



**HAL**  
open science

# Soil pH, Calcium Content and Bacteria as Major Factors Responsible for the Distribution of the Known Fraction of the DNA Bacteriophage Populations in Soils of Luxembourg

Perrine Florent, Henry-Michel Cauchie, Malte Herold, Stéphan Jacquet, Leslie Ogorzaly

## ► To cite this version:

Perrine Florent, Henry-Michel Cauchie, Malte Herold, Stéphan Jacquet, Leslie Ogorzaly. Soil pH, Calcium Content and Bacteria as Major Factors Responsible for the Distribution of the Known Fraction of the DNA Bacteriophage Populations in Soils of Luxembourg. *Microorganisms*, 2022, 10 (7), pp.1458. 10.3390/microorganisms10071458 . hal-04017700

**HAL Id: hal-04017700**

**<https://hal.science/hal-04017700v1>**

Submitted on 7 Mar 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



## Article

# Soil pH, Calcium Content and Bacteria as Major Factors Responsible for the Distribution of the Known Fraction of the DNA Bacteriophage Populations in Soils of Luxembourg

Perrine Florent <sup>1,2</sup> , Henry-Michel Cauchie <sup>1</sup> , Malte Herold <sup>1,2</sup> , Stéphan Jacquet <sup>3</sup> and Leslie Ogorzaly <sup>1,\*</sup>

<sup>1</sup> Environmental Research and Innovation Department (ERIN), Luxembourg Institute of Science and Technology (LIST), 4422 Belvaux, Luxembourg; perrine.florent@list.lu (P.F.); henry-michel.cauchie@list.lu (H.-M.C.); malte.herold@list.lu (M.H.)

<sup>2</sup> Faculté des Sciences, de la Technologie et de la Communication (FSTC), Doctoral School in Science and Engineering (DSSE), University of Luxembourg, 4365 Esch-sur-Alzette, Luxembourg

<sup>3</sup> INRAE, UMR CARRTEL, Université Savoie Mont Blanc, 74200 Thonon-les-Bains, France; stephan.jacquet@inrae.fr

\* Correspondence: leslie.ogorzaly@list.lu; Tel.: +352-2758885-5069

**Abstract:** Bacteriophages participate in soil life by influencing bacterial community structure and function, biogeochemical cycling and horizontal gene transfer. Despite their great abundance, diversity, and importance in microbial processes, they remain little explored in environmental studies. The influence of abiotic factors on the persistence of bacteriophages is now recognized; however, it has been mainly studied under experimental conditions. This study aimed to determine whether the abiotic factors well-known to influence bacteriophage persistence also control the natural distribution of the known DNA bacteriophage populations. To this end, soil from eight study sites including forests and grasslands located in the Attert River basin (Grand Duchy of Luxembourg) were sampled, covering different soil and land cover characteristics. Shotgun metagenomics, reference-based bioinformatics and statistical analyses allowed characterising the diversity of known DNA bacteriophage and bacterial communities. After combining soil properties with the identified DNA bacteriophage populations, our *in-situ* study highlighted the influence of pH and calcium cations on the diversity of the known fraction of the soil DNA bacteriophages. More interestingly, significant relationships were established between bacteriophage and bacterial populations. This study provides new insights into the importance of abiotic and biotic factors in the distribution of DNA bacteriophages and the natural ecology of terrestrial bacteriophages.

**Keywords:** DNA bacteriophages; soil; metagenomics; bacteria; soil properties



**Citation:** Florent, P.; Cauchie, H.-M.; Herold, M.; Jacquet, S.; Ogorzaly, L. Soil pH, Calcium Content and Bacteria as Major Factors Responsible for the Distribution of the Known Fraction of the DNA Bacteriophage Populations in Soils of Luxembourg. *Microorganisms* **2022**, *10*, 1458. <https://doi.org/10.3390/microorganisms10071458>

Academic Editors: Huixin Li, Li Xu and Hui Cao

Received: 23 June 2022

Accepted: 16 July 2022

Published: 19 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Soil is a complex compartment where biodiversity and interactions between microbes, animals and plants contribute to its life and optimal functioning. Generally, plants provide organic matter as well as nutrients to soil (e.g., nitrogen, carbon, phosphorous) and are responsible for its structure through the development of the rhizosphere [1,2]. Soil-dwelling animals help to degrade and mix organic matter throughout the soil and participate in the recycling of nutrients from dead biomass [3,4]. Microbes, including fungi, archaea and bacteria, play a key role in organic matter decomposition, element cycling and nutrient availability to plants [5–7]. This microbial activity, particularly due to the quantity, diversity and functional roles of prokaryotes, is considered central to soil processes [8–10]. Viruses are also biological entities fully part of soil biodiversity with essential but relatively unexplored functions [11].

Since the beginning of research on viral diversity in soils in the 1970s [11,12], knowledge of soil viruses has been expanded well, especially for bacteriophages (bacteria-infecting viruses), whose abundance can reach up to  $10^{10}$  particles per gram of soil [13,14].

Bacteriophages are ubiquitous in the environment and can be found in agricultural soils, forest soils (approximately  $10^9$  bacteriophages per gram of dry soil), wetlands, or even cold deserts (approximately  $10^3$  bacteriophages per gram of dry soil) [13]. Bacteriophages seem to be more abundant in wetlands than in forests due to high soil water content, but less abundant in agricultural soils than in forests where the soil temperature is lower [13]. In fact, many abiotic factors may be involved in the distribution, abundance and diversity of bacteriophages in soils. However, most of them have only been explored to investigate their effect on bacteriophage persistence. Temperature is recognised as an important, if not the main, factor in viral inactivation [15]. The ideal temperature range to prevent viral inactivation is considered to be 4 to 10 °C, where bacteriophages are generally stable at pH values between 5 and 9, but can be degraded under strongly alkaline (pH > 11) or acidic (pH < 2) conditions [16–18]. These extreme pH conditions can inactivate virus particles through the denaturation of their proteins or nucleic acids [18–21]. In addition, soil texture, especially clay type, alters the detection of bacteriophage as it strongly influences adsorption and retention phenomena in the soil. Indeed, clay soils present a very dense structure with a low permeability and a large specific surface area, which is auspicious for higher adsorption–desorption reactions with bacteriophages [22]. In addition, due to hydrophobicity interactions and the presence of electrical charges, bacteriophages tend to adsorb more onto clay surfaces than onto sandy or silty soil particles [23]. This process also helps to protect bacteriophages from other inactivating factors [11,24].

However, knowledge of the influence of these abiotic factors on bacteriophage persistence has been mainly gained under experimental conditions using bacteriophage models (e.g.,  $\Phi$ X174, MS2), which may not be transferable when studying environmental samples [19,21,25,26]. Indeed, the natural bacteriophage diversity and the impact of such abiotic factors remain largely underexplored in terrestrial compared to aquatic environments [27–29]. Regarding the studies exploring the diversity of bacteriophages on in situ terrestrial ecosystems, virus-like particles (VLPs) or viral operational taxonomic units (vOTUs) have dominantly been studied [30–33]. While the former does not allow the identification of viruses, the latter is rather a proxy for viral species. Therefore, the use of metagenomics in determining the taxonomical classification of viruses and their respective diversity would not only allow completing the knowledge about the influence of abiotic factors on the bacteriophage species communities but also explore the relationship with the bacterial communities.

Coevolution between bacteriophages and their bacterial hosts is widely recognised, with reciprocal influences [11,34]. Bacteriophages are unable to self-replicate and therefore must infect one or more bacterial hosts to replicate and spread. As agents of bacterial mortality, bacteriophages are likely to have a significant influence on bacterial community dynamics, structure and function, as well as biogeochemical cycling, horizontal gene transfer processes and the pre-decomposition of the organic matter [35,36]. However, viral-mediated mortality induces interest at the expense of understanding the influence of bacteria on bacteriophage occurrence in soils [37–39].

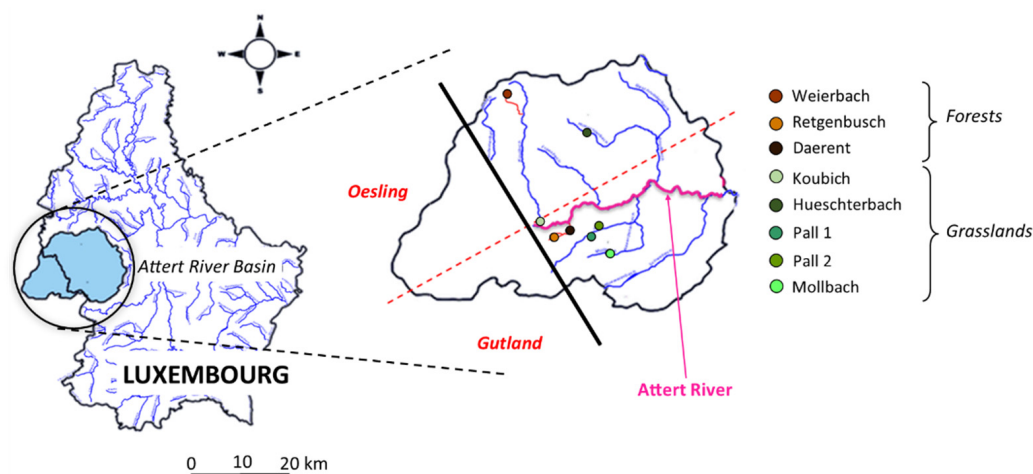
With the advent of viral metagenomics, databases increasingly include new viral sequences, which participate in the rapid expansion of our understanding of the virosphere. Based on the viruses already identified, the influence of selected abiotic factors (i.e., soil properties, vegetation and land use) on the distribution of the known DNA bacteriophages can be assessed in different types of soil. Therefore, in this study, we used a combination of approaches and methods (i.e., physicochemical analysis, metagenomics, reference-based bioinformatics, direct counting and statistics) to obtain an original picture of the natural viral and bacterial diversities and their ecological interactions with soil properties in terrestrial environments of Luxembourg. The aim of the study was to determine whether the abiotic factors well known to influence bacteriophage persistence also control the natural distribution of the known DNA bacteriophage populations in the studied soil samples, and to analyse the role of bacterial populations in this phenomenon. Two main hypotheses were tested: (i) abiotic factors (i.e., pH, cations) well known to affect bacteriophage survival also

have an impact on their diversity and distribution in terrestrial environments, and (ii) the bacterial community present in the soil influences the relationship between bacteriophages and soil properties.

## 2. Materials and Methods

### 2.1. Studied Area

The Attert River basin area reaches 310 km<sup>2</sup> and covers a wide range of bedrock geologies and land use [40], ultimately being representative of Luxembourg's main physiographic characteristics. Its northern boundary lies on schist bedrock (belonging to the Ardennes massif; Oesling region), while its southern part exhibits alternating sedimentary layers of marls and sandstone (representative of the eastern limit of the Paris basin; Gutland region). The Attert River originates near Nobressart (Belgium) and extends over 38 km from west to east, joining the Alzette River basin in Bissen, Luxembourg. The study is focused on the Luxembourg side of the Attert River basin (Figure 1, Table 1) where 8 sub-basins have been selected compiling 3 forested sites and 5 grassland sites. All of them are located near one of the tributaries of the Attert River.



**Figure 1.** Map of the Attert River basin with the location of the eight sampling sites (modified from Geoportail.lu). Brown colours: forest sites; green colours: grasslands. With an oversimplified scheme, the plain black line represents the border between the Luxembourg and Belgian sides of the basin while the dashed red line is the limit between Oesling and Gutland regions. The Attert River is highlighted in pink, the affluents of the Attert River are in blue and the affluents of affluents, in orange.

**Table 1.** List of the samples with their GPS coordinates.

Sample Sites	Sample ID	Vegetation	GPS Coordinates
Weierbach	WPF12092019	Forest	(N) 49.831937 (E) 5.80341
Retgenbusch	RPF12092019	Forest	(N) 49.743436 (E) 5.834341
Daerent	DPF12092019	Forest	(N) 49.748226 (E) 5.856064
Koubich	KPF12092019	Grassland	(N) 49.751942 (E) 5.826879
Hueschterbach	HPF12092019	Grassland	(N) 49.809914 (E) 5.863031
Pall 1	P1PF12092019	Grassland	(N) 49.74617 (E) 5.871577
Pall 2	P2PF12092019	Grassland	(N) 49.748935 (E) 5.881184
Mollbach	MPF12092019	Grassland	(N) 49.7347405 (E) 5.888961

The first forest site is the Weierbach sub-basin located in the northwest of the Attert River basin, in the Oesling region. The two other forest sub-basins are Retgenbusch and Daerent, located on the same tributary of the Attert River basin in the southwestern part of the river (Gutland region). All forest sites have distinct vegetation types; Weierbach is a beech forest, while Daerent is an oak forest and Retgenbusch is a spruce forest.

The first grassland site, Hueschterbach, is the only sub-basin located in the Oesling region and is characterised by schist bedrock. The four other grassland sites are in the Gutland region: Mollbach a sub-basin located in the southwest of the Attert River basin; Pall located in the north of the Mollbach sub-basin, for which two distinct sites (Pall 1 and Pall 2) were chosen on the tributary of the same name; the sub-basin Koulbich located in the northwest of the Attert River basin is known for pasture activities of barley and wheat cultures.

## 2.2. Sample Collection

The field campaign was conducted on 12 September 2019. The topsoil was collected for all selected sites cited above. Five to seven replicates were randomly sampled within a delimited surface of 25 m<sup>2</sup>, with a soil auger (3 cm in diameter) on the first 20 cm of depth. To exclude the intra-site spatial variability and to ease laborious shotgun metagenomic and bioinformatic analyses, 5 to 7 soil replicates were pooled altogether to set up one sample per site, representing around 300 g per site [41–43]. All soil samples were split into two parts, either for physicochemical analyses or stored at −20 °C for molecular analyses.

## 2.3. Soil Characterisation

The characterisation of the soil properties includes first the pH measurements made on 10 g of soil according to the ISO 10,390 [44] for pH<sub>H2O</sub> and pH<sub>KCl</sub> and to VDLUFA A5.1.1 [45] for pH<sub>CaCl2</sub>. Then, the nutrients dosing of phosphorus (P<sub>2</sub>O<sub>5</sub>) and potassium (K<sub>2</sub>O) was performed according to the VDLUFA A6.2.1.1 [46] using 5 g of soil with 100 mL of Ca-acetate-lactate at pH = 4.1. On the other hand, the dosing of magnesium (Mg) and sodium (Na) was conducted following the VDLUFA A6.2.1.7 [47], using 10 g of soil with 100 mL of CaCl<sub>2</sub> (0.01 M). In addition, the carbon-to-nitrogen ratio (C/N) was calculated after quantifying the total organic carbon (C<sub>org</sub>) and the total nitrogen (N<sub>tot</sub>) both on 300 g of soil, according to ISO 10,694 [48] and ISO 13,878 [49], respectively. The cation exchangeable capacity (CEC) and cation saturation rate (S/T) were analysed according to the ISO 23,470 [50] on 2.5 g of soil. From the extract CoHex, calcium (Ca<sup>2+</sup>), magnesium (Mg<sup>2+</sup>), potassium (K<sup>+</sup>) and sodium (Na<sup>+</sup>) CoHex were analysed by ICP-OES (inductively coupled plasma–optical emission spectrometry). Finally, the soil granulometry was determined by a sieving method on 10 g of soil mixed with 100 mL of H<sub>2</sub>O<sub>2</sub> and 10 mL HCl [51].

## 2.4. Viral Quantification

### 2.4.1. Viral Elution and Concentration

After the optimisation of the elution and concentration protocol (Supplementary Materials, Section S1), 5 g of soil, previously stored at −20 °C, was placed at room temperature and then mixed with 15 mL of the eluent solution, i.e., 3% beef extract and 0.05 M glycine (at pH = 9.5) and shaken vigorously using an orbital shaker for 30 min at 250 rpm (rotation per minute) at 4 °C. The samples were then centrifuged (Eppendorf 5801R) at 5000 × *g* for 15 min on a fixed rotor at 4 °C. The supernatant was collected and underwent a 0.22 µm Millex<sup>®</sup> sterile filter unit with a Millipore Express<sup>®</sup> PES membrane (SLGP033RS, Merck Millipore Ltd., Tullagreen, Ireland), followed by ultracentrifugal filtration using an Amicon Ultra-15 50 kD (UFC905024, Merck, Darmstadt, Germany), upon the recommendations of the manufacturer. After centrifugation at 5000 × *g* for 15 min on a fixed rotor at 4 °C, the concentrate solutions, with a volume varying from 0.3 to 3 mL depending on the site, were finally transferred into 2 mL cryotubes and stored at 4 °C.

### 2.4.2. Viral-like Particles (VLPs) Counting

From the viral concentrates obtained after the viral elution and concentration steps, the abundances of VLPs were determined by flow cytometry (i.e., a benchtop FACSCalibur), after the samples were divided into 3 technical replicates. Briefly, after homogenisation by centrifugation, the viral sample was fixed using glutaraldehyde (0.25% final concentration) and mixed with autoclaved and <0.02 µm filtered TE buffer (0.1 mM Tris and 1 mM EDTA,

pH 8), SYBR® Green I (at a final concentration of  $5 \times 10^{-5}$ , Molecular Probes, Eugene, OR, USA). The samples underwent a 10 min heat treatment at 70 °C before the Flow Cytometry (FCM) analysis [52].

## 2.5. Identification of the Microbial Populations

### 2.5.1. Total DNA Extraction and Shotgun Metagenomics Sequencing

After a comparison of different protocols (Supplementary Materials, Section S2), DNA extraction was performed using the DNeasy® PowerSoil® Kit (Qiagen, Hilden, Germany) from 0.25 g of frozen soil (−20 °C) for each site, according to the manufacturer's recommendations. The DNA elution step was performed in 60 µL instead of 100 µL to increase the DNA concentration. The final DNA solutions were stored at −20 °C until further analyses.

From the total extracted DNA, library preparations were completed using the Nextera DNA Flex Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Paired-end sequence reads were generated using the Illumina NovaSeq 600 (2 × 150 bp).

### 2.5.2. Reference-Based Bioinformatics

The raw sequencing reads were quality controlled and Illumina adapters were trimmed with Trimmomatic 0.39 [53]. Assemblies were performed separately for each sample using SPAdes v3.15.4 [54] with the “meta” option and default parameters. Taxonomic classification of contigs was carried out against the complete non-redundant (nr) microbial protein sequence database (released version of April 2022) from NCBI using MMseqs2 v13.45111 [55] and the easy-taxonomy workflow that determines ancestors by searching translated ORFs (translation Table 11) [56]. Quality-filtered reads from each metagenome were mapped back to the assembled contigs using minimap2 v2.24 with the “sr” preset [57] and sorted and indexed with SAMtools v1.15.1 [58]. Finally, metagenomic profiles were exported as comma-separated value (CSV) files. Metagenomic abundances were calculated based on the number of reads mapping to each contig. Mapped reads were divided by contig length and the total number of mapped reads per sample [59,60].

## 2.6. Statistical Analyses

Statistical analyses were performed with RStudio, version 1.4.1106, and R 4.0.4 [61]. Figures were generated using the package ‘ggplot2’. Regarding the ecological diversity, both  $\alpha$  and  $\beta$ -diversities of bacteria and DNA bacteriophages were studied. The  $\alpha$ -diversity is the average number of distinct species groups at a local scale, considering the richness of a population [62]. As popular  $\alpha$ -diversity indexes, the Shannon index (H) and the species richness (S) were calculated for each site, using the R package ‘vegan’ [63]. Then, the evenness (E), which defines the distribution of the abundance of the species in a community, was calculated as follows [64,65]:

$$E = \frac{H}{\log(S)}$$

The  $\beta$ -diversity is all changes in the species composition of a population between several habitats in a region. This diversity considers both the richness and evenness of populations [66]. Using Bray–Curtis dissimilarities, a multivariate method, the RDA (redundancy analysis), was performed to compare the different sites as a function of the bacteriophage populations at the genus level, through the analysis of their variation from one another. To build the RDA, a selection of the 10 most abundant genera was first performed, as rare genera might generate noise in the analysis). Then, abiotic factors (i.e., sand, clay, silt content, soil pH, magnesium concentration, carbon-to-nitrogen ratio and  $\text{Ca}^{2+}$  concentration) were selected by removing the variables for which strong correlations ( $>r = 0.80$ ) were found to avoid co-variations, using the function ‘pairs’. In addition, all abiotic factors that significantly constructed the RDA were kept (i.e., pH and  $\text{Ca}^{2+}$ ). The determination of the explanatory variables for the RDA was conducted using ‘vegan’ and

correlations of the abiotic factors with each axis were defined using the package ‘psych’. Permutational multivariate analysis of variance (PERMANOVA) was conducted to establish the impact of some abiotic factors on the distribution of the DNA bacteriophages. From the calculation of the Bray–Curtis dissimilarities to the PERMANOVA and RDA, the package ‘vegan’ was used to conduct these analyses. Finally, a co-correspondence analysis (CoCA) was performed to assess the co-variation of the distribution of the total bacterial and bacteriophage genera obtained from the metagenomic analyses, among the different sites. This analysis is based on contingency table data and highlights co-similarities between two groups of variables (here, bacterial species data and DNA bacteriophage genera data). CoCA was performed using the package ‘coca’, combined with a Mantel test using Pearson correlations.

### 3. Results

#### 3.1. Soil Description

The granulometric analysis of the soil revealed similar textures between the sites (Table 2). Indeed, selected sites are mainly silty, especially Mollbach and Retgenbusch, for which the silt content is around 52–58%. However, Weierbach and Pall 1 are clay loam with approximately 42% silt, while Daerent and Huechterbach are sandy loam with 53% sand. Concerning Pall 2, the proportion for the different textures are rather equal (28% clay, 38% silt, 34% sand). Koulbich is a loam soil, with a similar proportion of silt and sand ( $\approx 40\%$ ) but little clay (19%). The gravimetric water content (GWC) was the highest for Weierbach (24.6%) and Pall 1 (20.9%) and the lowest for Koulbich (10.8%).

The physical-chemical properties of soils are reported in Table 2. Within the forest sites, the soils of Weierbach and Retgenbusch are highly acidic (respectively, pH = 4.1 and 4.5) and their respective cation exchange capacity and calcium concentration are quite low (CEC of Retgenbusch = 2.6 milliequivalent per 100 g (meq per 100 g) and  $\text{Ca}^{2+}$  = 1.6 meq per 100 g, CEC of Weierbach = 8.5 meq per 100 g and  $\text{Ca}^{2+}$  = 2.1 meq per 100 g). Their nutrient content was also relatively poor, especially for phosphorus ( $\text{P}_2\text{O}_5$ ) and potassium ( $\text{K}_2\text{O}$ ), suggesting that these soils are not very fertile. In addition, the cation saturation rate (S/T) of Retgenbusch is superior to 85% ( $\text{S}/\text{T} = 101\%$ ), which can explain the high calcium content in the soil, while the soil of Weierbach shows a regular calcium intake ( $20\% \leq \text{S}/\text{T} = 42\% \leq 85\%$ ). The soil of Weierbach is distinguished by a slow organic matter decomposition (carbon-to-nitrogen ratio or C/N = 18 and high  $\text{C}_{\text{org}}$  content), displayed by a need for nitrogen to ensure an efficient organic decomposition. The soil of Daerent is fertile (CEC = 17.1 meq per 100 g, pH = 6.3,  $\text{K}_2\text{O}$  = 21 mg per 100 g of dry soil) and has a maximum decomposition rate (C/N = 10). Regarding the grassland sites, Pall 1, Pall 2 and Mollbach soils are all fertile (CEC = 23.7, 17.2, 14.3 meq per 100 g, respectively) with a high concentration of calcium ( $\text{Ca}^{2+}$  = 16.6, 15.1 and 13.1 meq per 100 g), while Koulbich and Hueschterbach are not quite (CEC = 7.4 and 8.5 meq/100 g, respectively) with lower calcium concentrations ( $\text{Ca}^{2+}$  = 5.8 and 8.7 meq per 100 g). However, all of them are saturated in calcium, with a cation saturation rate superior to 100%. Finally, the grassland soils have a rapid decomposition (C/N  $\approx 8.5$ ) with a maximum organic decomposition for Pall 1 (C/N = 9.8) and have a favourable pH (pH  $\approx 6$ –6.5) and mineral nutrient for plant growth.

**Table 2.** Physical-chemical properties of the eight studied soils: GWC: gravimetric water content, C<sub>org</sub>: organic carbon, N<sub>tot</sub>: total nitrogen, CEC: cation exchangeable capacity, S/T: cation saturation rate, Ca<sup>2+</sup>: calcium concentration.

Site Names	GWC	pH	C/N	C <sub>org</sub>	N <sub>tot</sub>	Mg	Na	P <sub>2</sub> O <sub>5</sub>	K <sub>2</sub> O	CEC	Ca <sup>2+</sup>	S/T	Textural Class	Clay	Silt	Sand
	%	-	-	%	%	mg per 100 g ds	mg per 100 g ds	mg per 100 g ds	mg per 100 g ds	Meq kg <sup>-1</sup>	Meq kg <sup>-1</sup>	%		%	%	%
Weierbach	24.6	4.1	18	9.7	0.54	3	2	5	13	85	21	42	CL	23.9	42.2	33.9
Retgenbusch	13.7	4.5	8.6	1.9	0.22	3	1	3	9	26	16	101	A	27.2	58.5	14.3
Daerent	15.8	6.3	10	3.9	0.39	46	2	4	21	171	126	107	SL	53.9	28.0	18.1
Koulbich	10.8	5.7	8.3	1.9	0.23	12	3	9	11	74	58	105	L	37.7	43.6	18.6
Hueschterbach	14.7	6.4	8.8	2.2	0.25	11	3	20	16	85	87	126	L	52.0	29.6	18.4
Pall 1	20.9	6.8	9.8	5.1	0.52	17	2	6	17	237	166	100	CL	24.2	42.6	33.2
Pall 2	16.3	6.2	8.8	3.6	0.41	31	3	7	15	172	151	115	CL	33.9	38.1	28.1
Mollbach	18.3	6.0	8.3	3.1	0.23	52	5	13	21	143	131	112	SL	28.1	52.5	19.4

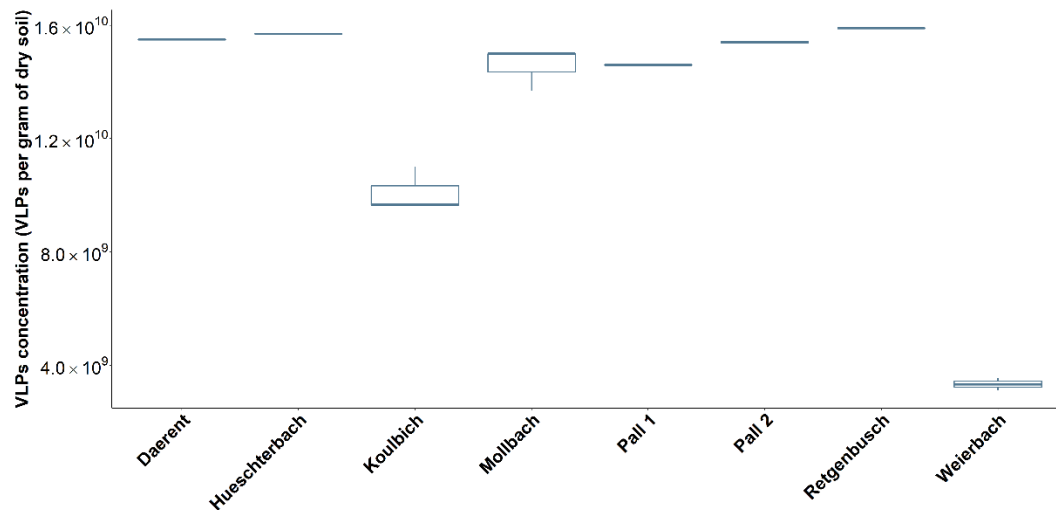
ds = dry soil, meq kg: milliequivalent per kilogram (quantity of positive charges per kilogram); textural class: CL = clay loam; L = silt loam; SL = sandy loam; L = loam [67].



### 3.2. Description of the DNA-Bacteriophage Populations

#### 3.2.1. VLPs Enumeration

The average concentration of VLPs found in every soil was about  $10^{10}$  VLPs per gram of dry soil (Figure 2). The highest concentrations were recorded for the soils of Retgenbusch, Hueschterbach, Daerent and Pall 2 with approximately  $1.5 \times 10^{10}$  VLPs per gram of dry soil. The lowest concentration was for the soil of Weierbach with  $3.3 \times 10^9$  VLPs per gram of dry soil. and displayed a 0.48 log difference with the lowest VLPs concentration (after the Weierbach), Koulbich and 0.67 log with Hueschterbach, the highest VLPs concentration. Finally, Koulbich showed on average a lower concentration of 0.18 log compared to all the other sites.



**Figure 2.** Boxplot of the virus-like particles (VLPs) concentration per gram of dry soil for all studied sites.

#### 3.2.2. Taxonomic Classification

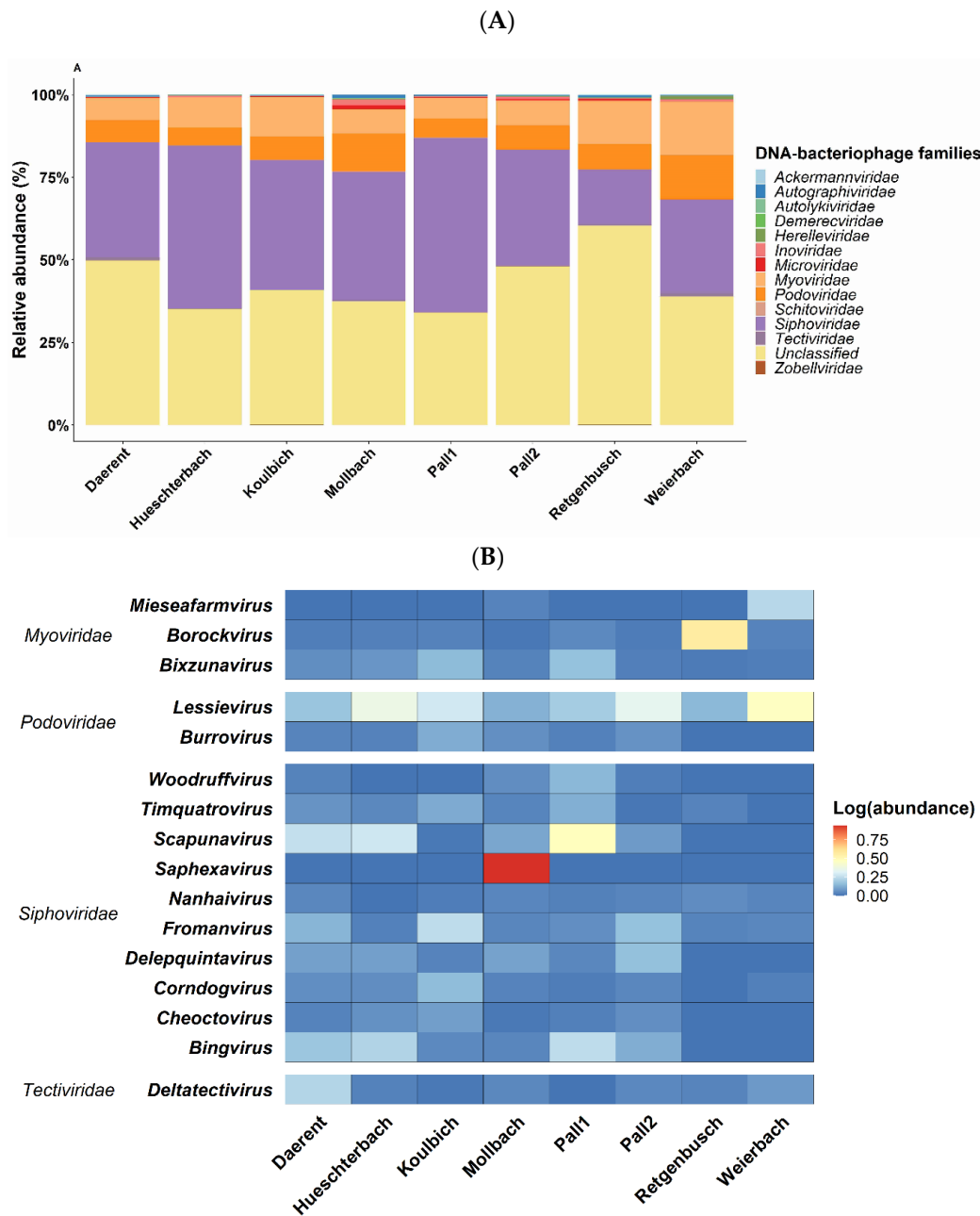
The statistics of the shotgun metagenomic sequencing are reported in Table 3.

The shotgun sequencing produced between 40 to 57 million paired reads per sample generating an average of 6.2 million contigs after assembly. From 75 to 99% of the total raw reads mapped back to the contigs. Of these mapped reads, between 0.01% and 0.04% were virus-affiliated reads (Supplementary Table S1), corresponding to a total of 37 known viral families for all sites included. Among these families, 13 DNA bacteriophage families were identified and represented between 89% and 96% of all virus-affiliated reads (Figure 3A). The DNA bacteriophage composition was strongly characterised by the *Caudovirales* order with between 75% and 93% of the total metagenomics identification. This high proportion was mainly due to the high abundance of *Siphoviridae* (13.9% to 49%), *Myoviridae* (5% to 12%) and *Podoviridae* (4% to 10%). *Caudovirales* gathers all the identified DNA bacteriophage families, except for *Autolykiviridae* (unclassified order) in Daerent and Mollbach soils, *Microviridae* (Petitvirales) and *Tectiviridae* (order Kalamavirales) both present in every site. By investigating each site, *Siphoviridae* was found to always be the most identified family, especially in the soils of Pall 1, Hueschterbach, Mollbach, Koulbich and Pall 2 (with, respectively, 49.58%, 44.83%, 35.15% and 34.80%) while its abundance was the lowest for Retgenbusch (13.9%). Regarding *Myoviridae* and *Podoviridae*, the former was highly abundant in Koulbich (10.65%), Retgenbusch (10.9%) and Weierbach (12.3%) while the latter was rather highly present in the soils of Weierbach (10.2%) and Mollbach (10.3%). It is noteworthy that most of the assigned DNA bacteriophages from the libraries belonged to the unclassified *Caudovirales* (between 29% and 51%). Of the two ssDNA bacteriophage families, i.e., *Inoviridae* and *Microviridae*, both were detected by the metagenomics with a low abundance in all sites (between 0.09% and 1.6% for the former and between 0.06% and 1% for the latter), while all the previously mentioned families are dsDNA bacteriophages.

**Table 3.** Statistics of the metagenomic sequencing: raw data, assembly and mapping on the eight studied soils.

Sites	Number of Read Pairs	Yield in Mbp	Average Quality (Q-Score)	Number of Total Mapped Read Pairs	Percentage of Mapped Read Pairs	Number of Total Contigs	GC Content	N50
Daerent	45,506,199	12,929	35.6	36,248,481	79.66%	5,469,310	63.70%	271
Hueschterbach	42,222,600	13,046	35.88	36,987,700	87.60%	5,727,441	63.95%	269
Koulbich	64,930,676	18,587	35.53	49,069,306	75.57%	7,792,136	63.18%	268
Mollbach	50,983,302	14,522	35.63	35,064,164	68.77%	6,306,081	63.40%	264
Pall 1	46,905,743	13,397	35.62	40,072,879	85.43%	5,980,767	64.72%	273
Pall 2	56,510,613	16,136	35.6	48,535,078	85.89%	6,883,923	63.40%	272
Retgenbusch	49,316,386	13,878	35.59	46,692,284	94.68%	5,677,103	60.01%	295
Weierbach	57,936,000	16,650	35.22	57,518,714	99.28%	5,898,678	60.95%	323

Mbp = millions of base pairs; GC = guanine cytosine; N50 = the sequence length of the shortest contig at 50% of the total genome length; total references to total microbial metagenomes (i.e., eukaryote, archaea, bacteria and viruses).



**Figure 3.** (A) Bar charts of the relative abundance (in %) of DNA bacteriophage families from the known fraction of all studied sites. (B) Heatmap of the most occurring DNA bacteriophage genera identified (selection of genera with a log-abundance above 0.24) as a function of the sites. The gradient represents the log-transformed abundance, where red colour represents the most abundant genera and blue, the least ones.

To visualise the composition of the DNA bacteriophage population, a heatmap on the log-transformation abundances, at the genus level, was performed on the 10 most abundant genera, all sites included (Figure 3B). Genus *Lessievirus* (*Podoviridae*) was detected in every site, and more specifically in the soil of Weierbach, for which *Mieseafarmvirus* (*Myoviridae*) was also present. The soil of Mollbach was specifically represented by *Saphexavirus* (*Siphoviridae*), while in Retgenbusch, *Borockvirus* (*Myoviridae*) was the most abundant genus. Pall 1 was found to be mainly represented by *Scapunavirus* (*Siphoviridae*) followed by *Bingvirus* (*Siphoviridae*). *Scapunavirus* was also found in a slightly higher abundance in Daer-

ent and Hueschterbach. Koulbich and Pall 2 did not display high abundances for specific genera, but still showed the presence of *Fromanvirus* (*Siphoviridae*) besides *Lessievirus*.

### 3.3. Impact of the Soil Properties on Known DNA-Bacteriophage Distribution

#### 3.3.1. $\alpha$ -Diversity: Known DNA-Bacteriophage Diversity within Sites

The DNA bacteriophage  $\alpha$ -diversity of the known microbial fraction of each soil was investigated using the Shannon index, species evenness and species richness (Table 4). All sites displayed similar diversity (around  $H = 2$ ) where Retgenbusch was the least diverse soil ( $H = 2.31$ ) and Weierbach the most one ( $H = 2.83$ ). Indeed, the soil of Weierbach presented the lowest number of DNA bacteriophage species ( $S = 91$ ) and the highest evenness ( $E = 0.63$ ). A one-way analysis of variance (one-way ANOVA) was performed on the Shannon index to analyse the hypothetical impact of the selected soil properties on the DNA bacteriophage diversity within the sites. As a result, none of the soil properties showed a significant effect on the DNA bacteriophage alpha-diversity.

**Table 4.** Shannon diversity index (H), evenness (E) and richness (S) on genus level of DNA bacteriophage from the known fraction of the population from the soil of the studied sites.

Sites	Shannon Index (H)	Evenness (E)	Species Richness (S)
Daerent	2.39	0.48	147
Hueschterbach	2.54	0.51	140
Koulbich	2.58	0.51	149
Mollbach	2.79	0.55	155
Pall 1	2.65	0.52	160
Pall 2	2.62	0.51	174
Retgenbusch	2.31	0.49	117
Weierbach	2.83	0.63	91

#### 3.3.2. $\beta$ -Diversity: Known DNA-Bacteriophage Diversity between Sites

The  $\beta$ -diversity analysis allowed comparing the DNA bacteriophage subset populations from one site to another. The PERMANOVA test (Table 5) performed on the DNA bacteriophage genera abundance did not reveal a significant effect of the vegetation (i.e., forest and grasslands) on the distribution of the DNA bacteriophages ( $p$ -value = 0.075,  $\alpha = 0.05$ ).

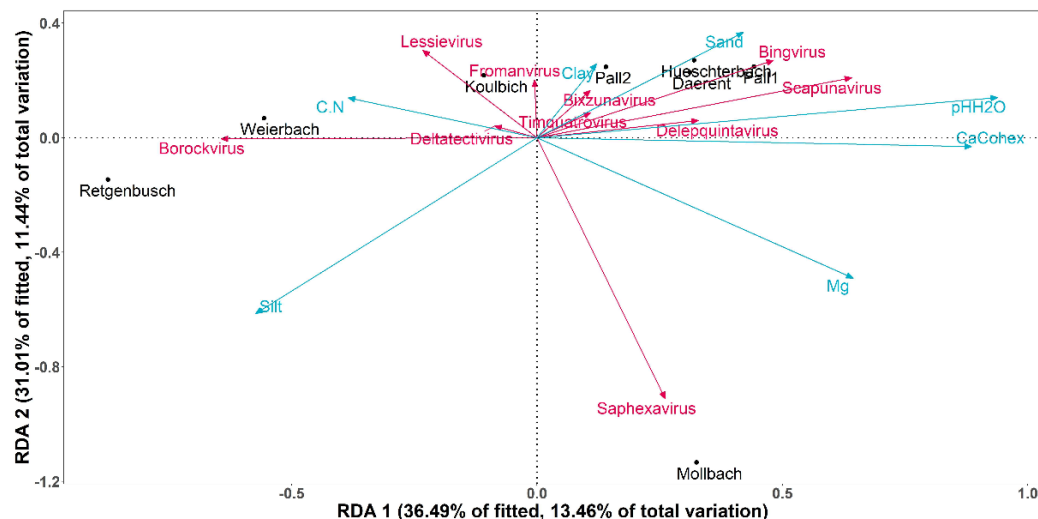
**Table 5.** PERMANOVA tests to analyse the influence of the selected abiotic factors on the distribution of the known DNA bacteriophage genera ( $\alpha = 0.05$ ).

Environmental Variables	$p$ -Value
Vegetation	0.075
Ca <sup>2+</sup>	0.008 **
Carbon-to-nitrogen ratio	0.349
pH	0.003 **
Silt content	0.125
Sand content	0.764
Clay content	0.901
Mg content	0.083
GWC	0.846

Significance: \*\*  $p < 0.01$ .

However, soil pH and calcium concentration were associated with the differences in the DNA bacteriophage compositions between the sites ( $p$ -value = 0.003 and 0.008, respectively,  $\alpha = 0.05$ ). A redundancy analysis (afterwards called RDA; Figure 4) was carried out to analyse the variations of bacteriophage genera depending on the soil properties of the sites. The RDA was built over the two first dimensions, covering together 67.5% of the fitted variation. Among all the abiotic factors, soil pH was the explanatory variable that had a

significant weight on the construction of the RDA ( $p$ -value = 0.008,  $\alpha$  = 0.05). On the one hand, the axis RDA1 was significantly correlated at 88.4% with soil pH ( $p$ -value < 0.01,  $\alpha$  = 0.05) and at 79.2% with calcium cation ( $p$ -value < 0.01,  $\alpha$  = 0.05). The RDA2, on the other hand, was not significantly correlated with any of the tested soil properties.

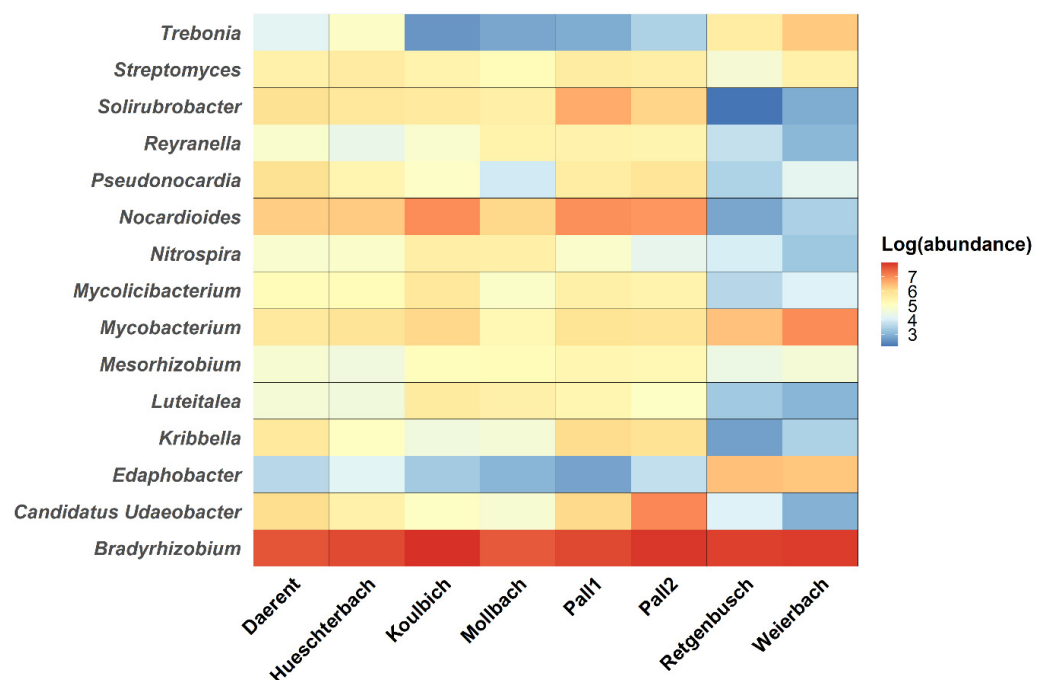


**Figure 4.** Biplot of redundancy analysis (RDA) on the 10 most abundant genera of DNA bacteriophage in relation to the soil properties of each site. The pink vectors represent the DNA bacteriophage genera with a Hellinger transformation, the blue vectors display the selected soil properties and the sites are shown in black. The RDA was built over the two first dimensions, with 36.49% of the fitted variations explained by RDA1 and 31.01% by RDA2. (C.N = carbon to nitrogen ratio; Mg = magnesium concentration; pH<sub>H<sub>2</sub>O</sub> = soil pH measured in water, CaCohex = calcium cation available).

Along the RDA1 axis, both soil pH and calcium concentration explained the sites distributions. Indeed, Weierbach and Retgenbusch were displayed apart from the other sites, as they were the most acidic soils (pH = 4.1 and 4.5, respectively) and had the lowest concentration in calcium ( $\text{Ca}^{2+}$  = 21 and 16 positive charges per kg of soil, respectively). The RDA showed a particular correlation between the sites of Weierbach and Retgenbusch and the genus *Borockvirus*. On the contrary, Pall 1, Huescherbach and Daerent presented the highest soil pH (pH = 6.8, 6.4 and 6.3, respectively) with relatively high concentrations of calcium ( $\text{Ca}^{2+}$  = 166, 126 and 87 positive charges per kg of soil, respectively). In terms of DNA bacteriophage genera, these soil properties (i.e., pH and  $\text{Ca}^{2+}$ ) showed a negative correlation with *Bingvirus* and *Scapunavirus*. Finally, Mollbach showed a particular distribution, related to the magnesium content, as indeed its soil presented the highest concentration of magnesium (Mg = 52 mg/100 g of soil). Additionally, *Saphexavirus* displayed a positive correlation specifically to the magnesium concentration.

### 3.4. Known Bacterial Populations and Connections with the Known DNA-Bacteriophage Communities

As reported in Supplementary Table S1, bacterial sequences represented 74 to 77% of the total microbial sequences identified in the libraries. First, bacterial populations of all sites were mainly composed of the following phyla, *Acidobacteria*, *Actinobacteria*, *Proteobacteria* and *Verrucomicrobia* (see Supplementary Figure S1). As for the DNA bacteriophages, the taxonomical composition of the bacterial populations was then analysed at the genus level by selecting the most abundant bacteria (abundance above 1000 mapped reads per species, all sites included). Retgenbusch and Weierbach showed similar patterns in bacterial genera with the lowest soil bacterial diversity ( $H$  = 2.78 and 2.89, respectively) (see Supplementary Table S2) and similar bacterial composition as is visible in Figure 5.

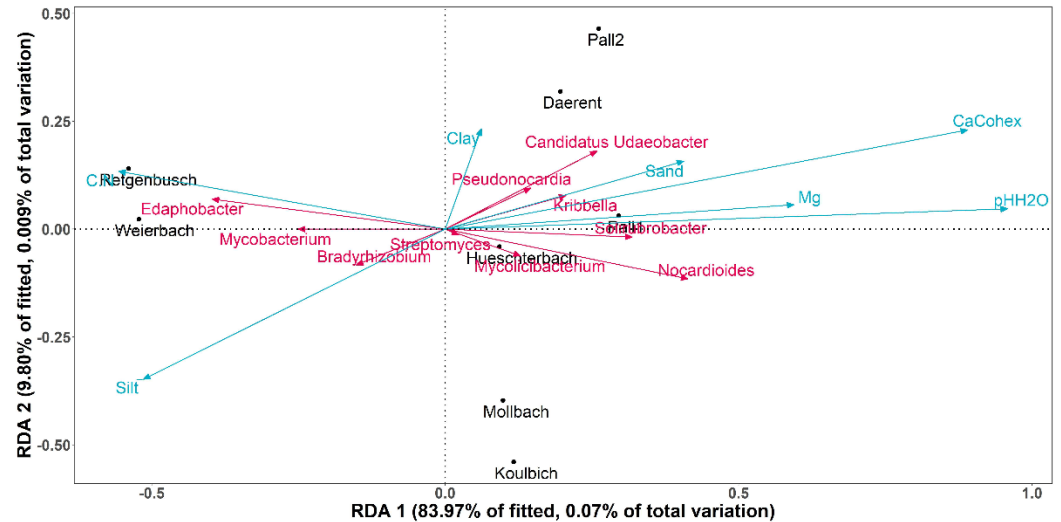


**Figure 5.** Heatmap of the most occurrent bacterial genera identified (selection with a log-abundance above 3) as a function of the sites. The gradient represents the log-transformed abundance of the genera, where red colour represents the most abundant genera and blue, the least ones.

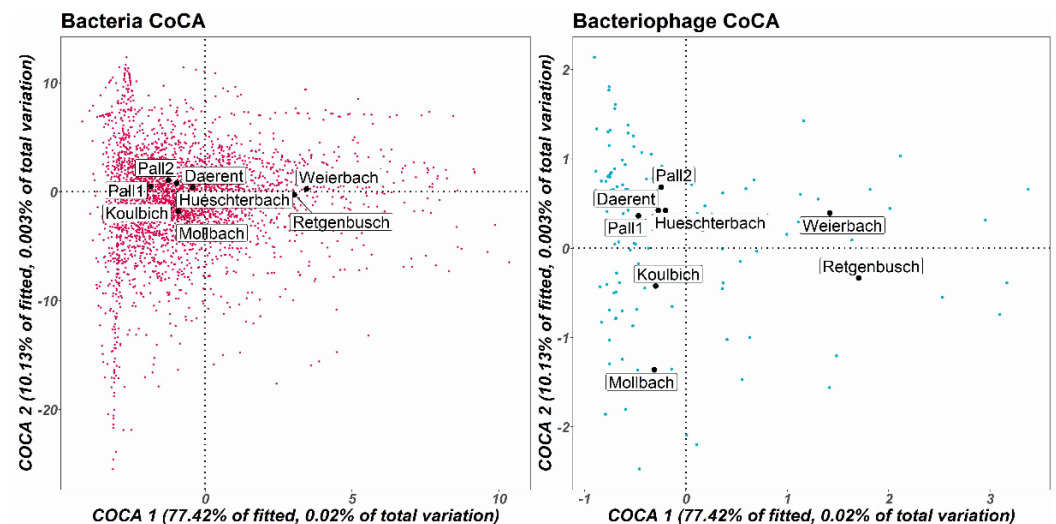
In addition, these two sites showed differences in bacterial compositions from the other sites. Indeed, Retgenbusch and Weierbach showed higher abundance for the genera *Trebonia*, *Edaphobacter* and *Mycobacterium*, for which the relative abundances were lower in the other soils. Regarding the other genera, with the exception of *Streptomyces* and *Bradyrhizobium*, all of them showed high abundance for all sites except for Retgenbusch and Weierbach. Finally, *Streptomyces* and *Bradyrhizobium*, were detected in every soil with the highest abundance reported for the latter. The differences in bacterial community composition between the sites were studied through a PERMANOVA test, performed on the bacteria genera abundance. As noticed for the DNA bacteriophages, the vegetation (i.e., forest and grasslands) did not significantly impact the bacteria distribution ( $p$ -value = 0.121,  $\alpha$  = 0.05). In addition, soil pH and calcium concentration also displayed significant effect on the bacterial compositions between the sites ( $p$ -value = 0.006 and 0.002, respectively,  $\alpha$  = 0.05). Contrary to the observations on the DNA bacteriophages, magnesium concentration played a significant role on the bacterial composition ( $p$ -value = 0.027,  $\alpha$  = 0.05). As for the DNA bacteriophage investigation, RDA was carried out to analyse the variations of bacterial genera depending on the same selected soil properties of the sites (Figure 6). The RDA was built over the two first dimensions, covering together 93.77% of the fitted variation. Among all the abiotic factors, soil pH was the sole explanatory variable that had a significant weight on the construction of the RDA ( $p$ -value = 0.01,  $\alpha$  = 0.05). On one hand, the axis RDA1 was significantly correlated at 92.16% with soil pH ( $p$ -value < 0.01,  $\alpha$  = 0.05) and at 79.21% with calcium ( $p$ -value < 0.01,  $\alpha$  = 0.05). On the other hand, the RDA2 was not significantly correlated with any of the tested soil properties. Weierbach and Retgenbusch were found apart from all the other sites, and both related to only *Edaphobacter* and *Mycobacterium*.

As DNA bacteriophages and bacteria are known for being exclusively linked to one another, the covariation of these two communities was analysed using a symmetric CoCA. This ordination method was built over the two first dimensions that covered 67.4% of the total covariation (Figure 7). It allows exploring the distribution of the whole known bacterial and DNA bacteriophage subset populations at the genera level by analysing both simultaneously. The patterns for both bacterial and DNA bacteriophage CoCA were similar and confirmed the distinct group formed by Weierbach and Retgenbusch. The

latter was displayed offset from the other sites while Daerent, Pall 1, Pall 2, Mollbach and Hueschterbach were similar to each other on both bacterial and bacteriophage CoCA. A Mantel test using Pearson correlation carried on DNA bacteriophage and bacterial genera confirmed a strong positive correlation ( $R$ -value = 0.75, significance = 0.001).



**Figure 6.** Biplot of redundancy analysis (RDA) on the 10 most abundant genera of bacteria in relation to the soil properties of each site. The pink vectors represent the bacterial genera, the blue vectors display the selected soil properties and the sites are shown in black. The RDA was built over the two first dimensions, with 83.97% of the fitted variations explained by RDA1 and 9.80% by RDA2. (C.N = carbon to nitrogen ratio; Mg = magnesium content; pHH<sub>2</sub>O = soil pH measured in water, CaCohex = calcium cation available).



**Figure 7.** Symmetric co-correspondence analysis (Sym-CoCA) of both total bacteria and DNA bacteriophage community compositions. The bacterial genera are represented in pink and bacteriophage genera, in blue. Here, they were built over the two first dimensions, where 77.42% of the fitted covariation between the two matrices was explained by COCA1 and 10.13% by COCA2. The sites are displayed in black.

#### 4. Discussion

In our study, a viral concentration of approximately  $10^9$  to  $10^{10}$  VLPs per gram of dry soil was found, which is consistent with previous works [14,68,69]. Although VLP counts are not always considered to be a good representative of virus concentrations, the counts performed here were carried out in the same manner as previously on freshwater

sediments using the FACSCalibur FCM table. This method was previously validated and provided similar VLP counts to those obtained by EFM [70]. A good correlation was found, highlighting the high sensitivity of the device and thus, allowing obtaining unambiguous signatures for VLPs by discriminating noise from VLPs (virus-like particles) to PLPs (prokaryote-like particles). It is noteworthy, however, that flow cytometry analysis does not allow identifying viruses or discriminating bacteriophages from the rest of the VLPS [71]. To draw a complete picture of the viral community, a taxonomic identification of the known DNA bacteriophage populations was complementarily performed through shotgun metagenomics and reference-based bioinformatics. The taxonomy was assigned using the nr database released in April 2022; however, it is noteworthy that the viral taxonomy built by the International Committee on Taxonomy of Viruses (ICTV) is nowadays updated according to the genetic similarities between bacteriophages, including reclassification of the families observed in the current study, such as *Siphoviridae*, *Myoviridae* and *Podoviridae*.

Our study showed *Caudovirales* as the most frequently identified order in all sites, while *Siphoviridae*, *Podoviridae* and *Myoviridae* were found as the most abundant families observed within the identified DNA bacteriophages, which is consistent with most of the literature that studied viral soil diversity [13,26,29]. However, the high proportion of these three families is rather the reflection of the taxonomy made available on the databases up to now, where most of the bacteriophage species are classified into these families [72]. For this reason, it is important to consider low taxonomic levels in diversity studies. In the current study, the viral genera obtained from the libraries did not display specific patterns related to the ecosystems (i.e., forests vs. grasslands). The three selected forest sites presented different vegetal covers (i.e., beech, spruce and oak) with diverse types of soil textures. The forest sites might not be sufficiently similar to constitute a common pattern in DNA bacteriophage diversity. However, several patterns in the known DNA bacteriophage distribution are distinguishable by considering the specific soil parameters. The soil pH as well as the calcium concentration significantly drove the ecological distribution of the known DNA bacteriophage communities among the different studied soils. More interestingly, the distribution of both DNA bacteriophage and bacterial communities in diverse types of soil were found to be strongly linked to one another and their distributions were explained by the same abiotic factors, i.e., soil pH and calcium concentration. Additionally, a complementary effect of magnesium concentration solely on the bacterial distribution was detected. The differences in microbial diversity were highlighted especially for the soils of Weierbach and Retgenbusch. Indeed, these two sites did not only display an acidic pH (around 4) with low calcium availability (<5 meq per 100 g of soil), but also low abundance and diversity of the known microbial populations at the genus level. While studying the known DNA bacteriophage populations, these same three soils were represented by the significant presence of one specific DNA bacteriophage genus. Indeed, *Saphexavirus*, *Borockovirus* and *Lessievirus* were highly abundant in the soil of Mollbach, Retgenbusch and Weierbach, respectively, compared to the other studied soils. Some studies reported a dominance of small ssDNA viruses rather than dsDNA-tailed bacteriophages in the soils [73–75]. In the present study, ssDNA bacteriophages (*Microviridae* and *Inoviridae*) were also detected, but in low frequency in comparison to dsDNA bacteriophages. Most ssDNA bacteriophages detected from environmental metagenomes belonged to the *Microviridae* family, for which all members infect a narrow range of hosts [76]. However, most ssDNA bacteriophages remain scarcely characterised within the taxonomy as *Inoviridae* represents 11% and *Microviridae* 2% of the ICTV database [77], reflecting a potential bias in the determination of their abundance in the soil samples.

Calcium is an important component participating in soil life. It plays a role in the soil structure, firstly as a binding agent in the aggregation of soil particles, secondly as a macronutrient in plant nutrition and thirdly as a factor facilitating soil permeability [78]. Then, calcium has a biological role in promoting an optimal environment to enhance soil microorganism activities [79]. Finally, calcium is highly related to soil pH. Indeed, soils with low pH generally display low calcium availability. As high concentrations of hydrogen ions



displace calcium concentration from binding sites, it results in the leaching of “free” calcium, particularly in the more porous soils [80,81]. Such soils, such as the ones of Weierbach and Retgenbusch, i.e., low pH and calcium concentration, did display a low microbial diversity among the bacterial and DNA bacteriophage populations characterised in this study. On the contrary, Pall 1, Pall 2 and Daerent with much higher calcium concentrations (between 10 and 20 meq per 100 g of soil) and an ideal range of soil pH (between 6 and 7) showed higher microbial diversity. In laboratory experiments, divalent cations (such as calcium and magnesium) were found to interact with bacteriophage MS2 by increasing its attachment to the surface of river natural organic matter [82]. In soil, adsorption onto surface particles is the most important process implied in the bacteriophages-soil interactions [11]. This interaction is highly dependent upon the type of cations found in the environment. Calcium cation (positively charged) is highly involved in the adsorption of clay-humus complex (CHC), negatively charged on its surface, through electrostatic forces [83,84]. Bacteriophages can then adsorb onto the CHC by the presence of the cations on the surface. Indeed, bacteriophages are charged on their surface; however, this charge is highly dependent on pH. At neutral pH (7), most bacteriophages are negatively charged and can interact with the available cations [85]. Reducing the pH will induce some bacteriophage surfaces to turn into positive charges, according to their isoelectric point (pI). Once positive, bacteriophages can compete with cations to bind with the clay charges of the CHC. Indeed, bacteriophages tend to be more retained in soils rich in clay. The latter has a lower specific surface, due to large particles and a greater ion-exchange capacity, which highly adsorb bacteriophages [86]. Therefore, the combination between calcium cation and pH can play a role in the bacteriophage dynamic in soils. However, not only did these factors show a significant effect on the DNA bacteriophage distribution, but they also did on the bacterial distribution. Additionally, our results showed a strong positive correlation between both bacterial and DNA bacteriophages distribution among the sites, suggesting a strong interaction. Therefore, soil pH seems like the most significant factor in leading the known DNA bacteriophage distributions, directly or/and indirectly through the shaping of bacterial communities. Calcium cation availability might rather play a role in the dynamic of DNA bacteriophages, through their adsorption with the soil particles.

The advances in metagenomics and bioinformatics have enabled deeper investigations for the identification of viruses [87,88]; however, some technical issues remain in the analysis of soil viruses [89]. As viruses tend to adsorb on soil particles, virus elution and concentration steps are often required before performing metagenomics. Nevertheless, this step generally ends in low recovery efficiencies [90,91], as noticed in our results during the optimisation of the elution and concentration protocol. Indeed, even though the viral recovery was satisfactory to conduct a viral quantification using flow cytometry, it resulted in low DNA yield after nucleic acid extraction. Therefore, it affected the good quality of viral sequences obtained after conducting the shotgun metagenomics, leading us to carry out the DNA extraction directly from the soil samples. As the analysis is solely relying on virus-affiliated reads that can be obtained from total DNA metagenomes, few viral contigs were obtained with N50 comprising between 200 and 300 per library. This can limit the interpretation of relationships between viruses and other microbes within the same samples.

Much remains to be understood about the diversity, dynamics and roles of bacteriophages in soils and metagenomics approaches have only recently enabled increasing our knowledge, particularly under in situ conditions. In laboratory experiments, some physicochemical properties, especially pH and ion strength, influenced bacteriophage inactivation [92,93]. Our results seemed to show the effect of these similar factors on bacteriophage diversity in soils; however, it is noteworthy that the study was focused on the known fraction of the microbial populations in the diverse selected soils. The OTU-based approach is less dependent on reference databases but works by clustering sequences according to similarity thresholds (often set at a sequence similarity level above 97%) [94–96]. This approach does not identify the taxonomic species but is rather a proxy for them and

has shown to be phylogenetically incoherent [94–96]. Therefore, further analyses using complementary molecular and bioinformatics methods, including untargeted approaches, could be applied to obtain a more complete picture of bacteriophage dynamics. Additionally, the investigation of the influence of certain abiotic factors on bacteriophage diversity based on what has already been characterised could open perspectives for the use of these species as, for instance, biotechnologies or in the development of new culture models for laboratory experiments.

## 5. Conclusions

In this study, the natural diversity and distribution of DNA bacteriophage populations in soils were explored as well as their relationship with the soil-dwelling bacterial populations. Through the complementary use of metagenomics, reference-based bioinformatics and multidimensional analyses, we identified the known fraction of DNA bacteriophage communities in forest and grassland soils and deciphered the influence of some soil properties and microbial communities on the DNA bacteriophage occurrence. Both soil pH and calcium content were observed to influence the distribution of DNA bacteriophage and bacterial populations. A strong positive correlation between DNA bacteriophage and bacterial communities was confirmed, suggesting in addition to soil pH and calcium content, bacteria are also involved. Therefore, our results add to the growing body of evidence that bacteriophages are likely to play key roles in soil ecology.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms10071458/s1>, Figure S1: Bar charts of the relative abundance (in %) of bacterial phylum (selected genera with an abundance above 100 occurrences) from the known fraction of all studied sites; Figure S2: Per base sequence quality from FastQC for read forward 1 (A) and read reverse 2 (B) of the sequencing of the Weierbach soil; Table S1: Percentages of relative abundances identified for each taxonomic domain; Table S2: Shannon diversity index (H), evenness (E) and species richness (S) of bacteria populations in the soil of the studied sites; Table S3: Recovery of bacteriophage  $\Phi$ X174 comparing three methods for mechanical viral detachment from the Weierbach soil; Table S4: Recovery of bacteriophage  $\Phi$ X174 from the Weierbach soil comparing six different eluent solutions; Table S5: Recovery of bacteriophage  $\Phi$ X174 from the Weierbach soil after concentration step, comparing Amicon with two different cut-off points (i.e., 30 kDa and 50 kDa); Table S6: Recovery of bacteriophage  $\Phi$ X174 at each step of the elution and concentration protocol for each study site; Table S7: Summary of each step from the DNase treatment to the library preparation from viral concentrates and soil sample of the Weierbach site. References [97–102] are cited in the supplementary materials.

**Author Contributions:** Conceptualization, P.F., L.O. and H.-M.C.; methodology, P.F., L.O., M.H., S.J. and H.-M.C.; bioinformatics, M.H.; flow cytometry: S.J.; validation, P.F., L.O., M.H., S.J. and H.-M.C.; writing—original draft preparation, P.F.; writing—review and editing, P.F., L.O., M.H., S.J. and H.-M.C.; visualization, P.F.; supervision, L.O. and H.-M.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Luxembourg National Research Fund (FNR) for granting and supporting the PRIDE program containing the Doctoral Training Unit Hydro-CSI project (PRIDE15/10623093—HYDRO-CSI).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The raw sequence data generated and analysed within this study are available in the National Centre for Biotechnology Information Databases (Sequence Read Archive) under the BioProject number PRJNA771214.

**Acknowledgments:** The authors would like to acknowledge the Luxembourg National Research Fund (FNR) for granting and supporting the PRIDE program containing the Doctoral Training Unit Hydro-CSI project (PRIDE15/10623093—HYDRO-CSI). The authors would also like to thank Thomas Mion and Magdalena Calusinska for their valuable assistance in the bioinformatics pipeline by providing us with their expertise in this field. The soil analyses could not have been done without the cooperation of Lionel Leydet (Administration des services techniques de l’agriculture, ASTA). We would finally thank our colleagues, Christophe Hissler and Nathalie Goeders for discussing with us regarding, respectively, the soil data and statistics analyses.

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

## References

- Lange, M.; Eisenhauer, N.; Sierra, C.A.; Bessler, H.; Engels, C.; Griffiths, R.I.; Mellado-Vázquez, P.G.; Malik, A.A.; Roy, J.; Scheu, S.; et al. Plant Diversity Increases Soil Microbial Activity and Soil Carbon Storage. *Nat. Commun.* **2015**, *6*, 6707. [[CrossRef](#)]
- Vezzani, F.M.; Anderson, C.; Meenken, E.; Gillespie, R.; Peterson, M.; Beare, M.H. The Importance of Plants to Development and Maintenance of Soil Structure, Microbial Communities and Ecosystem Functions. *Soil Tillage Res.* **2018**, *175*, 139–149. [[CrossRef](#)]
- Eldridge, D.; James, A. Soil-disturbance by Native Animals Plays a Critical Role in Maintaining Healthy Australian Landscapes. *Ecol. Manag. Restor.* **2009**, *10*, S27–S34. [[CrossRef](#)]
- Lou, X.; Zhao, J.; Lou, X.; Xia, X.; Feng, Y.; Li, H. The Biodegradation of Soil Organic Matter in Soil-Dwelling Humivorous Fauna. *Front. Bioeng. Biotechnol.* **2021**, *9*, 1406. [[CrossRef](#)] [[PubMed](#)]
- Odelade, K.A.; Babalola, O.O. Bacteria, Fungi and Archaea Domains in Rhizospheric Soil and Their Effects in Enhancing Agricultural Productivity. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3873. [[CrossRef](#)]
- Rashid, M.I.; Mujawar, L.H.; Shahzad, T.; Almeelbi, T.; Ismail, I.M.I.; Oves, M. Bacteria and Fungi Can Contribute to Nutrients Bioavailability and Aggregate Formation in Degraded Soils. *Microbiol. Res.* **2016**, *183*, 26–41. [[CrossRef](#)]
- Buerkert, A.; Joergensen, R.G.; Ludwig, B.; Schlecht, E. Chapter 18—Nutrient and Carbon Fluxes in Terrestrial Agro-Ecosystems. In *Marschner’s Mineral Nutrition of Higher Plants*, 3rd ed.; Academic Press: San Diego, CA, USA, 2012; pp. 473–482. ISBN 978-0-12-384905-2.
- Ortiz, A.; Sansinenea, E. The Role of Beneficial Microorganisms in Soil Quality and Plant Health. *Sustainability* **2022**, *14*, 5358. [[CrossRef](#)]
- Beule, L.; Vaupel, A.; Moran-Rodas, V.E. Abundance, Diversity, and Function of Soil Microorganisms in Temperate Alley-Cropping Agroforestry Systems: A Review. *Microorganisms* **2022**, *10*, 616. [[CrossRef](#)]
- Armon, R. Soil Bacteria and Bacteriophages. In *Biocommunication in Soil Microorganisms*; Witzany, G., Ed.; Springer: Heidelberg, Germany, 2011; Volume 23, pp. 67–112. ISBN 978-3-642-14511-7.
- Kimura, M.; Jia, Z.J.; Nakayama, N.; Asakawa, S. Ecology of Viruses in Soils: Past, Present and Future Perspectives. *Soil Sci. Plant Nutr.* **2008**, *54*, 1–32. [[CrossRef](#)]
- Coyne, M.S. A Cartoon History of Soil Microbiology. *J. Nat. Resour. Life Sci. Educ.* **1996**, *25*, 30–36. [[CrossRef](#)]
- Williamson, K.E.; Fuhrmann, J.J.; Wommack, K.E.; Radosevich, M. Viruses in Soil Ecosystems: An Unknown Quantity Within an Unexplored Territory. *Annu. Rev. Virol.* **2017**, *4*, 201–219. [[CrossRef](#)] [[PubMed](#)]
- Williamson, K.E.; Corzo, K.A.; Drissi, C.L.; Buckingham, J.M.; Thompson, C.P.; Helton, R.R. Estimates of Viral Abundance in Soils Are Strongly Influenced by Extraction and Enumeration Methods. *Biol. Fertil. Soils* **2013**, *49*, 857–869. [[CrossRef](#)]
- Bertrand, I.; Schijven, J.F.; Sánchez, G.; Wyn-Jones, P.; Ottoson, J.; Morin, T.; Muscillo, M.; Verani, M.; Nasser, A.; de Roda Husman, A.M.; et al. The Impact of Temperature on the Inactivation of Enteric Viruses in Food and Water: A Review. *J. Appl. Microbiol.* **2012**, *112*, 1059–1074. [[CrossRef](#)] [[PubMed](#)]
- Corapcioglu, M.Y.; Vogel, J.R.; Munster, C.L.; Pillai, S.D.; Dowd, S.; Wang, S. Virus Transport Experiments in a Sandy Aquifer. *Water Air Soil Pollut.* **2006**, *169*, 47–65. [[CrossRef](#)]
- Fauvel, B.; Cauchie, H.M.; Gantzer, C.; Ogorzaly, L. Influence of Physico-Chemical Characteristics of Sediment on the in Situ Spatial Distribution of F-Specific RNA Phages in the Riverbed. *FEMS Microbiol. Ecol.* **2019**, *95*, fiy240. [[CrossRef](#)]
- Jurczak-Kurek, A.; Gasiór, T.; Nejman-Faleńczyk, B.; Bloch, S.; Dydecka, A.; Topka, G.; Necel, A.; Jakubowska-Deredas, M.; Narajczyk, M.; Richert, M.; et al. Biodiversity of Bacteriophages: Morphological and Biological Properties of a Large Group of Phages Isolated from Urban Sewage. *Sci. Rep.* **2016**, *6*, 34338. [[CrossRef](#)]
- Feng, Y.Y.; Ong, S.L.; Hu, J.Y.; Tan, X.L.; Ng, W.J. Effects of PH and Temperature on the Survival of Coliphages MS2 and Qβ. *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 549–552. [[CrossRef](#)]
- Nishide, M.; Tsujimoto, K.; Uozaki, M.; Ikeda, K.; Yamasaki, H.; Koyama, A.H.; Arakawa, T. Effects of Electrolytes on Virus Inactivation by Acidic Solutions. *Int. J. Mol. Med.* **2011**, *27*, 803–809. [[CrossRef](#)]
- Taj, M.K.; Ling, J.X.; Bing, L.L.; Qi, Z.; Taj, I.; Hassani, T.M.; Samreen, Z.; Yunlin, W. Effect of Dilution, Temperature and PH on the Lysis Activity of T4 Phage against E. Coli BL21. *J. Anim. plant Sci.* **2014**, *24*, 1252–1255.
- Cuadros, J. Clay Minerals Interaction with Microorganisms: A Review. *Clay Miner.* **2017**, *52*, 235–261. [[CrossRef](#)]
- Park, J.A.; Kang, J.K.; Kim, J.H.; Kim, S.B.; Yu, S.; Kim, T.H. Bacteriophage Removal in Various Clay Minerals and Clay-Amended Soils. *Environ. Eng. Res.* **2015**, *20*, 133–140. [[CrossRef](#)]

24. Harvey, R.W.; Ryan, J.N. Use of PRD1 Bacteriophage in Groundwater Viral Transport, Inactivation, and Attachment Studies. *FEMS Microbiol. Ecol.* **2004**, *49*, 3–16. [[CrossRef](#)] [[PubMed](#)]
25. Williamson, K.E.; Radosevich, M.; Wommack, K.E. Abundance and Diversity of Viruses in Six Delaware Soils. *Appl. Environ. Microbiol.* **2005**, *71*, 3119–3125. [[CrossRef](#)] [[PubMed](#)]
26. Leroy, M.; Prigent, M.; Dutertre, M.; Confalonieri, F.; Dubow, M. Bacteriophage Morphotype and Genome Diversity in Seine River Sediment. *Freshw. Biol.* **2008**, *53*, 1176–1185. [[CrossRef](#)]
27. Parmar, K.; Dafale, N.; Pal, R.; Tikariha, H.; Purohit, H. An Insight into Phage Diversity at Environmental Habitats Using Comparative Metagenomics Approach. *Curr. Microbiol.* **2018**, *75*, 132–141. [[CrossRef](#)] [[PubMed](#)]
28. Mohiuddin, M.; Schellhorn, H.E. Spatial and Temporal Dynamics of Virus Occurrence in Two Freshwater Lakes Captured through Metagenomic Analysis. *Front. Microbiol.* **2015**, *6*, 960. [[CrossRef](#)] [[PubMed](#)]
29. Gong, Z.; Liang, Y.; Wang, M.; Jiang, Y.; Yang, Q.; Xia, J.; Zhou, X.; You, S.; Gao, C.; Wang, J.; et al. Viral Diversity and Its Relationship With Environmental Factors at the Surface and Deep Sea of Prydz Bay, Antarctica. *Front. Microbiol.* **2018**, *9*, 1–17. [[CrossRef](#)]
30. Narr, A.; Nawaz, A.; Wick, L.Y.; Harms, H.; Chatzinotas, A. Soil Viral Communities Vary Temporally and along a Land Use Transect as Revealed by Virus-like Particle Counting and a Modified Community Fingerprinting Approach (FRAPD). *Front. Microbiol.* **2017**, *8*, 1975. [[CrossRef](#)]
31. 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](#)]
32. Cao, M.-M.; Liu, S.-Y.; Bi, L.; Chen, S.-J.; Wu, H.-Y.; Ge, Y.; Han, B.; Zhang, L.-M.; He, J.-Z.; Han, L.-L. Distribution Characteristics of Soil Viruses Under Different Precipitation Gradients on the Qinghai-Tibet Plateau. *Front. Microbiol.* **2022**, *13*, 848305. [[CrossRef](#)]
33. Han, L.L.; Yu, D.T.; Bi, L.; Du, S.; Silveira, C.; Cobián Güemes, A.G.; Zhang, L.M.; He, J.Z.; Rohwer, F. Distribution of Soil Viruses across China and Their Potential Role in Phosphorous Metabolism. *Environ. Microbiome* **2022**, *17*, 6. [[CrossRef](#)]
34. Koskella, B.; Brockhurst, M.A. Bacteria–Phage Coevolution as a Driver of Ecological and Evolutionary Processes in Microbial Communities. *FEMS Microbiol. Rev.* **2014**, *38*, 916–931. [[CrossRef](#)]
35. Fazzino, L.; Anisman, J.; Chacón, J.M.; Heineman, R.H.; Harcombe, W.R. Lytic Bacteriophage Have Diverse Indirect Effects in a Synthetic Cross-Feeding Community. *ISME J.* **2020**, *14*, 123–134. [[CrossRef](#)] [[PubMed](#)]
36. Sullivan, M.B.; Weitz, J.S.; Wilhelm, S. Viral Ecology Comes of Age. *Environ. Microbiol. Rep.* **2017**, *9*, 33–35. [[CrossRef](#)] [[PubMed](#)]
37. Klimenko, A.I.; Matushkin, Y.G.; Kolchanov, N.A.; Lashin, S.A. Bacteriophages Affect Evolution of Bacterial Communities in Spatially Distributed Habitats: A Simulation Study. *BMC Microbiol.* **2016**, *16*, S10. [[CrossRef](#)]
38. Ashelford, K.E.; Day, M.J.; Fry, J.C. Elevated Abundance of Bacteriophage Infecting Bacteria in Soil. *Appl. Environ. Microbiol.* **2003**, *69*, 285–289. [[CrossRef](#)] [[PubMed](#)]
39. Braga, L.P.P.; Spor, A.; Kot, W.; Breuil, M.C.; Hansen, L.H.; Setubal, J.C.; Philippot, L. Impact of Phages on Soil Bacterial Communities and Nitrogen Availability under Different Assembly Scenarios. *Microbiome* **2020**, *8*, 52. [[CrossRef](#)]
40. Marx, S.; Flammang, F. *La Cartographie Des Sols Au Grand-Duché de Luxembourg Version Provisoire\_V4*; Ministère de L’agriculture, de la Viticulture et du Développement Rural: Luxembourg, 2015.
41. Heffernan, A.L.; Aylward, L.L.; Toms, L.M.L.; Sly, P.D.; Macleod, M.; Mueller, J.F. Pooled Biological Specimens for Human Biomonitoring of Environmental Chemicals: Opportunities and Limitations. *J. Expo. Sci. Environ. Epidemiol.* **2014**, *24*, 225–232. [[CrossRef](#)]
42. Caudill, S.P. Characterizing Populations of Individuals Using Pooled Samples. *J. Expo. Sci. Environ. Epidemiol.* **2010**, *20*, 29–37. [[CrossRef](#)]
43. Roy, K.; Ghosh, D.; DeBruyn, J.M.; Dasgupta, T.; Wommack, K.E.; Liang, X.; Wagner, R.E.; Radosevich, M. Temporal Dynamics of Soil Virus and Bacterial Populations in Agricultural and Early Plant Successional Soils. *Front. Microbiol.* **2020**, *11*, 1494. [[CrossRef](#)]
44. AFNOR. *ISO 10390: Sols, Biodéchets Traités et Boues—Détermination Du PH*, 3rd ed.; Association Française de Normalisation: Geneva, Switzerland, 2021.
45. VDLUFA. *Methodenbuch A5.1.1: Bestimmung des PH-Wertes*; Verlag: Darmstadt, Germany, 2016.
46. VDLUFA. *Methodenbuch A6.2.1.1: Bestimmung von Phosphor und Kalium im Calcium-Acetat-Lactat-Auszug*; Verlag: Darmstadt Germany, 2012.
47. VDLUFA. *Methodenbuch A6.2.1.7: Bestimmung von Pflanzenverfügbarem Kalium und Natrium im Calciumchloridauszug auf Gewichtsbasis*; Verlag: Darmstadt, Germany, 1997.
48. AFNOR. *ISO 10694: Qualité du Sol—Dosage du Carbone Organique et du Carbon Total Après Combustion Sèche (Analyse Élémentaire)*, 1st ed.; Association Française de Normalisation: Geneva, Switzerland, 1995.
49. AFNOR. *ISO 13878: Qualité du Sol—Détermination de la Teneur Totale en Azote par Combustion Sèche (“Analyse Élémentaire”)*, 1st ed.; Association Française de Normalisation: Geneva, Switzerland, 1998.
50. AFNOR. *ISO 23470: Qualité Du Sol—Détermination de La Capacité D’échange Cationique (CEC) Effective et des Cations Échangeables à L’aide d’une Solution de Trichlorure de Cobaltihexammine*, 1st ed.; Association Française de Normalisation: Geneva, Switzerland, 2007.
51. Marx, S.; Flammang, F. *Cartographie des Sols au Grand-Duché de Luxembourg*; Ministère de L’agriculture, de la Viticulture et du Développement Rural: Luxembourg, 2018.
52. Jacquet, S.; Dorigo, U.; Personnic, S. A Few Tests Prior to Flow Cytometry and Epifluorescence Analyses of Freshwater Bacterioand Virioplankton Communities. *Flow Cytom. Princ. Methodol. Appl.* **2013**, *1*, 1–30.

53. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
54. Nurk, S.; Meleshko, D.; Korobeynikov, A.; Pevzner, P.A. MetaSPAdes: A New Versatile Metagenomic Assembler. *Genome Res.* **2017**, *27*, 824–834. [[CrossRef](#)] [[PubMed](#)]
55. Steinegger, M.; Söding, J. MMseqs2 Enables Sensitive Protein Sequence Searching for the Analysis of Massive Data Sets. *Nat. Biotechnol.* **2017**, *35*, 1026–1028. [[CrossRef](#)]
56. Mirdita, M.; Steinegger, M.; Breitwieser, F.; Söding, J.; Levy Karin, E. Fast and Sensitive Taxonomic Assignment to Metagenomic Contigs. *Bioinformatics* **2021**, *37*, 3029–3031. [[CrossRef](#)]
57. Li, H. Minimap2: Pairwise Alignment for Nucleotide Sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [[CrossRef](#)]
58. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve Years of SAMtools and BCFtools. *Gigascience* **2021**, *10*, giab008. [[CrossRef](#)]
59. Palermo, C.N.; Fulthorpe, R.R.; Saati, R.; Short, S.M. Metagenomic Analysis of Virus Diversity and Relative Abundance in a Eutrophic Freshwater Harbour. *Viruses* **2019**, *11*, 792. [[CrossRef](#)]
60. Zhao, Y.; Li, M.C.; Konaté, M.M.; Chen, L.; Das, B.; Karlovich, C.; Williams, P.M.; Evrard, Y.A.; Doroshov, J.H.; McShane, L.M. TPM, FPKM, or Normalized Counts? A Comparative Study of Quantification Measures for the Analysis of RNA-Seq Data from the NCI Patient-Derived Models Repository. *J. Transl. Med.* **2021**, *19*, 269. [[CrossRef](#)]
61. RStudio Team. *RStudio: Integrated Development for R*. RStudio; PBC: Boston, MA, USA, 2020; Available online: <http://www.rstudio.com/> (accessed on 11 July 2022).
62. Thukral, A.K. A Review on Measurement of Alpha Diversity in Biology. *Agric. Res. J.* **2017**, *54*, 10. [[CrossRef](#)]
63. Oksanen, J.; Blanchet, F.G.; Friendly, M.; Kindt, R.; Legendre, P.; McGlinn, D.; Minchin, P.R.; O'Hara, R.B.; Simpson, G.L.; Solymos, P.; et al. vegan: Community Ecology Package. R Package Version 2.5-7. 2020. Available online: <https://CRAN.R-project.org/package=vegan> (accessed on 11 July 2022).
64. Mulder, C.P.H.; Bazeley-White, E.; Dimitrakopoulos, P.G.; Hector, A.; Scherer-Lorenzen, M.; Schmid, B. Species Evenness and Productivity in Experimental Plant Communities. *Oikos* **2004**, *107*, 50–63. [[CrossRef](#)]
65. Moore, J.C. Diversity, Taxonomic versus Functional. In *Encyclopedia of Biodiversity*, 2nd ed.; Academic Press: Cambridge, MA, USA, 2013; pp. 648–656. ISBN 9780123847195.
66. Socolar, J.B.; Gilroy, J.J.; Kunin, W.E.; Edwards, D.P. How Should Beta-Diversity Inform Biodiversity Conservation? *Trends Ecol. Evol.* **2016**, *31*, 67–80. [[CrossRef](#)] [[PubMed](#)]
67. USDA (United States Department of Agriculture, Natural resources conservation service, Soil Texture Calculator. Available online: [https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2\\_054167](https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/?cid=nrcs142p2_054167) (accessed on 18 July 2022).
68. Pratama, A.A.; van Elsland, J.D. The 'Neglected' Soil Virome—Potential Role and Impact. *Trends Microbiol.* **2018**, *26*, 649–662. [[CrossRef](#)]
69. Williamson, K.E.; Wommack, K.E.; Radosevich, M. Sampling Natural Viral Communities from Soil for Culture-Independent Analyses. *Appl. Environ. Microbiol.* **2003**, *69*, 6628–6633. [[CrossRef](#)] [[PubMed](#)]
70. Duhamel, S.; Jacquet, S. Flow Cytometric Analysis of Bacteria- and Virus-like Particles in Lake Sediments. *J. Microbiol. Methods* **2006**, *64*, 316–332. [[CrossRef](#)]
71. Colombet, J.; Billard, H.; Viguès, B.; Balor, S.; Boulé, C.; Geay, L.; Benzerara, K.; Menguy, N.; Ilango, G.; Fuster, M.; et al. Discovery of High Abundances of Aster-Like Nanoparticles in Pelagic Environments: Characterization and Dynamics. *Front. Microbiol.* **2019**, *10*, 2376. [[CrossRef](#)] [[PubMed](#)]
72. Turner, D.; Kropinski, A.M.; Adriaenssens, E.M. A Roadmap for Genome-Based Phage Taxonomy. *Viruses* **2021**, *13*, 506. [[CrossRef](#)]
73. Kim, K.H.; Chang, H.W.; Nam, Y.D.; Roh, S.W.; Kim, M.S.; Sung, Y.; Jeon, C.O.; Oh, H.M.; Bae, J.W. Amplification of Uncultured Single-Stranded DNA Viruses from Rice Paddy Soil. *Appl. Environ. Microbiol.* **2008**, *74*, 5975–5985. [[CrossRef](#)]
74. Reavy, B.; Swanson, M.M.; Cock, P.J.A.; 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. Microbiol.* **2015**, *81*, 3934–3945. [[CrossRef](#)]
75. Grujčić, V.; Nuy, J.K.; Salcher, M.M.; Shabarova, T.; Kasalický, V.; Boenigk, J.; Jensen, M.; Simek, K. Cryptophyta as Major Bacterivores in Freshwater Summer Plankton. *ISME J.* **2018**, *12*, 1668–1681. [[CrossRef](#)]
76. Cherwa, J.E.; Fane, B.A. Microviridae: Microviruses and Gokushoviruses. *eLS* **2011**, *5*, 1–10. [[CrossRef](#)]
77. Székely, A.J.; Breitbart, M. Single-Stranded DNA Phages: From Early Molecular Biology Tools to Recent Revolutions in Environmental Microbiology. *FEMS Microbiol. Lett.* **2016**, *363*, 27. [[CrossRef](#)]
78. Wuddivira, M.N.; Camps-Roach, G. Effects of Organic Matter and Calcium on Soil Structural Stability. *Eur. J. Soil Sci.* **2007**, *58*, 722–727. [[CrossRef](#)]
79. Zavarzin, G.A. Microbial Geochemical Calcium Cycle. *Microbiology* **2002**, *71*, 1–17. [[CrossRef](#)]
80. Whittinghill, K.A.; Hobbie, S.E. Effects of PH and Calcium on Soil Organic Matter Dynamics in Alaskan Tundra. *Biogeochemistry* **2011**, *111*, 569–581. [[CrossRef](#)]
81. Ng, J.F.; Ahmed, O.H.; Jalloh, M.B.; Omar, L.; Kwan, Y.M.; Musah, A.A.; Poong, K.H. Soil Nutrient Retention and PH Buffering Capacity Are Enhanced by Calciprill and Sodium Silicate. *Agronomy* **2022**, *12*, 219. [[CrossRef](#)]
82. Pham, M.; Mintz, E.A.; Nguyen, T.H. Deposition Kinetics of Bacteriophage MS2 to Natural Organic Matter: Role of Divalent Cations. *J. Colloid Interface Sci.* **2009**, *338*, 1–9. [[CrossRef](#)]

83. Brydon, J.E.; Sowden, F.J. A Study of the Clay-Humus Complexes of a Chernozemic and a Podzol Soil. *J. Soil Sci.* **2011**, *39*, 136–143. [[CrossRef](#)]
84. Xu, X.; Li, Y.; Hu, X.; Xie, G.; Xu, H.; Gao, M.; Zhang, X.; Zhang, R.; Tang, C.; Hu, X. Effect of Humic Acid on the Adsorption/Desorption Behaviors of Trivalent Chromium on Calcium Modified Montmorillonite and Kaolinite. *ChemistrySelect* **2022**, *7*, e202104302. [[CrossRef](#)]
85. Sadeghi, G.; Behrends, T.; Schijven, J.F.; Hassanizadeh, S.M. Effect of Dissolved Calcium on the Removal of Bacteriophage PRD1 during Soil Passage: The Role of Double-Layer Interactions. *J. Contam. Hydrol.* **2013**, *144*, 78–87. [[CrossRef](#)]
86. McLeod, M.; Aislabie, J.; Smith, J.; Fraser, R.; Roberts, A.; Taylor, M. Viral and Chemical Tracer Movement through Contrasting Soils. *J. Environ. Qual.* **2001**, *30*, 2134–2140. [[CrossRef](#)] [[PubMed](#)]
87. Fernandez-Cassi, X.; Rusinol, M.; Martinez-Puchol, S. Viral Concentration and Amplification from Human Serum Samples Prior to Application of Next-Generation Sequencing Analysis. In *The Human Virome: Methods and Protocols*; Springer: New York, NY, USA, 2018; Volume 1838, pp. 173–188. ISBN 9781493986828.
88. Zhang, Y.Z.; Shi, M.; Holmes, E.C. Using Metagenomics to Characterize an Expanding Virosphere. *Cell* **2018**, *172*, 1168–1172. [[CrossRef](#)]
89. Trubl, G.; Hyman, P.; Roux, S.; Abedon, S.T. Coming-of-Age Characterization of Soil Viruses: A User's Guide to Virus Isolation, Detection within Metagenomes, and Viromics. *Soil Syst.* **2020**, *4*, 23. [[CrossRef](#)]
90. Baer, A.; Kehn-Hall, K. Viral Concentration Determination Through Plaque Assays: Using Traditional and Novel Overlay Systems. *J. Vis. Exp.* **2014**, *93*, e52065. [[CrossRef](#)] [[PubMed](#)]
91. Van Twest, R.; Kropinski, A.M. *Bacteriophages Enrichment from Water and Soil*; Clokie, M.R.J., Kropinski, A.M., Eds.; Humana Press: Totowa, NJ, USA, 2009; Volume 501, ISBN 1402037902.
92. Walshe, G.E.; Pang, L.; Flury, M.; Close, M.E.; Flintoft, M. Effects of PH, Ionic Strength, Dissolved Organic Matter, and Flow Rate on the Co-Transport of MS2 Bacteriophages with Kaolinite in Gravel Aquifer Media. *Water Res.* **2010**, *44*, 1255–1269. [[CrossRef](#)]
93. Furiga, A.; Pierre, G.; Glories, M.; Aimar, P.; Roques, C.; Causserand, C.; Berge, M. Effects of Ionic Strength on Bacteriophage MS2 Behavior and Their Implications for the Assessment of Virus Retention by Ultrafiltration Membranes. *Appl. Environ. Microbiol.* **2011**, *77*, 229–236. [[CrossRef](#)] [[PubMed](#)]
94. Fierer, N.; Breitbart, M.; Nulton, J.; Salamon, P.; Lozupone, C.; Jones, R.; Robeson, M.; Edwards, R.A.; Felts, B.; Rayhawk, S.; et al. Metagenomic and Small-Subunit rRNA Analyses Reveal the Genetic Diversity of Bacteria, Archaea, Fungi, and Viruses in Soil. *Appl. Environ. Microbiol.* **2007**, *73*, 7059–7066. [[CrossRef](#)]
95. Nguyen, N.P.; Warnow, T.; Pop, M.; White, B. A Perspective on 16S rRNA Operational Taxonomic Unit Clustering Using Sequence Similarity. *NPJ Biofilms Microbiomes* **2016**, *2*, 16004. [[CrossRef](#)]
96. Schloss, P.D.; Westcott, S.L. Assessing and Improving Methods Used in Operational Taxonomic Unit-Based Approaches for 16S rRNA Gene Sequence Analysis. *Appl. Environ. Microbiol.* **2011**, *77*, 3219–3226. [[CrossRef](#)]
97. Chen, L.; Xun, W.; Sun, L.; Zhang, N.; Shen, Q.; Zhang, R. Effect of Different Long-Term Fertilization Regimes on the Viral Community in an Agricultural Soil of Southern China. *Eur. J. Soil Biol.* **2014**, *62*, 121–126. [[CrossRef](#)]
98. Doan, T.T.; Bouvier, C.; Bettarel, Y.; Bouvier, T.; Henry-des-Tureaux, T.; Janeau, J.L.; Lamballe, P.; Van Nguyen, B.; Jouquet, P. Influence of Buffalo Manure, Compost, Vermicompost and Biochar Amendments on Bacterial and Viral Communities in Soil and Adjacent Aquatic Systems. *Appl. Soil Ecol.* **2014**, *73*, 78–86. [[CrossRef](#)]
99. Swanson, M.M.; Fraser, G.; Daniell, T.J.; Torrance, L.; Gregory, P.J.; Taliany, M. Viruses in Soils: Morphological Diversity and Abundance in the Rhizosphere. *Ann. Appl. Biol.* **2009**, *155*, 51–60. [[CrossRef](#)]
100. Ateba, C.N.; Akindolire, M.A. Isolation and Characterisation of Bacteriophages with Lytic Activity Against Virulent Escherichia Coli O157:H7: Potential Bio-Control Agents. *Preprints* **2019**, 2019010132. [[CrossRef](#)]
101. Gupta, V.; Saxena, H.M. Isolation and Characterization of BpL1, a Broad Acting Lytic Bacteriophage against Brucella. *Int. J. Curr. Microbiol. Appl. Sci.* **2017**, *6*, 2486–2496. [[CrossRef](#)]
102. Hyman, P. Phages for Phage Therapy: Isolation, Characterization, and Host Range Breadth. *Pharmaceuticals* **2019**, *12*, 35. [[CrossRef](#)]