LLNL-JRNL-823248

Subsurface Planktonic Microbial Communities Reflect Regional-Scale Groundwater Hydraulic Connectivity

N. Merino, T. Jackson, J. H. Campbell, A. B. Kersting, J. Sackett, M. Zavarin, S. D. Hamilton-Brehm, D. P. Moser, J. C. Fisher, J. C. Bruckner

June 4, 2021

Science of the Total Environment

Disclaimer

Science of the Total Environment

Subsurface Microbial Communities as a Tool for Characterizing Regional-Scale Groundwater Flow

--Manuscript Draft--

Manuscript Number:	STOTEN-D-22-02506R2
Article Type:	Research Paper
Section/Category:	
Keywords:	fractured-rock aquifer, subsurface microbial ecology, microbial community assembly, regional-scale aquifer microbial community, deep biosphere
Corresponding Author:	Nancy Merino, Ph.D. LLNL: Lawrence Livermore National Laboratory UNITED STATES
First Author:	Nancy Merino, Ph.D.
Order of Authors:	Nancy Merino, Ph.D.
	Tracie R. Jackson
	James H. Campbell
	Annie B. Kersting
	Joshua Sackett
	Jenny C. Fisher
	James C. Bruckner
	Mavrik Zavarin
	Scott D. Hamilton-Brehm
	Duane P. Moser
Abstract:	Subsurface microbial community distribution patterns are influenced by biogeochemical and groundwater fluxes and may inform hydraulic connections along groundwater-flow paths. This study examined the regional-scale microbial community of the Death Valley Regional Flow System and evaluated whether subsurface communities can be used to identify groundwater-flow paths between recharge and discharge areas. Samples were collected from 36 sites in three groundwater basins: Pahute Mesa–Oasis Valley (PMOV), Ash Meadows (AM), and Alkali Flat–Furnace Creek Ranch (AFFCR). Microbial diversity within and between communities varied by location, and communities were separated into two overall groups that affiliated with the AM and PMOV/AFFCR basins. Network analysis revealed patterns between clusters of common microbes that represented groundwaters with similar geochemical conditions and largely corroborated hydraulic connections between recharge and discharge areas. Null model analyses identified deterministic and stochastic ecological processes contributing to microbial community assemblages. Most communities were more different than expected and governed by dispersal limitation, geochemical differences, or undominating processes. However, certain communities from sites located within or near the Nevada National Security Site were more similar than expected and dominated by homogeneous dispersal or selection. Overall, the (dis)similarities between the microbial communities of DVRFS recharge and discharge areas supported previously documented hydraulic connections between: (1) Spring Mountains and Ash Meadows; (2) Frenchman and Yucca Flat and Amargosa Desert; and (3) Amargosa Desert and Death Valley. However, only a portion of the flow path between Pahute Mesa and Oasis Valley could be supported by microbial community analyses, likely due to well-associated artifacts in samples from the two Oasis Valley sites. This study demonstrates the utility of combining microbial data with hydrologic, geologic, and water-chemistry information to comprehensively characterize groundwater systems, highlighting both strengths and limitations of this approach.
Response to Reviewers:	Reviewer #4: I appreciate the changes made by the authors. From my side, the paper

can now be accepted for publication in STOTEN.

We thank the reviewer for their comments!

Reviewer #5: The revised version of this manuscript has notably improved in clarity and readability, after having considered all points raised by me and the other 3 reviewers.

I still have some suggestions for 'cosmetic' amendments, as follows:

We thank the reviewer for their ‘cosmetic’ amendments and have made the changes.

- The concept of "microbial community as a tool for…" sounds awkward and it is only left in the title. I suggest to change the title in a statement title by reflecting what reported in the text (see e.g. L64, L117-121, L420-421, L430-431, L556-557, L578-579, L594-596 of the clean version). I would also avoid to repeat the same concept so many times along the text.

The authors discussed changing the title at great length, but in the end, all the authors agreed that the original title best represents the concept and focus of the paper.

L420-421 has been deleted.

L556-557 has been modified to (new line 558-560) “This study demonstrates that the regional-scale groundwater microbial community is a relevant data source; however, there were many limitations, as noted in Section 3.1, and future studies are needed. For example, the findings…”

- The graphical abstract is the less attractive figure among those presented. I would suggest to use the current figure 7 including some microbial data (e.g. the nine microbial clusters in figure 5?).

Graphical abstract has been updated. Figure 7 was not used because it only represents a portion of the DVRFS.

- The highlight #1 is likely out of the character limit

Highlight #1 modified to “The subsurface microbial community is consistent with known hydraulic connections.”

- The conclusive part was split in two sections (5. Future outlook, 6. Conclusion). I understand the Authors' reply, yet there are repetitions and redundancies between the new sections and in comparison to the previous text of the conclusion. Despite this is a subjective issue, I would suggest to shorten the conclusive part with a clearer focus on the outcomes of this study.

Section 5 and 6 were modified to reduce repetitions and redundancies, with the conclusive part having more focus on the outcomes of the study. (new lines 561-614)

Leaving aside the few minor issues, the manuscript is very interesting and fluently readable. It also provides many original data to the current knowledge in the field.

Thank you!

Lawrence Livermore National Laboratory

Nancy Merino, Ph.D.
Postdoctoral Scientist
Biosciences & Biotechnology Division
Physical and Life Sciences Directorate
Lawrence Livermore National Laboratory
7000 East Avenue, L-452
Livermore, CA 94550
925-424-7605 | merino4@llnl.gov

January 31, 2022

Dr. Warish Ahmed
Environmental Contaminant Mitigation & Biotechnologies (ECMB)
CSIRO Land & Water
Associate Editor, *Science of the Total Environment*

Dear Dr. Warish Ahmed,

Please find enclosed our manuscript entitled, “Subsurface Planktonic Microbial Communities as a Tool for Characterizing Regional-Scale Groundwater Flow”. We request you to consider this manuscript for publication in *Science of the Total Environment*. All authors are aware of and accept responsibility for this manuscript, which has not been previously published and is not under consideration by other journals.

Our manuscript evaluates whether planktonic groundwater microbial communities can be used in conjunction with hydrogeochemical analyses to better understand groundwater-flow paths and processes occurring in complex groundwater systems. As the global hydrologic cycle shifts with climate change, sustainable groundwater management will become more important; and new tools such as microbiome analyses may improve assessments of hydrologic connectivity and groundwater quality and quantity. However, to date, the groundwater microbiome is often an overlooked data source. In our study, we examined the microbial communities of a well-studied, regional-scale groundwater system known as the Death Valley Regional Flow System (DVRFS). To our knowledge, this is a unique and unprecedented study that describes a regional-scale groundwater microbiome from a system with advanced hydrologic models and a vast amount of historical hydrogeochemical data. We have thoroughly examined the microbiome using multiple methods, including community distribution patterns, network analysis, and community assembly ecological null models. We demonstrate consistency between the microbiome and several groundwater-flow paths identified by groundwater-flow conceptualizations, even providing additional evidence for a highly debated groundwater-flow path within the study area. We have written this manuscript to cross disciplines, including microbial ecology, hydrology, and biogeochemistry, as demonstrated by our cross-disciplinary author list.

Thank you for your consideration and we look forward to reviewers’ comments and your decision on this manuscript.

Thank you,

Nancy Merino

Subsurface Microbial Communities as a Tool for Characterizing Regional-Scale Groundwater Flow

Nancy Merino[1*], Tracie R. Jackson[2], James H. Campbell[3], Annie B. Kersting[1], Joshua Sackett[4,5,6], Jenny C. Fisher[4,7], James C. Bruckner[4], Mavrik Zavarin[1], Scott D. Hamilton-Brehm[4,8], Duane P. Moser[4,5*]

[1] Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, 94550, United States
[2] Nevada Water Science Center, U.S. Geological Survey, Boulder City, NV, 89005, United States
[3] Department of Natural Sciences, Northwest Missouri State University, Maryville, MO, 64468, United States
[4] Division of Earth and Ecosystems Sciences, Desert Research Institute, Las Vegas, NV, 89119, United States
[5] Division of Hydrologic Sciences, Desert Research Institute, Las Vegas, NV, 89119, United States
[6] School of Life Sciences, University of Nevada, Las Vegas, NV, 89154, United States
[7] Biology Department, Indiana University Northwest, Gary IN, 46408, United States (present address)
[8] Department of Microbiology, Southern Illinois University Carbondale, Carbondale, IL, 62901, United States (present address)

***Correspondence:**
Nancy Merino
Lawrence Livermore National Lab
7000 East Ave, B361/R1940
Livermore, CA 94550
925-424-7605
merino4@llnl.gov

Duane Moser
Division of Hydrologic Sciences
Desert Research Institute
Las Vegas, NV, 89119
702-862-5534
Duane.Moser@dri.edu

Keywords (max 6): fractured-rock aquifer, subsurface microbial ecology, microbial community assembly, regional-scale aquifer microbial community, deep biosphere

Abbreviations not standard to the field: Alkali Flat–Furnace Creek Ranch (AFFCR), Ash Meadows (AM), Death Valley Regional Flow System (DVRFS), Nevada National Security Site (NNSS), Pahute Mesa–Oasis Valley (PMOV), Underground Test Area (UGTA)

Reviewers/Editor comments:

Reviewer #2: This study describes the analysis of a large SSU amplicon dataset to characterize microbial communities inhabiting several aquifers of the Death Valley Regional Flow System. The groundwater samples were collected in the recharge or discharge areas of these aquifers and the authors sought to determine whether the analysis of the structure of the microbial communities could provide useful information on groundwater circulations in the DVRFS.
This paper provides an interesting contribution to the understanding of the complex interplays between groundwater circulations and the distribution of microorganisms in the subsurface although the numerous « sampling limitations » as mentioned by the authors limit the conclusions that can be drawn from this study.

The authors thank the reviewer for appreciating the contribution and considering the sampling limitations while reviewing the manuscript.

I also wonder about a critical step of the bioinformatics analysis of the amplicons. It is stated in the material and methods that « To reduce bias from sequence batch effects, the amplicon sequence variants (ASVs) were clustered into OTUs in QIIME 2 ». I don't understand why the authors decided to cluster the ASVs into OTUs. The combination of the two types of methods, one after the other, seems very unusual and even pointless to me. Indeed, as explained in the paper by Callahan et al "ASV methods are explicitly intended to replace OTUs as the atomic unit of analysis" (https://urldefense.us/v3/__http://dx.doi.org/10.1038/ismej.2017.119__;!!G2kpM7uM-TzIFchu!gURHAeANasbWyg-pWN0gVPAKGcftVKzwFNv0JtGvGuR9HzP-H7VS5XJEjJkTI4-U$). In addition, I think that clustering ASVs into OTUs could increase the observed similarity between microbial communities since ASV methods allow delineating groups of amplicons with a finer resolution than that of OTUs with a 3 % identity threshold. At last, according to the authors, the combinations of ASVs and OTUs might reduce the bias associated with the "sequence batch". However, the authors also explain that the sequence batch effect impacted ordination results so their "double approach" does not seem to be very effective to prevent bias from the sequence batch effect. Could the authors better explain the reasons for clustering ASV into OTUs ? How did they control that their approach really allowed to reduce the bias associated with the two "sequence batches" without "inflating" the observed similarity between microbial communities ?

As the reviewer notes, ASVs provide a finer resolution than OTUs and would have been the preferred choice for conducting microbial community analysis. However, because each sequencing batch may have inherently different error rates when processing through DADA2, each batch was processed through DADA2 separately and unique ASVs per batch were obtained. DADA2 uses a parametric error model and according to DADA2 developers “every amplicon dataset has a different set of error rates” (https://benjjneb.github.io/dada2/tutorial.html). After each batch had been processed through DADA2, we tried the advice provided in this link (https://github.com/benjjneb/dada2/issues/716) to merge the ASVs from the two sequencing batches, but we were unsuccessful. For a separate project, we have been submitting the same DNA samples to the company in different sequencing batches and were also unsuccessful to merge ASVs. Therefore, we conducted a literature search to understand the reliability of using OTUs. We found that several studies (e.g., references #1–6 below) observed similar overall results and conclusions between ASVs and OTUs, although alpha diversity may be overestimated with OTUs. Because this study’s objective was to evaluate the microbial community on a regional-scale and provide a broad overview of the regional-scale microbiome, we decided to use the traditional approach of clustering OTUs at the 97% identity level.

We thank the reviewer for this comment because these nuances were not clearly represented behind the phrase “To reduce bias from sequence batch effects”. Thus, the text has been clarified (new lines 214–222, p. 8, last paragraph): “Because there were two different sequencing batches, samples were separated by batch and processed through DADA2 separately to account for potential error rate model differences. This resulted in amplicon sequence variants (ASVs) unique to each batch, which may have been caused by technical or biological differences (e.g., sample preparation, human processing, PCR amplification, sequencing errors, and multiple 16S rRNAs per cell). However, many ASVs from both batches were closely-related, as determined by inspection of a phylogenetic tree and sequence alignment, and were subsequently clustered into OTUs in QIIME 2 (47) at the 97% identity level with the function vsearch cluster-features-de-novo. OTU clustering at the 97% identity level can provide comparable overall results as ASVs on a broad-scale (48–53).”

1. Glassman, S.I. and Martiny, J.B.H. (2018) Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. *mSphere* **3**: e00148-18.
2. Joos, L., Beirinckx, S., Haegeman, A., Debode, J., Vandecasteele, B., Baeyen, S., et al. (2020) Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. *BMC Genomics* **21**: 733.
3. Louca, S., Mazel, F., Doebeli, M., and Parfrey, L.W. (2019) A census-based estimate of Earth’s bacterial and archaeal diversity. *PLOS Biology* **17**: e3000106.
4. Moossavi, S., Atakora, F., Fehr, K., and Khafipour, E. (2020) Biological observations in microbiota analysis are robust to the choice of 16S rRNA gene sequencing processing algorithm: case study on human milk microbiota. *BMC Microbiology* **20**: 290.
5. Nearing, J.T., Douglas, G.M., Comeau, A.M., and Langille, M.G.I. (2018) Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. *PeerJ* **6**: e5364.
6. Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A.H., Nieuwdorp, M., and Levin, E. (2020) Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLOS ONE* **15**: e0227434.

I am also concerned by the status of the Oasis Valley samples. At the end of section 4.4, the authors extensively explain that the OV samples are potentially contaminated and not representative of the groundwater microbiome. However, in sections 5 and 6 the authors discuss the connections between Oasis valley and other sites based on the microbial data. If the authors do have any doubt about the OV samples, I would advise them to remove these samples from the study. As it is now, the paper is confusing on the status of the OV samples and this weakens its conclusions. Morever, if the OV samples are contaminated by soil-associated bacteria, the existence of a cluster of common OTU shared by the OV and SM samples (Fig. 5) raises the issue of the contamination of the SM samples. I believe that's what the authors meant with the phrase "The co-occurrence of 'common OTUs' […] suggests potential contamination by surface level microbes" (L383-385). However, I think, this issue should be more clearly stated in the manuscript.

The OV samples were used in this manuscript to demonstrate the analytical process in assessing a ‘bad’ groundwater sample through the lens of combining geochemical and microbial data together. PCA of the geochemical data identified that only OV1 was clearly distinct from all other samples while beta-diversity ordinations demonstrated OV1 and OV2 were both microbially different. Further examination showed that there were many soil-associated and aerobic microbes. This has been clarified in the text on new line numbers 532-543 (p. 20, paragraph 1): “The lack of microbial similarity between Pahute Mesa and Oasis Valley does not reflect regional groundwater-flow paths in the PMOV basin (**Figure 1**; 37, 38). However, the two Oasis Valley samples may not be representative of regional groundwater in this location as supported by the presence of many soil-associated and aerobic microbes. Moreover, the geochemistry of OV1 is distinct from other locations (**Figure 2**), with relatively high TOC concentrations (37 mg-C/L). The combination of microbial (**Figure 4**; taxonomic composition) and geochemical (**Figure 2**) data suggest that the two Oasis Valley samples are likely more impacted by land-surface-associated factors, owing to their relatively shallow depths and siting within the alluvium of the Amargosa River floodplain. Hence, these sites are probably disconnected hydrologically from the regional aquifer and conceptually can be regarded as hyporheic. This observation demonstrates the utility in supplementing geologic and hydrologic datasets with microbial community data for a comprehensive evaluation of groundwater samples.”

However, as the reviewer (and other reviewers) indicates, the inclusion of OV samples in the network is confusing. Therefore, we re-analyzed the network without OV samples because the network is supposed to represent ‘common OTUs’ and the OV samples are distinct, as shown through geochemical and beta-diversity analyses. The following new line numbers have been updated: 263-264 (p. 10, last paragraph), 388-398 (p. 15, first paragraph), 438-441 (p. 16, last paragraph), 456-458 (p. 17, paragraph 2), 464-466 (p. 17, last paragraph), 499-501 (p. 19, paragraph 1).

At last, I am puzzled by the rarefaction curves showed in Fig S1. All the curves grow very rapidly, in a very similar way and most of them plateau for about the same sample size whatever the number of OTUs found in the samples. It seems really unusual to me and I would like the authors to provide new rarefaction curves made with other tools than ranacapa (this R package seems to be poorly documented).

Fig S1 and the text (new line 226, p. 9 paragraph 1) have been updated with rarefaction curves generated by iNEXT (Hsieh et al., 2016). The result is the same as ranacapa. The iNEXT publication has been cited 1897 times as of April 1, 2022.

Hsieh, T. C., Ma, K. H., and Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods Ecol. Evol.* 7, 1451–1456. doi:10.1111/2041-210X.12613.

Beside that, I have other comments that, I hope, will help improve the manuscript.

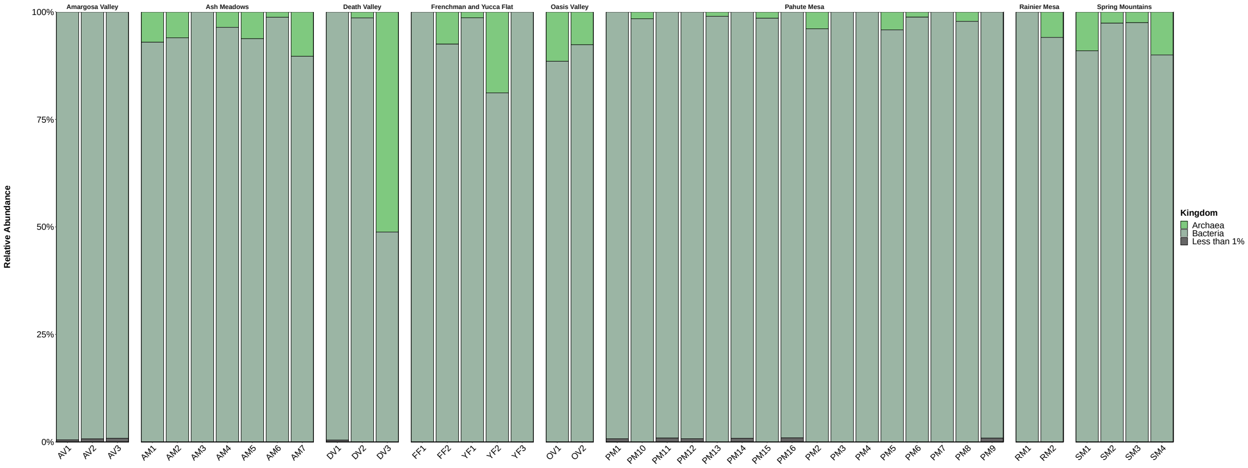
L210 : (and elsewhere in the text) : « To reduce bias from sequence batch effects […] ». To my knowledge, the terms "sequence batch" or "sequencing batch" are quite unusual. The authors should replace the word "batch" with "run" in the manuscript.

We intend to use the term “Sequence batch” in this manuscript to align with the definition for “batch effect”: “technical variation or non-biological differences between measurements of different groups of samples” (Nyamundanda 2017). Although there are other types of “batches” in this study (e.g., sampling method), the authors specifically wanted to understand the impact of the different “sequencing runs”. Because the two “batches” (definition here: “a quantity or consignment of goods [submitted] at one time”) of samples were submitted a couple years apart to a sequencing company for each “sequencing run”, there could be other differences that arise from changes (human or logistic) in the company’s protocols. We have discussed the terminology and decided to continue using “sequence batch”.

Nyamundanda, G., Poudel, P., Patil, Y. *et al.* A Novel Statistical Method to Diagnose, Quantify and Correct Batch Effects in Genomic Studies. *Sci Rep* **7,** 10849 (2017). https://doi.org/10.1038/s41598-017-11110-6

L317 : « […] the highest abundances (51%) were observed in sample DV3 (Inyo-BLM 1) ». The huge abundance of Archaea in DV3 is quite exceptional. Are the authors really sure that their results for DV3 do not suffer from any technical (PCR ?) bias ?

This manuscript did not examine the bias associated with different primers and used the typical universal primers for bacterial/archaeal communities for every sample. Therefore, any PCR bias towards Archaea would be observed in other samples as well, but this did not occur, as shown in this barplot:



On line 52 of the Supplementary Methods, we specified the exact sequences of the primers we used (515F/806R). Interestingly, a recent study found that this primer may underestimate archaeal populations (Pausan et al. 2019).

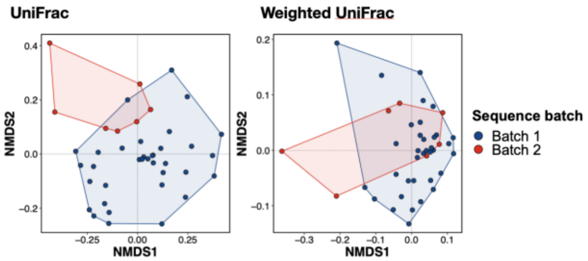
Pausan, M.R., Csorba, C., Singer, G., Till, H., Schöpf, V., Santigli, E., et al. (2019) Exploring the Archaeome: Detection of Archaeal Signatures in the Human Body. *Frontiers in Microbiology* **10**.

L353-354 : « However, this result may be attributed to a noticeable separation of samples by sequence batch for CLR (Figure S7E) and PhILR PCA (Figure S7H) ». Could the authors explain why the sequence batch has en effect for CLR PCA and PhiLR PCA but not DEICODE RPCA ? Does the sequence batch could also bias the NMDS ordinations in figure 4 ?

We can only speculate on why there is clear separation of sequence batch for CLR and PhILR PCA but not for DEICODE RPCA. We have clarified the text on new lines 373-375 (p. 14, last paragraph): “Sequence batch did not seem to impact DEICODE RPCA (**Figure S7B**), likely because the DEICODE algorithm utilizes the geometric mean of log-transformed nonzero data and conducts matrix completion (61).”

There is some sequence batch bias for UniFrac (see figure below), but not for weighted UniFrac. For the manuscript, we decided to focus on Weighted UniFrac and provided Table 1 to demonstrate the impact of sequence batch bias on location-type and geochemical-type variables, as described in the manuscript. For UniFrac, we did PERMANOVA of location-type and geochemical-type variables, and similar to

Weighted UniFrac, we found that sequence batch did not significantly (P = 0.053, R2 = 0.026) influence location-type variables; however, sequence batch did significantly (P = 0.006; R2 = 0.035) influence geochemical-type variables. In response to another reviewer, we have also included CCA ordination for further description of geochemical-type variables and manuscript updated (i.e., Fig 4, Section 3.4, Discussion section).



L418 : «[...] the overgrowth of microbial biofilms at biogeochemical hotspots (39, 86). » In my opinion, the reference #86 (Bochet et al.) is inappropriate. In their paper, Bochet et al . describe a large accumulation of microbial mat at some depths and an apparent clogging of the studied borehole. However, the clogging did not prevent the upstream flow of water in the borehole and the authors did not provide any evidence that the formation of microbial mat at biogechemical hotspots could reduced groundwater flow in the fracture network of the aquifer.

Replaced reference #86 with the following (new line #: 427, p. 16 paragraph 2):

1. Baveye, P., Vandevivere, P., Hoyle, B.L., DeLeo, P.C., and de Lozada, D.S. (1998) Environmental Impact and Mechanisms of the Biological Clogging of Saturated Soils and Aquifer Materials. *Critical Reviews in Environmental Science and Technology* **28**: 123–191.
2. Lumban Gaol, C., Ganzer, L., Mukherjee, S., and Alkan, H. (2021) Investigation of clogging in porous media induced by microorganisms using a microfluidic application. *Environmental Science: Water Research & Technology* **7**: 441–454.

L420 : « […] individual lineages may exhibit extraordinary genomic stability, possibly on the order of millions of years, particularly in the very deep subsurface. ». I think this sentence and the associated citation (#87) are irrelevant here. The reviewed paper by Merino et al. focuses on processes that control microbial community diversity. On the contrary, the paper by Becraft et al. considers genome evolutionary processes in populations of a single bacterial species over much longer time scale.

This sentence has been removed.

L431-433 : « In the AM basin, the microbial community is similar between the recharge area (Spring Mountains) and Ash Meadows discharge area (Figure 4). This result was corroborated with network analysis (Figure 5) [...] ». Looking at the network in fig 5 and S8, the AM and SM-OV clusters seems to be largely disconnected. There seems to be only 3 common OTUs connecting the two clusters. Furthermore the SM-OV cluster looks like quite artificial as I explain above. I think the authors should not use the network as another evidence of the similarity of the AM and SM samples.

Since the new network figure 5 and S8 without OV maintains the AM and SM cluster connectivity, this sentence has been clarified with "(e.g., 'common OTUs' in cluster SM)" (new line 441, p. 16 last paragraph).

L506 : « Inyo-BLM 1 (DV3) and Nevares Deep Well 2 (DV1) have similar geochemical conditions, including temperature [...] ». Further down the article (L516-518), it states «Community differences may be impacted by geochemical factors, such as higher pressures and temperatures at depth in Inyo-BLM 1 (DV3) compared to the shallow Nevares wells (DV1/DV2) ». It has to be one or the other. It cannot be both. Could the authors explain and resolve this contradiction ?

This has been clarified in new line 524-525 (p. 19, last paragraph): "Community differences may be impacted by **unmeasured** geochemical factors, such as higher pressures ~~and temperatures~~ at depth in Inyo-BLM 1 (DV3) compared to the shallow Nevares wells (DV1/DV2)."

L510 : « […] with more similarity between Nevares wells (DV1/DV2) compared to Inyo-BLM1 (DV3).» Taxonomic profils the relative abundance of phyla (Fig. S5) for samples DV1, DV2 and DV3 are very different. There seems to be as much difference between DV1 and DV3 and DV2 and DV3 than between DV1 and DV2. Although, beta diversity measures are based on OTU and not on taxa composition, it's a bit hard to believe that DV1 and DV2 really are more similar to each other than to DV3. Could the authors provide more details on the level of similarity between DV1 and DV2 compared to DV3 ?

This has been clarified on new line 521 (p. 19, last paragraph): "In contrast, 256 OTUs were present in both Nevares communities (DV1/DV2)."

Figure 5 and S8 : Changes in edge thickness and darkness are not immediately visible and should be made more obvious

The edge thickness and darkness were made more obvious.

Reviewer #3: This study examined the geochemical and microbial community data at regional-scale (Death Valley Regional Flow System, DVRFS) and evaluated whether planktonic microbial communities can be used to identify groundwater-flow paths between recharge and discharge areas.

Beta diversity using NMDS for weighted UniFrac distances seemed to be relatively distinguished by location. Geochemical data were separated by rock type (carbonate vs volcanic).
Network analysis revealed some patterns of hydraulic connections between 1) Spring Mountains and Ash Meadows, 2) near the NNSS sites (Pahute Mesa, Rainier Mesa, and Frenchman and Yucca Flat) and Amargosa Valley, and 3) Amargosa Valley and Death Valley. In addition, null model analyses identified ecological processes; 1) dispersal limitation and variable selection for Oasis Valley, Death Valley and Ash Meadows/Spring Mountains, and 2) homogeneous dispersal or selection for NNSS locations and Amargosa Valley.

Although the groundwater flow in some areas such as the upper area Pahute Mesa and Oasis Valley site is limited and the community similarity is low, NMDS ordination, network analysis, and ecological null models demonstrated that Death Valley microbial communities are more related to Amargosa Valley and NNSS locations, rather than to Ash Meadows and Spring Mountains.

Taken together, the regional groundwater-flow paths show that most of the groundwater moves westward through the Amargosa Valley and discharges at Furnace Creek in Death Valley. The microbial community results associated with the groundwater flow path to Death Valley appear to support compartmentalized groundwater systems. The Death Valley microbial community better supports hydraulic connections proposed in the basin-fill conceptualization, rather than the deep-carbonate flow conceptualization.

This research provides comprehensive results and interpretation of geochemical and microbial community data of wide regional scale site (DVRFS). It is considerable that authors conducted various statistical and analytical methods to validate the results and reduce the sample biases along different factors like sequencing, sampling methods and timepoints. And the interpretation of community data associated with network analysis of clusters and ecological processes of each location to regional flow-path system looks reasonable that it gives implications for field scale study, to understand how communities change along the selective pressure of certain environments. This study also provides important information for strategy to manage groundwater quality.

The authors thank the reviewer for their comments/suggestions and for appreciating the various statistical and analytical methods used in this study.

I made some comments or suggestions below.

1. Graphical abstract - better to draw a more complete picture with CSM.

The authors do not understand this comment and have contacted the editor for further advice. Please clarify.

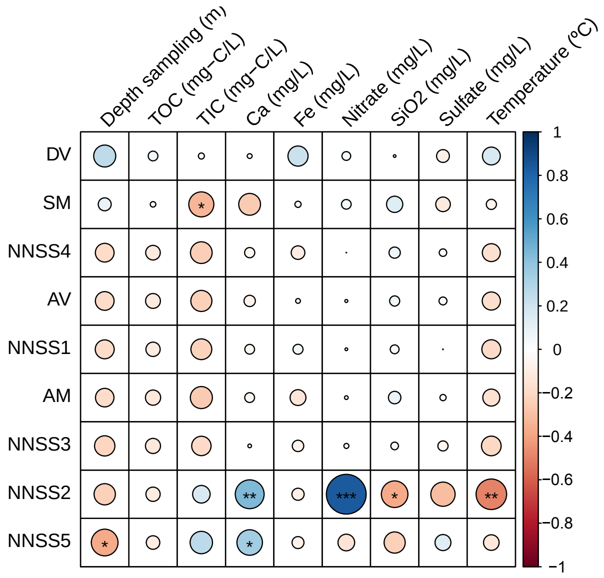
2. Have you considered O2 level (or other redox sensitive parameters) in groundwater as a geochemical parameter? I saw many groundwater chemistry data in the supplementary tables.

Yes, we considered O2 and dissolved oxygen concentrations, which are both in Table S2. However, it was decided that the dissolved oxygen concentrations may be unreliable and the O2 concentrations were only measured for a few sites.

3. For network analysis, common OTUs were used, and characteristics were classified by location. It would be easy for readers to understand the rationale or criteria for using common OTUs. How about network analysis considering groundwater geochemical factors or highly correlated factors such as depth and rock type?

We previously evaluated the network for association of each cluster with quantitative groundwater geochemical measurements using eigengenes, such as used by Lemonnier et al. (2020). However, there was not significant association ($P < 0.05$) except for three out of nine clusters (see figure below; Note: the network was updated to reflect comments from other reviewers and this reviewer's comment #6). In addition, because location was a strong variable for beta-diversity ordination, we decided to focus on location in the network, in which the clusters were clearly separated by location (Figure S8).

Lemonnier, C., Perennou, M., Eveillard, D., Fernandez-Guerra, A., Leynaert, A., Marié, L., et al. (2020) Linking Spatial and Temporal Dynamic of Bacterioplankton Communities With Ecological Strategies Across a Coastal Frontal Area. *Frontiers in Marine Science* **7**.



4. Since it is a site related to past nuclear tests, information related to this specificity is not shown in this manuscript. Considering the flow paths of groundwater, there might be a location affected by the tests. Please check the related footprint on the DV site?

The impact of past nuclear tests on the groundwater microbiome was evaluated in the preparation for this manuscript. However, we do not report our analyses because of the complexity in demonstrating the impact of nuclear tests on the microbial community. There is a lot of nuance associated with nuclear tests, and measurements that would help implicate an impact on the microbiome were not recorded (or not publicly available) for every site, such as the area and size of explosion and release of radionuclides. In addition, we can only speculate on this topic because a comparative "pristine sample" is difficult to obtain (e.g., groundwater in the same recharge or discharge zone but not impacted by a nuclear test).

However, as described in the manuscript, tritium was not a significant variable describing alpha- and beta-diversity (Figure 3, S6, Table 1).

5. Why not use RDA or dbRDA? (or other constrained ordination). Rather than comparing PCA and NMDS separately, through analysis such as RDA or dbRDA, geochemical variables and community data can be compared together for better overall interpretation/assessment.

We especially thank the reviewer for this comment. We previously evaluated constrained ordinations but did not observe any good relationship between geochemical variables and community data. However, upon re-visiting constrained ordinations, we found very nice relationships and have decided to include CCA in the manuscript! Figure 4 has been updated with Figure 4C, along with new lines 240-241 (p. 9 paragraph 3), 353-370 (p. 13, last paragraph), 446-448 (p. 17, paragraph 1), 490-494 (p. 18, last paragraph), 517 (p. 19, last paragraph), 547 (p. 20, last paragraph). Our scripts have also been updated.

6. I think that samples from Oasis Valley (OV) make the results more biased to interpret. As the authors discussed at Section 5.5, OV samples reveals lack of similarity compared to the recharge area Pahute Mesa (PM). It is valid that these results can denote that PMOV basin flow path is not related to microbial community. However, considering the distinct geochemistry of OV1 (Figure 2) and beta diversity results with different methods (Figure S7), OV seems to be affected by contamination of groundwater or the bias by sequence batch. Also, at network analysis (Figure 5), the cluster with OV and Spring Mountain (SM) looks little bit not understandable because they are located far and not related to same recharge-discharge groups. Therefore, it might be clearer if the OV data is discarded at network analysis.

We re-analyzed the network without OV samples. The following new line numbers have been updated: 263-264 (p. 10, last paragraph), 388-398 (p. 15, first paragraph), 438-441 (p. 16, last paragraph), 456-458 (p. 17, paragraph 2), 464-466 (p. 17, last paragraph), 499-501 (p. 19, paragraph 1).

7. For Section 5.3, the relationship between Death Valley (DV) microbial community and NNSS locations looks valid. However, turnover between DV and Amargosa Valley (AV) seems to be unclear. The authors mentioned the common OTUs in DV and AV at network analysis (line 493-494), but considering ratio reads (Figure S8), the co-occurrence looks very low between DV and AV (I'm not sure why top-left section of AV cluster is dominated with YF and DV). It seems DV microbial community is more correlated with NNSS2 cluster or location Frenchman & Yucca Flat.

With the re-analyzed network, cluster DV became more closely associated with NNSS1 and NNSS2. We have revised the text on new lines 499-501 (p. 19, first paragraph): "Network analysis indicates that cluster DV is closely associated with the NNSS clusters, particularly NNSS1 and NNSS2, and 'common OTUs' observed in Death Valley are also found in NNSS clusters and cluster AV (**Figure 5**)."

Reviewer #4: The submission by Merino et al tackles a scientifically interesting topic, i.e. can the hydrological cycle in the subsurface be assessed by means of microbial community information. In fact, the concept of the overall study as well as the methods applied are solid. The weak points of the study are openly mentioned and discussed by the authors, i.e. (1) there is not many sampling points for the huge area studied, (2) there hardly any replicate samples in time for the individual samplinf spots (no chance to evaluate seasonality as apossible driver of physical-chemical conditions and microbial community composition, (3) the different subsurface habitats have been sampled using different sampling methods (pumping vs. bailer, water from the aquifer vs. water from a well) and collecting very different sample volumes, (4) the samples have been collected over a rather long time period of several years, and finally (5) collection of samples for microbiological analysis
was not necessarily accompanied by water-chemistry analysis (which is very surprising). However, these limitations, can not be annulled, need to be accepted, and are sufficiently discussed in the paper.

We thank the reviewer for their suggestions and acknowledging the limitations of the study.

There is a few minor issues, I ask the authors to consider when revising the paper.
1. Is there really 'endemic' microbial communities. See Introduction (P4L83) and Discussion (P21L566). I would be very careful in using this terminology. The term 'endemic' is used in macroecology for individual species that have a very restriced range of occurence. It cannot be used for an entire community unless all community members are not found elsewhere, which is hard to prove with microbes. In microbial ecology, endemic microbial species (taxonomic units) have been reported repeatedly before these sequences have been detected also elsewhere with some delay.

The term ‘endemic’ has been removed throughout the manuscript.

2. P4L91 How can travel times be deliniated from water chemistry? do you mean tritium signatures, or concentrations of novel gases? Please provide a reference here.

References added (new line 93, p. 4, paragraph 2).
9. W. W. Woessner, *Groundwater-Surface Water Exchange* (GW Project, 2020).
10. P. Cook, *Introduction to Isotopes and Environmental Tracers as Indicators of Groundwater Flow* (GW Project, 2020).
11. J. F. Devlin, *Groundwater Velocity* (Groundwater Project, 2020).

3. P11L284 A study that emphasizes the dynamics of microbial community composition is Pilloni et al. 2019; Microorganisms 7, 46

The reference has been added and described on new lines 290-294 (p. 11, paragraph 3).

4. P15L388 Where does the evidence come from that a contamination during sampling might have caused the difference in community composition with Oasis Valley?

This sentence has been removed because we updated the network in response to other reviewers’ comments. We describe how Oasis Valley is a distinct sample from the others in Section 4.5 (previously labeled as 5.5), in which we describe Oasis Valley microbial communities as likely not representative of a groundwater microbiome sample because of the presence of soil-associated and aerobic microbes and the separation of Oasis Valley sample(s) geochemically (Figure 2) and microbially (Figure 4).

5. The classical way groundwater ecosystems are evaluated microbiologically is based on water samples ignoring the majority of microbial biomass and biodiversity living attached to the rock or sediment matrix. As mentioned in the paper, there is an interaction between these two 'worlds', the porewater and the sediment surface world, both characterized by different selective forces with respect to community assembly. A paper that looked into the mechanisms of community assembly on sediment surface is Fillinger et al. 2018, Environ. Microbiol. Doi:10.1111/1462-2920.14463. It might be of interest for the authors.

We thank the reviewer for this paper! It is definitely of interest and very cool to see the utilization of ecological null models for evaluating assembly processes on sediment surfaces. I am intrigued that the assembly can happen reproducibly, and that time was not an important factor. As noted in the conclusion of Fillinger et al. 2018, it will be interesting to do a similar study but with -omics to gain a more comprehensive picture. I think it will also be of interest to investigate for other surfaces like the phycosphere, plastisphere, marine snow, etc.

6. P21L568 It cannot be excluded that individual microbial species introduced to the subsurface with rain recharging the aquifer get lost with distance to the recharge area because their living conditions are not met in the subsurface environment. Similarly, rare but autochthonous species (OTUs) that are not seen in the groundwater samples from the recharge zones with the method applied (16S rRNA amplicon sequencing) gain in importance along the subterranean flow path. In this case the microbial community composition might fail as 'tracer' of hydrological connectivity. This point might be discussed.

This point is a great discussion topic. A couple sentences were added, new lines 568-572 (p. 21, paragraph 2): “Moreover, groundwater microbial community samples collected along a groundwater-flow path and at various depths can identify potential candidates for use as a microbial ‘tracer’ by providing insight into the persistence of specific microbial species. This spatial data also will help to confirm that ‘common OTUs’ in recharge and discharge zones can be used for helping characterize groundwater flow.”

My vote for this submission is minor revision.

Thank you!

Reviewer #5: This study provides original field data on the phylogenetic composition of the groundwater microbial community within the Death Valley Regional Flow System. Samples were collected from groundwater recharge and discharge areas at a regional scale. The 16S rRNA gene profiling (amplicon sequencing) showed that community similarity was higher between closer hydrologically-connected areas. Along with water chemistry, microbial community profiles provided supportive evidences to inform on the hydraulic connections and water origin.

The reported findings are in line with the current knowledge and will provide a good contribution to the specialized literature. However, in my opinion, several major issues should be addressed to improve the relevance and attractiveness of this study for the international scientific audience. Please, find below the major and minor points that attracted my attention.

We thank the reviewer for their comments and suggestions to improve the study for an international scientific audience.

Major points

1- The study aims should be more explicitly detailed at the end of the introduction (e.g., The aims of this study were to (i)…, (ii)…, (iii)…). What reported at L123-124 is too generic. The research hypotheses should be also much clearer from the beginning. There are several statements across the text recalling different study aims (eg., L277-280, L346-348, L392, L411, L423, L480, L527, L549, L558). Overall, it is difficult to focus on the major outcomes.

The study aims have been made more explicit on new lines 120-132 (p. 5, paragraph 2): "This study assesses the potential of using subsurface microbial community patterns as a supplemental dataset to characterize regional-scale groundwater flow. The Death Valley Regional Flow System (DVRFS) microbial community was used as a test case because, over the past few decades, hundreds of boreholes and wells have been drilled in the DVRFS to monitor water levels and water quality. Thus, the DVRFS provides a unique opportunity to identify novel microbial community patterns on a regional-scale and to determine whether microbial community patterns are consistent with regional-scale groundwater flow-paths (25). The aims of this study were to (1) evaluate the groundwater bacterial and archaeal community distribution patterns with alpha- and beta-diversity analyses and network analysis, (2) identify the assembly processes dominating the community patterns using ecological null models, and (3) use the microbial community patterns and assembly processes to assess any (dis)similarities along groundwater-flow paths. To achieve these aims, we collected samples from recharge and discharge areas and hypothesized that microbial (dis)similarities, co-occurrences, and assembly patterns may infer and support the most recent groundwater conceptualization (25)."

2- My impression is that too many data and interpretations are provided without a clear link to the study aims. I fully understand the sampling and experimental effort necessary to obtain such a solid dataset, but only the data supporting the research hypothesis should be presented and interpreted. At L145, it is stated that 5 groundwater flow paths are evaluated, while it was not possible using the available data (see e.g., L280-282, section 5.5, L562-565). This is confusing for the reader and emphasizes (minor) results that only have a local relevance.

The amount of data and interpretation is appropriate for this study. We have clarified the study aims in response to point #1 and have linked each analyses with the study aims.

For L280-282, this sentence refers to temporal changes, which we clearly state in the sentence that our study does not focus on the "temporal dynamics of subsurface microbial communities". This sentence has also been recommended to include by a couple reviewers.

For L145 (and section 5.5, L562-565), the OV samples in this manuscript were used to demonstrate the analytical process in assessing a 'bad' groundwater sample through the lens of combining geochemical and microbial data together. PCA of the geochemical data identified that only OV1 was clearly distinct from all other samples while beta-diversity ordinations demonstrated OV1 and OV2 were both microbially different. Further examination showed that there were many soil-associated and aerobic microbes. This has been clarified in the text on new line numbers 532-543. In addition, in response to other reviewers, we have re-analyzed the network without OV samples because the network is supposed to represent 'common OTUs' and the OV samples are distinct, as shown through geochemical and beta-diversity analyses. The following new line numbers have been updated: 263-264 (p. 10, last paragraph), 388-398 (p. 15, first paragraph), 438-441 (p. 16, last paragraph), 456-458 (p. 17, paragraph 2), 464-466 (p. 17, last paragraph), 499-501 (p. 19, paragraph 1).

3- A plethora of multivariate statistical tests are presented without a clear rationale (see L226-231; L350-364). Finally, the info provided by NMDS, PCAs, PERMANOVA, ANOSIM, etc seem redundant. This is again confusing. I understood that hydraulic connectivity reflected on either the water chemistry or microbial profiles, but this point goes lost among other results of minor relevance and local interest at the end.

These ordinations are important for microbial ecologists, and the microbial ecology literature continues to have very strong debates about the different approaches for evaluating microbial communities. Therefore, it is necessary to include the ordinations and statistical tests to convince microbial ecologists. The text has been clarified on new lines 237 (p. 9, paragraph 3) and 368-370 (p. 14, paragraph 2) for rationale.

4- Samples were collected between 2008 and 2014 (L163). How long the filters were stored at -80°C? When was the DNA extraction performed? The sequences were analyzed using QIIME2 released in 2018 (L210). Please, clarify the timeline of the analytical activities.

The timeline has been clarified in the Supplementary info:
SI New line 46: "DNA was extracted within a few weeks to months after sample collection,"
SI New line 90: "In 2020, raw sequences were first processed"

5- The result section should be dried out from interpretations and references, which has to be moved to the discussion. There are also many methodological details to be presented in the methods section (see e.g., L274-286, L291, L346-364, L392-394, 406-407).

The following example lines provided by the reviewer are needed to orient or convince the reader and have not been moved:

- L274-286 represents section 4.1, where we clearly outline our study limitations to orient the reader. This section cannot be moved to the methods or discussion section because it represents a necessary introduction to our dataset and must be positioned at the beginning of the Results section.
- L291 is necessary to convince readers, especially those that have worked on the DVRFS, that our geochemical dataset is representative of previous reports.

- L346-349 represents the overall conclusion of the NMDS ordinations.
- L350-364 explains the other ordinations used in the dataset in comparison against the NMDS ordinations. This is a technical interpretation based on the figures and necessary in this particular location of the manuscript for convincing microbial ecologists.
- L 392-394 this is necessary to orient the reader about ecological null models.

L 406-407 has been moved to the discussion section 4.5 (previously labeled as 5.5)

6- The conclusion section is verbose. This part is intended to resume the take-home message(s) of this study into 4-5 sentences. Please, recast accordingly.

The conclusion section has been split into two sections. New section 5 “Future Outlook” represents the second paragraph of the original conclusion section. The first paragraph of the conclusion section remains the same.

Minor points
7- I would not use the terms "planktonic" and "subsurface planktonic", but replace those with "groundwater". The term "microbiome" should be also replaced with "microbial community".
“Planktonic” has been replaced with “groundwater” or “subsurface” where appropriate.
“microbiome” has been replaced with “microbial community”

8- Too many acronyms are found in the abstract, highlights, intro and across the text in general. Please, limit the use of unconventional acronyms.
Acronyms have been reduced in the abstract and highlights. Two acronyms have been removed in the manuscript: DOE, USGS. All other acronyms are necessary in the context of the study and the number of times used throughout the manuscript. Acronyms used in this manuscript include those that describe (1) study site, (2) methods, tools, or software used, or (3) conventional acronyms used in microbial ecology.

9- L96-98. This is a scholastic sentence. It can be removed.
This sentence is necessary for historical context and for readers unfamiliar with genome sequencing.

10- Section 2 should be presented as a subsection in section 3. In general, it seems to me that the style of this ms does not conform with that described within the authors' guidelines of this journal.
Section 2 is now Section 3.1

11- L187. Radioactive samples? Please, explain or amend.
This has been amended on new line 192-193 (p. 8, paragraph 1): “bailed samples or samples containing radioactive elements”

12- L252. …coefficient of variation ≤ 3. How was it calculated?
This has been amended on new line 265 (p. 10, paragraph 2): “community standard deviation divided by community mean”

13- L328. What does "largely" refer to?
This has been changed to “mostly” (new line 341, p. 13, paragraph 2).

14- L416-418. This sentence sounds too speculative.
The sentence has been modified (new line 425-427, p. 1, paragraph 2): “Fluid conduits also can become

clogged over time and inhibit microbial dispersal, such as from tectonic influences (86–88), secondary mineral infilling, or the overgrowth of microbial biofilms at biogeochemical hotspots (42, 89, 90)."

15- L527-529. This is a speculation.
This sentence has been removed.

16- L544-546, 573-576. Please, remove these speculations. Further studies are always needed.
L544-546 have been removed.
L573-576 have been incorporated into the new section 5 "Future Outlook"

Reviewer #4: I appreciate the changes made by the authors. From my side, the paper can now be accepted for publication in STOTEN.

We thank the reviewer for their comments!

Reviewer #5: The revised version of this manuscript has notably improved in clarity and readability, after having considered all points raised by me and the other 3 reviewers.

I still have some suggestions for 'cosmetic' amendments, as follows:

We thank the reviewer for their ‘cosmetic’ amendments and have made the changes.

- The concept of "microbial community as a tool for…" sounds awkward and it is only left in the title. I suggest to change the title in a statement title by reflecting what reported in the text (see e.g. L64, L117-121, L420-421, L430-431, L556-557, L578-579, L594-596 of the clean version). I would also avoid to repeat the same concept so many times along the text.

The authors discussed changing the title at great length, but in the end, all the authors agreed that the original title best represents the concept and focus of the paper.

L420-421 has been deleted.

L556-557 has been modified to (new line 558-560) “This study demonstrates that the regional-scale groundwater microbial community is a relevant data source; however, there were many limitations, as noted in Section 3.1, and future studies are needed. For example, the findings…”

- The graphical abstract is the less attractive figure among those presented. I would suggest to use the current figure 7 including some microbial data (e.g. the nine microbial clusters in figure 5?).

Graphical abstract has been updated. Figure 7 was not used because it only represents a portion of the DVRFS.

- The highlight #1 is likely out of the character limit

Highlight #1 modified to “The subsurface microbial community is consistent with known hydraulic connections.”

- The conclusive part was split in two sections (5. Future outlook, 6. Conclusion). I understand the Authors' reply, yet there are repetitions and redundancies between the new sections and in comparison to the previous text of the conclusion. Despite this is a subjective issue, I would suggest to shorten the conclusive part with a clearer focus on the outcomes of this study.

Section 5 and 6 were modified to reduce repetitions and redundancies, with the conclusive part having more focus on the outcomes of the study. (new lines 561-614)

Leaving aside the few minor issues, the manuscript is very interesting and fluently readable. It also provides many original data to the current knowledge in the field.

Thank you!

Subsurface Microbial Communities as a Tool for Characterizing Regional-Scale Groundwater Flow

Nancy Merino[1*], Tracie R. Jackson[2], James H. Campbell[3], Annie B. Kersting[1], Joshua Sackett[4,5,6], Jenny C. Fisher[4,7], James C. Bruckner[4], Mavrik Zavarin[1], Scott D. Hamilton-Brehm[4,8], Duane P. Moser[4,5*]

[1] Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, 94550, United States
[2] Nevada Water Science Center, U.S. Geological Survey, Boulder City, NV, 89005, United States
[3] Department of Natural Sciences, Northwest Missouri State University, Maryville, MO, 64468, United States
[4] Division of Earth and Ecosystems Sciences, Desert Research Institute, Las Vegas, NV, 89119, United States
[5] Division of Hydrologic Sciences, Desert Research Institute, Las Vegas, NV, 89119, United States
[6] School of Life Sciences, University of Nevada, Las Vegas, NV, 89154, United States
[7] Biology Department, Indiana University Northwest, Gary IN, 46408, United States (present address)
[8] Department of Microbiology, Southern Illinois University Carbondale, Carbondale, IL, 62901, United States (present address)

***Correspondence:**
Nancy Merino
Lawrence Livermore National Lab
7000 East Ave, B361/R1940
Livermore, CA 94550
925-424-7605
merino4@llnl.gov

Duane Moser
Division of Hydrologic Sciences
Desert Research Institute
Las Vegas, NV, 89119
702-862-5534
Duane.Moser@dri.edu

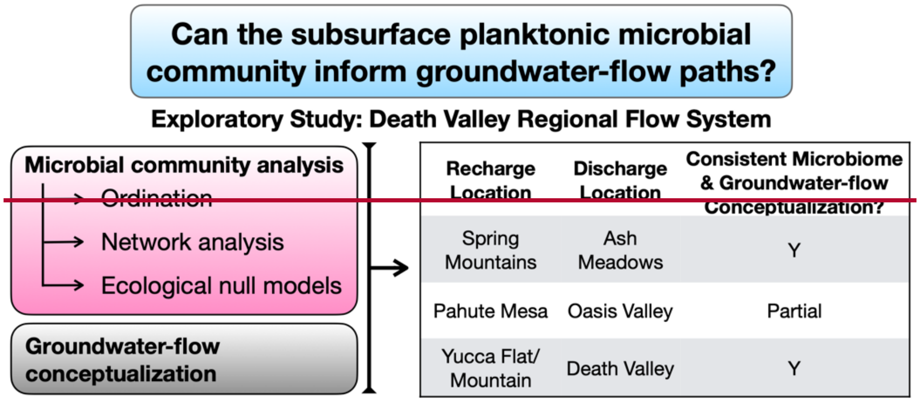
Keywords (max 6): fractured-rock aquifer, subsurface microbial ecology, microbial community assembly, regional-scale aquifer microbial community, deep biosphere

Abbreviations not standard to the field: Alkali Flat–Furnace Creek Ranch (AFFCR), Ash Meadows (AM), Death Valley Regional Flow System (DVRFS), Nevada National Security Site (NNSS), Pahute Mesa–Oasis Valley (PMOV), Underground Test Area (UGTA)

Highlights

1. The ~~groundwater~~ subsurface microbial community ~~of the Death Valley Regional Flow System~~ is consistent with known hydraulic connections.
2. Location plays a major role in microbial community variation.
3. Network analysis of common microbes corroborated recharge and discharge areas.
4. Ecological null models provided insight into community assembly patterns.

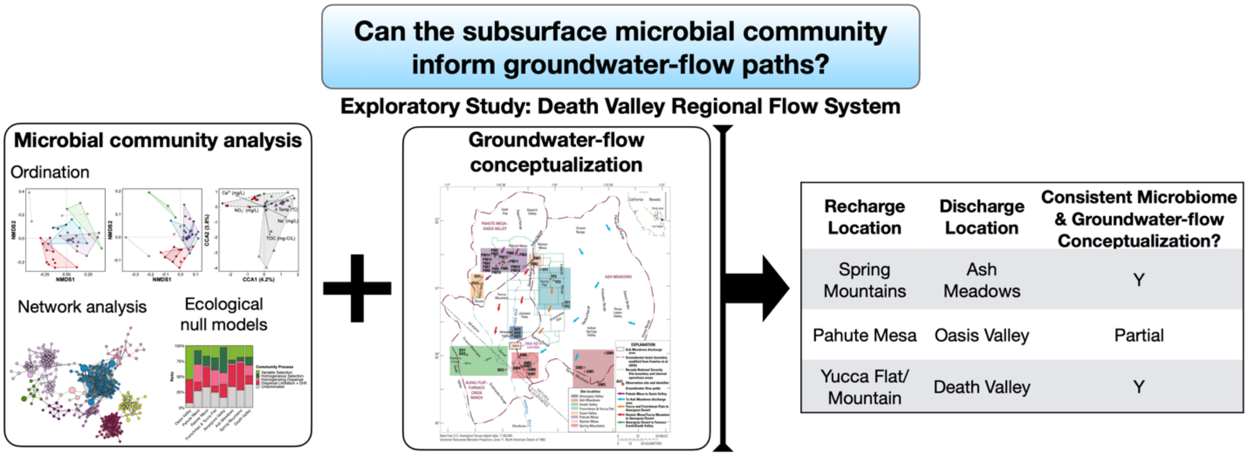
Graphical abstract



Can the subsurface planktonic microbial community inform groundwater-flow paths?

Exploratory Study: Death Valley Regional Flow System

Recharge Location	Discharge Location	Consistent Microbiome & Groundwater-flow Conceptualization?
Spring Mountains	Ash Meadows	Y
Pahute Mesa	Oasis Valley	Partial
Yucca Flat/ Mountain	Death Valley	Y



Can the subsurface microbial community inform groundwater-flow paths?

Exploratory Study: Death Valley Regional Flow System

Recharge Location	Discharge Location	Consistent Microbiome & Groundwater-flow Conceptualization?
Spring Mountains	Ash Meadows	Y
Pahute Mesa	Oasis Valley	Partial
Yucca Flat/ Mountain	Death Valley	Y

Abstract

Subsurface microbial community distribution patterns are influenced by biogeochemical and groundwater fluxes and may inform hydraulic connections along groundwater-flow paths. This study examined the regional-scale microbial community of the Death Valley Regional Flow System and evaluated whether subsurface communities can be used to identify groundwater-flow paths between recharge and discharge areas. Samples were collected from 36 sites in three groundwater basins: Pahute Mesa–Oasis Valley (PMOV), Ash Meadows (AM), and Alkali Flat–Furnace Creek Ranch (AFFCR). Microbial diversity within and between communities varied by location, and communities were separated into two overall groups that affiliated with the AM and PMOV/AFFCR basins. Network analysis revealed patterns between clusters of common microbes that represented groundwaters with similar geochemical conditions and largely corroborated hydraulic connections between recharge and discharge areas. Null model analyses identified deterministic and stochastic ecological processes contributing to microbial community assemblages. Most communities were more different than expected and governed by dispersal limitation, geochemical differences, or undominating processes. However, certain communities from sites located within or near the Nevada National Security Site were more similar than expected and dominated by homogeneous dispersal or selection. Overall, the (dis)similarities between the microbial communities of DVRFS recharge and discharge areas supported previously documented hydraulic connections between: (1) Spring Mountains and Ash Meadows; (2) Frenchman and Yucca Flat and Amargosa Desert; and (3) Amargosa Desert and Death Valley. However, only a portion of the flow path between Pahute Mesa and Oasis Valley could be supported by microbial community analyses, likely due to well-associated artifacts in samples from the two Oasis Valley sites. This study demonstrates the utility of combining microbial data with hydrologic, geologic, and water-chemistry information to comprehensively characterize groundwater systems, highlighting both strengths and limitations of this approach.

1 Introduction

Groundwater plays an essential role in providing water for societal needs and maintaining ecosystem health. As the global hydrologic cycle shifts with climate change, groundwater quality and quantity are increasingly affected, necessitating sustainable groundwater-management practices. These practices require a regional-scale understanding of the groundwater system. Regional-scale groundwater flow is typically characterized using geologic, water-level, water-chemistry, and aquifer-testing data (1, 2). However, groundwater systems also are impacted by and serve as distinct habitats for microbial communities (3–5). Thus far, groundwater microbial communities remain an underutilized data source to understand hydraulic connections between recharge and discharge areas.

Groundwater flow is governed largely by recharge and discharge rates and the hydraulic properties of the saturated rocks. Groundwater-flow paths are derived from hydraulic gradients between areas of recharge (e.g., precipitation in high topographic regions) and discharge (e.g., springs in low topographic regions) (6–8). Geologic units also exert control on groundwater-flow pathways. For example, the juxtaposition of permeable geologic units against impermeable units downgradient results in a hydraulic barrier to groundwater flow. The sources of groundwater and travel times along groundwater-flow paths can be informed by water-chemistry data (9–11).

The groundwater microbial community also can be used to characterize groundwater systems and validate groundwater-flow paths. This concept dates back to the 1890s when studies began using microbes as tracers to understand subsurface pathogen transport and to delineate groundwater-flow paths through injection and recovery experiments (12). With the advent of the genomics era and the reducing costs of high-throughput sequencing, it may now be possible to use the subsurface microbial community to aid in groundwater characterization. Unattached (planktonic) microbes may be transported through aquifers along groundwater-flow paths through pore spaces and fractures (13–15). Several abiotic and biotic factors impact subsurface microbial movement (16, 17), including groundwater-flow rates, hydraulic barriers, pore size, microbial adhesion processes, and the physiological state of microbial cells.

Subsurface microbial transport may impact microbial community assemblages through space and time and could result in more similar or dissimilar communities. Microbial community assembly can be evaluated with ecological null models (18–20) (see **Table S1** for terms and definitions used for ecological null models) that may inform groundwater characterization. For example, high microbial-dispersal rates (i.e., ‘homogenizing dispersal’) may provide a signature for tracking hydraulic connections through the co-occurrence of microorganisms in recharge and discharge areas. In contrast, hydraulic barriers may prevent or impede microbial transport and mixing (‘dispersal limitation’), such as in the hard-rock aquifers of Brittany (France) (21), causing microbial communities to diverge over time or along flow paths through stochastic ecological processes (22). Biogeochemical and redox conditions also drive community assembly, selecting for adaptable microorganisms (21, 23, 24), and are known as deterministic processes that occur when conditions are consistent (‘homogeneous selection’) or fluctuating (‘variable selection’). Consistent environmental conditions lead to less divergence in the community, whereas fluctuating conditions can perturb microbial abundances and diversity. Ecological null models can identify the dominating ecological processes that impact microbial community assembly between two communities. Along with other analytical approaches to analyze microbial community data, the groundwater microbial community could be an advantageous tool to supplement other geologic and hydrologic datasets to provide a comprehensive understanding of regional groundwater flow.

This study assesses the potential of using subsurface microbial community patterns as a supplemental dataset to characterize regional-scale groundwater flow. The Death Valley Regional Flow System (DVRFS) microbial community was used as a test case because, over the past few decades, hundreds of boreholes and wells have been drilled in the DVRFS to monitor water levels and water quality. Thus, the DVRFS provides a unique opportunity to identify novel microbial community patterns on a regional-scale and to determine whether microbial community patterns are consistent with regional-scale groundwater flow-paths (25). The aims of this study were to (1) evaluate the groundwater bacterial and archaeal community distribution patterns with alpha- and beta-diversity analyses and network analysis, (2) identify the assembly processes dominating the community patterns using ecological null models, and (3) use the microbial community patterns and assembly processes to assess any (dis)similarities along groundwater-flow paths. To achieve these aims, we collected samples from recharge and discharge areas and hypothesized that microbial (dis)similarities, co-occurrences, and assembly patterns may infer and support the most recent groundwater conceptualization (25).

2 Materials and Methods

2.1 Site Description

The DVRFS is a 100,000 km^2 region that is part of the Great Basin physiographic province in southern Nevada and California (**Figure 1**). The study area includes three groundwater basins in the DVRFS: Pahute Mesa–Oasis Valley (PMOV), Ash Meadows (AM), and Alkali Flat–Furnace Creek Ranch (AFFCR) (**Figure 1**) (25). The Nevada National Security Site (NNSS, formerly Nevada Test Site) intersects these three groundwater basins. Land-surface elevations range from –86 to 3,600 m relative to sea level. Principal geologic units are grouped into four categories: Paleozoic carbonate rocks, Tertiary volcanic rocks, Cenozoic basin fill, and undifferentiated, low-permeability siliciclastic and granitic rocks (25).

Long-term water-level and water-quality monitoring have occurred in the study area. A total of 828 underground nuclear tests were detonated beneath the NNSS from 1951–1992 (26). Underground nuclear testing prompted a long-term U.S. Department of Energy environmental management program, known as the Underground Test Area (UGTA) Activity, to assess and monitor potential groundwater transport of radionuclides. Extensive hydrogeologic investigations (25, 27, 28) also were conducted by the U.S. Geological Survey to evaluate water levels and groundwater-withdrawal rates. Several groundwater studies (25, 29–41) have evaluated flow paths in the DVRFS.

This study evaluates groundwater-flow paths using the most recent groundwater conceptualization of the DVRFS (25, 42) (**Figure 1**). The five regional groundwater-flow paths evaluated in this study are: (1) Pahute Mesa to Oasis Valley; (2) Spring Mountains to Ash Meadows discharge area; (3) Frenchman and Yucca Flat to Amargosa Desert; (4) Rainier Mesa/Yucca Mountain to Amargosa Desert; and (5) Amargosa Desert to Death Valley (25) (**Figure 1**). In the PMOV basin, Pahute Mesa is the primary recharge area contributing to discharge from Oasis Valley (42). In the AM basin, most recharge is derived from the Spring Mountains and Sheep Range, with lesser recharge amounts from the Groom, Pahranagat, Desert, Pintwater, and Spotted Ranges. Groundwater from the Spring Mountains and other ranges in the AM basin moves southward and westward to high-volume springs of the Ash Meadows discharge area, the largest oasis in the Mojave Desert (43). Groundwater from Yucca and Frenchman Flats moves into the Amargosa Desert through the well AD-4 corridor north of Ash Meadows discharge area (**Figure 1**). The well AD-4 corridor hydraulically connects transmissive carbonate rocks in the AM basin with transmissive basin fill in the Amargosa Desert (see **Section 4.3** for more details). Groundwater from Rainier Mesa and the Yucca

Mountain area moves southward into the Amargosa Desert. Most groundwater in the Amargosa Desert moves westward and discharges at Furnace Creek in Death Valley.

2.2 Sample Collection and Geochemical Analysis

Between 2008–2014, a total of 42 samples were collected from wells, mine vent holes, or springs. These samples were collected from 36 sites across the DVRFS (**Figure 1, Table S2**). The number of sites per location varied (**Table S2**) based on accessibility [Amargosa Valley (sites [n] = 1, samples [s] = 3), Ash Meadows discharge area (n = 7, s = 7), Death Valley (n = 3, s = 3), Frenchman Flat (n = 2, s = 2), Oasis Valley (n = 2, s = 2), Pahute Mesa (n = 12, s = 16), Rainier Mesa (n = 2, s = 2), Spring Mountains (n = 4, s = 4), and Yucca Flat (n = 3, s = 3)]. The location Amargosa Valley is within the Amargosa Desert. For well samples, groundwater generally was pumped at high rates (hundreds of L/min) for hours-to-days prior to collection of geochemical and microbial samples to minimize wellbore artifacts. In a few cases, well water samples were obtained by lower rate “jack pumps” (in-line submerged pump) or from static water columns with discrete sampler deployments (i.e., “bailers”). The mine water samples (U12n-10 [RM1] and U12n-vent2 [RM2]) were collected from flooded workings by bailer from hundreds of meters overhead via vertical ventilation holes. Spring samples from Ash Meadows and Spring Mountains were collected by pumping with a peristaltic pump and sterile platinum-cured silicone tubing (Masterflex LS-15) inserted as deeply into the spring orifice as possible via a 24’ (7.3m) telescoping probe to minimize surface associated influences. For the site in the Amargosa Valley (4PD), water was pumped continuously over a few weeks, with samples collected on days 1, 9, and 23. At four sites on Pahute Mesa (ER-20-8 [PM15, PM16], ER-EC-13 [PM10, PM5], ER-EC-15 [PM4, PM7, PM8], and PM-3 [PM1, PM2]), samples were collected via pumping from discrete piezometers screened at multiple different depths, as listed in **Table S2**, because there may be microbial community differences with depth (44).

Cells for DNA extraction were collected on 0.22 µm Sterivex^{TM} filters (EMD Millipore, U.S.A.) with the filtrate retained for geochemical analysis and the filters transported on dry ice and stored at -80ºC. The DNA extraction and sequencing methods are described in **Supplementary Methods section 1.1**. The method and volume of filtration for cells varied depending on the circumstance: (1) ranging from tens to hundreds of liters collected via a pressurized sampling manifold (~0.5 bar) at the wellhead for most pumped samples; or (2) two to four liters filtered offsite for bailed samples or samples containing radioactive elements. Nevares Deep Well 2 (DV1) is artesian and was sampled at ambient pressure (0.4 bar) via a dedicated sampling port at the wellhead. Most geochemical analyses reported here were processed as part of the UGTA groundwater sampling and analysis program (45). Otherwise, the geochemical analyses were performed at the DRI Water Analysis Laboratory (Reno, NV) or at Princeton University T.C.O. (Princeton, NJ) (44), as described in **Supplementary Methods section 1.2**.

2.3 Statistical Analyses of the Geochemical Data

The geochemical dataset (**Table S2, color scheme in Table S3**) used in this paper contains information on location, depth, and geochemical measurements. Several geochemical parameters contained missing and left-censored (i.e., below detection limit) values. To reduce the number of missing values, some missing values were substituted with geochemical data collected from the same site on a different sampling date. Substituted data were obtained from the UGTA Activity Environmental Management Project database (45). Although many of these values were measured between the 1970s to 1990s, the geochemical parameters have remained generally stable over the past few decades (45). A piper diagram was created using the Geochemist's Workbench Release 11.0.8 (Aqueous Solutions LLC, USA) with a subset of the geochemical data (Ca^{2+}, Cl^{-}, K^{+}, Mg^{2+}, Na^{+}, SO_4^{2-}, HCO_3^{-}). Principal component analysis (PCA) also was used to visualize the sample variation through dimensionality reduction, as described in **Supplementary Methods section 1.3**.

2.4 Quality control of the 16S rRNA amplicon sequences

Raw sequences were processed through QIIME 1 and DADA2 v1.12.1 (46), as described in **Supplementary Methods section 1.4**. Because there were two different sequencing batches, samples were separated by batch and processed through DADA2 separately to account for potential error rate model differences. This resulted in amplicon sequence variants (ASVs) unique to each batch, which may have been caused by technical or biological differences (e.g., sample preparation, human processing, PCR amplification, sequencing errors, and multiple 16S rRNAs per cell). However, many ASVs from both batches were closely-related, as determined by inspection of a phylogenetic tree and sequence alignment, and were subsequently clustered into OTUs in QIIME 2 (47) at the 97% identity level with the function vsearch cluster-features-de-novo. OTU clustering at the 97% identity level can provide comparable overall results as ASVs on a broad-scale (48–53). Taxonomy was assigned to the clustered OTUs using the QIIME 2 function feature-classifier classify-sklearn with a pretrained-classifier SILVA v138 database for OTUs from 515F/806R region of 16S rRNA sequences (54–57). The OTU and taxonomy table were then imported to phyloseq v1.16.2 (58), and contaminant OTUs removed, as described in **Supplementary Methods section 1.4**. Based on rarefaction curves generated using the R package iNEXT (59), sufficient sequencing depth was achieved for each sample (**Figure S1**).

2.5 Microbial community analyses

Alpha diversity (diversity within samples) indices were computed with phyloseq (function estimate_richness) and picante v1.8.2 (60) using a phylogenetic tree generated as described in **Supplementary Methods section 1.5**. The means between each categorical variable were evaluated using the non-parametric Mann-Whitney test with false discovery rate p-value adjustment. Phylogenetic relatedness also was determined by calculating the standardized effect size (SES) of Faith's PD (phylogenetic diversity), MPD (mean pairwise distance), and MNTD (mean nearest neighbor phylogenetic distance) (see **Table S1** for definitions), as described in Supplementary Methods.

Beta diversity (diversity between samples) was evaluated using five approaches for comparison: (1) non-metric multidimensional scaling (NMDS) ordination on unweighted and weighted UniFrac distances, (2) DEICODE robust Aitchison principal-component analysis (RPCA) (61) in QIIME 2 with auto-RPCA, (3) centered-log ratio (CLR) PCA (62), (4) phylogenetic isometric-log ratio (PhILR) PCA (63), and (5) canonical correspondence analysis (CCA). Permutational analysis of variance (PERMANOVA) marginal tests with 999 permutations were conducted with R function adonis2 on the NMDS ordinations and homogeneity was checked with R function betadisper.

Deterministic and stochastic ecological processes were evaluated using the β-nearest taxon index (βNTI) and Raup-Crick (Bray-Curtis) (RCBC) metrics, following null model analyses (18–20, 22, 64, 65) (see **Table S1** for definitions). Only for these metrics, communities were rarefied to the lowest sequencing depth. βNTI is used to identify deterministic processes by quantifying phylogenetic turnover. First, species were randomly shuffled across the tips of the phylogeny, also known as the between-community mean-nearest-taxon-distance (βMNTD) metric, to obtain a null distribution after repeating the shuffling 999 times. Subsequently, the significance is evaluated by calculating the difference between the observed βMNTD and mean null distribution, with βNTI < –2 (less phylogenetic turnover; homogeneous selection) and βNTI > 2 (more phylogenetic turnover; variable selection) as significant (18). Insignificant βNTI suggests that the observed compositional differences are due to stochastic ecological processes (66), further determined by RCBC. RCBC is based off of community composition and OTU abundances, and follows Raup–Crick (67) to probabilistically assemble local communities, followed by quantification with Bray–Curtis (68) to obtain a null distribution after repeating 9,999 times (19, 65, 69). The RCBC null distribution is then standardized between -1 and +1 and significant values determined as RCBC < –0.95 (homogeneous dispersal) and RCBC > 0.95 (dispersal limitation). Ecological null models also assumes that phylogenetic relatedness is associated with ecological niche differences (18, 24, 64), which was confirmed for the DVRFS microbial community using Mantel correlograms, following Dini-Andreote *et al.* (2015). A significant (P < 0.05) positive correlation was observed across relatively short phylogenetic distances (**Figure S2**).

Network analysis was performed using SPIEC-EASI (SParse InversE Covariance Estimation for Ecological Association Inference; 86) on commonly observed OTUs among all samples, excluding Oasis Valley samples, to ensure robust results. ‘Common OTUs’ were defined as those with low coefficient of variation across samples (≤ 3; community standard deviation divided by community mean) (58) and observed at least two times in more than two samples (466 total OTUs). The glasso (sparse inverse covariance selection) and Meinshausen–Buhlmann (MB; neighborhood selection) method were evaluated with the Stability Approach to Regularization Selection (StARS) and bounded-StARS (bstars) selection criteria. Similar results were observed for both selection criteria. Since a wider degree distribution range was observed for glasso (**Figure S3**), the glasso bstars approach was chosen for subsequent analysis in igraph (71). OTUs with weak connection (edge weight < 0.02) were removed. Various clustering algorithms were performed to define clusters within the network, and the unsupervised Louvain clustering algorithm (72, 73) was chosen based on the highest modularity score of 0.694 amongst all the algorithms. The 16 clusters were then manually inspected; clusters with only two connections were removed and nearby connected clusters with the same overall location associated with OTUs were combined. Cytoscape (https://cytoscape.org/) was used for visualization and network analysis.

2.6 Data Availability

The 16S rRNA gene amplicon sequences were deposited at GenBank under the accession KEWZ00000000. The version described in this paper is the first version, KEWZ01000000. Scripts used to analyze and create figures of the geochemistry and microbial community are available at https://github.com/LLNL/2022_DVRFS_microbiome.

3 Results

3.1 Sample description, study limitations, and biases

A total of 42 samples were collected from 36 sites (**Figure 1, Table S2–S3**) that encompass the AM, AFFCR, and PMOV groundwater basins. Samples were collected opportunistically based on site availability, and most samples were collected from recharge and discharge areas within their respective groundwater basins. Due to our sampling limitations, the influence of seasonality, changes in recharge events, or hydraulic dynamics could not be determined on the temporal dynamics of subsurface microbial communities. This has been documented in previous studies of the Hainich Critical Zone Exploratory fractured aquifer system (74), a landfill-leachate contaminated aquifer (75), and a tar-oil contaminated aquifer (76). However, the very long residence times and deep flow paths of water in this system would argue against short-term effects other than from pumping (25). There also are sample biases in this study, including the number of samples per location, sample collection method, and two sequencing batches. The specifics of these biases are described in **section 2.2** and were considered when analyzing the microbial community data.

3.2 Geochemistry

The major-ion chemistry of study sites were grouped into three broad categories: Ca-Mg-HCO_3, Na-HCO_3, and NaCl dominated waters (**Figure 2A**), which were determined from a piper diagram (**Figure**

S4) and concur with previous reports (32, 35, 36, 41, 77). Principal component analysis (PCA) indicates that Ca-Mg-HCO_3-type waters cluster together, whereas the Na-HCO_3-type and NaCl-type waters group together. Notably, all 42 samples distinctly clustered into carbonate and volcanic rock types (**Figure 2A**). Rock types were assigned based on the predominant rock type found at the depth sampled. Most samples also clustered by location, with the largest variation observed for samples collected from Yucca Flat, Rainier Mesa, and Oasis Valley (**Figure 2B**). The water chemistry of Yucca Flat samples (YF1–YF3) is highly variable because of different hydrogeologic settings. YF1 (UE-2ce-WW-Nash), YF2 (U-3cn5-Bilby), and YF3 (UE-3E-4-Aleman) were sampled from wells completed in carbonate rock, carbonate and volcanic rocks, and volcanic rocks, respectively. Rainer Mesa samples (RM1 [U12n-10], RM2 [U12n-vent2]) likely do not cluster together because of the water geochemistry. RM1 has higher SO_4^{2-}, Na^+, and SiO_2^- concentrations (influenced by a local granite intrusion and pyrite oxidation leading to high sulfate) compared to RM2 (**Figure 2C, Table S2**). Oasis Valley samples (OV1, OV2) likely do not cluster together because of higher SO_4^{2-} and total organic carbon (TOC) concentrations observed in OV1.

3.3 Description of groundwater bacterial and archaeal groundwater microbial community

The groundwater microbial community consisted of 5,124 unique operational taxonomic units (OTUs) that were identified from 1,267,990 reads (**Tables S4–S5**). The final number of reads for OTUs per sample ranged from 2,790–123,420 (average = 30,190) (**Table S4**). Because a higher number of reads was observed for samples sequenced in batch 2 (average = 90,352; batch size = 7) than batch 1 (average = 18,158; batch size = 35), subsequent analyses considered sequence batch effects. Among the OTUs identified, only 43 OTUs were present in more than ten sites. The most commonly observed taxa include those from genera *Pseudomonas*, *Curvibacter*, *Phenylobacterium, Thiobacillus,* and *Hydrogenophaga*. Although community structure varied from site to site, the overall community was composed mainly of Bacteria (49–100%), and most sites were dominated by the phyla Proteobacteria (average = 42%), Bacteroidota (15%), Firmicutes (13%), and Desulfobacterota (11%) (**Figure S5**). Within Proteobacteria, only Alphaproteobacteria and Gammaproteobacteria classes were observed. The archaeal relative abundances averaged 7% for all sites, and the highest abundances (51%) were observed in sample DV3 (Inyo-BLM 1).

Diversity within communities (alpha diversity) largely varied based on location and location type (**Figure 3 and S6**). Although there were significant differences ($P < 0.05$) between the means of some categories (e.g., springs and wells) (**Figure S6**), the standardized effect size (SES) of all alpha diversity metrics suggests that the differences could be affected by species richness (number of OTUs in each sample) (**Figure 3**) (78–80). SES also can describe the phylogenetic relatedness within communities, and for most communities sampled, the SES-PD and SES-MNTD were negative and significantly different than the null model ($P < 0.05$), indicating that OTUs within each community were more closely related than expected by chance.

3.4 Microbial community variability by location

Differences in community composition between samples (beta diversity) was mostly explained by location. PERMANOVA by location- and geochemical-type variables revealed that location significantly (PERMANOVA $P = 0.003$; ANOSIM $P = 0.001$) explained the most variance of any factor (~27%) for the NMDS weighted UniFrac (**Table 1**). NMDS ordination of unweighted (**Figure 4A**) and weighted UniFrac (**Figure 4B**) distances both identified that Oasis Valley microbial communities are distinct from other study-area microbial communities. The following clusters appear to group together for both ordinations, even when removing potentially biased samples (e.g., sequence batch): (1) Ash Meadows and Spring Mountains, and (2) Pahute Mesa and Frenchman and Yucca Flat. Rainier Mesa and Death Valley microbial communities differ between unweighted and weighted UniFrac NMDS ordinations, suggesting that both OTU relative abundances and phylogeny are important factors for the sites sampled. In addition to location, sampling depth relative to the water table ($P = 0.001$) significantly contributed to the variation but explained only ~6%.

Geochemical-type variables minimally explained community variation (**Table 1, Figure 4C**). For NMDC ordinations, rock type is the most significant contributor (~9%; $P = 0.002$), followed by TOC (~6%, $P = 0.026$) and temperature (~6%, $P = 0.002$). Unlike location-type variables, the PERMANOVA of geochemical-type variables is significantly influenced by sequence batch runs ($P = 0.039$) and sampling-method approaches ($P = 0.020$). However, constrained ordination using CCA (**Figure 4C**) confirmed that rock type, TOC, and temperature are significant ($P < 0.05$) factors determining the microbial community variation. Oasis Valley samples were removed before performing CCA because of their distinct geochemical and microbial compositions. Although CCA only explained ~8% of the total community variation, the microbial communities could be clustered by rock type and location. Notably, similar to the geochemical ordination (**Figure 2**), the two rock type clusters (carbonate and volcanic) followed along the vectors for calcium, nitrate, and sodium. Within location, CCA also revealed that the community variation may be explained by certain geochemical-type variables (e.g., **Figure 4C**; Pahute Mesa communities appear to vary along a temperature gradient). Thus, the combined results of CCA and NMDS ordinations indicate that, on a regional-scale, although microbial community variation can be explained more by location, geochemical conditions play an important role.

Other ordinations, including DEICODE RPCA (61), CLR PCA, and PhILR PCA (63) (**Figure S7**), that take into account the sparse, compositional nature of microbial community datasets (62) were evaluated to compare against the NMDS and CCA ordinations. Consistent with the NMDS ordinations, Oasis Valley communities were identified as a separate group for all three ordinations (**Figure S7A, S7D, S7G**). However, this result may be attributed to a noticeable separation of samples by sequence batch for CLR (**Figure S7E**) and PhILR PCA (**Figure S7H**). Sequence batch did not seem to impact DEICODE RPCA (**Figure S7B**), likely because the DEICODE algorithm utilizes the geometric mean of log-transformed nonzero data and conducts matrix completion (61). After removing potentially biased samples (i.e., Batch 2 and tunnel-collected samples), both CLR and PhILR PCA ordinations also identified the two groups observed for NMDS ordinations: (1) Ash Meadows and Spring Mountains, and (2) Pahute Mesa and Frenchman and Yucca Flat (**Figure S7F** and **S7I**). In contrast, these two groups were not observed for DEICODE RPCA (**Figure S7C**) and may be the result of the likely highly-ranked (61) nature of the microbial community dataset, in which samples contain few similar microbes depending on location or other unknown factors not included in the metadata. Highly-ranked datasets also might be caused by unforeseen local-scale geochemical gradients that impact the microbial community and influence the resultant RPCA ordination (61). In contrast, the unweighted and weighted UniFrac NMDS ordination may accurately represent the microbial community and has been demonstrated to achieve high clustering accuracy (61, 81, 82).

Network analysis of the ‘common OTUs’ identified nine clusters and revealed that co-occurring taxa generally subsisted within the same location (**Figure 5, Figure S8** depicts the ratio of reads per location for each OTU). For network analysis, OV samples were removed because beta-diversity and geochemical analyses indicate that OV1 and OV2 were microbially and geochemically distinct from all other locations. Five clusters (NNSS 1–5) represent OTUs that co-occurred within or near the NNSS sites (Pahute Mesa, Rainier Mesa, and Frenchman and Yucca Flat). Each NNSS 1–5 cluster contains different ‘common OTUs’ (**Table S7**). Cluster AV (Amargosa Valley) was categorized separately from the NNSS clusters since the majority of OTUs co-occurred only in Amargosa Valley. In addition, there is a high degree (average = 22.8) of connectivity within cluster AV, which may be attributed to the collection of three samples pumped at the same site (4PD [AV1–3]) over time from 1 to 23 days. This suggests there is some degree of temporal consistency within this site, and potentially in other sites. In comparison, the degree of connectivity for the five NNSS clusters average between 3.2–7.7 (**Table S6**). Cluster AM also has a high average degree (19.3) of connectivity and three nodes connect cluster AM to the rest of the network via cluster SM.

3.5 Ecological processes that influence microbial community variation

Null models (see **Table S1** for definitions) were used to identify the ecological processes that drove microbial community assembly and could influence community variation. This approach has been applied to communities in surface water (83), soil (64, 83, 84), and the subsurface (18–20, 24, 65, 83, 85, 86). There are two null modeling steps to identify deterministic and stochastic ecological processes: βNTI and RCBC, as described in **section 2.5**. Overall, stochastic and undominated processes governed the microbial communities most (**Figure 6 inset**), with deterministic processes only contributing ~25% (except for Oasis Valley at ~50%). Although certain locations had limited samples in our study, the assembly processes, taken as a whole, support the clusters identified by ordination analyses. The microbial communities of Spring Mountains, Ash Meadows, and Oasis Valley were governed largely by variable selection (βNTI > 2), dispersal limitation ($|\beta\text{NTI}| < 2$ and RCBC > 0.95), and undominated processes ($|\beta\text{NTI}| < 2$ and $|\text{RCBC}| < 0.95$) when compared against other communities (**Figure 6**). For microbial communities within and near the NNSS, homogeneous dispersal ($|\beta\text{NTI}| < 2$ and RCBC < -0.95) and homogeneous selection (βNTI < -2) were more prominent. For example, Pahute Mesa communities were more similar (homogeneous selection) and had less turnover (homogenizing dispersal) than expected by chance. This signifies high dispersal rates and comparable geochemical conditions across Pahute Mesa that forced local communities to be more similar.

4 Discussion

~~This study assesses the potential of using microbial community patterns to characterize regional-scale groundwater flow.~~ In this study, we hypothesize that groundwater flows through hydraulically connected fractures and enables the dispersal of planktonic microorganisms, sometimes over considerable distances. Along a given flow path, geochemical conditions evolve with rock-water interactions or mixing of disparate fluids within interconnected pore spaces. These conditions can impart selective pressures on transported microorganisms and influence the local community assembly patterns. Fluid conduits also can become clogged over time and inhibit microbial dispersal, such as from tectonic influences (87–89), secondary mineral infilling, or the overgrowth of microbial biofilms at biogeochemical hotspots (42, 90, 91). Separated communities may change with time into disparate compositions because of stochastic processes (e.g., genetic or ecological drift) (22). Ecological processes can be quantified (18–20), and in combination with microbial diversity and co-occurrence patterns, we assess whether regional microbial community patterns are consistent with regional groundwater-flow paths (**Figure 1**). Microbial community patterns are compared to the following regional groundwater-flow paths (25): (1) Spring Mountains to Ash Meadows discharge area; (2) Frenchman & Yucca Flat to Amargosa Desert; (3) Rainier Mesa to Amargosa Desert; (4) Pahute Mesa to Oasis Valley; and (5) Amargosa Desert to Death Valley (**Figure 1**).

4.1 Microbial community dispersal and evolution from Spring Mountains to Ash Meadows

The Spring Mountains (recharge area) to Ash Meadows (discharge area) groundwater-flow path (**Figure 1**) likely transports planktonic microorganisms. In the AM Basin, the microbial community is similar between the recharge area (Spring Mountains) and Ash Meadows discharge area (**Figure 4**). This result was corroborated with network analysis (**Figure 5**), which identified ‘common OTUs’ co-occurring in the Spring Mountains and Ash Meadows discharge area (e.g., ‘common OTUs’ in cluster SM). The separation of ‘common OTUs’ into two clusters (AM and SM; clusters previously described in **section 3.4**) suggests that other factors impacted the communities observed in each location. Indeed, null model analyses identified three major ecological forces that can explain differences between the Ash Meadows and Spring Mountains communities: selection processes, dispersal limitation, and undominating processes (**Figure 6**). This is not surprising given the differences in elevation and geochemistry (**Figure 2**) that arise from the evolution of groundwater along the flow path. Notably, these two communities are separated in the CCA ordination (**Figure 4C**), likely due to differences in calcium and nitrate. In addition, the most-recent groundwater characterization study of the DVRFS (25) identified other sources that contribute to Ash Meadows discharge, including the Sheep, Desert, Pintwater, and Spotted Ranges (**Figure 1**).

4.2 Microbial community similarities within the NNSS and at Amargosa Valley

Microbial communities at sites within and near the NNSS were relatively similar (Amargosa Valley, Frenchman and Yucca Flat, Pahute Mesa, Rainier Mesa; **Figure 4**), despite locations occurring within different groundwater basins (**Figure 1**). Network analysis (**Figure 5**) also identified co-occurring ‘common OTUs’ within these locations, and five disparate NNSS clusters of ‘common OTUs’ were identified (Cluster NNSS 1–5). The five NNSS clusters connected to the network via cluster AV (Amargosa Valley; within Amargosa Desert), cluster DV (Death Valley), or cluster AM (Ash Meadows).

Similarities between the microbial communities of NNSS sites with Amargosa Valley suggest the communities have adapted to similar environmental conditions or have dispersed via a groundwater flow connection towards Amargosa Valley. Many communities between these locations were dominated by homogenizing dispersal or homogeneous selection (**Figure 6**), demonstrating the communities are more similar than expected by chance. Network analysis also grouped ‘common OTUs’ with putatively diverse metabolisms within cluster AV. The four most abundant OTUs in cluster AV were putative sulfur-oxidizing (e.g., *Thioalkalispiraceae, Hydrogenophilaceae*), methanotroph (e.g., *Methylomonas*), and iron-oxidizing (e.g., *Gallionellaceae*) microbes. Some OTUs are likely metabolically flexible, capable of mixotrophic or facultative anaerobic growth, such as those within family *Hydrogenophilaceae.* The diversity of potential metabolisms present in cluster AV is consistent with the geochemical mixing of multiple groundwater sources in the Amargosa Desert. Groundwater in the Amargosa Desert is derived largely from the AFFCR and AM groundwater basins (25), in which three recharge areas converge in the Amargosa Desert: (1) Yucca Mountain and nearby upland areas in the AFFCR groundwater basin; (2) Yucca Flat in the AM groundwater basin; and (3) infiltration from the Amargosa River and Fortymile Wash in the AFFCR basin (**Figure 1**). This mixed groundwater then flows toward the Furnace Creek discharge area in Death Valley.

4.3 The Death Valley microbial community likely supports the basin-fill flow conceptualization

There are two conceptualizations for groundwater flow towards Death Valley: (1) the deep-carbonate flow conceptualization; and (2) the basin-fill flow conceptualization (**Figure 7**). Previous groundwater studies suggested that Death Valley discharge was sourced from a deep-carbonate flow path that passed beneath the Ash Meadows discharge area and Amargosa Desert (36, 92, 93) (**Figure 7**). In contrast, the most recent groundwater study by Halford and Jackson (2020) determined that the basin-fill flow conceptualization best explains water levels, water chemistry, and aquifer-testing data. The basin-fill flow conceptualization interprets groundwater flow as moving through shallow basin-fill, volcanic, and carbonate rocks toward Furnace Creek–the principal discharge area in Death Valley (**Figure 7**).

The Death Valley microbial community better supports hydraulic connections proposed in the basin-fill conceptualization, rather than the deep-carbonate flow conceptualization. The Death Valley microbial community consists of samples collected from three carbonate wells: DV1 (Nevares Deep Well 2) and DV2 (Nevares) in the Furnace Creek discharge area; and DV3 (Inyo-BLM 1) in the central Amargosa Desert (**Figure 1**). For the Death Valley microbial community to support the deep-carbonate flow conceptualization, microbial communities are expected to be related between Death Valley, Ash Meadows, and the Spring Mountains. Instead, NMDS ordination (**Figure 4A and 4B**), network analysis (**Figure 5**), and ecological null models (**Figure 6**) suggest that Death Valley microbial communities are more related to Amargosa Valley and NNSS locations, rather than to Ash Meadows and Spring Mountains. Although there may be some similarity between Death Valley communities and Ash Meadows in the CCA ordination (**Figure 4C**), this analysis only accounts for ~8% of the total community variation. Moreover, homogeneous dispersal or selection contributed to community turnover between Death Valley, NNSS locations, and Amargosa Valley (**Figure 6**; **Table S8**). This suggests that these communities were more similar than expected by random chance. In contrast, dispersal limitation and variable selection played a role in community turnover between Death Valley and Ash Meadows/Spring Mountains, which suggests a limited hydraulic connection between these areas. Network analysis indicates that cluster DV is closely associated with the NNSS clusters, particularly NNSS1 and NNSS2, and ‘common OTUs’ observed in Death Valley are also found in NNSS clusters and cluster AV (**Figure 5**).

4.4 The Death Valley microbial community may support groundwater compartmentalization into shallow and deep zones

Halford and Jackson (2020) posit that the DVRFS is compartmentalized into two parts: (1) a shallow, high-transmissivity part within 500 m of the water table where nearly all flow occurs; and (2) a deep, less active, low-transmissivity part that has limited interaction with the shallow part (25). The Death Valley microbial community appears to support a compartmentalized groundwater system.

Nevares wells (DV1/DV2) have ~100 m sampling depths and are conceptualized as occurring within the active, high-transmissivity part of the groundwater system. In contrast, Inyo-BLM 1 (DV3) was sampled at ~600 m depth, and is in low-transmissivity rock (~9–84 m^2 per day) (25, 94); thus, Inyo-BLM 1 (DV3) is conceptualized as occurring within the deep, less active, low-transmissivity part of the groundwater system. Inyo-BLM 1 (DV3) and Nevares Deep Well 2 (DV1) have similar geochemical conditions (**Figure 2, Table S2**), including temperature and sulfate concentrations. Despite similar aqueous chemistries, Thomas *et al.* (2013) suggested hydraulic isolation between the microbial communities of Inyo-BLM 1 (DV3) and Nevares Well and Spring.

In this study, NMDS and CCA ordinations (**Figure 4**) appears to concur with Thomas *et al.* (2013), with more similarity between Nevares wells (DV1/DV2) compared to Inyo-BLM1 (DV3). There also are large phylogenetic and compositional differences. Only eight OTUs were present in both Nevares Deep Well 2 (DV1) and Inyo-BLM 1 (DV3); and only two OTUs were present in both Nevares (DV2) and Inyo-BLM 1 (DV3). In contrast, 256 OTUs were present in both Nevares communities (DV1/DV2). The Inyo-BLM 1 (DV3) community consisted of a large population of putative methanogens (~50%), followed by sulfate reducers (~31%). In contrast, probable sulfate reducers composed most of the population at Nevares (DV1/DV2). Community differences may be impacted by unmeasured geochemical factors, such as higher pressures at depth in Inyo-BLM 1 (DV3) compared to the shallow Nevares wells (DV1/DV2). These deterministic differences were not apparent using ecological null models (**Figure 6**), and more samples are required to confirm the null models between Inyo-BLM 1 (DV3) and Nevares wells (DV1/DV2).

4.5 Lack of similarity between Pahute Mesa and Oasis Valley

The microbial community of Oasis Valley is compositionally and phylogenetically distinct from all other study-area locations (**Figure 4**). Oasis Valley OTUs represented ~39% of the total OTUs sampled in this study, of which ~85% were observed only in Oasis Valley. The lack of microbial similarity between Pahute Mesa and Oasis Valley does not reflect regional groundwater-flow paths in the PMOV basin (**Figure 1**; 37, 38). However, the two Oasis Valley samples may not be representative of regional groundwater in this location as supported by the presence of many soil-associated and aerobic microbes. Moreover, the geochemistry of OV1 is distinct from other locations (**Figure 2**), with relatively high TOC concentrations (37 mg-C/L). The combination of microbial (**Figure 4**; taxonomic composition) and geochemical (**Figure 2**) data suggest that the two Oasis Valley samples are likely more impacted by land-surface-associated factors, owing to their relatively shallow depths and siting within the alluvium of the Amargosa River floodplain. Hence, these sites are probably disconnected hydrologically from the regional aquifer and conceptually can be regarded as hyporheic. This observation demonstrates the utility in supplementing geologic and hydrologic datasets with microbial community data for a comprehensive evaluation of groundwater samples.

The Pahute Mesa microbial community consists of wells that were sampled within the Pahute Mesa recharge area (PM3; PM6; PM11; PM12) and wells immediately downgradient of the recharge area (PM1; PM2; PM4; PM5; PM7–PM10; PM13–PM16) (**Figure 1**). Pahute Mesa communities are similar by NMDS and CCA ordinations (**Figure 4**), and many ‘common OTUs’ co-occur within Pahute Mesa (cluster NNSS 1) (**Figure 5**). Moreover, the two ecological processes dominating between Pahute Mesa communities include homogeneous dispersal and homogeneous selection (**Figure 6**). Microbial dispersal may be influences by the flow system at Pahute Mesa, which is dominated by high groundwater-flow velocities (39, 42). Groundwater flow in this area occurs through highly transmissive volcanic rocks, and tritium plumes from Pahute Mesa have only traveled less than 4 km beyond the NNSS borders since 2021, with an advective transport velocity of ~84 m/yr (25). Taken together, these observations suggest that microbial dispersal in the PMOV basin is possible at least within the Pahute Mesa recharge location.

5 Future Outlook

This study demonstrates that the regional-scale groundwater microbial community is a relevant data source; however, there were many limitations, as noted in Section 3.1, and future studies are needed~~This study indicates that the subsurface microbial community may be used to supplement geologic and hydrologic data in characterizing groundwater flow~~. For example, t~~T~~he findings in this study suggest there could be a detection limit to identifying similar microbial communities between recharge and discharge areas, which may be correlated with groundwater-flow rates, transmissivity, or time. Areas with relatively fast-flowing groundwater and high transmissivity are more likely to have similar microbial communities (e.g., Ash Meadows and Spring Mountains). In contrast, areas with low transmissivity (e.g., deep-carbonate flow towards Death Valley) may impede large microbial migration from recharge to discharge areas, such that deterministic factors and genetic evolution outcompete dispersal. Future studies that ~~can~~ quantify and correlate microbial community patterns with hydrogeologic factors and perturbations (e.g., groundwater-flow rate, transmissivity, time-scale) ~~will provide groundbreaking assessments towards the use of subsurface microbial communities for hydraulic connectivity~~are needed. Temporal replicates will also help establish the microbial community variation within each site and may reveal potential seasonal-, pumping-, or recharge-associated perturbations. Moreover, groundwater microbial community samples collected along a groundwater-flow path and at various depths can identify potential candidates for use as a microbial ‘tracer’ by providing insight into the persistence of specific microbial species. This spatial data also will help to confirm that ‘common OTUs’ in recharge and discharge zones can be used for helping characterize groundwater flow. While our study approach was cost-effective (amplicon sequencing), it was limited to microbial abundance information and the bacterial/archaeal community. Future studies can combine other approaches, such as obtaining microbial functions/activity and metabolites, to provide additional insight into hydrobiogeochemical dynamics and fluxes.

6 Conclusion

~~This study determined that~~The DVRFS microbial community patterns ~~have the potential to aid characterization of regional-scale groundwater flow~~were mostly consistent with the regional-scale groundwater-flow conceptualization. ~~Most samples were collected in recharge or discharge areas, and similarities or differences between samples were used to determine microbial consistency with groundwater-flow paths.~~ Overall, microbial communities within recharge and discharge areas connected by a flow path were similar (e.g., Spring Mountains and Ash Meadows), and location was the most significant variable differentiating between the communities. Notably, communities within and near the NNSS were similar, although groundwater flow from Pahute Mesa towards other NNSS sites is limited. Network analysis also demonstrated that 'common OTUs' clustered together by location, and in particular, clusters of 'common OTUs' found within and near the NNSS were connected via a cluster composed of Amargosa Valley OTUs. These OTUs represented a range of putative metabolisms, indicative of the mixing of various groundwater sources at Amargosa Valley, which is consistent with the most recent groundwater conceptualization of the DVRFS. Ecological null model analyses also identified locations in which communities were relatively similar due to deterministic and stochastic processes, and largely corroborated the other microbial community analyses conducted in this study. However, the microbial community patterns contradicted the hydraulic connection between Pahute Mesa (recharge) and Oasis Valley (discharge) in the PMOV groundwater basin, probably reflecting a lack of direct hydrologic connectivity between the upgradient regional flow system and discharge zone wells that were available for sampling. Overall, this exploratory study demonstrates that regional-scale groundwater microbial community patterns can be used to supplement geologic and hydrologic data in characterizing groundwater flow.

7 Acknowledgments

We thank Navarro-Interra for coordination with NNSS field sampling and on-site measurements. We thank Rachel Lindvall and Sarah Roberts (LLNL); Ron Hershey and John Healey (Desert Research Institute); Jeff Sanchez and Robert Zella (U.S. Geological Survey); and Robert Goodwin, Russ Shelton, Jeff Wurtz, and others from the UGTA Program for onsite logistical support at NNSS. Special thanks to Bill Wilborn, former UGTA sub-project director, for his interest and support of our work. Thanks to Terry Fisk, Richard Friese, Josh Hoines, Genne Nelson, and Kevin Wilson (U.S. National Park Service); Alisa Lembke and the Inyo County Planning Commission; John Bredehoeft and Michael King (The Hydronamics Group, LLC); Darrell Lacy, Jamie Walker, John Klenke, Levi Kryder and others from the Nye County Nuclear Waste Repository Program Office; John Healey, Brad Lyles, Patty Montgomery, Molly Devlin, and Brittany Kruger (DRI); and the Nature Conservancy for site access and logistical assistance with the Death Valley, Amargosa Valley and Oasis Valley sites. Thanks to the DRI Water Lab, T.C. Onstott (Princeton University) and Barbara Sherwood Lollar (University of Toronto) for supplemental chemistry analysis. We also thank Yongqin Jiao (LLNL) for invaluable discussions; and Jeff Kimbrel and Juliet Johnston (LLNL), the Bioinformatics Virtual Coordination Network (BVCN), especially Haley Sapers and Liz Suter, and the LLNL Data Science Institute Consulting Service (DSICS), especially Jason Bernstein and Aram Avila-Herrera, for their advice in analyzing the microbial communities.

8 Funding Sources

This research was partially funded by the Office of Biological and Environmental Research of the U.S. Department of Energy (DOE) as part of the Subsurface Biogeochemical Research Program under Work Proposal Number SCW1053, *Biogeochemistry of Actinides* and under DOE Contract Number SC0005306, *Radiochemically-Supported Microbial Communities*. Partial funding was also provided by the U.S. DOE Environmental Remediation Sciences Program under Grant Number FG02-07ER64406, *Characterizations of Microbial Communities in Subsurface Nuclear Blast Cavities of the Nevada Test Site*); NSF GoLife DEB 1441646, C*ollaborative research: Untangling the Deep Genealogy of Microbial Dark Matter,* and NSF EPSCOR OIA 1826734, *RII Track-2 FEC: Single Cell Genome-to-Phenome: Integrating Genome and Phenome Analyses of Individual Microbial Cells in Complex Microbiomes;* NASA Astrobiology Institute's *Life Underground* project (NNA13AA92A); NWRPO (10-023), *Characterization of Indigenous Microorganisms from Nye County Wells and Assessment of their Utility as Hydrologic Tracers;* The Hydrodynamic Group, *Characterization of Indigenous Microorganisms from Deep Wells and Assessment of their Utility as Hydrologic Tracers.* This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (IM#: LLNL-JRNL-823248). Thanks also to the Alfred P. Sloan Foundation's Deep Carbon Observatory (DCO) for sequencing and postdoc support (APS 2013-10-03) through their Census of Deep Life (CoDL) and Deep

Life Community (DLC) programs. JS was partially supported by a Nevada NASA Space Grant Consortium Graduate Fellowship (NNX15AI02H).

Author Contributions: NM analyzed the 16S rRNA gene amplicon sequences and compiled the geochemical data. DPM served as principal investigator for the projects under which the samples were collected. SHB, JCB, JS, and JCF contributed to the collection, processing, and extraction of the samples. All authors wrote the manuscript.

Competing Interest Statement: The authors have no competing interests.

References

1. S. P. Neuman, Trends, prospects and challenges in quantifying flow and transport through fractured rocks. *Hydrogeol. J.* **13**, 124–147 (2005).

2. M. Ye, L. Wang, K. F. Pohlmann, J. B. Chapman, Evaluating Groundwater Interbasin Flow Using Multiple Models and Multiple Types of Data. *Groundwater* **54**, 805–817 (2016).

3. F. H. Chapelle, The significance of microbial processes in hydrogeology and geochemistry. *Hydrogeol. J.* **8**, 41–46 (2000).

4. T. M. Flynn, *et al.*, Functional microbial diversity explains groundwater chemistry in a pristine aquifer. *BMC Microbiol.* **13**, 146 (2013).

5. C. Griebler, F. Malard, T. Lefébure, Current developments in groundwater ecology—from biodiversity to ecosystem function and services. *Curr. Opin. Biotechnol.* **27**, 159–167 (2014).

6. J. Tóth, A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res. 1896-1977* **68**, 4795–4812 (1963).

7. J. Tóth, Groundwater as a geologic agent: An overview of the causes, processes, and manifestations. *Hydrogeol. J.* **7**, 1–14 (1999).

8. R. A. Freeze, P. A. Witherspoon, Theoretical analysis of regional groundwater flow: 2. Effect of water-table configuration and subsurface permeability variation. *Water Resour. Res.* **3**, 623–634 (1967).

9. W. W. Woessner, *Groundwater-Surface Water Exchange* (GW Project, 2020) (April 4, 2022).

10. P. Cook, *Introduction to Isotopes and Environmental Tracers as Indicators of Groundwater Flow* (GW Project, 2020) (April 4, 2022).

11. J. F. Devlin, *Groundwater Velocity* (Groundwater Project, 2020) (April 4, 2022).

12. R. W. Harvey, Microorganisms as tracers in groundwater injection and recovery experiments: a review. *FEMS Microbiol. Rev.* **20**, 461–472 (1997).

13. M. A. Walvoord, *et al.*, Groundwater flow and geochemistry in the Southeastern San Juan Basin: Implications for microbial transport and activity. *Water Resour. Res.* **35**, 1409–1424 (1999).

14. S. Amalfitano, *et al.*, Groundwater geochemistry and microbial community structure in the aquifer transition from volcanic to alluvial areas. *Water Res.* **65**, 384–394 (2014).

15. Y. Zhang, *et al.*, Microbial Community Composition in Deep-Subsurface Reservoir Fluids Reveals Natural Interwell Connectivity. *Water Resour. Res.* **56**, e2019WR025916 (2020).

16. T. R. Ginn, *et al.*, Processes in microbial transport in the natural subsurface. *Adv. Water Resour.*, 26 (2002).

17. C. P. Gerba, I. L. Pepper, D. T. Newby, "Microbial Transport in the Subsurface" in *Environmental Microbiology*, Third, (Elsevier, 2015), pp. 319–337.

18. J. C. Stegen, X. Lin, A. E. Konopka, J. K. Fredrickson, Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* **6**, 1653–1664 (2012).

19. J. C. Stegen, *et al.*, Quantifying community assembly processes and identifying features that impose them. *ISME J.* **7**, 2069–2079 (2013).

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

21. S. B. Maamar, *et al.*, Groundwater Isolation Governs Chemistry and Microbial Community Structure along Hydrologic Flowpaths. *Front. Microbiol.* **6** (2015).

22. J. Zhou, D. Ning, Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiol. Mol. Biol. Rev.* **81** (2017).

23. L. A. Hug, *et al.*, Aquifer environment selects for microbial species cohorts in sediment and groundwater. *ISME J.* **9**, 1846–1856 (2015).

24. L. Fillinger, K. Hug, C. Griebler, Selection imposed by local environmental conditions drives differences in microbial community composition across geographically distinct groundwater aquifers. *FEMS Microbiol. Ecol.* **95** (2019).

25. K. J. Halford, T. R. Jackson, "Groundwater Characterization and Effects of Pumping in the Death Valley Regional Groundwater Flow System, Nevada and California, with Special Reference to Devils Hole" (U.S. Geological Survey, 2020).

26. U.S. DOE, "United States Nuclear Tests: July 1945 through September 1992" (U.S. Department of Energy, 2015).

27. M. T. Moreo, K. J. Halford, R. J. La Camera, R. J. Laczniak, "Estimated Ground-Water Withdrawals from the Death Valley Regional Flow System, Nevada and California, 1913-98" (Geological Survey (U.S.), 2003).

28. M. T. Moreo, L. Justet, "Update to the Ground-Water Withdrawals Database for the Death Valley Regional Ground-Water Flow System, Nevada and California, 1913-2003" (Geological Survey (U.S.), 2008).

29. I. J. Winograd, F. J. Pearson, Major carbon 14 anomaly in a regional carbonate aquifer: Possible evidence for megascale channeling, South Central Great Basin. *Water Resour. Res.* **12**, 1125–1143 (1976).

30. I. J. Winograd, A. C. Riggs, T. B. Coplen, The relative contributions of summer and cool-season precipitation to groundwater recharge, Spring Mountains, Nevada, USA. *Hydrogeol. J.* **6**, 77–93 (1998).

31. J. M. Thomas, *et al.*, “Geochemical and Isotopic Interpretations of Groundwater Flow in the Oasis Valley Flow System, Southern Nevada” (U.S. Department of Energy, 2003) https:/doi.org/10.2172/806667 (March 26, 2020).

32. J. M. Thomas, *et al.*, Using Water Chemistry, Isotopes and Microbiology to Evaluate Groundwater Sources, Flow Paths and Geochemical Reactions in the Death Valley Flow System, USA. *Procedia Earth Planet. Sci.* **7**, 842–845 (2013).

33. K. Anderson, S. Nelson, A. Mayo, D. Tingey, Interbasin flow revisited: The contribution of local recharge to high-discharge springs, Death Valley, CA. *J. Hydrol.* **323**, 276–302 (2006).

34. M. Bushman, “Contribution of Recharge Along Regional Flow Paths to Discharge at Ash Meadows, Nevada,” Brigham Young University. (2008).

35. W. R. Belcher, M. S. Bedinger, J. T. Back, D. S. Sweetkind, Interbasin flow in the Great Basin with special reference to the southern Funeral Mountains and the source of Furnace Creek springs, Death Valley, California, U.S. *J. Hydrol.* **369**, 14 (2009).

36. W. R. Belcher, D. S. Sweetkind, C. C. Faunt, M. T. Pavelko, M. C. Hill, “An update of the Death Valley regional groundwater flow system transient model, Nevada and California” (U.S. Geological Survey, 2017) (October 13, 2019).

37. M. Bushman, S. T. Nelson, D. Tingey, D. Eggett, Regional groundwater flow in structurally-complex extended terranes: An evaluation of the sources of discharge at Ash Meadows, Nevada. *J. Hydrol.* **386**, 118–129 (2010).

38. J. M. Fenelon, D. S. Sweetkind, R. J. Laczniak, “Groundwater Flow Systems at the Nevada Test Site, Nevada: A Synthesis of Potentiometric Contours, Hydrostratigraphy, and Geologic Structures” (U.S. Geological Survey, 2010) (February 13, 2020).

39. J. M. Fenelon, K. J. Halford, M. T. Moreo, “Delineation of the Pahute Mesa–Oasis Valley groundwater basin, Nevada” (U.S. Geological Survey, 2016).

40. S. T. Nelson, A. L. Mayo, The role of interbasin groundwater transfers in geologically complex terranes, demonstrated by the Great Basin in the western United States. *Hydrogeol. J.* **22**, 807–828 (2014).

41. S. R. Warix, L. K. Rademacher, Z. P. Meyers, M. D. Frisbee, Groundwater geochemistry and flow in the Spring Mountains, NV: Implications for the Death Valley Regional Flow System. *J. Hydrol.* **580**, 124313 (2020).

42. T. R. Jackson, J. M. Fenelon, R. L. Paylor, “Groundwater flow conceptualization of the Pahute Mesa–Oasis Valley Groundwater Basin, Nevada—A synthesis of geologic, hydrologic, hydraulic-property, and tritium data” (U.S. Geological Survey, 2021).

43. R. J. Laczniak, *et al.*, “Ground-water discharge determined from estimates of evapotranspiration, Death Valley regional flow system, Nevada and California” (2001) https:/doi.org/10.3133/wri014195.

44. D. P. Moser, *et al.*, “Radiochemically-supported microbial communities. A potential mechanism for biocolloid production of importance to actinide transport” (Desert Research Institute, Las Vegas, NV (United States), 2014) https:/doi.org/10.2172/1213048 (May 25, 2021).

45. I. M. Farnham, The Nature and State of Groundwater Contamination at the Nevada National Security Site: What Have We Learned from Decades of Groundwater Analysis? in *WM2020*, (2020).

46. B. J. Callahan, *et al.*, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

47. E. Bolyen, *et al.*, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

48. S. Louca, F. Mazel, M. Doebeli, L. W. Parfrey, A census-based estimate of Earth’s bacterial and archaeal diversity. *PLOS Biol.* **17**, e3000106 (2019).

49. S. I. Glassman, J. B. H. Martiny, Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. *mSphere* **3**, e00148-18 (2018).

50. L. Joos, *et al.*, Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. *BMC Genomics* **21**, 733 (2020).

51. S. Moossavi, F. Atakora, K. Fehr, E. Khafipour, Biological observations in microbiota analysis are robust to the choice of 16S rRNA gene sequencing processing algorithm: case study on human milk microbiota. *BMC Microbiol.* **20**, 290 (2020).

52. J. T. Nearing, G. M. Douglas, A. M. Comeau, M. G. I. Langille, Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. *PeerJ* **6**, e5364 (2018).

53. A. Prodan, *et al.*, Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLOS ONE* **15**, e0227434 (2020).

54. C. Quast, *et al.*, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).

55. P. Yilmaz, *et al.*, The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* **42**, D643–D648 (2014).

56. N. A. Bokulich, *et al.*, Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 90 (2018).

57. N. Bokulich, M. Robeson, B. Kaehler, M. Dillon, bokulich-lab/RESCRIPt: 2020.6.1 (2020) https:/doi.org/10.5281/zenodo.3945228 (July 31, 2020).

58. P. J. McMurdie, S. Holmes, phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217 (2013).

59. T. C. Hsieh, K. H. Ma, A. Chao, iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods Ecol. Evol.* **7**, 1451–1456 (2016).

60. S. W. Kembel, *et al.*, Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464 (2010).

61. C. Martino, *et al.*, A Novel Sparse Compositional Technique Reveals Microbial Perturbations. *mSystems* **4** (2019).

62. G. B. Gloor, J. M. Macklaim, V. Pawlowsky-Glahn, J. J. Egozcue, Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* **8** (2017).

63. J. D. Silverman, A. D. Washburne, S. Mukherjee, L. A. David, A phylogenetic transform enhances analysis of compositional microbiota data. *eLife* **6**, e21887 (2017).

64. F. Dini-Andreote, J. C. Stegen, J. D. van Elsas, J. F. Salles, Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci.* **112**, E1326–E1332 (2015).

65. R. E. Danczak, *et al.*, Ecological Assembly Processes Are Coordinated between Bacterial and Viral Communities in Fractured Shale Ecosystems. *mSystems* **5** (2020).

66. O. J. Hardy, Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *J. Ecol.* **96**, 914–926 (2008).

67. D. M. Raup, R. E. Crick, Measurement of Faunal Similarity in Paleontology. *J. Paleontol.* **53**, 1213–1227 (1979).

68. J. R. Bray, J. T. Curtis, An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecol. Monogr.* **27**, 325–349 (1957).

69. J. M. Chase, N. J. B. Kraft, K. G. Smith, M. Vellend, B. D. Inouye, Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* **2**, art24 (2011).

70. Z. D. Kurtz, *et al.*, Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLOS Comput. Biol.* **11**, e1004226 (2015).

71. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal* **Complex Systems**, 10 (2006).

72. A. Clauset, M. E. J. Newman, C. Moore, Finding community structure in very large networks. *Phys. Rev. E* **70**, 066111 (2004).

73. V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).

74. L. Yan, *et al.*, Groundwater bacterial communities evolve over time in response to recharge. *Water Res.* **201**, 117290 (2021).

75. D. Abiriga, A. Jenkins, K. Alfsnes, L. S. Vestgarden, H. Klempe, Spatiotemporal and seasonal dynamics in the microbial communities of a landfill-leachate contaminated aquifer. *FEMS Microbiol. Ecol.* **97**, fiab086 (2021).

76. G. Pilloni, *et al.*, Dynamics of Hydrology and Anaerobic Hydrocarbon Degrader Communities in A Tar-Oil Contaminated Aquifer. *Microorganisms* **7**, 46 (2019).

77. J. R. Harrill, M. S. Bedinger, "Death Valley Regional Groundwater Flow System, Nevada and California—Hydrogeologic Framework and Transient Groundwater Flow Model" (U.S. Geological Survey, 2010) (March 11, 2020).

78. S. Pavoine, A. Gasc, M. B. Bonsall, N. W. H. Mason, Correlations between phylogenetic and functional diversity: mathematical artefacts or true ecological and evolutionary processes? *J. Veg. Sci.* **24**, 781–793 (2013).

79. N. G. Swenson, *Functional and Phylogenetic Ecology in R*, 1st Ed. (Springer, 2014).

80. B. Sandel, Richness-dependence of phylogenetic diversity indices. *Ecography* **41**, 837–844 (2018).

81. S. Weiss, *et al.*, Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5** (2017).

82. R. J. Wright, M. G. I. Langille, T. R. Walker, Food or just a free ride? A meta-analysis reveals the global diversity of the Plastisphere. *ISME J.* **15**, 789–806 (2021).

83. J. Wang, *et al.*, Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J.* **7**, 1310–1321 (2013).

84. C. Liu, *et al.*, Long-term nitrogen addition affects the phylogenetic turnover of soil microbial community responding to moisture pulse. *Sci. Rep.* **7** (2017).

85. E. D. Beaton, *et al.*, Local and Regional Diversity Reveals Dispersal Limitation and Drift as Drivers for Groundwater Bacterial Communities from a Fractured Granite Formation. *Front. Microbiol.* **7** (2016).

86. R. E. Danczak, M. D. Johnston, C. Kenah, M. Slattery, M. J. Wilkins, Microbial Community Cohesion Mediates Community Turnover in Unperturbed Aquifers. *mSystems* **3** (2018).

87. S. Uprety, *et al.*, The Effect of the 2015 Earthquake on the Bacterial Community Compositions in Water in Nepal. *Front. Microbiol.* **8** (2017).

88. H. Kim, *et al.*, Impact of earthquake on the communities of bacteria and archaea in groundwater ecosystems. *J. Hydrol.* **583**, 124563 (2020).

89. S. Morimura, X. Zeng, N. Noboru, T. Hosono, Changes to the microbial communities within groundwater in response to a large crustal earthquake in Kumamoto, southern Japan. *J. Hydrol.* **581**, 124341 (2020).

90. C. Lumban Gaol, L. Ganzer, S. Mukherjee, H. Alkan, Investigation of clogging in porous media induced by microorganisms using a microfluidic application. *Environ. Sci. Water Res. Technol.* **7**, 441–454 (2021).

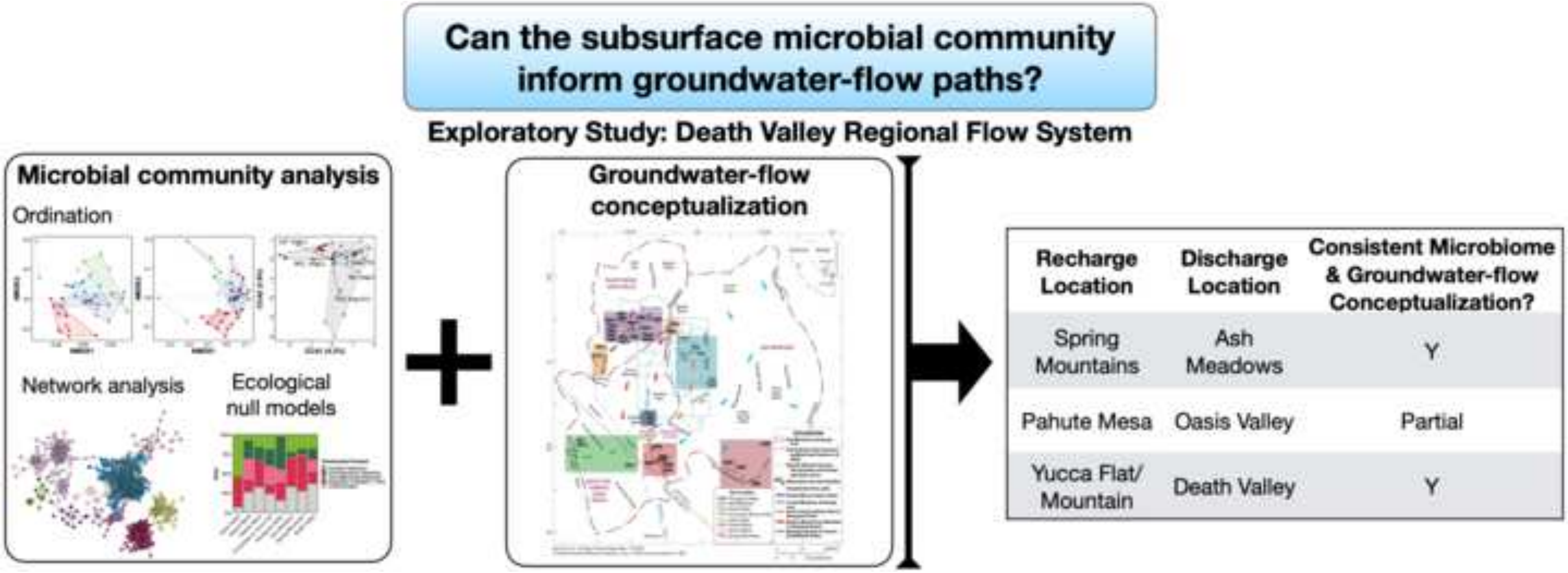
91. P. Baveye, P. Vandevivere, B. L. Hoyle, P. C. DeLeo, D. S. de Lozada, Environmental Impact and Mechanisms of the Biological Clogging of Saturated Soils and Aquifer Materials. *Crit. Rev. Environ. Sci. Technol.* **28**, 123–191 (1998).

92. C. C. Faunt, J. B. Blainey, M. C. Hill, F. A. D'Agnese, G. M. O'Brien, "Death Valley regional groundwater flow system, Nevada and California-Hydrogeologic framework and transient groundwater flow model" (2010) https:/doi.org/10.3133/pp1711.

93. J. Bredehoeft, M. King, Potential contaminant transport in the regional Carbonate Aquifer beneath Yucca Mountain, Nevada, USA. *Hydrogeol. J.* **18**, 775–789 (2010).

94. P. A. Cutillo, J. D. Bredehoeft, Estimating Aquifer Properties from the Water Level Response to Earth Tides. *Groundwater* **49**, 600–610 (2011).

Graphical Abstract



Can the subsurface microbial community inform groundwater-flow paths?

Exploratory Study: Death Valley Regional Flow System

Microbial community analysis

Ordination

Network analysis

Ecological null models

+

Groundwater-flow conceptualization

Recharge Location	Discharge Location	Consistent Microbiome & Groundwater-flow Conceptualization?
Spring Mountains	Ash Meadows	Y
Pahute Mesa	Oasis Valley	Partial
Yucca Flat/ Mountain	Death Valley	Y

Highlights

1. The subsurface microbial community is consistent with known hydraulic connections.
2. Location plays a major role in microbial community variation.
3. Network analysis of common microbes corroborated recharge and discharge areas.
4. Ecological null models provided insight into community assembly patterns.

Subsurface Microbial Communities as a Tool for Characterizing Regional-Scale Groundwater Flow

Nancy Merino[1*], Tracie R. Jackson[2], James H. Campbell[3], Annie B. Kersting[1], Joshua Sackett[4,5,6], Jenny C. Fisher[4,7], James C. Bruckner[4], Mavrik Zavarin[1], Scott D. Hamilton-Brehm[4,8], Duane P. Moser[4,5*]

[1] Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, 94550, United States
[2] Nevada Water Science Center, U.S. Geological Survey, Boulder City, NV, 89005, United States
[3] Department of Natural Sciences, Northwest Missouri State University, Maryville, MO, 64468, United States
[4] Division of Earth and Ecosystems Sciences, Desert Research Institute, Las Vegas, NV, 89119, United States
[5] Division of Hydrologic Sciences, Desert Research Institute, Las Vegas, NV, 89119, United States
[6] School of Life Sciences, University of Nevada, Las Vegas, NV, 89154, United States
[7] Biology Department, Indiana University Northwest, Gary IN, 46408, United States (present address)
[8] Department of Microbiology, Southern Illinois University Carbondale, Carbondale, IL, 62901, United States (present address)

***Correspondence:**
Nancy Merino
Lawrence Livermore National Lab
7000 East Ave, B361/R1940
Livermore, CA 94550
925-424-7605
merino4@llnl.gov

Duane Moser
Division of Hydrologic Sciences
Desert Research Institute
Las Vegas, NV, 89119
702-862-5534
Duane.Moser@dri.edu

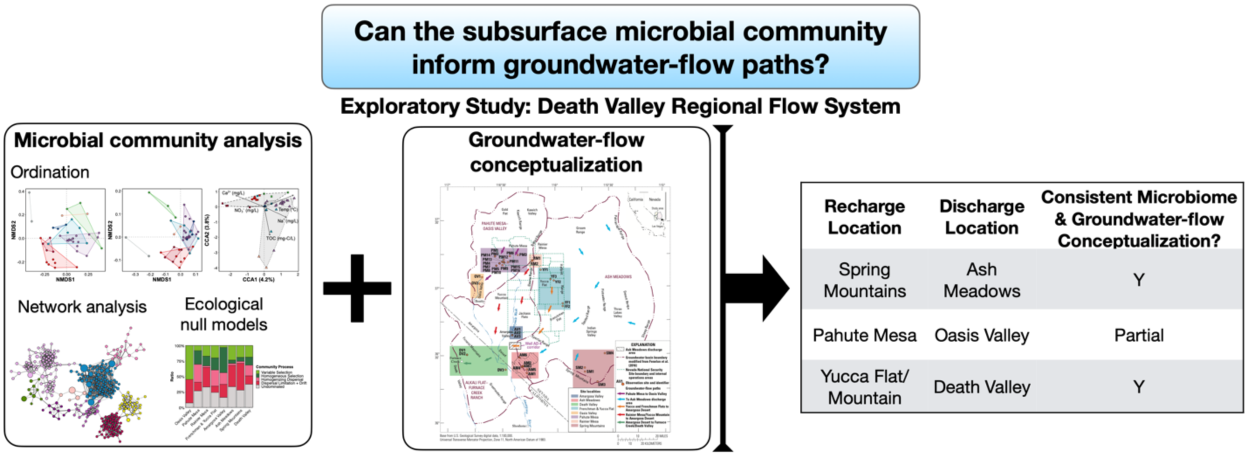
Keywords (max 6): fractured-rock aquifer, subsurface microbial ecology, microbial community assembly, regional-scale aquifer microbial community, deep biosphere

Abbreviations not standard to the field: Alkali Flat–Furnace Creek Ranch (AFFCR), Ash Meadows (AM), Death Valley Regional Flow System (DVRFS), Nevada National Security Site (NNSS), Pahute Mesa–Oasis Valley (PMOV), Underground Test Area (UGTA)

Highlights

1. The subsurface microbial community is consistent with known hydraulic connections.
2. Location plays a major role in microbial community variation.
3. Network analysis of common microbes corroborated recharge and discharge areas.
4. Ecological null models provided insight into community assembly patterns.

Graphical abstract



Recharge Location	Discharge Location	Consistent Microbiome & Groundwater-flow Conceptualization?
Spring Mountains	Ash Meadows	Y
Pahute Mesa	Oasis Valley	Partial
Yucca Flat/ Mountain	Death Valley	Y

Abstract

Subsurface microbial community distribution patterns are influenced by biogeochemical and groundwater fluxes and may inform hydraulic connections along groundwater-flow paths. This study examined the regional-scale microbial community of the Death Valley Regional Flow System and evaluated whether subsurface communities can be used to identify groundwater-flow paths between recharge and discharge areas. Samples were collected from 36 sites in three groundwater basins: Pahute Mesa–Oasis Valley (PMOV), Ash Meadows (AM), and Alkali Flat–Furnace Creek Ranch (AFFCR). Microbial diversity within and between communities varied by location, and communities were separated into two overall groups that affiliated with the AM and PMOV/AFFCR basins. Network analysis revealed patterns between clusters of common microbes that represented groundwaters with similar geochemical conditions and largely corroborated hydraulic connections between recharge and discharge areas. Null model analyses identified deterministic and stochastic ecological processes contributing to microbial community assemblages. Most communities were more different than expected and governed by dispersal limitation, geochemical differences, or undominating processes. However, certain communities from sites located within or near the Nevada National Security Site were more similar than expected and dominated by homogeneous dispersal or selection. Overall, the (dis)similarities between the microbial communities of DVRFS recharge and discharge areas supported previously documented hydraulic connections between: (1) Spring Mountains and Ash Meadows; (2) Frenchman and Yucca Flat and Amargosa Desert; and (3) Amargosa Desert and Death Valley. However, only a portion of the flow path between Pahute Mesa and Oasis Valley could be supported by microbial community analyses, likely due to well-associated artifacts in samples from the two Oasis Valley sites. This study demonstrates the utility of combining microbial data with hydrologic, geologic, and water-chemistry information to comprehensively characterize groundwater systems, highlighting both strengths and limitations of this approach.

1 Introduction

Groundwater plays an essential role in providing water for societal needs and maintaining ecosystem health. As the global hydrologic cycle shifts with climate change, groundwater quality and quantity are increasingly affected, necessitating sustainable groundwater-management practices. These practices require a regional-scale understanding of the groundwater system. Regional-scale groundwater flow is typically characterized using geologic, water-level, water-chemistry, and aquifer-testing data (1, 2). However, groundwater systems also are impacted by and serve as distinct habitats for microbial communities (3–5). Thus far, groundwater microbial communities remain an underutilized data source to understand hydraulic connections between recharge and discharge areas.

Groundwater flow is governed largely by recharge and discharge rates and the hydraulic properties of the saturated rocks. Groundwater-flow paths are derived from hydraulic gradients between areas of recharge (e.g., precipitation in high topographic regions) and discharge (e.g., springs in low topographic regions) (6–8). Geologic units also exert control on groundwater-flow pathways. For example, the juxtaposition of permeable geologic units against impermeable units downgradient results in a hydraulic barrier to groundwater flow. The sources of groundwater and travel times along groundwater-flow paths can be informed by water-chemistry data (9–11).

The groundwater microbial community also can be used to characterize groundwater systems and validate groundwater-flow paths. This concept dates back to the 1890s when studies began using microbes as tracers to understand subsurface pathogen transport and to delineate groundwater-flow paths through injection and recovery experiments (12). With the advent of the genomics era and the reducing costs of high-throughput sequencing, it may now be possible to use the subsurface microbial community to aid in groundwater characterization. Unattached (planktonic) microbes may be transported through aquifers along groundwater-flow paths through pore spaces and fractures (13–15). Several abiotic and biotic factors impact subsurface microbial movement (16, 17), including groundwater-flow rates, hydraulic barriers, pore size, microbial adhesion processes, and the physiological state of microbial cells.

Subsurface microbial transport may impact microbial community assemblages through space and time and could result in more similar or dissimilar communities. Microbial community assembly can be evaluated with ecological null models (18–20) (see **Table S1** for terms and definitions used for ecological null models) that may inform groundwater characterization. For example, high microbial-dispersal rates (i.e., ‘homogenizing dispersal’) may provide a signature for tracking hydraulic connections through the co-occurrence of microorganisms in recharge and discharge areas. In contrast, hydraulic barriers may prevent or impede microbial transport and mixing (‘dispersal limitation’), such as in the hard-rock aquifers of Brittany (France) (21), causing microbial communities to diverge over time or along flow paths through stochastic ecological processes (22). Biogeochemical and redox conditions also drive community assembly, selecting for adaptable microorganisms (21, 23, 24), and are known as deterministic processes that occur when conditions are consistent (‘homogeneous selection’) or fluctuating (‘variable selection’). Consistent environmental conditions lead to less divergence in the community, whereas fluctuating conditions can perturb microbial abundances and diversity. Ecological null models can identify the dominating ecological processes that impact microbial community assembly between two communities. Along with other analytical approaches to analyze microbial community data, the groundwater microbial community could be an advantageous tool to supplement other geologic and hydrologic datasets to provide a comprehensive understanding of regional groundwater flow.

This study assesses the potential of using subsurface microbial community patterns as a supplemental dataset to characterize regional-scale groundwater flow. The Death Valley Regional Flow System (DVRFS) microbial community was used as a test case because, over the past few decades, hundreds of boreholes and wells have been drilled in the DVRFS to monitor water levels and water quality. Thus, the DVRFS provides a unique opportunity to identify novel microbial community patterns on a regional-scale and to determine whether microbial community patterns are consistent with regional-scale groundwater flow-paths (25). The aims of this study were to (1) evaluate the groundwater bacterial and archaeal community distribution patterns with alpha- and beta-diversity analyses and network analysis, (2) identify the assembly processes dominating the community patterns using ecological null models, and (3) use the microbial community patterns and assembly processes to assess any (dis)similarities along groundwater-flow paths. To achieve these aims, we collected samples from recharge and discharge areas and hypothesized that microbial (dis)similarities, co-occurrences, and assembly patterns may infer and support the most recent groundwater conceptualization (25).

2 Materials and Methods

2.1 Site Description

The DVRFS is a 100,000 km^2 region that is part of the Great Basin physiographic province in southern Nevada and California (**Figure 1**). The study area includes three groundwater basins in the DVRFS: Pahute Mesa–Oasis Valley (PMOV), Ash Meadows (AM), and Alkali Flat–Furnace Creek Ranch (AFFCR) (**Figure 1**) (25). The Nevada National Security Site (NNSS, formerly Nevada Test Site) intersects these three groundwater basins. Land-surface elevations range from –86 to 3,600 m relative to sea level. Principal geologic units are grouped into four categories: Paleozoic carbonate rocks, Tertiary volcanic rocks, Cenozoic basin fill, and undifferentiated, low-permeability siliciclastic and granitic rocks (25).

Long-term water-level and water-quality monitoring have occurred in the study area. A total of 828 underground nuclear tests were detonated beneath the NNSS from 1951–1992 (26). Underground nuclear testing prompted a long-term U.S. Department of Energy environmental management program, known as the Underground Test Area (UGTA) Activity, to assess and monitor potential groundwater transport of radionuclides. Extensive hydrogeologic investigations (25, 27, 28) also were conducted by the U.S. Geological Survey to evaluate water levels and groundwater-withdrawal rates. Several groundwater studies (25, 29–41) have evaluated flow paths in the DVRFS.

This study evaluates groundwater-flow paths using the most recent groundwater conceptualization of the DVRFS (25, 42) (**Figure 1**). The five regional groundwater-flow paths evaluated in this study are: (1) Pahute Mesa to Oasis Valley; (2) Spring Mountains to Ash Meadows discharge area; (3) Frenchman and Yucca Flat to Amargosa Desert; (4) Rainier Mesa/Yucca Mountain to Amargosa Desert; and (5) Amargosa Desert to Death Valley (25) (**Figure 1**). In the PMOV basin, Pahute Mesa is the primary recharge area contributing to discharge from Oasis Valley (42). In the AM basin, most recharge is derived from the Spring Mountains and Sheep Range, with lesser recharge amounts from the Groom, Pahranagat, Desert, Pintwater, and Spotted Ranges. Groundwater from the Spring Mountains and other ranges in the AM basin moves southward and westward to high-volume springs of the Ash Meadows discharge area, the largest oasis in the Mojave Desert (43). Groundwater from Yucca and Frenchman Flats moves into the Amargosa Desert through the well AD-4 corridor north of Ash Meadows discharge area (**Figure 1**). The well AD-4 corridor hydraulically connects transmissive carbonate rocks in the AM basin with transmissive basin fill in the Amargosa Desert (see **Section 4.3** for more details). Groundwater from Rainier Mesa and the Yucca

Mountain area moves southward into the Amargosa Desert. Most groundwater in the Amargosa Desert moves westward and discharges at Furnace Creek in Death Valley.

2.2 Sample Collection and Geochemical Analysis

Between 2008–2014, a total of 42 samples were collected from wells, mine vent holes, or springs. These samples were collected from 36 sites across the DVRFS (**Figure 1, Table S2**). The number of sites per location varied (**Table S2**) based on accessibility [Amargosa Valley (sites [n] = 1, samples [s] = 3), Ash Meadows discharge area (n = 7, s = 7), Death Valley (n = 3, s = 3), Frenchman Flat (n = 2, s = 2), Oasis Valley (n = 2, s = 2), Pahute Mesa (n = 12, s = 16), Rainier Mesa (n = 2, s = 2), Spring Mountains (n = 4, s = 4), and Yucca Flat (n = 3, s = 3)]. The location Amargosa Valley is within the Amargosa Desert. For well samples, groundwater generally was pumped at high rates (hundreds of L/min) for hours-to-days prior to collection of geochemical and microbial samples to minimize wellbore artifacts. In a few cases, well water samples were obtained by lower rate "jack pumps" (in-line submerged pump) or from static water columns with discrete sampler deployments (i.e., "bailers"). The mine water samples (U12n-10 [RM1] and U12n-vent2 [RM2]) were collected from flooded workings by bailer from hundreds of meters overhead via vertical ventilation holes. Spring samples from Ash Meadows and Spring Mountains were collected by pumping with a peristaltic pump and sterile platinum-cured silicone tubing (Masterflex LS-15) inserted as deeply into the spring orifice as possible via a 24' (7.3m) telescoping probe to minimize surface associated influences. For the site in the Amargosa Valley (4PD), water was pumped continuously over a few weeks, with samples collected on days 1, 9, and 23. At four sites on Pahute Mesa (ER-20-8 [PM15, PM16], ER-EC-13 [PM10, PM5], ER-EC-15 [PM4, PM7, PM8], and PM-3 [PM1, PM2]), samples were collected via pumping from discrete piezometers screened at multiple different depths, as listed in **Table S2**, because there may be microbial community differences with depth (44).

Cells for DNA extraction were collected on 0.22 µm Sterivex^{TM} filters (EMD Millipore, U.S.A.) with the filtrate retained for geochemical analysis and the filters transported on dry ice and stored at -80ºC. The DNA extraction and sequencing methods are described in **Supplementary Methods section 1.1**. The method and volume of filtration for cells varied depending on the circumstance: (1) ranging from tens to hundreds of liters collected via a pressurized sampling manifold (~0.5 bar) at the wellhead for most pumped samples; or (2) two to four liters filtered offsite for bailed samples or samples containing radioactive elements. Nevares Deep Well 2 (DV1) is artesian and was sampled at ambient pressure (0.4 bar) via a dedicated sampling port at the wellhead. Most geochemical analyses reported here were processed as part of the UGTA groundwater sampling and analysis program (45). Otherwise, the geochemical analyses were performed at the DRI Water Analysis Laboratory (Reno, NV) or at Princeton University T.C.O. (Princeton, NJ) (44), as described in **Supplementary Methods section 1.2**.

2.3 Statistical Analyses of the Geochemical Data

The geochemical dataset (**Table S2, color scheme in Table S3**) used in this paper contains information on location, depth, and geochemical measurements. Several geochemical parameters contained missing and left-censored (i.e., below detection limit) values. To reduce the number of missing values, some missing values were substituted with geochemical data collected from the same site on a different sampling date. Substituted data were obtained from the UGTA Activity Environmental Management Project database (45). Although many of these values were measured between the 1970s to 1990s, the geochemical parameters have remained generally stable over the past few decades (45). A piper diagram was created using the Geochemist's Workbench Release 11.0.8 (Aqueous Solutions LLC, USA) with a subset of the geochemical data (Ca^{2+}, Cl^{-}, K^{+}, Mg^{2+}, Na^{+}, SO_4^{2-}, HCO_3^{-}). Principal component analysis (PCA) also was used to visualize the sample variation through dimensionality reduction, as described in **Supplementary Methods section 1.3**.

2.4 Quality control of the 16S rRNA amplicon sequences

Raw sequences were processed through QIIME 1 and DADA2 v1.12.1 (46), as described in **Supplementary Methods section 1.4**. Because there were two different sequencing batches, samples were separated by batch and processed through DADA2 separately to account for potential error rate model differences. This resulted in amplicon sequence variants (ASVs) unique to each batch, which may have been caused by technical or biological differences (e.g., sample preparation, human processing, PCR amplification, sequencing errors, and multiple 16S rRNAs per cell). However, many ASVs from both batches were closely-related, as determined by inspection of a phylogenetic tree and sequence alignment, and were subsequently clustered into OTUs in QIIME 2 (47) at the 97% identity level with the function vsearch cluster-features-de-novo. OTU clustering at the 97% identity level can provide comparable overall results as ASVs on a broad-scale (48–53). Taxonomy was assigned to the clustered OTUs using the QIIME 2 function feature-classifier classify-sklearn with a pretrained-classifier SILVA v138 database for OTUs from 515F/806R region of 16S rRNA sequences (54–57). The OTU and taxonomy table were then imported to phyloseq v1.16.2 (58), and contaminant OTUs removed, as described in **Supplementary Methods section 1.4**. Based on rarefaction curves generated using the R package iNEXT (59), sufficient sequencing depth was achieved for each sample (**Figure S1**).

2.5 Microbial community analyses

Alpha diversity (diversity within samples) indices were computed with phyloseq (function estimate_richness) and picante v1.8.2 (60) using a phylogenetic tree generated as described in **Supplementary Methods section 1.5**. The means between each categorical variable were evaluated using the non-parametric Mann-Whitney test with false discovery rate p-value adjustment. Phylogenetic relatedness also was determined by calculating the standardized effect size (SES) of Faith's PD (phylogenetic diversity), MPD (mean pairwise distance), and MNTD (mean nearest neighbor phylogenetic distance) (see **Table S1** for definitions), as described in Supplementary Methods.

Beta diversity (diversity between samples) was evaluated using five approaches for comparison: (1) non-metric multidimensional scaling (NMDS) ordination on unweighted and weighted UniFrac distances, (2) DEICODE robust Aitchison principal-component analysis (RPCA) (61) in QIIME 2 with auto-RPCA, (3) centered-log ratio (CLR) PCA (62), (4) phylogenetic isometric-log ratio (PhILR) PCA (63), and (5) canonical correspondence analysis (CCA). Permutational analysis of variance (PERMANOVA) marginal tests with 999 permutations were conducted with R function adonis2 on the NMDS ordinations and homogeneity was checked with R function betadisper.

Deterministic and stochastic ecological processes were evaluated using the β-nearest taxon index (βNTI) and Raup-Crick (Bray-Curtis) (RCBC) metrics, following null model analyses (18–20, 22, 64, 65) (see **Table S1** for definitions). Only for these metrics, communities were rarefied to the lowest sequencing depth. βNTI is used to identify deterministic processes by quantifying phylogenetic turnover. First, species were randomly shuffled across the tips of the phylogeny, also known as the between-community mean-nearest-taxon-distance (βMNTD) metric, to obtain a null distribution after repeating the shuffling 999 times. Subsequently, the significance is evaluated by calculating the difference between the observed βMNTD and mean null distribution, with βNTI < –2 (less phylogenetic turnover; homogeneous selection) and βNTI > 2 (more phylogenetic turnover; variable selection) as significant (18). Insignificant βNTI suggests that the observed compositional differences are due to stochastic ecological processes (66), further determined by RCBC. RCBC is based off of community composition and OTU abundances, and follows Raup–Crick (67) to probabilistically assemble local communities, followed by quantification with Bray–Curtis (68) to obtain a null distribution after repeating 9,999 times (19, 65, 69). The RCBC null distribution is then standardized between -1 and +1 and significant values determined as RCBC < –0.95 (homogeneous dispersal) and RCBC > 0.95 (dispersal limitation). Ecological null models also assumes that phylogenetic relatedness is associated with ecological niche differences (18, 24, 64), which was confirmed for the DVRFS microbial community using Mantel correlograms, following Dini-Andreote *et al.* (2015). A significant ($P < 0.05$) positive correlation was observed across relatively short phylogenetic distances (**Figure S2**).

Network analysis was performed using SPIEC-EASI (SParse InversE Covariance Estimation for Ecological Association Inference; 86) on commonly observed OTUs among all samples, excluding Oasis Valley samples, to ensure robust results. ‘Common OTUs’ were defined as those with low coefficient of variation across samples (≤ 3; community standard deviation divided by community mean) (58) and observed at least two times in more than two samples (466 total OTUs). The glasso (sparse inverse covariance selection) and Meinshausen–Buhlmann (MB; neighborhood selection) method were evaluated with the Stability Approach to Regularization Selection (StARS) and bounded-StARS (bstars) selection criteria. Similar results were observed for both selection criteria. Since a wider degree distribution range was observed for glasso (**Figure S3**), the glasso bstars approach was chosen for subsequent analysis in igraph (71). OTUs with weak connection (edge weight < 0.02) were removed. Various clustering algorithms were performed to define clusters within the network, and the unsupervised Louvain clustering algorithm (72, 73) was chosen based on the highest modularity score of 0.694 amongst all the algorithms. The 16 clusters were then manually inspected; clusters with only two connections were removed and nearby connected clusters with the same overall location associated with OTUs were combined. Cytoscape (https://cytoscape.org/) was used for visualization and network analysis.

2.6 Data Availability

The 16S rRNA gene amplicon sequences were deposited at GenBank under the accession KEWZ00000000. The version described in this paper is the first version, KEWZ01000000. Scripts used to analyze and create figures of the geochemistry and microbial community are available at https://github.com/LLNL/2022_DVRFS_microbiome.

3 Results

3.1 Sample description, study limitations, and biases

A total of 42 samples were collected from 36 sites (**Figure 1, Table S2–S3**) that encompass the AM, AFFCR, and PMOV groundwater basins. Samples were collected opportunistically based on site availability, and most samples were collected from recharge and discharge areas within their respective groundwater basins. Due to our sampling limitations, the influence of seasonality, changes in recharge events, or hydraulic dynamics could not be determined on the temporal dynamics of subsurface microbial communities. This has been documented in previous studies of the Hainich Critical Zone Exploratory fractured aquifer system (74), a landfill-leachate contaminated aquifer (75), and a tar-oil contaminated aquifer (76). However, the very long residence times and deep flow paths of water in this system would argue against short-term effects other than from pumping (25). There also are sample biases in this study, including the number of samples per location, sample collection method, and two sequencing batches. The specifics of these biases are described in **section 2.2** and were considered when analyzing the microbial community data.

3.2 Geochemistry

The major-ion chemistry of study sites were grouped into three broad categories: Ca-Mg-HCO_3, Na-HCO_3, and NaCl dominated waters (**Figure 2A**), which were determined from a piper diagram (**Figure**

S4) and concur with previous reports (32, 35, 36, 41, 77). Principal component analysis (PCA) indicates that Ca-Mg-HCO_3-type waters cluster together, whereas the Na-HCO_3-type and NaCl-type waters group together. Notably, all 42 samples distinctly clustered into carbonate and volcanic rock types (**Figure 2A**). Rock types were assigned based on the predominant rock type found at the depth sampled. Most samples also clustered by location, with the largest variation observed for samples collected from Yucca Flat, Rainier Mesa, and Oasis Valley (**Figure 2B**). The water chemistry of Yucca Flat samples (YF1–YF3) is highly variable because of different hydrogeologic settings. YF1 (UE-2ce-WW-Nash), YF2 (U-3cn5-Bilby), and YF3 (UE-3E-4-Aleman) were sampled from wells completed in carbonate rock, carbonate and volcanic rocks, and volcanic rocks, respectively. Rainer Mesa samples (RM1 [U12n-10], RM2 [U12n-vent2]) likely do not cluster together because of the water geochemistry. RM1 has higher SO_4^{2-}, Na^+, and SiO_2^- concentrations (influenced by a local granite intrusion and pyrite oxidation leading to high sulfate) compared to RM2 (**Figure 2C, Table S2**). Oasis Valley samples (OV1, OV2) likely do not cluster together because of higher SO_4^{2-} and total organic carbon (TOC) concentrations observed in OV1.

3.3 Description of groundwater bacterial and archaeal groundwater microbial community

The groundwater microbial community consisted of 5,124 unique operational taxonomic units (OTUs) that were identified from 1,267,990 reads (**Tables S4–S5**). The final number of reads for OTUs per sample ranged from 2,790–123,420 (average = 30,190) (**Table S4**). Because a higher number of reads was observed for samples sequenced in batch 2 (average = 90,352; batch size = 7) than batch 1 (average = 18,158; batch size = 35), subsequent analyses considered sequence batch effects. Among the OTUs identified, only 43 OTUs were present in more than ten sites. The most commonly observed taxa include those from genera *Pseudomonas*, *Curvibacter*, *Phenylobacterium, Thiobacillus,* and *Hydrogenophaga*. Although community structure varied from site to site, the overall community was composed mainly of Bacteria (49–100%), and most sites were dominated by the phyla Proteobacteria (average = 42%), Bacteroidota (15%), Firmicutes (13%), and Desulfobacterota (11%) (**Figure S5**). Within Proteobacteria, only Alphaproteobacteria and Gammaproteobacteria classes were observed. The archaeal relative abundances averaged 7% for all sites, and the highest abundances (51%) were observed in sample DV3 (Inyo-BLM 1).

Diversity within communities (alpha diversity) largely varied based on location and location type (**Figure 3 and S6**). Although there were significant differences (P < 0.05) between the means of some categories (e.g., springs and wells) (**Figure S6**), the standardized effect size (SES) of all alpha diversity metrics suggests that the differences could be affected by species richness (number of OTUs in each sample) (**Figure 3**) (78–80). SES also can describe the phylogenetic relatedness within communities, and for most communities sampled, the SES-PD and SES-MNTD were negative and significantly different than the null model (P < 0.05), indicating that OTUs within each community were more closely related than expected by chance.

3.4 Microbial community variability by location

Differences in community composition between samples (beta diversity) was mostly explained by location. PERMANOVA by location- and geochemical-type variables revealed that location significantly (PERMANOVA P = 0.003; ANOSIM P = 0.001) explained the most variance of any factor (~27%) for the NMDS weighted UniFrac (**Table 1**). NMDS ordination of unweighted (**Figure 4A**) and weighted UniFrac (**Figure 4B**) distances both identified that Oasis Valley microbial communities are distinct from other study-area microbial communities. The following clusters appear to group together for both ordinations, even when removing potentially biased samples (e.g., sequence batch): (1) Ash Meadows and Spring Mountains, and (2) Pahute Mesa and Frenchman and Yucca Flat. Rainier Mesa and Death Valley microbial communities differ between unweighted and weighted UniFrac NMDS ordinations, suggesting that both OTU relative abundances and phylogeny are important factors for the sites sampled. In addition to location, sampling depth relative to the water table (P = 0.001) significantly contributed to the variation but explained only ~6%.

Geochemical-type variables minimally explained community variation (**Table 1, Figure 4C**). For NMDC ordinations, rock type is the most significant contributor (~9%; P = 0.002), followed by TOC (~6%, P = 0.026) and temperature (~6%, P = 0.002). Unlike location-type variables, the PERMANOVA of geochemical-type variables is significantly influenced by sequence batch runs (P = 0.039) and sampling-method approaches (P = 0.020). However, constrained ordination using CCA (**Figure 4C**) confirmed that rock type, TOC, and temperature are significant (P < 0.05) factors determining the microbial community variation. Oasis Valley samples were removed before performing CCA because of their distinct geochemical and microbial compositions. Although CCA only explained ~8% of the total community variation, the microbial communities could be clustered by rock type and location. Notably, similar to the geochemical ordination (**Figure 2**), the two rock type clusters (carbonate and volcanic) followed along the vectors for calcium, nitrate, and sodium. Within location, CCA also revealed that the community variation may be explained by certain geochemical-type variables (e.g., **Figure 4C**; Pahute Mesa communities appear to vary along a temperature gradient). Thus, the combined results of CCA and NMDS ordinations indicate that, on a regional-scale, although microbial community variation can be explained more by location, geochemical conditions play an important role.

Other ordinations, including DEICODE RPCA (61), CLR PCA, and PhILR PCA (63) (**Figure S7**), that take into account the sparse, compositional nature of microbial community datasets (62) were evaluated to compare against the NMDS and CCA ordinations. Consistent with the NMDS ordinations, Oasis Valley communities were identified as a separate group for all three ordinations (**Figure S7A, S7D, S7G**). However, this result may be attributed to a noticeable separation of samples by sequence batch for CLR (**Figure S7E**) and PhILR PCA (**Figure S7H**). Sequence batch did not seem to impact DEICODE RPCA (**Figure S7B**), likely because the DEICODE algorithm utilizes the geometric mean of log-transformed nonzero data and conducts matrix completion (61). After removing potentially biased samples (i.e., Batch 2 and tunnel-collected samples), both CLR and PhILR PCA ordinations also identified the two groups observed for NMDS ordinations: (1) Ash Meadows and Spring Mountains, and (2) Pahute Mesa and Frenchman and Yucca Flat (**Figure S7F** and **S7I**). In contrast, these two groups were not observed for DEICODE RPCA (**Figure S7C**) and may be the result of the likely highly-ranked (61) nature of the microbial community dataset, in which samples contain few similar microbes depending on location or other unknown factors not included in the metadata. Highly-ranked datasets also might be caused by unforeseen local-scale geochemical gradients that impact the microbial community and influence the resultant RPCA ordination (61). In contrast, the unweighted and weighted UniFrac NMDS ordination may accurately represent the microbial community and has been demonstrated to achieve high clustering accuracy (61, 81, 82).

Network analysis of the ‘common OTUs’ identified nine clusters and revealed that co-occurring taxa generally subsisted within the same location (**Figure 5, Figure S8** depicts the ratio of reads per location for each OTU). For network analysis, OV samples were removed because beta-diversity and geochemical analyses indicate that OV1 and OV2 were microbially and geochemically distinct from all other locations. Five clusters (NNSS 1–5) represent OTUs that co-occurred within or near the NNSS sites (Pahute Mesa, Rainier Mesa, and Frenchman and Yucca Flat). Each NNSS 1–5 cluster contains different ‘common OTUs’ (**Table S7**). Cluster AV (Amargosa Valley) was categorized separately from the NNSS clusters since the majority of OTUs co-occurred only in Amargosa Valley. In addition, there is a high degree (average = 22.8) of connectivity within cluster AV, which may be attributed to the collection of three samples pumped at the same site (4PD [AV1–3]) over time from 1 to 23 days. This suggests there is some degree of temporal consistency within this site, and potentially in other sites. In comparison, the degree of connectivity for the five NNSS clusters average between 3.2–7.7 (**Table S6**). Cluster AM also has a high average degree (19.3) of connectivity and three nodes connect cluster AM to the rest of the network via cluster SM.

3.5 Ecological processes that influence microbial community variation

Null models (see **Table S1** for definitions) were used to identify the ecological processes that drove microbial community assembly and could influence community variation. This approach has been applied to communities in surface water (83), soil (64, 83, 84), and the subsurface (18–20, 24, 65, 83, 85, 86). There are two null modeling steps to identify deterministic and stochastic ecological processes: βNTI and RCBC, as described in **section 2.5**. Overall, stochastic and undominated processes governed the microbial communities most (**Figure 6 inset**), with deterministic processes only contributing ~25% (except for Oasis Valley at ~50%). Although certain locations had limited samples in our study, the assembly processes, taken as a whole, support the clusters identified by ordination analyses. The microbial communities of Spring Mountains, Ash Meadows, and Oasis Valley were governed largely by variable selection ($\beta\text{NTI} > 2$), dispersal limitation ($|\beta\text{NTI}| < 2$ and $\text{RCBC} > 0.95$), and undominated processes ($|\beta\text{NTI}| < 2$ and $|\text{RCBC}| < 0.95$) when compared against other communities (**Figure 6**). For microbial communities within and near the NNSS, homogeneous dispersal ($|\beta\text{NTI}| < 2$ and $\text{RCBC} < -0.95$) and homogeneous selection ($\beta\text{NTI} < -2$) were more prominent. For example, Pahute Mesa communities were more similar (homogeneous selection) and had less turnover (homogenizing dispersal) than expected by chance. This signifies high dispersal rates and comparable geochemical conditions across Pahute Mesa that forced local communities to be more similar.

4 Discussion

In this study, we hypothesize that groundwater flows through hydraulically connected fractures and enables the dispersal of planktonic microorganisms, sometimes over considerable distances. Along a given flow path, geochemical conditions evolve with rock-water interactions or mixing of disparate fluids within interconnected pore spaces. These conditions can impart selective pressures on transported microorganisms and influence the local community assembly patterns. Fluid conduits also can become clogged over time and inhibit microbial dispersal, such as from tectonic influences (87–89), secondary mineral infilling, or the overgrowth of microbial biofilms at biogeochemical hotspots (42, 90, 91). Separated communities may change with time into disparate compositions because of stochastic processes (e.g., genetic or ecological drift) (22). Ecological processes can be quantified (18–20), and in combination with microbial diversity and co-occurrence patterns, we assess whether regional microbial community patterns are consistent with regional groundwater-flow paths (**Figure 1**). Microbial community patterns are compared to the following regional groundwater-flow paths (25): (1) Spring Mountains to Ash Meadows discharge area; (2) Frenchman & Yucca Flat to Amargosa Desert; (3) Rainier Mesa to Amargosa Desert; (4) Pahute Mesa to Oasis Valley; and (5) Amargosa Desert to Death Valley (**Figure 1**).

4.1 Microbial community dispersal and evolution from Spring Mountains to Ash Meadows

The Spring Mountains (recharge area) to Ash Meadows (discharge area) groundwater-flow path (**Figure 1**) likely transports planktonic microorganisms. In the AM Basin, the microbial community is similar between the recharge area (Spring Mountains) and Ash Meadows discharge area (**Figure 4**). This result was corroborated with network analysis (**Figure 5**), which identified ‘common OTUs’ co-occurring in the Spring Mountains and Ash Meadows discharge area (e.g., ‘common OTUs’ in cluster SM). The separation of ‘common OTUs’ into two clusters (AM and SM; clusters previously described in **section 3.4**) suggests that other factors impacted the communities observed in each location. Indeed, null model analyses identified three major ecological forces that can explain differences between the Ash Meadows and Spring Mountains communities: selection processes, dispersal limitation, and undominating processes (**Figure 6**). This is not surprising given the differences in elevation and geochemistry (**Figure 2**) that arise from the evolution of groundwater along the flow path. Notably, these two communities are separated in the CCA ordination (**Figure 4C**), likely due to differences in calcium and nitrate. In addition, the most-recent groundwater characterization study of the DVRFS (25) identified other sources that contribute to Ash Meadows discharge, including the Sheep, Desert, Pintwater, and Spotted Ranges (**Figure 1**).

4.2 Microbial community similarities within the NNSS and at Amargosa Valley

Microbial communities at sites within and near the NNSS were relatively similar (Amargosa Valley, Frenchman and Yucca Flat, Pahute Mesa, Rainier Mesa; **Figure 4**), despite locations occurring within different groundwater basins (**Figure 1**). Network analysis (**Figure 5**) also identified co-occurring ‘common OTUs’ within these locations, and five disparate NNSS clusters of ‘common OTUs’ were identified (Cluster NNSS 1–5). The five NNSS clusters connected to the network via cluster AV (Amargosa Valley; within Amargosa Desert), cluster DV (Death Valley), or cluster AM (Ash Meadows).

Similarities between the microbial communities of NNSS sites with Amargosa Valley suggest the communities have adapted to similar environmental conditions or have dispersed via a groundwater flow connection towards Amargosa Valley. Many communities between these locations were dominated by homogenizing dispersal or homogeneous selection (**Figure 6**), demonstrating the communities are more similar than expected by chance. Network analysis also grouped ‘common OTUs’ with putatively diverse metabolisms within cluster AV. The four most abundant OTUs in cluster AV were putative sulfur-oxidizing (e.g., *Thioalkalispiraceae, Hydrogenophilaceae*), methanotroph (e.g., *Methylomonas*), and iron-oxidizing (e.g., *Gallionellaceae*) microbes. Some OTUs are likely metabolically flexible, capable of mixotrophic or facultative anaerobic growth, such as those within family *Hydrogenophilaceae.* The diversity of potential metabolisms present in cluster AV is consistent with the geochemical mixing of multiple groundwater sources in the Amargosa Desert. Groundwater in the Amargosa Desert is derived largely from the AFFCR and AM groundwater basins (25), in which three recharge areas converge in the Amargosa Desert: (1)

Yucca Mountain and nearby upland areas in the AFFCR groundwater basin; (2) Yucca Flat in the AM groundwater basin; and (3) infiltration from the Amargosa River and Fortymile Wash in the AFFCR basin (**Figure 1**). This mixed groundwater then flows toward the Furnace Creek discharge area in Death Valley.

4.3 The Death Valley microbial community likely supports the basin-fill flow conceptualization

There are two conceptualizations for groundwater flow towards Death Valley: (1) the deep-carbonate flow conceptualization; and (2) the basin-fill flow conceptualization (**Figure 7**). Previous groundwater studies suggested that Death Valley discharge was sourced from a deep-carbonate flow path that passed beneath the Ash Meadows discharge area and Amargosa Desert (36, 92, 93) (**Figure 7**). In contrast, the most recent groundwater study by Halford and Jackson (2020) determined that the basin-fill flow conceptualization best explains water levels, water chemistry, and aquifer-testing data. The basin-fill flow conceptualization interprets groundwater flow as moving through shallow basin-fill, volcanic, and carbonate rocks toward Furnace Creek–the principal discharge area in Death Valley (**Figure 7**).

The Death Valley microbial community better supports hydraulic connections proposed in the basin-fill conceptualization, rather than the deep-carbonate flow conceptualization. The Death Valley microbial community consists of samples collected from three carbonate wells: DV1 (Nevares Deep Well 2) and DV2 (Nevares) in the Furnace Creek discharge area; and DV3 (Inyo-BLM 1) in the central Amargosa Desert (**Figure 1**). For the Death Valley microbial community to support the deep-carbonate flow conceptualization, microbial communities are expected to be related between Death Valley, Ash Meadows, and the Spring Mountains. Instead, NMDS ordination (**Figure 4A and 4B**), network analysis (**Figure 5**), and ecological null models (**Figure 6**) suggest that Death Valley microbial communities are more related to Amargosa Valley and NNSS locations, rather than to Ash Meadows and Spring Mountains. Although there may be some similarity between Death Valley communities and Ash Meadows in the CCA ordination (**Figure 4C**), this analysis only accounts for ~8% of the total community variation. Moreover, homogeneous dispersal or selection contributed to community turnover between Death Valley, NNSS locations, and Amargosa Valley (**Figure 6**; **Table S8**). This suggests that these communities were more similar than expected by random chance. In contrast, dispersal limitation and variable selection played a role in community turnover between Death Valley and Ash Meadows/Spring Mountains, which suggests a limited hydraulic connection between these areas. Network analysis indicates that cluster DV is closely associated with the NNSS clusters, particularly NNSS1 and NNSS2, and ‘common OTUs’ observed in Death Valley are also found in NNSS clusters and cluster AV (**Figure 5**).

4.4 The Death Valley microbial community may support groundwater compartmentalization into shallow and deep zones

Halford and Jackson (2020) posit that the DVRFS is compartmentalized into two parts: (1) a shallow, high-transmissivity part within 500 m of the water table where nearly all flow occurs; and (2) a deep, less active, low-transmissivity part that has limited interaction with the shallow part (25). The Death Valley microbial community appears to support a compartmentalized groundwater system.

Nevares wells (DV1/DV2) have ~100 m sampling depths and are conceptualized as occurring within the active, high-transmissivity part of the groundwater system. In contrast, Inyo-BLM 1 (DV3) was sampled at ~600 m depth, and is in low-transmissivity rock (~9–84 m^2 per day) (25, 94); thus, Inyo-BLM 1 (DV3) is conceptualized as occurring within the deep, less active, low-transmissivity part of the groundwater system. Inyo-BLM 1 (DV3) and Nevares Deep Well 2 (DV1) have similar geochemical conditions (**Figure 2, Table S2**), including temperature and sulfate concentrations. Despite similar aqueous chemistries, Thomas *et al.* (2013) suggested hydraulic isolation between the microbial communities of Inyo-BLM 1 (DV3) and Nevares Well and Spring.

In this study, NMDS and CCA ordinations (**Figure 4**) appears to concur with Thomas *et al.* (2013), with more similarity between Nevares wells (DV1/DV2) compared to Inyo-BLM1 (DV3). There also are large phylogenetic and compositional differences. Only eight OTUs were present in both Nevares Deep Well 2 (DV1) and Inyo-BLM 1 (DV3); and only two OTUs were present in both Nevares (DV2) and Inyo-BLM 1 (DV3). In contrast, 256 OTUs were present in both Nevares communities (DV1/DV2). The Inyo-BLM 1 (DV3) community consisted of a large population of putative methanogens (~50%), followed by sulfate reducers (~31%). In contrast, probable sulfate reducers composed most of the population at Nevares (DV1/DV2). Community differences may be impacted by unmeasured geochemical factors, such as higher pressures at depth in Inyo-BLM 1 (DV3) compared to the shallow Nevares wells (DV1/DV2). These deterministic differences were not apparent using ecological null models (**Figure 6**), and more samples are required to confirm the null models between Inyo-BLM 1 (DV3) and Nevares wells (DV1/DV2).

4.5 Lack of similarity between Pahute Mesa and Oasis Valley

The microbial community of Oasis Valley is compositionally and phylogenetically distinct from all other study-area locations (**Figure 4**). Oasis Valley OTUs represented ~39% of the total OTUs sampled in this study, of which ~85% were observed only in Oasis Valley. The lack of microbial similarity between Pahute Mesa and Oasis Valley does not reflect regional groundwater-flow paths in the PMOV basin (**Figure 1**; 37, 38). However, the two Oasis Valley samples may not be representative of regional groundwater in this location as supported by the presence of many soil-associated and aerobic microbes. Moreover, the geochemistry of OV1 is distinct from other locations (**Figure 2**), with relatively high TOC concentrations (37 mg-C/L). The combination of microbial (**Figure 4**; taxonomic composition) and geochemical (**Figure 2**) data suggest that the two Oasis Valley samples are likely more impacted by land-surface-associated factors, owing to their relatively shallow depths and siting within the alluvium of the Amargosa River floodplain. Hence, these sites are probably disconnected hydrologically from the regional aquifer and conceptually can be regarded as hyporheic. This observation demonstrates the utility in supplementing geologic and hydrologic datasets with microbial community data for a comprehensive evaluation of groundwater samples.

The Pahute Mesa microbial community consists of wells that were sampled within the Pahute Mesa recharge area (PM3; PM6; PM11; PM12) and wells immediately downgradient of the recharge area (PM1; PM2; PM4; PM5; PM7–PM10; PM13–PM16) (**Figure 1**). Pahute Mesa communities are similar by NMDS and CCA ordinations (**Figure 4**), and many ‘common OTUs’ co-occur within Pahute Mesa (cluster NNSS 1) (**Figure 5**). Moreover, the two ecological processes dominating between Pahute Mesa communities include homogeneous dispersal and homogeneous selection (**Figure 6**). Microbial dispersal may be influences by the flow system at Pahute Mesa, which is dominated by high groundwater-flow velocities (39, 42). Groundwater flow in this area occurs through highly transmissive volcanic rocks, and tritium plumes from Pahute Mesa have only traveled less than 4 km beyond the NNSS borders since 2021, with an advective transport velocity of ~84 m/yr (25). Taken together, these observations suggest that microbial dispersal in the PMOV basin is possible at least within the Pahute Mesa recharge location.

5 Future Outlook

This study demonstrates that the regional-scale groundwater microbial community is a relevant data source; however, there were many limitations, as noted in Section 3.1, and future studies are needed. For example, the findings in this study suggest there could be a detection limit to identifying similar microbial communities between recharge and discharge areas, which may be correlated with groundwater-flow rates, transmissivity, or time. Areas with relatively fast-flowing groundwater and high transmissivity are more likely to have similar microbial communities (e.g., Ash Meadows and Spring Mountains). In contrast, areas with low transmissivity (e.g., deep-carbonate flow towards Death Valley) may impede large microbial migration from recharge to discharge areas, such that deterministic factors and genetic evolution outcompete dispersal. Future studies that quantify and correlate microbial community patterns with hydrogeologic factors and perturbations (e.g., groundwater-flow rate, transmissivity, time-scale) are needed. Temporal replicates will also help establish the microbial community variation within each site and may reveal potential seasonal-, pumping-, or recharge-associated perturbations. Moreover, groundwater microbial community samples collected along a groundwater-flow path and at various depths can identify potential candidates for use as a microbial ‘tracer’ by providing insight into the persistence of specific microbial species. This spatial data also will help to confirm that ‘common OTUs’ in recharge and discharge zones can be used for helping characterize groundwater flow. While our study approach was cost-effective (amplicon sequencing), it was limited to microbial abundance information and the bacterial/archaeal community. Future studies can combine other approaches, such as obtaining microbial functions/activity and metabolites, to provide additional insight into hydrobiogeochemical dynamics and fluxes.

6 Conclusion

The DVRFS microbial community patterns were mostly consistent with the regional-scale groundwater-flow conceptualization. Overall, microbial communities within recharge and discharge areas connected by a flow path were similar (e.g., Spring Mountains and Ash Meadows), and location was the most significant variable differentiating between the communities. Notably, communities within and near the NNSS were similar, although groundwater flow from Pahute Mesa towards other NNSS sites is limited. Network analysis also demonstrated that ‘common OTUs’ clustered together by location, and in particular, clusters of ‘common OTUs’ found within and near the NNSS were connected via a cluster composed of Amargosa Valley OTUs. These OTUs represented a range of putative metabolisms, indicative of the mixing of various groundwater sources at Amargosa Valley, which is consistent with the most recent groundwater conceptualization of the DVRFS. Ecological null model analyses also identified locations in which communities were relatively similar due to deterministic and stochastic processes, and largely corroborated the other microbial community analyses conducted in this study. However, the microbial community patterns contradicted the hydraulic connection between Pahute Mesa (recharge) and Oasis Valley (discharge) in the PMOV groundwater basin, probably reflecting a lack of direct hydrologic connectivity between the upgradient regional flow system and discharge zone wells that were available for sampling. Overall, this exploratory study demonstrates that regional-scale groundwater microbial community patterns can be used to supplement geologic and hydrologic data in characterizing groundwater flow.

7 Acknowledgments

We thank Navarro-Interra for coordination with NNSS field sampling and on-site measurements. We thank Rachel Lindvall and Sarah Roberts (LLNL); Ron Hershey and John Healey (Desert Research Institute); Jeff Sanchez and Robert Zella (U.S. Geological Survey); and Robert Goodwin, Russ Shelton, Jeff Wurtz, and others from the UGTA Program for onsite logistical support at NNSS. Special thanks to Bill Wilborn, former UGTA sub-project director, for his interest and support of our work. Thanks to Terry Fisk, Richard Friese, Josh Hoines, Genne Nelson, and Kevin Wilson (U.S. National Park Service); Alisa Lembke and the Inyo County Planning Commission; John Bredehoeft and Michael King (The Hydronamics Group, LLC); Darrell Lacy, Jamie Walker, John Klenke, Levi Kryder and others from the Nye County Nuclear Waste Repository Program Office; John Healey, Brad Lyles, Patty Montgomery, Molly Devlin, and Brittany Kruger (DRI); and the Nature Conservancy for site access and logistical assistance with the Death Valley, Amargosa Valley and Oasis Valley sites. Thanks to the DRI Water Lab, T.C. Onstott (Princeton University) and Barbara Sherwood Lollar (University of Toronto) for supplemental chemistry analysis. We also thank Yongqin Jiao (LLNL) for invaluable discussions; and Jeff Kimbrel and Juliet Johnston (LLNL), the

Bioinformatics Virtual Coordination Network (BVCN), especially Haley Sapers and Liz Suter, and the LLNL Data Science Institute Consulting Service (DSICS), especially Jason Bernstein and Aram Avila-Herrera, for their advice in analyzing the microbial communities.

8 Funding Sources

This research was partially funded by the Office of Biological and Environmental Research of the U.S. Department of Energy (DOE) as part of the Subsurface Biogeochemical Research Program under Work Proposal Number SCW1053, *Biogeochemistry of Actinides* and under DOE Contract Number SC0005306, *Radiochemically-Supported Microbial Communities*. Partial funding was also provided by the U.S. DOE Environmental Remediation Sciences Program under Grant Number FG02-07ER64406, *Characterizations of Microbial Communities in Subsurface Nuclear Blast Cavities of the Nevada Test Site*); NSF GoLife DEB 1441646, C*ollaborative research: Untangling the Deep Genealogy of Microbial Dark Matter,* and NSF EPSCOR OIA 1826734, *RII Track-2 FEC: Single Cell Genome-to-Phenome: Integrating Genome and Phenome Analyses of Individual Microbial Cells in Complex Microbiomes;* NASA Astrobiology Institute's *Life Underground* project (NNA13AA92A); NWRPO (10-023), *Characterization of Indigenous Microorganisms from Nye County Wells and Assessment of their Utility as Hydrologic Tracers;* The Hydrodynamic Group, *Characterization of Indigenous Microorganisms from Deep Wells and Assessment of their Utility as Hydrologic Tracers.* This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (IM#: LLNL-JRNL-823248). Thanks also to the Alfred P. Sloan Foundation's Deep Carbon Observatory (DCO) for sequencing and postdoc support (APS 2013-10-03) through their Census of Deep Life (CoDL) and Deep Life Community (DLC) programs. JS was partially supported by a Nevada NASA Space Grant Consortium Graduate Fellowship (NNX15AI02H).

Author Contributions: NM analyzed the 16S rRNA gene amplicon sequences and compiled the geochemical data. DPM served as principal investigator for the projects under which the samples were collected. SHB, JCB, JS, and JCF contributed to the collection, processing, and extraction of the samples. All authors wrote the manuscript.

Competing Interest Statement: The authors have no competing interests.

References

1. S. P. Neuman, Trends, prospects and challenges in quantifying flow and transport through fractured rocks. *Hydrogeol. J.* **13**, 124–147 (2005).

2. M. Ye, L. Wang, K. F. Pohlmann, J. B. Chapman, Evaluating Groundwater Interbasin Flow Using Multiple Models and Multiple Types of Data. *Groundwater* **54**, 805–817 (2016).

3. F. H. Chapelle, The significance of microbial processes in hydrogeology and geochemistry. *Hydrogeol. J.* **8**, 41–46 (2000).

4. T. M. Flynn, *et al.*, Functional microbial diversity explains groundwater chemistry in a pristine aquifer. *BMC Microbiol.* **13**, 146 (2013).

5. C. Griebler, F. Malard, T. Lefébure, Current developments in groundwater ecology—from biodiversity to ecosystem function and services. *Curr. Opin. Biotechnol.* **27**, 159–167 (2014).

6. J. Tóth, A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res. 1896-1977* **68**, 4795–4812 (1963).

7. J. Tóth, Groundwater as a geologic agent: An overview of the causes, processes, and manifestations. *Hydrogeol. J.* **7**, 1–14 (1999).

8. R. A. Freeze, P. A. Witherspoon, Theoretical analysis of regional groundwater flow: 2. Effect of water-table configuration and subsurface permeability variation. *Water Resour. Res.* **3**, 623–634 (1967).

9. W. W. Woessner, *Groundwater-Surface Water Exchange* (GW Project, 2020) (April 4, 2022).

10. P. Cook, *Introduction to Isotopes and Environmental Tracers as Indicators of Groundwater Flow* (GW Project, 2020) (April 4, 2022).

11. J. F. Devlin, *Groundwater Velocity* (Groundwater Project, 2020) (April 4, 2022).

12. R. W. Harvey, Microorganisms as tracers in groundwater injection and recovery experiments: a review. *FEMS Microbiol. Rev.* **20**, 461–472 (1997).

13. M. A. Walvoord, *et al.*, Groundwater flow and geochemistry in the Southeastern San Juan Basin: Implications for microbial transport and activity. *Water Resour. Res.* **35**, 1409–1424 (1999).

14. S. Amalfitano, *et al.*, Groundwater geochemistry and microbial community structure in the aquifer transition from volcanic to alluvial areas. *Water Res.* **65**, 384–394 (2014).

15. Y. Zhang, *et al.*, Microbial Community Composition in Deep-Subsurface Reservoir Fluids Reveals Natural Interwell Connectivity. *Water Resour. Res.* **56**, e2019WR025916 (2020).

16. T. R. Ginn, *et al.*, Processes in microbial transport in the natural subsurface. *Adv. Water Resour.*, 26 (2002).

17. C. P. Gerba, I. L. Pepper, D. T. Newby, “Microbial Transport in the Subsurface” in *Environmental Microbiology*, Third, (Elsevier, 2015), pp. 319–337.

18. J. C. Stegen, X. Lin, A. E. Konopka, J. K. Fredrickson, Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* **6**, 1653–1664 (2012).

19. J. C. Stegen, *et al.*, Quantifying community assembly processes and identifying features that impose them. *ISME J.* **7**, 2069–2079 (2013).

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

21. S. B. Maamar, *et al.*, Groundwater Isolation Governs Chemistry and Microbial Community Structure along Hydrologic Flowpaths. *Front. Microbiol.* **6** (2015).

22. J. Zhou, D. Ning, Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiol. Mol. Biol. Rev.* **81** (2017).

23. L. A. Hug, *et al.*, Aquifer environment selects for microbial species cohorts in sediment and groundwater. *ISME J.* **9**, 1846–1856 (2015).

24. L. Fillinger, K. Hug, C. Griebler, Selection imposed by local environmental conditions drives differences in microbial community composition across geographically distinct groundwater aquifers. *FEMS Microbiol. Ecol.* **95** (2019).

25. K. J. Halford, T. R. Jackson, “Groundwater Characterization and Effects of Pumping in the Death Valley Regional Groundwater Flow System, Nevada and California, with Special Reference to Devils Hole” (U.S. Geological Survey, 2020).

26. U.S. DOE, “United States Nuclear Tests: July 1945 through September 1992” (U.S. Department of Energy, 2015).

27. M. T. Moreo, K. J. Halford, R. J. La Camera, R. J. Laczniak, “Estimated Ground-Water Withdrawals from the Death Valley Regional Flow System, Nevada and California, 1913-98” (Geological Survey (U.S.), 2003).

28. M. T. Moreo, L. Justet, “Update to the Ground-Water Withdrawals Database for the Death Valley Regional Ground-Water Flow System, Nevada and California, 1913-2003” (Geological Survey (U.S.), 2008).

29. I. J. Winograd, F. J. Pearson, Major carbon 14 anomaly in a regional carbonate aquifer: Possible evidence for megascale channeling, South Central Great Basin. *Water Resour. Res.* **12**, 1125–1143 (1976).

30. I. J. Winograd, A. C. Riggs, T. B. Coplen, The relative contributions of summer and cool-season precipitation to groundwater recharge, Spring Mountains, Nevada, USA. *Hydrogeol. J.* **6**, 77–93 (1998).

31. J. M. Thomas, *et al.*, "Geochemical and Isotopic Interpretations of Groundwater Flow in the Oasis Valley Flow System, Southern Nevada" (U.S. Department of Energy, 2003) https:/doi.org/10.2172/806667 (March 26, 2020).

32. J. M. Thomas, *et al.*, Using Water Chemistry, Isotopes and Microbiology to Evaluate Groundwater Sources, Flow Paths and Geochemical Reactions in the Death Valley Flow System, USA. *Procedia Earth Planet. Sci.* **7**, 842–845 (2013).

33. K. Anderson, S. Nelson, A. Mayo, D. Tingey, Interbasin flow revisited: The contribution of local recharge to high-discharge springs, Death Valley, CA. *J. Hydrol.* **323**, 276–302 (2006).

34. M. Bushman, "Contribution of Recharge Along Regional Flow Paths to Discharge at Ash Meadows, Nevada," Brigham Young University. (2008).

35. W. R. Belcher, M. S. Bedinger, J. T. Back, D. S. Sweetkind, Interbasin flow in the Great Basin with special reference to the southern Funeral Mountains and the source of Furnace Creek springs, Death Valley, California, U.S. *J. Hydrol.* **369**, 14 (2009).

36. W. R. Belcher, D. S. Sweetkind, C. C. Faunt, M. T. Pavelko, M. C. Hill, "An update of the Death Valley regional groundwater flow system transient model, Nevada and California" (U.S. Geological Survey, 2017) (October 13, 2019).

37. M. Bushman, S. T. Nelson, D. Tingey, D. Eggett, Regional groundwater flow in structurally-complex extended terranes: An evaluation of the sources of discharge at Ash Meadows, Nevada. *J. Hydrol.* **386**, 118–129 (2010).

38. J. M. Fenelon, D. S. Sweetkind, R. J. Laczniak, "Groundwater Flow Systems at the Nevada Test Site, Nevada: A Synthesis of Potentiometric Contours, Hydrostratigraphy, and Geologic Structures" (U.S. Geological Survey, 2010) (February 13, 2020).

39. J. M. Fenelon, K. J. Halford, M. T. Moreo, "Delineation of the Pahute Mesa–Oasis Valley groundwater basin, Nevada" (U.S. Geological Survey, 2016).

40. S. T. Nelson, A. L. Mayo, The role of interbasin groundwater transfers in geologically complex terranes, demonstrated by the Great Basin in the western United States. *Hydrogeol. J.* **22**, 807–828 (2014).

41. S. R. Warix, L. K. Rademacher, Z. P. Meyers, M. D. Frisbee, Groundwater geochemistry and flow in the Spring Mountains, NV: Implications for the Death Valley Regional Flow System. *J. Hydrol.* **580**, 124313 (2020).

42. T. R. Jackson, J. M. Fenelon, R. L. Paylor, “Groundwater flow conceptualization of the Pahute Mesa–Oasis Valley Groundwater Basin, Nevada—A synthesis of geologic, hydrologic, hydraulic-property, and tritium data” (U.S. Geological Survey, 2021).

43. R. J. Laczniak, *et al.*, “Ground-water discharge determined from estimates of evapotranspiration, Death Valley regional flow system, Nevada and California” (2001) https:/doi.org/10.3133/wri014195.

44. D. P. Moser, *et al.*, “Radiochemically-supported microbial communities. A potential mechanism for biocolloid production of importance to actinide transport” (Desert Research Institute, Las Vegas, NV (United States), 2014) https:/doi.org/10.2172/1213048 (May 25, 2021).

45. I. M. Farnham, The Nature and State of Groundwater Contamination at the Nevada National Security Site: What Have We Learned from Decades of Groundwater Analysis? in *WM2020*, (2020).

46. B. J. Callahan, *et al.*, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

47. E. Bolyen, *et al.*, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

48. S. Louca, F. Mazel, M. Doebeli, L. W. Parfrey, A census-based estimate of Earth’s bacterial and archaeal diversity. *PLOS Biol.* **17**, e3000106 (2019).

49. S. I. Glassman, J. B. H. Martiny, Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. *mSphere* **3**, e00148-18 (2018).

50. L. Joos, *et al.*, Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. *BMC Genomics* **21**, 733 (2020).

51. S. Moossavi, F. Atakora, K. Fehr, E. Khafipour, Biological observations in microbiota analysis are robust to the choice of 16S rRNA gene sequencing processing algorithm: case study on human milk microbiota. *BMC Microbiol.* **20**, 290 (2020).

52. J. T. Nearing, G. M. Douglas, A. M. Comeau, M. G. I. Langille, Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. *PeerJ* **6**, e5364 (2018).

53. A. Prodan, *et al.*, Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLOS ONE* **15**, e0227434 (2020).

54. C. Quast, *et al.*, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).

55. P. Yilmaz, *et al.*, The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* **42**, D643–D648 (2014).

56. N. A. Bokulich, *et al.*, Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 90 (2018).

57. N. Bokulich, M. Robeson, B. Kaehler, M. Dillon, bokulich-lab/RESCRIPt: 2020.6.1 (2020) https:/doi.org/10.5281/zenodo.3945228 (July 31, 2020).

58. P. J. McMurdie, S. Holmes, phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217 (2013).

59. T. C. Hsieh, K. H. Ma, A. Chao, iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods Ecol. Evol.* **7**, 1451–1456 (2016).

60. S. W. Kembel, *et al.*, Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464 (2010).

61. C. Martino, *et al.*, A Novel Sparse Compositional Technique Reveals Microbial Perturbations. *mSystems* **4** (2019).

62. G. B. Gloor, J. M. Macklaim, V. Pawlowsky-Glahn, J. J. Egozcue, Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* **8** (2017).

63. J. D. Silverman, A. D. Washburne, S. Mukherjee, L. A. David, A phylogenetic transform enhances analysis of compositional microbiota data. *eLife* **6**, e21887 (2017).

64. F. Dini-Andreote, J. C. Stegen, J. D. van Elsas, J. F. Salles, Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci.* **112**, E1326–E1332 (2015).

65. R. E. Danczak, *et al.*, Ecological Assembly Processes Are Coordinated between Bacterial and Viral Communities in Fractured Shale Ecosystems. *mSystems* **5** (2020).

66. O. J. Hardy, Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. *J. Ecol.* **96**, 914–926 (2008).

67. D. M. Raup, R. E. Crick, Measurement of Faunal Similarity in Paleontology. *J. Paleontol.* **53**, 1213–1227 (1979).

68. J. R. Bray, J. T. Curtis, An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecol. Monogr.* **27**, 325–349 (1957).

69. J. M. Chase, N. J. B. Kraft, K. G. Smith, M. Vellend, B. D. Inouye, Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* **2**, art24 (2011).

70. Z. D. Kurtz, *et al.*, Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLOS Comput. Biol.* **11**, e1004226 (2015).

71. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal* **Complex Systems**, 10 (2006).

72. A. Clauset, M. E. J. Newman, C. Moore, Finding community structure in very large networks. *Phys. Rev. E* **70**, 066111 (2004).

73. V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).

74. L. Yan, *et al.*, Groundwater bacterial communities evolve over time in response to recharge. *Water Res.* **201**, 117290 (2021).

75. D. Abiriga, A. Jenkins, K. Alfsnes, L. S. Vestgarden, H. Klempe, Spatiotemporal and seasonal dynamics in the microbial communities of a landfill-leachate contaminated aquifer. *FEMS Microbiol. Ecol.* **97**, fiab086 (2021).

76. G. Pilloni, *et al.*, Dynamics of Hydrology and Anaerobic Hydrocarbon Degrader Communities in A Tar-Oil Contaminated Aquifer. *Microorganisms* **7**, 46 (2019).

77. J. R. Harrill, M. S. Bedinger, "Death Valley Regional Groundwater Flow System, Nevada and California—Hydrogeologic Framework and Transient Groundwater Flow Model" (U.S. Geological Survey, 2010) (March 11, 2020).

78. S. Pavoine, A. Gasc, M. B. Bonsall, N. W. H. Mason, Correlations between phylogenetic and functional diversity: mathematical artefacts or true ecological and evolutionary processes? *J. Veg. Sci.* **24**, 781–793 (2013).

79. N. G. Swenson, *Functional and Phylogenetic Ecology in R*, 1st Ed. (Springer, 2014).

80. B. Sandel, Richness-dependence of phylogenetic diversity indices. *Ecography* **41**, 837–844 (2018).

81. S. Weiss, *et al.*, Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5** (2017).

82. R. J. Wright, M. G. I. Langille, T. R. Walker, Food or just a free ride? A meta-analysis reveals the global diversity of the Plastisphere. *ISME J.* **15**, 789–806 (2021).

83. J. Wang, *et al.*, Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J.* **7**, 1310–1321 (2013).

84. C. Liu, *et al.*, Long-term nitrogen addition affects the phylogenetic turnover of soil microbial community responding to moisture pulse. *Sci. Rep.* **7** (2017).

85. E. D. Beaton, *et al.*, Local and Regional Diversity Reveals Dispersal Limitation and Drift as Drivers for Groundwater Bacterial Communities from a Fractured Granite Formation. *Front. Microbiol.* **7** (2016).

86. R. E. Danczak, M. D. Johnston, C. Kenah, M. Slattery, M. J. Wilkins, Microbial Community Cohesion Mediates Community Turnover in Unperturbed Aquifers. *mSystems* **3** (2018).

87. S. Uprety, *et al.*, The Effect of the 2015 Earthquake on the Bacterial Community Compositions in Water in Nepal. *Front. Microbiol.* **8** (2017).

88. H. Kim, *et al.*, Impact of earthquake on the communities of bacteria and archaea in groundwater ecosystems. *J. Hydrol.* **583**, 124563 (2020).

89. S. Morimura, X. Zeng, N. Noboru, T. Hosono, Changes to the microbial communities within groundwater in response to a large crustal earthquake in Kumamoto, southern Japan. *J. Hydrol.* **581**, 124341 (2020).

90. C. Lumban Gaol, L. Ganzer, S. Mukherjee, H. Alkan, Investigation of clogging in porous media induced by microorganisms using a microfluidic application. *Environ. Sci. Water Res. Technol.* **7**, 441–454 (2021).

91. P. Baveye, P. Vandevivere, B. L. Hoyle, P. C. DeLeo, D. S. de Lozada, Environmental Impact and Mechanisms of the Biological Clogging of Saturated Soils and Aquifer Materials. *Crit. Rev. Environ. Sci. Technol.* **28**, 123–191 (1998).

92. C. C. Faunt, J. B. Blainey, M. C. Hill, F. A. D'Agnese, G. M. O'Brien, "Death Valley regional groundwater flow system, Nevada and California-Hydrogeologic framework and transient groundwater flow model" (2010) https:/doi.org/10.3133/pp1711.

93. J. Bredehoeft, M. King, Potential contaminant transport in the regional Carbonate Aquifer beneath Yucca Mountain, Nevada, USA. *Hydrogeol. J.* **18**, 775–789 (2010).

94. P. A. Cutillo, J. D. Bredehoeft, Estimating Aquifer Properties from the Water Level Response to Earth Tides. *Groundwater* **49**, 600–610 (2011).

Table 1. PERMANOVA of the weighted UniFrac beta diversity ordination using location and geochemical parameters.

Interest	[X][a]	Sum sqrs	R^2	Pseudo *F* statistic	P-value	Betadisper[b] P-value	ANOSIM[b] R-value	ANOSIM P-value
Location	**Location**	**0.488**	**0.266**	**2.104**	**0.003**	**0.168**	**0.418**	**0.001**
	Depth (m)[c]	**0.105**	**0.058**	**3.180**	**0.001**	**0.672**		
	Sampling method[d]	0.063	0.035	0.957	0.416	0.464	0.080	0.169
	Sequence batch[e]	0.039	0.021	1.177	0.244	0.662	0.190	0.091
	Residual	0.994	0.543					
	Total	1.831	1					
Geochemistry	**Rock type[f]**	**0.158**	**0.086**	**4.661**	**0.002**	**0.056**	**0.296**	**0.001**
	Temperature (°C)	**0.101**	**0.055**	**2.990**	**0.002**	**< 0.001**		
	TOC (mg-C/L)	**0.108**	**0.059**	**3.190**	**0.026**	**0.290**		
	pH	0.046	0.025	1.363	0.148	0.477		
	Tritium[g]	0.040	0.022	1.178	0.244	0.115	–0.117	0.887
	Sampling method[d]	**0.122**	**0.067**	**1.806**	**0.020**	**0.464**	**0.080**	**0.169**
	Sequence batch[e]	**0.063**	**0.034**	**1.847**	**0.039**	**0.662**	**0.190**	**0.091**
	Residual	1.119	0.611					
	Total	1.831	1					

[a] PERMANOVA was conducted on the weighted UniFrac beta diversity distribution (**Figure 4B**). [X] refers to the formula: adonis2([data] ~ [X], perm = 999, by = “margin”). [X] considers all listed variables for each ‘Interest’. [data] is the distance matrix derived from the weighted UniFrac. For example, adonis2([data] ~ Location + Depth + SamplingMethod + SequenceBatch, permutations = 999, by = “margin”). Bolded variables are significant ($P < 0.05$ for PERMANOVA and/or ANOSIM).

[b] Betadisper (homogeneity condition) and ANOSIM (analysis of similarity for categorical variables) R functions were used on individual variables.

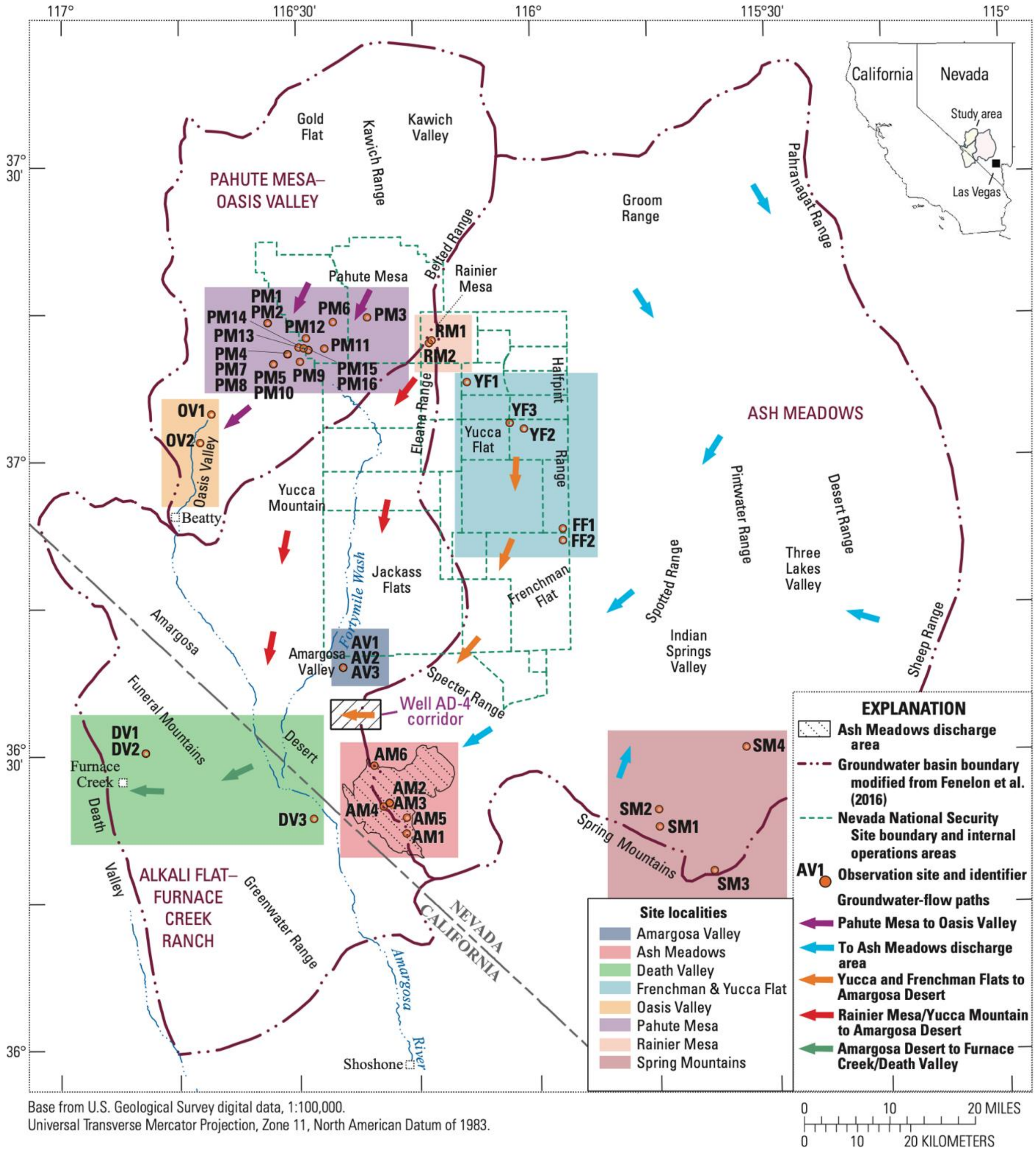
[c] Sample collection depth from the water table.

[d] Samples were collected by pump, jack pumps (in-line submerged pump), or bailer. For location variables, ‘sampling method’ was used as a proxy for ‘location type’ (well, spring, or tunnel) due to overlaps in metadata for spring and groundwater sampling methods.

[e] The majority of samples were sequenced in Batch 1 and seven samples sequenced in Batch 2 (**Table S2**).

[f] Samples were categorized by rock type based on host rock (carbonate or volcanic) at the depth sampled.

[g] Tritium concentrations were split into ‘High’ (≥100 Bq/L) and ‘Low’ (<100 Bq/L) categories



Base from U.S. Geological Survey digital data, 1:100,000.
Universal Transverse Mercator Projection, Zone 11, North American Datum of 1983.

Figure 1. Map of groundwater basins, locations, sampling sites, and groundwater-flow paths within the study area.

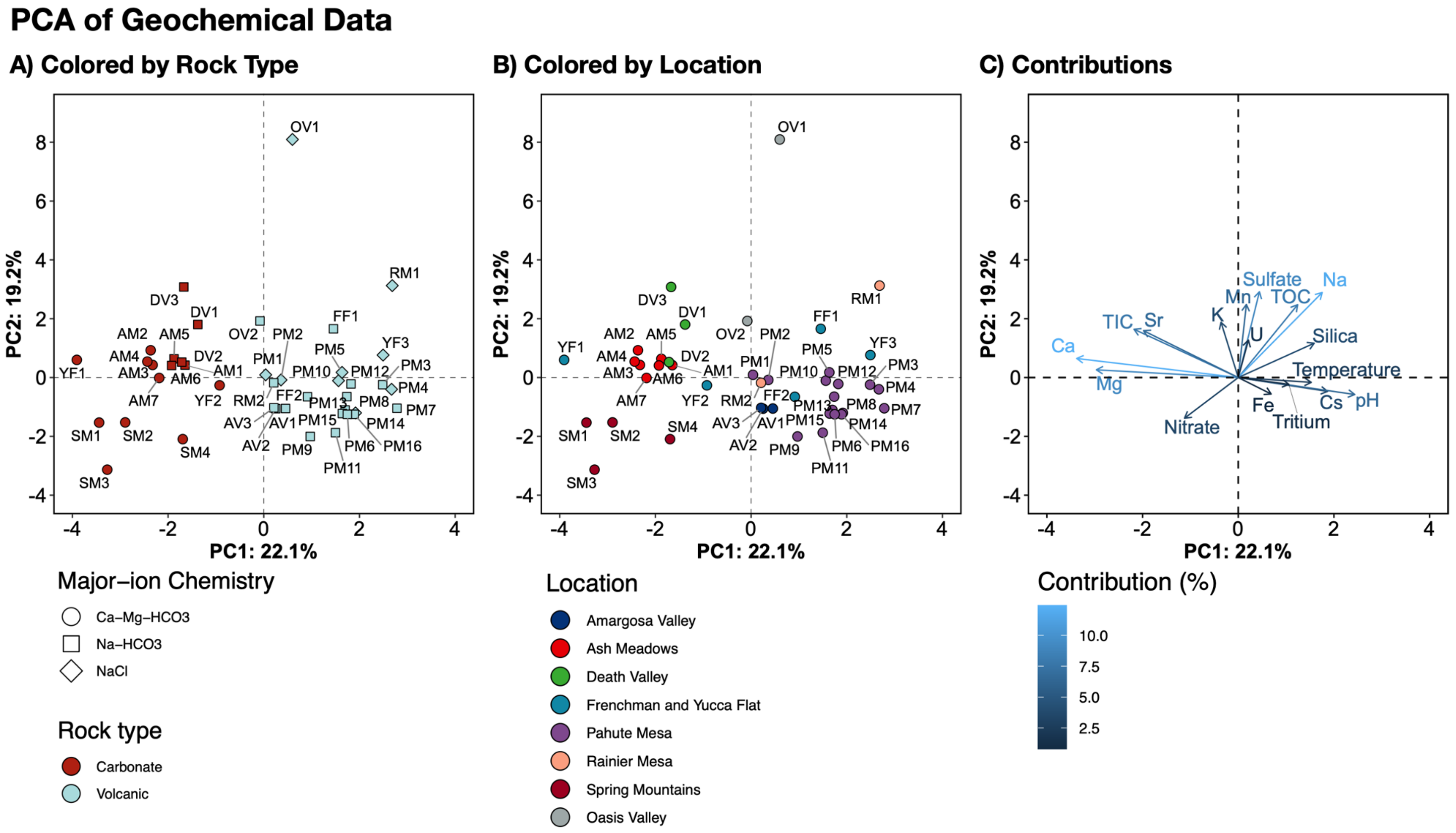


Figure 2. Principal components analysis (PCA) of geochemical data. (**A**) Variation of samples by major-ion chemistry and rock type. (**B**) Variation of samples by location. (**C**) Contribution of each geochemical variable (n=17) used in the PCA.

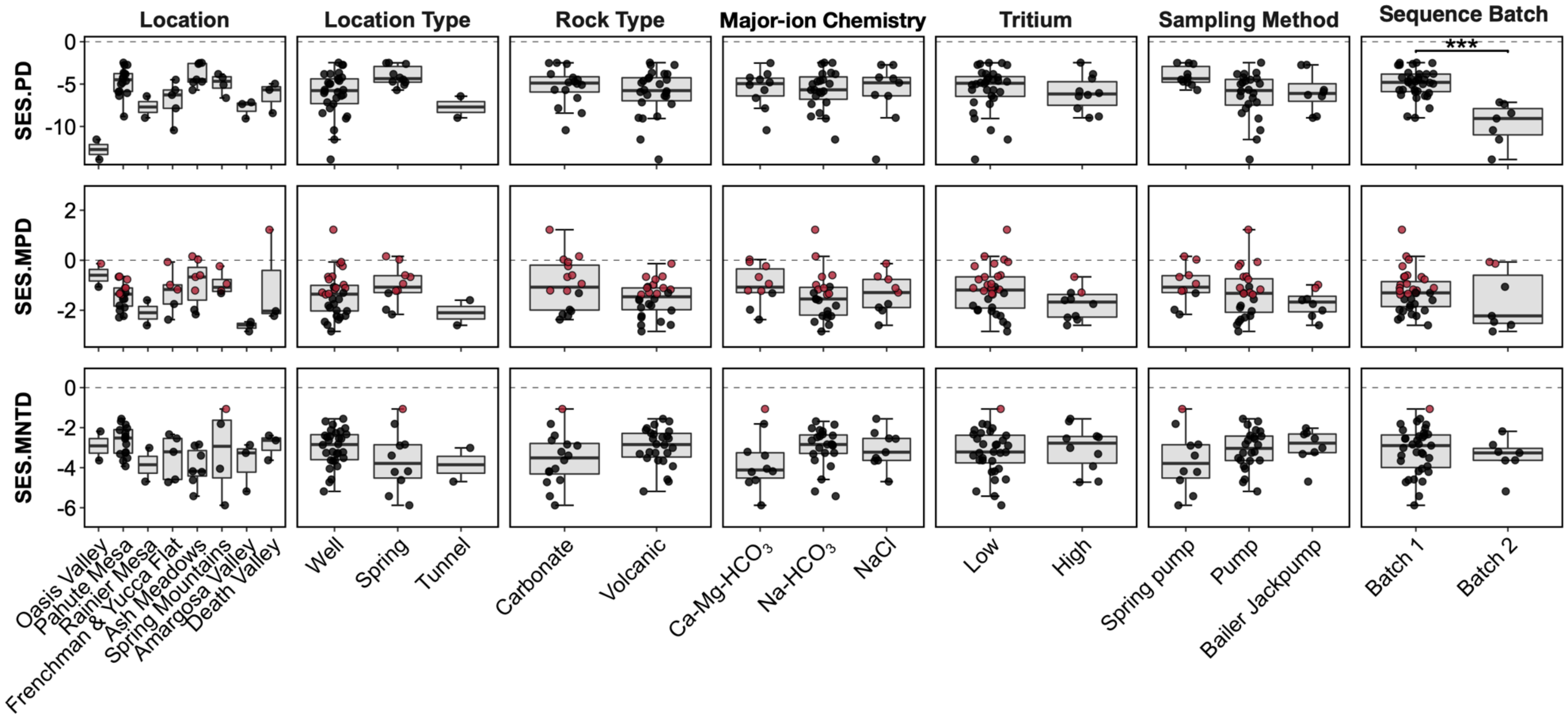


Figure 3. Standardized effect size (SES) of Faith's PD (SES-PD), mean pairwise distance (SES-MPD), and mean nearest neighbor phylogenetic distance (SES-MNTD). Asterisks indicate p-value significance: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Black circles indicate samples with communities that are significantly ($P < 0.05$) different from the null distribution, whereas red circles are not significantly different from the null distribution.

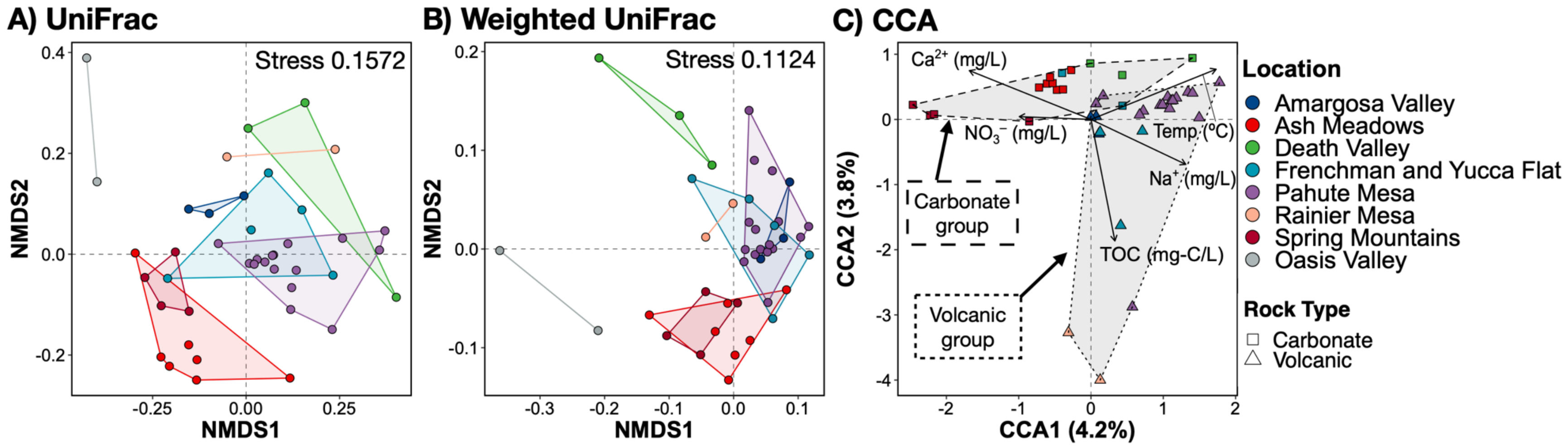


Figure 4. NMDS (Non-metric multidimensional scaling) ordinations of the planktonic microbiome. Beta diversity was evaluated using NMDS ordination [**(A)** unweighted UniFrac and **(B)** Weighted UniFrac] and **(C)** CCA ordination [with OV samples removed; vectors show the significant ($P < 0.05$) geochemical variables contributing to the variance]. Samples are colored by location. Rock type also is specified for CCA ordination.

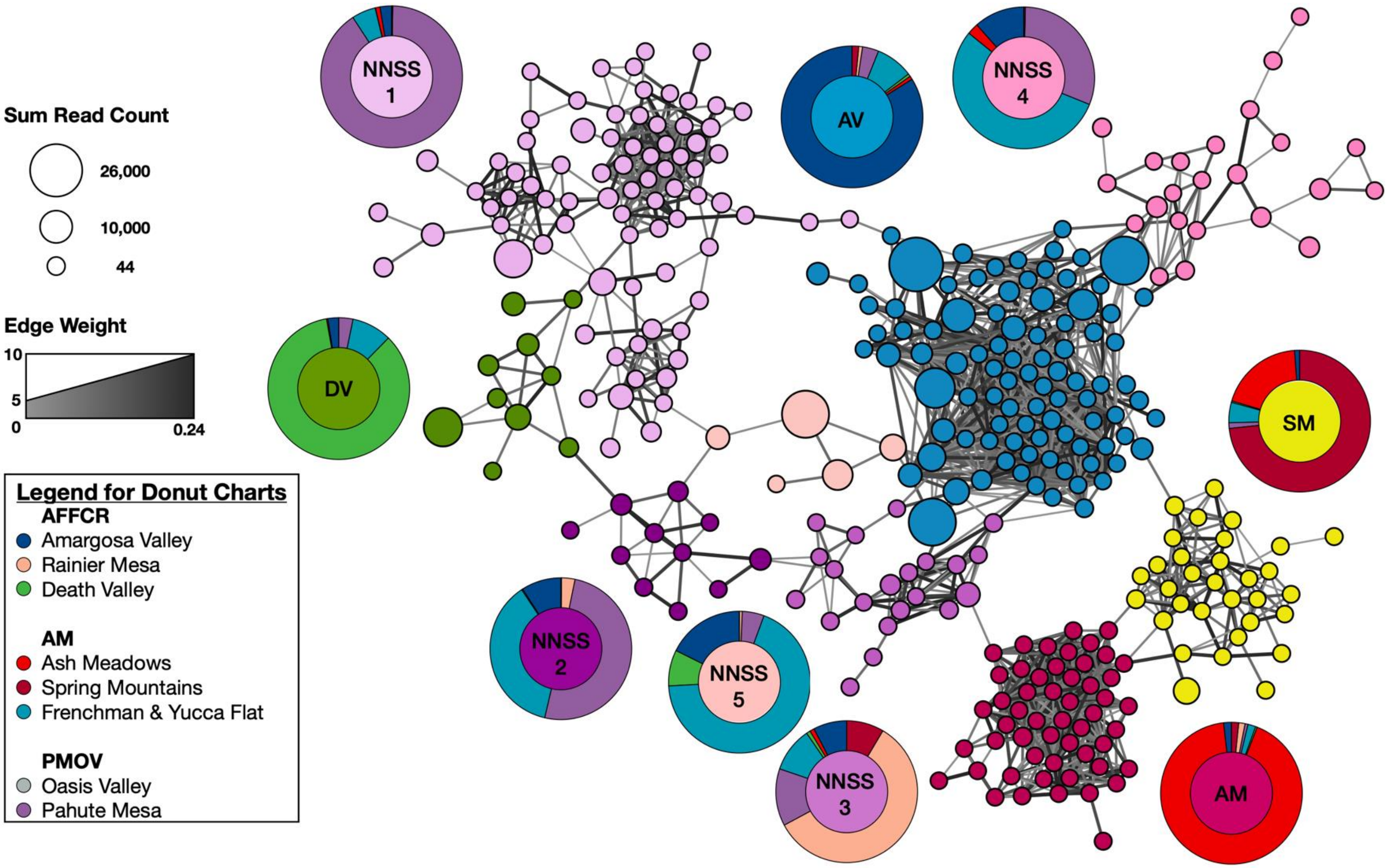


Figure 5. Co-occurring 'common OTUs' cluster by location. Nodes are colored by cluster and represent the overall location associated with a set of OTUs. Nine clusters were identified: NNSS 1 = Nevada National Security Site 1 (light purple; representative of Pahute Mesa communities), NNSS 2 (dark purple; Frenchman & Yucca Flat and Pahute Mesa), NNSS 3 (purple; Rainier Mesa), NNSS 4 (pink; Frenchman & Yucca Flat and Pahute Mesa), AV = Amargosa Valley (blue), AM = Ash Meadows (red), SM & OV = Spring Mountains and Oasis Valley (yellow), DV = Death Valley (green), and OV = Oasis Valley (grey). The ratio of reads per location for each cluster is represented by a donut chart, and **Figure S8** depicts the ratio of reads per location for each OTU (node). The size of the node is based on the total number of reads for that OTU. Edge weights are shown as thickness and darkness of the connection between two nodes (e.g., thicker and darker edges indicate a stronger connection).

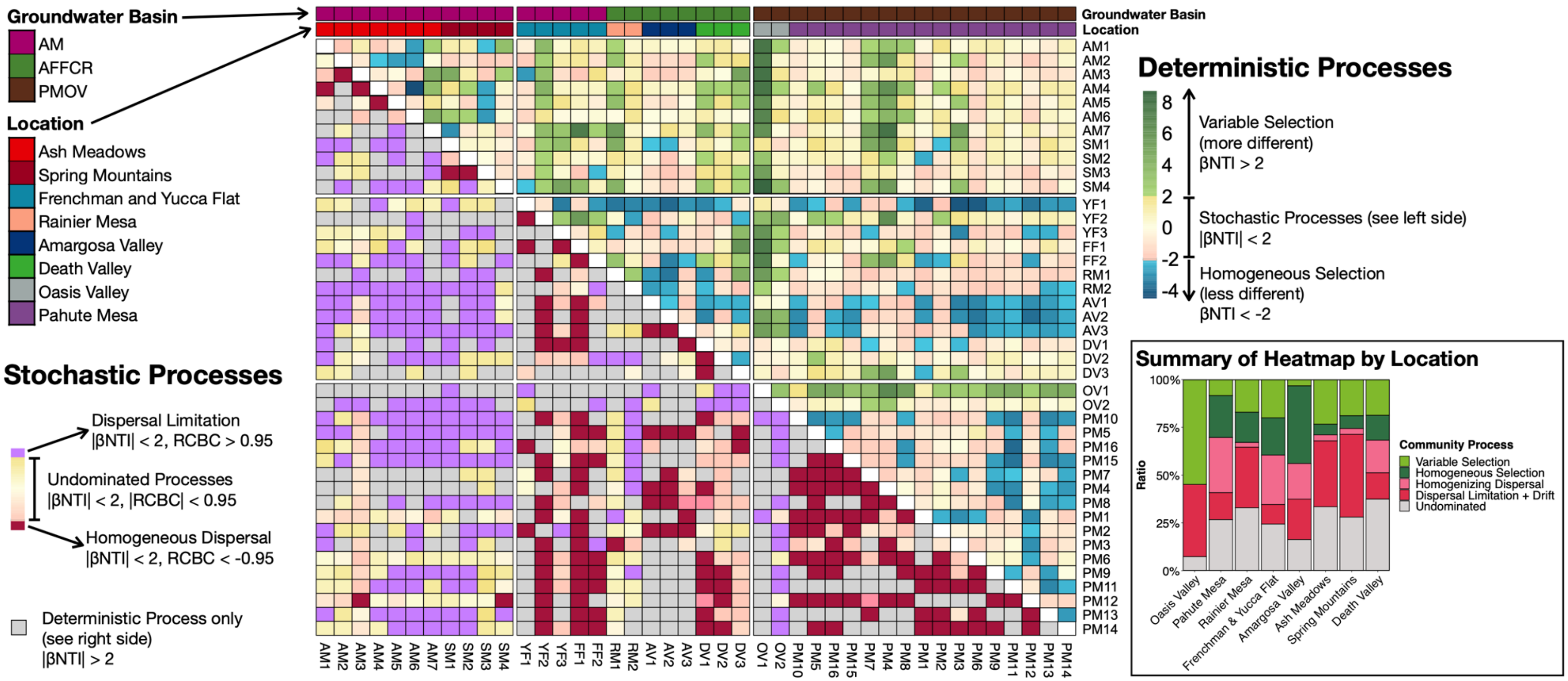


Figure 6. Heatmap of ecological processes dominating microbial community composition between locations. Deterministic and stochastic community assembly processes were quantified using null models (28–30). Deterministic processes (right side of heatmap) include variable selection (green; βNTI > 2) and homogeneous selection (blue; βNTI < –2). When |βNTI| < 2, the phylogenetic relatedness between two communities did not differ significantly than expected by chance, and stochastic processes dominate (see left side of heatmap). Stochastic processes include homogenizing dispersal (red; |βNTI| < 2 and RCBC < –0.95), dispersal limitation and drift (purple; |βNTI| < 2 and RCBC > 0.95), and undominated (|RCBC| < 0.95) processes (**Table S8**; **Table S1** for definitions). Ratios of the ecological processes were obtained for each overall location (inset).

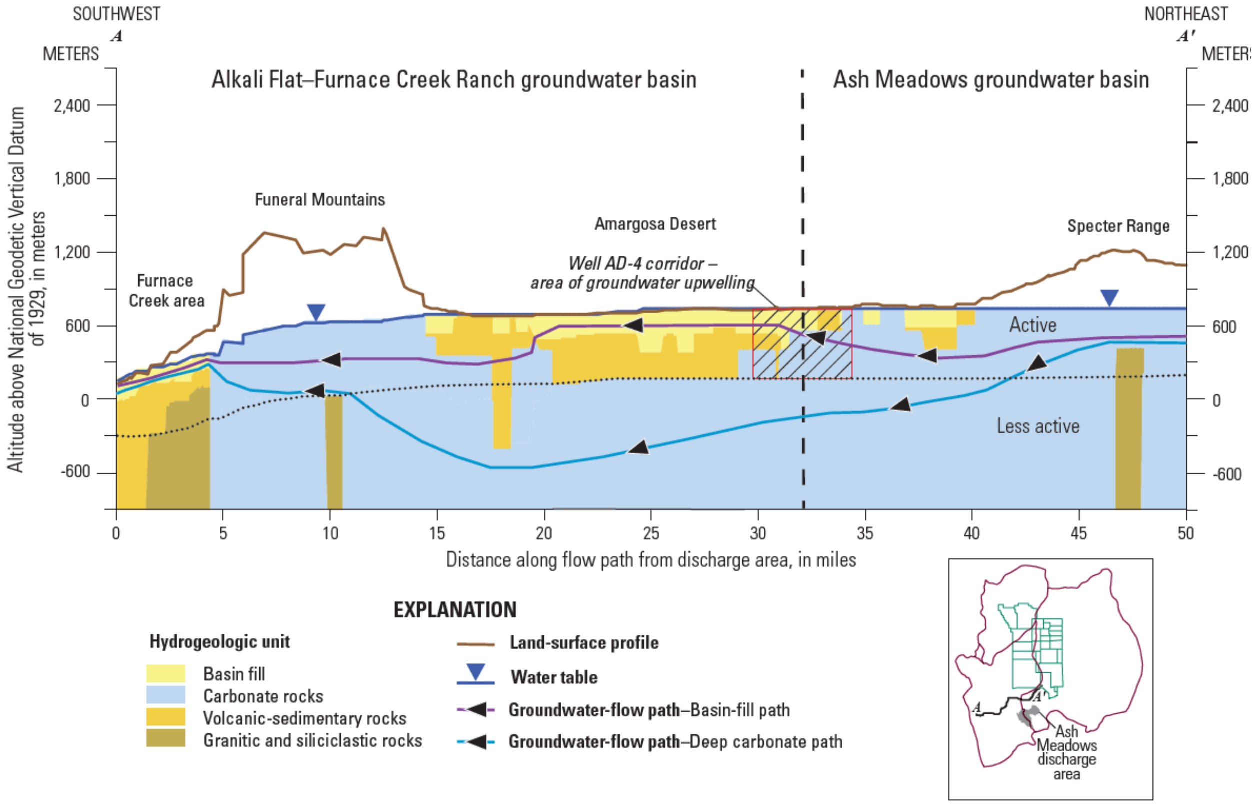


Figure 7. Conceptualized groundwater-flow paths towards Death Valley (Furnace Creek area). Figure modified from Halford and Jackson (2020).

Table 1. PERMANOVA of the weighted UniFrac beta diversity ordination using location and geochemical parameters.

Interest	[X][a]	Sum sqrs	R^2	Pseudo *F* statistic	P-value	Betadisper[b] P-value	ANOSIM[b] R-value	ANOSIM P-value
Location	**Location**	**0.488**	**0.266**	**2.104**	**0.003**	**0.168**	**0.418**	**0.001**
	Depth (m)[c]	**0.105**	**0.058**	**3.180**	**0.001**	**0.672**		
	Sampling method[d]	0.063	0.035	0.957	0.416	0.464	0.080	0.169
	Sequence batch[e]	0.039	0.021	1.177	0.244	0.662	0.190	0.091
	Residual	0.994	0.543					
	Total	1.831	1					
Geochemistry	**Rock type[f]**	**0.158**	**0.086**	**4.661**	**0.002**	**0.056**	**0.296**	**0.001**
	Temperature (°C)	**0.101**	**0.055**	**2.990**	**0.002**	**< 0.001**		
	TOC (mg-C/L)	**0.108**	**0.059**	**3.190**	**0.026**	**0.290**		
	pH	0.046	0.025	1.363	0.148	0.477		
	Tritium[g]	0.040	0.022	1.178	0.244	0.115	–0.117	0.887
	Sampling method[d]	**0.122**	**0.067**	**1.806**	**0.020**	**0.464**	**0.080**	**0.169**
	Sequence batch[e]	**0.063**	**0.034**	**1.847**	**0.039**	**0.662**	**0.190**	**0.091**
	Residual	1.119	0.611					
	Total	1.831	1					

[a] PERMANOVA was conducted on the weighted UniFrac beta diversity distribution (**Figure 4B**). [X] refers to the formula: adonis2([data] ~ [X], perm = 999, by = “margin”). [X] considers all listed variables for each ‘Interest’. [data] is the distance matrix derived from the weighted UniFrac. For example, adonis2([data] ~ Location + Depth + SamplingMethod + SequenceBatch, permutations = 999, by = “margin”). Bolded variables are significant (P < 0.05 for PERMANOVA and/or ANOSIM).

[b] Betadisper (homogeneity condition) and ANOSIM (analysis of similarity for categorical variables) R functions were used on individual variables.

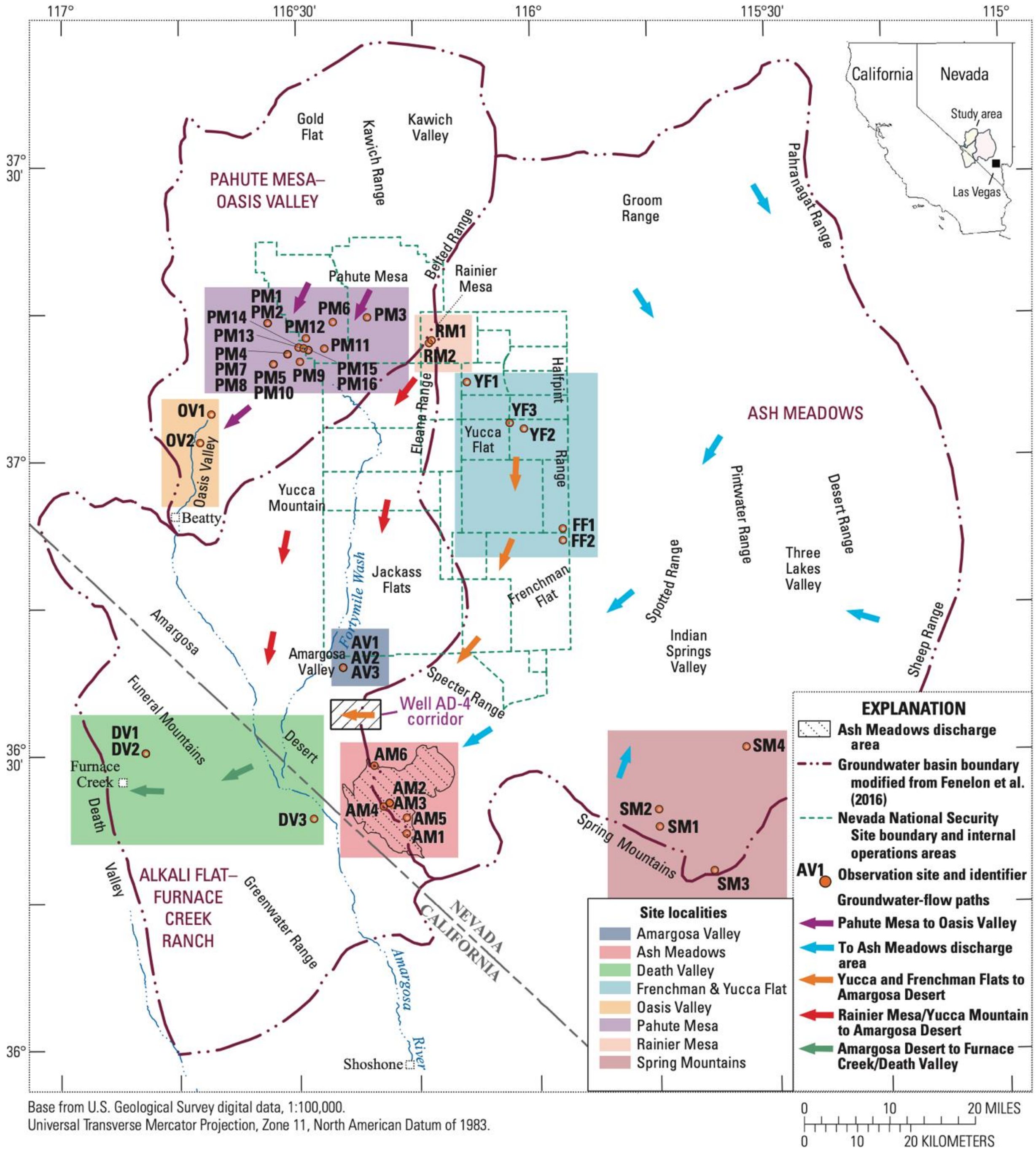
[c] Sample collection depth from the water table.

[d] Samples were collected by pump, ~~pump jacks~~jack pumps (in-line submerged pump), or bailer. For location variables, ‘sampling method’ was used as a proxy for ‘location type’ (well, spring, or tunnel) due to overlaps in metadata for spring and groundwater sampling methods.

[e] The majority of samples were sequenced in Batch 1 and seven samples sequenced in Batch 2 (**Table S2**).

[f] Samples were categorized by rock type based on host rock (carbonate or volcanic) at the depth sampled.

[g] Tritium concentrations were split into ‘High’ (≥100 Bq/L) and ‘Low’ (<100 Bq/L) categories



Base from U.S. Geological Survey digital data, 1:100,000.
Universal Transverse Mercator Projection, Zone 11, North American Datum of 1983.

Figure 1. Map of groundwater basins, locations, sampling sites, and groundwater-flow paths within the study area.

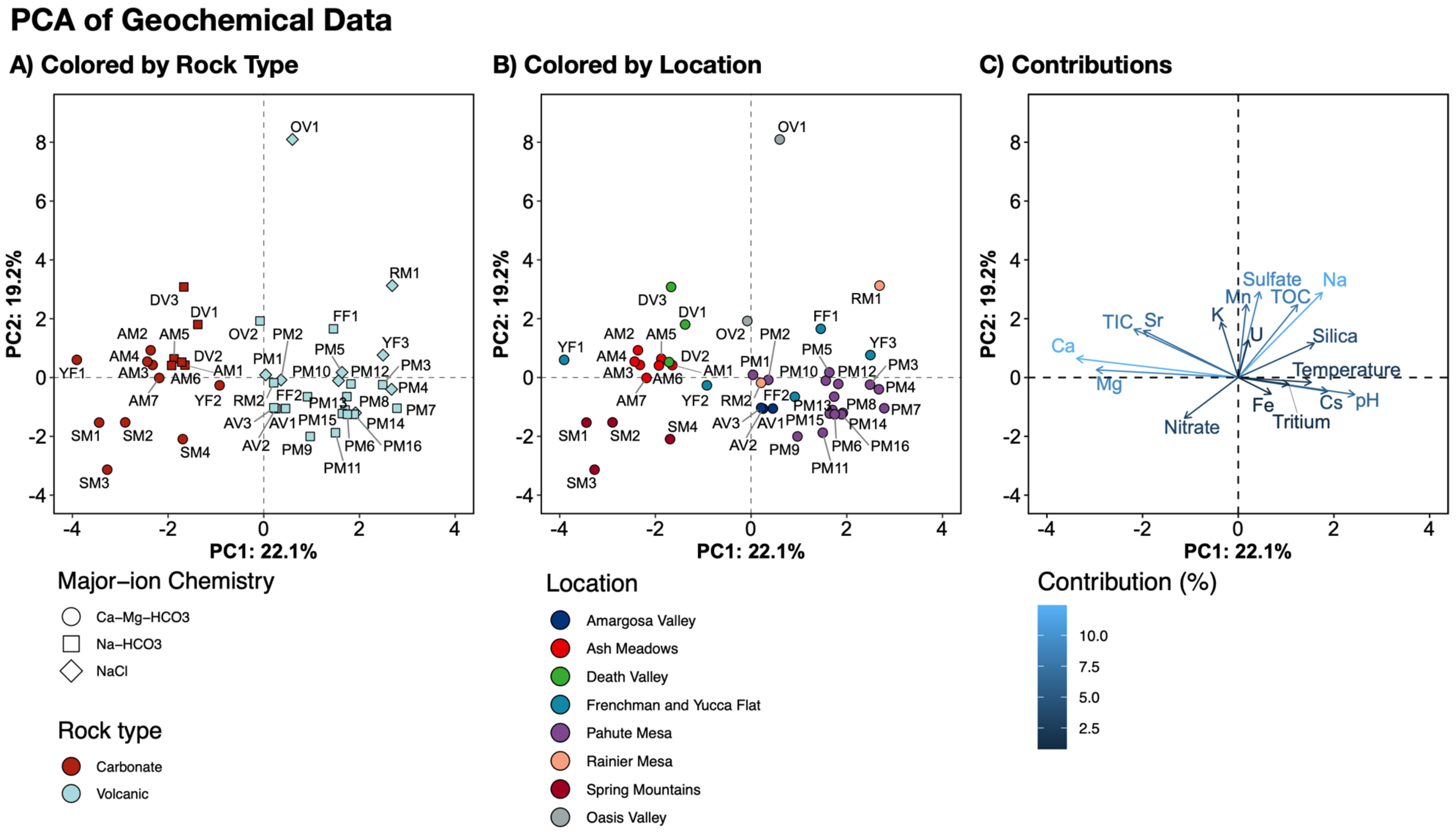


Figure 2. Principal components analysis (PCA) of geochemical data. (**A**) Variation of samples by major-ion chemistry and rock type. (**B**) Variation of samples by location. (**C**) Contribution of each geochemical variable (n=17) used in the PCA.

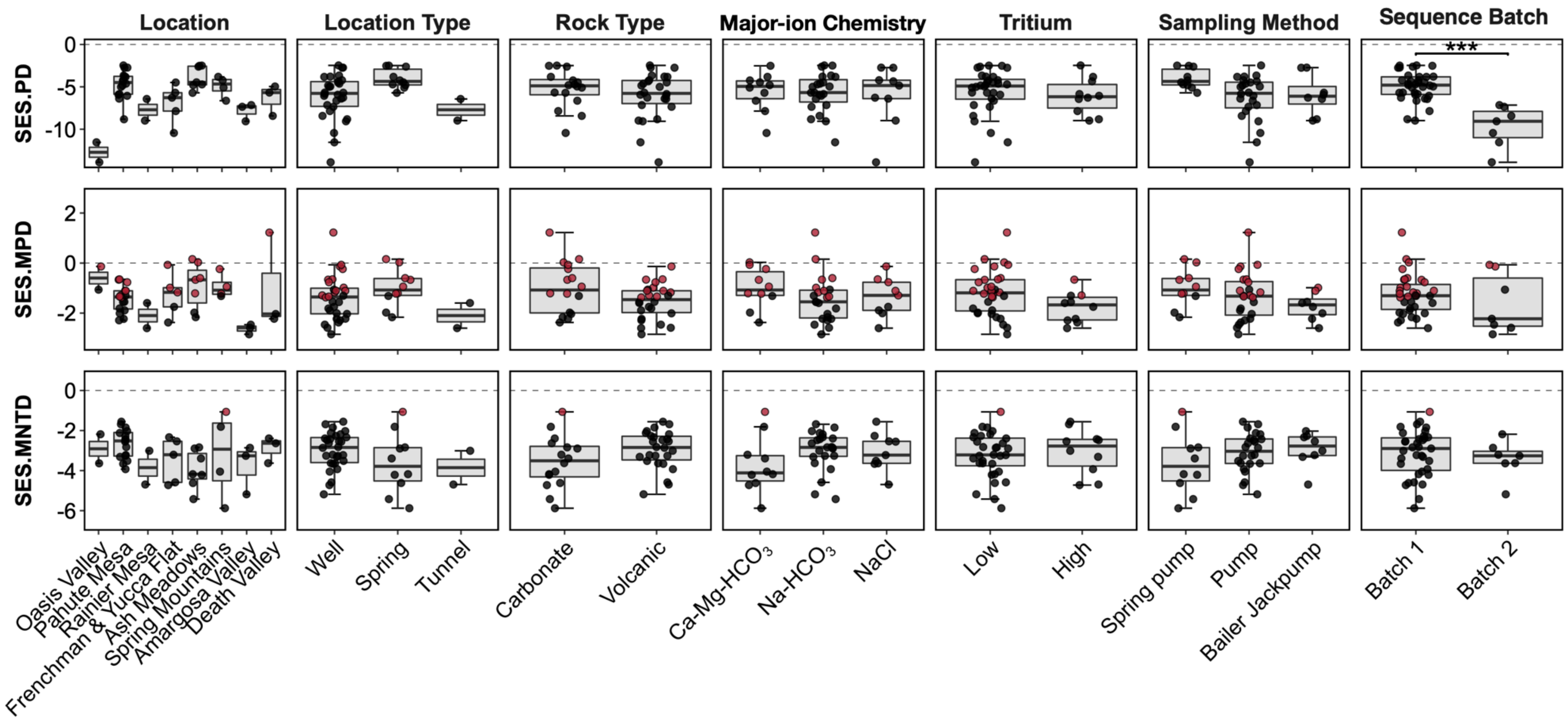
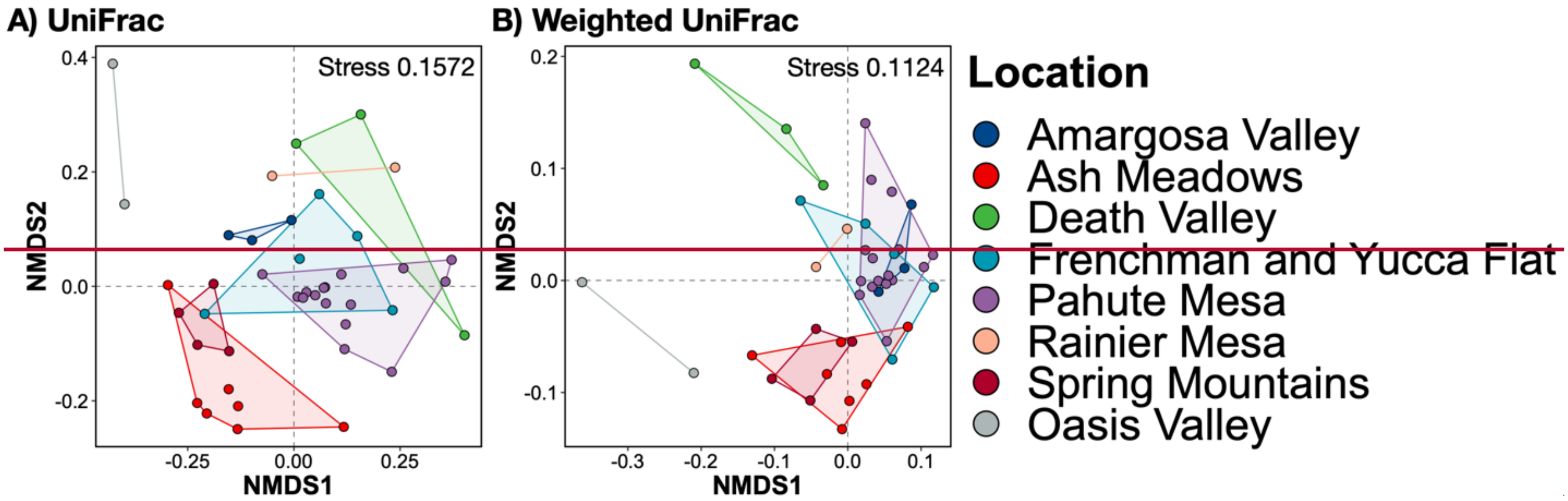


Figure 3. Standardized effect size (SES) of Faith's PD (SES-PD), mean pairwise distance (SES-MPD), and mean nearest neighbor phylogenetic distance (SES-MNTD). Asterisks indicate p-value significance: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Black circles indicate samples with communities that are significantly ($P < 0.05$) different from the null distribution, whereas red circles are not significantly different from the null distribution.



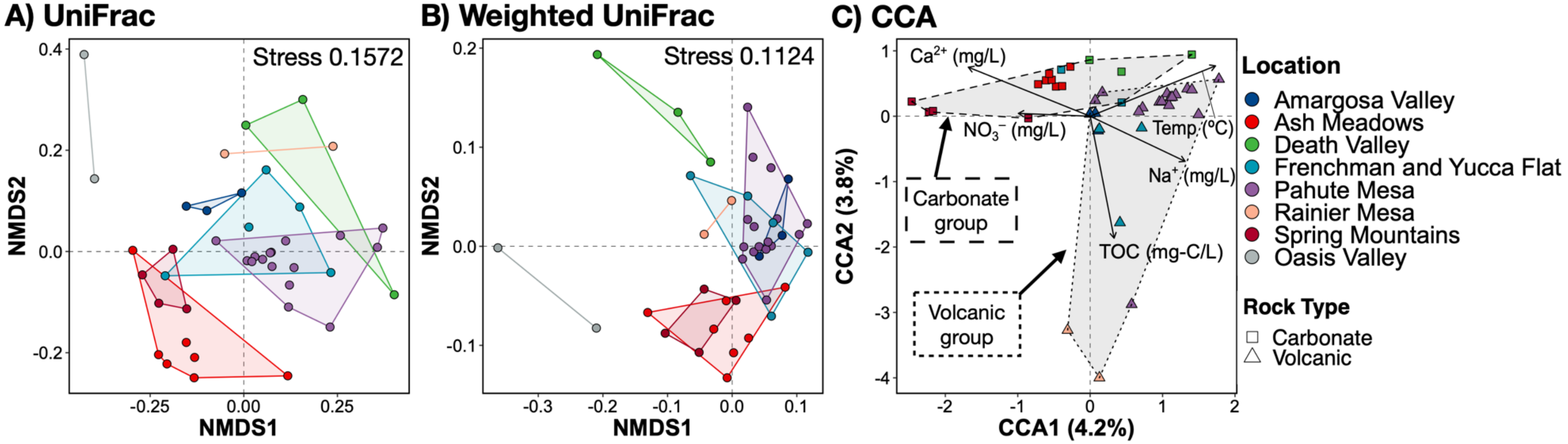
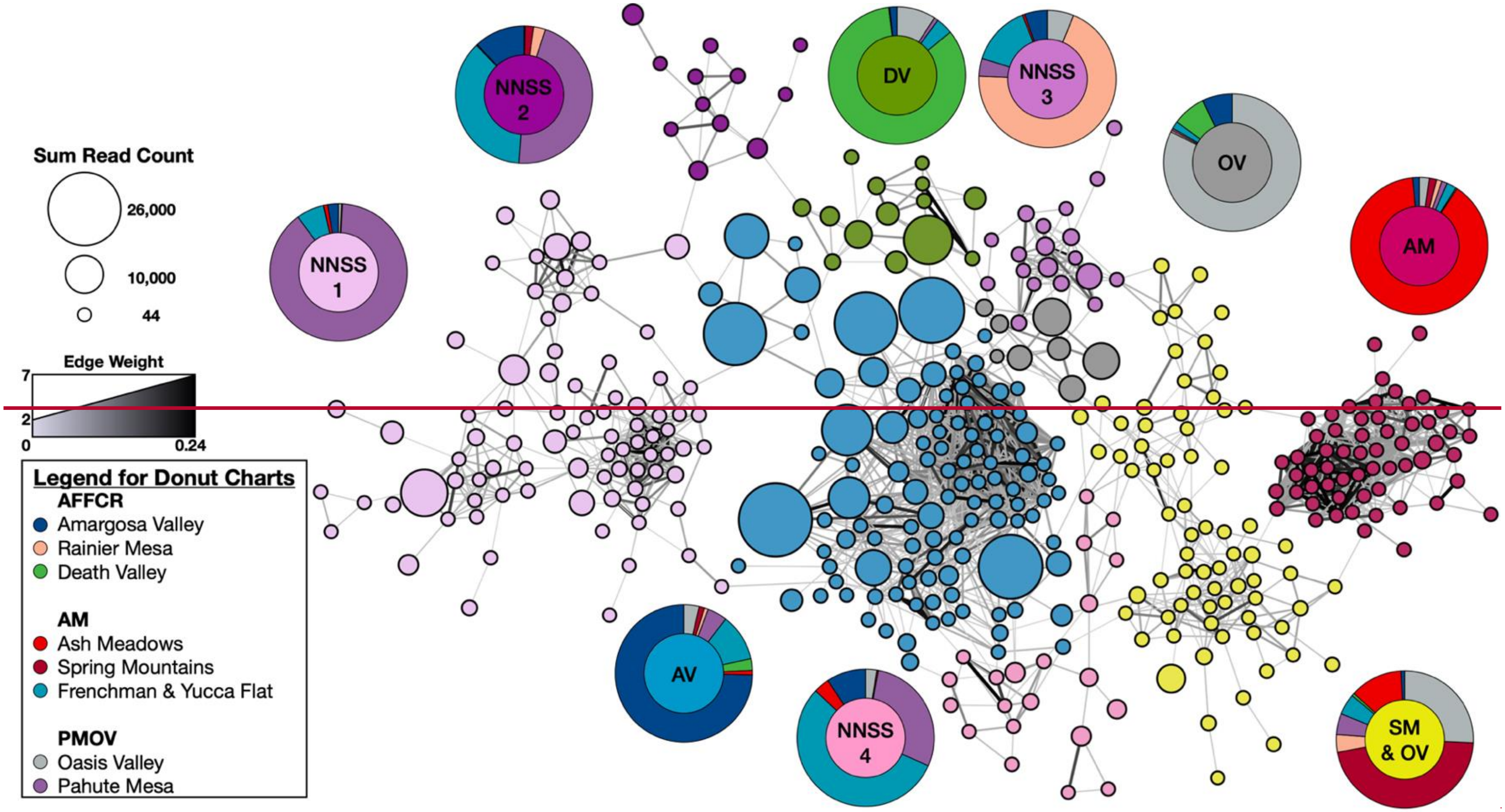


Figure 4. NMDS (Non-metric multidimensional scaling) ordinations of the planktonic microbiome. Beta diversity was evaluated using NMDS ordination ~~for the following distances:~~[**(A)** unweighted UniFrac and **(B)** Weighted UniFrac] and **(C)** CCA ordination [with OV samples removed; vectors show the significant ($P < 0.05$) geochemical variables contributing to the variance]. Samples are colored by location. Rock type also is specified for CCA ordination.



Sum Read Count
26,000
10,000
44
Edge Weight
7
2
0
0.24
Legend for Donut Charts
AFFCR
Amargosa Valley
Rainier Mesa
Death Valley
AM
Ash Meadows
Spring Mountains
Frenchman & Yucca Flat
PMOV
Oasis Valley
Pahute Mesa
NNSS 1
NNSS 2
DV
NNSS 3
OV
AM
AV
NNSS 4
SM & OV

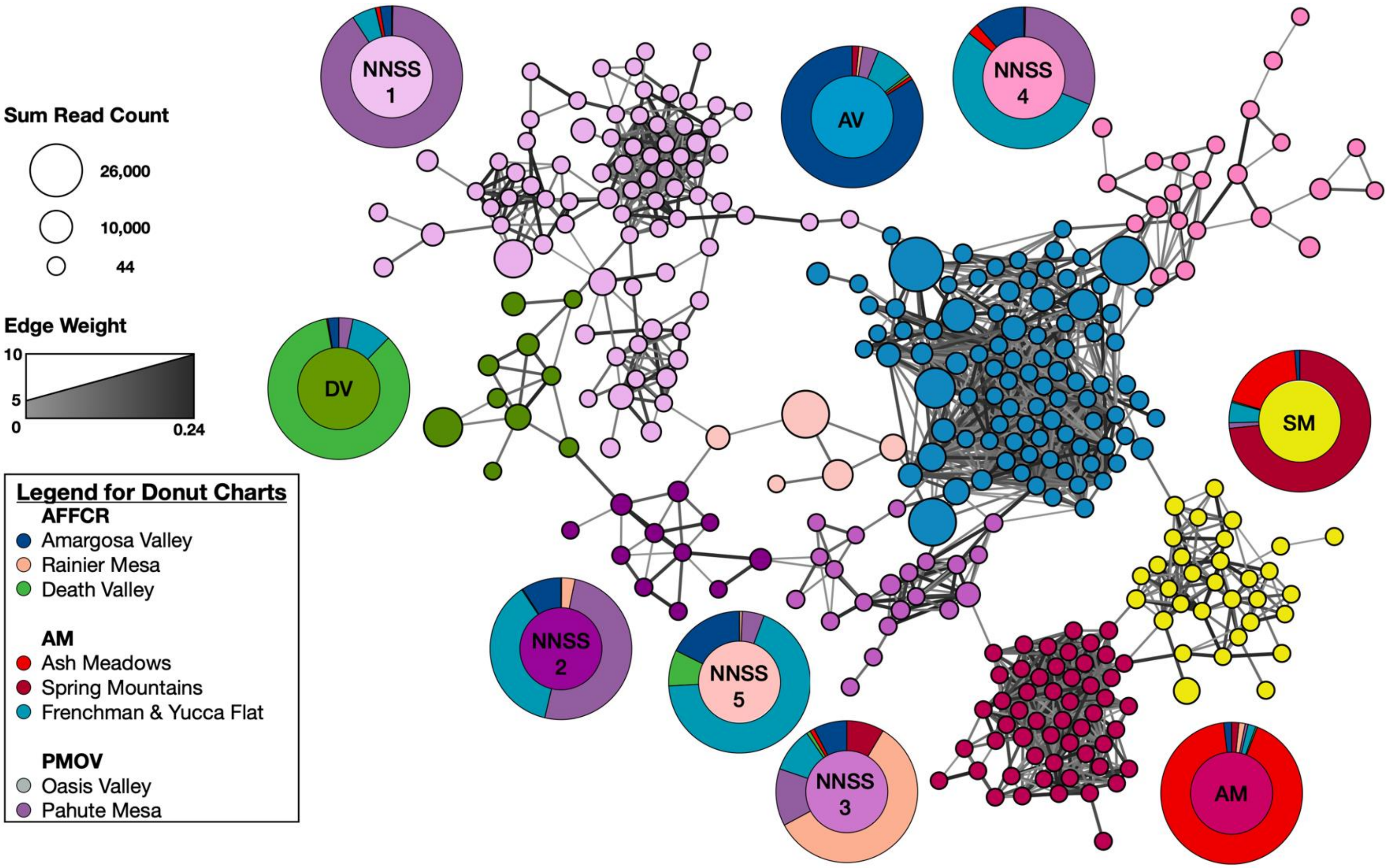


Figure 5. Co-occurring 'common OTUs' cluster by location. Nodes are colored by cluster and represent the overall location associated with a set of OTUs. Nine clusters were identified: NNSS 1 = Nevada National Security Site 1 (light purple; representative of Pahute Mesa communities), NNSS 2 (dark purple; Frenchman & Yucca Flat and Pahute Mesa), NNSS 3 (purple; Rainier Mesa), NNSS 4 (pink; Frenchman & Yucca Flat and Pahute Mesa), AV = Amargosa Valley (blue), AM = Ash Meadows (red), SM & OV = Spring Mountains and Oasis Valley (yellow), DV = Death Valley (green), and OV = Oasis Valley (grey). The ratio of reads per location for each cluster is represented by a donut chart, and **Figure S8** depicts the ratio of reads per location for each OTU (node). The size of the node is based on the total number of reads for that OTU. Edge weights are shown as thickness and darkness of the connection between two nodes (e.g., thicker and darker edges indicate a stronger connection).

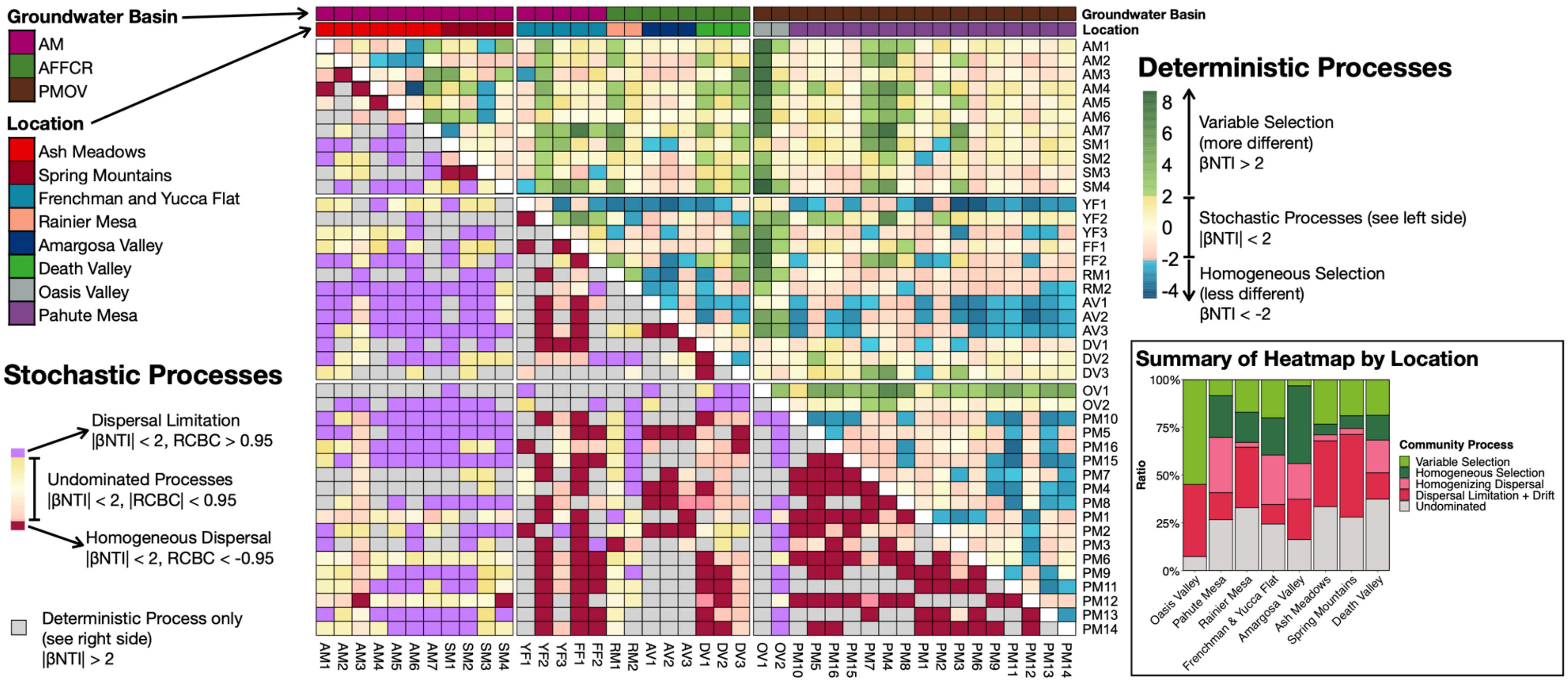


Figure 6. Heatmap of ecological processes dominating microbial community composition between locations. Deterministic and stochastic community assembly processes were quantified using null models (28–30). Deterministic processes (right side of heatmap) include variable selection (green; βNTI > 2) and homogeneous selection (blue; βNTI < –2). When |βNTI| < 2, the phylogenetic relatedness between two communities did not differ significantly than expected by chance, and stochastic processes dominate (see left side of heatmap). Stochastic processes include homogenizing dispersal (red; |βNTI| < 2 and RCBC < –0.95), dispersal limitation and drift (purple; |βNTI| < 2 and RCBC > 0.95), and undominated (|RCBC| < 0.95) processes (**Table S8**; **Table S1** for definitions). Ratios of the ecological processes were obtained for each overall location (inset).

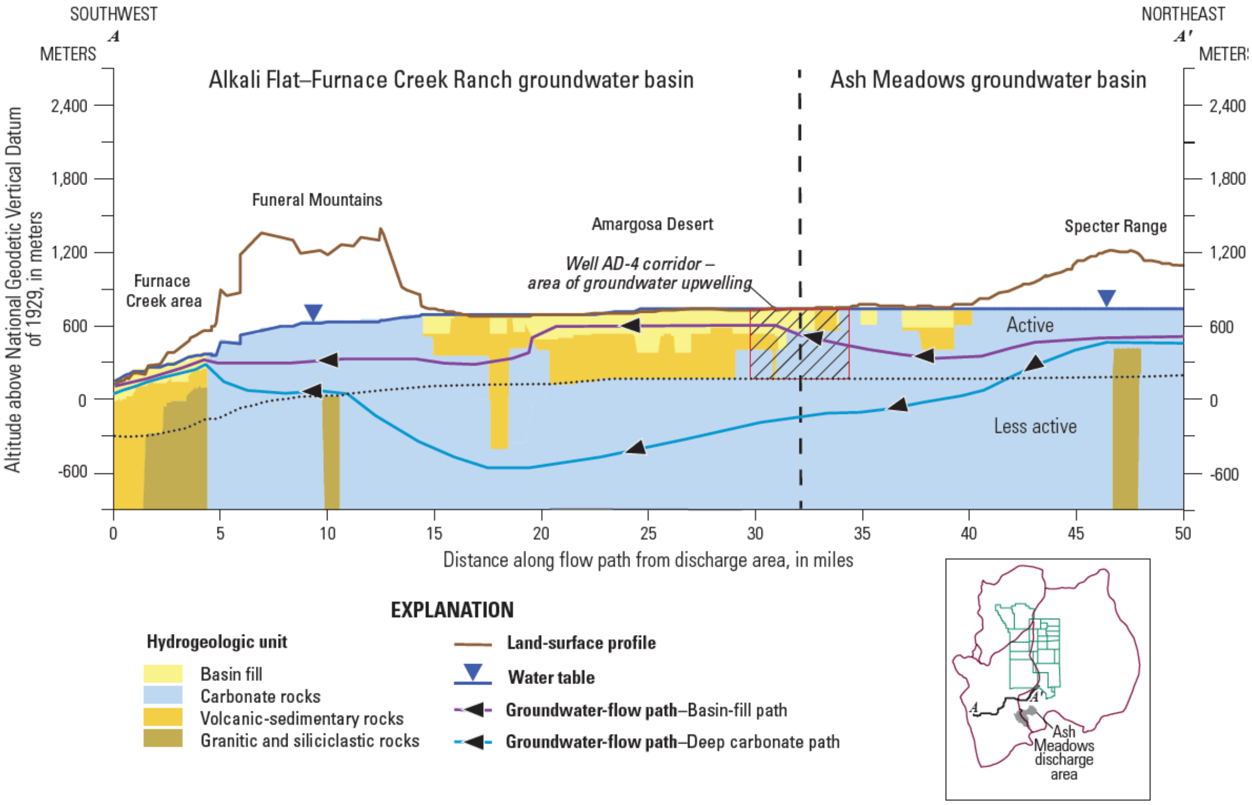


Figure 7. Conceptualized groundwater-flow paths towards Death Valley (Furnace Creek area). Figure modified from Halford and Jackson (2020).

Figure 4 - new

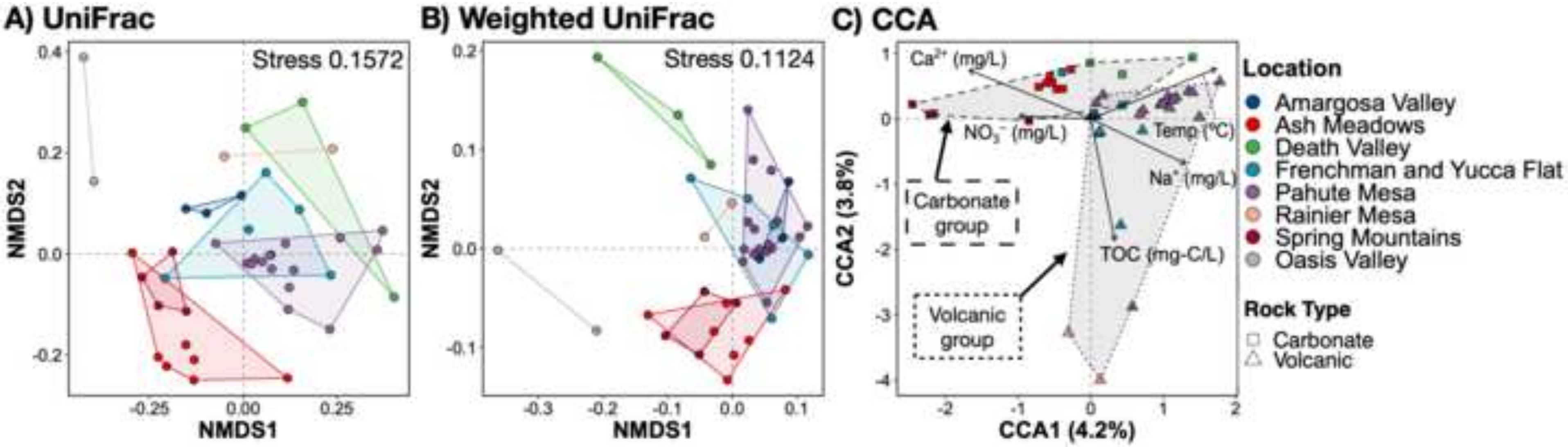

A) UniFrac
Stress 0.1572
NMDS2
NMDS1
0.4
0.2
0.0
-0.2
-0.25
0.00
0.25
B) Weighted UniFrac
Stress 0.1124
NMDS2
NMDS1
0.2
0.1
0.0
-0.1
-0.3
-0.2
-0.1
0.0
0.1
C) CCA
CCA2 (3.8%)
CCA1 (4.2%)
1
0
-1
-2
-3
-4
-2
-1
0
1
2
Ca²⁺ (mg/L)
NO₃⁻ (mg/L)
Temp (°C)
Na⁺ (mg/L)
TOC (mg-C/L)
Carbonate group
Volcanic group
Location
Amargosa Valley
Ash Meadows
Death Valley
Frenchman and Yucca Flat
Pahute Mesa
Rainier Mesa
Spring Mountains
Oasis Valley
Rock Type
Carbonate
Volcanic

Figure 5 - new

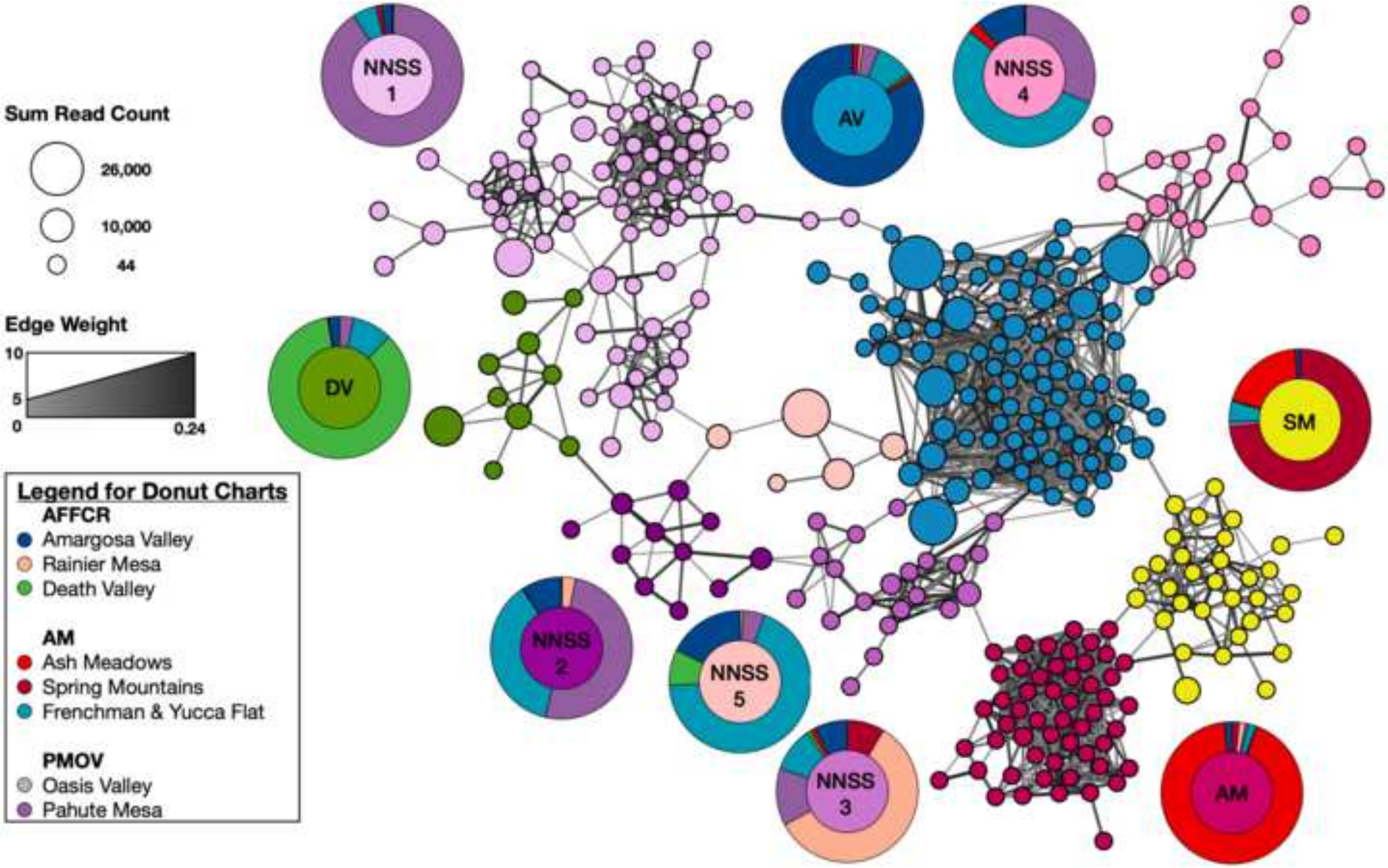
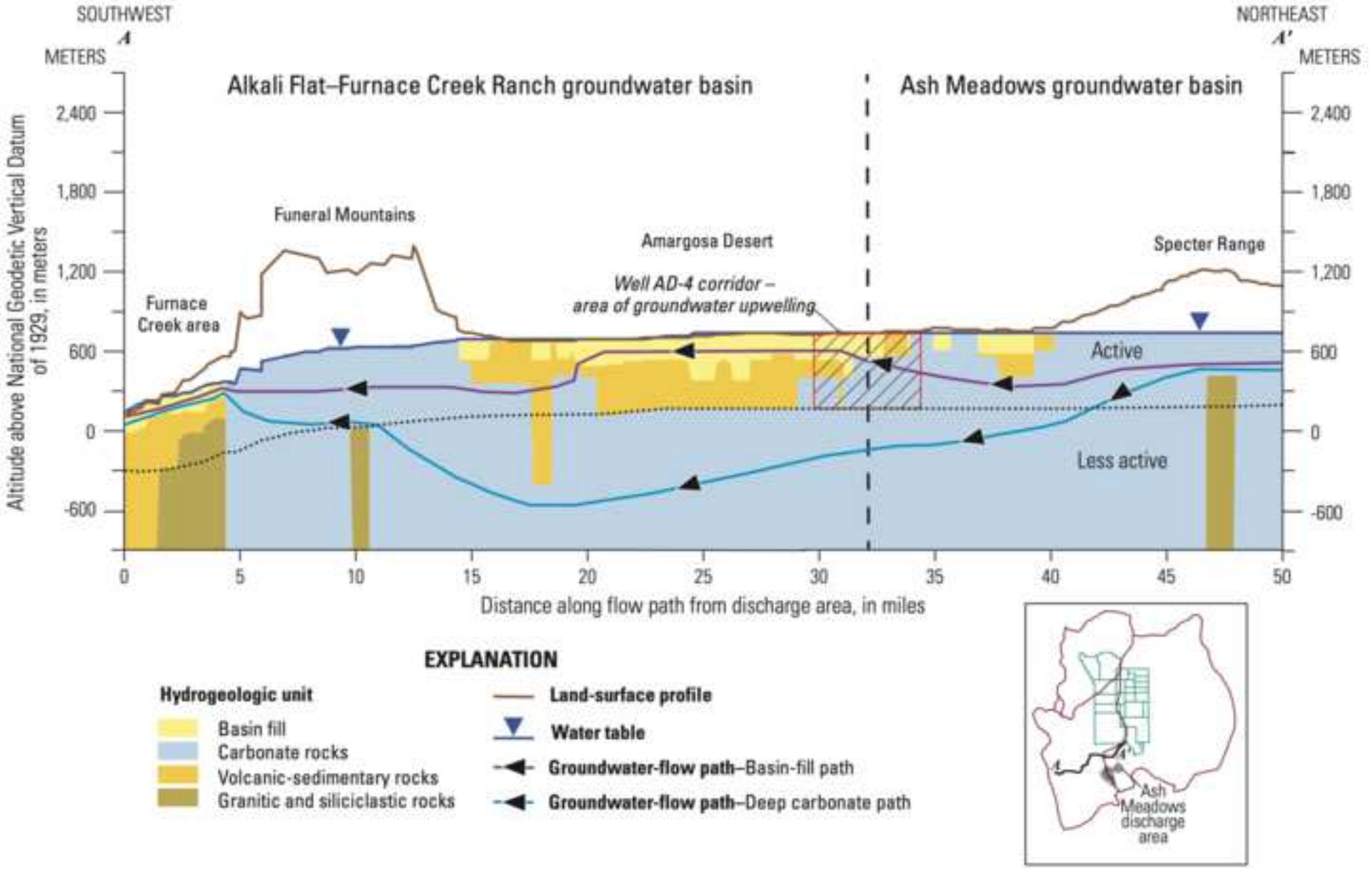
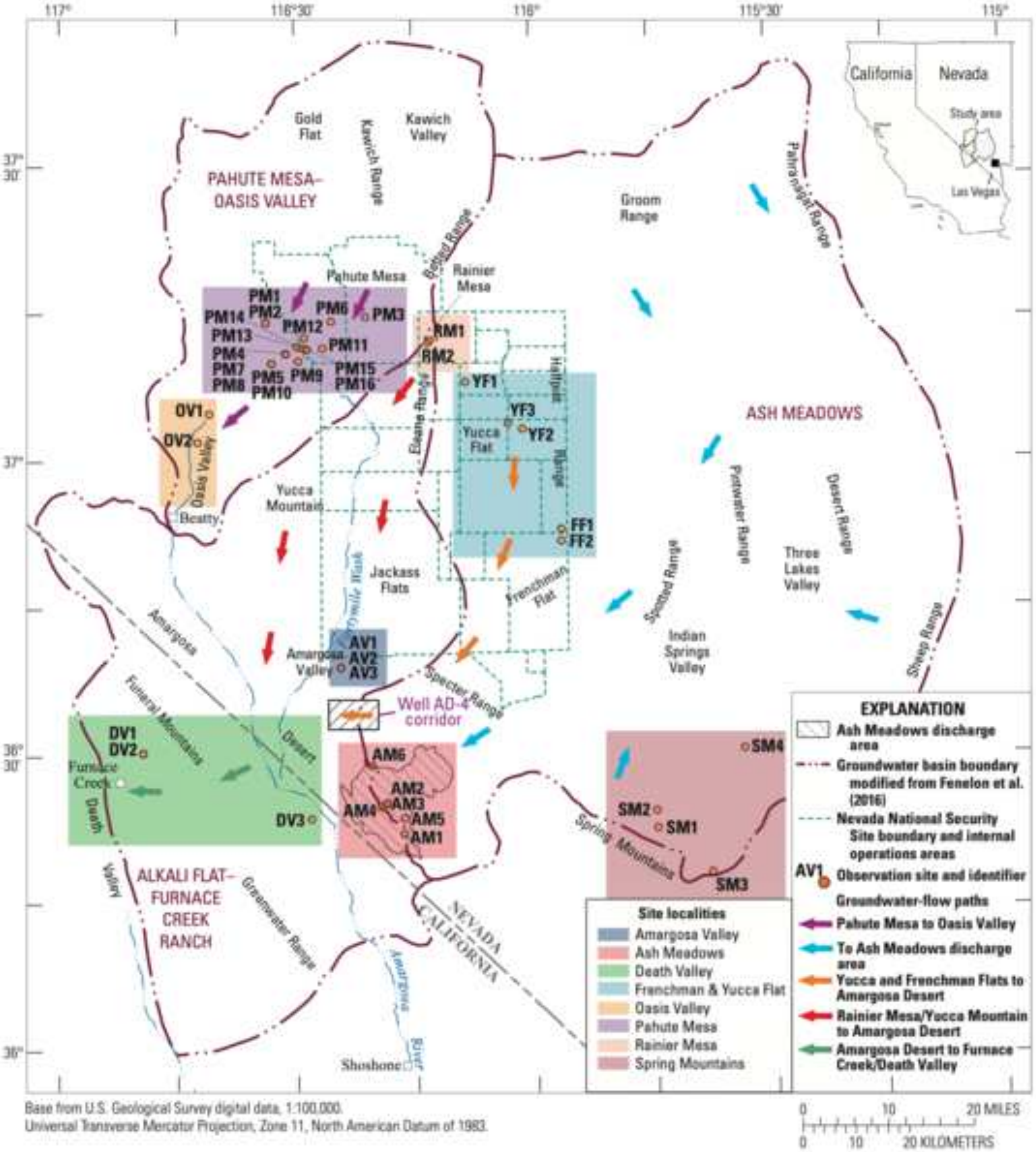

Sum Read Count
26,000
10,000
44
Edge Weight
10
5
0
0.24
Legend for Donut Charts
AFFCR
Amargosa Valley
Rainier Mesa
Death Valley
AM
Ash Meadows
Spring Mountains
Frenchman & Yucca Flat
PMOV
Oasis Valley
Pahute Mesa
NNSS 1
AV
NNSS 4
DV
SM
NNSS 2
NNSS 5
NNSS 3
AM

Figure 7


SOUTHWEST
A
METERS
NORTHEAST
A'
METERS
Alkali Flat–Furnace Creek Ranch groundwater basin
Ash Meadows groundwater basin
2,400
1,800
1,200
600
0
-600
Altitude above National Geodetic Vertical Datum of 1929, in meters
Funeral Mountains
Furnace Creek area
Amargosa Desert
Well AD-4 corridor – area of groundwater upwelling
Specter Range
Active
Less active
0
5
10
15
20
25
30
35
40
45
50
Distance along flow path from discharge area, in miles
EXPLANATION
Hydrogeologic unit
Basin fill
Carbonate rocks
Volcanic-sedimentary rocks
Granitic and siliciclastic rocks
Land-surface profile
Water table
Groundwater-flow path–Basin-fill path
Groundwater-flow path–Deep carbonate path
A
A'
Ash Meadows discharge area



Base from U.S. Geological Survey digital data, 1:100,000.
Universal Transverse Mercator Projection, Zone 11, North American Datum of 1983.

Figure 2

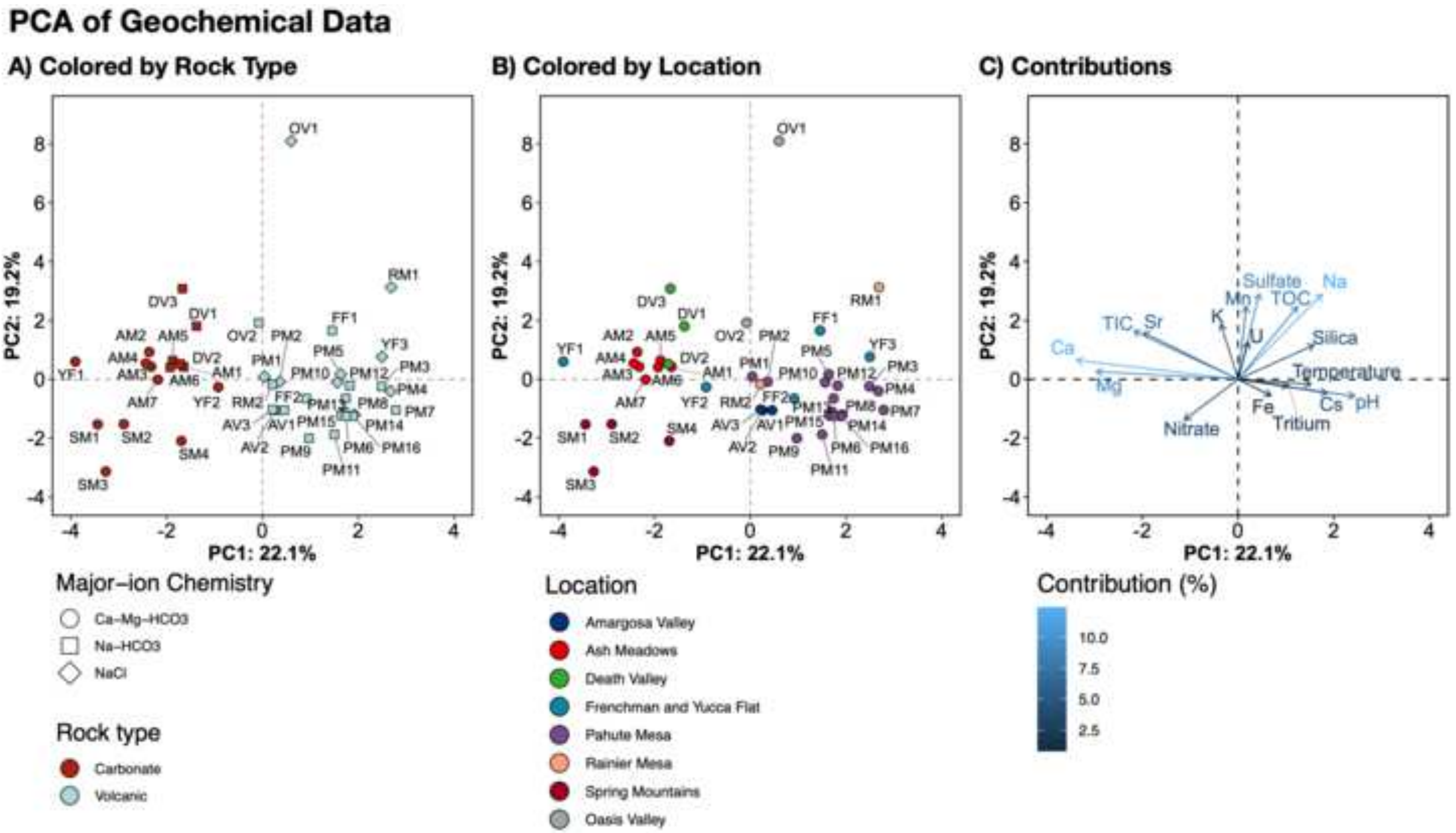
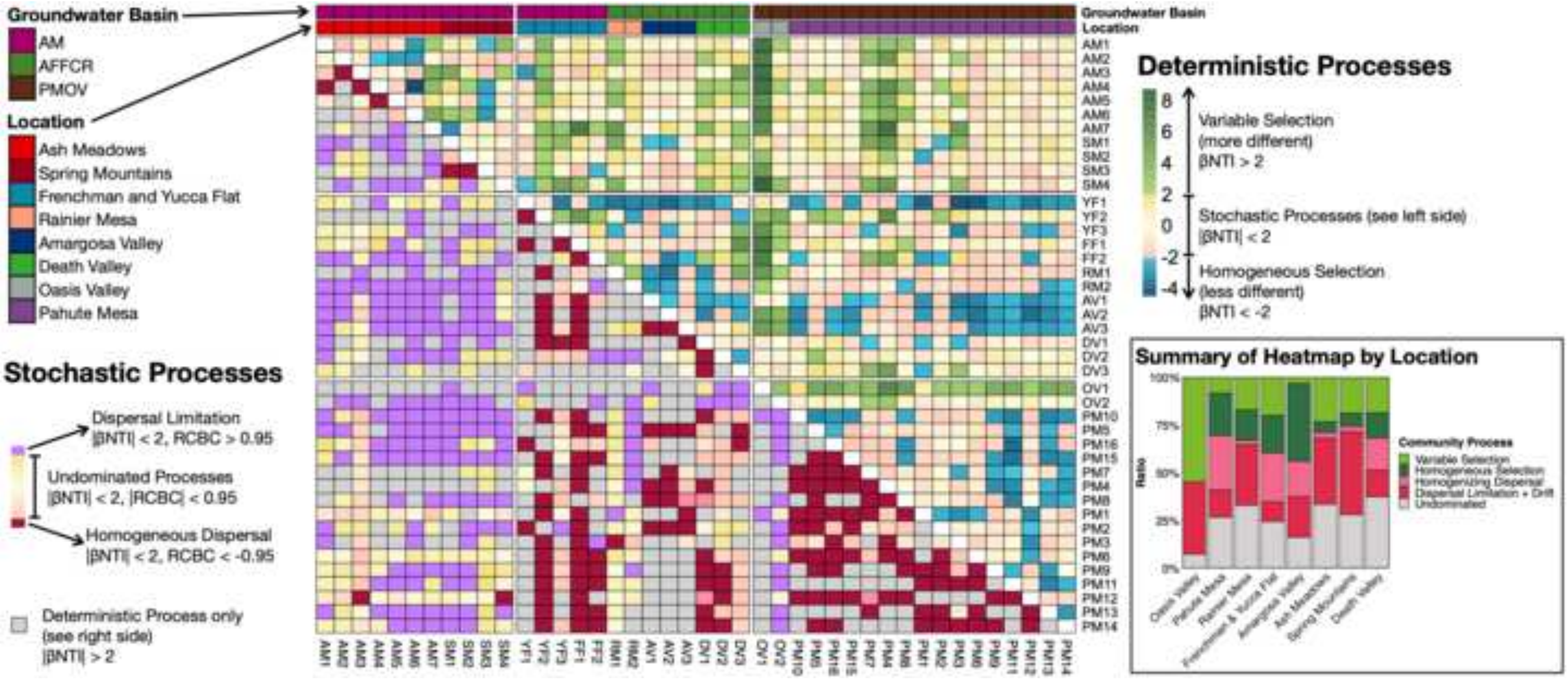
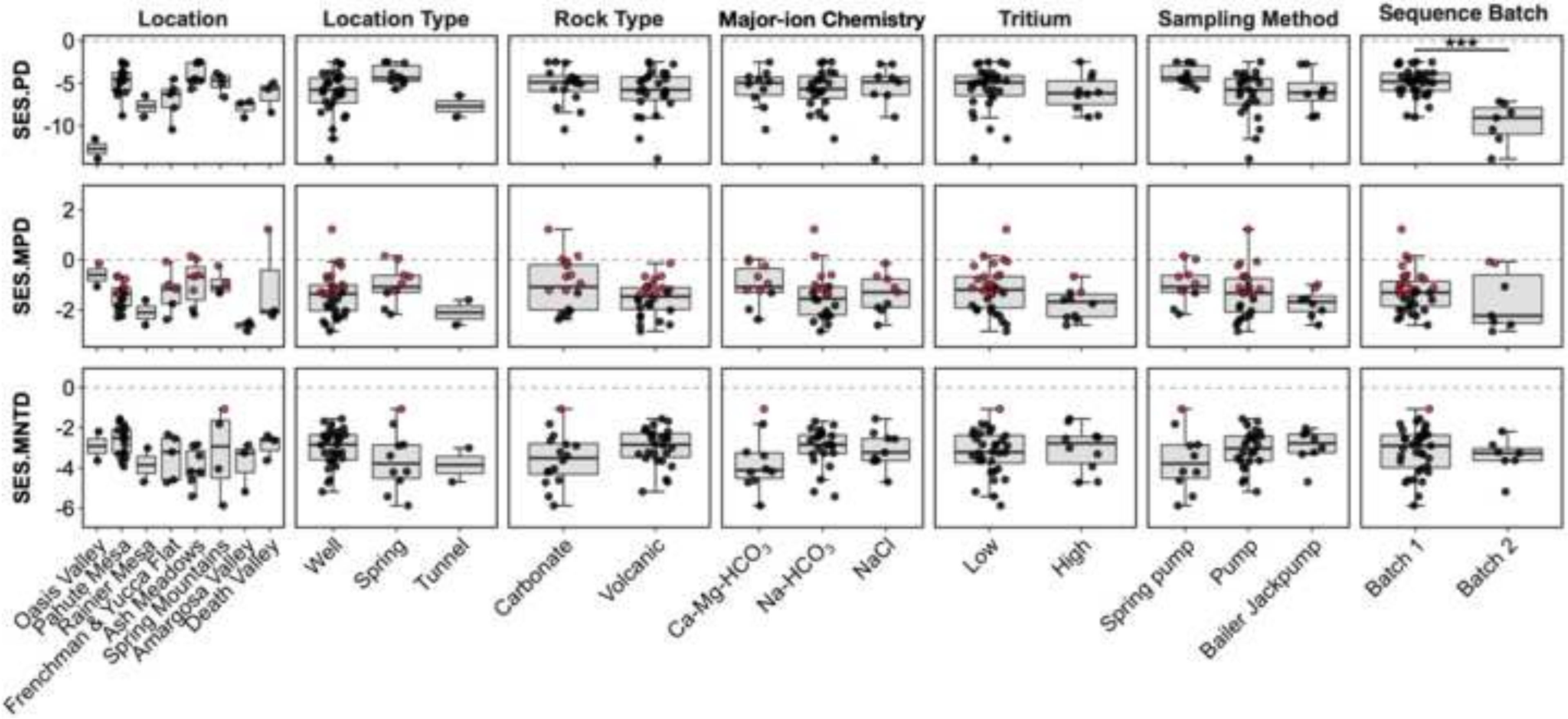

PCA of Geochemical Data
A) Colored by Rock Type
B) Colored by Location
C) Contributions
PC2: 19.2%
PC1: 22.1%
8
6
4
2
0
-2
-4
OV1
RM1
DV3
DV1
FF1
AM2
AM5
OV2
PM2
YF3
PM5
AM4
DV2
PM1
AM1
PM10
PM12
PM3
YF1
AM3
AM6
PM4
AM7
YF2
RM2
FF2
PM13
PM8
PM7
AV3
AV1
PM15
PM14
SM1
SM2
SM4
AV2
PM9
PM6
PM16
PM11
SM3
Sulfate
Na
Mn
TOC
K
TIC
Sr
U
Silica
Ca
Mg
Temperature
Fe
Cs
pH
Nitrate
Tribum
Major–ion Chemistry
Ca–Mg–HCO3
Na–HCO3
NaCl
Rock type
Carbonate
Volcanic
Location
Amargosa Valley
Ash Meadows
Death Valley
Frenchman and Yucca Flat
Pahute Mesa
Rainier Mesa
Spring Mountains
Oasis Valley
Contribution (%)
10.0
7.5
5.0
2.5

Figure 6


Groundwater Basin
AM
AFFCR
PMOV
Location
Ash Meadows
Spring Mountains
Frenchman and Yucca Flat
Rainier Mesa
Amargosa Valley
Death Valley
Oasis Valley
Pahute Mesa
Stochastic Processes
Dispersal Limitation
|βNTI| < 2, RCBC > 0.95
Undominated Processes
|βNTI| < 2, |RCBC| < 0.95
Homogeneous Dispersal
|βNTI| < 2, RCBC < -0.95
Deterministic Process only
(see right side)
|βNTI| > 2
Groundwater Basin
Location
AM1
AM2
AM3
AM4
AM5
AM6
AM7
SM1
SM2
SM3
SM4
YF1
YF2
YF3
FF1
FF2
RM1
RM2
AV1
AV2
AV3
DV1
DV2
DV3
OV1
OV2
PM10
PM5
PM16
PM15
PM7
PM4
PM8
PM1
PM2
PM3
PM6
PM9
PM11
PM12
PM13
PM14
Deterministic Processes
8
6
4
2
0
-2
-4
Variable Selection
(more different)
βNTI > 2
Stochastic Processes (see left side)
|βNTI| < 2
Homogeneous Selection
(less different)
βNTI < -2
Summary of Heatmap by Location
100%
75%
50%
25%
0%
Ratio
Oasis Valley
Pahute Mesa
Rainier Mesa
Frenchman & Yucca Flat
Amargosa Valley
Ash Meadows
Spring Mountains
Death Valley
Community Process
Variable Selection
Homogeneous Selection
Homogenizing Dispersal
Dispersal Limitation + Drift
Undominated


Location
Location Type
Rock Type
Major-ion Chemistry
Tritium
Sampling Method
Sequence Batch
***
SES.PD
0
-5
-10
SES.MPD
2
0
-2
SES.MNTD
0
-2
-4
-6
Oasis Valley
Pahute Mesa
Rainier Mesa
Frenchman & Yucca Flat
Ash Meadows
Spring Mountains
Amargosa Valley
Death Valley
Well
Spring
Tunnel
Carbonate
Volcanic
Ca-Mg-HCO₃
Na-HCO₃
NaCl
Low
High
Spring pump
Pump
Bailer Jackpump
Batch 1
Batch 2

Revised supplementary material - no changes marked

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Author Contributions: NM analyzed the 16S rRNA gene amplicon sequences and compiled the geochemical data. DPM served as principal investigator for the projects under which the samples were collected. SHB, JCB, JS, and JCF contributed to the collection, processing, and extraction of the samples. All authors wrote the manuscript.