

# Single cell genomics indicates horizontal gene transfer and viral infections in a deep subsurface Firmicutes population

Jessica M. Labonté<sup>1</sup>, Erin K. Field<sup>1</sup>, Maggie Lau<sup>2</sup>, Dylan Chivian<sup>3</sup>, Esta Van Heerden<sup>4</sup>, K. Eric Wommack<sup>5</sup>, Thomas L. Kieft<sup>6</sup>, Tullis C. Onstott<sup>2</sup> and Ramunas Stepanauskas<sup>1\*</sup>

<sup>1</sup> Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, USA, <sup>2</sup> Department of Geosciences, Princeton University, Princeton, NJ, USA, <sup>3</sup> Lawrence Berkeley National Laboratory, Berkeley, CA, USA, <sup>4</sup> Department of Microbial, Biochemical and Food Biotechnology, University of the Free State, Bloemfontein, South Africa, <sup>5</sup> Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA, <sup>6</sup> Department of Biology, New Mexico Institute of Mining and Technology, Socorro, NM, USA

## OPEN ACCESS

### Edited by:

Manuel Martinez Garcia,  
University of Alicante, Spain

### Reviewed by:

Jennifer F. Biddle,  
University of Delaware, USA  
Marina G. Kalyuzhanaya,  
San Diego State University, USA

### \*Correspondence:

Ramunas Stepanauskas,  
Bigelow Laboratory for Ocean Sciences, 60 Bigelow Drive, East Boothbay, ME 04544, USA  
rstepanauskas@bigelow.org

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

Received: 04 February 2015

Paper pending published:  
05 March 2015

Accepted: 08 April 2015

Published: 22 April 2015

### Citation:

Labonté JM, Field EK, Lau M, Chivian D, Van Heerden E, Wommack KE, Kieft TL, Onstott TC and Stepanauskas R (2015) 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

A major fraction of Earth's prokaryotic biomass dwells in the deep subsurface, where cellular abundances per volume of sample are lower, metabolism is slower, and generation times are longer than those in surface terrestrial and marine environments. How these conditions impact biotic interactions and evolutionary processes is largely unknown. Here we employed single cell genomics to analyze cell-to-cell genome content variability and signatures of horizontal gene transfer (HGT) and viral infections in five cells of *Candidatus Desulfurudis audaxviator*, which were collected from a 3 km-deep fracture water in the 2.9 Ga-old Witwatersrand Basin of South Africa. Between 0 and 32% of genes recovered from single cells were not present in the original, metagenomic assembly of *Desulfurudis*, which was obtained from a neighboring subsurface fracture. We found a transposable prophage, a retrorv, multiple clustered regularly interspaced short palindromic repeats (CRISPRs) and restriction-modification systems, and an unusually high frequency of transposases in the analyzed single cell genomes. This indicates that recombination, HGT and viral infections are prevalent evolutionary events in the studied population of microorganisms inhabiting a highly stable deep subsurface environment.

**Keywords:** single cell genomics, *Desulfurudis*, evolution, horizontal gene transfer (HGT), transposases, transposable phages, population genomics, terrestrial deep subsurface

## Introduction

Deep subsurface microorganisms constitute a significant fraction of the living biomass on our planet (Whitman et al., 1998; Kallmeyer et al., 2012), but our knowledge of these microorganisms remains very limited, due to the difficulties in accessing these environments without introducing microbial contamination and to our inability to isolate the indigenous microbiota through cultivation (Colwell and D'Hondt, 2013). Some deep subsurface fracture communities have been physically isolated from other habitable environments for up to tens of millions of years and are characterized by extreme energy limitation. Existing data suggest that deep subsurface microbial communities differ from those of surface environments

in taxonomic composition, scarce energy sources, and energy production affected by a complete lack of light (Stevens and McKinley, 1995; Lin et al., 2006). They often have extremely low metabolic rates and generation times that span hundreds to thousands of years (Phelps et al., 1994; Jørgensen and D'Hondt, 2006; Lin et al., 2006; Onstott et al., 2014).

The deep fracture water of the Kaapvaal Craton, South Africa is a mixture of ancient hydrothermal fluid and meteoric water (Lippmann et al., 2003; Onstott et al., 2006; Lippmann-Pipke et al., 2011) that contains abiogenic hydrocarbons and H<sub>2</sub>-rich gasses (Ward et al., 2004; Lollar et al., 2006, 2008, 2014) and harbors unusual microbial communities and metazoans (Takai et al., 2001a,b; Moser et al., 2005b; Borgonie et al., 2011). A composite genome assembly of an indigenous firmicute *Candidatus Desulforudis audaxviator* MP104C was obtained from metagenomic reads from fracture water collected from a 2.8 km-deep borehole in the Mponeng gold mine (Chivian et al., 2008). The gene content of MP104C revealed a chemotactic sulfate-reducing bacterium capable of fixing N<sub>2</sub> and CO<sub>2</sub> while utilizing the chemical products of water radiolysis as its ultimate energy source (Chivian et al., 2008). Some of its key metabolic features, such as N<sub>2</sub> fixation, cobalamin synthesis, and sulfite reduction appear to be products of ancient horizontal gene transfer (HGT) events from other bacteria and archaea. The presence of multiple transposases and CRISPRs in the MP104C genome provided evidence of HGT and interactions with phages. Surprisingly, >99.9% of the metagenomic reads from the Mponeng sample mapped to the composite genome assembly, and only 32 positions in the 2.35 Mbp genome contained single nucleotide polymorphisms. This indicated a unique, one-species ecosystem with one of the lowest reported rates of mutation. Ribosomal RNA sequences that are near-identical to MP104C have been detected in many subsurface sites outside the Witwatersrand Basin, including in the Fennoscandian Shield in Finland (Itävaara et al., 2011), a deep saline geothermal aquifer in Germany (Lerm et al., 2013), and the Juan de Fuca crustal basalt fluids (Jungbluth et al., 2013). This suggests that *D. audaxviator* is highly adapted to conditions of the immense deep subsurface environment and has a cosmopolitan distribution. However, to the best of our knowledge, 16S rRNA sequences have remained the only type of information about *Desulforudis* from locations outside the Mponeng fracture, leaving its genetic diversity and associated evolutionary and ecological implications largely unknown.

Single cell genomics (SCG) is a novel technology offering recovery of genomic information from individual, uncultivated cells (Stepanauskas, 2012). SCG provides quantitative information of genomic variability in natural microbial communities, allowing for the study of gene exchange among cells and genome rearrangements within a cell. Such information is hard to obtain using other methods, such as metagenomics, where genome assemblies are consensus from a multitude of cells that are assumed to be clonal (Dupont et al., 2012; Iverson et al., 2012; Narasingarao et al., 2012).

To better understand the evolutionary mechanisms of natural *Desulforudis* populations, we performed genomic sequencing of individual cells collected from a 3.14 km-deep borehole in the Tau Tona mine, which is located near the site where *Ca.*

*D. audaxviator* MP104C was obtained. We sequenced five single amplified genomes (SAGs) carrying SSU rRNA gene sequences with >99.5% identity to MP104C. Our results show that, despite a highly stable environment and extremely low cell abundance, microbial populations still engage in horizontal gene transfer, a process usually believed to be a strategy for fast adaptations to changing environmental conditions.

## Materials and Methods

### Field Sample Collection

For single cell genomics, water samples were collected from the borehole DPH5057-TT109-Bh1 in the Tau Tona gold mine, located in the north margin of the Witwatersrand Basin, near Carletonville, South Africa, at 3.14 km depth below the surface. This borehole penetrated ~25 m ahead of a tunnel that was advancing through the seismically-active, Pretorius Fault Zone (Heesakkers et al., 2011) when it struck fracture water. The low microbial abundance (8000 cells mL<sup>-1</sup>), elevated pH (8–9) and temperature (48–49°C), and high reductive potential (340–370 mV) in the collected Tau Tona fracture water were typical of the deep subsurface of the Kaapvaal Craton (Magnabosco et al., 2014); while the <sup>14</sup>C content of the DIC suggested a subsurface residence time of ~21 kyr (Supplementary Table 1).

An acid-washed, autoclaved stainless steel manifold with multiple sampling ports was connected to the valve that had been installed on the borehole by the mine. Sampling ports had a valve to adjust water flow rate to meet various sampling requirements. The port was opened to flush out the water for ~5 min before any sampling activities were undertaken. Unfiltered borehole water was collected in a 50 mL Falcon tube and transported to the laboratory on ice. Within 6 h, 1 mL aliquots were transferred to cryovials containing glyTE cryoprotectant (5% glycerol and 1x TE buffer pH 8.0, final concentrations), mixed gently, and kept frozen at –80°C until processing.

### Single Cell Sorting, Whole Genome Amplification, Sequencing, and Assembly

Single cell sorting, whole-genome amplification, PCR-amplification and sequencing of the small subunit ribosomal RNA (SSU rRNA) genes, as well as the shotgun sequencing and *de novo* assembly of the selected SAGs were performed at the Bigelow Laboratory Single Cell Genomics Center (scgc.bigelow.org), as described previously (Stepanauskas and Sieracki, 2007; Swan et al., 2011; Martinez-Garcia et al., 2012; Field et al., 2015). PCR-amplified SSU rRNA gene sequences (~800–900 bp) of SAGs were edited using Sequencher v4.7 (Gene Codes) and compared with previously deposited sequences using the RDP v10 Classifier (SSU rRNA) (Wang et al., 2007) and National Center for Biotechnology Information (NCBI) BLAST (Altschul et al., 1990) nucleotide database (nt). 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, ≥1200 bp of good quality) from the RDP pipeline (Cole et al., 2014). 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)], as determined with jModelTest 2 (Darriba et al., 2012). Five SAGs having SSU rRNA genes with more than 99.5% identity to that of *Ca. D. audaxviator* were selected for genomic sequencing.

### Comparison of Genomes and Identification of Unique SAG Genes

Average nucleotide identity (ANI) was calculated using JSpecies (Richter and Rosselló-Móra, 2009) with the ANIb parameters. Comparisons of putative proteins from SAG genes were performed with MUMmer using the promer algorithm (Delcher et al., 2002). The SAGs were compared to each other and to the metagenomic assembly *Ca. D. audaxviator* MP104C (GenBank accession number NC\_010424). Contigs that were similar to the MP104C genome but contained gaps were manually inspected and aligned using the Geneious Aligner (Biomatters, Auckland, New Zealand) with a cost matrix of 70% similarity (5.0/-4.0) to identify the specific regions of putative horizontal gene transfer. Gene annotations from the Integrated Microbial Genomes (IMG) (Markowitz et al., 2009, 2012) pipeline (<http://img.jgi.doe.gov>) were used to identify transferred genes and possible modes of horizontal gene transfer (tRNA recombination, transposition and homologous recombination).

### Functional Classification of Proteins

The predicted genes were translated and used to search against the NCBI non-redundant database (nr). The protein functions were classified with the COG analyzer (Tatusov et al., 2000) in MEGAN5 (Huson et al., 2011), where each “gene” was considered a “read.” Transposases were identified using the IMG annotations, and classified with the ACLAME database (Leplae et al., 2004).

### Metagenomic Fragment Recruitment

Previously published metagenomes from boreholes in Tau Tona (Lau et al., 2014; Magnabosco et al., 2014), Mponeng (Chivian et al., 2008), and Masimong (Lau et al., 2014; Magnabosco et al., 2014) gold mines were used to look at the geographic distribution of *Desulforudis* (Supplementary Table 2). BLAST+ v2.2.28 (Camacho et al., 2009) was used to recruit metagenome reads to each SAG assembly using BLASTn with default parameter values, except for the following: -eval 0.0001 -soft\_masking true -lcase\_masking -xdrop\_gap 150. Fragment recruitment results, along with whole genome comparisons were displayed using Circos v0.66 (Krzywinski et al., 2009).

### Calculation of HGT Rate

Protein sequences from each sequenced SAG were compared with the NCBI nr database using BLASTp with an *e*-value cut off of 0.001, and the results were searched for “paradoxical” best hits, i.e., best hits to homologs not in *Ca. D. audaxviator* MP104C, as in Koonin (2001). All automatically detected “paradoxical” best hits were manually checked to eliminate possible false positives.

### Phylogenetic Analysis of the Virion Morphogenesis Protein

Analysis of the gene content of the SAGs revealed a transposable phage in SAG AC-310-N13. Sequences similar to the virion morphogenesis protein of the phage were obtained via BLASTx against the NCBI non-redundant database (nr). Protein sequences were aligned using MUSCLE (Edgar, 2004), and the alignment was manually edited in Geneious (Kearse et al., 2012). Phylogeny was performed using phyML (Guindon et al., 2010) implemented in Geneious 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). Trees were viewed with FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

### Accession Numbers

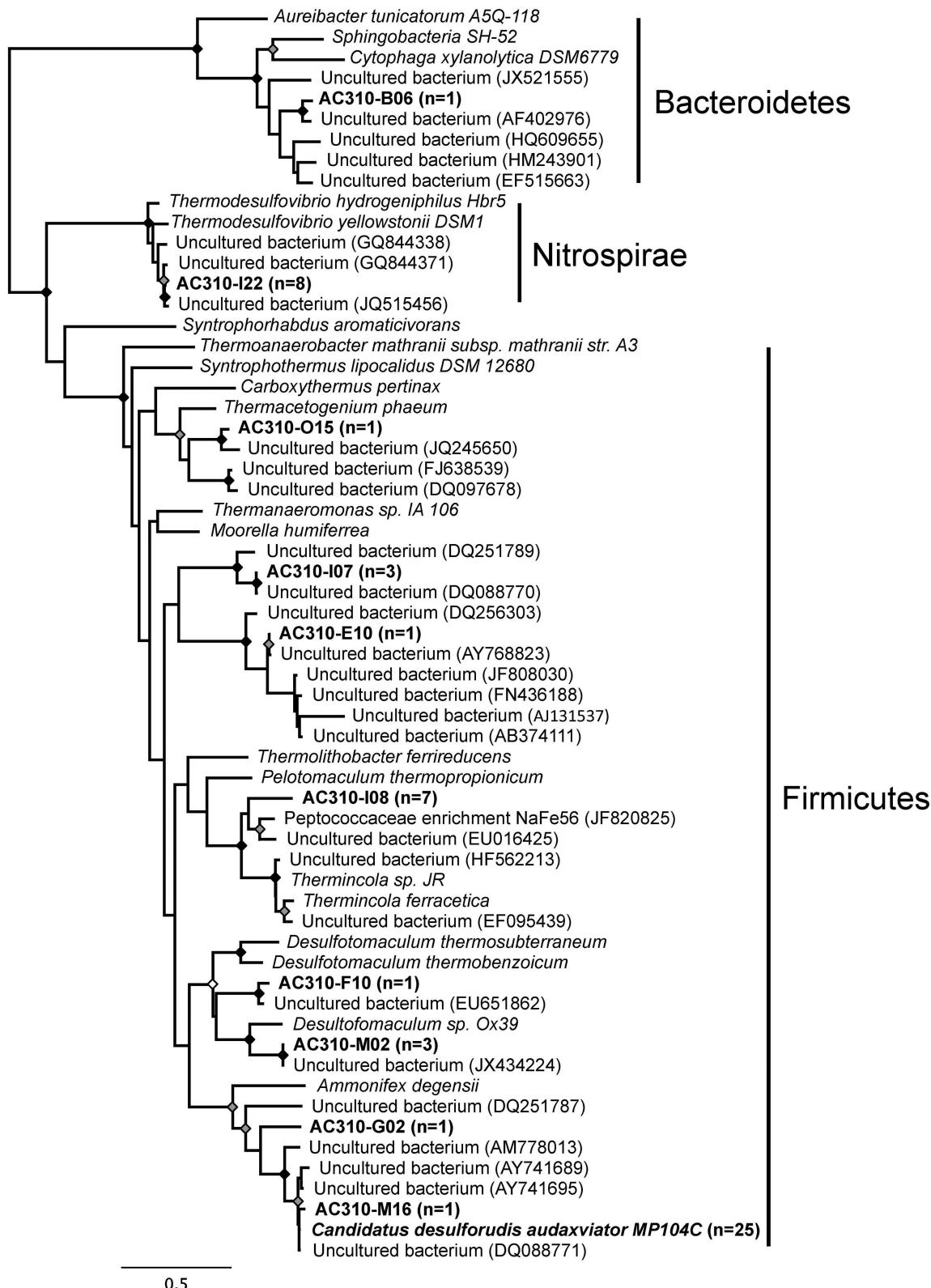
The annotated genomes are available in the Integrated Microbial Genomes database (<https://img.jgi.doe.gov>) under the IMG Taxon ID numbers 2596583550 (AC-310-A06), 2596583551 (AC-310-E02), 2600254919 (AC-310-N13), 2596583553 (AC-310-O10), and 2596583554 (AC-310-P15). The unassembled reads of single amplified genomes are available through the SCGC public data portal, <http://data.bigelow.org/~scgc/>.

## Results and Discussion

### Generation, Identification, and Genomic Sequencing of Single Amplified Genomes (SAGs)

Small subunit rRNA gene sequences were obtained from 52 of the 315 (17%) multiple displacement amplification reactions containing individual cells. This percentage was slightly lower than 19–39% obtained from marine (Swan et al., 2011) and freshwater (Martinez-Garcia et al., 2012) bacterioplankton using the same techniques. The retrieved SSU rRNA gene sequences revealed SAGs belonging to Firmicutes (43 SAGs), Nitrospirae (8 SAGs) and Bacteroidetes (1 SAG) (Figure 1). The SSU rRNA genes formed 10 clusters with >99% sequence identity to each other and to sequences that were previously retrieved from diverse subsurface environments (Figure 1). Of the 52 identified SAGs, 25 contained SSU rRNA genes that were 99.5–100% identical to that of *Ca. D. audaxviator* MP104C. This shows that close relatives of *Desulforudis* inhabit multiple subsurface environments and can coexist with other microbial species, in support of prior, metagenomic evidence (Moser et al., 2003, 2005a; Lin et al., 2006).

Of the five *Desulforudis* SAGs that were genetically sequenced, AC-310-E02 and AC-310-O10 had SSU rRNA genes that were 100% identical to MP104C, while AC-310-A06, AC-310-N13, and AC-310-P15 had SSU rRNA genes that were 99.5% identical to MP104C. Between 66 and 129 kbp of DNA sequence was recovered from each SAG in contigs larger than 2 kbp, corresponding to 4–8% of the expected genome size (Table 1). This SAG genome recovery was significantly lower ( $p < 0.001$ , *t*-test) than the 55% average genome recovery reported in a recent single cell genomics study of surface ocean bacterioplankton, which employed comparable laboratory and computational procedures (Swan et al., 2013). The unusually low genome recovery



**FIGURE 1 | Maximum likelihood phylogeny of the SSU rRNA gene sequences from SAGs (in bold) and closest relatives in GenBank.** The number of single cells with identical SSU rRNA gene sequences is indicated in parentheses. The tree was generated using maximum likelihood, with 100

bootstrap replicates, using the GTR model with a gamma distribution (+G), estimated rates of variation among sites and a proportion of invariable sites (+I). Bootstrap replicates 100%, ≥95%, and ≥85% are shown by black, gray, and white dots, respectively, at the nodes.

**TABLE 1 | Genomic sequence recovery from single cells of *Desulforudis*.**

Genome	# raw reads	Assembly length (bp)	Largest contig (bp)	# contigs	GC content (%)	# genes	Genome recovery (%)
AC-310-A06	9,248,293	66,621	18,148	10	54.4	84	4.1
AC-310-E02	9,194,185	129,196	16,829	16	56.9	148	7.8
AC-310-N13	14,367,041	78,324	12,465	14	57.3	105	3.6
AC-310-O10	8,678,532	79,711	20,659	11	55.0	105	6.4
AC-310-P15	8,362,251	107,785	24,300	11	56.0	119	4.7

Only contigs longer than 2000 bp were included in this analysis. The % genome recovery estimate is based on the assumption that genomes of all *Desulforudis* have the same length as MP104C, i.e., 2,349,476 bp (Chivian et al., 2008). NA, Not applicable.

from *Desulforudis* SAGs may have been caused by multiple factors, such as incomplete cell lysis; chemical interactions of DNA with other molecules that prevented amplification; or inefficient multiple displacement amplification resulting from the high GC content (61%) of *D. audaxviator*. Another plausible explanation is that the analyzed *Desulforudis* cells were endospores or other resting stages that were difficult to lyse. Inside an endospore, DNA is protected by small acid-soluble proteins (Setlow, 1995) that may not be denatured with the alkaline lysis of the cells used in this study. Further support for this hypothesis is provided by the presence of sporulation genes in *Ca. D. audaxviator* MP104C (Chivian et al., 2008) and the unusually dim fluorescence of Tau Tona microbial cells with the SYTO-9 nucleic acid stain (Supplementary Figure 1), which is typical for hard-walled microbial resting stages (Setlow et al., 2002). We are not aware of studies of microbial sporulation in the Kaapvaal Craton or other terrestrial subsurface environments. However, about half of the microbial cells within deep marine sediments are reported to be in a dormant state (Lennon and Jones, 2011; Lomstein et al., 2012; Hoehler and Jørgensen, 2013). Some regions of MP104C appear to be preferentially recovered in SAGs (inner circles in Figure 2). A possible explanation is that DNA follows a specific folding pattern inside cells or endospores, with only some genome regions being accessible to MDA. Another plausible explanation is preferential recovery of low %GC regions in SAGs. The lower %GC of the SAG assemblies (54–56% GC) relative to the 61% GC of the MP104C genome is consistent with the latter hypothesis.

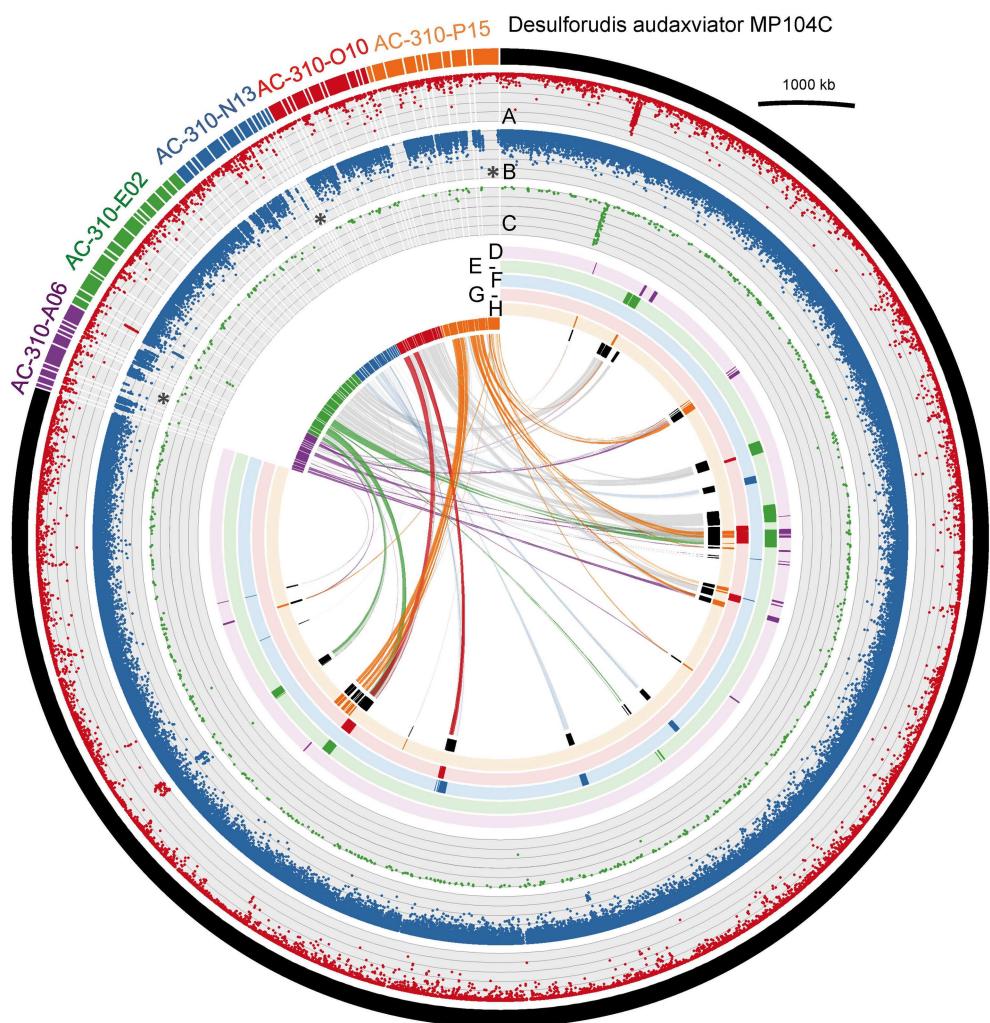
## Genome Content Variation among *Desulforudis* Cells

The average nucleotide identity (ANI) among the five sequenced SAGs and the MP104C genome ranged between 88.5 and 99.9% (Supplementary Table 3). This demonstrates a significant divergence of homologous genome regions among the analyzed *Desulforudis* cells, despite their 16S rRNA genes being ≥99.5% identical to each other. Interestingly, 10 SAG contigs (out of a total of 62) had no homology to MP104C (Supplementary Table 4). Another 22 SAG contigs contained both conserved and novel genome regions, when compared to MP104C (showcased as links in Figure 2, and Supplementary Figure 2). The significant recruitment of metagenomic fragments from Tau Tona but not Mponeeng and Masimong by some of the SAG-specific genome regions demonstrated their geographic endemism (asterisks in Figure 2 and Supplementary Table 5).

Between 0 and 32% of genes found in the sequenced SAGs were absent in the *Ca. D. audaxviator* MP104C genome (Supplementary Table 6), suggesting that they are products of horizontal gene transfer (HGT). In other taxa of Firmicutes, such as *Bacillus subtilis* and *B. halodurans*, the percentage of newly acquired genes has been estimated to ~9%, while rates of up to 20% have been observed in Spirochaetes (Koonin, 2001). The majority of SAG genes (80%) with no homology to *Ca. D. audaxviator* MP104C were similar to genes found in other Firmicutes, such as *Desulfotomaculum* sp. and *Pelotomaculum* sp. (Figure 3). Interestingly, many of these novel regions were in close proximity to transposases (9 cases), tRNA (2 cases), recombinases (2 cases), CRISPRs (1 case), and a prophage (1 case), all of which are indicative of recombination and HGT (Ochman et al., 2000; Gogarten et al., 2002; Brüssow et al., 2004; Marrafini and Sontheimer, 2010) (Supplementary Figure 2, and Supplementary Table 7). These results indicate a high prevalence of HGT events among autochthonous bacterial populations of the deep subsurface, despite their extremely slow metabolism and generation times that can be hundreds to thousands of years (Phelps et al., 1994; Jørgensen and D'Hondt, 2006; Lin et al., 2006).

The recovered regions of the five sequenced SAGs encoded for 20 transposases, representing 3.6% of all genes (Supplementary Table 8). The MP104C genome encodes 35 transposases, accounting for only 1.5% of all genes, which is still above the 0.83% average found in all sequenced microbial genomes and metagenomes (Aziz et al., 2010). This suggests that transposition may be significant in the evolution of *Desulforudis* natural populations. Transposases are among the most abundant genes in nature (Aziz et al., 2010) and are best known for their role in the HGT of antibiotic resistance genes (Scott, 2002; Whittle et al., 2002) and the relocation of regulatory elements (Shapiro, 2010). The only other study of transposase genes in a subsurface environment that we are aware of, in the Lost City deep hydrothermal system, also found an unusually high abundance of these genes, with more than 8% of metagenomic shotgun reads having a significant similarity to transposases (Brazelton and Baross, 2009).

To identify the genes that may have been acquired by transposition, we looked at neighborhoods of transposases on the SAG contigs and classified them based on their COG function (Supplementary Table 9). Most of the genes (57%) were not assigned or had no hits to the database. Intriguingly, most identified genes were involved in signal transduction (30%



**FIGURE 2 | Comparison of the sequenced SAGs AC-310-A06, AC-310-E02, AC-310-N13, AC-310-O10, and AC-310-P15 (colored outer circle segments) to the metagenomic assembly *Ca. Desulforudis audaxviator* MP104C (black outer circle segment).** Circles A–D show results of the metagenomic fragment recruitment from Tau Tona, Mponeng and Masimong fracture samples, respectively. The six colored inner circles indicate MP104C genes recovered from

each SAG. The inner-most circle (in black) indicates the presence of MP104C genes in any of the five SAGs. The links represent similarities based on MUMmer comparisons, using a 90% nucleic acid identity threshold. Asterisks indicate regions of low metagenomic fragment recruitment. Metagenomic fragment recruitment was performed with BLASTn, and hits with 75–100% DNA identity to references are displayed.

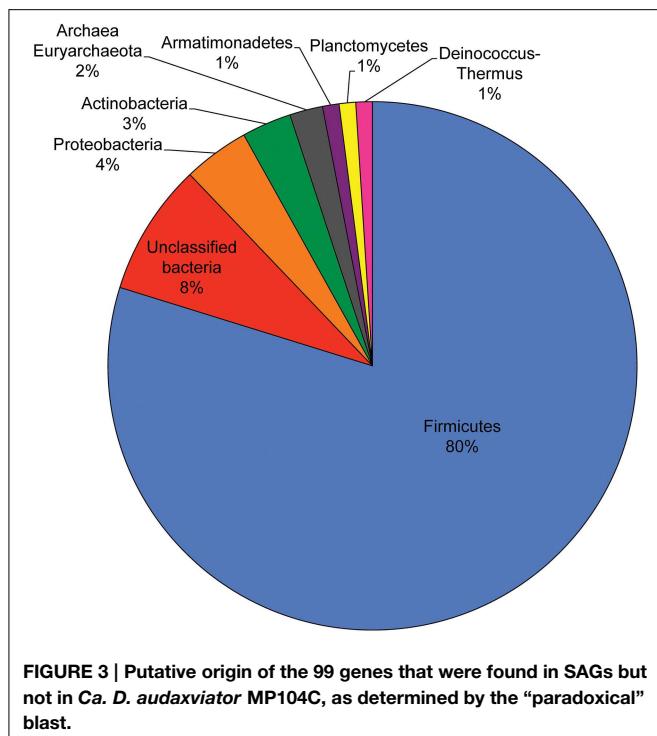
or the identified genes) and transcription/transcription regulation (11%) (Figure 4; Supplementary Table 9). This suggests that transposition in *Desulforudis* populations may be primarily involved in the evolution of regulation and communication rather than new metabolic capabilities.

The apparent prevalence of HGT in an environment with extremely low cell abundance and slow growth is puzzling. A plausible explanation may be provided by the possibility that biofilm-associated rather than free-living cells dominate subsurface microbiome (Taylor and Wirsén, 1999; Cozzarelli et al., 2007). Densely populated biofilms may provide microenvironments that enhance horizontal gene transfer (Molin and Tolker-Nielsen, 2003) and may serve as a source of cells and spores to the fracture water that was analyzed in this study. Further studies

are required to determine the presence and role of biofilms in the Witwatersrand Basin.

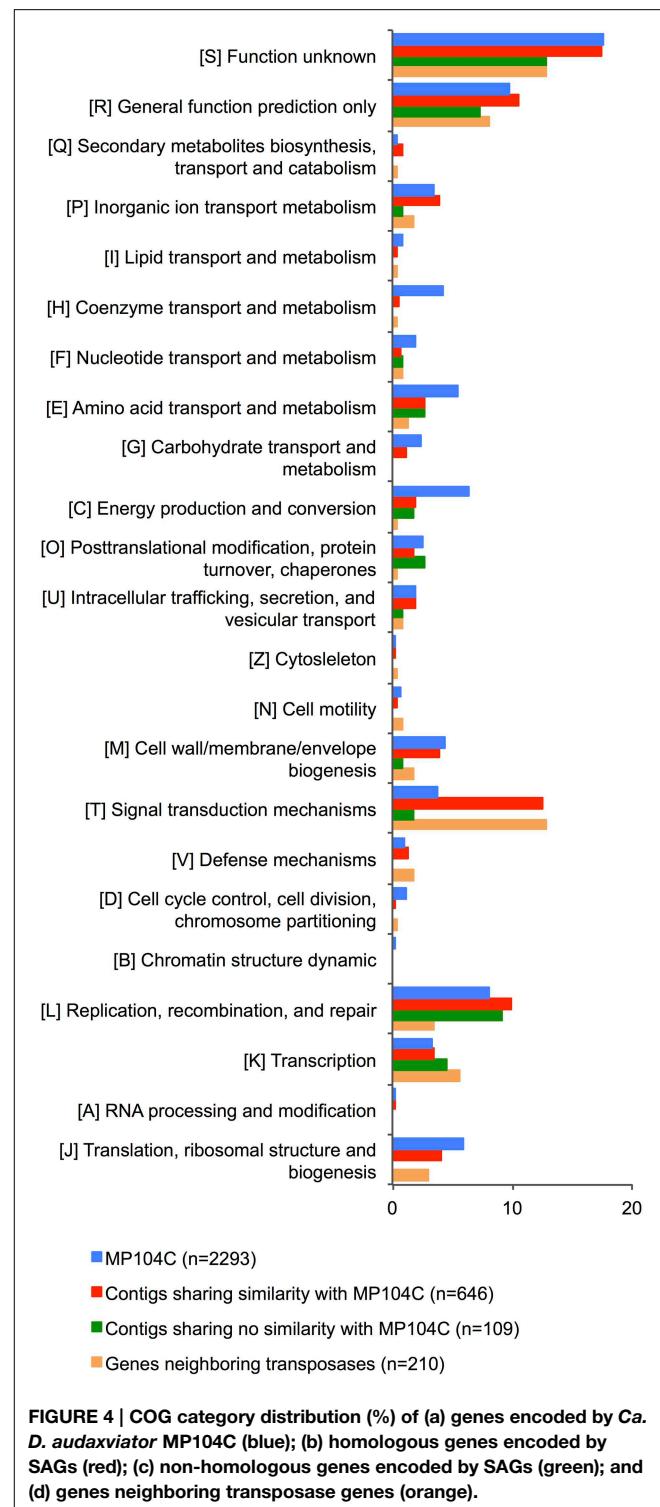
### Signatures of *Desulforudis* Interactions with Phages and Mobile Genetic Elements

Two contigs in AC-310-N13 (NODE\_2 and NODE\_13) contain phage-like sequences with no homology to MP104C (Supplementary Table 4). The presence of genes encoding a virion morphogenesis protein, a major capsid protein, and a terminase protein that are distantly related to Mu-like phages, and the co-occurrence of bacteria- and phage-like genes on NODE\_2 indicate a Mu-like transposable prophage (Figure 5). Transposable phages are believed to be temperate phages. They reproduce by transposition as part of their infectious cycle,



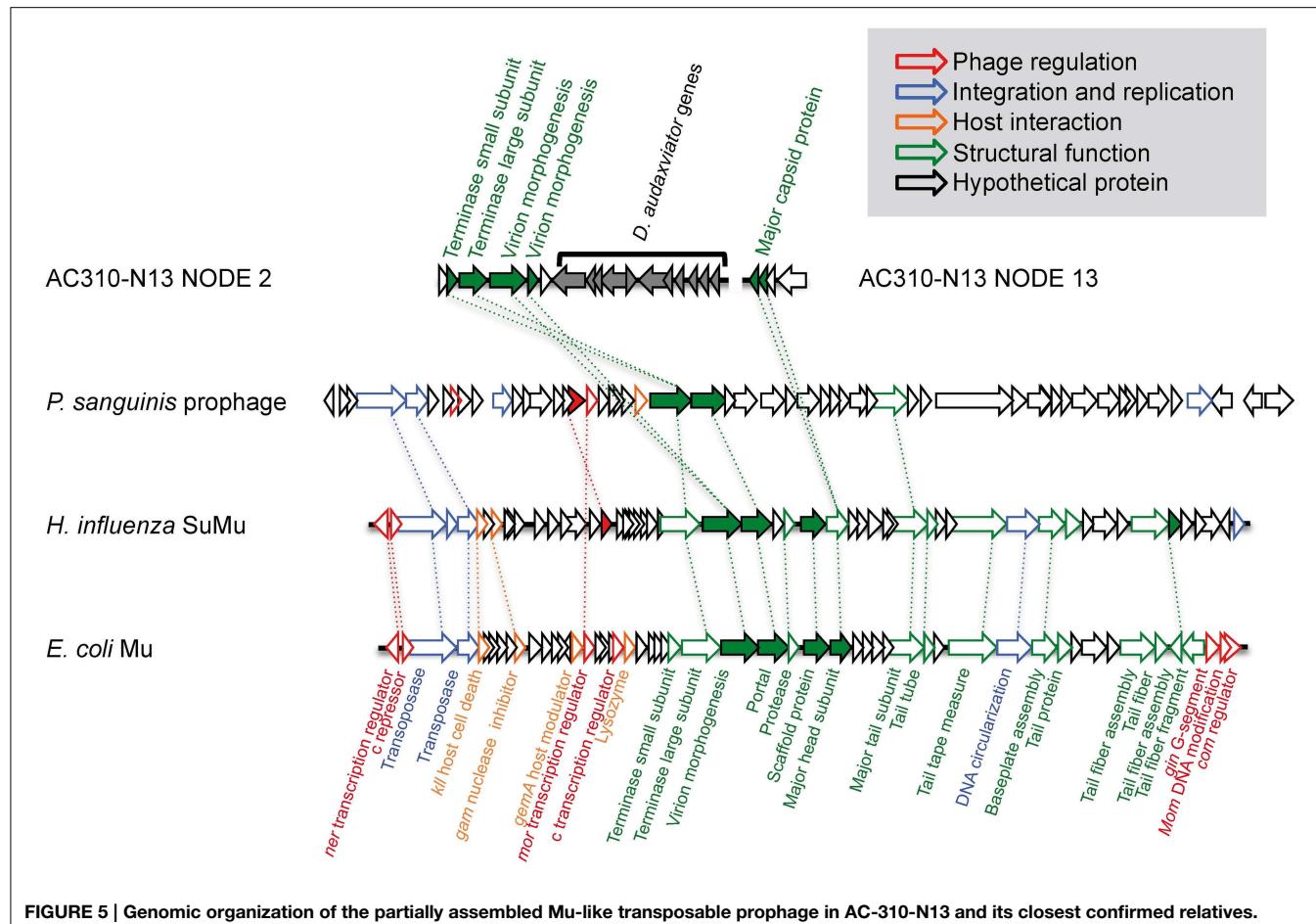
in this way facilitating HGT and impacting the evolution of their hosts (Wang et al., 2004). Very few transposable phages have been isolated and sequenced to date. However, a recent re-analysis of genomic databases has revealed Mu-like prophages within the genomes of several Firmicutes and Proteobacteria (Toussaint, 2013). The virion morphogenesis protein in AC-310-N13 is distantly related to prophages in other subsurface bacteria and non-subsurface Firmicutes, indicating co-evolution of these phages and their hosts (Supplementary Figure 3). The low nutrient concentrations observed in the deep subsurface are expected to be unfavorable for lytic phage infections (Fuhrman, 1999). Therefore, lysogeny may be the preferred lifestyle for deep subsurface phages. At least one prior study has indicated the prevalence of lysogeny in deep hydrothermal vents (Williamson et al., 2008). A selective advantage of lysogenic phage-host systems may be the immunity induced by them against other phages (Canchaya et al., 2003). Prophages sometimes increase the fitness of their host through lysogenic conversion or the transfer of useful genetic information (Brüssow et al., 2004). To the best of our knowledge, our study is the first to demonstrate the presence of transposable phages in the deep subsurface and their interactions with indigenous microorganisms.

We detected a CRISPR and associated *cas* genes in AC-310-E02 (Supplementary Figure 4). The same loci were also present in the MP104C genome, with some differences in the CRISPR spacers. Three additional CRISPR loci were also present in MP104C. However, none of the identified CRISPR spacers were homologous to the Mu-like viral sequences in AC-310-N13. Furthermore, genes encoding a type III restriction-modification system and a restriction endonuclease were found



**FIGURE 4 | COG category distribution (%) of (a) genes encoded by *Ca. D. audaxviator* MP104C (blue); (b) homologous genes encoded by SAGs (red); (c) non-homologous genes encoded by SAGs (green); and (d) genes neighboring transposase genes (orange).**

in AC-310-E02, while a type I restriction modification protein was found in AC-310-N13. The presence of these phage defense mechanisms provides further evidence for the importance of phage-host interactions in the ultra-deep fractures of the Kaapvaal Craton. While several prior studies report microscopic observations of viral particles in terrestrial subsurface



**FIGURE 5 |** Genomic organization of the partially assembled Mu-like transposable prophage in AC-310-N13 and its closest confirmed relatives.

(Kyle et al., 2008), the prevalence of lysogeny in hydrothermal systems (Williamson et al., 2008; Engelhardt et al., 2013), and the presence of prophage-like sequences in subsurface cultured isolates (Coombs, 2009), our study may be the first to identify transposable phage interactions with their hosts in indigenous, predominant microorganisms of the deep terrestrial subsurface.

Remarkably, one of the SAG contigs with no homology to MP104C, AC-310-P15 NODE\_5, carries a gene for a retron-type RNA-directed DNA polymerase. Retrons are small genetic elements that encode for a reverse transcriptase (RT) similar to the RT found in retroviruses and other retroelements. Retrons remain poorly characterized, but are found in many bacterial genomes and could play a role in the evolution and fitness of their host (Lampson et al., 2005; Simon and Zimmerly, 2008). To the best of our knowledge, this is the first record of retroids in deep subsurface environments.

## Conclusions

Our study confirms that close relatives (>99.5% identity of the 16S rRNA gene) of the firmicute *Candidatus Desulfuridus audaxviator* inhabit multiple deep subsurface fractures of the Kaapvaal Craton in South Africa. Unexpectedly, we found that

*D. audaxviator* coexists with other bacterial species and undergoes frequent horizontal gene transfer and viral infections. Our findings of a prophage, a retrorvirus, multiple CRISPRs and restriction-modification systems, and an unusually high frequency of transposases suggest that recombination, HGT and viral infections may play a role in the evolution of indigenous microorganisms in the deep subsurface, despite their extremely low cell abundance, slow metabolism and long generation times.

## Acknowledgments

We thank M. Pullin, S. Hendrickson, O. Kuloyo, J. G. Vermeulen, L. Snyder, L. Li, E. Cason, and P. Bennett for assistance in field sample collection. We also thank the management and staff of AngloGold Ashanti Tau Tona for their logistical support during sample collection. This study was made possible by the financial support of the Deep Carbon Observatory (to RS, TO, TK, and EV), the U.S. National Science Foundation (OCE-1136488 and OCE-1232982 to RS; EAR-0948659 to TO; EAR-0948335 and EAR-1141435 to TK), and the National Geographic Society (8990-11 to TO), the NASA EPSCoR/New Mexico Space Grant Consortium (to TK). Metagenomic sequencing was performed at the Marine Biological Laboratory and the

TIA Metagenomics Platform of the University of the Free State, and we are grateful for the assistance of Mitch Sogin, Susan Huse, Joseph Vineis, Andrew Voorhis, and Hilary Morrison. Finally, we thank Kathryn Elder and Sue Handwerk of the NOSAMS facility for processing our <sup>14</sup>C samples.

## References

- Altschul, S., Gish, W., and Miller, W. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Aziz, R. K., Breitbart, M., and Edwards, R. A. (2010). Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Res.* 38, 4207–4217. doi: 10.1093/nar/gkq140
- Borgonie, G., García-Moyano, A., Litthauer, D., Bert, W., Bester, A., van Heerden, E., et al. (2011). Nematoda from the terrestrial deep subsurface of South Africa. *Nature* 474, 79–82. doi: 10.1038/nature09974
- Brazelton, W. J., and Baross, J. A. (2009). Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. *ISME J.* 3, 1420–1424. doi: 10.1038/ismej.2009.79
- Brüssow, H., Canchaya, C., and Hardt, W. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68, 560–602. doi: 10.1128/MMBR.68.3.560-602.2004
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421
- Canchaya, C., Proux, C., Fournous, G., Bruttin, A., and Brüssow, H. (2003). Prophage genomics prophage genomics. *Microbiol. Mol. Biol. Rev.* 67, 238–276. doi: 10.1128/MMBR.67.2.238-276.2003
- Chivian, D., Brodie, E., Alm, E., Culley, D., Dehal, P., DeSantis, T., et al. (2008). Environmental genomics reveals a single-species ecosystem deep within Earth. *Science* 322, 275–278. doi: 10.1126/science.1155495
- 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
- Colwell, F. S., and D'Hondt, S. (2013). Nature and extent of the deep biosphere. *Rev. Mineral. Geochem.* 75, 547–574. doi: 10.2138/rmg.2013.75.17
- Coombs, J. M. (2009). Potential for horizontal gene transfer in microbial communities of the terrestrial subsurface. *Horiz. Gene Transf.* 532, 413–433. doi: 10.1007/978-1-60327-853-9\_24
- Cozzarelli, I. M., Weiss, J. V., Hurst, C. J., Crawford, R. L., Garland, J. L., Lipson, D. A., et al. (2007). “Biogeochemistry of aquifer systems,” in *Manual of Environmental Microbiology*, eds C. J. Hurst, R. L. Crawford, J. L. Garland, D. A. Lipson, A. L. Mills, and L. D. Stetzenbach (Washington, DC: ASM Press), 843–859.
- Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772. doi: 10.1038/nmeth.2109
- Delcher, A. L., Phillippy, A., Carlton, J., and Salzberg, S. L. (2002). Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res.* 30, 2478–2483. doi: 10.1093/nar/30.11.2478
- Dupont, C. L., Rusch, D. B., Yoosoph, S., Lombardo, M.-J., Richter, R. A., Valas, R., et al. (2012). Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J.* 6, 1186–1199. doi: 10.1038/ismej.2011.189
- 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
- Engelhardt, T., Sahlberg, M., Cypionka, H., and Engelen, B. (2013). Biogeography of Rhizobium radiobacter and distribution of associated temperate phages in deep seafloor sediments. *ISME J.* 7, 199–209. doi: 10.1038/ismej.2012.92
- Field, E. K., Sczyrba, A., Lyman, A. E., Harris, C. C., Woyke, T., Stepanauskas, R., et al. (2015). Genomic insights into the uncultivated marine Zetaproteobacteria at Loihi Seamount. *ISME J.* 9, 857–870. doi: 10.1038/ismej.2014.183
- Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* 399, 541–548. doi: 10.1038/21119
- Gogarten, J., Doolittle, W., and Lawrence, J. (2002). Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* 19, 2226–2238. doi: 10.1093/oxfordjournals.molbev.a004046
- 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
- Heesakkers, V., Murphy, S., and Reches, Z. (2011). Earthquake rupture at focal depth, Part I: structure and rupture of the pretorius fault, TauTona Mine, South Africa. *Pure Appl. Geophys.* 168, 2395–2425. doi: 10.1007/s00024-011-0354-7
- Hoehler, T. M., and Jørgensen, B. B. (2013). Microbial life under extreme energy limitation. *Nat. Rev. Microbiol.* 11, 83–94. doi: 10.1038/nrmicro2939
- Huson, D. H., Mitra, S., Ruscheweyh, H., Weber, N., and Schuster, S. C. (2011). Integrative analysis of environmental sequences using MEGAN4. *Genome. Res.* 21, 1552–1560. doi: 10.1101/gr.120618.111
- Itävaara, M., Nyssönen, M., Kapanen, A., Nousiainen, A., Ahonen, L., and Kukkonen, I. (2011). Characterization of bacterial diversity to a depth of 1500 m in the Outokumpu deep borehole, Fennoscandian Shield. *FEMS Microbiol. Ecol.* 77, 295–309. doi: 10.1111/j.1574-6941.2011.01111.x
- Iverson, V., Morris, R. M., Frazer, C. D., Berthiaume, C. T., Morales, R. L., and Armbrust, E. V. (2012). Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota. *Science* 335, 587–590. doi: 10.1126/science.1212665
- Jørgensen, B. B., and D'Hondt, S. (2006). Ecology. A starving majority deep beneath the seafloor. *Science* 314, 932–934. doi: 10.1126/science.1133796
- Jungbluth, S. P., Grote, J., Lin, H.-T., Cowen, J. P., and Rappé, M. S. (2013). Microbial diversity within basement fluids of the sediment-buried Juan de Fuca Ridge flank. *ISME J.* 7, 161–172. doi: 10.1038/ismej.2012.73
- Kallmeyer, J., Pockalny, R., Adhikari, R. R., Smith, D. C., and D'Hondt, S. (2012). Global distribution of microbial abundance and biomass in sub-seafloor sediment. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16213–16216. doi: 10.1073/pnas.1203849109
- 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
- Koonin, E. (2001). Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* 55, 709–742. doi: 10.1146/annurev.micro.55.1.709
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1345. doi: 10.1101/gr.092759.109
- Kyle, J. E., Eydal, H. S. C., Ferris, F. G., and Pedersen, K. (2008). Viruses in granitic groundwater from 69 to 450 m depth of the Aspö hard rock laboratory, Sweden. *ISME J.* 2, 571–574. doi: 10.1038/ismej.2008.18
- Lampson, B. C., Inouye, M., and Inouye, S. (2005). Retrons, msDNA, and the bacterial genome. *Cytogenet. Gen. Res.* 110, 491–499. doi: 10.1159/000084982
- Lau, M. C. Y., Cameron, C., Magnabosco, C., Brown, C. T., Schilkey, F., Grim, S., et al. (2014). Phylogeny and phylogeography of functional genes shared among seven terrestrial subsurface metagenomes reveal N-cycling and microbial evolutionary relationships. *Front. Microbiol.* 5, 1–17. doi: 10.3389/fmicb.2014.00531
- Lennon, J. T., and Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* 9, 119–130. doi: 10.1038/nrmicro2504
- Leplae, R., Hebrant, A., Wodak, S. J., and Toussaint, A. (2004). ACLAME: a CLASification of Mobile genetic Elements. *Nucleic Acids Res.* 32, D45–D49. doi: 10.1093/nar/gkh084

## Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.00349/abstract>

- Lerm, S., Westphal, A., Miethling-Graff, R., Alawi, M., Seibt, A., Wolfgramm, M., et al. (2013). Thermal effects on microbial composition and microbiologically induced corrosion and mineral precipitation affecting operation of a geothermal plant in a deep saline aquifer. *Extremophiles* 17, 311–327. doi: 10.1007/s00792-013-0518-8
- Lin, L.-H., Wang, P.-L., Rumble, D., Lippmann-Pipke, J., Boice, E., Pratt, L. M., et al. (2006). Long-term sustainability of a high-energy, low-diversity crustal biome. *Science* 314, 479–482. doi: 10.1126/science.1127376
- Lippmann, J., Stute, M., Torgersen, T., Moser, D. P., Hall, J. A., Lin, L., et al. (2003). Dating ultra-deep mine waters with noble gases and  $^{36}\text{Cl}$ , Witwatersrand Basin, South Africa. *Geochim. Cosmochim. Acta* 67, 4597–4619. doi: 10.1016/S0016-7037(03)00414-9
- Lippmann-Pipke, J., Erzinger, J., Zimmer, M., Kujawa, C., Boettcher, M., Heerden, E., et al. (2011). Geogas transport in fractured hard rock—correlations with mining seismicity at 3.54 km depth, TauTona gold mine, South Africa. *Appl. Geochem.* 26, 2134–2146. doi: 10.1016/j.apgeochem.2011.07.011
- Lollar, B. S., Lacrampe-Couloume, G., Slater, G. F., Ward, J., Moser, D. P., Gihring, T. M., et al. (2006). Unravelling abiogenic and biogenic sources of methane in the Earth's deep subsurface. *Chem. Geol.* 226, 328–339. doi: 10.1016/j.chemgeo.2005.09.027
- Lollar, B. S., Lacrampe-Couloume, G., Voglesonger, K., Onstott, T. C., Pratt, L. M., and Slater, G. F. (2008). Isotopic signatures of  $\text{CH}_4$  and higher hydrocarbon gases from Precambrian Shield sites: a model for abiogenic polymerization of hydrocarbons. *Geochim. Cosmochim. Acta* 72, 4778–4795. doi: 10.1016/j.gca.2008.07.004
- Lollar, B. S., Onstott, T. C., and Ballantine, C. J. (2014). The contribution of the Precambrian continental lithosphere to global  $\text{H}_2$  production. *Nature* 516, 379–382. doi: 10.1038/nature14017
- Lomstein, B. A., Langerhuus, A. T., D'Hondt, S., Jørgensen, B. B., and Spivack, A. J. (2012). Endospore abundance, microbial growth and necromass turnover in deep sub-seafloor sediment. *Nature* 484, 101–104. doi: 10.1038/nature10905
- Magnabosco, C., Tekere, M., Lau, M. C. Y., Linage, B., Kuloyo, O., Erasmus, M., et al. (2014). Comparisons of the composition and biogeographic distribution of the bacterial communities occupying South African thermal springs with those inhabiting deep subsurface fracture water. *Front. Microbiol.* 5:679. doi: 10.3389/fmicb.2014.00679
- Markowitz, V. M., Chen, I.-M. A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., et al. (2012). IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res.* 40, D115–D122. doi: 10.1093/nar/gkr1044
- Markowitz, V. M., Mavromatis, K., Ivanova, N. N., Chen, I.-M. A., Chu, K., and Kyrpides, N. C. (2009). IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 25, 2271–2278. doi: 10.1093/bioinformatics/btp393
- Marraffini, L. A., and Sontheimer, E. J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11, 181–190. doi: 10.1038/nrg2749
- Martinez-Garcia, M., Swan, B. K., Poulton, N. J., Gomez, M. L., Masland, D., Sieracki, M. E., et al. (2012). High-throughput single-cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. *ISME J.* 6, 113–123. doi: 10.1038/ismej.2011.84
- Molin, S., and Tolker-Nielsen, T. (2003). Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Curr. Opin. Biotechnol.* 14, 255–261. doi: 10.1016/S0958-1669(03)00036-3
- Moser, D. P., Gihring, T., Fredrickson, J. K., Brockman, F. J., Balkwill, D., Döllhopf, M. E., et al. (2005a). *Desulfotomaculum* spp. and *Methanobacterium* spp. dominate 4–5 km deep fault. *Appl. Environ. Microbiol.* 71, 8773–8783. doi: 10.1128/AEM.71.12.8773-8783.2005
- Moser, D. P., Gihring, T. M., Brockman, F. J., Fredrickson, J. K., Balkwill, D. L., Michael, E., et al. (2005b). *Desulfotomaculum* and *Methanobacterium* spp. dominate a 4- to 5-kilometer-deep fault. *Appl. Environ. Microbiol.* 71, 8773–8783. doi: 10.1128/AEM.71.12.8773
- Moser, D. P., Onstott, T. C., Fredrickson, J. K., Brockman, F. J., Balkwill, D. L., Drake, G. R., et al. (2003). Temporal shifts in the geochemistry and microbial community structure of an ultradepth mine borehole following isolation. *Geomicrobiol. J.* 20, 517–548. doi: 10.1080/713851170
- Narasingarao, P., Podell, S., Ugalde, J. A., Brochier-Armanet, C., Emerson, J. B., Brocks, J. J., et al. (2012). De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *ISME J.* 6, 81–93. doi: 10.1038/ismej.2011.78
- Ochman, H., Lawrence, J., and Groisman, E. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304. doi: 10.1038/35012500
- Onstott, T. C., Lin, L.-H., Davidson, M., Mislowack, B., Borcsik, M., Hall, J., et al. (2006). The origin and age of biogeochemical trends in deep fracture water of the Witwatersrand Basin, South Africa. *Geomicrobiol. J.* 23, 369–414. doi: 10.1080/01490450600875688
- Onstott, T. C., Magnabosco, C., Aubrey, A. D., Burton, A. S., Dworkin, J. P., Elsila, J. E., et al. (2014). Does aspartic acid racemization constrain the depth limit of the subsurface biosphere? *Geobiology* 12, 1–19. doi: 10.1111/gbi.12069
- Phelps, T., Murphy, E., Pfiffner, S., and White, D. (1994). Comparison between geochemical and biological estimates of subsurface microbial activities. *Microb. Ecol.* 335–349. doi: 10.1007/BF00662027
- 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
- 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
- Scott, K. (2002). The role of conjugative transposons in spreading antibiotic resistance between bacteria that inhabit the gastrointestinal tract. *Cell. Mol. Life Sci.* C 59, 2071–2082. doi: 10.1007/s000180200007
- Setlow, B., Loshon, C. A., Genest, P. C., Cowan, A. E., Setlow, C., and Setlow, P. (2002). Mechanisms of killing spores of *Bacillus subtilis* by acid, alkali and ethanol. *J. Appl. Microbiol.* 92, 362–375. doi: 10.1046/j.1365-2672.2002.01540.x
- Setlow, P. (1995). Mechanisms for the prevention of damage to DNA in spores of *Bacillus* species. *Annu. Rev. Microbiol.* 49, 29–54. doi: 10.1146/annurev.mi.49.100195.000333
- Shapiro, J. A. (2010). Mobile DNA and evolution in the 21st century. *Mob. DNA* 1:4. doi: 10.1186/1759-8753-1-4
- Simon, D. M., and Zimmerly, S. (2008). A diversity of uncharacterized reverse transcriptases in bacteria. *Nucleic Acids Res.* 36, 7219–7229. doi: 10.1093/nar/gkn867
- Stepanauskas, R. (2012). Single cell genomics: an individual look at microbes. *Curr. Opin. Microbiol.* 15, 613–620. doi: 10.1016/j.mib.2012.09.001
- Stepanauskas, R., and Sieracki, M. E. (2007). Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9052–9057. doi: 10.1073/pnas.0700496104
- Stevens, T. O., and McKinley, J. P. (1995). Lithoautotrophic microbial ecosystems in deep basalt aquifers. *Science* 270, 450–455. doi: 10.1126/science.270.5235.450
- Swan, B. K., Martinez-Garcia, M., Preston, C. M., Sczyrba, A., Woyke, T., Lamy, D., et al. (2011). Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* 333, 1296–1300. doi: 10.1126/science.1203690
- Swan, B. K., Tupper, B., Sczyrba, A., Lauro, F. M., Martinez-Garcia, M., Gonzalez, J. M., et al. (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11463–11468. doi: 10.1073/pnas.1304246110
- Takai, K. E. N., Moser, D. P., Flaun, M. D. E., Onstott, T. C., and Fredrickson, J. K. (2001a). Archaeal diversity in waters from deep South African gold mines. *Society* 67, 5750–5760. doi: 10.1128/AEM.67.21.5750-5760.2001
- Takai, K., Moser, D. P., Onstott, T. C., Spoelstra, N., Pfiffner, S. M., Dohnalkova, A., et al. (2001b). *Alkaliphilus transvaalensis* gen. nov., sp. nov., an extremely alkaliphilic bacterium isolated from a deep South African gold mine. *Int. J. Syst. Evol. Microbiol.* 51, 1245–1256. doi: 10.1099/00207713-51-4-1245
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33
- Taylor, C. D., and Wirsén, C. O. (1999). Rapid microbial production of filamentous sulfur mats at hydrothermal vents. *Appl. Environ. Microbiol.* 65, 2253–2255.
- Toussaint, A. (2013). Transposable Mu-like phages in Firmicutes: new instances of divergence generating retroelements. *Res. Microbiol.* 164, 281–287. doi: 10.1016/j.resmic.2013.01.008
- Wang, P. W., Chu, L., and Guttman, D. S. (2004). Complete sequence and evolutionary genomic analysis of the *pseudomonas aeruginosa* transposable bacteriophage D3112. *J. Bacteriol.* 186, 400–410. doi: 10.1128/JB.186.2.400-410.2004

- Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07
- Ward, J. A., Slater, G. F., Moser, D. P., Lin, L. H., Lacrampe-Couloume, G., Bonin, A. S., et al. (2004). Microbial hydrocarbon gases in the Witwatersrand Basin, South Africa: implications for the deep biosphere. *Geochim. Cosmochim. Acta* 68, 3239–3250. doi: 10.1016/j.gca.2004.02.020
- Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Perspective Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6578–6583.
- Whittle, G., Shoemaker, N., and Salyers, A. (2002). The role of *Bacteroides* conjugative transposons in the dissemination of antibiotic resistance genes. *Cell. Mol. Life Sci. C* 59, 2044–2054. doi: 10.1007/s000180200004
- 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

**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 © 2015 Labonté, Field, Lau, Chivian, Van Heerden, Wommack, Kieft, Onstott 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) or licensor 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.