Zea mays genotype influences microbial and viral rhizobiome community structure

Pooja Yadav[1], Amanda Quattrone[2], Yuguo Yang[1], Jacob Owens[1,a], Rebecca Kiat[1], Thirumurugen Kuppusamy[1], Sabrina E. Russo[1,3], Karrie A. Weber*[1,4,5]

[1]School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE

[2]Complex Biosystems, University of Nebraska-Lincoln, Lincoln, NE

[3]Center for Plant Science Innovation, University of Nebraska-Lincoln, Lincoln, NE

[4]Department of Earth and Atmospheric Sciences, University of Nebraska-Lincoln, Lincoln, NE

[5]Daugherty Water for Food Global Institute, University of Nebraska, Lincoln, NE

Current affiliations:

[a]University of Nebraska-Medical Center, Lincoln, NE

Corresponding Author:
*Karrie A. Weber
232 Manter Hall
University of Nebraska—Lincoln
Lincoln, NE 68588-0118
(402)472-2739
kweber@unl.edu

Competing interests

Authors declare no competing interest.

Abstract

Plant genotype is recognized to contribute to variations in microbial community structure in the rhizosphere, soil adherent to roots. However, the extent to which the viral community varies has remained poorly understood and has the potential to contribute to variation in soil microbial communities. Here we cultivated replicates of two different genotypes of *Zea mays parviglumis* and *Z. mays* genotype B73 in a greenhouse and harvested the rhizobiome (rhizoplane and rhizosphere) to identify the abundance of cells and viruses as well as apply 16S rRNA gene amplicon sequencing and genome resolved metagenomics to identify the rhizobiome microbial and viral community. Our results demonstrate that viruses exceeded microbial abundance in the rhizobiome of *parviglumis* and B73 with a significant variation in both, the microbial and viral community between the two genotypes. Of the viral contigs identified only 4.5% (n =7) of total viral contigs were shared between the two genotypes, demonstrating that plants even at the level of genotype can significantly alter the surrounding soil viral community. An auxiliary metabolic gene associated with glycoside hydrolase (GH5) degradation was identified in one viral metagenome-assembled genome (vMAG) identified in the B73 rhizobiome infecting Propionibacteriaceae (Actinobacteriota) further demonstrating the viral contribution in metabolic potential for carbohydrate degradation and carbon cycling in the rhizosphere. This variation demonstrates the potential of plant genotype to contribute to microbial and viral heterogeneity in soil systems and harbor genes capable of contributing to carbon cycling in the rhizosphere.

Introduction

Plant roots create a carbon and nutrient-rich environment in soils supporting microbiota in the zone adjacent to the root epidermis and the soil adhering to plant roots termed the rhizoplane and rhizosphere or together referred to as the rhizobiome [1]. It is within this region where a diversity of microbiota including Bacteria, Archaea, Eukarya and viruses are recognized as members of the rhizobiome contributing to below ground processes [2]. The microbial diversity of this environment has the potential to be impacted by viral infection where viral mediated cell lysis is recognized to impact microbial abundance in soils [3]. Viruses are abundant in soils ranging from 10^7 to 10^9 viruses g^{-1} of soil [4-6] and virus abundance can often exceed that of microbes [7]. To date a few studies have virus abundance and predation in the rhizosphere [8, 9], but the impact of these processes in this dynamic region in soils remains poorly studied.

Virus infection can result in host cell lysis through a lytic or lysogenic infection thereby, reducing host cell abundance. The lysis of these microbes results in the release of cell lysate which serves as an easily accessible carbon and nutrient source supporting the growth of other microbial members [10]. Hence, viral infections of microbial host cells can influence the metabolic functions of the microbial community. This may occur due to shifts in taxonomic composition and change in metabolic potential caused by differential viral predation or through the acquisition of auxiliary metabolic genes (AMGs) [11-15] resulting from viral infection. AMGs are viral acquired genes from their host, not required in viral replication but allow viruses to directly manipulate host metabolism [16]. Some well-known examples of AMGs acquired by viruses have been inferred to play a role carbon [15], sulfur [17], and nitrogen [18] cycling. Together, these changes in relative abundance of microbial composition and introduction of functional genes like AMGs can alter the metabolic potential of microbial community by metabolically reprogramming their hosts, and/or

expression of AMGs [9, 15, 19, 20]. While, the abundance of viruses in the rhizosphere has been reported before [8, 9, 21], the implication of viral impacts on the host abundance and metabolic potential remains understudied. Given the absence of universal marker genes across viruses, metagenomic studies are required to shed light on the role of viruses in rhizobiome and, how viruses impact the abundance of the host cells, and their potential to reprogram host metabolism [22].

In the carbon-rich region surrounding the plant root, the rhizobiome community structure varies between plant species [23] and genotype [24], owing to the influence of the plant, including variation in root exudate profiles [25]. The variation in root exudate profile secreted by host genotype and the resulting substrate-driven or bottom-up shifts in the microbial community structure are based on differential production of root exudates and community utilization [26, 27]. In addition to substrate driven controls on rhizobiome structure, viruses are recognized to infect microorganisms in the rhizosphere [8, 9]. Viruses require metabolically active hosts for replication and thus the patterns of viral communities are linked to active host cells [28-31]. While physiochemical heterogeneity in soils can alter viral abundance where it is recognized that soil moisture and pH can influence viral community structure [32-34], variation in plant roots also has the potential to contribute to viral community structure. In soils viruses impact host abundance and metabolic potential by infecting metabolically active organisms driving carbon-cycling [15, 35], and play a role in biogeochemical processes in the rhizosphere [9, 36]. The region surrounding the root is one are where microorganisms are recognized to be metabolically active owing to the production of carbon-rich plant root exudates.

Here we test the hypothesis that the microbial and viral rhizobiome community varies between two *Z. mays* genotypes, *Z. mays ssp. parviglumis* (teosinte) and *Z. mays* genotype B73 (domesticated corn). Seeds of each genotype were sterilized, germinated under axenic conditions and transplanted into a homogenized soil matrix and cultivated to the V4-V8 growth stage in a greenhouse under controlled conditions. Plants were harvested for phenotyping and rhizobiome samples were collected to enumerate cells and viruses as well as identify microbial community members using 16S rRNA gene amplicon sequencing and genome resolved metagenomics from shotgun metagenome sequence data. Variation within the microbial and viral communities between *Z. mays* genotypes was statistically determined. The virus linked microbial hosts in the rhizobiome of the two genotypes rendered differences in host abundance. Assembled and annotated viral contigs were screened for functional genes defined as AMGs and included the potential to degrade carbohydrates. Here we report differential microbial and viral community assemblage in the rhizobiome between two *Z. mays* genotypes, *parviglumis* and B73.

Materials and methods

Seed source, germination, and Z. mays growth

Seeds of *Z. mays parviglumis* and B73 were obtained from the University of Nebraska, Lincoln in-house collection [37]. Twenty-four *parviglumis* and B73 seeds were sterilized and germinated in sterilized petri-plates. From among the 24 seedlings, 12 seedlings were transplanted into 1.8 L pots containing homogenized sieved prairie soil mixed with autoclaved agricultural-field sandy-soil (60:40; mass/mass; Supplementary Information) [38]. Plants were cultivated in a greenhouse with an average 13-hour day length for 22 days at 31°C with relative humidity between

60.2 - 97.6 %. Every three days plants were watered and fertilized with 25 mL of 25% Hoagland's solution [39]. Six biological replicates at the same growth stage (V4-V8) were selected for harvesting.

Rhizobiome harvesting and sample collection

Plant roots and rhizobiome were harvested by gently cutting the pot to minimize soil and root disruption and carefully separated from bulk soil as previously described [37]. The roots were submerged in a known volume of sterile, nucleic acid and enzyme-free 1X phosphate buffered saline (PBS) and sonicated (Branson 450D, 30% amplitude, 0.3 s duty cycle) for one minute, three times at 30 seconds intervals [40, 41] to harvest the rhizobiome, rhizoplane and rhizosphere. The rhizobiome suspension was centrifuged (800 X g). A known volume of the supernatant was filtered through 0.45 µm PVDF and 0.2 µm PVDF membrane, and fixed with formaldehyde (0.5 v/v final concentration) for cell and virus enumeration [42, 43]. The remaining suspension was centrifuged (3000 X g for 10 minutes) and supernatant was strained through a 40 mm sterile cell strainer (Fisherbrand™ product #22363547) to remove plant cells and root debris. The supernatant was collected on 0.2 µm polycarbonate membranes (product #GTTP04700) through vacuum filtration (≤ 2.5 bar). The rhizosphere pellet was collected by another round of centrifugation at 10,000 X g for 20 minutes. The membrane and pellet were flash frozen in liquid nitrogen and stored at -80 °C. Following rhizobiome collection root phenotype characteristics were determined (Supplementary Information).

Nucleic acid extraction, 16S rRNA gene amplicon, and shotgun metagenomic sequence library preparation

All reagents and materials used in sample processing were treated or prepared in 0.1 % diethylpyrocarbonate or RNase Away. Nucleic acids were extracted from a known mass of the rhizosphere pellet (10% w/w of total rhizosphere and rhizoplane) and the membrane collected cells using a modified Griffith's method as previously described [43, 44]. DNA was quantified and purity was checked as previously described (Westrop et al. 2023). A quantified mass of DNA extracted from the pellet and the filter (2.5 % of total DNA (ng) were proportionally pooled to represent the total rhizobiome community. The DNA quantity was normalized per gram rhizosphere and sent for sequencing.

The 16S rRNA gene amplicon sequencing library was prepared by the University of Minnesota Genomics Center, Minnesota using V4 and V5 regions of the 16S rRNA gene primers 515F and 806R as described [45, 46]. The samples were paired-end sequenced (2 X 300 bp) using Illumina MiSeq600 cycle v3 kit to a sequencing depth of 72,000 reads per sample. The 16S rRNA gene amplicon sequence data was analyzed using the DADA2 package to obtain amplicon sequence variants (ASVs) as described [47] and calculate diversity measures (Supplementary Information). The ASVs were filtered for spurious sequences as described and was analyzed using Phyloseq [48].

The shotgun metagenome sequence library was prepared using the Nextera XT DNA library kit (Illumina, California, USA) as previously described [43]. The average final library size was 300 bp with an insert of 180 bp and a read length configuration of 150 PE for 40 M PE reads per sample (20 M in each direction). Sequence data were quality checked (Supplementary Information) and clean reads were used to generate de novo co-assemblies using Megahit [49].

Assembled contigs were binned using MetaBat2 (v2.15) [50], MaxBin2 (v2.2) [51], and CONCOCT (v1.1) [52] and dereplicated using DAS_Tool (v1.1) [53]. The bins were quality checked using CheckM (v1.1) [54] and manually curated using Anvi'o (v7.0) [55] as described [43]. The final set of MAGs were taxonomically identified and used to estimate relative abundance as described (Supplementary Information).

Putative viral contigs were identified using VirSorter2 (v2.1) [56] and DeepVirFinder [57] (Supplementary Information). A pairwise comparison with ≥ 95 % average nucleotide identity (ANI) across ≥ 85 % alignment coverage [58, 59] was used to assess genome similarity in the viral community of *parviglumis* and B73. Viral contigs ≥ 10 Kbp are described as a vMAG [60]. Viral genome coverage was normalized as performed for MAGs and used as a proxy for relative abundance as previously described [15, 58, 61].

Virus-host linkages were established utilizing a combination of three different methods, clustered regularly interspaced short palindromic repeats (CRISPR) sequences [62], genome similarities [63], and tetranucleotide frequencies (threshold distance < 0.001) between host and viral contig (Supplementary Information) [15, 63-66]. Host genome metabolic potential determined using METABOLIC (v4.0) [67], DRAM (v1.2) [68], and FeGenie (v1.0) [69] at the MAG level. Viral contribution in the host metabolic processes was investigated by screening viral encoded AMGs in *parviglumis* and B73 using DRAM-v.

Statistical analyses

All statistical analyses were conducted in R version 4.1.2 (2021-11-01). Differences between *Z. mays* genotype in bacterial and virus abundances were tested using t-tests assuming

unequal variances. The correlation of cells and virus abundances were analyzed using Mantel's test and linear regression (Supplementary Information). Relationships between the microbial and viral community structure were analyzed using principal coordinate analysis (PCoA) with Bray-Curtis dissimilarities as implemented in the vegan package [70]. Differences between *Z. mays* genotypes in the microbial and viral community composition were evaluated using permutational multivariate analysis of variance (perMANOVA) using adonis2.

Results

Virus to cell ratio in the parviglumis and B73 rhizobiome

Virus abundance exceeded cell abundance in the rhizobiome collected from both *parviglumis* and B73 (VCR of 2.24 ± 0.21 and 1.52 ± 0.12, respectively) (Fig. 1) and was positively correlated with cell abundance (*parviglumis* $p = 0.001$; B73 $p = 0.008$). Cell abundance in the *parviglumis* and B73 rhizobiome totaled (1.34 ± 0.30) x 10^7 cells g^{-1} and (1.54 ± 0.16) x 10^7 cells g^{-1} respectively but was not statistically different between genotypes ($p > 0.05$, Fig. 1a). Similarly, *parviglumis* and B73 rhizobiome virus abundance of *parviglumis* (2.89 ± 0.55) x 10^7 g^{-1} and B73 (2.28 ± 0.20) x 10^7 g^{-1} rhizosphere (Fig. 1b) exceeded cell abundance and was not statistically different between genotypes ($p > 0.05$). Cell and virus abundance did not differ significantly between genotypes, even when standardized by plant phenotypic parameters including root length, surface area, volume, density, fresh weight, or biomass (Supplementary Fig. S1 and Fig. S2). Cell abundance gram^{-1} of rhizobiome collected from *parviglumis* and B73 was positively correlated with the rhizosphere mass. While cell or virus abundance gram^{-1} of rhizobiome collected was not

statistically different between genotypes, the VCR significantly varied ($p = 0.02$) between genotypes indicating variation between genotypes.

Microbial community composition of parviglumis and B73 rhizobiome

The rhizobiome microbial community collected from *parviglumis* and B73 revealed 6,317 unique ASVs from clean reads of *parviglumis* (range 34,397 to 96,390) and B73 (range 29,779 to 75,780). The ASVs were broadly classified into 1 archaeal phylum and 17 bacterial phyla (Fig. 2a). The core rhizobiome consisted of Verrucomicrobiota, Proteobacteria, and Actinobacteriota with an abundance in similar proportions accounting for a combined 72.64% and 74.97% of the rhizobiome community in *parviglumis* and B73, respectively. The microbial community identified by 16S rRNA gene amplicon sequencing was statistically different between genotypes. (Fig. 2b; $R^2 = 0.098$, $p = 0.001$) where *Z. mays* genotype accounted for 10.8% of the variation in the microbial community structure (Fig. 2b).

Genome assembly of metagenomic sequence data resulted in MAGs similar to taxa identified at ASV level where reads represented more than 2.5% of the microbial community (Fig. 2c). Biological replicates of the shotgun metagenome sequenced data yielded clean reads in the range of 22.85 to 24.12 million for *parviglumis* and 23.35 to 27.30 million for B73. These data rendered 50 and 61 rhizobiome metagenome-assembled genomes (MAGs) for *parviglumis* and B73, respectively. The *parviglumis* rhizobiome MAGs averaged 42.91 ± 2.49% completeness, totaling 15 medium-quality MAGs and 35 low-quality MAGs and B73 MAGs averaged estimated completeness of 43.60 ± 2.23%, comprising one high-quality, 19 medium-quality, and 41 low-quality MAGs. Nine and 11 phyla in the archaeal and bacterial domains were identified in *parviglumis* and B73, respectively (Fig. 3). Similar to 16S rRNA sequencing data, the

Thermoproteota was the only archaeal MAG constructed from the rhizobiome of both genotypes. The bacterial MAGs reconstructed from *parviglumis* rhizobiome were Acidobacteriota, Actinobacteriota, Bacteroidota, Desulfobacterota, Gemmatimondadota, Patescibacteria, Planctomycetota, Verrucomicrobiota, and unclassified bacteria. In addition to these taxa, two additional phyla, Chloroflexota and Desulfobacterota, were also reconstructed from B73 rhizobiome (Fig. 2c and 3).

As observed in 16S rRNA gene sequence data, taxa identified from MAGS revealed variance in the microbial community structure between genotypes (Fig. 2d; $p = 0.001$, $R^2 = 0.99$). The PCoA revealed 88.1% and 2.3% of the variance on axis 1 and 2, respectively (Fig. 2d). High variability of PCoA axis 1 demonstrated the variation in beta diversity between the two genotypes. This further highlighted the differences in the microbial community structure of *Z. mays* genotypes.

Similar to the 16S rRNA gene amplicon sequence data, MAGs also identified Proteobacteria and Actinobacteriota as the two most abundant rhizobiome phyla of both *parviglumis* and B73 rhizobiome (Fig. 3). Analyses of the B73 rhizobiome 16S rRNA gene amplicon data revealed that in addition to taxa identified in the Proteobacteria and Actinobacteriota, Verrucomicrobiota taxa were abundant community members (Fig. 2a, 2c, and 3). Within the Proteobacteria, community members in the family Xanthobacteraceae (18.00%) were the most abundant in the *parviglumis* rhizobiome. This result is different relative to the most abundant taxa identified at the family level in the B73 rhizobiome, where taxa within the Chthoniobacterales UBA10450 in the Verrucomicrobiota (16.39%) were the most abundant in B73 followed by Xanthobacteraceae (13.11%), Chloroflexota CSP1-4 (1.62%) and Binatia UBA 9968

in Desulfobacterota (1.63%) (Fig. 2c and 3). In addition to microbial taxa, viral contigs were also identified from the shotgun metagenome sequence data (Fig. 3).

Rhizobiome viral community structure variation between parviglumis and B73

Identification of putative viruses from rhizobiome shotgun metagenome sequence data detected 329 and 488 viral contigs in *parviglumis* and B73, respectively. Further screening of viral contigs yielded 67 and 86 contigs with in the rhizobiome of *parviglumis* and B73, respectively. The average nucleotide identity (ANI) distribution not only displayed a high level of strain heterogeneity but also revealed a wide range of viral genome variations. The genome size of putative viruses identified in the parviglumis rhizobiome ranged from 1.50 – 40.54 Kbp with 5 vMAGs and B73 viral contigs ranged from 1.51 – 31.18 Kbp with only 1 vMAG reconstructed. Viral contig clustering between two rhizobiomes resulted in the identification of only 7 (4.5%) viral contigs identified in both *parviglumis* and B73 (2 clustered > 3 kb and 5 clustered < 3kb) (Fig. 4a). Rhizobiome viral community structure collected from each genotype was significantly different ($R^2 = 0.999, p = 0.005$) (Fig. 4b).

Taxonomy of rhizobiome viral contigs could be assigned to 38.80% in *parviglumis* and 30.23% in B73 viral contigs. Putative viruses were taxonomically classified as Caudovirales and among the annotated viral contigs, the family *Siphoviridae* (61.53%), *Myoviridae* (15.38%), and *Podoviridae* (11.53%) were the most abundant in the viral community of the *parviglumis* rhizobiome. The *Siphoviridae* (76.92%), *Podoviridae* (11.53%), and *Myoviridae* (7.69%) were prevalent in the B73 rhizobiome. Contigs classified in the families of *Ackermannviridae* (3.84%), *Autographviridae* (3.84%), and *Chaseviridae* (3.84%) were unique to *parviglumis* (Fig. 4c). One unclassified archaeal dsDNA virus, Halovirus (3.84%) was identified in the rhizobiome of B73

(Fig. 4c) but was not identified in *parviglumis* rhizobiome. Over two-thirds of the viral contigs identified remained unclassified and could not be assigned to a specific viral taxa.

Virus-host linkages

Virus-host linkages could be established for 31.34% of *parviglumis* and 36.04% of B73 viral contigs. Phyla Thermoproteota, Acidobacteriota, Actinobacteriota, and Patescibacteria were commonly identified in both genotypes and were associated with the highest number of viral contigs in *parviglumis* and B73 rhizobiome. Few viral contigs were coarsely associated with unclassified bacteria (4.76% in *parviglumis* and 9.67% in B73). Thermoproteota family Nitrososphaeraceae was the only archaeal host identified in *Z. mays* rhizobiome. Of the identified hosts, 19.04% and 9.67% of viral contigs were linked to Nitrososphaeraceae in *parviglumis* and B73, respectively (Fig. 5a). Among the bacterial taxa Vicinamibacterales UBA2999, Micrococcaceae, Saccharimonadaceae, Xanthobacteraceae, and some unclassified bacteria were identified as common hosts in *parviglumis* and B73 (Fig. 5a). Viruses linked to Acidobacteriae UBA2999, Koribacteraceae, and Streptomycetaceae were unique to *parviglumis*. Rhizobiaceae, Propionibacteriaceae, Binatia UBA9968, Alphaproteobacteria UBA1301, Sphingomonadaceae, Chthoniobacterales UBA10450 were unique hosts in the rhizobiome of B73. None of the viral contigs were found to be infecting more than a single MAG in an established virus-host correlation indicating the virus host specificity at the MAG level. Although Verrucomicrobiota member Chthoniobacterales UBA10450 was highly abundant in the rhizobiome of both genotypes, it served as a virus host only in the B73 rhizobiome. Together these results indicate the variation of the community in the rhizobiome with host specificity.

A total of 64 and 98 putative CRISPR spacer regions were detected in *parviglumis* and B73 infected MAGs, respectively. Among all the identified CRISPR spacer sequences in both genotypes, only a single CRISPR spacer sequence provided a positive hit to one of B73 MAGs. This corresponded to B73 MAG Xanthobacteraceae VAZQ01 sp005883115 with a 100% nucleotide identity match. The direct repeats flanking the CRISPR spacer region compared against the infected MAGs rendered an e-value $< 10^{-10}$ confirming infection in Xanthobacteraceae VAZQ01 sp005883115 in B73 rhizobiome.

The average virus to host abundance ratio ranged from 0.41 to 8.38 in rhizobiome *parviglumis* and B73. Viruses infecting the Micrococcaceae (Actinobacteriota) were lower in abundance than hosts in rhizobiome collected from both genotypes. Viruses infecting Koribacteraceae (Acidobacteriota) in *parviglumis* and Chthoniobacterales UBA10450 (Verrucomicrobiota) in B73 were also lower than host cell. Virus abundance equaled host cell abundance of Vicinamibacterales UBA2999 (Acidobacteriota) in *parviglumis* only. Rhizobiome viruses exceeded the host cells, Nitrososphaeraceae (Thermoproteota), Xanthobacteraceae (Proteobacteria), and Saccharimonadaceae (Patescibacteria) in both genotypes(Fig. 5b).

Metabolic potential of microbial and viral rhizobiome

The metabolic potential of the microbial community revealed the ability of the community to degrade carbohydrates as well as short chain fatty acids (Fig. 6). Bacterial hosts such as Xanthobacteraceae (Proteobacteria) not only showed the capability to degrade polyphenols, starch, short chain fatty acids, alcohols, chitins, and acetate utilization but were also capable of central carbon, nitrogen, sulfur, and iron metabolism (Fig. 6).

Host cells Acidobacteriota (Acidobacteriae UBA7541, Vicinamibacterales UBA2999, Koribacteraceae) and Actinobacteriota (Micrococcaceae, Propionibacteriaceae) harbored genes

that associated with polysaccharide degradation enabling central carbon metabolisms such as glycolysis and the Kreb's cycle. Interestingly, the metabolic potential for CO_2 fixation, utilizing the dicarboxylate-hydroxybutyrate cycle, hydroxypropionate-hydroxybutylate cycle, and reductive citrate cycle (Arnon-Buchanan cycle) were identified in MAGs corresponding to archaea Thermoproteota and bacteria Actinobacteriota, Bacteriodota, Proteobacteria, Verrucomicrobiota in both genotypes. Genes associated with the Calvin Cycle were not identified in any of the MAGs. The metabolic potential to degrade polyphenol and cleavage of arabinan and fucose cleavage was restricted to the Koribacteraceae identified in the *parviglumis* rhizobiome (Fig. 6a). The ability to degrade chitin was specific to Desulfobacterota (Binatia UBA9968) and Verrucomicrobiota (Chthoniobacterales UBA10450) with the potential to degrade beta-mannan in the B73 rhizobiome (Fig. 6b). Archaea Thermoproteota indicated the metabolic potential to convert acetate into methane, carbohydrate degradation and detoxification of aromatic compounds, and ammonia oxidation. Additionally, *parviglumis* MAGs Thermoproteota (Nitrososphaeraceae) and Proteobacteria (Xanthobacteraceae) were capable of denitrification. No B73 MAGs showed any role in nitrification and similarly, no *parviglumis* or B73 viral infected MAGs showed any role in sulfur metabolism. The microbial community in both the *parviglumis* and B73 viral infected MAGs showed a role in siderophore transport to support iron acquisition in soils.

One auxiliary metabolic gene was identified in a viral contig collected from the rhizobiome of B73 whereas AMGs were not identified in viral contigs recovered from *parviglumis*. Analysis of the viral encoded AMG identified this gene as a glycoside hydrolase family 5 (GH5) capable of degrading amorphous cellulose, xyloglucan, xylan, beta-mannan, mixed-linkage glucans backbone cleavage, and chitin (Fig. 6b). This virus was linked to the host Propionibacteriaceae

(Actinobacteriota) in the B73 rhizobiome. These results indicate that rhizobiome viruses can harbor genes with the potential to produce labile mono and oligosaccharides.

Discussion

Here we demonstrated that both microbial and viral community structure varies between two *Z. mays* genotypes, *parviglumis* and B73, cultivated in the same homogenized soil under the same environmental conditions. While it is recognized the *Z. mays* genotype influences the rhizosphere microbial community [71], to our knowledge this is the first study to demonstrate the variation of the rhizobiome viral community structure between plant genotypes. Plants release root exudates below ground, creating a diverse chemical milieu that impacts rhizobiome assemblage and structure [46, 72, 73]. The root exudate composition is also recognized to vary among maize genotypes [25], thus further promoting variation in the rhizosphere community. Our results are consistent with prior research demonstrating *Z. mays* genotype plays a factor influencing the selection of rhizobiome microbial community [71]. Here we identified Proteobacteria, Actinobacteriota, and Verrucomicrobiota constituted the core microbiome of *parviglumis* and B73 consistent with a prior study [46] and were more abundant relative to the Acidobacteriota and Desulobacterota. Interestingly, Acidobacteriota were associated with more viral contigs which suggests an active infection could result in lower cell abundance. The relatively lower abundant microbial taxa in *Z. mays* rhizobiome such as Chloroflexota and Desulfobacterota (relative abundance < 2.5% in 16S rRNA gene data) were reconstructed in B73, but we were unable to recover MAGs from *parviglumis*. This could be a result of low relative abundance and/or lysed host cells not detected in sequencing. Alternatively the result could also be a result of differential root-exudate production between genotypes [25] and recruitment of these taxa to the root. Microbial community structure and viral host cells are recognized to influence viral community

structure [31, 74, 75]. Small but significant differences in rhizobiome microbial community composition in *parviglumis* and B73 manifested dramatic differences in the viral community of these two *Z. mays* genotypes.

Viruses were abundant in the rhizobiome of *parviglumis* and B73 rhizobiome outnumbering cells. The virus to cell ratio resulted in a VCR >1.5, 3 times more than previously reported in the wheat rhizosphere (VCR = 0.27) [8]. Viral contigs recovered from the shotgun metagenome sequence data supported variation of the viral community between genotypes revealing the presence of a distinct viral community in the rhizobiome of *parviglumis* and B73. We recognize that members of the viral community composition reported in this study may represent an underestimate of viral diversity in the *Z. mays* rhizobiome due to sampling constraints because reported data in this study are limited to viruses associated with the host genome, adsorbed to particulate matter, and greater than 0.2 μm in size. Prior research has demonstrated significant variations in both microbial and viral community due to spatial distribution and/or soil heterogeneity [59, 76, 77], the close clustering of the *Z. mays* biological replicates on PCoA plot demonstrated that soil heterogeneity among replicates didn't contribute to significant differences in microbial and viral community structure. Rather the genotype of the plant played a significant role in the pronounced variation in the rhizobiome viral community where only seven viral contigs were shared between the two plant genotypes. This finding is also supported by pairwise comparison clustering of viral contigs which yielded only 7 similar viral contigs and validated that *parviglumis* and B73 comprised contrastingly different viral communities while grown in the same conditions using the same soil inoculum. Thus in addition to soil physiochemical factors, plant species and genotype can further contribute to the spatial variation of viral communities observed in soils.

Viral community variation in parviglumis and B73

Virus-host specificity and host range abundance play an important role in viral selection in the rhizosphere [8]. The viral contigs recovered from the *parviglumis* and B73 rhizobiome revealed host specificity. All viral contigs recovered were found infect a unique MAG. The strong correlation between cell and virus abundance is interconnected through prey-predator interactions. Repeated viral infections and host cell lysis give rise to virus-defense through immunity systems including the CRISPR-cas system found within some bacterial and archaeal lineages, thus, avoiding future viral infections [78]. The presence and 100% identity match to the CRISPR sequence in B73 MAG suggested the adaptation of a defense mechanism against viral infection by Xanthobacteraceae. This active defense mechanism of bacteria exploiting CRISPR sequences could be a possible explanation for the lower number of viral contigs in Xanthobacteraceae MAGs hence, the high abundance of Xanthobacteraceae in rhizobiome. The small genome size of the ultra-small bacteria Saccharimonadaceae (Patescibacteria) may not be able to comprise CRISPR sequences thus rendering the cell more susceptible to viral infection [79, 80]. However, the Patescibacteria adopted the strategy to evade viral infection by deleting common membrane phage protein receptors [81]. Additionally, the choice of viral reproduction strategy between lytic and lysogeny can affect the host community dynamics resulting in distinct fates of host cells [4, 82, 83]. Viral host specificity can thus play a role influencing microbial diversity [75]. Viral predation results in altered community structure [84], can increase the release biologically available nutrients, and can consequently impact ecosystem productivity [85]. Here we observed rhizobiome viral contigs linked to the host cells which demonstrated host preference which could subsequently impact host abundance as demonstrated by the variation in the ration of viruses abundance to host

abundance. Viral infection found to be impacting host metabolic output not only by impacting host abundance but also through acquiring glycoside hydrolases AMG capable of complex polysaccharide degradation [15, 86].

Impact of virus-host interactions on metabolic potential

Viruses that exhibit a lytic life cycle cause cell-mediated cell lysis liberating labile sources of carbon and nutrients [87] which can stimulate microbial population growth from the bottom up potentially altering community diversity and metabolic output [88, 89]. In a stable isotope probing (SIP) study performed in soils, the phyla Acidobacteriota, Actinobacteriota, Chloroflexota, Proteobacteria, Patescibacteria, Verrucomicrobiota were identified as active viral hosts involved in soil carbon cycling [80]. Taxa identified in our study were also identified to harbor the metabolic potential to drive carbon cycling near plant roots. Acidobacteriota (Acidobacteriae UBA7541, Vicinamibacterales UBA2999, Koribacteraceae), Actinobacteriota (Micrococcaceae, Propionibacteriaceae, Streptomycetaceae), Patescibacteria (Saccharimonadaceae), Proteobacteria (Alphaproteobacteria UBA1301, Sphingomonadaceae, Rhizobiaceae, Xanthobacteraceae), Verrucomicrbiota (Chthoniobacterales UBA10450), with the addition archaeon Thermoproteota (Nitrososphaeraceae) are involved in central carbon metabolism. In addition to carbon biogeochemical cycling, members of the rhizobiome of both *parviglumis* and B73 had the potential for siderophore transport to support iron acquisition in soils. We did not detect the metabolic potential for nitrification and sulfur metabolism within taxa that were infected by viruses.

Viral encoded glycoside hydrolases AMGs have the potential to contribute to carbon metabolism in *Z. mays* rhizobiome through expression within the host cells. Here we observed the unclassified vMAG in the B73 rhizosphere that contained a glycoside hydrolase (GH5) which is

recognized to degrade amorphous cellulose, beta-mannan, chitin, mixed-linkage glucans, xylans, and xyloglucan. The acquisition of this AMG, glycoside hydrolases responsible for carbohydrate hydrolysis, by soil viruses is consistent with prior studies [15, 35, 80, 90] and demonstrates the potential within the rhizobiome. This vMAG that was linked to host Propionibacteriaceae (Actinobacteriota) that respond to a variety of carbon sources by secreting extracellular enzymes such as cellulases, amylases, and chitinases were found to be regulating carbohydrate catabolic pathways [91, 92]. These results indicate that virus-encoded glycoside hydrolases have the potential to contribute to the production of labile mono and oligosaccharides which can further serve as a carbon source for the host and other rhizobiome community members.

Conclusion

Together these results indicated that plant genotype plays a role in microbial assemblage in rhizobiome and viral community parallelly shifts in response to microbial community selection. Viral community structure is linked to compositional patterns of microbial communities based on plant genotypes. Thus, variation in below ground plant roots has the potential significantly influence heterogeneity of viruses within soils. The virus-host interaction showed differences in host abundance between two genotypes which has the capability to influence the metabolic potential of the rhizobiome. Viral AMGs are capable of carbohydrate degradation which have the potential to produce labile mono and oligosaccharides in the rhizosphere. These labile carbon sources could further promote microbial community proliferation. Therefore, to extent to which host-virus interaction alters host abundance, microbial community structure, and functions remains a critical question in rhizosphere microbial ecology and soil biogeochemical cycling.

Data availability

Experimental data and results summary have been uploaded as a supplement. DNA sequence data and corresponding metadata has been deposited to the NCBI under BioProject accession number PRJNA955173.

Author's contribution

PY and KAW designed the study. PY, AQ, YY, TK, SER, and KAW developed the data collection protocols. PY, AQ, YY, TK, JO, RK, SER, and KAW conducted the experiment and collected the data. PY and KAW analysed the data, with help from SER and AQ, PY and KAW wrote the initial draft. All authors edited the manuscript.

Competing interests

Authors declare no competing interest.

References

[1] Oburger E, Schmidt H. New methods to unravel rhizosphere processes. *Trends Plant Sci*. 2016;21**:** 243-255.

[2] Rout ME, Southworth D. The root microbiome influences scales from molecules to ecosystems: the unseen majority. *Am J Bot*. 2013;100**:** 1689-1691.

[3] Pratama AA, Terpstra J, de Oliveria ALM, Salles JF. The role of rhizosphere bacteriophages in plant health. *Trends Microbiol*. 2020;28**:** 709-718.

[4] Williamson KE, Fuhrmann JJ, Wommack KE, Radosevich M. Viruses in Soil Ecosystems: An Unknown Quantity Within an Unexplored Territory. *Annual review of virology*. 2017;4**:** 201-219.

[5] Trubl G, Roux S, Solonenko N, Li YF, Bolduc B, Rodriguez-Ramos J *et al*. Towards optimized viral metagenomes for double-stranded and single-stranded DNA viruses from challenging soils. *PeerJ*. 2019;7**:** e7265.

[6] Williamson KE, Corzo KA, Drissi CL, Buckingham JM, Thompson CP, Helton RR. Estimates of viral abundance in soils are strongly influenced by extraction and enumeration methods. *Biol Fert Soils*. 2013;49**:** 857-869.

[7] Williamson KE, Radosevich M, Wommack KE. Abundance and diversity of viruses in six Delaware soils. *Appl Environ Microbiol*. 2005;71**:** 3119-3125.

[8] Swanson MM, Fraser G, Daniell TJ, Torrance L, Gregory PJ, Taliansky M. Viruses in soils: morphological diversity and abundance in the rhizosphere. *Annals of Applied Biology*. 2009;155**:** 51-60.

[9] Bi L, Yu DT, Du S, Zhang LM, Zhang LY, Wu CF *et al*. Diversity and potential biogeochemical impacts of viruses in bulk and rhizosphere soils. *Environ Microbiol*. 2021;23**:** 588-599.

[10] Middelboe M, Jørgensen NO. Viral lysis of bacteria: an important source of dissolved amino acids and cell wall compounds. *J Mar Biol Assoc UK*. 2006;86**:** 605-612.

[11] Fuhrman JA. Marine viruses and their biogeochemical and ecological effects. *Nature*. 1999;399**:** 541-548.

[12] Suttle CA. Marine viruses--major players in the global ecosystem. *Nat Rev Microbiol*. 2007;5**:** 801-812.

[13] Wommack KE, Colwell RR. Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev*. 2000;64**:** 69-114.

[14] Torsvik V, Ovreas L, Thingstad TF. Prokaryotic diversity--magnitude, dynamics, and controlling factors. *Science (New York, NY)*. 2002;296**:** 1064-1066.

[15] Emerson JB, Roux S, Brum JR, Bolduc B, Woodcroft BJ, Jang HB *et al*. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat Microbiol*. 2018;3**:** 870-880.

[16] Hurwitz BL, Hallam SJ, Sullivan MB. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol* 2013;14**:** R123.

[17] Kieft K, Zhou Z, Anderson RE, Buchan A, Campbell BJ, Hallam SJ *et al*. Ecology of inorganic sulfur auxiliary metabolism in widespread bacteriophages. *Nat Commun*. 2021;12**:** 1-16.

[18] Wang S, Yang Y, Jing J. A Synthesis of Viral Contribution to Marine Nitrogen Cycling. *Front Microbiol*. 2022**:** 1452.

[19] Wu R, Smith CA, Buchko GW, Blaby IK, Paez-Espino D, Kyrpides NC *et al*. Structural characterization of a soil viral auxiliary metabolic gene product–a functional chitosanase. *Nat Commun*. 2022;13**:** 5485.

[20] Zheng X, Jahn MT, Sun M, Friman V-P, Balcazar JL, Wang J *et al*. Organochlorine contamination enriches virus-encoded metabolism and pesticide degradation associated auxiliary genes in soil microbiomes. *ISME J*. 2022;16**:** 1397-1408.

[21] Li Y, Sun H, Yang W, Chen G, Xu H. Dynamics of bacterial and viral communities in paddy soil with irrigation and urea application. *Viruses*. 2019;11**:** 347.

[22] Edwards RA, Rohwer F. Viral metagenomics. *Nat Rev Microbiol*. 2005;3**:** 504-510.

[23] Kuske CR, Ticknor LO, Miller ME, Dunbar JM, Davis JA, Barns SM *et al*. Comparison of soil bacterial communities in rhizospheres of three plant species and the interspaces in an arid grassland. *Appl Environ Microbiol*. 2002;68**:** 1854-1863.

[24] Semchenko M, Xue P, Leigh T. Functional diversity and identity of plant genotypes regulate rhizodeposition and soil microbial activity. *New Phytologist*. 2021;232**:** 776-787.

[25] Lopez-Guerrero MG, Wang P, Phares F, Schachtman DP, Alvarez S, van Dijk K. A glass bead semi-hydroponic system for intact maize root exudate analysis and phenotyping. *Plant methods*. 2022;18**:** 1-21.

[26] Paterson E, Gebbing T, Abel C, Sim A, Telfer G. Rhizodeposition shapes rhizosphere microbial community structure in organic soil. *New Phytol*. 2007;173**:** 600-610.

[27] Bulgarelli D, Schlaeppi K, Spaepen S, Van Themaat EVL, Schulze-Lefert P. Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol*. 2013;64**:** 807-838.

[28] Srinivasiah S, Lovett J, Ghosh D, Roy K, Fuhrmann JJ, Radosevich M *et al*. Dynamics of autochthonous soil viral communities parallels dynamics of host communities under nutrient stimulation. *FEMS microbiology ecology*. 2015;91.

[29] Ashelford KE, Day MJ, Bailey MJ, Lilley AK, Fry JC. In situ population dynamics of bacterial viruses in a terrestrial environment. *Appl Environ Microbiol*. 1999;65**:** 169-174.

[30] Ashelford KE, Norris SJ, Fry JC, Bailey MJ, Day MJ. Seasonal population dynamics and interactions of competing bacteriophages and their host in the rhizosphere. *Appl Environ Microbiol*. 2000;66**:** 4193-4199.

[31] Brum JR, Ignacio-Espinoza JC, Roux S, Doulcier G, Acinas SG, Alberti A *et al*. Patterns and ecological drivers of ocean viral communities. *Science* 2015;348**:** 1261498.

[32] Santos-Medellin C, Estera-Molina K, Yuan M, Pett-Ridge J, Firestone MK, Emerson JB. Spatial turnover of soil viral populations and genotypes overlain by cohesive responses to moisture in grasslands. *P Natl Acad Sci USA*. 2022;119**:** e2209132119.

[33] Trubl G, Solonenko N, Chittick L, Solonenko SA, Rich VI, Sullivan MB. Optimization of viral resuspension methods for carbon-rich soils along a permafrost thaw gradient. *PeerJ*. 2016;4**:** e1999.

[34] Lee S, Sorensen JW, Walker RL, Emerson JB, Nicol GW, Hazard C. Soil pH influences the structure of virus communities at local and global scales. *Soil Biol Biochem* 2022;166**:** 108569.

[35] Trubl G, Jang HB, Roux S, Emerson JB, Solonenko N, Vik DR *et al*. Soil viruses are underexplored players in ecosystem carbon processing. *MSystems*. 2018;3**:** e00076-00018.

[36] Gerba CP. Applied and theoretical aspects of virus adsorption to surfaces. *Adv Appl Microbiol*. 1984;30**:** 133-168.

[37] Quattrone A, Lopez-Guerrero M, Yadav P, Meier MA, Russo SE, Weber KA. Interactions between Root Hairs and the Soil Microbial Community Affect the Growth of Maize Seedlings. in review.

[38] Wolf DC, Skipper HD. Soil sterilization. In: Weaver RW, Angle S, Bottomley P, Bezdicek D, Smith S, Tabatabai A *et al* (eds). *Methods of Soil Analysis: Part 2 Microbiological and Biochemical Properties*. Soil Science Society of America, Inc.: Madison, Wisconsin.1994: pp 41-51.

[39] Hoagland DR, Arnon DI. The water-culture method for growing plants without soil. Document Number; California Agricultural Eexperiment Station, 1950:Circular-347

[40] Wommack KE, Williamson KE, Helton RR, Bench SR, Winget DM. Methods for the isolation of viruses from environmental samples. *Methods Mol Biol*. 2009;501**:** 3-14.

[41] Pan D, Watson R, Wang D, Tan ZH, Snow DD, Weber KA. Correlation between viral production and carbon mineralization under nitrate-reducing conditions in aquifer sediment. *ISME J*. 2014;8**:** 1691-1703.

[42] Pan D, Nolan J, Williams KH, Robbins MJ, Weber KA. Abundance and distribution of microbial cells and viruses in an alluvial aquifer. *Front Microbiol*. 2017;8**:** 1199.

[43] Westrop JP, Yadav P, Nolan P, Campbell KM, Singh R, Bone SE *et al*. Nitrate-Stimulated Release of Naturally Occurring Sedimentary Uranium. *Environ Sci Technol*. 2023;57**:** 4354–4366.

[44] Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA-and rRNA-based microbial community composition. *Appl Environ Microbiol*. 2000;66**:** 5488-5491.

[45] Gilbert JA, Jansson JK, Knight R. The Earth Microbiome project: successes and aspirations. *BMC Biol*. 2014;12**:** 1-4.

[46] Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangl JL *et al*. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *P Natl Acad Sci USA*. 2013;110**:** 6548-6553.

[47] Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13**:** 581-583.

[48] McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLos One*. 2013;8**:** e61217.

[49] Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 2015;31**:** 1674-1676.

[50] Kang DD, Li F, Kirton E, Thomas A, Egan R, An H *et al*. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7**:** e7359.

[51] Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 2016;32**:** 605-607.

[52] Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ *et al*. Binning metagenomic contigs by coverage and composition. *Nat Methods*. 2014;11**:** 1144–1146.

[53] Sieber CM, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG *et al*. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol*. 2018;3**:** 836-843.

[54] Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome R* 2015;25**:** 1043-1055.

[55] Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML *et al*. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ*. 2015;3**:** e1319.

[56] Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO *et al*. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome*. 2021;9**:** 1-13.

[57] Ren J, Song K, Deng C, Ahlgren NA, Fuhrman JA, Li Y *et al*. Identifying viruses from metagenomic data using deep learning. *Quant Biol*. 2020;8**:** 64-77.

[58] Jarett JK, Džunková M, Schulz F, Roux S, Paez-Espino D, Eloe-Fadrosh E *et al*. Insights into the dynamics between viruses and their hosts in a hot spring microbial mat. *ISME J*. 2020;14**:** 2527-2541.

[59] Santos-Medellin C, Zinke LA, Ter Horst AM, Gelardi DL, Parikh SJ, Emerson JB. Viromes outperform total metagenomes in revealing the spatiotemporal patterns of agricultural soil viral communities. *ISME J* 2021;15**:** 1956-1970.

[60] Adriaenssens EM, Wittmann J, Kuhn JH, Turner D, Sullivan MB, Dutilh BE *et al*. Taxonomy of prokaryotic viruses: 2017 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. *Arch Virol*. 2018;163**:** 1125-1129.

[61] Starr EP, Shi S, Blazewicz SJ, Probst AJ, Herman DJ, Firestone MK *et al*. Stable isotope informed genome-resolved metagenomics reveals that Saccharibacteria utilize microbially-processed plant-derived carbon. *Microbiome*. 2018;6**:** 1-12.

[62] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S *et al*. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007;315**:** 1709-1712.

[63] Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N *et al*. Uncovering Earth's virome. *Nature*. 2016;536**:** 425-430.

[64] Roux S, Brum JR, Dutilh BE, Sunagawa S, Duhaime MB, Loy A *et al*. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature*. 2016;537**:** 689-693.

[65] Ahlgren NA, Ren J, Lu YY, Fuhrman JA, Sun F. Alignment-free oligonucleotide frequency dissimilarity measure improves prediction of hosts from metagenomically-derived viral sequences. *Nucleic Acids R*. 2017;45**:** 39-53.

[66] Ahlgren NA, Fuchsman CA, Rocap G, Fuhrman JA. Discovery of several novel, widespread, and ecologically distinct marine Thaumarchaeota viruses that encode amoC nitrification genes. *ISME J*. 2019;13**:** 618-631.

[67] Zhou Z, Tran PQ, Breister AM, Liu Y, Kieft K, Cowley ES *et al*. METABOLIC: High-throughput profiling of microbial genomes for functional traits, biogeochemistry, and community-scale metabolic networks. *Microbiome*. 2022;10**:** 33.

[68] Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, Solden LM *et al*. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic acids research*. 2020;48**:** 8883-8900.

[69] Garber AI, Nealson KH, Okamoto A, McAllister SM, Chan CS, Barco RA *et al*. FeGenie: a comprehensive tool for the identification of iron genes and iron gene neighborhoods in genome and metagenome assemblies. *Front Microbiol*. 2020**:** 37.

[70]Author. Vegan: Community Ecology Package. R package. 2017. https://CRAN.R-project.org/package=vegan.

[71] Bouffaud M-L, Kyselková M, Gouesnard B, Grundmann G, Muller D, Moënne-Loccoz Y. Is diversification history of maize influencing selection of soil bacteria by roots? *Mol Ecol*. 2012;21**:** 195-206.

[72] Badri DV, Vivanco JM. Regulation and function of root exudates. *Plant Cell Environ*. 2009;32**:** 666-681.

[73] Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S *et al*. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat Microbiol*. 2018;3**:** 470-480.

[74] Liang X, Wang Y, Zhang Y, Zhuang J, Radosevich M. Viral abundance, community structure and correlation with bacterial community in soils of different cover plants. *Appl Soil Ecol*. 2021;168**:** 104138.

[75] Sandaa R-A, E. Storesund J, Olesin E, Lund Paulsen M, Larsen A, Bratbak G *et al*. Seasonality drives microbial community structure, shaping both eukaryotic and prokaryotic host–viral relationships in an arctic marine ecosystem. *Viruses*. 2018;10**:** 715.

[76] Hutchings MJ, John EA, Wijesinghe DK. Toward understanding the consequences of soil heterogeneity for plant populations and communities. *Ecology*. 2003;84**:** 2322-2334.

[77] Roux S, Emerson JB. Diversity in the soil virosphere: to infinity and beyond? *Trends Microbiol*. 2022;30**:** P1025-1035.

[78] Burstein D, Sun CL, Brown CT, Sharon I, Anantharaman K, Probst AJ *et al*. Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. *Nat Commun* 2016;7**:** 1-8.

[79] Tian R, Ning D, He Z, Zhang P, Spencer SJ, Gao S *et al*. Small and mighty: adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. *Microbiome*. 2020;8**:** 1-15.

[80] Trubl G, Kimbrel JA, Liquet-Gonzalez J, Nuccio EE, Weber PK, Pett-Ridge J *et al*. Active virus-host interactions at sub-freezing temperatures in Arctic peat soil. *Microbiome*. 2021;9**:** 1-15.

[81] Tian R, Ning D, He Z, Zhang P, Spencer SJ, Gao S *et al*. Small and mighty: adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. *Microbiome*. 2020;8**:** 51.

[82] Liang X, Zhang Y, Wommack KE, Wilhelm SW, DeBruyn JM, Sherfy AC *et al*. Lysogenic reproductive strategies of viral communities vary with soil depth and are correlated with bacterial diversity. *Soil Biol Biochem*. 2020;144**:** 107767.

[83] Brum JR, Hurwitz BL, Schofield O, Ducklow HW, Sullivan MB. Seasonal time bombs: dominant temperate viruses affect Southern Ocean microbial dynamics. *ISME J*. 2016;10**:** 437-449.

[84] Thingstad TF, Våge S, Storesund JE, Sandaa R-A, Giske J. A theoretical analysis of how strain-specific viruses can control microbial species diversity. *P Natl Acad Sci USA*. 2014;111**:** 7813-7818.

[85] Suttle CA. Viruses in the sea. *Nature*. 2005;437**:** 356-361.

[86] Emerson JB, Roux S, Brum JR, Bolduc B, Woodcroft BJ, Jang HB *et al*. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat Microbiol* 2018;3**:** 870-880.

[87] Tong D, Wang Y, Yu H, Shen H, Dahlgren RA, Xu J. Viral lysing can alleviate microbial nutrient limitations and accumulate recalcitrant dissolved organic matter components in soil. *ISME J*. 2023**:** 1-10.

[88] Middelboe M, Lyck PG. Regeneration of dissolved organic matter by viral lysis in marine microbial communities. *Aquat Microb Ecol*. 2002;27**:** 187-194.

[89] Liang C, Schimel JP, Jastrow JD. The importance of anabolism in microbial control over soil carbon storage. *Nat Microbiol*. 2017;2**:** 1-6.

[90] Starr EP, Shi S, Blazewicz SJ, Koch BJ, Probst AJ, Hungate BA *et al*. Stable-isotope-informed, genome-resolved metagenomics uncovers potential cross-kingdom interactions in rhizosphere soil. *Msphere*. 2021;6**:** e00085-00021.

[91] Genilloud O. Physiology of actinobacteria. In: Wink J, Mohammadipanah F, Hamedi J (eds). *Biology and Biotechnology of Actinobacteria*. Springer: Cham, Switzerland.2017: pp 151-180.

[92] Singh PR, Jain M. Studies on the cellulolytic bacteria and cellulose degradation in a cattle waste-fed biogas digester. *MIRCEN journal of applied microbiology and biotechnology*. 1986;2**:** 309-317.

1   **Figure legends**

2

3   **Fig. 1 Cell and virus abundance in the parviglumis and B73 rhizobiome. (a)** Cell abundance

4   per gram of soil in the rhizoplane and rhizosphere referred to a rhizobiome of biological replicates

5   of *parviglumis* (n=5) and B73 (n=6). (b) Virus abundance per gram of the rhizobiome soil. (c)

6   Virus to cell ratio (VCR) between *parviglumis* and B73. The brackets indicated a non-parametric

7   unpaired t-test with Welch's correction assuming unequal variances. The asterisk denoted a *p*-

8   value < 0.05 as statistically significant whereas *ns* denoted that the result was not significant.

9   Values reported are the mean of replicates and error bars represent the standard error of measure.

10

11   **Fig. 2 Microbial community composition of Z. mays parviglumis and B73 identified by 16S**

12   **rRNA gene amplicon sequence variants (ASVs) and metagenome-assembled genomes**

13   **(MAGs).** The microbial community of *parviglumis* and B73 represented (a) a percentage of the

14   relative abundance of the microbial community identified by ASV. (b) PCoA of ASV identified

15   rhizobiome from parviglumis (blue) and B73 (red) in biological and technical replicates with the

16   Bray-Curtis distance matrix. The variance of 56.1% on PCoA1 and 17.4% on PCoA2 (R^2 =0.1089,

17   *p*-value 0.047) as determined by perMANOVA. Ellipses denote 95% confidence interval. (c)

18   Relative abundance of MAGs determined using normalized genome coverage (d) PCoA of MAG

19   identified rhizobiome from *parviglumis* (blue) and B73 (red) with Bray-Curtis dissimilarities with

20   88.1% variance on PCoA1 and 2.3% variance on PCoA2 (R^2 =0.999, p = 0.001). The box plot

21   represents the interquartile range with whiskers ± 1.5 times of interquartile range and the center

22   depicts the median.

23

24   **Fig. 3 Metagenome assembled genomes (MAGs) and viral contigs from the Z. mays**

25   ***parviglumis* and B73 rhizobiome**. MAGs were reconstructed and viral contigs were identified

26   from three biological replicates from each genotype, *parviglumis* and B73 rhizobiome. The

27   number of genomes corresponding to the archaeal and bacterial taxa classified by the GTDB

28   database and clustered to the family level when possible is denoted in red and the relative

29   abundance of genomes is denoted in blue. The mean coverage depth of MAGs normalized against

30   the number of reads in the sample to calculate the relative abundance of the MAGs.

31

32   **Fig. 4 Identification of viral contigs within the *parviglumis* and B73 rhizobiome.** To assess the

33   similarities/differences of viral contigs in the *parviglumis* and B73 rhizobiome, pairwise

34   comparison with average nucleotide identity (ANI) > 95% with coverage > 85% of the viral

35   contigs of the shorter sequence was performed where (a) only 7 viral contigs were clustered >

36   95% ANI showing heterogeneity in identified viral contigs. (b) Ordination conducted with Bray-

37   Curtis dissimilarity on the relative abundance of the viral contigs explained 91.9% variance on

38   PCoA1 and 3.6% variance of PCoA2 (R^2 = 0.999 p = 0.005) as analyzed by permutational

39   multivariate analysis of variance (perMANOVA) conducted utilizing adonis2. (c) Taxonomic

40   classification of identified viral contigs comprised 6 different families of Caudovirales in

41   *parviglumis* and 3 families in B73 with an unclassified archaeal virus. The relative abundance of

42   the classified viral contigs was calculated as the proportion of the corresponding viral family

43   divided by the total classified viral contigs multiplied by 100.

44

45   **Fig. 5 Virus-host linkages identified in the *parviglumis* and B73 rhizobiome.** (a) Viral contigs

46   were linked to the hosts utilizing genome similarities, CRISPR sequences, and tetranucleotide

47   frequencies between the metagenome-assembled genomes (MAGs) and viral contigs. While the

48   hosts identified for the *parviglumis* and B73 viral contigs showed differences in the host

49   preferences, hosts couldn't be identified for more than 60% of viral contigs in both genotypes and

50   are identified in gray. (b) Virus to host abundance ratio was divided into 3 categories; less than 1,

51   approximately 1, and greater than 1 demonstrated variation in viral genome copy number to their

52   host in *parviglumis* and B73.The black vertical line separates the *Z. mays* genotypes, *parviglumis*

53   and B73. The viral contigs are color-coded according to the identified potential host. The host

54   taxonomy is identified at the family level (or lowest level of classification). Error bars represent

55   the standard error of measure.

56

57   **Fig. 6. Metabolic potential of the microbial community identified from the annotation of**

58   **MAGs and virus contigs in the *parviglumis* and B73 rhizobiome.** The rhizobiome metabolic

59   potential was determined using DRAM, METABOLIC, and FeGenie. The heatmap represents the

60   metabolic potential of the total metagenome-assembled genomes (MAGs) (top), viral infected

61   MAGs (middle), and contribution of viral auxiliary metabolic genes (AMGs) (bottom) of (a)

62   *parviglumis* and (b) B73 rhizobiome, respectively. Percentage of genomes were determined at the
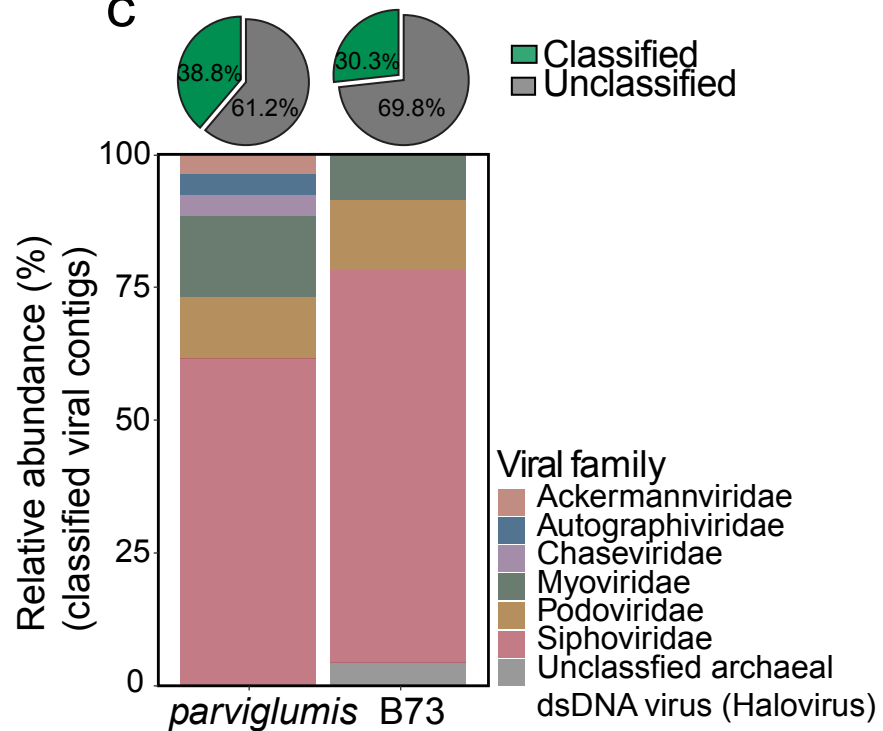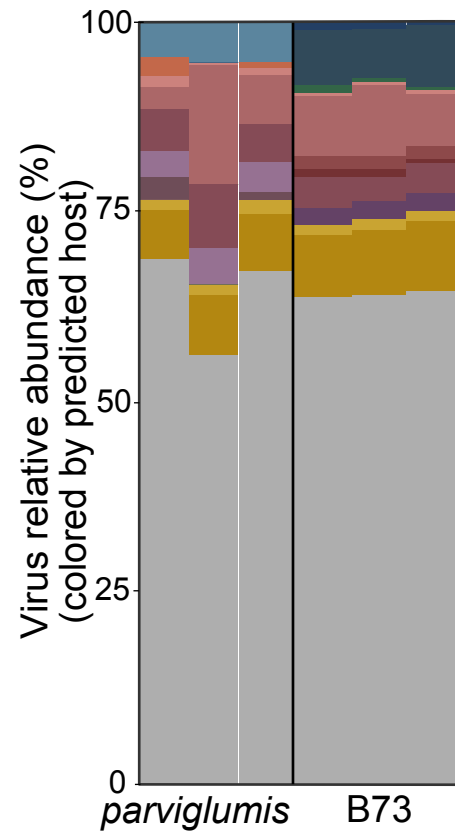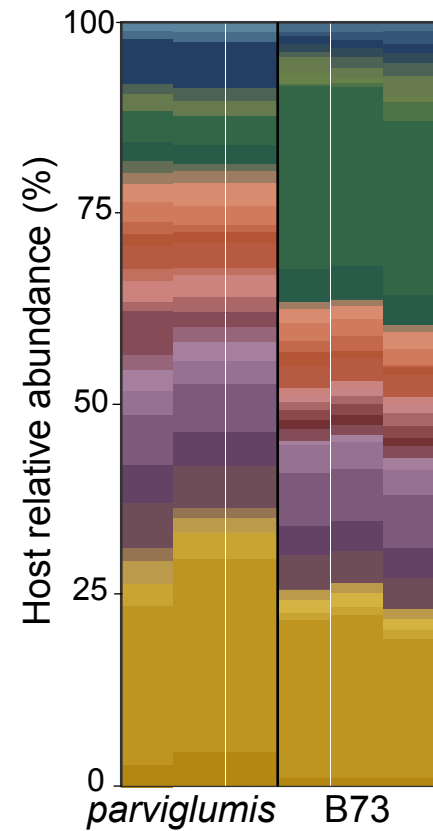
63   family level.

a

Host taxonomy

Acidobacteriaceae
Acidobacteriae UBA7541
Actinobacteriota HRBIN12
Alphaproteobacteria UBA1301
Binatia UBA9968
Burkholderiaceae
Chitinophagaceae
Chloroflexota CSP1-4
Chthoniobacteraceae
Chthoniobacterales UBA10450
Gaiellaceae
Gemmataceae
Gemmatimonadaceae
Isosphaeraceae
Jatrophihabitantaceae
Koribacteraceae
Kribbellaceae
Methyloligellaceae
Microbacteriaceae
Micrococcaceae
Mycobacteriaceae
Nitrososphaeraceae
Propionibacteriaceae
Rhizobiaceae
Saccharimonadaceae
Solirubrobacteraceae
Sphingomonadaceae
Streptomycetaceae
Streptosporangiaceae
Unclassified Bacteria
Unclassified Actinobacteriota
Unclassified Gammaproteobacteria
Unclassified Proteobacteria
Unclassified Rhizobales
Vicinamibacterales UBA2999
Xanthobacteraceae
Xanthomonadaceae
Unknown host

b