Breitbart M, Rohwer F.. Here a virus, there a virus, everywhere the same virus? Trends Microbiol 13: 278–284

Article *in* Trends in Microbiology · July 2005

DOI: 10.1016/j.tim.2005.04.003 · Source: PubMed

CITATIONS
326

READS
723

2 authors, including:

Mya Breitbart
University of South Florida
148 PUBLICATIONS **8,785** CITATIONS

SEE PROFILE

Review TRENDS in Microbiology Vol.13 No.6 June 2005

Here a virus, there a virus, everywhere the same virus?

Mya Breitbart[1] and Forest Rohwer[1,2]

[1]Biology Department, LS301, San Diego State University, 5500 Campanile Dr, San Diego, CA 92182 USA
[2]Center for Microbial Sciences, San Diego State University, 5500 Campanile Dr, San Diego, CA 92182 USA

There are an estimated 10^{31} viruses on Earth, most of which are phages that infect bacteria. Metagenomic analyses have shown that environmental viral communities are incredibly diverse. There are an estimated 5000 viral genotypes in 200 liters of seawater and possibly a million different viral genotypes in one kilogram of marine sediment. By contrast, some culturing and molecular studies have found that viruses move between different biomes. Together, these findings suggest that viral diversity could be high on a local scale but relatively limited globally. Also, by moving between environments, viruses can facilitate horizontal gene transfer.

Introduction

Until the 1970s, environmental microbiology consisted mainly of culturing microbes (see Glossary) from different environments. These studies suggested that most environments contained relatively few microbes. Based on these low numbers, it was thought that microbes and their associated activities probably had minor roles in global carbon and nutrient cycles. The introduction of direct counts by Hobbie *et al.* dramatically changed this view by showing that culturing had underestimated microbial abundances by ~100–1000-fold [1] (Figure 1a). Direct counts were complimented by measurements of microbial activity and production using radiolabeled compounds [2–5]. Together, these studies showed that microbes are major players in global carbon cycles*. In the ocean, for example, at least 50% of the CO_2 fixed by photosynthesis each day ends up supporting microbial respiration and the production of new cells (Figure 1c).

Abundance and importance of viruses in the environment

The discovery that there were millions of microbes in every milliliter of seawater and gram of sediment, begged the question, 'Who is eating them?' Early studies showed that protists, primarily nanoflagellates, were important grazers of marine microbes [6–10]. Viruses entered the picture in 1989, when Bergh *et al.* used transmission electron microscopy to show that there are ~10 million virus-like particles (VLPs) per ml of seawater [11]. Earlier studies by Anderson *et al.* [12,13] had also shown that viruses are extremely abundant in the marine environment. At the time, however, this observation did not make sense because the vast pool of environmental microbes had yet to be discovered. Most environmental VLPs are assumed to be phages, viruses that infect bacteria, because bacteria are the most common prey items. In addition, viral abundance often correlates strongly with

Glossary

Direct counts: The introduction of direct counts drastically changed environmental microbiology by revealing that culture-based techniques had underestimated the number of microbes present in the environment by ~100–1000-fold. For direct counts, a water sample is typically pulled through a 0.02 μm filter to concentrate the microbes and viruses onto one plane. An epifluorescent stain that binds to nucleic acid is then applied and the number of microbes and viruses are determined using epifluorescent microscopy. The fluorescent signal is being visualized, and not the actual particle, so viruses, which are smaller than the 200 nm limit possible with visible light can be observed.

Viruses, virus-like particles and phages: Viruses are biological entities that infect cells and replicate themselves. Outside the cell, viruses exist as free particles called virions, which are typically RNA or DNA genomes surrounded by a protein shell called a capsid. Viruses differ from other life forms by the fact that the virions are disassembled during their intercellular, replicative phase. Phages are viruses that infect Bacteria. The term virus-like particles (VLPs) is used to describe small dots observed when performing direct counts on environmental samples (Figure 1a), because it is not absolutely known if these dots represent viruses or other small pieces of nucleic acid.

Microbes, prokaryotes, Bacteria, Archaea and Eukarya: Microbes are any life form that can only be observed with a microscope, including single-celled organisms and viruses. Bacteria, Archaea and Eukarya are the three Domain names proposed by Woese, based on analyses of ribosomal RNA sequences [72]. Prokaryote is a term that refers to cellular organisms without a nucleus or organelles. Traditionally, prokaryote was used to designate Bacteria and Archaea but recent discoveries of organelles in these two Domains have rendered this term obsolete. For the purpose of this review, we have chosen to use the term microbe, most of which are Bacteria and Archaea.

Metagenomics: The term 'metagenomics' was first coined by Jo Handelsman *et al.* to describe 'the functional and sequence-based analysis of the collective microbial genomes contained in an environmental sample' [73].

Viral species, genotype and groups: A viral species was formally defined by van Regenmortel as '…a polythetic class of viruses that constitutes a replicating lineage and occupies a particular ecological niche' [74]. This is the definition used by the International Committee on Taxonomy of Viruses (ICTV) [52] and functionally requires that the virus be in culture. A viral 'genotype' was used by Breitbart *et al.* [35] because metagenomic data is not compatible with the formal definition of a viral species. Currently, genotype refers to *in silico* conditions that prevent different phage genomes from assembling together when they are randomly fragmented. As more data is accumulated, this definition will change until a metagenomic-based definition of a viral species is established. Groups and subgroups are used to designate large orderings of phages based on their genomes [26]. These designations are roughly equivalent to families and subfamilies in the ICTV system [52].

MORANS and ORFans: are single genes that become incorporated into phage or microbial genomes. MORON means 'more DNA' and is used to describe a short region of extra DNA, and are often identified based on their presence in one phage genome and absence in a closely related phage genome [55]. ORFans are protein-encoding regions with no detectable sequence similarity to proteins from other genomes [66].

Corresponding author: Rohwer, F. (forest@sunstroke.sdsu.edu).
Available online 30 April 2005

* The discussion presented here will focus on marine microbes, where many of the early studies were first performed. Many of the lessons from marine microbial ecology are directly applicable to other environments.

 0966-842X/$ - see front matter doi:10.1016/j.tim.2005.04.003

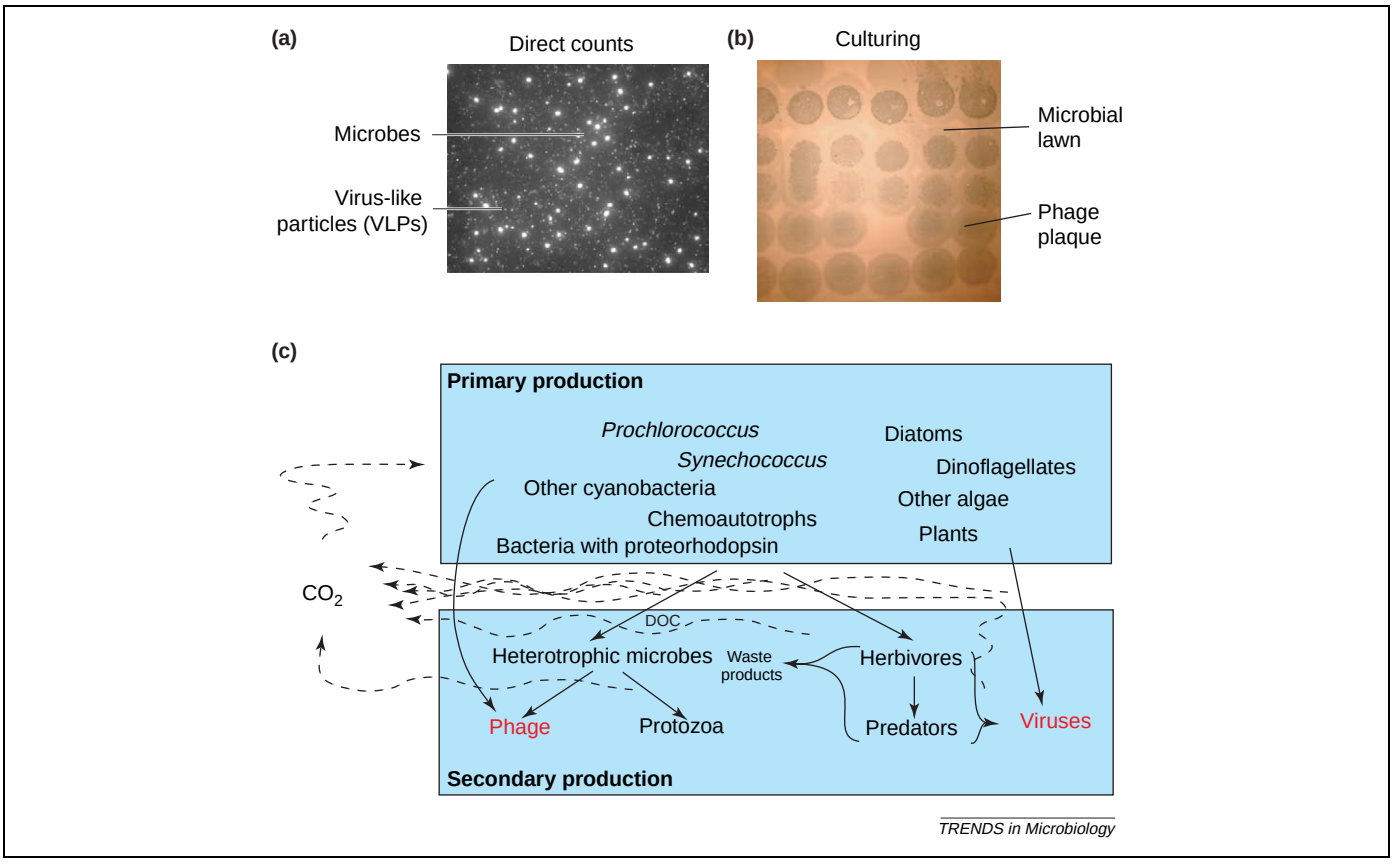

Figure 1. Methods and models used to study viral ecology **(a)** Epifluorescence micrograph showing microbes and virus-like particles (VLPs) from seawater stained with a nucleic acid stain. The microbes are the larger brighter spots, whereas the VLPs are small pinprick dots. **(b)** Phage isolation accomplished by plaquing on a bacterial lawn. A bacterial host is first cultured from the environment of interest and then plated as a lawn on top of an agar plate. Phage are spotted on top of this lawn and the clear areas represent regions where phage have lysed the host. **(c)** Ecological roles of viruses. Carbon dioxide is fixed into sugars by primary producers. The sugars and downstream metabolites are then used to make biomass (production), perform metabolism (respiration) or are released into the environment as dissolved organic carbon (DOC). Secondary producers, such as herbivores, eat the new primary producer biomass and heterotrophic microbes eat the DOC. Herbivores are then eaten by predators and the microbes are grazed by protists or viruses. A rough rule of thumb is that ~50% of the fixed carbon is shunted into heterotrophic microbes via DOC and approximately half of the new microbes are killed by viruses. This means that ~25% of the fixed carbon is respired as CO_2 due to viral lysis. Viruses also directly kill primary producers and other secondary producers.

microbial abundance, with a fairly constant ratio of 5–10 VLPs per bacteria [14]. Further studies showed that phages kill between 4–50% of the bacteria produced every day [15–17], with the rest being eaten by protists [18].

In this review we will discuss genomic and metagenomic approaches to studying environmental viral diversity. Evidence is presented that local viral diversity is high, whereas global diversity is relatively low because viruses are moving between environments. A model rectifying these apparent contradictions is proposed. Finally, the potential influence of viral movement on evolution is discussed in the context of microbial metagenomes.

Culture-based studies of viral diversity

Culture-based studies of phages are usually performed by 'plaquing'. First, the bacterial host is grown in liquid culture and then it is mixed with a sample containing the phage. This mixture of phage plus host is then resuspended into top agar and poured onto a plate of media. The top agar is loose enough that phage can diffuse through it and attack nearby bacteria. Where this occurs, the microbial lawn is killed off and a clear area appears. This is called a 'plaque' (Figure 1b). Similar techniques are used for culturing eukaryotic and archaeal viruses. Using this approach, it has been shown that at least one, and usually multiple, viruses can infect any given marine microbial host [19–23].

Culture-independent studies of viral diversity

Conserved gene studies of viral diversity

Studies of specific microbial species in the environment were severely limited by culturing bias until Pace *et al.* introduced 16S rDNA analyses [24]. Using PCR amplification and sequencing of the 16S rDNA locus, it has been shown that most microbes in the environment belong to uncultured groups (reviewed in [25]). The fact that most microbes have not been cultured has severely limited studies of viral diversity. As a further complication, there is not a single genetic element shared by all viruses [26], meaning that it is not possible to study total viral diversity using 16S rDNA-like approaches. However, whole genome comparisons have shown that there are conserved genes shared amongst all members within certain viral taxonomic groups. These conserved genes (or signature genes [26]) have been used to study diversity within known groups of viruses among cultured isolates, as well as in the environment. The diversity of phages infecting cyanobacteria, for example, has been studied extensively using sequences of structural proteins [27–30]. The diversity of

algal viruses and T7-like Podophages (a group of genetically related, tailed, dsDNA phages) has been examined by sequencing DNA polymerase genes [27,31–33], and a RNA-dependent RNA polymerase fragment has been used to identify at least two novel groups of marine picorna-like viruses that probably infect eukaryotic algae [34]. The vast majority of sequences found in these studies belong to novel subgroups that are not represented by cultured isolates. All of the conserved gene studies suggest that environmental viral diversity is high and essentially uncharacterized.

Metagenomic studies of viral diversity

Studies employing conserved genes do not enable the discovery of completely novel groups of viruses. To assess total dsDNA viral diversity, Breitbart *et al.* partially sequenced shotgun libraries from uncultured viral communities isolated from seawater [35], marine sediment [36] and human fecal matter [37]. Recently, Cann *et al.* partially sequenced a viral metagenome from equine feces [38]. Approximately 75% of the sequences in these viral metagenomes did not match any genes in the database (i.e. E-value >0.001 to the non-redundant GenBank database), suggesting that most viral diversity remains uncharacterized. Using the sequences that could be associated with a particular phage group, these studies demonstrated that certain phage groups (T7-like Podophage, λ-like Siphophage, and T4-like Myophage) were abundant in all the marine communities, whereas other, unrelated groups (SFI21-like and TP901-like Siphophage) were more common in fecal matter [37]. The phage groups most abundant in fecal matter are known to infect Gram-positive bacteria, whereas the marine phage groups infect Gram-negative bacterial hosts. This suggests that the most abundant phage groups in a given environment reflect the kinds of microbes found in that environment.

Sequencing viral metagenomic libraries also provides information about the underlying community structure (i.e. the number of genotypes and relative abundances; Figure 2). Occurrence of the same DNA sequence in different shotgun fragments means that the same viral genotype has been sampled multiple times. To take advantage of this information, a modified version of the Lander-Waterman algorithm [39] was developed to model uncultured viral communities based on metagenomic data [35]. These analyses show that uncultured viral communities are some of the most diverse communities ever observed. There might be, for example, up to a million different viral genotypes in one kilogram of marine sediment [36]. Together, these culture-independent studies of viral diversity show that the majority of viral diversity remains unknown, with most viruses belonging to novel groups without any cultured representatives.

Movement of viruses between biomes

If all environments have unique, endemic viral populations, then extrapolation from metagenomic data predicts global viral diversity to be ~100 million distinct viral genotypes [40]. Alternatively, viruses might be moving between environments. In this case, local diversity could be quite high but global diversity would be relatively limited. To differentiate between these two hypotheses, conserved genes have been used to show that identical, or nearly-identical, phage-encoded sequences are present in different biomes (e.g. marine versus freshwater versus soil) [27,33]. These phage-encoded sequences are so similar that they must have moved between environments within recent evolutionary time. For example, one phage-encoded DNA polymerase sequence, named HECTOR, was found in marine water, soil, rumen fluid, associated with corals and in solar saltern water [33]. Assuming an average burst size of 25 particles and an average half-life for phages of 48 h [14], a phage encoding the HECTOR sequence would need to complete a lytic cycle once every ten days (i.e. five half-lives) to survive. This phage would complete ~36 generations per year. The mutation rate for dsDNA phage-sized genomes is 10^{-7}–10^{-8} changes bp^{-1} per generation [41–43]. Therefore, it is expected that there would be a 1-bp change in the HECTOR sequence every ~525 years. The HECTOR sequence never differed by >3 bp over a 533-bp fragment [33], therefore, this sequence has moved between these environments within the last 1000–2000 years.

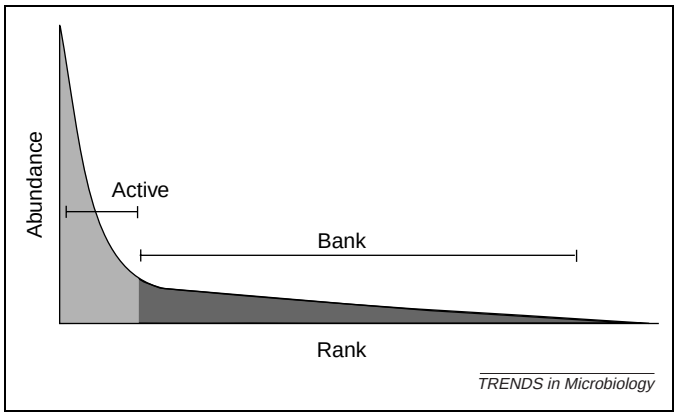

Figure 2. Example of a rank-abundance curve, which is a typical way of representing genotype distribution within a community. The most abundant genotype is ranked as 1, the second most abundant genotype is ranked as 2, and so on. Because the relative abundances of different viral genotypes are constantly changing, the curve represents a snapshot of the community in time. In the Bank model, only a few of the most abundant viral genotypes are in the Active fraction. As new prey items become dominant in response to changing environmental conditions, the viruses that can prey on those hosts also become abundant. The viruses that were previously in the Active fraction begin to decay and in the absence of new production become part of the bank fraction.

It is not known whether a complete HECTOR-encoding phage, or just this piece of DNA, is moving between environments (e.g. as suggested by the mosaic model [44]). Sano *et al.* [45] directly determined if viruses could move between biomes by incubating viral communities from soil, sediment and freshwater with marine microbial communities. Within 48 h, the viral communities were growing rapidly, showing that viruses can find hosts in different biomes. Two possible explanations for this observation are: (i) identical microbial hosts are found in the different environments; or (ii) viruses are not completely host-specific and can attack the different microbial hosts found in each environment. Currently, the second possibility seems more likely because cross-infecting viruses are relatively common [23,46,47]. If most viruses actually do attack multiple hosts, this will

dramatically reduce estimates of global viral diversity and change models of virus–host dynamics.

Viral community structure

Here we propose the Bank model (Figure 2), which could explain the observation that viruses are both globally distributed and have high diversity on the local scale. In the Bank model, only the most abundant viruses are active. The rest of the viruses are inactive, rare and form a potential population for recruitment, much like a seed-bank in plant populations. This distribution matches the observed rank-abundance curves predicted by modeling of marine metagenomic libraries, where the most abundant viral genotype comprises $<5\%$ of the total community and the vast majority of viral genotypes are extremely rare ($<0.01\%$ of the community) [35,36]. When an environment changes, different hosts grow and the viruses preying on these hosts move from the 'bank' into the 'active' fraction. The previously abundant and active viruses start to decay and enter the bank fraction. The Bank model predicts that the viruses in the active fraction should stay the same as long as the host populations remain susceptible to these viruses. Changing the active host population is predicted to cause a rapid change in the active viral fraction. In other words, you get what you culture. Supporting this model, culturing and genome size distribution studies provide evidence that specific viruses become abundant, decay to undetectable levels and then return on seasonal cycles [47–49].

In the Bank model, the active viral–host pairs are behaving in a 'Kill-the-Winner' fashion [50], where the most abundant host population is reduced by its viral predators. Population reduction of the most abundant host creates an open niche, which enables a new host to become abundant. In this manner, the identity of the most abundant host is constantly changing. The Bank and Kill-the-Winner models only vary in their treatment of the inactive fraction. In particular, how are the rare viruses maintained in the bank fraction? In the absence of new production, virions decay rapidly, with an average half-life of 48 h in the marine environment [14]. However, data from some viral decay studies show that virions display a rapid initial decay, which is followed by a period of slower decay, possibly because the remaining viral particles are more refractory to harmful environmental factors [15,51]. The viruses in the bank fraction were probably previously active and are now in various states of population decline. Alternatively, the rare viruses in the bank fraction might be supplied via constant production of a low level of virions (e.g. induction of proviruses) or allochthonous inputs.

Phage specialization to different environments

In general, officially recognized phage taxa [52] share similar suites of genes [26]. These genes are usually grouped together into modules [53] and are likely optimized to work together. Examples of modules are the virion assembly genes or DNA replication machinery. It is relatively easy for phages to mix and match these modules and produce viable recombinants [54].

Phage modules are 'fuzzy' entities and they often have other genes inserted into them. These genes have been termed MORONs (for more DNAs) and they can come from other phages or hosts [55]. Insertion of these extra DNA elements can occur via different illegitimate recombination mechanisms [56] that do not destroy the essentia workings of the module. Over evolutionary time, MORONs are deleted [57] or they can be maintained by positive selection. In the second case, the behavior of the phage and its host can be dramatically affected. A well-known example of this phenomenon is the acquisition of exotoxin genes by phages. These exotoxins convert the phage-infected microbes into pathogens [58].

Phages adapt to new environments by acquiring ecologically important genes as MORONs. The cyanobacteria *Synecococcus* and *Prochlorococcus* are major marine autotrophs. Photosynthesis by these two genera account for approximately one-third of the carbon fixed in the marine environment. Cyanophages infecting *Synecococcus* and *Prochlorococcus* have acquired genes involved in photosynthesis [59,60]. The *psbA* gene, which encodes the D1 protein, has been found in all three of the completely sequenced *Prochlorococcus* phage genomes and one *Synecococcus* phage genome [59]. D1 is a rate-limiting photosynthesis protein and host-derived D1 concentrations dramatically drop during phage infection [61]. Expression of the phage-encoded *psbA* gene enables the phage to maintain photosynthesis throughout the infection cycle, presumably providing the phage with energy. In the open ocean, the energy benefit associated with the acquisition of photosynthesis genes by cyanophages was probably a key evolutionary step. Similarly, phosphate is a major limiting nutrient in parts of the ocean and many marine phages encode enzymes involved in phosphate metabolism (e.g. *phoH*, RNA reductases, endonucleases) [62–64]. The photosynthesis and phosphate metabolism genes are embedded into phage genomes that have clear relationships to well-studied phages that infect *Escherichia coli*. The take-home message is that phages acquire ecologically important genes to adapt to new environments [62]. Phages can carry these genes between environments, contributing to local and global lateral gene transfer.

Relationships between metagenomes and mobile elements

A metagenome does not necessarily reveal the phenotype of the source community because the expression and interactions of the genes are dependent on their arrangements within genomes. The basic units of organization in genomes are operons in microbial genomes [65] and modules in phages (Figure 3) [53]. Although technically different, these units (or gene clusters) are similar, relatively autonomous and can move around while retaining their functionality. Distinct microbial and phage genomes are assembled by mixing and matching these clusters via homologous recombination. Interaction rates between different operons and modules are increased by professional mobile elements, including plasmids, pathogenicity islands and temperate phages (represented by prophages within microbial genomes). Unlike operons and modules, MORONs and ORFans are single genes that become incorporated into phage or

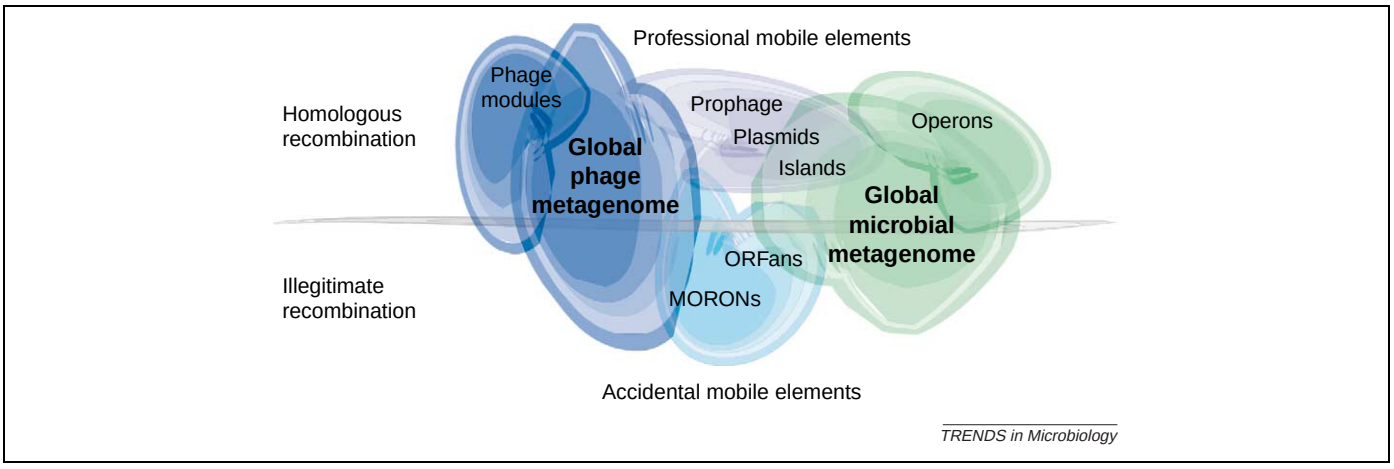

Figure 3. Relationships between metagenomes, operons, modules and mobile elements. Metagenomes are populations of genes within a particular sample. The basic arrangement of these genes are modules in phages and operons in microbes. Professional mobile elements facilitate the movement of modules and operons between and within genomes. Any gene can potentially become an accidental mobile element via illegitimate recombination. Examples of accidental genes include specialization genes like exotoxins, *psbA* and *phoH*.

microbial genomes via illegitimate recombination [66]. Usually these sequences are lost through deletion, but occasionally these genes are maintained by positive selection (e.g. specialization genes). The ecological and evolutionary barriers between phage and microbial metagenomes are blurred by these horizontal gene transfer events.

Conclusions

Over the past three decades, it has become apparent that viruses are ubiquitious, abundant and ecologically important in the environment. Considering that viruses are the most abundant biological entities on the planet, surprisingly little is known about their identity or diversity. Culture-independent methods have revealed a wealth of environmental viral diversity. For example, standard diversity indices based on mathematical models of metagenomic data predict that the viruses found in only 1 kg of nearshore marine surface sediment are more diverse than all the reptiles known on the planet [36]. Although local viral diversity is extremely high, viruses appear to be moving between environments, which constrains total global viral diversity and provides a conduit for horizontal gene transfer. Mathematical modeling from metagenomic data suggests that most viral genotypes are relatively rare, serving as a 'Bank' of viruses that can become abundant when their hosts become dominant (e.g. in response to changing environmental conditions).

Armed with new techniques for detecting and enumerating viruses, researchers have found viruses in a wide range of environments, including the deep sea [67], solar salterns (which are ten times saltier than the ocean) [68], acidic hot springs (>80°C with pH=3.0) [69], alkaline lakes (pH=10) [70], under >30 m of ice in polar lakes [71] and in the terrestrial subsurface (>2000 m deep; Wegley *et al.*, unpublished). Future research needs to examine the relationships between uncultured viruses found in different environments, as well as the spatial and temporal scales on which viral communities change. More than simply providing sequence data, metagenomic analyses can offer insights into biogeographical distributions, community structure and ecological dynamics. In addition, the crucial genes for microbial adaptation to a given environment are likely to be moved by phages and can therefore be identified by analyzing the genomic content of viral communities. Continued metagenomic studies of viral communities will enable the estimation of global viral diversity, as well as a deeper understanding of the impact of horizontal gene transfer on microbial diversity and evolution.

Acknowledgements

We thank the NSF, the Gordon and Betty Moore Foundation, and the Environmental Protection Agency STAR Fellowship for support.

References

1 Hobbie, J. *et al*. (1977) Use of nuclepore filters for counting bacteria by epifluorescence microscopy. *Appl. Environ. Microbiol.* 33, 1225–1228
2 Azam, F. and Hodson, R.E. (1977) Dissolved ATP in the sea and its utilization by marine bacteria. *Nature* 267, 696–698
3 Riemann, B. (1978) Differentiation between heterotrophic and photosynthetic plankton by size fractionation, glucose uptake, ATP and chlorophyll content. *Oikos* 31, 358–367
4 Karl, D.M. and Bossard, P. (1985) Measurement and significance of ATP and adenine nucleotide pool turnover in microbial cells and environmental samples. *J Microbiol Meth* 3, 125–139
5 Fuhrman, J.A. and Azam, F. (1982) Thymidine incorporation as a measure of heterotrophic bacterioplankton production in marine surface waters: evaluation and field results. *Mar. Biol.* 66, 109–120
6 Caron, D. and Goldman, J. (1988) Dynamics of protistan carbon and nutrient cycling. *J. Protozool.* 35, 247–249
7 Berman, T. *et al*. (1987) Nutrient flux between bacteria, bacterivorous nanoplanktonic protists and algae. *Mar. Microb. Food Webs* 2, 69–82
8 Sherr, B.F. *et al*. (1984) Abundance and productivity of heterotrophic nanoplankton in Georgia coastal waters. *J. Plankton Res.* 6, 195–202
9 Bjørnsen, P.K. *et al*. (1988) Trophic interactions between heterotrophic nanoflagellates and bacterioplankton in manipulated seawater enclosures. *Limnol. Oceanogr.* 33, 409–420
10 Fuhrman, J.A. and McManus, G.B. (1984) Do bacteria-sized marine eucaryotes consume significant bacterial production? *Science* 224, 1257–1260
11 Bergh, Ø. *et al*. (1989) High abundance of viruses found in aquatic environments. *Nature* 340, 467–468
12 Anderson, N.G., ed. (1967) *Isolation of viral particles from large volumes*. In *Transmission of viruses by the Water Route* (Berg, G., ed.), pp. 75–88, Interscience Publishers

13 Anderson, N.G. and Cline, G.B. (1967) New centrifugal methods for virus isolation. In *Methods in Virology* (Vol. II) (Maramorosch, M.K. and Koprowski, H., eds), pp. 137–178, Academic Press
14 Wommack, K. and Colwell, R. (2000) Virioplankton: Viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114
15 Heldal, M. and Bratbak, G. (1991) Production and decay of viruses in aquatic environments. *Mar. Ecol. Prog. Ser.* 72, 205–212
16 Steward, G.F. *et al*. (1992) Estimation of virus production in the sea: I. Method development. *Mar. Microb. Food Webs* 6, 57–78
17 Steward, G.F. *et al*. (1992) Estimation of virus production in the sea: II. Field results. *Mar. Microb. Food Webs* 6, 79–90
18 Fuhrman, J.A. and Noble, R.T. (1995) Viruses and protists cause similar bacterial mortality in coastal seawater. *Limnol. Oceanogr.* 40, 1236–1242
19 Moebus, K. and Nattkemper, H. (1981) Bacteriophage sensitivity patterns among bacteria isolated from marine waters. *Helgoländer Meeresunters* 34, 375–385
20 Moebus, K. (1991) Preliminary observations on the concentration of marine bacteriophages in the water around Helgoland. *Helgoländer Meeresunters* 45, 411–422
21 Moebus, K. (1992) Further investigation on the concentration of marine bacteriophages in the water around Helgoland with reference to the phage-host systems encountered. *Helgoländer Meeresunters* 46, 275–292
22 Suttle, C.A. and Chan, A.M. (1993) Marine cyanophage infecting oceanic and coastal strains of *Synechococcus*: abundance, morphology, cross-infectivity, and growth characteristics. *Mar. Ecol. Prog. Ser.* 92, 99–109
23 Sullivan, M. *et al*. (2003) Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* 424, 1047–1051
24 Lane, D. *et al*. (1985) Rapid determination of 16S rRNA sequences for phylogenetic analysis. *Proc. Natl. Acad. Sci. U. S. A.* 82, 6955–6959
25 Giovannoni, S. and Rappe, M. (2000) Evolution, diversity, and molecular ecology of marine prokaryotes. In *Microbial Ecology of the Oceans* (Kirchman, D.L., ed.), pp. 47–84, John Wiley and Sons
26 Rohwer, F. and Edwards, R. (2002) The phage proteomic tree: A genome based taxonomy for phage. *J. Bacteriol.* 184, 4529–4535
27 Short, C. and Suttle, C. (2005) Nearly identical bacteriophage structural gene sequences are widely distributed in marine and freshwater environments. *Appl. Environ. Microbiol.* 71, 480–486
28 Zhong, Y. *et al*. (2002) Phylogenetic diversity of marine cyanophage isolates and natural virus communities as revealed by sequences of viral capsid assembly protein gene g20. *Appl. Environ. Microbiol.* 68, 1576–1584
29 Fuller, N.J. *et al*. (1998) Occurence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. *Appl. Environ. Microbiol.* 64, 2051–2060
30 Dorigo, U. *et al*. (2004) Cyanophage diversity, inferred from g20 gene analyses, in the largest natural lake in France, Lake Bourget. *Appl. Environ. Microbiol.* 70, 1017–1022
31 Chen, F. *et al*. (1996) Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. *Appl. Environ. Microbiol.* 62, 2869–2874
32 Chen, F. and Suttle, C.A. (1996) Evolutionary relationships among large double-stranded DNA viruses that infect microalgae and other organisms as inferred from DNA polymerase genes. *Virology* 219, 170–178
33 Breitbart, M. and Rohwer, F. (2004) Global distribution of nearly identical phage-encoded DNA sequences. *FEMS Microbiol. Lett.* 236, 249–256
34 Culley, A. *et al*. (2003) High diversity of unknown picorna-like viruses in the sea. *Nature* 424, 1054–1057
35 Breitbart, M. *et al*. (2002) Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14250–14255
36 Breitbart, M. *et al*. (2004) Diversity and population structure of a nearshore marine sediment viral community. *Proc. R. Soc. Lond. B. Biol. Sci.* 271, 565–574
37 Breitbart, M. *et al*. (2003) Metagenomic analyses of an uncultured viral community from human feces. *J. Bacteriol.* 185, 6220–6223
38 Cann, A. *et al*. (2005) Analysis of the virus population present in equine faeces indicates the presence of hundreds of uncharacterized virus genomes. *Virus Genes* 30, 151–156
39 Lander, E.S. and Waterman, M.S. (1988) Genomic mapping by fingerprinting random clones: a mathematical analysis. *Genomics* 2, 231–239
40 Rohwer, F. (2003) Global phage diversity. *Cell* 113, 141
41 Drake, J.W. *et al*. (1998) Rates of Spontaneous Mutation. *Genetics* 148, 1667–1686
42 Drake, J.W. and Holland, J.J. (1999) Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13910–13913
43 Sniegowski, P.D. *et al*. (2000) The evolution of mutation rates: separating causes from consequences. *BioEssays* 22, 1057–1066
44 Hendrix, R.W. *et al*. (1999) Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2192–2197
45 Sano, E. *et al*. (2004) Movement of viruses between biomes. *Appl. Environ. Microbiol.* 70, 5842–5846
46 Jensen, E. *et al*. (1998) Prevalence of broad-host-range lytic bacteriophages of *Sphaerotilus natans*, *Escherichia coli*, and *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* 64, 575–580
47 Marston, M. and Sallee, J. (2003) Genetic diversity and temporal variation in the cyanophage community infecting marine *Synechococcus* species in Rhode Island's coastal waters. *Appl. Environ. Microbiol.* 69, 4639–4647
48 Wommack, K.E. *et al*. (1999) Hybridization analysis of Chesapeake Bay virioplankton. *Appl. Environ. Microbiol.* 65, 241–250
49 Wommack, K.E. *et al*. (1999) Population dynamics of Chesapeake Bay virioplankton: Total-community analysis by pulsed-field gel electrophoresis. *Appl. Environ. Microbiol.* 65, 231–240
50 Thingstad, T.F. and Lignell, R. (1997) Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquat. Microb. Ecol.* 13, 19–27
51 Mathias, C. *et al*. (1994) Seasonal variations of virus abundance and viral control of the bacterial production in a backwater system of the Danube river. *Appl. Environ. Microbiol.* 61, 3734–3740
52 Murphy, F.A. *et al*., eds (1995) *Virus Taxonomy: Sixth Report of the International Committee on Taxonomy of Viruses*, Springer-Verlag
53 Campbell, A. and Botstein, D. (1983) Evolution of the lambdoid phage. In *Lambda II* (Hendrix, R., ed.), pp. 365–380, Cold Spring Harbor Laboratory Press
54 Botstein, D. and Herskowitz, I. (1974) Properties of hybrids between *Salmonella* phage P22 and coliphage lambda. *Nature* 251, 584–589
55 Juhala, R.J. *et al*. (2000) Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *J. Mol. Biol.* 299, 27–51
56 Mosig, G. *et al*. (2001) Two recombination-dependent DNA replication pathways of bacteriophage T4, and their roles in mutagenesis and horizontal gene transfer. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8306–8311
57 Hendrix, R.W. *et al*. (2000) The origins and ongoing evolution of viruses. *Trends Microbiol.* 8, 504–508
58 Davis, B.M. and Waldor, M.K. (2002) Mobile genetic elements and bacterial pathogenesis. In *Mobile DNA II* (Craig, N.L., ed.), pp. 1040–1055, ASM Press
59 Mann, N. *et al*. (2003) Marine ecosystems: bacterial photosynthesis genes in a virus. *Nature* 424, 741
60 Lindell, D. *et al*. (2004) Photosynthesis genes in *Prochlorococcus* cyanophage. *Proc. Natl. Acad. Sci. U. S. A.* 101, 11013–11018
61 Bailey, S. *et al*. (2004) Cyanophage infection and photoinhibition in marine cyanobacteria. *Res. Microbiol.* 155, 720–725
62 Sullivan, M. *et al*. Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biol.* 3, e144
63 Rohwer, F. *et al*. (2000) The complete genomic sequence of the marine phage Roseophage SIO1 shares homology with non-marine phages. *Limnol. Oceanogr.* 42, 408–418
64 Chen, F. and Lu, J. (2002) Genomic sequence and evolution of marine cyanophage P60: A new insight on lytic and lysogenic phages. *Appl. Environ. Microbiol.* 68, 2589–2594
65 Jacob, F. and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–356
66 Fischer, D. and Eisenberg, D. (1999) Finding families for genomic ORFans. *Bioinformatics.* 15, 759–762
67 Danovaro, R. and Serresi, M. (2000) Viral density and virus-to-bacterium ratio in deep-sea sediments of the Eastern Mediterranean. *Appl. Environ. Microbiol.* 66, 1857–1861

68 Guixa-Boixareu, N. *et al*. (1996) Viral lysis and bacterivory as prokaryotic loss factors along a salinity gradient. *Aquat. Microb. Ecol.* 11, 21–227
69 Rice, G. *et al*. (2001) Viruses from extreme thermal environments. *Proc. Natl. Acad. Sci. U. S. A.* 98, 13341–13345
70 Jiang, S. *et al*. (2003) Abundance, distribution, and diversity of viruses in alkaline, hypersaline Mono Lake, California. *Microb Ecol.* 47, 9–17
71 Kepner, R.J. *et al*. (1998) Viruses in Antarctic lakes. *Limnol. Oceanogr.* 43, 1754–1761
72 Woese, C. *et al*. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U. S. A.* 87, 4576–4579
73 Riesenfeld, C. *et al*. (2004) METAGENOMICS: Genomic analysis of microbial communities. *Annu. Rev. Genet.* 38, 525–552
74 van Regenmortel, M. (1990) Virus species, a much overlooked but essential concept in virus classification. *Intervirology* 31, 241–254