

1 Differential structure and functional gene response to geochemistry associated with the
2 suspended and attached shallow aquifer microbiomes from the Illinois Basin, IL

3
4
5
6 Yiran Dong^{1,2}, Robert A. Sanford³, Lynn Connor⁴, Joanne Chee-Sanford⁴, Bracken T. Wimmer⁵,
7 Abbas Iranmanesh⁵, Liang Shi^{1,2}, Ivan G. Krapac⁵, Randall A. Locke II⁵, Hongbo Shao^{5*}

8
9
10
11 ¹ School of Environmental Studies, China University of Geosciences (Wuhan), China

12 ² State Key Laboratory of Biogeology and Environmental Geology, China University of
13 Geosciences (Wuhan), China

14 ³Department of Geology, University of Illinois Urbana-Champaign, USA

15 ⁴USDA-ARS, Urbana, IL, USA

16 ⁵ Illinois State Geological Survey, IL, USA

17
18
19
20
21
22 *Corresponding author. Mailing address: 615 E. Peabody Ave, MC-650, Champaign, IL, 61820.

23 Phone: +01-(217)300-3157. Email: hbshao@illinois.edu

27 **Author Contribution**

28 Y. D., R. A. S. and H. S. designed the experiments; Y. D., R. A. S. performed data analyses,
29 wrote, reviewed and edited the manuscripts. L. C. and J. C. conducted all the molecular work. B.
30 T. W., A. I., I. G. K, R. A. L., and H. S. collected samples and carried out geochemical analyses.
31 L. S., H. S. and R. A. L. validated and reviewed the manuscript. R. A. L. and H. S. acquired the
32 primary funding and supervised the project.

33

34

35

36 **Abstract**

37 Despite the clear ecological significance of the microbiomes inhabiting groundwater and
38 connected ecosystems, our current understanding of their habitats, functionality, and the
39 ecological processes controlling their assembly have been limited. In this study, an efficient
40 pipeline combining geochemistry, high-throughput Fluidigm™ functional gene amplification
41 and sequencing was developed to analyze the suspended and attached microbial communities
42 inhabiting five groundwater monitoring wells in the Illinois Basin, USA. The dominant taxa in
43 the suspended and the attached microbial communities exhibited significantly different spatial
44 and temporal changes in both alpha- and beta-diversity. Further analyses of representative
45 functional genes affiliated with N₂ fixation (*nifH*), methane oxidation (*pmoA*), and sulfate
46 reduction (*dsrB*, and *aprA*), suggested functional redundancy within the shallow aquifer
47 microbiomes. While more diversified functional gene taxa were observed for the suspended
48 microbial communities than the attached ones except for *pmoA*, different levels of changes over
49 time and space were observed between these functional genes. Notably, deterministic and
50 stochastic ecological processes shaped the assembly of microbial communities and functional
51 gene reservoirs differently. While homogenous selection was the prevailing process controlling
52 assembly of microbial communities, the neutral processes (e.g., dispersal limitation, drift and
53 others) were more important for the functional genes. The results suggest complex and changing
54 shallow aquifer microbiomes, whose functionality and assembly vary even between the spatially
55 proximate habitats and fractions. This research underscored the importance to include all the
56 interface components for a more holistic understanding of the biogeochemical processes in
57 aquifer ecosystems, which is also instructive for practical applications.

58

59 **Key words**

60 Shallow aquifer microbiomes, multiplex PCR, temporal-spatial affects, functional genes,

61 suspended vs attached microbial communities, community assembly

62

63 1. Introduction

64 Groundwater and connected ecosystems harbor a significant fraction of the microbial life
65 on Earth and are critical in many biogeochemical processes, including mineral weathering (Akob
66 and Küsel, 2011), nutrient transformation (Schwab et al., 2017) and contaminant remediation
67 (Benner et al., 1999; Hwang et al., 2006). Thus, they may impact the safety of large-scale
68 engineered projects (e.g., oil and gas production, storage of radioactive waste, and greenhouse
69 gas capture and sequestration) (Booker et al., 2017; Borton et al., 2018; Wilkins et al., 2014). In
70 these ecosystems, the flux of water and nutrients through the subsurface link the habitats and
71 drive microbial activity (Akob and Küsel, 2011; Danielopol et al., 2003; Gerth and Förstner,
72 2004; Lee and Lee, 2018). It has been well recorded that the aquifer microbiomes are highly
73 diverse and complex. Composition of aquifer microbial species cohorts are spatially and
74 temporally variable, potentially selected by their adaptation to the ambient environments (e.g.,
75 pore water chemistry, mineral composition, and available nutrients) (Akob and Küsel, 2011;
76 Griebler and Lueders, 2009; Hug et al., 2015).

77 Within the saturated, capillary fringe and vadose zones in subsurface, microorganisms
78 can reside in two distinct phases (i.e., planktonic and biofilm), and significant differences in
79 community composition, structure and activity between free-living and attached communities are
80 commonly accepted (Smith et al., 2018). While most studies on shallow aquifer microbiomes
81 have focused on easily accessible free-living (or suspended) organisms, a large proportion of
82 microbial biomass dwelling on mineral surfaces or in rocks have been relatively under-
83 investigated (Alfreider et al., 1997; Flemming and Wuertz, 2019; Hug et al., 2015; Lazar et al.,
84 2019). The community composition overlapped between the two habitats is typically less than
85 30 % (Barnhart et al., 2013; Flynn et al., 2008; Flynn et al., 2013). A study using microscopic

86 cell counts demonstrated that the organisms attached to groundwater sediments were orders of
87 magnitude higher than those suspended in the groundwater (Hazen et al., 1991). This finding has
88 been corroborated by many studies on pristine aquifers (e.g., sandy glaciers, sandstone,
89 limestone and marlstone formations), suggesting that the ecological and geochemical
90 significance of attached microbial communities need to be considered for these ecosystems
91 (Flynn et al., 2008; Flynn et al., 2013; Flynn et al., 2012; Hug et al., 2015; King et al., 2017; Lee
92 and Lee, 2018).

93 Evaluation of functionalities based on a variety of genetic, genomic, and multi-omic
94 analyses demonstrated complex microbe-environment interactions in aquifer. Some studies
95 suggest that differences in microbial communities do not necessarily correlate with the
96 ecosystem processes due to functional redundancy (Fields et al., 2006; Zelaya et al., 2019). In
97 other studies, however, abundance and/or expression of the specific functional genes associated
98 with elemental cycling showed active response to the chemical gradients or external
99 perturbation. These are typically more pronounced in the contaminated sites,²⁴ oil and gas
100 reservoirs, where both organic or inorganic substrates (e.g., hydrocarbon, methane,
101 metals/metalloids, nitrate, sulfate, and ammonium) were significantly higher than the pristine
102 background (Fields et al., 2006; He et al., 2018; Schwab et al., 2017; Wegner et al., 2019; Xu et
103 al., 2010). Another study also implied a strong relationship between the abundance of different
104 C-cycling functional genes and their specific enzymatic activities in different soil samples,
105 implying that activity of enzymes may be predicted by the abundance of associated functional
106 genes (Trivedi et al., 2016).

107 It is broadly recognized that assembly of microbial communities is simultaneously
108 influenced by deterministic (e.g., homogenous/heterogenous selection) and stochastic processes

109 (e.g., dispersal limitation, homogenizing dispersal, drift and others) (Danczak et al., 2018; Ning
110 et al., 2019; Ning et al., 2020). Several studies on this topic in relation to aquifer microbiomes
111 have been done in systems with spatial or temporal changes, or with different levels of
112 perturbation (Danczak et al., 2018; Fillinger et al., 2019; Yan et al., 2020; Zelaya et al., 2019). In
113 most studies, deterministic processes imposed strong selection for community assembly, while
114 their relative contribution was case dependent and associated with ambient geochemistry,
115 hydrodynamic conductivity, connectivity or distance between the habitats (Gulmann et al., 2015;
116 Yan et al., 2020; Zelaya et al., 2019). Comparatively, stronger selection was suggested for
117 assembly of the attached microbial communities (Graham et al., 2017). However, our current
118 understanding of the sediment-associated subsurface microbial community assembly is limited.

119 In this work, we applied an *in-situ* incubation technique with representative aquifer
120 material derived from five shallow groundwater monitoring wells selected at a large-scale carbon
121 capture and storage (CCS) demonstration project (Finley, 2014). The specific objectives were to
122 enhance the understanding of: (1) microbial population distribution between suspended and
123 attached phases of the shallow aquifer ecosystems; (2) the environmental factors that drive
124 temporal and spatial changes of subsurface microbiome composition; (3) the microbial genetic
125 response to dynamic subsurface geochemical conditions; and (4) the contributions of
126 deterministic vs stochastic processes to the assembly of the microbial communities and
127 functional gene reservoirs. Heterogeneous spatial distribution and temporal changes in the
128 microbial community structure and functional genes as well as their assemblies were
129 systematically investigated and evaluated in the five wells.

130

131 **2. Methods and Material**

132 2.1 Sampling sites

133 Samples for this study were collected from a large-scale CCS demonstration project,
134 Illinois Basin – Decatur Project (IBDP) at Decatur, Illinois, USA (Finley, 2014). In this project,
135 one million tons of CO₂ were injected into a deep sandstone formation during 2011-2015. Within
136 the shallow groundwater monitoring net (see section S1.1 in the Supplementary Information
137 (SI)), five monitoring wells were selected (Table 1 and Fig. S1). Wells G101, G102, G103, and
138 G104 are screened in the Upper Pennsylvanian bedrock with the intervals in the range of 41.4-
139 43.4 m below land surface. Wells G101 and G102 were installed 15 m northeast and northwest,
140 respectively, of the CO₂ injection well (CCS1); well G103 was installed 61 m northwest of
141 CCS1, and G104 was approximately 610 m northwest of CCS1 (Fig. S1). Well 04UG is screened
142 in the Upper Glasford Formation, an unconsolidated Quaternary glacial deposit. This well has a
143 screened interval from 18.6 to 20.1 m below land surface and is about 18 m southwest of G104
144 (Fig. S1 and Table 1). The Upper Pennsylvanian bedrock is dominated by quartz and clay with
145 small amounts of dolomite (0.9 %) and pyrite. Comparatively, sediment in silty Glasford
146 Formation consists of significantly higher amount of quartz (55.5-58.5 %) and dolomite (16.9-
147 22.8 %) than the Upper Pennsylvanian bedrock (Table 1). Although the depth of G104 was
148 similar to the other three G wells, its hydraulic conductivity was about one order of magnitude
149 higher than the other wells. Based on Darcy's law using observed water levels and calculated
150 hydraulic conductivity values, the groundwater flow rates in four bedrock wells were estimated
151 to be 1.0-1.5 meters/year (Coady and Mehnert, 2011; Locke and Iranmanesh, 2015).

152

153 2.2 Sample collection and sample pretreatment

154 For this study, both suspended and attached organisms were sampled to better represent
155 the native groundwater microbiomes (Table S1). A series of “traps” were prepared by filling
156 nylon bags with autoclaved aquifer material previously collected from same boreholes and at the
157 same screened interval under study. The aquifer material was manually crashed and the particles
158 with size ranging from 0.2-1.6 mm were collected and rinsed with nanopure water. The air-dried
159 aquifer material was then autoclaved, aliquoted and packed with nylon bags that had been
160 sequentially washed with 70 % ethanol and sterile nanopure water. The “traps” were lowered
161 into each selected well at the level of the well screen for microbial attachment (Flynn et al.,
162 2008). These *in-situ* sediment microcosms were deployed on October 17, 2017, and remained
163 there for 3-6 months so that active microbial populations could colonize the sand in the traps.
164 Water and “trap” samples were collected initially during sediment trap deployment and
165 subsequently at each sampling time (i.e., 3 and 6 months).

166 Before filtering planktonic cells from groundwater or retrieving *in situ* “traps” for
167 attached microbe sampling, groundwater was pumped out of the well using an electric
168 submersible pump. During pumping, the physical and geochemical parameters (e.g., pH,
169 temperature, electrical potential and conductivity) were monitored overtime until all the readings
170 were stabilized. Groundwater samples for geochemical analyses were filtered in the field using
171 0.45 µm flow-through cartridge filters (Geotech Environmental Equipment Inc., CO, USA). For
172 analysis of dissolved inorganic carbon (DIC), 3 mL of groundwater was collected using a
173 degassed syringe, then injected into a stoppered 70 mL serum bottle previously purged using O₂-
174 free, ultrapure N₂ gas and 1 mL 50 % (w/w) phosphoric acid. The samples for dissolved organic
175 carbon were stored in amber glass bottles and preserved using H₂SO₄ with the final concentration
176 0.5 % (v/v). The samples were stored on ice in a cooler before transported back to the laboratory

177 and kept refrigerated before use.

178 The planktonic cells in groundwater were filtered from three liters of groundwater using a
179 90 mm nylon membrane with pore size 0.2 μm . Duplicate planktonic cell samples were collected
180 for each site and at each sampling campaign. Meanwhile, duplicate *in situ* “traps” were retrieved
181 from each well at the three- and six- month sampling campaigns. All the microbial samples (i.e.,
182 filter membranes and sediment traps) were individually packed in separate Whirl-Pak[®] sampling
183 bags (Nasco, WI, USA) and promptly transferred into a cooler filled with dry ice. On the same
184 day, all the microbiological samples were transported back to the laboratory at the Illinois State
185 Geological Survey, Urbana, IL. Upon arrival, they were transferred to a -20°C freezer and
186 stored for future analyses.

187

188 2.3 Geochemical Analyses

189 Geochemical parameters such as pH, specific conductance, dissolved oxygen, oxidation-
190 reduction potential (Eh) of the produced water were measured *in situ* (Dong et al., 2014b). A
191 total of 24 cations were quantified using US EPA method 200.7 (EPA, 1994). The concentration
192 of ferrous iron [Fe(II)] and sulfide were measured as described previously (Shao and Butler,
193 2007). The anions (i.e., Cl^- , Br^- , F^- , SO_4^{2-} , NO_3^- , PO_4^{3-}) were measured using a Dionx[™] ICS-
194 1500 ion chromatography (Thermo Fisher Scientific, MI, USA) (Pfaff, 1993). Alkalinity, total
195 dissolved solids, ammonium-N, dissolved inorganic carbon (DIC), and nonvolatile organic
196 carbon were analyzed using the methods SM 2320B, SM 2540C, US EPA 350.1, and SM 5310B,
197 respectively (Locke and Iranmanesh, 2015; Locke et al., 2018).

198

199 2.4 Molecular analyses

200 Genomic DNA was extracted from the filter membranes or a 0.5 g aliquot of sand from
201 the sediment traps using a modified phenol: chloroform extraction method (Dong et al., 2014a).
202 Unless mentioned, a negative control (sterile water) was put through the same extraction process
203 to ensure no contamination was introduced during DNA preparation. The quality and integrity of
204 the extracted DNA were assessed as described (Dong et al., 2014b).

205 Amplicon-sequencing was performed to understand the spatial and temporal distribution
206 of microbial communities and their metabolic potential. Based on the preliminary screening (see
207 section S1.2 in SI), a total of 12 pairs of primers targeting the functional genes associated with
208 microbial C, N and S metabolism and the 16S rRNA genes for both Bacteria and Archaea were
209 selected for multiplex PCR amplification (Table S2). The amplicons for the targeted ribosomal
210 RNA and the functional genes were amplified and barcoded using a Fluidigm Access Array and
211 BioMark™ HD System (Fluidigm Corporation, CA, USA) following the manufacturer's
212 recommendations (see section S1.3 in SI). Multiplex PCR amplification and sequencing were
213 conducted in the Roy J. Carver Biotechnology Center at the University of Illinois, Urbana-
214 Champaign. Bioinformatic and statistical analyses were performed as described in S1.4 of the SI.

215

216 2.5 Nucleotide sequence accession numbers

217 The sequences reported in this study have been deposited in NCBI GenBank with the BioProject
218 number PRJNA706996.

219

220 3. Results

221 3.1 Dynamic geochemical conditions in the sampling wells

222 Although the wells under investigation were not far away from each other (Fig. S1),
223 distinguishing geochemical properties were observed between them. Most of groundwater
224 constituents exhibited significant differences between the wells (ANOVA: $p < 0.05$) except for
225 sulfide, Fe(II), NO_3^- -N and non-volatile organic carbon (NVOC) (ANOVA: $p = 0.10-0.83$) (Fig. 1
226 and Table S3). All the groundwater samples exhibited circumneutral pH (Fig. 1a). Well 04UG
227 screened in the shallower Upper Glasford Formation showed significantly higher and dynamic
228 dissolved oxygen (DO) and redox potential (Eh) than the four bedrock wells (G101-G104) (Figs.
229 1b-1c). In addition, the concentrations of Ca^{2+} , Mg^{2+} , and SO_4^{2-} in the G101-G104 samples were
230 significantly lower than that in well 04UG, while S(-II) was lower in Well 04UG than the others
231 (Fig. 1g and 1h, Table S3). In terms of nitrogen compounds, relatively low concentrations of
232 NH_4^+ was consistently detected in all the wells, while NO_3^- was below detection limit in most
233 samples except for 04UG and G102 at sporadic time points (Fig. 1j and 1i). Spatial heterogeneity
234 in chemical composition was also observed between G101-G104, the proximate wells screening
235 the same geological formation (Fig. 1). In each of the four wells, the geochemical conditions
236 were relatively stable during the 6 months of monitoring period, except for dissolved sulfide and
237 Fe(II) concentrations (Figs. 1h-1i).

238

239 3.2 Distinct microbial structure and taxonomic composition associated with sample types

240 Alpha-diversity profiles based on Inverse Simpson Index (Invsimpson) showed that at the
241 same time point, the overall OTU diversity and evenness were significantly higher for the
242 suspended microbial communities than their attached counterparts (T-test, $p < 0.025$). In addition,
243 among all the wells, the shallower 04UG exhibited the most variable diversity for the suspended
244 microbial communities over six months. In comparison, significantly lower diversity and more

245 stable microbial communities with the Invsimpson values consistently less than 40 were
246 observed for the attached fraction in contrast to the highly variable suspended ones (Fig. 2).

247 The suspended and attached microbial communities exhibited distinct microbial
248 population composition. Although *Proteobacteria* were the most dominant phylum in both
249 communities, they were significantly enriched in the sediment matrix (Fig. S2). The orders
250 *Bacteroidales*, *Omnitrophales*, *Sphingobacteriales* and *Victivallales* along with several
251 *Proteobacterial* orders (e.g., *Methyloccoccales*, *Burkholderiales*, *Desulfobacterales*,
252 *Methylophiales*, *Gallionellales*) accounted for the majority of the suspended microbial
253 communities. The dominant orders in the attached samples, however, were exclusively affiliated
254 with *Proteobacteria* (e.g., *Desulfobacterales*, *Desulfobulbales*, *Geobacterales*, *Methyloccoccales*
255 and *Desulfuromonadales*) (Fig. 3). Due to the limitation of the pore size, our work were focused
256 on the microorganisms with typical size (several hundred nanometers to several micrometers).
257 But the organisms in lower nanometer size (e.g., CPR bacteria and DPANN archaea) detected in
258 earlier studies on other aquifers (He et al., 2021; Luef et al., 2015) were potentially overlooked.

259 Comparison of the 20 most dominant OTUs in the suspended and attached microbial
260 communities showed distinct patterns depending on sample fraction (Fig. 4). Only three OTUs
261 (i.e., an unclassified *Desulfobulbaceae*, a *Methylobacter*, and an unclassified *Methylococcaceae*)
262 were among the major phylotypes (>1 %) in both sample fractions, while the others were
263 significant in only one sample fraction. For example, *Sulfurospirillum*, *Desulfocapsa*, *Geothrix*
264 and multiple *Geobacter* spp. were the major OTUs in the attached microbial communities, while
265 *Methylococcaceae*, *Methylophilus*, *Sideroxydans* and *Methylovolum* were among the most
266 significant free-living phylotypes (Fig. 4). In addition, cumulatively relative abundance for the
267 core OTUs that were detected in at least half of the samples was significantly higher in the

268 attached microbial communities compared to that in the suspended ones (Fig. S3 and Section
269 S2.1 in the SI). This suggested that the shared core phylotypes made up the majority of the
270 populations in the attached microbial communities, while the transient ones accounted for the
271 majority in the suspended communities.

272

273 3.3 Spatial and temporal microbial distribution

274 The microbial communities inhabiting the wells showed different extents of dissimilarity
275 depending on the sample source. The shallower 04UG communities were most dissimilar from
276 those found in wells G101-G104, which was supported by a Bray-Curtis analysis model (Figs. 5a
277 and S4). The samples from 04UG were 85.88-89.53 % (suspended) and 73.6-77.06 % (attached)
278 in dissimilarity from those derived from G101-G104 (Fig. S4). In comparison, lower
279 dissimilarities of 63.88-75.92 % (suspended) and 55.74-63.67 % (attached) were observed
280 among the samples taken from wells G101-G104 completed in the same formation (Fig. S4). It
281 should be noted that although well G104 was somewhat distant from the wells G101-G103, the
282 community dissimilarity between these samples was comparable.

283 The changes in microbial community composition after three and six months were
284 observed within each well. The greatest change was observed in the shallower 04UG well,
285 consistent with the highly variable alpha diversity in this well (Figs. 2). Using SIMPER analysis,
286 the contribution of the specific populations to the quarterly changes in the microbial community
287 composition in each well was assessed and the major taxa that contributed to >1 % of the
288 observed change were identified. These major taxa contributed 34.6-57.5 % to the changes at
289 different time points seen with the suspended microbial communities, compared with 43.1-
290 67.7 % in the attached ones (Tables S5 and S6). Specifically, the prevailing OTUs in the attached

291 microbial communities (e.g., *Geobacter* and unclassified *Deltaproteobacteria*) are among the
292 major taxa contributing the greatest to the changes in the biofilm (Table S6).

293

294 3.4 Environmental drivers for subsurface microbial structure along space and time

295 Canonical correspondence analysis (CCA) showed that generally the microbial
296 communities were clustered based on the sampling type and well (Fig. 5). The suspended and the
297 attached microbial communities from the shallower and more oxidized 04UG well samples were
298 significantly different from those for the other wells (PERMANOVA, $p < 0.05$). For the
299 suspended microbial community composition, factors such as SO_4^{2-} , DO, Eh and nitrate were
300 important for well 04UG, while Na^+ , Cl^- and S(-II) were most important for wells G101-G103
301 (Fig. 5b). Comparatively, similar driving factors were observed for the attached microbial
302 communities except that the deeper G101-G103 were also influenced by Fe(II) and alkalinity
303 (Fig., 5c).

304

305 3.5 Functional genes

306 Among the 12 functional genes tested, *dsrB*, *pmoA*, *aprA*, and *nifH* associated with
307 cycling of C, N, S in the aquifer ecosystems were selected for further analyses (see Section S2.2
308 in SI). Analyses of these functional gene sequences showed that the majority were assigned to a
309 relatively small number of OTUs. For example, the top 5 OTUs of *dsrB* and *nifH* genes
310 comprised >69.5 and 69.7 % of the sequencing reads for the suspended and attached fractions,
311 respectively. The least diversity was observed with *pmoA*, with 5 and 6 OTUs comprising ~95 %
312 of the sequencing reads in the suspended and attached samples, respectively. The closest
313 phylogenetic references identified for these dominant genotypes illustrated the diversity of

314 potential functional populations found in different wells or at different time points (Figs. 6, S5-
315 S7).

316 Functional gene composition changed differentially between wells and with time. For the
317 functional genes identified for the suspended microbial communities, the pair-wise dissimilarity
318 for samples derived from the same well ranged from 36.6 to 93.7 % with a relatively stronger
319 effect along time changes observed for *aprA* and *dsrB*. In contrast, the changes for the attached
320 microbial communities between three and six months were generally lower than the suspended
321 counterparts (Fig. 7). A different pattern was observed for *pmoA*. A stable composition with less
322 than 40 % dissimilarity was observed for all the suspended microbial communities, lower than
323 any other functional gene. However, *pmoA* gene composition was more variable in the attached
324 communities for most of the wells, with the highest dissimilarity index 82.6 % in Well 04UG
325 (Fig. 7). Further assessment of *pmoA* genes indicated that more than 90 % of the changes were
326 derived from the large change in abundance of the most dominant OTUs instead of the overall
327 gene-variant composition.

328 PCoA analyses, based on the Bray-Curtis dissimilarity model, between the samples
329 showed distinct shifts in the functional genes based mainly by when and where they were taken,
330 or the fraction under study (Fig. 8). Overall, the distribution of the functional genes exhibited
331 significant difference between the sampling wells (PERMANOVA, $F=3.2-11.0$, $p<0.0002$).
332 Specifically, the 04UG samples exhibited less similarity with the other wells for all the 4
333 functional genes in both suspended and attached phases, suggesting its relative independence
334 (PERMANOVA, $p<0.05$) (Fig. 8). Interestingly, the dissimilarity in the functional genes
335 between the suspended and attached microbial communities showed different patterns. For *aprA*,
336 *dsrB* and *nifH*, the functional gene composition for wells (G101-G104) were generally separated

337 by the sample type (Figs. 8a-8c). Unlikely, the low diversity *pmoA* gene was sample-dependent
338 as the suspended and attached phases from the same well were more closely associated than the
339 sample-based difference (Fig. 8d).

340

341 3.6 Ecological drivers of microbial community and potential functional population assembly

342 To understand the assembly processes that determine the formation of local communities,
343 the distribution of microbial populations and potential community functionalities (i.e. designated
344 by distribution of functional genes), iCAMP, a phylogenetic bin-based null model was applied
345 (Ning et al., 2020). This method overcomes the constraints imposed by classical statistical
346 approaches based on the pairwise turnovers of the whole community composition. Instead, the
347 iCAMP model differentiates the microbial groups/functional genes on finer biological
348 organization levels (Ning et al., 2020). Based on 16S rRNA gene sequence analysis, the
349 deterministic process homogenous selection was the most significant and was responsible for
350 69 % and 83 % of the suspended and the attached community composition, respectively.
351 Notably, higher contributions from the stochastic process of dispersal limitation was suggested
352 for the suspended microbial communities than the attached ones (22 % vs. 4 %) (Fig. 9a).

353 For the functional genes, homogenous selection was much less than that observed for the
354 overall community assembly based on 16S rRNA gene sequence analysis. The exception to this
355 was *nifH* in the attached samples (Fig. 9). For *dsrB*, *pmoA* and *aprA*, relative contribution of
356 homogeneous selection was typically less than 30 %. Instead, the stochastic processes (e.g., drift
357 and others) played significant roles for assembly of the functional gene pools. While
358 significantly higher contribution was observed for the suspended samples compared with the
359 attached ones for *dsrB* and *nifH*, similar contribution was identified between the two fractions for

360 *pmoA* and *aprA* (Figs. 9b-9e). The importance of dispersal limitation varied depending on the
361 functional genes, with the highest contribution of this process observed for *dsrB* in the attached
362 samples (Fig. 9b). Comparable contribution by dispersal limitation was observed for *nifH* in the
363 suspended samples and *aprA* gene-communities in both fractions (41-46 %) (Figs. 9c and 9e). In
364 contrast, this process did not appear to be an important driver for assembly of *nifH* genes in the
365 attached samples and *pmoA* gene pools in both fractions (Fig. 9c).

366 Stochasticity was calculated to assess the impacts by the stochastic processes (e.g.,
367 birth/death, speciation/extinction and immigration) (Ning et al., 2020). The stochasticity was
368 much lower for the 16S rRNA genes than the functional genes. In terms of individual sampling
369 well, homogenous selection contributed most for assembly of the microbial communities with
370 the exception for 04UG in that dispersal limitation played comparable role (Fig. S8a).
371 Comparatively, drift and other processes were consistently important for assembly of the
372 functional gene reservoirs, while the significance of dispersal limitation or homogenous selection
373 varied depending on the types of the gene (Fig. S8b-S8e).

374

375 **4. Discussion**

376 An understanding of the stability, functionality, resilience and assembly of subsurface
377 microbiomes is critical for both fundamental ecological questions and engineering applications
378 (Fields et al., 2006). The integration of geochemistry, multiplex PCR, high-throughput
379 sequencing of amplified genes and statistical analyses used in the present study provide a
380 comprehensive investigation of a few heterogeneous and dynamic shallow aquifer ecosystems.
381 Our work presents fine resolution for the shallow aquifer microbiomes in that suspended and
382 attached microbial communities distinctively respond to the varying degrees of geochemical

383 fluctuation over space and time, even when they are in close proximity. The most similar work
384 was the study on the shallow and pristine confined aquifer of US DOE ENIGMA-Oak Ridge
385 Field Research Center, which provides valuable information about the high spatiotemporal
386 variability at the scale at tens of meters (Zelaya et al., 2019). With the focus solely on the
387 suspended microbial communities, however, this earlier work did not show the distinct patterns
388 for the attached microbial communities. In a series of recent studies on the Hainich Critical Zone
389 Exploratory (CZE) in central Germany, differential microbial structure and functional
390 populations between groundwater, fracture surface and rock matrices were identified, which
391 provide insights of the fraction-dependent microbial communities dwelling aquifer (Lazar et al.,
392 2019; Schwab et al., 2017; Wegner et al., 2019; Yan et al., 2020). The linkage between microbial
393 structure, functional groups and the metabolic potential suggested in our study underscores the
394 significance to consider both suspended and attached microbial communities. This will allow
395 more accurate understanding of the subsurface ecosystems and provide means to predict the
396 biogeochemical processes.

397

398 4.1 Distinct planktonic and attached microbial distribution

399 The significantly different suspended and attached microbial communities in terms of
400 diversity, composition, and features (e.g., core and prevailing taxa) (Figs. 2-5a, S2-S4) may be
401 shaped by the unique micro-environments in the two fractions. Typically, pristine subsurface
402 habitats are oligotrophic with limited carbon, nitrogen and energy sources in groundwater (Akob
403 and Küsel, 2011; Hazen et al., 1991; Schwab et al., 2017; Wegner et al., 2019). However, flow
404 conditions, nutrient availabilities and motility/dispersal can create distinct micro-environments
405 for the free-living and the biofilm-associated organisms. In comparison, the capacities to retain

406 water, sorb nutrients and protect cells against rapidly changing geochemistry by biofilm establish
407 superior habitat for colonization of actively growing organisms and mutualism compared with
408 those characterized in groundwater for long-time scale (Stoodley et al., 2002).

409 Our observations of partial overlapped but distinctly distributed free-living and surface-
410 associated species are in line with other work on subsurface ecosystems (Flynn et al., 2008;
411 Flynn et al., 2013; Zhou et al., 2012), *in-situ* bioreactors (Christensen et al., 2018; King et al.,
412 2017), and microcosms (Barnhart et al., 2013). Similar with the reports by Flynn et al. (2008 and
413 2012), no more than 7.8 to 14.1 % of the OTUs are shared between the fractions but account for
414 25-67 % of the sequence abundance in our work. In addition, the major members for the attached
415 communities are strikingly low in the groundwater and *vice versa*. Such preferential enrichment
416 of specific species in different fractions of the subsurface agrees with a metagenomic study at the
417 U. S. DOE uranium-contaminated shallow groundwater remediation test site in Rifle, CO, in
418 which the dominant suspended phylotypes are only the rare species in the attached fraction (Fig.
419 4) (Hug et al., 2015). Interestingly, in the earlier studies conducted on the adjacent Mahomet
420 Aquifer, a confined pristine aquifer located about 20 miles northeast of our site, the suspended
421 and attached microbial communities exhibited comparable microbial diversity (Flynn et al.,
422 2008; Flynn et al., 2013; Flynn et al., 2012). This inconsistency with our observation of lower
423 diversity in attached microbial communities than that in the suspended counterparts (Fig. 2 and
424 3) is unlikely due to insufficient colonization (see section S3.1 in SI). Instead, the geological
425 formation in these earlier studies (e.g., Banner Formation) differs from the Glasford and Upper
426 Pennsylvanian Formations in this work, where the geochemistry and microbial structure are not
427 expected to be identical. In addition, as demonstrated in many earlier studies, a significant
428 fraction of the aquifer microbiomes is attached to the sedimentary porous media rather than in

429 the free-living phase (Alfreider et al., 1997; Flemming and Wuertz, 2019; Griebler and Lueders,
430 2009; Hug et al., 2015). If such scenario occurs in our systems too, the main phylotypes of the
431 less diversified attached communities are generally more preferentially enriched in the whole
432 system.

433 Assessment of the preferentially enriched taxa in the suspended versus attached microbial
434 communities suggests that the two fractions may exert varying biogeochemical roles. For
435 example, the phylotypes of known methanotrophs (e.g., *Methylophilus*, *Meliobacter*,
436 *Methylomonas*, and *Methylovulum*) that oxidize C1 compounds (e.g., methane and/or methanol)
437 (Iguchi et al., 2011; Locke et al., 2018; Oshkin et al., 2016; Warttinen et al., 2006) are among
438 the prevailing and core ecotypes in the free-living microbial communities. The *Sideroxydans*-
439 affiliated OTU is phylogenetically related with the species capable of oxidizing iron with natural
440 humic acids or using CO₂ as the carbon source (Hädrich et al., 2019; Wessel et al., 2014). In
441 comparison, the major attached taxa show close phylogenetic relation with some known sulfur
442 metabolizers growing on sulfate, thiosulfate and/or sulfite (e.g., *Desulfobacterales*,
443 *Desulfobulbaceae*, *Desulfocapsa*, and *Desulfobacula* spp.) (Muyzer and Stams, 2008).
444 Additionally, the potential metal-reducers (e.g., *Geobacter*, *Geothrix* and *Rhodoferax* spp.)
445 (Finneran et al., 2003; Lovley and Walker, 2019; Mehta-Kolte and Bond, 2012) are among the
446 major phylotypes in the attached fraction. The potential fraction-based biogeochemical processes
447 suggested by uneven microbial distribution also imply geochemical gradient between
448 groundwater and biofilm.

449

450 4.2 Spatial and temporal changes of the subsurface microbiomes

451 Our observations of spatial and time-dependent variation in microbial communities

452 inhabiting proximate subsurface niches are reflective of the intimate interactions between
453 groundwater and the mineral components in the sediments. These interactions create physical
454 and chemical variation ranging from micrometer to kilometer scales, and thus mediate the
455 biogeochemical processes (e.g., weathering, nitrogen cycling and carbon cycling) driven by
456 groundwater microbiomes (Akob and Küsel, 2011; Griebler and Lueders, 2009). Unlike the
457 contaminated sites with elevated concentrations of metals/metalloids, halogenated solvents or
458 hydrocarbons (Fields et al., 2006; He et al., 2018; Xu et al., 2010), the environmental variation in
459 pristine subsurface ecosystems primarily comprise of natural geochemical/biogeochemical
460 changes. Studies have demonstrated that the hydrochemical factors (e.g., groundwater residence
461 time, geochemistry, nutrient supply, dispersal of microbes and pulse of seasonal recharge) can
462 create spatially heterogenous and temporally changing environmental conditions in pristine
463 subsurface ecosystems. These can directly and indirectly influence the structure, metabolic
464 potential and activity of the organisms residing these habitats (Griebler and Lueders, 2009;
465 Gulmann et al., 2015; Hug et al., 2015; Lin et al., 2011; Nemergut et al., 2013; Smith et al.,
466 2018; Wegner et al., 2019; Zelaya et al., 2019; Zhou et al., 2012).

467 The heterogeneous and dynamically changing suspended microbiomes observed in this
468 study are reflective of the complex subsurface environmental conditions derived from the
469 introduction of nutrients, changes of water geochemistry or migration of microbial populations
470 from the surface flux recharge or between geological strata (Smith et al., 2018). For example,
471 pronounced water level changes induced by hydrological dynamics in groundwater can lead to
472 seasonal patterns of the sediment microbial communities in the proximate oligotrophic porous
473 aquifer (Lin et al., 2011; Zhou et al., 2012). In this study, wells G101-G104 monitor groundwater
474 within a relatively confined geological stratum, which has not seen much influence from recent

475 meteoric water recharge and has limited water transmissivity with the overlaying horizons. In
476 comparison, the shallower 04UG well with oxygenated water is probably subject to active
477 recharge from the surface or the overlying Tiskilwa Formation (Locke et al., 2018; Watson et al.,
478 2020). In addition, although wells G101-G104 monitor the same geological formation and are
479 similar in the depth, the distinct geochemical properties for G104 (Table S3) suggest that it is
480 likely more influenced by hydraulic connections to shallower formation(s) (Wimmer et al.,
481 2021). For G101-G103 that are only tens of meters distant with each other, the differing
482 hydraulic conductivity may explain heterogenous groundwater retention time and water-rock
483 interactions associated with the variable geochemical conditions (Dong et al., 2014b; Wegner et
484 al., 2019; Zelaya et al., 2019).

485

486 4.3 Functionalities for the subsurface microbiomes as suggested by the functional genes

487 Relationship between the functionalities and microbial community structure has been an
488 important ecological question for subsurface ecosystems. The functionalities of the microbiota
489 shaped by inter-related microbial pools, nutrient input, energy resource and the ambient
490 environmental conditions are implied in this work. The consistent presence of *nifH* genes
491 suggests that N₂ fixation is an important nitrogen source. Our observations that NO₃⁻
492 concentrations are below detection limit in most samples, and low concentrations of NH₄⁺ were
493 consistently detected in all the samples (Fig. 1) differ from other aquifer ecosystems with active
494 NO₃⁻ or NH₄⁺ introduction (e.g., agricultural and contaminated sites), where nitrification,
495 denitrification, dissimilatory nitrate reduction to ammonia (DNRA), or ammonia oxidation were
496 often identified (Küsel et al., 2016; Stoliker et al., 2016; Wells et al., 2018). For example, the
497 aquifers with different hydrochemical settings and land use along the Hainich groundwater

498 transect in Germany have groundwater microbiomes that are mainly driven by nitrification,
499 denitrification and anaerobic ammonia oxidation [anammox], while N₂ fixation is marginal
500 (Wegner et al., 2019). Field and laboratory incubation studies also suggest that under anoxic
501 conditions and with limiting labile organic compounds, anammox and denitrification potentially
502 exist, while nitrification may be enhanced when oxic eutrophic surface water is introduced.
503 Under carbon-rich conditions, DNRA would become predominant once O₂ was depleted
504 (Schwab et al., 2017; Stoliker et al., 2016; Wegner et al., 2019). In our study, there was no
505 evidence for active NO₃⁻ or NH₄⁺ introduction from the overlaying formations. Thus, N₂ fixation
506 may outcompete other N-metabolizing processes in the deeper and N-limiting environments to
507 form more bioavailable N substrates.

508 The consistent and abundant functional genes of sulfate reduction (e.g., *dsrB* and *aprA*)
509 and methane oxidation (*pmoA*) suggest these two processes may be among the energy sources at
510 the IBDP site. In the Decatur area, it has been documented that drift gas containing 22.5-99.5 %
511 methane formed in the glacially deeply buried soils from decay of peats and organic-rich silts.
512 Since these glacial deposits reside within a confined saturated formation, they are anaerobic and
513 thus favorable for methanogenesis. Significant methane levels were measured from vacuum
514 wells at locations near the IBDP site (Meents, 1960). In Mahomet Aquifer located about 20 miles
515 northeast of our studied site, variable concentrations of methane were reported in the shallow
516 groundwater wells (Flynn et al., 2008; Flynn et al., 2013; Flynn et al., 2012; Yamamoto et al.,
517 2019). Coal solids are also found in the Pennsylvanian Formation and may provide subsurface
518 microbial communities with organic carbon to generate CH₄ (Adams et al., 1984; Strapoc et al.,
519 2008). These suggest that methane production is common in the glacial deposit of Decatur area.
520 The lack of *mcrA* PCR products in the preliminary experiment to assess feasibility of the primer

521 sets (see section S1.2 in the SI), detection of *pmoA* genes, and prevalence of methane oxidizers
522 in the suspended microbial communities suggest that aerobic methane oxidization may be critical
523 to explain the fate of methane, even in the anoxic wells.

524 Different strategies are implied by the suspended versus attached microbiota to maintain
525 the overall functional stability. Earlier studies suggested that the microbiome functionalities in
526 groundwater are stable and may be uncoupled from microbial community structure due to
527 functional redundancy (Fields et al., 2006; Zelaya et al., 2019). With the attached communities in
528 our study, however, the structure and richness for the functional genes follow generally similar
529 stability (Fig. 6-7, S5-S7). Therefore, we propose that there are possible resiliency in
530 functionality by the suspended communities, while the conserved functionalities for the attached
531 biofilm communities are more associated with the relatively stable microbial structure. This can
532 be explained from several perspectives. Firstly, the source of microbial populations may vary
533 between the free-living and attached communities. As the free-living communities represent the
534 microbial membership from a much larger physical environment than the sediment, the
535 possibility to introduce the new taxa with more diversified genetic traits with the flowing
536 groundwater than the selectively colonized biofilm species is expected (Flynn et al., 2008; Flynn
537 et al., 2013; Hug et al., 2015). Secondly, the resistance to the external perturbation between the
538 two fractions differs. Disturbance by external chemical stressors may be milder for the attached
539 microbial cohorts than those within the groundwater. Thirdly, the attached microbial
540 communities may form biofilm that provides stability for cells, mediates surface adhesion, and
541 serves as a scaffold for cells, enzymes, and antibiotics to attach (Algburi et al., 2017; Flemming
542 and Wingender, 2010). The suspended microbiota, in contrast, are void of such protection and

543 may be more vulnerable to external disturbances (Ben Maamar et al., 2015; Flynn et al., 2008;
544 Zelaya et al., 2019).

545

546 4.4 Influential ecological processes for community and functioning assembly

547 It is well accepted that both stochastic and deterministic processes operate simultaneously
548 in the assembly of local communities in terms of diversity, distribution and succession (Ning et
549 al., 2019; Stegen et al., 2013; Stegen et al., 2015). The prevailing contribution by homogenous
550 selection on the community assembly and the variable stochasticity for microbial consortia
551 derived from different wells (Fig. S8) are in line with the previous aquifer studies covering wide
552 geographical area (e.g., Hainich Critical Zone Exploratory in Germany, the Licking Aquifer and
553 ENIGMA-Oak Ridge Field Research Center, TN in U. S.) (Danczak et al., 2018; Fillinger et al.,
554 2019; Herrmann et al., 2019; Yan et al., 2020; Zelaya et al., 2019). In these earlier studies,
555 increasing selection stress and stronger competition cost for colonization were generally
556 associated with the aquifers with lower flow velocity. In contrast, in the areas with higher flow
557 velocity, hydraulic conductivity, nutrient input or uniform hydrochemistry, stochastic processes
558 dominate and typically result in more diversified microbial communities with higher turnover
559 rates (Danczak et al., 2018; Hädrich et al., 2019; Yan et al., 2020). Therefore, it was
560 hypothesized that more interconnected communities experiencing lower turnover are susceptible
561 to homogenizing selection. In contrast, the less inter-connected microbial communities
562 experiencing higher turnover are more likely affected by stochastic processes (Danczak et al.,
563 2018). In addition, our observation of more significant contribution by dispersal limitation to the
564 suspended microbial communities than the attached ones (Fig. 9a) suggests proliferation of
565 specific taxonomic groups in the attached fraction induced by fine-scale hydrodynamic niche

566 differentiation and geochemistry (Besemer et al., 2012; Besemer et al., 2009). As active
567 microbial competition in porous medium can select against biofilm formation and result in
568 dispersal limitation (Coyte et al., 2016), this stochastic biological process may not be significant
569 for attached microbial consortia.

570 Our work also provides finer resolution about the patterns for assembly of potential
571 functional populations based on the functional gene analyses, which is not evident in the overall
572 community composition. The strong stochasticity for functional genes (Figs. 9 and S8) is
573 supported by the “competitive lottery model” that ecological niches are colonized randomly from
574 a pool of species with similar ecological function. In this hypothesis integrating both neutral and
575 functional aspects, available space within the niche is expected to be colonized by whichever
576 suitable species happens to arrive and inhabit first (Burke et al., 2011; Munday, 2004). Based on
577 the model, the functional populations can be reflective of the taxonomic groups. Alternatively, a
578 given function may be broadly distributed across a variety of phylotypes instead of a specific
579 taxonomic group (Burke et al., 2011; Munday, 2004). In a previously published work on the
580 microbial communities associated with a group of marine green macroalgae *Ulva australis*, the
581 specific functional genes were assigned to a broad range of bacteria, which differed from sample
582 to sample. It suggests that the core functions are not restricted to the particular taxonomic groups
583 (Burke et al., 2011). Such scenario occurs for some of our OTUs for the functional genes under
584 study. Other OTUs, however, are shared between samples or sample types (Figs. 6, S5-S7). It
585 suggests that the taxa with specific trophic or ecological properties are able to occupy the niches
586 within an ecosystem. Whoever gets there first wins the “lottery” for that specific space (Munday,
587 2004). However, their changes overtime may reshape the potential functional populations within

588 the niches driven by random components (e.g., dispersal limitation, drift or other processes) as
589 determined (Figs. 9 and S8).

590

591 **5. Conclusions and Environmental Significance**

592 Heterogeneous microbial distribution and changes along different time points in response
593 to the environmental disturbance at the high turnover rates are observed within systems tens of
594 meters apart and at the frequency of months. While distinct microbial structure and diversity
595 between groundwater versus sediment fractions have been widely observed in the subsurface
596 ecosystems (Flynn et al., 2008; Flynn et al., 2013; Smith et al., 2018), this work further suggests
597 that the difference in possible repercussion for stability and resilience in terms of biochemical
598 potential may vary between specific functional populations and metabolic processes. More
599 importantly, identification of different controlling ecological processes for assembly of the
600 microbial communities and functional genes implies considerable biological patterns that cannot
601 be directly inferred from the species diversity or the free-living microbial consortia (Burke et al.,
602 2011; Schwab et al., 2017). Therefore, the weight of the free-living and attached microbial
603 communities to the overall biogeochemical processes needs careful consideration and assessment
604 when studying the aquifer microbiomes. As quantitative analyses have not been included in the
605 present study, the follow-up work will be focused on applying qPCR and other meta-omic
606 approaches to quantitatively analyze the fraction-dependent biomass distribution, active
607 populations and expressed functional genes to explore the active microbe-environment
608 interactions in terms of the changes along space and time.

609 In addition to the fundamental significance, this work also presents potential instruction
610 for practical applications, such as biomonitoring, bioremediation and bioaugmentation.

611 Systematic understanding of the fraction-dependent microbial distribution, the physiological
612 potential and their assemblies will facilitate the determination of critical factors that relate the
613 community structure to the functions beyond what is just available from the simple analyses
614 (King et al., 2017). For example, the suspended microorganisms are more sensitive to
615 geochemical changes and thus can be applied as the indicators for natural or anthropogenic
616 environmental perturbation (e.g., contaminant leakage). Meanwhile, remediation of regulated
617 metals/metalloids may be relied on the development of attached metal-metabolizing populations.
618 In conclusion, a comprehensive understanding about subsurface microbiomes will facilitate
619 designing of the appropriate engineered strategies and hydrobiogeochemical models that instruct
620 contaminant assessment, monitoring, and treatment.
621

622 **6. Acknowledgement**

623 The Midwest Geological Sequestration Consortium is funded by the U.S. Department of
624 Energy through the National Energy Technology Laboratory (NETL) via the Regional Carbon
625 Sequestration Partnership Program (contract number DE-FC26-05NT42588) and by a cost share
626 agreement with the Illinois Department of Commerce and Economic Opportunity, Office of Coal
627 Development, through the Illinois Clean Coal Institute. YD and LS were supported by National
628 Natural Science Foundation of China under the contract 41877321, 92051111 and 91851211. We
629 appreciate the technical support from Mark Band (University of Illinois Urbana Champaign) and
630 instructive discussion with Ye Deng (Research Center for Eco-Environmental Sciences, Chinese
631 Academy of Sciences, China) and Chris Field (University of Illinois Urbana Champaign).

632

633 **Supporting Information Available**

634

635 Table 1. Summary of the shallow groundwater monitoring wells under investigation

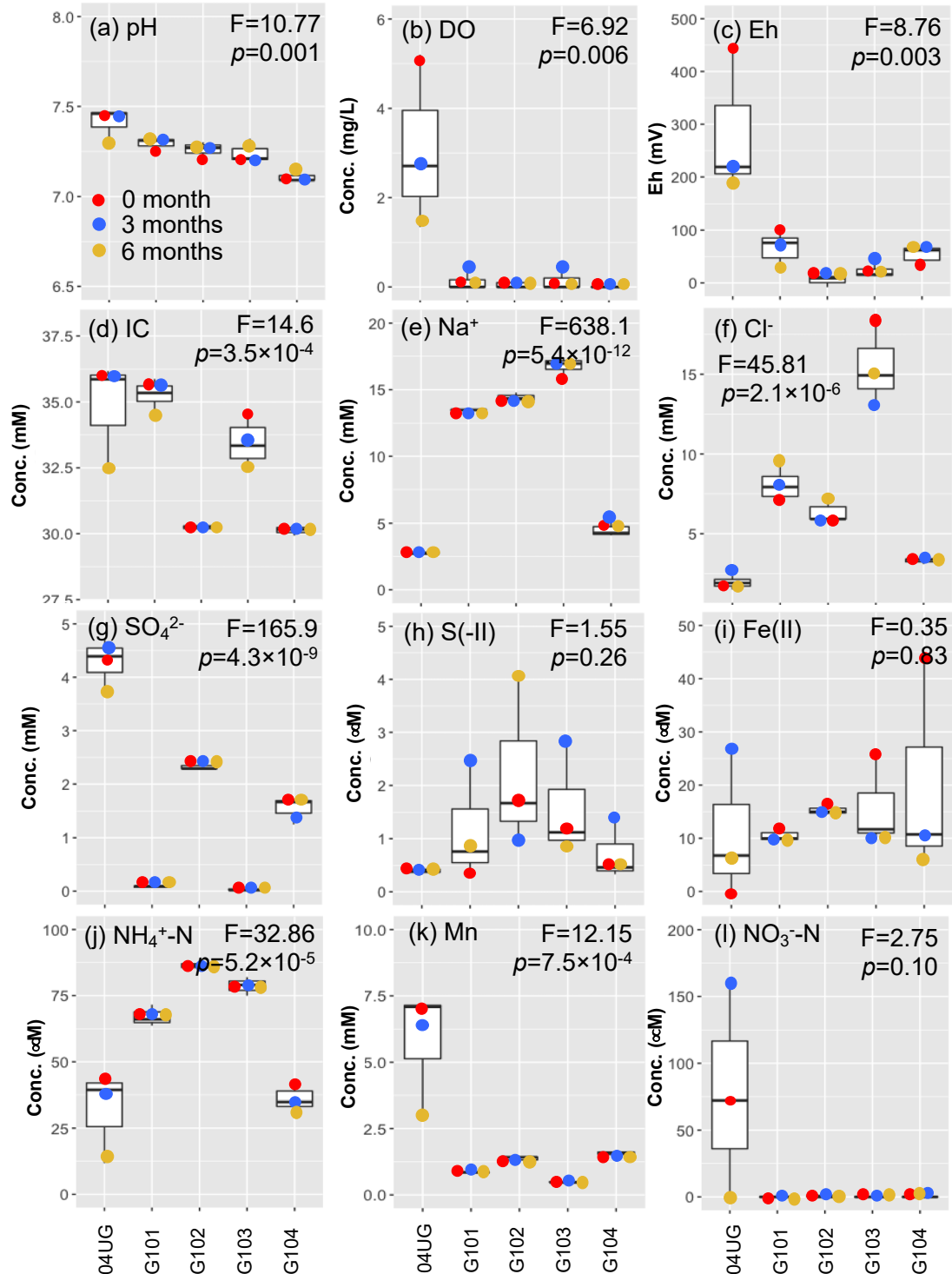
	04UG	G101	G102	G103	G104
Monitored formation	Glasford	Upper Pennsylvanian bedrock	Upper Pennsylvanian bedrock	Upper Pennsylvanian bedrock	Upper Pennsylvanian bedrock
Depth (m)	18.6-20.1	41.6–43.1	41.9–43.4	41.6–43.2	41.0–42.6
Surface elevation (m)	207.8	205.2	205.4	205	207.8
Groundwater level (m) ^a	192-202 ^b	186-189	186-189	187-189	190-192
Hydraulic conductivity	5.89×10 ⁻⁶ - 6.64×10 ⁻⁵	4.11×10 ⁻⁶ - 1.37×10 ⁻⁵	^c	1.69×10 ⁻⁶ - 3.51×10 ⁻⁵	5.09×10 ⁻⁷ - 3.51×10 ⁻⁶
Sediment mineralogy (%) ^d	Clay	3.9-5.3		8.8-12.2	
	Quartz	55.5-58.5		76.0-78.2	
	K-feldspar	5.1-6.8		3.3-4.4	
	P-feldspar	4.0-6.2		5.3-12.9	
	Calcite	3.3-4.7		2.0-2.1	
	Dolomite	16.9-22.8		0.9	
	Siderite	0.7-2.7		1.9-3.4	
	Pyrite	0.9-1.2		1.0-1.6	

636 ^a. Groundwater level elevation was based on the monitoring since February, 2009 (for 04UG) and July, 2010 (for
637 G101, G102, G103, and G104);

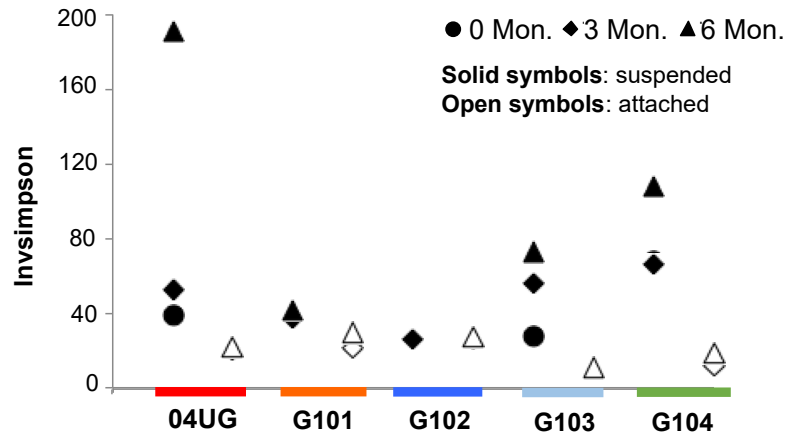
638 ^b. Due to compromised well protector, surface runoff entered the wells and resulted in significant (~10 m)
639 fluctuation in groundwater level;

640 ^c. Hydraulic conductivity was not estimated for compliance well MVA-G102 because of logistical issues in
641 conducting the slug and bailer tests;

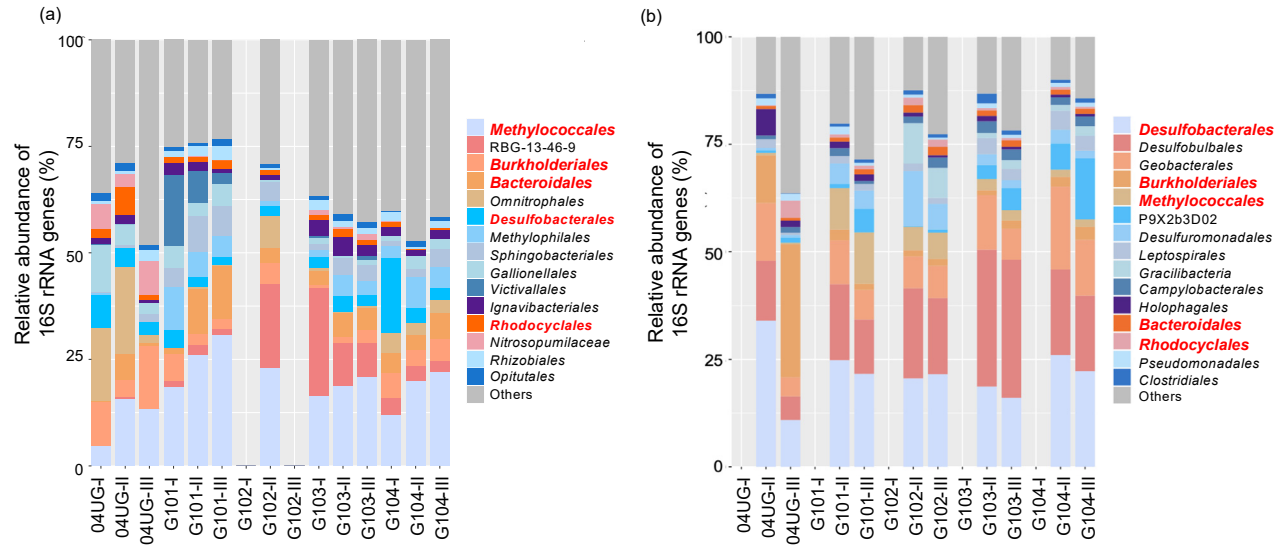
642 ^d. The sediment mineralogy was reported as the range for the wells targeting the same geological formation.



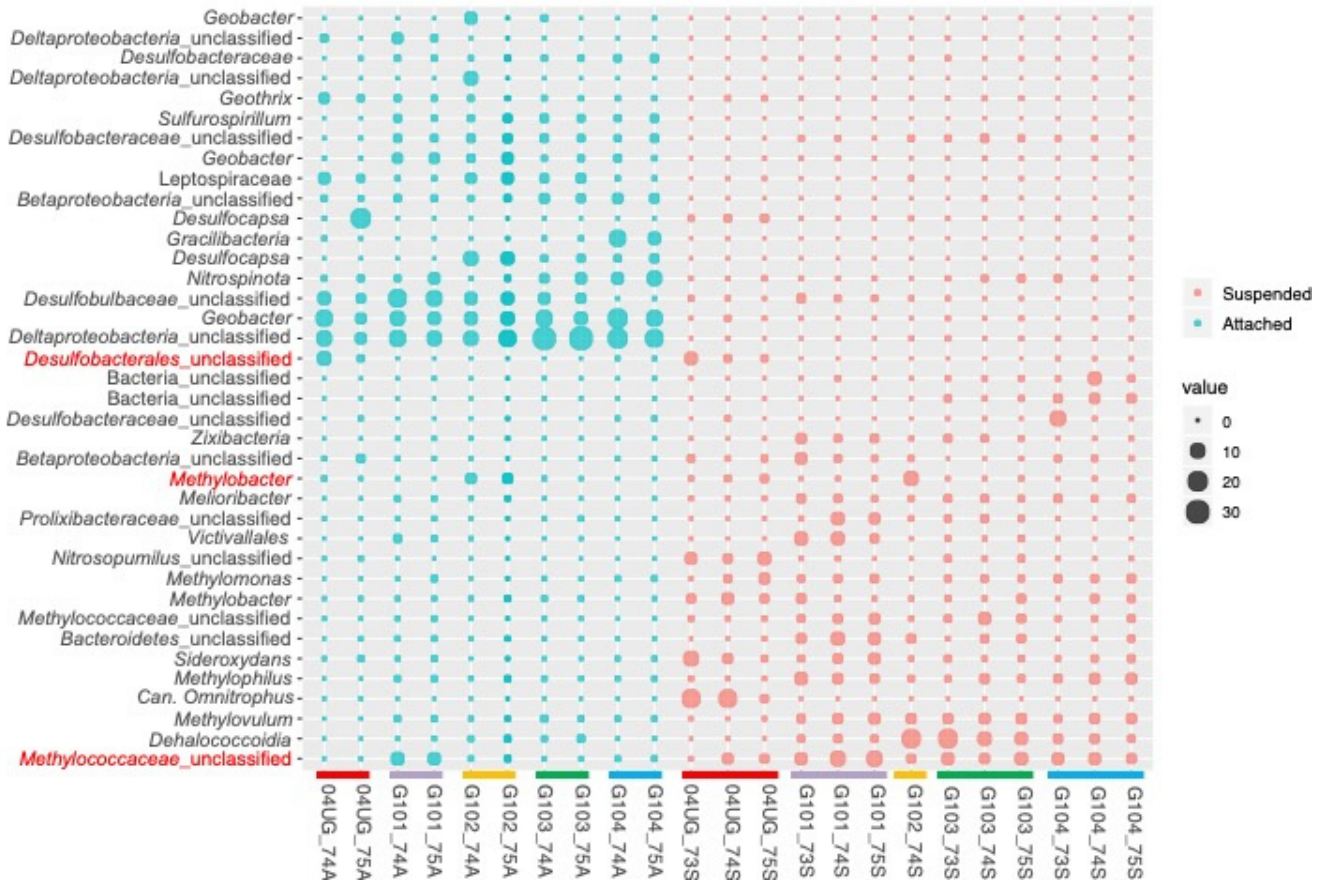
643
 644 Figure 1. Boxplots of temporal changes in within-well geochemistry for 3 sampling time points in
 645 the 5 wells. The median values were denoted by the dark horizontal line within each box. The
 646 variation between different wells for each geochemical parameter was calculated using ANOVA
 647 and the F and p values for the analysis were listed at the top right corner in the subfigures. The
 648 individual value was presented as the colored circles with red, blue and yellow indicating T1, T2
 649 and T3 that were 0, 3 and 6 months, respectively.



650
 651 Figure 2. Alpha diversity based on 16S rRNA genes for the suspended and attached microbial
 652 communities derived from the different sampling wells. Inverse Simpson values were calculated
 653 based on the high throughput sequencing of the V4 hypervariable region of the 16S rRNA genes.
 654 The circle, diamond and triangles indicated the three sampling timepoints at 0, 3 and 6 months,
 655 respectively. No sequences were obtained in some of the samples (e.g., the suspended samples
 656 from G102 at time 0 and 6 months and all the time 0 attached samples) and thus were not
 657 included.



658
 659 Figure 3. The top 15 orders of microorganisms identified in the suspended (a) and attached (b)
 660 microbial communities based on high-throughput sequencing of the V4 hypervariable region of
 661 the 16S rRNA genes. “Other” included the OTUs with relatively minor orders or the ones that
 662 could not be classified at the order level. RBG-13-46-9 is affiliated to Class *Dehalococcoidia* of
 663 Phylum *Chloroflexi*. The orders labelled in red were shared by the attached and suspended
 664 microbial communities. Void of the histograms for some samples meant the biomass was too low
 665 to extract enough DNA for PCR amplification. For the attached microbial communities, the
 666 average values for the duplicate samples were presented.
 667



669

670

671

672

673

674

675

676

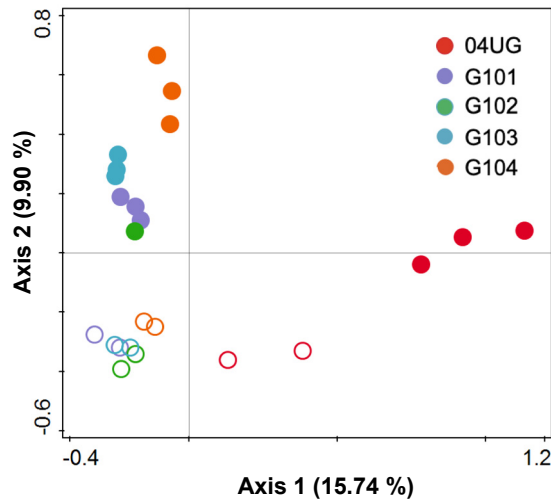
677

678

679

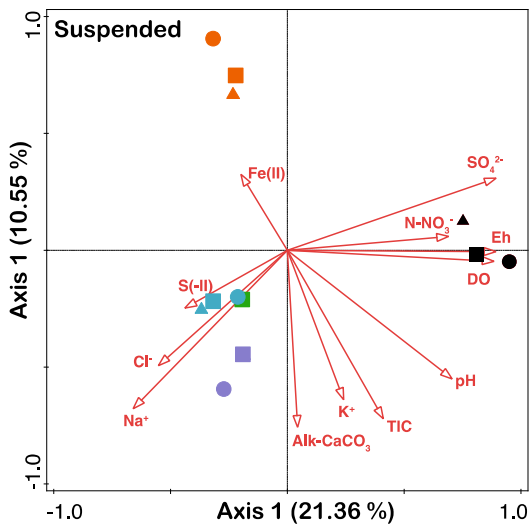
Figure 4. The distinct distribution of the top 20 OTUs identified in the suspended (red) and attached (mint) microbial communities along spatial changes and at different time point. The taxonomic affiliation for the non-redundant OTUs (n=38) at the lowest identified phylogenetic level was provided. The size of the bubbles was proportional to the relative abundance of the OTUs in individual samples. Three OTUs affiliated to the unclassified *Desulfobacteriales*, *Methylobacter*, and unclassified *Methylococcaceae* were labeled in red because they were among the top 20 OTUs in both suspended and attached microbial consortia. The colored bars on the bottom of the diagram illustrated the samples collected from different wells (red: G04UG; purple: G101; yellow: G102; green: G103; blue: G104).

680 (a)

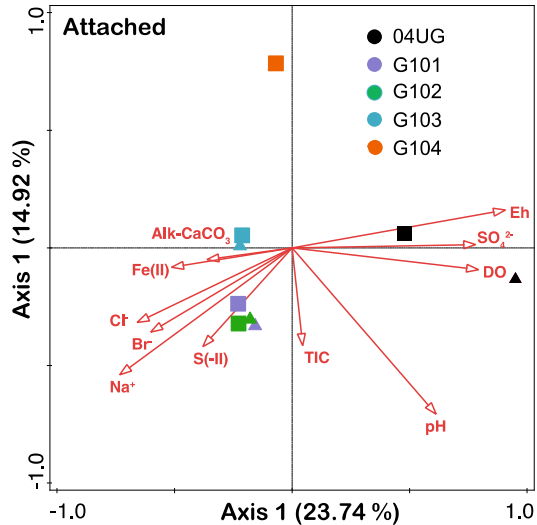


681
682

(b)

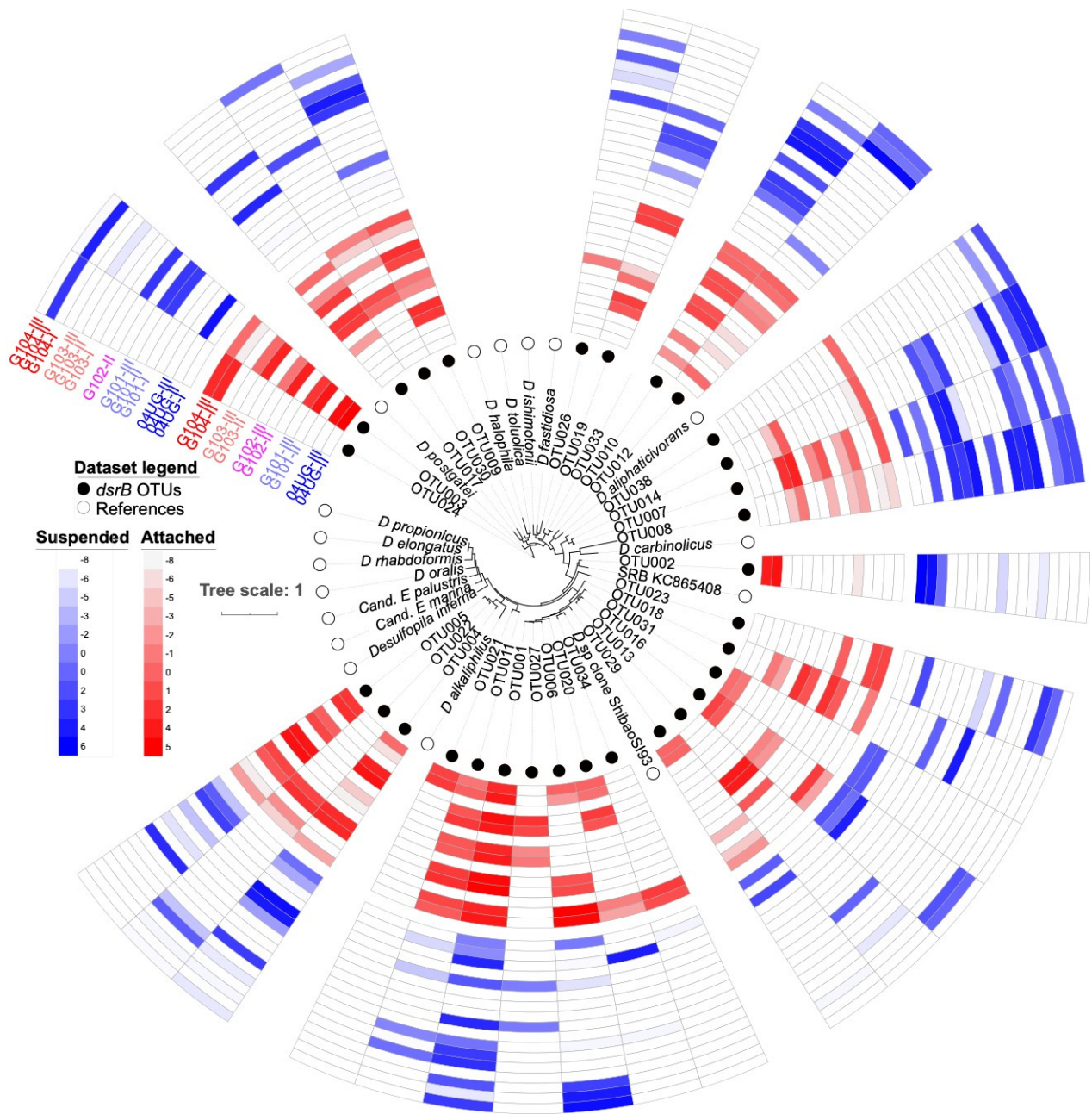


(c)

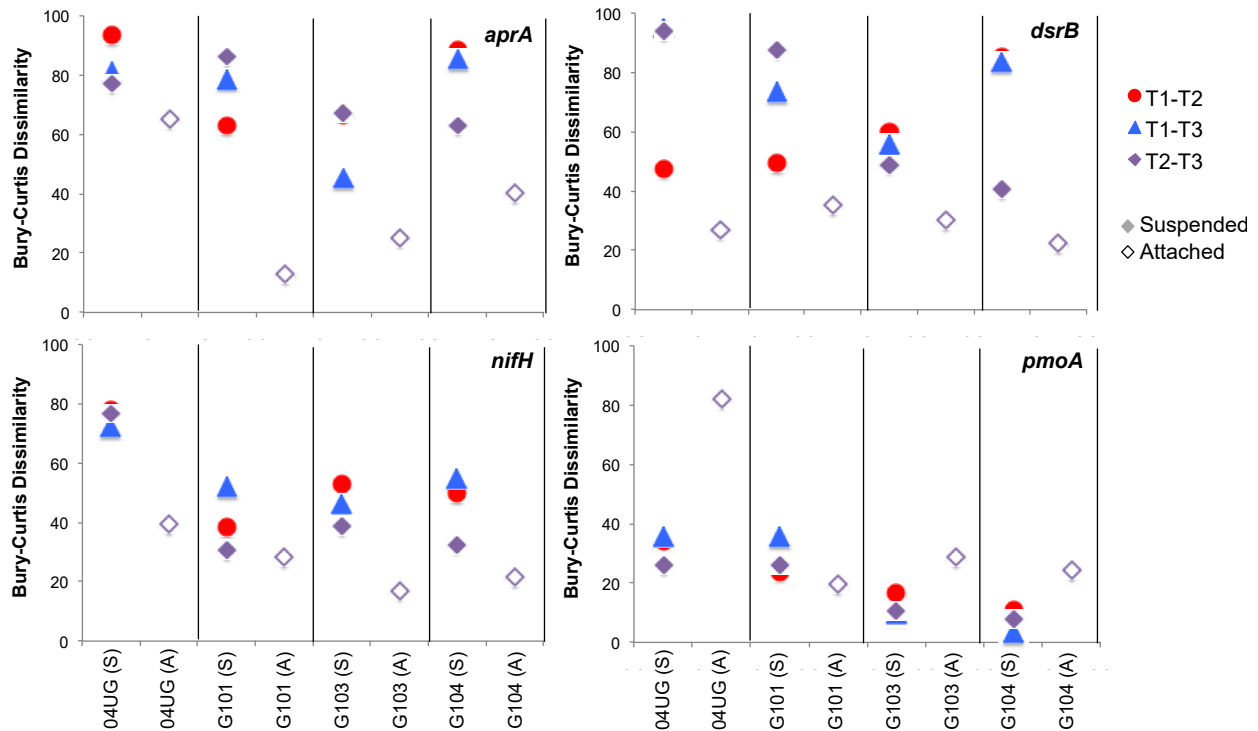


683
684
685
686
687
688
689
690
691
692
693
694

Figure 5. Distinct suspended and attached microbial communities identified in the shallow groundwater wells under investigation (a); CCA analyses of the suspended (b) and attached (c) microbial communities mediated by environmental factors. In (a), closed and open symbols indicated suspended and attached microbial communities, respectively. In (b) and (c), the samples collected at different time points were presented in different symbols: circle: time 0; square: 3 months; triangle: 6 months. The geochemical parameters were selected based on VIF analyses that removed the redundant ones. In the diagrams, some of the plots were too close to be visualized separately. Each plot represented average of duplicates if available.

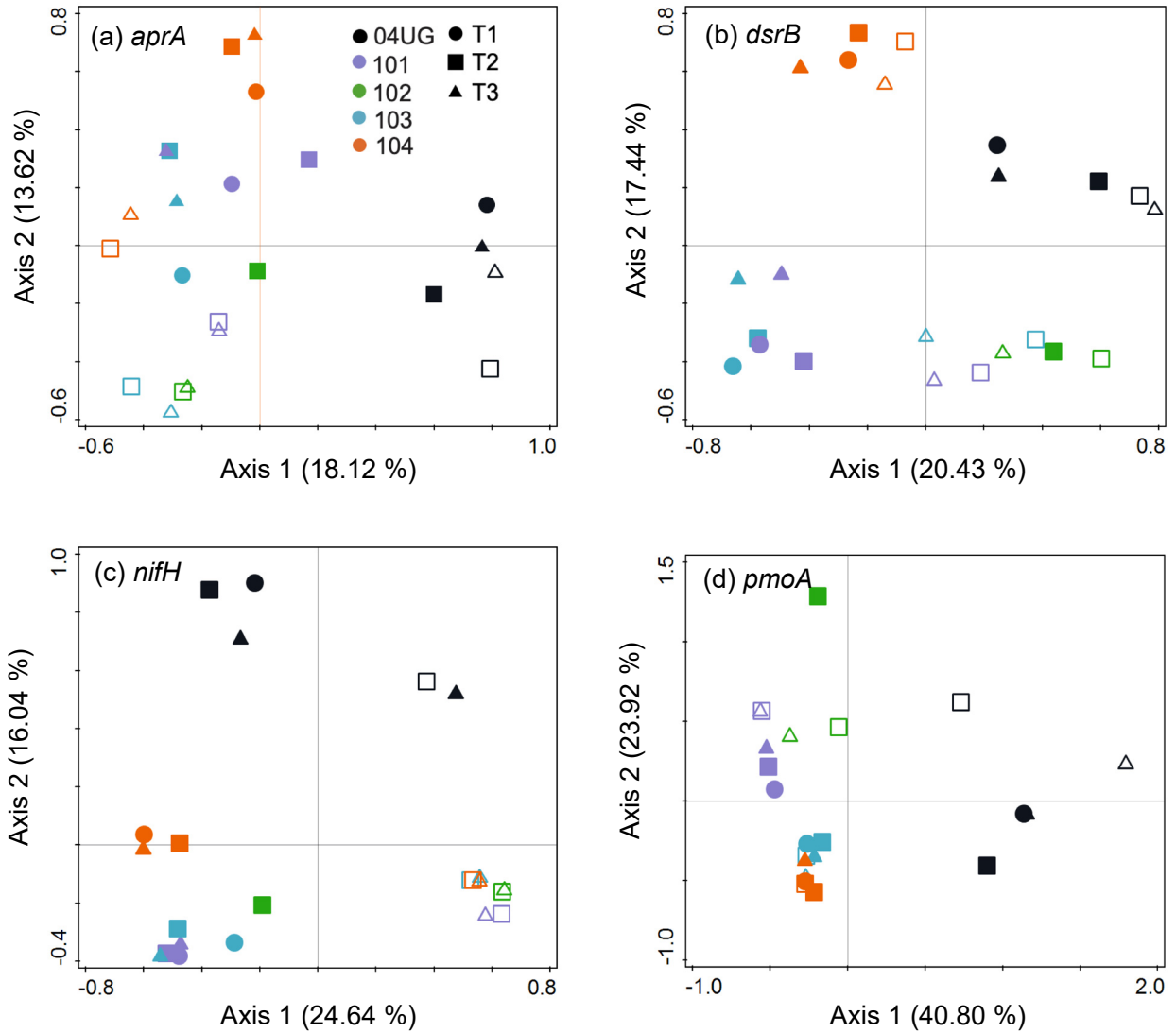


695
696 Figure 6. Distribution of *dsrB* in the suspended (blue) and attached microbial communities (red)
697 and the phylogenetic tree for the detected *dsrB* genes. The phylotypes under consideration
698 comprised of the non-redundant top 5 abundant OTUs identified in each sample. The reference
699 *dsrB* genes were identified via BLAST against the NCBI-non redundant database and the gene
700 sequences for the cultivated species were preferentially selected. *D. postgatei*: *Desulfobacter*
701 *postgatei* 2ac9 strain DSM 2034 (AF418198); *D. inferna*: *Desulfopila inferna* strain
702 JS_SRB250Lac (FJ548990); *D. oralis*: *Desulfobulbus oralis* strain HOT041/ORNL (CP021255);
703 *D. infernum*: *Desulfacinum infernum* strain DSM 9756 (AF482454); *D. toluolica*: *Desulfobacula*
704 *toluolica* Tol2 (FO203503); *D. fastidiosa*: *Desulfofaba fastidiosa* (AY268892); *Cand. E. marina*:
705 *Candidatus Electrothrix marina* isolate AarhusBay_A5 (KU844010).
706



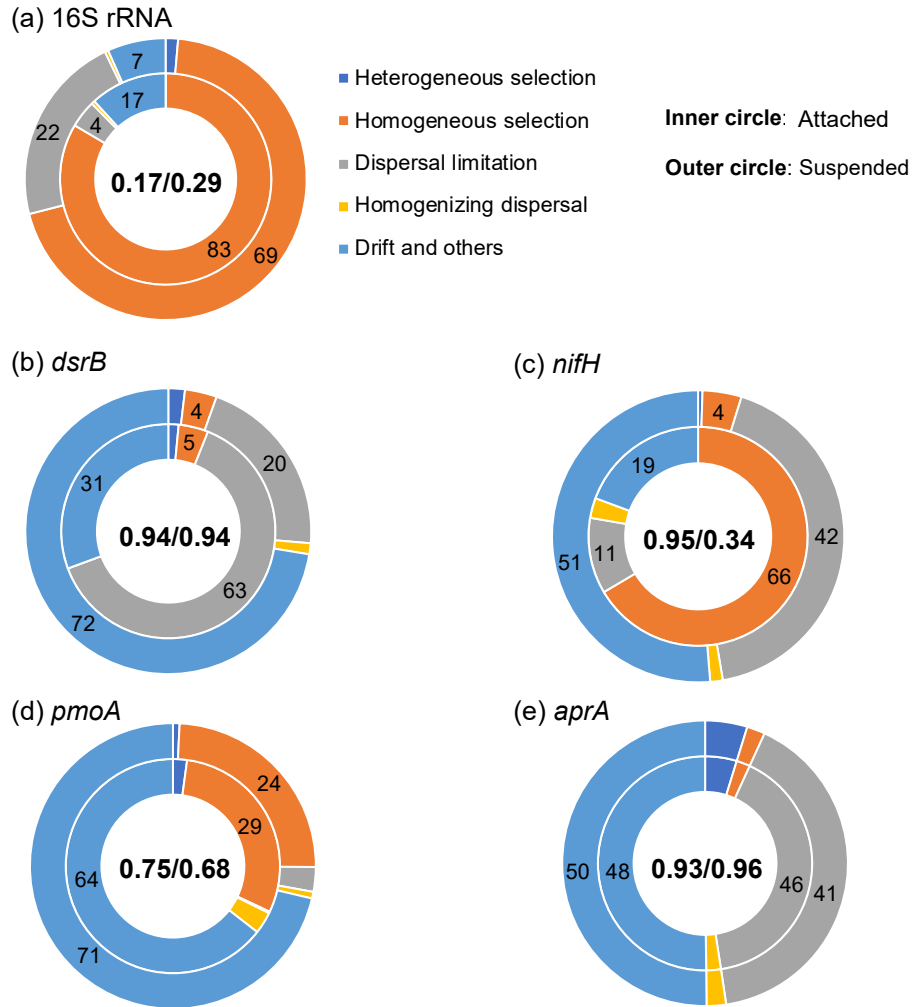
707

708 Figure 7. Bray-Curtis dissimilarity in percentage of the representative functional gene
 709 composition identified in different wells. The functional gene OTUs were determined based on
 710 0.03 cutoff. The solid and open symbols represented the analyses for the suspended and attached
 711 microbial communities, respectively. For the *pmoA* gene in 04UG, significantly higher Bray-
 712 Curtis dissimilarity was obtained potentially due to the fluctuation of limited *pmoA* OTUs. T1-T2,
 713 T2-T3 and T1-T3 indicated pair-wise Bray-Curtis dissimilarity between the samples collected at 0
 714 versus 3 months, 3 versus 6 months, and 0 versus 6 months, respectively.
 715



716
 717 Figure 8. PCoA analyses based on Bray-Curtis dissimilarity of the functional genes identified in
 718 the suspended (solid symbols) and attached (open symbols) microbial communities collected
 719 from different wells and at different time points. Each plot represented average of duplicates if
 720 available. Some of the samples were too close to be visualized. For well G102, no sequencing
 721 reads were available for the suspended microbial samples collected at time 0 and 6 months due to
 722 low biomass.

723



724

725 Figure 9. The relative importance of different ecological processes on the assembly of the
 726 microbial communities (a) and functional genes (b-e) based on the estimation using iCAMP.
 727 The values inside the rings indicated the stochasticity for the 16S rRNA genes (a) or the
 728 functional genes (b-e) determined in the attached versus the suspended microbial communities.
 729 The values labelled on the rings indicated the relative contribution of different ecological
 730 processes in percentage.
 731

References

- Adams, M. A., Eddy, G. E., Hewitt, J. L., Kirr, J. N. and Rightmire, C. T. (1984) Pennsylvanian geology, coal, and coalbed methane resources of the Illinois Basin—Illinois, Indiana, and Kentucky. in *Coalbed Methane Resources of the United States*. (eds) Rightmire, C.T., Eddy, G.E. and Kirr, J.N., American Association of Petroleum Geologists, Tulsa, OK, pp. 105-134.
- Akob, D. M. and Küsel, K. 2011. Where microorganisms meet rocks in the Earth's critical zone. *Biogeosciences* 8(12), 3531-3543.
- Alfreider, A., Krössbacher, M. and Psenner, R. 1997. Groundwater samples do not reflect bacterial densities and activity in subsurface systems. *Water Res.* 31(4), 832-840.
- Algburi, A., Comito, N., Kashtanov, D., Dicks, L. M. T. and Chikindas, M. L. 2017. Control of biofilm formation: antibiotics and beyond. *Appl. Environ. Microbiol.* 83, e00165-00117.
- Barnhart, E. P., De Leon, K. B., Ramsay, B. D., Cunningham, A. B. and Fields, M. W. 2013. Investigation of coal-associated bacterial and archaeal populations from a diffusive microbial sampler (DMS). *Int. J. Coal Geol.* 115, 64-70.
- Ben Maamar, S., Aquilina, L., Quaiser, A., Pauwels, H., Michon-Coudouel, S., Vergnaud-Ayraud, V., Labasque, T., Roques, C., Abbott, B. W. and Dufresne, A. 2015. Groundwater isolation governs chemistry and microbial community structure along hydrologic flowpaths. *Front. Microbiol.* 6, 1457.
- Benner, S. G., Blowes, D. W., Gould, W. D., Herbert, R. B. and Ptacek, C. J. 1999. Geochemistry of a permeable reactive barrier for metals and acid mine drainage. *Environ. Sci. Technol.* 33(16), 2793-2799.

Besemer, K., Peter, H., Logue, J. B., Langenheder, S., Lindstrom, E. S., Tranvik, L. J. and Battin, T. J. 2012. Unraveling assembly of stream biofilm communities. *ISME J.* 6(8), 1459-1468.

Besemer, K., Singer, G., Hodl, I. and Battin, T. J. 2009. Bacterial community composition of stream biofilms in spatially variable-flow environments. *Appl. Environ. Microbiol.* 75(22), 7189-7195.

Booker, A. E., Borton, M. A., Daly, R. A., Welch, S. A., Nicora, C. D., Hoyt, D. W., Wilson, T., Purvine, S. O., Wolfe, R. A., Sharma, S., Mouser, P. J., Cole, D. R., Lipton, M. S., Wrighton, K. C. and Wilkins, M. J. 2017. Sulfide generation by dominant *Halanaerobium* microorganisms in hydraulically fractured shales. *mSphere* 2(4).

Borton, M. A., Hoyt, D. W., Simon, R., Daly, R. A., Welch, S. A., Nicora, C. D., Purvine, S., Eder, E. K., Hanson, A. J., Sheets, J. M., Morgan, D. M., Wolfe, R. A., Sharma, S., Carr, T. R., Cole, D. R., Mouser, P. J., Lipton, M. S., Wilkins, M. J. and Wrighton, K. C. 2018. Coupled laboratory and field investigations resolve microbial interactions that underpin persistence in hydraulically fractured shales. *PNAS* 115(28), 6585-6594.

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S. and Thomas, T. 2011. Bacterial community assembly based on functional genes rather than species. *PNAS* 108(34), 14288-14293.

Christensen, G. A., Moon, J., Veach, A. M., Mosher, J. J., Wymore, A. M., van Nostrand, J. D., Zhou, J. Z., Hazen, T. C., Arkin, A. P. and Elias, D. A. 2018. Use of in-field bioreactors demonstrate groundwater filtration influences planktonic bacterial community assembly, but not biofilm composition. *Plos One* 13(3).

Coady, S. and Mehnert, E. 2011. Estimating Aquifer Properties for IBDP Wells G101, G102, G103, G104. Illinois State Geological Survey, Champaign, IL, pp. 22.

Coyte, K. Z., Tabuteau, H., Gaffney, E. A., Foster, K. R. and Durhama, W. M. 2016. Microbial competition in porous environments can select against rapid biofilm growth. PNAS 114(2), 161-170.

Danczak, R. E., Johnston, M. D., Kenah, C., Slattery, M. and Wilkins, M. J. 2018. Microbial community cohesion mediates community turnover in unperturbed aquifers. mSystems 3(4).

Danielopol, D. L., Griebler, C., Gunatilaka, A. and Notenboom, J. 2003. Present state and future prospects for groundwater ecosystems. Environ. Conserv. 30(2), 104-130.

Dong, Y., Kumar, C. G., Chia, N., Kim, P.-J., Miller, P. A., Price, N. D., Cann, I. K. O., Flynn, T. M., Sanford, R. A., Krapac, I. G., Locke, R. A., Hong, P.-Y., Tamaki, H., Liu, W.-T., Mackie, R. I., Hernandez, A. G., Wright, C. L., Mikel, M. A., Walker, J. L., Sivaguru, M., Fried, G., Yannarell, A. C. and Fouke, B. W. 2014a. *Halomonas sulfidaeris*-dominated microbial community inhabits a 1.8 km-deep subsurface Cambrian Sandstone reservoir. Environ. Microbiol. 16(6), 1695-1708.

Dong, Y., Sanford, R. A., Locke, R. A., Cann, I. K., Mackie, R. I. and Fouke, B. W. 2014b. Fe-oxide grain coatings support bacterial Fe-reducing metabolisms in 1.7-2.0 km-deep subsurface quartz arenite sandstone reservoirs of the Illinois Basin (USA). Front. Microbiol. 5, 511.

EPA 1994. Determination of Metals and Trace Elements in Water and Wastes by Inductively Coupled Plasma-Atomic Emission Spectrometry. Cincinnati, OH, pp. 1-58.

Fields, M. W., Bagwell, C. E., Carroll, S. L., Yan, T., Liu, X., Watson, D. B., Jardine, P. M., Criddle, C. S. and Hazen, T. C. 2006. Phylogenetic and functional biomarkers as indicators of bacterial community responses to mixed-waste contamination. Environ. Sci. Technol. 40(8), 2601-2607.

- Fillinger, L., Hug, K. and Griebler, C. 2019. Selection imposed by local environmental conditions drives differences in microbial community composition across geographically distinct groundwater aquifers. *FEMS Microbiol. Ecol.* 95(11).
- Finley, R. J. 2014. An overview of the Illinois Basin–Decatur Project. *GREENH GASES* 4(5), 571-579.
- Finneran, K. T., Johnsen, C. V. and Lovley, D. R. 2003. *Rhodoferax ferrireducens* sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III). *Int. J. Syst. Evol. Microbiol.* 53, 669-673.
- Flemming, H.-C. and Wingender, J. 2010. The biofilm matrix. *Nat. Rev. Microbiol.* 8, 623-633.
- Flemming, H.-C. and Wuertz, S. 2019. Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.* 17(4), 247-260.
- Flynn, T. M., Sanford, R. A. and Bethke, C. M. 2008. Attached and suspended microbial communities in a pristine confined aquifer. *Water Resour. Res.* 44, W07425.
- Flynn, T. M., Sanford, R. A., Ryu, H., Bethke, C. M., Levine, A. D., Ashbolt, N. J. and Santo Domingo, J. W. 2013. Functional microbial diversity explains groundwater chemistry in a pristine aquifer. *BMC Microbiol.* 13, 146.
- Flynn, T. M., Sanford, R. A., Santo Domingo, J. W., Ashbolt, N. J., Levine, A. D. and Bethke, C. M. 2012. The active bacterial community in a pristine confined aquifer. *Water Resour. Res.* 48, W09510.
- Gerth, J. and Förstner, U. 2004. Fitness for aquatic systems long-term forecast: key to groundwater protection. *Environ. Sci. Pollut. Res.* 11(1), 49-56.

Graham, E. B., Crump, A. R., Resch, C. T., Fansler, S., Arntzen, E., Kennedy, D. W., Fredrickson, J. K. and Stegen, J. C. 2017. Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. *Environ. Microbiol.* 19(4), 1552-1567.

Griebler, C. and Lueders, T. 2009. Microbial biodiversity in groundwater ecosystems. *Freshwater Biol.* 54(4), 649-677.

Gulmann, L. K., Beaulieu, S. E., Shank, T. M., Ding, K., Seyfried, W. E. and Sievert, S. M. 2015. Bacterial diversity and successional patterns during biofilm formation on freshly exposed basalt surfaces at diffuse-flow deep-sea vents. *Front. Microbiol.* 6.

Hädrich, A., Taillefert, M., Akob, D. M., Cooper, R. E., Litzba, U., Wagner, F. E., Nietzsche, S., Ciobota, V., Rösch, P., Popp, J. and Küsel, K. 2019. Microbial Fe(II) oxidation by *Sideroxydans lithotrophicus* ES-1 in the presence of Schlöppnerbrunnen fen-derived humic acids. *FEMS Microbiol. Ecol.* 95(4).

Hazen, T. C., Jiménez, L., de Victoria, G. L. and Fliermans, C. B. 1991. Comparison of bacteria from deep subsurface sediment and adjacent groundwater. *Microbiol. Ecol.* 22(1), 293-304.

He, C., Keren, R., Whittaker, M. L., Farag, I. F., Doudna, J. A., Cate, J. H. D. and Banfield, J. F. 2021. Genome-resolved metagenomics reveals site-specific diversity of episympiotic CPR bacteria and DPANN archaea in groundwater ecosystems. *Nat. Microbiol.* 6(3), 354-365.

He, Z., Zhang, P., Wu, L., Rocha, A. M., Tu, Q., Shi, Z., Wu, B., Qin, Y., Wang, J., Yan, Q., Curtis, D., Ning, D., Van Nostrand, J. D., Wu, L., Yang, Y., Elias, D. A., Watson, D. B., Adams, M. W. W., Fields, M. W., Alm, E. J., Hazen, T. C., Adams, P. D., Arkin, A. P. and Zhou, J. 2018. Microbial functional gene diversity predicts groundwater contamination and ecosystem functioning. *mBios* 9(1).

Herrmann, M., Wegner, C. E., Taubert, M., Geesink, P., Lehmann, K., Yan, L., Lehmann, R., Totsche, K. U. and Küsel, K. 2019. Predominance of *Cand. Patescibacteria* in groundwater is caused by their preferential mobilization from soils and flourishing under oligotrophic conditions. *Front. Microbiol.* 10, 1047.

Hug, L. A., Thomas, B. C., Brown, C. T., Frischkorn, K. R., Williams, K. H., Tringe, S. G. and Banfield, J. F. 2015. Aquifer environment selects for microbial species cohorts in sediment and groundwater. *ISME J.* 9(8), 1846-1856.

Hwang, C., Wu, W. M., Gentry, T. J., Carley, J., Carroll, S. L., Schadt, C., Watson, D., Jardine, P. M., Zhou, J., Hickey, R. F., Criddle, C. S. and Fields, M. W. 2006. Changes in bacterial community structure correlate with initial operating conditions of a field-scale denitrifying fluidized bed reactor. *Appl. Microbiol. Biotechnol.* 71(5), 748-760.

Iguchi, H., Yurimoto, H. and Sakai, Y. 2011. *Methylovulum miyakonense* gen. nov., sp. nov., a Type I methanotroph isolated from forest soil. *Int. J. Syst. Evol. Microbiol.* 61, 810-815.

King, A. J., Preheim, S. P., Bailey, K. L., Robeson, M. S., 2nd, Roy Chowdhury, T., Crable, B. R., Hurt, R. A., Jr., Mehlhorn, T., Lowe, K. A., Phelps, T. J., Palumbo, A. V., Brandt, C. C., Brown, S. D., Podar, M., Zhang, P., Lancaster, W. A., Poole, F., Watson, D. B., M, W. F., Chandonia, J. M., Alm, E. J., Zhou, J., Adams, M. W., Hazen, T. C., Arkin, A. P. and Elias, D. A. 2017. Temporal dynamics of in-field bioreactor populations reflect the groundwater system and respond predictably to perturbation. *Environ. Sci. Technol.* 51(5), 2879-2889.

Küsel, K., Totsche, K. U., Trumbore, S. E., Lehmann, R., Steinhäuser, C. and Herrmann, M. 2016. How deep can surface signals be traced in the critical zone? Merging biodiversity with biogeochemistry research in a central German Muschelkalk landscape. *Front. Earth Sci.* 4, 2296-6463

Lazar, C. S., Lehmann, R., Stoll, W., Rosenberger, J., Totsche, K. U. and Küsel, K. 2019. The endolithic bacterial diversity of shallow bedrock ecosystems. *Sci Total Environ* 679, 35-44.

Lee, J. H. and Lee, B. J. 2018. Microbial reduction of Fe(III) and SO_4^{2-} and associated microbial communities in the alluvial aquifer groundwater and sediments. *Microbial. Ecol.* 76(1), 182-191.

Lin, X., Kennedy, D., Peacock, A., McKinley, J., Resch, C. T., Fredrickson, J. and Konopka, A. 2011. Distribution of microbial biomass and the potential for anaerobic respiration in Hanford Site 300 Area subsurface sediment. *Appl. Environ. Microbiol.*, 759-767.

Locke, R. A. and Iranmanesh, A. 2015. Quarterly groundwater report for Illinois EPA Underground Injection Control Permit Number UIC-012-ADM. Illinois State Geological Survey, Champaign, IL, pp. 15.

Locke, R. A., Shao, H., Iranmanesh, A. and Wimmer, B. T. 2018. Groundwater Compliance Report For US EPA Underground Injection Control Permits IL-115-6A-0001 and IL-115-6A-0002. Illinois State Geological Survey, pp. 37.

Lovley, D. R. and Walker, D. J. F. 2019. *Geobacter* protein nanowires. *Front. Microbiol.* 10.

Luef, B., Frischkorn, K. R., Wrighton, K. C., Holman, H. Y., Birarda, G., Thomas, B. C., Singh, A., Williams, K. H., Siegerist, C. E., Tringe, S. G., Downing, K. H., Comolli, L. R. and Banfield, J. F. 2015. Diverse uncultivated ultra-small bacterial cells in groundwater. *Nat. Commun.* 6, 6372.

Meents, W. F. 1960. Glacial-drift Gas in Illinois. Illinois State Geological Survey, Champaign, IL, pp. 58.

Mehta-Kolte, M. G. and Bond, D. R. 2012. *Geothrix fermentans* secretes two different redox-active compounds to utilize electron acceptors across a wide range of redox potentials. *Appl. Environ. Microbiol.* 78(19), 6987-6995.

Munday, P. L. 2004. Competitive coexistence of coral-dwelling fishes: The lottery hypothesis revisited. *Ecology* 85(3), 623-628.

Muyzer, G. and Stams, A. J. M. 2008. The ecology and biotechnology of sulphate-reducing bacteria. *Nat. Rev. Microbiol.* 6(6), 441-454.

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P. and Ferrenberg, S. 2013. Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77(3), 342-356.

Ning, D., Deng, Y., Tiedje, J. M. and Zhou, J. 2019. A general framework for quantitatively assessing ecological stochasticity. *PNAS* 116(34), 16892-16898.

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., Yang, Y., Arkin, A. P., Firestone, M. K. and Zhou, J. 2020. A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat. Commun.* 11(1), 4717.

Oshkin, I. Y., Belova, S. E., Danilova, O. V., Miroshnikov, K. K., Rijpstra, W. I. C., Sinninghe Damste, J. S., Liesack, W. and Dedysh, S. N. 2016. *Methylovulum psychrotolerans* sp. nov., a cold-adapted methanotroph from low-temperature terrestrial environments, and emended description of the genus *Methylovulum*. *Int. J. Syst. Evol. Microbiol.* 66(6), 2417-2423.

Pfaff, J. D. 1993. Determination of Inorganic Anions by Ion Chromatography. U. S. EPA, Cincinnati, OH, pp. 1-28.

Schwab, V. F., Herrmann, M., Roth, V. N., Gleixner, G., Lehmann, R., Pohnert, G., Trumbore, S., Küsel, K. and Totsche, K. U. 2017. Functional diversity of microbial communities in pristine aquifers inferred by PLFA- and sequencing-based approaches. *Biogeosciences* 14(10), 2697-2714.

Shao, H. B. and Butler, E. C. 2007. The influence of iron and sulfur mineral fractions on carbon tetrachloride transformation in model anaerobic soils and sediments. *Chemosphere* 68(10), 1807-1813.

Smith, H. J., Zelaya, A. J., De Leon, K. B., Chakraborty, R., Elias, D. A., Hazen, T. C., Arkin, A. P., Cunningham, A. B. and Fields, M. W. 2018. Impact of hydrologic boundaries on microbial planktonic and biofilm communities in shallow terrestrial subsurface environments. *FEMS Microbiol. Ecol.* 94(12).

Stegen, J. C., Lin, X. J., Fredrickson, J. K., Chen, X. Y., Kennedy, D. W., Murray, C. J., Rockhold, M. L. and Konopka, A. 2013. Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7(11), 2069-2079.

Stegen, J. C., Lin, X. J., Fredrickson, J. K. and Konopka, A. E. 2015. Estimating and mapping ecological processes influencing microbial community assembly. *Front. Microbiol.* 6.

Stoliker, D. L., Repert, D. A., Smith, R. L., Song, B. K., LeBlanc, D. R., McCobb, T. D., Conaway, C. H., Hyun, S. P., Koh, D. C., Moon, H. S. and Kent, D. B. 2016. Hydrologic controls on nitrogen cycling processes and functional gene abundance in sediments of a groundwater flow-through lake. *Environ. Sci. Technol.* 50(7), 3649-3657.

Stoodley, P., Sauer, K., Davies, D. G. and Costerton, J. W. 2002. Biofilms as complex differentiated communities. *Annu. Rev. Microbiol.* 56, 187-209.

Strapoc, D., Picardal, F. W., Turich, C., Schaperdoth, I., Macalady, J. L., Lipp, J. S., Lin, Y. S., Ertefai, T. F., Schubotz, F., Hinrichs, K. U., Mastalerz, M. and Schimmelmann, A. 2008. Methane-producing microbial community in a coal bed of the Illinois Basin. *Appl. Environ. Microbiol.* 74(8), 2424-2432.

Trivedi, P., Delgado-Baquerizo, M., Trivedi, C., Hu, H. W., Anderson, I. C., Jeffries, T. C., Zhou, J. Z. and Singh, B. K. 2016. Microbial regulation of the soil carbon cycle: evidence from gene-enzyme relationships. *ISME J.* 10(11), 2593-2604.

Wartiainen, I., Hestnes, A. G., McDonald, I. R. and Svenning, M. M. 2006. *Methylobacter tundripaludum* sp. nov., a methane-oxidizing bacterium from Arctic wetland soil on the Svalbard islands, Norway (78 degrees N). *Int. J. Syst. Evol. Microbiol.* 56(1), 109-113.

Watson, A., Peterson, E. W., Malone, D. and Tranel, L. 2020. Geology and aquifer sensitivity of quaternary glacial deposits overlying a portion of the Mahomet buried bedrock valley aquifer system. *Hydrology* 7(4), 69.

Wegner, C. E., Gaspar, M., Geesink, P., Herrmann, M., Marz, M. and Küsel, K. 2019. Biogeochemical regimes in shallow aquifers reflect the metabolic coupling of the elements nitrogen, sulfur, and carbon. *Appl. Environ. Microbiol.* 85(5).

Wells, N. S., Kappelmeyer, U. and Knöllera, K. 2018. Anoxic nitrogen cycling in a hydrocarbon and ammonium contaminated aquifer. *Water Res.* 142(1), 373-382.

Wessel, A. K., Arshad, T. A., Fitzpatrick, M., Connell, J. L., Bonnecaze, R. T., Shear, J. B. and Whiteley, M. 2014. Oxygen Limitation within a Bacterial Aggregate. *mBio* 5(2).

Wilkins, M. J., Daly, R. A., Mouser, P. J., Trexler, R., Sharma, S., Cole, D. R., Wrighton, K. C., Biddle, J. F., Denis, E. H., Fredrickson, J. K., Kieft, T. L., Onstott, T. C., Peterson, L., Pfiffner, S. M., Phelps, T. J. and Schrenk, M. O. 2014. Trends and future challenges in sampling the deep terrestrial biosphere. *Front. Microbiol.* 5.

Wimmer, B. T., Iranmanesh, A., Carman, C. H., Garner, D., Patterson, C. G. and Locke II, R. A. 2021. Illinois Basin-Decatur Project: Shallow Groundwater Monitoring Program. Illinois State Geological Survey, pp. 31.

Xu, M., Wu, W. M., Wu, L., He, Z., Van Nostrand, J. D., Deng, Y., Luo, J., Carley, J., Ginder-Vogel, M., Gentry, T. J., Gu, B., Watson, D., Jardine, P. M., Marsh, T. L., Tiedje, J. M., Hazen, T., Criddle, C. S. and Zhou, J. 2010. Responses of microbial community functional structures to pilot-scale uranium in situ bioremediation. *ISME J.* 4(8), 1060-1070.

Yamamoto, K., Hackley, K. C., Kelly, W. R., Panno, S. V., Sekiguchi, Y., Sanford, R. A., Liu, W. T., Kamagata, Y. and Tamaki, H. 2019. Diversity and geochemical community assembly processes of the living rare biosphere in a sand-and-gravel aquifer ecosystem in the Midwestern United States. *Sci. Rep.* 9(1), 13484.

Yan, L., Herrmann, M., Kampe, B., Lehmann, R., Totsche, K. U. and Küsel, K. 2020. Environmental selection shapes the formation of near-surface groundwater microbiomes. *Water Res.* 170, 115341.

Zelaya, A. J., Parker, A. E., Bailey, K. L., Zhang, P., Van Nostrand, J., Ning, D., Elias, D. A., Zhou, J., Hazen, T. C., Arkin, A. P. and Fields, M. W. 2019. High spatiotemporal variability of bacterial diversity over short time scales with unique hydrochemical associations within a shallow aquifer. *Water Res.* 164, 114917.

Zhou, Y. X., Kellermann, C. and Griebler, C. 2012. Spatio-temporal patterns of microbial communities in a hydrologically dynamic pristine aquifer. *FEMS Microbiol. Ecol.* 81(1), 230-242.

