The phage in the dirt: Studies in the ecology of terrestrial bacteriophages

Jessica Lee
MBL Microbial Diversity, summer 2008
Introduction

Probably the most numerous biological entities on earth, viruses also play important roles in the ecology of microbial communities and in global biogeochemical cycling (5, 13). We are only just beginning to build a rudimentary model of viral ecology in aquatic systems, but our understanding of phage ecology in soil lags far behind. This may be primarily due to our lesser understanding of microbial soil systems in general; the heterogeneity of the soil matrix, characterized by microenvironments, makes it more complicated to study the bacteria and archaea that live there than in the (by comparison) homogeneous and easily-defined water column. In addition, viruses adhere to soil particles: in most environments, over 90% of viruses are adsorbed to the soil matrix (8). Factors influencing strength and degree of adsorption include ionic strength and composition, pH, presence of organic matter, soil water content, and physical characteristics of the virus (15). Extensive research has been done on the movement of viruses through soils as it relates to human-pathogenic viruses, but there is little knowledge about how the soil matrix affects phage/host interaction.

The quantity and activity of soil phages

For example, recent efforts to quantify virus particles in soils have found phages to be more abundant in soils where bacterial hosts are more abundant, for instance in organic-rich soils (approximately 80-390 x 10^7 virus-like particles per gram dry weight of soil) than in poor soils (23-64 x 10^7) (15). But when compared to the much wider variation in bacterial abundance (up to 10,000-fold), viral abundance actually remains relatively constant among ecosystems. Consequently, the Viral to Bacterial Ratio (VBR) varies widely (up to 8000-fold), much more widely than in aquatic systems where bacterial abundance varies less and viral abundance tracks more closely with that of the hosts (12). From this observation, it seems evident that there must be features of the soil environment that lead to significant differences in phage/host population dynamics: perhaps viral activity is lower and the population responds more slowly to changes in host abundance (12); perhaps viruses are more adapted to lead a lysogenic lifestyle (14). At the same time, it is important to acknowledge that our definition of an “active” viral particle is still unclear, and that it may be true that viral particles live longer in soils due their physical properties (15). It is also unclear how influenced our sampling is by our ability to extract viral particles from soil, which is a difficult process and has been shown to generate bias among different viral types (16).
The diversity of phages

Culture-independent methods have led to impressive estimates of viral diversity in marine systems. Based on current knowledge of marine phage diversity and host specificity, (11) proposes an estimate of 10 unique phage types per unique microbial species. He further calculates that if there exist 6 million free-living microbial species and 3–5 million associated with eukaryotes, leading to approximately 10 million microbial species, there should therefore be 100 million unique phage genotypes on earth (11). However, our estimate of global phage diversity will depend strongly on our increasing understanding of phage diversity in different biomes, of which the terrestrial is much more undersampled than the aquatic. Recent PCR-based studies in the diversity of the T7-like DNA polymerase gene in different biomes has in fact found sequences which seem to be pervasive across a wide range of biomes, suggesting cross-infectivity of phages in widely differing hosts, or at least very (pervasive) lateral gene transfer (1). It remains to be demonstrated whether entire viruses move between viruses or just pieces of their DNA moving, but if the former is true it could mean that we are estimating phage genetic diversity based on an incorrect understanding of host specificity—and that while local diversity is high, global diversity may after all be relatively low (2).

The goal of this project was certainly not to answer all of the above questions but to initiate experiments that begin to address them. A host-phage system native to soil in Woods Hole, MA, was isolated and characterized. The phage was used in a microcosm experiment to measure the adsorption of viral particles to organic-rich sandy soil. Finally, a gene library was created to survey the diversity of gp23, the capsid gene of T4-like phages, in the soils of three different Woods Hole sites.

Methods

Isolation of host bacteria

All potential bacterial host strains were isolated from topsoil in the courtyard of the Loeb laboratory at the Marine Biological Laboratory in Woods Hole, MA. 1 g soil was suspended in 5 mL distilled H$_2$O by vortexing, then boiled for 10 minutes. The suspension was plated on Nutrient Agar (Lawrence, KS) and incubated at 30°C until colonies were observed. Several single colonies were picked and streaked for isolation; isolates were named JAL-S1 through JAL-S13. Once isolated, all host strains were maintained at room
temperature on LB agar (15% agar) plates. Overnight cultures and exponential-phase cultures were grown in liquid LB at 30°C with shaking; overnight refers to >8 hrs incubation and exponential phase to approximately 3 hrs.

**Determination of host 16S phylogeny**

To determine the phylogenetic relationship among all the spore-forming isolates, 16S ribosomal RNA sequences were sequenced using colony PCR, specific to the bacterial 16S gene, on each strain as follows: a single colony was picked into 75 µL 100% dimethyl sulfoxide (DMSO) and physically macerated using a sterile mini-pestle in an eppendorf tube. 1 µL of the colony-DMSO mixture was used as template in a PCR reaction containing 1x Promega PCR Master Mix (Promega, Madison, WI) and 1.2 pmol each primer 16S_8F and 16S_1492R, for a total reaction volume of 15 µL. Primers were removed and dNTPs were inactivated in the PCR product by treatment with ExoSAP-IT enzyme (USB corporation, Cleveland, OH), in a reaction containing 0.25 µL enzyme, 1.5 µL PCR product, and 3.25 µL H2O. Sequencing was done on a ABI 3730XL sequencer by the Keck facility at the Josephine Bay Paul Center (Woods Hole, MA).

The 16S sequences were then aligned with the SILVA 16S database using ARB (9) and placed in a RAXML maximum likelihood tree using the same program.

**Isolation of bacteriophage infecting Bacillus hosts**

**Enrichment:** To isolate phages infecting the different Bacillus isolates, soil from three locations in Woods Hole was used: sand from the MBL beach, topsoil from the rhizosphere of woodland adjacent to the Shining Sea bikepath, and sediment from School Street Marsh (Fig. 1). The soils were mixed together, 1 part woodland soil with 2 parts each beach sand and marsh sediment by mass. For each host isolate, 30 g of soil mix was added to a culture flask with 30 mL LB and 1 mL overnight culture of the chosen host. The enrichments were incubated, shaking, at 30ºC overnight. After removal from the incubator, the flasks were allowed to sit, undisturbed, 20 minutes, to settle the soil, and 1.5 mL of supernatant was removed and centrifuged to further separate soil from liquid. The supernatant was filtered through a 0.2 uM syringe filter to remove remaining particles and bacterial cells. The lysate was diluted in LB broth and dilutions of each lysate at 10⁰, 10⁻², 10⁻⁴, 10⁻⁶, and 10⁻⁸ were spotted in 5 µL volumes onto a lawn of its respective host. Lawns were generated by mixing
100 μL exponential-phase host culture with 3 mL LB-top agar (6.0% agar) at 45°C and pouring onto an LB-agar plate. The spot assays were incubated at 30°C overnight and checked for clearings indicating lysis by phages.

**isolation:** Enrichments which produced phages were plaque-purified as follows: lysate was diluted in LB to 10^6, 10^7, and 10^8. For each dilution, 100 μL lysate dilution was mixed with 100 μL exponential-phase host culture and 3 mL LB-top agar, poured over an LB-agar plate, and incubated at 30°C overnight. When individual plaques appeared, a sterile Pasteur pipet was used to pick a single plaque into 100 μL LB. The plaque-LB mixture was vortexed and let to sit for approximately 6 hours at room temperature.

To amplify the plaque-purified phage, the plaque mixture was mixed again, undiluted, with 100 μL exponential-phage host culture and 3 mL and plated; extensive or total lysis was observed after overnight incubation. 2 mL LB was overlaid on the plate, and scraped off with all the top agar and allowed to sit at room temperature several hours. The extract was centrifuged 5 minutes at 4500 rpm and the supernatant filtered through a 0.2 μM syringe filter.

The resulting phage extract was put through the entire plaque-purification and amplification procedure once more.

Only one extract, named P1C, (resulting from one double-plaque-purified plaque) was ultimately used for all further characterization experiments.

![Figure 1. Map of sample sites in Woods Hole, MA.](image)
**Phage lysate titrations**

The concentration of plaque-forming units per mL lysate (pfu/mL) was calculated using a plaque assay similar to the one for plaque purification: lysate was diluted in LB to 10^6, 10^7, and 10^8 concentrations. For each dilution, 100 µL lysate was mixed with 100 µL exponential-phase host culture and 3 mL LB-top agar, and poured over an LB-agar plate. After incubation overnight at 30°C, the number of plaques on each plate was counted and used to calculate the total concentration of pfu/mL in the original undiluted lysate; when all three plates were countable, the mean of all three calculations was taken.

**Host range of P1C**

The infectivity of P1C was tested against other *Bacillus* isolates. A lawn of host cells was created by mixing 100 µL exponential-phase host with 3 mL LB-top agar at 45°C and pouring on top of an LB-agar plate and allowing to set. 5 µL undiluted phage lysate was spotted on top of the lawn and incubated overnight at 30°C. A clearing was observed on susceptible hosts; no clearing was observed on resistant hosts (Fig. 2).

**Transmission Electron Microscopy**

Undiluted, double-plaque-purified lysate of phage P1C was examined by transmission electron microscopy. 10 µL lysate was spotted onto a grid, negative-stained with 10uL saturated uranyl acetate solution (0.025g/mL), and dried at room temperature for 1 hour before being viewed using a Zeiss 10CA transmission electron microscope.

**Extraction of phage DNA**

Phage P1C genomic DNA was extracted as follows: 3.6 mL lysate was filtered twice through a 0.2 µM syringe filter to remove bacterial cells and frozen at -20°C for 2 hours to facilitate lysis of phage capsids. To obtain maximum possible concentration of DNA, reagents from the Promega Wizard Genomic DNA purification kit were used as follows: the lysate was combined with 1.2 mL Protein Precipitation solution and incubated on ice 5 minutes, then centrifuged 3 minutes at 13,200 rpm to pellet proteins, and the supernatant
was transferred to new tubes and combined with 3.6 mL isopropanol to precipitate DNA. To concentrate the DNA in one eppendorf tube, 1.5 mL of the solution was centrifuged at 13,200 rpm, the supernatant discarded, the tube topped up with more isopropanol-DNA mixture, and the process repeated until all DNA had been pelleted together. The DNA was washed by adding 600 µL 70% ethanol, centrifuging, and removing the supernatant. It was then air-dried for 30 minutes at room temperature, then rehydrated in 50 µL Rehydration Solution for 1 hr at 65ºC.

Restriction digest of phage DNA

Phage DNA was digested with MspI restriction endonuclease, which cuts twice at (CCGG). 1 µL of MSP-I enzyme (New England Biolabs, Ipswich, MA) was combined with 1X NEBuffer 2 (New England Biolabs) and 1 µg P1C genomic DNA extract in a total volume of 30 µL, for 1 hour at 37ºC. The products were run for 40 minutes on a 1% agarose gel at 100V, with 6 µL 1 Kb DNA ladder (Invitrogen, Carlsbad, CA).

Soil substrate: sourcing and sterilization

Soil for phage particle assays was a sandy loam obtained from the top 10 cm of rhizosphere in the woodland adjacent to the Shining Sea bike path (Fig 1). The soil was sieved using a 20mm sieve and autoclaved once a day for three days, the first time 25 minutes on liquid cycle and the following times 25 minutes on dry. After each autoclaving, the soil was assayed for remaining viable bacterial cells and spores by resuspending 3 g soil in 5 mL Freshwater Base medium (17.1 mM NaCl, 1.97 mM MgCl₂•6H₂O, 0.15 mM CaCl₂•2H₂O, 6.71 mM KCl), vortexing, plating 100 µL of undiluted and 10⁻² diluted supernatant on Nutrient Agar plates (Difco), and incubating at 30ºC for 2 days. After the third autoclaving, no more than one colony was found on any plate.

Phage particle retention on soil

To assay the retention of phage particles on soil over time, 1.5 mL P1C lysate was combined with 15 g autoclaved woodland soil in a sterile culture flask, for an initial concentration of 5x10⁸ pfu/g soil and incubated, in the dark, at room temperature, and lightly covered in foil, for 5 hours. Each hour, the phage particles recoverable from soil were assayed as follows: 1 g inoculated soil was combined with 1 mL LB and vortexed 10 seconds.
The mixture was centrifuged at 1,500 rpm for 2 minutes in order to settle the soil enough to allow removal of the supernatant by pipet; the solution remained very cloudy and still contained suspended soil particles. The solution was diluted $10^{-2}$ and $10^{-3}$ in LB, and each dilution was used in a plaque assay (as described for lysate titration, above) and the plaques produced were counted to estimate the number of phage particles recovered in solution. No replicates were conducted for this experiment—only one soil sample with two dilutions per timepoint was taken and one plate per dilution plated.

*Total DNA extraction from soil*

Soils from three sites were used for the generation of phage gene clone libraries: MBL beach sand, bikepath woodland soil, and School Street Marsh sediment (Fig 1). Total DNA was extracted using the MoBio PowerSoil DNA Kit (Carlsbad, CA), with the following modifications to increase concentration of final extract: for each site, two 0.5 g samples were processed in separate tubes until the membrane-binding step, at which point all extract from each site was pooled onto one membrane. The DNA from each sample was eluted in 60 µL H₂O. Concentration was measured using a Nanodrop (Wilmington, DE).

*Phage gene amplification*

To amplify the T4-like gp23 gene from environmental soil extracts, PCR was conducted using a modification of the procedure given in (4). Briefly, a mixture consisting of 1x PCR buffer (Invitrogen), 1.5 mM MgCl₂, 0.80 mM dNTP mix (Invitrogen), 2.0 µM each primer, 1 unit Taq DNA polymerase (Invitrogen), and between 10 and 70 ng template DNA. The primers used were MZIA1bis (5′-GATATTTGIGGIGTTCA GCCIATGA-3′) and MZIA6 (5′-CGCGGGTGATTTCCAGCATGATTTC-3′). The reactions were cycled on a PTC-200 thermocycler (Bio-Rad Laboratories, Hercules, CA) at 94°C for 1 min 30 sec; 30 cycles of (94°C for 45 sec, 50°C for 1 min, and 72°C for 45 sec); and finally 72°C for 5 min. Two reactions from each site were pooled to obtain a total of 100 µL product per site.

To amplify the T7-like DNA Polymerase gene by PCR, the procedure given by (1) was used. The reaction mixture consisted of 1x Taq buffer (Invitrogen), 1 mM MgCl₂, 200 µM dNTP mix (Invitrogen), 1 µM each primer, and 1 unit Taq polymerase (Invitrogen), for a total volume of 50 µL. The primers used were T7DPol230F (5′-ARG ARMRIAAYGGIT-3′) and T7DPol510R (5′-GTRTGDATRTCICC-3′). The mixtures were cycled on a PTC-200
thermocycler (Bio-Rad) for 94 °C for 5 min; 30 cycles of (94 °C for 1 min, 50–0.5 °C/cycle for 1 min, 72 °C for 2 min); and 72 °C for 10 min. In addition, a parallel reaction was conducted using the following thermocycler conditions: 94 °C for 5 min; 35 cycles of (94 °C for 1 min, 45°C for 1 min, 72 °C for 2 min); and 72 °C for 10 min. Both reactions generated product and the products were pooled to obtain 100µL from each site.

Clone library generation and treeing

The PCR products from each site were processed using the QIAquick PCR Purification Kit (QIAGEN, Valencia, CA) and eluted in 25 µL H₂O. 3 µL of the resulting product was then used in a 5 µL cloning reaction using the TOPO TA Cloning kit by manufacturer’s directions (Invitrogen). 32 clones from each site were picked for sequencing. Clone libraries were sequenced by the Keck facility at the Josephine Bay Paul Center (Woods Hole, MA).

Because the specificity of the PCR amplifications of both genes was so low, sequenced clones were screened for homology to a published gene (T4 gp23 or T7 DNA polymerase) by using BLASTx. Sequences either hit the target gene or they did not; those that did not were discarded from the dataset. Sequences were aligned with other gp23 sequences published in GeneBank using MAFFT (Multiple Alignment using Fast Fourier Transform) with a 200PAM scoring matrix for nucleotides and penalties of 1.53 for gap open and 0.12 for gap extension. A maximum likelihood tree was generated using RAxML with a GTR+Gamma model.

Results

Isolation and characterization of a phage-host system from Woods Hole soils

Twelve spore-forming isolates, showing 16S sequence similarity to a range of Bacillus species (Fig. 3), were isolated from organie-rich soil in the rhizosphere of trees in the courtyard of Loeb Laboratory, Marine Biological Laboratory, Woods Hole, MA. All twelve were used as potential hosts in an enrichment for bacteriophage from soil from a range of locations in Woods Hole; unfortunately, a lytic phage were isolated successfully on only one of the twelve host strains. The host was named S13 and the phage P1C. The host strain is closely related to Lysinibacillus sphaericus (Fig. 3). Figure 4 shows morphology of the colonies and of the microbial cells. Like the other bacterial isolates used
in this study, it is an aerobic, mesophilic, spore-forming bacterium. However, \textit{L. sphaericus} has been particularly well-studied in the past decades because 9 of its 49 recognized serotypes produce mosquitocidal toxins (7) [Howard 1973](10). In 2008 the genome of \textit{L. sphaericus} C3-41 was published; the genome consists of 4,639,821 bp, with a 177,642 bp plasmid. Features possibly relevant to phage-host interactions include one CRISPR with three spacers, and three putative prophages. However, given that C3-41 has been in culture for years, during which time evolution tends to favor the loss of CRISPR spacers and alter the expression of mobile genetic elements such as prophages, the relevance of these features to the host-phage system isolated in this experiment is uncertain.

Figure 2. 16S rRNA tree of bacterial isolates used in this study.
The isolated phage was characterized by several means.

**Plaque morphology on LB agar:** P1C was double-plaque purified from small (ca. 2mm) plaques. However, during subsequent titrations, larger, turbid plaques with opaque centers occasionally appeared (Fig. 5). It is unclear whether these are an indication that the lysate actually consisted of a heterogeneous mix of phages, or if a prophage contained in the host cells used for the assay was occasionally induced to lysis.

**Host specificity and resistance to the phage:** four strains of S13 with resistance to P1C were isolated (named P1C-R1 through P1C-R4). In addition, it was found that ten of the twelve *Bacillus* isolates used in this experiment were resistant to the isolated phage; only S13 and its closest relative, S4, were susceptible. No experiments were conducted on the resistant strains.

**Phage particle morphology:** Transmission electron microscopy was used to produce negative-stained images of phage particles. All particles that resembled known phages were of the same morphology: an icosahedral head of approximately 80 nm in diameter, with a tail approximately 120 nm long; tail fibers were probably present but barely perceptible (Fig 6a-
d). It resembles known myoviruses, which are dsDNA phages. Other unidentified objects were observed in the field of view, such as polygonal structures, slightly larger and smaller than phage heads but lacking tails; also long filaments approximately 10 nm wide but up to millimeters in length (Figs. 6e, 6f). The morphological identification of phage is of course limited to the observer’s familiarity with phage morphotypes.

Figure 6 (a-d). Transmission electron microscopy images of P1C phages in LB. Samples were negatively stained with uranyl acetate.
Nucleic acid and genome size determination: dsDNA was extracted from the phage lysate using a genomic DNA extraction kit, in order to demonstrate that the phage contains dsDNA. A concentration of 86 ng/µL was obtained from 3600 µL of lysate at a titer of $5 \times 10^8$ pfu/mL; even if there was residual host DNA present in the lysate, it is unlikely that it would have been the sole source of such a high yield, so it is most likely that this is a dsDNA phage.
In addition, the DNA was restriction digested using MspI to produce small bands with the hope that the sizes could be summed to calculate the total genome size (Fig 7a). Unfortunately, the bands were too numerous to be resolved accurately. This is primarily due to the fact that the site cut by MspI, GCCG, is probably extremely common in the phage genome—as indicated by the in-silico digest of the T4 genome (Fig 7b). However, it also provides no reassurance that the extract does not contain a significant amount of host DNA. To assess whether any viable host cells or spores remained in the freeze-thawed, double-filtered lysate, 200 µL were plated on an LB-agar plate and incubated at 30ºC. After three days, 54 colonies formed, indicating that viable host material had indeed survived the treatment. In addition, it is possible that free-floating host DNA from lysed cells remained in the lysate and passed through the 0.2 µM syringe filter to remain in the final DNA extract. It is evident that further studies will require a more precise method for phage purification, for instance CsCl density gradient separation.

Phage retention on soil

The retention of phage particles on a organic-rich sandy loam soil was measured by inoculating soil with solution of a known phage titer and assaying the number of plaque-forming units (pfus) recoverable at regular intervals over a five-hour period. Even at the initial timepoint, with no incubation, the number of pfus recovered from the inoculated soil was reduced by 10² (Fig. 8). Thereafter, the recoverable phage titer fell continually, approximately at the rate of (y = 2x10⁶ x e⁻³.3312x, R² = 0.9575).

![Figure 8. Recovery of plaque-forming units from sterile soil inoculated with PIC lysate. The soil was incubated at room temperature for 5 hours and assayed once each hour by suspending 1 g soil in 1 g LB, vortexing, and plating 100 uL of supernatant with a lawn of S13 host.](image-url)
Unfortunately, because this was only a pilot experiment there were no replicates conducted, but this finding merits further research into the forces governing the rate at which viral particles adsorb to soil upon contact in a system with no fluid flow. It is possible that the decreasing pfu recovery was due not only to phage adsorption to the soil but also by inactivation of phage particles in some other manner, for instance by reversible contact. At the least, it is probable that similar experiments conducted on different soils or on different viruses would yield different results; as mentioned in the introduction, a range of physical parameters are known to influence phage adsorption to soil.

**Phage gene diversity in three Woods Hole soil environments**

Both gp23 (the T4-like capsid gene) and the T7 DNA polymerase were PCR-amplified in DNA extracted from three different soils found in Woods Hole: beach sand from the MBL Dock, rhizosphere sandy loam from the woodland adjacent to the Shining Sea Bike Path, and anaerobic sediment from School Street Marsh. Amplification of both genes in all three samples produced only weak products with significant smearing, indicating the amplification of many different-sized genes and possibly also nonspecific amplification. Clone libraries were produced for both genes, using 32 clones from each site. However, the DNA polymerase clone library produced no sequences that resembled phage DNA polymerases. When submitted to NCBI BLAST, most matched either the cloning vector, the lacZ gene (which is adjacent to the cloning site on the vector), or an unidentified hypothetical bacterial gene. Most were also very short (<300 bp) reads; most likely, the smallest amplicons were preferentially cloned, and the DNA polymerase amplicons were not abundant enough in the PCR product to be represented in the 96-clone library. Gel purification would have been the simplest method for reducing the number of short reads in the product mix.

A similar problem was encountered with the gp23 library; however, in total 37 gp23-like sequences were recovered out of the 96 clones sequenced. Of these, 24 were from the MBL Dock, 10 from School Street Marsh, and 3 from the Bike Path. These sequences were aligned with gp23 sequences from published phage genomes in GeneBank as well as with environmental sequences published by two previous gp23 diversity surveys (4, 6), and plotted in a maximum likelihood tree (Fig. 9).
Figure 9. Maximum likelihood tree of gp23 clones sequenced from three sites in Woods Hole. Green, woodland adjacent to the Shining Sea Bike Path. Blue, beach sand from the MBL Dock. Brown, anaerobic sediment from the School Street Marsh.

Soil type played some role in predicting the phylogenetic relationship among gp23 clones: the Marsh sequences formed three distinct and relatively-deep-branching clusters, while two of the Bike path sequences formed their own cluster as well. However, the Dock sequences fell into 11 distinct clusters of varying population. Both Dock and Marsh sequences were no more similar among themselves than they were to each other. It might have been expected that Dock clones, having regular exposure to the marine environment, would group preferentially with gp23 sequences of cultured marine phages and
environmental sequences; and that Marsh sequences would group with clones isolated from a rice paddy site, as both are freshwater, anaerobic systems. However, neither grouping was observed. In addition, one of the Bike Path sequences was most closely related to a Dock sequence.

Many of the genome-sequenced phages grouped together, at a long branch length from the rest of the tree. No clones from Woods Hole were found to be closely related.

**Discussion**

It may be worthy of note that the cloning reaction was most successful with sand from the MBL Dock and least successful with soil from the Bike Path woodland. These primers have been used with great success with marine samples, and even once before with anaerobic freshwater sediments (rice paddy) similar to those in School Street Marsh, but have never before been used for aerobic, organic-rich topsoil. The bias in amplification success may be an indication of the bias of the primers, which were originally tested on *E. coli* phage T4 (a T-even), on *Aeromonas* phage Aeh1 (a Schizo-T-even), and the *Synechococcus* cyanophage S-PM2 (an ExoT-even); the Bike Path sequences that were recovered in this experiment were not closely related to any of those three. However, the scarcity of gp23 sequences in the Bike Path soil may also be an indication that not many T4-like phages exist there—or that not many phages exist there at all. The yield of DNA extraction was in fact much higher from the Bike Path soil than from any of the other soils (208 ng DNA/g soil, compared to 69 ng/g for mud and 18 ng/g for sand), so it is possible that phages simply may be low in abundance relative to other organic matter.

The observation that soil type does not determine phylogenetic relationship in Woods Hole gp23 clones—that sequence clusters from beach, marsh, and woodland environments are distributed evenly among all branches of the tree—supports the hypothesis of Breitbart and Rohwer that phages are promiscuous across biomes. Because *gp23* is a structurally essential gene, it is less likely that the gene would be transferred laterally on its own, and more likely that the phages themselves cross-infect hosts of different soil and water environments, or that the same hosts are present in different environments. If the former is true it may re-shape our understanding of phage-host specificity.
Not until we have a better grasp of the phage diversity present in many kinds of soils can we understand how the ecology of viral populations is shaped by environment. We are only just now starting to accumulate knowledge of that diversity.

This project was a very short-term foray into many avenues of research on phage in soils that deserve further work. The characterization of the phage P1C alone will take several more experiments using classical phage biology techniques. However, the study of phage populations in the environment requires the development and optimization of many new techniques, because the science of studying phages that are adsorbed to soils is still in its infancy. The extraction of phage particles from soil, which until now has been done primarily for the purposes of microscopy and enumeration (3, 16), could greatly benefit the genetic studies by allowing the concentration of phage DNA separate from other interfering genetic material, perhaps making the generation of a DNA polymerase clone library. It would also allow the study of phage-encoded host genes, which would otherwise be impossible to study if host DNA were part of the template used for PCR.

As mentioned earlier, the process of phage particle retention on soil has been well-studied, but we know little about how this would influence phage-host dynamics. This study found that phage particles, even when combined with soil, may remain free-floating and viable within a short period before becoming inactivated. Further studies must be done to determine whether adsorbed phage particles are ever recoverable again. Even more interesting would be an investigation into evolutionary adaptations of terrestrial phages to cope with immobilization on soil. Given the ubiquity of phage gene types across different biomes, however, such adaptations may not even exist.

Finally, anaerobiosis is another poorly understood factor in phage biology, but an important one for terrestrial phages because of the ubiquity of oxygen limitation in soils. Whether there exist unique phages specific for obligately anaerobic hosts, and whether phages of facultative aerobes can maintain the same infectivity when their hosts switch metabolisms (and possibly some physiological features), presents rich material for research. We are not wanting for questions in the world of phages in soil.
References


2. **Breitbart, M., and F. Rohwer.** 2005. Here a virus, there a virus, everywhere the same virus? Trends in Microbiology **13:**278-284.


Acknowledgements

I owe my utmost thanks to:

Tom Schmidt and Bill Metcalf, the brains behind the summer Microbial Diversity course;
the teaching assistants, especially Dave, Dagmar, and Stephanie;
Louie Kerr, for his help in generating beautiful TEM images;
Curtis Suttle, for his encouragement and advice on phages;
Bob Haselkorn, Peter Geiduscheck, William Studier, for providing positive controls for the phage gene PCRs;
Pat and Charles Robertson, for the generous scholarship that allowed me to attend MBL;
the Stanford School of Earth Sciences McGee grant, for its financial help as well;
my classmates, for being there to talk through every experimental up and down;
S13 and P1C, for growing quickly and almost predictably. One couldn’t ask for a better set of experimental organisms