forests

MDPI

Article

Diversity and Potential Interactions of Soil Viruses and Host Bacteria under Different Land Use Patterns

Yuting Yan [1,2], Danting Yu [1,2,*], Lili Han [3,4], Chengyu Yuan [1,2] and Jizheng He [2,5]

1 Fujian Provincial Key Laboratory for Subtropical Resources and Environment, Fujian Normal University, Fuzhou 350117, China
2 Key Laboratory of Humid Subtropical Eco-Geographical Process (Ministry of Education), Fujian Normal University, Fuzhou 350117, China
3 State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China
4 University of the Chinese Academy of Sciences, Beijing 100049, China
5 Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville, VIC 3010, Australia
* Correspondence: dty@fjnu.edu.cn; Tel.: +86-157-5082-0719

Abstract: Viruses, as the most abundant entities on earth, play an important role in shaping bacterial communities, mediating gene transfer between host cells, and promoting biogeochemical cycles. Yet, soil viruses remain understudied, as there is a lack of information about the mechanisms of community construction, interactions between viruses and host bacteria, and ecological functions. To expand our understanding of soil viruses, we investigated six viromes across three land use types in northeast and southwestern China, including agricultural and forest soils. We analyzed viral and bacterial community composition and explored their interactions. We utilized metagenomic sequencing technology and high-throughput 16S rRNA gene sequencing to study viral and bacterial communities. Twenty-four viral families were detected in six viromes including sixteen dsDNA virus families and eight ssDNA virus families. Viral and bacterial communities were significantly different between various land use types, with viruses having greater disparities. The composition of bacterial communities in soils across different land use types was inconsistent with their viral communities. We identified abundant auxiliary carbohydrate-active enzyme genes from viromes. The results revealed that soil viral communities differ by land use type and that viruses could regulate bacterial carbon cycling processes by encoding auxiliary metabolic genes.

Keywords: viruses; virome; land use; auxiliary metabolic genes

Citation: Yan, Y.; Yu, D.; Han, L.; Yuan, C.; He, J. Diversity and Potential Interactions of Soil Viruses and Host Bacteria under Different Land Use Patterns. *Forests* **2023**, *14*, 342. https://doi.org/10.3390/f14020342

Academic Editor: Paulo A. Zaini

Received: 13 January 2023
Revised: 3 February 2023
Accepted: 6 February 2023
Published: 9 February 2023

1. Introduction

The soil environment is likely the most complex and competitive biome on Earth, harboring huge numbers of living organisms. The organisms belowground, especially microorganisms, play significant roles in nutrient cycles and energy flow. Soil biotic community composition has a strong relationship with soil health and fertility [1,2]. These microorganisms are constantly exposed to infection by viruses present in the surrounding environment.

Viruses, which lack independent metabolisms, are ubiquitous in nature and living organisms can be infected by one or several viruses. In soil, the majority of viruses are phages, which infect bacteria [3]. Phages have four typical life cycles: lytic, lysogenic, pseudolysogeny, and chronic infection [4]. In the lytic cycle, virulent phages infect host cells and phage progeny are released by destroying the host cell. The lytic cycle has a large impact on bacteria through mechanisms such as influencing bacterial mortality, shaping bacterial community diversity, accelerating the transformation of nutrients from living organisms to dissolved and particulate organic matter, and leading to generalized transduction by erroneous encapsidation [5–7]. Lysogeny is common among soil bacteria,

with temperate phages existing as prophages by integrating their DNA into host genomes or being carried as plasmids [8,9]. Temperate phages can impact the fitness and phenotype of the host and cause specialized transduction by lysogenic conversion [9,10]. Some viruses act as agents of transduction, mistakenly packaging microbial genetic material into their capsid and then transferring it to a new host cell; this behavior may enhance the fitness and competitiveness of their hosts by improving microbial adaptation to the environment, increasing the phage's chances of reproduction. Pseudolysogeny is an unstable state, where phages exist in host bacterial cells without multiplying or synchronizing their replication with the host [4]. In the chronic infection cycle, phages are shed from the host cell without obvious cell death [4]. Viruses, especially phages, can impact microbial composition and function, complicating soil nutrient cycles and food web dynamics.

Viral abundance in soil can be measured by transmission electron microscopy and epifluorescence microscopy. The amount of viruses is about 10^7~10^9 per gram (dry weight) of soil [11–13]. Viruses belonging to the families *Myoviridae*, *Podoviridae*, *Siphoviridae*, and *Microviridae* are widespread in the soil ecosystem [11,12,14–18]. Despite the copious amounts of viruses in soil, their potential interactions with host microorganisms are poorly understood. Research on phage infection in soil using electron microscopy revealed that the frequency of visibly infected bacterial cells ranged from 8.9% to 48% in sheep-grazed pastures and rice field soils [19,20]. The true phage infection rate is certainly underestimated because the virus particles were only detected after assembly in host cells and temperate phages cannot be detected in the lysogenic cycle. Recent studies on soil viruses in the permafrost thaw gradient found that viruses infected bacteria (*Acidobacteria*, *Verrucomicrobia*, *Deltaproteobacteria*) with key biogeochemical functions; since viruses impact host complex carbon degradation via encoding glycoside hydrolase genes, viruses may influence soil carbon cycles [16]. The recent identification of more auxiliary metabolic genes (AMGs) in viromes suggests that viruses may impact host-mediated biogeochemistry and ecosystem function [21]. Auxiliary carbohydrate-active enzyme (CAZyme) genes involved in manipulating carbon cycles were also detected in mangrove and agricultural soil viromes [22,23]. Host bacteria can survive in different environments with the help of soil viruses, which carry AMGs associated with energy acquisition, stress tolerance, and the degradation of xenobiotics [24].

In order to meet human need, a large proportion of natural landscapes have been converted to human-dominated lands. Land use practices can have long-term impacts on soil microbial communities and ecosystem services [25,26]. However, the activities and microbial processes—especially of viruses—across different land use practices are understudied. Here, we utilized a metagenomic approach to characterize viral communities in soils under different land use practices and analyzed viral and bacterial diversity to explore interactions between viruses and their host microorganisms. The aims of this study were to gain a better understanding of the distribution and diversity of viruses in soils under different land use practices, to investigate the interactions between viruses and bacteria, and to explore the potential roles of viruses in soil ecosystem services.

2. Materials and Methods

2.1. Sample Sites

Soil sample sites under different land use types including forest, paddy field, and dryland were located in the major agricultural areas of China (Jilin and Sichuan provinces). Three soil samples were collected in Summer 2016 from typical Chinese dryland-farming regions in the Jilin province, and the other three soil samples were obtained in Summer 2017 from Chinese paddy agricultural areas in the Sichuan province (Figure S1, Table S1). The forest soil samples (NE_F, SW_F) were collected from natural forests without human disturbance, belonging to the temperate continental monsoon and sub-tropical humid monsoon climates, respectively. For each sampling site, triplicate soil samples (0~20 cm depth) were collected, homogenized, and divided into two fractions. One fraction was stored at 4 °C to measure soil chemical properties, while the other was utilized for high

throughput analysis. The soil's chemical properties of organic matter content (OM), total nitrogen (TN), available phosphorus (AP), and available potassium (AK) were produced by standard methods at the Chinese Academy of Agricultural sciences (Beijing, China). The measurements of electric conductivity (EC) value and pH were determined by a pH-meter (Professional Meter PP-20, Sartorius, Germany).

2.2. Sample Collection for Viral Metagenomes

We collected the virus-like particles (VLPs) from soil samples as previously described [17]. Briefly, at least 500 g of soil was suspended in 1.5 L glycine buffer (250 mM; pH = 8.5) and viruses were physically dispersed by shaking for 30 min on ice. Then, the solution was centrifuged at 3500 rpm to separate and collect the supernatant; this process was repeated 3 times to improve viral recovery. The total supernatant was passed through a 0.45-µm, 0.2-µm tangential flow filter (GE Healthcare Life Sciences, Pittsburgh, PA, USA) to remove nonvirus particles, then concentrated by 30-KDa tangential flow filter. After initial concentration, less than 100 mL enrichment was obtained; this was concentrated again, this time using a 30-KDa centrifugal ultrafiltration tube (Merck Millipore Ltd., Tullagreen, Ireland) until the volume was about 1 mL. DNase I (Thermo Fisher Scientific, Carlsbad, USA) was added to the solution (10 units DNaseI/100 µL) and incubated at 37 °C for 1 h. The presence of free and bacterial DNA contaminants were checked by 16S rRNA gene amplification with primers E9F and U1510R [14]. The viral DNA was extracted by AllPrep PowerViral DNA/RNA kit (Qiagen, Hilden, Germany), cleaned up using DNA Clean & Concentrator kit (Zymo Research, Orange, CA, USA), and amplified using REPLI-g Mini Kit (Qiagen, Hilden, Germany). After fragmentation of viral genomes, six viral DNA libraries were constructed with TruSeq™ DNA Sample Prep Kit (Illumina, San Diego, CA, USA) and sequenced via Illumina NovaSeq at Shanghai Majorbio Bio-pharm Biotechnology Co., Ltd. (Shanghai, China), generating 300 bp paired-end reads.

2.3. Assembly and Identification of Viral Contigs

After sequencing, a fastp tool with useful quality control and data-filtering features filtered low quality reads and checked for primer/adapter contamination to achieve high-quality reads data [27]. The clean reads were then assembled by metaSPAdes, retaining contigs > 1000 bp [28]. The contigs from the six samples were merged into a fasta file and clustered by CD-HIT at the 95% level (http://www.bioinformatics.org/cd-hit/, accessed on 1 March 2022) to generate a non-redundant set of contigs. These contigs were aligned via BLASTx against viral sequences from the reference sequence database at NCBI [29], phage sequences from the PHASTE web server [30], and viral protein sequences from the NCBI-nr Database in March 2022. Viral taxonomic information was generated after accession numbers were matched with the taxonomy database. The viruses were then divided into eight groups by their host ranges.

2.4. Analysis of Viromes

Viral contigs that contained the genes for integrases, site-specific recombinases, and transposases matching those found in the Pfam and KEGG databases were selected and used to predicted putative temperate phages [31]. Auxiliary carbohydrate-active enzyme (CAZyme) genes from viral contigs were identified on the dbCAN meta server (http://bcb.unl.edu/dbCAN2/index.php, accessed on 20 May 2022), a web server for automated carbohydrate-active enzyme annotation. Annotations were based on the dbCAN CAZyme domain HMM database, CAZy database, and PPR library, using HMMER, DIAMOND, and Hotpep with default parameter values.

2.5. Phage Hosts Prediction

In order to explore putative hosts of phages at the genome level, a CRISPR-Cas spacer database was constructed from bacterial genomes from the NCBI RefSeq database using the CRISPR Recognition Tool (CRT, http://www.room220.com/crt/, accessed on 20 May

2022). The sequence similarity between viral contigs and the spacers in bacterial genomes was generated by Blastn-short 2.9.0 (E-value $\leq 10^{-5}$, bitscore ≥ 45). Connections between bacterial hosts and phages were developed at the genus level and the host range of viruses was identified by Virus–host DB (https://www.genome.jp/virushostdb/view/, accessed on 20 May 2022), ViralZone (https://viralzone.expasy.org/, accessed on 20 May 2022), and manual research of the published literature.

2.6. Bacterial Community Analyses

The total soil microbial DNA was extracted from 0.25 g soil by PowerSoil DNA isolation kit (Qiagen, Hilden, Germany) and libraries were generated by high-throughput sequencing on an Illumina MiSeq at Shanghai Majorbio Bio-pharm Biotechnology Co., Ltd. (Shanghai, China) after polymerase chain reaction (PCR) amplification of the 16S rRNA gene (806R/338F). Based on USEARCH software 10.0.240, raw sequences were truncated, merged, and filtered to remove redundant sequences. The clean sequences were clustered and operational taxonomic units (OTUs) were generated at a 97% similarity level. The OTUs were aligned via BLAST against the silva_97 database and further bioinformatics analysis was performed by R 3.5.1.

2.7. Network Analysis

In order to further explore the relationship between viruses and bacterial hosts, five additional viromes from previous studies of agricultural soil samples [17,18] were included in network analysis using the CoNet plug-in in Cytoscape v3.5.0 [32]. To reduce the complexity of calculation and build a clear relationship between viruses and bacteria, the viruses used in the analysis were either phages (viruses that infect bacteria) or viruses without exact host information. The correlation scores were calculated using Pearson correlation, Spearman correlation, mutual information, Bray–Curtis dissimilarity, and Kullback–Leibler dissimilarity with automatic threshold settings and networks optimized by randomization (iterations: 100, routine: edgeScores, resampling: shuffle_rows/bootstrap, p-value threshold ≤ 0.05).

3. Results

3.1. Diversity and Distribution of Viral Communities

This study collected six soil samples, namely NE_P, NE_M, NE_F, SW_P, SW_M, and SW_F. Overall, we obtained 108.2 million raw reads. After quality control, the six viromes ranged in size from 0.8 to 9.0 million reads, and a total of 23,433 viral contigs (shortest contig length: 1.0 kb, longest contig length: 8.6 Mb) were successfully assembled.

Twenty-four viral families were identified, of which seventeen were dsDNA virus families and seven were ssDNA virus families. As shown in Figure 1, five of the six viromes (all but the SW_M virome) contained a small number of human viruses (0.01%~4.64%) coming from families *Circoviridae*, *Coronaviridae*, *Genomoviridae*, *Hepadnaviridae*, *Herpesviridae*, *Poxviridae*, and *Retroviridae*. The abundance of human viruses in the SW_P virome (4.64%) was the highest. Agricultural activities might introduce human viruses to the local environment. A moderate quantity of non-human vertebrate (2.70%~18.45%) viruses and invertebrate viruses (1.72%~55.76%) were also detected. These viruses were mainly from the families *Circoviridae* and *Genomoviridae*. Animal viruses infecting vertebrates and invertebrates were more abundant in the three NE viromes. Most notably, 72.02% of the viruses in the NE_M virome were classified as animal viruses. The SW_P virome—which originated in paddy soil—had the highest proportion (11.69%) of plant viruses. In contrast, the two viromes from forest soil had the lowest proportion of plant viruses. The representative viral groups were the families *Alphasatellitidae* and *Nanoviridae*. Viruses of eukaryotic microorganisms were more common in the NE_F (1.83%) and SW_F (1.91%) viromes. In the other four viromes, all from soil under human activity, less than 0.03% of viruses were annotated as viruses of eukaryotic microorganisms. The percentage of phages ranged between 17.80% and 72.09% in the six viromes. The NE_F (57.30%) and SW_F (72.1%) vi-

romes had higher percentages of phages than the viromes which derived from agricultural soils. The phages mainly belonged to the families *Microviridae*, *Myoviridae*, *Podoviridae*, and *Siphoviridae*, with a small fragment of *Inoviridae* also observed. Fungi viruses were only detected in the SW_P virome (1.23%) and the viruses were classified into the *Genomoviridae* family. This viral group was similar to plant viruses, but *Gemycircularvirus* and *Rhizidiovirus* were identified as DNA viruses targeting fungi.

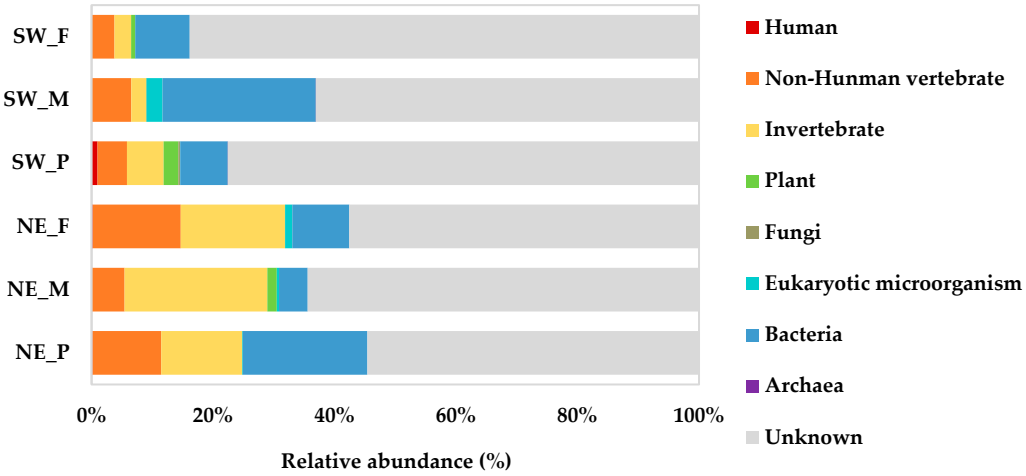

Figure 1. The distribution of viral contigs by host range in six soil viromes. According to the host range on ViralZone, viruses were classified into eight groups. In view of increased attention on the viruses which can infect humans, these viruses were independent from vertebrate viruses. Eukaryotic host: human, non-human vertebrate, invertebrate, plant, fungi, eukaryotic microorganism. Prokaryotic host: bacteria, archaea.

Viral distribution differed between agricultural and forest soils, and the high prevalence of unique viral species found in viromes resulted in a rather low similarity of viral communities (Figure 2). The relative abundance of common species was rather low between agricultural soils, but higher between agricultural and forest soils. The three NE viromes shared seventeen identical viral species. The NE_P and NE_M viromes contained thirty-four of the same species, but there were only ten and six identical viral species between the NE_F virome and the viromes collected from paddy and dryland soils, respectively. The three SW viromes showed a similar trend, with the SW_P and SW_M viromes sharing a relatively higher percentage of identical viral species. SW_F only had thirteen and eight identical viral species with SW_P and SW_M, respectively. The common viral species of the NE viromes and SW viromes were similar, with eleven viral species widespread in the six viromes; they belonged to the families *Circoviridae*, *Cruciviridae*, and *Microviridae*. The *Microviridae* and *Circoviridae* families were widespread in the viral communities. The viral species were richer in the SW viromes than in the NE viromes. Distance-based redundancy analysis (db-RDA) with Bray–Curtis distance revealed that biotic and abiotic factors (especially AP) driven by spatial and geographic variation and human activities influenced viral species richness (Figure 3).

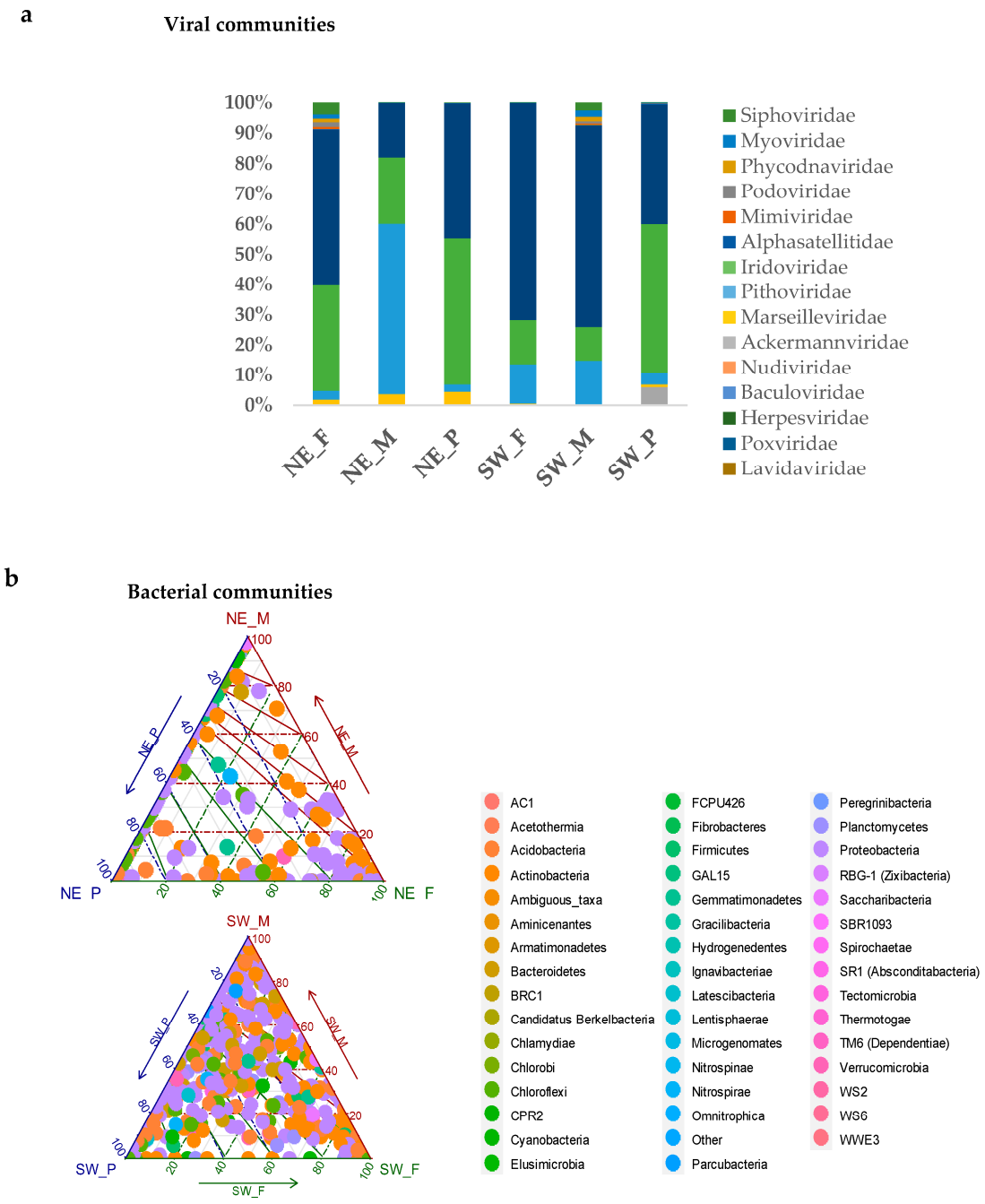

Figure 2. *Cont.*

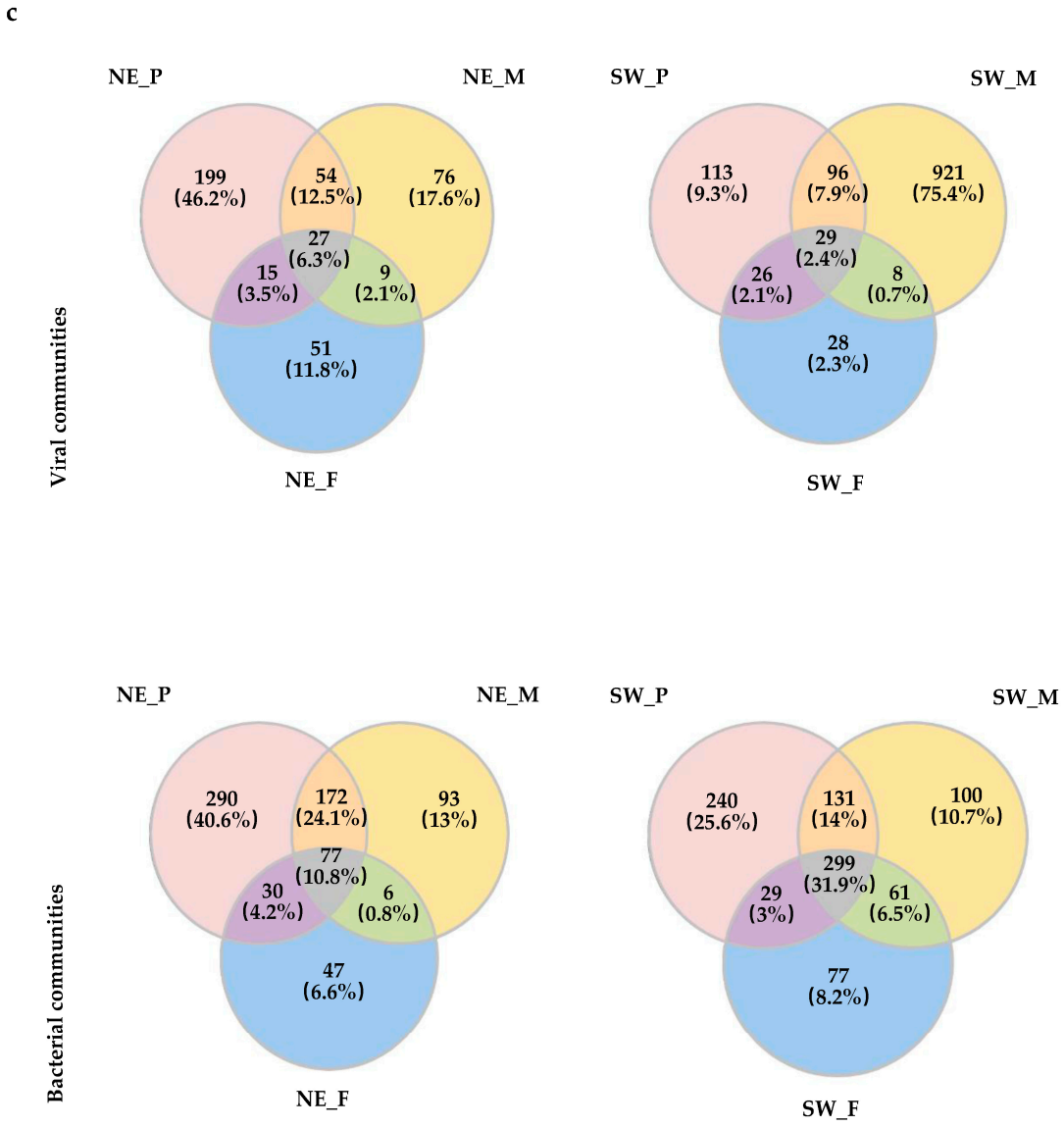

Figure 2. Overview of viral and bacterial communities. (**a**) The distribution of six soil virome communities. (**b**) Ternary plot depicting compartment at species level within different land use patterns. (**c**) Venn diagram at species level. The numbers refer to the number of species and the numbers in parentheses refer to percentage.

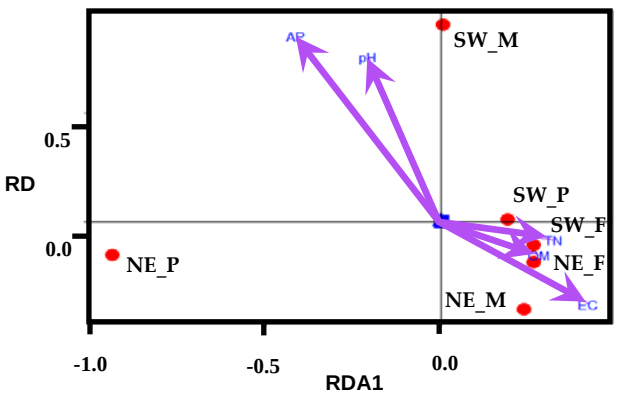

Figure 3. Redundancy analysis on the relationship between soil chemical properties and viral communities.

Temperate phages can undergo virion-productive or lysogenic cycles. Most temperate phages encode integrase, site-specific recombinases, and transposases for integration and excision. The previous work on soil viromes has indicated that lysogenic infection is common in soils [9,33]. Table S2 shows that only twenty-six viral contigs contained target proteins and that these contigs were all derived from the NE_P and SW_P viromes.

3.2. Distribution of Bacterial Populations

Community similarity analysis showed that the bacterial communities had experienced obvious differentiation (Figure 2). The Venn diagrams of bacterial communities showed that many species of bacteria were shared between agricultural soils, but there were fewer common species between forest and agricultural soils. The most abundant bacterial groups belonged to the phyla *Actinobacteria*, *Proteobacteria*, *Chloroflexi*, and *Acidobacteria*.

3.3. The Linkages between Viruses and Potential Hosts

The majority of bacterial and phage species in the six soil samples did not have established infection relationships based on the Virus–host database. Of the bacterial species with known phage information, NE bacterial species mainly belonged to the *Actinobacteria*, *Firmicutes*, and *Proteobacteria* phyla, while SW bacterial species came from the *Actinobacteria*, *Bacteroidetes*, and *Proteobacteria* phyla (Figure 4). The *Streptomyces* genus had a higher percentage (8.08%~43.16%) in the six bacterial libraries, but the *Streptomyces* phages (*Siphoviridae*) accounted for a tiny proportion in the six viromes. The *Bacillus* genus was predominant in NE_P (40.56%) and NE_M (66.08%) libraries. In contrast, the *Bacillus* phages (belonging to *Myoviridae*, *Podoviridae*, and *Siphoviridae*) had a higher proportion in the NE_F (21.21%) and SW_M (32.92%) viromes. The percentage of the *Stenotrophomonas* genus was as high as 51.02% in the SW_M library and the related phages (belonging to *Myoviridae* and *Siphoviridae*) only accounted for 1.74% of the SW_M viromes. The related abundance of *Myoviridae* and *Siphoviridae*, which infected *Bacillus*, was highest in the SW_M virome and lowest in NE_M and SW_F viromes; however, the relative abundance of their bacterial hosts did not appear to follow any observable tendency. Overall, the trend of the bacterial hosts and related phages was not consistent since the majority of the bacteria had no phage information.

To further study the relationship between phages and potential hosts and examine the impacts of viruses on their hosts, we investigated the CRISPR/Cas system of bacteria from the six soil samples. Only seventy viral contigs had robust hits on the CRISPR/Cas spacer database. The bacterial hosts were from the phyla *Actinobacteria*, *Bacteroidetes*, *Chlamydiae*, *Chlorobi*, *Firmicutes*, *Planctomycetes*, *Proteobacteria*, and *Verrucomicrobia*, with several viral species linked to more than one bacterial genus (Table S3). Through analysis of CRISPR-Cas spacer matches in the six soil samples, we found *Microviridae* not only infect *Proteobacteria* and *Chlamydiae*, but may also infect *Actinobacteria*, *Bacteroidetes*, *Chlorobi*, *Firmicutes*, *Planctomycetes*, and *Verrucomicrobia*.

The genus *Aeromonas* was identified as the host of Aeromonas phage 62AhydR11PP. Five Apis mellifera-associated microviruses were predicted to infect *Megamonas*, *Prosthecochloris*, *Tyzzerella*, *Legionella*, and *Coprobacillus*; the phages might derive from the gut microbiota of the honey bee. The hosts of Blackfly microvirus SF02 were identified to be nine genera from *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, and *Chlorobi*. *Gokushovirinae* were widespread in different habitats [12,34], but their hosts were still unknown. We found the spacers recovered from *Enterobacter*, *Legionella*, *Phascolarctobacterium*, *Candidimonas*, and *Orbus* may result from infection by three species of *Gokushovirinae* (Bog1183_53, Fen672_31, GAIR4). Sequences of bacteria that may suffer from the infection of multiple *Microviridae* species included Clostridium_spiroforme, Clostridiales_bacterium_VE202-01, Enterobacter_roggenkampii, Legionella_pneumophila, Nitrosospira_sp._Nsp5, Nitrosospira_multiformis, and Prosthecochloris_sp._ZM_2. Because the CRISPR/Cas adaptive immune system is widespread in nearly all archaea and about half of all bacteria [35,36],

the high diversity of viruses and bacteria in local environments might explain the small number of CRISPR spacers deriving from viral infection found in bacterial genomes.

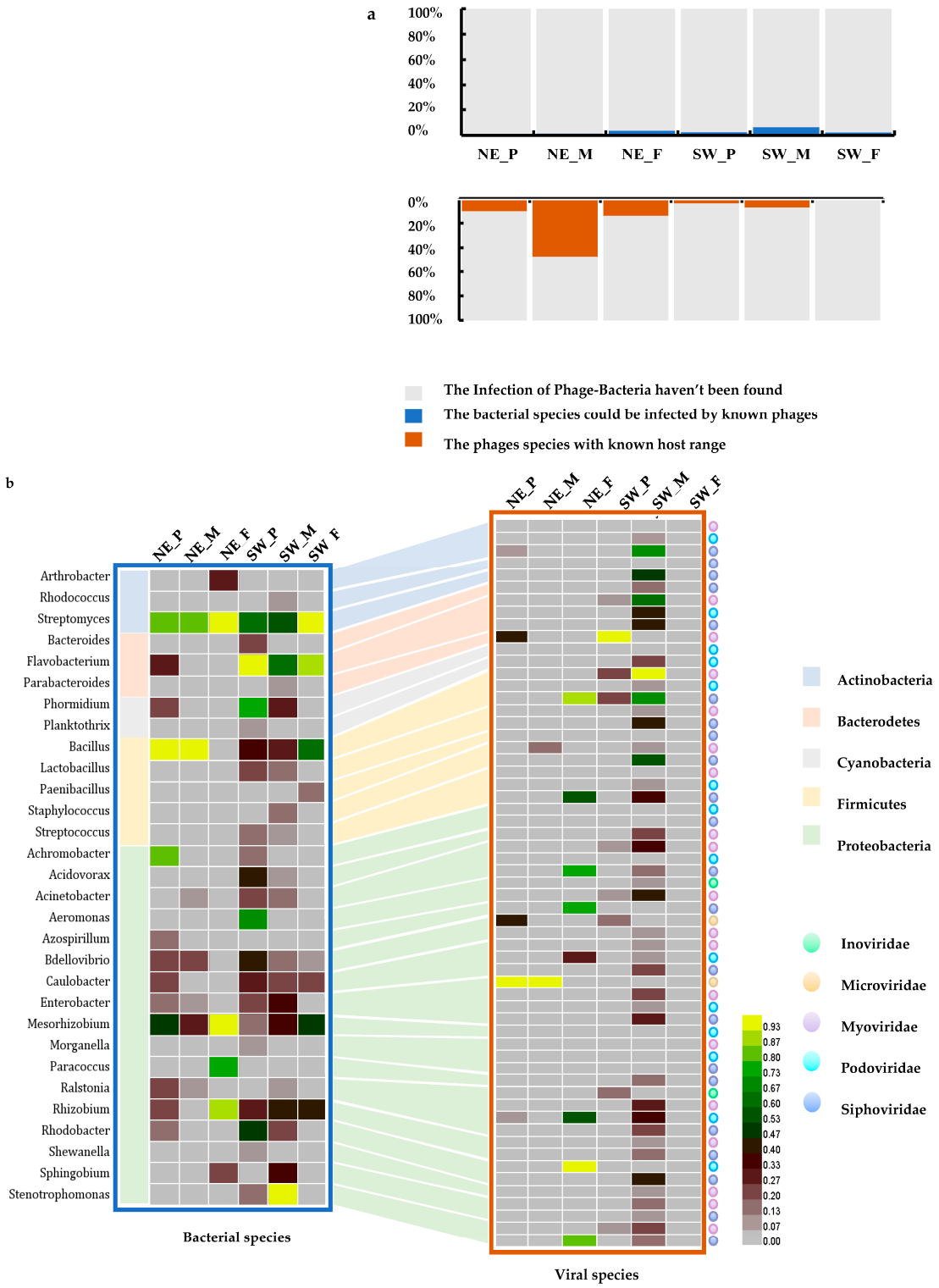

Figure 4. The relationship between bacteria and phages in the six viromes. (**a**) Proportion of bacterial genera and phage species that have a clear relationship of infection based on Virus–host database. (**b**) The distribution of bacterial genera and related phages species. On the left is a bacterial heat map and five different colors represent phylum level.

3.4. Co-Occurrence of Viruses and Bacteria Based on Network Analysis

For the purpose of determining the complex connections between viruses and bacteria in soil samples, multiple networks were constructed at the species level and networks were visualized in Gephi. (Figure 5). Nodes represented bacterial genera or phage species and connecting lines (edges) were colored by source. The initial network contained 518 nodes and 2331 edges, with the percentage of virus nodes only accounting for 12.55% and 10.51% of edges linking viruses to bacteria. In total, 61.97% of species were in the inter-connected subnetwork containing only bacteria, revealing the strong correlations between bacteria. After removing the internal connections of bacteria, the secondary network included 209 nodes and 866 edges. Despite the large numbers of viruses that have been detected in six viromes, only sixty-five viral species were assigned to viral groups in the network and all viral species mainly belonged to the families Microvirida and Cruciviridae. This simplified network showed a clearer representation of the weak correlations between bacteria and viruses.

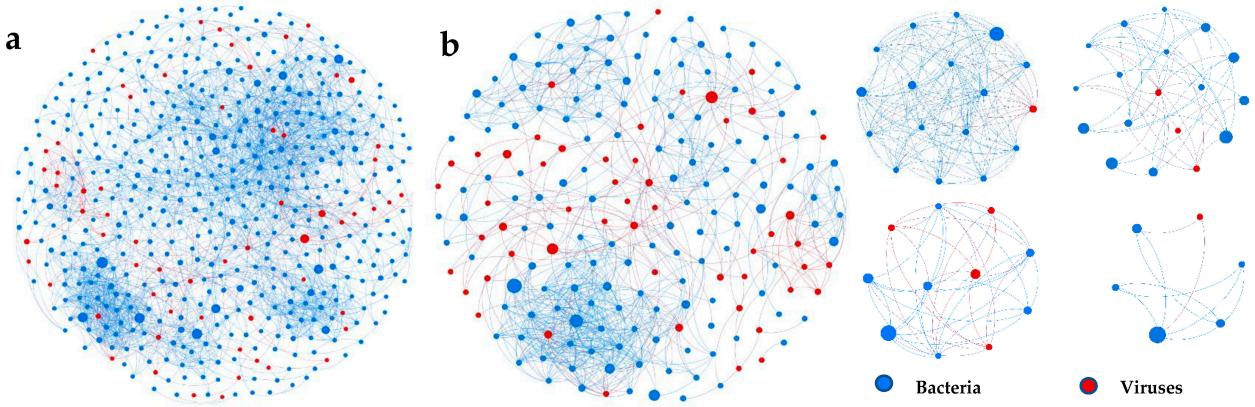

Figure 5. The relationship between bacteria and phages in the six viromes. (**a**) Proportion of bacterial genera and phage species that have a clear relationship of infection based on Virus–host database. (**b**) The distribution of bacterial genera and related phages species.

3.5. AMGs Carried within Viromes

A total of 933 viral contigs spanning 100 CAZyme genes were identified by dbCAN meta server. The detected genes were derived from five CAZymes classes: Auxiliary Activities (AAs), Carbohydrate Esterases (CEs), Glycoside Hydrolases (GHs) Glycosyl-Transferases (GTs), and Polysaccharide Lyases (PLs), as well as one module: Carbohydrate-Binding Modules (CBMs). SW_M virome contained 94 CAZymes, with most of them belonging to GHs and GTs. The viral contigs from the NE_P, NE_F, and SW_P viromes were affiliated with CBMs, GHs, GTs, and PLs at a low level.

4. Discussion

We focused on soil viral and bacterial communities in three land use types in order to explore the influence of human agricultural activities on distribution and diversity of viruses and bacteria. The results suggested that common species accounted for only a small proportion of the total population and that viral communities differed remarkably between local environments. The agricultural effects on viral communities and the induction of prophages from bacteria were not obvious. We speculate that agricultural activities increase the rates of virus–host collisions to some extent, resulting in high infection probability. In the lytic lifestyle, a mass of progeny virions is released after package completion, increasing viral richness but decreasing viral evenness in agricultural soils. We speculated that the dynamics of viral communities in soil environments may be impacted by the following situations: (i) under natural conditions, the virions suffer from various physical, chemical, and biotic factors and are lost or inactive, which decreases diversity; (ii) the infection of lytic

phages has a negative impact on biofilm stability and might cause biofilm dispersal, driving the lifecycle of the biofilm forward [37]; (iii) the complex soil environment and bacterial resistance to phages limit further infections. In time, with bacterial community recovery and natural deaths of viruses, viral diversity returns to what it once was; (iv) agricultural activities increase the collision rates of phages and their hosts, resulting in higher infection rates. This might promote the propagation of populations of scarce or foreign virions, maintaining or increasing viral diversity. As a result of complex agricultural activities, the reality is more complicated. Multiple processes interact to influence the soil viral and bacterial community directly or indirectly, and the agricultural effects are still unresolved.

Agricultural activities had positive effects on viral richness, which may be the result of the concerted action of multiple factors. As we know, soil communities are likely the most complex and diverse habitats on earth, with millions of species and billions of individual organisms possibly existing in a single ecosystem [38]. The activities of bacteria, fungi, archaea, viruses, and protists generate soil microbial networks with key roles in carbon and nitrogen cycling, structure maintenance, and ecosystem services [38,39]. In natural environments, the majority of bacteria are believed to reside in biofilm matrices [40], where they are protected from phage infection [41]. The semi-discrete and heterogeneous environments found in soil might further enhance the abiotic ecological pressures on virus–host interactions [42]. Terrestrial ecosystems suffer from increasing human-driven practices, which have altered microbial diversity and ecosystem function [43]. Anthropogenic nutrient inputs can stimulate soil biofilm formation and microaggregate development [44]. Agricultural irrigation and soil cultivation have significant effects on soil chemical properties and soil structure [26]. Unique characterization of viruses and the sophisticated relationships between infection and resistance make the impact of viral infection on the bacterial community complex. We posit that viruses tend to exit their infection lifestyle and exist as virions in soil environments, as frequent viral infection is inhibited by the semi-discrete surroundings and the existence of biofilm-limited infection by lytic viruses, the resistance to co-occurring viruses in bacterial cells, and the high viral pleiotropic costs of deleterious mutations. The dominant viruses are those which repeatedly infected hosts previously, but their host develops resistance from the frequent viral infection; this leads to coevolution in bacterial and viral communities. Even agricultural activities arguably increase the infection rates. The diversity of viruses was not as high as we predicted due to bacterial resistance, so infections were more likely to occur in scarce viruses and their hosts. Overall, we found that viral richness and evenness in agricultural soils were higher than in forest soils.

Although it is logical to look at bacteria and phages as predator–prey interactions, such that the distribution of bacteria limits the distribution of phages and phage infection shapes bacterial community structure, developing correlations between viral and bacterial community structures is still quite difficult [45]. In our study, bacterial populations were not observably correlated with viral infection. Virus-mediated selection pressure was complex and could not be explained by a single lifestyle or hypothesis.

Virus-mediated horizontal gene transfer might have positive influences on host fitness and competitiveness. For example, photosynthesis genes carried by cyanophages augmented host photosynthesis during infection [46–48]. We identified abundant CAZyme genes from viral contigs, most of which were glycoside hydrolases and glycosyl-transferases. Virally encoded glycoside hydrolase and glycosyl-transferases may potentially augment the breakdown of complex carbohydrates and accelerate saccharide synthesis to increase energy production, promote metabolism, and boost host metabolism during viral infection [49]. The CAZyme genes in six viromes were different from those in mangrove soil [22] and agricultural soil [23]. The CAZyme genes encoded by viruses may be concerned with the frequency of viral infection and local host metabolic level.

Viruses are ubiquitous in nature and can be found wherever life exists. Viral function and ecology have attracted much attention in the past ten years. Much is still unexplored, such as the detailed relationship of infection between viruses and hosts, how viruses

influence the dynamic of host communities, and co-evolution resulting from viral infection and host resistance. The results in our study illustrate the high viral diversity in the soil environment and the complex relationship between phages and bacteria. It should be noted that this study has examined only DNA viruses. The next important step is to also analyze RNA viruses, which will assist in the exploration of the viral community and assessment of ecological risk.

5. Conclusions

This study showed that the structures of soil viral and bacterial communities were influenced by land use types. Agricultural activities may have positive effects on the survival of rare viruses. The emergence of animal viruses existing in soils revealed the potential threat of viruses to soil and public health. Viruses can regulate microbial metabolic processes by encoding AMGs. These findings provide insight into the response of soil viruses to land use types.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/f14020342/s1, Table S1: Details of sampling sites. Table S2: The viral contigs containing genes encoding integrase, site-specific recombinases, and transposases in viromes based on Pfam and KEGG. Table S3: Phage–bacteria linkages between viral contigs and bacterial genomes by CRISPR spacer.

Author Contributions: Conceptualization, D.Y.; methodology, D.Y. and Y.Y.; formal analysis, D.Y. and Y.Y.; investigation, D.Y., Y.Y. and C.Y.; writing—original draft preparation, D.Y. and Y.Y.; writing—review and editing, D.Y., L.H. and J.H.; supervision, D.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 41907028, and the Natural Science Foundation of Fujian Province, China, grant number 2021J01180.

Data Availability Statement: Most of the data presented in this study are contained within the article and in the Supplementary Materials. Data not shown in the article are available on request from the corresponding author.

Acknowledgments: We would like to thank Joseph Elliot at the University of Kansas for his assistance with English language and grammatical editing of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sanchez-Moreno, S. Biodiversity and soil health: The role of the soil food web in soil fertility and suppressiveness to soil-borne diseases. *Acta Hortic.* **2018**, *1196*, 95–104. [CrossRef]
2. Altieri, M.A.; Nicholls, C.I. Soil fertility management and insect pests: Harmonizing soil and plant health in agroecosystems. *Soil Till. Res.* **2003**, *72*, 203–211. [CrossRef]
3. Kuzyakov, Y.; Mason-Jones, K. Viruses in soil: Nano-scale undead drivers of microbial life, biogeochemical turnover and ecosystem functions. *Soil Biol. Biochem.* **2018**, *127*, 305–317. [CrossRef]
4. Clokie, M.R.; Millard, A.D.; Letarov, A.V.; Heaphy, S. Phages in nature. *Bacteriophage* **2011**, *1*, 31–45. [CrossRef] [PubMed]
5. Wilhelm, S.W.; Suttle, C.A. Viruses and Nutrient Cycles in the Sea—Viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* **1999**, *49*, 781–788. [CrossRef]
6. Enault, F.; Briet, A.; Bouteille, L.; Roux, S.; Sullivan, M.B.; Petit, M.A. Phages rarely encode antibiotic resistance genes: A cautionary tale for virome analyses. *Isme J.* **2017**, *11*, 237–247. [CrossRef]
7. Suttle, C.A. Viruses in the sea. *Nature* **2005**, *437*, 356–361. [CrossRef]
8. Davies, E.V.; James, C.E.; Williams, D.; O'Brien, S.; Fothergill, J.L.; Haldenby, S.; Paterson, S.; Winstanley, C.; Brockhurst, M.A. Temperate phages both mediate and drive adaptive evolution in pathogen biofilms. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 8266–8271. [CrossRef]
9. Williamson, K.E.; Radosevich, M.; Smith, D.W.; Wommack, K.E. Incidence of lysogeny within temperate and extreme soil environments. *Environ. Microbiol.* **2007**, *9*, 2563–2574. [CrossRef]
10. Stern, A.; Sorek, R. The phage-host arms race: Shaping the evolution of microbes. *Bioessays* **2011**, *33*, 43–51. [CrossRef]
11. Swanson, M.M.; Fraser, G.; Daniell, T.J.; Torrance, L.; Gregory, P.J.; Taliansky, M. Viruses in soils: Morphological diversity and abundance in the rhizosphere. *Ann. Appl. Biol.* **2009**, *155*, 51–60. [CrossRef]

12. Reavy, B.; Swanson, M.M.; Cock, P.J.; Dawson, L.; Freitag, T.E.; Singh, B.K.; Torrance, L.; Mushegian, A.R.; Taliansky, M. Distinct circular single-stranded DNA viruses exist in different soil types. *Appl. Environ. Microb.* **2015**, *81*, 3934–3945. [CrossRef]
13. Williamson, K.E.; Wommack, K.E.; Radosevich, M. Sampling natural viral communities from soil for culture-independent analyses. *Appl. Environ. Microb.* **2003**, *69*, 6628–6633. [CrossRef]
14. Adriaenssens, E.M.; Van Zyl, L.; De Maayer, P.; Rubagotti, E.; Rybicki, E.; Tuffin, M.; Cowan, D.A. Metagenomic analysis of the viral community in Namib Desert hypoliths. *Environ. Microbiol.* **2015**, *17*, 480–495. [CrossRef]
15. Adriaenssens, E.M.; Kramer, R.; Van Goethem, M.W.; Makhalanyane, T.P.; Hogg, I.; Cowan, D.A. Environmental drivers of viral community composition in Antarctic soils identified by viromics. *Microbiome* **2017**, *5*, 83. [CrossRef]
16. Emerson, J.B.; Roux, S.; Brum, J.R.; Bolduc, B.; Woodcroft, B.J.; Jang, H.B.; Singleton, C.M.; Soden, L.M.; Naas, A.E.; Boyd, J.A.; et al. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat. Microbiol.* **2018**, *3*, 870–880. [CrossRef]
17. Yu, D.T.; Han, L.L.; Zhang, L.M.; He, J.Z. Diversity and Distribution Characteristics of Viruses in Soils of a Marine-Terrestrial Ecotone in East China. *Microb. Ecol.* **2017**, *75*, 375–386. [CrossRef]
18. Han, L.L.; Yu, D.T.; Zhang, L.M.; Shen, J.P.; He, J.Z. Genetic and functional diversity of ubiquitous DNA viruses in selected Chinese agricultural soils. *Sci. Rep.* **2017**, *7*, 45142. [CrossRef]
19. Bowatte, S.; Newton, P.C.D.; Takahashi, R.; Kimura, M. High frequency of virus-infected bacterial cells in a sheep grazed pasture soil in New Zealand. *Soil Biol. Biochem.* **2010**, *42*, 708–712. [CrossRef]
20. Takahashi, R.; Bowatte, S.; Taki, K.; Ohashi, Y.; Asakawa, S.; Kimura, M. High frequency of phage-infected bacterial cells in a rice field soil in Japan. *Soil Sci. Plant Nutr.* **2011**, *57*, 35–39. [CrossRef]
21. Trubl, G.; Jang, H.B.; Roux, S.; Emerson, J.B.; Solonenko, N.; Vik, D.R.; Solden, L.; Ellenbogen, J.; Runyon, A.T.; Bolduc, B.; et al. Soil Viruses Are Underexplored Players in Ecosystem Carbon Processing. *Msystems* **2018**, *3*, e00076-18. [CrossRef]
22. Jin, M.; Guo, X.; Zhang, R.; Qu, W.; Gao, B.L.; Zeng, R.Y. Diversities and potential biogeochemical impacts of mangrove soil viruses. *Microbiome* **2019**, *7*, 58. [CrossRef]
23. Bi, L.; Yu, D.-T.; Du, S.; Zhang, L.-M.; Zhang, L.-Y.; Wu, C.-F.; Xiong, C.; Han, L.-L.; He, J.-Z. Diversity and potential biogeochemical impacts of viruses in bulk and rhizosphere soils. *Environ. Microbiol.* **2021**, *23*, 588–599. [CrossRef]
24. Zheng, X.X.; Jahn, M.T.; Sun, M.M.; Friman, V.P.; Balcazar, J.L.; Wang, J.F.; Shi, Y.; Gong, X.; Hu, F.; Zhu, Y.G. Organochlorine contamination enriches virus-encoded metabolism and pesticide degradation associated auxiliary genes in soil microbiomes. *Isme J.* **2022**, *16*, 1397–1408. [CrossRef]
25. Foley, J.A.; DeFries, R.; Asner, G.P.; Barford, C.; Bonan, G.; Carpenter, S.R.; Chapin, F.S.; Coe, M.T.; Daily, G.C.; Gibbs, H.K.; et al. Global consequences of land use. *Science* **2005**, *309*, 570–574. [CrossRef]
26. Lacerda, G.V.; Noronha, M.F.; Cabral, L.; Delforno, T.P.; de Sousa, S.T.P.; Fernandes, P.I.; Melo, T.S.; Oliveira, V.M. Land Use and Seasonal Effects on the Soil Microbiome of a Brazilian Dry Forest. *Front. Microbiol.* **2019**, *10*, 648. [CrossRef]
27. Chen, S.F.; Zhou, Y.Q.; Chen, Y.R.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, 884–890. [CrossRef]
28. Nurk, S.; Meleshko, D.; Korobeynikov, A.; Pevzner, P.A. metaSPAdes: A new versatile metagenomic assembler. *Genome Res.* **2017**, *27*, 824–834. [CrossRef]
29. O'Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufo, S.; McVeigh, D.H.R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; Astashyn, A.; et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [CrossRef]
30. Arndt, D.; Grant, J.R.; Marcu, A.; Sajed, T.; Pon, A.; Liang, Y.J.; Wishart, D.S. PHASTER: A better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* **2016**, *44*, W16–W21. [CrossRef]
31. Shkoporov, A.N.; Clooney, A.G.; Sutton, T.D.S.; Ryan, F.J.; Daly, K.M.; Nolan, J.A.; McDonnell, S.A.; Khokhlova, E.V.; Draper, L.A.; Forde, A.; et al. The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. *Cell Host Microbe* **2019**, *26*, 527–541.e5. [CrossRef]
32. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef]
33. Ogunseitan, O.A.; Sayler, G.S.; Miller, R.V. Application of DNA probes to analysis of bacteriophage distribution patterns in the environment. *Appl. Environ. Microbiol.* **1992**, *58*, 2046–2052. [CrossRef]
34. Quaiser, A.; Dufresne, A.; Ballaud, F.; Roux, S.; Zivanovic, Y.; Colombet, J.; Sime-Ngando, T.; Francez, A.J. Diversity and comparative genomics of Microviridae in Sphagnum- dominated peatlands. *Front. Microbiol.* **2015**, *6*, 375. [CrossRef]
35. Jansen, R.; Embden, J.D.; Gaastra, W.; Schouls, L.M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **2002**, *43*, 1565–1575. [CrossRef]
36. Terns, M.P.; Terns, R.M. CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* **2011**, *14*, 321–327. [CrossRef]
37. McDougald, D.; Rice, S.A.; Barraud, N.; Steinberg, P.D.; Kjelleberg, S. Should we stay or should we go: Mechanisms and ecological consequences for biofilm dispersal. *Nat. Rev. Microbiol.* **2011**, *10*, 39–50. [CrossRef]
38. Bardgett, R.D.; van der Putten, W.H. Belowground biodiversity and ecosystem functioning. *Nature* **2014**, *27*, 505–511. [CrossRef]
39. Barrios, E. Soil biota, ecosystem services and land productivity. *Ecol. Econ.* **2007**, *64*, 269–285. [CrossRef]
40. Costerton, J.W.; Cheng, K.; Geesey, G.G.; Ladd, T.I.; Nickel, J.C.; Dasgupta, M.; Marrie, T.J. Bacterial biofilms in nature and disease. *Annu. Rev. Microbiol.* **1987**, *41*, 435–464. [CrossRef]

41. Vidakovic, L.; Singh, P.K.; Hartmann, R.; Nadell, C.D.; Drescher, K. Biofilm architecture confers individual and collective mechanisms of viral protection. *Nat. Microbiol.* **2018**, *3*, 26–31. [CrossRef]
42. Zablocki, O.; Adriaenssens, E.M.; Cowan, D. Diversity and Ecology of Viruses in Hyperarid Desert Soils. *Appl. Environ. Microb.* **2016**, *82*, 770–777. [CrossRef]
43. Ding, G.C.; Piceno, Y.M.; Heuer, H.; Weinert, N.; Dohrmann, A.B.; Carrillo, A.; Andersen, G.I.; Castellanos, T.; Tebbe, C.C.; Smalla, K. Changes of soil bacterial diversity as a consequence of agricultural land use in a semi-arid ecosystem. *PLoS ONE* **2013**, *8*, e59498. [CrossRef]
44. Wu, Y.C.; Cai, P.; Jiang, X.X.; Niu, X.K.; Ji, D.D.; Ashry, N.M.; Gao, C.H.; Huang, Q.Y. Soil biofilm formation enhances microbial community diversity and metabolic activity. *Environ. Int.* **2019**, *132*, 105116. [CrossRef]
45. Campbell, A. Conditions for the existence of bacteriophage. *Evolution* **1961**, *15*, 153–165. [CrossRef]
46. Brussow, H.; Canchaya, C.; Hardt, W.D. Phages and the evolution of bacterial pathogens: From genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 560–602. [CrossRef]
47. Lindell, D.; Jaffe, J.D.; Johnson, Z.I.; Church, G.M.; Chisholm, S.W. Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* **2005**, *438*, 86–89. [CrossRef]
48. Lindell, D.; Jaffe, J.D.; Coleman, M.L.; Futschik, M.E.; Axmann, I.M.; Rector, T.; Kettler, G.; Sullivan, M.B.; Steen, R.; Hess, W.R.; et al. Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* **2007**, *449*, 83–86. [CrossRef]
49. Anderson, C.L.; Sullivan, M.B.; Fernando, S.C. Dietary energy drives the dynamic response of bovine rumen viral communities. *Microbiome* **2017**, *5*, 155. [CrossRef]