



# Microbial Communities in a Serpentinizing Aquifer Are Assembled through Strong Concurrent Dispersal Limitation and Selection

Lindsay I. Putman,<sup>a,b\*</sup> Mary C. Sabuda,<sup>c,d</sup> William J. Brazelton,<sup>e</sup> Michael D. Kubo,<sup>f,g</sup> Tori M. Hoehler,<sup>g</sup> Tom M. McCollom,<sup>h</sup> Dawn Cardace,<sup>i</sup> Matthew O. Schrenk<sup>a,b</sup>

<sup>a</sup>Department of Earth and Environmental Sciences, Michigan State University, East Lansing, Michigan, USA

<sup>b</sup>Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan, USA

<sup>c</sup>Department of Earth and Environmental Sciences, University of Minnesota—Twin Cities, Minneapolis, Minnesota, USA

<sup>d</sup>BioTechnology Institute, University of Minnesota—Twin Cities, St. Paul, Minnesota, USA

<sup>e</sup>Department of Biology, University of Utah, Salt Lake City, Utah, USA

<sup>f</sup>SETI Institute, Mountain View, California, USA

<sup>g</sup>Exobiology Branch, NASA Ames Research Center, Moffett Field, California, USA

<sup>h</sup>Laboratory for Atmospheric and Space Physics, UCB 600, University of Colorado—Boulder, Boulder, Colorado, USA

<sup>i</sup>Department of Geosciences, University of Rhode Island, Kingston, Rhode Island, USA

**ABSTRACT** In recent years, our appreciation of the extent of habitable environments in Earth's subsurface has greatly expanded, as has our understanding of the biodiversity contained within. Most studies have relied on single sampling points, rather than considering the long-term dynamics of subsurface environments and their microbial populations. One such habitat are aquifers associated with the aqueous alteration of ultramafic rocks through a process known as serpentinization. Ecological modeling performed on a multiyear time series of microbiology, hydrology, and geochemistry in an ultrabasic aquifer within the Coast Range Ophiolite reveals that community assembly is governed by undominated assembly (i.e., neither stochastic [random] nor deterministic [selective] processes alone govern assembly). Controls on community assembly were further assessed by characterizing aquifer hydrogeology and microbial community adaptations to the environment. These analyses show that low permeability rocks in the aquifer restrict the transmission of microbial populations between closely situated wells. Alpha and beta diversity measures and metagenomic and metatranscriptomic data from microbial communities indicate that high pH and low dissolved inorganic carbon levels impose strong environmental selection on microbial communities within individual wells. Here, we find that the interaction between strong selection imposed by extreme pH and enhanced ecological drift due to dispersal limitation imposed by slow fluid flow results in the undominated assembly signal observed throughout the site. Strong environmental selection paired with extremely low dispersal in the subsurface results in low diversity microbial communities that are well adapted to extreme pH conditions and subject to enhanced stochasticity introduced by ecological drift over time.

**IMPORTANCE** Microbial communities existing under extreme or stressful conditions have long been thought to be structured primarily by deterministic processes. The application of macroecology theory and modeling to microbial communities in recent years has spurred assessment of assembly processes in microbial communities, revealing that both stochastic and deterministic processes are at play to different extents within natural environments. We show that low diversity microbial communities in a hard-rock serpentinizing aquifer are assembled under the influence of strong selective processes imposed by high pH and enhanced ecological drift that occurs as the result

**Citation** Putman LI, Sabuda MC, Brazelton WJ, Kubo MD, Hoehler TM, McCollom TM, Cardace D, Schrenk MO. 2021. Microbial communities in a serpentinizing aquifer are assembled through strong concurrent dispersal limitation and selection. *mSystems* 6:e00300-21. <https://doi.org/10.1128/mSystems.00300-21>.

**Editor** Angela D. Kent, University of Illinois at Urbana-Champaign

**Copyright** © 2021 Putman et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Matthew O. Schrenk, [schrenkm@msu.edu](mailto:schrenkm@msu.edu).

\* Present address: Lindsay I. Putman, Department of Biological Sciences, Michigan Technological University, Houghton, Michigan, USA.

**Received** 5 April 2021

**Accepted** 19 August 2021

**Published** 14 September 2021

of dispersal limitation due to the slow movement of water in the low permeability aquifer. This study demonstrates the important roles that both selection and dispersal limitation play in terrestrial serpentinites, where extreme pH assembles a microbial metacommunity well adapted to alkaline conditions and dispersal limitation drives compositional differences in microbial community composition between local communities in the subsurface.

**KEYWORDS** serpentinization, groundwater, community assembly, extreme pH, microbial ecology, deep subsurface, hard-rock aquifer

Recent estimates of the distribution of biomass on Earth indicate that bacteria are the second largest reservoir of biomass (~70 gigatons of carbon) and that the majority of bacterial and archaeal biomass is hosted within subsurface environments (~64 gigatons of carbon) (1). In the last decade, several deep subsurface environments have been studied and characterized in detail, allowing us to improve our understanding of the composition and function of microbial communities in this biome and the important roles they play in global biogeochemical cycling (2–4). One type of deep subsurface habitat that has become the subject of intensive study over the last decade is serpentinized peridotite. Serpentinization is a geochemical reaction that hydrates ultramafic rock (peridotite) (5) and generates high pH, reducing fluids with abundant hydrogen (6). In recent years, the microbiology of serpentinizing systems has been studied in a variety of marine (7–9) and terrestrial (6, 10–16) environments. This work has confirmed the presence of endemic, low diversity, microbial communities capable of tolerating the extreme pH (15, 16) and utilizing available hydrogen (17, 18), carbon monoxide (17, 19), methane (14, 20), acetate (6, 14, 21), formate (18–21), and sulfur compounds (22) for both dissimilatory and assimilatory metabolic processes. Combined, this work has provided ample evidence that distinct communities of microorganisms inhabit serpentinizing environments and that they play active roles in these ecosystems.

To date, the ecological forces that structure microbial communities and drive turnover within and between microbial communities in serpentinizing systems have yet to be investigated. Community assembly describes the stochastic (dispersal and ecological drift) and deterministic (environmental and/or biological selection) ecological processes that structure observed microbial communities (23, 24). Advances in DNA sequencing technology in the last 2 decades (25), paired with the more recent implementation of community assembly theory and modeling of sequencing data from microbial ecosystems, have provided new insights into the mechanisms that structure microbial communities in a variety of conditions (24, 26–36). Assessments of community assembly in groundwater systems have shown that selection and dispersal processes play an extremely important role in structuring both the environmental resource landscape as well as microbial communities (29, 31, 32, 34, 37). Due to accessibility, much of this work has been performed in neutral pH, near surface, groundwaters with connectivity to the surface (31, 32, 37). Relatively little work has been done in isolated hard-rock aquifers within the deep subsurface (29, 34). Old groundwaters (>100 years), which are often hosted within hard-rock aquifers in the deep subsurface, are a substantial reservoir of globally available freshwater (38). As these systems are likely dispersal limited, it is important that we understand the ecological forces that structure microbial communities in hard-rock aquifers (29, 34) to better account for the controls on microbial community composition and function within these large reservoirs of valuable water resources (39).

In this study, we apply a null modeling ecological framework (28) that uses the  $\beta$ -nearest taxon index ( $\beta$ NTI) and the abundance-weighted Raup-Crick (RCbray) metric to quantify microbial community assembly processes in an isolated serpentinizing aquifer. We apply the modeling framework to the longest running and highest resolution time series data set collected from a serpentinizing system thus far, to better understand the forces that drive assembly and microbial community turnover within these

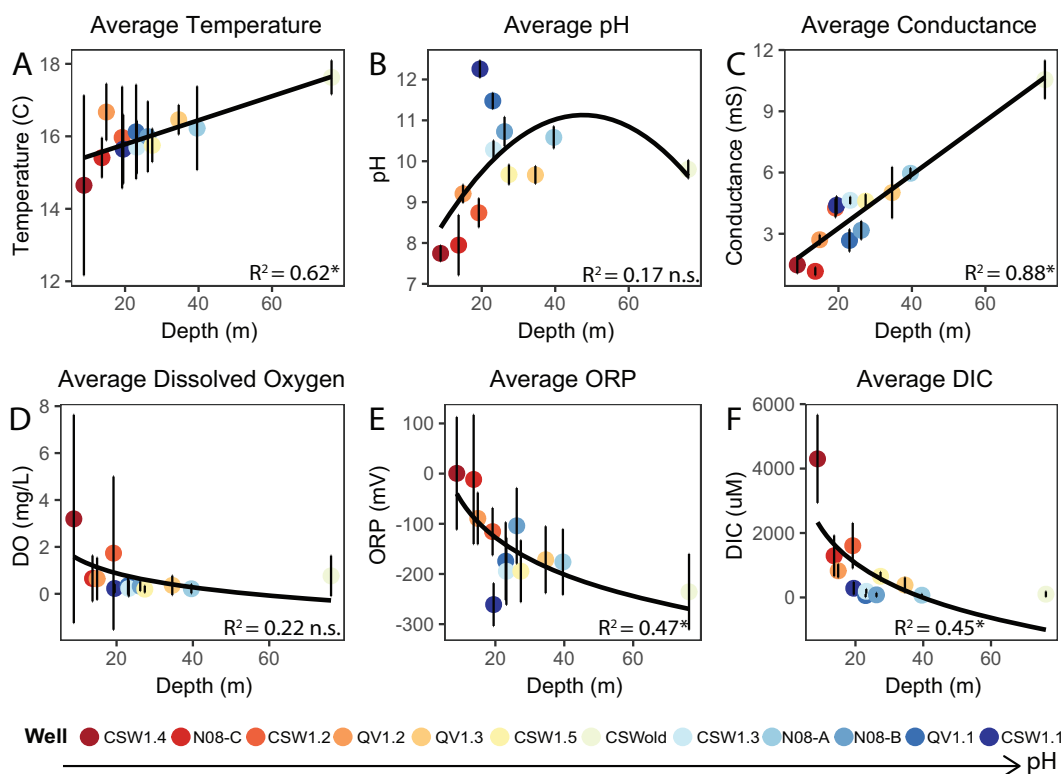
unique systems. Additionally, we integrate an aqueous geochemistry data set collected alongside microbial samples to better understand environmental drivers of microbial community composition. We perform basic hydrologic modeling to estimate aquifer hydraulic properties to better understand movement of water and the physical environment that microbial communities experience in the subsurface. Our data show that pH imposes a strong selective force on microbial community composition but that, overall, slow fluid flow and poor connectivity in the subsurface isolates microbial communities from one another. This isolation introduces a significant amount of stochasticity by enhancing ecological drift within and between these alkaliphilic microbial communities.

## RESULTS

**Geochemistry.** Fluids were sampled from 12 wells of various depths (7 to 76 m) that access a high pH aquifer in the serpentinizing ophiolite at the Coast Range Ophiolite Microbial Observatory (CROMO) in Northern California (40). Fluids were sampled two or three times per year over the course of 6 years. Sampled fluids, which are pumped under positive pressure via permanently installed submersible pumps, traveled directly through a chamber with a multiprobe meter to measure temperature, pH, dissolved oxygen (DO), oxidation-reduction potential (ORP), and specific conductance (11, 15). Luer-lock syringes were attached directly to the pump outflow tubing for anoxic sampling of fluids to support subsequent chemical analyses as described in Text S1 in the supplemental material. Analysis of the fluid chemistry (see Table S1 at <https://doi.org/10.6084/m9.figshare.14983851>) over the duration of the study shows that CROMO fluids generally become more reducing, increase in temperature, salinity, and pH, and decrease in DO and dissolved inorganic carbon (DIC) concentrations with depth (Fig. 1A to F).

**Ecological modeling.** DNA and RNA were extracted from microbial communities that were captured on 0.22- $\mu\text{m}$  Sterivex filter cartridges (Millipore, Billerica, MA, USA) during sampling trips from 2011 to 2017 as described below and in Text S1. The following diversity and ecological modeling results are based on a data set generated from gene amplicon sequencing of the V4 region of the 16S rRNA gene. The data set utilized for these analyses consists of 104 samples, 5,974,056 sequence reads, and 13,444 operational taxonomic units (OTUs) clustered at a 3% distance threshold as described below. The 16S rRNA gene amplicon data set represents 51 different phyla, and community composition is largely dominated by members of the *Betaproteobacteriales* and *Clostridiales* orders (see Fig. S1 in the supplemental material). The OTU count table (CROMO\_Filtered\_FINAL\_counts.xlsx) utilized for the following analyses and ecological modeling are published on Figshare ([https://figshare.com/projects/Community\\_Assembly\\_in\\_Serpentinizing\\_Ophiolites/101648](https://figshare.com/projects/Community_Assembly_in_Serpentinizing_Ophiolites/101648)).

Ecological modeling of the microbial communities was carried out to characterize the relative contributions of different community assembly processes, using the framework developed by Stegen and colleagues (28). This framework employs null modeling techniques to generate randomized communities that can be compared with observed microbial community composition to determine whether microbial communities are more or less similar to each other than would be expected if communities assembled by random chance. First,  $\beta\text{NTI}$  values are calculated to differentiate between deterministic (selective) and stochastic (random) processes. Significant  $\beta\text{NTI}$  values indicate that deterministic processes are responsible for observed differences between microbial communities in a given pairwise comparison (37). Nonsignificant  $\beta\text{NTI}$  values indicate that stochastic processes are responsible for observed differences between microbial communities (28). Stochastic assembly processes are further investigated using the RCbray metric to characterize the roles of dispersal and ecological drift (28). A detailed description of each community assembly process associated with ecological modeling results (28) can be found in Table 1.  $\beta\text{NTI}$  (CROMO\_weighted\_bNTI\_matrix.csv) and RCbray (CROMO\_RCbray\_matrix.xlsx) matrices used for the following analyses are available on Figshare ([https://figshare.com/projects/Community\\_Assembly\\_in\\_Serpentinizing\\_Ophiolites/101648](https://figshare.com/projects/Community_Assembly_in_Serpentinizing_Ophiolites/101648)).

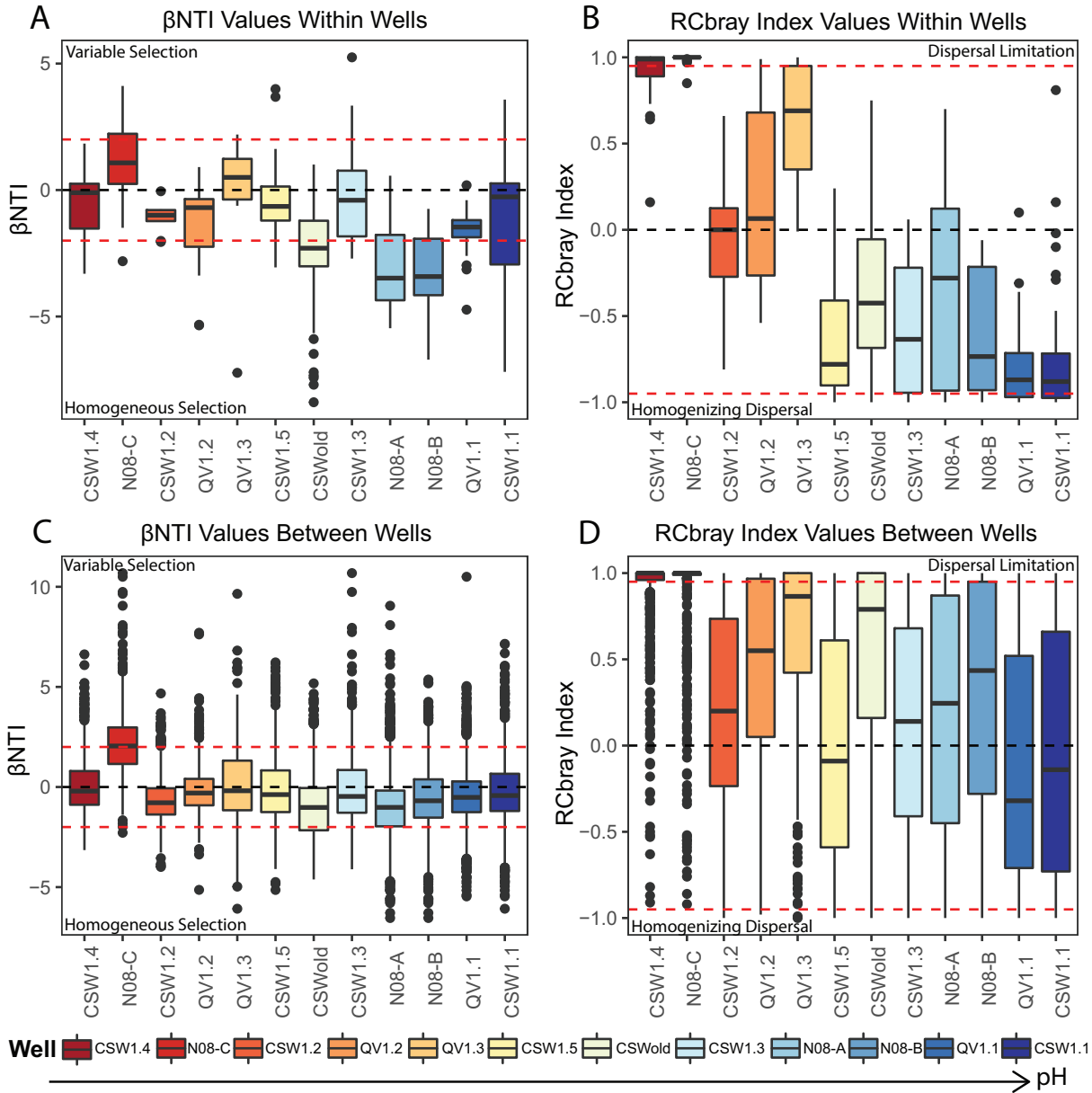


**FIG 1** Average measurements for temperature (A), pH (B), specific conductance (C), dissolved oxygen (DO) (D), oxidation reduction potential (ORP) (E), and dissolved inorganic carbon (DIC) (F) in each well plotted against depth. Averages were calculated using data from 2011 to 2017.  $R^2$  coefficients from regression models are listed in the bottom right corner of each plot, where an asterisk indicates a significant correlation ( $P \leq 0.05$ ) and n.s. indicates a nonsignificant correlation. Wells are organized from lowest to highest average pH in the legend.

Community assembly processes were assessed within individual wells over time (Fig. 2A and B) and between wells over time (Fig. 2C and D). Within individual wells, selection processes do play an important role and can account for 9 to 66% of microbial community turnover observed between samples as a function of time (see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>). However, individual wells with more neutral pH conditions (7.5 to 9) display stochastic assembly signatures ( $|\beta\text{NTI}| < 2$ ), with some influence of variable selection

**TABLE 1** Description of community assembly processes associated with ecological modeling results

Process	Description	Deterministic or stochastic	Model result
Variable selection	Communities are more different from each other than can be expected by random chance (i.e., different physical/chemical conditions in samples drive differing community composition).	Deterministic	$\beta\text{NTI} > 2$
Homogeneous selection	Communities are more similar to each other than can be expected by random chance (i.e., similar physical/chemical conditions in samples drive community composition to be similar).	Deterministic	$\beta\text{NTI} < -2$
Dispersal limitation	Microbial communities between samples are unable to interact due to separation by space or time. When communities cannot interact, variation in composition caused by ecological drift over time results in communities that are less similar to each other than expected by chance.	Stochastic	$ \beta\text{NTI}  < 2$ and $\text{RCbray} > 0.95$
Homogenizing dispersal	Microbial communities between samples can freely interact. Free and easy mixing between communities results in communities that are more similar to each other than expected by chance.	Stochastic	$ \beta\text{NTI}  < 2$ and $\text{RCbray} < -0.95$
Undominated	While both deterministic and stochastic processes are at play, differences in observed community composition between samples cannot be explained by either selection or random processes (i.e., nonsignificant $\beta\text{NTI}$ and $\text{RCbray}$ result).	Both	$ \beta\text{NTI}  < 2$ and $ \text{RCbray}  < 0.95$



**FIG 2** Boxplots showing the distribution of  $\beta$ NTI (A) and RCbray (B) values within individual wells over time and the distribution of  $\beta$ NTI (C) and RCbray (D) values between wells over time. Boxplots represent pairwise comparisons within individual wells over time (A and B), and pairwise comparisons between samples from each individual well and all other wells (C and D). The dashed black line in plots represents the null expectation. Dashed red lines in each plot represent significance thresholds for both the  $\beta$ NTI and RCbray metrics. Plotted RCbray values (B and D) do not include values from pairwise comparisons that have significant  $\beta$ NTI values. See Table 1 for further clarification and interpretation of model results. Wells are organized from lowest to highest average pH in boxplots. Boxplots display the distribution of the data, where the main colored box displays the interquartile range (25th to 75th percentile) of the data. Solid lines within the boxes represent the median of the data. Whiskers on the plot display the maximum and minimum expected values of the data distribution, and black colored points represent outlier data points with respect to the plotted distribution.

( $\beta$ NTI > 2) between certain time points (Fig. 2A). The neutral pH well N08-C shows a particularly strong influence of variable selection (25%) over time, indicating that local geochemical conditions may fluctuate enough within the well to drive greater than expected microbial community turnover (see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>). Moderate pH (9 to 10.5) wells are primarily influenced by stochastic processes over time with some influence of both variable and homogeneous selection processes between certain time points (Fig. 2A). Extreme pH (10.5 to 12) wells become more strongly influenced by homogeneous selection ( $\beta$ NTI < -2) (Fig. 2A), although stochastic processes still play a

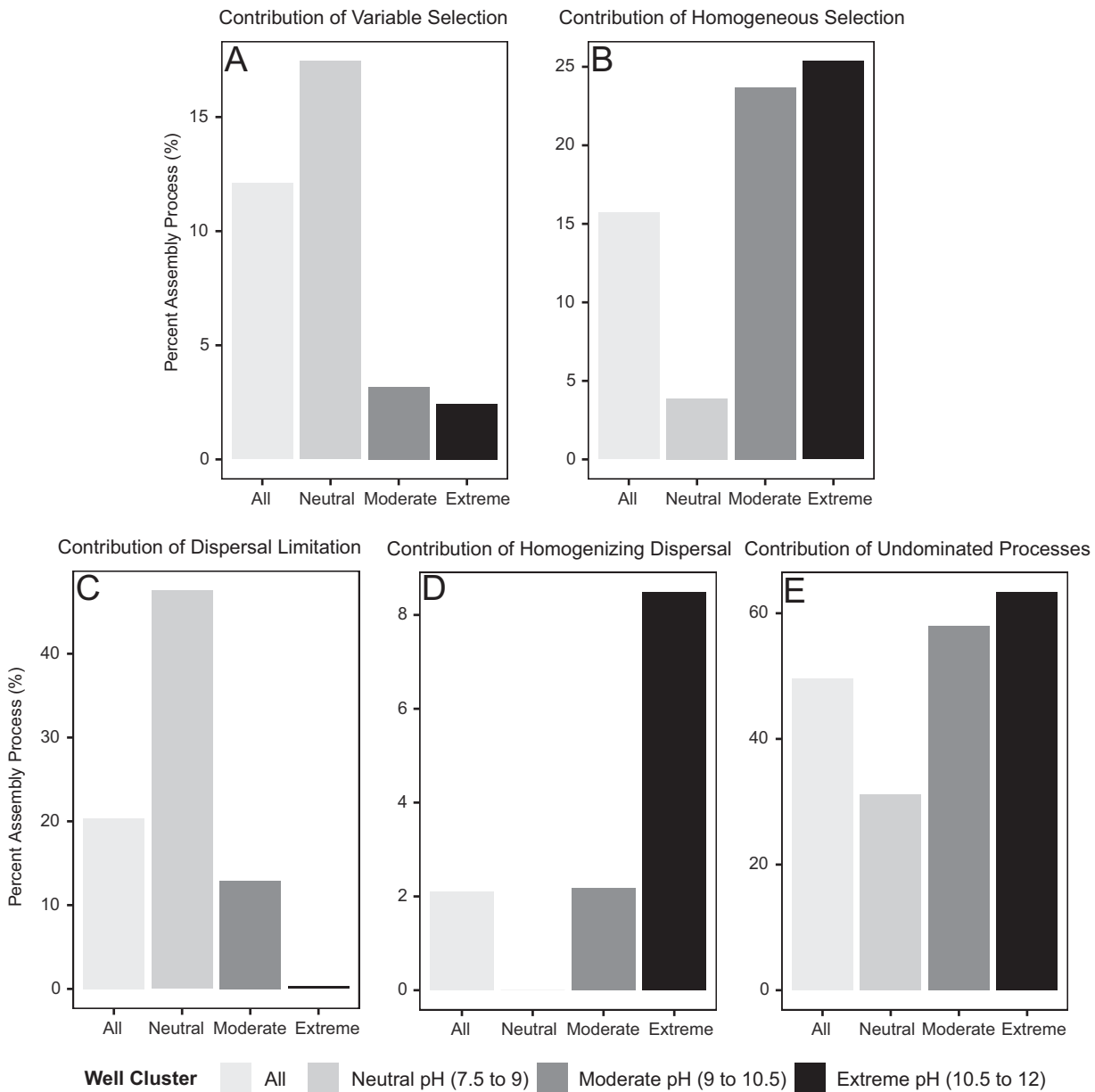
significant role within these wells, accounting for 34 to 91% of assembly within an individual well (see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>). In general, no single assembly process was able to explain variation in microbial community composition between time points (14 to 84% undominated processes) (Fig. 2B; see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>), indicating that no single assembly process can explain variation in microbial community composition between time points (undominated assembly).

In pairwise comparisons between individual wells,  $\beta$ NTI mean and median values lie close to zero, indicating that stochastic processes dominate (Fig. 2C). Nevertheless, selection processes still have some influence on between-well comparisons. Variable selection strongly influences pairwise comparisons between the neutral pH (7.5 to 9) wells CSW1.4 and N08-C with other higher pH (9 to 12) wells at the site, while homogeneous selection plays a stronger role in pairwise comparisons between the deepest wells at the site N08-A (39.6 m) and CSWold (76.2 m) and other medium to deep wells (20 to 35 m) with moderate to extreme pH (9 to 12) (Fig. 2C). Further characterization of stochastic processes at play in comparisons between wells reveals that dispersal limitation plays a larger role compared to within-well comparisons, but that overall, no single assembly process dominates between-well comparisons (undominated assembly) (Fig. 2B and D). Mean and median RCbray values trend more toward dispersal limitation, with numerous samples within the distribution of pairwise comparisons crossing the significance threshold (RCbray > 0.95) (Fig. 2D). A small proportion of between-well pairwise comparisons cross the significance threshold for being under the influence of homogenizing dispersal (RCbray < -0.95), suggesting that microbial communities associated with those specific wells are mixing and interacting with each other (Fig. 2D). These comparisons are primarily associated with the two partially cased wells at the site, CSW1.1 and QV1.1, and nearby wells at similar depths. Discussion of the increased influence of homogenizing dispersal in wells CSW1.1 and QV1.1 (87) can be found in Text S1.

Mantel tests were used to assess correlations between ecological modeling metrics and measured environmental parameters (see Table S3 at <https://doi.org/10.6084/m9.figshare.14983860>). Environmental variables were transformed into Euclidean distance matrices so they could be correlated with ecological modeling matrices.  $\beta$ NTI values were significantly correlated with pH ( $R = 0.29$ ), ORP ( $R = 0.23$ ), and DIC ( $R = 0.14$ ) (Fig. S2). This correlation indicates that greater differences in environmental measurements of pH, ORP, and DIC are associated with variable selection (i.e., more positive values of  $\beta$ NTI). Samples that are more different from each other with respect to pH, ORP, and DIC conditions are likely to be influenced by variable selection, whereas samples with similar pH, ORP, and DIC conditions are likely to be influenced by homogeneous selection. RCbray values were significantly correlated with well depth ( $R = 0.32$ ), temperature ( $R = 0.15$ ), pH ( $R = 0.27$ ), specific conductance ( $R = 0.34$ ), ORP ( $R = 0.30$ ), and DIC ( $R = 0.33$ ) (Fig. S3). That is, samples with greater differences in environmental conditions are more likely to be influenced by dispersal limitation (high RCbray), while samples with more similar environmental conditions are more likely to show influences of homogenizing dispersal. The relative strength of selection processes on microbial communities seems to be tightly linked to pH, ORP, and DIC conditions, while the influence of stochastic processes is correlated with a wider range of environmental conditions.

Finally, the contribution of each community assembly process was quantified by compiling the number of significant pairwise comparisons for each process and dividing these numbers by the total number of pairwise comparisons. Overall, half of the comparisons could not be attributed to a single assembly process (50% undominated processes), with the other half distributed among dispersal limitation (20%), homogeneous selection (16%), variable selection (12%), and homogenizing dispersal (2%) (Fig. 3; see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>). Due to the importance of pH in structuring ecological niches in the system (Fig. S4), community assembly processes were also quantified for all pairwise comparisons between neutral pH (7.5 to 9) wells (CSW1.4, N08-C, CSW1.2, and QV1.2), moderate pH (9 to 10.5) wells (QV1.3, CSW1.5, CSWold, and CSW1.3) and extreme pH (10.5 to 12) wells (N08-A, N08-B,





**FIG 3** Bar plots showing the overall contribution of variable selection (A), homogeneous selection (B), dispersal limitation (C), homogenizing dispersal (D), and undominated processes (E) for all wells, neutral pH wells (pH 7.5 to 9), moderate pH wells (pH 9 to 10.5), and extreme pH wells (pH 10.5 to 12). The contribution (percentage) of each assembly process was calculated within each grouping of samples by dividing the number of significant pairwise comparisons for each assembly process by the total number of pairwise comparisons within the sample grouping. Pairwise comparisons include within-well comparisons and between-well comparisons for all wells and wells within each pH category.

QV1.1, and CSW1.1). When separated into discrete pH ranges, trends in the influence of community assembly processes with pH are evident. As pH increases, the role of variable selection decreases (from 17% to 2%), the role of homogeneous selection increases (from 4% to 25%), the role of dispersal limitation decreases (from 48% to 0%), the role of homogenizing dispersal increases (from 0% to 8%), and the role of undominated processes increase (from 31% to 63%) (Fig. 3A to E; see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>). The increasing role of homogeneous selection indicates that microbial communities are more like each other than expected by chance as pH conditions become more extreme (Fig. 3B). The shift from dispersal limitation (Fig. 3C) to homogenizing dispersal (Fig. 3D) with increasing pH indicates that microbial communities are less isolated from one another, and interact more, at high pH.

**Physical controls on microbial dispersal.** Ecological modeling results revealed that CROMO microbial communities are overwhelmingly assembled through undominated processes (50% undominated processes; see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>), where neither stochastic nor deterministic processes entirely govern assembly (Table 1). Since the modeling results indicate that half the data analyzed in the model are governed by a mixture of stochastic and deterministic processes that cannot be well defined by the model (Table 1), alternative assessments of dispersal and environmental selection (following section) were performed to better understand physicochemical drivers of assembly at CROMO. Connectivity of the aquifer to the surface and estimation of aquifer hydraulic properties were assessed to better understand dispersal within the subsurface.

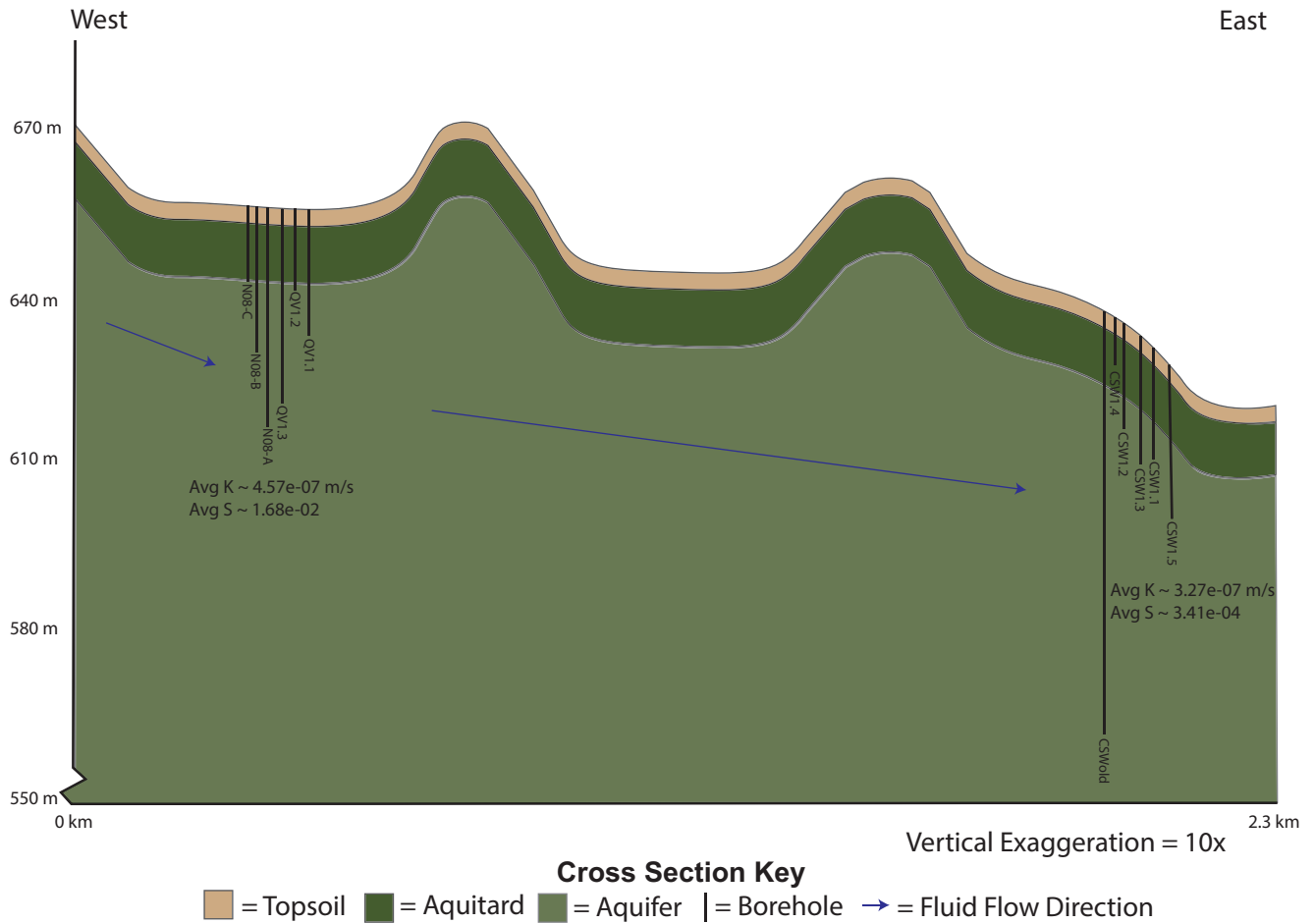
Characterization of the CROMO aquifer's connectivity to the surface and aquifer hydraulic properties was carried out for this study using tritium analyses and pumping tests as described below. Previous work by Ortiz and colleagues (41) indicated that the main aquifer at CROMO is confined. Tritium analyses (see Table S4 at <https://doi.org/10.6084/m9.figshare.14983863>) performed on samples collected in May of 2017 were below detection (0.8 tritium unit [TU]) in all wells except for a shallow well, N08-C, where tritium was detected at 0.9 TU.

Pumping tests used to estimate aquifer hydraulic conductivity ( $K$ ) indicate that water flows through the subsurface at a low rate ( $K = 10^{-7}$  m/s) (Fig. 4; see Table S5 at <https://doi.org/10.6084/m9.figshare.14983866>). Estimates of  $K$  and other hydraulic properties (see Table S5 at <https://doi.org/10.6084/m9.figshare.14983866>) are in line with estimates obtained at the Samail ophiolite, Sultanate of Oman (42), and Koniambo massif in New Caledonia (43).

**Adaptations to environmental conditions.** Given the distinct geochemical trends observed at CROMO (Fig. 1), a permutational multivariate analysis of variance (PERMANOVA) analysis was performed on Bray-Curtis dissimilarities and measured environmental data to determine how much variance in the community dissimilarity data can be explained by measured environmental data (Table 2). Results from this analysis show that well location (i.e., corresponding subsurface location) ( $R^2 = 0.468$ ) is the primary driver of differences in microbial community composition. Significant differences in microbial community composition are also seen over time ( $R^2 = 0.029$ ) and are driven by differences in DIC ( $R^2 = 0.012$ ), DO ( $R^2 = 0.011$ ), and pH ( $R^2 = 0.010$ ) to a lesser extent (Table 2). PERMANOVA results also indicate that a significant amount of variation in microbial community composition is unexplained (residuals  $R^2 = 0.44$ ) (Table 2). It is important to note that community composition changes related to changes in time, DIC, DO, and pH may already be accounted for as a portion of the large amount of variance explained by individual well location. This observation could explain why the amount of variation described by each individual variable is quite low ( $R^2$  of 0.01 to 0.029). We performed an additional PERMANOVA analysis without the well location variable to further explore this result. While less of the overall variance in community composition is explained (residuals  $R^2 = 0.69$ ), more variance in community composition is associated with geochemical metadata, especially pH ( $R^2=0.13$ ) (see Table S6 at <https://doi.org/10.6084/m9.figshare.14983869>). The increased variance accounted for in the data when considering the well location variable likely includes both measured and unmeasured physicochemical data specific to each well location (Table 2).

Since pH appears to play an important role in structuring ecological niches at CROMO (Fig. S4; Table 2; see Table S6 at <https://doi.org/10.6084/m9.figshare.14983869>), as well as microbial communities in a variety of environments (24, 33), the influence of pH in structuring CROMO microbial communities was further assessed. While DO and DIC concentrations were also found to drive changes in community composition (Table 2), DO was not found to be significantly correlated with ecological model  $\beta$ NTI values, which assess the role of environmental selection (Fig. S2; see Table S3 at <https://doi.org/10.6084/m9.figshare.14983860>). DIC was significantly correlated with  $\beta$ NTI values but is not assessed here, since pH has been shown to govern the





**FIG 4** Cross section with 10× vertical exaggeration showing the topographic profile of the landscape around the CROMO wells. Well locations are scattered close to the cross section line, so well locations were estimated in their approximate order from west to east. Wells are drawn to accurate depth based on the surface elevation within the profile. The depths of the topsoil and aquitard layers were estimated based on geophysical mapping done by Ortiz and colleagues (41). Given that the same units were present at both the CSW and QV clusters and were about the same thickness, we made the simplifying assumption that the topsoil and aquitard layers are laterally continuous and of even thickness between well cluster locations. We estimated that the main aquifer was 76.2 m thick. It needed to accommodate the depth of CSWold, but we do not have a good estimate as to the depth where basement rock begins. The general direction of groundwater flow is shown within the diagram, and average estimates for hydraulic conductivity (*K*) and storativity (*S*) at each well cluster are included from Table S5 at <https://doi.org/10.6084/m9.figshare.14983866>.

amount of available DIC in serpentinizing systems (44). Due to this, environmental selection due to changes in DIC availability should ultimately be controlled by pH conditions.

Community richness (number of unique taxa in a community) and evenness (how close in numerical abundance taxa in a community are; measured by Pielou’s evenness index) decrease as pH increases (Fig. 5A and B). This result supports previous reports that pH is a major driver of microbial community diversity at the site (15), as has been seen in numerous other environments (24, 33). To further assess this point, metagenomic and metatranscriptomic data collected during the 2011 to 2013 field campaigns and again in 2016 were screened for genes associated with adaptations to high pH (see Table S7 at <https://doi.org/10.6084/m9.figshare.14372030>). It is known through physiological studies of alkaliphilic cultures at high pH that a proton (H<sup>+</sup>) gradient exists from cell cytoplasm to cell exterior that can drive loss of protons from the cell (45). To combat this issue, microorganisms must actively transport protons back into the cell (45). The use of cation/proton antiporters, which can transport sodium (Na<sup>+</sup>) or potassium (K<sup>+</sup>) ions out of the cell while transporting protons into the cell, as well as Na<sup>+</sup>-pumping V-type ATPases, which pump Na<sup>+</sup> ions across the cell membrane to generate ATP, are common mechanisms used by microorganisms growing in alkaline pH

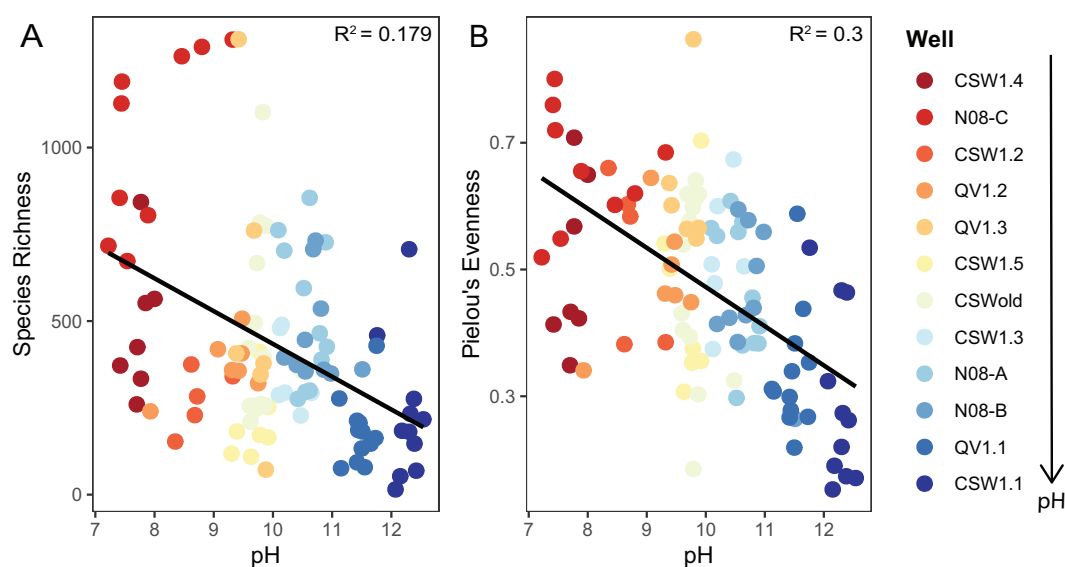
**TABLE 2** Results from PERMANOVA analysis on Bray-Curtis dissimilarities and associated sample metadata

Variable	PERMANOVA results					
	Degrees of freedom	Sum of squares	Mean squares	F score	R <sup>2a</sup>	P value <sup>b</sup>
Well	11	16.239	1.47626	8.0442	0.46783	<b>0.001</b>
Day	1	1.018	1.01783	5.5461	0.02932	<b>0.001</b>
Temp (°C)	1	0.145	0.14524	0.7914	0.00418	0.702
pH	1	0.355	0.35524	1.9357	0.01023	<b>0.028</b>
Conductance (mS)	1	0.274	0.27357	1.4907	0.00788	0.113
DO (mg/liter)	1	0.376	0.37644	2.0512	0.01084	<b>0.014</b>
ORP (mV)	1	0.274	0.27357	1.4907	0.00788	0.09
DIC (μM)	1	0.431	0.43122	2.3497	0.01242	<b>0.006</b>
Residuals	85	15.599	0.18352		0.4494	
Total	103	34.711			1	

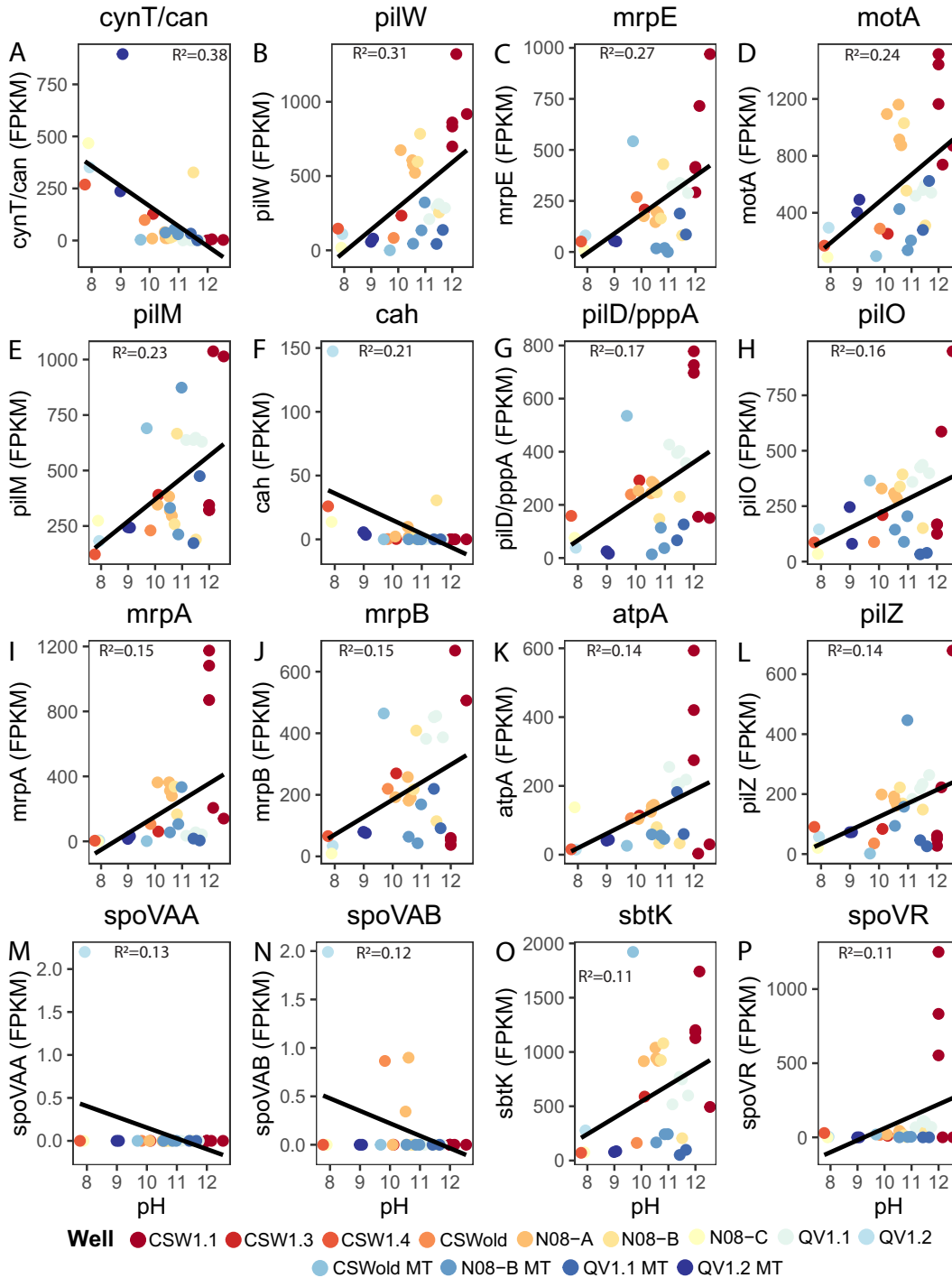
<sup>a</sup>R<sup>2</sup> values indicate the percent variation in community dissimilarity that can be described by an individual variable.

<sup>b</sup>Significant P values are bolded.

conditions (45, 46). We find that CROMO metagenomes have abundant Na<sup>+</sup>/H<sup>+</sup> and K<sup>+</sup>/H<sup>+</sup> antiporters, and Na<sup>+</sup>-pumping V-type ATPases (see Table S7 at <https://doi.org/10.6084/m9.figshare.14372030>). The abundance and transcription of genes encoding carbonic anhydrases *cynT/can* and *cah* and sporulation genes *spoVAA* and *spoVAB* decrease with increasing pH (Fig. 6A, F, M, and N), while the abundance and transcription of genes associated with motility, pili for motility and adhesion, Na<sup>+</sup>/H<sup>+</sup> antiporters, Na<sup>+</sup>-dependent V-type ATPases, Na<sup>+</sup>-dependent bicarbonate transporters, and the sporulation gene *spoVR* increase with increasing pH (Fig. 6B to E, G to L, and O and P). The abundance and transcription of genes associated with motility and adhesion, Na<sup>+</sup>/H<sup>+</sup> antiporters, Na<sup>+</sup>-dependent bicarbonate transporters, Na<sup>+</sup>-pumping V-type ATPases, sporulation, and survival-related transcription factors increase with increasing salinity and depth (Fig. S5A to R). Despite a great abundance of sporulation genes observed in wells deep (>27 m) within the aquifer, poor transcription of these genes paired with evidence for the transcription of other maintenance and metabolic genes (20, 22) indicate that microbial community members were not actively employing this strategy



**FIG 5** Microbial community richness (A) and evenness (B) plotted against pH. Points represent individual microbial community samples, which are colored by well location. Wells are organized from lowest to highest average pH in the legend. Linear regression lines are plotted in black. Significant ( $P \leq 0.05$ ) regression  $R^2$  values are displayed in the upper right corner of plots.



**FIG 6** Significant linear regression plots of the normalized abundance of metagenomic and metatranscriptomic data in fragments per kilobase per million reads (FPKM) for the genes *cynT/can* (A), *pilW* (B), *mrpE* (C), *motA* (D), *pilM* (E), *cah* (F), *pilD/pppA* (G), *pilO* (H), *mrpA* (I), *mrpB* (J), *atpA* (K), *pilZ* (L), *spoVAA* (M), *spoVAB* (N), *sbtK* (O), and *spoVR* (P) plotted against pH. Solid black lines represent a linear trend line. Linear regression  $R^2$  values are included at the top of each plot. Samples of metagenomic and metatranscriptomic data are colored by well location. Metatranscriptomic samples are labeled with an MT following the well name in the legend. Functions associated with genes are as follows: motility/adhesion (*pilW*, *motA*, *pilM*, *pilD/pppA*, *pilO*, and *pilZ*),  $\text{Na}^+/\text{H}^+$  antiporters (*mrpE*, *mrpA*, and *mrpB*), sporulation (*spoVAA*, *spoVAB*, and *spoVR*),  $\text{Na}^+$ -dependent bicarbonate transporters (*sbtK*),  $\text{Na}^+$ -dependent V-type ATPase (*atpA*), carbonic anhydrase (*cynT/can* and *cah*).

to cope with extreme environmental conditions at the time of sampling. The abundance and transcription of genes associated with sporulation,  $\text{Na}^+/\text{H}^+$  antiporters, and  $\text{Na}^+$ -pumping V-type ATPases decrease with increasing ORP (Fig. S6A to F). The abundance and transcription of genes associated with carbonic anhydases increase with increasing

concentrations of DIC (Fig. S6G and L), while the abundance of genes associated with  $\text{Na}^+/\text{H}^+$  antiporters and with motility and adhesion decrease with increasing concentrations of DIC (Fig. S6H to K). The abundance and transcription of genes associated with motility and adhesion, sporulation,  $\text{Na}^+$ -pumping V-type ATPases,  $\text{Na}^+$ -dependent bicarbonate transporters, and  $\text{K}^+/\text{H}^+$  symporters increase with increasing temperature (Fig. S6M to V). Linear regression model results for significant and insignificant regressions can be found in Table S8 (<https://doi.org/10.6084/m9.figshare.14983872>) and Table S9 (<https://doi.org/10.6084/m9.figshare.14939202>), respectively.

In addition to the community diversity, metagenomic, and metatranscriptomic evidence provided, phylogenetic analysis of these communities using the nearest taxon index (a within sample, or alpha diversity, variation of the  $\beta$ NTI metric used for ecological modeling) indicate that CROMO microbial communities are composed of closely related organisms and demonstrate significant phylogenetic clustering (Fig. S7). Phylogenetic clustering indicates that environmental filtering and the conservation of adaptive traits have likely occurred as the result of habitat specialization (47).

## DISCUSSION

Given the extreme pH conditions at CROMO, we anticipated that deterministic processes would be dominant, as has been observed in other systems under extreme or stressful environmental conditions (33, 48–52). Ecological modeling results revealed that community assembly at CROMO instead occurs through a combination of deterministic and stochastic processes that cannot be well defined (undominated assembly; see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>). Because the model cannot define the stochastic and deterministic processes at work when undominated assembly governs pairwise comparisons, we employed alternative methods to assess the roles of dispersal (Fig. 4; see Table S4 at <https://doi.org/10.6084/m9.figshare.14983863>) and selection (Fig. 5 and 6 and Fig. S5, S6, and S7; Table 2) within the CROMO aquifer. Here, we explore the combined results and aim to describe the stochastic and deterministic processes at work within the undominated fraction of data identified by the ecological model. We discuss the roles of dispersal and selection in the CROMO aquifer, incorporating results and interpretations from the multiple data streams reported here, as well as previous reports and interpretations of community assembly relevant to observations in this study.

While most assembly processes at CROMO cannot be well defined (undominated assembly), dispersal limitation (20%) and homogeneous selection (16%) are the two primary quantifiable processes at work within the aquifer (Fig. 3; see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>). Evidence of dispersal limitation is clear when looking at the hydrogeology of the system (Fig. 4). Tritium levels below detection in well fluids (see Table S4 at <https://doi.org/10.6084/m9.figshare.14983863>) indicate that the main aquifer is not well connected to the surface and that modern recharge (fluid <50 years in age) does not contribute appreciably to the subsurface reservoir (38, 53, 86). When considered alongside hydraulic conductivity estimates (see Table S5 at <https://doi.org/10.6084/m9.figshare.14983866>), these results indicate that the CROMO aquifer is a confined system receiving no detectable modern recharge and is characterized by poor connectivity between wells and low rates of fluid flow (Fig. 4). Differences in compositional beta diversity (Bray-Curtis dissimilarity; Table 2) provide additional evidence of dispersal limitation. The mean Bray-Curtis dissimilarity at the site when considering all pairwise comparisons is extremely high (0.8), indicating that compositional diversity of microbial communities have little overlap (54, 55). Differences in well location (and conditions—see results relevant to Table 2 and Table S6 at <https://doi.org/10.6084/m9.figshare.14983869>) alone account for nearly 50% of observed variation in community dissimilarity, while changes in community composition over time follow at a much lower level of explained variance (~3%) (Table 2). Additionally, 44% of the variation in microbial community composition could not be explained by environmental variables during PERMANOVA analysis. The unexplained variation in community composition could be the result of stochastic

changes in community composition due to ecological drift. Isolation of microbial communities due to dispersal limitation could certainly introduce enough ecological drift within each well location to generate such high levels of microbial community compositional dissimilarity (56–59). Environmental heterogeneity and differences in community size could be another potential explanation for the observed high beta diversity (58). However, hydrologic data paired with the fact that the null ecological modeling method accounts for differences in alpha diversity (60) indicate that model results are accurate and that the high observed beta diversity is the result of ecological drift.

The influence of selection is also clear throughout the site. The metacommunity (average sitewide NTI = 1.73), as well as local communities sampled through time, show significant phylogenetic clustering, indicating that microorganisms within the aquifer are highly adapted to environmental conditions (see Fig. S7 in the supplemental material). While it is unlikely that this observation is solely due to high pH adaptations, pH is known to be a significant driver of microbial community composition (33, 61), and high pH conditions and the physiological adaptations required to survive under these conditions likely played a significant role in structuring the observed microbial communities at CROMO. Microbial community diversity measures are also significantly affected by pH conditions (Fig. 5), and metagenomic and metatranscriptomic data indicate that pH homeostasis mechanisms used in alkaline conditions are prevalent and actively transcribed within the wells (Fig. 6; see Table S7 at <https://doi.org/10.6084/m9.figshare.14372030>). The increased abundance and transcription of genes associated with motility at high pH, including flagellar synthesis genes, and type IV pili (Fig. 6B, D and E, G and H, and L) potentially represent a means to improve access to both limited available DIC and oxidants in the deep subsurface or to enable the attachment of organisms to minerals to form biofilms and/or aid in electron transport to mineral surfaces. Improved access to substrates via motility or adhesion and the capability to form biofilms that can lower pH within the microenvironment (62) may be key survival strategies that allow microbial populations to persist under hyperalkaline conditions in the deep subsurface. Additionally, factors such as DO concentrations (Table 2), redox conditions (Fig. S6A to F), salinity (Fig. S5A to I), temperature (Fig. S6M to V), and DIC concentrations (Table 2; Fig. S6G to L), which are likely controlled by pH, and other unmeasured variables likely play a smaller role in driving community composition differences across the wells (Table 2; see Table S6 at <https://doi.org/10.6084/m9.figshare.14983869>). While ecological modeling results did not identify homogeneous selection as the dominant assembly process across the site (see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>), other measures of microbial community diversity (Fig. 5; see Table S2 at <https://doi.org/10.6084/m9.figshare.14983857>; see Table S6 at <https://doi.org/10.6084/m9.figshare.14983869>), phylogeny (Fig. S4 and S7), and physiology (Fig. 6; see Table S7 at <https://doi.org/10.6084/m9.figshare.14372030>) indicate that pH plays a major role in structuring microbial community composition in the CROMO aquifer.

Given the strong evidence for selection and dispersal limitation within the physicochemical and microbial community data, the preponderance of undominated processes at the site was initially surprising. However, recent studies of soil (33) and a fractured shale aquifer (34) using the Stegen et al. (28) ecological modeling framework provide some insight. Tripathi and colleagues (33) showed that while pH plays a significant role in structuring soil microbial communities, selection imposed by pH accounted for only ~17% of observed assembly. As in our study, this low value indicates that other assembly processes still play a substantial role in structuring microbial communities under extreme pH conditions. Recent work by Danczak and colleagues (34) in a fractured shale aquifer highlighted how strong concurrent homogeneous selection and variable selection counteract to result in an undominated community assembly signal. While this same mechanism is not observed at CROMO, results from Danczak and colleagues (34) demonstrate that an undominated assembly signal can be the result of strong counteracting processes. Extreme pH and slow fluid flow at CROMO consistently impose strong counteracting deterministic and

stochastic processes, resulting in an undominated assembly signal observed across the site (Fig. 3E).

Our observations at CROMO are also supported by recent computational and experimental work that has shown that dispersal conditions and microbial community size affect the importance of selection in structuring microbial communities (30, 59). Strong environmental selection can result in low diversity and low biomass communities that are inherently prone to enhanced drift as theorized by Vellend (23) and observed in computational modeling experiments by Evans and colleagues (30). Ecological drift in communities assembled under strong selection can be further enhanced under low dispersal conditions (30, 59), resulting in microbial communities composed of a small number of highly abundant species, with large variation in the birth and death rates of rare taxa within the communities (59). Microbial communities characterized in serpentinizing systems thus far have notoriously low diversity and are primarily composed of a few dominant taxa and many highly variable low abundance species (6, 15, 16, 22, 44), which likely makes serpentinite microbial communities inherently prone to enhanced ecological drift.

Finally, we propose a more nuanced mechanism of assembly within the serpentinizing subsurface, relevant to the geological history (63) and long groundwater residence times in the region (64). Poor connectivity to the surface (see Table S4 at <https://doi.org/10.6084/m9.figshare.14983863>) and slow fluid flow in the subsurface (Fig. 4; see Table S5 at <https://doi.org/10.6084/m9.figshare.14983866>) indicate that dispersal of microorganisms into and out of the region is extremely slow. Microorganisms present within the serpentinizing subsurface were therefore potentially transported to the terrestrial surface when the ophiolite was obducted and continued to grow and persist (22, 65) or were introduced to the system by persisting hundreds or thousands of years along slow regional groundwater flow pathways, similar to observations of community assembly patterns in deep seafloor sediments (52). In either case, extreme pH conditions buffered by the presence of serpentine minerals present throughout much of the ophiolite (66) impose strong environmental filtering and likely select for well-adapted alkaliphilic microorganisms long before they reach the subsurface localities observed in this study (52). Within this geologic context, we suggest that homogeneous selection imposed by extreme pH structures a low diversity alkaliphilic metacommunity within the larger region of the CROMO aquifer from which the observed local communities in this study are assembled. Within observed local communities, homogeneous selection maintains the persistence of dominant alkaliphilic microbial community members. At the same time, poor dispersal between local communities and variability in the birth and death rates of the large pool of rare community members results in enhanced ecological drift over space and time (23, 30, 59). Overall, these results highlight the important roles and complex interplay that occurs between selection, dispersal, and ecological drift in structuring subsurface microbial communities. Given the great heterogeneity observed in subterranean environments, it is critical to continue to assess community assembly processes in these systems to better constrain how different physical and environmental conditions structure microbial communities within the expansive subsurface biosphere.

## MATERIALS AND METHODS

**Site description and sample collection.** CROMO is located within the McLaughlin Natural Reserve near Lower Lake, CA. This site lies within the Coast Range Ophiolite and consists of four wells drilled >30 years ago and eight wells drilled in August of 2011. The latter set of wells was drilled using techniques designed to minimize and quantify subsurface contamination (40). Initial characterization of CROMO fluids indicated a strong influence of serpentinization (low Eh, pH > 11) and microbial communities that share similarity to other characterized serpentinite springs (11, 15, 40). The main aquifer is confined by a well-cemented aquitard (41), and tritium analyses, described below, indicate no mixing of modern surface water within the aquifer (53, 67, 86). CROMO has been sampled two or three times per year since the site was established in 2011, resulting in a high-resolution time series data set of geochemistry and microbial biodiversity. The sampling procedure utilized to collect fluids from CROMO wells has been previously described (11, 15). Geochemical and microbial sample collection methods are described in detail in Text S1 in the supplemental material.



**Topographic profile and cross section construction.** A topographic profile and cross section of the study area (Fig. 4) were created using ArcMap (Esri, Redlands, CA, USA), and Adobe Illustrator (Adobe, San Jose, CA, USA). Well head locations and elevations were obtained as described in Text S1 and drawn into the cross section at their proper location and depth. Well head coordinate and elevation data are listed in Table S10 (<https://doi.org/10.6084/m9.figshare.14983893>). A detailed description of topographic profile methods and assumptions made in the construction of the cross section can be found in Text S1. Detailed maps displaying the location of McLaughlin Reserve in California, a geologic map of the CROMO site, and the location of each well cluster at the site can be found in the article by Ortiz et al. (41).

**Estimation of aquifer properties.** Aquifer properties were estimated using the program AQTESOLV (68). Drawdown/Recovery analysis was performed within AQTESOLV to obtain estimates of transmissivity ( $T$ ), and storativity ( $S$ ) to calculate hydraulic conductivity ( $K$ ) using  $T$  and aquifer thickness ( $b$ ) according to the following equation:  $T = K \times b$  (see Table S5 at <https://doi.org/10.6084/m9.figshare.14983866>). Detailed description of data collection, model parameters (83), and fitting model solutions (84, 85) can be found in Text S1.

**Tritium sample collection.** Water samples for tritium analyses were collected in June 2017. Water was collected into 0.5-liter narrow-mouth high-density polyethylene bottles. Samples were air-tight, and bottle mouths were wrapped with parafilm to prevent evaporation or leakage from the bottles. Samples were stored at room temperature and shipped to University of Waterloo Environmental Isotope Laboratory (UW-EIL, Waterloo, ON, Canada). Detailed methods of tritium analyses performed by UW-EIL are described in Text S1. Measured tritium values for each well are included in Table S4 (<https://doi.org/10.6084/m9.figshare.14983863>).

**Geochemical analyses.** Analytical methods used to measure DIC have been previously described (11, 15). Briefly, fluid samples were acidified to convert all forms of DIC to carbon dioxide ( $\text{CO}_2$ ), which could partition into the headspace of the sealed sample vial. Subsequently, the concentration of  $\text{CO}_2$  in the headspace gas was quantified by gas chromatography with flame ionization detection (FID) (SRI 8610C; SRI Instruments, Torrance, CA, USA), using an inline methanizer to convert liberated  $\text{CO}_2$  to methane prior to passage through the FID. All analyses were performed in duplicate. Geochemical data used for analyses can be found in Table S1 (<https://doi.org/10.6084/m9.figshare.14983851>).

**Extraction of DNA and RNA.** Extractions of DNA and RNA from 0.22- $\mu\text{m}$  Sterivex filter cartridges were performed as previously described by Twing and colleagues (15) and Sabuda and colleagues (22), respectively. A brief description can be found in Text S1.

**Sample preparation, sequencing, and data analysis of metagenomes and metatranscriptomes.** Metagenomic and metatranscriptomic sequences were previously reported (15, 20, 22, 88). Sample preparation and sequencing of metagenomes and metatranscriptomes were carried out as previously described (22). The assemblies and predicted protein annotations reported here were performed as previously described by Sabuda and colleagues (22).

**16S rRNA gene amplicon sequencing.** Throughout the course of the project, samples were submitted for amplicon sequencing of the V4 region of the 16S rRNA gene at three different sequencing centers: Department of Energy Joint Genome Institute (JGI), Marine Biological Laboratory's (MBL) Josephine Bay Paul Center, and the Michigan State University (MSU) Genomics Core Facility (see Table S11 at <https://doi.org/10.6084/m9.figshare.14983896>). Sequences generated by the JGI were reported by Twing and colleagues (15). Sequences generated by the MBL were reported by Crespo-Medina and colleagues (11). Samples sequenced at the MSU Genomics Core were submitted for sequencing of the V4 region of the 16S rRNA gene. Amplification, quantification, and sequencing procedures performed by the MSU Genomics Core have been previously described (22). Blanks and extraction blanks collected alongside samples in this data set could not be quantified or amplified and were not submitted for sequencing.

**16S rRNA sequence processing.** Sequences generated by JGI, MBL, and the MSU Genomics Core were processed using mothur v1.39.5 (69) as previously described (11, 15, 22). Quality-trimmed fasta files and count tables for sequences from each sequencing center were concatenated together and clustered into OTUs at a 3% distance threshold using the *de novo* distance-based greedy clustering (DGC) method as implemented in mothur v1.39.5 (70). *De novo* clustering has successfully been used to compile and reanalyze 16S rRNA data from multiple sources for large meta-analysis studies of the human microbiome (71, 72). Continued discussion on the successful use of DGC to cluster sequences from different sequencing centers and analyses of sequences clustered using this method are included in Text S1. Clustered OTUs were aligned to the SILVA SSURef alignment (v132), and taxonomic classifications were assigned using mothur.

Following the successful merger of the 16S rRNA data sets (23,994 OTUs and 6,210,850 reads), count data from sample replicates were averaged and rounded to the nearest whole count number to avoid statistical issues that can arise from pseudoreplication (73, 74). Singletons, which rounded to zero when present within a set of replicates, were removed (8,526 OTUs), and sequences identified as eukaryotes (58 OTUs and 1,475 reads), archaea (138 OTUs and 7,271 reads), mitochondria (46 OTUs and 1,295 reads), chloroplasts (94 OTUs and 7,881 reads), and unknown (842 OTUs and 9,127 reads) by SILVA were removed. Following this process, the 16S rRNA sequences were screened for potential contaminants associated with DNA extraction reagents and for human skin and feces-derived organisms previously identified in deep subsurface samples by Sheik and colleagues (75). Following analysis of the taxonomy and distribution of each potential contaminant within the data set, sequences that were likely contaminants were removed from the data set (792 OTUs and 111,576 reads). In addition to this cleanup, two human contaminant microorganisms, *Simkania negevensis* (76) and *Akkermansia muciniphila* (77), were removed that were identified at high abundance in a handful of samples. These organisms from families cvE6 and *Akkermansiaceae* accounted for a maximum of 39% and 48% of the reads in the affected

samples, respectively, and 1.6% of total reads from the data set. There were no patterns in sample contamination by these two organisms when looking at sample collection or sample extraction logs. A similar approach to removing contaminant sequences from a deep subsurface data set was successfully used by Fullerton and colleagues (78) and Sabuda and colleagues (88). The final 16S rRNA data set used for analysis consisted of 13,444 OTUs and retained 96% of the original reads (5,974,056 reads). Following data set filtering, a phylogenetic tree was generated using FastTree v2.1.3 (79).

Excel files containing raw count table data, averaged and contaminant filtered data, and count and taxonomy information for removed OTUs (<https://doi.org/10.6084/m9.figshare.14879535> and <https://doi.org/10.6084/m9.figshare.14371964>) are available on Figshare along with R code used to identify and screen potential contaminant OTUs ([https://figshare.com/projects/Community\\_Assembly\\_in\\_Serpentinizing\\_Ophiolites/101648](https://figshare.com/projects/Community_Assembly_in_Serpentinizing_Ophiolites/101648)).

**Statistical analyses.** Basic statistical analyses and data exploration of the community data and associated metadata were performed using the R packages phyloseq and vegan (80, 81). These packages were used to combine and organize 16S rRNA tag sequencing data with corresponding sample metadata and to perform basic data visualization and data analysis. Correlations between the ecological modeling matrices and environmental variables were performed using Mantel tests, using the command `mantel.rtest()` from the R package `ade4` (82). PERMANOVA analyses were performed in `vegan` using the `adonis()` function (28). PERMANOVA results were used to define significant relationships between community data and environmental variables. Sample richness was calculated using the function `specnumber()` from the R package `vegan` (81). Pielou's evenness was calculated by dividing the Shannon diversity index, calculated using the function `diversity()` in `vegan` (81), by sample richness, which was previously calculated. Geochemical data, richness and Pielou's evenness data, and metagenomic and metatranscriptomic data were plotted in `ggplot()`, and linear regression models were calculated using the `lm()` function.

**Ecological modeling.** Ecological modeling of community and phylogenetic turnover within the system was performed according to the framework developed by Stegen and colleagues (28). A basic assumption of the framework is that there is significant phylogenetic signal (i.e., closely related organisms have similar habitat preferences) at short phylogenetic distances, making the use of the  $\beta$ NTI metric, which assesses phylogenetic turnover between close relatives, an ideal measure for the framework (60). Mantel correlograms were used to test for phylogenetic signal at different phylogenetic distance classes. Positive, significant correlations at short phylogenetic distances were observed when pH environmental niches were assessed (see Fig. S4 in the supplemental material), indicating that the community assembly framework developed by Stegen and colleagues (28) was an appropriate choice for the data set analyzed here.

The framework developed by Stegen and colleagues uses two different metrics to quantify the contributions of selection, dispersal, and ecological drift that contribute toward an observed assembled community (28). This model can quantify all processes except for speciation/diversification as has been outlined in Vellend's seminal paper that conceptualizes the processes of community assembly (23). The Stegen et al. framework (28) first assesses the role of selection by assessing changes in phylogenetic distance using beta-mean nearest taxon distance and  $\beta$ NTI for all pairwise comparisons within the data set. Once the roles of selective processes have been assessed, pairwise comparisons that were not significant with the  $\beta$ NTI metric are assessed with the RCbray metric to look at the role of the stochastic processes of dispersal and drift (28). Detailed description of the model and how each metric is calculated are included within Text S1.

All R scripts, excel, and csv files used in analyses are available on Figshare at [https://figshare.com/projects/Community\\_Assembly\\_in\\_Serpentinizing\\_Ophiolites/101648](https://figshare.com/projects/Community_Assembly_in_Serpentinizing_Ophiolites/101648).

**Data availability.** The 16S rRNA gene sequence data used in this work are publicly available in the NCBI Sequence Read Archive (SRA) under the BioProject accession number [PRJNA690585](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA690585). CROMO metagenome sequences previously published by Twing and colleagues (15) are publicly available in the JGI IMG/M database under the project identifiers (IDs) 1021918, 1021921, 1021924, and 1021927 and in the MG-RAST database under the following sample IDs: 4569549.3, 4569550.3, 4569551.3, and 4569552.3. Metagenome sequences previously published by Seyler and colleagues (20) are publicly available under the BioProject IDs: [PRJNA410019](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410019), [PRJNA410020](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410020), [PRJNA410022](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410022), [PRJNA410035](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410035), [PRJNA410037](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410037), [PRJNA410553](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410553), [PRJNA410555](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410555), [PRJNA410036](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410036), [PRJNA410024](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410024), [PRJNA410028](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410028), [PRJNA410025](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410025), [PRJNA410023](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410023), [PRJNA410027](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410027), and [PRJNA410026](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA410026). Metagenomic sequences previously published by Sabuda and colleagues (88) are publicly available in the NCBI SRA database under the following accession IDs: [SRX9385611](https://www.ncbi.nlm.nih.gov/sra/SRX9385611), [SRX9385612](https://www.ncbi.nlm.nih.gov/sra/SRX9385612), and [SRX9385613](https://www.ncbi.nlm.nih.gov/sra/SRX9385613). CROMO metatranscriptome sequences previously published by Sabuda and colleagues (22) are publicly available in the SRA under the following accession IDs: [SRX3339504](https://www.ncbi.nlm.nih.gov/sra/SRX3339504), [SRX3339503](https://www.ncbi.nlm.nih.gov/sra/SRX3339503), [SRX3339089](https://www.ncbi.nlm.nih.gov/sra/SRX3339089), [SRX3331179](https://www.ncbi.nlm.nih.gov/sra/SRX3331179), [SRX3331177](https://www.ncbi.nlm.nih.gov/sra/SRX3331177), [SRX3330963](https://www.ncbi.nlm.nih.gov/sra/SRX3330963), [SRX3330943](https://www.ncbi.nlm.nih.gov/sra/SRX3330943), and [SRX3330753](https://www.ncbi.nlm.nih.gov/sra/SRX3330753).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**TEXT S1**, DOCX file, 0.1 MB.

**FIG S1**, EPS file, 0.9 MB.

**FIG S2**, TIF file, 0.3 MB.

**FIG S3**, TIF file, 0.4 MB.

**FIG S4**, EPS file, 0.6 MB.

**FIG S5**, EPS file, 1.5 MB.

**FIG S6**, EPS file, 1.7 MB.

**FIG S7**, EPS file, 1.2 MB.

**FIG S8**, EPS file, 1.5 MB.

**FIG S9**, EPS file, 1.5 MB.

## ACKNOWLEDGMENTS

We thank Cathy Koehler and Paul Aigner, Codirectors of the University of California—Davis McLaughlin Reserve for their support and help in establishing CROMO and continued field work at the site. Thanks to Kate Fullerton (University of Tennessee, Knoxville) for advice and examples of the contaminant filtering process and Brent Heerspink (MSU) for guidance and feedback on hydrological modeling work. Special thanks to Katrina Twing and Melitza Crespo-Medina for their involvement in this project and sample collection during the early years of this project.

This work was supported by the NASA Astrobiology Institute CAN-7 Rock Powered Life grant NNA15BB02A.

L.I.P. and M.O.S. created and designed the study. L.I.P., M.C.S., W.J.B., M.D.K., T.M.M., D.C., and M.O.S. performed field sampling at CROMO. M.D.K. processed and analyzed all geochemical data. W.J.B. prepared and assembled metagenomic and metatranscriptomic data. M.O.S. analyzed and extracted relevant data from metagenomes and metatranscriptomes. L.I.P. performed 16S rRNA processing, groundwater modeling, ecological modeling and statistical analyses of modeling results and associated geochemical data. T.M.M. and T.M.H. provided critical edits and insight to geochemical dynamics of the system. L.I.P. wrote the manuscript and created all final figures and tables with input and discussion from M.C.S., W.J.B., M.D.K., T.M.H., T.M.M., D.C., and M.O.S. All authors critically evaluated the manuscript.

## REFERENCES

- Bar-On YM, Phillips R, Milo R. 2018. The biomass distribution on Earth. *Proc Natl Acad Sci U S A* 115:6506–6511. <https://doi.org/10.1073/pnas.1711842115>.
- Flynn TM, Sanford RA, Ryu H, Bethke CM, Levine AD, Ashbolt NJ, Santo Domingo JW. 2013. Functional microbial diversity explains groundwater chemistry in a pristine aquifer. *BMC Microbiol* 13:146. <https://doi.org/10.1186/1471-2180-13-146>.
- 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>.
- Lau MCY, Kieft TL, Kuloyo O, Linage-Alvarez B, van Heerden E, Lindsay MR, Magnabosco C, Wang W, Wiggins JB, Guo L, Perlman DH, Kyin S, Shwe HH, Harris RL, Oh Y, Yi MJ, Purtschert R, Slater GF, Ono S, Wei S, Li L, Lollar BS, Onstott TC. 2016. An oligotrophic deep-subsurface community dependent on syntrophy is dominated by sulfur-driven autotrophic denitrifiers. *Proc Natl Acad Sci U S A* 113:E7927–E7936. <https://doi.org/10.1073/pnas.1612244113>.
- McCormack TM, Bach W. 2009. Thermodynamic constraints on hydrogen generation during serpentinization of ultramafic rocks. *Geochim Cosmochim Acta* 73:856–875. <https://doi.org/10.1016/j.gca.2008.10.032>.
- Kohl L, Cumming E, Cox A, Rietze A, Morrissey L, Lang SQ, Richter A, Suzuki S, Neelson KH, Morrill PL. 2016. Exploring the metabolic potential of microbial communities in ultra-basic, reducing springs at The Cedars, CA, USA: experimental evidence of microbial methanogenesis and heterotrophic acetogenesis. *J Geophys Res Biogeosci* 121:1203–1220. <https://doi.org/10.1002/2015JG003233>.
- Schrenk MO, Kelley DS, Bolton SA, Baross JA. 2004. Low archaeal diversity linked to subsurface geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ Microbiol* 6:1086–1095. <https://doi.org/10.1111/j.1462-2920.2004.00650.x>.
- Brazelton WJ, Mehta MP, Kelley DS, Baross JA. 2011. Physiological differentiation within a single-species biofilm fueled by serpentinization. *mBio* 2:e00127-11. <https://doi.org/10.1128/mBio.00127-11>.
- Quéméneur M, Bes M, Postec A, Mei N, Hamelin J, Monnin C, Chavagnac V, Payri C, Pelletier B, Guentas-Dombrowsky L, Gérard M, Pisapia C, Gérard E, Ménez B, Ollivier B, Erauso G. 2014. Spatial distribution of microbial communities in the shallow submarine alkaline hydrothermal field of the Prony Bay, New Caledonia. *Environ Microbiol Rep* 6:665–674. <https://doi.org/10.1111/1758-2229.12184>.
- Tiago I, Verissimo A. 2013. Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization. *Environ Microbiol* 15:1687–1706. <https://doi.org/10.1111/1462-2920.12034>.
- Crespo-Medina M, Twing KI, Kubo MDY, Hoehler TM, Cardace D, McCollom T, Schrenk MO. 2014. Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5:604. <https://doi.org/10.3389/fmicb.2014.00604>.
- Sánchez-Murillo R, Gazel E, Schwarzenbach EM, Crespo-Medina M, Schrenk MO, Boll J, Gill BC. 2014. Geochemical evidence for active tropical serpentinization in the Santa Elena Ophiolite, Costa Rica: an analog of a humid early Earth? *Geochem Geophys Geosyst* 15:1783–1800. <https://doi.org/10.1002/2013GC005213>.
- Woycheese KM, Meyer-Dombard DR, Cardace D, Argayosa AM, Arcilla CA. 2015. Out of the dark: transitional subsurface-to-surface microbial diversity in a terrestrial serpentinizing seep (Manleluag, Pangasinan, the Philippines). *Front Microbiol* 6:44. <https://doi.org/10.3389/fmicb.2015.00044>.
- Brazelton WJ, Thornton CN, Hyer A, Twing KI, Longino AA, Lang SQ, Lilley MD, Früh-Green GL, Schrenk MO. 2017. Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy. *PeerJ* 5:e2945. <https://doi.org/10.7717/peerj.2945>.
- Twing KI, Brazelton WJ, Kubo MDY, Hyer AJ, Cardace D, Hoehler TM, McCollom T, Schrenk MO. 2017. Serpentinization-influenced groundwater harbors extremely low diversity microbial communities adapted to high pH. *Front Microbiol* 8:308. <https://doi.org/10.3389/fmicb.2017.00308>.
- Rempfert KR, Miller HM, Bompard N, Nothaft D, Matter JM, Kelemen P, Fierer N, Templeton AS. 2017. Geological and geochemical controls on subsurface microbial life in the Samail Ophiolite, Oman. *Front Microbiol* 8:56. <https://doi.org/10.3389/fmicb.2017.00056>.
- Brazelton WJ, Nelson B, Schrenk MO. 2012. Metagenomic evidence for H<sub>2</sub> oxidation and H<sub>2</sub> production by serpentinite-hosted subsurface microbial communities. *Front Microbiol* 2:268. <https://doi.org/10.3389/fmicb.2011.00268>.
- Suzuki S, Kuenen JG, Schipper K, van der Velde S, Ishii S, Wu A, Sorokin DY, Tenney A, Meng X, Morrill PL, Kamagata Y, Muyzer G, Neelson KH. 2014. Physiological and genomic features of highly alkaliphilic hydrogen-

- utilizing Betaproteobacteria from a continental serpentinizing site. *Nat Commun* 5:3900. <https://doi.org/10.1038/ncomms4900>.
19. Fones EM, Colman DR, Kraus EA, Nohaft DB, Poudel S, Rempfert KR, Spear JR, Templeton AS, Boyd ES. 2019. Physiological adaptations to serpentinization in the Samail Ophiolite, Oman. *ISME J* 13:1750–1762. <https://doi.org/10.1038/s41396-019-0391-2>.
  20. Seyler LM, Brazelton WJ, McLean C, Putman LI, Hyer A, Kubo MDY, Hoehler T, Cardace D, Schrenk MO. 2020. Carbon assimilation strategies in ultrabasic groundwater: clues from the integrated study of a serpentinization-influenced aquifer. *mSystems* 5:e00607-19. <https://doi.org/10.1128/mSystems.00607-19>.
  21. Crespo-Medina M, Twing KI, Sánchez-Murillo R, Brazelton WJ, McCollom TM, Schrenk MO. 2017. Methane dynamics in a tropical serpentinizing environment: the Santa Elena Ophiolite, Costa Rica. *Front Microbiol* 8:916. <https://doi.org/10.3389/fmicb.2017.00916>.
  22. Sabuda MC, Brazelton WJ, Putman LI, McCollom TM, Hoehler TM, Kubo MDY, Cardace D, Schrenk MO. 2020. A dynamic microbial sulfur cycle in a serpentinizing continental ophiolite. *Environ Microbiol* 22:2329–2345. <https://doi.org/10.1111/1462-2920.15006>.
  23. Vellend M. 2010. Conceptual synthesis in community ecology. *Q Rev Biol* 85:183–206. <https://doi.org/10.1086/652373>.
  24. Nemerugut 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>.
  25. Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S, Rappuoli R. 2008. Microbiology in the post-genomic era. *Nat Rev Microbiol* 6:419–430. <https://doi.org/10.1038/nrmicro1901>.
  26. Chase JM, Kraft NJB, Smith KG, Vellend M, Inouye BD. 2011. Using null models to disentangle variation in community dissimilarity from variation in  $\alpha$ -diversity. *Ecosphere* 2:art24. <https://doi.org/10.1890/ES10-00117.1>.
  27. Zhou J, Liu W, Deng Y, Jiang Y, Xue K, He Z, Van Nostrand JD, Wu L, Yang Y, Wang A. 2013. Stochastic assembly leads to alternative communities with distinct functions in a bioreactor microbial community. *mBio* 4:e00584-12. <https://doi.org/10.1128/mBio.00584-12>.
  28. 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>.
  29. Beaton ED, Stevenson BS, King-Sharp KJ, Stamps BW, Nunn HS, Stuart M. 2016. Local and regional diversity reveals dispersal limitation and drift as drivers for groundwater bacterial communities from a fractured granite formation. *Front Microbiol* 7:1933. <https://doi.org/10.3389/fmicb.2016.01933>.
  30. 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>.
  31. 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>.
  32. Danczak RE, Johnston MD, Kenah C, Slattery M, Wilkins MJ. 2018. Microbial community cohesion mediates community turnover in unperturbed aquifers. *mSystems* 3:e00066-18. <https://doi.org/10.1128/mSystems.00066-18>.
  33. Tripathi BM, Stegen JC, Kim M, Dong K, Adams JM, Lee YK. 2018. Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J* 12:1072–1083. <https://doi.org/10.1038/s41396-018-0082-4>.
  34. Danczak RE, Daly RA, Borton MA, Stegen JC, Roux S, Wrighton KC, Wilkins MJ. 2020. Ecological assembly processes are coordinated between bacterial and viral communities in fractured shale ecosystems. *mSystems* 5:e00098-10. <https://doi.org/10.1128/mSystems.00098-20>.
  35. Liu W, Graham EB, Zhong L, Zhang J, Li W, Li Z, Lin X, Feng Y. 2020. Dynamic microbial assembly processes correspond to soil fertility in sustainable paddy agroecosystems. *Funct Ecol* 34:1244–1256. <https://doi.org/10.1111/1365-2435.13550>.
  36. Liu W, Graham EB, Dong Y, Zhong L, Zhang J, Qiu C, Chen R, Lin X, Feng Y. 2021. Balanced stochastic versus deterministic assembly processes benefit diverse yet uneven ecosystem functions in representative agroecosystems. *Environ Microbiol* 23:391–404. <https://doi.org/10.1111/1462-2920.15326>.
  37. 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>.
  38. Gleeson T, Befus KM, Jasechko S, Luijendijk E, Cardenas MB. 2016. The global volume and distribution of modern groundwater. *Nature Geosci* 9:161–167. <https://doi.org/10.1038/ngeo2590>.
  39. Lall U, Josset L, Russo T. 2020. A snapshot of the world's groundwater challenges. *Annu Rev Environ Resour* 45:171–196. <https://doi.org/10.1146/annurev-environ-102017-025800>.
  40. Cardace D, Hoehler T, McCollom T, Schrenk M, Carnevale D, Kubo M, Twing K. 2013. Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling objectives and preliminary outcomes. *Sci Drill* 16:45–55. <https://doi.org/10.5194/sd-16-45-2013>.
  41. Ortiz E, Tominaga M, Cardace D, Schrenk MO, Hoehler TM, Kubo MD, Rucker DF. 2018. Geophysical characterization of serpentinite hosted hydrogeology at the McLaughlin Natural Reserve, Coast Range Ophiolite. *Geochem Geophys Geosyst* 19:114–131. <https://doi.org/10.1002/2017GC007001>.
  42. Dewandel B, Lachassagne P, Boudier F, Al-Hattali S, Ladouche B, Pinault J, Al-Suleimani Z. 2005. A conceptual hydrogeological model of ophiolite hard-rock aquifers in Oman based on a multiscale and a multidisciplinary approach. *Hydrogeol J* 13:708–726. <https://doi.org/10.1007/s10040-005-0449-2>.
  43. Jeanpert J, Iseppi M, Adler PM, Genthon P, Sevin B, Thovert J-F, Dewandel B, Join J-L. 2019. Fracture controlled permeability of ultramafic basement aquifers. Inferences from the Koniambo massif, New Caledonia. *Eng Geol* 256:67–83. <https://doi.org/10.1016/j.enggeo.2019.05.006>.
  44. Schrenk MO, Brazelton WJ, Lang SQ. 2013. Serpentinization, carbon, and deep life. *Rev Mineral Geochem* 75:575–606. <https://doi.org/10.2138/rmg.2013.75.18>.
  45. Krulwich TA, Sachs G, Padan E. 2011. Molecular aspects of bacterial pH sensing and homeostasis. *Nat Rev Microbiol* 9:330–343. <https://doi.org/10.1038/nrmicro2549>.
  46. Miller HM, Matter JM, Kelemen P, Ellison ET, Conrad ME, Fierer N, Ruchala T, Tominaga M, Templeton AS. 2016. Modern water/rock reactions in Oman hyperalkaline peridotite aquifers and implications for microbial habitability. *Geochim Cosmochim Acta* 179:217–241. <https://doi.org/10.1016/j.gca.2016.01.033>.
  47. Fine PVA, Kembel SW. 2011. Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography* 34:552–565. <https://doi.org/10.1111/j.1600-0587.2010.06548.x>.
  48. Chase JM. 2007. Drought mediates the importance of stochastic community assembly. *Proc Natl Acad Sci U S A* 104:17430–17434. <https://doi.org/10.1073/pnas.0704350104>.
  49. Ofiteru ID, Lunn M, Curtis TP, Wells GF, Criddle CS, Francis CA, Sloan WT. 2010. Combined niche and neutral effects in a microbial wastewater treatment community. *Proc Natl Acad Sci U S A* 107:15345–15350. <https://doi.org/10.1073/pnas.1000604107>.
  50. Caruso T, Chan Y, Lacap DC, Lau MCY, McKay CP, Pointing SB. 2011. Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *ISME J* 5:1406–1413. <https://doi.org/10.1038/ismej.2011.21>.
  51. Valverde A, Makhalanyane TP, Cowan DA. 2014. Contrasting assembly processes in a bacterial metacommunity along a desiccation gradient. *Front Microbiol* 5:668. <https://doi.org/10.3389/fmicb.2014.00668>.
  52. Starnawski P, Bataillon T, Ettema TJG, Jochum LM, Schreiber L, Chen X, Lever MA, Polz MF, Jørgensen BB, Schramm A, Kjeldsen KU. 2017. Microbial community assembly and evolution in seafloor sediment. *Proc Natl Acad Sci U S A* 114:2940–2945. <https://doi.org/10.1073/pnas.1614190114>.
  53. Newman BD, Osenbrück K, Aeschbach-Hertig W, Solomon DK, Cook P, Rózański K, Kipfer R. 2010. Dating of “young” groundwaters using environmental tracers: advantages, applications, and research needs. *Isotopes Environ Health Stud* 46:259–278. <https://doi.org/10.1080/10256016.2010.514339>.
  54. 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>.
  55. Graham EB, Crump AR, Resch CT, Fansler S, Arntzen E, Kennedy DW, Fredrickson JK, Stegen JC. 2016. Coupling spatiotemporal community assembly processes to changes in microbial metabolism. *Front Microbiol* 7:1949. <https://doi.org/10.3389/fmicb.2016.01949>.
  56. 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>.



57. Graham EB, Stegen JC. 2017. Dispersal-based microbial community assembly decreases biogeochemical function. *Processes* 5:65. <https://doi.org/10.3390/pr5040065>.
58. Liu J, Vellend M, Wang Z, Yu M. 2018. High beta diversity among small islands is due to environmental heterogeneity rather than ecological drift. *J Biogeogr* 45:2252–2261. <https://doi.org/10.1111/jbi.13404>.
59. Fodelianakis S, Valenzuela-Cuevas A, Barozzi A, Daffonchio D. 2021. Direct quantification of ecological drift at the population level in synthetic bacterial communities. *ISME J* 15:55–66. <https://doi.org/10.1038/s41396-020-00754-4>.
60. 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>.
61. Nemergut D, Shade A, Violle C. 2014. When, where and how does microbial community composition matter? *Front Microbiol* 5:497–494. <https://doi.org/10.3389/fmicb.2014.00497>.
62. Yin W, Wang Y, Liu L, He J. 2019. Biofilms: the microbial “protective clothing” in extreme environments. *Int J Mol Sci* 20:3423. <https://doi.org/10.3390/ijms20143423>.
63. Shervais JW, Murchey BL, Kimbrough DL, Renne PR, Hanan B. 2005. Radiometric and biostratigraphic age relations in the Coast Range Ophiolite, northern California: implications for the tectonic evolution of the Western Cordillera. *Geol Soc Am Bull* 117:633–653. <https://doi.org/10.1130/B25443.1>.
64. Fehn U, Peters EK, Tullai-Fitzpatrick S, Kubik PW, Sharma P, Teng RTD, Gove HE, Elmore D. 1992. <sup>129</sup>I and <sup>36</sup>Cl concentrations in waters of the eastern Clear Lake area, California: residence times and source ages of hydrothermal fluids. *Geochim Cosmochim Acta* 56:2069–2079. [https://doi.org/10.1016/0016-7037\(92\)90330-L](https://doi.org/10.1016/0016-7037(92)90330-L).
65. Inagaki F, Hinrichs K-U, Kubo Y, Bowles MW, Heuer VB, Hong W-L, Hoshino T, Ijiri A, Imachi H, Ito M, Kaneko M, Lever MA, Lin Y-S, Methé BA, Morita S, Morono Y, Tanikawa W, Bihan M, Bowden SA, Elvert M, Glombitza C, Gross D, Harrington GJ, Hori T, Li K, Limmer D, Liu C-H, Murayama M, Ohkouchi N, Ono S, Park Y-S, Phillips SC, Prieto-Mollar X, Purkey M, Riedinger N, Sanada Y, Sauvage J, Snyder G, Susilawati R, Takano Y, Tasumi E, Terada T, Tomaru H, Trembath-Reichert E, Wang DT, Yamada Y. 2015. Exploring deep microbial life in coal-bearing sediment down to ~2.5 km below the ocean floor. *Science* 349:420–421. <https://doi.org/10.1126/science.aaa6882>.
66. Sleep NH, Meibom A, Fridriksson T, Coleman RG, Bird DK. 2004. H<sub>2</sub>-rich fluids from serpentinization: geochemical and biotic implications. *Proc Natl Acad Sci U S A* 101:12818–12823. <https://doi.org/10.1073/pnas.0405289101>.
67. Kralik M. 2015. How to estimate mean residence times of groundwater. *Proc Earth Planet Sci* 13:301–306. <https://doi.org/10.1016/j.proeps.2015.07.070>.
68. Duffield GM. 2007. AQTESOLV for Windows user’s guide, version 4.5. Hydrosolve Inc, Reston, VA.
69. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541. <https://doi.org/10.1128/AEM.01541-09>.
70. Westcott SL, Schloss PD. 2015. De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational taxonomic units. *PeerJ* 3:e1487. <https://doi.org/10.7717/peerj.1487>.
71. Koren O, Knights D, Gonzalez A, Waldron L, Segata N, Knight R, Huttenhower C, Ley RE. 2013. A guide to enterotypes across the human body: meta-analysis of microbial community structures in human microbiome datasets. *PLoS Comput Biol* 9:e1002863. <https://doi.org/10.1371/journal.pcbi.1002863>.
72. Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. 2017. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun* 8:1784. <https://doi.org/10.1038/s41467-017-01973-8>.
73. Davies GM, Gray A. 2015. Don’t let spurious accusations of pseudoreplication limit our ability to learn from natural experiments (and other messy kinds of ecological monitoring). *Ecol Evol* 5:5295–5304. <https://doi.org/10.1002/ece3.1782>.
74. Lazić SE, Mellor JR, Ashby MC, Munafo MR. 2020. A Bayesian predictive approach for dealing with pseudoreplication. *Sci Rep* 10:2366. <https://doi.org/10.1038/s41598-020-59384-7>.
75. Sheik CS, Reese BK, Twing KI, Sylvan JB, Grim SL, Schrenk MO, Sogin ML, Colwell FS. 2018. Identification and removal of contaminant sequences from ribosomal gene databases: lessons from the Census of Deep Life. *Front Microbiol* 9:840. <https://doi.org/10.3389/fmicb.2018.00840>.
76. Vouga M, Baud D, Greub G. 2017. *Simkania negevensis*, an insight into the biology and clinical importance of a novel member of the *Chlamydiales* order. *Crit Rev Microbiol* 43:62–80. <https://doi.org/10.3109/1040841X.2016.1165650>.
77. Geerlings SY, Kostopoulos I, de Vos WM, Belzer C. 2018. *Akkermansia muciniphila* in the human gastrointestinal tract: when, where, and how? *Microorganisms* 6:75. <https://doi.org/10.3390/microorganisms6030075>.
78. Fullerton KM, Schrenk MO, Yücel M, Manini E, Basili M, Rogers TJ, Fattorini D, Di Carlo M, d’Errico G, Regoli F, Nakagawa M, Vetriani C, Smedile F, Ramírez C, Miller H, Morrison SM, Buongiorno J, Jessen GL, Steen AD, Martínez M, Maarten de Moor J, Barry PH, Giovannelli D, Lloyd KG. 2021. Effect of tectonic processes on biosphere-geosphere feedbacks across a convergent margin. *Nat Geosci* 14:301–306. <https://doi.org/10.1038/s41561-021-00725-0>.
79. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26:1641–1650. <https://doi.org/10.1093/molbev/msp077>.
80. McMurdie PJ, Holmes S. 2013. PhyloSeq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. <https://doi.org/10.1371/journal.pone.0061217>.
81. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlenn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2019. vegan: community ecology package. R package version 2.5-5. <https://cran.r-project.org/package=vegan>.
82. Chessel D, Dufour AB, Thioulouse J. 2004. The ade4 package - I: one-table methods. *R News* 4:5–10.
83. Marechal JC, Dewandel B, Subrahmanyam K. 2004. Use of hydraulic tests at different scales to characterize fracture network properties in the weathered-fractured layer of a hard rock aquifer. *Water Resour Res* 40:W11508. <https://doi.org/10.1029/2004WR003137>.
84. Theis CV. 1935. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Eos Trans AGU* 16:519–524. <https://doi.org/10.1029/TR016i002p00519>.
85. Hantush MS. 1961. Aquifer tests on partially penetrating wells. *J Hydraul Div* 87:171–195. <https://doi.org/10.1061/JYCEAJ.0000639>.
86. Phillips FM, Castro MC. 2013. Groundwater dating and residence-time measurements, p 361–400. *In* Treatise on geochemistry, 2nd ed. Elsevier Science, Amsterdam, Netherlands. <https://doi.org/10.1016/B978-0-08-095975-7.00513-1>.
87. Jones I, Lerner DN. 1995. Level-determined sampling in an uncased borehole. *J Hydrol* 171:291–317. [https://doi.org/10.1016/0022-1694\(95\)06015-B](https://doi.org/10.1016/0022-1694(95)06015-B).
88. Sabuda MC, Putman LI, Hoehler TM, Kubo MD, Brazelton WJ, Cardace D, Schrenk MO. 2021. Biogeochemical gradients in a serpentinization-influenced aquifer: implications for gas exchange between the subsurface and atmosphere. *J Geophys Res Biogeosci* 126:e2020JG006209. <https://doi.org/10.1029/2020JG006209>.