



Single Cell Genomics-Based Analysis of Gene Content and Expression of Prophages in a Diffuse-Flow Deep-Sea Hydrothermal System

Jessica M. Labonté^{1,2*}, Maria Pachiadaki^{1,3}, Elizabeth Fergusson¹, Jesse McNichol^{3†}, Ashley Grosche⁴, Lara K. Gulmann³, Costantino Vetriani^{4,5}, Stefan M. Sievert³ and Ramunas Stepanauskas^{1*}

OPEN ACCESS

Edited by:

Rui Zhang,
Xiamen University, China

Reviewed by:

Ian P. G. Marshall,
Aarhus University, Denmark
Nikolai Ravin,
Research Center of Biotechnology
of the Russian Academy of Sciences,
Russia

*Correspondence:

Jessica M. Labonté
labontej@tamug.edu
Ramunas Stepanauskas
rstepanauskas@bigelow.org

† Present address:

Jesse McNichol,
Department of Biological Sciences,
University of Southern California,
Los Angeles, CA, United States

Specialty section:

This article was submitted to
Aquatic Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 20 February 2019

Accepted: 21 May 2019

Published: 12 June 2019

Citation:

Labonté JM, Pachiadaki M, Fergusson E, McNichol J, Grosche A, Gulmann LK, Vetriani C, Sievert SM and Stepanauskas R (2019) Single Cell Genomics-Based Analysis of Gene Content and Expression of Prophages in a Diffuse-Flow Deep-Sea Hydrothermal System. *Front. Microbiol.* 10:1262. doi: 10.3389/fmicb.2019.01262

¹ Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, United States, ² Department of Marine Biology, Texas A&M University at Galveston, Galveston, TX, United States, ³ Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, United States, ⁴ Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ, United States, ⁵ Department of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ, United States

Phage–host interactions likely play a major role in the composition and functioning of many microbiomes, yet remain poorly understood. Here, we employed single cell genomics to investigate phage–host interactions in a diffuse-flow, low-temperature hydrothermal vent that may be reflective of a broadly distributed biosphere in the seafloor. We identified putative prophages in 13 of 126 sequenced single amplified genomes (SAGs), with no evidence for lytic infections, which is in stark contrast to findings in the surface ocean. Most were distantly related to known prophages, while their hosts included bacterial phyla Campylobacterota, Bacteroidetes, Chlorobi, Proteobacteria, Lentisphaerae, Spirochaetes, and Thermotogae. Our results suggest the predominance of lysogeny over lytic interaction in diffuse-flow, deep-sea hydrothermal vents, despite the high activity of the dominant Campylobacteria that would favor lytic infections. We show that some of the identified lysogens have co-evolved with their host over geological time scales and that their genes are transcribed in the environment. Functional annotations of lysogeny-related genes suggest involvement in horizontal gene transfer enabling host's protection against toxic metals and antibacterial compounds.

Keywords: prophages, single cell genomics, phage life cycle, hydrothermal vent, lysogeny

INTRODUCTION

Hydrothermal vents represent oases of high biological productivity in the deep-sea and are the first environments where chemolithoautotrophy was identified as a primary source of organic carbon production (Corliss et al., 1979; Jannasch and Mottl, 1985). Deep-sea hydrothermal vents are characterized by the emission of chemically-altered fluids that form through seawater–rock interactions at high temperature and pressure, resulting in fluids that are enriched in reduced chemical species such as H₂, H₂S, Fe²⁺, and CH₄, with the composition and concentration strongly dependent upon the geological setting (Kelley et al., 2002). The mixing of high-temperature hydrothermal fluids with seawater creates gradients of pH, temperature and chemicals, both above

and below the seafloor (Baross and Hoffman, 1985), generating microenvironments suitable for diverse microbial communities (Huber et al., 2003; Schrenk et al., 2003). Vent prokaryotes have adapted unique strategies to survive under these dynamic physical–chemical conditions (Nakagawa and Takai, 2008), likely dating back to the Archean (Kelley et al., 2002). Metagenomic studies of diffuse-flow deep-sea vents have revealed a high proportion of genes related to prophages and horizontal gene transfer (Anderson et al., 2011a,b). Additional evidence for the higher abundance of prophages in hydrothermal systems, as compared to the surrounding deep-sea water, has been provided by the experimental induction of prophages with mitomycin C (Williamson et al., 2008). These studies suggest that viruses in general, and temperate phages in particular, may play important roles in the ecology and evolution of microorganisms in hydrothermal systems. However, due to the technical challenges of cultivation-based isolation of the predominant lineages of microorganisms and their infecting agents, our current understanding of the phage–host interactions in hydrothermal systems remains rudimentary.

Phages can follow diverse life cycles. They can enter the lytic cycle, releasing virions following cell lysis (virulent phage), and some can enter the lysogenic cycle (temperate phage), in which case their genome becomes integrated inside the one of their hosts. Given the right conditions, functional prophages can then re-enter the lytic cycle. Other life strategies such as chronic infections and pseudolysogeny exist but remain poorly understood. Phages, especially temperate phages, can serve as vectors for the horizontal transfer of genes between bacterial species or even genera via transduction (Canchaya et al., 2003). By establishing a quasi-stable relationship with their hosts through lysogeny, temperate phages may benefit the host population by increasing the host fitness through the expression of new phenotypes, a process called lysogenic conversion (Levin and Lenski, 1983). Other advantages of carrying prophages may include the immunity of the host against phage infections of related viruses through the prophage repressor and superinfection exclusion genes (Canchaya et al., 2003). Lysogeny may be very common among isolated strains of bacteria, with an estimated 60–70% of sequenced genomes containing prophage-like sequences (Casjens, 2003; Paul, 2008). In environmental samples, the proportion of lysogens, or bacterial genomes where prophages have been identified in, highly depends on the environment and varies between 0 and 100% (Paul, 2008).

To gain a better understanding of the hosts, genomic content, and roles of phages and prophages in diffuse-flow hydrothermal systems, we analyzed genomic sequences of 126 single amplified genomes (SAGs) of Bacteria and Archaea from the Crab Spa vent on the East Pacific Rise at 9°N (McNichol et al., 2016, 2018). We identified 13 phage sequences and found no evidence for lytic infections, suggesting that they are all prophages. The genome content of the identified prophages suggests complex evolutionary histories and potential roles in the adaptation of their hosts to the dynamic hydrothermal vent environment.

MATERIALS AND METHODS

Generation and Sequencing of Single Amplified Genomes (SAGs)

Fluid samples were collected from Crab Spa (9°50.3981 N, 104°17.4942 W), a diffuse-flow deep-sea hydrothermal vent located on the East Pacific Rise at a water depth of 2,506 m (McNichol et al., 2016, 2018), during oceanographic cruises AT15-38 (2008) and AT26-10 (2014). In 2008, vent fluids for SAG analyses were collected with titanium water samplers (Von Damm et al., 1985). In 2014, additional batch incubations under simulated *in situ* conditions were performed with isobaric gas-tight samplers (Seewald et al., 2001) after amendments with nitrate, hydrogen and/or oxygen at either *in situ* temperature or 50°C (McNichol et al., 2016, 2018). SAGs of planktonic Bacteria and Archaea were generated [using cell sorting and multiple displacement amplification (MDA)] and identified (via sequencing of their SSU rRNA genes) from un-manipulated field samples (2008, 2014) and from isobaric incubations (2014), following previously described procedures (Stepanuskas et al., 2017). A total of 126 SAGs were selected for genomic sequencing as follows (see **Supplementary Table 1**): (a) 20 *Sulfurovum* SAGs with identical SSU rRNA gene sequences (defined as “*Sulfurovum* population”) belonging the class Campylobacteria (formerly Epsilonproteobacteria, Waite et al., 2017, 2018) – the predominant class of bacteria in this and many other hydrothermal vents (Sievert and Vetriani, 2012; McNichol et al., 2018); (b) 26 SAGs representing a wide spectrum of Campylobacteria (defined as “Diverse Campylobacteria”); (c) 32 SAGs representing other, phylogenetically diverse lineages of bacteria and archaea that are known to be ubiquitous in hydrothermal vents or other subsurface environments (defined as “Other community members”); and (d) 48 SAGs representing deep branches of the tree of life with no or few cultured representatives (defined as “Microbial dark matter”). These SAGs were sequenced using Illumina technology and *de novo* assembled as previously described (Stepanuskas et al., 2017). The generation, identification, and genomic sequencing of SAGs were performed at Bigelow Laboratory Single Cell Genomics Center.¹

Phage Identification

Phage contigs were identified using a combination of viral marker genes [viral genes and families (*phage, *virus, virion, prophage, terminase, capsid, head, tail, fiber, baseplate, portal, lysis, structural, T4, lambda, mu, lambdoid, podo*, myovir*, siphovir* excluding integrase and transposase as they are also found in bacterial genomes and can be unrelated to phages)] tRNAs (which are common prophage integration sites), DNA sequence anomalies (GC skew and tetramer frequencies), and metagenomic fragment recruitment from viral and bacterial metagenomes, as previously described (Labonté et al., 2015b). The identified viral contigs were then manually confirmed using diverse comparative genomics tools (Labonté et al., 2015b). To

¹scgc.bigelow.org

obtain additional evidence, when possible, whole genome synteny comparisons were performed with EasyFig for Mac version 2.1 (Sullivan et al., 2011), with tBLASTx and the filtering of small hits and annotations option.

Phylogenetic Analyses

For the phage portal protein genes, similar sequences were recovered from the GenBank nr database (E-value < 10^{-5} ; identity > 35%). Sequences were aligned with MUSCLE (Edgar, 2004) implemented in Geneious V6.1.8. Phylogeny was performed using PhyML V3.1 (Guindon et al., 2010) with 100 bootstrap replicates and the LG model with a gamma distribution (+G), estimated rates of variation among sites and a proportion of invariable sites (+I). For the terminase gene, the sequences were recovered from the GenBank nr database, and aligned based on their amino acid charges with PROMALS (Pei et al., 2008). Phylogeny was performed using PhyML implemented in Geneious with 100 approximate likelihood ratio test (aLTR) replicates and the LG model with a gamma distribution (+G), estimated rates of variation among sites and a proportion of invariable sites (+I). For the SSU rRNA phylogeny, SSU rRNA sequences were first compared with previously deposited sequences using the RDP v10 Classifier (SSU rRNA) (Cole et al., 2014). Using SINA (Pruesse et al., 2012), the SAG SSU rRNA gene sequences were aligned with sequences selected with the RDP Seqmatch (SSU rRNA gene sequences from isolates, $\geq 1,200$ bp of good quality) from the RDP pipeline (Cole et al., 2014) using the rRNA SILVA reference alignment. A maximum likelihood tree of the SAG SSU rRNA gene sequences (100 bootstrap replicates) was constructed using phyML v3.1 (Guindon et al., 2010) with the best model [GTR model with a gamma distribution (+G), estimated rates of variation among sites and a proportion of invariable sites (+I)].

Average Nucleotide Identity Calculations

For the host genomes, average nucleotide identity (ANI) was calculated using JSpecies (Richter and Rosselló-Móra, 2009) with the ANIb parameters. Because the prophage genomes were too short for JSpecies, the genomes were aligned with the Geneious Aligner with a cost matrix of 65% similarity (Kearse et al., 2012) and the ANI was calculated as the percentage of identity in the distance matrix.

Large Volume Pump (LVP) Sampling

Filtration of hydrothermal vent fluids was carried out *in situ* with McLane large volume pump (LVP) (McLane, Falmouth, MA, United States) equipped with a hose and nozzle to collect biomass for subsequent “omic” analyses from the diffuse-flow vent sites Crab Spa (LVP2 and LVP4, 9°50.3981 N, 104°17.4942 W) and Teddy Bear (LVP5 and LVP8, 9°50.0 N, 104°17.51 W) during the AT26-10 cruise in 2014. The LVP was dropped as a free-falling instrument from R/V *Atlantis*. Once on the seafloor, it was moved by ROV *Jason* and came back to the surface within 40 min where it was picked up and brought on deck. Once on deck, the filters were taken out of the thermally-insulated filter housing, cut in sections, and stored at -80°C (the filter sections for RNA were soaked in RNAlater overnight at 4°C prior to freezing). The time

interval between the end of the pumping and preserving the filter on board the ship was at most 2 h. The amount of hydrothermal fluid pumped through the filter was 422, 345, 326, and 257 l for LVP2, 4, 5, and 8, respectively.

DNA Extraction From Large Volume Pump Filters

DNA was extracted from one $\frac{1}{4}$ ($\frac{1}{8}$ for LVP4) of the LVP filters (GF-75 combusted glass fiber filter with a pore size of $0.3\ \mu\text{m}$ on top of polyethersulfone membrane filter with a pore size of $0.2\ \mu\text{m}$) using a phenol:chloroform:isoamyl alcohol and chloroform extraction method (Green and Sambrook, 2017). Total DNA was extracted from LVP filters by adding 6 mL lysis buffer [0.73 M sucrose, 40 mM EDTA (pH 8.0), 50 mM Tris (pH 8.3)] to each filter while thawing on ice. Samples were subject to three freeze/thaw cycles using liquid nitrogen and a water bath pre-heated to 40°C . One hundred and twenty microliters of lysozyme solution (5% lysozyme in lysis buffer, $0.2\ \mu\text{m}$ filter sterilized) was added to the filter and the mixture was incubated at 37°C rotating for 45 min. Three hundred microliters of proteinase K solution (1% proteinase K in lysis buffer) and 200 μL SDS (20% solution, $0.2\ \mu\text{m}$ filter sterilized) were added to each filter, and samples were incubated at 55°C rotating for 2 h. Lysate was transferred to 50 mL sterile Teflon centrifuge tubes. Samples were extracted three times using an equal volume to the aqueous phase of phenol:chloroform:isoamyl (25:24:1; pH 8.0) and twice using an equal volume of chloroform:isoamyl alcohol (24:1). Amicon Ultra-15 30k filters were used to wash and concentrate the extracted DNA. Purified samples were washed twice with sterile TE buffer. The pure DNA was stored at -80°C until further analysis. The amount of DNA extracted from the filters was 468, 315, 71, and 52 μg for LVP2, 4, 5, and 8, respectively. Libraries for sequencing were created with Nextera XT (Illumina) reagents following manufacturer's instructions, except for purification steps, which were done with column cleanup kits (QIAGEN), and library size selection, which was done with BluePippin (Sage Science, Beverly, MA, United States), with a target size of 500 ± 50 bp. Libraries were sequenced with NextSeq 500 (Illumina) in 2×150 bp mode using v.1 reagents. The obtained sequence reads were quality-trimmed with Trimmomatic v0.32 (Bolger et al., 2014) using the following settings: -phred33 LEADING:0 TRAILING:5 SLIDINGWINDOW:4:15 MINLEN:36. Paired metagenomic reads were joined using flash version 1.2.11 and the following parameters: -x 0.05 -m 20 -M 150 (Magoč and Salzberg, 2011).

Transcriptomics of the Crab Spa Diffuse-Flow Vent

Total RNA was extracted from the LVP filters using the same protocol as for the DNA extraction. The only difference in the extraction protocol was that RNA samples were extracted twice with saturated phenol (pH 4.3) and once with chloroform:isoamyl alcohol (24:1). DNA was removed using a TURBO DNA-free kit (AM1907, Ambion) according to manufacturer's protocol. Samples were screened along with no template controls and a positive *Escherichia coli* control for contaminating bacterial

DNA using 8F (5'-AGA GTT TGA TCC TGG CTC AG-3') and 518R primers (5'-CCGTCAATTCMTTTRAGTTT-3', IDT) as previously described (Vetriani et al., 1999). DNase-treated RNA (1 µg reactions) was then amplified by *in vitro* transcription using the MessageAmp II aRNA Amplification Kit (Ambion). The aRNA generated was poly-A tailed and precipitated according to the kit instructions using NH₄OAc and ethanol. RNA (0.2–5 µg) was resuspended in 5 µL DEPC-treated water and reverse-transcribed to double-stranded cDNA using SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen, Carlsbad, CA, United States). Samples were subsequently stored at –20°C until further analyses.

cDNA was quantified with a Qubit® fluorometer using a dsDNA BR Assay kit (Q32850, Thermo Fisher), and quality was assessed using High Sensitivity D1000 ScreenTape (5067-5584, Agilent) and reagents (5067-5585, Agilent) on an Agilent 2200 TapeStation. Sequencing libraries were prepared using an Ovation Ultralow DR Multiplex System (0331-32, NuGEN) with some modifications to the manufacturer's protocol. Briefly, cDNA was sheared in a total volume of 130 µl in a microTUBE AFA Fiber Snap-Cap (520045, Covaris) using a Covaris M220 focused-ultrasonicator with the following settings: Peak Incident Power = 50; Duty Factor = 2.0; Cycles per burst = 200; Treatment Time = 160 s. Sheared cDNA was cleaned up using a MinElute Reaction Cleanup kit (28204, QIAGEN). End repair, ligation and amplification steps were performed according to manufacturer's protocol, but ligation and amplified library purification steps were performed with a QIAquick PCR Purification kit (28104, QIAGEN). Amplified libraries were size-selected on a BluePippin (Sage Sciences) using a 1.5% agarose gel cassette (BDF1510, Sage Sciences) with a size range of 300–500 bp. The size-selected libraries were cleaned up using a MinElute Reaction Cleanup kit, assessed for quality on an Agilent 2200 TapeStation using D1000 ScreenTape and reagents, and quantified with a KAPA Library Quantification Kit (KK4824, KAPA). Libraries were multiplexed and sequenced on an Illumina NextSeq 500 sequencer using a High Output v1 NextSeq kit (FC-404-1004, Illumina).

Fragment Recruitment of Metagenomic Reads and Characterization of Metagenomic Islands

Metagenomes from the Crab Spa hydrothermal vent were used to search for regions of low fragment recruitment, also known as metagenomic islands, where the metagenomic reads would not go above 75% similarity and would cover less than 50% of the region. BLAST+ v2.2.28 (Camacho et al., 2009) was used to recruit the raw metagenomic reads to each SAG assembly using BLASTn with default parameter values, except for the following: -evalue 0.00001 -soft_masking true -lcase_masking -xdrop_gap 150. Genes present in the metagenomic islands were identified using BLASTx against the GenBank non-redundant (NR) database. The metagenomes were *de novo* assembled using MEGAHIT (Li et al., 2015). VirSorter (Roux et al., 2015) was used to identify viral contigs using both the virome and the RefSeq databases.

To reveal the expressed transcripts of each SAG, fragment recruitment analysis of the paired metatranscriptomic reads against the open reading frame (ORF) was performed, using DIAMOND (Buchfink et al., 2015) at a 95% amino acid identity cut-off. The total counts per ORF were normalized to account for differences in ORF length by calculating the Fragments per ORF Kilobase per Million Mapped fragments (FPKM).

RESULTS AND DISCUSSION

Identification of Phage-Like Genome Regions

We searched for phage-like DNA sequences in *de novo* assemblies of 126 single amplified genomes (SAGs) of Bacteria and Archaea (Supplementary Table 1) that were obtained from discharged fluids of Crab Spa, a low temperature, diffuse-flow hydrothermal vent. The sequenced SAGs originated from a library of 3,190 SAGs, which was created from both field samples and laboratory experiments, during which the indigenous community was exposed to an elevated temperature and/or amendments of nitrate, oxygen or hydrogen (Sievert and Vetriani, 2012; McNichol et al., 2016, 2018; Supplementary Table 1). A total of 1,176 SAGs could be identified using the 16S rRNA gene. In agreement with prior studies of diverse hydrothermal vents, including Crab Spa (Sievert and Vetriani, 2012; McNichol et al., 2016, 2018), the SAG library was dominated at ~84% by Campylobacteria, including *Sulfurimonas* (13%), *Arcobacter* (11%), *Sulfurovum* (6%), *Campylobacteraceae* (3%), and *Thioreductor* (2%) (Supplementary Figure 1), which are believed to be predominantly chemoautotrophic primary producers that can use reduced sulfur compounds and/or hydrogen gas as electron donors in combination with nitrate, oxygen and/or sulfur as electron acceptors (Yamamoto and Takai, 2011; Sievert and Vetriani, 2012; Akerman et al., 2013; Vetriani et al., 2014; McNichol et al., 2018).

Our analysis revealed phage-like genome regions in 13 of the 126 sequenced SAGs (Table 1 and Figure 1). Campylobacteria SAGs AD-702-P16 and AC-213-G04 were found to contain sequences with some homology to a *Siphoviridae* prophage previously identified in *Nitratiruptor* sp. SB155, a campylobacterium that was isolated from an actively venting sulfide mound in the Iheya North hydrothermal field, Japan (Nakagawa et al., 2005, 2007) (Figure 1A). Thermotogae SAGs AD-726-C10 and AD-726-D06 contained sequences related to a *Siphoviridae* prophage previously found in *Marinitoga piezophila*, a bacterium belonging to Thermotogae that was isolated from the Grandbonum deep-sea hydrothermal vent site at 13°N on the East Pacific Rise (Lucas et al., 2012) (Figure 1B). Spirochaetes SAGs AC-213-D02 and AC-213-I23 contained genome regions with homology to a *Siphoviridae* prophage from *Treponema primitia*, a spirochaete previously isolated from termite hindguts (Grabner and Breznak, 2004) (Figure 1C). Finally, genome regions from seven phylogenetically diverse SAGs contained retron-like phage sequences (Figure 1D) and sequences with limited similarity in genome organization to previously characterized phages,

TABLE 1 | SAG-associated viral sequences.

SAG full name	SAG assembly size	Viral contig	Viral contig length	IMG contig ID	Sample source*
<i>Thioreductor</i> sp. SCGC AC-213-G04	560,551	NODE_9	17,784	2634579743	<i>In situ</i>
		NODE_31	5,905	2634579765	<i>In situ</i>
<i>Nautilia</i> sp. SCGC AD-702-P16	1,229,711	NODE_14	22,365	2616651450	NO ₃ and H ₂ amendment, 50°C
Thermotogae bacterium SCGC AD-726-C10	1,382,712	NODE_3	51,223	N/A	<i>In situ</i>
		NODE_21	17,560	N/A	
Thermotogae bacterium SCGC AD-702-D06	1,610,077	NODE_15	25,047	2616650705	NO ₃ and H ₂ amendment, 50°C
		NODE_39	12,417	2616650729	
Spirochaetes bacterium SCGC AC-213-D02	975,315	NODE_11	26,275	2616648138	<i>In situ</i>
Spirochaetes bacterium SCGC AC-213-I23	700,071	NODE_26	10,176	2616648280	<i>In situ</i>
<i>Arcobacter</i> sp. SCGC AD-684-A04	1,687,778	NODE_18	28,252	2616649007	H ₂ amendment, 24°C
Deltaproteobacteria bacterium SCGC AD-676-A15	1,092,595	NODE_5	36,763	2634580936	<i>In situ</i>
Deltaproteobacteria bacterium SCGC AD-133-I14	1,218,259	NODE_15	17,408	2634580581	<i>In situ</i>
Chlorobi bacterium SCGC AD-682-M19	669,902	NODE_31	28,515	2634581363	NO ₃ and H ₂ amendment, 24°C
Lentisphaerae bacterium SCGC AD-698-F20	1,000,236	NODE_1	91,859	2634581988	Control, 24°C
Proteobacteria bacterium SCGC AD-698-K10	949,159	NODE_14	22,010	2634582163	Control, 24°C
Bacteroidetes bacterium SCGC AD-726-B06	1,992,123	NODE_65	10,682	2616651820	<i>In situ</i>

**In situ* refers to fluid samples retrieved from the natural environment and immediately preserved. Other conditions refer to ~24 h incubations at *in situ* pressure as previously described (McNichol et al., 2016).

suggesting the recovery of novel phage types (**Figure 1E**). It is important to point out that some prophages in the analyzed cells may have remained undetected, due to their insufficient homology to the available reference genomes, and because of the incomplete genome recovery from SAGs (**Supplementary Table S1**). Moreover, the functionality of the phages cannot be determined given the incompleteness of the recovered prophage genomes.

Indications of Lysogeny

Many of the identified phage-like regions in SAGs were related to previously described prophages, indicating that they may constitute temperate phages. Two contigs containing phage-like sequences, Deltaproteobacteria AD-676-A15_NODE_5 and Thermotogae bacterium AD-726-C10_NODE_3, encoded both viral and host genes, providing direct evidence for the viral integration into the host's chromosome (**Table 1** and **Figure 1**). Additionally, phage-like genomic regions in SAGs Lentisphaerae AD-698-F20 and *Nautilia* sp. AD-702-P16 were flanked by tRNA genes, which are common sites of prophage integration (Lowe and Eddy, 1996). Furthermore, phage-like genomic regions encoded recombinases in SAGs Chlorobi AD-682-M19 and Thermotogae AD-726-C10, integrases in *Nautilia* sp. AD-702-P16, and transposases in five SAGs (Spirochaetes AD-213-D02, Thermotogae AD-702-D06, Thermotogae AD-726-C10, Chlorobi AD-682-M19, and Lentisphaerae AD-698-F20), providing additional indications for potential lysogeny (**Table 1** and **Figure 1**). The *de novo* assembly of prophage integration sites from short fragment reads is notoriously difficult, due to their frequent co-location with repeat regions (Zerbino and Birney, 2008), and is further impaired by the incomplete genome recovery from single cells (Labonté et al., 2015a). Thus, the relationship between the kinetics of single

cell MDA reaction and the recovery of host genome has been employed as a complementary source of evidence when discriminating lytic infections and lysogens in single cell genomics studies (Labonté et al., 2015b). Typically, a negative correlation between the MDA critical point (MDA Cp) and the genome recovery is observed with SAGs originating from cells that undergo lytic infections, forming outliers in this relationship. This is due to the infection-induced degradation of host genome while maintaining high, phage-dominated content of total DNA in a cell (Labonté et al., 2015b). In this study, none of the SAGs with phage-like sequences formed obvious outliers from this correlative pattern as it was observed in surface oceanic samples where lytic infections were more prevalent (**Figure 2**). The fact that phage sequences share homology with other prophages, the observation of complete prophages within a contig with flanked host sequences, and the MDA Cp kinetics, all suggest that most if not all of the detected phage-like genomic regions represent prophages. Thus, multiple lines of evidence suggest that the identified phages are temperate phages, or prophages.

Several other reports that rely on different research tools also suggest high prevalence of prophages in hydrothermal systems, as compared to the overlying water column (Williamson et al., 2008; Anderson et al., 2011a). Likewise, our single cell genomics analysis of *Desulforudis audaxviator* revealed an abundance of lysogens and no evidence for lytic infections in this lineage that inhabits deep, terrestrial subsurface environment (Labonté et al., 2015a). In contrast, the application of identical methodology to sort and amplify the cells, along with the same analytical tools on surface ocean bacterioplankton revealed lytic infections in a third of the analyzed SAGs while providing no indications of lysogens (Labonté et al., 2015b). This is in good agreement with our current understanding of the frequency and types of phage infections in marine bacterioplankton (Weinbauer et al.,

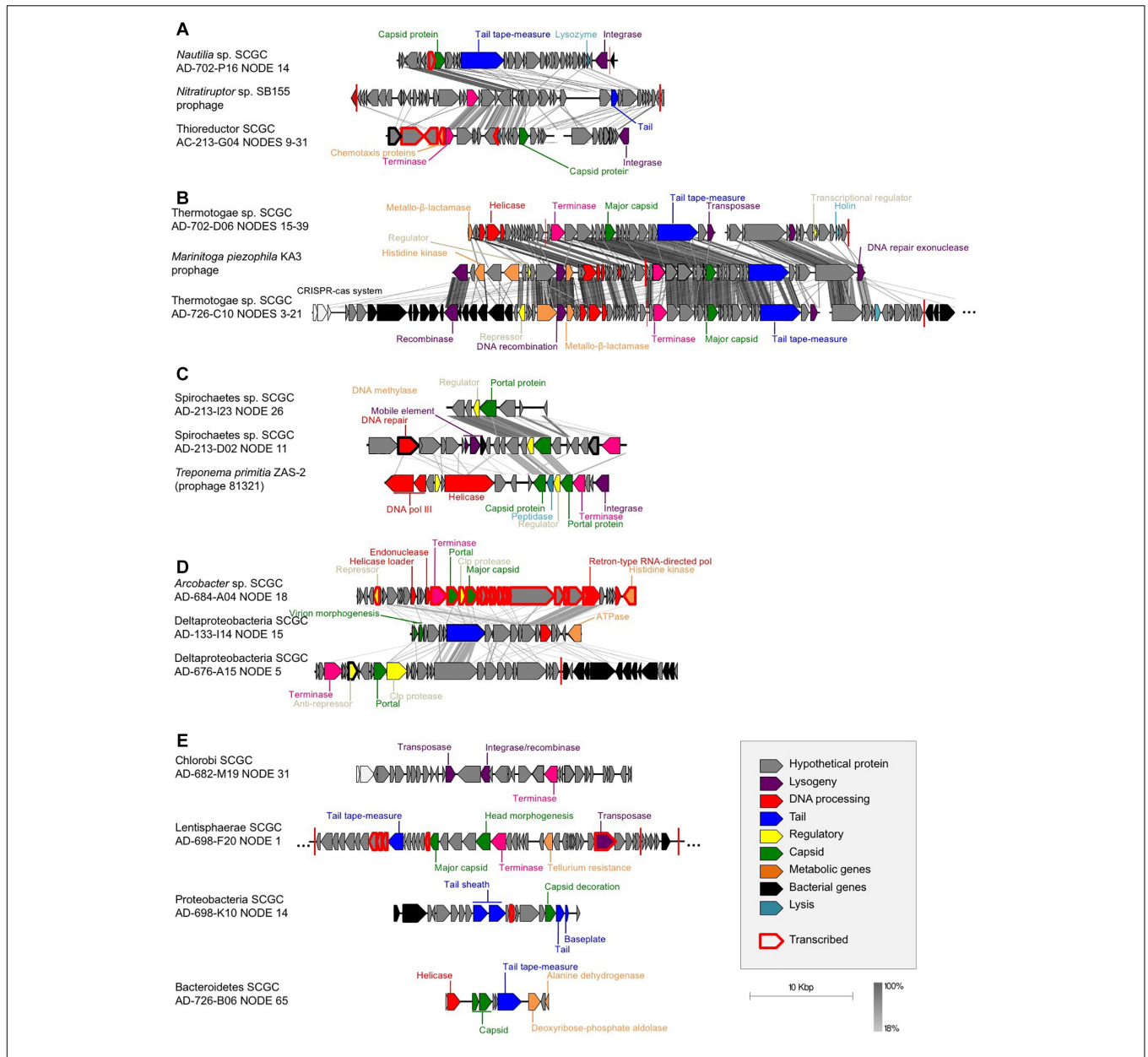
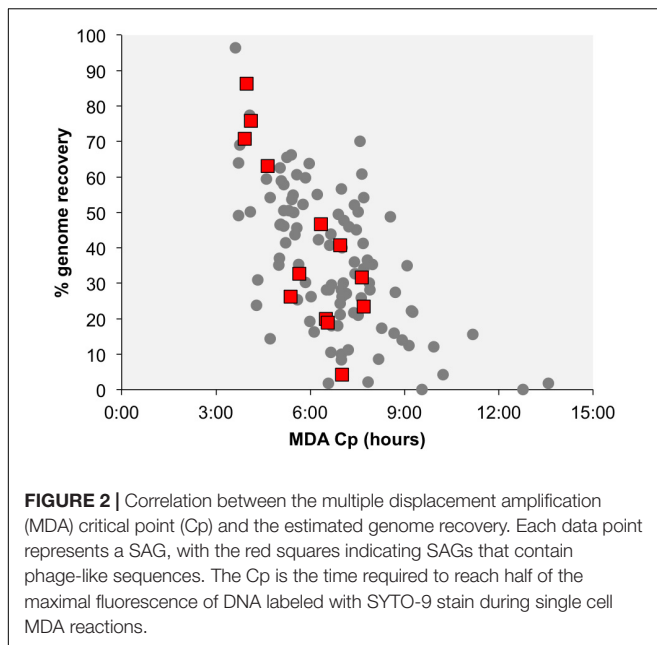


FIGURE 1 | Gene synteny between the identified SAG prophages and their closest relatives from cultured isolates: **(A)** Campylobacteria, **(B)** Thermotogae (*M. piezophila* KA3, accession # NC_016751, positions 301,979-343,554), **(C)** Spirochaetes (*T. primitia* ZAS-2, accession # NC_015578, positions 2,202,498-2,225,178), **(D)** retron-like prophages, and **(E)** prophages with no synteny with previously sequenced phages. tBLASTx was used to identify homologous regions. Each arrow represents a gene, with black arrows representing bacterial genes, showing the putative site of integration. Red vertical bars represent tRNA genes. Gray scale legend indicates DNA sequence identity between genomic sequences. Genes in red bold were found in the metatranscriptomes, which demonstrates their expression in the environment.

2003; Paul, 2008). To confirm the absence of lytic infections, we searched the Crab Spa metagenomic data (LPV2 and LVP4) to identify viral contigs using VirSorter (Roux et al., 2015). Using the virome database, no viral contig was identified. Using the RefSeq database, we found no viral contig in LPV4 (out of 103,028 contigs) and two small contigs (466 and 831 bp) in LPV2 (57,013). The absence of a strong viral signal can also point to the absence of lytic infections.

One hypothesis postulates that lysogeny is more prevalent than lytic infections in starving cells, which may not permit efficient production of phage particles (Wommack and Colwell, 2000; Weinbauer, 2004; Paul, 2008). This is, however, unlikely in the Crab Spa hydrothermal system where Campylobacteria actively divide, fix carbon, and produce biomass (McNichol et al., 2018). Another hypothesis suggests that lysogeny is a response to changes in the cell's environment and



physiology (Maurice et al., 2013), which appears relevant to deep-sea hydrothermal vents where the mixing of hydrothermal fluid with seawater creates rapidly changing conditions. However, this hypothesis is also unlikely in this particular case as the cells sampled here are adapted to these conditions and were actively growing in this mixing zone below the seafloor, most likely by forming biofilms, exploiting the available energy created by the mixing of reduced and oxidized fluids to perform chemosynthesis.

Campylobacteria represented ~90% of the SAG library, yet very few infections were observed. Since Campylobacteria were highly active (McNichol et al., 2018), a lytic infection at the time of sampling could lead to cell lysis at the time of cell preservation aboard the research vessel for subsequent single cell sorting. Perhaps the active growth of Campylobacteria explains why phages in a state of lytic infection were not found in SAGs. Infection by a lytic phage is fast under optimal conditions but could proceed slowly in starving cells. The low frequency of infected cells may reduce the probability of sorting infected cells. This could explain why no lytic infections were observed in the cells.

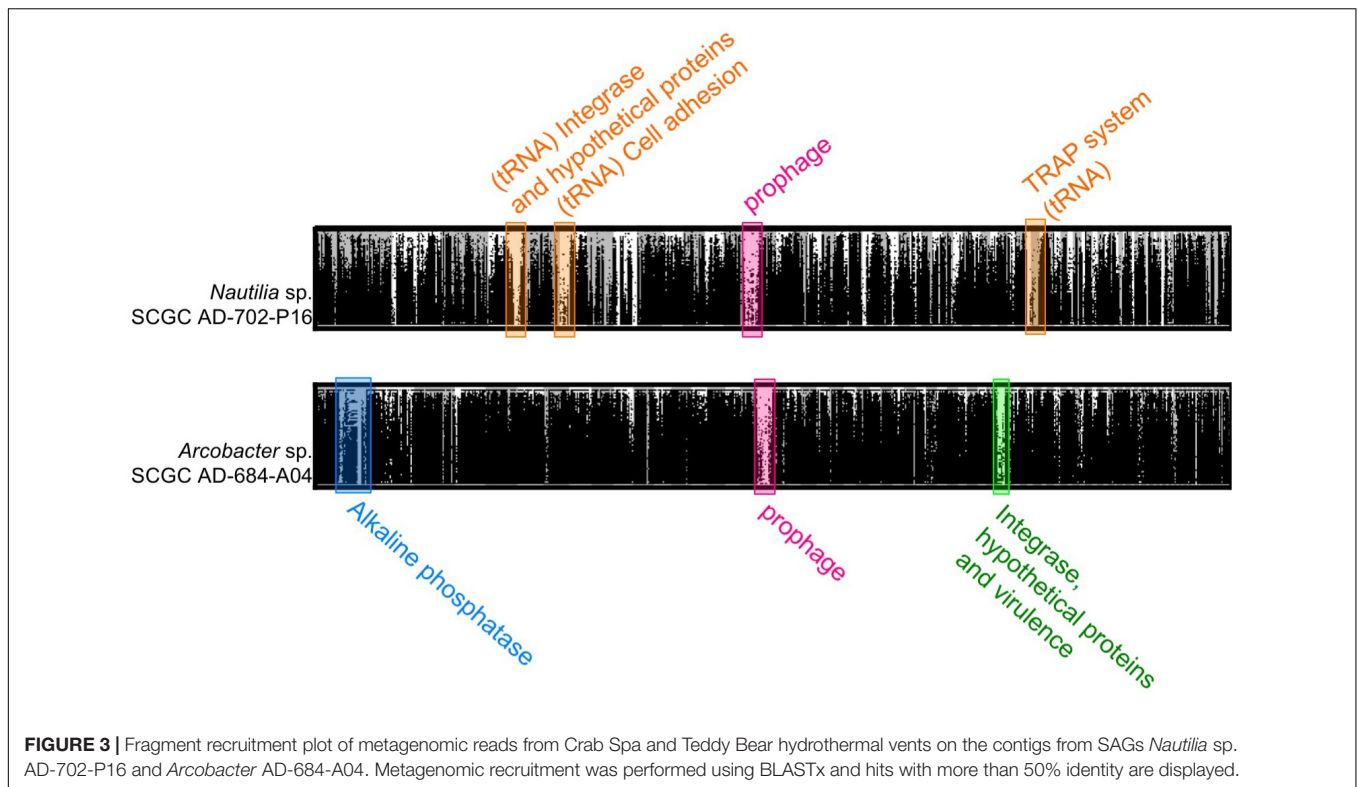
CRISPR/Cas systems have been shown to be prevalent in diffuse-flow hydrothermal system, likely providing resistance to phage infections. In one study, they were even able to associate phages to their host using CRISPR spacer sequence matches in metagenomic data (Anderson et al., 2011b). As many as 55 of the 126 (44%) analyzed SAGs, including 36 (29% of total) Campylobacteria (Supplementary Table 1), contained CRISPR/Cas, the immune system of bacteria and archaea (Boyaval et al., 2007). The prevalence of CRISPR/Cas in the Crab Spa vent may be yet another factor contributing to the resistance of these cells to phage infections.

In order to determine how conserved the putative prophages are in diffuse-flow hydrothermal vents of the study area, we

used the prophage-containing SAGs as references to recruit fragments of two metagenomes from the Crab Spa vent and two metagenomes from a nearby vent named Teddy Bear. We found prophages to be poor metagenome recruiters, as compared to other genome regions (Figure 3). This suggests that the identified prophages are present only in a small fraction of related cells that inhabit the studied hydrothermal vents. Such regions of low metagenomic recruitment are known as metagenomic islands (Rodríguez-Valera et al., 2009) and usually consist of mobile DNA elements and prophages. In the campylobacterial SAGs AD-684-A04 and AD-702-P16, which were highly represented in the metagenomes, additional metagenomic islands encoded for a cell adhesion protein, a Tripartite ATP-independent periplasmic (TRAP) transporter system, an alkaline phosphatase, an integrase, and genes involved in virulence. Products of some of these genes are potential phage recognition sites as they are genes expressed at the surface of the cells, suggesting a role in phage evasion. Previous studies also observed enrichment in phage recognition sites in metagenomic islands in marine genomes, suggesting strain-specificity in phage-interacting genes (Rodríguez-Valera et al., 2009). Here, we show that prophages and infection-related genes are a predominant type of metagenomic island in the studied hydrothermal vent.

Potential Impact of Identified Prophages on Their Hosts

Metatranscriptomic analysis of the Crab Spa fluid microbiome demonstrated that multiple prophage-encoded genes are transcribed in this environment (Figure 1; genes in bold red). The prophage of the *Arcobacter* SAG AD-684-A04 was fully transcribed, suggesting that this prophage was inducing lytic infections in some members of the vent microbial community. Four other prophages also contained some transcribed genes. Among them were transposases and integrases, suggesting that phage-mediated horizontal gene transfer and recombination are contributing to microbial adaptation and evolution. Several Crab Spa prophages contained genes for metabolic features that may benefit their hosts. It is common for prophages to express one gene or small gene cassettes when in a dormant state. These genes, referred to as “morons” may contribute to host fitness. Some genes identified in the prophages could provide a fitness advantage to their host. In one example, a prophage in the *Thioreductor* sp. SAG AC-213-G04 encodes two genes involved in energy chemotaxis (Accession numbers ACM93792 and ACM93615; Supplementary Table 2), potentially allowing for the movement of an organism toward an energy source, as previously reported for a prophage of a *Pseudoalteromonas* strain isolated from Arctic sea ice (Yu et al., 2014). In another example, a prophage in *Lentisphaerae* SAG AD-698-F20 encodes a putative tellurium resistance gene (Accession number KKT02083 and RAST annotations; Supplementary Table 2). Tellurium is a toxic element that is present in relatively high concentrations in some hydrothermal vents (Csotonyi et al., 2006). Therefore, tellurium resistance may provide a significant advantage for a vent bacterium. While many strains resistant to metalloids have been isolated from



hydrothermal vents (Rathgeber et al., 2002), this may be the first report of a tellurium resistance gene on a prophage. Yet another example is putative metallo-beta-lactamase genes on prophages of Thermotogae SAGs AD-702-D06 (accession number WP_014295825) and AD-726-C10 (accession number WP_003366035), which may be involved in a multi-drug resistance of the host (Palzkill, 2013). It is plausible that additional host-relevant features are encoded by the many hypothetical proteins on the Crab Spa prophages, the function of which remains unknown. Collectively, these data suggest complex and active roles of prophages in the Crab Spa microbial community.

We observed consistency in genome organization and host phylogeny among prophages retrieved from Spirochaetes and Thermotogae. There is partial genome synteny and the monophyly of portal proteins of prophages in Spirochaetes SAGs from hydrothermal vents and the *T. primitia* strain ZAS-2 from termite guts (Figures 1, 4). This implies phage–host co-evolution that predates the last common ancestor of Spirochaetes residing in these distant habitats. Considering that the two Spirochaetes SAGs and *T. primitia* share only 82% identity in their SSU rRNA gene, and assuming that 1% divergence of the 16S rRNA gene takes an average of 50 million years (Ochman and Wilson, 1987), it is plausible that the last common ancestor of these cells lived around one billion years ago (Figure 5). The higher average nucleotide similarity among hosts (63–64%) than the conserved regions of prophages (43–47%) shows that the divergence of these prophages outpaced the divergence of their hosts. The same is true for Thermotogae, where the average

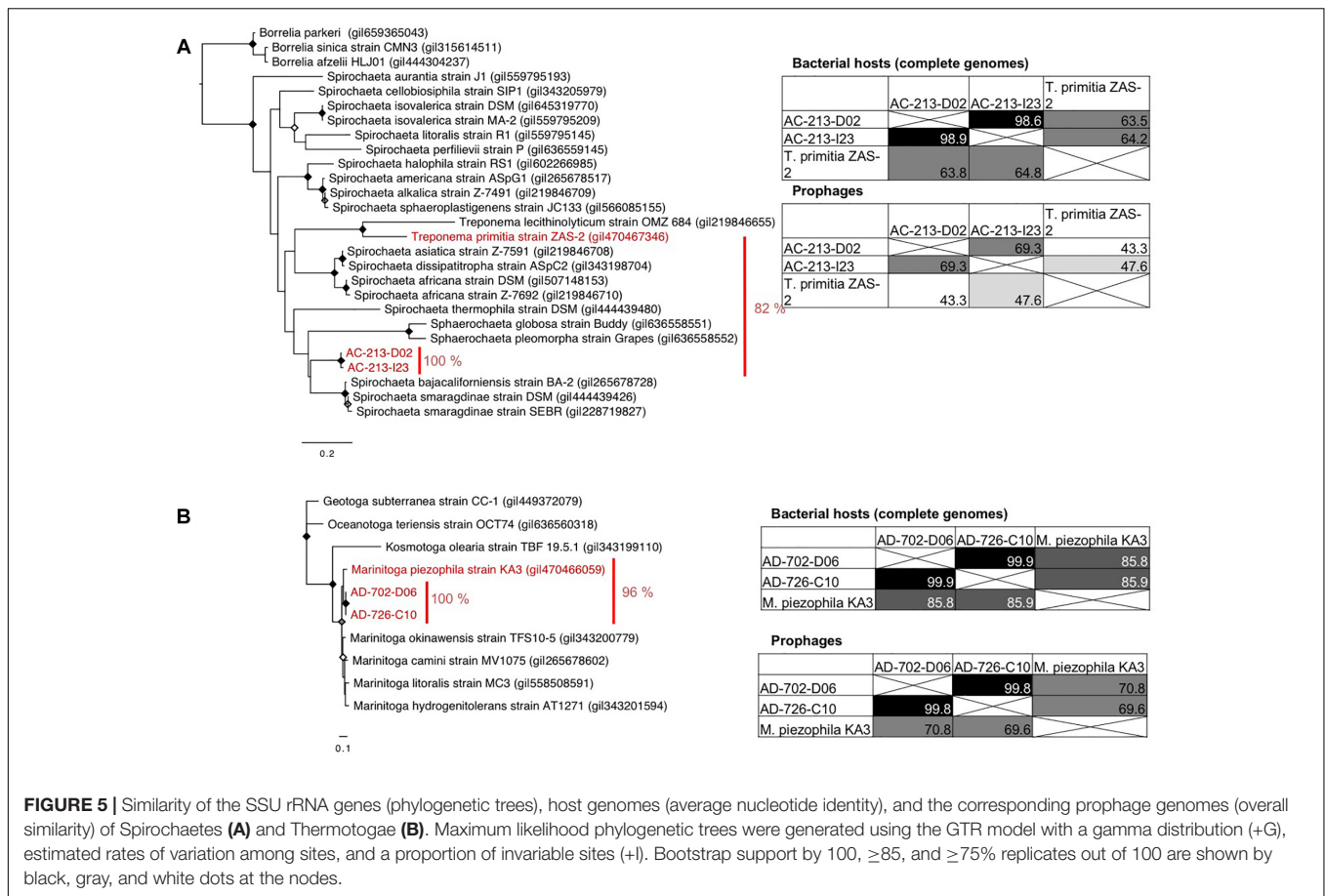
nucleotide similarity between SAGs AD-702-D06 and AD-726-C10 on the one hand and the *M. piezophila* strain KA3 on the other hand (~85%) was greater than the nucleotide similarity of the shared genes of their prophages (~70%) (Figure 5). Notably, all three Thermotogae originate from hydrothermal vents on the East-Pacific Rise and there is a 96% identity in the SSU rRNA gene sequence between SAGs and the isolate, suggesting a more recent common ancestor, as compared to Spirochaetes. These examples of prophage–host co-evolution over geological time scales provide further support for strong co-dependence of the detected prophages and their hosts, but also that the mutation rate and selection pressure differ between prophages and their hosts.

CONCLUSION

Our results suggest the predominance of lysogenic over lytic phage–host interaction in a diffuse-flow, deep-sea hydrothermal vent. This is in line with prior findings from other hydrothermal vents and contrasts with the prevalence of lytic infections in surface ocean bacterioplankton, likely reflecting differences in the stability of these two ecosystems. Resistance to phage infection in the dominant members of the community such as Campylobacteria could be due to a high prevalence of CRISPR regions. We show that some of the identified prophages have co-evolved with their host over geological time scales and that their genes are transcribed in the environment. Functional annotations of lysogeny genes suggest involvement in horizontal gene



FIGURE 4 | Phylogenetic analysis of the portal protein from Spirochaetes SAGs (red), Spirochaetes isolates (bold) and other bacteria and phages, showing segregation of the sequences based on their host. The tree was generated using maximum likelihood, with 100 bootstrap replicates, using the LG model with a gamma distribution (+G), estimated rates of variation among sites and a proportion of invariable sites (+I). >99% bootstrap support is indicated as black circles at nodes. Scale bar represents the number of amino acid substitution per site.



transfer and in protection of the host against toxic metals and anti-bacterial compounds.

DATA AVAILABILITY

Viral contig sequences, phylogenetic trees and alignments are available in **Supplementary Material** of this publication. SAG genomic sequences are available on the D.O.E. Joint Genome Institutes IMG portal, <https://img.jgi.doe.gov> (**Supplementary Table 1** for accession numbers). Metagenomes LVP2 and LVP4 and metatranscriptomes CV88 and LPV4 are available on the MG-RAST server under the accession numbers mgm4679649.3, mgm4679741.3, mgm4773053.3, and mgm4773182.3, respectively.

AUTHOR CONTRIBUTIONS

JL performed the phage identification and analysis, prepared the figures, and wrote the manuscript. MP analyzed the SAG genomes, performed the metagenomic and transcriptomic fragment recruitment, and participated in the manuscript preparation. SMS performed the SAG sampling in 2008 and the LVP sampling. EF generated the SAGs used in this study. JM performed the incubations from which SAGs were

generated and participated in the manuscript preparation. AG and LG performed the DNA and RNA extractions for the metagenomics and transcriptomics analyses, and participated in the manuscript preparation. CV, SMS, and RS designed the experiments and participated in the manuscript preparation. RS lead the generation of the SAGs, helped in data analysis and participated in the manuscript preparation.

FUNDING

This work was supported by the U.S. National Science Foundation's Dimensions of Biodiversity Program [OCE-1136488 (to RS), OCE-1136727 (to SMS) and OCE-1136451 (to CV)], as well as DEB-1441717 and OCE-1335810 (to RS), and the DOE JGI CSP project 1477.

ACKNOWLEDGMENTS

We thank the staff of the Bigelow Laboratory Single Cell Genomics Center for the generation of the single cell and metagenomics data. We thank the officers, crew, and pilots of the R/V *Atlantis*, HOV *Alvin*, and ROV *Jason* for their expert help at sea and their outstanding efforts acquiring the samples needed

for this study. We also thank the scientific parties of AT15-38 and AT26-10 for support, with special thanks to Sean Sylva, Jeff Seewald, and François Thomas for help with the shipboard lab experiments during AT26-10 as well as Craig Taylor and Miriam Sollich for their help with the LVP deployments and recovery.

REFERENCES

- Akerman, N. H., Butterfield, D. A., and Huber, J. A. (2013). Phylogenetic diversity and functional gene patterns of sulfur-oxidizing subsurface epsilon proteobacteria in diffuse hydrothermal vent fluids. *Front. Microbiol.* 4:185. doi: 10.3389/fmicb.2013.00185
- Anderson, R. E., Brazelton, W. J., and Baross, J. A. (2011a). Is the genetic landscape of the deep subsurface biosphere affected by viruses? *Front. Microbiol.* 2:219. doi: 10.3389/fmicb.2011.00219
- Anderson, R. E., Brazelton, W. J., and Baross, J. A. (2011b). Using CRISPRs as metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol. Ecol.* 77, 120–133. doi: 10.1111/j.1574-6941.2011.01090.x
- Baross, J. A., and Hoffman, S. E. (1985). Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Orig. Life Evol. Biosph.* 15, 327–345. doi: 10.1007/BF01808177
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712. doi: 10.1126/science.1138140
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST plus: architecture and applications. *BMC Bioinformatics* 10:1. doi: 10.1186/1471-2105-10-421
- Canchaya, C., Proux, C., Fournous, G., Bruttin, A., and Brüssow, H. (2003). Prophage genomics. *Microbiol. Mol. Biol. Rev.* 67, 238–276. doi: 10.1128/MMBR.67.2.238-276.2003
- Casjens, S. (2003). Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.* 49, 277–300. doi: 10.1046/j.1365-2958.2003.03580.x
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244
- Corliss, J. B., Dymond, J., Gordon, L. I., Edmond, J. M., von Herzen, R. P., Ballard, R. D., et al. (1979). Submarine thermal springs on the galapagos rift. *Science* 203, 1073–1083. doi: 10.1126/science.203.4385.1073
- Csotonyi, J. T., Stackebrandt, E., and Yurkov, V. (2006). Anaerobic respiration on tellurate and other metalloids in bacteria from hydrothermal vent fields in the Eastern Pacific Ocean. *Appl. Environ. Microbiol.* 72, 4950–4956. doi: 10.1128/AEM.00223-06
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Graber, J. R., and Breznak, J. A. (2004). Physiology and nutrition of *Treponema primitia*, an H₂/CO₂-acetogenic spirochete from termite hindguts physiology and nutrition of *Treponema primitia*, an H₂/CO₂-acetogenic spirochete from termite hindguts. *Appl. Environ. Microbiol.* 70, 1307–1314. doi: 10.1128/AEM.70.3.1307
- Green, M. R., and Sambrook, J. (2017). Isolation of high-molecular-weight DNA using organic solvents. *Cold Spring Harb. Protoc.* 2017, 356–359. doi: 10.1101/pdb.prot093450
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Huber, J. A., Butterfield, D. A., and Baross, J. A. (2003). Bacterial diversity in a subsurface habitat following a deep-sea volcanic eruption. *FEMS Microbiol. Ecol.* 43, 393–409.
- Jannasch, H. W., and Mottl, M. J. (1985). Geomicrobiology of deep-sea hydrothermal vents. *Science* 229, 717–725. doi: 10.1126/science.229.4715.717
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Kelley, D. S., Baross, J. A., and Delaney, J. R. (2002). Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annu. Rev. Earth Planet. Sci.* 30, 385–391. doi: 10.1146/annurev.earth.30.091201.141331
- Labonté, J. M., Field, E. K., Lau, M., Chivian, D., Van Heerden, E., Wommack, K. E., et al. (2015a). Single cell genomics indicates horizontal gene transfer and viral infections in a deep subsurface firmicutes population. *Front. Microbiol.* 6:349. doi: 10.3389/fmicb.2015.00349
- Labonté, J. M., Swan, B. K., Poulos, B., Luo, H., Koren, S., Hallam, S. J., et al. (2015b). Single-cell genomics-based analysis of virus–host interactions in marine surface bacterioplankton. *ISME J.* 9, 2386–2399. doi: 10.1038/ismej.2015.48
- Levin, B. R., and Lenski, R. E. (1983). “Coevolution in bacteria and their viruses and plasmids,” in *Coevolution*, eds D. J. Futuyama and M. Slatkin (Sunderland, MA: Sinauer Associates), 99–127.
- Li, D., Liu, C. M., Luo, R., Sadakane, K., and Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033
- Lowe, T. M., and Eddy, S. R. (1996). TRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955–964. doi: 10.1093/nar/25.5.0955
- Lucas, S., Han, J., Lapidus, A., Cheng, J. F., Goodwin, L. A., Pitluck, S., et al. (2012). Complete genome sequence of the thermophilic, piezophilic, heterotrophic bacterium *Marinitoga piezophila* KA3. *J. Bacteriol.* 194, 5974–5975. doi: 10.1128/JB.01430-12
- Magoč, T., and Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507
- Maurice, C. F., Bouvier, C., De Wit, R., and Bouvier, T. (2013). Linking the lytic and lysogenic bacteriophage cycles to environmental conditions, host physiology and their variability in coastal lagoons. *Environ. Microbiol.* 15, 2463–2475. doi: 10.1111/1462-2920.12120
- McNichol, J., Stryhanyuk, H., Seewald, J. S., Sievert, S. M., Thomas, F., McNichol, J., et al. (2018). Primary productivity below the seafloor at deep-sea hot springs. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6756–6761. doi: 10.1073/pnas.1804351115
- McNichol, J., Sylva, S. P., Thomas, F., Taylor, C. D., Sievert, S. M., and Seewald, J. S. (2016). Assessing microbial processes in deep-sea hydrothermal systems by incubation at in situ temperature and pressure. *Deep. Res. Part I Oceanogr. Res. Pap.* 115, 221–232. doi: 10.1016/j.dsr.2016.06.011
- Nakagawa, S., and Takai, K. (2008). Deep-sea vent chemoautotrophs: diversity, biochemistry and ecological significance. *FEMS Microbiol. Ecol.* 65, 1–14. doi: 10.1111/j.1574-6941.2008.00502.x
- 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. doi: 10.1111/j.1462-2920.2005.00856.x
- Nakagawa, S., Takai, Y., Shimamura, S., Reysenbach, A.-L., Takai, K., and Horikoshi, K. (2007). Deep-sea vent epsilon-proteobacterial genomes provide insights into emergence of pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12146–12150. doi: 10.1073/pnas.0700687104
- Ochman, H., and Wilson, A. C. (1987). Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* 26, 74–86. doi: 10.1007/BF02111283

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.01262/full#supplementary-material>

- Palzkill, T. (2013). Metallo- β -lactamase structure and function. *Ann. N.Y. Acad. Sci.* 1277, 91–104. doi: 10.1111/j.1749-6632.2012.06796.x
- Paul, J. H. (2008). Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J.* 2, 579–589. doi: 10.1038/ismej.2008.35
- Pei, J., Kim, B. H., and Grishin, N. V. (2008). PROMALS3D: A tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* 36, 2295–2300. doi: 10.1093/nar/gkn072
- Pruesse, E., Peplies, J., and Glöckner, F. O. (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28, 1823–1829. doi: 10.1093/bioinformatics/bts252
- Rathgeber, C., Yurkova, N., Stackebrandt, E., Beatty, J. T., and Yurkov, V. (2002). Isolation of tellurite- and selenite-resistant bacteria from hydrothermal vents of the Juan de Fuca ridge in the Pacific Ocean: isolation of tellurite- and selenite-resistant bacteria from hydrothermal vents of the Juan de Fuca ridge in the Pacific Ocean. *Appl. Environ. Microbiol.* 68, 4613–4622. doi: 10.1128/AEM.68.9.4613
- Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106
- Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pasić, L., Thingstad, T. F., Rohwer, F., et al. (2009). Explaining microbial population genomics through phage predation. *Nat. Rev. Microbiol.* 7, 828–836. doi: 10.1038/nrmicro2235
- Roux, S., Enault, F., Hurwitz, B. L., and Sullivan, M. B. (2015). VirSorter: mining viral signal from microbial genomic data. *PeerJ* 3:e985. doi: 10.7717/peerj.985
- Schrenk, M. O., Kelley, D. S., Delaney, J. R., and Baross, J. A. (2003). Incidence and diversity of microorganisms within the walls of an active deep-sea sulfide chimney. *Appl. Environ. Microbiol.* 69, 3580–3592. doi: 10.1128/AEM.69.6.3580-3592.2003
- Seewald, J. S., Doherty, K. W., Hammar, T. R., and Liberatore, S. P. (2001). A new gas-tight isobaric sampler for hydrothermal fluids. *Deep Sea Res. Part I: Oceanogr. Res. Pap.* 49, 189–196. doi: 10.1016/s0967-0637(01)00046-2
- Sievert, S., and Vetriani, C. (2012). Chemoautotrophy at deep-sea vents: past, present, and future. *Oceanography* 25, 218–233. doi: 10.5670/oceanog.2012.21
- Stepanauskas, R., Fergusson, E. A., Joseph, B., Poulton, N. J., Tupper, B., Labonté, J. M., et al. (2017). Improved genome recovery and integrated cell-size analyses of individual, uncultured microbial cells and viral particles. *Nat. Commun.* 8:84. doi: 10.1038/s41467-017-00128-z
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039
- Vetriani, C., Jannasch, H. W., MacGregor, B. J., Stahl, D. A., and Reysenbach, A.-L. (1999). Population structure and phylogenetic characterization of marine benthic Archaea in Deep-Sea sediments population structure. *Appl. Environ. Microbiol.* 65, 4375–4384.
- Vetriani, C., Voordeckers, J. W., Crespo-Medina, M., O'Brien, C. E., Giovannelli, D., and Lutz, R. A. (2014). Deep-sea hydrothermal vent Epsilon proteobacteria encode a conserved and widespread nitrate reduction pathway (Nap). *ISME J.* 8, 1510–1521. doi: 10.1038/ismej.2013.246
- Von Damm, K. L., Edmond, J. M., Grant, B., Measures, C. I., Walden, B., and Weiss, R. F. (1985). Chemistry of submarine hydrothermal solutions at 21 °N. East Pacific Rise. *Geochim. Cosmochim. Acta* 49, 2197–2220. doi: 10.1016/0016-7037(85)90222-4
- Waite, D. W., Vanwonterghem, I., Rinke, C., Parks, D. H., Zhang, Y., Takai, K., et al. (2017). Comparative genomic analysis of the class Epsilonproteobacteria and proposed reclassification to epsilonbacteraeota (phyl. nov.). *Front. Microbiol.* 8:682. doi: 10.3389/fmicb.2017.00682
- Waite, D. W., Vanwonterghem, I., Rinke, C., Parks, D. H., Zhang, Y., Takai, K., et al. (2018). Addendum: comparative genomic analysis of the class Epsilonproteobacteria and proposed reclassification to epsilonbacteraeota (phyl. nov.). *Front. Microbiol.* 9:772. doi: 10.3389/fmicb.2018.00772
- Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28, 127–181. doi: 10.1016/j.femsre.2003.08.001
- Weinbauer, M. G., Brettar, I., and Höfle, M. G. (2003). Lysogeny and virus-induced mortality of bacterioplankton in surface, deep, and anoxic marine waters. *Limnol. Oceanogr.* 48, 1457–1465. doi: 10.4319/lo.2003.48.4.1457
- Williamson, S. J., Cary, S. C., Williamson, K. E., Helton, R. R., Bench, S. R., Winget, D., et al. (2008). Lysogenic virus-host interactions predominate at deep-sea diffuse-flow hydrothermal vents. *ISME J.* 2, 1112–1121. doi: 10.1038/ismej.2008.73
- Wommack, K. E., and Colwell, R. R. (2000). Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114. doi: 10.1128/MMBR.64.1.69-114.2000
- Yamamoto, M., and Takai, K. (2011). Sulfur metabolisms in epsilon- and gamma-proteobacteria in deep-sea hydrothermal fields. *Front. Microbiol.* 2:192. doi: 10.3389/fmicb.2011.00192
- Yu, Z.-C., Chen, X.-L., Shen, Q.-T., Zhao, D.-L., Tang, B.-L., Su, H.-N., et al. (2014). Filamentous phages prevalent in *Pseudoalteromonas* spp. confer properties advantageous to host survival in Arctic sea ice. *ISME J.* 9, 1–11. doi: 10.1038/ismej.2014.185
- Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

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.

Copyright © 2019 Labonté, Pachiadaki, Fergusson, McNichol, Grosche, Gulmann, Vetriani, Sievert and Stepanauskas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.