Supporting Information

Microbial Community Changes in Hydraulic Fracturing Fluids and Produced Water from Shale Gas Extraction

Authors
Arvind Murali Mohan[1,2], Angela Hartsock[1], Kyle J Bibby[3], Richard W Hammack[1], Radisav D Vidic[1,3] and Kelvin B Gregory[1,2*]

[1]National Energy Technology Laboratory, Pittsburgh, PA 15236

[2]Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213

[3]Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, PA 15261

***Corresponding author: Kelvin B Gregory**
Phone: 412-268-9811
Fax: 412-268-7813
email: kelvin@cmu.edu

Supporting information includes 3 tables and 3 figures and methods for clone library, pyrosequencing and qPCR techniques

Materials and Methods:

Composition of PCR Reaction Mixture and Thermal Cycling Conditions for Clone Libraries. The total volume of each PCR reaction mixture was 27 µL and contained: 0.5 µL (5U/µL) Taq Polymerase (Qiagen Valencia, CA), 0.5 µL (10 µM) each of forward and reverse primers, 2 µL (2 mM) of dNTP solution (Quiagen, Valencia, CA), 0.5 µL (12.5 g/L) of Bovine Serum Albumen (Applied Biosystems, Foster City, CA), 3 µL (2 mM) of MgCl2 (Quiagen, Valencia, CA), 5 µL of 5X Q solution (Quiagen, Valencia, CA), 2.5 µL of 10X PCR buffer (Quiagen, Valencia, CA), 3 µL of template DNA and made up to the total volume using PCR grade water. Thermal cycler conditions for amplification were, initial denaturation at 95°C for 3 min, 30 cycles of 94°C denaturation for 1 min, annealing at 55°C for 45 s, elongation at 72°C for 45 s and a final extension at 72°C for 7 min.

Composition of PCR Reaction Mixture and Thermal Cycling Conditions for Tag Encoded Pyrosequencing. The total volume of each PCR reaction mixture was 50µL and contained: 25µL of Hot Start TAQ Master Mix (Qiagen Valencia, CA), 0.4-0.5 µL of each of the primers (0.5 µM final concentration), 3 µL of template DNA and made up to the total volume using PCR grade water. Thermal cycler conditions for amplification were, initial denaturation at 95°C for 4 min, 30 cycles of 94°C denaturation for 1 min, annealing at 55°C for 1 min, elongation at 72°C for 1 min and a final extension at 72°C for 10 min .

Composition of PCR Reaction Mixture and Thermal Cycling Conditions for Tag Encoded Pyrosequencing. The total volume of the reaction mixture was 20 µL and contained: 9.5 µL TaqMan universal PCR Master Mix (Applied Biosystems, Foster City, CA), 1.2 µL (10 µM) of each forward and reverse primers, 0.5 µL (10 µM) of probe TM1389F, 1 µL of genomic DNA and made up to the total volume using PCR grade water. Thermal cycling conditions were an initial soak step at 50°C for 2 min followed by denaturation at 95°C for 10 min. Finally, 40 cycles of 15 s denaturation at 95°C, and 1 min of annealing and extension at 56°C were performed.

Table S1: Local Parsimony (p) scores and significance from pairwise comparison of bacterial 16S rRNA gene sequences from pyrosequencing in source water, fracturing fluid and Produced Water (PW) days 1, 7, 9 and 187.

Sample Pairs	Pars Score	Pars Sig
PW day 1- PW day 7	97	<0.0010
PW day 1- PW day 9	62	<0.0010
PW day 7- PW day 9	82	1
PW day 1- Fracturing Fluid	148	0.01
PW day 7- Fracturing Fluid	67	<0.0010
PW day 9- Fracturing Fluid	54	<0.0010
PW day 1- PW day 187	17	<0.0010
PW day 7- PW day 187	14	<0.0010
PW day 9- PW day 187	15	<0.0010
Fracturing Fluid- PW day 187	15	<0.0010
PW day 1- Source Water	128	0.278
PW day 7- Source Water	57	<0.0010
PW day 9- Source Water	49	<0.0010
Fracturing Fluid- Source Water	124	1
PW day 187- Source Water	14	<0.0010

Table S2: Diversity, richness estimates[1], and coverage of bacterial 16S rRNA genes in source water, fracturing fluid, produced water (PW) days 1, 7, 9 and 187 from clone library

Sample Type	No. of Sequences	$OTU_{0.03}$	Coverage %	Chao Richness	Abundance-based Coverage Estimation Richness	Shannon Diversity Index
Source water	75	33	75	57	64	2.91
Fracturing fluid	71	33	72	79	115	3.02
PW day 1	75	30	79	49	53	2.92
PW day 7	83	27	82	44	56	2.62
PW day 9	82	24	78	160	141	2.44
PW day 187	90	3	99	3	4	0.16

[1]Diversity and richness estimates are based on ≥97% sequence identity

Table S3: Relative abundance of minor bacterial classes (<2% of the bacterial population in each of the samples) in fracturing source water, fracturing fluid and produced water (PW) day 1, 7, 9 and 187 samples from pyrosequencing.

Class/ Sample	Source water	Fracturing Fluid	PW day 1	PW day 7	PW day 9	PW day 187
Deferribacteres		0.02%				
Fusobacteria	0.03%	0.02%				
Actinobacteria	0.10%	0.05%	0.01%			0.12%
Spirochaetes		0.02%				
Bacteroidetes_incertae_sedis		0.02%		0.17%		
Bacteroidia	0.71%	0.45%				
Sphingobacteria	0.85%	0.10%	0.21%			
Flavobacteria		1.49%				0.03%
Deltaproteobacteria	0.27%	1.00%	1.38%	0.70%	1.34%	0.03%
Cyanobacteria	0.75%	1.22%	0.72%			
Chloroplast	0.92%	0.17%				
Thermotogae	0.07%	0.15%	0.05%	0.04%	0.24%	
Clostridia	0.07%	0.52%				
Bacilli	0.10%	0.15%	0.21%			0.02%
Synergistia		0.30%	0.41%			
Anaerolineae		0.05%				
Mollicutes	0.03%					
Betaproteobacteria				0.02%	0.05%	0.01%
Negativicutes			0.03%			
Opitutae			0.01%	0.04%		
Flavobacteria				0.02%	0.10%	
Alphaproteobacteria				1.34%	1.05%	
Gammaproteobacteria						0.09%
Unclassified	1.80%	1.02%	1.68%	1.41%	0.76%	0.43%

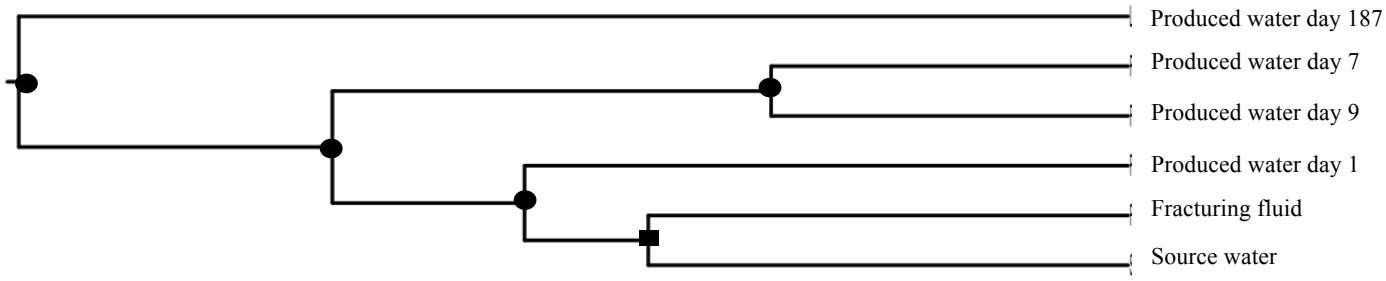

Figure S1(a): Comparisons of microbial communities in fracturing source water, fracturing fluid and produced water day 1, 7, 9 and 187 samples from pyrosequencing using weighted and normalized Unifrac metric. The symbols found on the nodes of the tree indicate the level of jackknife support for each node and are based on 100 replicates:
>99.9%, 90-99.9%

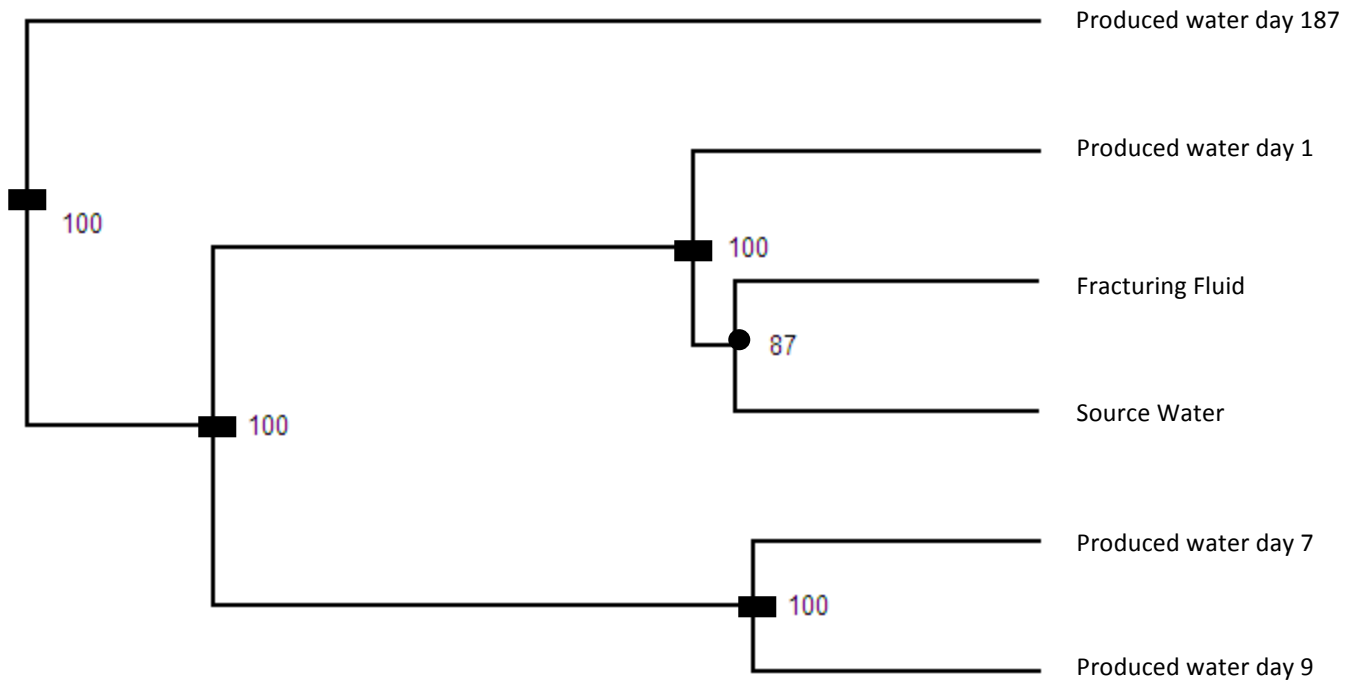

Figure S1 (b): Comparisons of bacterial sequences from source water, fracturing fluid and produced water day 1, 7, 9 and 187 samples from clone library results using weighted and normalized Unifrac metric. The symbols found on the nodes of the tree indicate the level of jackknife support for each node and are based on 100 replicates: >99%, 70-90%

Order (Class)	Sample name					
	SW	FF	PW day1	PW day7	PW day9	PW day 187
Rhodobacterales (Alphaproteobacteria)						
Sphingomonadales (Alphaproteobacteria)						
Caulobacterales (Alphaproteobacteria)						
Rhodospirillales (Alphaproteobacteria)						
Pseudomonadales (Gammaproteobacteria)						
Vibrionales (Gammaproteobacteria)						
Alteromonadales (Gammaproteobacteria)						
Chromatiales (Gammaproteobacteria)						
Campylobacterales (Epsilonproteobacteria)						
Burkholderiales (Betaproteobacteria)						
Thermoanaerobacterales (Clostridia)						
Halanaerobiales (Clostridia)						
Clostridiales (Clostridia)						
Bacteroidales (Bacteroidia)						
Flavobacteriales (Flavobacteria)						
Fusobacteriales (Fusobacteria)						
Bacillales (Bacilli)						
Lactobacillales (Bacilli)						

Legend (% of bacterial community)

0%	
>0 - 5%	
>5-10%	
>10-20%	
>20-30%	
>30-50%	
>95%	

Figure S2: Relative abundance of major bacterial orders (>3% of the community in at least one of the samples) with respect to the total bacterial community, in fracturing source water (SW), fracturing fluid (FF) and produced water (PW) day 1, 7, 9 and 187 samples from pyrosequencing results.

Order (Class)	SW	FF	PW day1	PW day7	PW day9	PW day187
Rhodobacterales (Alphaproteobacteria)						
Pseudomonadales (Gammaproteobacteria)						
Sphingomonadales (Alphaproteobacteria)						
Rhizobiales (Alphaproteobacteria)						
Orders within Chloroplast and Cyanobacteria						
Flavobacteriales (Flavobacteria)						
Methylophilales (Betaproteobacteria)						
Rhodospirillales (Alphaproteobacteria)						
Chromatiales (Gammaproteobacteria)						
Campylobacterales (Epsilonproteobacteria)						
Sphingobacteriales (Alphaproteobacteria)						
Caulobacterales (Alphaproteobacteria)						
Aeromonadales (Gammaproteobacteria)						
Alteromonadales (Gammaproteobacteria)						
Thiotrichales (Gammaproteobacteria)						
Vibrionales (Gammaproteobacteria)						
Burkholderiales (Betaproteobacteria)						
Thermoanaerobacterales (Clostridia)						
Bacteroidales (Bacteroidia)						
Clostridiales (Clostridia)						
Desulfovibrionales (Deltaproteobacteria)						
Fusobacteriales (Fusobacteria)						
Halanaerobiales (Clostridia)						
Bacillales (Bacilli)						
Oceanospirillales (Gammaproteobacteria)						
Lactobacillales (Bacilli)						
Xanthomonadales (Gammaproteobacteria)						
Unclassified						

Legend (% of bacterial community)

0%	
>0 - 5%	
>5-10%	
>10-20%	
>20-30%	
>30-50%	
>95%	

Figure S3: Relative abundance of bacterial orders with respect to the total bacterial community, in fracturing source water (SW), fracturing fluid (FF) and produced water (PW) day 1, 7, 9 and 187 samples from clone library results.