Water Research 164 (2019) 114917

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

High spatiotemporal variability of bacterial diversity over short time scales with unique hydrochemical associations within a shallow aquifer

Anna J. Zelaya [a, b], Albert E. Parker [a, c], Kathryn L. Bailey [d], Ping Zhang [e], Joy Van Nostrand [e], Daliang Ning [e], Dwayne A. Elias [d, 1], Jizhong Zhou [e, 1], Terry C. Hazen [f, 1], Adam P. Arkin [g, 1], Matthew W. Fields [a, b, *, 1]

[a] Center for Biofilm Engineering, Montana State University, Bozeman, MT, USA
[b] Department of Microbiology & Immunology, Montana State University, Bozeman, MT, USA
[c] Department of Mathematical Sciences, Montana State University, Bozeman, MT, USA
[d] Division of Environmental Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[e] Institute for Environmental Genomics, University of Oklahoma, Norman, OK, USA
[f] Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, USA
[g] Department of Bioengineering, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

ARTICLE INFO

Article history:
Received 28 March 2019
Received in revised form
24 July 2019
Accepted 25 July 2019
Available online 26 July 2019

Keywords:
Groundwater
Microbial community dynamics
Well construction
Functional gene array
Spatiotemporal
Time series
Shallow subsurface

ABSTRACT

Understanding microbial community structure and function within the subsurface is critical to assessing overall quality and maintenance of groundwater; however, the factors that determine microbial community assembly, structure, and function in groundwater systems and their impact on water quality remains poorly understood. In this study, three shallow wells (FW301, FW303, FW305) in a non-contaminated shallow aquifer in the ENIGMA-Oak Ridge Field Research Center (Oak Ridge, TN) were sampled approximately 3 times a week over a period of three months to measure changes in groundwater geochemistry and microbial diversity. It was expected that the sampled microbial diversity from two historic field wells (FW301, FW303) would be relatively stable, while diversity from a newer well (FW305) would be less stable over time. The wells displayed some degree of hydrochemical variability over time unique to each well, with FW303 being overall the most stable well and FW301 being the most dynamic based upon dissolved oxygen, conductivity, and nitrate. Community analysis via ss-rRNA paired-end sequencing and distribution-based clustering revealed higher OTU richness, diversity, and variability in groundwater communities of FW301 than the other two wells for diversity binned over all time points. Microbial community composition of a given well was on average > 50% dissimilar to any other well at a given time (days), yet, functional gene diversity as measured with GeoChip remained relatively constant. Similarities in community structure across wells were observed with respect to the presence of 20 shared bacterial groups in all samples in all wells, although at varying levels over the tested time period. Similarity percentage (SIMPER) analysis revealed that variability in FW301 was largely attributed to low abundance, highly-transient populations, while variability in the most hydrochemically stable well (FW303) was due to fluctuations in more highly abundant and frequently present taxa. Additionally, the youngest well FW305 showed a dramatic shift in community composition towards the end of the sampling period that was not observed in the other wells, suggesting possible succession events over time. Time-series analysis using vector auto-regressive models and Granger causality showed unique relationships between richness and geochemistry over time in each well. These results indicate temporally dynamic microbial communities over short time scales, with day-to-day population shifts in local community structure influenced by available source community diversity and local groundwater hydrochemistry.

* Corresponding author. 366 Barnard Hall, Center for Biofilm Engineering, Montana State University, Bozeman, MT, USA.
E-mail address: matthew.fields@biofilm.montana.edu (M.W. Fields).
[1] ENIGMA (http://enigma.lbl.gov/).

https://doi.org/10.1016/j.watres.2019.114917
0043-1354/

1. Introduction

Groundwater is an essential resource used by humans for drinking, agriculture, and industrial processes such as mining and energy production (Danielopol et al., 2003; Reilly et al., 2008). Currently, global groundwater resources are being threatened by overuse, depletion, and contamination from a variety of sources including agricultural seepage (*i.e.*, nitrates and pesticides), industrial processes, pharmaceuticals, personal care products, landfill leaching, and improperly managed wastewater treatment systems (Danielopol et al., 2003; Gerth and Forstner, 2004; Sui et al., 2015). Additionally, subsurface ecosystems are increasingly being disturbed via engineering efforts (*e.g.*, managed aquifer recharge systems (Lee and Lee, 2017), geothermal/cooling (van der Gun et al., 2016)), often without a thorough investigation of the microbial ecological networks that may be perturbed. Understanding the role of microbial communities in uncontaminated, pristine environments can help assessment of anthropological impacts on subsurface ecosystems and thus inform proper management of water resources.

While subsurface hydrology of groundwater aquifers has been well studied by hydrologists and hydrochemists (Chapelle, 2000), the microbiology and ecology of groundwater systems have only recently begun to be studied (Griebler and Lueders, 2009). Interest in groundwater microorganisms and associated ecological function has increased largely because of the emerging need to assess groundwater quality and fitness for human use, for example determining possible vectors for disease or microbial potential to naturally attenuate contaminated water systems (Anderson and Lovley, 1997; Chapelle, 2000; Langwaldt and Puhakka, 2000; Fields et al., 2006). Other studies have uncovered the significant roles that microorganisms play in groundwater processes, such as water geochemistry (Chapelle, 2000), rates of mineral weathering in aquifers (Akob and Küsel, 2011), and the fate and transport of metals and organic compounds (Anderson and Lovley, 1997; Fields et al., 2006; Hwang et al., 2009). While much of the attention given to shallow aquifer systems has been on the effects of contamination and other disturbances (Fields et al., 2006; Zinger et al., 2012; references therein), knowledge of the activity and composition of microorganisms in pristine groundwater environments have been far less studied (Sogin et al., 2006; Griebler and Lueders, 2009; Caporaso et al., 2010; Zinger et al., 2012), and understanding of the spatial and temporal dynamics in groundwater environments is still in its early stages (Smith et al., 2018, references therein).

he ENIGMA-Oak Ridge Field Research Center (EOR-FRC), established by the U.S. Department of Energy (DOE), consists of several contaminated areas as well as a non-contaminated background site. Multiple field wells in the background area access a shallow, oligotrophic aquifer. A clearer understanding of the natural variability in community structure in pristine aquifers is necessary to fully assess the impact of disturbances on aquifer environments. The described study analyzed the spatial and temporal variability of microbial communities in a pristine aquifer at the EOR-FRC and sought to answer the following questions: (1) How similar are the hydrochemical parameters of three wells in an uncontaminated aquifer system located in close proximity to one another over time? (2) How stable are the microbial communities within and across these wells over time? (3) Does potential functional capabilities of the microbial communities remain stable over time? (4) What is the relationship between the groundwater geochemistry and microbial community composition over space and short-time periods?

2. Materials and methods

2.1. Site location

The EOR-FRC background site covers an area of approximately 1.63 km^2 and is located in the Tennessee West Bear Creek Valley, about 2 km away from any contaminated areas (Fig. S1). The EOR-FRC background site is heavily wooded. The Bear Creek floodplain consists of perennial and ephemeral crosscutting streams which run through the area and feed into Bear Creek (Watson et al., 2004). No known contaminants have been disposed at the background area of the EOR-FRC throughout the history of DOE operations. The subsurface is comprised of underlying bedrock, interbedded shale and limestones, which weather into unconsolidated and lowly-permeable clay-rich saprolite (Watson et al., 2004). The saprolitic layer in turn is overlain by approximately 0.5−3 m of organic and highly permeable clay-rich soil (Solomon et al., 1992). The water table underneath the chosen field wells is considered shallow, lying approximately 6.6 m below the subsurface. Vertical flow of groundwater occurs in the upper porous layers, however hydrological studies of the area show that preferential flow paths are created along fractures in the underlying rock. Tracer studies reveal that the flow paths are poorly connected in 3 dimensions (Solomon et al., 1992). Therefore, while the three field wells chosen for this study tap into the same pristine aquifer, it is concluded that the three wells lie along separate and unconnected flow paths within the same watershed.

2.2. Groundwater sampling

The three field wells chosen were FW301, FW303, and FW305 and are considered shallow wells (<17 m depth) (Fig. S1). Installation of wells FW301 and FW303 was completed in December of 2000 and September of 2001, respectively, while FW305 was constructed on May 29 of 2013. Supplementary Table 4 provides detailed well characteristics. Filtered groundwater samples were collected approximately three times a week over approximately a three-month period (July 23 to October 8, 2013). After 2 well volumes were purged from each well, 4L of groundwater were filtered using a 10 μm filter (to catch large particulates) followed by a 0.2 μm Polyethersulfone (PES) membrane filter to collect biomass for nucleic acid extraction. The 0.2 μm filters were stored at −80 °C until time of nucleic acid extraction. Of the 84 total samples collected, 61 were chosen for paired-end barcoded sequencing (~20 samples per well). Additionally, a total of 12 environmental variables were measured (temperature, pH, oxidation/reduction potential [ORP], dissolved oxygen, conductivity, lactate, acetate, propionate, fluoride, chloride, nitrate, and sulfate) for every sampling time point following the protocol as outlined by King et al. (2017). DOC has previously been observed between 20 and 30 mg/l for this area and Fe^{2+} below 2 mg/l. Briefly, Temperature, pH, ORP, conductivity, and DO were measured using a Multiparameter Series Troll 9500 (In-Situ Inc.). Groundwater hydrochemical data was collected on the same days as biomass collection. Groundwater was collected (5 ml) and filtered (0.22 μm) and processed via a Dionex ICS 5000 + Dual Pump, Dual Column system (ThermoFisher Scientific; Waltham, MA) within 1 h of collection. Anions and organic acids were analyzed simultaneously using an AS11HC column with a KOH gradient of 0−60 mM, and sugars were analyzed on a CarboPac SA 10 column with an isocratic flow of 1 mM KOH, per the manufacturer's instructions. A 5-point calibration curve was performed using prepackaged standards from Dionex at the beginning of each run. Precipitation data (rainfall)

were also monitored and recorded daily. There were some dates where samples measured below the limits of detection for methods used. In these cases, the missing values were substituted using a method similar to the ½ limit of detection rule used in environmental applications (Environmental Protection Agency, 1996; Singh and Nocerino, 2002). Lactate was measured below detection for the majority of time-points (supplementary material); therefore, it was dropped from subsequent statistical analyses.

2.3. Nucleic acid extraction, sequencing, and downstream analysis

Nucleic acids were obtained using the modified miller method of nucleic acid extraction (Hazen et al., 2010). The SSU rRNA gene sequences were amplified using a primer pair targeting the V4 region (Forward primer, 515F, 5′-GTGCCAGCMGCCGCGGTAA-3'; reverse primer, 806R, 5′-GGACTACHVGGGTWTCTAAT-3′) (Caporaso et al., 2012). PCR amplification was performed in triplicate and prepared for sequencing according to the MiSeq™ Reagent Kit Preparation Guide using a 500-cycle v2 MiSeq reagent cartridge (Illumina, San Diego, CA, USA).

Raw sequence data was processed using an in-house pipeline (E. Alm, https://github.com/almlab/SmileTrain/wiki). Reads were merged, removed of primers, quality checked, de-replicated, and checked for chimeras. USEARCH was used to merge reads, filter low quality reads, and remove chimeras (Edgar, 2013). Operational Taxonomic Units (OTUs) were clustered using Distribution Based Clustering (Preheim et al., 2013). Taxonomic annotation of individual OTUs were obtained using the online version of the RDP 16S Classifier (Release 11, Update 5, accessed at https://pyro.cme.msu.edu/classifier/form.sprused on November 6, 2017) (Wang et al., 2007). Only classifications with an estimated confidence of >80% were used in downstream analyses. The OTUs below this cutoff were assigned as "unidentified Bacteria." Additionally, sequences classified as "Chloroplast" were removed from dataset. The resulting classified OTU data table was further filtered to remove rare OTUs. Rare OTUs were defined as OTUs that had a relative abundance of $<10^{-6}$ across all groundwater samples (Haegeman et al., 2013). The representative sequences were used to a build phylogenetic tree by FastTree2 (Price et al., 2009, 2010) after aligned by Clustal Omega (version 1.2.2) (Sievers et al., 2011).

2.4. Data analysis

Initial diagnostics (normality, residual plots) were performed using R statistical package (R Core Team, 2017). Richness and Diversity using Hill Numbers as well as Sorensen and Bray-Curtis pairwise dissimilarities were determined in R with package VEGAN (Oksanen et al., 2013). Correlation analysis and linear regression models were used to identify relationships among the hydrochemical and microbial data using lme4 package (Bates, 2010). Analysis of Variances (ANOVA) was performed to determine mean differences across weeks within each well and across wells. Constrained (Canonical Correspondence Analysis, or CCA) and unconstrained (Detrended Correspondence Analysis, or DCA) ordinations were carried out in R statistical package with VEGAN. Similarity percentage (SIMPER) analysis using PAST statistical package (Hammer et al., 2001) was used to determine specific OTUs contributing to weekly pair-wise dissimilarities.

2.5. GeoChip hybridization and analysis

GeoChip hybridization was performed as described with aliquots (5–10 ng for measurable DNA or 2–5 μl if no measurable DNA) of DNA amplified using the Templiphi kit (GE Healthcare) with the following modifications. Spermidine (0.1 mM) and single stranded binding protein (267 ng/mL) were added to improve the amplification efficiency and samples were amplified for 6 h (Wu et al., 2006). Amplified DNA (~2 μg) was labelled with mixed cyanine 3 using random primers (Life Technologies, random hexamers, 3 μg/μL) and Klenow (imer; San Diego, CA; 40 U/mL) at 37 °C for 6 h and then cleaned using a QIAquick purification kit (Qiagen) per the manufacturer's instructions and dried. Labelled DNA was rehydrated with hybridization buffer and hybridized for 20–22 h at 67 °C and then imaged (NimbleGen MS 200 microarray scanner) and data was extracted using the Agilent Feature Extraction program. Data normalization and quality filtering were performed with multiple steps (Liang et al., 2010). Spots were scored as positive and retained if the signal-to-noise ratio [SNR = (signal mean − background mean)/background standard deviation] was ≥2.0, the coefficient of variation (CV) of the background was <0.8, and the signal intensity was >200.

2.6. Time series data analysis

Time-series analyses were performed using vector autoregressive models (VAR package, Pfaff, 2007) and Granger causality in order to determine possible temporally lagged relationships between geochemistry and microbial diversity. Data was first imputed for every daily time-point using the method of Forsythe et al. (1977). Stationarity of detrended data was ensured via the Augmented Fuller Dickey Test (ADF test) in R using the "tseries" package (Trapletti and Hornik, 2018). A starting lag time (p) of p = 1 was used based on historic hydrologic data (Watson et al., 2004). The lag time of p = 1 represents a 2-day lag, as the data points represent every-other-day sampling. This starting lag time fell within the range of biologically relevant lag times based on calculated residence times for each well (Supplemental Table 2). Partial auto-correlation plots of the residuals were used to assess the need and fit of the VAR models (Clarke and Mirza, 2006). To determine significant relationships, Granger causality was used in R using the "aod" package (Lesnoff and Lancelot, 2012).

2.7. Mechanisms underlying community assembly

The relative roles of community assembly processes were quantified using a null-model-based framework proposed by Stegen et al., 2013, 2015. The influence of selection was estimated based on nearest taxon index between communities (βNTI). A turnover between two communities was considered to be governed by "variable" or "homogeneous selection" when the phylogenetic dissimilarity (βMNTD, β-mean nearest taxon distance) was significantly higher (βNTI>2) or lower (βNTI < −2) than null expectation (Fine and Kembel, 2011; Webb et al., 2008; Stegen et al., 2015). The turnovers with −2<βNTI<2 are not differentiable from random patterns, representing the influence of stochastic processes such as neutral dispersal, drift, etc. (Stegen et al., 2013). The relative role of a process in a well was measured as the percentage of temporal turnovers governed by this process (Stegen et al., 2013; King et al., 2017). The above null model analysis was performed on the rarefied data set.

3. Results

3.1. Hydrochemical behavior is unique to each well and is unstable over observed time period

Although the three field wells chosen for this study were in close proximity to each other, had the same underlying geology, and tapped into the same uncontaminated aquifer, all three wells showed differences in hydrochemical parameters to varying

degrees. Each well differed significantly from each other ($p < 0.01$) in several hydrochemical measurements (Fig. 1). Additionally, FW301 differed significantly from the other two wells in 7 of the 12 total hydrochemical parameters measured (Fig. 1, Figs. S2A–D).

FW301 showed the greatest degree of within-well variability in measured geochemistry over time, specifically for DO, conductivity, and nitrate (Fig. 1, Table S1, Fig. S3). Although FW303 had the highest levels of conductivity, pH, and nitrate, it was also the most hydrochemically stable of the three wells overtime (Fig. 1). Additionally, FW303 had the lowest measured concentrations of dissolved oxygen, although temporary increases were observed during weeks 7 and 8 (Fig. S3). FW305, a well that was cored 8 weeks prior to the start of this study, was more hydrochemically stable than FW301, and behaved uniquely in that there was a steady increase in dissolved oxygen over time, which coincided with greater pH instability towards the later weeks of the sampled time-period (Fig. S3). Measured groundwater geochemistry over the three-month time period was similar in all three wells for oxidation-reduction potential as well as for organic acids (acetate and propionate) (Figs. S2E–G).

3.2. Microbial diversity and taxonomic composition is not similar across wells of close proximity in the same aquifer

A total of 4091 groundwater OTUs were identified in all groundwater samples across all three wells over the tested time period. Of the total OTUs that passed quality filtering, 1497 (~37%) OTUs were classified as unidentifiable Bacteria (could not be identified at >80% confidence at the phylum level). The total sampled groundwater diversity included 32 phyla, 59 classes, and 153 families. Diversity profiles based on Hill's numbers, also known as effective species numbers (n_{eff}), showed that overall OTU richness and unevenness was significantly higher for FW301 (p-value = 0.001) (Fig. 2A). These trends were unaffected by sampling depth, as analysis using rarified samples show similar diversity profiles (Figs. S7 and S8). In addition, FW301 had a greater interquartile range at $q = 1$ (Exponential Shannon) and $q = 2$ (Inverse Simpson) compared to the other wells, suggesting increased temporal variability in FW301 with respect to abundant and predominant OTUs (Fig. 2A).

Based on relative abundance charts, all wells are largely dominated by Proteobacteria, with classes α-, β-, and γ-Proteobacteria having the highest representation (Fig. S4). At higher taxonomic resolution, each well is dominated by unique bacterial groups, with the exception of *Comamonadaceae*, which is highly represented in all three wells. FW303 is largely predominated by Burkholderiales (Fig. S4), and FW305 contains large representations of *Oxalobacteraceae* and *Rhodobacteraceae*. In agreement with diversity measures, FW301 is more highly diverse and does not have a clearly dominating family, instead being intermittantly predominated by *Comamonadacea* and various other proteobacteria over time (Fig. S4).

3.3. Microbial community structure is unique over time across wells

Weekly mean richness varies significantly in FW301 ($F(10,10) = 2.856$, $p = 0.057$) but not in FW303 ($F(10,9) = 0.845$, $p = 0.604$) or FW305 ($F(10,9) = 1.004$, $p = 0.502$), indicating that rare and transient taxa contribute to much of the variability measured in FW301 but not in the other two wells (Fig. 2). As q increases (effects of rare OTUs are removed), mean weekly diversity differs significantly in FW305 ($F(10,9) = 2.791$, $p = 0.0689$ for $q = 1$ and $F(10,9) = 3.589$, $p = 0.0338$ for $q = 2$) but not for FW301 or FW303. This indicates weekly differences in community structure based on both abundant and dominant communities over time in FW305 but not in the other two wells. These same results are obtained when performed on rarified samples. Standard mean errors (SEM) of weekly mean diversity for all three indices measured are high for FW301. Although ANOVA's did not show significant difference across weeks as q increased, high SEM indicates high within-week variability in FW301.

Only 20 bacterial taxa were observed in all groundwater samples across all 3 wells (Fig. 3, Table 1). As such, these taxa can be considered as 20 shared bacterial groups in the system over the tested time period. However, 14 of these 20 taxa were observed at low relative abundances in each well (<1%) on most sampled dates (Table 1; Fig. 4). In other words, only 6 of the shared groups (*Comamonadaceae*, Burkholderiales *incertae sedis*, unclassified γ-Proteobacteria, *Rhodocyclaceae*, *Oxalobacteraceae*, and *Sphingomonadaceae*) were moderately to highly abundant (>1%) and consistently present (in at least half the samples in each well). Additionally, 18 non-shared groups were observed to be highly abundant (>5%) in at least one groundwater sample over the tested time period (Fig. 3B). At the phylum level, each well experienced periodic blooms of Firmicutes and Deinococcus-Thermus; however, the dates and intensities of these blooms were unique in each well. These results suggest that while only 6 taxa were observed to consistently predominate groundwater samples, an additional 18 taxa transiently predominated over time in groundwater. Shifts in relative abundance of the 6 shared and consistently dominant bacterial groups combined with the transient presence of 18 additional taxa (that are dominant on some days but lowly abundant in others) across all wells contribute to a high degree of dissimilarity in bacterial groups in each well. The wells are between 56 and 79% dissimilar to each other with regard to the OTUs present at any given time in any given well (for OTUs >5%) as measured via

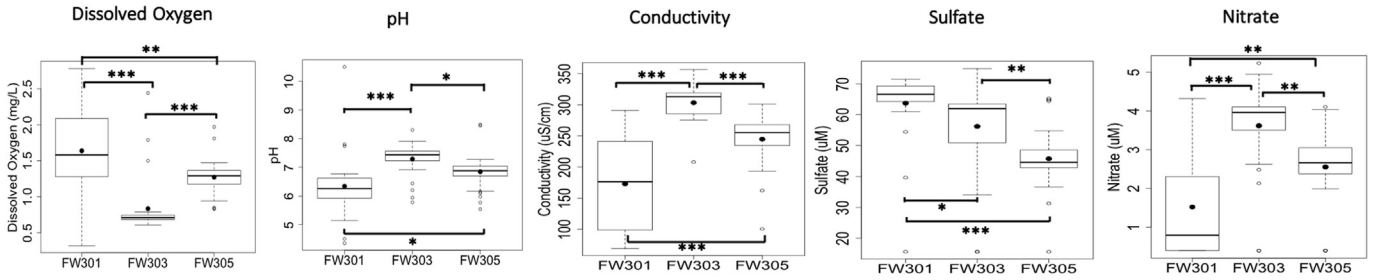

Fig. 1. Tukey-style boxplots of within-well geochemistry for 3 wells with $n = 28$ for each well. Mean and median are denoted by a filled-in dark circle and horizontal line, respectively, within each boxplot. Significance testing was carried out via a one-way ANOVA (DF 2). Analysis of variance showed that wells differed significantly in dissolved oxygen ($F(2,81) = 25.49$, $p = 2.59e\text{-}09$), pH ($F(2,81) = 9.814$, $p = 0.000153$), conductivity ($F(2,81) = 51.05$, $p = 4.52e\text{-}15$), sulfate ($F(2,81) = 16.18$, $p = 1.23e\text{-}06$), and nitrate ($F(2,80) = 24.82$, $p = 4.12e\text{-}09$). Post-hoc analysis (pairwise t-tests with holm adjustment) were used to differentiate differences between groups. All three wells differed significantly from each other. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

the Sorensen index (Fig. S5A), and on average about 80% dissimilar with regard to the spatial and temporal shifts in the abundances of shared OTUs as measured by the Bray-Curtis dissimilarity index (Fig. S5B).

Within each well, a set of OTUs were identified as core OTUs if they were present in every sampled time point obtained for a given well. The degree of transience was determined based on the percentage of all samples for which that OTU was not observed/missing/absent from the dataset of that well. Each well had a unique set of core OTUs, with FW301 having the greatest total number of core OTUs (106) and FW305 having the least (16 OTUs), (Fig. 5A–C).

3.4. Microbial community structure is highly variable within each well over time

Within-well variability of OTUs present and the proportional representation in the respective community were observed to be very different from sampling event to sampling event, and the degree of temporal variability within a single well was largely unexpected. Between sampling points, the OTUs present in any given well differed on average between 41 and 74% (Fig. S6, Table S1), which contributed to significant differences in within-well diversity in terms of richness in FW301 ($p\text{-}value = 0.05$) and in the inverse Simpson diversity measure in FW305 ($p\text{-}value = 0.034$) from week to week (Fig. 2B). While the 20 shared bacterial groups represented a level of stability in the system, temporal relative abundances of these taxa changed uniquely in each well (Fig. 2A). Daily cumulative relative abundances of the 20 shared groups fluctuated over time, ranging from 1.9 to 89% of the total daily measured groundwater diversity (Fig. 3). However, in FW303, the 20 groups maintained a cumulative relative abundance of 50% or greater in all samples. Therefore, in FW303, shared groups made up over half of the community on all sampled time-points, indicating a level of structural stability not observed in the other two wells.

SIMPER analysis was used to calculate the Bray-Curtis dissimilarities between samples based on specific populations within each sample, thus calculating the contributions of specific populations to the weekly changes in community structure in each well. In FW301, OTUs mostly classified as β-Proteobacteria (*e.g.*, Burkholderiales, *Comamonadaceae*, *Rhodocyclaceae*) contributed significantly to the weekly differences (Fig. S6, Table S1). However, FW301 was unique in that the cumulative abundance of contributing OTUs was typically less than 40% of the total measured abundance for any given date. This data indicates that the majority of OTUs contributing to weekly dissimilarities in FW301 were moderately to rarely abundant on any given day, with periodic blooms of different OTUs contributing largely to weekly dissimilarities. In FW303, changes in relative abundances of the consistently present Burkholderiales *incertae sedis* contributed highly to the weekly differences. Additionally, unlike in FW301, the majority of OTUs with the most contributions to weekly dissimilarities were often highly abundant. This indicates that the differences in this well were attributed to changes in the consistently abundant, as opposed to the rarely transient, populations. FW305 had the highest proportion of *Comamonadaceae* populations; however, a steady decline in relative abundances can be observed over time (Figs. 3A and 4, S6). *Rhodocyclaceae* and *Oxalobacteraceae* also contributed significantly to changes in community structure over time in FW305 (S6).

An unconstrained ordination revealed patterns in FW305 not observed in the other two wells, and distinct shifts in the bacterial taxa were observed that separated the sampled communities by time of sampling (Fig. 6A). Samples collected in July/August were clustered, followed by a sudden shift in the community in mid-August. Communities sampled in the late August and September cluster distinctly apart from those of the earlier months. The last two time points, sampled in October, are dramatically different in overall composition than any of the samples. Interestingly, FW301 and FW303 were cored approximately 20 years ago, while FW305 was cored 8 weeks prior to the start of this study. The unconstrained analysis may reveal unique successional dynamics occurring as the surrounding subsurface environment was disturbed during the coring and the bacterial community responds to altered flow and/or hydrochemical conditions.

3.5. Functional gene diversity displays less spatial and temporal dissimilarity

Based on ordination plots of functional gene arrays, the functional gene diversity (based on clusters of orthologous groups, or COGs) of the groundwater communities in the tested wells was more similar over time than the phylogenetic diversity (Fig. 6). Additionally, the relative abundances of observed COGs were less variable over time as compared to the SSU rRNA gene sequences (Fig. 6). The functional potential within and across samples remained relatively constant throughout the three-month study (Fig. 7 and Table 2). All three wells were enriched in COGs classified as metal homeostasis (~27%), stress (~16%), carbon cycling (~16%) and virulence (~15%). Previous metagenomic comparisons between uncontaminated and contaminated groundwater have shown enrichment in different COGs of presumptive cation/metal transporters (Hemme et al., 2015).

3.6. Changes in diversity relate to hydrochemical fluctuations uniquely in each well

Canonical correlation analysis (CCA) indicated potential hydrochemical drivers of community composition in each given well (Fig. 8). According to the CCA plots, changes in fluoride, conductivity, sulfate, and DO are the greatest drivers for differences in community composition between the three background wells, particularly FW301 and FW303. The results also suggested that acetate differentiated FW305. One limitation of correlation analysis is that it might not capture potential delays in response between changes in geochemistry and diversity within a well, and vice versa. Additionally, time-series data are often inherently dependent and therefore do not meet the requirement of statistical independence (Pfaff, 2008).

To account for potential effect of time delayed responses as well as issues with dependent data, vector autoregressive models (VAR) were used (Fig. 9, Table 3) to model the relationship between changes in richness and hydrochemistry (including measured precipitation) over time in each well. For FW301,the results of the VAR models indicate that both dissolved oxygen and chloride are significantly affected by increases in richness. In the case of dissolved oxygen, increases in richness corresponded to decreases of DO concentrations after 6 days (p-value, 0.035). On the other hand, increases in richness were modeled to precede increases in chloride with a delay of 2 days (p-value, 0.022). For FW303, ORP and conductivity had significant relationships with changes in richness. Increases in richness were modeled to occur 4 days after increases in conductivity (p-value, 0.00195). Interestingly, while the CCA plot showed that ORP only had a minor influence in driving community composition in any well, the VAR model indicated a strong bi-directional Granger causality between ORP and richness following a 4-day time-delay (Table 3).

Variables are considered to be cointegrated when long-run equilibria are displayed (Masih and Masih, 1996). Relationships in FW305 showed a high degree of cointegration (data not shown)

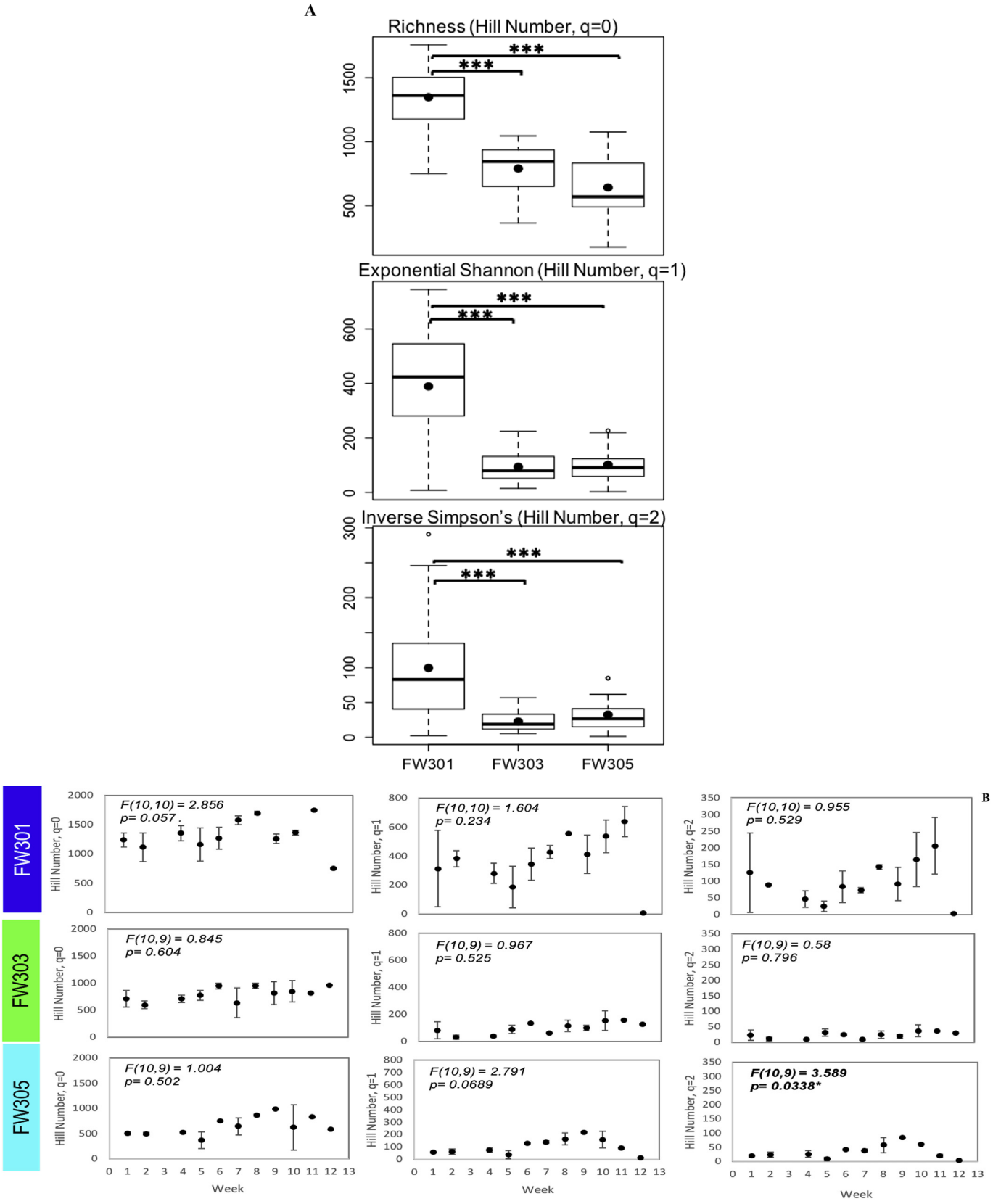

Fig. 2. A) Tukey-style boxplots of measured within-well groundwater diversity based on Hill Numbers for q = 0 (Richness), q = 1 (Exponential Shannon) and q = 2 (Inverse Simpson's). Mean (horizontal line) and median (filled-in dark circle) are denoted in each boxplot. Significance testing was carried out via a one-way ANOVA (N = 61). Post-hoc analysis (Tukey's Honest Significant Difference) was performed to determine differences between groups. FW301 had greater OTU richness, Exponential Shannon, and Inverse Simpson's than FW303 ($F(2,54) = 43.3$, p = 6.0e-12); ($F(2,54) = 33.14$, $p = 4.1e-10$; ($F(2,54) = 13.69$, $p = 0.0000347$, respectively), and FW305 ($F(2,54) = 43.3$, $p = 6.0e-12$; $F(2,54) = 33.14$, $p = 4.1e-10$; $F(2,54) = 13.69$, $p = 0.000332$, respectively). FW301 also has a greater interquartile range than FW303 and FW305 for Exponential Shannon (IQR = 259.2;

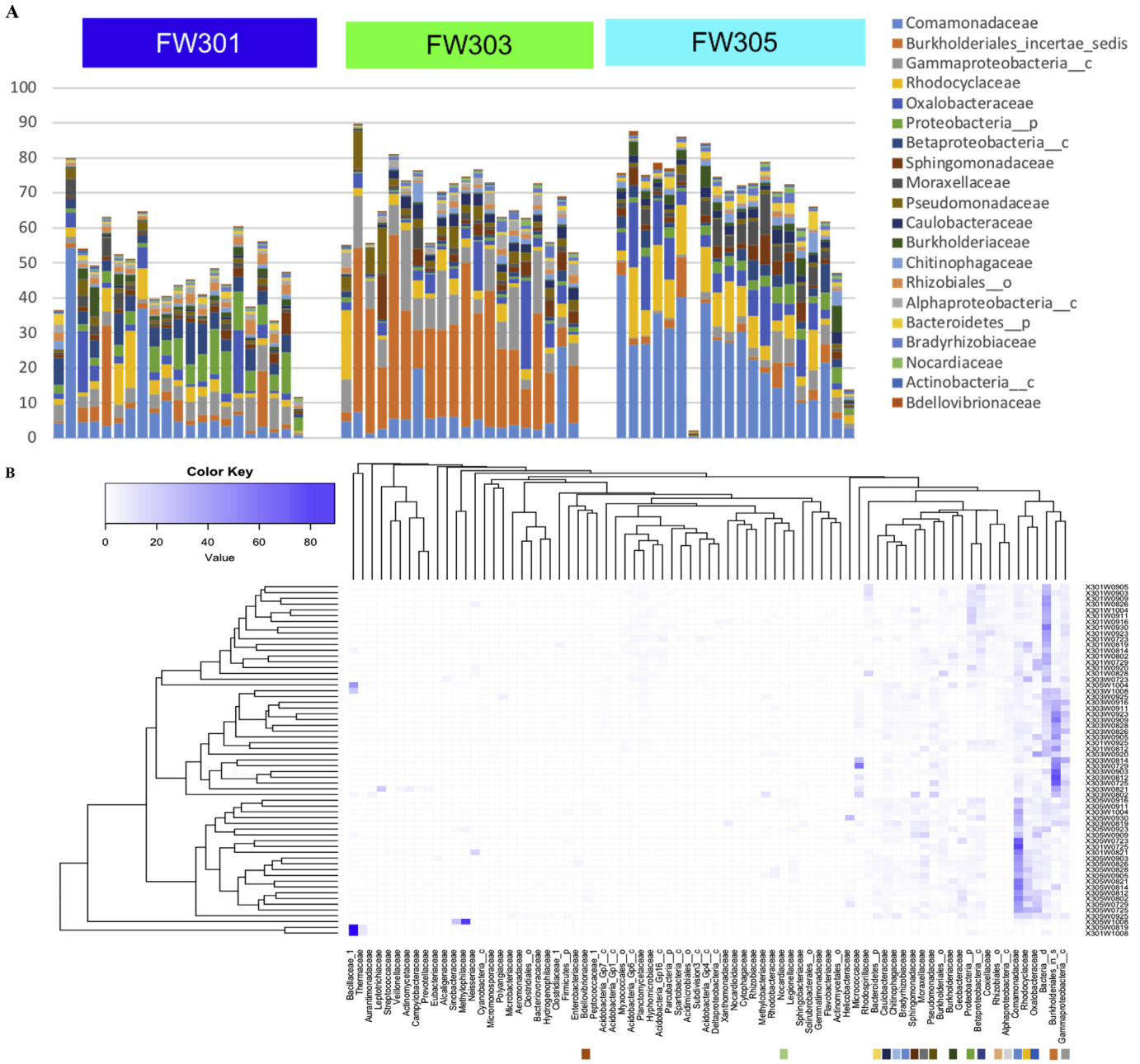

Fig. 3. A) 20 shared groups observed in all groundwater samples. Out of 218 total unique taxa identified, only 20 were found in all groundwater samples. (~1%). These 20 taxa account for 1.9–89.6% of the daily measured groundwater diversity. All 20 taxa were found in >1% abundance in at least one groundwater sample over time. B) Transiently abundant taxa found in >1% in at least one groundwater sample. 18 of these taxa are found at >5% relative abundance in at least one sampled time point.

such that traditional time-series analysis using VAR models would have resulted in spurious results (Pfaff, 2008). All modeled variables for FW305 were cointegrated with richness, with the exception of the chloride and fluoride anions. Future work includes development and analyses with more appropriate models (*e.g.*, Vector Error Correction Models) for the FW305 dataset.

3.7. Different ratios of deterministic and stochastic forces

In sampled groundwater of the tested wells, homogenous selection was the dominant calculated process, with the majority of βNTI comparisons being < -2, suggesting strong biotic or abiotic pressures worked to select for closely related taxa from one day to the next (Table 4). Based on the βNTI analysis, each well was also at

83.9; 64.1, respectively) and Inverse Simpson's (IQR = 95; 21.9; 26.3, respectively) but not richness (IQR = 307; 275; 324.5, respectively). This indicates higher temporal variability in FW301 for measured abundant and dominant OTUs, but not for measured richness. Results are similar for rarified samples. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. B) Plots show weekly means and standard error bars (SEM) for within-well groundwater diversity. A total of 2 days per week (n = 2) were analyzed, with the exception of weeks 9 and 11 where only one date was sampled in wells FW305 and FW303, respectively, and week 12 were only one date was sampled in all three wells. Week 3 shows no weekly mean because no samples were collected that week. Test of significance was carried out via a one-way ANOVA in R on all samples taken over 11 weeks for a total of 21 samples in well FW301, and 20 samples in the other two wells.

Table 1
Number of times shared bacterial taxa appeared at >1% over 3-month sampling period. On most days, shared taxa were lowly abundant in groundwater communities. Only 6 shared taxa appeared at >1% on more than half of the days sampled in all wells (bold).

Bacterial Family	FW301 (n = 21)	FW303 (n = 20)	FW305 (n = 20)
Comamonadaceae	20	20	19
Burkholderiales (*incertae sedis*)	18	20	12
γ-Proteobacteria (class)	20	20	13
Rhodocyclaceae	20	20	19
Oxalobacteraceae	17	16	18
Proteobacteria (phylum)	21	3	17
β-Proteobacteria (class)	20	8	16
Sphingomonadaceae	12	14	16
Moraxellaceae	9	95	16
Pseudomonadaceae	9	20	4
Caulobacteraceae	0	16	18
Burkholderiaceae	11	0	10
Chitinophagaceae	5	13	11
Rhizobiales (order)	18	10	1
α-Proteobacteria (class)	14	15	0
Bacteroidetes (phylum)	5	2	15
Bradyrhizobiaceae	1	12	13
Nocardiaceae	0	1	2
Actinobacteria (class)	0	0	0
Bdellovibrionaceae	0	0	2

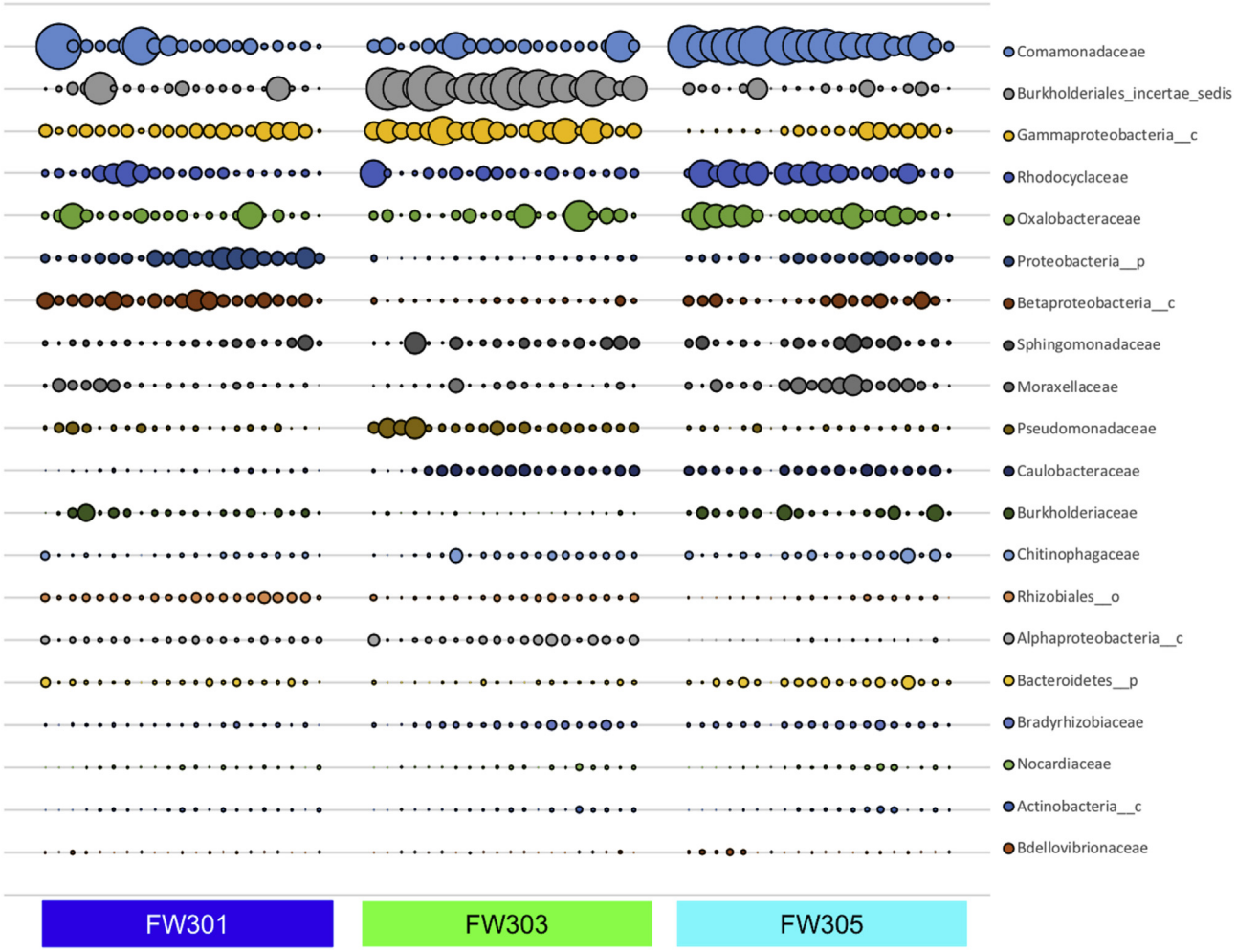

Fig. 4. Bubble plot of relative abundances at the lowest classifiable level for shared groups. Out of all unique taxonomic groups, only 20 were identified in all samples of all 3 wells. Over time, these taxa account for 1.9%–89.6% of the daily measured groundwater diversity.

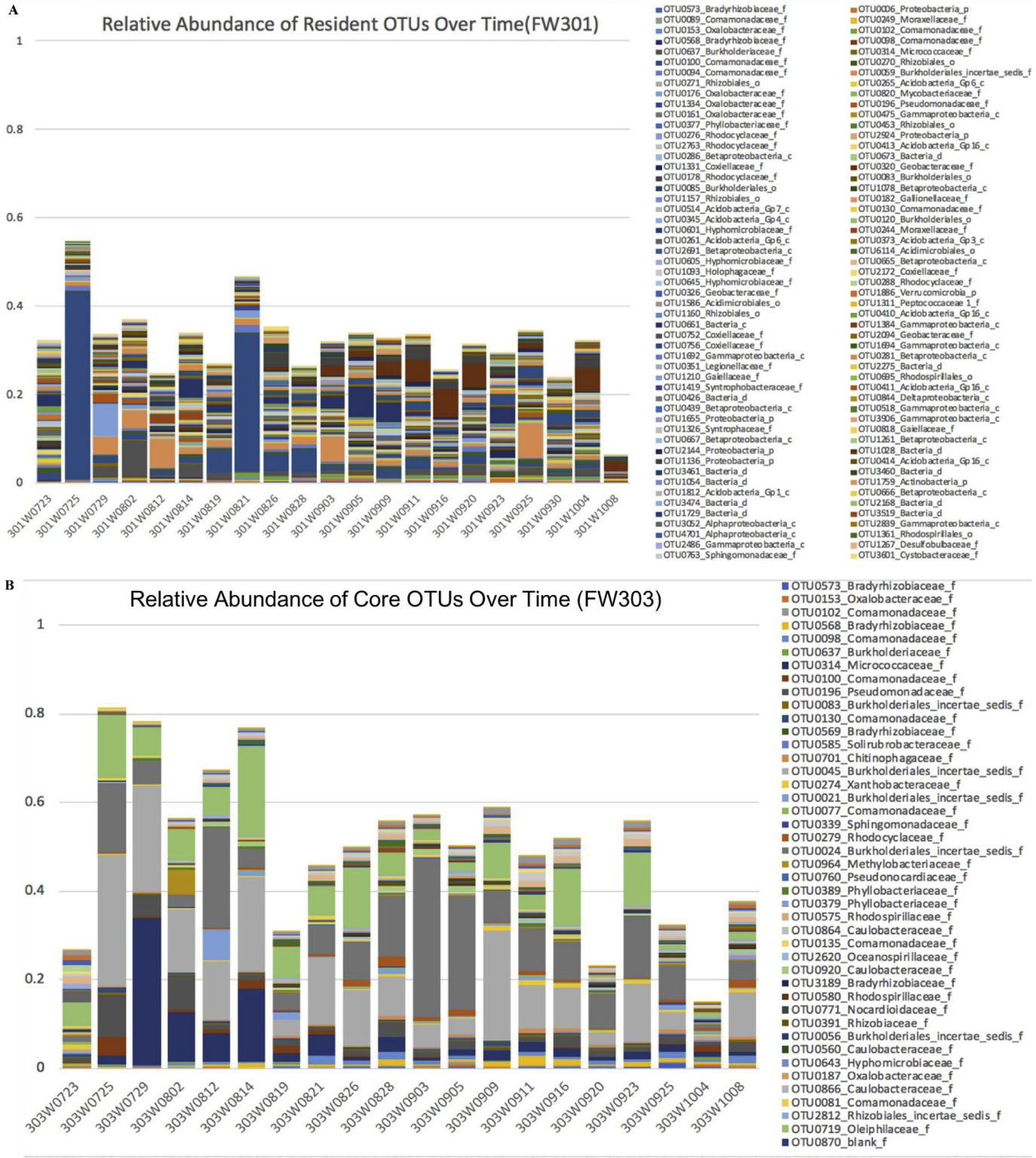

Fig. 5. Relative abundance bar charts of resident OTUs observed in A) FW301, B)FW303 and C) FW305 over time. OTUs were considered resident if they were observed in every sample analyzed for the three-month sampling period. While a total of 3372 (FW301), 2277 (FW303) and 2129 (FW305) OTUs were observed in each well, respectively, over the three-month sampling period only 106(FW301), 58(303), and 16(FW305) resident OTUs were observed over time in each well, respectively.

least partially influenced by both deterministic and stochastic processes, although each well was influenced by either process in different proportions (Fig. 10). Of all wells, FW301 was the most affected by stochastic processes (35%), which agrees with the high temporal variability observed in this well. FW305 was primarily governed by strong selective pressures (93% homogeneous selection) and was the well that was least influenced by undominated (stochastic) processes (7%). This result supports the idea that this

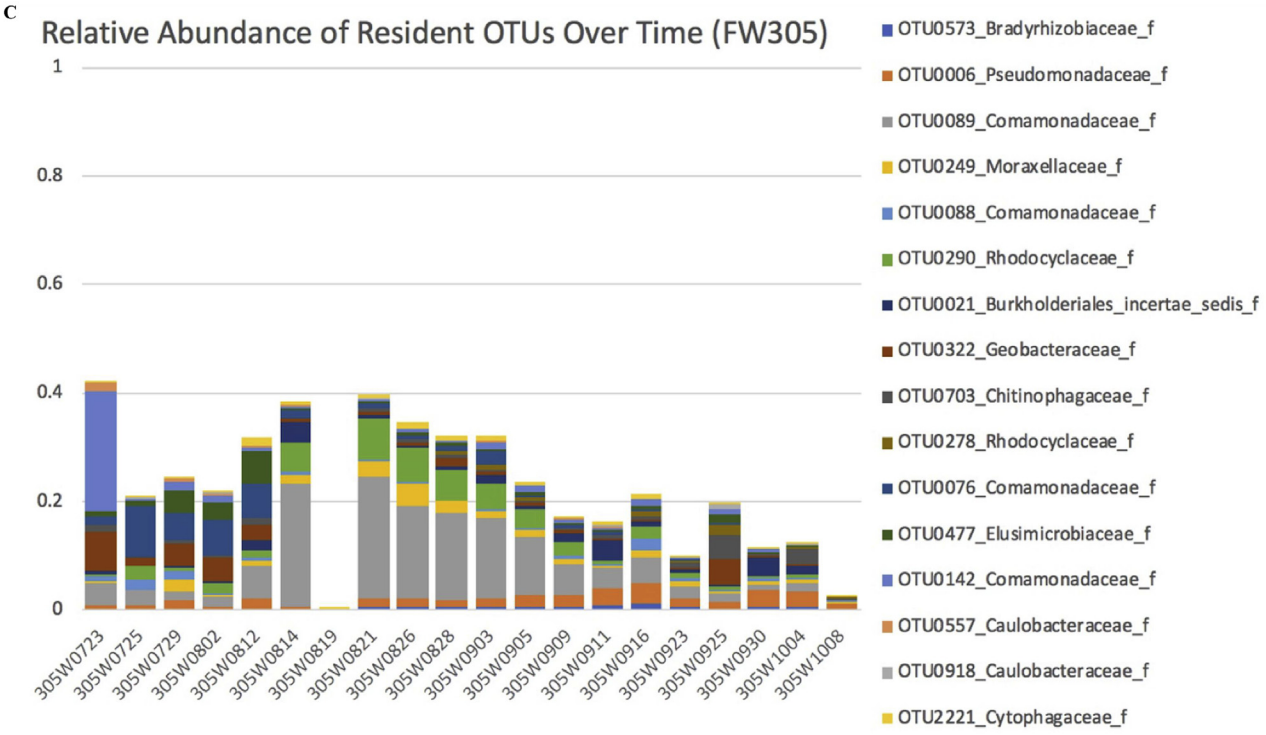

Fig. 5. (*continued*).

newly constructed well is undergoing deterministic selective pressures to shape microbial community composition.

4. Discussion

Groundwater systems and their indigenous microorganisms play a vital role in sustaining life, both above and below the surface through major contributions to the water cycle (Goldscheider et al., 2006). Understanding the factors that affect microbial community structure and function is critical to meet present and future global demands of water maintenance and utilization. However, the connection between microbial diversity, biogeohydrochemical processes, and the flow of energy and resources between shallow subsurface ecotones is poorly understood. In addition, shallow subsurface ecotones include both groundwater and porous medium, and therefore, both aqueous and solid media are important and unique niches that impact resource allocation. The current study focused on the variability of the aqueous phase of an uncontaminated, oligotrophic aquifer.

Despite having hydrological, chemical, and geological heterogeneity, most groundwater systems are thought to be constant and predictable due to the nature of oligotrophic environments through similar transition zones (Griebler and Lueders, 2009). Based upon these common sentiments, temporal hydrological events have been hypothesized to impact changes, particularly via redox shifts caused by hydraulic events over the time span of seasons (Vroblesky and Chapelle, 1994; Haack et al., 2004; Zhou et al., 2012). More recently, temporal dynamics have also been attributed to environmental disturbances, for example nutrient influxes (*e.g.*, hot spots and hot moments) at the interface between terrestrial and aquatic ecosystems (McClain et al., 2003), or in the case of water systems, the influx of additional water from the surrounding environment. For example, it has been shown that there can be direct transfer of nutrients and microbial populations within transitionally saturated and permanently saturated zones (Bougon et al., 2011; Yabusaki et al., 2017). Studies of the hyphoreic zone along the Columbia River in Washington, USA have shown that seasonal changes in river height and overflow due to increased rainfall result in increased diversity of aquifer microbial communities (Lin et al., 2012). However, the opposite effect on diversity has also been documented. Rising groundwater levels due to seasonal effects on water table recharge were shown to accompany a drop in bacterial diversity (attributed to dilution effects) in spring and summer months in an alpine oligotrophic porous aquifer (Zhou et al., 2012).

The described study focused on the temporal dynamics of an uncontaminated groundwater system with limited hydraulic connectivity between sampling wells into groundwater housed within the same underlying geological formations. Both geochemistry and microbial communities were uniquely dynamic over space and time with varying degrees of fluctuation intensities and sparsely shared populations. In fact, few bacterial groups were shared by the 3 wells over the measured time span. The results indicated a high degree of transience in groundwater bacterial taxa with the majority (approximately 93%) being transient over time.

In this study, rainfall did not correlate to measured changes in diversity or hydrochemical fluctuations in the tested wells (data not shown). However, it is worth noting that the study was conducted during the transition period between the drier summer months of July–August, and the fall months of September–October (all rain events below 2 cm except September 21). The high and low points for the background water table have historically been in the spring (March–April) and late fall (October), respectively (Kim et al., 2009; Revil et al., 2013), but the nature and duration of precipitation events has been changing (*i.e.*, climate change) and further work is needed to better understand the temporal and spatial impacts of perturbation. Given that recharge in this aquifer occurs primarily via heavy rainfall events and vertical percolation of excess water to the saturation zone (Solomon et al., 1992), which typically occurs during the wetter winter months, the results indicate a high degree

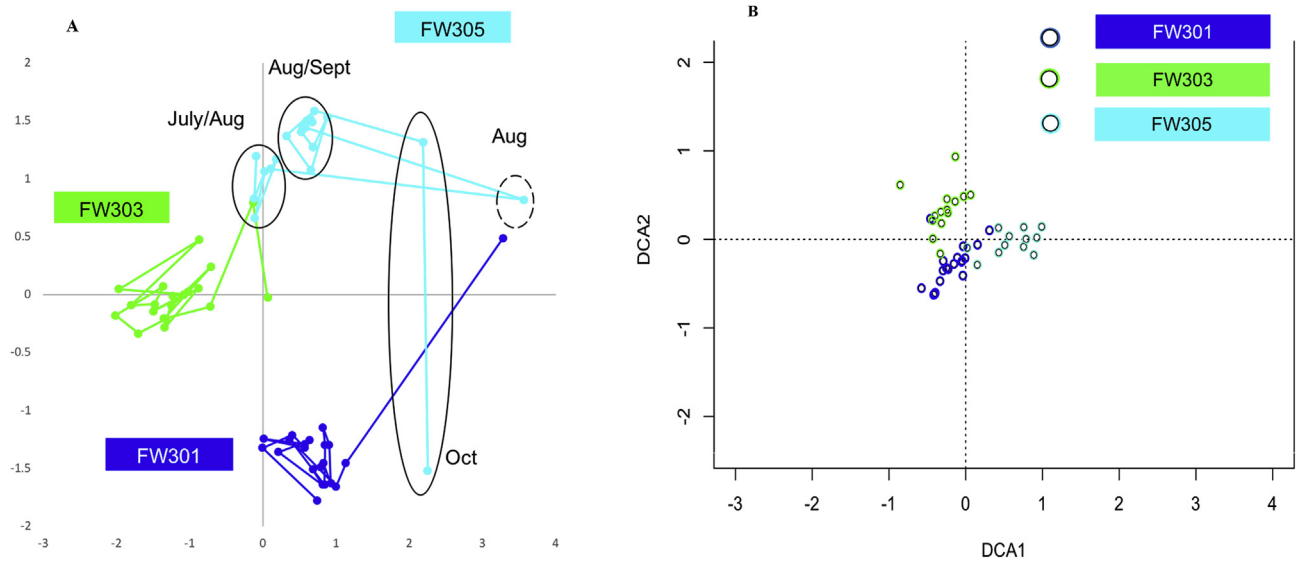

Fig. 6. A) Unconstrained ordination (DCA) of groundwater OTUs per sample. Solid lines connect consecutively sampled dates. Dashed circle represents a major shift in community composition in mid-August. B) Samples are ordinated based on results of Functional Gene Array (FGA). Samples are indicated by colored circles. Based on DCA, samples are much more similar with respect to potential function (FGA) than community structure.

of variable biodiversity during the dry period. Future work is underway to better understand microbial community dynamics during the hydrological year for the uncontaminated and contaminated carbonate-rock aquifers and the implications for ecosystem function.

Temporal dynamics of microbial communities have been studied in other environments, such as the human microbiome (Koenig et al., 2011), plant-flower (Shade et al., 2013a), surface soils (Wertz et al., 2007), and other aquatic systems (Hofle et al., 1999; Shade et al., 2007). A meta-analysis surveying the temporal variability across various biomes (Shade et al., 2013b) demonstrated that areas with high physicochemical instability (e.g., stream, human palm) varied considerably over time, while more *relatively* stable environments (*e.g.*, soil) varied less considerably, suggesting that microbial communities might vary predictably across environments over time depending on physical and chemical gradients. Additionally, environments likely undergoing succession (*e.g.*, infant gut, flower surfaces) display high levels of diversity and community composition shifts over time. In the described study, groundwater from a pristine aquifer previously believed to be hydrochemically stable displayed temporal changes of geochemistry (*e.g.*, ORP, DO, pH, and conductivity) to varying extents over short time scales (days) that were statistically significant (Fig. 1). For FW301, the VAR models indicated that increases in species richness corresponded to subsequent decreases in DO levels (6 days) and increases in chloride (2 days). A possible explanation could be altered population distributions are linked to microbial activity that impacts DO consumption as well as anion/cation balances between groundwater and sediments. In addition, fine-scale changes (micro-environments) in pH associated to microbial activity could also impact anion/cation interactions with sediments. For FW303, ORP and conductivity were predicted to have bidirectional causal relationships with species richness. Similar to FW301, presumptive microbial activity could be linked to the population distributions and thereby impact ORP and/or balance of anion/cations between groundwater and sediments. Further work is needed to delineate these relationships with actual bacterial activity, for example, bench-scale packed bed reactors in which these field parameters can be controlled.

It is typically assumed that low levels of biodiversity may allow a conditionally consistent ecosystem to function efficiently while greater biodiversity would be needed in fluctuating environments (Humbert and Dorigo, 2005). Spatiotemporal patterns in the distribution of bacterial communities have been observed along geochemical gradients, such as those along acid-mine drainage (Volant et al., 2014) and hydrologically dynamic environments (e.g. high groundwater-surface water mixing) such as hyphoreic zones of riverine systems (Stegen et al., 2016). Most of the work that has been done with pristine aquifers that are presumed to be more hydrodynamically stable suggests more constant microbial communities with low diversity mainly attributed to a more consistently oligotrophic environment (Griebler and Lueders, 2009). However, with greater depth of spatial and temporal analyses, our results indicate that a shallow groundwater system does experience hydrochemical and microbial taxa changes on short time frames. Although, GeoChip analysis of functional diversity predicted a functionally stable community, and this result could coincide with the notion of functional redundancy (Naeem and Li, 1997) in which high turnover of low abundance microbial groups could contribute significantly to maintenance of overall functions (Jousset et al., 2017). Moreover, different hydrochemical parameters were shown to have significant relationships with changes in richness and diversity (Figs. 8 and 9, Table 3), but these relationships were different for each well. In fact, although homogenous selection was estimated to be the dominant ecological process for the three tested wells, each well was predicted to have a unique ratio of deterministic and stochastic selection forces according βNTI analyses. Future work is needed to track baseline system parameters over time in order to better understand perturbation impacts (*e.g.*, extreme weather) and the impact on long-term system functions.

In the most newly constructed well (FW305), unique relationships observed suggest that this newly cored well may have been undergoing major shifts in community composition likely attributed to succession events following the coring event (Fig. 4). The microbial variability in FW305 occurred among the most dominant community members. The DCA ordination also showed evidence of major compositional shifts over the three-month sampling period.

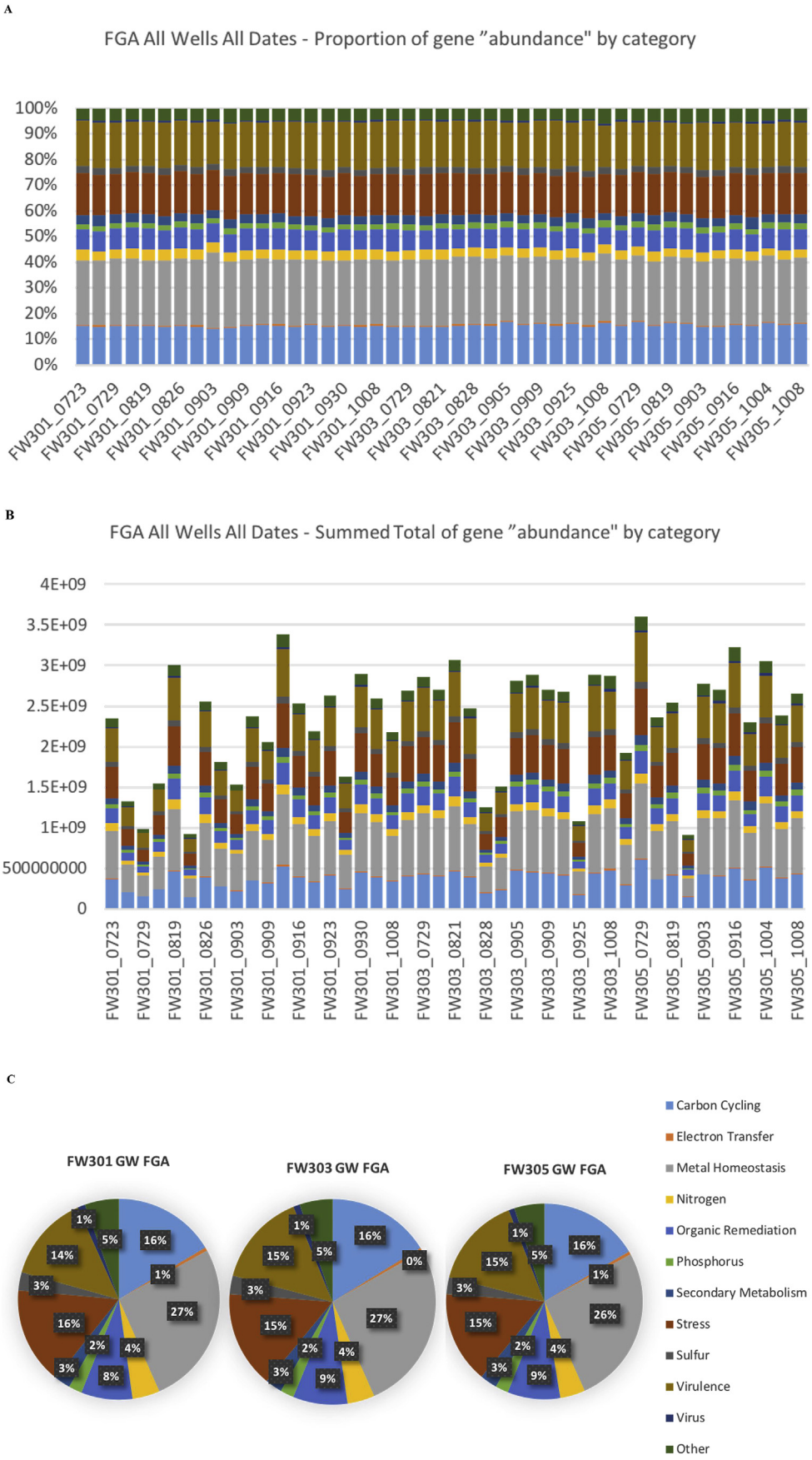

Fig. 7. Functional Gene Array (GeoChip). A) Proportion of individual genes (by category) remain consistent within and across wells. B) Summed total of gene signal abundance by category. C) Pie charts representing total proportion of COG's by category.

Table 2
Number of enriched clusters of orthologous groups (COGs) based on GeoChip data and presumptive function.

Potential Function	FW301	FW303	FW305
Carbon cycling	14,851	10,629	8126
Electron transfer	498	349	256
Metal homeostasis	23,990	17,478	12,954
Nitrogen	3978	2800	2034
Organic remediation	7449	5589	4277
Phosphorus	1934	1403	1054
Secondary metabolism	2268	1739	1388
Stress	14,226	9968	7467
Sulfur	2681	1914	1378
Virulence	12,635	9621	7436
Virus	1106	677	487
Other	4946	3381	2434
Total	**90,562**	**65,557**	**49,291**

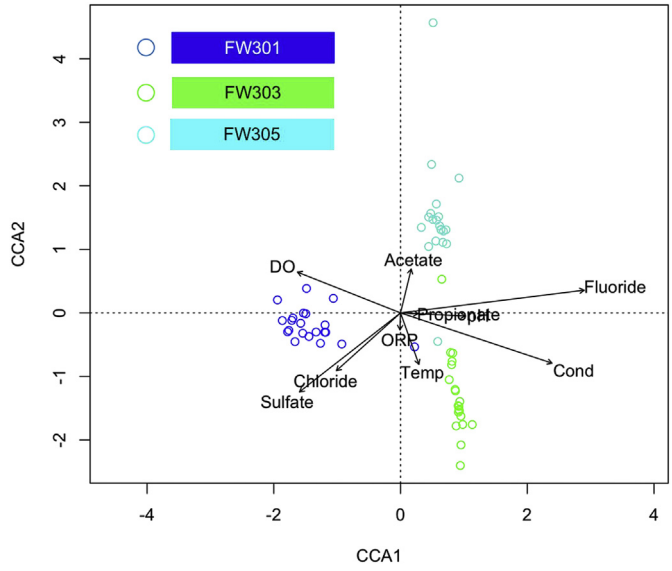

Fig. 8. Multivariate ordination showing associations between groundwater community structure and hydrochemical variables using Canonical correspondence analysis (CCA). Small black dots represent individual OTUs, larger colored circles represent samples for each well, and black vectors represent association between communities and geochemistry. Strength of association is inferred from vector length.

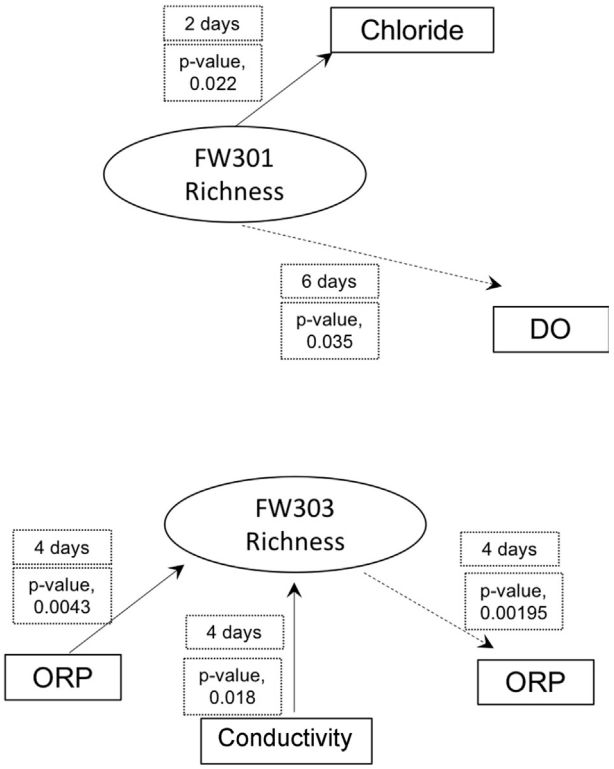

Fig. 9. Association networks based on results of Vector Autoregressive Models of diversity based on Hill numbers and hydrochemical relationships. Significant lagged relationships (p-value = ≤0.05) and the lag value in days are shown. Positive associations are represented by solid edges, while negative associations are represented by dashed edges. The direction of the arrow indicates the direction of the delayed relationships - the node preceding the arrow is ahead of the node which follows the arrow.

In addition, richness in FW305 was cointegrated with 10 of the 12 hydrochemical parameters measured. By definition, data that are co-integrated must display Granger causality in at least one direction (Pfaff, 2008). In other words, evidence of co-integration between variables "rules out Granger non-causality", implying either unidirectional or bidirectional Granger causality (Masih and Masih, 1996). Since FW305 was only 8 weeks old prior to the start of sampling, the co-integration observed in this well likely describe the long-term inter-connected relationships between hydrology, geology, and microbial populations that became disrupted due to well construction. Further work is needed to track temporal variability of planktonic and sediment associated microbial communities in established and new wells, including sediment surrogates (King et al., 2017; Smith et al., 2018) to better understand community dynamics impacted by coring.

Interestingly, the well with the highest overall hydrochemical and microbial variability, FW301, also has a screen interval roughly 3x larger than FW303 and 2x larger than FW305. As such, it is possible that the sampled formation water from FW301 originated from a greater range of depths of the vertical subsurface profile, which could have resulted in greater mixing of different waters and higher measured variability in both sampled bacterial diversity and hydrochemical parameters. The resulting greater sampling range may have resulted in the greater sampling of microorganisms originating from shallower depths of the soil profile, variable infiltration of total dissolved solids from the surrounding weathered saprolite, and/or higher hydrochemical perturbations (*e.g.*, DO, organic). Differences in environmental parameters across depth that would lead to heterogeneous observations corroborate the idea that changes in hydrology and groundwater chemistry impact groundwater microbial communities (Flynn et al., 2013). Future work will focus on soil sampling across the transition zones and water sampling from wells with discrete depth intervals across longer time periods in order to further elucidate temporal relationships between hydrochemistry and bacterial diversity.

The existence and ubiquitous nature of rare microbial taxa in all types of environments has been well documented (Sogin et al., 2006; Bent and Forney, 2008; Fuhrman, 2009; Pedrós-Alió, 2012), and more recent analysis of temporal studies (air, water, soil, human) have also documented the presence of rare taxa that occasionally bloom to dominant members of the community or remain at low abundance (Shade et al., 2014; Jousset et al., 2017). The "conditionally rare taxa" (CRT) were recently estimated to comprise up to 28% of community membership of a variety of different ecosystems, represent a broad diversity of bacterial and archaeal lineages, and explain large amounts of temporal community dissimilarities (Shade et al., 2014). While the presence of periodically abundant groups were observed in all three wells, FW301 had a much higher contribution of CRT to the dominant taxa than the

Table 3
Results of VAR model and Granger causality using the Wald Test for a given lag (p) chosen that eliminated autocorrelation for FW301 (a), FW303 (b), and FW305 (c) for species richness and geochemistry. Blank spaces for FW303 and FW305 resulted because the VAR model could not be performed due to non-stationarity. ORP, oxidation-reduction potential; DO, dissolved oxygen.

(a)

Richness "Granger Causes"	p-value	Richness is "Granger Caused"	p-value
Temperature	0.08	Temperature	0.50
Conductivity	0.36	Conductivity	0.26
ORP	0.07	ORP	0.37
Nitrate	0.78	Nitrate	0.10
Sulfate	0.01	Sulfate	0.99
pH	0.5	pH	0.11
DO	**0.0003**	DO	0.09
Fluoride	0.58	Fluoride	0.16
Chloride	**0.04**	Chloride	0.99
Precipitation	0.90	Precipitation	0.65

(b)

Richness "Granger Causes"	p-value	Richness is "Granger Caused"	p-value
Temperature	0.64	Temperature	0.76
Conductivity	0.18	**Conductivity**	**0.02**
ORP	**0.001**	**ORP**	**0.006**
Nitrate	0.09	Nitrate	0.96
Sulfate	0.06	Sulfate	0.93
pH	0.72	pH	0.62
DO		DO	
Fluoride	0.64	Fluoride	0.96
Chloride	0.40	Chloride	0.61
Precipitation	0.98	Precipitation	0.78

(c)

Richness "Granger Causes"	p-value	Richness is "Granger Caused"	p-value
Temperature		Temperature	
Conductivity		Conductivity	
ORP (oxidation-reduction potential)		ORP	
Nitrate		Nitrate	
Sulfate		Sulfate	
pH		pH	
DO (dissolved oxygen)		DO	
Fluoride	0.84	Fluoride	0.54
Chloride	0.82	Chloride	0.68
Precipitation		Precipitation	

Table 4
Percent contributions of different community assembly processes as calculated by βMNTD and βNTI.

Well	Mean βNTI	βNTI Stad. Dev.	Homogenous Selection	Variable Selection	Stochastic Processes	% Total Process
FW301	−2.28	1.05	65%	0%	35%	100%
FW303	−2.95	1.80	76.3%	1.6%	22.1%	100%
FW305	−4.84	1.54	93%	0%	7%	10%

other two wells, coinciding with higher overall diversity and hydrochemical instability over the sampled time period. Whether this is ecologically significant or a consequence of screen interval length is currently unknown; however it may suggest an important role for rare taxa in the localized area of the aquifer. In contrast, sampled groundwater from a newer well, FW305, showed higher temporal variability in the dominant taxa, indicating selective pressures in this recently disturbed system.

Given the unexpected extent of temporal variability in transient bacterial taxa for uncontaminated groundwater based upon commonly used pyrotag data, the potential contribution of 'active' and 'inactive' cells is not known. However, the observed community dissimilarity for each well was constant over the tested time period, and the composition of the diversity changed dependent upon the well. In addition, similar predominant bacterial groups (Proteobacteria, Actinobacteria, and Bacteroidetes, Acidobacteria, Chloroflexi, and Nitrospirae) were observed as in previous groundwater studies (Griebler and Lueders, 2009; Stegen et al., 2013; Wegner et al., 2019). Moreover, analysis of the same samples with the Geochip demonstrated that functional gene diversity was more constant over the same time period. It should be noted that Geochip sensitivity (microarray) is different compared to direct sequence comparisons for the SSU rRNA genes and future work should include functional gene diversity for selected processes of interest and/or metagenomes. In addition, current work is underway to track general and specific microbial activity over time for groundwater systems to relate biodiversity and ecosystem function. Further investigations into the relationship between microbial diversity, ephemeral populations, transient hydrochemical flux, and the contributions to overall microbial activity in groundwater systems are currently underway.

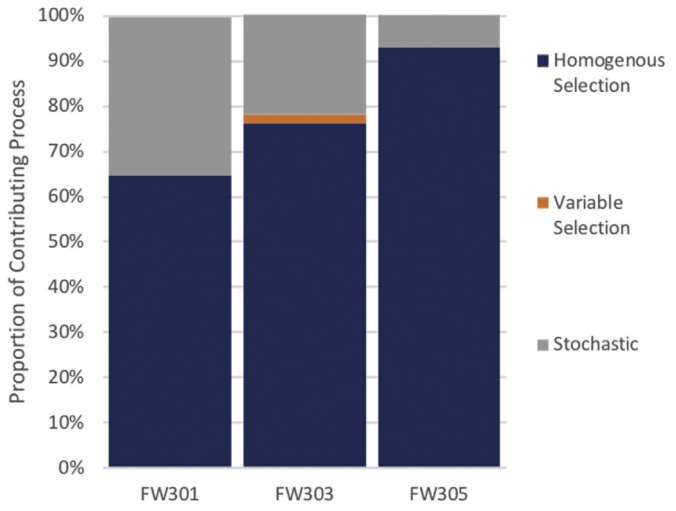

Fig. 10. Relative influence of community assembly processes on temporal turnover of within-well bacterial assemblages. Processes were calculated βNTI as described in materials and methods.

5. Conclusions

Freshwater resources are becoming jeopardized at alarming rates. Full assessment of uncontaminated groundwater composition, both biotic and abiotic components, over space and time are necessary to establish expected variability over both short- and long-term scales. Without such baselines, observed changes to groundwater composition could be erroneously attributed to anthropogenic effects or other disturbance events. This study showed for the first time that groundwater composition of an uncontaminated aquifer varies dynamically and uniquely over short time scales (days) and that each of the wells had a unique 'core' bacterial community. The different abundant and rare OTUs contributed to community dissimilarity respective to each well but homogenous selection was the dominant ecological force calculated for each well. VAR models suggested that DO, Cl^-, ORP, and conductivity had Granger causality with OTU richness.

Declaration of interests

None.

Acknowledgments

This material is funded by ENIGMA- Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231. The authors would like to thank Dr. H.J. Smith for helpful comments as well as the anonymous reviewers.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.watres.2019.114917.

References

Akob, D.M., Küsel, K., 2011. Where microorganisms meet rocks in the Earth's critical zone. Biogeosciences 8, 3531–3543.
Anderson, R.T., Lovley, D.R., 1997. Ecology and biogeochemistry of *in situ* groundwater bioremediation. Adv. Microb. Ecol. 15, 289–350.
Bates, D.M., 2010. Lme4: Mixed-Effects Modeling with R. Springer. http://lme4.r-forge.r-project.org/book/front.pdf.
Bent, S.J., Forney, L.J., 2008. The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. ISME J. 2 (7), 689–695.
Bougon, N., Aquilina, L., Molenat, J., et al., 2011. Influence of depth and time on diversity of free-living microbial community in the variably saturated zone of a granitic aquifer. FEMS Microbiol. Ecol. 80, 98–113.
Caporaso, J.G., Lauber, C.L., Walters, W.A., et al., 2010. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. 108, 4516–4522.
Caporaso, J.G., Lauber, C.L., Walters, W.A., et al., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6 (8), 1621–1624.
Chapelle, F.H., 2000. The significance of microbial processes in hydrogeology and geochemistry. Hydrogeol. J. 8 (1), 41–46.
Clarke, J.A., Mirza, S., 2006. A comparison of some common methods for detecting Granger noncausality. J. Stat. Comput. Simul. 76 (3), 207–231.
Danielopol, D.L., Griebler, C., Gunatilaka, A., Notenboom, J., 2003. Present state and future prospects for ground- water ecosystems. Environ. Conserv. 30, 104–130.
Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–998.
Environmental Protection Agency, 1996. Guidance for Data Quality Assessment. Office of Research and Development, US EPA.
Fields, M.W., Bagwell, C.E., Carroll, S.L., et al., 2006. Phylogenetic and functional biomarkers as indicators of bacterial community responses to mixed-waste contamination. Environ. Sci. Technol. 40, 2601–2607.
Fine, P.V.A., Kembel, S.W., 2011. Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. Ecography 34, 552–565.
Flynn, T.M., Sanford, R.A., Ryu, H., 2013. Functional microbial diversity explains groundwater chemistry in a pristine aquifer. BMC Microbiol. 13, 146.
Forsythe, G.E., Malcolm, M.A., Moler, C.B., 1977. Computer Methods for Mathematical Computations. Prentice-Hall Professional Technical Reference.
Fuhrman, J.A., 2009. Microbial community structure and its functional implications. Nature 459, 193–199.
Gerth, J., Forstner, U., 2004. Part IV: fitness for aquatic systems long-term forecast: key to groundwater protection. Environ. Sci. Pollut. Res. 11, 49–56.
Goldscheider, N., Hunkeler, D., Rossi, P., 2006. Microbial biocenoses in pristine aquifers and an assessment of investigation methods. Hydrogeol. J. 14, 926–941.
Griebler, C., Lueders, T., 2009. Microbial biodiversity in groundwater ecosystems. Freshw. Biol. 54, 649–677.
Haack, S.K., Fogarty, L.R., West, T.G., et al., 2004. Spatial and temporal changes in microbial community structure associated with recharge-influenced chemical gradients in a contaminated aquifer. Environ. Microbiol. 6, 438–448.
Haegeman, B., Hamelin, J., Moriarty, J., et al., 2013. Robust estimation of diversity in theory and practice. ISME J. 7 (6), 1092–1101.
Hammer, O., Harper, D., Ryan, P.D., 2001. PAST: paleontological statistics software package for education and data analysis. Palaeontol. Electron. 4 (1), 9.
Hazen, T.C., Dubinksy, E.A., DeSantis, T.Z., et al., 2010. Deep-Sea oil plume enriches indigenous oil-degrading Bacteria. Science 330 (6001), 204–208.
Hemme, C.L., Tu, Q., Shi, Z., et al., 2015. Comparative metagenomics reveals impact of contaminants on groundwater microbiomes. Front. Microbiol. 6, 1205.
Hofle, M.G., Haas, H., Dominik, K., 1999. Seasonal dynamics of bacterioplankton community structure in a eutrophic lake as determined by 5S rRNA analysis. Appl. Environ. Microbiol. 65, 3164–3174.
Humbert, J.F., Dorigo, U., 2005. Biodiversity and aquatic ecosystem functioning: a mini review. Aquatic Ecosyst Health 8, 367–374.
Hwang, C., Wu, W., Gentry, T.J., et al., 2009. Bacterial community succession during *in situ* uranium bioremediation: spatial similarities along controlled flow paths. ISME J. 3, 47–64.
Jousset, A., Bienhold, C., Chatzinotas, A., et al., 2017. Where less may be more: How the rare biosphere pulls ecosystem strings. ISME J. 11, 853–862.
Kim, Y.-J., Moon, J.-W., Roh, Y., Brooks, S.C., 2009. Mineralogical characterization of saprolite at the FRC background site in Oak Ridge, Tennessee. Environ. Geol. 58, 1301–1307.
King, A., Preheim, S.P., Bailey, K.L., et al., 2017. Temporal dynamics of in-field bioreactor populations reflect the groundwater system and respond predictably to perturbation. Environ. Sci. Technol. 51, 2879–2889.
Koenig, J.E., Spor, A., Scalfone, N., et al., 2011. Succession of microbial consortia in the developing infant gut microbiome. Proc. Natl. Acad. Sci. U.S.A. 108, 4578–4585.
Langwaldt, J.H., Puhakka, J.A., 2000. On-site biological remediation of contaminated groundwater: a review. Environ. Pollut. 107, 187–197.
Lee, J.H., Lee, B.J., 2017. Microbial reduction of Fe (III) and SO_4^{2-} and associated microbial communities in the alluvial aquifer groundwater and sediments. Microb. Ecol. 76, 182–191.
Lesnoff, M., Lancelot, R., 2012. aod: analysis of Overdispersed Data. R package version 1.3. http://cran.r-project.org/package=aod.
Liang, Y.T., He, Z.H., Wu, L.Y.,Y., et al., 2010. Development of a common oligonucleotide reference standard for microarray data normalization and comparison across different microbial communities. Appl. Environ. Microbiol. 76 (4), 1088–1094.
Lin, X., McKinley, J., Resch, C.T., et al., 2012. Spatial and temporal dynamics of the

microbial community in the Hanford unconfined aquifer. ISME J. 6 (9), 1665–1676.
Masih, A.M.M., Masih, R., 1996. Energy consumption, real income and temporal causality: results from a multi-country study based on cointegration and error-correction modeling techniques. Energy Econ. 18, 165–183.
McClain, M.E., Boyer, E.W., Dent, C.L., 2003. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystem. Ecosystems 6, 301–312.
Naeem, S., Li, S., 1997. Biodiversity enhances ecosystem reliability. Nature 390, 507–509.
Oksanen, J., Blanchet, F.G., Kindt, R., et al., 2013. Vegan: Community Ecology Package. R-Package Version 2.0-10.
Pedrós-Alió, C., 2012. The rare bacterial biosphere. Ann. Rev. Mar. Sci. 4, 449–466.
Pfaff, B., 2007. Package "vars". http://cran.r-project.org/web/packages/vars/index.html.
Pfaff, B., 2008. Analysis of Integrated and Cointegrated Time Series with R. Springer. https://doi.org/10.1007/978-0-387-75967-8.
Preheim, S.P., Perrotta, A.R., Martin-Platero, A.M., et al., 2013. Distribution-based clustering: using ecology to refine the operational taxonomic unit. Appl. Environ. Microbiol. 79 (21), 6593–6603.
Price, M.N., Dehal, P.S., Arkin, A.P., 2009. FastTree: vomputing large minimum evolution trees with Profiles instead of a distance matrix. Mol. Biol. Evol. 26 (7), 1641–1650.
Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2-approximately maximum-likelihood trees for large alignments. PLoS One 5 (3) e9490. https://doi.org/10.1371/journal.pone.0009490.
R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Reilly, T.E., Dennehy, K.F., Alley, W.M., Cunningham, W.L., 2008. Groundwater availability in the United States. U.S. Geol. Surv. Circ. 1323, 70p. http://pubs.usgs.gov/circ/1323/.
Revil, A., Skold, M., Karaoulis, M., Schmutz, M., Hubbard, S.S., Mehlhorn, T.L., Watson, D.B., 2013. Hydrogeophysical investigations of the former S-3 ponds contaminant plumes, Oak Ridge integrated field Research challenge site. Tennessee. Geophys. 78, EN29–EN41.
Shade, A., Kent, A.D., Jones, S.E., 2007. Interannual dynamics and phenology of bacterial communities in a eutrophic lake. Limnol. Oceanogr. 52, 487–494.
Shade, A., McManus, P.S., Handelsman, J., 2013a. Unexpected diversity during community succession in the apple flower microbiome. mBio 4, e00602–e00612.
Shade, A., Caporaso, J.G., Handelsman, J., 2013b. A meta-analysis of changes in bacterial and archaeal communities with time. ISME J. 7 (8), 1493–1506.
Shade, A., Jones, S.E., Caporaso, J.G., et al., 2014. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. mBio 5 e01371-14.
Sievers, F., Wilm, A., Dineen, D., et al., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539-539.
Singh, A., Nocerino, J., 2002. Robust estimation of mean and variance using environmental data sets with below detection limit observations. Chemometr. Intell. Lab. Syst. 60, 69–86. https://doi.org/10.1016/S0169-7439(01)00186-1.
Smith, H.J., Zelaya, A.J., De León, K.B., et al., 2018. Impact of hydrologic boundaries on microbial planktonic and biofilm communities in shallow terrestrial subsurface environments. FEMS Microbiol. Ecol. 94 (12), fiy191.
Sogin, M.L., Morrison, H.G., Huber, J.A., et al., 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere. Proc. Natl. Acad. Sci. U.S.A. 32, 12115–12120.
Solomon, D.K., Moore, G.K., Toran, L.E., Dreier, R.B., McMaster, W.M., 1992. Status Report: A Hydrologic Framework for the Oak Ridge Reservation. DOE Environmental Sciences Division. https://www.researchgate.net/publication/236370038_.
Stegen, J.C., Lin, X., Fredrickson, J.K., et al., 2013. Quantifying community assembly processes and identifying features that impose them. ISME J. 7 (11), 2069–2079.
Stegen, J.C., Lin, X., Fredrickson, J.K., Konopka, A.E., 2015. Estimating and mapping ecological processes influencing microbial community assembly. Front. Microbiol. 6, 370.
Stegen, J.C., Fredrickson, J.K., Wilkins, M.J., et al., 2016. Groundwater-surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. Nat. Commun. 7, 11237.
Sui, Q., Cao, X., Lu, S., Zhao, W., Qiu, Z., Yu, G., 2015. Occurrences, sources and fate of pharmaceuticals and personal care products in the groundwater: a review. Emerg. Contam. 1, 14–24.
Trapletti, A., Hornik, K., 2018. Tseries: time Series analysis and computational finance. R package version 0.10-46. https://CRAN.R-project.org/package=tseries.
van der Gun, J., Aureli, A., Merla, A., 2016. Enhancing groundwater governance by making the linkage with multiple uses of the subsurface space and other subsurface resources. Water 8, 222.
Volant, A., Bruneel, O., Desoeuvre, A., et al., 2014. Diversity and spatiotemporal dynamics of bacterial communities, physicochemical and other drivers along an acid mine drainage. FEMS Microbiol. Ecol. 90 (1), 247–263.
Vroblesky, D.A., Chapelle, F.H., 1994. Temporal and spatial changes of terminal electron-accepting processes in a petroleum hydrocarbon-contaminated aquifer and the significance for contaminant biodegradation. Water Resour. Res. 30, 1561–1570.
Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. 73 (16), 5261–5267.
Watson, D.B., Kostka, J.E., Fields, M.W., Jardine, P.M., 2004. The Oak Ridge Field Research Center Conceptual Model. DOE Environmental Sciences Division. https://public.ornl.gov/orifc/FRC-conceptual-model.pdf.
Webb, C.O., Ackerly, D.D., Kembel, S.W., 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. Bioinformatics 24, 2098–2100.
Wegner, C.E., Gaspar, M., Geesink, P., et al., 2019. Biogeochemical regimes in shallow aquifers reflect the metabolic coupling of the elements nitrogen, sulfur, and carbon. Appl. Environ. Microbiol. 85 (5) e02346-18.
Wertz, S., Degrange, V., Prosser, J.I., et al., 2007. Decline of soil microbial diversity does not influence the resistance and resilience of key soil microbial functional groups following a model disturbance. Environ. Microbiol. 9, 2211–2219.
Wu, L., Liu, X., Schadt, C.W., Zhou, J., 2006. Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Appl. Environ. Microbiol. 72, 4931–4941.
Yabusaki, S.B., Wilkins, M.J., Fang, Y., et al., 2017. Water table dynamics and biogeochemical cycling in a shallow, variably-saturated floodplain. Environ. Sci. Technol. 51, 3307–3317.
Zhou, Y., Kellermann, C., Griebler, C., 2012. Spatio-temporal patterns of microbial communities in a hydrologically dynamic pristine aquifer. FEMS Microbiol. Ecol. 81, 230–242.
Zinger, L., Gobet, A., Pommiers, T., 2012. Two Decades of Describing the Unseen Majority of Aquatic Microbial Diversity, vol. 21, pp. 1878–1896.