Predominance and Metabolic Potential of *Halanaerobium* spp. in Produced Water from Hydraulically Fractured Marcellus Shale Wells

Daniel Lipus,^{a,b} Amit Vikram,^b Daniel Ross,^{a,c} Daniel Bain,^d Djuna Gulliver,^a Richard Hammack,^a Kyle Bibby^{b,e}

National Energy Technology Laboratory (NETL), Pittsburgh, Pennsylvania, USA^a; Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania, USA^b; AECOM, Pittsburgh, Pennsylvania, USA^c; Department of Geology and Environmental Science, University of Pittsburgh, Pittsburgh, Pennsylvania, USA^d; Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA^e

ABSTRACT Microbial activity in the produced water from hydraulically fractured oil and gas wells may potentially interfere with hydrocarbon production and cause damage to the well and surface infrastructure via corrosion, sulfide release, and fouling. In this study, we surveyed the microbial abundance and community structure of produced water sampled from 42 Marcellus Shale wells in southwestern Pennsylvania (well age ranged from 150 to 1,846 days) to better understand the microbial diversity of produced water. We sequenced the V4 region of the 16S rRNA gene to assess taxonomy and utilized quantitative PCR (qPCR) to evaluate the microbial abundance across all 42 produced water samples. Bacteria of the order *Halanaerobiales* were found to be the most abundant organisms in the majority of the produced water samples, emphasizing their previously suggested role in hydraulic fracturing-related microbial activity. Statistical analyses identified correlations between well age and biocide formulation and the microbial community, in particular, the relative abundance of *Halanaerobiales*. We further investigated the role of members of the order *Halanaerobiales* in produced water by reconstructing and annotating a *Halanaerobium* draft genome (named MDAL1), using shotgun metagenomic sequencing and metagenomic binning. The recovered draft genome was found to be closely related to the species *H. congolense*, an oil field isolate, and *Halanaerobium* sp. strain T82-1, also recovered from hydraulic fracturing produced water. Reconstruction of metabolic pathways revealed *Halanaerobium* sp. strain MDAL1 to have the potential for acid production, thiosulfate reduction, and biofilm formation, suggesting it to have the ability to contribute to corrosion, souring, and biofouling events in the hydraulic fracturing infrastructure.

IMPORTANCE There are an estimated 15,000 unconventional gas wells in the Marcellus Shale region, each generating up to 8,000 liters of hypersaline produced water per day throughout its lifetime (K. Gregory, R. Vidic, and D. Dzombak, *Elements* 7:181–186, 2011, <https://doi.org/10.2113/gselements.7.3.181>; J. Arthur, B. Bohm, and M. Layne, *Gulf Coast Assoc Geol Soc Trans* 59:49–59, 2009; <https://www.marcellusgas.org/index.php>). Microbial activity in produced waters could lead to issues with corrosion, fouling, and souring, potentially interfering with hydraulic fracturing operations. Previous studies have found microorganisms contributing to corrosion, fouling, and souring to be abundant across produced water samples from hydraulically fractured wells; however, these findings were based on a limited number of samples and well sites. In this study, we investigated the microbial community structure in produced water samples from 42 unconventional Marcellus Shale wells, confirming

Received 19 September 2016 Accepted 28 January 2017

Accepted manuscript posted online 3 February 2017

Citation Lipus D, Vikram A, Ross D, Bain D, Gulliver D, Hammack R, Bibby K. 2017. Predominance and metabolic potential of *Halanaerobium* spp. in produced water from hydraulically fractured Marcellus Shale wells. *Appl Environ Microbiol* 83:e02659-16. <https://doi.org/10.1128/AEM.02659-16>.

Editor Alfons J. M. Stams, Wageningen University

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to Kyle Bibby, bibbykj@pitt.edu.

the dominance of the genus *Halanaerobium* in produced water and its metabolic potential for acid and sulfide production and biofilm formation.

KEYWORDS Marcellus Shale, hydraulic fracturing, microbial ecology, produced water, corrosion, fouling, sulfide, *Halanaerobium*, metagenome

Oil and gas are now produced from previously unproductive (unconventional) hydrocarbon reservoirs due to the widespread use of horizontal drilling in conjunction with multistage, high-volume hydraulic fracturing. Hydraulic fracturing uses up to 25 million liters of water per well as the working fluid to fracture and increase the permeability of the hydrocarbon-containing formation (1, 2). Approximately 10% to 60% of the injected fluid returns to the surface after hydraulic fracturing as produced water and is characterized by total dissolved solids (TDS) with concentrations as high as 300,000 mg/liter (1, 2). Wells continue to generate produced water throughout their operational lifetimes, generating up to 8,000 liters of saline to hypersaline wastewater per day (1). Biological activity is generally undesirable during produced-water holding prior to reuse or disposal, as well as during well operation (3–5). Microorganisms in produced water may have the potential to produce acids and sulfides, leading to corrosion and gas souring (3, 6–9), and to form biofilms, resulting in clogging and fouling events (6, 9–11). Microbial activity is controlled by biocide addition during the fracture process and produced-water holding, and there is concern about the ecological impacts of biocides due to inadvertent release (12). Understanding the microbial ecology of produced water is a critical component of controlling undesirable microbial activity.

Previous studies have investigated the microbial ecology of Marcellus Shale produced water, reporting a rapid transition of the microbial community in produced water from an aerobic surface water microbial community to a fermentative, anaerobic community (6, 7, 10, 13, 14). Notably, bacteria of the genus *Halanaerobium* (order *Halanaerobiales*) have been shown to be predominant members of the produced water microbial community (6, 10, 13–15), representing a potential operational concern due to its fermentative (16–19), thiosulfate-reducing (15, 19–22) nature. Previous taxonomic characterization of microbial communities in produced water from the Marcellus Shale was performed on the basis of a sum total of six wells, with a maximum of three per study, all less than 18 months following fracture (6, 7, 10, 13, 14, 23, 81). Akob et al. evaluated 13 Pennsylvania shale gas wells (12 Marcellus Shale wells and 1 Burket Shale well) for the presence of anaerobic fermenters, methanogens, and H₂S-producing bacteria using culture-based enrichment tests, but they reported 16S rRNA gene taxonomy data only for the Burket sample (7). As there are more than 15,000 unconventional wells in the Marcellus Shale region alone (<https://www.marcellusgas.org/index.php>) and as produced water and well infrastructure are expected to be managed and maintained for 30 years or more, investigation of additional sites and well ages is necessary to confirm observed microbiological trends.

In addition, operational and geochemical factors, including the produced water salinity and the biocide composition used in the hydraulic fracturing fluid, may influence the microbial ecology of produced water. Salinity has been suggested to be a major factor controlling the bacterial community composition and activity in a variety of aquatic environments (24–26). Biocide composition likely affects the microbial ecology due to its wide variety of application approaches, as many different types and treatment combinations of biocides have been used with limited efficacy during hydraulic fracturing (27). No study has yet investigated the influence of these factors on produced water microbial community structure.

The objective of this study was to analyze the microbial abundance and community structure in produced water sampled from a greater number of wells, well sites, and well ages than previously considered. An additional goal was to identify possible correlations between the produced water microbial community and biocide composition, well age, and salinity as measured by levels of total dissolved solids (TDS).

Furthermore, this study aimed to specifically evaluate the abundance and metabolic potential of members of the order *Halanaerobiales*, which previous studies suggested to be among the most abundant produced water organisms. Ultimately, enhanced understanding of produced water microbial ecology will inform microbial monitoring and control efforts in produced water, leading to enhanced protection of well infrastructure.

RESULTS

Sample background and geochemistry. Produced water samples from 42 hydraulically fractured, horizontal Marcellus Shale gas wells, representing 18 well sites, in southwestern Pennsylvania in June 2014 were analyzed. In this study, the term “well site” refers to a single well pad. A well pad is a single site consisting of multiple wells extending laterally into the same formation. None of the wells sampled in this study had previously been remediated for fouling or souring issues. The production ages of analyzed wells ranged from 150 to 1,846 days, and TDS concentrations for the 42 samples ranged between 38,000 and 223,000 mg/liter (see Table S1 in the supplemental material). Sulfate concentrations were found to be below the detectable limit in 26 samples and low across the remaining samples, with a maximum concentration of 30 mg/liter. Further inorganic ion composition data are shown in Table S3.

Microbial abundance. Microbial abundance for each sample is reported in Table S3. Microbial abundance in the produced water samples as determined by quantitative PCR (qPCR) ranged between 1.5×10^5 and 2.1×10^8 16S rRNA gene copies per ml of produced water, within the range of previously reported values for produced water from the Marcellus Shale (13, 28). Correlation analysis based on Spearman rank coefficients and linear regression analysis suggested microbial abundance to be positively related to well age (Table S1, Table S4, and Table S5) and did not reveal any correlations between microbial abundance and TDS (Table S1, Table S4, and Table S5) and biocide treatment combination (see Fig. S2 in the supplemental material) (t test: all $P < 0.05$).

16S rRNA gene analysis and bacterial community structure determination. The members of the *Firmicutes* phylum, specifically, members of the orders *Halanaerobiales* and *Clostridiales*, were the dominant species across all produced water samples (Fig. 1). The *Halanaerobiales* order was identified in all produced water samples and was the most dominant order in 40 of the 42 samples (Fig. 1). For most *Halanaerobiales* sequences, we were not able to achieve classification below the family level using the RDP classifier in QIIME (Fig. S3). Within *Halanaerobiales*, sequences were affiliated with the genus *Halanaerobium* (up to 8% of all sequences in a single sample), unassigned *Halobacteroidaceae* (up to 21%), or unassigned *Halanaerobiaceae* (up to 99%) (Fig. S3). Annotation of 16S rRNA gene sequences using an alternative annotation strategy (MG-RAST) assigned the majority of unassigned *Halanaerobiaceae* reads to the genus *Halanaerobium*. *Clostridiales* spp. were also observed in all produced water samples and were the second most abundant group behind *Halanaerobiales* (Fig. 1). Abundant *Clostridiales* taxa included members of the families *Clostridiaceae* (up to 20% of sequences in a single sample), *Acidaminobacteraceae* (up to 21%), and *Lachnospiraceae* (up to 30%) (Fig. S3). We were not able to classify *Clostridiales* sequences below the family level. Other abundant (>5%) orders included *Pseudomonadales* identified in 41 samples (up to 25%), *Bacteroidales* identified in 40 samples (up to 32%), *Campylobacteriales* identified in 40 samples (up to 13%), *Bacillales* identified in 34 samples (up to 29%), *Oceanospirillales* identified in 32 samples (up to 14%), and *Desulfovibrionales* identified in 17 samples (up to 6%). Within the order *Pseudomonadales*, most sequences (up to 24% of all sequences in a single sample) were affiliated with the genus *Pseudomonas*. Within the order *Campylobacteriales*, most sequences (up to 8%) were affiliated with the genus *Arcobacter*, and most sequences (up to 10%) within the *Oceanospirillales* were assigned as unclassified *Halomonadaceae* (Fig. S3). Sequences identified as *Archaea* were detected in 17 samples as representing members of the orders *Methanosarcinales* (up to 0.5% of all sequences in a single sample) and *Metha-*

FIG 1 Relative abundances of microbial orders for 42 Marcellus Shale produced water samples analyzed using 16S rRNA gene sequencing in this study. Sites in this description refer to individual well pads. Orders that were less than 2% abundant in all samples were classified as minor orders and are summarized in Fig. S3.

nomicrobiales (up to 1.1%). All minor orders with abundances below 2% are summarized in Fig. S4.

We used Spearman rank coefficients and analysis of similarities (ANOSIM) to correlate the relative abundances of *Halanaerobiales* spp. with operational parameters (Table S6). Results identified an inverse correlation between the relative abundances of *Halanaerobiales* spp. and well age ($P < 0.001$). *Halanaerobiales* relative abundances were also plotted against well age and TDS concentration for linear regression analysis, revealing no trends (Fig. S5). We identified a correlation between biocide treatment combination and *Halanaerobiales* abundance (t test $P = 0.018$; ANOSIM $P = 0.008$), as samples from wells treated with biocide treatment combination 1 were associated with higher relative abundances of *Halanaerobiales*. Biocide treatment combination 1 was 2,2-dibromo-3-nitropropionamide (DBNPA) based, and biocide treatment combination 2 was tetrahydro-3,5-dimethyl-2H-1,3,5-thiadiazine-2-thione (Dazomet) based (Table S2). Both DBNPA and Dazomet are nonoxidizing biocides. Further mechanistic explanation for this observation will require additional analysis.

Alpha diversity. The produced water microbial community alpha diversity was calculated by determination of the number of operational taxonomic units (OTUs), the

Chao1 richness index (a measure of sample richness, i.e., number of OTUs), and the Shannon index (a measure of both sample richness and the distribution of OTUs). Alpha diversity parameters were assessed per 2,000 sequences to enable even cross-sample comparisons. All sequences were used for the three samples with less than 2,000 OTU-assigned sequences (Table S7). The number of observed OTUs ranged from 6 to 187; Chao1 richness analysis predicted richness values between 7 and 237; and the Shannon diversity ranged between 0.54 and 4.62 (Table S7). Correlations between alpha diversity parameters (number of OTUs, Chao1 index, and Shannon index) and well age, TDS, and microbial abundance were assessed using Spearman rank coefficients and linear regression analysis. Results suggested a moderate correlation between Shannon diversity and well age, but they did not reveal any other correlations (Table S8 and Fig. S6 to S9). Alpha diversity parameters were also plotted against biocide treatment combinations, and no clustering by treatment combination was observed (Fig. S10, *t* test: all $P > 0.05$).

Beta diversity. NMDS (nonmetric multidimensional scaling) ordination analysis did not reveal any clustering by well site or any correlations between well sites and operational parameters (Fig. S11). Weighted UniFrac distances were used to build a tree with the unweighted pair group method using average linkages (UPGMA), and branches were labeled by TDS concentration, well age, and composition of the applied biocide (Fig. S12). No clustering was observed for TDS concentration or well age (Fig. S12). Analysis of similarities (ANOSIM) did not suggest any correlations for TDS concentration or well age (all $P > 0.05$). We observed clustering by biocide treatment combination (Fig. S12 and Table S2), and ANOSIM revealed a difference in microbial community composition between the samples treated with biocide treatment combination 1 and those treated with biocide treatment combination 2 ($P = 0.004$).

Metagenome sequencing. Sample “site 13, well 2” was selected for shotgun metagenomic sequencing to further investigate the metabolic potential of members of the genus *Halanaerobium*, found to be abundant across all analyzed produced water samples. On the basis of 16S rRNA gene data, sample “site 13, well 2” was 99.1% *Halanaerobiales* and characterized by a low overall diversity (12 OTUs). Shotgun metagenome sequencing generated 6,089,871 trimmed, high-quality reads. Trimmed sequence reads were then mapped against available *Halanaerobium* reference genome sequences (Fig. 2A). The best mapping results were achieved for *Halanaerobium* sp. strain T82-1 (GenBank accession number [LSBN00000000.1](#)) (19), a draft genome recovered from produced water (68% mapped reads and 99% coverage), and for *Halanaerobium saccharolyticum* strain DSM 6643 (accession number [NZ_CAU100000000.1](#)) (50% mapped reads and 81% coverage) (Fig. 2A).

Sequence reads were then assembled into 446 contiguous sequences (“contigs”) using metaSPAdes. The minimum contig length was 5,000 bp, the maximum length was 96,535 bp, and the N50 was 15,595 bp. Taxonomy assignment performed with PhyloPythiaS using the 2013 generic model with default settings (29) identified 82% of contigs as belonging to the genus *Halanaerobium*. 16S rRNA gene prediction identified four 16S rRNA genes that were between 1,023 bp and 1,260 bp in length (30). Extraction and phylogenetic analysis of 16S rRNA genes showed the extracted genes to be identical to each other (average nucleotide identity [ANI], 100% across 1,023 bp) and closely related to the 16S rRNA gene of the species *Halanaerobium congolense* (ANI, 99% across 1,023 bp) (Fig. 2B). Phylogenetic analysis showed that the 16S rRNA genes did not cluster as closely with the only other available produced water *Halanaerobium* isolate (*Halanaerobium* sp. strain DL-01) (ANI, 98% across 1,023 bp) (Fig. 2B) (15). No 16S rRNA gene region is currently available or could be identified for the recently published produced water draft genome of *Halanaerobium* sp. strain T82-1 (19).

Metagenomic contig binning and annotation. Metagenomic binning resulted in one *Halanaerobium* draft genome, named *Halanaerobium* sp. strain MDAL1, containing 129 contigs with a total size of 2,389,586 bp and a GC content of 34.2%, consistent with previously sequenced *Halanaerobium* genomes with sizes between 2.3 and 2.9 million

FIG 2 (A) Mapping results for metagenomic reads against other available *Halanaerobium* genomes. (B) Unrooted phylogenetic tree showing the relationship between the 16S rRNA genes recovered from metagenomic contigs and the 16S rRNA genes of other *Halanaerobium* reference sequences downloaded from NCBI. Bar, 4 nucleotide substitutions per 1,000 nucleotides.

bp and GC contents between 30.3% and 33.3% (Table S9) (16, 31, 32). The *Halanaerobium* draft genome was the only bin obtained from the metagenome sequences. Contigs that could not be binned were affiliated with the genus *Acetohalobium* and with the class *Clostridia* or could not be classified. More than half (53.9%) of the original sample sequence reads were successfully mapped back to the *Halanaerobium* sp. strain MDAL1 draft genome. The *Halanaerobium* sp. strain MDAL1 draft genome bin was found to be 83% complete using 898 *Halanaerobium* marker genes in CheckM (33). Annotation identified 2,219 gene coding sequences (CDs) and 23 RNA sequences, representing 304 SEED subsystems for the recovered *Halanaerobium* draft genome. Phylogenomic analysis of *Halanaerobium* genomes and calculations of average nucleotide identities (ANI) and average amino acid identities (AAI) in comparison to other

TABLE 1 Average nucleotide identity and amino acid identity between recovered produced water *Halanaerobium* sp. strain MDAL1 draft genome and other available *Halanaerobium* genomes^a

Genome	GenBank accession no.	ANI (%)	AAI (%)
<i>Halanaerobium</i> sp. strain MDAL1 draft genome	MIJU00000000.1		
<i>Halanaerobium</i> sp. strain T82-1	LSBN00000000.1	98.48	93.82
<i>Halanaerobium saccharolyticum</i> subsp. <i>saccharolyticum</i> DSM 6643	NZ_CAUI00000000.1	84.62	82.04
<i>Halanaerobium praevalens</i> DSM 2228	CP002175.1	83.93	73.73
<i>Halanaerobium hydrogeniformans</i>	CP002304.1	82.65	67.59

^aANI, average nucleotide identity; AAI, amino acid identity.

available *Halanaerobium* genomes suggested the recovered *Halanaerobium* sp. strain MDAL1 draft genome to be closely related to that of *Halanaerobium* sp. strain T82-1 (ANI = 98.48%; AAI = 93.82%) (Table 1).

Of particular interest was the metabolic potential for fermentation pathways, sulfur metabolism, and biofilm formation, as acid production, sulfide production, and biofouling are undesirable and therefore of high interest to the hydraulic fracturing industry and broader oil and gas industry (1, 12). In addition, we evaluated the presence of genes involved in stress response mechanisms, as these processes have previously been shown to enable increased resistance to biocides in produced water (26).

Genes associated with mixed acid fermentation were identified in the *Halanaerobium* sp. strain MDAL1 draft genome and include *ldh*, which encodes a lactate dehydrogenase responsible for the conversion of pyruvate to lactate; *ptaA*, encoding a phosphate acetyltransferase that converts pyruvate to acetate; and *adh*, which encodes an alcohol dehydrogenase involved in the fermentation of simple sugars into ethanol (Table S10). Furthermore, a gene encoding pyruvate formate lyase Pfl involved in the transformation of pyruvate to hydrogen or carbon dioxide was identified (Table S10). BLASTx analysis confirmed the presence of the described fermentation genes in *Halanaerobium* sp. strain T82-1 (99% identity). The discovery of these genes allowed reconstruction of putative metabolic pathways for the conversion of pyruvate into the fermentation products lactate, acetate, ethanol, hydrogen, and carbon dioxide, confirming the potential of *Halanaerobium* to contribute to acid production in produced water.

The production of sulfides during hydraulic fracturing operations can lead to gas souring and is therefore a significant concern to hydraulic fracturing operations (3, 34). While we did not identify any classical sulfate reduction genes (e.g., *dsrAB*, *aps*), our analysis revealed several genes involved in thiosulfate reduction, a process previously reported to contribute to sulfide production and observed across various anaerobic, halophilic, and thermophilic bacterial taxa (22, 35). Identified genes included an unclassified rhodanese-like gene, a thiosulfate sulfurtransferase rhodanese gene, an unclassified sulfurtransferase rhodanese gene, and anaerobic sulfite reductase genes *asrA*, *asrB*, and *asrC* in the *Halanaerobium* sp. strain MDAL1 draft genome (Table S10). BLASTx analysis revealed the unclassified rhodanese-like gene to share 100% homology with the genes encoding the previously described thiosulfate reduction RdlA rhodanese proteins in *Halanaerobium* sp. strain T82-1, *H. congolense*, and *H. saccharolyticum* (19, 21, 35). Furthermore, a Trk-type sulfate permease and a putative ABC-type sulfate-like transporter were identified, allowing the reconstruction of a putative thiosulfate reduction pathway (Table S10). Thiosulfate is transported into the cell by a sulfate ABC-type transporter and converted into adenylyl sulfate by the Rdl rhodanese protein and into sulfide by a sulfurtransferase and the AsrABC complex. The utilization of thiosulfate and production of hydrogen sulfide has recently been reported for *Halanaerobium* DL-01 isolated from produced water (15). In the same study, researchers found *Halanaerobium* DL-01 to produce acetate when thiosulfate was used as an electron acceptor (15). Our results, together with observations from previous studies, suggest that *Halanaerobium* species identified in produced water have the ability to reduce thiosulfate and poten-

tially produce sulfides and organic acids. These characteristics have also been described for *Halanaerobium* species isolated from different environments (21, 31).

Biofilm formation during or after hydraulic fracturing can cause damage to the hydraulic fracturing infrastructure and lead to clogging, interfering with hydraulic fracturing operations (27, 36). Genes encoding proteins suggested to be involved in biofilm formation processes were identified and included genes encoding sporulation two-component response regulator Spo0A, associated with surface attachment initiation, and glycosyl transferase group 2 family protein Glt2 and diguanylate cyclase AdrA, which have both been associated with exopolysaccharide production (Table S10) (37, 38). Furthermore, we identified several flagellin and motility genes, including *fliC*, *flhA*, *motA*, and *motB*, which have been suggested to be important for cellular attachment and the initial stages of biofilm formation (39, 40).

Produced water contains elevated salinity levels and the presence of heavy metals, resulting in osmotic, oxidative, and periplasmic stress responses to protect the cell (41–43). Multiple genes associated with salinity tolerance were identified in the *Halanaerobium* sp. strain MDAL1 draft genome (Table S10). The genome encodes the Trk and Ktr transmembrane transporter complexes, including the potassium and sodium uptake proteins TrkA, TrkH, and KtrA (44–47). The presence of the *trkH* and *trkA* genes supports previously reported characteristics for *Halanaerobium* spp., which have been shown to use a “salt-in strategy” to counter osmotic stress through the uptake of potassium from the environment by utilizing membrane proteins of the Trk and Ktr complexes (20, 43, 45). In addition, glycine and betaine ABC transport protein ProX, L-proline glycine betaine ABC transport system permease protein ProW, and high-affinity betaine transport system OpuA were identified (Table S9) (48, 49). BLASTx analysis of *proW*, *proX*, and *opuU* gene sequences revealed close homology to genes of *Halanaerobium* sp. strain T82-1 (>99% identity). Other identified genes with a potentially important role in microbial stress response under produced water conditions included *perR*, a redox-sensitive transcriptional regulator; *sor*, encoding a superoxide reductase; and unidentified genes putatively encoding rubredoxin, glutaredoxin, and rubredoxin. These genes are potentially involved in oxidative stress response pathways, which have been suggested to be triggered by a secondary response to osmotic stress as well as to the high concentration of cations in produced water (50). In addition, sequences that included periplasmic stress response gene *ompH* (51) and genes encoding the universal stress protein UspA, the heat shock protein GrpE, Hsp20, and the heat shock chaperones GroES and GroEL were identified (Table S10). We also discovered gene sequences encoding chemotaxis-related proteins MotA and MotB and the flagellin assembly protein family Flg, Fli, and, Flh in the *Halanaerobium* sp. strain MDAL1 draft genome (Table S10).

DISCUSSION

This study evaluated microbial communities in 42 produced water samples taken from wells of various ages within the southwest Pennsylvania region of the Marcellus Shale and investigated the metabolic potential of members of the dominant produced water genus, *Halanaerobium*. The goal of this study was to improve the understanding of the microbial community associated with hydraulic fracturing operations by analyzing samples from the largest number and greatest diversity of wells to date. Furthermore, the recovery and annotation of a metagenome-assembled *Halanaerobium* sp. strain MDAL1 draft genome bin provided new insights into the metabolic potential of microbial populations in produced water, specifically highlighting acid and sulfide production. Ultimately, improved understanding of produced water microbial ecology will enable the development of better produced water management strategies to minimize corrosion, produced water souring, and fouling events and to encourage produced water reuse.

***Halanaerobiales* dominates Marcellus Shale produced water.** Microbial community structure analysis revealed bacteria of the phylum *Firmicutes* to be dominant across all produced water samples analyzed, constituting as much as 99% of the microbial

Study:	This Study 2016	Struchtemeyer 2012	Murali Mohan 2013	Strong 2013	Wuchter 2013	Cluff 2014
Region:	Marcellus Shale	Barnett Shale	Marcellus Shale	Bakken and Marcellus Shale	Antrim Shale	Marcellus Shale
Well ages:	150 - 1846 days	1 - 60 days	1 - 187 days	~540 days	~150 days	0 - 328 days
<i>Halanaerobiales</i>						
<i>Clostridiales</i>						
<i>Campylobacteriales</i>						
<i>Rhodobacterales</i>						
<i>Bacteroidales</i>						
<i>Bacillales</i>						
<i>Desulfotribionales</i>						
<i>Erysipelotrichales</i> *						
<i>Enterobacteriales</i>						
<i>Rhizobacteriales</i>						
<i>Oceanospirillales</i>						
<i>Lactobacillales</i>						
<i>Pseudomonadales</i>						
<i>Actinomycetales</i>						

*Observed in petroleum reservoir, formation waters and subsurface hot springs

¹Describes highest observed value

Abundance ¹ :			
< 1%		20% - 30%	
1% - 5%		30% - 40%	
5% - 10%		40% - 50%	
10% - 20%		> 50 %	

FIG 3 Relative abundances of identified Marcellus Shale produced water microbial orders identified in this study and across previous produced water studies. Reported values represent the highest observed relative abundance in any single sample. Citations of studies in the upper panel are as follows: Struchtemeyer 2012, reference 28; Murali Mohan 2013, 13; Strong 2013, 14; Wuchter 2013, 23; Cluff 2014, 10.

community. Within the phylum *Firmicutes*, the majority of the sequences were associated with the orders *Halanaerobiales* (up to 99.8% total relative abundance) and *Clostridiales* (up to 79.2% total relative abundance). *Halanaerobiales* are fermentative, obligate anaerobic halophiles that have previously been shown to be abundant in produced water following the flowback period in the Marcellus Shale (10, 13, 20) and the Antrim Shale (23). Members of the order *Halanaerobiales* have been isolated from conventional oil wells (21) and have been identified in a variety of other ecosystems such as the Dead Sea, marine salterns, cyanobacterial mats, and hypersaline lakes (20, 52). Phylogenetic analysis of recovered 16S rRNA genes suggested the recovered draft genome to be closely related to that of the genus *H. congolense* (21), which was isolated from an oil field and shares multiple functional characteristics with the *Halanaerobium* sp. strain MDAL1 draft genome.

Clostridiales sequences were identified in all but eight produced water samples, constituting more than 20% of the microbial community in six samples. Similarly to *Halanaerobiales*, the order *Clostridiales* comprises fermentative, obligate anaerobes, some of which are spore forming (53). Unlike *Halanaerobiales*, the order *Clostridiales* consists of many families with a diverse range of characteristics. Members of the acetogenic families *Clostridiaceae*, *Lachnospiraceae*, and *Acidaminobacteraceae* were abundant in our samples. These families include putative sulfate-reducing genera (54), such as *Desulfotomaculum*, which was identified in our study. Within the *Clostridiaceae*, members of the moderately halophilic, acid-producing, biofilm-forming genus *Clostridium* were also identified (55, 56).

We compared the most abundant orders identified (Fig. 3) with findings reported in previous produced water microbial ecology studies (10, 13, 14, 23, 28). Our data suggested *Halanaerobiales* to be the most dominant order across Marcellus Shale produced water samples. This observation (Fig. 3) supports previous studies (10, 13), which have shown *Halanaerobiales* to exist at lower abundances in wells after a short operational lifetime and to account for as much as 99% of the population in wells older than 6 months. Sequences affiliated with the *Clostridiales*, *Campylobacteriales*, *Rhodobacterales*, *Bacillales*, *Pseudomonadales*, *Oceanospirillales*, and *Bacteroidales* were found to be abundant in this study (Fig. 3) and have also been identified in previous produced

water studies (10, 13, 14, 19). Comparison to previous studies showed that all orders identified in our study at greater than 2% relative abundances have previously been detected in produced water, with the exception of members of the order *Erysipelotrichales*, which have not been identified in produced water from unconventional wells. *Erysipelotrichales* has been previously identified in petroleum reservoirs, formation waters, and subsurface hot springs (57, 58).

Influence of well age, salinity, and biocide treatment combination on produced water microbial communities. We evaluated the correlation between TDS, well age, and biocide application in the fracturing fluid with respect to the microbial community composition across all samples, the occurrence of the bacterial order *Halanaerobiales*, the microbial diversity within each sample, and the microbial abundance to identify operational or geochemical factors that impact microbial community structure in produced water. Previous studies have suggested TDS and well age to influence microbial community composition in produced water (10, 13); however, we did not observe any correlations between TDS concentration and the microbial ecology in produced water and identified only a moderate correlation between well age and the Shannon index, a measure of diversity within each sample.

In this study, we were particularly focused on the abundance of *Halanaerobiales*, as members of this taxon were shown to be predominant across all samples. We identified a moderate inverse correlation between *Halanaerobiales* relative abundance and well age, a result standing in contrast to previous data suggesting *Halanaerobiales* relative abundance to increase with well age; however, those studies investigated only a limited number of samples with well ages at the lower end of the time spectrum analyzed in this study (10, 13). We also identified a correlation between *Halanaerobiales* abundance and biocide treatment combination. A DBNPA-based biocide treatment combination correlated with a higher *Halanaerobiales* relative abundance. The biocide treatment combinations also contained additives such as polyethylene glycol and sodium hydroxide, serving other functions in the hydraulic fracturing fluid (59–61). Future research efforts are necessary to develop mechanistic explanations for these findings.

Functional potential of *Halanaerobium* from produced water. Microbial acid and sulfide production is of high interest to the oil and gas industry, due to microbe-influenced corrosion, fouling, and gas souring (3, 7, 10, 11). Taxonomic analysis of produced water allowed identification of multiple microbial taxa that are potentially involved in these processes, such as members of the orders *Halanaerobiales* and *Clostridiales*. Metagenomic investigation of *Halanaerobiales* acid production pathways identified putative metabolic fermentation pathways for lactate, acetate, ethanol, hydrogen, and carbon dioxide, consistent with previous reports for *Halanaerobium* species isolated from diverse environments (16, 20, 21, 31, 32, 62, 63). In particular, these findings agree with data recently reported for a similar produced water *Halanaerobium* draft genome (strain T82-1) with the capacity to produce ethanol, hydrogen, and acetate as fermentation products (19). In addition, our analysis identified potential sulfate-reducing bacteria. Specifically, sequences identified as corresponding to members of the sulfide-producing order *Desulfovibrionales* were more abundant than previously suggested (13, 14, 23). Reconstruction of putative sulfide production pathways in the recovered *Halanaerobium* sp. strain MDAL1 draft genome revealed the metabolic potential for thiosulfate reduction via a rhodanese thiosulfate reductase (Rdl). This pathway, utilizing thiosulfate or elemental sulfur instead of sulfate, has been previously described for *H. congolense*, a *Halanaerobium* species isolated from an oil field, and for *Halanaerobium* sp. strain T82-1 and *Halanaerobium* sp. strain DL-01, recently recovered from produced water (6, 19, 21, 35). These results suggest that multiple metabolic pathways, some of which would not be detected by current sulfate-reducing bacterium tests, have the potential to contribute to microbial sulfide production in the produced water environment. Genetic evidence for thiosulfate reduction also suggests the need to evaluate thiosulfate concentrations in produced water in future research efforts to generate additional geochemical support for these processes.

Fouling incidents in hydraulic fracturing infrastructure are also commonly attributed to microbial activity (1, 2). This study revealed several putative biofilm-forming microbial taxa to exist in produced water; in particular, the genera *Pseudomonas* and *Clostridium* have previously been suggested to be involved in biofilm formation (26, 64–66). In addition, several genes involved in biofilm formation processes were identified in the *Halanaerobium* sp. strain MDAL1 draft genome, suggesting the biofilm formation potential of produced water *Halanaerobium*. Our data therefore confirm recent work which reported genes for biofilm formation pathways to exist in draft genomes of *Halanaerobium* from fractured shale formations, despite their absence in other currently available *Halanaerobium* genomes (19). In addition, these findings also confirm previous studies that have reported the presence of *Halanaerobium* in biofilms found in hydrocarbon environments (67).

Finally, draft genome analysis revealed the potential for diverse stress response mechanisms in produced water *Halanaerobium*. Produced water *Halanaerobium* populations overcome osmotic stress through the uptake of potassium (i.e., salt-in strategy) and the utilization of osmoprotectants. In addition, we observed the genetic potential for motility; oxidative stress protection mechanisms characterized by rubredoxin, glutaredoxin, and superoxide reductase activity; and several heat shock- and periplasmic-stress-associated genes, enabling *Halanaerobium* survival in the saline, heavy metal-rich produced water environment. These findings are of particular interest as the stress response in microorganisms exposed to produced water has been shown to lead to enhanced biocide resistance and should be taken into consideration when evaluating biocide application strategies (26).

Study implications. An enhanced understanding of produced water microbial ecology is critical to limit corrosion, fouling, and souring issues; protect well infrastructure; minimize unnecessary biocide application; and encourage produced water recycling. This report represents the largest sampling and most extensive characterization of unconventional produced water microbial ecology to date. Recent studies have analyzed unconventional produced water microbiology data based on samples obtained from one or two Marcellus Shale well pads, producing valuable information limited by the small number of included sites (6, 10, 13, 19). The broader sampling effort in this produced water microbial ecology study allowed the confirmation of general trends observed during those previous temporal studies of a smaller number of wells, specifically, the predominance of the putative biofilm-forming and fermentative *Halanaerobiales* species. In addition, correlation analysis revealed a statistically significant influence of fracturing fluid biocide composition on the produced water microbial community and, in particular, on *Halanaerobiales* abundance.

Finally, this study evaluated the metabolic potential of *Halanaerobium* in produced water by successfully recovering a *Halanaerobium* draft genome and comparing it to other available *Halanaerobium* genomes. Annotation revealed genetic potential for several fermentation pathways, thiosulfate reduction, biofilm formation, and a diverse stress response, suggesting that *Halanaerobium* sp. strain MDAL1 contributes to acid and sulfide production in produced water. These genetic traits have also been previously observed in other *Halanaerobium* isolates and draft genomes, in particular, the species *Halanaerobium* sp. strain T82-1, *H. congolense*, and *Halanaerobium* sp. strain DL-01 (15, 19, 21).

In conclusion, this study confirmed the dominance of halophilic, fermentative microorganisms, in particular, *Halanaerobium*, across a broad sampling of produced water samples from the Marcellus Shale. Correlation analysis results suggest that TDS concentrations have little influence on the microbial ecology in produced water and that the biocide treatment combination may affect the abundance of *Halanaerobiales* in produced water. This study was one of the first efforts to evaluate the metabolic potential of microorganisms associated with hydraulic fracturing operations, supporting the role of *Halanaerobium* as a major contributor to microbial activity and a source for corrosion, souring, and biofouling in hydraulic fracturing infrastructure.

MATERIALS AND METHODS

Samples. Sample parameters reported in this study were based on previously suggested guidelines for the investigation of wastewater from unconventional shale gas extraction (68). All Marcellus Shale produced water samples were obtained during one sampling day in June 2014 from wells in southwest Pennsylvania. Samples were directly taken from the gas-water separator, collected in sterile 200-ml bottles, kept on ice during transport, and stored at -80°C within 24 h of sampling. The gas-water separator represented the closest available sampling port to the production well. Samples analyzed in this study represented 42 wells from 18 different well sites in the Marcellus Shale region in southwest Pennsylvania with production ages of 150 to 1,846 days (see Table S1 in the supplemental material).

Chemical analysis. Cation and trace element concentrations were determined in filtered, diluted subsamples using inductively coupled plasma mass spectrometry (ICP-MS) (NeXION 300X ICP-MS). Chloride concentrations were assessed using Thermo Scientific ICS-1100 ion chromatography (Thermo Fisher Scientific, Waltham, MA). Total dissolved solid (TDS) concentrations were determined based on measured cation and anion concentrations. Biocide utilization data for the 42 wells were obtained from FracFocus (<https://fracfocus.org/>). Nine different biocide treatment combinations were used across the analyzed wells, with biocide treatment combination 1 (19 wells) and biocide treatment combination 2 (13 wells) used in the majority of the wells (Table S2). Biocides were part of the fracturing fluid and applied during hydraulic fracturing.

DNA extraction, PCR, and sequencing. For each sample, 30 to 50 ml of produced water was centrifuged ($15,900 \times g$) to collect biomass. This volume was used because it represents a volume range successfully applied by our group and other research groups for DNA extraction from produced water (14, 28). Collected biomass was then digested with 10 μl of 20 mg/ml lysozyme for 30 min at 37°C followed by DNA extraction using a MoBio PowerSoil kit (Carlsbad, CA) according to the manufacturer's instructions. DNA from all samples was amplified using V4 region 16S rRNA gene primers as described previously (69, 70), cleaned using AMPure beads (Beckman Coulter, Pasadena, CA), and subjected to amplicon sequencing on an Illumina sequencer (Illumina, San Diego, CA). Detailed methodology can be found in the supplemental material.

qPCR. The microbial abundance for 42 produced water samples was determined using SYBR green-based quantitative PCR (qPCR) using primers described previously by Maeda et al. (71). Detailed methodology can be found in the supplemental material.

16S rRNA gene data processing. 16S rRNA gene sequences from all samples were analyzed using QIIME version 1.7.0 (70). Beta diversity was assessed by calculating weighted UniFrac distances (72). Alpha diversity was assessed by determining the number of operational taxonomic units (OTUs) and the Chao1 and Shannon indices per 2,000 sequences. DNA sequences were also annotated and deposited on MG-RAST (accession number 4696241.3). Detailed methodology and a description of fastq headers for sample sequences can be found in Fig. S1 in the supplemental material.

Correlation analysis. Spearman rank coefficients correlating taxonomy and diversity measures with TDS, well age, and bacterial abundance were calculated using R (73). Additionally, correlations were investigated using linear regression analysis. Differences in diversity and taxonomy by biocide treatment combination were assessed using two-tailed t tests. Microbial diversity between well sites was assessed by calculating weighted UniFrac distances and visualized using NMDS (nonmetric multidimensional scaling) ordination. Analysis of similarities (ANOSIM) based on Bray-Curtis and Euclidean distances was used to investigate statistical differences in community structures in R and Past (73, 74). Furthermore, samples were clustered based on analysis of weighted UniFrac distances with the unweighted pair group method using average linkages (UPGMA) in QIIME (70). The detailed methodology of the correlation analysis can be found in the supplemental material.

Metagenome library preparation. DNA from one produced water sample (site 13, well 2) was selected for metagenome sequencing. The metagenome sequencing library was processed using Nextera XT (Illumina, San Diego, CA) according to the manufacturer's instructions (see detailed workflow in the supplemental material). The same DNA extract used for 16S rRNA gene sequencing was used for metagenome sequencing. DNA libraries were normalized using Illumina bead technology, quantified using a Qubit fluorometer (Life Technologies, Carlsbad, CA), diluted to a concentration of 20 to 40 pM, and sequenced using a 300-cycle V2 kit on an Illumina MiSeq sequencer (Illumina, San Diego, CA).

Quality control and assembly. Sequence data were quality trimmed ($<Q30$, <100 bp) using CLC Genomics workbench version 8.5.1 (CLC Bio, Aarhus, Denmark) and assembled into contiguous sequences (contigs) using metaSPAdes version 3.5.1 (75). rRNA genes in assembled contigs were predicted using RNAmmer (30). Phylogenetic analysis based on extracted 16S rRNA gene sequences was performed using CLC workbench version 8.5.1 (CLC Bio, Aarhus, Denmark).

Reference genome mappings. *Halanaerobium* reference genomes available in June 2016 were downloaded from NCBI using CLC Genomic workbench version 8.5.1 (CLC Bio, Aarhus, Denmark) (Table 1). Trimmed sequence data were then mapped against each reference genome using the "map reads to reference" tool with default settings.

Binning, genome annotation, and functional gene mapping. *Halanaerobium* contigs were recovered from metagenomic data based on taxonomy using PhyloPythiaS (29) and on tetranucleotide frequency, differential coverage, and marker genes using Maxbin (76). The resulting *Halanaerobium* draft genome was assessed for completeness using CheckM (33).

KEGG orthology terms (77) were assigned to contigs and exported using RAST. KEGG pathways were determined using the KEGG mapper tool (77, 78). Functional gene sequences of interest were downloaded from RAST and evaluated using BLASTx or mapped against selected reference genomes (obtained from NCBI) using CLC workbench version 8.5.1 with default parameters.

Accession number(s). The generated *Halanaerobium* draft genome (*Halanaerobium* sp. strain MDAL1) was uploaded to NCBI GenBank (GenBank accession number [MIJU000000001](https://doi.org/10.1093/genbank/MIJU000000001); BioProject accession number [PRJNA341965](https://doi.org/10.1093/bioinformatics/PRJNA341965)) and RAST to be annotated using the SEED database (79, 80). The draft genome can be accessed under RAST accession number 6666666.207575.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/AEM.02659-16>.

SUPPLEMENTAL FILE 1, PDF file, 1.1 MB.

ACKNOWLEDGMENTS

We acknowledge Hector Moreno for his contributions toward this study.

This technical effort was performed under the RES contract RES1000027/183U as part of the National Energy Technology Laboratory's Regional University Alliance (NETL-RUA), a collaborative initiative of NETL, and was supported by the Oak Ridge Institute of Science and Education (ORISE).

REFERENCES

- Gregory K, Vidic R, Dzombak D. 2011. Water management challenges associated with the production of shale gas by hydraulic fracturing. *Elements* 7:181–186. <https://doi.org/10.2113/gselements.7.3.181>.
- Arthur J, Bohm B, Layne M. 2009. Hydraulic fracturing considerations for natural gas wells of the Marcellus Shale. *Gulf Coast Assoc Geol Soc Trans* 59:49–59.
- Gaspar J, Davis D, Camacho C, Alvarez PJ. 2016. Biogenic versus thermogenic H₂S source determination in Bakken wells: considerations for biocide application. *Environ Sci Technol Lett* 3:127–132. <https://doi.org/10.1021/acs.estlett.6b00075>.
- Mauter M, Alvarez P, Burton A, Cafaro D, Chen W, Gregory K, Jiang G, Li Q, Pittock J, Reible D. 2014. Regional variation in water-related impacts of shale gas development and implications for emerging international plays. *Environ Sci Technol* 48:8298–8306. <https://doi.org/10.1021/es405432k>.
- Gaspar J, Mathieu J, Yang Y, Tomson R, Leyris JD, Gregory KB, Alvarez PJ. 2014. Microbial dynamics and control in shale gas production. *Environ Sci Technol Lett* 1:465–473. <https://doi.org/10.1021/ez5003242>.
- Mouser PJ, Borton M, Darrah TH, Hartsock A, Wrighton KC. 2016. Hydraulic fracturing offers view of microbial life in the deep terrestrial subsurface. *FEMS Microbiol Ecol* 92:fiw166. <https://doi.org/10.1093/femsec/fiw166>.
- Akob D, Cozzarelli I, Dunlap D, Rowan E, Lorah M. 2015. Organic and inorganic composition and microbiology of produced waters from Pennsylvania shale gas wells. *Appl Geochem* 60:116–125. <https://doi.org/10.1016/j.apgeochem.2015.04.011>.
- Jan-Roblero J, Posadas A, de la Serna JZD, García R, Hernández-Rodríguez C. 2008. Phylogenetic characterization of bacterial consortia obtained of corroding gas pipelines in Mexico. *World J Microbiol Biotechnol* 24:1775–1784. <https://doi.org/10.1007/s11274-008-9674-x>.
- Vikram A, Lipus D, Bibby K. 2016. Metatranscriptome analysis of active microbial communities in produced water samples from the Marcellus Shale. *Microb Ecol* 72:571–581. <https://doi.org/10.1007/s00248-016-0811-z>.
- Cluff M, Hartsock A, MacRae J, Carter K, Mouser P. 2014. Temporal changes in microbial ecology and geochemistry in produced water from hydraulically fractured Marcellus Shale gas wells. *Environ Sci Technol* 48:6508–6517. <https://doi.org/10.1021/es501173p>.
- Davis JP, Struchtemeyer CG, Elshahed MS. 2012. Bacterial communities associated with production facilities of two newly drilled thermogenic natural gas wells in the Barnett Shale (Texas, USA). *Microb Ecol* 64: 942–954. <https://doi.org/10.1007/s00248-012-0073-3>.
- Kahrilas G, Blotvogel J, Stewart P, Borch T. 2015. Biocides in hydraulic fracturing fluids: a critical review of their usage, mobility, degradation, and toxicity. *Environ Sci Technol* 49:16–32.
- Murali Mohan A, Hartsock A, Bibby K, Hammack R, Vidic R, Gregory K. 2013. Microbial community changes in hydraulic fracturing fluids and produced water from shale gas extraction. *Environ Sci Technol* 47: 13141–13150. <https://doi.org/10.1021/es402928b>.
- Strong LC, Gould T, Kasinkas L, Sadowsky MJ, Aksan A, Wackett LP. 2013. Biodegradation in waters from hydraulic fracturing: chemistry, microbiology, and engineering. *J Environ Eng* 140:B4013001-1–B4013001-8. [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000792](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000792).
- Liang R, Davidova IA, Marks CR, Stamps BW, Harriman BH, Stevenson BS, Duncan KE, Suflita JM. 2016. Metabolic capability of a predominant *Halanaerobium* sp. in hydraulically fractured gas wells and its implication in pipeline corrosion. *Front Microbiol* 7:988. <https://doi.org/10.3389/fmicb.2016.00988>.
- Brown SD, Begemann MB, Mormile MR, Wall JD, Han CS, Goodwin LA, Pitluck S, Land ML, Hauser LJ, Elias DA. 2011. Complete genome sequence of the haloalkaliphilic, hydrogen-producing bacterium *Halanaerobium hydrogeniformans*. *J Bacteriol* 193:3682–3683. <https://doi.org/10.1128/JB.05209-11>.
- Sikorski J, Lapidus A, Chertkov O, Lucas S, Copeland A, Del Rio TG, Nolan M, Tice H, Cheng J-F, Han C. 2010. Complete genome sequence of *Acetohalobium arabaticum* type strain (Z-7288T). *Stand Genomic Sci* 3:57. <https://doi.org/10.4056/signs.1062906>.
- Zavarzin G, Zhilina T, Pusheva M. 1994. Halophilic acetogenic bacteria, p 432–444. In Drake H (ed), *Acetogenesis*. Springer, Berlin, Germany.
- Daly RA, Borton MA, Wilkins MJ, Hoyt DW, Kountz DJ, Wolfe RA, Welch SA, Marcus DN, Trexler RV, MacRae JD, Krzycki JA, Cole DR, Mouser PJ, Wrighton KC. 2016. Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing in shales. *Nat Microbiol* 2016:16146. <https://doi.org/10.1038/nmicrobiol.2016.146>.
- Oren A. 2014. The order *Halanaerobiales*, and the families *Halanaerobiaceae* and *Halobacteroidaceae*, p 153–177. In Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), *The prokaryotes: Firmicutes and Tenericutes*. Springer, Berlin, Germany.
- Ravot G, Magot M, Ollivier B, Patel B, Ageron E, Grimont P, Thomas P, Garcia JL. 1997. *Haloanaerobium congolense* sp. nov., an anaerobic, moderately halophilic, thiosulfate- and sulfur-reducing bacterium from an African oil field. *FEMS Microbiol Lett* 147:81–88. <https://doi.org/10.1111/j.1574-6968.1997.tb10224.x>.
- Ravot G, Ollivier B, Magot M, Patel B, Crolet J, Fardeau M, Garcia J. 1995. Thiosulfate reduction, an important physiological feature shared by members of the order *Thermotogales*. *Appl Environ Microbiol* 61: 2053–2055.
- Wuchter C, Banning E, Mincer TJ, Drenzek NJ, Coolen MJ. 2013. Microbial diversity and methanogenic activity of Antrim Shale formation waters from recently fractured wells. *Front Microbiol* 4:367. <https://doi.org/10.3389/fmicb.2013.00367>.
- Tang X, Xie G, Shao K, Chen Y, Gao G. 2012. Influence of salinity on the bacterial community composition in Lake Bosten, a large oligosaline lake in arid northwestern China. *Appl Environ Microbiol* 78:4748–4751. <https://doi.org/10.1128/AEM.07806-11>.
- Simachew A, Lanzén A, Gessesse A, Øvreås L. 2016. Prokaryotic community diversity along an increasing salt gradient in a soda ash concentration pond. *Microb Ecol* 71:326–338. <https://doi.org/10.1007/s00248-015-0675-7>.
- Vikram A, Lipus D, Bibby K. 2014. Produced water exposure alters

- bacterial response to biocides. *Environ Sci Technol* 48:13001–13009. <https://doi.org/10.1021/es5036915>.
27. Struchtemeyer CG, Morrison MD, Elshahed MS. 2012. A critical assessment of the efficacy of biocides used during the hydraulic fracturing process in shale natural gas wells. *Int Biodeter Biodegr* 71:15–21. <https://doi.org/10.1016/j.ibiod.2012.01.013>.
 28. Struchtemeyer CG, Elshahed MS. 2012. Bacterial communities associated with hydraulic fracturing fluids in thermogenic natural gas wells in North Central Texas, USA. *FEMS Microbiol Ecol* 81:13–25. <https://doi.org/10.1111/j.1574-6941.2011.01196.x>.
 29. McHardy AC, Martín HG, Tsigirgos A, Hugenholtz P, Rigoutsos I. 2007. Accurate phylogenetic classification of variable-length DNA fragments. *Nat Methods* 4:63–72. <https://doi.org/10.1038/nmeth976>.
 30. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <https://doi.org/10.1093/nar/gkm160>.
 31. Kivistö A, Larjo A, Ciranna A, Santala V, Roos C, Karp M. 2013. Genome sequence of *Halanaerobium saccharolyticum* subsp. *saccharolyticum* strain DSM 6643T, a halophilic hydrogen-producing bacterium. *Genome Announc* 1:e00187-13. <https://doi.org/10.1128/genomeA.00187-13>.
 32. Ivanova N, Sikorski J, Chertkov O, Nolan M, Lucas S, Hammon N, Deshpande S, Cheng J-F, Tapia R, Han C. 2011. Complete genome sequence of the extremely halophilic *Halanaerobium praevalens* type strain (GSLT). *Stand Genomic Sci* 4:312. <https://doi.org/10.4056/signs.1824509>.
 33. Parks D, Imelfort M, Skennerton CT, Hugenholtz P, Tyson G. 2014. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *PeerJ PrePrints* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
 34. An A, An D, Elliott A, Menon P, Shen Y, Voordouw G, Cote DE, Miner K. 2015. Advances in tools to monitor souring and corrosion in oil and gas fields, 2015. SPE International Symposium on Oilfield Chemistry, Society of Petroleum Engineers. <https://doi.org/10.2118/173805-MS>.
 35. Ravot G, Casalot L, Ollivier B, Loison G, Magot M. 2005. *rdIA*, a new gene encoding a rhodanese-like protein in *Halanaerobium congolense* and other thiosulfate-reducing anaerobes. *Res Microbiol* 156:1031–1038. <https://doi.org/10.1016/j.resmic.2005.05.009>.
 36. Vengosh A, Jackson RB, Warner N, Darrah TH, Kondash A. 2014. A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environ Sci Technol* 48:8334–8348. <https://doi.org/10.1021/es405118y>.
 37. Whiteley CG, Lee D-J. 2015. Bacterial diguanylate cyclases: structure, function and mechanism in exopolysaccharide biofilm development. *Biotechnol Adv* 33:124–141. <https://doi.org/10.1016/j.biotechadv.2014.11.010>.
 38. Stanley NR, Lazazzera BA. 2004. Environmental signals and regulatory pathways that influence biofilm formation. *Mol Microbiol* 52:917–924. <https://doi.org/10.1111/j.1365-2958.2004.04036.x>.
 39. Pratt LA, Kolter R. 1998. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* 30:285–293. <https://doi.org/10.1046/j.1365-2958.1998.01061.x>.
 40. Belas R. 2014. Biofilms, flagella, and mechanosensing of surfaces by bacteria. *Trends Microbiol* 22:517–527. <https://doi.org/10.1016/j.tim.2014.05.002>.
 41. Touati D. 2000. Iron and oxidative stress in bacteria. *Arch Biochem Biophys* 373:1–6. <https://doi.org/10.1006/abbi.1999.1518>.
 42. Kempf B, Bremer E. 1998. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch Microbiol* 170:319–330. <https://doi.org/10.1007/s002030050649>.
 43. Oren A. 2013. Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. *Front Microbiol* 4:315. <https://doi.org/10.3389/fmicb.2013.00315>.
 44. Kraegeloh A, Amendt B, Kunte HJ. 2005. Potassium transport in a halophilic member of the bacteria domain: identification and characterization of the K⁺ uptake systems TrkH and TrkI from *Halomonas elongata* DSM 2581T. *J Bacteriol* 187:1036–1043. <https://doi.org/10.1128/JB.187.3.1036-1043.2005>.
 45. Oren A. 2002. Halophilic microorganisms and their environments, p 202–231. Springer, New York, NY.
 46. Oren A. 2008. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Systems* 4:2. <https://doi.org/10.1186/1746-1448-4-2>.
 47. Kunte H. 2005. K⁺ transport and its role for osmoregulation in a halophilic member of the Bacteria domain: characterization of the K⁺ uptake systems from *Halomonas elongata*, p 287–300. In *Adaptation to life at high salt concentrations in Archaea, Bacteria, and Eukarya*. Springer, Berlin, Germany.
 48. Hoffmann T, Wensing A, Brosius M, Steil L, Völker U, Bremer E. 2013. Osmotic control of *opuA* expression in *Bacillus subtilis* and its modulation in response to intracellular glycine betaine and proline pools. *J Bacteriol* 195:510–522. <https://doi.org/10.1128/JB.01505-12>.
 49. Bouvier J, Bordes P, Romeo Y, Fourçans A, Bouvier I, Gutierrez C. 2000. Characterization of OpuA, a glycine-betaine uptake system of *Lactococcus lactis*. *J Mol Microbiol Biotechnol* 2:199–205.
 50. Cray JA, Stevenson A, Ball P, Bankar SB, Eleutherio EC, Ezeji TC, Singhal RS, Thevelein JM, Timson DJ, Hallsworth JE. 2015. Chaotropicity: a key factor in product tolerance of biofuel-producing microorganisms. *Curr Opin Biotechnol* 33:228–259. <https://doi.org/10.1016/j.copbio.2015.02.010>.
 51. Bartlett D, Chi E. 1994. Genetic characterization of ompH mutants in the deep-sea bacterium *Photobacterium* sp. strain SS9. *Arch Microbiol* 162:323–328. <https://doi.org/10.1007/BF00263779>.
 52. Ollivier B, Caumette P, Garcia J-L, Mah R. 1994. Anaerobic bacteria from hypersaline environments. *Microbiol Rev* 58:27–38.
 53. Paredes-Sabja D, Setlow P, Sarker MR. 2011. Germination of spores of *Bacillales* and *Clostridiales* species: mechanisms and proteins involved. *Trends Microbiol* 19:85–94. <https://doi.org/10.1016/j.tim.2010.10.004>.
 54. Widdel F, Pfennig N. 1977. A new anaerobic, sporing, acetate-oxidizing, sulfate-reducing bacterium, *Desulfotomaculum acetoxidans*. *Arch Microbiol* 112:119–122. <https://doi.org/10.1007/BF00446665>.
 55. Detkova E, Pusheva M. 2006. Energy metabolism in halophilic and alkaliphilic acetogenic bacteria. *Microbiology* 75:1–11. <https://doi.org/10.1134/S0026261706010012>.
 56. Müller V, Frerichs J. 20 September 2013. Acetogenic bacteria. In eLS. John Wiley & Sons Ltd., Chichester, United Kingdom. <https://doi.org/10.1002/9780470015902.a0020086.pub2>.
 57. Lavalleur HJ, Colwell FS. 2013. Microbial characterization of basalt formation waters targeted for geological carbon sequestration. *FEMS Microbiol Ecol* 85:62–73. <https://doi.org/10.1111/1574-6941.12098>.
 58. Wang L, Ke W, Sun X, Liu J, Gu J, Mu B. 2014. Comparison of bacterial community in aqueous and oil phases of water-flooded petroleum reservoirs using pyrosequencing and clone library approaches. *Appl Microbiol Biotechnol* 98:4209–4221. <https://doi.org/10.1007/s00253-013-5472-y>.
 59. Arthur JD, Bohm B, Layne M. 2008. Hydraulic fracturing considerations for natural gas wells of the Marcellus Shale. The Groundwater Protection Council 2008 Annual Forum, 21–24 September, Cincinnati, OH.
 60. Montgomery C. 2013. Fracturing fluids. Abstr ISRM Int Conf Eff Sustain Hydraul Fracturing, International Society for Rock Mechanics, abstr ISRM-ICHF-2013-035. <https://www.onepetro.org/conference-paper/ISRM-ICHF-2013-035>.
 61. FracFocus Chemical Discovery Registry. 2014. Chemical use in hydraulic fracturing: what chemicals are used. Ground Water Protection Council, Interstate Oil and Gas Compact Commission, Oklahoma City, OK. <http://fracfocus.org/chemical-use/what-chemicals-are-used>. Accessed 10 June 2015.
 62. Abdeljabbar H, Cayol J-L, Hania WB, Boudabous A, Sadfi N, Fardeau M-L. 2013. *Halanaerobium sehlinense* sp. nov., an extremely halophilic, fermentative, strictly anaerobic bacterium from sediments of the hypersaline lake Sebkha. *Int J Syst Evol Microbiol* 63:2069–2074. <https://doi.org/10.1099/ijs.0.040139-0>.
 63. Bhupathiraju VK, McInerney MJ, Woese CR, Tanner RS. 1999. *Haloanaerobium kushneri* sp. nov., an obligately halophilic, anaerobic bacterium from an oil brine. *Int J Syst Bacteriol* 49(Pt 3):953–960. <https://doi.org/10.1099/00207713-49-3-953>.
 64. Drenkard E, Ausubel FM. 2002. *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature* 416:740–743. <https://doi.org/10.1038/416740a>.
 65. Mann EE, Wozniak DJ. 2012. *Pseudomonas* biofilm matrix composition and niche biology. *FEMS Microbiol Rev* 36:893–916. <https://doi.org/10.1111/j.1574-6976.2011.00322.x>.
 66. Đapa T, Leuzzi R, Ng YK, Baban ST, Adamo R, Kuehne SA, Scarselli M, Minton NP, Serruto D, Unnikrishnan M. 2013. Multiple factors modulate biofilm formation by the anaerobic pathogen *Clostridium difficile*. *J Bacteriol* 195:545–555. <https://doi.org/10.1128/JB.01980-12>.
 67. Neria-González I, Wang E, Ramírez F, Romero J, Hernández-Rodríguez C. 2006. Characterization of bacterial community associated to biofilms of

- corroded oil pipelines from the southeast of Mexico. *Anaerobe* 12: 122–133. <https://doi.org/10.1016/j.anaerobe.2006.02.001>.
68. Bibby K, Brantley S, Reible D, Linden K, Mouser P, Gregory K, Ellis B, Vidic R. 2013. Suggested reporting parameters for investigations of wastewater from unconventional shale gas extraction. *Environ Sci Technol* 47: 13220–13221. <https://doi.org/10.1021/es404960z>.
 69. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6:1621–1624. <https://doi.org/10.1038/ismej.2012.8>.
 70. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JL. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f.303>.
 71. Maeda H, Fujimoto C, Haruki Y, Maeda T, Kokeguchi S, Petelin M, Arai H, Tanimoto I, Nishimura F, Takashiba S. 2003. Quantitative real-time PCR using TaqMan and SYBR Green for *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis*, *Prevotella intermedia*, *tetQ* gene and total bacteria. *FEMS Immunol Med Microbiol* 39:81–86. [https://doi.org/10.1016/S0928-8244\(03\)00224-4](https://doi.org/10.1016/S0928-8244(03)00224-4).
 72. Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71: 8228–8235. <https://doi.org/10.1128/AEM.71.12.8228-8235.2005>.
 73. R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org>.
 74. Hammer Ø, Harper D, Ryan P. 2009. PAST—PALaeontological STATistics, ver. 1.89. University of Oslo, Oslo, Norway.
 75. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 76. Wu Y-W, Tang Y-H, Tringe SG, Simmons BA, Singer SW. 2014. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* 2:26. <https://doi.org/10.1186/2049-2618-2-26>.
 77. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28:27–30. <https://doi.org/10.1093/nar/28.1.27>.
 78. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40(Database issue):D109–D114. <https://doi.org/10.1093/nar/gkr988>.
 79. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
 80. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.
 81. Mohan AM, Bibby KJ, Lipus D, Hammack RW, Gregory KB. 2014. The functional potential of microbial communities in hydraulic fracturing source water and produced water from natural gas extraction characterized by metagenomic sequencing. *PLoS One* 9:e107682. <https://doi.org/10.1371/journal.pone.0107682>.