Metagenomic insights into the dominant Fe(II) oxidizing Zetaproteobacteria from an iron mat at Lō‘ihi, Hawai‘i

Esther Singer1, John F. Heidelberg2, Ashita Dhillon3 and Katrina J. Edwards1,2,*

1 Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA
2 Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA
3 Genzyme Corporation, Allston, MA, USA

INTRODUCTION
Microbial Fe(II) oxidation is widespread in the deep-sea and is thought to play a significant role in rock and mineral weathering (Edwards et al., 2004). The microbial influence on the global iron cycle, and consequently on other linked biogeochemical cycles, such as carbon, has remained poorly understood, although several studies have recently started to investigate key habitats with large potential to host Fe(II)-oxidizing bacteria (Fe(II)-OB; Edwards et al., 2003; McAllister et al., 2004; Rassa et al., 2009; Emerson and Moyer, 2010; Emerson, 2012). Among these habitats are 125,000 seamounts worldwide, which often host extensive hydrothermal vent systems and Fe-rich mats of microbial origin (Vissell et al., 2010; McAllister et al., 2011).

Compared to the most commonly studied marine hydrothermal systems, located at mid-ocean ridge (MOR) spreading centers, the Lō‘ihi Seamount hydrothermal system has unique fluid chemistries. For example, vent fluids at Lō‘ihi are highly enriched in CO2, CH4, NH4, PO4, Fe, and Mn, but depleted in H2S (Karl et al., 1986; Sedwick et al., 1992; Wheat et al., 2000; Edwards et al., 2004; Glazer and Rouxel, 2009). The high concentration of dissolved CO2 buffers the hydrothermal fluids at a lower pH (3.3–3.5) compared to what is most commonly observed at MOR systems (Sedwick et al., 1992; Edwards et al., 2004). The low sulfide concentration results in high hydrothermal Fe concentrations and extensive iron oxide mats by comparison to many MOR systems, where Fe predominately occurs as FeS instead. The summit of Lō‘ihi intersects the Oxygen Minimum Zone (OMZ). The low O2 concentration in bottom seawater at ∼1,000 m depth (O2 ∼50 μM) associated with this OMZ and the high Fe concentrations (up to 500 μM) in warm hydrothermal fluids (below 100°C) present beneficial conditions for FeOB, who have to compete with the abiotic oxidation of Fe (Glazer and Rouxel, 2009).

The Lō‘ihi Seamount supports abundant FeOB, and is dominated by Fe(II)-oxidizing Zetaproteobacteria, as shown in various studies (Emerson and Moyer, 2002; Rassa et al., 2009). The isolation of the first Zetaproteobacterium Mariprofundus ferrooxidans PV-1 from an iron mat at a cool (23°C) diffuse vent site at the Lō‘ihi Seamount, has initiated evaluation of the ecological significance of this class in biocorrosion (Emerson et al., 2007; Weiss et al., 2007; Singer et al., 2011). Since then Zetaproteobacteria have been predominantly found in diverse marine environments and their involvement in microbially mediated Fe(II) oxidation is widely accepted (Handley et al., 2010; Dang et al., 2011; Meyer-Dombard et al., 2012). Continued discoveries of Zetaproteobacteria in marine Fe(II) oxidizing niches raise the question whether Zetaproteobacteria could be the dominant marine FeOB. However, cultivation of these FeOB present various difficulties to date. Molecular and functional assays to study the biogeochemistry and ecology of Fe(II) oxidation, the molecular mechanism of Fe(II) oxidation, and the cultivation of environmentally representative groups of Zetaproteobacteria all have remained elusive and have consequently called for cultivation-independent molecular techniques. Full-length Zetaproteobacteria 16S rRNA sequences have so far been published from 11 regions in the world oceans and show that this class appears to follow a strong biogeographic distribution (McAllister et al., 2011). The distribution of operational
taxonomic units (OTUs) appears to be more strongly correlated with geographic occurrence than with environmental parameters, such as temperature, pH, or total Fe concentration (McAlister et al., 2011). The genome of *Mariprofundus ferrooxydans* PV-1 has provided insights into the genomic underpinnings of Fe(II)- oxidizing *Zetaproteobacteria* (Singer et al., 2011), but analysis of Fe(III)-associated iron oxyhydroxide stalks suggests that PV-1 plays a minor role at most sites (Emerson and Moyer, 2010). This study discusses the (meta-)genomic content of new *Zetaproteobacteria* lineages dominant at the Lōʻihi Seamount, Loh clone SPL-4, and Loh clone SPL-7, delivers genomic and proteomic comparisons to PV-1, and attempts to evaluate the environmental potential of *Zetaproteobacteria* at Fe-rich seamounts in the global oceans.

**MATERIALS AND METHODS**

**ENVIRONMENTAL SAMPLE SOURCE**

Iron-oxide-coated microbial biomats were collected by suction sampler using the submersible vehicle *Pico's V* at the Spillway site (Marker 34; 18.91 N, 155.26 W, 1271.95 m, $T_{max} = 83^\circ$C) at the Lōʻihi Seamount, Hawai‘i, in 2003. DNA was extracted using a phenol-chloroform extraction and purified using cesium chloride density gradient centrifugation as described by (Moore and Dowhan, 2002). The extracted and purified DNA was used to construct a library of $\sim$8,000 fosmids (each $\sim$35 kbp) with the Copy Control™ Fosmid Library Production kit (Epigenic, Madison, WI, USA) according to the manufacturer’s protocol and stored at $\sim$80°C. DNA was extracted at random from 384 fosmids (total library size $\sim$13.4 Mbp) with the BACMAX™ DNA Purification Kit (Epigenic, Madison, WI, USA) in 2011. The extraction protocol was amended by extending the incubation with the Plasmid DNA Safe mix from 20 min to 24 h under the addition of extra Plasmid-Safe DNase and ATP to minimize contaminating genomic DNA from the *Escherichia coli* fosmid vector.

**SEQUENCING AND ANNOTATION**

DNA concentrations from individual fosmids were measured on a Quibit 2.0 fluorometer and pooled at equal molar amounts. The pooled DNA was sequenced using conventional whole-genome shotgun sequencing on a Roche 454 GS FLX Titanium sequencer at the Core Genomics Center at the University of Pennsylvania and yielded 1,492,332 reads ($\sim$816 Mbp). Quality control was performed with Prinseq v0.20.1 (maximum length: 450 bp; minimum quality: 25; maximum homopolymer length: 9 bp; maximum N-tails: 1 bp) (Schmieder and Edwards, 2011) and fosmid vectors were discarded with a custom biopython script (Cock et al., 2009). The remaining metagenome ($\sim$423 Mbp) was assembled with Mira using settings: *de novo*, genome, accurate, 454 (Chevreux et al., 1999) and resulted in 3,988 contigs. These 3,988 contigs were further assembled using *Genious Pro* v. 5.4.6 (Drummond et al., 2011) leading to 2,865 final contigs with minimum length of 200 bp, average length of 1,956 bp, and average coverage of $\geq$25x. Annotation of these final 2,865 contigs was performed on the RAST (Rapid Annotation using Subsystem Technology) server version 4.0 (Aziz et al., 2008). Manual curation was performed with the SEED Viewer v. 2.0 (Overbeek, 2005). rRNA sequences were predicted using BLASTN in CAMERA (Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis; Sun et al., 2011). Alignment and phylogenetic tree construction of 16S rRNA sequences were performed with the Geneious aligner and PHYML tree builder using the Jukes-Cantor substitution model and 100 bootstraps in Geneious (Guindon and Gascuel, 2003). This Whole-Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AMFO00000000. The version described in this paper is the first version, AMFO01-1000000.

**PHYLOGENETIC BINNING**

RAST-annotated genes were aligned against the NCBI non-redundant database using the BLASTX algorithm ($E \leq 10^{-5}$) for community structure analysis. Final contigs of *Zetaproteobacteria* origin were determined by alignment against the genome of *Mariprofundus ferrooxydans* PV-1 (AASN01000000) using the BLASTX algorithm: Genes with best BLASTX matches to the genome of PV-1 average to 62.15% average amino acid identity (AAid) ($E \leq 10^{-5}$). On this basis we selected all genes with BLASTX hits of $\geq$60% AAid to PV-1 proteins to be preliminarily binned as *Zetaproteobacteria* genes. Contigs exclusively harboring best BLASTX matches to PV-1 homologs at $\geq$60% AAid were classified as *Zetaproteobacteria* contigs without further analysis. Contigs including genes homologous to PV-1 genes, but also to other bacterial classes were only included in our Zeta-subset if they passed the following requirements: (1) *Zetaproteobacteria* genes dominated in overall number and AAid or (2) remained non-*Zetaproteobacteria* genes could not be attributed to a single phylogenetic class, which therefore excluded the possibility of chimeras (Tiuwingen et al., 2011). This screening analysis resulted in 853 genes on 86 contigs with 85.57% AAi to PV-1. Ambiguous contigs, which may still belong to the genome of this *Zetaproteobacterium*, but did not pass our screening filters discussed above, were not included in the discussion of this study.

**DNA AND PROTEIN SEQUENCE ANALYSIS**

DNA sequence synteny was evaluated with the Artemis Comparson Tool (ACT; Carver et al., 2005). Protein subcellular localization analysis was performed with PSORT (Nakas and Maruyama, 1999) and Gieg-mPloc (Shen and Chou, 2010). Signal peptides and distinction between the general export pathway (Sec) and the twin-arginine translocase (Tat) mechanism were predicted on the basis of Hidden Markov Models (HMMs) using PHRED-TAT (Bagos et al., 2010). Motif analysis was performed in the Pfam database using MOTIF Search on the GenomewNet network of Kyoto University available at (http://www.genome.jp).

**RESULTS**

**PHYLOGENY**

In total, the dataset harbors five environmental 16S rRNA sequences ($\geq$500 bp), two of which are nearly full-length *Zetaproteobacteria* sequences (Table 1; Figure 1). Based on this phylogenetic marker, *Zetaproteobacteria* are well represented at Marker 34, Lō‘ihi, which is in line with earlier studies based on polymerase chain reaction (PCR)-dependent 16S rRNA analyses

---

*Singer et al.* *New Zetaproteobacteria from Lō‘ihi, Hawai‘i*
Table 1 | Best BLASTN hits of 16S rRNA sequences (>500 bp) in the Lō'ihi iron mat metagenome (retrieved 01/2013).

| Closest relative in GenBank | Class                   | Accession no. | Contig Length [bp] |
|-----------------------------|------------------------|---------------|--------------------|
| SPL OTU 1 clone 10          | Zetaproteobacteria      | JF320176      | 592                |
| Uncultured bacterium clone Poh_5 | Zetaproteobacteria     | JF320718      | 1645               |
| Methylomicrobium alcaliphilum | Gammaproteobacteria    | F0080060      | 202                |
| Methylbacterium psychrophilus | Gammaproteobacteria    | NR_025016     | 2858               |
| Uncultured bacterium clone Ld1-14 | Actinobacteria       | G0246409      | 2406               |

Zetaproteobacteria relatives are marked in bold letters. 16S rRNA sequences from potential fosmid vector contamination (Escherichia coli) are not listed here.

FIGURE 1 | Maximum-likelihood tree with 100 bootstrap cycles of full-length 16S rRNA gene sequences of all known Zetaproteobacteria species to date and representatives from other Proteobacteria classes as stated in (McAllister et al., 2011). Coloring is by geographic origin within the Pacific Ocean: red – Lō'ihi Seamount, blue – Vaiakuu Seamount/Tonga Arc/East Lau Spreading Center/Kermadec Arc, green – Southern Mariana Trough. GenBank accession numbers for published sequences are shown in parentheses. The scale bar represents 9 nucleotide substitutions per 100 positions.
(Moyer et al., 1995; Emerson and Moyer, 2002; Rassu et al., 2009). rRNA sequences from other FeOB were not detected. On the basis of their isolation from the Spillway (SPL) site at Lo‘ihi, their phylogenetic placement in OTU groups 4 and 7, and because they have not been cultured to date, we named the new Zetaproteobacteria lineages Loh clone SPL-4 and Loh clone SPL-7 (Figure 1). All Zetaproteobacteria genes presented here are assumed to belong to these two lineages.

16S rRNA genes from SPL-4 and SPL-7 are 91% identical to each other, and 90 and 93% identical to the 16S rRNA gene of *Mariprofundus ferrooxydans* PV-1, respectively. Closest relatives from OTU groups 4 and 7 have been sampled in the Northern and Southern Pacific, but were previously not considered dominant OTUs at Lo‘ihi (McAllister et al., 2011). Clones of OTU 7 have so far only been detected at the Lo‘ihi Seamount (Marker 34) representing ~5% of all local Zetaproteobacteria; clones from OTU 4 have been isolated from Markers 48 and 57 at Lo‘ihi (~~8% of local Zetaproteobacteria), the Tonga Arc, and the East Lau Spreading Center, and so far show a wider distribution in the Pacific oceans compared to clones from OTU 7 according to (McAllister et al., 2011). Although the biodiversity and biogeochemistry study by McAllister et al. (2011) is the most comprehensive to date, it is important to note that their findings completely rely on results from clone libraries, the results of which are likely to be significantly biased, for instance by Taq DNA polymerase errors and PCR template concentrations (Chandler et al., 1997; Acinas et al., 2005). Our findings could hence provide a portrait genomic traits that are shared by other Lo‘ihi strains and may be applicable to more ubiquitous Zetaproteobacteria.

**BIOGEOCHEMICALLY RELEVANT GENES**

Zetaproteobacteria genes were analyzed with respect to metabolic potential and environmental significance to the iron mat environment at Lo‘ihi. We also attempted a broader description of the nature of the Zetaproteobacteria by comparative (meta-)genomics with the genome of *Mariprofundus ferrooxydans* PV-1.

**NEW Zetaproteobacteria GENE FUNCTIONS**

Genes of potential biogeochemical relevance, which have not been described for a Zetaproteobacterium before, encode for nitrite reductase (contigs 68 and 2306, Figure 2). The nitrite reductase gene cluster on contig 68 is similar in gene content and synteny to those in the genomes of denitrifying *Thiobacillus denitrificans*, *Pseudomonas aeruginosa*, *P. stutzeri*, and *Sulfobacillus* (Rinaldo and Cutruzzolà, 2007), e.g., NirS-encoded cytochrome cd, nitrite reductase (cd,NIR) shares 82% amino acid similarity (AASim) with nitrite reductase from the versatile *Thiobacillus denitrificans* ATCC 25239 (gb|AAZ96030.1). Contig 2306 encodes copper-containing nitrite reductase NirK (CupNIR) with 80% and 70% AASim to NirK in the ammonia-oxidizing *Nitrososoccus halophilus* and *Nitrosomonas europaea*, respectively. nirK occurs in clusters next to genes encoding for cytochromes c, which have 68% and 66% AASim to NcgB and NcgC from *Nitrosomonas europaea*, as well as next to two genes encoding for multicopper oxidases type 3, which have 63% and 68% AASim to NcgA.

The two types of dissimilatory NRs containing either heme cd1 or two types of Cu centers as prosthetic groups, encoded by nirS and nirK, respectively, have not been shown to coexist in the same bacterial species, assuming that both encode nitrite reducing activity (Cutruzzolà, 1999; Jones et al., 2008). Hence either nitrite reductases are not conserved among different Zetaproteobacteria spp. or nirK and ngkABC are rather involved in nitrite detoxification, such as in *Nitrosomonas europaea* (Beaumont et al., 2002). Nevertheless, nitrite reduction may be coupled to Fe(II) oxidation and could render certain Zetaproteobacteria spp. facultative anaerobes. This would allow these strains to inhabit a wider range of environments than PV-1, which is known to be a strict microaerophile and can typically acquire energy only at <5% of air-saturated values (Emerson and Merrill Floyd, 2003; Weiss et al., 2007).

It is difficult to infer evolutionary paths of nitrite reductases, because neither nirS nor nirK are reliable phylogenetic markers (with nirS following 165 rRNA phylogenies more congruently). Since both, nirS and nirK in the Zetaproteobacteria genes share most comparable AASim to respective functional genes within the Gammaproteobacteria, it remains elusive, which of these types of nitrite reductase is more representative for the Zetaproteobacteria and whether one or both types of nitrite reductases were introduced by horizontal gene transfer (HGT).

Nitrate and nitrite reduction are encoded by various other organisms in this Lo‘ihi mat environment. Among these organisms are the hydrothermal vent-adapted, thermophilic, strictly aerobic *Marinithermus hydrothermalis*, the sulfur-oxidizing endosymbionts of * Riftia pachyptila* (vent Ph05) and of *Teucria phyclo- nana* (vent Tica), which both have very similar physiologies (Gardebrecht et al., 2011), the purple sulfur bacterium *Thiocy- ties violascens* DSM 198 (Kämpf and Pfennig, 1980), and the widely distributed *Thiobacillus denitrificans*, a facultative anaerobe, which couples inorganic sulfur oxidation as well as anaerobic oxidation of Fe(II) to denitrification at circumneutral pH.
Acidithiobacillus ferrooxidans both MobB, including the Fe$_4$S$_4$ region, are predicted to face Based on their negative charge at pH 7, the soluble parts of in the SPL-strains, contain T at signal sequences and are predicted differences among MobB within the This indicates that there are sequence and potentially structural potentials appear similar between strains. COMPARATIVE GENOMICS WITH Mariprofundus ferrooxydans PV-1 Iron oxidation The first Zetaproteobacteria gene candidates assumed to be involved in neutrophilic microaerophilic Fe(III) oxidation were detected via protein extraction from an Fe(II)-oxidizing PV-1 cell culture (Singer et al., 2011). The extracted molybdopterin oxidoreductase Fe$_4$S$_4$ region (MobB) and most of the surrounding gene cluster (ZP_01451010–ZP_01451022) was also identified in our dataset (contigs 12, 296) and shows that gene synteny is well conserved, however, split over two contigs (Figure 3). Cytochromes and MobB (contig 296) are more conserved than succinate dehydrogenases (contig 12). MobB in the SPL-strains shares 82% AASim with MobB in PV-1, while it is 59% and 57% similar to MobB in Gallionella capsiferriformans and Sideroxydans lithotrophicus, respectively. This indicates that there are sequence and potentially structural differences among MobB within the Zetaproteobacteria and in comparison to other FeOB. Both MobB proteins, in PV-1 as well as in the SPL-strains, contain T signal sequences and are predicted to harbor transmembrane helices located in the inner membrane. Based on their negative charge at pH 7, the soluble parts of both MobB, including the Fe$_4$S$_4$ region, are predicted to face the periplasm unlike previously described (Singer et al., 2011). MobB may therefore accept electrons shuttled from the outer membrane to the periplasm during Fe(III) oxidation as depicted in our revised conceptual iron oxidation model (Figure 4). Outer membrane cytochromes are likely involved in the import of electrons from ferrous iron into the cells of FeOB, such as in Acidithiobacillus ferrooxidans (Bird et al., 2011). Both, PV-1 and the SPL lineages, also harbor genes encoding type IV bioenergetics proteins (PfAMAOQ) (contig 292 and 1144), which may aid in the direct contact of Fe-species, such as in Goshater spp. (Mehta et al., 2005). As neither transcription factors, nor promoters were found in the immediate gene neighborhood vicinity of the MobB gene cluster, it remains unclear if it is in fact actively transcribed in the organism or if transcription is dependent on other genes, which are elsewhere in the genome, for instance involved in redox sensing. Despite the disconnection between contigs 296 and 12 in the SPL-strains, succinate dehydrogenases may still be part of the electron transport chain shuttling electrons during Fe(II) oxidation, however, may not necessarily be transcribed together with the MobB gene cluster. Carbon fixation Genes from our SPL lineages encode for Form Iaq ribulose-1,5-biphosphate carboxylase (RuBisCo) large subunit (contig 280) with 77% AASim to Mariprofundus ferrooxydans PV-1 (ZP_01453129) and two RuBisCo activation proteins CbbQ and CbbO (contig 280), which are 72% and 88% similar to respective genes in PV-1 (ZP_01451217, ZP_01451218), respectively. Form Iaq appears predominantly in obligate chemolithotrophs and functions best in niches with medium to low CO$_2$ concentrations (0.1–1%) and O$_2$ present (Badger and Bek, 2007). Form II RuBisCo proteins, which are present in our dataset, could not be unambiguously allocated to the Zetaproteobacteria genes, although RuBisCo large and small chain proteins (contig 135) are 87% and 90% similar to respective proteins encoded on the PV-1 genome (ZP_01453295–96). Associated RuBisCo activation proteins CbbQ and CbbO (contig 135) are 81 and 66% similar to CbbQ (ZP_01453297) and CbbO (ZP_01453298) in PV-1. Form II RuBisCo enzymes have a low discrimination threshold against O$_2$ as an alternative substrate, poor affinity for CO$_2$, and therefore potentially take over when the organism moves to a high-CO$_2$ (1.5%) and low-O$_2$ environment (Badger and Bek, 2007). At Marker 34, temperature differences between ambient seawater (2.6°C) and hydrothermal efflux (27°C) may create turbulent eddies in the water column, which would expose cells to oscillating anaerobic and microaerobic conditions, where CO$_2$ levels are variable (ranging from 2 mM to 20 mM) and dependent on positioning within the chemocline interface (Badger and Bek, 2007; Glazer and Rouxel, 2009). Utilization of both forms of RuBisCo proteins could thus enable SPL-4 and SPL-7 to optimize the acquisition of carbon under a wider range of CO$_2$ and O$_2$ concentrations inside and outside the mat in this dynamic system. Sulfide oxidation Sulfide oxidation is encoded by a sulfide-quinone reductase (contig 212), which is 83% AASim to sulfide-quinone reductase in PV-1 (ZP_01453072). The presence of a transcriptional regulator two genes further upstream (most closely related to ZP_01451744) suggests that sulfide-quinone oxidoreductase is actively transcribed. Homologs of sulfide-quinone reductases from the SPL-strains are most closely related to genes in Leptospirillum spp., as well as S. lithothrophicus. Neither of these organisms was isolated from deep-sea hydrothermal vents, but they are associated with mats dominated by FeOB (Edwards et al., 2000; Emerson et al., 2007; Golightly et al., 2009).
FIGURE 3 | Artemis Comparison Tool (ACT) analysis of gene neighborhoods around Molybdopterin oxidoreductase Fe₄S₄ region (MobB) on contigs 296 and 12 from the SPL Zetaproteobacteria and in Mariprofundus ferrooxydans PV-1.

FIGURE 4 | Conceptual iron oxidation model at neutral pH in Zetaproteobacteria revised from (Singer et al., 2011). Proteins potentially involved in energy acquisition via Fe(II) oxidation through the outer and inner membrane as predicted from genomic and protein structure and localization analysis. Besides Molybdopterin oxidoreductase Fe₄S₄ region, proteins extracted from Fe(II) oxidizing Zetaproteobacteria cultures are discussed in (Barco et al., in preparation).
The Lo‘ihi Seamount is mostly depleted of sulfide (e.g., no sulfur phases were observed in 2006), which is why Fe-oxides are the most common form of Fe(III) minerals (Clauer and Rowan, 2009). An elevation in dissolved sulfide concentration has been observed at Lo‘ihi during an eruption in 1996 (Davis et al., 2003), however, incidents like that represent the exception rather than the norm. It was recently shown that the domain of Zetaproteobacteria OTUs is more dependent on geographic factors, such as distance than environmental chemistry (McAllister et al., 2011). Because sulfide oxidation is an unfavorable metabolism at Lo‘ihi, the maintenance of sulfide and sulfite oxidation genes should be an evolutionarily unstable strategy. Hence it remains to be determined if all Zetaproteobacteria are capable of sulfide oxidation and if these sulfide-oxidizing organisms have been introduced along with other genes from FeOB, which may experience high sulfide concentrations more frequently.

**DISCUSSION**

Zetaproteobacteria IN THE GLOBAL OCEANS

This study has provided novel insights into the physiology, ecology, and genetics of novel Zetaproteobacteria strains cycling iron and carbon at a deep-sea hydrothermal vent environment. Loh clones SPL-4 and SPL-7 belong to OTUs 4 and 7, which were previously estimated to account for ~13% of the Zetaproteobacteria present at the Lo‘ihi Seamount based on 16S clone library data (McAllister et al., 2011). In our metagenomic database, which was created from samples of the same environment and not amplified prior to sequencing, these strains were dominant. The discrepancy between PCR- and fosmid-based techniques exemplifies the outcome of barely predictable biases, which should be taken into account, especially when studying microbial diversity in Fe(II) oxidizing environments. In addition, the fosmid kits used in our study were accompanied by major difficulties upon insert retrieval and led to a loss of 50% of the library, primarily due to vector contamination (see Methods section). This loss could have been reduced if whole genome sequencing on environmental DNA had been chosen over the use of fosmid libraries. Thereby a significant amount of genomic content missing from our database could have possibly been retrieved and would have potentially revealed further molecular and functional aspects of the sucetuous Zetaproteobacteria lifestyle beyond those discussed in this study. Annotation of Zetaproteobacteria genes was difficult sometimes as current databases are skewed toward Gammaproteobacteria, however, the combination of screening filters, including BLAST searches, tetranucleotide patterns and taxonomic classification models based on GliMMer interpolated context models (ICMs) was tested and enables reliable detection of most Zetaproteobacteria genes.

Comparative genomics of SPL-4 and SPL7 with *Mariprofundus ferrooxydans* PV-1, show that iron transporters, mat-specific genes, diverse oxygen-level dependent forms of RuBisCo, sulfide-quinone reductases, transmembrane phosphate uptake transporters, and heavy metal efflux pumps are well conserved and support genome-wide relatedness of these deep-sea hydrothermal vent Zetaproteobacteria as predicted from their geographical closeness. Furthermore, the SPL-strains also show parallels in metabolic potential and gene relatedness to other hydrothermal vent-native organisms, especially to endosymbionts of *R. pachyptila* and the physiologically similar endosymbiont of *Tevnia jerichonana*. Shared metabolic functions include nitrite reduction, sulfide oxidation, as well as genes typically found in mats, e.g., pilus assembly genes. Interesting is the presence of five transposases in the Zetaproteobacteria contigs, four of which are most closely related to *Leptopirillum* spp. Unfortunately, these transposases are mostly located next to genes of unknown function and do not reveal information about which (if any) functions are potentially transferred between iron-oxidizing organisms. However, the transposed potential carried between FeOB may be of future research interest as genomic similarities between iron oxidation pathways among FeOB are scarce.

Although PV-1 and the SPL-strains are not closely related on the basis of 16S rDNA, the present genomic data have revealed that key metabolic pathways are mostly conserved as operon structures. However, gene and protein sequences between the analyzed lineages can be evolutionary divergent, for example *Mobh*, assumed to play a role in Fe(II) oxidation, as well as sulfide-quinone reductases, are conserved at 82% and 83% AASim. Succinate dehydrogenases, which are well conserved in gene syntenies to respective genes in PV-1 only share 65% AASim. This indicates that despite shared geographical origin, similar adaptation strategies, and parallels in metabolic potential, Zetaproteobacteria genomes may still differ significantly, for instance in %ANI. Knowledge about how much metabolic potential can vary between Zetaproteobacteria of different geographical origin and phylogenetic affiliation, and about their global role in Fe(II) oxidation, will become available upon the expansion of the Zetaproteobacteria database.

**ACKNOWLEDGMENTS**

We thank Dr. Craig Moyer and Dr. J. Gis Kuenen for samples and valuable comments as well as John Yu and Rohan Sachdeva for bioinformatics support in this project. This work was supported by the Center for Dark Energy Biosphere Investigations (C-DEBI) (53-4809-7461 to Esther Singer), NSF Microbial Observatories Program (Award MCB-0653265 to Katrina J. Edwards), and NSF Science and Technology Center OCE 0844578 (Katrina J. Edwards).

**REFERENCES**

Azuma, M., Edae, T., Takamatsu, Y., and Dohi, T. (2001). PCR-induced sequence artifacts and biases: insights from comparison of two 16S rDNA clone libraries constructed from the same sample. Appl. Environ. Microbiol. 71, 8945–8951.

Ant, R. K., Barud, D., Best, A. A., Diyogho, M., Dues, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:375. doi:10.1186/1471-2164-9-75.

Badger, M. R., and Bink, E. J. (2007). Multiple Rambus forms in proteobacteria: their functional significance in relation to CO2 assimilation by the CBB cycle. J. Exp. Bot. 59, 1252–1261.

Bouras, P. G., Nikolaou, E. F., Latypov, Y. D., and Tsang, K. D. (2010). Combined prediction of Tat and Sec signal peptides with hidden Markov models. Bioinformatics 26, 2011–2017.

Beaumont, H. J. E., Hounsen, N., Gayraud-Senin, L. A., Arp, D. J., Akani, D. M., Hooper, A. B., et al. (2002). Nitrite reductase of *Allochromatium* euastreae is not essential for production of extracellular nitrogen oxides and confers tolerance to nitrite. J. Bacteriol. 184, 2507–2510.
Bellon, H., Chain, P., Latais, T. K., Chakraborty, A., Latimer, F., Richardson, F., et al. (2004). The genome sequence of the obligately chemolithoautotrophic, facultatively anaerobic bacterium *Thiothrix denitrificans*. J. Bacteriol. 186, 1473–1480.

Bird, L., Benfey, V., and Newman, D. (2011). Bioinformatic challenges of microbial iron metabolism. Trends Microbiol. 19, 355–360.

Baron, T. J., Benthien, K. M., Berri-
man, M., Ramadhany, M. A., Bure-
sell, R. G., and Parikh, I. (2005). A CT heterologous comparison tool. Bioinformatics 21, 3422–3423.

Chandler, D., Fredriksen, I., and Brodachen, F. (1997). Effect of PCR template concentration on the composition and purification of total community 16S rDNA clone libraries. Mol. Ecol. 6, 475–482.

Chevremu, B., Weizer, T., and Sahai, S. (1999). “Genome sequence assembly using trace signals and additional sequence information,” in *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB)*, 45–56.

Cook, P. A., Arav, T., Chang, J. T., Chapman, R. A., Cox, G. J., Dalle, A., et al. (2009). Rosboro: freely available Python tools for computational molecular biology and bioinformatics. PLoS ONE 4, e6673. doi: 10.1371/journal.pone.0025386

Carver, T. J., Rutherford, K. M., Berri-
man, M., and Rogers, D. R. (2004). Neu-
tral iron-oxidizing, chemolithoautotrophic *Ferritrophicum radici-
coli* from a microbial mat at an active, hydrothermal vent system (Tutum Bay, Papua New Guinea). J. Bacteriol. 186, 3059–3074.

Emerson, D., Rentz, J. A., Lilburn, T ., et al. (2007). Char-
acterization of a microbial mat at an active, hydrothermal vent (Tutum Bay, Papua New Guinea). J. Bacteriol. 186, 3059–3074.

Edwards, K. J., Bond, D. L., Druschel, G., Nakai, K., and Horton, P . (1999). “Genome sequence assembly using trace signals and additional sequence information,” in *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB)*, 45–56.

Emerson, D., Rentz, J. A., Lilburn, T ., et al. (2007). Characterization of a microbial mat at an active, hydrothermal vent (Tutum Bay, Papua New Guinea). J. Bacteriol. 186, 3059–3074.

Edwards, K. J., Bond, D. L., Druschel, G., Nakai, K., and Horton, P . (1999). “Genome sequence assembly using trace signals and additional sequence information,” in *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB)*, 45–56.

Emerson, D., Rentz, J. A., Lilburn, T ., et al. (2007). Characterization of a microbial mat at an active, hydrothermal vent (Tutum Bay, Papua New Guinea). J. Bacteriol. 186, 3059–3074.

Edwards, K. J., Bond, D. L., Druschel, G., Nakai, K., and Horton, P . (1999). “Genome sequence assembly using trace signals and additional sequence information,” in *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB)*, 45–56.

Emerson, D., Rentz, J. A., Lilburn, T ., et al. (2007). Characterization of a microbial mat at an active, hydrothermal vent (Tutum Bay, Papua New Guinea). J. Bacteriol. 186, 3059–3074.

Edwards, K. J., Bond, D. L., Druschel, G., Nakai, K., and Horton, P . (1999). “Genome sequence assembly using trace signals and additional sequence information,” in *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB)*, 45–56.

Emerson, D., Rentz, J. A., Lilburn, T ., et al. (2007). Characterization of a microbial mat at an active, hydrothermal vent (Tutum Bay, Papua New Guinea). J. Bacteriol. 186, 3059–3074.
New Zetaproteobacteria from Lo‘ihi, Hawai‘i
sampling of hydrothermal fluids from Lo‘ihi seamount after the 1996
event. J Geophys Res 105, 19353–
19367.
Conflict of Interest Statement: The
authors declare that the research was
conducted in the absence of any
commercial or financial relationships
that could be construed as a potential
conflict of interest.
Received: 05 November 2012; accepted:
22 February 2013; published online: 19
March 2013.
Citation: Singer E, Heidelberg JF, Dhillon
A and Edwards KJ (2013) Metage-
nomic insights into the dominant Fe(II)
oxidizing Zetaproteobacteria from an
iron mat at Lo‘ihi, Hawai‘i. Front.
Microbiol 4:52. doi: 10.3389/fmicb.2013.
00052
This article was submitted to Frontiers
in Extreme Microbiology, a specialty of
Frontiers in Microbiology.
Copyright © 2013 Singer, Heidel-
berg, Dhillon and Edwards. This
is an open-access article distributed
under the terms of the Creative
Commons Attribution License, which
permits use, distribution and repro-
duction in other forums, provided the
original authors and source are cred-
ted and subject to any copyright notices
concerning any third-party graphics
etc.