*, 2011) from borehole fluid collected at the Southern Mariana Trough (Kato *et al.*, 2009) but none from Lō'ihi. In fact, these two ZetaOtu9s accounted for the majority of diversity borehole fluids and ZetaOtu15 was only found at this site. Is it possible that ZetaOtu9 and/or ZetaOtu15 are representative of the predominant sheath-formers in veiled mats? Sheaths reminiscent of those observed in Lō'ihi veils were abundant in several MAR samples dominated by ZetaOtu9 (Otu3 in Scott *et al.*, 2015), although the larger, veil-like mat structures found at Lō'ihi (Fig. 5 and Supporting Information Fig. S3) were not observed at the MAR. Instead, MAR mats abundant in sheaths were more compact and localized around vent orifices (Scott *et al.*, 2015). It remains unclear whether sheath-dominated mats at the MAR are functionally similar to veiled mats at Lō'ihi or rather that ZetaOtu9 proliferates under similar environmental conditions and that different MED nodes result from genetic drift.

Curiously, veil samples were also enriched in node 1668 – an unclassified Gammaproteobacteria – and contained 99.7% of all reads from this node (Fig. 4 and Supporting Information Fig. S3, Table S2). BLAST analysis returned perfect hits to four environmental clones, all found previously by Fleming *et al.* from veiled mats at Lō'ihi (Fleming

et al., 2013), suggesting that node 1668 is an important component of veiled iron mats. The role of this organism in veils and why it appears completely restricted to these habitats at Lō'ihi remains a mystery.

Community changes along physicochemical gradients

Transect sampling provided intriguing clues about how communities respond to shifting temperature and/or geochemical gradients over small spatial scales. Marked changes in composition were observed in close proximity to venting compared with communities further away. In particular, we found that ZetaOtu6 was abundant near points of active venting and decreased precipitously with distance (Fig. 5 and Supporting Information Figs S4 and S5). Based on our analyses ZetaOtu6 appeared restricted to mats in close proximity of active venting. Over 93% of all reads classified as ZetaOtu6 in our dataset were found in these samples. Similar patterns were observed for ZetaOtu26 at Marker 38 on the South Hiolo Ridge – abundant near the vent and then decreasing to near extinction in the other two samples (Supporting Information Fig. S5). In fact, ZetaOtu26 was only identified from a vent-associated sample (Supporting Information Fig. S5). These findings suggest a potential connection between higher temperatures and both ZetaOtu6 and ZetaOtu26 distribution. Conversely, the abundance of ZetaOtu14 was greatest in samples furthest from active venting (Supporting Information Figs S4 and S5) and we observed striking transitions in ZetaOtu2 abundance across a small spatial scale (<20 cm) (Supporting Information Fig. S5). Completely absent in the sample closest to the vent (Supporting Information Fig. S5), ZetaOtu2 increased to over 50% of Zetaproteobacteria reads in the samples furthest away from the vent.

The range of values for temperature (21.3–7.2°C) and Fe(II) (131–30 µM) do not seem extreme for this site but do represent a 3- to 4-fold difference respectively. This means that ZetaOtu diversity may be correlated to shifts in physicochemical properties such as Fe(II) concentrations and temperature. Factors such as flow velocity over the face of the mat, or dynamic changes in overall vent chemistry may also play a role. Recent work by Jesser *et al.* showed marked differences in arsenic resistance genes between samples collected only a few centimeters apart in Lō'ihi mats (Jesser *et al.*, 2015), suggesting that toxic metals/metalloids resistance may also be important drivers shaping local community diversity. Based on these findings we also expected to see differences in community composition with mat depth, however no substantial changes in community composition from depth profiles were observed from three different sites at the North Hiolo Ridge and Pohaku (data not shown).

Mats forming closer to vent orifices exhibited a diverse set of macroscopic mat morphotypes and contained a variety of filamentous biogenic iron oxides – the structural components of mats. Some Lō'ihi samples contained abundant helical stalks, while others had a variety of thin, filamentous forms that were more challenging to associate with a unified structure. As filaments become more mineralized, they lose fine details that normally allow for easy identification. An exception to this may be the two samples identified as a white streamer morphotype that proliferated at vent orifices along the North Hiolo Ridge (Supporting Information Figs S3 and S4). These two samples contained the majority of reads from node 1039, an unclassified Epsilonproteobacteria that has only been identified from Lō'ihi (based on BLAST analysis) (Rassa *et al.*, 2009; McAllister *et al.*, 2011) and was rare in the rest of our dataset. These samples also contained a number of nodes not common in other samples, including two unclassified bacteria (nodes 1422 and 0283), SAR406 (0178) and five unclassified Gammaproteobacteria (66–95%). Interestingly two of these Gammaproteobacteria nodes (0621 and 0616) appear to be the closest relatives of the veil associate described above (node 1668). Our analysis supports earlier findings showing that streamers are likely the product of sulfur-oxidizing bacteria (Moyer *et al.*, 1995; Emerson and Moyer, 2010; Jesser *et al.*, 2015). Consistent with this explanation, vent fluids at North Hiolo Ridge tend to carry higher concentrations of hydrogen sulfide (5–45 µM) compared with other Lō'ihi sites where sulfide concentrations are <10 µM (Glazer and Rouxel, 2009).

Implications for an iron mat microbiome

The work presented here confirms that microbial iron mats contain a unique microbiome and expands our knowledge about important phylogenetic and putative functional groups in these systems. Clearly, the key metabolism of this microbiome is lithotrophic iron-oxidation, and Zetaproteobacteria are likely the key players in this process. Whether other iron-oxidizers from different phylogenetic groups flourish in these systems is unknown at this time. Members of the Gammaproteobacteria are possible candidates, especially the veil associate, node 1668. Several nodes belonging to the Chromatiales were found in iron mats, but had a patchy distribution and low relative abundances. A study of the microbial community enriched on a biocathode found an abundance of Chromatiales (Wang *et al.*, 2015). Metagenomic and metaproteomic analyses suggested that this organism was capable of growth via extracellular electron transfer, and the proposed pathway proposed bears a striking similarity to an iron-oxidation pathway proposed in *M. ferrooxydans* (Barco *et al.*, 2015).

We found 35 non-Zetaproteobacteria MED nodes that exhibited signs of broad geographic distribution though many were in low abundance at one or both geographic

locations. Interestingly, nodes from Methylococcales were well represented in several samples from the MAR and Lō'ihi, including veil and non-veil mats. The vast majority of these bacteria use methane as carbon and energy source (Bowman, 2014). Unfortunately, we did not measure methane in vent fluids in areas of community profiling. Historically, methane is known to be present in vent fluids at Lō'ihi, and studies of shown activity for methanotrophy based on radiotracer analysis although mats themselves were not tested (Karl *et al.*, 1989). There are several phylogenetic groups of methanotrophs; therefore, it is interesting that all the putative methanotrophs identified in iron mats belong only to the Methylococcales, including Methylococcaceae and Crenotrichaceae. As far as we know, there is no geochemical interaction between iron metabolism (oxidation or reduction) and methane; however, it is possible that these Methylococcales are somehow better adapted to high-iron environments.

A well-represented clade in non-veil mats was Kiloniellaceae (Supporting Information Figs S3–S5). Kiloniellaceae (Alphaproteobacteria) have been found associated with macroalgae (Wiese *et al.*, 2009) and sponges (Cleary *et al.*, 2013), though iron mats are notable for being depauperate in macrofauna. Scale worms however, and smaller nematode-like worms, have been observed in iron mats (D. Emerson, unpublished observations). Either members of the Kiloniellaceae are free-living, or could be indicators of an unknown host. Members of the Flavobacteriales were enriched in non-veil communities (Fig. 4), accounting for approximately 6% of total reads and 21 MED nodes. A recent study showed evidence for an association between Bacteroidetes and filamentous sulfur-oxidizing bacteria at a hydrothermal vent where Bacteroidetes were growing organotrophically on exudates produced by the lithoautotrophic sulfur-oxidizer (Stokke *et al.*, 2015). It is possible similar associations occur between iron-oxidizers and Bacteroidetes in iron mats.

Exploring Zetaproteobacteria diversity

These findings raise interesting questions about factors determining the distribution of different Zetaproteobacteria groups. Since ferrous iron is assumed to be the primary energy source for both cosmopolitan and endemic strains, other traits may play a role in determining the successful ecological strategies of cosmopolitan Zetaproteobacteria. For example, are cosmopolitan strains better adapted for transport to other areas, perhaps by sustaining a dormant state? Based on our molecular analysis alone, it remains unclear whether cosmopolitan Zetaproteobacteria actually produce extracellular structures. The fact they are found in mats dominated by sheaths or stalks suggests they do not produce either of these structures, otherwise we would expect these ZetaOtu to be associated with only one

morphotype, but not both. Perhaps instead they are secondary colonizers of mats. One common morphotype in mats that attaches to either sheaths or stalks is an organism that produces unique Y-shaped branched filaments (Chan *et al.*, 2016). It is possible that at least one of the cosmopolitan ZetaOtu is the Y-shaped FeOB. Furthermore, what do ZetaOtu and oligotypes (MED nodes) reveal in terms of population biology and ecological function for Zetaproteobacteria? Do oligotypes correspond to functionally distinct groups adapted to different niches? The fact that particular oligotypes tend to dominate specific mat types – despite their close proximity to other mats – suggests that oligotypes represent important functional differences within/between populations. Given that oligotypes from the same ZetaOtu differentiate across large geographic space may suggest that genetic drift also plays an important role in structuring Zetaproteobacteria diversity.

In addition to ZetaHunter classification, we also assessed all major Zetaproteobacteria nodes using BLAST against the *nr* database, which included 16S rRNA gene sequences from single amplified genomes (SAGs) recovered from mats at Lō'ihi (Field *et al.*, 2014) and the MAR (Scott *et al.*, 2015) (Scott, unpublished data). Looking only at perfect matches (100% identity, 100% coverage), we assessed the number of hits to each major ZetaNode, including the number of hits from sequenced SAGs and pure culture genomes (Table 2). Four nodes returned no perfect BLAST matches – including the two predominant veil ZetaOtu (9 and 15) – however, all remaining 13 nodes hit multiple 16S SAGs, 10 of which were represented by one or more single amplified genome. Only one dominant node (2155, ZetaOtu9) had top BLAST hits to pure culture isolates, and these were two organisms are capable of growth on either Fe(II) or H₂ (D. Emerson, unpublished data). Remarkably, the prototypical pure culture isolates of marine FeOB (M34, PV1, JV1) were the top hit for one node represented by just 155 reads (node 1770, Fig. 3 and Supporting Information Fig. S6).

Next we looked at average amino acid identities (AAI) between SAGs within a node to determine whether node assignment was a good proxy for genome similarity and hence potential ecological function. Though these results are preliminary and based on limited data, we observed some interesting trends, perhaps the most telling coming from AAI comparisons among ZetaOtu2 nodes. AAI values for values within a node were high for the two cases where we had multiple SAGs (2103 & 2094) (Table 2). When we compared AAI values across nodes within ZetaOtu2 – first between 2103 and 2094 – we saw values decrease slightly to 89%. When we compared SAGs from these nodes to the SAG in node 1194, the AAI values fell below 75%. Phylogenetic analysis showed that nodes 2103 and 2094 were in distinct, yet closely related clades, while 1194 was distantly related to both (Supporting Information Fig. S6).

Similarly, when we compared the five genomes from node 2155 (ZetaOtu9) to two SAGs classified in 1971 (ZetaOtu9), the AAI values dropped below 80%. This is finding is also supported by the distinct phylogenetic placement of node 2155 versus 1971.

There were two cases where AAI values within nodes were quite low – nodes 2110 (ZetaOtu1) and 2067 (ZetaOtu14) (Table 2). Curiously, in both cases the top SAG BLAST hits were classified as different ZetaOtus than the host nodes. The two hits from 2110 (ZetaOtu1) were classified as ZetaOtu14 and the two hits from 2067 (ZetaOtu14) were classified as ZetaOtu1. This was also the case for several 16S SAGs where we did not have genomes. However, when 16S rRNA sequences from these SAGs were truncated to encompass only the V4–V5 region we then observed agreement with MED node classification. Given that we did not observe this pattern for any other ZetaOtu, these findings suggest that the V4–V5 region does not have the variability to resolve ZetaOtu1 and ZetaOtu14. Phylogenetic analysis shows that ZetaOtu1 and ZetaOtu14 are not monophyletic groups and often placed together in single clades (Supporting Information Fig. S6). The overall results of this analysis however appear promising and may yield more robust results by coupling longer reads with supervised oligotyping.

MED analysis enhanced our ability to detect diversity patterns that were unresolved at a 97% identity OTU-level alone. ZetaHunter OTU classification was able to detect diversity unique to Lō'ihi (ZetaOtus 1, 4, 6, & 10) and the MAR (ZetaOtu17) however it could not resolve biogeographic differences within more widespread ZetaOtus (ZetaOtus 2, 9, & 15). ZetaHunter classification, however, provides a broad framework to assess the organization of Zetaproteobacteria diversity while MED provides the precision to assess important population-level difference. By coupling these two approaches, we were able to build a more complete picture of the distribution of different Zetaproteobacteria groups.

Concluding remarks

Iron-oxidizing bacteria (FeOB) are widespread in nature and generate thick microbial mats composed of iron-oxyhydroxides. These structural edifices increase spatial heterogeneity, alter biogeochemistry, and modulate nutrient availability of the surrounding environment. This work, combined with previous studies at Lō'ihi and other hydrothermal iron mats, demonstrates that there is a cohesive microbiome associated with iron mats and Zetaproteobacteria dominate these ferrous-fueled ecosystems. Whether other persistently associated taxa simply respond to changing resources, or actually play a role in mat biogenesis, remains an open question. We showed that specific lineages are consistently present in the microbiome of

iron mats and exhibit distribution patterns that appear attributable to niche specialization – this included a Gammaproteobacteria that was only found in veiled mats. Among the Zetaproteobacteria, there were gross differences in community composition between veil mats (dominated by sheath-forming morphotypes) and non-veil mats (dominated by stalk- and other filament-forming morphotypes). At the same time, cosmopolitan ZetaOtus were found at all sites, suggesting they possess traits that allow them to adapt to differences at both local and broad geographic settings. MED allowed for better biogeographic discrimination of cosmopolitan ZetaOtus and also revealed fine-scale changes within contiguous mats – diversity that may be driven by relatively subtle shifts in environmental factors. Understanding the environmental and ecological causation of fine-scale community diversity remains an outstanding challenge for microbial ecology.

Experimental procedures

Study overview, site description & sampling

Sampling of iron mats was conducted from March 17 to April 1 2013, aboard the R/V *Thompson* (cruise No. TN293). The DSV *Jason II*, equipped with a precision microbial mat sampler (Breier *et al.*, 2012), was used for the collection of 231 individual syringes of mat samples across six dives (J2671–J2676). The spatial resolution for collecting samples with the mat sampler was approximately 1 cm, either vertically or horizontally. Primary sampling locations within the Lō'ihi Seamount system were North Hiolo Ridge (18° 54' 23.3" N 155° 15' 25.0" W), South Hiolo Ridge (18° 54' 20.6" N 155° 15' 25.5" W) and Pohaku (18° 54' 4.8" N 155° 15' 29.505" W) (Fig. 1A). Roughly 20 ml of veiled or non-veiled iron mats were collected per sample (Fig. 1B), and replicate samples from the same physical location were combined into a single sample. Material reserved for community analysis was stored shipboard at –80°C and transported to the lab on dry ice. For a summary of individual samples, including sampling site and marker location, see Supporting Information Table S1. Chemistry measurements were conducted using an in-situ electrochemical analyzer, ISEA, (ISEA-III, Analytical Instrument Systems, Flemington, NJ, USA) as described elsewhere (Chan *et al.*, 2016). Temperature measurements in vent chimneys were made by pushing the *Jason II* temperature probe directly into the vent orifice; other temperature measurements in microbial mats were made using a high sensitivity thermocouple incorporated into the ISEA analyzer wand. High resolution mapping of the Lō'ihi summit was achieved with the AUV *Sentry* (Fig. 1A). The National Deep Submergence Facility at Woods Hole Oceanographic Institution operates the DSV *Jason II* and AUV *Sentry*.

16S rRNA gene pyrosequencing processing & analysis

In this study we selected 41 samples for 454-pyrosequencing analyses. Samples were processed following previously published protocols (Scott *et al.*, 2015). From each mat sample, approximately 250 mg (wet weight) of mat material was used

for DNA extraction from each sample using a Mo Bio Power-Soil® DNA Extraction Kit (Mo Bio Laboratories, Carlsbad, CA, USA), modified to include an initial phenol:chloroform:isoamyl alcohol (PCI) step. We found that adding an initial PCI treatment increased final DNA yield from an average <5ng/μl to concentrations typically >15 ng/μl (data not shown). Briefly, 200 μl of bead solution was removed from each bead tube and replaced with 200 μl of 25:24:1 PCI (Sigma-Aldrich, St. Louis MO, USA). Samples were then extracted using the manufacture recommended protocol and sent to Research and Testing Laboratory (Lubbock TX, USA) for pyrosequencing. We targeted the V4–V5 hypervariable region (*E. coli* positions 531–997) using 530F (5'-GTG CCA GCM GCN GCG G-3') and 1100R (5'-GGG TTN CGN TCG TTG-3') following established protocols (Dowd *et al.*, 2008). See Supporting Information Table S1 for barcode sequences of individual samples.

Sequence processing was performed using mothur v.1.35.0 (Schloss *et al.*, 2009) following previously published methodology (Schloss *et al.*, 2011). Specifically, we used mothur to remove primer and barcode sequences from all reads, as well as any short reads (<250 bp), reads containing more than six homopolymers, reads with any ambiguities and/or any chimeric reads detected by UCHIME (Edgar *et al.*, 2011). All quality filtered reads were aligned against a mothur-compatible recreation (mothur.org/wiki/Silva_reference_files; last accessed 09.11.2016) of the SILVA-SEED (SILVA v123) reference alignment (Quast *et al.*, 2012). Pyrotag reads were classified against the Greengenes reference taxonomy (McDonald *et al.*, 2012). For comparative purposes we also include data from iron mat samples previously collected at the Mid-Atlantic Ridge (MAR) (Scott *et al.*, 2015). All pyrosequencing libraries are deposited at the European Nucleotide Archive under the sample accessions numbers ERS1089449–ERS1089534, study accession number PRJEB10821.

Community analysis

Minimum Entropy Decomposition (MED) (Eren *et al.*, 2014) analysis was performed on the complete dataset of 126 051 reads from 46 samples (38 Lō'ihi and 8 MAR) using the following screening parameters: minimum substantive abundance, 26.0; minimum entropy value for decomposition, 0.0965; maximum number of discriminants to use for decomposition, 4; maximum variation allowed in each node, 3 nucleotides. Per node normalization was conducted before decomposition based on the node size of the most abundant sequence in the dataset (Eren *et al.*, 2014). Screening removed 16 801 outliers resulting in 109 250 total reads in the final analysis. Alignments contained 570 characters with an average read length of 268 base pairs without gaps. Bray–Curtis dissimilarity coefficient (Bray and Curtis, 1957) was used to calculate distance metrics for non-metric multidimensional scaling (NMDS) analyses. We then used linear discriminant analysis (LDA) effect size (LEfSe, default parameters) (Segata *et al.*, 2011) to identify differentially abundant biomarkers (MED nodes) across geographic regions (Lō'ihi vs. MAR), sites (North Hiolo Ridge, South Hiolo Ridge and Pohaku) (Fig. 1A) and mat structure (veils vs. non-veils) (Fig. 1B). LEfSe uses LDA to estimate the effect size of differentially abundant biomarkers to explain differences between sample types (Segata *et al.*, 2011). Per-sample

normalization of the sum of the values was used to account for different sample coverage.

In addition, all reads from Zetaproteobacteria nodes were classified further using ZetaHunter (<https://github.com/mooreryan/ZetaHunter>, last accessed 09.03.2016) – a command line script designed to assign SSU rRNA gene sequences to Zetaproteobacteria OTUs (97% identity) (Schloss *et al.*, 2009; Edgar *et al.*, 2011; Kopylova *et al.*, 2012) defined by a reference database with an underlying phylogenetic structure (SILVA v123) (Pruesse *et al.*, 2012; Quast *et al.*, 2012). In the absence of a well-defined taxonomy, OTU binning is currently the accepted method for classifying diversity of Zetaproteobacteria (McAllister *et al.*, 2011), henceforth referred to as ZetaOtus.

Acknowledgements

We thank the captain and crew of the R/V *Thompson*, the operations team of the ROV *Jason II*, and the science party for assistance with sample collection. We also thank J. Delton Hanson and Jeremy Wilkinson from Research and Testing Laboratory (Lubbock, TX) for sequencing, John A. “Chip” Breier for assistance and advice with the precision mat sampler, and the operations team of the AUV *Sentry*, especially Carl Kaiser, for help with bathymetric maps. We are also grateful to Sean McAllister and Ryan Moore for assistance with ZetaHunter – including implementation and data interpretation. Finally, we thank two anonymous reviewers for their careful reading of this manuscript as well as their insightful comments and suggestions. This work was funded in part by NSF grant OCE-1155754. The authors have no conflict of interest to declare.

Data accessibility

All pyrosequencing libraries were deposited at the European Nucleotide Archive under the sample accessions numbers ERS1089449–ERS1089534, study accession number PRJEB10821.

Author contributions

JJS, BTG and DE designed research, performed research, contributed reagents and analytical tools, and analyzed data. JJS and DE wrote the paper.

References

- Barco, R.A., Emerson, D., Sylvan, J.B., Orcutt, B.N., Jacobson Meyers, M.E., Ramírez, G.A., *et al.* (2015) New insight into microbial iron oxidation as revealed by the proteomic profile of an obligate iron-oxidizing chemolithoautotroph. *Appl Environ Microb* **81**: 5927–5937.
- Bastian, M., Heymann, S., and Jacomy, M. (2009) Gephi: an open source software for exploring and manipulating networks. *ICWSM* **8**: 361–362.
- Bowman, J.P. (2014) The family *Methylococcaceae*. In *The Prokaryotes*. Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., and Thompson, F. (eds). New York, NY: Springer, pp. 411–440.

- Bray, R.J., and Curtis, J.T. (1957) An ordination of upland forest communities of southern Wisconsin. *Ecol Monogr* **27**: 325–349.
- Breier, J.A., Gomez-Ibanez, D., Reddington, E., Huber, J.A., and Emerson, D. (2012) A precision multi-sampler for deep-sea hydrothermal microbial mat studies. *Deep-Sea Res Pt I* **70**: 83–90.
- Chan, C.S., Fakra, S.C., Emerson, D., Fleming, E.J., and Edwards, K.J. (2010) Lithotrophic iron-oxidizing bacteria produce organic stalks to control mineral growth: implications for biosignature formation. *ISME J* **5**: 717–727.
- Chan, C.S., McAllister, S.M., Leavitt, A.H., Glazer, B.T., Krepski, S.T., and Emerson, D. (2016) The architecture of iron microbial mats reflects the adaptation of chemolithotrophic iron oxidation in freshwater and marine environments. *Front Microbiol* **7**: 796.
- Cleary, D.F.R., Becking, L.E., de Voogd, N.J., Pires, A.C.C., Polónia, A.R.M., Egas, C., and Gomes, N.C.M. (2013) Habitat- and host-related variation in sponge bacterial symbiont communities in Indonesian waters. *FEMS Microbiol Ecol* **85**: 465–482.
- Dang, H., Chen, R., Wang, L., Shao, S., Dai, L., Ye, Y., *et al.* (2011) Molecular characterization of putative biocorroding microbiota with a novel niche detection of Epsilon- and Zetaproteobacteria in Pacific Ocean coastal seawaters. *Environ Microbiol* **13**: 3059–3074.
- Denef, V.J., Mueller, R.S., and Banfield, J.F. (2010) AMD biofilms: using model communities to study microbial evolution and ecological complexity in nature. *ISME J* **4**: 599–610.
- Dowd, S.E., Callaway, T.R., Wolcott, R.D., Sun, Y., McKeenan, T., Hagevoort, R.G., and Edrington, T.S. (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiol* **8**: 125.
- Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.
- Edwards, K.J., Glazer, B.T., Rouxel, O.J., Bach, W., Emerson, D., Davis, R.E., *et al.* (2011) Ultra-diffuse hydrothermal venting supports Fe-oxidizing bacteria and massive uranium deposition at 5000 m off Hawaii. *ISME J* **5**: 1748–1758.
- Emerson, D. (2016) The irony of iron-biogenic iron oxides as an iron source to the ocean. *Front Microbiol* **6**: 1502.
- Emerson, D., Fleming, E.J., and McBeth, J.M. (2010) Iron-oxidizing bacteria: an environmental and genomic perspective. *Annu Rev Microbiol* **64**: 561–583.
- Emerson, D., and Moyer, C.L. (2002) Neutrophilic Fe-oxidizing bacteria are abundant at the Loihi Seamount hydrothermal vents and play a major role in Fe oxide deposition. *Appl Environ Microb* **68**: 3085–3093.
- Emerson, D., and Moyer, C.L. (2010) Microbiology of seamounts: common patterns observed in community structure. *Oceanography* **23**: 148–163.
- Emerson, D., Rentz, J.A., Lilburn, T.G., Davis, R.E., Aldrich, H., Chan, C., and Moyer, C.L. (2007) A novel lineage of Proteobacteria involved in formation of marine Fe-oxidizing microbial mat communities. *PLoS One* **2**: e667.
- Emerson, D., and Weiss, J.V. (2004) Bacterial iron oxidation in circumneutral freshwater habitats: findings from the field and the laboratory. *Geomicrobiol J* **21**: 405–414.
- Eren, A.M., Morrison, H.G., Lescault, P.J., Reveillaud, J., Vineis, J.H., and Sogin, M.L. (2014) Minimum entropy decomposition: unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME J* **9**: 968–979.
- Field, E.K., Sczyrba, A., Lyman, A.E., Harris, C.C., Woyke, T., Stepanauskas, R., and Emerson, D. (2014) Genomic insights into the uncultivated marine Zetaproteobacteria at Loihi Seamount. *ISME J* **9**: 857–870.
- Fitzsimmons, J.N., Boyle, E.A., and Jenkins, W.J. (2014) Distal transport of dissolved hydrothermal iron in the deep South Pacific Ocean. *Proc Natl Acad Sci USA* **111**: 16654–16661.
- Fleming, E.J., Davis, R.E., McAllister, S.M., Chan, C.S., Moyer, C.L., Tebo, B.M., and Emerson, D. (2013) Hidden in plain sight: discovery of sheath-forming, iron-oxidizing Zetaproteobacteria at Loihi Seamount, Hawaii, USA. *FEMS Microbiol Ecol* **85**: 116–127.
- Garcia, M.O., Caplan-Auerbach, J., De Carlo, E.H., Kurz, M.D., and Becker, N. (2006) Geology, geochemistry and earthquake history of Lōihi Seamount, Hawaii's youngest volcano. *Chem Erde-Geochem* **66**: 81–108.
- Glazer, B.T., and Rouxel, O.J. (2009) Redox speciation and distribution within diverse iron-dominated microbial habitats at Loihi Seamount. *Geomicrobiol J* **26**: 606–622.
- Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., and Andersson, A.F. (2011) Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* **5**: 1571–1579.
- Jesser, K.J., Fullerton, H., Hager, K.W., and Moyer, C.L. (2015) Quantitative PCR analysis of functional genes in iron-rich microbial mats at an active hydrothermal vent system (Lōihi Seamount, Hawaii). *Appl Environ Microb* **81**: 2976–2984.
- Karl, D.M., Brittain, A.M., and Tilbrook, B.D. (1989) Hydrothermal and microbial processes at Loihi Seamount, a mid-plate hot-spot volcano. *Deep-Sea Res* **36**: 1655–1673.
- Kato, S., Yanagawa, K., Sunamura, M., Takano, Y., Ishibashi, J.I., Kakegawa, T., *et al.* (2009) Abundance of Zetaproteobacteria within crustal fluids in back-arc hydrothermal fields of the Southern Mariana Trough. *Environ Microbiol* **11**: 3210–3222.
- Kopylova, E., Noe, L., and Touzet, H. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217.
- McAllister, S.M., Davis, R.E., McBeth, J.M., Tebo, B.M., Emerson, D., and Moyer, C.L. (2011) Biodiversity and emerging biogeography of the neutrophilic iron-oxidizing Zetaproteobacteria. *Appl Environ Microb* **77**: 5445–5457.
- 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.
- Melton, E.D., Swanner, E.D., Behrens, S., Schmidt, C., and Kappler, A. (2014) The interplay of microbially mediated and abiotic reactions in the biogeochemical Fe cycle. *Nat Rev Microbiol* **12**: 797–809.
- Meron, D., Buia, M.C., Fine, M., and Banin, E. (2012) Changes in microbial communities associated with the sea anemone *Anemonia viridis* in a natural pH gradient. *Microb Ecol* **65**: 269–276.

- Meron, D., Rodolfo-Metalpa, R., Cuning, R., Baker, A.C., Fine, M., and Banin, E. (2012) Changes in coral microbial communities in response to a natural pH gradient. *ISME J* **6**: 1775–1785.
- Moyer, C.L., Dobbs, F.C., and Karl, D.M. (1995) Phylogenetic diversity of the bacterial community from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. *Appl Environ Microb* **61**: 1555–1562.
- Pruesse, E., Peplies, J., and Glockner, F.O. (2012) SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al.* (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.
- Rassa, A.C., McAllister, S.M., Safran, S.A., and Moyer, C.L. (2009) Zeta-Proteobacteria dominate the colonization and formation of microbial mats in low-temperature hydrothermal vents at Loihi Seamount, Hawaii. *Geomicrobiol J* **26**: 623–638.
- Resing, J.A., Sedwick, P.N., German, C.R., Jenkins, W.J., Moffett, J.W., Sohst, B.M., and Tagliabue, A. (2015) Basin-scale transport of hydrothermal dissolved metals across the South Pacific Ocean. *Nature* **523**: 200–203.
- Rubin-Blum, M., Antler, G., Tsadok, R., Shemesh, E., Austin, J.A., Coleman, D.F., *et al.* (2014) First evidence for the presence of iron oxidizing Zetaproteobacteria at the Levantine continental margins. *PLoS One* **9**: e91456.
- Schloss, P., Gevers, D., and Westcott, S.L. (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One* **6**: e27310.
- 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 Microb* **75**: 7537–7541.
- Scott, J.J., Breier, J.A., Luther, G.W., III, and Emerson, D. (2015) Microbial iron mats at the Mid-Atlantic Ridge and evidence that Zetaproteobacteria may be restricted to iron-oxidizing marine systems. *PLoS One* **10**: e0119284.
- Segata, N., Izard, J., Waldron, L., and Gevers, D. (2011) Metagenomic biomarker discovery and explanation. *Genome Biol* **12**: R60.
- Staudigel, H., Moyer, C.L., Garcia, M.O., Malahoff, A., Clague, D.A., and Koppers, A.A.P. (2010) Lō'ihi Seamount. *Oceanography* **23**: 72–73.
- Stokke, R., Dahle, H., Roalkvam, I., Wissuwa, J., Daae, F.L., Tooming-Klunderud, A., *et al.* (2015) Functional interactions among filamentous Epsilonproteobacteria and Bacteroidetes in a deep-sea hydrothermal vent biofilm. *Environ Microbiol* **17**: 4063–4077.
- Toner, B.M., Santelli, C.M., Marcus, M.A., Wirth, R., Chan, C.S., McCollom, T., *et al.* (2009) Biogenic iron oxyhydroxide formation at mid-ocean ridge hydrothermal vents: Juan de Fuca Ridge. *Geochim Cosmochim Acta* **73**: 388–403.
- Wang, Z., Leary, D.H., Malanoski, A.P., Li, R.W., Hervey, W.J., IV, Eddie, B.J., *et al.* (2015) A previously uncharacterized, nonphotosynthetic member of the Chromatiaceae is the primary CO₂-fixing constituent in a self-regenerating biocathode. *Appl Environ Microb* **81**: 699–712.
- Wiese, J., Thiel, V., Gartner, A., Schmaljohann, R., and Imhoff, J.F. (2009) *Kiloniella laminariae* gen. nov., sp. nov., an alphaproteobacterium from the marine macroalga *Laminaria saccharina*. *Int J Syst Evol Microbiol* **59**: 350–356.
- Wilkins, D., Yau, S., Williams, T.J., Allen, M.A., Brown, M.V., DeMaere, M.Z., *et al.* (2013) Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol Rev* **37**: 303–335.
- Wong, B. (2011) Color blindness. *Nat Methods* **8**: 441–441.

Supporting information

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

Fig. S1. Topological network map showing the decomposition of MED nodes and taxonomic distribution of all final nodes. Node size is proportional to the total number of reads contained within a node. Color scheme here and throughout adapted from (Wong, 2011).

Fig. S2. Network analysis of MED nodes. Network clustering of samples and highly connected MED nodes (degree >20). Networks are the same except colored by (A) site and (B) mat structure. Size is proportional to the number of reads in either MED nodes or samples. Clustering is more apparent by structure (B) rather than site (A). The most connected MED nodes in the network (shared by the most samples) were all Zetaproteobacteria, specifically nodes 2086 (ZetaOtu2), 2059 (ZetaOtu28), 2110 (ZetaOtu1), 2067 (ZetaOtu14), 2094 (ZetaOtu2) and 2103 (ZetaOtu2).

Fig. S3. Marker 39 whole community analysis. Same site as Fig. 5 but highlighting shifts in non-Zetaproteobacteria nodes. (A) Comparison of MED node profiles between three veil and six non-veil samples collected around a series of diffuse flow vents at the Texture Garden site. Though all samples had an abundance of Zetaproteobacteria, veil samples had contained MED node 1668 (unclassified Gammaproteobacteria) while the non-veil samples were enriched in node 1537 (Flavobacteria). Similar patterns were found in the community profiles of veil (V, top) and non-veil (NV, bottom) samples from (C) M39_B and (D) M39_A. Arrows indicate the direction of fluid flow with associated fluid temperature. Scale bar, 10cm. (D) Overview of site. **Insert** Counterclockwise rotated and close-up view of M39_B where the veil morphotype covers the upper portion of the formation.

Fig. S4. North Hiolo Ridge (Marker 31 area) iron mat communities assessed along a vertical transect beginning at the vent (#5, sample 676-BM1-D6) and moving up the face of the formation in the direction of flow. White, streamer-like structures similar to those produced by sulfur-oxidizing bacteria were observed within the vent orifice (39.6°C). **Top bar graph:** For simplicity, total community diversity is presented as broad taxonomic groups, however there were important changes in individual MED nodes that occurred along the transect. Node 0283 (unclassified Bacteria) decreased precipitously from 26% of community reads in #5 to undetected in #1. Two unclassified Gammaproteobacteria nodes (0621 and 0859) each decreased from 15% of total community in #5 to ~0.5% further up the face. Conversely, we observed a corresponding increase in Flavobacteria (node 1532) moving up the face. **Bottom bar graph:** Breakdown of major ZetaOtu distribution. Sample #5 contained relatively few Zetaproteobacteria reads however

marked changes in diversity were observed moving up the face. Specifically, ZetaOtu6 was relatively abundant in #4 and #3 but absent in samples associated with cooler temperatures. Conversely, ZetaOtu14 increased with distance from active venting. Scale bar, 10cm.

Fig. S5. South Hiolo Ridge (Marker 38) iron mat communities assessed along a vertical transect beginning at the vent (sample #3). Both temperature and Fe(II) concentrations decreased with distance from the vent. **Top bar graph:** Total community diversity presented as broad taxonomic groups showed an increase in Bacteroidetes and Alphaproteobacteria with distance from vent. This pattern is similar to changes seen along other transects. **Bottom bar graph:** Substantial changes were observed in Zetaproteobacteria diversity, specifically the abundance of ZetaOtu6 and ZetaOtu26 were greatest near the vent while ZetaOtu2 and ZetaOtu14 increased with distance. Arrows indicate the direction of fluid flow with associated fluid temperature. Scale bar, 10cm.

Fig. S6. Maximum likelihood phylogenetic tree (see separate 91 document) of 16S rRNA gene sequences showing the placement of major Zetaproteobacteria MED nodes. Phylogenetic inference was generated using IQ-TREE (Nguyen *et al.*, 2014). Full-length sequences were truncated to the size of the pyrotag reads and the final alignment contained 295 columns and 155 total sequences, including outgroups. The best-fit model, TIM3+G4, was chosen according to Bayesian Information Criterion (Schwarz, 1978). Circles indicate MED nodes and colors correspond to the ZetaHunter OTU classification, as in Fig. 4. The colors of vertical bars also correspond to ZetaHunter classifi-

cation. Stars indicate single amplified genomes (SAGs) and arrows indicate the placement of pure culture isolates. Leaf labels are as follows: sequence ID, NCBI accession number (for MSAG_AD_336_N22, MSAG_AD_336_O11, DIS1, TAG1, SV108, and M39 the IMG genome ID is provided instead), and ZetaHunter OTU designation. ZetaHunter classification was based on V4–V5 for pyrotag reads and near full length sequences for SAGs, pure cultures and environmental clones. ZetaHunter classification for two MED nodes did not agree with the classification of their nearest neighbors. The majority of strains closely related to node 2067 (ZetaOtu14) were classified as ZetaOtu1. The opposite was observed for node 2110 (ZetaOtu1) where many of its neighbors were classified as ZetaOtu14. When these sequences were truncated to encompass only the V4–V5 region we then observed agreement with MED node classification. These switches are color highlighted in the tree according to the ZetaOtu classification of their V4–V5 region. We did not observe this pattern for any other ZetaOtu or MED node.

Table S1. Separate file. Summary of sample details including sample name, barcode(s), geographic location, collection site, structure type, number of reads, number of MED nodes and 16S rRNA diversity indices. Also include is the Figure number for samples that were analyzed in more detail.

Table S2. Separate file. Summary table of all 334 MED nodes identified in this study as well as their Greengenes lineage and abundance across all samples. For Zetaproteobacteria nodes the ZetaHunter classification is also given. Totals for geographic and mat structure are also provided.