Microbial Community Cohesion Mediates Community Turnover in Unperturbed Aquifers

Robert E. Danczak,[a] Michael D. Johnston,[b] Chris Kenah,[c] Michael Slattery,[c] Michael J. Wilkins[a,b]

[a]Department of Microbiology, Ohio State University, Columbus, Ohio, USA
[b]School of Earth Sciences, Ohio State University, Columbus, Ohio, USA
[c]Ohio Environmental Protection Agency, Columbus, Ohio, USA

ABSTRACT Microbial ecological processes are frequently studied in the presence of perturbations rather than in undisturbed environments, despite the relatively stable conditions dominating many microbial habitats. To examine processes influencing microbial community structuring in the absence of strong external perturbations, three unperturbed aquifers in Ohio (Greene, Licking, and Athens) were sampled over 2 years and analyzed using geochemical measurements, 16S rRNA gene sequencing, and ecological modeling. Redox conditions ranging from highly reducing to more oxidizing distinguished aquifer geochemistry across the three locations. Distinct microbial communities were present in each aquifer, and overall community structure was related to geochemistry, although community composition was more similar between the Athens and Licking locations. The ecological processes acting upon microbial assemblages within aquifers were varied; geochemical changes affected the Athens location, while time or some unknown factor affected Greene County. Stochastic processes, however, dominated the Licking aquifer, suggesting a decoupling between environmental fluctuations and community development. Although physicochemical differences might be expected to drive variable selection, dispersal limitation (inability to mix) explained differences between Athens and Licking. Finally, community complexity as measured by "cohesion" indicated that less-interconnected communities experienced higher turnover and were more likely to be affected by stochastic processes. Conversely, more-interconnected communities experienced lower turnover and susceptibility to homogenizing selection. Based upon these data, we support the hypothesis that unperturbed environments house dynamic microbial communities due to external and internal forces.

IMPORTANCE Many microbial ecology studies have examined community structuring processes in dynamic or perturbed situations, while stable environments have been investigated to a lesser extent. Researchers have predicted that environmental communities never truly reach a steady state but rather exist in states of constant flux due to internal, rather than external, dynamics. The research presented here utilized a combined null model approach to examine the deterministic and stochastic processes responsible for observed community differences in unperturbed, groundwater ecosystems. Additionally, internal dynamics were investigated by relating a recently published measure of community complexity (cohesion) to ecological structuring processes. The data presented here suggest that communities that are more cohesive, and therefore more complex, are more likely affected by homogenizing selection, while less-complex communities are more susceptible to dispersal. By understanding the relationship between internal dynamics and community structuring processes, insight about microbial population development in natural systems can be obtained.

KEYWORDS aquifers, cohesion, microbial ecology, null modeling

The terrestrial subsurface is estimated to be the largest reservoir of life on Earth (1, 2). Given that biomass density generally decreases with depth, shallow aquifer systems that underlie much of the continental land masses likely host a significant fraction of this microbial life (2). Numerous studies have revealed extensive microbially diverse populations in these ecosystems, with recent work significantly expanding the microbial portion of the tree of life by over 50% (3–6). However, despite our understanding of diversity in shallow subsurface systems, the majority of research on ecological drivers of community structuring and turnover has focused on dynamic environments or community disturbances and subsequent recoveries (7–10). For example, microbial communities in shallow aquifers and adjacent hyporheic mixing zones in Colorado and Washington were influenced by river water intrusion, leading to shifts in redox conditions and carbon availability (11–15). Similar processes have also been observed in engineered ecosystems where artificial perturbations in groundwater ecosystems lead to changes in both community structure and functionality (9, 16, 17). Across these and other dynamic ecosystems, community stability can refer to both resilience when recovering from environmental disturbances (8) and resistance to change in the presence of a perturbation (18).

Despite this focus on environments that undergo perturbations, critical information regarding ecological structuring processes can be obtained from systems where physical and chemical conditions are less dynamic (19). In more stable systems, it has been predicted that no environmental community truly reaches a steady state but rather exists in constant flux (20). This statement has been supported by much previous work, as described in a meta-analysis by Shade et al. (18), and also applies to more dynamic environments (12, 21). In all these instances, internal community dynamics such as speciation and microbe-microbe interactions may play a larger role in assemblage development than any external pressure. This effect was observed in a laboratory chemostat study, where unperturbed communities followed the same trajectory as perturbed communities, suggesting that internal community dynamics played an unexpected role in turnover (22). Additionally, Graham et al. demonstrated that differences in community structure do not necessarily relate to ecosystem processing and that development of some functional groups can be uncoupled from the surrounding conditions (23). Graham and Stegen suggested that communities assembled primarily through stochastic processes would have lower biogeochemical function and, by extension, be less related to surrounding geochemistry (24). Despite the implications of these studies, less work has focused on this potential decoupling between environmental variability and community development. Here, we propose that shallow subsurface regions may represent one such ecosystem where naturally occurring ecological structuring processes can be investigated in the absence of large-scale perturbations.

Over the course of 2 years, we examined microbial community dynamics in groundwater monitoring wells located in three counties across central and southern Ohio (Athens, Greene, and Licking). Each well has a history of being geochemically unperturbed and presents an opportunity to study community stability, here defined as "constancy through time" (25) and measured by low turnover, in an environmental setting. Using a combination of biogeochemical measurements and ecological calculations, we investigated whether the dynamic equilibrium of microbial communities in these wells is unrelated to the surrounding physicochemical environment. Overall, our data suggest that internal dynamics (e.g., microbe-microbe interactions) of aquifer microbial communities may significantly contribute to overall community turnover in these systems.

RESULTS

Monitoring wells experience different degrees of geochemical stability and are significantly distinct. Three groundwater monitoring wells were sampled quarterly over the course of 2 years (Fig. 1A). The aquifers in Athens and Greene Counties were both approximately 10 to 15 ft beneath the surface and shared similar physical
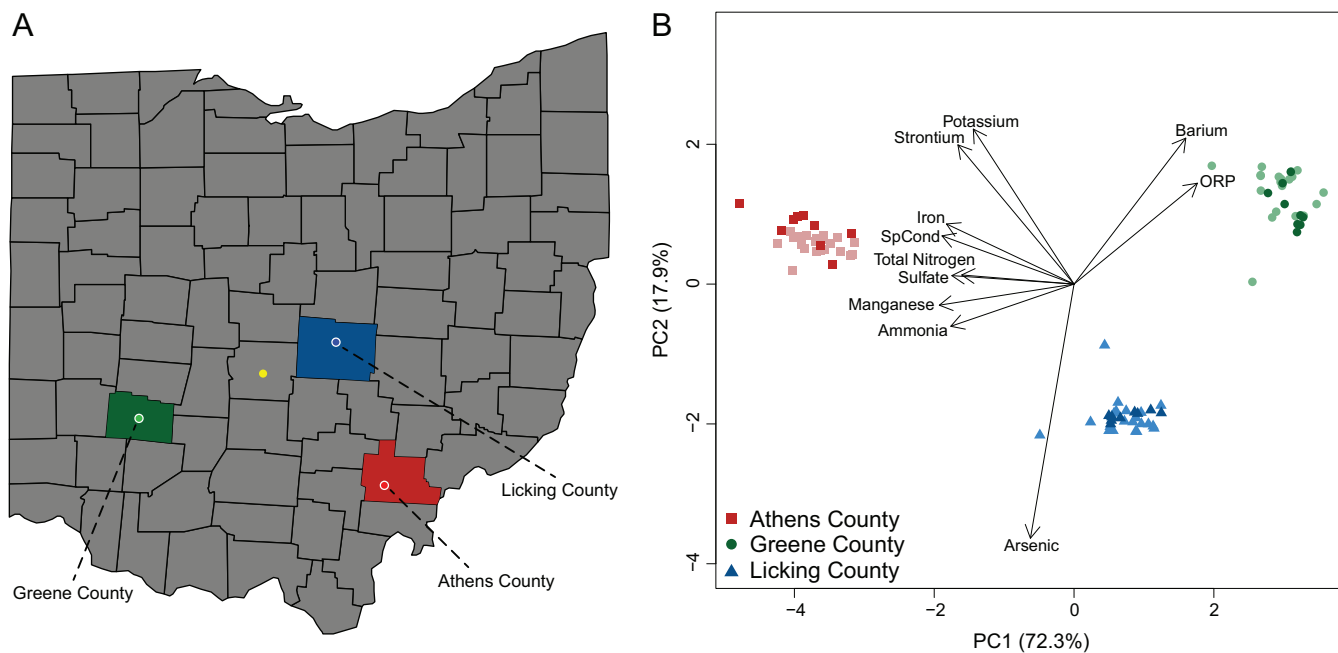
FIG 1 A map of Ohio indicating the three field sites (A) and a PCA of the geochemistry (B). Lighter colors in panel B are geochemical measurements taken before biological samples were collected. Arrows on panel B indicate the loading of various geochemical parameters. The yellow dot represents the location of Columbus, OH, for reference.

characteristics (comprised of mostly brown clay, sand, and gravel). In contrast, the Licking County aquifer was deeper (~26 ft.) and was predominantly comprised of sand and gravel. Likely due to these formational differences, hydraulic conductivity (as inferred from well production) was higher in the Licking County well.

In addition to differences in physical characteristics, each aquifer featured unique geochemical profiles (Fig. 1B). The Greene County aquifer was the most oxidized environment (per high oxidation-reduction potential [ORP] measurements), the Licking aquifer exhibited intermediate redox characteristics and contained the highest dissolved arsenic concentrations (5 to 7 μg/liter), while the Athens County aquifer was the most reduced, containing high concentrations of dissolved iron (6 to 9 mg/liter). Despite the abundance of reduced species at the Athens location, both $\partial^{18}O$ and ∂D groundwater measurements indicated that this location was more hydrologically connected to the surface than either of the other two aquifers over the course of the sampling period (see Fig. S1 in the supplemental material). This suggests significant redox buffering capacity within the Athens aquifer, enabling the accumulation of reduced species despite potential inputs of more oxic surface fluids. Despite this heterogeneity across sampling locations, temporal geochemical patterns were moderately stable over the 2-year sampling period, with few measured parameters deviating more than two standard deviations from their average concentrations (see Fig. S1 and Table S1 in the supplemental material). While not determined statistically, the Athens aquifer appeared to experience minor natural variation and was the least stable system, followed by the Greene aquifer, and then the most stable system, the Licking aquifer.

Microbial communities within the wells are complex, varied, and stable. Microbial communities retained on 0.2-μm-pore filters were analyzed from each aquifer via 16S rRNA gene sequencing after 97% clustering. Mirroring geochemical measurements, Bray-Curtis dissimilarity results indicated that the communities from each aquifer were significantly distinct from each other (Table 1), suggesting linkages between physicochemical differences and microbial populations that are supported by significant permuted Procrustes and Mantel tests (a P value of 0.001 for both). Further mirroring the aquifer geochemistry, both time-resolved alpha diversity and beta dis-

TABLE 1 Pairwise and overall PERMANOVA statistics for the various beta diversity measurements used in this study[a]

	Athens County	Greene County	Licking County
Athens County		Bray: *0.001* wUni: *0.001* uwUni: *0.001*	Bray: *0.001* wUni: *0.001* uwUni: *0.001*
Greene County	Bray: **18.87** wUni: **32.07** uwUni: **17.16**		Bray: *0.001* wUni: *0.001* uwUni: *0.001*
Licking County	Bray: **14.82** wUni: **15.36** uwUni: **8.00**	Bray: **15.69** wUni: **19.95** uwUni: **14.22**	
Overall Bray-Curtis F-stat: 16.35, p-value: 0.001			
Overall wUniFrac F-stat: 21.38, p-value: 0.001			
Overall uwUniFrac F-stat: 14.23, p-value: 0.001			

[a]Bray-Curtis is delineated by "Bray," weighted UniFrac by "wUni," and unweighted UniFrac by "uwUni." Values in boldface are *F* statistics for the specific pairwise PERMANOVA comparisons, while values in italics are the corresponding *P* values.

persion analyses revealed no seasonal trends or strong variability, respectively, in community composition within each well (see Fig. S2 in the supplemental material). Although Shannon's diversity (*H*) measurements initially indicated that the communities in the Greene County aquifer were the most diverse, followed by Athens County and Licking County (Fig. 2A), the incorporation of phylogenetic information into
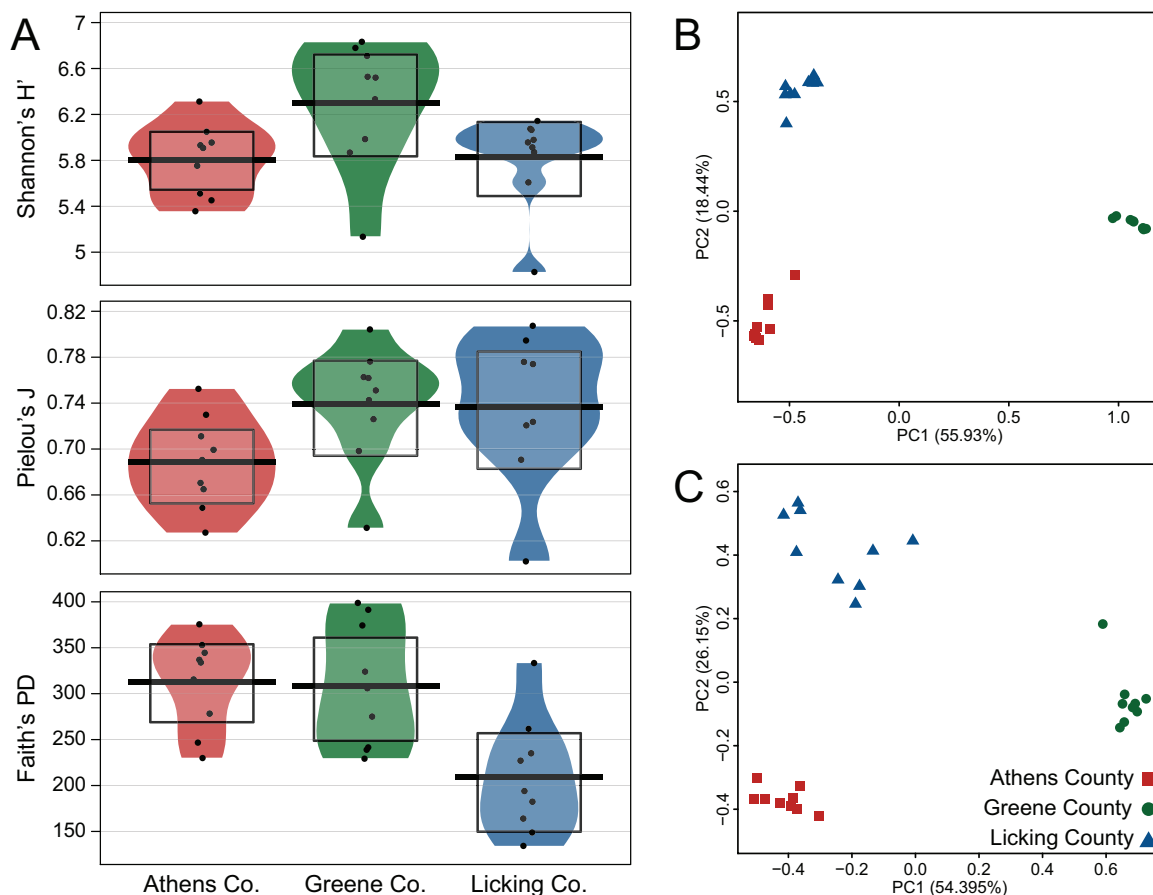

FIG 2 (A) Alpha diversity metrics (Shannon's *H'*, Pielou's *J*, and Faith's PD). (B and C) Unweighted (B) and weighted (C) UniFrac measurements for each sample. Statistical differences are listed in Table 1.

diversity estimates via Faith's phylogenetic diversity (PD) revealed that the Athens and Greene populations were equally diverse (Fig. 2A). As inferred from Pielou's evenness index (J'), the discrepancy between Shannon's and Faith's diversity estimates was due to a more even distribution of operational taxonomic units (OTUs) in the Greene aquifer (Fig. 2A).

Both weighted and unweighted UniFrac analyses were utilized to examine changes in community phylogenetic structure through time and space. Unweighted measurements consider only the phylogenetic distance between community members, while weighted measurements incorporate both phylogenetic distance and abundance information. Therefore, differences between these metrics can reveal detailed relationships between microbial communities. UniFrac results revealed that although communities from each aquifer were significantly different from the others (Table 1; Fig. 2B and C), greater similarity existed between the Athens and Licking locations. However, greater differences existed among and between locations within the weighted data set than in the unweighted (Table 1). This suggests that while many microbial lineages are shared between the aquifers, weighted UniFrac-inferred differences between locations can be significantly attributed to abundances of these phylogenies across the communities (Fig. S2). As an example, among roughly 600 OTUs shared across these different wells, an uncultured OTU within the *Rhodospirillaceae* that represented over 11% of the community across all time points in the Athens aquifer represented less than 0.05% of OTUs in either of the other wells. To investigate the ecological processes driving these compositional differences between the aquifers, null modeling tools were employed to assess the relative influence of selective and stochastic factors constraining microbial population turnover.

Ecological processes controlling microbial community development within aquifers vary between locations. Ecological modeling with the β-nearest taxon index (βNTI) and Raup-Crick (Bray-Curtis) (RC_{BC}) models was utilized to examine differences in selective pressures between locations and determine potential explanations for community variation (Table 2). This approach uses randomized community structures to assess whether measured communities are more similar or dissimilar than would be expected purely by chance. For example, communities that are more different than would be expected by random chance (βNTI > 2) occur due to variable selection, which typically results from fluctuating geochemical conditions (i.e., influx of organic carbon) (14). Communities that are more similar than by random chance alone (βNTI < −2) would occur due to homogenizing selection, which occurs under constant conditions (26). Finally, if communities are as different as expected by random chance ($|\beta$NTI$| < 2$), stochastic processes dictate community structure. These stochastic processes can be further distinguished as dispersal limitation (greater than expected turnover; RC_{BC} > 0.95), homogenizing dispersal (less than expected turnover; RC_{BC} < −0.95), and undominated (ecological processes unclear; $|RC_{BC}| < 0.95$). By combining these metrics, we are able to measure turnover between two communities and identify whether selective (homogenizing/variable) or stochastic (homogenizing dispersal/dispersal limitation) processes dominate (14, 27).

Within-aquifer comparisons for the biomass retained on the filter indicated that the microbial community in the Greene aquifer was dominated by homogenizing selective processes (i.e., communities were more similar than expected by random chance), while the communities in the Athens and Licking aquifers were subject to a mixture of homogenizing selection and homogenizing dispersal (Fig. 3). Homogenizing selection refers to conditions that force microbial communities to be more similar than would be expected by random chance, while homogenizing dispersal describes conditions where low community turnover is explained by high levels of dispersal that mix microbial populations. Rank-based mantel correlations were calculated to determine potential drivers of homogenizing selection across all sampling locations (see Table S2 in the supplemental material). These correlations examined whether microbial community turnover was directly related to temporal shifts in geochemistry. Results indicated that community turnover within the Athens aquifer was significantly positively related to

TABLE 2 Definitions of the various terms used throughout this article[a]

Term	Definition
βNTI	β-Nearest taxon index, which uses phylogenetic information to assemble random populations of microorganisms to measure whether compared communities are more or less different than random chance; differentiates deterministic and stochastic processes
RC_BC	Raup-Crick (Bray-Curtis), which builds null communities probabilistically based upon OTU abundances within a sample to determine whether communities are more or less different than random chance alone; differentiates stochastic processes
Variable selection	The result of communities being more different than would be expected by random chance alone (e.g., geochemical conditions are very different between samples); deterministic process
Homogenizing selection	The result of communities being less different than would be expected by random chance alone (e.g., geochemical conditions are identical/different between samples); deterministic process
Dispersal limitation	Indicates populations between samples are unable to interact, due to separation by either time or space; stochastic process
Homogenizing dispersal	Indicates populations between samples are freely able to interact, potentially due to significant hydrological connectivity; stochastic process
Ecological drift	Occurs when communities are as different as random chance alone; primarily occurs due to random mutations or variations in generation times
Undominated	Indicates no single ecological process is capable of explaining the observed results
Cohesion	A metric that measures the complexity/interconnectedness of a given community as measured by the degree of cocorrelation
Positive cohesion	The cohesion result for those community members that were positively correlated; unknown interpretation
Negative cohesion	The cohesion result for those community members that were negatively correlated; potentially significantly related to community turnover and complexity
Interconnectedness	A term used to reference the degree of negative cohesion within a community

[a]See Zhou and Ning (37) and Herren and McMahon (28) for greater detail.

larger changes in sulfate, sodium, and strontium concentrations, while turnover in the Greene aquifer was only related to longer periods of time between samplings (Table S2). Total nitrogen was the only geochemical factor in the Licking aquifer that was significantly, but weakly, correlated with the observed turnover.

Many differences between aquifer communities are linked to stochastic ecological processes. While differences in community structures between spatially distinct aquifers were inferred from Bray-Curtis and UniFrac measurements, the broad-scale ecological processes responsible for driving these differences (and similarities) were analyzed using the same modeling framework described above. Variable selection dominated turnover between the Greene aquifer and other locations, while turnover between Licking and Athens aquifers was subject to a mix of variable selection and dispersal limitation (Fig. 3). Every comparison between samples from the Greene aquifer and either the Athens or Licking aquifers resulted in higher turnover than would be expected by random chance (βNTI > 2). This suggests that the distinct physico-chemical conditions in the Greene aquifer (more oxidizing) exert a stronger influence on the local microbial community than conditions in either Athens or Licking aquifers (more reduction at both locations). Turnover measurements between the Athens and Licking aquifers typically signified that samples were as different as expected by random chance, suggesting that ecologically stochastic processes were driving community divergence. Specifically, Raup-Crick analyses indicated that dispersal limitation was dominant, meaning that the communities within the two spatially distinct aquifers
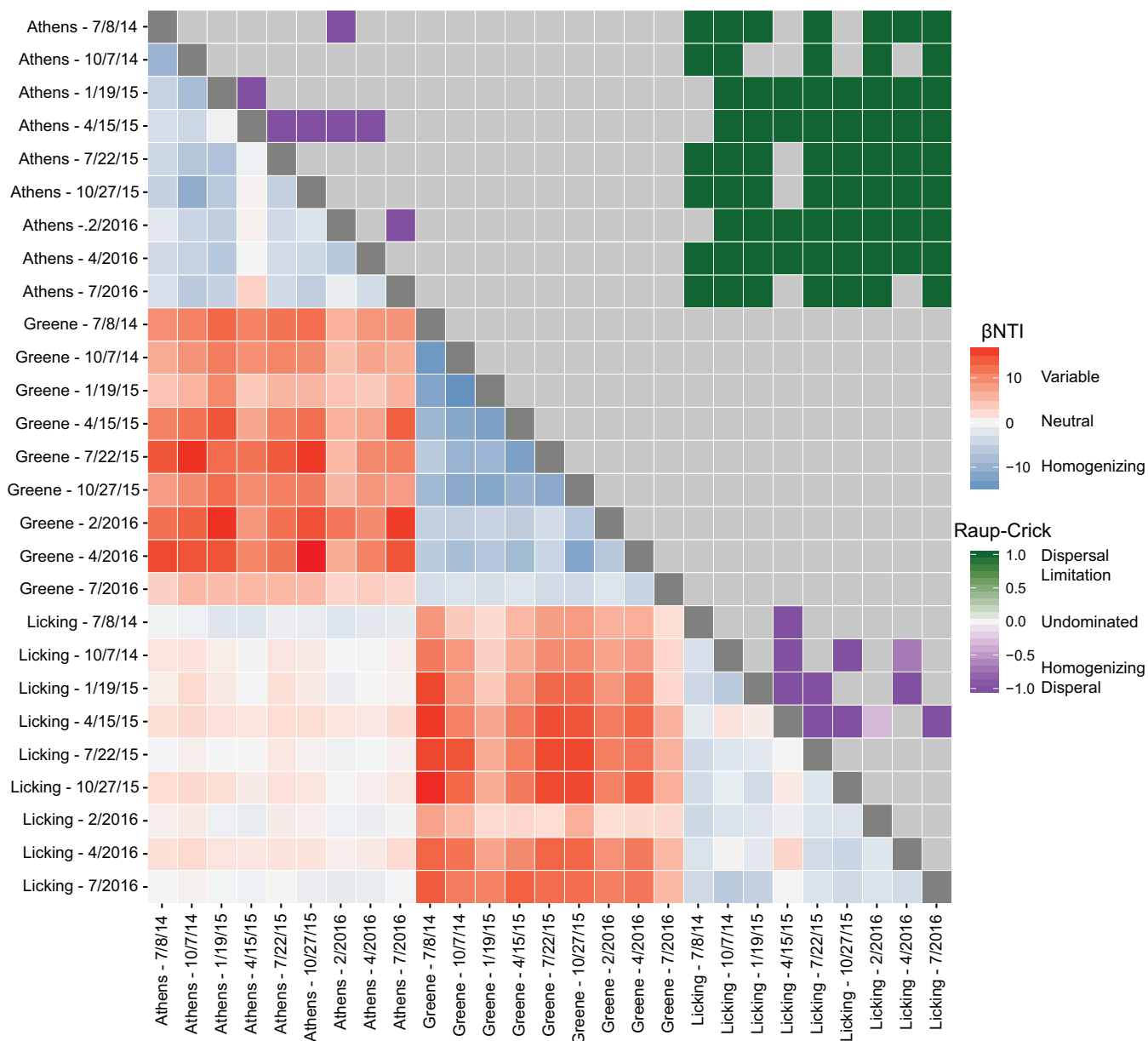
FIG 3 Heat map representing the βNTI values within and between wells (red and blue; lower triangle) and Raup-Crick (Bray-Curtis) values (purple and green; upper triangle). Gray boxes represent comparisons that had significant βNTI values (e.g., $|\beta$NTI$| > 2$) and therefore do not need Raup-Crick values. The accompanying legends provide the corresponding interpretations for each result.

were different as a consequence of distance (i.e., inability to mix) rather than the geochemical characteristics. Therefore, the ecological drivers identified here using βNTI and RC$_{BC}$ measurements offer insights into the mechanisms behind the community clustering patterns identified by UniFrac analyses (see Fig. S3 in the supplemental material).

Many geochemical parameters, including barium, arsenic, manganese, calcium, and sulfate, were significantly related to the between-location βNTI values ($r > 0.6$, $P = 1 \times 10^{-4}$), indicating that larger environmental differences significantly drive differences in community composition between sites. In the absence of ecological modeling, however, community differences between Athens and Licking samples may have been attributed to these physicochemical differences alone. Given that variable selection would lead to more dissimilarity, the closer clustering between Athens and Licking aquifer samples per UniFrac analyses is likely driven by dispersal limitation. Conversely,
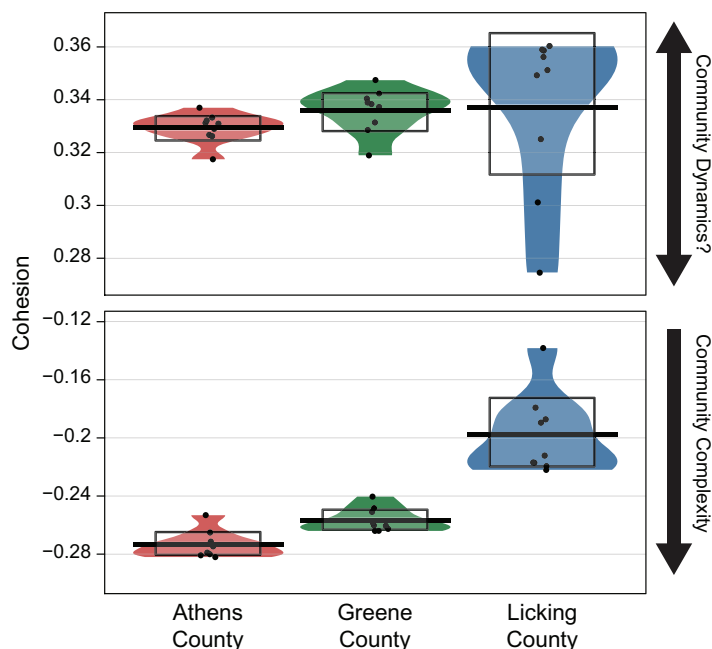
FIG 4 The cohesion metric for each well. The top panel illustrates the positive cohesion measurement, while the bottom panel illustrates the negative measurement. Potential interpretations of results are illustrated by arrows adjacent to each panel. Two-sided Mann-Whitney *U* tests indicated that each location was significantly different from the others (Bonferroni corrected *P* values of <0.01).

the action of variable selection processes on community structure results in Greene biomass samples clustering separately (Fig. S3).

Community complexity is significantly related to community turnover. In addition to determining the relationship between external physicochemical factors and microbial populations, we examined if the degree of complexity (extent of interconnected taxa) within a community affected turnover. Complexity itself is difficult to measure directly, so a new metric proposed by Herren and McMahon called "cohesion" was utilized (28). In brief, cohesion measures complexity by calculating the abundance-weighted pairwise correlations of every taxon in a given community, yielding a series of positive and negative cohesion values that are subsequently treated separately. Given that negative cohesion values have been related to greater levels of community complexity, the Athens and Greene aquifers feature similar cohesion values and therefore contained the most interconnected or cohesive communities, followed by the Licking aquifer (Fig. 4). These results mirror the community alpha diversity measurements (Fig. 2A) and weighted gene coexpression network analysis (WGCNA) results, inferred from the edge-to-node ratio of each network (see Fig. S4 in the supplemental material). Conversely, positive cohesion did not demonstrate any strong trends (Fig. 4), as previously demonstrated by Herren and McMahon (28).

Both Spearman-based correlations and multivariate comparisons between cohesion results and turnover metrics (Bray-Curtis, UniFrac, and βNTI) were performed to determine how complexity related to community turnover. Our results, as measured by Bray-Curtis dissimilarity, revealed significant relationships between negative cohesion and overall community turnover (Mantel $r = 0.498$, $P = 0.0001$), which supports previous research (28). Weaker correlations with positive cohesion were also detected (Mantel $r = 0.285$, $P = 0.001$). Expanding these analyses to incorporate phylogenetic information, Mantel correlations between both UniFrac measurements and cohesion supported the Bray-Curtis observations, with stronger relationships in the weighted UniFrac fraction (negative Mantel $r = 0.426$, $P = 0.0001$; positive Mantel $r = 0.310$, $P = 0.0003$) than the unweighted UniFrac fraction (negative Mantel $r = 0.306$, $P = 0.0009$; positive Mantel $r = 0.149$, $P = 0.019$). These Bray-Curtis and UniFrac correlations are

only partially supported by multivariate analyses where only negative cohesion was significantly related to turnover (see Fig. S5 in the supplemental material). Taken together, not only did community complexity, as measured particularly by negative cohesion, influence community turnover taxonomically (i.e., Bray-Curtis), but it additionally had a significant impact on the phylogenetic structure, suggesting a deterministic influence.

In order to directly determine if cohesion influences community assembly processes deterministically, cohesion was correlated to βNTI. Both negative cohesion and positive cohesion were related to the βNTI, although a stronger correlation was observed with the negative cohesion (negative Mantel $r = 0.200$, $P = 0.006$; positive Mantel $r = 0.142$, $P = 0.022$). Furthermore, multivariate analyses demonstrated a similar pattern to the Bray-Curtis and UniFrac results, showing significant relationship to Licking community structure (Fig. S5). Whereas previous analyses linked external parameters (e.g., geochemical variables) to the observed community turnover, this technique measures the interconnectedness of taxa in each sample, independent from external effects. Given that negative cohesion is weakly but significantly related to Licking community structure (which had the least-negative cohesion values), these results suggest that more-interconnected and therefore more-complex communities (with more negative cohesion values) may be less susceptible to high community turnover, regardless of the presence or absence of external perturbation by environmental factors. This is likely due to critical community processes (e.g., metabolite exchange and competition) being controlled by multiple interconnected taxa in a more complex community. As such, community structures are more difficult to disrupt, resulting in lower rates of turnover (29, 30).

DISCUSSION

Microbial community assembly processes were examined over a 2-year period in three separate Ohio aquifers. Despite being unperturbed (Fig. 1B; Fig. S1), significant correlations between community turnover (as measured by βNTI) and geochemical parameters within location existed, indicating that small variations in the surrounding chemical environment can exert a measurable effect on microbial communities, as previously hypothesized with this technique (13). Sulfate concentrations in the Athens aquifer, for example, never varied more than 25% from the median (or average) concentration over the 2 years of sampling but were significantly correlated to community turnover ($r = 0.64$, $P = 0.002$). Given that sulfate is a terminal electron acceptor utilized prominently in reducing environments, these concentration fluctuations may select for putative sulfate reducers (31, 32). In the Greene aquifer, the observation that longer periods of time between sampling efforts were linked to turnover suggests that a changing unmeasured variable (e.g., a geochemical parameter) may have been responsible for this relationship (Table S1). Alternatively, the process of ecological drift (community change via stochastic differences in rates of cell division and death) could account for this trend, but this signal (if it exists) is masked by the strong homogenizing selection. A similar observation has previously been made in a dynamic hyporheic zone environment (15). Lastly, samples from the Licking aquifer likely lacked strong correlations between microbial community turnover and geochemical parameters as these communities were subject to the highest degree of stochastic processes, specifically homogenizing dispersal, overwhelming deterministic processes (Fig. 3). This may be linked to the inferred higher hydraulic conductivity (and associated lithofacies composition) in this aquifer, which could enable greater mixing between microbial populations and thus drive homogenizing dispersal via potentially higher dispersal rates (33).

While stochastic processes (i.e., dispersal limitation) governed many βNTI comparisons between locations, Mantel correlations suggested that numerous physicochemical parameters were significantly related to community turnover ($|\beta$NTI$| > 2$). Of these significant relationships, correlations with barium, arsenic, manganese, calcium, and sulfate concentrations, along with oxidation-reduction potential (ORP) variations, each yielded r values above 0.6 and P values of 1×10^{-4}, indicating that they may play
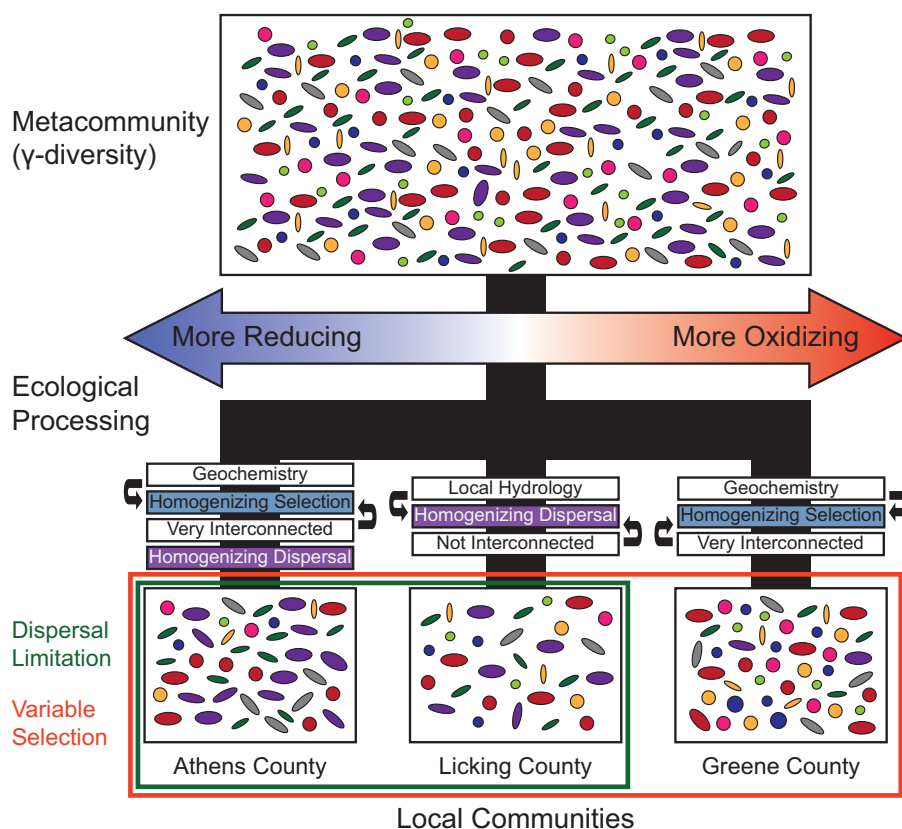
FIG 5 Summary figure illustrating our hypotheses. The metacommunity represents the total group of possible organisms that could assemble into our local communities. The black lines illustrate the total "ecological processing" that generates the local communities, with the blue, white, and purple boxes indicating specific ecological filters within a given location. Arrows adjacent to the colored bars represent interactions that enhance specific processes. For example, stable geochemistry enhances homogenizing selection. The red-blue gradient arrow represents a range of redox conditions from reducing to oxidizing. The red and green boxes indicate the measured ecological processes that must occur between the wells to obtain the observed results (dispersal limitation between the Athens and Licking locations and variable selection between the Greene aquifer and the other two locations).

outsized roles in the observed community differences (Table S2). Differences in redox conditions (as inferred from ORP measurements) govern the biogeochemical reactions that can occur in a given location and likely play a significant role in community structuring (34). In reducing groundwater environments, sulfate and redox active metals, like arsenic and manganese, are often utilized by many microorganisms as terminal electron acceptors (32, 35, 36). The presence or absence of these chemical species could drive variable selection through the development of niches for certain taxa, allowing them to outcompete less-specialized community members (37).

We propose that the direct coupling between changes in geochemistry and microbial community structure is only supported in systems in which external parameters fluctuate beyond a given magnitude (Fig. 5). Similar behavior has been observed in systems that could be considered "multistable" (38), but given that this system does not enter into an alternative steady state, this conclusion is unlikely. While whole microbial communities in a given aquifer were related to geochemical profiles (e.g., PROTEST, Mantel), links to community turnover as measured by βNTI were often absent. Although turnover could be linked to environmental drivers in the Athens aquifer and time at the Greene location, only one weak relationship was apparent for the microbial community in the Licking aquifer, which experienced intermediate redox conditions. Additional cohesion measurements showing low levels of community complexity support our inference that populations in the Licking aquifer were subject to the greatest degree of internal ("endogenous") dynamics of the three ecosystems (20). In

contrast, the temporal environmental changes within the Athens and Greene aquifers were strong enough to deterministically constrain community turnover and limit any changes due to stochastic processes. Given the gradient of redox conditions found across these three sampling locations, we hypothesize that communities in more strongly oxic (Greene) or reduced (Athens) systems are driven toward a particular specialized but interconnected population with minor susceptibility to geochemical fluctuations and overall stability (24, 39). However, intermediate redox conditions (e.g., suboxic) such as those measured in the Licking aquifer may lead to a more generalized population (24), which when combined with homogenizing dispersal results in a less interconnected ("cohesive") microbial community in which internal dynamics play a greater role than environmental pressures on community structure. Given that dispersal-based assembly may lead to lesser biogeochemical function within microbial communities (23, 24), a positive-feedback loop could potentially develop. In such a system, stochastically assembled communities with few interconnections have diminished influence over surrounding redox conditions, subsequently leading to less interconnected communities.

Conclusions. Using a combination of biogeochemical measurements and ecological calculations, we analyzed the microbial community structure in three Ohio aquifers over the course of 2 years. Despite the absence of major perturbations in these ecosystems, complex microbial dynamics across these aquifer environments highlighted the importance of stochastic ecological processes, revealing community turnover independent of physicochemical changes in the subsurface (20). In addition, while some microbial lineages represented by ~600 OTUs were shared across all three locations (Fig. S2), variations in selective pressures and stochastic processes altered the distribution of these microorganisms.

A multifaceted approach incorporating both biogeochemical relationships and ecological modeling has enabled a better understanding of the drivers behind complex microbial community development. The data presented here help support "stability" measurements conducted elsewhere (22), which revealed that no microbial community exists in an absolutely steady state. Rather, communities appear to exist in flux, regardless of the small changes in the surrounding environmental conditions. While stochastic processes may exert significant control over population dynamics, even small shifts in physicochemical parameters can drive additional community changes. Additional future insights may be provided by microbial community information beyond the marker gene. Metagenomics and associated "omics" tools (transcriptomics and shotgun proteomics) provide additional information on functional potential, inferred relationships between taxa, and activity. The incorporation of these data into this ecological framework would potentially enable a clearer view of the effects of both external and internal dynamics on community turnover. Overall, these results suggest that interplay of internal and external dynamics contributes to microbial community assembly, with more-complex community structures being more likely to be subjected to homogenizing selection and less-complex communities more susceptible to dispersal.

MATERIALS AND METHODS

Sample collection. Groundwater samples were collected on a quarterly basis over a 2-year period from July 2014 to July 2016 from the Ohio Department of Natural Resources (ODNR) Observation Well Network used to monitor groundwater level fluctuations in three different counties (Fig. 1A). These three wells located in Athens County (39.4417°E, −82.2178°N), Greene County (39.5792°E, −84.0167°N), and Licking County (40.1467°E, −82.4197°N) are situated within separate buried valley aquifers—valleys that have been back-filled with glacial sands and gravels and some till.

Sample collection followed previously established protocols (40). At each sampling, wells were purged of more than 250 liters of water to ensure that aquifer-derived water was being collected (dedicated pumps were placed at the screened interval for the ODNR wells). Approximately 38 liters of postpurge groundwater was pumped sequentially through a 142-mm-diameter 0.2- then 0.1-μm-pore Supor PES membrane filter (Pall Corporation, NY). Filters were then immediately flash frozen and kept on dry ice until they were returned to the Ohio State University, where they were stored at −80°C until DNA extraction. During sampling, oxidation-reduction potential (ORP), temperature, and pH were measured using a handheld Myron Ultrameter II (Myron L Company, CA). Groundwater samples (1 per site per date) were analyzed according to Ohio EPA standard operating protocol (SOP) and testing methods, which can

be found in the supplemental material (Table S1). Cations were measured both on an ICP-OES Optima 73000DV and ICP-MS ELAN 9000 (PerkinElmer, Inc., MA), while anions were measured using a Konelab Aquakem 20 (Thermo, Fisher Scientific, MA). Water isotopes were measured using a Picarro water isotope analyzer model L1102-i (Picarro, Inc., CA); two-sided Mann-Whitney U tests with Bonferroni corrections were used to determine pairwise significant differences between locations. All geochemical measurements and Ohio EPA SOPs are in Table S1.

DNA extraction, sequencing, and processing. DNA was extracted in duplicate from roughly a quarter of each 0.2-μm Supor PES membrane filter by using the Powersoil DNA isolation kit (Mo Bio Laboratories, Inc., Carlsbad, CA). Final DNA concentrations were determined by using a Qubit fluorimeter (Invitrogen, Carlsbad, CA).

To generate 16S rRNA gene data, the V4 region of 16S rRNA genes was amplified and sequenced using the universal bacterial/archaeal primer set 515F/806R on an Illumina MiSeq instrument at Argonne National Laboratory according to the Earth Microbiome Project standard protocol (41). The resulting reads were chimera checked and *de novo* clustered at 97%, with singletons removed through the QIIME pipeline (V1.7.0), with tree generation and assigned taxonomy using the SILVA v132 database (42). The OTU table is provided as Table S3 in the supplemental material.

Data analyses. All statistical analyses were performed using R v3.3.2 (43). Duplicate samples were averaged together before any analysis was performed. Alpha diversity within these microbial community samples (16S rRNA gene data) was determined according to Shannon diversity calculations (H') (44) (*diversity*, vegan package v2.4.4) (45), Pielou's evenness (J) (H'/log species number), and Faith's phylogenetic diversity (PD) (*pd*, picante package v1.6.2) (46, 47). Pairwise Bray-Curtis metrics were calculated to obtain a taxonomic measure of beta diversity (*vegdist*, vegan package v2.4.4) (45, 48). Both weighted and unweighted UniFrac values were calculated using the OTU data sets (GUniFrac package v1.0) to obtain a phylogenetic measure of beta diversity (49, 50). For both UniFrac metrics, a principal-component analysis (PCA) was used to visualize the community relationships (*prcomp*). Significant differences between aquifers were determined using global and pairwise permutational analysis of variance (PERMANOVA) (*Adonis*, vegan package v2.4.4) comparisons on UniFrac measurements (45). Hierarchical clustering (unweighted pair group method using average linkages [UPGMA]) was used to better visualize these significant relationships (*hclust*). Significant relationships between UniFrac community structure and geochemical profiles were analyzed using both a permuted Procrustes (PROTEST) and Mantel test (*protest* and *mantel*, vegan package v2.4.4) (45).

Ecological modeling. In order to investigate potential ecological drivers within and between these three aquifers, ecological modeling was performed following the protocol outlined by Stegen et al. (33). Per this protocol, a phylogenetic signal was first found to establish a link between phylogeny and ecology using oxidation-reduction potential as the environmental variable for calculating niche differences (see Fig. S6 in the supplemental material). Next, β-mean nearest taxon distance (βMNTD) was calculated for each possible pairwise comparison between samples in order to capture underlying phylogenetic contributions to community composition. Using 999 community randomizations to create null models, the β-nearest taxon index (βNTI) was calculated to determine the deviation of the observed βMNTD from the null βMNTD. Resulting βNTI values were used to examine the phylogenetic turnover in each aquifer, providing insight into whether deterministic (i.e., selection) or stochastic (i.e., random) processes shaped community composition. If a $|\beta$NTI$|$ value exceeds 2, a deterministic process is responsible for differences between microbial communities in two samples; if a $|\beta$NTI$|$ value is less than 2, a stochastic process explains observed differences in microbial community composition between two samples (12). Stochastic processes might include ecological drift (random changes in organismal abundance), enhanced passive cell mobilization, or limitations in cell dispersal within a particular habitat (51). Conversely, deterministic forces include niche-based factors, such as environmental fluctuations that might select for specific microorganisms (e.g., localized abundances of certain nutrients) or specific species interactions, which might drive microbial community structure toward a particular composition (37). Deterministic processes can be further categorized as either variable selection or homogenizing selection. Variable selection occurs when two communities are more dissimilar than would be expected by random chance and occurs if βNTI is greater than 2. Typically, these processes occur when the environmental conditions between the compared communities are very different (e.g., large differences in organic carbon type), resulting in different compositions (14). If βNTI is less than -2, communities are more similar than could occur by random chance and homogenizing selection is considered the dominant process. This type of selection typically occurs in situations in which relatively constant environmental conditions push community structure toward a common composition, such as in the case of microbial community succession in geochemically stable soil environments (26). Correlations between βNTI measurements and differences in geochemical data were then calculated, following a similar protocol to that of Stegen et al. (34). First, differential tables for each environmental variable were calculated to obtain pairwise comparisons for each sample. Then, correlations were performed using the Spearman rank correlation specification in the Mantel test (*mantel*, ecodist package v2.0.1) (52). Due to an aberration that prevents more than 20,000 total OTUs from being read through the *cophenetic* command in the picante package, the abundance data were rarefied to 8,000 sequences for βNTI analyses only (the sample with the lowest sequence count).

In addition to βNTI, Raup-Crick (Bray-Curtis) (RC_{BC}) measurements were used to examine stochastic processes according to Stegen et al. (27). Using 9,999 iterations per pairwise comparison, null communities were probabilistically generated based upon observed OTU abundances across all three aquifers. The null distributions of Bray-Curtis values were then compared to observed Bray-Curtis comparisons to determine deviations and, in turn, significances. These deviations are then normalized to vary between

+1 and −1, resulting in the RC_{BC} metric. If the $|RC_{BC}|$ is > 0.95, the turnover between the compared communities was the result of either dispersal limitation or homogenizing dispersal. Dispersal limitation (RC_{BC} > 0.95) is measured when turnover is greater than expected by drift alone and occurs in situations where communities are unable to mix (i.e., dispersal between sampling locations is impossible). Homogenizing dispersal (RC_{BC} < −0.95) occurs in situations where organisms are freely able to move throughout a given environment and is the result of turnover that is less than expected by drift alone. Finally, if $|RC_{BC}|$ is < 0.95, the comparison is assumed to be the result of undominated processes and suggests that no single ecological process was able to explained the observed community variation (33). Given that RC_{BC} is particularly useful in measuring differences in stochastic processes, these results are only presented in situations where βNTI was insignificant ($|\beta$NTI$|$ < 2). Table S4 in the supplemental material contains the βNTI and RC_{BC} values.

Cohesion. Cohesion was calculated within aquifers according to the protocol outlined by Herren and McMahon (28). Between-aquifer cohesion measurements were not performed. Cohesion approximately measures complexity by calculating the total abundance-weighted pairwise correlations of every taxon in a given community. Briefly, this technique uses Pearson correlation measurements to determine the interconnectedness of a community. The strength of each pairwise correlation was in turn verified using null modeling as follows. The null model was iterated through each OTU (deemed the "focal taxon"), randomizing relative abundances within community using observed OTU distributions, aside from the focal taxon. Pearson correlations were then performed between this focal taxon and all other randomized members. This process was repeated 200 times within each iteration, and the resulting distribution of correlation values is considered the "expected" relationship if random interactions occurred. These expected correlations were then subtracted from the observed correlations to obtain a "connectedness" measurement, with positive and negative "connectedness" values being separated due to the nature of relative abundance data (when one organism becomes more abundant, another must become less abundant). These connectedness measurements were then weighted by taxon abundance and summed to yield both a positive and negative cohesion metric. Two-sided Mann-Whitney U tests with Bonferroni corrections were used to determine pairwise significant differences in cohesion between locations. The cohesion values were transformed into a differential table as described above. Spearman rank correlations (using the Mantel test) between the aquifer cohesion differential table and Bray-Curtis, UniFrac, and βNTI values were performed to measure the relationship between community interconnectedness and turnover (Table S2). Additionally, the cohesion values were related to Bray-Curtis, UniFrac, and βNTI using a multivariate environmental fit (*envfit*, vegan package v.2.4.4) (45). Weighted gene coexpression analysis (WGCNA) (53) was additionally performed and qualitatively related to cohesion metrics (WGCNA package v1.51) (54, 55).

All R scripts and extra metadata used in this study are available on GitHub at https://github.com/danczakre/AquiferEcology.

All terms used throughout this study can be found in Table 2.

Accession number(s). The data generated during 16S rRNA gene sequencing can be obtained from the NCBI SRA database under accession no. SRX2896383.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mSystems.00066-18.

FIG S1, PDF file, 0.3 MB.
FIG S2, PDF file, 0.5 MB.
FIG S3, PDF file, 0.1 MB.
FIG S4, PDF file, 1.9 MB.
FIG S5, PDF file, 0.2 MB.
FIG S6, PDF file, 0.1 MB.
TABLE S1, XLSX file, 0.1 MB.
TABLE S2, XLSX file, 0.1 MB.
TABLE S3, TXT file, 16 MB.
TABLE S4, XLSX file, 0.1 MB.

ACKNOWLEDGMENT

REFERENCES

1. Whitman WB, Coleman DC, Wiebe WJ. 1998. Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A 95:6578–6583. https://doi.org/10.1073/pnas.95.12.6578.
2. McMahon S, Parnell J. 2014. Weighing the deep continental biosphere. FEMS Microbiol Ecol 87:113–120. https://doi.org/10.1111/1574-6941.12196.
3. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF. 2015. Unusual biology across a

group comprising more than 15% of domain Bacteria. Nature 523: 208–211. https://doi.org/10.1038/nature14486.

4. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K, Suzuki Y, Dudek N, Relman DA, Finstad KM, Amundson R, Thomas BC, Banfield JF. 2016. A new view of the tree of life. Nat Microbiol 1:16048. https://doi.org/10.1038/nmicrobiol.2016.48.

5. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, Thomas BC, Singh A, Wilkins MJ, Karaoz U, Brodie EL, Williams KH, Hubbard SS, Banfield JF. 2016. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun 7:13219. https://doi.org/10.1038/ncomms13219.

6. Parks DH, Rinke C, Chuvochina M, Chaumeil P-A, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW. 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat Microbiol 2:1533–1542. https://doi.org/10.1038/s41564-017-0012-7.

7. Shade A, Read JS, Welkie DG, Kratz TK, Wu CH, McMahon KD. 2011. Resistance, resilience and recovery: aquatic bacterial dynamics after water column disturbance. Environ Microbiol 13:2752–2767. https://doi.org/10.1111/j.1462-2920.2011.02546.x.

8. Shade A, Read JS, Youngblut ND, Fierer N, Knight R, Kratz TK, Lottig NR, Roden EE, Stanley EH, Stombaugh J, Whitaker RJ, Wu CH, McMahon KD. 2012. Lake microbial communities are resilient after a whole-ecosystem disturbance. ISME J 6:2153–2167. https://doi.org/10.1038/ismej.2012.56.

9. Hug LA, Thomas BC, Brown CT, Frischkorn KR, Williams KH, Tringe SG, Banfield JF. 2015. Aquifer environment selects for microbial species cohorts in sediment and groundwater. ISME J 9:1846–1856. https://doi.org/10.1038/ismej.2015.2.

10. Nemergut DR, Shade A (ed). 2015. The causes and consequences of microbial community structure. Frontiers SA Media, Lausanne, Switzerland.

11. Lin X, McKinley J, Resch CT, Kaluzny R, Lauber CL, Fredrickson J, Knight R, Konopka A. 2012. Spatial and temporal dynamics of the microbial community in the Hanford unconfined aquifer. ISME J 6:1665–1676. https://doi.org/10.1038/ismej.2012.26.

12. Stegen JC, Lin X, Konopka AE, Fredrickson JK. 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. ISME J 6:1653–1664. https://doi.org/10.1038/ismej.2012.22.

13. Danczak RE, Sawyer AH, Williams KH, Stegen JC, Hobson C, Wilkins MJ. 2016. Seasonal hyporheic dynamics control coupled microbiology and geochemistry in Colorado River sediments. J Geophys Res Biogeosci 121:2976–2987. https://doi.org/10.1002/2016JG003527.

14. Stegen JC, Fredrickson JK, Wilkins MJ, Konopka AE, Nelson WC, Arntzen EV, Chrisler WB, Chu RK, Danczak RE, Fansler SJ, Kennedy DW, Resch CT, Tfaily M. 2016. Groundwater-surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. Nat Commun 7:11237. https://doi.org/10.1038/ncomms11237.

15. Graham EB, Crump AR, Resch CT, Fansler S, Arntzen E, Kennedy DW, Fredrickson JK, Stegen JC. 2017. Deterministic influences exceed dispersal effects on hydrologically connected microbiomes. Environ Microbiol 19:1552–1567. https://doi.org/10.1111/1462-2920.13720.

16. Zhou J, Deng Y, Zhang P, Xue K, Liang Y, Van Nostrand JD, Yang Y, He Z, Wu L, Stahl DA, Hazen TC, Tiedje JM, Arkin AP. 2014. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proc Natl Acad Sci U S A 111:E836–E845. https://doi.org/10.1073/pnas.1324044111.

17. Handley KM, Wrighton KC, Miller CS, Wilkins MJ, Kantor RS, Thomas BC, Williams KH, Gilbert JA, Long PE, Banfield JF. 2015. Disturbed subsurface microbial communities follow equivalent trajectories despite different structural starting points. Environ Microbiol 17:622–636. https://doi.org/10.1111/1462-2920.12467.

18. Shade A, Caporaso JG, Handelsman J, Knight R, Fierer N. 2013. A meta-analysis of changes in bacterial and archaeal communities with time. ISME J 7:1493–1506. https://doi.org/10.1038/ismej.2013.54.

19. Song HS, Renslow RS, Fredrickson JK, Lindemann SR. 2015. Integrating ecological and engineering concepts of resilience in microbial communities. Front Microbiol 6:1298. https://doi.org/10.3389/fmicb.2015.01298.

20. Konopka A, Lindemann S, Fredrickson J. 2015. Dynamics in microbial communities: unraveling mechanisms to identify principles. ISME J 9:1488–1495. https://doi.org/10.1038/ismej.2014.251.

21. Evans S, Martiny JBH, Allison SD. 2017. Effects of dispersal and selection on stochastic assembly in microbial communities. ISME J 11:176–185. https://doi.org/10.1038/ismej.2016.96.

22. Fernandez-Gonzalez N, Huber JA, Vallino JJ. 2016. Microbial communities are well adapted to disturbances in energy input. mSystems 1:e00117-16. https://doi.org/10.1128/mSystems.00117-16.

23. Graham EB, Knelman JE, Schindlbacher A, Siciliano S, Breulmann M, Yannarell A, Beman JM, Abell G, Philippot L, Prosser J, Foulquier A, Yuste JC, Glanville HC, Jones DL, Angel R, Salminen J, Newton RJ, Bürgmann H, Ingram LJ, Hamer U, Siljanen HMP, Peltoniemi K, Potthast K, Bañeras L, Hartmann M, Banerjee S, Yu RQ, Nogaro G, Richter A, Koranda M, Castle SC, Goberna M, Song B, Chatterjee A, Nunes OC, Lopes AR, Cao Y, Kaisermann A, Hallin S, Strickland MS, Garcia-Pausas J, Barba J, Kang H, Isobe K, Papaspyrou S, Pastorelli R, Lagomarsino A, Lindström ES, Basiliko N, Nemergut DR. 2016. Microbes as engines of ecosystem function: when does community structure enhance predictions of ecosystem processes? Front Microbiol 7:214. https://doi.org/10.3389/fmicb.2016.00214.

24. Graham EB, Stegen JC. 2017. Dispersal-based microbial community assembly decreases biogeochemical function. Processes 5:65. https://doi.org/10.3390/pr5040065.

25. Worm B, Duffy JE. 2003. Biodiversity, productivity and stability in real food webs. Trends Ecol Evol 18:628–632. https://doi.org/10.1016/j.tree.2003.09.003.

26. Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF. 2015. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. Proc Natl Acad Sci U S A 112:E1326–E1332. https://doi.org/10.1073/pnas.1414261112.

27. Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, Rockhold ML, Konopka A. 2013. Quantifying community assembly processes and identifying features that impose them. ISME J 7:2069–2079. https://doi.org/10.1038/ismej.2013.93.

28. Herren CM, McMahon KD. 2017. Cohesion: a method for quantifying the connectivity of microbial communities. ISME J 11:2426–2438. https://doi.org/10.1038/ismej.2017.91.

29. Coyte KZ, Schluter J, Foster KR. 2015. The ecology of the microbiome: networks, competition, and stability. Science 350:663–666. https://doi.org/10.1126/science.aad2602.

30. Herren CM, Baym M. 2018. Stronger connectivity of the resident gut microbiome lends resistance to invading bacteria. bioRxiv https://doi.org/10.1101/261750.

31. Anderson RT, Vrionis HA, Ortiz-Bernad I, Resch CT, Long PE, Dayvault R, Karp K, Marutzky S, Metzler DR, Peacock A, White DC, Lowe M, Lovley DR. 2003. Stimulating the in situ activity of Geobacter species to remove uranium from the groundwater of a uranium-contaminated aquifer. Appl Environ Microbiol 69:5884–5891. https://doi.org/10.1128/AEM.69.10.5884-5891.2003.

32. Williams KH, Long PE, Davis JA, Wilkins MJ, N'Guessan AL, Steefel CI, Yang L, Newcomer D, Spane FA, Kerkhof LJ, McGuinness L, Dayvault R, Lovley DR. 2011. Acetate availability and its influence on sustainable bioremediation of uranium-contaminated groundwater. Geomicrobiol J 28:519–539. https://doi.org/10.1080/01490451.2010.520074.

33. Stegen JC, Lin X, Fredrickson JK, Konopka AE. 2015. Estimating and mapping ecological processes influencing microbial community assembly. Front Microbiol 6:370. https://doi.org/10.3389/fmicb.2015.00370.

34. Stegen JC, Konopka A, McKinley JP, Murray C, Lin X, Miller MD, Kennedy DW, Miller EA, Resch CT, Fredrickson JK. 2016. Coupling among microbial communities, biogeochemistry, and mineralogy across biogeochemical facies. Sci Rep 6:30553. https://doi.org/10.1038/srep30553.

35. Aklujkar M, Coppi MV, Leang C, Kim BC, Chavan MA, Perpetua LA, Giloteaux L, Liu A, Holmes DE. 2013. Proteins involved in electron transfer to Fe(III) and Mn(IV) oxides by Geobacter sulfurreducens and Geobacter uraniireducens. Microbiology 159:515–535. https://doi.org/10.1099/mic.0.064089-0.

36. Giloteaux L, Holmes DE, Williams KH, Wrighton KC, Wilkins MJ, Montgomery AP, Smith JA, Orellana R, Thompson CA, Roper TJ, Long PE, Lovley DR. 2013. Characterization and transcription of arsenic respiration and resistance genes during in situ uranium bioremediation. ISME J 7:370–383. https://doi.org/10.1038/ismej.2012.109.

37. Zhou J, Ning D. 2017. Stochastic community assembly: does it matter in microbial ecology? Microbiol Mol Biol Rev 81:e00002-17. https://doi.org/10.1128/MMBR.00002-17.

38. Gonze D, Lahti L, Raes J, Faust K. 2017. Multi-stability and the origin of microbial community types. ISME J 11:2159–2166. https://doi.org/10.1038/ismej.2017.60.

39. Shade A, Peter H, Allison SD, Baho DL, Berga M, Bürgmann H, Huber DH, Langenheder S, Lennon JT, Martiny JB, Matulich KL, Schmidt TM, Handelsman J. 2012. Fundamentals of microbial community resistance and

resilience. Front Microbiol 3:417. https://doi.org/10.3389/fmicb.2012.00417.

40. Danczak RE, Johnston MD, Kenah C, Slattery M, Wrighton KC, Wilkins MJ. 2017. Members of the Candidate Phyla Radiation are functionally differentiated by carbon- and nitrogen-cycling capabilities. Microbiome 5:112. https://doi.org/10.1186/s40168-017-0331-1.

41. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6:1621–1624. https://doi.org/10.1038/ismej.2012.8.

42. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596. https://doi.org/10.1093/nar/gks1219.

43. R Development Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org.

44. Shannon CE. 1948. A mathematical theory of communication. Bell Syst Tech J 27:379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

45. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2017. vegan: community ecology package. R package version 2.4.4. R Foundation for Statistical Computing, Vienna, Austria. https://cran.r-project.org/web/packages/vegan/index.html.

46. Kembel SW, Ackerly DD, Blomberg SP, Cornwell WK, Cowan PD, Helmus MR, Morlon H, Webb CO. 2014. Package "picante." R Foundation for Statistical Computing, Vienna, Austria. https://cran.r-project.org/web/packages/picante/picante.pdf.

47. Faith DP, Baker AM. 2007. Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. Evol Bioinform Online 2:121–128. https://doi.org/10.1177/117693430600200007.

48. Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr 27:325–349. https://doi.org/10.2307/1942268.

49. Lozupone CA, Hamady M, Kelley ST, Knight R. 2007. Quantitative and qualitative diversity measures lead to different insights into factors that structure microbial communities. Appl Environ Microbiol 73:1576–1585. https://doi.org/10.1128/AEM.01996-06.

50. Chen J. 2012. GUniFrac: generalized UniFrac distances. R package version 1.0. R Foundation for Statistical Computing, Vienna, Austria. https://cran.r-project.org/web/packages/GUniFrac/GUniFrac.pdf.

51. Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S. 2013. Patterns and processes of microbial community assembly. Microbiol Mol Biol Rev 77:342–356. https://doi.org/10.1128/MMBR.00051-12.

52. Goslee S, Urban D. 2007. The ecodist package for dissimilarity-based analysis of ecological data. J Stat Softw 22:1–19. https://doi.org/10.18637/jss.v022.i07.

53. Zhang B, Horvath S. 2005. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 4:Article17. https://doi.org/10.2202/1544-6115.1128.

54. Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559. https://doi.org/10.1186/1471-2105-9-559.

55. Langfelder P, Horvath S. 2012. Fast R functions for robust correlations and hierarchical clustering. J Stat Softw 46:i11.