

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

Spatial turnover of soil viral populations and genotypes overlain by cohesive responses to moisture in grasslands

### Permalink

<https://escholarship.org/uc/item/1z06d8p5>

### Journal

Proceedings of the National Academy of Sciences of the United States of America, 119(45)

### ISSN

0027-8424

### Authors

Santos-Medellín, Christian  
Esterá-Molina, Katerina  
Yuan, Mengting  
[et al.](#)

### Publication Date

2022-11-08

### DOI

10.1073/pnas.2209132119

Peer reviewed



# Spatial turnover of soil viral populations and genotypes overlain by cohesive responses to moisture in grasslands

Christian Santos-Medellín<sup>a</sup>, Katerina Estera-Molina<sup>b</sup>, Mengting Yuan<sup>b</sup>, Jennifer Pett-Ridge<sup>c,d</sup>, Mary K. Firestone<sup>b</sup>, and Joanne B. Emerson<sup>a,1</sup>

Edited by Edward DeLong, University of Hawai'i at Manoa, Honolulu, HI; received May 26, 2022; accepted October 9, 2022

Viruses shape microbial communities, food web dynamics, and carbon and nutrient cycling in diverse ecosystems. However, little is known about the patterns and drivers of viral community composition, particularly in soil, precluding a predictive understanding of viral impacts on terrestrial habitats. To investigate soil viral community assembly processes, here we analyzed 43 soil viromes from a rainfall manipulation experiment in a Mediterranean grassland in California. We identified 5,315 viral populations (viral operational taxonomic units [vOTUs] with a representative sequence  $\geq 10$  kbp) and found that viral community composition exhibited a highly significant distance–decay relationship within the 200-m<sup>2</sup> field site. This pattern was recapitulated by the intrapopulation microheterogeneity trends of prevalent vOTUs (detected in  $\geq 90\%$  of the viromes), which tended to exhibit negative correlations between spatial distance and the genomic similarity of their predominant allelic variants. Although significant spatial structuring was also observed in the bacterial and archaeal communities, the signal was dampened relative to the viromes, suggesting differences in local assembly drivers for viruses and prokaryotes and/or differences in the temporal scales captured by viromes and total DNA. Despite the overwhelming spatial signal, evidence for environmental filtering was revealed in a protein-sharing network analysis, wherein a group of related vOTUs predicted to infect actinobacteria was shown to be significantly enriched in low-moisture samples distributed throughout the field. Overall, our results indicate a highly diverse, dynamic, active, and spatially structured soil virosphere capable of rapid responses to changing environmental conditions.

soil virome | soil microbiome | distance–decay relationship | relic DNA

With an estimated area of 52.5 million km<sup>2</sup> (1), grasslands are major contributors to the cycling (2) and storage (3) of soil organic carbon. Soil microorganisms play key roles in these biogeochemical processes (4, 5), and, by infecting soil microbiota (6, 7), viruses likely have substantial impacts on carbon dynamics (8). The potential importance of viruses in soils (9–12), together with their measured high abundance— $10^7$  to  $10^{10}$  virus-like particles per gram of soil (9)—and improvements in our ability to sequence and track soil viral genomes (8, 13) have led to a renewed flurry of investigations into soil viral ecology (14–21). However, despite a new appreciation for the vast diversity of soil viruses (14–18, 22), little is known about the factors that govern soil viral community assembly, precluding a robust and predictive understanding of viral impacts on terrestrial ecosystem processes.

Soils are physically, chemically, and biologically heterogeneous (23). The intricate network of aggregates and pore spaces that constitutes the soil matrix (24) not only sustains a varying landscape of edaphic properties but also restricts the movement of microorganisms (25). Such environmental gradients and dispersal limitations often lead to the spatial structuring of microbial diversity (26). For example, distance–decay of community similarity, a biogeographical relationship in which communities become more compositionally dissimilar as the spatial distance between them increases (27), has been reported for soil bacteria across multiple scales and habitats (28–30). Yet, despite evidence of local adaptation of soil bacteriophages to host strains at the centimeter scale (31), the extent to which the soil virosphere is also spatially structured has not been thoroughly explored. An understanding of the compositional turnover of viral communities across space is, therefore, necessary to begin unraveling the spatial constraints of host–virus interactions in soil.

Given its multifaceted role as a resource, solvent, and transport medium, water is a central regulator of the activity, abundance, and dispersal of soil microorganisms (32). In Mediterranean climate grasslands, where wet winters follow dry summers, seasonal rainfalls dictate the compositional dynamics of soil microbiomes (33). Thus, the forecasted alteration of precipitation patterns due to climate change (34) could impact soil trophic networks and their contributions to the biogeochemical processes in these habitats (35). Rainfall manipulation experiments have shown that reduced precipitation can reshape soil bacterial community composition (36, 37). Whether these shifts are coupled to changes in the soil virosphere remains unknown, although recent observations suggest that water

## Significance

Through infection and mortality of microbiota, viruses have myriad impacts on host metabolism, evolution, and Earth's biogeochemical cycles. The sheer abundance of soil viruses hints at their likely importance in terrestrial ecosystems, yet the factors shaping soil viral diversity are poorly understood. Here, we show that grassland viral communities exhibited robust spatial structuring across one field, suggesting strong dispersal limitations at the local scale. Further, a shift in viral community composition accompanied a decrease in soil moisture, whereby phages predicted to infect abundant actinobacteria were enriched. Thus, despite spatial turnover, viruses responded cohesively to changing environmental conditions, suggesting the potential for a predictive understanding of soil virosphere dynamics and impacts on terrestrial processes.

Author affiliations: <sup>a</sup>Department of Plant Pathology, University of California, Davis, CA 95616; <sup>b</sup>Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720; <sup>c</sup>Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550; and <sup>d</sup>Life & Environmental Sciences Department, University of California, Merced, CA 95343

Author contributions: C.S.-M., J.P.-R., M.K.F., and J.B.E. designed research; C.S.-M., K.E.-M., and M.Y. performed research; C.S.-M. analyzed data; and C.S.-M. and J.B.E. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2022 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

<sup>1</sup>To whom correspondence may be addressed. Email: jbemerson@ucdavis.edu.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2209132119/-/DCSupplemental>.

Published November 2, 2022.

availability could be a major driver of soil viral community assembly. For example, a comparison of three distinct grassland sites revealed a significant correlation between soil moisture and viral richness (21), and a laboratory study identified a substantial shift in viral diversity triggered by the wetting of dry biocrust soil (20). As such, characterizing the compositional response of soil viral communities to reduced precipitation can help us understand the potential impact of a changing environment on host–virus interactions.

In this study, we generated viral-size-fraction metagenomes (viromes) to profile the double-stranded DNA (dsDNA) viruses inhabiting a Mediterranean grassland exposed to rainfall-exclusion treatments. The comprehensive access to soil viral diversity enabled by this viromics approach (14, 38) allowed us to characterize the spatial turnover of viral populations and genotypes at a local scale and dissect viral community responses to changes in soil moisture. Complementary analyses of 16S ribosomal RNA (rRNA) gene amplicon sequencing data from the same samples allowed for comparisons of community assembly patterns for viruses and their bacterial and archaeal hosts. Our results revealed a spatially structured soil virosphere that can respond cohesively to reduced precipitation.

## Results and Discussion

To characterize dsDNA viral diversity and investigate viral community compositional patterns in Mediterranean grasslands, we collected surface (0 to 15 cm) soil samples from a field site at the Hopland Research and Extension Center in northern California (Fig. 1A). Soils were harvested from 22 subplots distributed across 15 experimental plots arranged in two separate blocks (*SI Appendix, Fig. 1 A and B*). These plots have been maintained since 2017 with either 100% or 50% of the average historical precipitation via rainfall-excluding shelters and controlled irrigation (39) (*SI Appendix, Fig. 1 C and D*). Samples were collected from densely rooted locations (*SI Appendix, Fig. 1 B*) at two time points (March and April, T1 and T2, respectively) during the 2020 growing season of *Avena barbata* (slender wild oat), the naturalized annual grass that dominates the site (*SI Appendix, Fig. 1 E*).

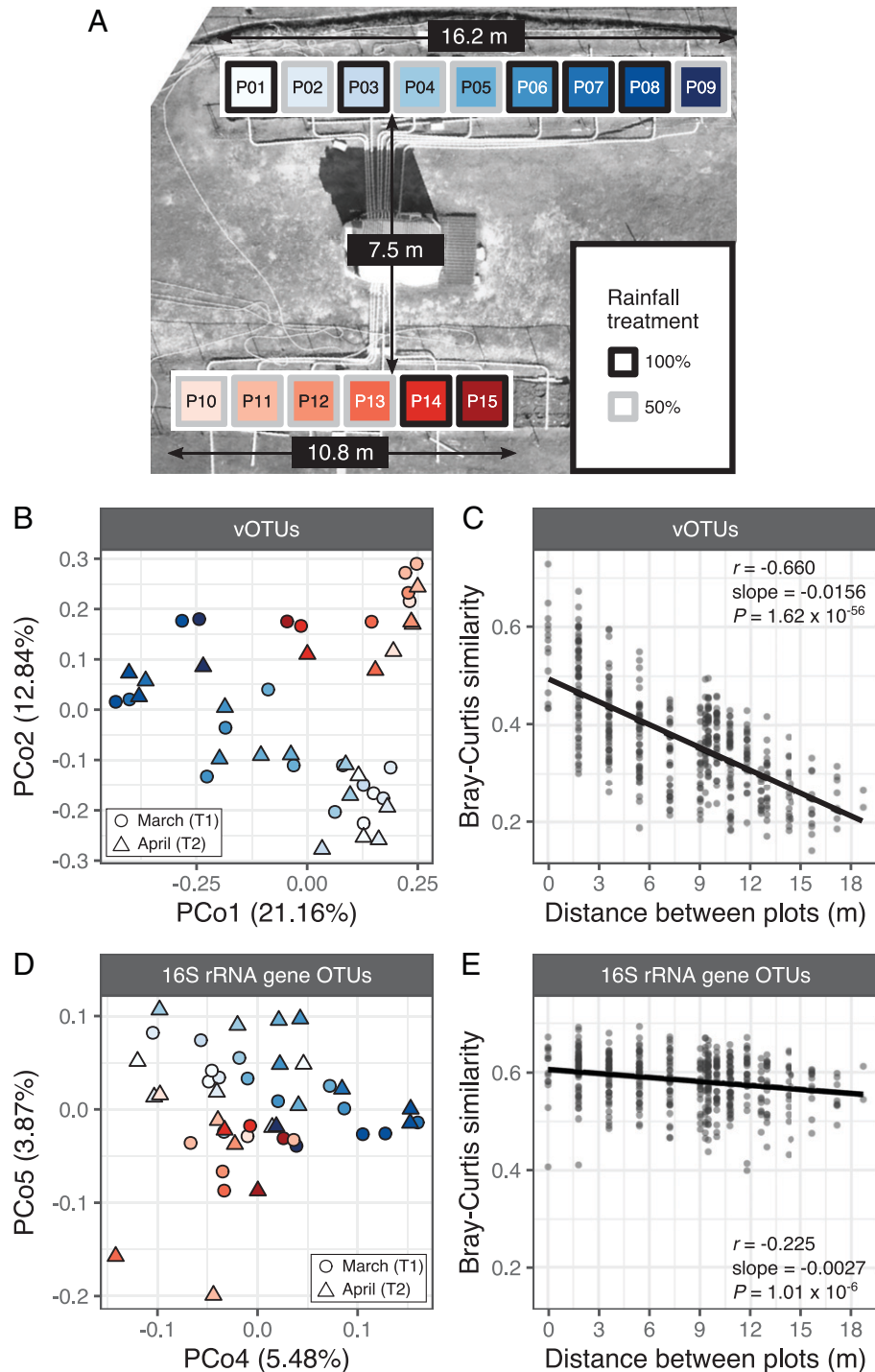
Soil viral community composition was profiled via 44 viral-size-fraction metagenomes. These viromes were generated from frozen soil samples, and their processing did not include the DNase digestion typically performed to remove extracellular DNA contaminants prior to virion DNA extraction (see *Materials and Methods*). Yet, despite potential overrepresentation of microbial sequences, 18,040 out of 30,238 contigs assembled across viromes were identified as viral by VIBRANT (40), a viral enrichment comparable to previous reports from non-DNase-treated viromes generated from fresh soil samples (38). Viral contig clustering at  $\geq 95\%$  average nucleotide identity (ANI) across 85% of the alignment fraction yielded 6,088 approximately species-level viral operational taxonomic units (vOTUs) that served as references for read recruitment to establish vOTU relative abundances (41). After removing vOTUs exclusively detected in single viromes and excluding one virome due to poor vOTU recovery, the final dataset consisted of 43 viromes and 5,315 vOTUs (*Dataset S1*).

**Viral Community Composition Followed a Stronger Distance–Decay Relationship than Bacterial and Archaeal Communities at the Field Scale.** Viral community beta-diversity patterns were largely explained by the spatial arrangement of the field plots, as evidenced by a longitudinal gradient captured by the first axis of a principal coordinates analysis (PCoA) and the separation of upper and lower field blocks along the second PCoA axis (Fig. 1B). In contrast, no meaningful clustering based on time point or

watering treatment was observed along these two axes (*SI Appendix, Fig. 2 A and B*). A permutational multivariate analysis of variance (PERMANOVA) further confirmed the predominant role of spatial structuring on viral community composition relative to other experimental factors (*SI Appendix, Table 1*). Additionally, we identified a significant negative correlation between viral Bray–Curtis similarity and spatial distance between plots (Fig. 1C), indicating that distance–decay relationships were a key driver of viral community composition. These trends were driven, in part, by substantial differences in vOTU detection patterns across viromes: of 5,135 vOTUs, 50% were detected in 9 or fewer of the 43 viromes (*SI Appendix, Fig. 3A*). Moreover, the percentage of vOTUs shared between pairs of viromes declined steeply as spatial separation increased (*SI Appendix, Fig. 3B*). This strong spatial structuring of viral diversity within one field is consistent with prior work in agricultural soils (14) and over larger spatial distances (42), suggesting that distance–decay relationships could be a conserved feature of the soil virosphere. For example, a recent viromic survey of five different natural reserves, including grasslands and other habitats, in northern California found that more than 90% of the dsDNA vOTUs were exclusively found in individual locations (43). Similarly, minimal overlap in soil RNA viral community composition was observed across five sites, including four grasslands, within a 3-km transect (42). Thus, the observed turnover of viral populations at the local scale in this study (Fig. 1B) may translate to stark compositional distinctions at the regional scale. Future studies with a broader spatial and temporal range will be needed to fully assess the prevalence of these biogeographical patterns.

To assess whether bacterial and archaeal communities displayed similar spatial patterns at our field site, we performed 16S rRNA gene amplicon profiling on total DNA extracted from the same soil samples used to generate the viromes. In contrast to the strong spatial patterns in the viral communities, collection time point was the main factor shaping prokaryotic beta-diversity, as indicated by a significant PERMANOVA (*SI Appendix, Table 2*) and a clear distinction between March and April samples along the first axis of a PCoA (*SI Appendix, Fig. 2 C and D*). While spatial structuring was also detected in the bacterial and archaeal communities (*SI Appendix, Table 2*), its effect was only evident along the fourth and fifth PCoA axes (Fig. 2D). Further, even though spatial distance was significantly negatively correlated with microbial community Bray–Curtis similarity (Fig. 2E), this association was not as pronounced as for the viral communities. Specifically, the spatial turnover rate of community similarity (the slope of the distance–decay relationship) was 5.8 times higher for viruses than for bacteria and archaea. Similar observations of spatial succession of viral but not prokaryotic communities along an 18-m gradient in an agricultural field (14) raise the possibility that stronger spatial structuring in viral relative to bacterial and archaeal communities (Fig. 1 B and D) might be a generalizable pattern across soils.

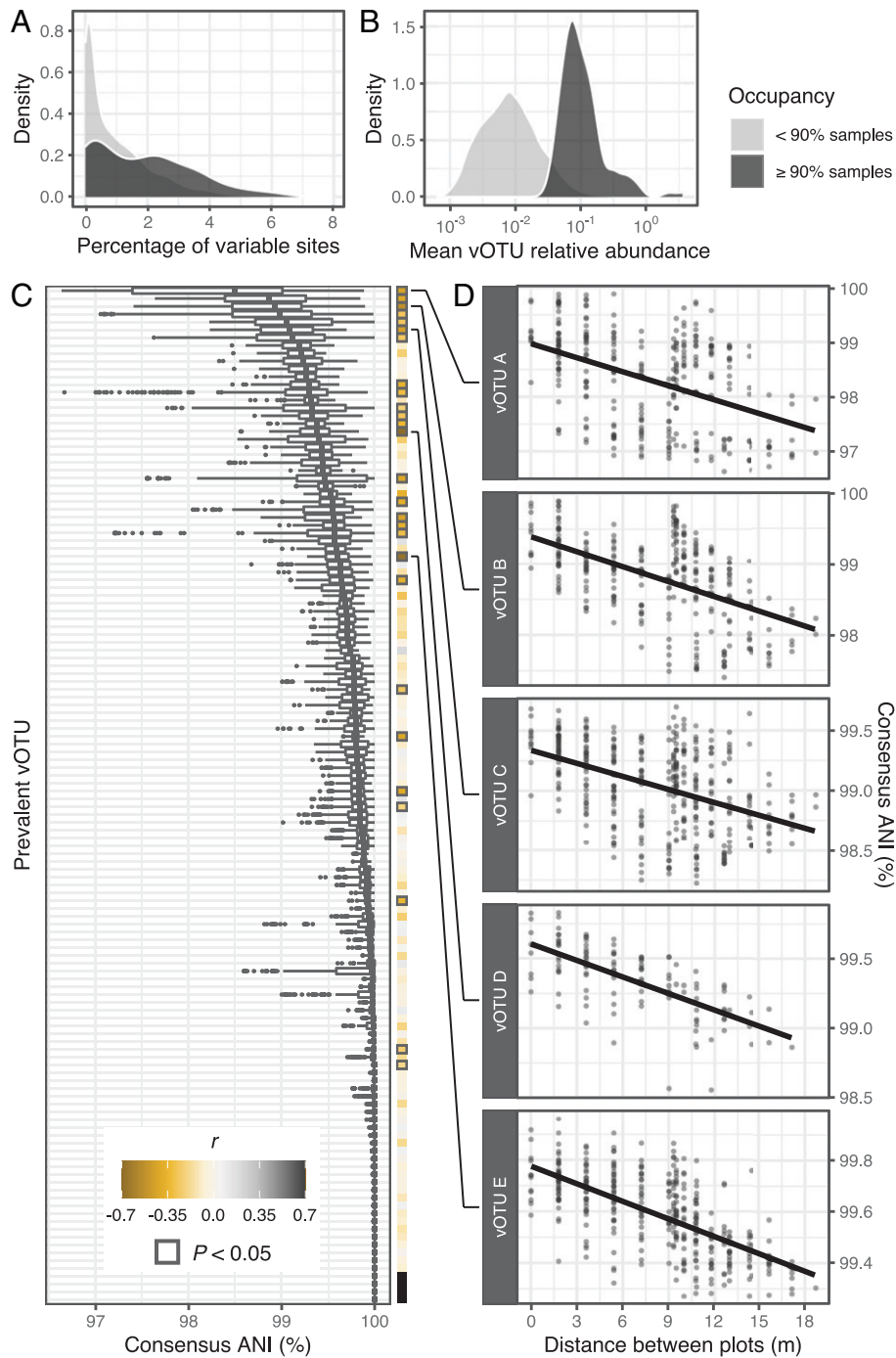
The observed differences in distance–decay patterns between viral and prokaryotic communities suggest that the underlying assembly processes governing spatial structuring at local scales could be differentially impacting these two components of the soil microbiome. For example, viruses and cellular microorganisms likely experience distinct dispersal limitations linked to size, adsorption specificities, and transport mechanisms (44, 45). Some of the observed spatial gradients in the abiotic environment in our study (*SI Appendix, Fig. 4A*), such as spatially structured soil calcium concentrations (*SI Appendix, Fig. 4B*), are consistent with this scenario, as attachment to soil surfaces in the presence of  $\text{Ca}^{2+}$  has been shown to be significantly higher for viral particles than for bacteria (45), potentially impacting their relative



**Fig. 1.** Spatial structuring of viral and prokaryotic communities in a Mediterranean grassland. (A) Aerial view of the field site. Colored squares mark the locations of the sampled plots within the upper (blue) and lower (red) blocks. Square outlines indicate the rainfall manipulation regime assigned to each plot. Differences in font color are for legibility only. (B and D) Unconstrained analysis of principal coordinates performed on (B) vOTU and (D) 16S rRNA gene OTU Bray-Curtis dissimilarities. B displays the first and second axes and D displays the fourth and fifth axes, as they best captured the spatial structuring in (B) viral and (D) bacterial and archaeal communities. Color reflects the plot from which the sample was collected and corresponds to the gradient palette in A. Point shape represents the collection time point. Axis labels indicate the percentage of total variance explained. (C and E) Relationship between Bray-Curtis similarity and spatial distance in (C) viral communities and (E) bacterial and archaeal communities. Each point represents a pair of samples, and the spatial distance between them was measured as the length of the line connecting the centers of the corresponding plots. Pairs of samples involving different time points were excluded from these analyses. Trend lines display the least squares linear regression model. Inset statistics correspond to the Pearson's correlation coefficient ( $r$ ), the linear regression slope, and the associated  $P$  value.

movement in soil. Environmental filtering, whereby abiotic and/or biotic factors influence the distribution of microbial populations through selective pressure (26), may also be particularly relevant for viral community assembly. For instance, edaphic properties can directly affect viral viability and transport (11, 46), while viral reliance on hosts for replication renders the hosts

themselves to be unavoidable environmental filters (47). Interestingly, viral beta-diversity was significantly correlated with both the abiotic environment and prokaryotic community composition (*SI Appendix, Fig. 5 A and B*), suggesting that even the dampened spatial structuring of edaphic properties (*SI Appendix, Fig. 4A*) and prokaryotic beta-diversity (*Fig. 1E*) could have contributed to



**Fig. 2.** Spatial structuring of viral population microdiversity. (A and B) Kernel density plots showing the distributions of (A) microdiversity (measured as the percentage of polymorphic sites in a vOTU sequence) and (B) mean relative abundance within prevalent ( $\geq 90\%$  occupancy) and nonprevalent ( $< 90\%$  occupancy) vOTUs. (C) Distributions of ANIs for each prevalent vOTU, calculated between pairs of sample-specific vOTU consensus sequences. Each box plot corresponds to a single vOTU, and the y axis is in rank order (ascending from top to bottom) of the median ANI value for each vOTU. Boxes display the median and interquartile range (IQR), and data points farther than  $1.5 \times$  IQR from box hinges are plotted as outliers. The heat map on the right shows the Pearson's correlation coefficients between consensus ANI and spatial distance. Bold outlines indicate a significant  $P$  value ( $< 0.05$ ) for the correlation after multiple comparisons correction (Holm algorithm). Filled black squares correspond to vOTUs with no variation across samples (i.e., all ANIs were equal to 1). (D) The top five vOTUs with the most significant correlations (lowest  $P$  values) between consensus ANI and spatial distance. Each point represents a pair of samples, and the spatial distance between them was measured as the length of the line connecting the centers of the corresponding plots. Pairs of samples involving different time points were excluded from these analyses. The trend line displays the least squares linear regression model. Note that vOTUs are defined in part by sharing  $\geq 95\%$  ANI (see *Materials and Methods*), so within-vOTU ANI values will necessarily be  $\geq 95\%$  ANI. Also note subtle differences in the y axis range across graphs.

the amplified spatial structuring of viruses in this study (Fig. 1C). Consistent with biotic environmental selection, RNA viral communities in grasslands differed significantly in the presence of plant litter and across soil compartments (16), and dsDNA viral communities differed along a permafrost thaw gradient (18), with

patterns similar to those of their host communities in both cases (16, 48). Differences in the strength of the spatial patterns between viruses and prokaryotes could also be related to differences in the integrated temporal scales captured by DNA pools in viromes compared to total DNA (49). For example, fast viral



particle decay rates and the large burst sizes (11, 50, 51) characteristic of viral replication could amplify the signal of recent viral infections in the viromes, while relic DNA (52) and DNA from dormant biota (53) could mask the signal from active microbes in total DNA. Together, our results suggest that dispersal, abiotic, and biotic factors could all contribute to soil viral community assembly patterns and, potentially, to differences between soil host and viral biogeography.

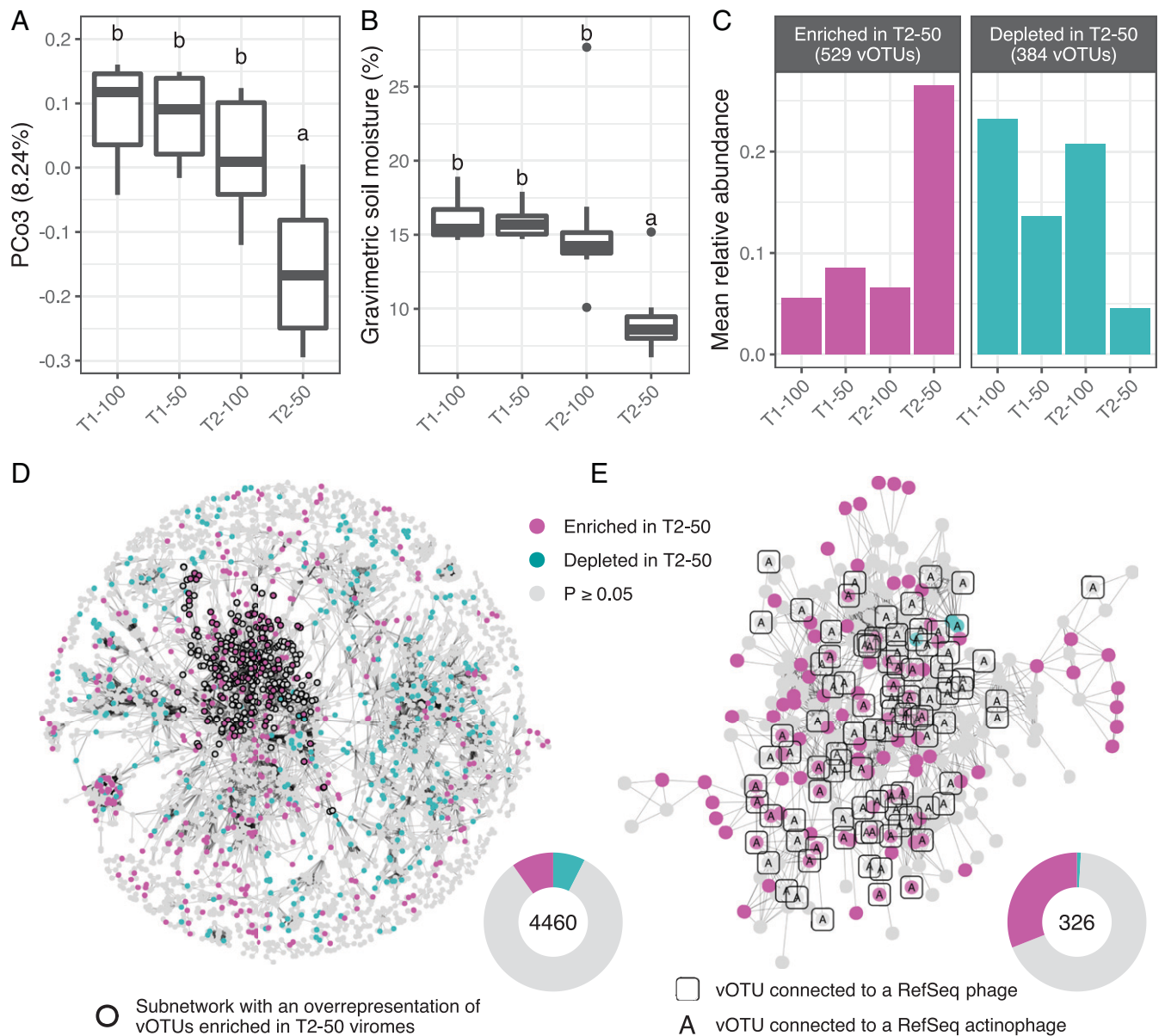
**Genomic Microdiversity of Viral Populations Tended to Be Spatially Structured.** In addition to environmental filtering and dispersal, diversification (i.e., the generation of novel genetic variation) can contribute to diversity patterns in microbial communities (26, 54, 55). To explore the role of spatial structuring on viral genotypic heterogeneity across our field site, we profiled within-population genomic variation. Briefly, using inStrain (56), we scanned all mapped reads assigned to individual vOTUs and identified polymorphic sites. Then, to assess intersample vOTU genomic similarities, we reconstructed sample-specific consensus vOTU sequences and performed pairwise ANI comparisons. Given that most vOTUs were detected in a limited number of viromes (*SI Appendix, Fig. 3A*), we restricted this analysis to a subset of 130 vOTUs that were detected in at least 90% of the viromes. This set of prevalent vOTUs had high levels of intrapopulation heterogeneity (Fig. 2*A*) and also consisted of some of the most abundant viral community members (Fig. 2*B*). The ANI distributions revealed a wide range of genomic variation among dominant allelic variants: While some prevalent vOTUs had mean pairwise variant similarities close to 0.95 ANI (the threshold used to define a viral population), others appeared nearly clonal across samples (Fig. 2*C*). Moreover, the microdiversity of many vOTUs was spatially structured within our field site. For 21% of the prevalent vOTUs and 54% of the 26 vOTUs that displayed the most variation (prevalent vOTUs with median ANIs <99.5%), genomic similarity displayed a significant negative correlation with spatial distance, indicating that the predominant allelic variants tended to diverge with increasing distance (Fig. 2*D* and *SI Appendix, Table 3*). Together, these results show that viral community composition and the genetic makeup of viral populations exhibited significant distance–decay relationships across our field site.

**Low-Moisture Soils Had Significantly Different Viral Communities and Were Enriched in Putative Actinophages.** Although spatial structuring emerged as the predominant driver of soil viral diversity patterns (Fig. 1*B* and *C*), viral community composition was also shaped by the experimental precipitation treatments. In particular, the third axis of our PCoA, which accounted for 8.24% of variance in the dataset, captured a significant distinction between the April viromes from 50% precipitation plots (“T2-50” samples) and the rest of the viromes (Fig. 3*A*). Gravimetric soil moisture contents were also significantly lower for these T2-50 samples (Fig. 3*B*), a distinction that likely reflects the differential precipitation exclusion patterns preceding each collection time point. In particular, both 50% and 100% plots were fully exposed to rainfall during the month leading up to the first sample collection; in contrast, 50% treatment plots underwent a 24-d-long precipitation exclusion immediately before the second sample collection (*SI Appendix, Fig. 1 C and D*). These differences in recent precipitation likely explain why T1-50 samples had similar soil moisture to the 100% precipitation treatment samples from both time points, whereas T2-50 samples had lower soil moisture and correspondingly distinct viral communities. Together, these

trends suggest that viral communities were directly or indirectly structured by changes in soil moisture during the growing season.

We next identified vOTUs with significant responses to these soil moisture patterns to assess potential commonalities among them. An indicator species analysis revealed 529 vOTUs that were significantly enriched in T2-50 viromes relative to the rest of the viromes and 384 vOTUs that were significantly depleted (Fig. 3*C*). As functional traits can be phylogenetically conserved in bacteria (57), we assessed whether analyses at higher levels of viral genome conservation might reveal a similarly cohesive response to soil moisture. Using vConTACT2 (58), we constructed a network of vOTUs (nodes), in which each edge indicated a significant overlap in predicted protein contents between a pair of vOTUs. We then adapted an algorithm designed to assess local overrepresentation of traits in biological networks (59) to characterize the network distribution of soil-moisture-responding vOTUs. Briefly, for each vOTU, we identified a local neighborhood of all vOTUs that could be reached, either directly or indirectly, via an edge path with a length shorter than the first percentile of all pairwise node distances in the network. After discarding all local neighborhoods with fewer than 10 vOTUs, we recovered 2,865 subnetworks of highly interconnected nodes with a median size of 39 vOTUs, allowing us to consider many more sizeable groups of related vOTUs than a standard vConTACT2 analysis of “genus-level” viral clusters (VCs) (15, 60), as there were only 24 VCs with at least 10 vOTUs in this dataset. Next, we performed hypergeometric tests to assess the overrepresentation of indicator vOTUs enriched or depleted in T2-50 samples within each network neighborhood. A total of 108 neighborhoods showed a significant overabundance of vOTUs consistently enriched in T2-50 samples, with 26 to 67% of vOTUs in these neighborhoods displaying this trait, compared to only 10% of vOTUs in the whole network (*SI Appendix, Fig. 6 A–C*). This pattern contrasted with the lack of substantial network aggregation of vOTUs depleted in T2-50 samples, as only four small, local neighborhoods displayed a significant, albeit weak, overrepresentation of this trait (*SI Appendix, Fig. 6 A and B*). Interestingly, all of the significantly T2-50 enriched trait neighborhoods were constrained to a single region in the protein-sharing network, indicating that a relatively cohesive group of related vOTUs tended to be enriched in T2-50 samples (Fig. 3*D*). Further, the indicator vOTUs within this subnetwork covered a range of detection patterns across viromes (occupancies) and were spatially distributed across the field site (*SI Appendix, Fig. 7 A and B*). This suggests that, despite the strong spatial structuring of viral communities overall, this group of genomically related vOTUs responded cohesively to changes in soil moisture, regardless of their field plot locations.

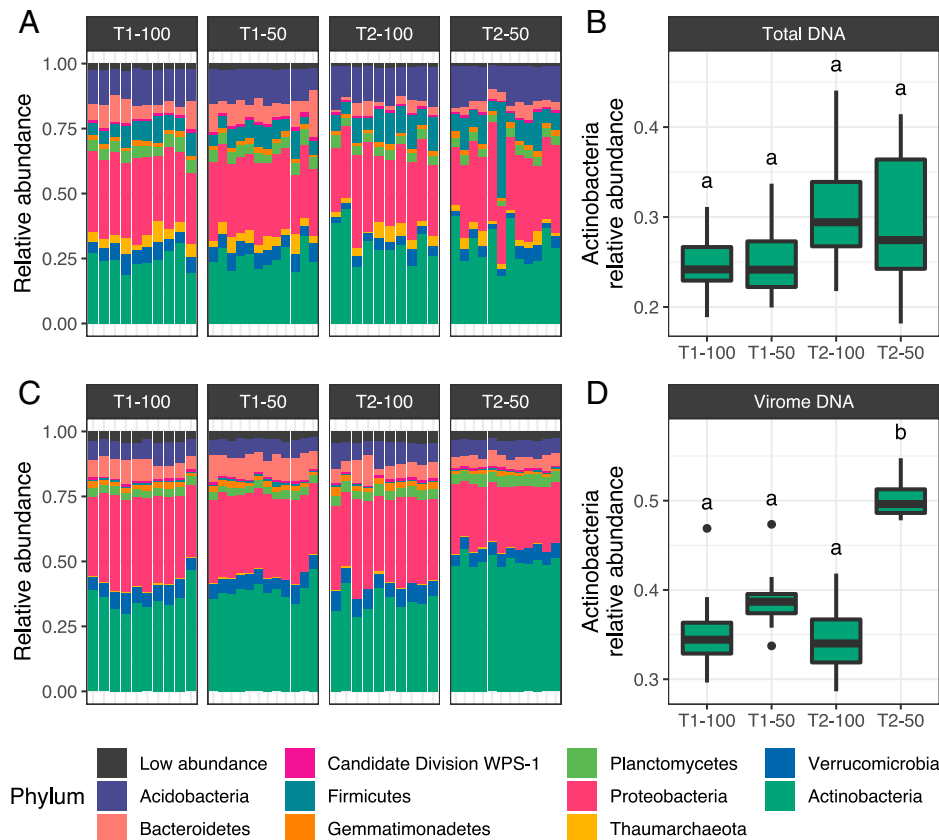
To further explore the subnetwork with a significant overrepresentation of low-moisture (T2-50)-enriched vOTUs, we performed a second protein-sharing network analysis, with all prokaryotic viral genomes in the NCBI RefSeq database. We identified edge connections between vOTUs in the low-moisture trait subnetwork and RefSeq viral genomes to assess network neighborhood trends in viral and host taxonomy (*SI Appendix, Fig. 8 A–C*). Of 326 vOTUs in the subnetwork, 96 were connected to at least one RefSeq viral genome, all of which were classified as Siphoviridae or as undefined viruses from the order Caudovirales (*SI Appendix, Fig. 8B*), both taxonomic classifications currently under consideration to be replaced by monophyletic genome-based families (61). More interestingly, all 191 RefSeq viral genomes connected to a trait subnetwork vOTU were isolated from Actinobacteria hosts, suggesting that the low-moisture-responsive vOTU subnetwork



**Fig. 3.** Viral community trends associated with soil moisture content. In A–C, samples are grouped along the x axis by collection time point (T1 and T2) and precipitation regime (100% and 50%). (A) Distribution of scores along the third axis of a PCoA performed on vOTU Bray–Curtis dissimilarities. The y axis label indicates the percentage of total variance explained. The first two axes of the same analysis are shown in Fig. 1B. (B) Gravimetric soil moisture contents. Boxes display the median and interquartile range (IQR), and data points farther than 1.5× IQR from box hinges are plotted as outliers. In A and B, different letters indicate significantly different sample groupings ( $P < 0.05$ ), as determined by two-tailed Tukey’s range tests. (C) Summed mean relative abundances of the sets of vOTUs detected as indicator species differentiating T2-50 communities from the rest of the viromes. Facets distinguish indicator vOTUs that were relatively enriched or depleted, respectively, in T2-50 viromes. (D) Gene-sharing network displaying significant overlaps in predicted protein content (edges) between vOTUs (nodes). Node color shows whether a vOTU was an indicator species enriched or depleted in T2-50 samples or not an indicator species (defined by  $P$  values below or above 0.05, respectively, from an indicator value permutation test). Bold outlines highlight a subnetwork of all local neighborhoods with a significant overrepresentation of vOTUs enriched in T2-50 viromes (SI Appendix, Fig. S6 B and C). (E) Zoomed-in version of the subnetwork highlighted in D. Nodes surrounded by squares correspond to vOTUs with a significant overlap in their predicted protein contents with any of 971 RefSeq phage genomes, according to the network analysis shown in SI Appendix, Fig. S8. All such RefSeq phage genomes with significant links to this subnetwork were from phages isolated on Actinobacteria hosts, indicated by tagging vOTU nodes linked to RefSeq actinophages with the letter “A.” In D and E, inset donut plots on the lower right show the total number of vOTUs in the displayed network (center), along with the proportions of the indicator and nonindicator vOTUs in that network (fractions of the circle). Network visualization layouts were generated with the Fruchterman–Reingold algorithm.

was largely composed of actinobacteriophages (Fig. 3E). In contrast, only 38% of all 971 vOTUs associated with RefSeq genomes across the entire network were exclusively linked to an actinobacteriophage (SI Appendix, Fig. 8C), indicating a substantial concentration of putative actinobacteriophages in the subnetwork. These results suggest that low soil moisture could have increased the activity of actinobacteria, in turn driving increased predation by actinophages.

**The Relative Enrichment of Putative Actinophages in Low-Moisture Soils Coincided with an Increase in Relic DNA from Actinobacteria.** Many actinobacteria are drought-resistant members of soil microbiomes that can increase their activity and abundance under low-moisture conditions across multiple environments (36, 62–64), including Mediterranean grasslands (33). While actinobacteria were among the most abundant members in the 16S rRNA gene amplicon profiles, there were



**Fig. 4.** Abundance patterns of actinobacteria in total and relic DNA profiles. (A and C) Phylum abundances in 16S rRNA gene profiles from (A) total DNA 16S rRNA gene amplicon libraries and (C) virome DNA libraries. Each stacked bar plot corresponds to a sample, and the 10 most abundant phyla are colored. All other phyla are grouped in the “Low abundance” category. (B and D) Relative abundances of actinobacteria in (B) total DNA 16S rRNA gene amplicon libraries and (D) virome DNA libraries. Samples are organized by collection time point (T1 and T2) and precipitation treatment regime (100% and 50%). Boxes display the median and interquartile range (IQR), and data points farther than 1.5× IQR from box hinges are plotted as outliers. Letters above boxes indicate significantly different groupings ( $P < 0.05$ ), as determined by pairwise Wilcoxon’s rank-sum tests. For C and D, abundances were normalized to the number of reads classified as 16S rRNA genes in each virome profile; a complementary analysis with abundances normalized to the total number of reads in each virome profile is provided in *SI Appendix, Fig. S10*.

no significant differences in their relative abundances across watering treatments or time points (Fig. 4 A and B). Further, even though the first axis of a PCoA captured a microbial community compositional shift from March to April (*SI Appendix, Fig. 2D*), there was no clear distinction between T2-50 prokaryotic communities and the rest of the samples (*SI Appendix, Fig. 9*). Given that microbial community sensitivity to an environmental disturbance is linked to the intensity and duration of the stressor (65), the absence of a significant effect of watering treatment on prokaryotic community composition could stem from the temporal scale of the dry-down captured in this study. In particular, the 24-d-long rainfall exclusion that 50% treatment plots experienced between the two collection time points (*SI Appendix, Fig. 1C*) might not have been long enough to detect the compositional shifts previously observed during more prolonged desiccation periods (33). Additionally, given that rainfall-exclusion treatments in our site started in 2017, treatment legacy effects could have also dampened the microbiome response under low-moisture conditions, as preexposure to drought can increase the resistance of soil bacteria to future desiccation events (36).

While the effects of low moisture observed in the virosphere were not recapitulated by bacterial and archaeal communities overall or by actinobacterial relative abundances specifically, it is possible that the presence of genetic material from dead and dormant cells in the total DNA profiles could have concealed underlying ecological dynamics driven by physiologically active

microorganisms (52, 53). Because of its high abundance in soils (66), extracellular DNA from dead cells can introduce substantial biases in estimates of microbial abundance, especially when the turnover rate of this relic DNA is disrupted by environmental perturbations, such as bacteriophage blooms (49). To consider the relic DNA pool in our samples more directly, we recovered reads classified as 16S rRNA gene fragments from virome profiles. Given that viral enrichment in viromes was achieved via 0.22- $\mu\text{m}$  filtration prior to DNA extraction, any bacterial and archaeal sequences present in these libraries likely originated from relic DNA or small (<0.22  $\mu\text{m}$ ) microbial cells (38, 67). Interestingly, the relative abundance of Actinobacteria 16S rRNA gene reads recovered from T2-50 viromes was significantly higher than in any other group of samples (Fig. 4 C and D and *SI Appendix, Fig. 10*). This increase in (presumably) free actinobacteria DNA, coupled with the enrichment of putative actinophages in the T2-50 subnetworks (Fig. 3E), suggests that actinobacteria hosts may have experienced higher infection and lysis under lower-moisture conditions, a scenario compatible with the capacity of environmental relic DNA to retain signatures of recent viral infections (68). Interestingly, actinobacteria can display enhanced transcriptional activity under drought conditions (33, 62), suggesting that the potential increase in actinobacteria–actinophage interactions observed in this study could stem from a heightened metabolic state in actinobacteria under low-moisture conditions. Future studies dissecting the functional changes that soil microbiomes



undergo during reduced precipitation could help us understand the mechanisms enabling these dynamics.

## Conclusions

Here, we show that grasslands harbor an active and highly dynamic soil virosphere that is structured over space and can respond to a changing environment. The high degree of spatial turnover that we observed—within one field site during one growing season—suggests dispersal limitations for viral populations and genotypes on scales of meters and months, hinting at the potential spatial patterning of host–virus interactions in soil. Moreover, the disparity in distance–decay relationships between viral and prokaryotic communities could reflect potential differences in the assembly processes shaping these two components of the soil microbiome. The compositional shift triggered by reduced precipitation further indicates that, despite the underlying spatial structuring, groups of genomically related viruses can respond cohesively to environmental conditions, such as decreases in soil moisture, presumably by way of their hosts. Finally, the coupled enrichment of putative actinophages and relic DNA from actinobacteria under low-moisture conditions raises the possibility that reduced precipitation increased the infection and lysis of a prevalent, drought-responsive group of soil microorganisms. In summary, soil viral community assembly seems to be tightly coupled to the heterogeneous and dynamic biotic and abiotic landscape of the local environment, and it will be interesting to see how these patterns scale over more extensive temporal and spatial distances.

## Materials and Methods

**Field Experiment and Sample Collection.** Samples were collected as part of a rainfall manipulation field experiment (39) at the University of California Hopland Research and Extension Center (39° 00' 14.6" N, 123° 05' 09.1" W). The field site contained 16 plots (1.8 × 1.8 m) arranged in two separate blocks 7.5 m apart: a 16.2-m-long upper block with nine plots and a 12.6-m-long lower block with seven plots (SI Appendix, Fig. 1A). Plot boundaries were delimited by 1-m-deep vertical plastic liners, installed in the spring of 2017, that limited water transfer between adjacent soils. Each plot contained eight circular subplots (40-cm diameter) delineated by 15-cm-deep polyvinyl chloride collars (SI Appendix, Fig. 1B). Starting in 2017 and continuing until 2020, plots were exposed to two precipitation regimes, where the amount of water received by each plot was adjusted to match 100% or 50% of the average historical precipitation at the site. Differential watering was achieved by the periodic deployment of rainfall-excluding shelters (SI Appendix, Fig. 1C) and by controlled irrigation of individual plots (SI Appendix, Fig. 1D). For this study, soils were harvested from 22 subplots distributed across 15 of the 16 plots (SI Appendix, Fig. 1A). All of these subplots were segmented in two halves by a 15-cm-deep Plexiglas divider (SI Appendix, Fig. 1B), and they were located within a 60-cm radius from the center of each plot. Sample collections were performed on 13 March and 14 April 2020 (T1 and T2, respectively) during the active growth phase of *A. barbata* (SI Appendix, Fig. 1E). At each time point, half of each subplot was destructively harvested. Samples were processed by removing any visible roots, homogenizing the soil, and storing the soil at –80 °C until further processing. For soil moisture measurements, separate fresh soil subsamples were collected and processed immediately.

**Virome DNA Extraction, Library Construction, and Shotgun Sequencing.** Due to the COVID-19 2020 lockdown, we could not perform virome extractions on fresh samples as intended and instead stored soils at –80 °C until processing. Soil virions were enriched through filtration and concentration prior to DNA extraction, following a modified version of a previously published protocol (69). For each sample, 10 g of soil were resuspended in 10 mL of protein-supplemented phosphate-buffered saline solution (PPBS: 2% bovine serum albumin, 10% phosphate-buffered saline, 1% potassium citrate, and 150 mM

MgSO<sub>4</sub>). To elute virions, soil suspensions were vortexed until homogenized, placed on an orbital shaker (10 min, 400 rpm, 4 °C), and centrifuged (10 min, 3,095 × *g*, 4 °C). Supernatants were recovered and stored briefly at 4 °C, while pellets were resuspended in 10 mL of fresh PPBS for back-extraction of the remaining soil. This process was repeated for a total of three rounds of extraction of the same soil. Supernatants from the same sample were then pooled and centrifuged three times (10 min, 10,000 × *g*, 4 °C), retaining the supernatant and discarding the pellet each time to remove residual soil particles. Purified supernatants were then filtered through a 0.22-μm polyethersulfone membrane to remove cells. Eluted virions in the filtrate were concentrated via ultracentrifugation (2 h 25 min, 32,000 × *g*, 4 °C) in an Optima LE-80K ultracentrifuge with a 50.2 Ti rotor (Beckman-Coulter). Supernatants were removed, and pellets were resuspended in 100 μL of ultrapure water. As previously shown (38, 70), the DNase treatment step that serves to remove free DNA at this stage is not compatible with samples stored frozen (we suspect that this is because freezing compromises virions), so we were unable to perform a DNase treatment. We have previously shown that non-DNase-treated soil viromes still successfully enrich the viral signal relative to total metagenomes and capture the same ecological trends as DNase-treated viromes from the same samples (38).

DNA was extracted from the viral fraction with the DNeasy PowerSoil Pro kit (Qiagen), following the manufacturer's protocol, with the addition of a 10-min incubation at 65 °C prior to the bead-beating step. Shotgun metagenomic libraries were constructed with the DNA Hyper Prep kit (Kapa Biosystems-Roche), and paired-end sequencing (150 bp) was performed on the NovaSeq S4 platform (Illumina).

### Total DNA Extraction, Amplicon Library Construction, and Sequencing.

Total DNA was extracted from 0.25 g of soil with the DNeasy PowerSoil Pro kit (Qiagen), following the manufacturer's instructions, with the addition of a 10-min incubation at 65 °C prior to the bead-beating step. Construction of amplicon libraries followed a previously described dual-indexing strategy (71, 72). Briefly, universal primers 515F and 806R were used to target the V4 region of the 16S rRNA gene. Amplifications were performed with the Platinum Hot Start PCR Master Mix (Thermo Fisher) following the Earth Microbiome Project's PCR protocol (73): an initial denaturation step at 94 °C for 3 min, 35 cycles of 94 °C for 45 s, 50 °C for 60 s, and 72 °C for 90 s, and a final extension step at 72 °C for 10 min. To account for any potential amplification of reagent contaminants (74), we used a DNA-free control (molecular-grade water processed through the same DNA extraction protocol) as a template to generate a blank library. Libraries were cleaned with AmpureXP magnetic beads (Beckman Coulter), quantified (Qubit 4 fluorometer; Thermo Fisher), and pooled in equimolar concentrations. Paired-end sequencing (250 bp) was performed on the MiSeq platform (Illumina).

**Soil Chemistry and Moisture Measurements.** Soil moisture was calculated as the ratio of mass of water per mass of dry soil. While soil moisture was originally measured for all samples, data for a subset of 11 March samples (five from 100% plots and six from 50% plots) were lost and could not be included in downstream analyses. Soil chemistry profiling was performed by Ward Laboratories: Soil pH and soluble salts were measured using a 1:1 soil:water suspension; soil organic matter was measured as the percentage weight loss on ignition; nitrate was measured via a KCl extraction; potassium, calcium, magnesium, and sodium were measured via an ammonium acetate extraction; zinc, iron, manganese, and copper were measured via a DTPA extraction; phosphorus was measured via the Olsen method; and sulfate was measured via a Mehlich-3 extraction. Soil chemistry measurements were only performed on the set of 22 soil samples collected in March (T1).

### Bioinformatic Processing.

**Virome processing.** We used Trimmomatic v0.33 (75) to remove Illumina adapter sequences and quality-trim reads (minimum q-score of 30 evaluated on 4-base sliding windows; minimum read length of 50) and BBDuk v38.82 (76) to remove PhiX sequences. Next, we generated de novo assemblies of individual libraries with MEGAHIT v1.2.9 (77) in metalarge mode (–k-min 27 –k-max 127 –k-step 10), using a contig minimum size threshold of 10,000 bp. Assembled contigs were then classified as viral with VIBRANT v1.2.1 (40) in virome mode. Consistent with established best practices (41), the resulting viral contigs were

dereplicated into nonredundant vOTUs with dRep v3.2.2 (78), using the following parameters: a threshold of  $\geq 95\%$  ANI across  $\geq 85\%$  alignment fraction ( $-sa = 0.95$ ,  $-nc = 0.85$ ), single-linkage algorithm for hierarchical clustering ( $-clusterAlg = single$ ), and filtered nucmer alignments for secondary clustering comparisons ( $-S\_algorithm = ANImf$ ). Representative sequences were selected based exclusively on length ( $-N50W = 0$ ,  $sizeW = 1$ ). Competitive read recruitment against the dereplicated database of vOTUs was performed with Bowtie 2 v2.4.2 (79) in sensitive mode, and the resulting alignments were sorted and indexed with SAMtools v1.11 (80). We used CoverM v0.5.0 (<https://github.com/wood/CoverM>) to generate two vOTU coverage tables: one displaying the trimmed mean coverage ( $-m = trimmed\_mean$ ) and the other displaying the absolute number of mapped reads ( $-m = count$ ). In both cases, all vOTUs with  $< 75\%$  horizontal coverage were discarded ( $-min-covered-fraction = 0.75$ ). We filtered out 773 vOTUs that were exclusively detected in single viromes and removed one virome due to poor vOTU recovery (136 vOTUs compared to a median of 1,562 vOTUs). The final dataset consisted of 43 viromes and 5,315 vOTUs.

**Microdiversity profiling.** Intrapopulation genetic diversity was characterized with inStrain v1.4.0 (56). First, the Bowtie 2 alignments described above were parsed with the profile module to identify divergent sites within the set of mapped reads assigned to each vOTU. Variants were only called if a site had a minimum coverage of five reads. We then used the compare module to calculate average nucleotide identities between sample-specific consensus sequences, which were reconstructed based on the most common allele detected at each variant site. Pairwise comparisons were considered for downstream analyses only if more than 25% of the vOTU sequence length was covered by the profile module in both samples ( $percent\_genome\_compared > 0.25$ ).

**Gene-sharing network construction.** We used Prodigal v2.6.3 (81) in metagenome mode to predict protein content for each dereplicated vOTU and used the resulting amino acid file to construct a gene-sharing network with vConTACT2 v0.9.19 (58). The protein alignment step was performed with Diamond (82), and the protein cluster step was calculated with the MCL algorithm (83). The NCBI RefSeq database of bacterial and archaeal viral genomes (v85) was included as a reference. Layouts used to visualize the resulting network were calculated with the Fruchterman-Reingold algorithm implemented in the GGally package (84).

**Detection and classification of 16S rRNA gene fragments in virome libraries.** As previously described (14), we used SortMeRNA v4.2.0 (85) against representative versions of the bacterial and archaeal SILVA database v132 (86) to recover reads containing 16S rRNA gene sequences from the set of quality-filtered virome reads. We assigned taxonomy with the RDP classifier (87) using the RDP database v18 (88) as reference. A count table was generated from the resulting hierarchical file with the hier2phyloseq() function from the RDPutils package (89).

**Processing of 16S rRNA gene amplicon libraries.** Assembly of paired-end reads into single sequences was performed with PANDAseq v2.9 (90), followed by chimeric sequence removal with usearch v6.1 (91). OTU clustering was performed at a 97% sequence identity threshold with the QIIME (92) implementation of UCLUST v1.2.22 (91) following the open reference protocol against the SILVA database v132 (86). For consistency with 16S rRNA gene analysis performed on viromes, representative sequences were reannotated with the RDP classifier (87) using the RDP database v18 (88) as reference. After discarding singletons and OTUs also detected in the blank library, the final dataset consisted of 53,854 OTUs.

**Data Analysis.** All statistical analyses were conducted using R v3.6.3 (93). Unless otherwise noted, all viral analyses were performed on the trimmed mean coverage vOTU table. For vOTU and 16S OTU profiles, Bray-Curtis dissimilarities were calculated on log-transformed relative abundances with the vegdist() function from vegan v2.5-7 (94). PERMANOVAs were performed with the adonis() function from vegan v2.5-7 (94). To calculate the environmental distance, we first computed the z-score for each soil chemistry variable and then used the dist() function to determine the Euclidean distances between pairs of samples. Principal coordinates analyses were performed with the pcoa() function from ape v5.4-1 (95). Pearson's correlation tests evaluating the association of spatial distance with Bray-Curtis similarity, community overlap, environmental distance, edaphic variables, and vOTU microdiversity were performed using the cor.test() function with the alternative parameter set to "two.tailed." The associated linear

regression slope was calculated with the lm() function. In all cases, spatial distance between pairs of samples was measured as the length of the line connecting the centers of the corresponding plots. To remove any effect of time point on our spatial correlation analyses, we excluded all pairwise comparisons between samples collected at different time points. For correlation analyses involving multiple comparisons (edaphic variables and microdiversity), *P* values were corrected with the Holm algorithm. Indicator species analysis was performed with the multipatt() function from indicpecies v1.7.9 (96). For this analysis, we divided the dataset into two groups, one with the T2-50 viromes and the other with the rest of the samples, and we identified vOTUs significantly associated with each group. We used the lm() function to fit linear models evaluating the effect of collection time point and watering treatment on beta-diversity (as captured by individual principal coordinates) and gravimetric soil moisture. We then used the glht() function from the multcomp package (97) to perform Tukey's range tests. We used the pairwise.wilcox.test() function to perform pairwise Wilcoxon rank-sum tests to assess the effect of collection time point and watering treatment on the relative abundances of actinobacteria 16S rRNA gene profiles from total DNA and virome DNA. To determine the relative enrichment of vOTUs along the horizontal field transect, we performed a differential abundance analysis with DESeq2 (98), using vOTU nonnormalized count tables as input. In particular, we used the DESeq() function to implement negative binomial generalized models to test the effect of the position of each plot on the abundance of individual vOTUs and used the effect size to rank each viral population. All plots were generated with ggplot2 (99).

**Local neighborhood enrichment.** To assess whether vOTUs detected as indicator species of T2-50 samples tended to share similar genomic attributes, we adapted a previously described algorithm designed to systematically assess the distribution of traits in biological networks (59). This algorithm consists of two main steps: 1) For each node in the network, determine a local neighborhood comprised of all nodes that can be directly or indirectly reached via an edge path with a length shorter than a defined threshold, and 2) for each local neighborhood, assess the overrepresentation of a particular attribute among its members. In this study, we used the gene-sharing network generated by vConTACT2 (58), in which nodes represent vOTUs, edges indicate a significant overlap in the predicted content between vOTUs, and edge scores denote the statistical significance of the associated overlap (expressed as  $-\log_{10} P$  value). To determine the distance threshold for local neighborhoods, we first calculated the length of the weighted shortest path for each possible pair of nodes in the network and then identified the first percentile. We performed this step with the distances() function from the igraph package (100), using the reciprocal of the edge scores assigned by vConTACT2 as edge weights. We explored the distribution of the following node attributes across the network: 1) enrichment or 2) depletion in T2-50 samples. To assess the overrepresentation of each of these traits in each of the local neighborhoods, we performed hypergeometric tests using the phyper() function with the "lower.tail" parameter set to false. Local neighborhoods with less than 10 nodes were not considered for the overrepresentation analyses. Multiple comparisons correction was performed with the Holm algorithm.

**Data, Materials, and Software Availability.** All raw sequences have been deposited in the NCBI Sequence Read Archive under the BioProject accession PRJNA818793 (101). The database of dereplicated vOTUs is available at <https://zenodo.org/record/7076890#Y1BMPEzMJPY> (102). All scripts and intermediate files are available at <https://github.com/cmsantosm/HoplandViromes> (103).

**ACKNOWLEDGMENTS.** We thank Donald Herman, Aaron Chew, Anne Kakouridis, Steven Blazewicz, Ella Sieradzki, Rachel Hestrin, Alison Smith, Troy McWilliams, Thomas Seward, and John Bailey for setting up and maintaining the rainfall manipulation experiment. We thank Kateryna Zhalnina, Jack Hagen, David Sanchez, Cynthia Mancilla, and Tasnim Ahmed for their help with field harvests. This work was supported by an award from the US Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research, Genomic Science Program (grant DE-SC0020163 to M.K.F., J.P.-R., and J.B.E.). Research at Lawrence Livermore National Laboratory was conducted under the auspices of DOE contract DE-AC52-07NA27344. Shotgun metagenomic library construction and high-throughput sequencing were performed by the DNA Technologies and Expression Analysis Core at the University of California, Davis Genome Center, supported by NIH Shared Instrumentation grant 1S100D010786-01.

1. Food and Agriculture Organization of the United Nations, *Grasslands of the World* (FAO, 2005).
2. H. A. Henry, D. J. Gibson, J. A. Newman, Biogeochemical cycling in grasslands under climate change. *Grassl. Clim. Change* **1**, 115 (2019).
3. C. Terrer *et al.*, A trade-off between plant and soil carbon storage under elevated CO<sub>2</sub>. *Nature* **591**, 599–603 (2021).
4. R. D. Bardgett, C. Freeman, N. J. Ostle, Microbial contributions to climate change through carbon cycle feedbacks. *ISME J.* **2**, 805–814 (2008).
5. C. Gougoulias, J. M. Clark, L. J. Shaw, The role of soil microbes in the global carbon cycle: Tracking the below-ground microbial processing of plant-derived carbon for manipulating carbon dynamics in agricultural systems. *J. Sci. Food Agric.* **94**, 2362–2371 (2014).
6. E. P. Starr *et al.*, Stable-isotope-informed, genome-resolved metagenomics uncovers potential cross-kingdom interactions in rhizosphere soil. *MSphere* **6**, e0008521 (2021).
7. S. Lee *et al.*, Methane-derived carbon flows into host-virus networks at different trophic levels in soil. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2105124118 (2021).
8. J. B. Emerson, Soil viruses: A new hope. *ISME J.* **4**, e00120-19 (2019).
9. K. E. Williamson, J. J. Fuhrmann, K. E. Wommack, M. Radosevich, Viruses in soil ecosystems: An unknown quantity within an unexplored territory. *Annu. Rev. Virol.* **4**, 201–219 (2017).
10. A. A. Pratama, J. D. van Elsland, The 'neglected' soil virome—Potential role and impact. *Trends Microbiol.* **26**, 649–662 (2018).
11. Y. Kuzaykov, K. Mason-Jones, Viruses in soil: Nano-scale undead drivers of microbial life, biogeochemical turnover and ecosystem functions. *Soil Biol. Biochem.* **127**, 305–317 (2018).
12. N. Fierer, Embracing the unknown: Distinguishing the complexities of the soil microbiome. *Nat. Rev. Microbiol.* **15**, 579–590 (2017).
13. G. Trubl, P. Hyman, S. Roux, S. T. Abedon, Coming-of-age characterization of soil viruses: A user's guide to virus isolation, detection within metagenomes, and viromics. *Soil Syst.* **4**, 23 (2020).
14. C. Santos-Medellin *et al.*, Viromes outperform total metagenomes in revealing the spatiotemporal patterns of agricultural soil viral communities. *ISME J.* **15**, 1956–1970 (2021).
15. A. M. Ter Horst *et al.*, Minnesota peat viromes reveal terrestrial and aquatic niche partitioning for local and global viral populations. *Microbiome* **9**, 233 (2021).
16. E. P. Starr, E. E. Nuccio, J. Pett-Ridge, J. F. Banfield, M. K. Firestone, Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 25900–25908 (2019).
17. G. Trubl *et al.*, Active virus-host interactions at sub-freezing temperatures in Arctic peat soil. *Microbiome* **9**, 208 (2021).
18. J. B. Emerson *et al.*, Host-linked soil viral ecology along a permafrost thaw gradient. *Nat. Microbiol.* **3**, 870–880 (2018).
19. G. Trubl *et al.*, Soil viruses are underexplored players in ecosystem carbon processing. *mSystems* **3**, e00076-18 (2018).
20. M. W. Van Goethem, T. L. Swenson, G. Trubl, S. Roux, T. R. Northen, Characteristics of wetting-induced bacteriophage blooms in biological soil crust. *MBio* **10**, e02287-19 (2019).
21. R. Wu *et al.*, DNA viral diversity, abundance, and functional potential vary across grassland soils with a range of historical moisture regimes. *MBio* **12**, e0259521 (2021).
22. F. Schulz *et al.*, Hidden diversity of soil giant viruses. *Nat. Commun.* **9**, 4881 (2018).
23. R. Tecon, D. Or, Biophysical processes supporting the diversity of microbial life in soil. *FEMS Microbiol. Rev.* **41**, 599–623 (2017).
24. R. L. Wilpiseski *et al.*, Soil aggregate microbial communities: Towards understanding microbiome interactions at biologically relevant scales. *Appl. Environ. Microbiol.* **85**, e00324-19 (2019).
25. A. Erktan, D. Or, S. Scheu, The physical structure of soil: Determinant and consequence of trophic interactions. *Soil Biol. Biochem.* **148**, 107876 (2020).
26. D. R. Nemerugut *et al.*, Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* **77**, 342–356 (2013).
27. J. C. Nekola, P. S. White, The distance decay of similarity in biogeography and ecology. *J. Biogeogr.* **26**, 867–878 (1999).
28. E. J. Sayer *et al.*, Grassland management influences spatial patterns of soil microbial communities. *Soil Biol. Biochem.* **61**, 61–68 (2013).
29. X.-B. Wang *et al.*, Habitat-specific patterns and drivers of bacterial β-diversity in China's drylands. *ISME J.* **11**, 1345–1358 (2017).
30. M. Feng *et al.*, Interpreting distance-decay pattern of soil bacteria via quantifying the assembly processes at multiple spatial scales. *MicrobiologyOpen* **8**, e00851 (2019).
31. M. Vos, P. J. Birkett, E. Birch, R. I. Griffiths, A. Buckling, Local adaptation of bacteriophages to their bacterial hosts in soil. *Science* **325**, 833 (2009).
32. J. P. Schimel, Life in dry soils: Effects of drought on soil microbial communities and processes. *Annu. Rev. Ecol. Syst.* **49**, 409–432 (2018).
33. R. L. Barnard, C. A. Osborne, M. K. Firestone, Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. *ISME J.* **7**, 2229–2241 (2013).
34. V. Masson-Delmotte *et al.*, Eds. "Global warming of 1.5 C" (IPCC, 2018).
35. D. Naylor *et al.*, Soil microbiomes under climate change and implications for carbon cycling. *Annu. Rev. Environ. Resour.* **45**, 29–59 (2020).
36. N. J. Bouskill *et al.*, Pre-exposure to drought increases the resistance of tropical forest soil bacterial communities to extended drought. *ISME J.* **7**, 384–394 (2013).
37. A. S. Amend *et al.*, Microbial response to simulated global change is phylogenetically conserved and linked with functional potential. *ISME J.* **10**, 109–118 (2016).
38. J. W. Sorensen *et al.*, DNase treatment improves viral enrichment in agricultural soil viromes. *mSystems* **6**, e0061421 (2021).
39. C. Fossum *et al.*, Belowground allocation and dynamics of recently fixed plant carbon in a California annual grassland. *Soil Biol. Biochem.* **165**, 108519 (2022).
40. K. Kieft, Z. Zhou, K. Anantharaman, VIBRANT: Automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* **8**, 90 (2020).
41. S. Roux *et al.*, Minimum information about an uncultivated virus genome (MIUViG). *Nat. Biotechnol.* **37**, 29–37 (2019).
42. L. S. Hillary, E. M. Adriaenssens, D. L. Jones, J. E. McDonald, RNA-viromics reveals diverse communities of soil RNA viruses with the potential to affect grassland ecosystems across multiple trophic levels. *ISME Communications* **2**, 1–10 (2022).
43. D. M. Durham *et al.*, Substantial differences in soil viral community composition within and among four Northern California habitats. *ISME Communications* **2**, 1–5 (2022).
44. M. J. Choudoir, K. M. DeAngelis, A framework for integrating microbial dispersal modes into soil ecosystem ecology. *iScience* **25**, 103887 (2022).
45. S. Sasidharan *et al.*, Transport and retention of bacteria and viruses in biochar-amended sand. *Sci. Total Environ.* **548–549**, 100–109 (2016).
46. M. Kimura, Z.-J. Jia, N. Nakayama, S. Asakawa, Ecology of viruses in soils: Past, present and future perspectives. *Soil Sci. Plant Nutr.* **54**, 1–32 (2008).
47. B. Koskella, M. A. Brockhurst, Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **38**, 916–931 (2014).
48. B. J. Woodcroft *et al.*, Genome-centric view of carbon processing in thawing permafrost. *Nature* **560**, 49–54 (2018).
49. J. T. Lennon, M. E. Muscarella, S. A. Placella, B. K. Lehmkuhl, How, when, and where relic DNA affects microbial diversity. *MBio* **9**, e00637-18 (2018).
50. R. Takahashi *et al.*, High frequency of phage-infected bacterial cells in a rice field soil in Japan. *Soil Sci. Plant Nutr.* **57**, 35–39 (2011).
51. M. Haldal, G. Bratbak, Production and decay of viruses in aquatic environments. *Mar. Ecol. Prog. Ser.* **72**, 205–212 (1991).
52. P. Carini *et al.*, Effects of spatial variability and relic DNA removal on the detection of temporal dynamics in soil microbial communities. *MBio* **11**, e02776-19 (2020).
53. K. J. Locey *et al.*, Dormancy dampens the microbial distance-decay relationship. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190243 (2020).
54. J. C. Ignacio-Espinoza, N. A. Ahlgren, J. A. Fuhrman, Long-term stability and Red Queen-like strain dynamics in marine viruses. *Nat. Microbiol.* **5**, 265–271 (2019).
55. A. Crits-Christoph, M. R. Olm, S. Diamond, K. Bouma-Gregson, J. F. Banfield, Soil bacterial populations are shaped by recombination and gene-specific selection across a grassland meadow. *ISME J.* **14**, 1834–1846 (2020).
56. M. R. Olm *et al.*, inStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. *Nat. Biotechnol.* **39**, 727–736 (2021).
57. J. B. H. Martiny, S. E. Jones, J. T. Lennon, A. C. Martiny, Microbiomes in light of traits: A phylogenetic perspective. *Science* **350**, aac9323 (2015).
58. H. Bin Jang *et al.*, Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat. Biotechnol.* **37**, 632–639 (2019).
59. A. Baryshnikova, Systematic functional annotation and visualization of biological networks. *Cell Syst.* **2**, 412–421 (2016).
60. G. Bonilla-Rosso, T. Steiner, F. Wichmann, E. Bexkens, P. Engel, Honey bees harbor a diverse gut virome engaging in nested strain-level interactions with the microbiota. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 7355–7362 (2020).
61. D. Turner, A. M. Kropinski, E. M. Adriaenssens, A roadmap for genome-based phage taxonomy. *Viruses* **13**, 506 (2021).
62. L. Xu *et al.*, Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4284–E4293 (2018).
63. K. Felsmann *et al.*, Soil bacterial community structure responses to precipitation reduction and forest management in forest ecosystems across Germany. *PLoS One* **10**, e0122539 (2015).
64. C. Santos-Medellin *et al.*, Prolonged drought impacts lasting compositional changes to the rice root microbiome. *Nat. Plants* **7**, 1065–1077 (2021).
65. A. Shade *et al.*, Fundamentals of microbial community resistance and resilience. *Front. Microbiol.* **3**, 417 (2012).
66. P. Carini *et al.*, Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat. Microbiol.* **2**, 16242 (2016).
67. A. M. Nicolas *et al.*, Soil candidate phyla radiation bacteria encode components of aerobic metabolism and co-occur with nanoarchaea in the rare biosphere of rhizosphere grassland communities. *mSystems* **6**, e0120520 (2021).
68. C. Corinaldesi, M. Tangherlini, G. M. Luna, A. Dell'anno, Extracellular DNA can preserve the genetic signatures of present and past viral infection events in deep hypersaline anoxic basins. *Proc. Biol. Sci.* **281**, 20133299 (2014).
69. P. C. Göller, J. M. Haro-Moreno, F. Rodriguez-Valera, M. J. Loessner, E. Gómez-Sanz, Uncovering a hidden diversity: Optimized protocols for the extraction of dsDNA bacteriophages from soil. *Microbiome* **8**, 17 (2020).
70. J. B. Emerson *et al.*, Dynamic viral populations in hypersaline systems as revealed by metagenomic assembly. *Appl. Environ. Microbiol.* **78**, 6309–6320 (2012).
71. J. G. Caporaso *et al.*, Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U.S.A.* **108** (suppl. 1), 4516–4522 (2011).
72. J. Edwards, C. Santos-Medellin, V. Sundaresan, Extraction and 16S rRNA sequence analysis of microbiomes associated with rice roots. *Bio Protoc.* **8**, e2884 (2018).
73. L. R. Thompson *et al.*, Earth Microbiome Project Consortium, A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017).
74. S. J. Salter *et al.*, Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).
75. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
76. B. Bushnell, BBTools software package. <https://sourceforge.net/projects/bbmap/> (2014).
77. D. Li, C.-M. Liu, R. Luo, K. Sadakane, T.-W. Lam, MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
78. M. R. Olm, C. T. Brown, B. Brooks, J. F. Banfield, dRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
79. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
80. H. Li *et al.*, 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
81. D. Hyatt *et al.*, Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
82. B. Buchfink, C. Xie, D. H. Huson, Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).

83. S. M. van Dongen, "Graph clustering by flow simulation," PhD thesis, Utrecht University, Utrecht, The Netherlands (2000).
84. B. Schloerke, *et al.*, GGally: Extension to "ggplot2" (2018).
85. E. Kopylova, L. Noé, H. Touzet, SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
86. 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).
87. Q. Wang, G. M. Garrity, J. M. Tiedje, J. R. Cole, Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
88. J. R. Cole *et al.*, Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642 (2014).
89. J. Quensen, RDPutils: R Utilities for Processing RDPTool Output (2018).
90. A. P. Masella, A. K. Bartram, J. M. Truszkowski, D. G. Brown, J. D. Neufeld, PANDAseq: Paired-end assembler for illumina sequences. *BMC Bioinformatics* **13**, 31 (2012).
91. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
92. J. G. Caporaso *et al.*, QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
93. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2018).
94. J. Oksanen, *et al.*, vegan: Community Ecology Package (2018).
95. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
96. M. De Cáceres, P. Legendre, Associations between species and groups of sites: Indices and statistical inference. *Ecology* **90**, 3566–3574 (2009).
97. T. Hothorn, *et al.*, Multcomp: Simultaneous inference in general parametric models (R package version, 1–3) (2014).
98. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
99. H. Wickham, ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag, New York, 2016).
100. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal, Complex Systems* **1695**, 1–9 (2006).
101. C. Santos-Medellín, Spatial Structuring and Response to Moisture of Soil Viral Communities in Grasslands. NCBI: Sequence Read Archive. <https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA818793>. Deposited 23 March 2022.
102. C. Santos-Medellín, Viral metagenomic sequences from Mediterranean grassland soils. Zenodo. <https://zenodo.org/record/7076890#Y1BMPEzMJJPY>. Deposited 14 September 2022.
103. C. Santos-Medellín, Code and data from: Spatial turnover of soil viral populations and genotypes overlain by cohesive responses to moisture in grasslands. GitHub. <https://github.com/cmsantosm/HoplandViromes>. Deposited 14 September 2022.