

## ORIGINAL ARTICLE

# The metatranscriptome of a deep-sea hydrothermal plume is dominated by water column methanotrophs and lithotrophs

Ryan A Lesniewski<sup>1,2</sup>, Sunit Jain<sup>1</sup>, Karthik Anantharaman<sup>1</sup>, Patrick D Schloss<sup>2</sup> and Gregory J Dick<sup>1,3,4</sup>

<sup>1</sup>Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA;

<sup>2</sup>Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI, USA;

<sup>3</sup>Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA and <sup>4</sup>Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

Microorganisms mediate geochemical processes in deep-sea hydrothermal vent plumes, which are a conduit for transfer of elements and energy from the subsurface to the oceans. Despite this important microbial influence on marine geochemistry, the ecology and activity of microbial communities in hydrothermal plumes is largely unexplored. Here, we use a coordinated metagenomic and metatranscriptomic approach to compare microbial communities in Guaymas Basin hydrothermal plumes to background waters above the plume and in the adjacent Carmen Basin. Despite marked increases in plume total RNA concentrations (3–4 times) and microbially mediated manganese oxidation rates (15–125 times), plume and background metatranscriptomes were dominated by the same groups of methanotrophs and chemolithoautotrophs. Abundant community members of Guaymas Basin seafloor environments (hydrothermal sediments and chimneys) were not prevalent in the plume metatranscriptome. *De novo* metagenomic assembly was used to reconstruct genomes of abundant populations, including Marine Group I archaea, *Methylococcaceae*, SAR324 *Deltaproteobacteria* and SUP05 *Gammaproteobacteria*. Mapping transcripts to these genomes revealed abundant expression of genes involved in the chemolithotrophic oxidation of ammonia (*amo*), methane (*pmo*) and sulfur (*sox*). Whereas *amo* and *pmo* gene transcripts were abundant in both plume and background, transcripts of *sox* genes for sulfur oxidation from SUP05 groups displayed a 10–20-fold increase in plumes. We conclude that the biogeochemistry of Guaymas Basin hydrothermal plumes is mediated by microorganisms that are derived from seawater rather than from seafloor hydrothermal environments such as chimneys or sediments, and that hydrothermal inputs serve as important electron donors for primary production in the deep Gulf of California.

*The ISME Journal* (2012) 6, 2257–2268; doi:10.1038/ismej.2012.63; published online 14 June 2012

**Subject Category:** geomicrobiology and microbial contributions to geochemical cycle

**Keywords:** metagenomics; metatranscriptomics; deep sea; vents; plume; hydrothermal

## Introduction

Deep-sea hydrothermal plumes occur throughout the mid-ocean ridges where hot fluids from hydrothermal vents are injected into the cold waters of the deep sea. Here, hydrothermal fluids laden with reduced electron donors for microbial growth (for example, H<sub>2</sub>S, H<sub>2</sub>, Fe<sup>2+</sup>, Mn<sup>2+</sup>, CH<sub>4</sub> and NH<sub>4</sub><sup>+</sup>) rapidly mix with oxic seawater, yielding geochemical gradients that fuel microbial chemosynthesis.

Hydrothermal plumes have long been recognized as sites of elevated microbial activity that is of geochemical significance from two perspectives (Cowen *et al.*, 1986; Winn *et al.*, 1986). First, chemically reduced fluids emanating from vents fuel chemosynthetic carbon fixation that may be a significant source of carbon to deep ocean waters and sediments (McCollum, 2000; Bennett *et al.*, 2011a, b). Second, microbially catalyzed chemical reactions and their end products influence the fate and reactivity of hydrothermal inputs in the oceans. For example, plume microorganisms mediate the oxidation of Fe, Mn and methane (Cowen *et al.*, 1986; de Angelis *et al.*, 1993; Mandernack and Tebo, 1993; Dick *et al.*, 2009b), and stabilize dissolved Fe(II) with organic chelators (Toner *et al.*, 2009; Sander and Koschinsky, 2011). Fe and Mn oxides

Correspondence: GJ Dick, Department of Earth and Environmental Sciences, University of Michigan, 2534 CC Little Building, 1100 N. University Avenue, Ann Arbor MI, 48109-1005, USA.

E-mail: gdick@umich.edu

Received 15 December 2011; revised 24 April 2012; accepted 9 May 2012; published online 14 June 2012

formed in plumes impact global biogeochemical budgets of elements such as P, As, Cr and V through scavenging reactions (Kadko, 1993). Thus, plume geomicrobial processes influence hydrothermal sources and sinks of critical oceanic micronutrients (Elderfield and Schultz, 1996; Tagliabue *et al.*, 2010).

Despite the significance of microorganisms in deep-sea hydrothermal vent plumes from a geochemical standpoint, plumes have attracted little microbiologically focused research compared with the intensive study of the microbiology of seafloor and subseafloor hydrothermal environments and animal symbioses. Thus, little is known about the nature of plume microbial communities or their relationship to those of the seafloor or surrounding deep seawater. Several studies have reported the SUP05 group of sulfur-oxidizing *Gammaproteobacteria* in deep-sea hydrothermal plumes in widespread geographical locations (Sunamura *et al.*, 2004; Dick and Tebo, 2010; German *et al.*, 2010). Other prevalent microbial groups identified in plumes include *Epsilonproteobacteria* (Sunamura *et al.*, 2004; Nakagawa *et al.*, 2005; German *et al.*, 2010), ammonia-oxidizing *Betaproteobacteria* (Lam *et al.*, 2008), methanotrophs (Naganuma *et al.*, 2004; Lam *et al.*, 2008) and Marine Group I (MGI) archaea (Takai *et al.*, 2004). Although these studies have expanded our knowledge of the major microbial groups present in plumes, it remains unclear which microbes are metabolically active in the plume environment, how they are linked to biogeochemical processes of interest and whether they are derived from seafloor environments (for example, chimneys and sediments) or ambient deep seawater.

Here, we investigate microbial communities in deep-sea hydrothermal vent plumes of the Guaymas Basin and in surrounding background waters of the deep Gulf of California. The Guaymas Basin hydrothermal system is overlain by organic-rich sediments that modify the chemistry of ascending hydrothermal fluids, leading to plumes that are enriched in methane and ammonium (Von Damm *et al.*, 1985). Guaymas Basin plumes feature prominent turbidity anomalies due to the production of biogenic Mn oxidation oxides, which are catalyzed by microorganisms that have yet to be identified (Campbell *et al.*, 1988; Dick *et al.*, 2009b). Despite microbially mediated Mn oxidation rates in Guaymas Basin plumes that are orders of magnitude higher than background deep Gulf of California waters (Dick *et al.*, 2009b), microbial community membership and structure are strikingly similar in plume and background, suggesting that plume biogeochemistry is underpinned either (i) at fine taxonomic scales or (ii) by shifts in metabolic activity (as opposed to shifts in community structure) (Dick and Tebo, 2010). In the current study, we evaluate these two hypotheses using high-throughput shotgun sequencing of community

DNA and RNA collected from Guaymas Basin plume and background samples. We present evidence from parallel metagenomic and metatranscriptomic analyses that enhanced microbial activity in the Guaymas Basin plume is primarily due to stimulation of particular methanotroph populations, with secondary contributions from ammonia-oxidizing archaea and sulfur-oxidizing SUP05 *Gammaproteobacteria*.

## Materials and methods

### *Sample collection and extraction of nucleic acids*

Samples were collected on three cruises aboard the *R/V New Horizon* in 2004 and 2005 as described previously (Dick *et al.*, 2009b; Dick and Tebo, 2010). Plumes were recognized by turbidity anomalies due to enrichment of particles in plumes (data presented in Dick *et al.*, 2009b) and confirmed by measurement of Mn concentrations by inductively coupled plasma mass spectrometry. Details of the location and environmental conditions of samples are provided in Table 1. One background sample (Bkgrd-1) was taken from just above the plume, and the other (Bkgrd-2) from 100 km away in neighboring Carmen Basin, which does not have hydrothermal venting. Water samples were collected in 10-l bottles by CTD-Rosette, pressure filtered with N<sub>2</sub> gas immediately (once on deck) onto 0.2- $\mu$ m 142-mm polycarbonate membranes and preserved immediately in RNAlater (Ambion, Austin, TX, USA). As the half-life of RNA molecules is short (<30 min), degradation of RNA during collection is a concern, especially for deep-sea research. Such degradation was minimized in our collection because samples were kept under *in situ* conditions (cold and dark) during CTD retrieval and then processed immediately on deck. Although we cannot exclude the possibility that some changes in the RNA pool occurred during sample processing, these artifacts should affect all samples evenly and are circumvented to some extent by the comparative nature of this study (plume versus background).

DNA was extracted from  $\frac{1}{4}$  filters as described previously (Dick and Tebo, 2010). RNA was extracted from  $\frac{1}{4}$  filters with the mirVana mRNA Isolation kit (Ambion), treated with DNase I to remove DNA, and concentrated and re-purified using the RNeasy MinElute Kit (Qiagen, Valencia, CA, USA). Total RNA was quantified by RiboGreen (Invitrogen, Carlsbad, CA, USA). RNA amplification by random priming and complementary DNA (cDNA) synthesis was performed as described previously (Frias-Lopez *et al.*, 2008; Shi *et al.*, 2011).

### *DNA sequencing and processing*

Purified genomic DNA and cDNA were prepared for pyrosequencing using standard protocols (454 Life Sciences, Roche Applied Sciences, Branford, CT, USA). Shotgun sequencing of DNA and cDNA was performed

**Table 1** Summary of samples and sequencing

Sample	Cast	Date	Latitude Longitude	Depth (m)	tMn <sup>a</sup> (nM)	dMn <sup>b</sup> (nM)	O <sub>2</sub> (μM)	CH <sub>4</sub> (μM) <sup>c</sup>	NH <sub>4</sub> <sup>+</sup> (μM) <sup>c</sup>	H <sub>2</sub> S (μM) <sup>c</sup>	T (°C)	Mn ox. rate <sup>e</sup> (nM hr <sup>-1</sup> )	Total RNA <sup>g</sup> (ng l <sup>-1</sup> )	No. sequence reads <sup>h</sup>	
														DNA	cDNA
Plume-1	11-2 #14	18 July 2004	27°0.823 111°24.654	1996	356	145	27.4	44-66	10-14	3-5	3.0	5.00 <sup>f</sup>	ND	576 187	-
Plume-2	11-1 #8	11 July 2004	27°1.852 111°24.000	1775	257	59	27.0	18-26	4-6	1-2	3.0	0.60	ND	563 818	-
Plume-3	21-6 #2	5 Feb. 2005	27°2.240 111°21.790	1963	315	62	26.1	18-28	4-6	1-2	2.9	2.35	209	-	664 240
Plume-4	12-27a #1	16 July 2004	27°30.360 111°20.818	1950	288	40	27.7	12-18	2-3	0-1	3.0	1.70	261	-	894 665
Bkgrd-1	12-8 #12	13 July 2004	27°29.174 111°21.844	1600	55	5	28.5	0-0.7 <sup>d</sup>	0-0.2 <sup>d</sup>	0	3.0	0.04	67	358 335	514 607
Bkgrd-2	34-2 #7	4 Aug. 2005	26°22.755 110°43.433	1900	21	7	46.5	0-0.7 <sup>d</sup>	0-0.2 <sup>d</sup>	0	2.6	0.04	70	406 533	504 086

Abbreviations: Aug., August; Feb., February; ND, not determined; ox., oxidation.

<sup>a</sup>Concentration of total manganese.

<sup>b</sup>Concentration of dissolved manganese.

<sup>c</sup>Estimated from end-member vent fluid concentrations (Von Damm *et al.*, 1985) and dilution factor calculated from [dMn]. These calculated concentrations match well with data from GB plumes from Lam (2004) (11.2–31.6 μM CH<sub>4</sub>; up to 3 μM NH<sub>4</sub><sup>+</sup>). Ranges indicate the calculated concentration ± 20%.

<sup>d</sup>Background data are ranges of above-sill depths in GB from Lam (2004).

<sup>e</sup>Mn oxidation rate has been shown to be microbially mediated; data from Dick *et al.* (2009b) and Dick and Tebo (2010).

<sup>f</sup>This rate experiment was inadvertently conducted at room temperature rather than at 4 °C as the others were. We present the rate (5.00 nM hr<sup>-1</sup>) estimated from the observed rate at room temperature (10.00 nM hr<sup>-1</sup>) and typical temperature response curves measured in other samples.

<sup>g</sup>Average of two independent extractions for each sample.

<sup>h</sup>Number of pyrosequencing reads retained after quality filtering.

with 454 Titanium pyrosequencing at the Dedicated Marine Microbiology DNA Sequencing Pipeline at Pennsylvania State University. Pyrosequencing reads were trimmed and quality-filtered with the mothur software package (Schloss *et al.*, 2009) according to the particular analysis as described below.

#### *De novo genomic assembly and annotation*

All four metagenomic data sets were combined into a single *de novo* metagenomic assembly using MIRA (Cheverux *et al.*, 2004) with parameters as follows: (-job = denovo.genome,accurate,454; -notraceinfo; -CL:pec = no; -GE:not = 8; -AS:urd = no; -SK:bph = 12:pr = 80 454\_SETTINGS; -AS:mrl = 50; -CO:mrpg = 3; -AL:mrs = 80). The resulting contigs were tentatively assigned into taxonomic groups by binning with tetranucleotide frequency with emergent self-organizing maps (ESOM) as described previously (Dick *et al.*, 2009a). Bins were then manually evaluated for content and accuracy through analysis of the distribution of conserved phylogenetic markers (Ciccarelli *et al.*, 2006), taxonomic profiling by BLAST, and where possible, through comparison with reference genomes. Assembled contigs were submitted to the DOE Joint Genome Institutes (JGI) Integrated Microbial Genomes website (<http://merced.jgi-psf.org/cgi-bin/mer/main.cgi>) for gene calling and annotation.

#### *rRNA classification and OTU analysis*

Transcripts containing ribosomal RNA (rRNA) genes were identified by BLASTn (evalue 1e-5 or bitscore

cutoff of 50) to a manually curated version of the SILVA databases for SSU and LSU (V-104, <http://www.arb-silva.de>) and to the RFAM database for 5/5.8S (Griffiths-Jones *et al.*, 2003). SSU-containing reads were filtered to remove all reads containing any ambiguities or any homopolymers longer than 10 base pairs (bp).

Two methods of SSU rRNA community analysis were undertaken. First, all SSU transcripts from bacteria and archaea were classified using mothur's kmer classifier (Schloss *et al.*, 2009) with the RDP (version 6) and SILVA taxonomy files ([http://www.mothur.org/wiki/Taxonomy\\_outline](http://www.mothur.org/wiki/Taxonomy_outline)). Second, for finer-scale analysis, SSU transcripts were compared with custom databases containing specific regions of the bacterial, archaeal and eukaryotic SSU gene. These databases were constructed from (i) SSU genes identified in the metagenomic assembly from this study; (ii) publicly available SSU sequences from Guaymas Basin deep-sea plumes and seafloor environments (Edgcomb *et al.*, 2002; Dhillon *et al.*, 2003; Pagé *et al.*, 2008; Dick and Tebo, 2010); and (iii) SILVA's non-redundant database (V-104), which was quality screened in mothur (minlength = 1000, maxlength = 1850, maxambig = 5 and maxhomoP = 5). Specific regions of the SSU gene were chosen for each domain according to the coverage of available sequences, which reflects different primers sets typically used for the three domains: positions 50–600 for bacteria, 300–800 for archaea (Supplementary Figure S5) and 550–1100 for Eukarya. To remove redundancy in this combined database, sequences were then clustered together at 97% identity using UCLUST (Edgar, 2010).

The resulting database was used to recruit the SSU rRNA from the Guaymas Basin metagenomic and metatranscriptomic data sets at high percent identity (>95% over >90 bp,  $e$ -value  $<e^{-40}$ ). The BLAST hit with highest bit score for each query was then used to calculate the number of transcripts recruited to each subject per data set. In cases where transcripts hit multiple subjects with equal bit score, scores were weighted based on number of subjects hit (score = 1/(number of subjects)). These scores were then normalized by the total number of rRNA sequences found for each domain in each sample.

#### Messenger RNA (mRNA) analysis

After removal of rRNA sequences, remaining transcripts were assigned to reference sequences through BLASTn to several databases. First, transcripts were compared against a database of called genes and intergenic regions from the *de novo* metagenomic assembly described in this study. Assignments were made for top hits with >95% sequence ID over >40 bp and  $e$ -value  $>e^{-10}$ . These transcripts were removed from subsequent reference searches. Second, to make assignments to more distant reference homologs, transcripts were searched with BLASTx against the NCBI non-redundant database and the Global Ocean Sampling expedition (Yooseph *et al.*, 2007). Reference sequences were identified for top hits with  $e$ -value  $<e^{-3}$ . In cases of ties, the reference sequence most frequently identified among the Guaymas Basin transcripts was assigned, as described previously (Shi *et al.*, 2011). In addition to this BLAST-based analysis, transcripts were also analyzed with MGRAST and the SEED database (Meyer *et al.*, 2008). Hits with an  $e$ -value  $<e^{-10}$ , identity >30% and overlap >40 bp were used to tally metabolic profiles of each sample.

## Results

To investigate the enhanced microbial activity of Guaymas Basin hydrothermal plumes (Dick *et al.*, 2009b), we analyzed four plume samples and two background seawater samples from deep waters of the Gulf of California (Table 1). As hydrothermal influence extends throughout deep waters of Guaymas Basin (Dick *et al.*, 2009b), background samples were chosen from two different areas outside of the deep basin; one from just above the plume and basin sill depth (Bkgrd-1), and one from neighboring Carmen Basin at a depth corresponding to the Guaymas Basin plume (Bkgrd-2) (Dick *et al.*, 2009b). Turbidity and Mn concentration, which are tracers of hydrothermal activity (Baker *et al.*, 1995), were used to identify plumes and verify that the background samples were not under hydrothermal influence. The temperature range of samples was narrow (2.6–3.0 °C) as was O<sub>2</sub> concentration

(26–29 μM) for all samples except the Carmen Basin background (46 μM) (Supplementary Table S1). No samples showed prominent temperature anomalies, thus the plumes described here represent highly dilute hydrothermal fluid that has been well mixed with deep seawater. However, large differences in geochemistry and microbial activity between plume and background were apparent as elevated concentrations of Mn (5–7 ×) and total RNA (3–4 ×) and increased rates of microbially mediated Mn oxidation (15–125 ×) (Table 1).

#### Metagenomic sequencing, *de novo* assembly and binning

Shotgun sequencing of total community DNA and RNA (cDNA) from plume and background samples yielded a total of 395 and 561 megabase (Mb) pairs of sequence for the metatranscriptomic and metagenomic data sets, respectively. The average length of quality-trimmed sequence reads was 284 bp for metagenomic data sets and 156 bp for metatranscriptomic data sets (for which smaller fragments were intentionally sequenced to capture small RNAs (sRNAs) (Shi *et al.*, 2009)). All four metagenomic data sets were used in a combined assembly to reconstruct dominant genomes from the plume and background environments. A total of 1.43 million DNA sequence reads were incorporated into contigs (67% of the data set), yielding a total of 13 151 contigs and 25 Mb of consensus sequence >1 kb in length. Among these many contigs of highly variable length and genomic coverage, a subset was of substantial length (1826 contigs >2.5 kb; maximum contig length = 65.3 kb), indicating considerable genomic assembly of the dominant community members. Manual inspection of contigs showed that polymorphisms were rare (for example, Supplementary Figure S1), indicating that the metagenomic assembly effectively distinguished genotypes at fine scales of genetic divergence.

Contigs tentatively assigned to taxonomic groups through a combination of unsupervised clustering of tetranucleotide frequency patterns (Dick *et al.*, 2009a) and genomic coverage revealed several well-defined genomic bins, some of which clustered with reference genomes (Supplementary Figure S2). Analysis of phylogenetic marker genes and taxonomic profiling of called genes showed that these bins represented four major microbial groups. Most reads assembled into genomes of the SUP05 group of *Gammaproteobacteria*, followed by the MGI archaea, the SAR324 group of *Deltaproteobacteria* and members of the *Methylococcaceae* (*Gammaproteobacteria*) (Table 2).

#### SSU rRNA transcript-based survey of active taxonomic groups in plumes and background

rRNA was not removed from community RNA before sequencing, thus metatranscriptomic data sets were

**Table 2** Summary of metagenomic bins

Genome bin	Total consensus sequence (Mb)	Avg. contig coverage	Max. contig length (kb)	Avg. contig length (kb)	Avg. GC%	Total reads (P:B) <sup>a</sup>
SUP05 ( $\gamma$ – proteobacteria)	2.9	12 ×	65.2	8.8	40	112 076 (1.5)
Marine Group I archaea <sup>b</sup>	2.9	5 ×	17.4	2.4	31	64 872 (1.4)
SAR324 ( $\delta$ – proteobacteria)	2.5	6 ×	25.9	5.4	42	53 268 (1.2)
Methylococcales ( $\gamma$ – proteobacteria)	0.9	6 ×	10.6	4.0	42	16 607 (3.0)
Unassigned	1.2	6 ×	11.0	3.5	41	22 017 (1.4)

Abbreviations: avg., average; max., maximum.

<sup>a</sup>Indicates plume:background ratio.

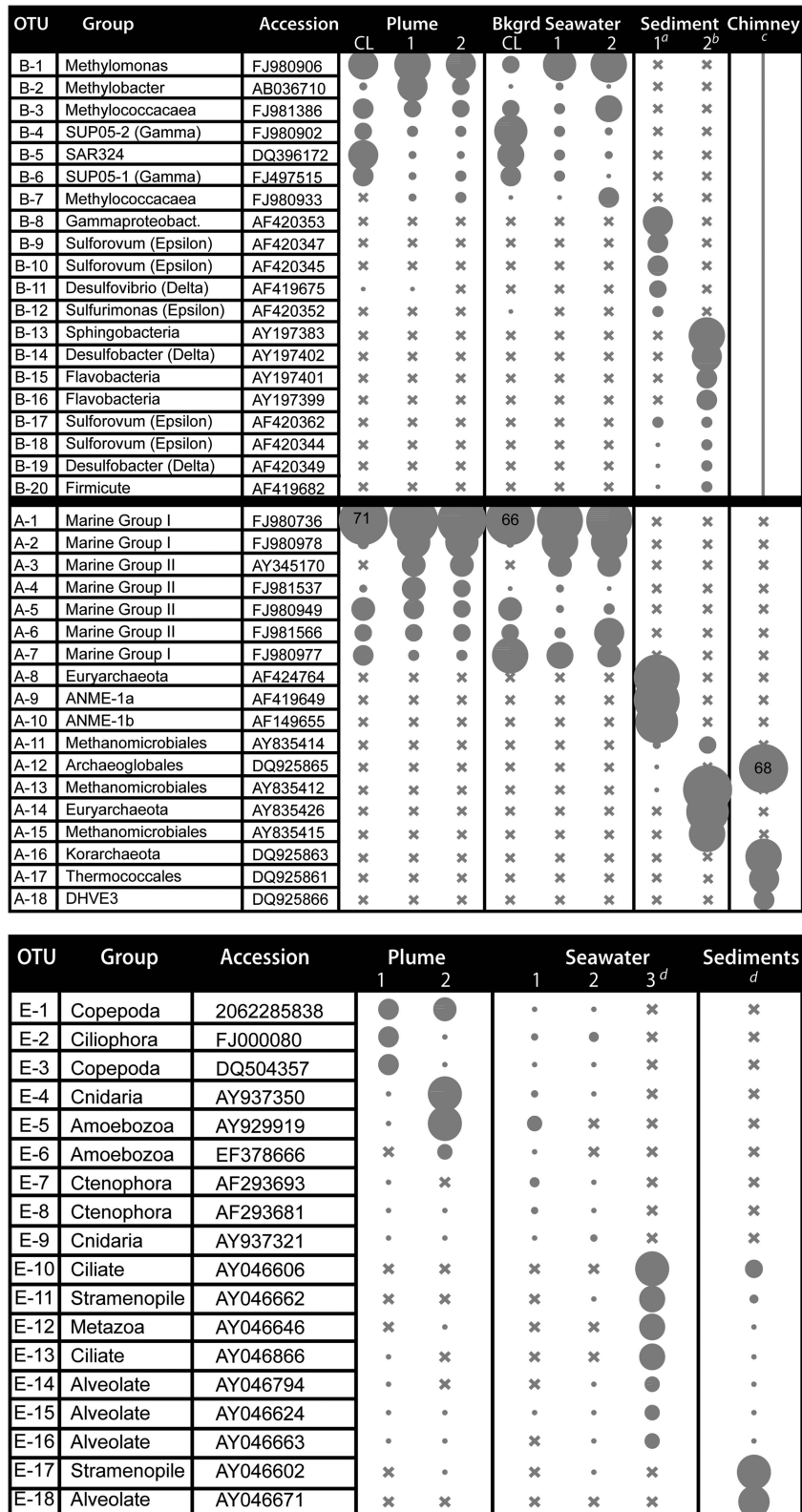
<sup>b</sup>Includes contigs identified through BLAST to *Nitrosopumulus maritimus* as described in Baker *et al.* (2012).

dominated by rRNA sequences (76–91% of all transcripts; Supplementary Table S1; Supplementary Figure S3), consistent with rRNA accounting for the vast majority of total cellular RNA. These rRNA transcript sequences potentially hold valuable information on what taxonomic groups are metabolically active in plume and background environments. Bacteria dominated rRNA transcripts in all plume and background data sets, making up nearly 85% of total rRNA reads in metatranscriptomic data sets, whereas Eukarya and archaea accounted for ~10% and 5%, respectively (Supplementary Table S1, Supplementary Figure S4). Using kmer searching against the RDP database, only ~65% of 16S sequences could be classified at domain level and nearly one-third of these could not be classified further (Supplementary Figure S4). The most abundant classified clade was *Gammaproteobacteria*, composing 18–35% of all classified sequences. At this coarse level of classification, there were no clear trends between plume and background samples, thus no prominent shifts were apparent in the taxonomic identify of active microorganisms in plume versus background. Searches against the SILVA, greengenes and NCBI databases produced similar results.

To analyze the metatranscriptomic data on finer taxonomic levels, we developed a custom BLAST-based approach for comparison of SSU-containing transcripts to custom and publicly available databases (see Materials and methods). High-stringency recruitment of transcripts to database OTUs resulted in assignment of 42 439 bacterial, 2971 eukaryotic and 2616 archaeal rRNA transcripts, with high average sequence similarity between transcript and reference OTU sequences (99.66% for bacteria, 99.68% for archaea and 99.43% for Eukarya). Seven OTUs dominated bacterial SSU transcript sequences (Figure 1), accounting for >30% of all plume transcripts assigned to bacteria. Six of these most active plume OTUs are *Gammaproteobacteria*, including four different members of the *Methylococcales* (OTUs B-1, B-2, B-3 and B-7) and two of the SUP05 group (OTUs B-4 and B-6). OTU B-5 belongs to the uncultivated SAR324 cluster of *Deltaproteobacteria*. These bacterial OTUs that dominate the plume metatranscriptome were also abundant in transcript data sets from background

seawater samples, in most cases in roughly equal abundance. One notable exception was OTU B-2, an uncultivated group of *Methylobacter*-like sequences, which was consistently more abundant (on average 5 ×) in the plume samples compared with the background (Figure 1). Plume metatranscriptome-dominant bacterial OTUs were not detected in previous bacterial diversity surveys of Guaymas Basin hydrothermal sediments (Figure 1) (Teske *et al.*, 2002; Dhillon *et al.*, 2003). Guaymas Basin sediments are dominated by *Epsilonproteobacteria* (Teske *et al.*, 2002) and *Deltaproteobacteria* (Dhillon *et al.*, 2003) (OTUs B-9 through B-20); however, these OTUs accounted for <0.1% of bacterial SSU transcripts in plume data sets and in many cases were not detected (Figure 1). No known Mn(II)-oxidizing bacteria, including those isolated previously from Guaymas Basin (Dick *et al.*, 2006; Dick and Tebo, 2010), were represented in the OTUs that recruited abundant transcripts.

SSU transcripts assigned to archaea and Eukarya were less abundant than those of bacteria in all metatranscriptomic data sets (Supplementary Figure S4). Similar to results from bacteria, the abundant archaeal and eukaryotic OTUs in the metatranscriptomic data sets were also present in background samples (Figure 1). Archaeal OTUs had similar transcript abundances in plume versus background with the notable exceptions of A-4 and A-5, Marine Group II archaea (MGII), which were overrepresented in plumes. Eukaryotic OTUs showed more variable transcript abundance across plume and background samples (Figure 1). No archaeal or eukaryotic OTUs previously identified in Guaymas Basin seafloor environments, including sediments (Edgcomb *et al.*, 2002; Teske *et al.*, 2002; Dhillon *et al.*, 2003) or high-temperature vent chimney colonization experiments (Pagé *et al.*, 2008), were among the dominant transcript types in plume or background samples. Likewise, no dominant archaeal members from plume and background metatranscriptomes were found in the sediment or chimney data sets. Archaeal OTUs that dominated the metatranscriptome were MGI and II archaea (Figure 1). Abundant eukaryotic OTUs were most similar to a variety of groups, including Copepoda, Ciliophora, Cnidaria and Amoebozoa



**Figure 1** Abundance of SSU rRNA transcripts for bacterial (top), archaeal (middle) and eukaryotic (bottom) OTUs in Guaymas Basin plumes (1 and 2) and background seawater (1 and 2). Also shown are data from DNA-based SSU rRNA gene clone libraries from Dick and Tebo (2010) (CL), hydrothermal sediment ((a): data from Teske *et al.* (2002); (b): data from Dhillon *et al.* (2003); (d): data from Edgcomb *et al.* (2002)) and hydrothermal vent chimneys ((c): data from Pagé *et al.* (2008)). Size of circles is scaled according to the number of transcripts recruited to each OTU as a percentage of bacterial/archaeal SSU transcripts (or percentage of clone libraries for previous studies). The largest circles indicate 30–40% of bacterial/archaeal SSU transcripts, with the exception of three samples, for which percentages are indicated by numbers in the circles. Smallest circles indicate 0.1–1.5% of bacterial/archaeal SSU transcripts. Crosses indicate zero transcripts were recruited.

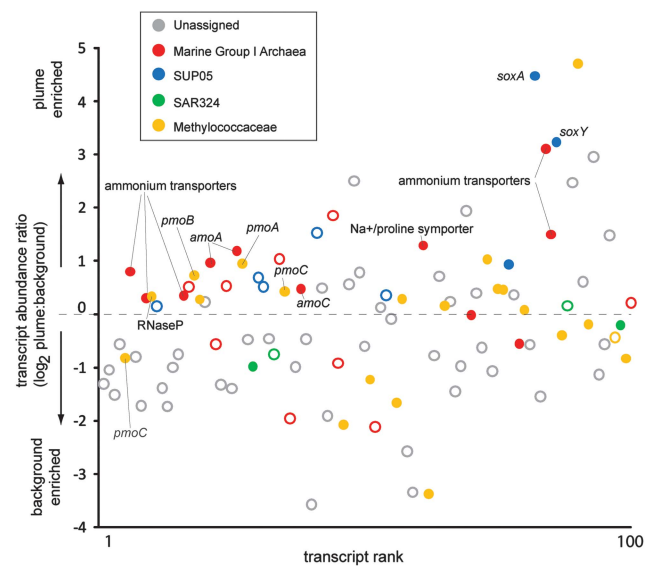
(Figure 1), which are common in marine environments (Caron *et al.*, 2009; Countway *et al.*, 2010; Sauvadet *et al.*, 2010). Many eukaryotic SSU rRNA transcripts displayed sequence similarities below 95% ID to reference sequences, suggesting an abundance of novel microbial eukaryotes.

#### Functional gene transcripts reveal major energy metabolisms in plume and background

Two main strategies were taken to characterize the 489 883 non-rRNA transcripts: (1) functional annotation via MG-RAST (Meyer *et al.*, 2008) and (2) BLAST comparison with public and custom databases. Only 58 647 of the 489 883 non-rRNA transcripts (12%) were assigned functions by MG-RAST (Supplementary Figure S6 and S7). Protein biosynthesis and central carbohydrate metabolism dominated the general functions in all transcriptomic data sets (>10%), with ammonia assimilation, RNA processing and modification, and metabolism of amino acids also contributing significantly (Supplementary Figure S7). Other notable and significant functional categories represented in the metatranscriptome were one-carbon metabolism and carbon fixation (Supplementary Figure S7). No consistent or major shifts in functional categories between plume and background were observed at this coarse functional level.

BLAST-based analysis of protein-coding transcripts gave clear signals of the most active energy metabolisms in plume and background communities. Genes for oxidation of ammonia and methane were among the most abundant transcripts, including genes for ammonium transporters and ammonia monooxygenase (*amoA* and *amoC*) highly similar to those from *Nitrosopumilus maritimus* (Walker *et al.*, 2010) of the MGI archaea, and genes for particulate methane monooxygenase (*pmoA*, *pmoB* and *pmoC*) most similar to genes from cultured *Methylobacter*, *Methylomonas* and *Methylochromium* species (Figure 2, Supplementary Table S2) (Nakamura *et al.*, 2007). The *pmo* transcripts were also highly similar to sequences retrieved from the Santa Monica Basin and North Fiji hydrothermal vent field (up to 99% nucleotide sequence identity; (Redmond *et al.*, 2010; Tavormina *et al.*, 2010)), and hydrocarbon plumes from the Deepwater Horizon oil spill in the Gulf of Mexico (99–100% nucleotide sequence identity; (Kessler *et al.*, 2011)). These sequences form the widely distributed OPU3 group of monooxygenases from uncultivated organisms, which are thought to have important roles in the oxidation of methane in the marine water column (Tavormina *et al.*, 2010).

The most abundant protein-coding transcripts identified, including the *amo* and *pmo* genes, are present in comparable relative abundances in both plume and background (Figure 2). However, several genes showed a markedly higher transcript abundance in the plume environment, including an MGI



**Figure 2** Abundance of putative functional transcripts in plume versus background samples. Vertical axis is the  $\log_2$  ratio of plume to background transcript abundance, normalized for gene length and total number of reads in each metatranscriptome data set. Positive values indicate enrichment within the plume while negative values indicate enrichment in background. Horizontal axis represents genes ranked by total number of transcripts recruited. Ranks 1 to 100 are shown; rank 1 recruited 1997 transcripts and rank 100 recruited 37 transcripts. Color of circle represents the genomic bin of the gene as shown in the legend. Closed circles indicate that a putative function has been assigned, whereas open circles indicate that gene is of unknown function.

archaea ammonium transporter (8.6 times increase) and SUP05 genes for the *sox* sulfur oxidation system (10–20 times increase) (Figure 2). Notably absent from the abundant protein-coding transcripts were genes encoding known Mn-oxidizing enzymes such as the multicopper oxidases MnxG, CumA and MofA (for example, Dick *et al.*, 2008b) or the heme peroxidase MopA (for example, Dick *et al.*, 2008a; Anderson *et al.*, 2009). Although expression of novel multicopper oxidases was observed (mRNA transcript abundance ranks 186, 282, 818 and 856; data not shown in Figure 2), this diverse family of enzymes oxidizes a wide range of substrates (Tebo *et al.*, 2004), thus further investigation is required to evaluate their potential link to Mn oxidation.

In order to capture potential sRNAs that have recently been suggested to have novel and important ecological roles in marine bacteria (Shi *et al.*, 2009), total community RNA was not size-selected before cDNA library preparation. Of the 489 883 non-rRNA transcripts, 22 989 mapped to portions of the genome annotated as intergenic regions and 1049 mapped to transfer RNAs (Supplementary Figure S3). Of the 35 intergenic regions that recruited at least 40 transcripts, 18 had matches to the Rfam database (Griffiths-Jones *et al.*, 2003). These regions annotated as conserved sRNA families (for example RNase P, transfer-messenger RNA, signal recognition partial RNA, 6S RNA and U5)

and were previously observed in metagenomic (Zhu *et al.*, 2007) and metatranscriptomic (Shi *et al.*, 2009) data (Supplementary Table S3). Seven other highly transcribed intergenic regions showed similarity to intergenic regions of other complete genome sequences, and the remaining ten showed no similarity to NCBI-nt database entries. The abundance of putative sRNAs identified in our deep Guaymas Basin data sets (7% of unassigned transcripts) is lower than that identified from Hawaii Ocean Time-series (> 28%) (Shi *et al.*, 2009). No clear shifts in the abundance of sRNA transcripts were observed between plume and background (Supplementary Table S3).

## Discussion

The Guaymas Basin hydrothermal plume shows enhanced microbial activity in the form of total RNA concentrations and microbially mediated Mn oxidation rates compared with surrounding seawater in the deep Gulf of California (Table 1). We used metagenomics and metatranscriptomics to assess the diversity and function of microbial communities underpinning this difference between plume and background. Through the combined use of *de novo* genomic assembly, tetra-ESOM binning and manual curation, metagenomic bins were defined and used to recruit metatranscriptomic reads to genes encoding proteins and rRNAs in a phylogenetically resolved manner, thus providing windows into the active metabolisms and taxonomic groups operating in the Guaymas Basin plume. These are some of the first insights into the active microorganisms and metabolisms of deep-sea hydrothermal plumes, which are globally distributed and have significant but poorly understood roles in the biogeochemistry of the deep oceans.

Results reveal an active hydrothermal plume microbial community that is strikingly similar to that of the surrounding deep waters of the Gulf of California, and quite distinct from seafloor communities inhabiting hydrothermal sediments and chimneys (Figure 1). This indicates that seafloor environments are not a source of metabolically active microorganisms that could account for the 3–4-fold increase in total RNA observed in Guaymas Basin plumes relative to background (Table 1), and implies an ecological boundary between the seafloor and hydrothermal plume. In contrast, many of the OTUs that dominate the plume metatranscriptome (putative methylotrophs, SUP05 and MGI) are also well-represented in background seawater, indicating a strong link between plume and ambient deep-sea communities. Surprisingly, only a few of the plume-dominant OTUs are consistently overrepresented in plume metatranscriptomes versus background, suggesting a relatively even enhancement of metabolism across the community rather than dramatic stimulation of specific community members. The

microbial groups most enriched in the plume metatranscriptome were a *Methylobacter*-like group (OTU B-2) and several groups of MGII archaea (OTUs A-4 and A-5), none of which have cultured representatives. We infer that the *Methylobacter* group, which is also overrepresented in DNA-based clone libraries (Dick and Tebo, 2010) and metagenomic data sets (Table 2), likely thrives in the plume by taking advantage of higher methane concentrations (up to 31.6  $\mu\text{M}$  compared with 16–65 nM in background (Lam, 2004)). Little is known about the physiology or metabolism of the enigmatic MGII archaea (*Euryarchaeota*) despite their moderate abundance in both surface and deep marine waters (Perenthaler *et al.*, 2002; DeLong *et al.*, 2006; Martin-Cuadrado *et al.*, 2008). The recent reconstruction of a MGII genome from surface seawater metagenomic data suggested a photo-heterotrophic lifestyle focused on degradation of protein and lipids (Iverson *et al.*, 2012). The MGII OTUs that are enriched in GB plumes share only 86–87% 16S rRNA gene sequence identity with the population described by Iverson *et al.* (2012), and we found no evidence for proteorhodopsin genes in our metagenomic or metatranscriptomic data sets, thus the GB MGII populations may be quite distinct from their surface water relatives. However, if the heterotrophic characteristics of surface MGII groups do indeed extend to deep-sea MGII populations such as those in our data sets, the plume enhancement observed here may signal a heterotrophic response to enhanced primary production in the plume.

The metagenome and mRNA metatranscriptome further highlight the importance of methane, ammonia and sulfur as electron donors for microbial growth in the plume (Figure 2). The high expression of MGI archaea genes for ammonia oxidation shows that it is among the most active metabolisms in Guaymas Basin deep-sea hydrothermal plumes. MGI are abundant in mesopelagic waters (Karner *et al.*, 2001) and are also widespread in shallow marine environments where they have important roles in nitrification (Francis *et al.*, 2005). Our results show that these organisms are actively oxidizing ammonia in a deep-sea hydrothermal plume environment that is highly enriched in ammonium due to interactions between hydrothermal fluids and sediments (Von Damm *et al.*, 1985). Interestingly, although the ammonia-oxidizing archaea were prevalent in plume and background metatranscriptomes, neither evidence for ammonia-oxidizing bacteria nor nitrite-oxidizing bacteria was detected (Baker *et al.*, 2012). Deeper genomic analysis revealed that these populations of MGI archaea share many genomic characteristics with the cultured *N. maritimus* (Walker *et al.*, 2010), but they also express genes not present in the *N. maritimus* genome, including genes for urea utilization (Baker *et al.*, 2012).

Prevalent expression of methane monooxygenases in the Guaymas Basin plume and surrounding



waters is consistent with vigorous microbial methane oxidation that has been observed in dispersing hydrothermal plumes (de Angelis *et al.*, 1993; Cowen *et al.*, 2002). Although abundant methanotrophic populations are typically restricted to environments with elevated methane, we observed molecular markers of methanotrophy not only in hydrothermal plumes of Guaymas Basin, but also in background waters above the plume (400 m above the seafloor) and nearly 100 km away from hydrothermal sources, in the Carmen Basin. This suggests that seafloor inputs of hydrocarbons have had a major and geographically widespread impact on the types and activities of microorganisms in the deep Gulf of California, leading to microbial communities with an enormous capacity for consumption of seafloor-derived methane. These results are consistent with prevalent diffusive seepage of methane from sediments, potentially including low-temperature surface sediments away from active hydrothermal venting (Lizarralde *et al.*, 2011; Biddle *et al.*, 2012). Interestingly, we recovered abundant transcript sequences encoding methane monooxygenases that were 100% identical to methanotrophs recently reported from the Deep-water Horizon oil spill (Kessler *et al.*, 2011). Future work is needed to quantify methane consumption rates in the deep Gulf of California; such data will be useful to understand the extent to which microorganisms modulate seafloor methane fluxes, which has an important role in regulating global climate change.

The uncultivated SUP05 group of *Gammaproteobacteria* is also abundant and active in both plume and background environments, as evidenced by two highly represented OTUs in the plume metatranscriptome (B-4 and B-6) and by metagenomic contigs that had the highest coverage and greatest length of the entire community (Table 2). SUP05 is abundant in oxic/anoxic interfacial environments such as deep-sea hydrothermal plumes (Sunamura *et al.*, 2004; Dick and Tebo, 2010; German *et al.*, 2010), oxygen minimum zones (Stevens and Ulloa, 2008; Lavik *et al.*, 2009; Walsh *et al.*, 2009; Canfield *et al.*, 2010) and symbioses with deep-sea bivalves (Newton *et al.*, 2007). Our data set supports previous genomic studies that showed that members of this group are chemolithoautotrophic, obtaining energy by coupling sulfur oxidation to reduction of nitrate and/or oxygen (Newton *et al.*, 2007; Walsh *et al.*, 2009). The increased abundance of SUP05 *sox* gene transcripts in the Guaymas Basin plume—among the most plume-enriched genes in the metatranscriptome—provides evidence that SUP05 responds to reduced sulfur compounds commonly found within the plume environment (Figure 2). The prevalence and activity of this group in the deep sea deserves further attention, especially in light of its potential impact on the global cycling of carbon, sulfur, nitrogen and important greenhouse gases (Walsh *et al.*, 2009; Canfield *et al.*, 2010).

In addition to the taxonomic and functional insights that this study has provided, the prevalence of highly transcribed genes of unknown function is also noteworthy. This was true for abundant transcripts in the background samples, many of which were hypothetical proteins of unknown function (Figure 2). Despite rapid microbially mediated Mn(II) oxidation rates measured in GB plumes, the fact that no known Mn(II)-oxidizing bacteria or enzymes were identified in metatranscriptomic data sets highlights gaps in knowledge regarding the identity of environmentally relevant Mn(II)-oxidizing microbes. We suspect that the responsible organisms and genes are present in our data set but have not yet been linked to Mn(II) oxidation. The physiological roles and biogeochemical implications of many of the sRNAs identified in this study are also unclear. Despite growing recognition of their abundance in the environment, little is known about microbial sRNAs. Given their novelty and putative roles in regulating microbial ecological processes such as carbon metabolism and nutrient acquisition (Shi *et al.*, 2009), these sRNAs warrant further study.

## Conclusions

Overall, this study demonstrates that microbial activity in Guaymas Basin hydrothermal plumes is dominated by water column microorganisms rather than seafloor-derived microorganisms. Metatranscriptomics points to methanotrophy, ammonia oxidation and sulfur oxidation as the main energetic pathways that fuel plume productivity. Growing evidence suggests that the three chemolithoautotrophic groups that dominate Guaymas hydrothermal plumes (SUP05, MGI archaea and SAR324) are in fact ubiquitous in the deep oceans (Karner *et al.*, 2001; Suzuki *et al.*, 2004; DeLong *et al.*, 2006; Agogue *et al.*, 2008; Aristegui *et al.*, 2009; Swan *et al.*, 2011), where they contribute to dark carbon fixation that is more prevalent than previously recognized (Karl *et al.*, 1984; Herndl *et al.*, 2005; Reinthaler *et al.*, 2010). Our results imply that these populations are also major factors mediating the biogeochemical cycling of carbon, nitrogen and sulfur in dispersing deep-sea hydrothermal plumes, which occur throughout the ~60 000 km mid-ocean ridge system and are recognized as sites of significant organic carbon production and trace metal transformations (McCollum, 2000; Toner *et al.*, 2009; Dick *et al.*, 2009b; Bennett *et al.*, 2011a, b). In light of the geographically widespread occurrence of deep-sea hydrothermal plumes and the significant quantity of electron donors for chemosynthesis they inject into the deep sea, this study highlights the need to better estimate the extent to which hydrothermal vents inputs influence microbial communities in the deep oceans.

## Acknowledgements

We thank Bradley M Tebo for providing the samples that were used in this study. This project is funded in part by the Gordon and Betty Moore Foundation and the National Science Foundation (OCE 1029242). We also thank the University of Michigan Rackham Graduate School Faculty Research Fellowship Program for their support and Kathryn Iverson for assistance with bioinformatic analyses. DNA and cDNA sequencing was conducted in the laboratory of Stephan Schuster at Penn State University with the kind assistance of Lynn Tomsho. Cody Sheik, Meng Li and Brett Baker provided helpful comments and revisions to the manuscript.

## Data deposit

Sequence data for metagenomic and metatranscriptomic reads, contigs and genes have been submitted to the JGI-IMG and GenBank databases under accession numbers IMG 2061766003 and GenBank BioProject IDs 77837 and 72707. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AJXC00000000. The version described in this paper is the first version, AJXC01000000.

## References

- Agogue H, Brink M, Dinasquet J, Herndl GJ. (2008). Major gradients in putatively nitrifying and non-nitrifying Archaea in the deep North Atlantic. *Nature* **456**: 788–791.
- Anderson CR, Johnson HA, Caputo N, Davis RE, Torpey JW, Tebo BM. (2009). Mn(II) oxidation is catalyzed by heme peroxidases in 'Aurantimonas manganooxydans' strain SI85-9A1 and *Erythrobacter* sp. strain SD-21. *Appl Environ Microbiol* **75**: 4130–4138.
- Aristegui J, Gasol JM, Duarte CM, Herndl GJ. (2009). Microbial oceanography of the dark ocean's pelagic realm. *Limnol Oceanogr* **54**: 1501–1529.
- Baker BJ, Lesniewski RA, Dick GJ. (2012). Genome-enabled transcriptomics reveals archaeal populations that drive nitrification in a deep-sea hydrothermal plume. *ISME J*; e-pub ahead of print 14 June 2012; doi:10.1038/ismej.2012.64.
- Baker ET, German CR, Elderfield H. (1995). Hydrothermal plumes over spreading-center axes: global distributions and geological inferences. In: Humphris SE, Zierenberg RA, Mullineaux LS, Thomson RE (eds). *Seafloor Hydrothermal Systems: Physical, Chemical, Biological, and Geological Interactions*. American Geophysical Union: Washington, DC, pp 47–71.
- Bennett SA, Hansman RL, Sessions AL, Nakamura K, Edwards KJ. (2011a). Tracing iron-fueled microbial carbon production within the hydrothermal plume at the Loihi seamount. *Geochimica et Cosmochimica Acta* **75**: 5526–5539.
- Bennett SA, Statham PJ, Green DRH, Bris NL, McDermott JM, Prado F *et al.* (2011b). Dissolved and particulate organic carbon in hydrothermal plumes from the East Pacific Rise, 9°50'N. *Deep-Sea Res I* **58**: 922–931.
- Biddle JF, Cardman Z, Mendlovitz H, Albert DB, Lloyd KG, Boetius A *et al.* (2012). Anaerobic oxidation of methane at different temperature regimes in Guaymas Basin hydrothermal sediments. *ISME J* **6**: 1018–1031.
- Campbell AC, Gieskes JM, Lupton JE, Lonsdale PF. (1988). Manganese geochemistry in the Guaymas Basin, Gulf of California. *Geochim Cosmochim Acta* **52**: 345–357.
- Canfield DE, Stewart FJ, Thamdrup B, De Brabandere L, Dalsgaard T, Delong EF *et al.* (2010). A cryptic sulfur cycle in oxygen-minimum-zone waters off the Chilean coast. *Science* **330**: 1375–1378.
- Caron DA, Countway PD, Savai P, Gast RJ, Schnetzer A, Moorthi SD *et al.* (2009). Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Appl Environ Microbiol* **75**: 5797–5808.
- Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P. (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**: 1283–1287.
- Countway PD, Vigil PD, Schnetzer A, Moorthi SD, Caron DA. (2010). Seasonal analysis of protistan community structure and diversity at the USC Microbial Observatory (San Pedro Channel, North Pacific Ocean). *Limnol Oceanogr* **55**: 2381–2396.
- Cowen JP, Massoth GJ, Baker ET. (1986). Bacterial scavenging of Mn and Fe in a mid- to far-field hydrothermal particle plume. *Nature* **322**: 169–171.
- Cowen JP, Wen X, Popp BN. (2002). Methane in aging hydrothermal plumes. *Geochim Cosmochim Acta* **66**: 3563–3571.
- de Angelis MA, Lilley MD, Olson EJ, Baross JA. (1993). Methane oxidation in deep-sea hydrothermal plumes of the Endeavor Segment of the Juan de Fuca Ridge. *Deep-Sea Res* **40**: 1169–1186.
- DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard N-U *et al.* (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496–503.
- Dhillon A, Teske A, Dillon J, Stahl DA, Sogin ML. (2003). Molecular characterization of sulfate-reducing bacteria in the Guaymas Basin. *Appl Environ Microbiol* **69**: 2765–2772.
- Dick GJ, Lee YE, Tebo BM. (2006). Manganese(II)-oxidizing *Bacillus* spores in Guaymas Basin hydrothermal sediments and plumes. *Appl Environ Microbiol* **72**: 3184–3190.
- Dick GJ, Podell S, Johnson HA, Rivera-Espinoza Y, Bernier-Latmani R, McCarthy JK *et al.* (2008a). Genomic insights into Mn(II) oxidation by the marine alphaproteobacterium *Aurantimonas* sp. strain SI85-9A1. *Appl Environ Microbiol* **74**: 2646–2658.
- Dick GJ, Torpey JW, Beveridge TJ, Tebo BM. (2008b). Direct identification of a bacterial manganese(II) oxidase, the multicopper oxidase MnxG, from spores of several different marine *Bacillus* species. *Appl Environ Microbiol* **74**: 1527–1534.
- Dick GJ, Andersson AF, Baker BJ, Simmons SL, Thomas BC, Yelton AP *et al.* (2009a). Community-wide analysis of microbial genome sequence signatures. *Genome Biol* **10**: R85.
- Dick GJ, Clement BG, Fodrie FJ, Webb SM, Bargar JR, Tebo BM. (2009b). Enzymatic microbial Mn(II) oxidation and Mn biooxide production in the Guaymas Basin hydrothermal plume. *Geochim Cosmochim Acta* **73**: 6517–6530.
- Dick GJ, Tebo BM. (2010). Microbial diversity and biogeochemistry of the Guaymas Basin deep-sea hydrothermal plume. *Environ Microbiol* **12**: 1334–1347.

- Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.
- Edgcomb VP, Kysela DT, Teske A, Gomez AD, Sogin ML. (2002). Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proc Natl Acad Sci USA* **99**: 7658–7662.
- Elderfield H, Schultz A. (1996). Mid-ocean ridge hydrothermal fluxes and the chemical composition of the ocean. *Ann Rev Earth Planet Sci* **24**: 191–224.
- Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB. (2005). Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *Proc Natl Acad Sci USA* **102**: 14683–14688.
- Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW *et al.* (2008). Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA* **105**: 3805–3810.
- German CR, Bowen A, Coleman ML, Honig DL, Huber JA, Jakuba MV *et al.* (2010). Diverse styles of submarine venting on the ultraslow spreading Mid-Cayman Rise. *Proc Natl Acad Sci USA* **107**: 14020–14025.
- Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. (2003). Rfam: an RNA family database. *Nucleic Acids Res* **31**: 439–441.
- Herndl GJ, Reinthaler T, Teira E, van Aken H, Veth C, Pernthaler A *et al.* (2005). Contribution of Archaea to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* **71**: 2303–2309.
- Iverson V, Morris RM, Frazar CD, Berthiaume CT, Morales RL, Armbrust EV. (2012). Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota. *Science* **335**: 587–590.
- Kadko D. (1993). An assessment of the effect of chemical scavenging within submarine hydrothermal plumes upon ocean geochemistry. *Earth Planet Sci Lett* **120**: 361–374.
- Karl DM, Knauer GA, Martin JH, Ward BB. (1984). Bacterial chemolithotrophy in the ocean is associated with sinking particles. *Nature* **309**: 54–56.
- Karner MB, DeLong EF, Karl DM. (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507–510.
- Kessler JD, Valentine DL, Redmond MC, Du MR, Chan EW, Mendes SD *et al.* (2011). A persistent oxygen anomaly reveals the fate of spilled methane in the deep Gulf of Mexico. *Science* **331**: 312–315.
- Lam P, Cowen JP, Popp BN, Jones RD. (2008). Microbial ammonia oxidation and enhanced nitrogen cycling in the Endeavour hydrothermal plume. *Geochim Cosmochim Acta* **72**: 2268–2286.
- Lam TYP. (2004). Microbial ammonia oxidation in deep-sea hydrothermal plumes. University of Hawaii, Hawaii.
- Lavik G, Stuhmann T, Bruchert V, Van der Plas A, Mohrholz V, Lam P *et al.* (2009). Detoxification of sulphidic African shelf waters by blooming chemolithotrophs. *Nature* **457**: 581–584.
- Lizarralde D, Soule SA, Seewald JS, Proskurowski G. (2011). Carbon release by off-axis magmatism in a young sedimented spreading centre. *Nat Geosci* **4**: 50–54.
- Mandernack KW, Tebo BM. (1993). Manganese scavenging and oxidation at hydrothermal vents and in vent plumes. *Geochim Cosmochim Acta* **57**: 3907–3923.
- Martin-Cuadrado AB, Rodriguez-Valera F, Moreira D, Alba JC, Ivars-Martinez E, Henn MR *et al.* (2008). Hindsight in the relative abundance, metabolic potential and genome dynamics of uncultivated marine archaea from comparative metagenomic analyses of bathypelagic plankton of different oceanic regions. *ISME J* **2**: 865–886.
- McCollum TM. (2000). Geochemical constraints on primary productivity in submarine hydrothermal vent plumes. *Deep-Sea Res* **47**: 85–101.
- Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M *et al.* (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinform* **9**: 386.
- Naganuma T, Elsaied HE, Hayashi T. (2004). Molecular analysis of deep-sea hydrothermal vent aerobic methanotrophs by targeting genes of 16S rRNA and particulate methane monooxygenase. *Mar Biotechnol* **6**: 503–509.
- Nakagawa S, Takai K, Inagaki F, Hirayama H, Nunoura T, Horikoshi K *et al.* (2005). Distribution, phylogenetic diversity and physiological characteristics of epsilon-proteobacteria in a deep-sea hydrothermal field. *Environ Microbiol* **7**: 1619–1632.
- Nakamura T, Hoaki T, Hanada S, Maruyama A, Kamagata Y, Fuse H. (2007). Soluble and particulate methane monooxygenase gene clusters in the marine methanotroph *Methylomicrobium* sp strain NI. *FEMS Microbiol Lett* **277**: 157–164.
- Newton IL, Woyke T, Auchtung TA, Dilly GF, Dutton RJ, Fisher MC *et al.* (2007). The *Calyptogena magnifica* chemoautotrophic symbiont genome. *Science* **315**: 998–1000.
- Pagé A, Tivey MK, Stakes DS, Reysenbach AL. (2008). Temporal and spatial archaeal colonization of hydrothermal vent deposits. *Environ Microbiol* **4**: 874–884.
- Pernthaler A, Preston CM, Pernthaler J, DeLong EF, Amann R. (2002). Comparison of fluorescently labeled oligonucleotide and polynucleotide probes for the detection of pelagic marine bacteria and archaea. *Appl Environ Microbiol* **68**: 661–667.
- Redmond MC, Valentine DL, Sessions AL. (2010). Identification of novel methane-, ethane-, and propane-oxidizing bacteria at marine hydrocarbon seeps by stable isotope probing. *Appl Environ Microbiol* **76**: 6412–6422.
- Reinthal T, van Aken HM, Herndl GJ. (2010). Major contribution of autotrophy to microbial carbon cycling in the deep North Atlantic's interior. *Deep-Sea Res Pt II* **57**: 1572–1580.
- Sander SG, Koschinsky A. (2011). Metal flux from hydrothermal vents increased by organic complexation. *Nat Geosci* **4**: 145–150.
- Sauvadet AL, Gobet A, Guillou L. (2010). Comparative analysis between protist communities from the deep-sea pelagic ecosystem and specific deep hydrothermal habitats. *Environ Microbiol* **12**: 2946–2964.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.
- Shi Y, Tyson GW, DeLong EF. (2009). Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* **459**: 266–269.
- Shi Y, Tyson GW, Eppley JM, DeLong EF. (2011). Integrated metatranscriptomic and metagenomic

- analyses of stratified microbial assemblages in the open ocean. *ISME J* **5**: 999–1013.
- Stevens H, Ulloa O. (2008). Bacterial diversity in the oxygen minimum zone of the eastern tropical South Pacific. *Environ Microbiol* **10**: 1244–1259.
- Sunamura M, Higashi Y, Miyako C, Ishibashi J, Maruyama A. (2004). Two *Bacteria* phylotypes are predominant in the Suiyo Seamount hydrothermal plume. *Appl Environ Microbiol* **70**: 1190–1198.
- Suzuki MT, Preston CM, Beja O, de la Torre JR, Steward GF, DeLong EF. (2004). Phylogenetic screening of ribosomal RNA gene-containing clones in bacterial artificial chromosome (BAC) libraries from different depths in Monterey Bay. *Microb Ecol* **48**: 473–488.
- Swan BK, Martinez-Garcia M, Preston CM, Sczyrba A, Woyke T, Lamy D *et al.* (2011). Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* **333**: 1296–12300.
- Tagliabue A, Bopp L, Dutay JC, Bowie AR, Chever F, Jean-Baptiste P *et al.* (2010). Hydrothermal contribution to the oceanic dissolved iron inventory. *Nat Geosci* **3**: 252–256.
- Takai K, Oida H, Suzuki Y, Hirayama H, Nakagawa S, Nunoura T *et al.* (2004). Spatial distribution of marine crenarchaeota group I in the vicinity of deep-sea hydrothermal systems. *Appl Environ Microbiol* **70**: 2404–2413.
- Tavormina PL, Ussler W, Joye SB, Harrison BK, Orphan VJ. (2010). Distributions of putative aerobic methanotrophs in diverse pelagic marine environments. *ISME J* **4**: 700–710.
- Tebo BM, Bargar JR, Clement BG, Dick GJ, Murray KJ, Parker D *et al.* (2004). Biogenic manganese oxides: properties and mechanisms of formation. *Ann Rev Earth Planet Sci* **32**: 287–328.
- Teske A, Hinrichs K-U, Edgcomb V, Gomez A, Kysela D, Sylva SP *et al.* (2002). Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities. *Appl Environ Microbiol* **68**: 1994–2007.
- Toner BM, Fakra SC, Manganini SJ, Santelli CM, Marcus MA, Moffett JW *et al.* (2009). Preservation of iron(II) by carbon-rich matrices in a hydrothermal plume. *Nat Geosci* **2**: 197–201.
- Von Damm KL, Edmond JM, Measures CI, Grant B. (1985). Chemistry of submarine hydrothermal solutions at Guaymas Basin, Gulf of California. *Geochim Cosmochim Acta* **49**: 2221–2237.
- Walker CB, de la Torre JR, Klotz MG, Urakawa H, Pinel N, Arp DJ *et al.* (2010). Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA* **107**: 8818–8823.
- Walsh DA, Zaikova E, Howes CG, Song YC, Wright JJ, Tringe SG *et al.* (2009). Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Science* **326**: 578–582.
- Winn CD, Karl DM, Massoth GJ. (1986). Microorganisms in deep-sea hydrothermal plumes. *Nature* **320**: 744–746.
- Yoosuf S, Sutton G, Rusch DB, Halpern AL, Williamson SJ, Remington K *et al.* (2007). The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families. *PLoS Biol* **5**: e16.
- Zhu YL, Li Y, Pulkunat DK. (2007). Deciphering RNA structural diversity and systematic phylogeny from microbial metagenomes. *Nucleic Acids Res* **35**: 2283–2294.

Supplementary Information accompanies the paper on The ISME Journal website (<http://www.nature.com/ismej>)