MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, Dec. 2011, p. 610–635 Vol. 75, No. 4
1092-2172/11/$12.00 doi:10.1128/MMBR.00011-11

Genomics of Bacterial and Archaeal Viruses: Dynamics within the Prokaryotic Virosphere

Mart Krupovic,[1]* David Prangishvili,[1] Roger W. Hendrix,[2] and Dennis H. Bamford[3]

Institut Pasteur, Unité Biologie Moléculaire du Gène chez les Extrêmophiles, 25 rue du Dr. Roux, 75015 Paris, France[1]; Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260[2]; and Department of Biosciences and Institute of Biotechnology, Viikki Biocenter 2, P.O. Box 56 (Viikinkaari 5), University of Helsinki, FIN-00014 Helsinki, Finland[3]

INTRODUCTION

Over the past few years, the viruses of prokaryotes have been transformed in the view of microbiologists from simply being convenient experimental model systems into being a major component of the biosphere. They are the global champions of diversity, they constitute a majority of organisms on the planet, they have large roles in the planet's ecosystems, they exert a significant—some would say dominant—force on the evolution of their bacterial and archaeal hosts, and they have been doing this for billions of years, possibly for as long as there have been cells. This transformation in status or, rather, our expanded appreciation of the importance of these viruses in the biosphere is due to a few significant developments in both understanding and technology. (i) It has become clear that the population sizes of these viruses are astoundingly large. This realization grew out of electron microscopic enumerations of tailed phage virions in costal seawater, and numerous measurements in other environments have been made since then (18, 257, 267). A current estimate based on these measure-

* Corresponding author. Mailing address: Institut Pasteur, Unité Biologie Moléculaire du Gène chez les Extrêmophiles, Département de Microbiologie, 25 rue du Dr. Roux, 75015 Paris, France. Phone: 33 1 45 68 88 33. Fax: 33 1 45 68 88 34. E-mail: krupovic@pasteur.fr.

TABLE 1. Overview of bacterial virus families

Family	Capsid morphology	Additional feature(s)	Genome type	No. of complete genomes[a]	Example
Myoviridae	Icosahedral	Tail (contractile)	dsDNA, L	134	T4
Siphoviridae	Icosahedral	Tail (long noncontractile)	dsDNA, L	268	λ
Podoviridae	Icosahedral	Tail (short noncontractile)	dsDNA, L	98	T7
Tectiviridae	Icosahedral	Internal membrane	dsDNA, L	4	PRD1
Corticoviridae	Icosahedral	Internal membrane	dsDNA, C	1	PM2
Plasmaviridae	Pleomorphic	Enveloped	dsDNA, C	1	L2
Microviridae	Icosahedral	Nonenveloped	ssDNA, C	15	ϕX174
Inoviridae	Filamentous	Long flexible or short rigid	ssDNA, C	29	M13
Cystoviridae	Icosahedral	Enveloped, multilayered	dsRNA, L, S	5	ϕ6
Leviviridae	Icosahedral	Nonenveloped	ssRNA, L	7	MS2

[a] The number of complete genome sequences was obtained from GenBank and does not include the genome sequences available for different isolates of the same virus strain. Families *Myoviridae*, *Siphoviridae*, and *Podoviridae* are grouped into the order *Caudovirales*. L, linear; C, circular; S, segmented.

ments is that there are $\sim 10^{31}$ individual tailed phage virions in the global biosphere—enough to reach for 200 million light years if laid end to end—and measurements of population turnover suggest that it takes roughly 10^{24} productive infections per second to maintain the global population (108, 267). Numbers like these profoundly affect the ways in which we can think about viral population dynamics and evolution and about the roles that these viruses play in global ecology. We note that the numbers cited here, impressive though they are, substantially underestimate viral abundance in that they pertain only to the tailed phages, for which such numbers are most easily determined; they do not include any of the numerous other groups of bacterial and archaeal viruses or the viral genomes found in most prokaryotic cells as proviruses (35, 36, 197). (ii) Advances in DNA sequencing technology have led to dramatic qualitative improvements in how we understand the genetic structure of viral populations, the mechanisms of viral evolution, and the diversity of viral sequences. The majority of newly determined gene and protein sequences of these viruses has no relatives detectable in the public sequence databases, and analysis of metagenomic data provides strong evidence that there is more genetic diversity in the genes of the viruses of prokaryotes than in any other compartment of the biosphere. (iii) Facilitated by these conceptual and technical advances, studies of bacterial and archaeal viruses as important components of global biology have flourished. These viruses are revealed as important players in carbon and energy cycling in the oceans and other natural environments and as major agents in the ecology and evolution of their cellular hosts (34, 59, 235). (iv) The isolation and characterization of new viruses have accelerated (22, 207, 294). This has been especially important for the archaeal viruses, where the discovery of new viruses and of new virus types had lagged behind bacteriophage discovery (205, 213). For the bacteriophages, the isolation of newly discovered viruses has helped improve the still extremely sparse coverage of sequence diversity and the narrow phylogenetic range of hosts represented by current data (38, 39, 42, 65, 102, 187). (v) High-resolution structures determined for capsid proteins and other virion proteins, together with information about virion assembly mechanisms, have allowed surprising inferences about ancestral connections among genes whose DNA sequences and encoded protein sequences have diverged to the point that they are no longer detectably related. This has opened previously unavailable possibilities for an understanding of the early history of viruses (9, 12, 109, 145).

If there is one theme that characterizes the developments described above, it is "diversity": diversity of gene sequences; diversity of gene types; diversity of virus types; diversity of sequences within a type; diversity of inferred biochemical, genetic, and evolutionary mechanisms; and diversity of interactions with the host. While the work to date has been tremendously illuminating, it has also provided a tremendous challenge: how can we better understand the nature and extent of that diversity, the mechanisms by which it arises, and how the diversity influences the ways in which the viral population interacts with the rest of the biosphere? How can we bring conceptual order to this viral universe without losing the richness of its diversity? In this review, we attempt to give a progress report on where we have gotten on that path, with a focus on evolutionary genomics across all bacterial and archaeal viruses.

COMPARATIVE GENOMICS OF BACTERIAL VIRUSES

Based on the type of nucleic acid and virion morphology, bacterial viruses (also known as bacteriophages or phages) are classified by the ICTV (International Committee on Taxonomy of Viruses) into one order and 10 families (78). Genomes of bacterial viruses can be either RNA or DNA and may vary in size tremendously; e.g., the single-stranded RNA (ssRNA) chromosome of levivirus GA is less than 3.5 kb (116), while the double-stranded DNA (dsDNA) genome of myovirus G is 497.5 kb (110). Table 1 summarizes characteristics of bacterial virus families.

Tailed dsDNA Viruses

The tailed dsDNA phages, classified in the order *Caudovirales*, are the viruses that most people associate with the term bacteriophage. They constitute >90% of the phages reported in the literature (3) and may also make up a majority of all bacteriophages in nature, although the jury is still out on the latter point. Their lytic growth cycle is similar in general terms to what is seen for most viruses. The most prominent difference is perhaps the fact that they inject their DNA into the cell upon infection, leaving most of the virion proteins on the

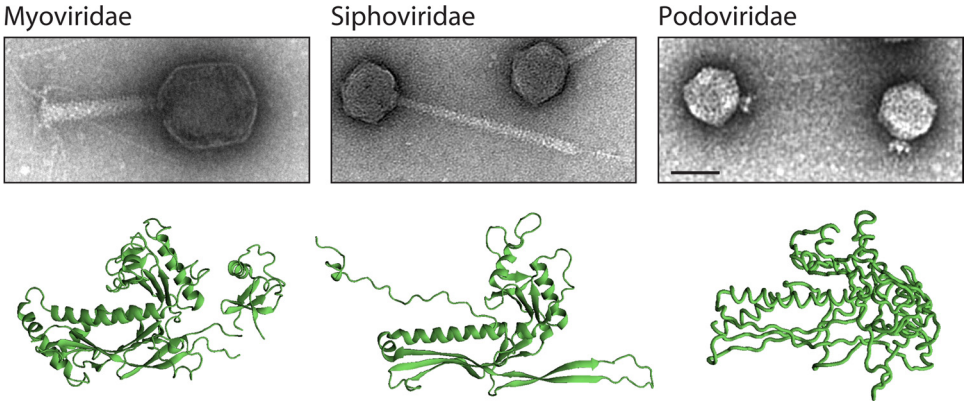

FIG. 1. Viruses of the order *Caudovirales*. Transmission electron micrographs of a T4-like virus, HK97, and P22 are shown to represent the families *Myoviridae*, *Siphoviridae*, and *Podoviridae*, respectively. The three-dimensional structures of the major capsid proteins of the corresponding viruses are shown underneath the electron micrographs. Protein Data Bank (PDB) accession numbers are 1YUE for T4 gp24 (82), 1OHG for HK97 gp5 (288), and 2XYY for P22 gp5 (43). (The electron micrograph of a T4-like virus is courtesy of Damien Maura and Laurent Debarbieux, Institut Pasteur. The micrographs of HK97 and P22 are reproduced from reference 158.) The bar is the same for all three micrographs, 50 nm.

outside, a circumstance that made the Hershey-Chase experiment (113) practical. Many of these viruses are "temperate," meaning that they have available an alternative life-style in which they coexist with the host as a "prophage," with the expression of most but not all phage genes shut off. In most cases, the prophage is integrated into the host chromosome (e.g., *Escherichia coli* phage λ), but some prophages are extrachromosomal plasmids, either circular (e.g., *E. coli* phage P1) or linear (e.g., *E. coli* phage N15). The lysogenic life cycle, as this association is called, is thought to exert an important influence on both the evolution of the virus and the evolution of the host. In the case of virus evolution, this is because any phage infecting a cell is likely to encounter prophage DNA with which it might potentially recombine. In the case of host evolution, one of the most efficient ways for novel genes to enter a bacterial genome is as part of a prophage, and there are many examples of prophage-harbored genes that alter the phenotype of the host (34, 35).

The virion capsid, known as the "head" for these viruses, is made of protein with no lipid. There is typically one major structural subunit (the major capsid protein) making up the protein lattice of the capsid, arranged in hexamers and pentamers, closely following the icosahedral symmetry and "quasiequivalent" packing predicted by Caspar and Klug (40). In some cases, the capsid is prolate, that is, elongated along one axis of the icosahedron (e.g., myovirus T4). There are also usually other proteins associated with the capsid in various roles, and all of these phages have a "portal" (called a "connector" in some phages), a dodecameric ring located at one vertex of the capsid. The portal, among other duties, marks the vertex of the capsid to which the tail will attach.

The tail is the structure that attaches the virion to the cell and engineers the entry of the DNA into the cell at the beginning of infection (208). For purposes of taxonomy, the order *Caudovirales* is divided into three families based on tail morphology (111): the *Siphoviridae*, with long noncontractile tails; the *Myoviridae*, with long contractile tails; and the *Podoviridae*, with stubby tails (Fig. 1). These divisions provide a convenient way for us to talk about the differences in the phages' tail morphologies, but as we explain below, they do not provide a guide to the biological properties or phylogenetic histories of other aspects of the virus.

The mechanisms by which the virions of these viruses are assembled from their protein and DNA parts have been under diligent study for nearly 5 decades. They are replete with examples of folding and assembly chaperones, the conformational regulation of assembly, size determination, and other examples of how biological structure is generated, and they have been reviewed extensively (75, 107, 122, 260). For current purposes, we will delineate only one feature of this process, DNA packaging, because it illuminates one of the evolutionary questions discussed below; that is, in these viruses, the procapsid shell is first assembled and then filled with DNA by a pump acting through the portal vertex. The DNA is packed to a very high density, comparable to the density of crystalline DNA. The DNA is packaged in this way at every turn of the lytic cycle, and each time, the capsid acts as a sort of Procrustean bed to limit the size of the viral genome.

As mentioned above, tailed phages represent the dominant morphotype among characterized bacterial viruses: of more than 5,500 prokaryotic viruses examined under an electron microscope by the end of 2006, 96% were tailed (3). This dominance is also reflected in the number of complete genome sequences for these viruses: tailed virus genomes constituted more than 80% (506 out of 607) of all complete genome sequences of prokaryotic viruses publicly available in the GenBank database (as of 6 August 2011). The number gets even higher when genomic sequences of proviruses are included. Most of the sequenced bacterial genomes, irrespective of their ecological niche, contain at least one prophage or its remnants (34–36, 197); in some cases, prophages might constitute as much as 10 to 20% of the bacterial chromosome (36). This abundance of genomic data allowed the performance of a series of informative in-depth comparative analyses that contributed tremendously to our current understanding of the genetic diversity and evolution of these viruses (39, 103, 112, 155, 156, 199).

Diversity in genome size and content. There is compelling evidence that all tailed bacterial viruses share a common ancestry (9, 82), at least insofar as the capsid is concerned. However, the capsid size and, consequently, viral genome length within this supergroup vary extensively. The extreme case is the nearly 50-fold difference between the 11.6-kb genome of podovirus P1 infecting mycoplasmas (272) and the genome of the giant myovirus G (497.5 kb [110]). Protein-coding genes are generally tightly packed in viral genomes and typically occupy >90% of the genome in tailed viruses (108). The functional versatility of tailed viruses with larger genomes is much greater than that of their smaller siblings. Such additional functions might include, for example, DNA replication and modification, nucleotide metabolism, carbon and nitrogen metabolism, photosynthesis, or, as in the case of the above-mentioned myovirus G, aminoacyl-tRNA synthesis (110, 264, 293).

A priori, there are two possibilities that could explain the genome size differences between evolutionarily related viruses: viruses with larger genomes arise from the smaller ones by the acquisition of new genes, or, vice versa, viruses with smaller genomes are descendants of an ancestor with a larger genome that underwent reduction. Although some investigators are in favor of the genome reduction hypothesis (225), recent sequence analyses indicate that tailed bacterial viruses with large genomes evolved from a much smaller ancestor (80). Three major mechanisms are accountable for such genome expansion: (i) multiple gene or genomic segment duplications, (ii) the lateral acquisition of cellular and viral genes, and (iii) the dissemination of diverse mobile genetic elements (MGEs), such as insertion sequences (80). The latter hypothesis is also more plausible when the two possibilities (genome expansion versus reduction) are compared in the framework of viral capsid architecture and genome packaging mechanisms (110). The difficulty with reducing the genome size in this context arises from the fact that capsids are limited to certain discrete sizes described by the "allowed" triangulation numbers of Caspar and Klug (40). For a genome to shrink to fit in the next smaller capsid size, it would require multiple independent deletions of the nonessential genes interspersed among the essential genes before the essential genes were sufficiently clustered to fit into a smaller capsid. Going the other direction is, conceptually at least, more plausible.

Sequence divergence. Tailed viruses are an extremely divergent supergroup; i.e., homologous nucleotide and protein sequences of these viruses often bear no recognizable similarity. For example, the major capsid protein (MCP) sequences of more distantly related tailed viruses often display only 10 to 20% identity in pairwise comparisons. Nevertheless, it became apparent that all these MCPs (for which structural information is available) have the same basic topology, exemplified by gp5 of HK97 (Fig. 1) (4, 9, 43, 72, 82, 288). This divergence of homologous sequences is due mainly to genetic drift. However, not all viral genes are equally prone to an accumulation of point mutations. A notable example is the large subunit of the terminase, an enzyme that cleaves a concatemeric viral DNA into genome-length units and powers their encapsidation into preformed procapsids, and is perhaps the most conserved genuine viral protein specific to tailed bacterial and archaeal viruses as well as herpesviruses (36, 149, 224). Although the overall level of sequence similarity between terminases encoded by distantly related tailed viruses can be extremely low, the enzymatic nature of these proteins demands the presence of several well-conserved motifs that do not tolerate substitutions (224).

Modular genome organization and mosaicism. Typically, genes of tailed viruses are organized into functional modules (Fig. 2); i.e., genes encoding proteins that function in the same process are clustered together and are regulated by common promoters (32). For example, genes of lambdoid viruses are usually organized into modules responsible for DNA packaging, virion morphogenesis, the establishment of lysogeny and host cell lysis, genome replication and recombination, as well as transcription regulation (Fig. 2) (32). Therefore, the viral chromosome can be seen as an assortment of such modules. The gene order in these functional modules is often well conserved, although their positions with respect to each other may vary in different virus groups (108). The module for virion structure and assembly in siphoviruses is a good example of conserved gene order (32, 36, 103). Although the sequence similarity between genes present in this module in distantly related viruses is often beyond recognition, the order of the genes is generally the same.

Comparative analyses revealed that genomes of tailed phages are highly mosaic with respect to each other (123, 199, 216). The continuity of the high-degree pairwise similarity between two viral genomes is often disrupted by regions that display only distant (or usually no) similarity to each other. A unit in such genetic mosaics is typically a single gene or, less frequently, parts of genes encoding protein domains. However, not all functional modules are equally susceptible to gene substitution/insertion/deletion. In general, genes encoding interacting proteins coevolve with each other. One such example is that of modules encoding virion structural proteins, especially those responsible for capsid formation (Fig. 2) (106). In some cases, viral genomes are shaped by matching entire functional modules, as in the case of siphovirus HK97, which shares the capsid assembly module with *Shigella flexneri* myovirus V (SfV), whereas the tail assembly module is clearly related to that of siphovirus λ (Fig. 2) (158).

As more relatively large (>100-kb) genomes have been determined, a more marked version of the division of genes into two categories with different evolutionary behaviors is seen. This has been most completely described for the T4-like phages and their cyanophage relatives (67, 187, 264) (see the section on noncore genes in marine cyanophage genomes below), and it has also been seen for a group of phages infecting Gram-positive hosts, including *Bacillus* phage SPO1 (261). The first group of genes, known as the core genes, is found in all members of a given group of phages. They include all of the essential genes, such as head and tail structural genes and DNA metabolism genes. The remainder of the genes, the noncore genes, is generally not found in all genomes in a group, and they are more likely to lack identified functions. Where functions are known for these genes, they typically provide accessory functions that appear to better adapt the phage to its ecological niche (56, 68, 264). In contrast to the dramatic horizontal movement of the noncore genes, the core genes of the T4 group were found to have congruent phylogenies (79)—that is, they had not swapped horizontally—and a similar result was seen for the SPO1 group (261). This is reminiscent of what

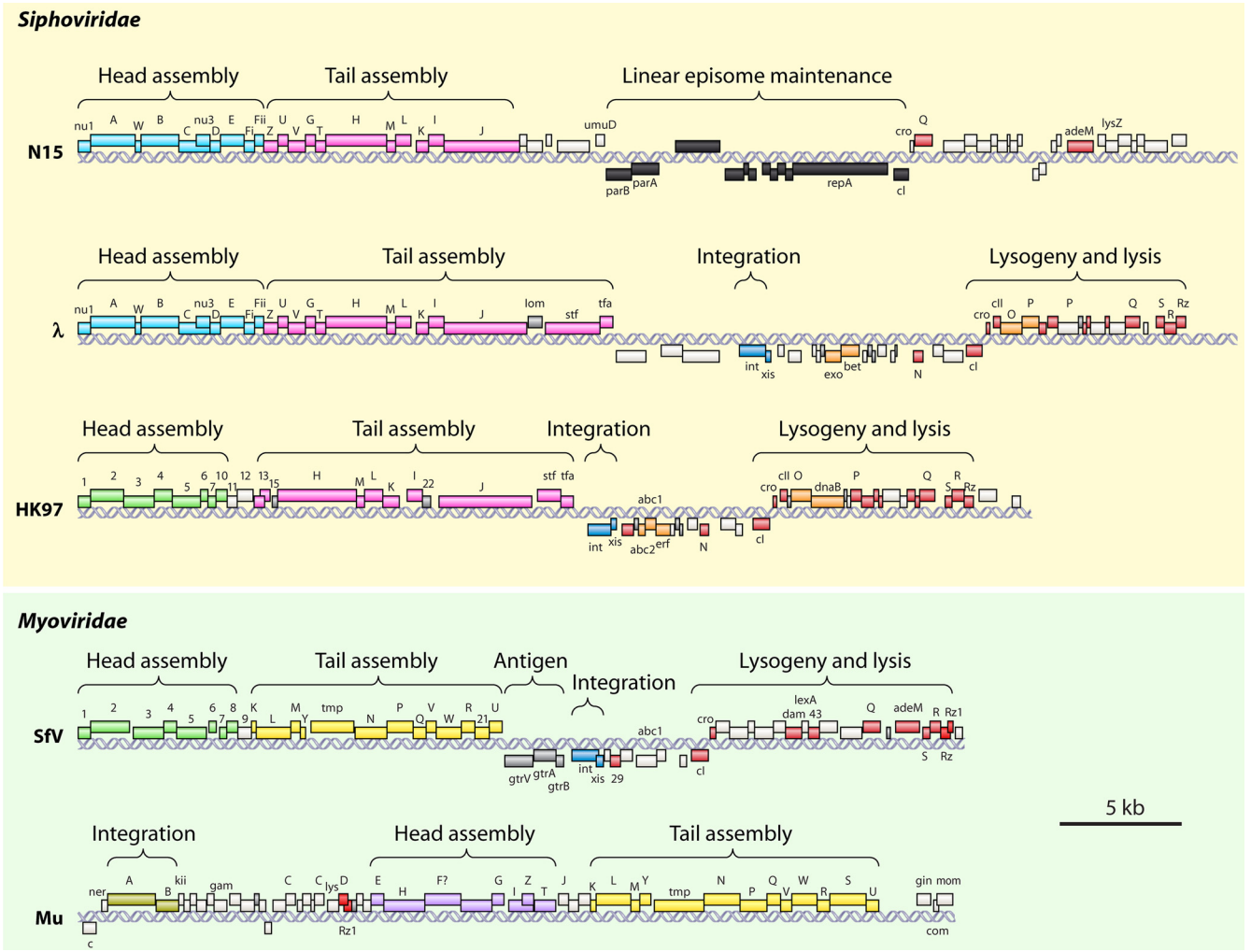

FIG. 2. Mosaic organization of the tailed dsDNA virus genomes. Homologous genes in cassettes are colored similarly; analogous but not homologous genes are shown in orange. To a first approximation, the head genes stay together as a single mosaic module, as do the tail genes. It should be noted, however, that the two modules are not strictly linked and can be exchanged between distantly related viruses. The early genes (right end of the map for all but phage Mu) are mosaic on a finer scale, with individual genes typically representing mosaic modules. (Reproduced from reference 158.)

is seen for the head genes or for the tail genes of smaller phages. In fact, a similar description of the smaller genomes as being composed of core genes and noncore genes can be made, with the principal difference being that the fraction of the genome occupied by noncore genes is smaller in the smaller genomes.

A case can be made that the noncore genes are where most of the evolutionary action occurs in this group of viruses. They move in and out of the genomes with unusual rapidity, judging by the great differences in the inventories of noncore genes between genomes that are otherwise extremely similar in sequence, and they presumably present the phage population with corresponding opportunities to occupy new ecological niches. More speculatively, the noncore genes may provide an opportunity to "invent" new genetic functions. They are, in fact, substantially smaller on average than the core genes (261), and it has been suggested that some of the noncore genes may encode single protein domains that could be joined by nonhomologous recombination into novel combinations with other such domains to generate novel functions (103). The latter speculation is made more plausible by the somewhat paradoxical idea that although the core genes of a phage are under strong purifying selection (287), the requirement of the DNA packaging mechanism that the head be filled with DNA suggests that some genes can be maintained in the genome even if they provide no selective benefit to the phage.

Noncore genes in cyanophage genomes. Due to their importance in the global ecosystem, marine tailed dsDNA viruses, especially those infecting cyanobacteria, have been studied extensively during the past decade, leading to the accumulation of a substantial amount of genomic information on these viruses. One of the most exciting findings that emerged from their analysis was the identification of a variety of host-derived functions encoded by the so-called noncore viral genes (see above). These genes are typically clustered within hyperplastic genomic regions, interspersed between the core genes, and are

thought to modulate virus-host interactions, assisting in the adaptation of viruses to new hosts and/or environments (53, 181). Virus-encoded components of the photosynthetic machinery represent one of the best examples of such fitness-conferring functions. Cyanomyoviruses and cyanopodoviruses often carry genes, *psbA* and *psbD*, encoding the key components of photosystem II (PSII), the proteins D1 and D2, respectively (165, 172, 265). More recently, an analysis of metagenomic data revealed that the entire suite of proteins composing PSI (7 proteins) might also be encoded by marine cyanophages (248). However, not all cyanophages possess photosynthetic genes, and those that do possess these genes appear to have acquired them on multiple independent occasions from their hosts (165, 180, 265). It was suggested that the production of viral PSII components might ensure stable energy production during the infection cycle, thereby securing successful virus replication (172). Indeed, the expression of the viral *psbA* genes was demonstrated experimentally (54, 164). Notably, whereas the level of expression of the host photosynthetic genes declined over the course of infection, the level of transcription of the viral counterparts steadily increased (164). *psbA* and *psbD* are not the only photosynthesis-related genes in cyanophage genomes: proteins which may potentially influence everything from adaptation to light stress, light harvesting, photosystem stability, and photosynthetic electron transport are sporadically encoded by different cyanophages (52, 53). Other noncore genes of cyanophages include those associated with carbon, phosphate, and nitrogen metabolism; the phosphate stress response; nucleotide metabolism; vitamin B_{12} and lipopolysaccharide biosynthesis (53, 264, 265, 287); and, as recently revealed through metagenomic data mining, antioxidation, the assembly of Fe-S clusters, and translation (249). Although the exact role of the proteins encoded by these noncore genes is currently not known, their widespread distribution in cyanophage genomes does suggest that they are likely to increase virus fitness and drive adaptation to particular hosts and environments (53). Furthermore, the abundance of "cellular" genes in viral genomes might have a significant impact on the evolutionary dynamics of these genes not only within the virus population but also in their hosts. Viral genomes generally evolve at a different pace than do those of cellular organisms. Consequently, the increased evolutionary rate of "cellular" genes during their stay on the viral chromosome followed by the reintroduction into the cellular genomic context might affect the fitness of the host (165). A similar host-to-virus-to-host genetic transfer loop was recently proposed to explain the rapid evolution of replicative minichromosome maintenance (MCM) helicases in the archaeal order *Methanococcales* (150). Photosynthesis-related genes do indeed seem to generate functionally new variants in viral genomes. For example, cyanomyovirus P-SSM2-encoded phycoerythrobilin synthase (PebS) single-handedly catalyzes a reaction for which uninfected host cells require two consecutive enzymes, PebA and PebB (58), while genes for the PSI components PsaJ and PsaF were found to be fused in the viral genomic context (248). Interestingly, certain clusters of noncore genes, e.g., those associated with carbon metabolism, in cyanophage genomes were suggested to be mobile, as they are (i) found in varied locations in different viral genomes, (ii) frequently located in the proximity of tRNA genes, and (iii) often encode endonucleases that might be directly involved in the mobility of associated regions (264). Further studies are required to unravel the exact molecular mechanisms of gene acquisition and genome evolution in (cyano)phages.

Mechanisms of evolution. The evolution of tailed virus genomes is driven mainly by illegitimate, rather than homologous, recombination occurring at essentially random positions within the genome (123, 199). Most new variants generated in such process are expected to be nonviable and are eliminated by natural selection. However, in rare circumstances, recombinant genomes may produce viable progeny and lead to productive infection. The success of such recombinants depends on (i) the preservation of essential viral functions and (ii) the maintenance of proper genome length. Therefore, the survivors usually bear neutral recombination joints. The genetic exchange might occur (i) between two viruses coinfecting a single bacterial cell or (ii) between a virus and a resident provirus. The latter scenario was reasoned to be more likely (101), because single infections are more frequent than double infections and because there is no immediate pressure on the genome length of the integrated provirus, nor is there an urgency to form functional progeny.

Although the most inventive, illegitimate recombination, as mentioned above, is expected to be the least successful; i.e., viable viruses are produced at a very low frequency. However, in the light of current information on the size and dynamics of the global population of tailed viruses (108, 267), even such rare events are, in fact, frequent enough to account for the observed diversity of tailed viruses. It has been estimated that 10^9 to 10^{10} fully functional novel genome variants are produced via nonhomologous recombination every second on a global scale (110). Once the newly created mosaic joints prove to be advantageous (or at least not harmful), they can rapidly spread within the population of closely related viruses by homologous recombination (106).

A recent analysis of two closely related T4-like myoviruses, phi1 and RB49, revealed that the diversity observed between the two viral genomes is generated largely by homologous recombination occurring between small regulatory cassettes that are abundantly present in the two genomes (52 in phi1 and 51 in RB49) (7). Interestingly, recombination between the nearby cassettes was found to create small circular DNA molecules that are incorporated into the virion and were hypothesized to function as vehicles of horizontal gene transfer. Similar cassette sequences were also found for other T4-like viruses, suggesting that homologous recombination may substantially contribute to the evolution of (these) viral genomes (7). In summary, the genomes of tailed bacterial viruses evolve by a combination of genetic drift, extensive illegitimate and homologous recombination, genomic segment duplications and inversions, as well as the acquisition of diverse mobile elements, such as inteins, transposons, and mobile homing endonuclease genes (110, 207).

Tailed phages and the classification conundrum. The mosaic nature of viral genomes, documented most extensively for tailed phages, has caused some consternation among scientists concerned with how to classify viruses. The current method of virus taxonomy, embodied in publications of the International Committee on Taxonomy of Viruses (78) and including the division of the tailed phages into three families based on tail

morphology, is a strictly hierarchical Linnaean system. The application of such a system to a mosaic population like the tailed phages inevitably leads to logical problems, because any single hierarchy necessarily misrepresents most genes of a virus, each of which has its own phylogenetic history (158). There have been attempts to devise classification systems that better represent the phylogenetic complexities of real viruses (e.g., see reference 162), which, however, suffer from the fact that the closer they come to representing these complexities, the more complex they become. The viruses, of course, are oblivious to our discussions of taxonomy and just keep on generating diversity. The way in which we describe viruses constrains how we think about them, and the real test of the utility of a classification system is how well it reveals and illuminates the biological relationships among the viruses. At present, we are caught between the Scylla of a system that is simple but misleading and the Charybdis of systems that are more faithful to the biology but dauntingly complex.

Tectiviridae

The characteristic tectiviral virion consists of an icosahedral protein capsid, which covers a protein-rich membrane vesicle. The latter encloses a linear dsDNA genome of ~15 kb (94). The genome possesses inverted terminal repeats and covalently attached terminal proteins that prime the genome replication carried out by the virus-encoded type B DNA polymerase. Tectiviruses can be divided into two groups: those infecting Gram-negative bacteria (exemplified by PRD1) and those of Gram-positive hosts (exemplified by Bam35).

Genome sequences are available for several PRD1-like as well as Bam35-like viruses (244, 255). Although representatives of the two groups are extremely similar to each other structurally (1, 157), their genomes share virtually no sequence conservation; only four proteins of Bam35 could be linked to those of PRD1 by use of bioinformatic methods (227). Within the groups, however, the genomes are highly similar. For example, a comparative analysis of the six available PRD1-like genomes revealed that they share 91.9% to 99.8% nucleotide sequence identity, despite the fact that the six viruses have been isolated independently at geographically remote locations worldwide (244). This is in sharp contrast to the genomics of tailed dsDNA viruses, where genetic diversity is immense (as discussed above). Bam35-like tectiviruses infecting different *Bacillus* species are also genetically closely related to each other although not to the extent of PRD1-like viruses (255). Bam35-like viruses infecting the same host were found to be genetically more closely related to each other than to related viruses infecting different hosts (255). This finding suggests that Bam35-like tectiviruses coevolve with their hosts. A notable difference between the PRD1-like and Bam35-like tectiviruses is that the latter possess the ability to lysogenize their hosts and replicate their genomes intracellularly as linear plasmids (87, 263). Indeed, *Bacillus cereus* cells contain a linear plasmid, pBClin15 (15.1 kb), which is closely related to the *Bacillus thuringiensis*-infecting tectivirus Bam35 (18 to 88% identity between the corresponding proteins, including those required for virion formation) (263). Unlike all characterized tectiviruses, the pBClin15 genome does not possess inverted terminal repeats and was suggested to represent a degenerating provirus that cannot give rise to virus particles. If true, this case would represent a relatively recent transition event during which one type of mobile genetic element (a virus) transformed into another (a plasmid).

Despite the lack of extensive sequence conservation, the overall genome organization of PRD1-like tectiviruses is very similar to that of their Bam35-like relatives (227). In both groups, the genomes follow a modular organization, also characteristic of tailed dsDNA viruses (32), with functionally related genes grouped into distinct clusters that are regulated by common promoters (244, 255). However, unlike tailed bacteriophages, the mosaic mode of genome evolution is not characteristic of tectiviruses, with the exception of the nonorthologous replacement of lysin-encoding genes in Bam35-like viruses (255) and, possibly, the acquisition of the entire lysis gene cluster (including those for endolysin, holin, and phage λ Rz/Rz1 homologues) by the ancestor of PRD1-like viruses from a tailed bacteriophage (146). It therefore appears that tectiviral genomes evolve mainly through the accumulation of point mutations and the seldom acquisition of new genes. The striking near identity between PRD1-like genomes was suggested to be a result of their "optimal genome-level organization and structure, in which any change decreases fitness" (244). The overall genome organization and virion architecture suggest that PRD1-like and Bam35-like tectiviruses had a common ancestor, which existed prior to the diversification of the Gram-negative and Gram-positive bacteria more than a billion years ago (157, 227).

Corticoviridae

Bacterial virus PM2 is the type member of the family *Corticoviridae* (11). Its overall virion morphology resembles that of tectiviruses (Table 1), except that the genome of PM2 is a highly supercoiled circular dsDNA molecule of ~10 kb (97, 174). The PM2 genome is replicated in proximity of the cytoplasmic membrane via a rolling-circle replication (RCR) mechanism initiated by the phage-encoded replication initiation protein (28, 174). The PM2 genome is organized into three operons (two early and one late) regulated by phage-encoded transcription factors (173). Interestingly, the organization of the two early PM2 operons was found to be similar to those required for the maintenance of *Pseudoalteromonas* plasmid pAS28 (124, 173). One of these two operons shares significant sequence similarity with the corresponding region in pAS28 and was therefore suggested to represent a recent horizontal acquisition in the PM2 genome from a pAS28-like plasmid (153, 173). Notably, the replication proteins in PM2 and pAS28 are nonhomologous, although their genes occupy equivalent positions within the two replicons.

All PM2 genes can be categorized into four functionally distinct modules: genes encoding (i) proteins responsible for structural components of the virion (proteins P1 to P10), (ii) proteins responsible for transcription regulation (P13 to P16), (iii) proteins responsible for genome replication (RCR initiation protein P12), and (iv) proteins involved in cell lysis (P17 and P18). A recent structural analysis revealed that the MCP and the penton protein of PM2 have the same general topology as the corresponding proteins of tectivirus PRD1 (2). In addi-

tion, the two viruses encode homologous genome-packaging ATPases (262).

PM2 infects marine *Pseudoalteromonas* species (57, 135) and is currently the sole assigned member of the *Corticoviridae* (11). The virus is strictly virulent under laboratory conditions; at the end of the infection cycle, the host cells are disrupted via a mechanism which differs considerably from those described for other dsDNA bacteriophages, including tectivirus PRD1 (146, 147). However, a bioinformatic survey revealed that PM2-like proviruses are relatively abundant in the genomes of aquatic bacteria: 4% of available aquatic bacterial genomes ($n = 269$) were positive for PM2-like proviruses (143). A comparative analysis of these proviruses revealed that genes encoding the MCP (gene *II*), structural protein P7 (gene *VII*), and the putative genome-packaging ATPase (gene *IX*) were invariably present in all proviral genomes. In contrast, genes of the other three functional groups (replication, transcription regulation, and host lysis) were found to be frequently replaced with functionally equivalent but nonhomologous genes from structurally unrelated viruses and plasmids (143). For example, 5 of the 13 identified PM2-like proviruses were found to encode typical endolysin genes, homologous to those encoded by numerous tailed dsDNA bacterial viruses. The PM2 genome, on the other hand, does not carry a gene for an apparent endolysin homologue (147). Similarly, some corticoviral elements code for transcriptional regulators homologous to those of PM2 (173), while others encode clearly distinct transcriptional factors. For example, a corticoelement of *Vibrio splendidus*, instead of PM2-like transcriptional regulators, encodes a homologue of a typical CI repressor (143). Most surprising is the observed diversity of replication proteins encoded by different PM2-like proviruses; corticoviral elements can rely on at least four distinct strategies for genome replication (as judged by the presence of four nonhomologous genome replication protein-coding genes in distinct corticoviruses). The latter observation suggests that the history of replication genes can be (and often is) independent from that of other viral functions (140). It therefore appears that any gene in corticoviral genomes, except for those encoding virion proteins, can be exchanged for functionally analogous but nonhomologous equivalents. Such an active role of horizontal gene transfer in the shaping of viral genomes is not typical for small viruses (<15 kb) (see below). Corticoviruses, with their relatively small genomes (~10 kb), might therefore represent an interesting model system where the genetic interplay between different mobile elements (viruses and plasmids) can be investigated. The recently developed genetic system for *Pseudoalteromonas*-PM2 should facilitate such studies (134). Another important line of research should focus on revealing the role that corticoviruses play in aquatic ecosystems. The relative abundance of corticoviral genomes integrated into the chromosomes of aquatic bacteria suggests that their actual role in the environment is yet to be uncovered.

Microviridae

The family *Microviridae* (*Micro*, Greek for "small") unites small icosahedral viruses with circular ssDNA genomes replicated via an RCR mechanism (Table 1). The type member of the family, microvirus ϕX174, has been extensively studied structurally, biochemically, and genetically (74, 75; for a recent review, see reference 46). Microviruses have been isolated from diverse hosts, such as enterobacteria (237), *Bdellovibrio* (27), *Spiroplasma* (48, 233), and *Chlamydia* (89, 242) (the latter three are intracellular parasitic bacteria).

Based on their genome and virion organizations, microviruses are divided into two distinct groups: those infecting enterobacteria and those replicating in obligate parasitic bacteria (Fig. 3) (89). The latter group was recently assigned by the ICTV into a subfamily, *Gokushovirinae* (*Gokusho*, Japanese for "very small"), within the family *Microviridae* (http://www.ictvonline.org/). As in the case of tectiviruses, the two groups share little genome and protein sequence conservation, but the overall genome organizations are similar (Fig. 3A). The ϕX174-like viruses possess slightly larger genomes than the viruses from the other group (5.3 to 6.2 kb versus 4.5 kb). The difference in genome size is reflected in the absence of genes for the major spike protein (G) and the external scaffolding protein (D) in gokushoviruses (Fig. 3A) (89). A large number of genomes (>40) from closely related ϕX174-like microvirus isolates infecting *Escherichia coli* has been sequenced (237). A phylogenetic analysis of these sequences suggested the relatively recent emergence of gene *D* in the ϕX174-like clade, with the subsequent spread of the gene in the population of related viruses. This analysis also revealed at least two other horizontal transfer events between the clades that probably occurred by homologous recombination (237). However, unlike in tailed dsDNA viruses, the illegitimate recombination was not found to have any substantial contribution to the genome evolution of microviruses.

Interestingly, a comparison of gokushovirus genome sequences revealed that *Bdellovibrio*-infecting virus ϕMH2K is as similar to some *Chlamydia* viruses (Chp1) as the *Chlamydia*-infecting gokushoviruses are to each other (27). In addition, it was noted that microviruses carry preserved open reading frames (ORFs) nested within overlapping genes (Fig. 3A). Point mutations were suggested to accrete in such ORFs until a gene encoding a beneficial function is produced (the mechanism known as overprinting), as might have been the case for lysis gene *E* and gene *K* of ϕX174 (27, 198). Consequently, it was suggested that occasional species jumping and genetic drift are the two major mechanisms behind the evolution of microviruses. Such inherently different modes of genome evolution for microviruses compared to those of tailed dsDNA bacteriophages were reasoned to be a result of (i) the strictly lytic life-style, (ii) a small genome size restrained by the capsid dimensions, and (iii) a low abundance of the double-stranded replicative form of the viral genome inside the host cell, which should reduce the frequency of recombination (27, 74, 237).

Metagenomic studies have recently uncovered the prevalence of ssDNA viruses in soil and aquatic environments (6, 66, 133, 168), and in a few cases, the complete circular genomes could be assembled (6, 273). Notably, the abundance of uncultured microviruses in certain environments equals or even exceeds that of tailed dsDNA viruses (6, 168). Phylogenetic analysis indicates that at the sequence level, these viruses are most closely related to *Chlamydia*- or *Bdellovibrio*-infecting gokushoviruses (Fig. 3B). Due to their abundance in marine environments, ssDNA viruses might play a significant ecological role. The identification and isolation of their hosts are

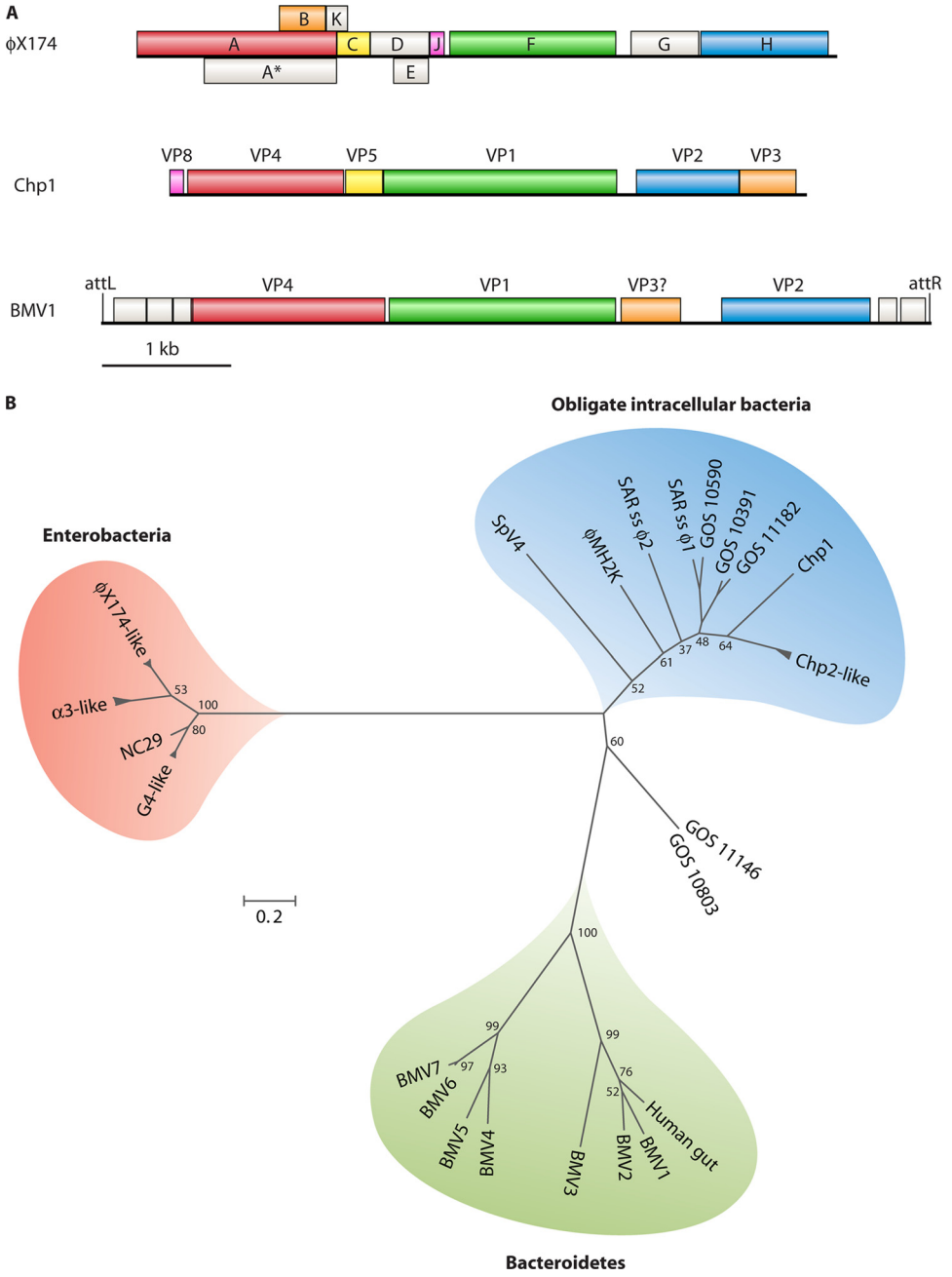

FIG. 3. Genomic organization and phylogenetic relationship between members of the family *Microviridae*. (A) Genome maps of representative members of three distinct groups of viruses within the *Microviridae*: *Escherichia coli*-infecting ϕX174 (genus *Microvirus*), *Chlamydia psittaci*-infecting Chp1 (genus *Chlamydiamicrovirus*, subfamily *Gokushovirinae*), and provirus BMV1 integrated into the genome of *Bacteroides* sp. strain 2-2-4. Homologous genes are colored similarly. attL and attR, left and right attachment sites, respectively. (B) Unrooted maximum likelihood phylogenetic tree of the ϕX174 gpF-like major capsid proteins representing the relationship between enterobacterium-infecting microviruses, obligate intracellular bacterium-infecting gokushoviruses, and putative proviruses integrated into the genomes of different species of the phylum *Bacteroidetes*. Putative marine microviruses SARssϕ1 and SARssϕ2 (273) cluster together with gokushoviruses. (For accession numbers of the proteins that were used to generate the phylogeny, please refer to reference 148.)

therefore of great importance. Interestingly, attempts to induce viruses from marine *Synechococcus* strains isolated from the Gulf of Mexico resulted in the production of icosahedral nontailed viruslike particles (VLPs) that contained ssDNA (176). However, a detailed characterization of the VLPs was not performed, leaving the question regarding their identity open.

Although temperate members of the *Microviridae* have not yet been characterized, the genomes of *Chlamydophila caviae* and *Chlamydia pneumoniae* contain gene fragments

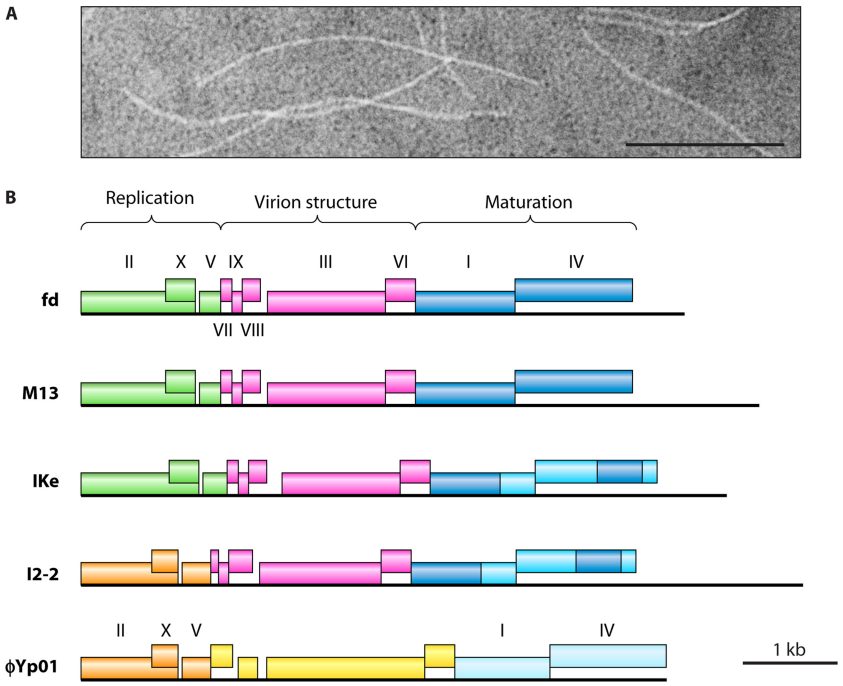

FIG. 4. Mosaic organization of ssDNA filamentous virus genomes. (A) Negative-contrast transmission electron micrograph of a *Pseudomonas aeruginosa*-infecting filamentous virus, phi05_2302 (88). (The micrograph is courtesy of Aušra Domanska.) Bar, 200 nm. (B) Homologous genes in cassettes are colored similarly; orthologous gene cassettes with various degrees of similarity are represented by various shades of the same color, whereas nonorthologous replacements are denoted by differently colored cassettes. The darker regions in genes I and IV of inoviruses IKe and I2-2 indicate homologous recombination events with M13/fd-like sequences. (Panel B is reproduced from reference 158.)

displaying sequence similarity to genes of *Chlamydia*-infecting gokushoviruses (230), suggesting that certain members of the *Microviridae* might be temperate. Indeed, seven putative proviruses, clearly related to gokushoviruses in their genome organizations, were recently identified in the genomes of anaerobic Gram-negative bacteria belonging to the phylum *Bacteroidetes* (148). Homologues of both structural proteins (VP1 and VP2) constituting the mature gokushovirus virions (51) are conserved in the proviral genomes (Fig. 3A). Phylogenetic analyses further revealed that the identified proviruses represent the third group within the *Microviridae* (Fig. 3B). A tentative subfamily name, *Alpavirinae* (*Alpa*, Sanskrit for "small" or "minute"), was proposed for this group of (pro)viruses (148). Interestingly, even though a temperate life-style is expected to increase the success of illegitimate recombination (due to the reasons discussed above for tailed dsDNA viruses), no apparent horizontal gene acquisition or exchange events could be detected in the proviruses of the *Bacteroidetes* (148). The identification of microvirus-like proviruses indicates that all families of bacterial viruses with circular (single-stranded or double-stranded) DNA genomes or replicating their genomes through a circular intermediate (e.g., tailed dsDNA viruses) have temperate members.

The association of microvirus-related proviruses with the *Bacteroidetes* and the abundance of microvirus-like sequences in diverse environments indicate that the diversity and host range of microviruses are still largely unknown. The isolation of new microvirus-host systems is therefore of great importance and will help to shed light on the ecological role, origin, and evolution of this group of viruses.

Inoviridae

Similarly to microviruses, inoviruses possess small circular ssDNA genomes (4.5 to 9.0 kb) replicated via an RCR mechanism, but their virions are built with helical rather than icosahedral symmetry (Fig. 4A and Table 1). The family contains two genera, *Inovirus* and *Plectrovirus*. Different aspects of inovirus biology as well as their biotechnological appeal have been extensively investigated and recently reviewed (105, 223), while plectroviruses have received only modest attention. Inoviruses were found to infect a wide range of Gram-negative bacteria as well as some Gram-positive species (49, 132, 241). Plectroviruses, on the other hand, infect mollicutes, such as *Spiroplasma* and *Acholeplasma* (171). Viruses from the two genera show virtually no genome or protein sequence similarity and differ considerably in morphology: inovirus virions are generally long and flexible (Fig. 4A), while those of plectroviruses are short rods (247). Unlike other bacterial viruses, with the exception of plasmavirus L2 (206), members of the *Inoviridae* do not lyse host cells but leave them by extrusion coupled to virion assembly (105, 175, 223). All characterized plectroviruses and some (but not all) inoviruses are capable of lysogenizing their hosts. This ability seems to have been acquired on multiple independent occasions, since not all members of the family are able to integrate into the host chromosome, and those that are able to do so utilize different genome integration strategies. *Spiro-*

plasma plectroviruses (SpVs) SpV1-R8A2B and SpV1-C74 and *Neisseria* inovirus Nf encode distinct transposases belonging to the IS*30*, IS*3*, and IS*110*/IS*492* families, respectively (125, 179). Interestingly, the transposase genes in the two plectroviruses, but not in inovirus Nf, have replaced an RCR initiation protein-coding gene, which is otherwise present in the related SpV-like virus SVTS2 (247). As a result, genome replication in the transposase-encoding plectroviruses occurs via a transposition mechanism, with the encapsidated genome representing a circular transposition intermediate (179). Another temperate inovirus, ϕRSM1, infecting *Ralstonia solanacearum*, encodes a recombinase of the resolvase/invertase family and utilizes a host tRNA gene for integration (126). An even more peculiar strategy for genome integration is utilized by CTXϕ-like inoviruses infecting *Vibrio* sp. (177) and possibly also inoviruses of *Xanthomonas* (163) and *Yersinia* sp. (50). Cholera toxin-encoding CTXϕ is arguably "smarter than the average phage" (23), since it hijacks cellular XerC/D enzymes for site-specific genome integration into the cellular DNA at the chromosome dimer resolution (*dif*) sites and does not encode a recombinase of its own (115, 177). The molecular mechanism of such an integration has been revealed in considerable detail (178, 274). Interestingly, a single *dif* site in *Vibrio cholerae* is targeted by several distinct mobile elements. In addition to CTXϕ, two different satellites, RS1ϕ (2.7 kb [62]) and TLCϕ (4.7 kb [239]), are integrated into the same locus, forming a composite lysogenic array. RS1ϕ and TLCϕ do not encode virion morphogenesis proteins, but with the aid of their respective helper viruses (CTXϕ or KSF-1ϕ for RS1ϕ and fs2ϕ for TLCϕ), the two satellites can be encapsidated into infectious filamentous virions (76, 77, 100). CTXϕ provirus, on the other hand, depends on RS1ϕ for normal reactivation and genome replication. RS1ϕ encodes an antirepressor, RstR, which promotes the transcription of viral genes required for CTXϕ production and the consequent transmission of both RS1ϕ and CTXϕ (60). Furthermore, the replication of the CTXϕ proviral genome via a rolling-circle-like mechanism occurs between two origins of replication. One is provided by the CTXϕ provirus itself, but the second one is encountered within the downstream RS1ϕ element, thereby leading to the production of hybrid viral genomes composed of DNA sequences contributed by both integrated elements (63, 185). Indeed, *V. cholerae* strains containing a single chromosomally integrated filamentous phage genome have never been isolated (223). A peculiar chimera was also obtained during the coinfection of *V. cholerae* with two distinct inoviruses, VGJϕ and CTXϕ (33). In this case, the double-stranded replicative forms of the two phage genomes recombined in a site-specific manner to produce chimeric dsDNA molecules, which subsequently generated ssDNA viral genomes consisting of the coding strand of VGJϕ and the noncoding strand of CTXϕ. Such genomes are next encapsidated by using the structural proteins of VGJϕ. Importantly, VGJϕ and CTXϕ recognize different receptors on the *V. cholerae* surface (mannose-sensitive hemagglutinin pilus and toxin-coregulated pilus, respectively). Therefore, VGJϕ acts as a generalized transduction vehicle, which delivers the CTXϕ genome into *V. cholerae* strains that are normally not susceptible to this cholera toxin-encoding phage (33). Studies of *Vibrio*-infecting inoviruses revealed the beauty and complexity of interactions between different viruses and other mobile genetic elements within a single cell. Further research is needed to more fully reveal the breadth and molecular details of such interactions that are of outstanding scientific interest and medical importance.

The genomes of inoviruses typically display a modular organization, with genes encoding proteins responsible for genome replication, virion morphogenesis, and structure grouped into clusters (Fig. 4B) (241). However, unlike in tailed viruses and tectiviruses, gene expression is not temporal but occurs concurrently and is regulated mainly by differences in the strengths and accessibilities of the ribosome-binding sites and unequal codon usage, etc. (241). Nevertheless, it appears that the evolution of inovirus genomes is governed by mechanisms that are very similar to those operating in tailed dsDNA viruses (Fig. 4B) (158). The differences in genome lengths observed among inoviruses result mainly from the lateral acquisition of new genes from diverse sources. For example, a homologue of the cellular outer membrane channel EpsD, which is used for virion extrusion, is encoded by *E. coli*-infecting inoviruses (gene IV in M13) but not in *Vibrio*-infecting viruses, such as CTXϕ, that utilize the cellular protein for the same purpose (61). Inovirus B5, infecting a Gram-positive bacterium, *Propionibacterium freudenreichii*, does not encounter the outer membrane during extrusion and consequently also does not encode an EpsD homologue (49). Other notable examples are the cholera toxin-encoding operon *ctxAB* of CTXϕ (279) and the reverse transcriptase gene in *Pseudomonas aeruginosa* temperate inovirus Pf4 (285). Members of the *Inoviridae* from both genera (*Inovirus* and *Plectrovirus*) seem to enjoy frequent nonorthologous replacements within the module responsible for genome replication (Fig. 4B) (158, 247, 258, 280). In addition, homologous recombination and intergenome rearrangements have been reported to play a role in the evolution of inoviruses (158, 200, 241).

Cystoviridae

Members of the family *Cystoviridae* possess tripartite dsRNA genomes enclosed in an icosahedral double-layered protein capsid, which is surrounded by a membrane envelope (209). The innermost capsid, the so-called polymerase complex or procapsid, is delivered into the cell interior upon infection and is essential for genome replication and transcription (183, 209). Cystoviruses infect phytopathogenic pseudomonads, such as *Pseudomonas syringae*. Five genome sequences of cystoviruses have been determined so far (Table 1), including that of bacteriophage ϕ6, the type member of the family (183).

Cystoviral genomes are organized in a modular fashion. However, unlike in prokaryotic DNA viruses, which, as a rule, possess monopartite genomes, functionally related proteins are generally encoded on different genomic segments. The largest of the three segments, segment L, encodes proteins of the procapsid, and segment M contains genes for the receptor-binding complex, whereas the smallest of the three segments, segment S, encodes the procapsid shell protein, the major membrane protein, and proteins responsible for host cell lysis (209). Gene expression is temporal and is achieved by a combination of unequal stability and efficiency of production of the mRNAs from different genomic segments (183). Different cystoviruses have developed distinct strategies to ensure dif-

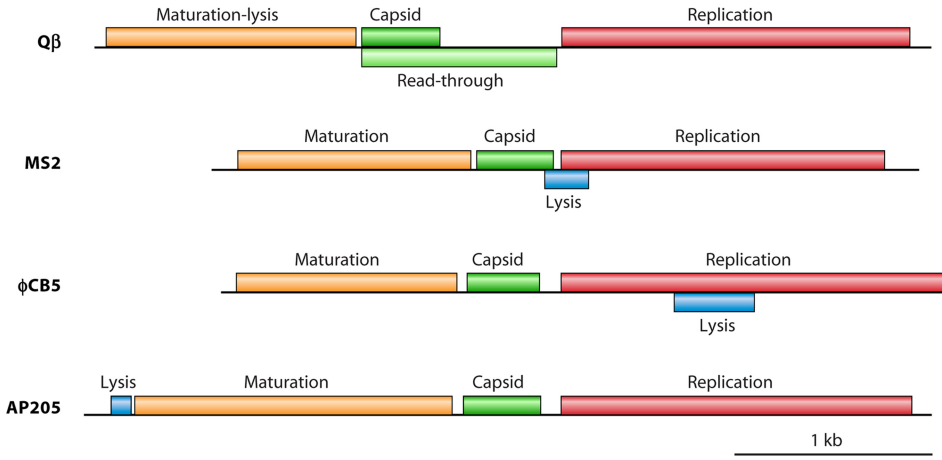

FIG. 5. Genomic organizations found in ssRNA viruses of the family *Leviviridae*. Note the various locations of the lysis protein-coding gene. During Qβ infection, cell lysis is performed by the maturation protein (19).

ferential levels of gene expression. For example, the transcription of the L segment in cystoviruses ϕ6 and ϕ2954 is activated by two different cellular proteins, YajQ and glutaredoxin 3, respectively, that bind to the procapsid, thereby modulating the activity of the polymerase P2 (217, 220). Another cystovirus, ϕ8, on the contrary, does not regulate the level of transcription but rather encodes a protein, Hb, which ensures the rapid degradation of the L segment by the host RNase R late in infection (221, 266).

The general mode of genome evolution in cystoviruses is similar to that of DNA viruses infecting bacteria, but the molecular basis is different. Illegitimate recombination is well documented for these viruses (182, 218). Although this type of recombination appears to be rare during normal ϕ6 replication, when "provoked" (e.g., under conditions that prevent minus-strand synthesis or for segments with 3′-terminal truncations), it occurs with high frequency (on the order of 50% within a distance of ~300 nucleotides) (182). Such recombination occurs by template switching during the synthesis of a nascent RNA strand within the polymerase complex and usually involves three identical bases at the crossover site but can range from 0 to 12 nucleotides (218). Homologous recombination was also reported to occur *in vitro* for cystovirus ϕ8, which is the most distant isolated relative of ϕ6 (193). However, this type of recombination appeared to be extremely rare both *in vitro* (recombination frequency, $\sim 10^{-7}$ [193]) and in wild populations of cystoviruses (251) and was reasoned to play only a minor role (if any) in the evolution of cystoviruses due to the low probability of the incorporation of two molecules of the same segment type into a single procapsid (183). Comparative analyses of cystoviral genomes have also revealed cases of intergenome rearrangements, such as the disruption of synteny in the L segment of ϕ8, where gene *7* is located at the 3′ terminus instead of the 5′ terminus, a position observed for this gene in all the other cystoviruses (266). In addition, cystoviruses are able to acquire genetic material from different sources, which seems to occur via the nonspecific incorporation of heterologous transcripts into the procapsid, followed by recombination with the viral RNA via template switching (193). Notably, the incorporation of native transcripts containing viral packaging signals was found to be only 10 times more efficient, suggesting that occasional incorporations of cellular transcripts or those of other viruses might be rather frequent. An analysis of the recently sequenced genome of cystovirus ϕ2954 (219) revealed another mechanism contributing to the evolution of cystoviruses and also common to other bacterial viruses. Gene *5*, encoding the muralytic enzyme P5 in all members of the *Cystoviridae*, is replaced in ϕ2954 by a nonorthologous gene, which is homologous to the *flgJ* gene of *P. syringae*, encoding a flagellar protein, FlgJ, with peptidoglycan hydrolase activity. The newly acquired FlgJ homologue in ϕ2954 is thus likely to be a functional equivalent of P5, found in other cystoviruses (219). Finally, genomic segment reassortment in wild populations of cystoviruses as well as in the laboratory setting occurs at extremely high frequencies, exceeding the rate of nucleotide substitutions (193, 251). Surprisingly, however, the rate of segment reassortment was found to vary significantly for viruses isolated at different geographical locations (190). This finding suggests that coinfections of the same host cell with multiple cystoviruses might occur frequently in their natural habitat and that once a new beneficial function is acquired, it will be quickly spread within the population by genomic segment reassortment.

Leviviridae

Leviviruses are among the smallest viruses known. They possess linear ssRNA genomes enclosed into icosahedral protein shells (276). Based on serological cross-reactivity, genome size, and organization, the family is divided into two genera: *Levivirus* and *Allolevivirus*. Leviviruses (type species MS2), with ~3.5-kb genomes, encode four proteins: capsid protein, maturation protein, lysis protein, and an RNA-dependent RNA polymerase (RdRp). Notably, the lysis protein is encoded by an ORF, which overlaps at its 5′ and 3′ regions with the capsid and RdRp genes, respectively (Fig. 5). The gene seemingly originated by a mechanism similar to that responsible for the emergence of the lysis gene of microvirus ϕX174 (see above); i.e., a preexisting genomic region accumulated mutations until a beneficial function was achieved (27, 137). Alloleviviruses

(type species Qβ) possess somewhat longer genomes (~4.2 kb) and, instead of the lysis protein, encode two variants of the capsid protein: a normal-sized variant and an extended version generated by an occasional (6% frequency) readthrough of the stop codon (276). Host cell lysis in allolevivirus infections is carried out by the maturation protein (Fig. 5), which inhibits cell wall biosynthesis (19).

Most of the isolated members of the *Leviviridae* infect *E. coli* and are F pilus specific. The genome sequences for 30 such viruses are currently available [15 for leviviruses and 15 for alloleviviruses; these viruses were isolated from different samples, such as wastewater, surface waters, swine lagoons, and chicken litter, etc., and could all propagate on an *E. coli* HS(pFamp)R host (64)]. Their analysis revealed no apparent recombination events or deviations from the general genome organization characteristic of leviviruses and alloleviviruses (85). Despite the low level of accuracy of the RdRp, genomes within each of the two genera of the *Leviviridae* share more than 50% nucleotide identity (85). Such sequence stability among the *E. coli*-infecting RNA viruses might be a result of their adaptation to the "needs of the day" (276); i.e., depending on the environmental conditions at a given moment, only certain viral genotypes are selected for and maintained within the population. Indeed, genome sequences of more distantly related ssRNA viruses infecting hosts other than *E. coli* display more divergence. For example, PRR1, infecting diverse Gram-negative bacteria carrying the broad-host-range incompatibility group P plasmid, was found to possess a levivirus-like genome organization, with all proteins but one being more similar at the sequence level to those of alloleviviruses (240). The complete genome sequence of virus AP205, infecting *Acinetobacter* species, revealed a genomic organization different from the two layouts consistently found for the ssRNA coliphages (137). The lysis gene of AP205 does not overlap with any other gene but is located upstream of the maturation gene (Fig. 5). The location of the lysis gene in ϕCb5, an unclassified member of the *Leviviridae* infecting *Caulobacter crescentus*, is yet different (127). The lysis gene of ϕCb5 is completely embedded within the replicase gene (Fig. 5), further suggesting that lysis genes in different members of the *Leviviridae* emerged on several independent occasions.

Most of the current understanding of the evolvability of leviviruses comes from results obtained during numerous *in vitro* studies (276), which revealed that leviviruses are capable of overcoming *in vitro*-introduced detrimental defects by a number of mechanisms. Such studies have indicated that (i) frequent base substitutions lead to adaptations to changing environmental conditions (20), (ii) deletions as well as duplications occurring by random recombination are readily introduced into the ssRNA of leviviruses in response to various deliberate changes to the genome sequence and structure (161, 192), and (iii) genomic defects can be repaired by illegitimate as well as less frequent homologous recombination occurring as a result of transesterification and template switching, respectively (47, 196).

Different Evolutionary Pathways of Bacterial Viruses

The general mechanisms that shape the genomes of bacterial viruses are often similar, but the extent of the contribution of different mechanisms seems to vary considerably when viruses from different families are compared. For example, illegitimate recombination seems to be the major force behind the diversity of tailed dsDNA bacteriophages but was not found to play a significant role in the evolution of microviruses (ssDNA genome), leviviruses (ssRNA), cystoviruses (dsRNA), or PRD1-like tectiviruses (dsDNA). The horizontal acquisition of new genes from various sources also does not seem to be equally prevalent among different virus groups. Viruses with small icosahedral capsids, such as microviruses and leviviruses, tend to accumulate mutations in the existing nucleotide sequence until a new useful function (e.g., lysis protein-coding gene) is created. Illegitimate recombination and the lateral capture of new genes seem to occur more frequently in temperate viruses than in their strictly virulent relatives. Further comparative genomic studies of different viral groups are likely to provide a deeper insight into the molecular details of viral genome evolution and illuminate new mechanisms in play.

COMPARATIVE GENOMICS OF ARCHAEAL VIRUSES

The domain *Archaea* consists of four currently recognized phyla: *Euryarchaeota*, *Crenarchaeota*, *Thaumarchaeota*, and *Korarchaeota* (30, 291). Crenarchaea and euryarchaea constitute the vast majority of characterized archaeal organisms. However, recent studies have indicated that both thaumarchaea and korarchaea are abundant and widely distributed in the environment (232, 246). For example, thaumarchaea are present in freshwater, soil, ocean, and hot spring habitats, where they might carry out the aerobic oxidation of ammonia, a key step of the nitrogen cycle (246). Unfortunately, only a few members of the *Thaumarchaeota* have been isolated and characterized, while not a single korarchaeon has so far been grown in pure culture (73). As a result, archaeal viruses have so far been isolated only on crenarchaeal and euryarchaeal hosts (205). The only piece of information on thaumarchaeal viruses comes from the analysis of a provirus related to tailed bacterial and euryarchaeal viruses that is integrated into the genome of the mesophilic thaumarchaeon "*Candidatus* Nitrososphaera viennensis" (152).

Viruses infecting organisms from the phyla *Euryarchaeota* and *Crenarchaeota* share virtually no sequence similarity and are often discussed separately (213, 214). All archaeal viruses characterized thus far contain DNA genomes, which can be either linear or circular. Table 2 summarizes characteristics of archaeal viruses. Most of the archaeal virus families contain only a few members, while some are represented by a single virus species (213). Therefore, only those virus groups that contain more than one member and for which evolutionary conclusions can be attempted from comparative genomic analyses will be discussed here.

Crenarchaeal Viruses

Comparative genomic analyses showed that crenarchaeal viruses from different families posses a small pool of common genes. However, some viruses, such as those from the *Globuloviridae* or "*Clavaviridae*," share no genes with other known viruses (184, 214). Most of the common genes encode proteins involved in genetic processes, such as DNA replication, recombination, or transcription regulation, and are likely to be ex-

TABLE 2. Overview of archaeal viruses[a]

Family	Capsid morphology	Additional feature(s)	Genome type	No. of complete genomes	Example
Crenarchaeal viruses					
Fuselloviridae	Spindle shaped	Short appendages	dsDNA, C	9	SSV1
Bicaudaviridae	Spindle shaped	Two tails	dsDNA, C	1	ATV
Guttaviridae	Droplet shaped	Multiple thin fibers	dsDNA, C		SNDV
Ampullaviridae	Bottle shaped	Short filaments at the broader end	dsDNA, L	1	ABV
Globuloviridae	Spherical	Helical nucleoprotein	dsDNA, L	2	PSV
Lipothrixviridae	Filamentous	Enveloped, flexible or rigid	dsDNA, L	8	AFV1
Rudiviridae	Stiff rods	Three terminal fibers	dsDNA, L	4	SIRV1
"*Clavaviridae*"	Rigid bacilliform	One pointed end, one rounded end	dsDNA, C	1	APBV1
NA	Icosahedral	Internal membrane	dsDNA, C	2	STIV
NA	Spindle shaped	One tail	dsDNA, C	1	STSV1
Euryarchaeal viruses					
Myoviridae	Icosahedral	Tail (contractile)	dsDNA, L	3	ϕCh1
Siphoviridae	Icosahedral	Tail (long noncontractile)	dsDNA, L	2	ψM2
Salterprovirus (genus)	Spindle shaped		dsDNA, L	2	His1
NA	Spindle shaped	Enveloped(?)	dsDNA, C	1	PAV1
NA	Oblate		dsDNA, C		A3 VLP
NA	Pleomorphic	Enveloped	ssDNA, C	1	HRPV-1
NA	Pleomorphic	Enveloped	dsDNA, C	1	HHPV-1
NA	Icosahedral	Internal membrane	dsDNA, L	1	SH1

[a] Crenarchaeal and euryarchaeal viruses not assigned (NA) to families but displaying features distinguishing them from viruses belonging to approved families are indicated. L, linear; C, circular; VLP, virus-like particle; ATV, *Acidianus* two-tailed virus; SNDV, *Sulfolobus neozealandicus* droplet-shaped virus; APBV1, *Aeropyrum pernix* bacilliform virus 1; STSV1, *Sulfolobus tengchongensis* spindle-shaped virus 1.

changed horizontally between unrelated viruses or viruses and their hosts (214).

***Fuselloviridae*.** Fuselloviruses possess lemon- or spindle-shaped virions with short filamentous appendages at one pointed end (Fig. 6) and infect hyperthermophilic hosts from the crenarchaeal genera *Sulfolobus* and *Acidianus* (213, 231). In addition to a spindle-like morphology, virions of fuselloviruses can display a pleomorphic appearance, as was found to be the case for the viruses *Sulfolobus* spindle-shaped virus 6 (SSV6) and *Acidianus* spindle-shaped virus 1 (ASV1) (morphology ranging from thin-cigar- to pear-like) (Fig. 6) (231). The circular dsDNA genomes of fuselloviruses (13.7 to 24.1 kb) are positively supercoiled and have been shown to integrate site specifically into the host chromosome (104, 186). Peculiarly, the integrase gene of fuselloviruses, which was found to be dispensable for virus propagation (55), is partitioned upon recombination with the host chromosome into two fragments that flank the integrated provirus (250). Thirteen fusellovirus genomes (9 viral and 4 proviral) are currently available (104, 231).

Comparative genomics has revealed that, despite a substan-

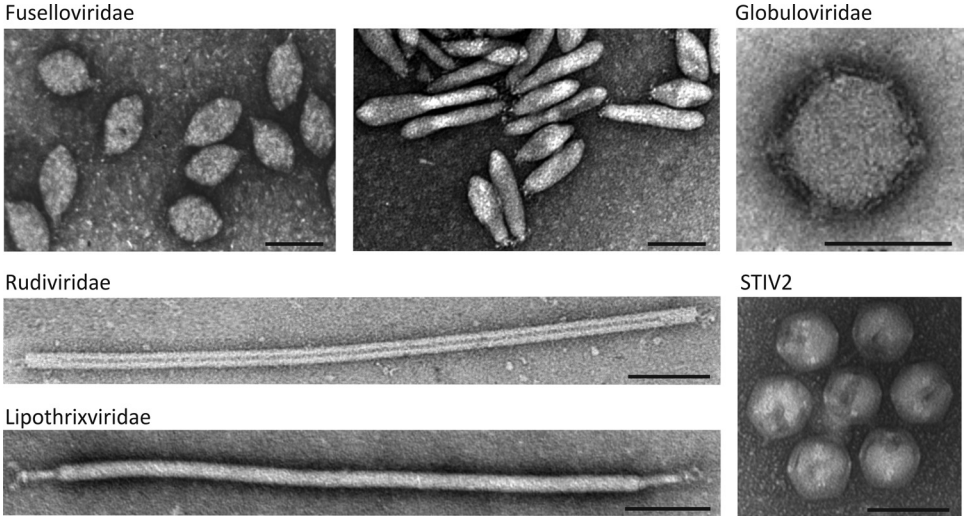

FIG. 6. Negative-contrast transmission electron micrographs of virions from different families of crenarchaeal viruses, including the *Fuselloviridae* (left, SSV1; right, SSV6), *Rudiviridae* (SIRV1), *Lipothrixviridae* (AFV1), *Globuloviridae* (PSV), and the unclassified virus STIV2. Bars, 100 nm.

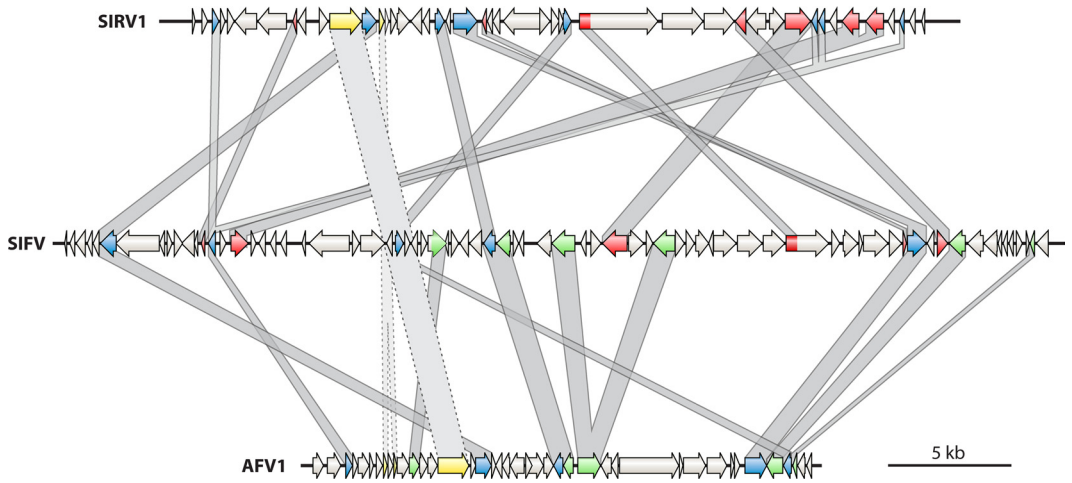

FIG. 7. Genomic relationship between linear archaeal viruses of the families *Rudiviridae* and *Lipothrixviridae*. Genes shared by *Sulfolobus islandicus* rod-shaped virus 1 (SIRV1) (*Rudiviridae*) and the lipothrixviruses *Sulfolobus islandicus* filamentous virus (SIFV) and *Acidianus* filamentous virus 1 (AFV1) are shaded blue. Genes restricted to the virus pairs SIRV1-SIFV, SIRV1-AFV1, and SIFV-AFV1 are shown in red, yellow, and green, respectively.

tial variation in genetic content, fuselloviruses share a set of 12 to 13 "core" genes (104, 231). A phylogenetic analysis of these genes pointed to the coevolution of fuselloviruses with their hosts in a biogeographic context (104). It also became apparent that homologous recombination is prevalent among fuselloviruses and seems to occur within protein-coding genes rather than intergenic regions (231). Although heterogeneity in the genome length of these viruses suggests that horizontal gene transfer plays a significant role in fusellovirus evolution, its extent has not yet been estimated.

Fuselloviruses display a peculiar relationship with pRN-like plasmids of *Sulfolobus* (166). The two types of mobile elements seem to exchange genes. For example, it was concluded that pRN-like plasmids pXZ1 and pSSVi developed the ability to integrate into the host chromosome by the lateral acquisition of the integrase genes from fuselloviruses (201, 283). Another plasmid, pSSVx, was found to be a virus-plasmid chimera, which, in addition to the typical pRN-specific gene content, contains two fuselloviral genes other than the integrase-coding gene (8). Strikingly, pSSVi and pSSVx, but not pXZ1, can be encapsidated into viruslike particles (VLPs) upon superinfection of plasmid-containing *Sulfolobus* cells with a fusellovirus, SSV1 or SSV2. This enables the dissemination of the plasmids in a viruslike fashion (8, 283). Notably, the plasmid-containing VLPs are smaller than the virions of helper fuselloviruses, as are their corresponding genetic materials (~15.5 kb for SSV1 and ~5.7 kb for pSSVi and pSSVx). Despite the difference in size, the composition and the shape of the VLPs appear to be the same as those of fusellovirus virions, suggesting that the particle size is determined by the size of the DNA (8). Since pSSVi and pSSVx do not share virus-specific genes, it was proposed that a specific nucleotide sequence feature (i.e., packaging signal) shared by pSSVi and pSSVx initiates the encapsidation of these plasmids into VLPs (283). Further molecular studies of the fusellovirus-mediated plasmid encapsidation mechanism should provide valuable information not only on the virus-plasmid interplay in archaea but also on genome packaging and the assembly of fuselloviruses.

***Lipothrixviridae* and *Rudiviridae*.** Helical archaeal viruses with linear dsDNA genomes are classified into two families, the *Rudiviridae* and *Lipothrixviridae* (213). While rudivirus virions are stiff and rodlike, those of lipothrixviruses are flexible filaments (Fig. 6). At least for rudiviruses, the length of the filamentous virion is proportional to the length of the genomic dsDNA (278), a phenomenon also characteristic of virions of bacterial inoviruses (ssDNA genome) (223) and plant-infecting filamentous viruses with ssRNA genomes (129). Unlike rudiviruses, lipothrixvirus filaments are enveloped (213). Despite these differences, rudiviruses share at least nine genes with lipothrixviruses (Fig. 7), suggesting that viruses from the two families descended from a common ancestor (202, 214). This proposal has recently been strengthened by high-resolution structural analysis, which revealed that the DNA-binding major capsid proteins of rudiviruses and lipothrixviruses share the same fold (93, 268). Interestingly, the fold of the major capsid protein is unique to these linear archaeal viruses and has not yet been observed for viruses from the other two domains of life (141). To acknowledge the evolutionary relationship between the two viral families, a new taxonomic order, the "*Ligamenvirales*" (*Ligamen*, Latin for "string" or "thread"), was recently proposed (D. Prangishvili and M. Krupovic, submitted for publication).

The genomes of both lipothrixviruses and rudiviruses are linear dsDNA molecules with inverted terminal repeats (205). However, only those of rudiviruses terminate in covalently closed hairpin structures with 5′-to-3′ linkages (24). The genome of the lipothrixvirus *Acidianus* filamentous virus 1 (AFV1), on the other hand, was found to possess terminal proteins attached to its 5′ ends (M. Pina, T. Basta, P. Forterre, and D. Prangishvili, unpublished data). Whereas all rudiviruses encode a protein that is structurally and functionally related to RCR initiation proteins (189), a candidate for the genome replication protein in lipothrixviruses is yet to be identified.

Consistently, the genome replication mechanisms in the two groups of archaeal linear viruses are likely to be different. It therefore appears that, as in the case of bacterial viruses (140, 143), the processes of genome replication and virion assembly are not coupled and evolve independently.

Based on the differences in the virion terminal structures as well as the variation in gene content and organization, the family *Lipothrixviridae* is subdivided into four genera: *Alphalipothrixvirus*, *Betalipothrixvirus*, *Gammalipothrixvirus*, and *Deltalipothrixvirus* (213). A comparative analysis of lipothrixviruses revealed that linear genomes of viruses from one genus can undergo intergenomic recombination with those from other genera (277). Lipothrixviruses and rudiviruses display a considerable variation in gene content and order (Fig. 7). The lack of apparent genome synteny between different linear archaeal viruses (even between members of the same family, e.g., *Lipothrixviridae*) (Fig. 7) is in sharp contrast to the preservation of the gene order in the morphogenetic module of tailed dsDNA viruses infecting bacteria and archaea (see above). Interestingly, transcriptional analysis of the rudiviruses *Sulfolobus islandicus* rod-shaped virus 1 (SIRV1) and SIRV2 revealed that there is little temporal regulation of viral gene expression in these viruses, with the transcription of all but one viral genes commencing soon after infection (130). Such a lack of a necessity to maintain a certain gene order and modular genome organization might be due to the much simpler virion assembly and structure of linear viruses than those of tailed icosahedral viruses. Whereas the head and tail modules of tailed bacterial viruses usually contain approximately 20 genes (Fig. 2), the rudivirus virions are composed of only four structural proteins, where the major capsid protein (accounting for 99% of virion proteins) can coassemble with dsDNA to form rudivirus-like particles (278).

It has been noticed that rudiviruses accumulate new unique genes, with no similarity matches in public databases, at the termini of their linear genomes (278). A similar distribution of conserved versus nonconserved genes has been observed for the linear dsDNA genomes of poxviruses (159). Comparative genomic analysis revealed that rudiviruses and lipothrixviruses not only exchange genes horizontally with each other but also are capable of gene acquisition from their hosts (202). For example, genes for dUTPase, flavin-dependent thymidylate synthase (ThyX), and Holliday junction resolvase were likely transferred into viral genomes from cellular organisms (202, 214).

Similarly to fuselloviruses, lipothrixvirus AFV1 is engaged in a peculiar, but not well-understood, relationship with the conjugative, integrative plasmid pAH1 (14). The infection of the *Acidianus* host with AFV1 leads to the exclusion of the episomal, but not the integrated, form of plasmid pAH1. The molecular details of such a competition are not known. However, it was suggested that a DNA-binding protein (128), the only gene product that is shared by both replicons, might be responsible for this phenomenon (14). Interestingly, the screening of over 300 *Sulfolobus* isolates for the presence of extrachromosomal elements did not reveal a single isolate that would simultaneously contain both replicons, a virus and a plasmid (14), suggesting that such virus-plasmid incompatibility might also occur in natural environments.

Spherical viruses of the family *Globuloviridae*. The family *Globuloviridae* is currently represented by two viral species, *Pyrobaculum* spherical virus (PSV) (99) and *Thermoproteus tenax* spherical virus 1 (TTSV1) (5). Electron microscopy and biochemical studies revealed that spherical PSV virions consist of a lipid envelope, which encloses a helical nucleocapsid containing a linear dsDNA genome with inverted terminal repeats (Fig. 6) (99). PSV and TTSV1 share 15 homologous genes, but none of the genes in the two viral genomes display similarity to genes carried by organisms outside the family *Globuloviridae*. The majority of the predicted genes in the PSV (44 out of 48 genes) and TTSV1 (34 out of 38) genomes are carried on the same DNA strand. However, none of the four PSV genes that are present on the cDNA strand are present in the TTSV1 genome (5), suggesting that genome rearrangements in the two genomes occurred independently.

None of the PSV or TTSV1 proteins have been studied functionally. In order to gain insights into the biology of globuloviruses, a structural genomics approach has been applied to PSV (188). As a result, high-resolution structures have been obtained for five PSV proteins. Two of these proteins were found to be potentially involved in DNA binding (winged helix-turn-helix domain in ORF165a and Zn-binding domain in ORF126), and one, ORF137, was found to be potentially involved in RNA binding (Sm-like domain). The remaining two proteins, ORF131 and ORF239, do not share structural similarity with proteins in the public databases (154). In order to understand the biology of these viruses, it is now of great importance to subject at least one member of the *Globuloviridae* to a detailed functional and structural characterization.

STIV-like viruses. The overall virion organization of *Sulfolobus* turreted icosahedral virus (STIV) resembles that of bacterial tectiviruses and corticoviruses (Fig. 6). The STIV virion consists of an icosahedrally organized proteinaceous capsid that surrounds a protein-rich lipid membrane enclosing the circular dsDNA genome of 17.6 kb (170, 234). Interestingly, the major capsid proteins of tectivirus PRD1, corticovirus PM2, and archaeal STIV were found to possess the same structural fold (1, 2, 17, 131, 145). In addition, the three viruses encode homologous genome-packaging enzymes (262).

Recently, the genome sequence of a close STIV relative, STIV2, was determined (98). Sequence comparisons revealed an overall similar genome organization for the two viruses; however, several insertions/deletions were also observed (98). Notably, STIV, but not STIV2, carries a gene which has homologues in nearly all rudiviruses. A small membrane protein encoded by this gene is implicated in virion release at the end of the infection cycle (222). The presence of homologous genes in rudiviruses and STIV is in accordance with the experimental evidence for similar virion release mechanisms utilized by the two groups of viruses (21, 31). Consequently, it was suggested that STIV acquired its virion release gene relatively recently from *Sulfolobus*-infecting rudiviruses (222). A similar horizontal flux of lysis genes was also described for evolutionarily unrelated viruses infecting bacteria, such as tectivirus PRD1 and tailed viruses (146).

STIV is one of the most extensively studied archaeal viruses (86). Importantly, a genetic system has recently been established for STIV (290). This development will now allow a detailed genetic analysis of the virus and promises that in the near future, *Sulfolobus*-STIV will become one of the most

well-understood virus-host systems in archaea. Functional studies of STIV proteins are of great interest and should definitely be prioritized.

The two STIV-like viruses infect the same crenarchaeal host. However, bioinformatic analysis revealed the presence of two putative euryarchaeal proviruses, *Thermococcus kodakarensis* virus 4 (TKV4) and *Methanococcus voltae* virus (MVV), encoding STIV-like double-β-barrel MCPs as well as homologous genome-packaging ATPases (139). This finding suggests that viruses structurally related to STIV were or (more likely) still are infecting organisms from the second major phylum of the *Archaea*, the *Euryarchaeota*.

Euryarchaeal Viruses

The diversity of virion morphotypes associated with euryarchaeal hosts is considered to be less pronounced than that of crenarchaeal viruses (212, 213). Indeed, all currently isolated viruses of euryarchaea fall into three morphological groups: (i) head and tail (resembling bacterial viruses of the order *Caudovirales*), (ii) tailless icosahedral, and (iii) spindle shaped/pleomorphic (212, 236). However, a recent sampling of the viral diversity in saltern ponds, a beloved habitat of halophilic archaea, unveiled that our current knowledge of the diversity of euryarchaeal viruses is far from complete; viral morphotypes previously thought to be specific to crenarchaeal viruses are also present in hypersaline environments (252).

Tailed dsDNA viruses of archaea. Although head-and-tail viruses represent the majority of reported halovirus isolates, they seem to constitute only a minor part of euryarchaeal viruses (213, 252). The isolation and genome characterization of tailed archaeal viruses have so far been restricted to organisms of only two taxonomic orders of *Archaea*: the *Halobacteriales* (136, 195, 270) and *Methanobacteriales* (169, 203). owever, proviruses clearly related to tailed dsDNA viruses have also been reported for organisms of the orders *Methanococcales*, *Methanosarcinales* (149), and *Archaeoglobales* (M. Krupovic, unpublished data). Notably, a putative provirus related to tailed bacterial and euryarchaeal viruses was recently identified in the genome of "*Candidatus* Nitrososphaera viennensis" (152), a mesophilic archaeon belonging to the newly established phylum *Thaumarchaeota* (29). Phylogenetic analyses suggested that the latter phylum branched off before the separation of the *Crenarchaeota* and *Euryarchaeota* (29, 256). Therefore, the widespread distribution of tailed dsDNA (pro) viruses in the *Euryarchaeota* (149) and the presence of a related provirus in the *Thaumarchaeota* (152) suggest that the association of these viruses with the *Archaea* is ancient, possibly predating the diversification of bacteria and archaea from their last common ancestor.

Comparative analyses of these (pro)viral genomes revealed that the evolution of tailed archaeal virus genomes is governed by the same mechanisms as those described for their bacterial counterparts (37, 112), i.e., diversification, gene acquisition from different sources, illegitimate recombination, and intergenome rearrangements (136, 149, 195, 203, 270). Comparative genomics has also uncovered a set of conserved genes (149), which was found to include genes for the major capsid protein, prohead protease, portal, and the terminase complex. The very same set of genes underlies the structural similarity between tailed bacteriophages and herpesviruses (9, 12, 44, 224). Notably, products of these genes in the archaeal viruses were found to share significant sequence similarity with the corresponding proteins of tailed bacteriophages. In addition, fold prediction and structural modeling experiments suggested that the major capsid proteins of tailed archaeal viruses adopt the same topology as that of the equivalent proteins of tailed bacterial viruses and herpesviruses (149). The molecular principles of virion assembly and maturation as well as the virion structure of tailed archaeal viruses are thus expected to be very similar to those of tailed bacterial viruses and eukaryotic herpesviruses, possibly testifying for a common ancestry of these viruses.

Spindle-shaped/pleomorphic viruses. Viruses with a spindle-shaped appearance have been isolated from hyperthermophilic and halophilic euryarchaeal hosts of the orders *Thermococcales* (*Pyrococcus abyssi* virus 1 [PAV1]) and *Halobacteriales* (viruses His1 and His2), respectively (15, 91). Similarly to crenarchaeal fuselloviruses (231), these viruses display plasticity in virion morphology, which ranges from spindle shaped to elongated and flattened (for PAV1 [91]) or oblate and pleomorphic (for His2 [15]). Similarly, preparations of the predominantly pleomorphic haloarchaeal virus *Haloarcula hispanica* pleomorphic virus 1 (HHPV-1) occasionally also contain spindle-shaped particles (236).

The circular dsDNA genome of PAV1 did not show similarity to other archaeal spindle-shaped viruses or, indeed, to any known virus (90). However, PAV1 shares several genes with plasmids from the *Thermococcales* (3 genes with *Thermococcus nautilus* plasmid pTN2 and an additional gene with *Pyrococcus* sp. strain JT1 plasmid pRT1) (254, 284), suggesting a dynamic genetic interaction between these different types of mobile elements (253). Haloarchaeal virus His1 was first assigned to the family *Fuselloviridae*, which includes crenarchaeal viruses of similar morphology (16). However, the isolation of the second spindle-shaped haloarchaeal virus, His2 (Fig. 8), and analyses of the two genomes made it obvious that the gross morphology is the only common characteristic between crenarchaeal fuselloviruses and euryarchaeal His1/2 viruses (15). Consequently, viruses His1 and His2 were reassigned to the genus *Salterprovirus*. Both viruses contain linear dsDNA genomes with inverted terminal repeats and terminal proteins (210). Both viruses encode type B DNA polymerases, and their genomes are likely to be replicated in a protein-primed manner (15, 210). The two viruses are only distantly related to each other, with generally no similarity between their corresponding protein sequences (15).

The genome sequence of the recently characterized pleomorphic haloarchaeal virus HHPV-1 unexpectedly revealed that a block of genes encoding virion structural proteins of HHPV-1 is colinear and shares significant similarity with that of the salterprovirus His2 (Fig. 8) but not that of His1 (204, 236). The predicted genes in the His2 genome are organized into two modules, which are seemingly transcribed in the opposite directions, toward the termini of the linear genome from a central region (Fig. 8) (15). Interestingly, His2 genes that share sequence similarity with those of HHPV-1 occupy almost the entirety of one of these modules (Fig. 8). The other module encodes proteins implicated in His2 genome replication (DNA polymerase and a number of proteins with potential zinc finger DNA-binding motifs [15]). Such a module is not present in the circular HHPV-1 genome, which instead encodes a putative rolling-circle replication (RCR) initiation protein at an equivalent

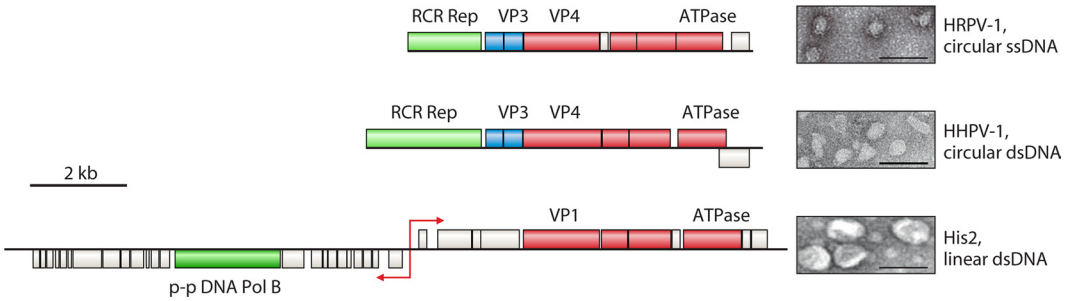

FIG. 8. Genomic relationship between euryarchaeal pleomorphic and spindle-shaped viruses with different nucleic acid types. The electron micrographs and the types of nucleic acids comprising the genomes of the corresponding viruses are indicated on the right. Boxes above and below the vertical lines represent genes transcribed from left to right and from right to left, respectively (with kinked arrows indicating the direction of transcription). Genes shared by the three viruses *Halorubrum* pleomorphic virus 1 (HRPV-1), *Haloarcula hispanica* pleomorphic virus 1 (HHPV-1), and haloarchaeal virus His2 are shown in red, while those restricted to HRPV-1 and HHPV-1 are in blue. Genes encoding genome replication proteins are in green. RCR Rep, rolling-circle replication initiation proteins; p-p DNA Pol B, protein-primed family B DNA polymerase. Note that RCR initiation proteins of HRPV-1 and HHPV-1 share only 9% pairwise identity (236). (The HRPV-1 micrograph is reproduced from reference 204 with permission from John Wiley & Sons. The HHPV-1 micrograph is reproduced from reference 236. The His2 micrograph is reproduced from reference 15 with permission from Elsevier.) Bars, 100 nm.

position (Fig. 8) (204, 236). It therefore seems that His2 or HHPV-1 donated a block of structural protein-coding genes to a linear plasmid and a circular plasmid, respectively, thereby transforming a plasmid into a virus. A similar evolutionary scenario was also proposed to be accountable for the origin of plant geminiviruses (151) as well as bacterial siphovirus N15 (228, 229).

The circular ssDNA genome of another haloarchaeal pleomorphic virus, *Halorubrum* pleomorphic virus 1 (HRPV-1), presented an even greater amazement. Comparative analyses revealed that it is remarkably similar to the dsDNA genome of HHPV-1 in both gene content and organization (Fig. 8), leaving no doubt about the evolutionary relationship between these ssDNA and dsDNA viruses (204, 236). In addition, the patterns of virion structural proteins were also very similar for HHPV-1 and HRPV-1 (236). It was also noticed that the protein pattern is reminiscent of that of the pleomorphic ssDNA mycoplasma virus L172 (71), suggesting that the three viruses might be evolutionarily related (204, 236). Regrettably, the genome sequence of L172 is not yet available. Interestingly, although both HRPV-1 and HHPV-1 appear to replicate their genomes via the rolling-circle mechanism, the two RCR initiation proteins display only 9% pairwise sequence identity. In contrast, the sequence identity between other proteins encoded by the two viruses is in the range of 22 to 55% (236). It therefore appears that the genes for genome replication proteins were acquired by HRPV-1, HHPV-1, and His2 from three distinct sources. Notably, a number of proviruses related to the three viruses were identified in the genomes of diverse species of the *Halobacteriales* (70, 204), pointing to a wide distribution of the pleomorphic viruses in haloarchaea and also suggesting that these proviruses might be actively participating in the recombination that shapes the genetic diversity of this viral group. Given the fact that the genetic similarity between the two salterproviruses His1 and His2 does not extend far beyond the genome replication machinery (15), the evolutionary relationship between the two viruses has to be reevaluated.

The example of HHPV-1/HRPV-1 suggests that there is a certain liberty in the choice of the replicative form of the genome that will be encapsidated into mature virions. A reminiscent situation can be observed with eukaryotic retrotranscribing viruses: members of the family *Retroviridae* encapsidate the ssRNA replicative intermediate, while hepadnaviruses (family *Hepadnaviridae*) seem to have chosen the dsDNA form (259, 295). Nevertheless, such a close relationship between viruses with different types of nucleic acids (dsDNA versus ssDNA) is unprecedented among prokaryotic viruses and shakes the foundation of the current virus taxonomy.

Icosahedral tailless viruses. Haloarchaeal virus SH1 is the only isolated representative of such a morphotype among euryarchaeal viruses (211). In its overall virion organization, SH1 resembles bacterial tectiviruses (Table 2; see also above). The linear dsDNA genome also has inverted terminal repeats and covalently attached terminal proteins (13, 210). Unexpectedly, however, differently from all other known viruses with such a genome type (with the exception of lipothrixviruses [see above]), SH1 does not encode an identifiable DNA polymerase (13). Although analyses of the SH1 protein sequences did not reveal much of a relationship to other archaeal viruses (13), it did point to a connection between SH1 and bacterial virus P23-77 (circular dsDNA genome), infecting *Thermus thermophilus* (120). The two viruses were found to share a block of genes encoding the putative genome-packaging ATPase and the two major capsid proteins (Fig. 9). In line with this finding, the evolutionary link between SH1 and P23-77 has also been deduced from structural analyses, which showed very similar organizations ($T = 28$) of their virions (Fig. 9) (117, 118). Notably, several putative proviruses related to SH1 and P23-77 were recently identified in the genomes of halophilic archaea and thermophilic bacteria (121). Further sampling and virus isolation efforts are required to more fully understand the relationship between these haloarchaeal and bacterial viruses.

GENOMIC RELATIONSHIP TO OTHER MOBILE GENETIC ELEMENTS

A number of viruses belonging to different families rely on plasmids for infectivity. The latter dependence is due to the fact that some plasmids encode cell envelope structures that

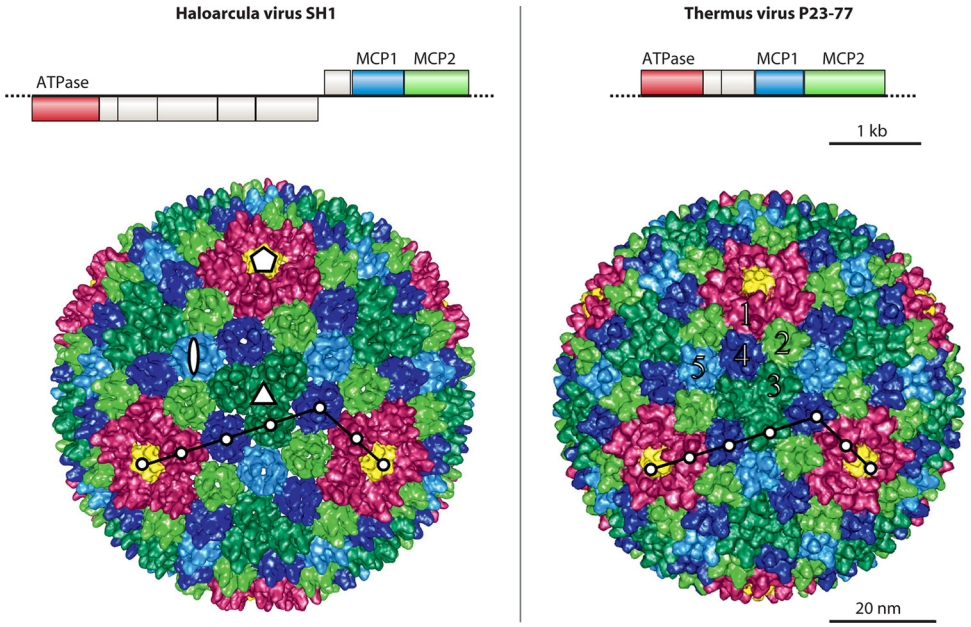

FIG. 9. Comparison of haloarchaeal virus SH1 and bacterial virus P23-77. Surface representations of the two virions are viewed along an icosahedral 3-fold symmetry axis. Symmetry axes are indicated with a white ellipse (2-fold), a triangle (3-fold), and a pentagon (5-fold). Symmetry-related capsomers are shown in the same color. The triangulation number (*T*) describes the geometrical arrangement of the capsomers and is the same for the SH1 and P23-77 capsids ($T = 28$). (The capsid representations are reproduced from reference 118 with permission from Elsevier.) Genomic regions encoding homologous ATPases and the two major capsid proteins (MCP1 and MCP2) of SH1 and P23-77 are depicted above their respective capsid representations.

are targeted by viruses during host cell recognition and the delivery of the genetic material into the cell interior. For example *E. coli*-infecting inoviruses and leviviruses adsorb to the tip and to the sides, respectively, of the conjugative pili encoded by the F plasmid (IncF group) (275, 289), while PRD1-like tectiviruses and levivirus PRR1 are known to use the cell envelope-associated DNA transfer complex encoded by the IncP group plasmids (95, 96, 191). Similarly, a tailed bacteriophage, podovirus J, was reported to adsorb to the sides of plasmid-encoded pili (26). Plasmids, on the other hand, might encode factors that abolish viral replication in a number of different ways (45, 69, 114, 282). Viruses return the favor by encoding antirestriction proteins or evolving toward resistance to the detrimental effects of these factors (25, 245, 281). There is therefore a constant struggle between viruses and plasmids inside the cells which drives the evolution of both the mobile elements and their hosts.

Comparative genomics has further revealed that there is a clear evolutionary link between viruses and plasmids. Although the latter were proposed to represent remnants of ancient viruses (and not vice versa) that have lost the ability to produce virions (83), the question regarding the directionality of descent remains open. Viruses and plasmids encode a number of common proteins that do not have homologues in cellular proteomes. Notable examples of such proteins are initiators of rolling-circle replication, superfamily III helicases, or primase/polymerase proteins (138, 167), which are specific to viruses and plasmids and do not participate in the replication of cellular chromosomes. The origin of other mobile genetic elements (MGEs), such as transposons, integrons, and genomic islands (271), might also stem from ancient viruses (84). Whatever the line of descent, a feature characteristic to all MGEs (viruses included) is their genomic modularity; i.e., MGEs might be viewed as assemblages of distinct functional modules (160, 194, 271). Certain modules are common to different types of MGEs and can be exchanged horizontally between them (e.g., between plasmids and viruses). Such a horizontal transfer occurs relatively frequently, as judged from the numerous examples of the nonorthologous replacement of genome replication and recombination protein-coding genes in various prokaryotic plasmids and viruses (e.g., see references 92, 143, 149, 194, and 286).

The conversion of one MGE type into another is, in principle, only one gene away (Fig. 10). A virus that loses the ability to build a virion (simple viruses encode a single protein type for capsid assembly [238]) can no longer be considered a virus and becomes a plasmid, and vice versa: a plasmid that learns how to build a capsid becomes a virus (Fig. 10). The latter scenario was suggested to lie at the origin of plant geminiviruses (151). Similarly, the domestication of a transposase gene by a viral genome (e.g., myovirus Mu) brings the virus closer to being a transposon, where the subsequent loss of the structural module would finalize the metamorphosis. Genomes of most simple viruses are composed of two functional modules, one for the virion structure and the other one for genome replication (142). While the genome replication modules in viruses are prone to horizontal transfer between distinct types of MGEs as well as their host cells (140, 150, 286), the virion structure module is unique to viruses, thereby defining their identity and distinguishing viruses from other types of MGEs (10, 145).

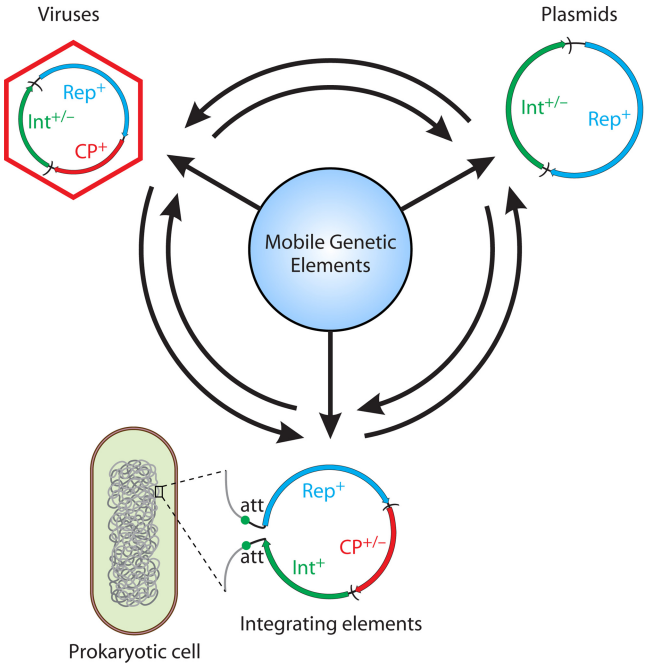

FIG. 10. Genomic relationship between different groups of mobile genetic elements. Genomic modularity is characteristic of different types of mobile genetic elements, such as viruses, plasmids, and integrating or transposing elements. Transitions from one type of an element to another (e.g., from a plasmid to a virus) appear to have occurred on multiple independent occasions during evolution by the acquisition or loss of genes that are specific to a given group of mobile elements. The genes that are dispensable for a given type of mobile elements are shown in parentheses. CP, Rep, and Int, genes for capsid, replication, and recombination (either integrases or transposases) proteins, respectively; *att*, attachment sites, which mark the borders of an integrated element.

IS THERE A SIGNAL OF VERTICAL INHERITANCE IN VIRAL GENOMES?

Viral genomes consist of several functional modules (structure, replication, gene expression, and host lysis, etc.), and their evolution proceeds by the shuffling of these modules by means of illegitimate recombination with other viruses (often evolutionarily unrelated), plasmids, or even the host chromosome. Such a recombination within a given viral group (or lineage) seems to be fairly unrestricted; i.e., virtually any gene or, in the case of genes for interacting proteins, blocks of genes can undergo lateral transfer from one virus to another (110). However, genomic analysis indicates that functionally different types of genes/modules are not equally prone to horizontal exchange between distinct viral lineages. Comparative genomics analyses of corticoelements provide a good illustration of this point. Three out of four functional modules comprising the genomes of corticoviruses are frequently exchanged for functionally equivalent but nonhomologous units from other viruses and plasmids (143). Most notable is the freelance nature of genes encoding replication proteins. Despite the fact that these genes are usually among the most conserved ones at the sequence level, their evolution seems to occur independently from that of other viral functions (140). So is there a signal of vertical inheritance in viral genomes? The answer would probably be "no" if viral identity (evolutionary history) is considered to be a sum of all the functional modules that constitute the viral genome, where each of these modules has an equal additive value. Such an approach usually brings us to compare viral proteins that are more conserved at the sequence level and to dismiss from the analysis those that have diverged to a greater extent. The latter group of proteins generally includes structural components of the virion that are under different evolutionary constraints than, for example, proteins with an enzymatic activity. Such selection is dictated largely by the inability to extract the phylogenetic signal from the more divergent proteins; i.e., the reason is technical and not biological. However, when talking about viruses, scientists usually consider them to be related if they share both common architectural and genomic characteristics. For example, all tailed dsDNA bacteriophages are considered to be evolutionarily related (9), despite the fact that different members of this supergroup utilize different genome replication and transcription strategies, nonhomologous host recognition and lysis proteins. The common denominator for all these viruses is the basic module for virion assembly and structure, which defines their identity. A tailed virus with an internal membrane and a head composed of double-β-barrel major capsid proteins has not yet been isolated, nor has a levivirus with the HK97-like MCP been characterized. However, we do know that some tailed viruses (e.g., phi29) utilize genome replication proteins homologous to those of tectiviruses (protein-primed type B DNA polymerase) (243); we also know that tectiviruses use the host lysis system that is characteristic of tailed bacterial viruses (146). Therefore, the only determinants that are consistently retained within a given group of viruses are those responsible for the construction of the virion. It should also be noted that the inability to reconstruct the phylogeny for viruses that diverged further back in time does not testify to the absence of a common evolutionary history. Consequently, despite the extensive shuffling of functional modules responsible for various genetic processes and interactions with the host, a set of genes defining the structural character of a virion is restricted to and preserved in a given virus lineage and therefore can be used to dissect deep vertical relationships between evolutionarily distant viruses.

WHERE ARE WE HEADING WITH THE GENOMICS OF PROKARYOTIC VIRUSES?

More Genomes for "Underrepresented" Viral Groups

Do the 600 complete genomes of prokaryotic viruses provide satisfactory information on their genetic diversity and evolution? Certainly not! Although comparative genomics did indeed help to draw a rather detailed and informative picture of some viral groups (e.g., mycobacteriophages [101, 207]), genomic data for the vast majority of bacterial and archaeal viruses are mostly missing, making this a challenge for the genome community. At the top of the list are viral families for which only a single virus isolate has been sequenced (Tables 1 and 2); these include most of the archaeal virus groups and the bacterial virus families *Plasmaviridae* and *Corticoviridae*.

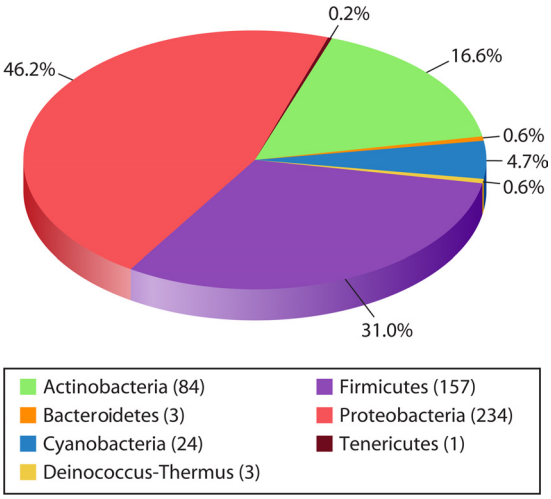

FIG. 11. Distribution of tailed bacterial viruses according to the taxonomic grouping of their hosts. Numerical values next to the names of bacterial phyla represent the exact numbers of genome sequences available for viruses infecting bacteria from a corresponding phylum. The numbers do not include the genome sequences available for closely related viruses that are considered different isolates of the same virus strain. The data were obtained from GenBank (June 2011).

More Genomes for Closely Related Viruses

The fine details of virus evolution are reachable through comparisons of closely related viral genomes. This is well illustrated by the depth of insights obtained from analyses of 80 complete genomes of mycobacteriophages (all isolated using the same host, *Mycobacterium smegmatis* mc^{2}155) (207) and 57 clusters of virion assembly genes from proteobacterium-infecting P22-like viruses (39). In addition, genomic and phenotypic comparisons of 10 *Pseudomonas*-infecting phiKMV-like podoviruses indicated that small "intraspecies" genomic variations can invoke significant phenotypic changes (41). We need to force such studies forward to better understand the pace and mechanisms of viral genome evolution as well as the dynamics of viral populations in the environment.

Wider Phylogenetic Coverage

Further research should also be directed toward achieving a more comprehensive coverage of viruses infecting phylogenetically diverse hosts. There are approximately 52 currently identifiable bacterial phyla, half of which (26) can be cultured under laboratory conditions (226). Our current understanding of the biology and genomics of tailed bacteriophages (a group of prokaryotic viruses with the highest number of complete genome sequences) is based on investigations of viruses infecting only a few of these bacterial phyla. More than 90% of the complete genome sequences of tailed viruses comes from those infecting hosts falling into only three different phyla (Fig. 11), namely, the *Proteobacteria*, *Firmicutes*, and *Actinobacteria* (mainly mycobacteria). The remaining 6% of available genome sequences are (unevenly) distributed among viruses infecting the *Cyanobacteria*, *Bacteroidetes*, *Deinococcus-Thermus*, and *Tenericutes* (Fig. 11). Tailed phages are often considered to be the most abundant organisms in the biosphere (81). However, the reality is that we do not know that, nor do we know if bacteria from other phyla are susceptible to tailed viruses. For example, only archaea belonging to the phyla *Euryarchaeota* and *Thaumarchaeota* appear to be infected by tailed dsDNA viruses (149, 152), while an association of such viruses (or related proviruses) with organisms from the phylum *Crenarchaeota* could not be demonstrated despite continuous efforts (205). Consequently, the restricted phylogenetic coverage of virus-host systems does not provide a global picture of the diversity and evolution of these viruses, and the characterization of viruses, both bacterial and archaeal, from phylogenetically diverse hosts is likely to reveal new surprises.

Extending Functional Studies to Poorly Understood Virus-Host Systems

Viral genomics studies rarely go far beyond *in silico* sequence analyses. At best, they are supplemented with characterizations of structural virion proteins by mass spectrometry and/or N-terminal protein sequencing. Genome annotations are based mainly on the information obtained from functional studies of the few model virus-host systems (mainly enterobacterium- and *Bacillus*-infecting viruses) that have already been selected at the dawn of molecular biology. An important line of research is to extend functional studies of viral proteins to less studied systems. The few attempts at this have already resulted in exciting findings, leading to a better understanding of virus-host interactions.

Interplay between Viruses and Other Mobile Genetic Elements

Numerous functional and genomics studies have revealed that there is an intimate interplay between viruses and plasmids both in archaea and in bacteria. The two types of elements not only share a gene pool (143, 229, 253) but also depend on each other (8, 60, 119, 283) or, on the contrary, compete (14) inside the host cell. A single cell thus represents a unique ecosystem where viruses, plasmids, integrons, and transposable elements all come together and where each one of these genetic elements tries to survive and stay fit. Studies of virus-plasmid relationships and interactions are of outstanding interest and should be continued in the future. It is now especially important to comprehend the molecular basis underlying the coexistence and coevolution of different types of MGEs.

Merging Culture-Dependent and Culture-Independent Approaches

The realization that viruses play an important role in the global ecosystem (59, 292) accompanied by the advances in high-throughput sequencing technologies stimulated numerous metagenomic studies of viromes (genomes of all the viruses that inhabit a particular environment) in diverse environments. Although such studies have generated impressive amounts of data and pointed out the immense genetic diversity in the virosphere, they were less successful in shedding light on the evolution of viruses and virus-host relationships in these habitats. This is in part due to the fact that the amount of useful information that can be extracted from virome analyses

depends largely on the existing database of reference viral genome sequences. To make the best use of the metagenomic data, culture-independent and culture-dependent techniques should be merged into a continuous pipeline, where culture-independent high-throughput approaches target specific virus-host systems for consequent isolation and characterization (144). The newly developed microfluidic digital PCR technique (269) might prove itself useful in achieving this goal. We envision that the field of prokaryotic virus genomics is yet to flourish and will lead to numerous fascinating discoveries in the future.

ACKNOWLEDGMENTS

This work was supported by the European Molecular Biology Organization (long-term fellowship ALTF 347-2010 to M.K.), the Agence Nationale de la Recherche (program Blanc to D.P.), the U.S. National Institutes of Health (grant GM51975 to R.W.H.), and the Academy of Finland Center of Excellence in Virus Research (2006 to 2011; grant 11296841 to D.H.B.).

REFERENCES

1. **Abrescia, N. G., et al.** 2004. Insights into assembly from structural analysis of bacteriophage PRD1. Nature **432:**68–74.
2. **Abrescia, N. G., et al.** 2008. Insights into virus evolution and membrane biogenesis from the structure of the marine lipid-containing bacteriophage PM2. Mol. Cell **31:**749–761.
3. **Ackermann, H. W.** 2007. 5500 phages examined in the electron microscope. Arch. Virol. **152:**227–243.
4. **Agirrezabala, X., et al.** 2007. Quasi-atomic model of bacteriophage T7 procapsid shell: insights into the structure and evolution of a basic fold. Structure **15:**461–472.
5. **Ahn, D. G., et al.** 2006. TTSV1, a new virus-like particle isolated from the hyperthermophilic crenarchaeote *Thermoproteus tenax*. Virology **351:**280–290.
6. **Angly, F. E., et al.** 2006. The marine viromes of four oceanic regions. PLoS Biol. **4:**e368.
7. **Arbiol, C., A. M. Comeau, M. Kutateladze, R. Adamia, and H. M. Krisch.** 2010. Mobile regulatory cassettes mediate modular shuffling in T4-type phage genomes. Genome Biol. Evol. **2:**140–152.
8. **Arnold, H. P., et al.** 1999. The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. Mol. Microbiol. **34:**217–226.
9. **Baker, M. L., W. Jiang, F. J. Rixon, and W. Chiu.** 2005. Common ancestry of herpesviruses and tailed DNA bacteriophages. J. Virol. **79:**14967–14970.
10. **Bamford, D. H.** 2003. Do viruses form lineages across different domains of life? Res. Microbiol. **154:**231–236.
11. **Bamford, D. H., and J. K. Bamford.** 2006. Lipid-containing bacteriophage PM2, the type-organism of *Corticoviridae*, p. 171–175. *In* R. Calendar (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.
12. **Bamford, D. H., J. M. Grimes, and D. I. Stuart.** 2005. What does structure tell us about virus evolution? Curr. Opin. Struct. Biol. **15:**655–663.
13. **Bamford, D. H., et al.** 2005. Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon *Haloarcula hispanica*. J. Virol. **79:**9097–9107.
14. **Basta, T., J. Smyth, P. Forterre, D. Prangishvili, and X. Peng.** 2009. Novel archaeal plasmid pAH1 and its interactions with the lipothrixvirus AFV1. Mol. Microbiol. **71:**23–34.
15. **Bath, C., T. Cukalac, K. Porter, and M. L. Dyall-Smith.** 2006. His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, *Salterprovirus*. Virology **350:**228–239.
16. **Bath, C., and M. L. Dyall-Smith.** 1998. His1, an archaeal virus of the *Fuselloviridae* family that infects *Haloarcula hispanica*. J. Virol. **72:**9392–9395.
17. **Benson, S. D., J. K. Bamford, D. H. Bamford, and R. M. Burnett.** 1999. Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. Cell **98:**825–833.
18. **Bergh, O., K. Y. Borsheim, G. Bratbak, and M. Heldal.** 1989. High abundance of viruses found in aquatic environments. Nature **340:**467–468.
19. **Bernhardt, T. G., I. N. Wang, D. K. Struck, and R. Young.** 2001. A protein antibiotic in the phage Qbeta virion: diversity in lysis targets. Science **292:**2326–2329.
20. **Betancourt, A. J.** 2009. Genomewide patterns of substitution in adaptively evolving populations of the RNA bacteriophage MS2. Genetics **181:**1535–1544.
21. **Bize, A., et al.** 2009. A unique virus release mechanism in the Archaea. Proc. Natl. Acad. Sci. U. S. A. **106:**11306–11311.
22. **Bize, A., et al.** 2008. Viruses in acidic geothermal environments of the Kamchatka Peninsula. Res. Microbiol. **159:**358–366.
23. **Blakely, G. W.** 2004. Smarter than the average phage. Mol. Microbiol. **54:**851–854.
24. **Blum, H., W. Zillig, S. Mallok, H. Domdey, and D. Prangishvili.** 2001. The genome of the archaeal virus SIRV1 has features in common with genomes of eukaryal viruses. Virology **281:**6–9.
25. **Bouchard, J. D., and S. Moineau.** 2004. Lactococcal phage genes involved in sensitivity to AbiK and their relation to single-strand annealing proteins. J. Bacteriol. **186:**3649–3652.
26. **Bradley, D. E., F. A. Sirgel, J. N. Coetzee, R. W. Hedges, and W. F. Coetzee.** 1982. Phages C-2 and J: IncC and IncJ plasmid-dependent phages, respectively. J. Gen. Microbiol. **128:**2485–2498.
27. **Brentlinger, K. L., et al.** 2002. *Microviridae*, a family divided: isolation, characterization, and genome sequence of phiMH2K, a bacteriophage of the obligate intracellular parasitic bacterium *Bdellovibrio bacteriovorus*. J. Bacteriol. **184:**1089–1094.
28. **Brewer, G. J.** 1978. Membrane-localized replication of bacteriophage PM2. Virology **84:**242–245.
29. **Brochier-Armanet, C., B. Boussau, S. Gribaldo, and P. Forterre.** 2008. Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. Nat. Rev. Microbiol. **6:**245–252.
30. **Brochier-Armanet, C., P. Forterre, and S. Gribaldo.** 2011. Phylogeny and evolution of the Archaea: one hundred genomes later. Curr. Opin. Microbiol. **14:**274–281.
31. **Brumfield, S. K., et al.** 2009. Particle assembly and ultrastructural features associated with replication of the lytic archaeal virus *Sulfolobus* turreted icosahedral virus. J. Virol. **83:**5964–5970.
32. **Brüssow, H., and F. Desiere.** 2001. Comparative phage genomics and the evolution of *Siphoviridae*: insights from dairy phages. Mol. Microbiol. **39:**213–222.
33. **Campos, J., et al.** 2003. Novel type of specialized transduction for CTX phi or its satellite phage RS1 mediated by filamentous phage VGJ phi in *Vibrio cholerae*. J. Bacteriol. **185:**7231–7240.
34. **Canchaya, C., G. Fournous, and H. Brussow.** 2004. The impact of prophages on bacterial chromosomes. Mol. Microbiol. **53:**9–18.
35. **Canchaya, C., C. Proux, G. Fournous, A. Bruttin, and H. Brussow.** 2003. Prophage genomics. Microbiol. Mol. Biol. Rev. **67:**238–276.
36. **Casjens, S.** 2003. Prophages and bacterial genomics: what have we learned so far? Mol. Microbiol. **49:**277–300.
37. **Casjens, S. R.** 2005. Comparative genomics and evolution of the tailed-bacteriophages. Curr. Opin. Microbiol. **8:**451–458.
38. **Casjens, S. R.** 2008. Diversity among the tailed-bacteriophages that infect the *Enterobacteriaceae*. Res. Microbiol. **159:**340–348.
39. **Casjens, S. R., and P. A. Thuman-Commike.** 2011. Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology **411:**393–415.
40. **Caspar, D. L., and A. Klug.** 1962. Physical principles in the construction of regular viruses. Cold Spring Harb. Symp. Quant. Biol. **27:**1–24.
41. **Ceyssens, P. J., et al.** 2011. Phenotypic and genotypic variations within a single bacteriophage species. Virol. J. **8:**134.
42. **Ceyssens, P. J., and R. Lavigne.** 2010. Bacteriophages of *Pseudomonas*. Future Microbiol. **5:**1041–1055.
43. **Chen, D. H., et al.** 2011. Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. Proc. Natl. Acad. Sci. U. S. A. **108:**1355–1360.
44. **Cheng, H., N. Shen, J. Pei, and N. V. Grishin.** 2004. Double-stranded DNA bacteriophage prohead protease is homologous to herpesvirus protease. Protein Sci. **13:**2260–2269.
45. **Cheng, X., W. Wang, and I. J. Molineux.** 2004. F exclusion of bacteriophage T7 occurs at the cell membrane. Virology **326:**340–352.
46. **Cherwa, J. E., and B. A. Fane.** 2011. *Microviridae*: microviruses and gokushoviruses. *In* Encyclopedia of life sciences. John Wiley & Sons, Ltd., Chichester, United Kingdom. doi:10.1002/9780470015902.a0000781.pub2.
47. **Chetverin, A. B., D. S. Kopein, H. V. Chetverina, A. A. Demidenko, and V. I. Ugarov.** 2005. Viral RNA-directed RNA polymerases use diverse mechanisms to promote recombination between RNA molecules. J. Biol. Chem. **280:**8748–8755.
48. **Chipman, P. R., M. Agbandje-McKenna, J. Renaudin, T. S. Baker, and R. McKenna.** 1998. Structural analysis of the *Spiroplasma* virus, SpV4: implications for evolutionary variation to obtain host diversity among the *Microviridae*. Structure **6:**135–145.
49. **Chopin, M. C., A. Rouault, S. D. Ehrlich, and M. Gautier.** 2002. Filamentous phage active on the gram-positive bacterium *Propionibacterium freudenreichii*. J. Bacteriol. **184:**2030–2033.
50. **Chouikha, I., L. Charrier, S. Filali, A. Derbise, and E. Carniel.** 2010. Insights into the infective properties of YpfPhi, the *Yersinia pestis* filamentous phage. Virology **407:**43–52.
51. **Clarke, I. N., et al.** 2004. Chlamydiaphage Chp2, a skeleton in the phiX174 closet: scaffolding protein and procapsid identification. J. Bacteriol. **186:**7571–7574.

52. **Clokie, M. R., and N. H. Mann.** 2006. Marine cyanophages and light. Environ. Microbiol. **8:**2074–2082.
53. **Clokie, M. R., A. D. Millard, and N. H. Mann.** 2010. T4 genes in the marine ecosystem: studies of the T4-like cyanophages and their role in marine ecology. Virol. J. **7:**291.
54. **Clokie, M. R., et al.** 2006. Transcription of a 'photosynthetic' T4-type phage during infection of a marine cyanobacterium. Environ. Microbiol. **8:**827–835.
55. **Clore, A. J., and K. M. Stedman.** 2007. The SSV1 viral integrase is not essential. Virology **361:**103–111.
56. **Comeau, A. M., C. Bertrand, A. Letarov, F. Tetart, and H. M. Krisch.** 2007. Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. Virology **362:**384–396.
57. **Cvirkaite-Krupovic, V., M. Krupovic, R. Daugelavicius, and D. H. Bamford.** 2010. Calcium ion-dependent entry of the membrane-containing bacteriophage PM2 into its *Pseudoalteromonas* host. Virology **405:**120–128.
58. **Dammeyer, T., S. C. Bagby, M. B. Sullivan, S. W. Chisholm, and N. Frankenberg-Dinkel.** 2008. Efficient phage-mediated pigment biosynthesis in oceanic cyanobacteria. Curr. Biol. **18:**442–448.
59. **Danovaro, R., et al.** 2011. Marine viruses and global climate change. FEMS Microbiol. Rev. **35:**993–1034.
60. **Davis, B. M., H. H. Kimsey, A. V. Kane, and M. K. Waldor.** 2002. A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. EMBO J. **21:**4240–4249.
61. **Davis, B. M., et al.** 2000. Convergence of the secretory pathways for cholera toxin and the filamentous phage, CTXphi. Science **288:**333–335.
62. **Davis, B. M., K. E. Moyer, E. F. Boyd, and M. K. Waldor.** 2000. CTX prophages in classical biotype *Vibrio cholerae*: functional phage genes but dysfunctional phage genomes. J. Bacteriol. **182:**6992–6998.
63. **Davis, B. M., and M. K. Waldor.** 2000. CTXphi contains a hybrid genome derived from tandemly integrated elements. Proc. Natl. Acad. Sci. U. S. A. **97:**8572–8577.
64. **Debartolomeis, J., and V. J. Cabelli.** 1991. Evaluation of an *Escherichia coli* host strain for enumeration of F male-specific bacteriophages. Appl. Environ. Microbiol. **57:**1301–1305.
65. **Desiere, F., S. Lucchini, C. Canchaya, M. Ventura, and H. Brussow.** 2002. Comparative genomics of phages and prophages in lactic acid bacteria. Antonie Van Leeuwenhoek **82:**73–91.
66. **Desnues, C., et al.** 2008. Biodiversity and biogeography of phages in modern stromatolites and thrombolites. Nature **452:**340–343.
67. **Desplats, C., C. Dez, F. Tetart, H. Eleaume, and H. M. Krisch.** 2002. Snapshot of the genome of the pseudo-T-even bacteriophage RB49. J. Bacteriol. **184:**2789–2804.
68. **Desplats, C., and H. M. Krisch.** 2003. The diversity and evolution of the T4-type bacteriophages. Res. Microbiol. **154:**259–267.
69. **Duckworth, D. H., J. Glenn, and D. J. McCorquodale.** 1981. Inhibition of bacteriophage replication by extrachromosomal genetic elements. Microbiol. Rev. **45:**52–71.
70. **Dyall-Smith, M. L., et al.** 2011. *Haloquadratum walsbyi*: limited diversity in a global pond. PLoS One **6:**e20968.
71. **Dybvig, K., J. A. Nowak, T. L. Sladek, and J. Maniloff.** 1985. Identification of an enveloped phage, mycoplasma virus L172, that contains a 14-kilobase single-stranded DNA genome. J. Virol. **53:**384–390.
72. **Effantin, G., P. Boulanger, E. Neumann, L. Letellier, and J. F. Conway.** 2006. Bacteriophage T5 structure reveals similarities with HK97 and T4 suggesting evolutionary relationships. J. Mol. Biol. **361:**993–1002.
73. **Elkins, J. G., et al.** 2008. A korarchaeal genome reveals insights into the evolution of the Archaea. Proc. Natl. Acad. Sci. U. S. A. **105:**8102–8107.
74. **Fane, B. A., et al.** 2006. ϕX174 et al., the *Microviridae*, p. 129–145. *In* R. Calendar (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.
75. **Fane, B. A., and P. E. Prevelige, Jr.** 2003. Mechanism of scaffolding-assisted viral assembly. Adv. Protein Chem. **64:**259–299.
76. **Faruque, S. M., et al.** 2002. RS1 element of *Vibrio cholerae* can propagate horizontally as a filamentous phage exploiting the morphogenesis genes of CTXphi. Infect. Immun. **70:**163–170.
77. **Faruque, S. M., et al.** 2003. CTXphi-independent production of the RS1 satellite phage by *Vibrio cholerae*. Proc. Natl. Acad. Sci. U. S. A. **100:**1280–1285.
78. **Fauquet, C. M., M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball (ed.).** 2005. Virus taxonomy. Eighth report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, CA.
79. **Filée, J., E. Bapteste, E. Susko, and H. M. Krisch.** 2006. A selective barrier to horizontal gene transfer in the T4-type bacteriophages that has preserved a core genome with the viral replication and structural genes. Mol. Biol. Evol. **23:**1688–1696.
80. **Filée, J., and M. Chandler.** 2008. Convergent mechanisms of genome evolution of large and giant DNA viruses. Res. Microbiol. **159:**325–331.
81. **Filée, J., F. Tetart, C. A. Suttle, and H. M. Krisch.** 2005. Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. Proc. Natl. Acad. Sci. U. S. A. **102:**12471–12476.
82. **Fokine, A., et al.** 2005. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. Proc. Natl. Acad. Sci. U. S. A. **102:**7163–7168.
83. **Forterre, P.** 2005. The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. Biochimie **87:**793–803.
84. **Forterre, P., and D. Prangishvili.** 2009. The great billion-year war between ribosome- and capsid-encoding organisms (cells and viruses) as the major source of evolutionary novelties. Ann. N. Y. Acad. Sci. **1178:**65–77.
85. **Friedman, S. D., F. J. Genthner, J. Gentry, M. D. Sobsey, and J. Vinje.** 2009. Gene mapping and phylogenetic analysis of the complete genome from 30 single-stranded RNA male-specific coliphages (family *Leviviridae*). J. Virol. **83:**11233–11243.
86. **Fulton, J., et al.** 2009. Genetics, biochemistry and structure of the archaeal virus STIV. Biochem. Soc. Trans. **37:**114–117.
87. **Gaidelytė, A., S. T. Jaatinen, R. Daugelavičius, J. K. Bamford, and D. H. Bamford.** 2005. The linear double-stranded DNA of phage Bam35 enters lysogenic host cells, but the late phage functions are suppressed. J. Bacteriol. **187:**3521–3527.
88. **Gaidelytė, A., M. Vaara, and D. H. Bamford.** 2007. Bacteria, phages and septicemia. PLoS One **2:**e1145.
89. **Garner, S. A., J. S. Everson, P. R. Lambden, B. A. Fane, and I. N. Clarke.** 2004. Isolation, molecular characterisation and genome sequence of a bacteriophage (Chp3) from *Chlamydophila pecorum*. Virus Genes **28:**207–214.
90. **Geslin, C., et al.** 2007. Analysis of the first genome of a hyperthermophilic marine virus-like particle, PAV1, isolated from *Pyrococcus abyssi*. J. Bacteriol. **189:**4510–4519.
91. **Geslin, C., et al.** 2003. PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, "*Pyrococcus abyssi*." J. Bacteriol. **185:**3888–3894.
92. **Gonnet, M., G. Erauso, D. Prieur, and M. Le Romancer.** 2011. pAMT11, a novel plasmid isolated from a *Thermococcus* sp. strain closely related to the virus-like integrated element TKV1 of the *Thermococcus kodakaraensis* genome. Res. Microbiol. **162:**132–143.
93. **Goulet, A., et al.** 2009. *Acidianus* filamentous virus 1 coat proteins display a helical fold spanning the filamentous archaeal viruses lineage. Proc. Natl. Acad. Sci. U. S. A. **106:**21155–21160.
94. **Grahn, A. M., S. J. Butcher, J. K. Bamford, and D. H. Bamford.** 2006. PRD1—dissecting the genome, structure and entry, p. 161–170. *In* R. Calendar (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.
95. **Grahn, A. M., R. Daugelavicius, and D. H. Bamford.** 2002. Sequential model of phage PRD1 DNA delivery: active involvement of the viral membrane. Mol. Microbiol. **46:**1199–1209.
96. **Grahn, A. M., J. Haase, E. Lanka, and D. H. Bamford.** 1997. Assembly of a functional phage PRD1 receptor depends on 11 genes of the IncP plasmid mating pair formation complex. J. Bacteriol. **179:**4733–4740.
97. **Gray, H. B., Jr., W. B. Upholt, and J. Vinograd.** 1971. A buoyant method for the determination of the superhelix density of closed circular DNA. J. Mol. Biol. **62:**1–19.
98. **Happonen, L. J., et al.** 2010. Familial relationships in hyperthermo- and acidophilic archaeal viruses. J. Virol. **84:**4747–4754.
99. **Haring, M., et al.** 2004. Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera *Pyrobaculum* and *Thermoproteus*: a novel virus family, the *Globuloviridae*. Virology **323:**233–242.
100. **Hassan, F., M. Kamruzzaman, J. J. Mekalanos, and S. M. Faruque.** 2010. Satellite phage TLCphi enables toxigenic conversion by CTX phage through dif site alteration. Nature **467:**982–985.
101. **Hatfull, G. F.** 2006. Mycobacteriophages, p. 602–620. *In* R. Calendar (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.
102. **Hatfull, G. F.** 2010. Mycobacteriophages: genes and genomes. Annu. Rev. Microbiol. **64:**331–356.
103. **Hatfull, G. F., et al.** 2010. Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. J. Mol. Biol. **397:**119–143.
104. **Held, N. L., and R. J. Whitaker.** 2009. Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. Environ. Microbiol. **11:**457–466.
105. **Hemminga, M. A., et al.** 2010. Viruses: incredible nanomachines. New advances with filamentous phages. Eur. Biophys. J. **39:**541–550.
106. **Hendrix, R. W.** 2003. Bacteriophage genomics. Curr. Opin. Microbiol. **6:**506–511.
107. **Hendrix, R. W.** 2005. Bacteriophage HK97: assembly of the capsid and evolutionary connections. Adv. Virus Res. **64:**1–14.
108. **Hendrix, R. W.** 2002. Bacteriophages: evolution of the majority. Theor. Popul. Biol. **61:**471–480.
109. **Hendrix, R. W.** 1999. Evolution: the long evolutionary reach of viruses. Curr. Biol. **9:**R914–R917.
110. **Hendrix, R. W.** 2009. Jumbo bacteriophages. Curr. Top. Microbiol. Immunol. **328:**229–240.
111. **Hendrix, R. W., and S. R. Casjens.** 2005. Caudovirales, p. 35–42. *In* C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger, and L. A. Ball (ed.), Virus taxonomy. Eighth report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, CA.

112. **Hendrix, R. W., M. C. Smith, R. N. Burns, M. E. Ford, and G. F. Hatfull.** 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc. Natl. Acad. Sci. U. S. A. **96:**2192–2197.
113. **Hershey, A. D., and M. Chase.** 1952. Independent functions of viral protein and nucleic acid in growth of bacteriophage. J. Gen. Physiol. **36:**39–56.
114. **Hoskisson, P. A., and M. C. Smith.** 2007. Hypervariation and phase variation in the bacteriophage 'resistome'. Curr. Opin. Microbiol. **10:**396–400.
115. **Huber, K. E., and M. K. Waldor.** 2002. Filamentous phage integration requires the host recombinases XerC and XerD. Nature **417:**656–659.
116. **Inokuchi, Y., et al.** 1986. The complete nucleotide sequence of the group II RNA coliphage GA. J. Biochem. **99:**1169–1180.
117. **Jäälinoja, H. T., et al.** 2008. Structure and host-cell interaction of SH1, a membrane-containing, halophilic euryarchaeal virus. Proc. Natl. Acad. Sci. U. S. A. **105:**8008–8013.
118. **Jaatinen, S. T., L. J. Happonen, P. Laurinmaki, S. J. Butcher, and D. H. Bamford.** 2008. Biochemical and structural characterisation of membrane-containing icosahedral dsDNA bacteriophages infecting thermophilic *Thermus thermophilus*. Virology **379:**10–19.
119. **Jalasvuori, M., V. P. Friman, A. Nieminen, J. K. Bamford, and A. Buckling.** 1 June 2011, posting date. Bacteriophage selection against a plasmid-encoded sex apparatus leads to the loss of antibiotic-resistance plasmids. Biol. Lett. [Epub ahead of print.] doi:10.1098/rsbl.2011.0384.
120. **Jalasvuori, M., et al.** 2009. The closest relatives of icosahedral viruses of thermophilic bacteria are among viruses and plasmids of the halophilic archaea. J. Virol. **83:**9388–9397.
121. **Jalasvuori, M., A. Pawlowski, and J. K. Bamford.** 2010. A unique group of virus-related, genome-integrating elements found solely in the bacterial family *Thermaceae* and the archaeal family *Halobacteriaceae*. J. Bacteriol. **192:**3231–3234.
122. **Johnson, J. E.** 2010. Virus particle maturation: insights into elegantly programmed nanomachines. Curr. Opin. Struct. Biol. **20:**210–216.
123. **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.
124. **Kato, J., et al.** 1998. Development of a genetic transformation system for an alga-lysing bacterium. Appl. Environ. Microbiol. **64:**2061–2064.
125. **Kawai, M., I. Uchiyama, and I. Kobayashi.** 2005. Genome comparison *in silico* in *Neisseria* suggests integration of filamentous bacteriophages by their own transposase. DNA Res. **12:**389–401.
126. **Kawasaki, T., et al.** 2007. Genomic characterization of the filamentous integrative bacteriophages ϕRSS1 and ϕRSM1, which infect *Ralstonia solanacearum*. J. Bacteriol. **189:**5792–5802.
127. **Kazaks, A., T. Voronkova, J. Rumnieks, A. Dishlers, and K. Tars.** 2011. Genome structure of *Caulobacter* phage phiCb5. J. Virol. **85:**4628–4631.
128. **Keller, J., et al.** 2007. Crystal structure of AFV3-109, a highly conserved protein from crenarchaeal viruses. Virol. J. **4:**12.
129. **Kelloniemi, J., K. Makinen, and J. P. Valkonen.** 2008. Three heterologous proteins simultaneously expressed from a chimeric potyvirus: infectivity, stability and the correlation of genome and virion lengths. Virus Res. **135:**282–291.
130. **Kessler, A., A. B. Brinkman, J. van der Oost, and D. Prangishvili.** 2004. Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon *Sulfolobus*. J. Bacteriol. **186:**7745–7753.
131. **Khayat, R., et al.** 2005. Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. Proc. Natl. Acad. Sci. U. S. A. **102:**18944–18949.
132. **Kim, A. Y., and H. P. Blaschek.** 1991. Isolation and characterization of a filamentous viruslike particle from *Clostridium acetobutylicum* NCIB 6444. J. Bacteriol. **173:**530–535.
133. **Kim, K. H., et al.** 2008. Amplification of uncultured single-stranded DNA viruses from rice paddy soil. Appl. Environ. Microbiol. **74:**5975–5985.
134. **Kivelä, H. M., S. Madonna, M. Krupovic, M. L. Tutino, and J. K. Bamford.** 2008. Genetics for *Pseudoalteromonas* provides tools to manipulate marine bacterial virus PM2. J. Bacteriol. **190:**1298–1307.
135. **Kivelä, H. M., R. H. Männistö, N. Kalkkinen, and D. H. Bamford.** 1999. Purification and protein composition of PM2, the first lipid-containing bacterial virus to be isolated. Virology **262:**364–374.
136. **Klein, R., et al.** 2002. *Natrialba magadii* virus phiCh1: first complete nucleotide sequence and functional organization of a virus infecting a haloalkaliphilic archaeon. Mol. Microbiol. **45:**851–863.
137. **Klovins, J., G. P. Overbeek, S. H. van den Worm, H. W. Ackermann, and J. van Duin.** 2002. Nucleotide sequence of a ssRNA phage from *Acinetobacter*: kinship to coliphages. J. Gen. Virol. **83:**1523–1533.
138. **Koonin, E. V., T. G. Senkevich, and V. V. Dolja.** 2006. The ancient virus world and evolution of cells. Biol. Direct **1:**29.
139. **Krupovic, M., and D. H. Bamford.** 2008. Archaeal proviruses TKV4 and MVV extend the PRD1-adenovirus lineage to the phylum *Euryarchaeota*. Virology **375:**292–300.
140. **Krupovic, M., and D. H. Bamford.** 2009. Does the evolution of viral polymerases reflect the origin and evolution of viruses? Nat. Rev. Microbiol. **7:**250.
141. **Krupovic, M., and D. H. Bamford.** 2011. Double-stranded DNA viruses: 20 families and only five different architectural principles for virion assembly. Curr. Opin. Virol. **1:**118–124.
142. **Krupovic, M., and D. H. Bamford.** 2010. Order to the viral universe. J. Virol. **84:**12476–12479.
143. **Krupovic, M., and D. H. Bamford.** 2007. Putative prophages related to lytic tailless marine dsDNA phage PM2 are widespread in the genomes of aquatic bacteria. BMC Genomics **8:**236.
144. **Krupovic, M., and D. H. Bamford.** 2011. Revealing virus-host interplay. Science **333:**45–46.
145. **Krupovic, M., and D. H. Bamford.** 2008. Virus evolution: how far does the double beta-barrel viral lineage extend? Nat. Rev. Microbiol. **6:**941–948.
146. **Krupovic, M., V. Cvirkaite-Krupovic, and D. H. Bamford.** 2008. Identification and functional analysis of the *Rz*/*Rz1*-like accessory lysis genes in the membrane-containing bacteriophage PRD1. Mol. Microbiol. **68:**492–503.
147. **Krupovic, M., R. Daugelavicius, and D. H. Bamford.** 2007. A novel lysis system in PM2, a lipid-containing marine double-stranded DNA bacteriophage. Mol. Microbiol. **64:**1635–1648.
148. **Krupovic, M., and P. Forterre.** 2011. *Microviridae* goes temperate: microvirus-related proviruses reside in the genomes of *Bacteroidetes*. PLoS One **6:**e19893.
149. **Krupovic, M., P. Forterre, and D. H. Bamford.** 2010. Comparative analysis of the mosaic genomes of tailed archaeal viruses and proviruses suggests common themes for virion architecture and assembly with tailed viruses of bacteria. J. Mol. Biol. **397:**144–160.
150. **Krupovic, M., S. Gribaldo, D. H. Bamford, and P. Forterre.** 2010. The evolutionary history of archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of mobile genetic elements. Mol. Biol. Evol. **27:**2716–2732.
151. **Krupovic, M., J. J. Ravantti, and D. H. Bamford.** 2009. Geminiviruses: a tale of a plasmid becoming a virus. BMC Evol. Biol. **9:**112.
152. **Krupovic, M., A. Spang, S. Gribaldo, P. Forterre, and C. Schleper.** 2011. A thaumarchaeal provirus testifies for an ancient association of tailed viruses with archaea. Biochem. Soc. Trans. **39:**82–88.
153. **Krupovic, M., et al.** 2006. Genome characterization of lipid-containing marine bacteriophage PM2 by transposon insertion mutagenesis. J. Virol. **80:**9270–9278.
154. **Krupovic, M., M. F. White, P. Forterre, and D. Prangishvili.** Postcards from the edge: structural genomics of archaeal viruses. Adv. Virus Res., in press.
155. **Kwan, T., J. Liu, M. Dubow, P. Gros, and J. Pelletier.** 2006. Comparative genomic analysis of 18 *Pseudomonas aeruginosa* bacteriophages. J. Bacteriol. **188:**1184–1187.
156. **Kwan, T., J. Liu, M. DuBow, P. Gros, and J. Pelletier.** 2005. The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. Proc. Natl. Acad. Sci. U. S. A. **102:**5174–5179.
157. **Laurinmäki, P. A., J. T. Huiskonen, D. H. Bamford, and S. J. Butcher.** 2005. Membrane proteins modulate the bilayer curvature in the bacterial virus Bam35. Structure **13:**1819–1828.
158. **Lawrence, J. G., G. F. Hatfull, and R. W. Hendrix.** 2002. Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. J. Bacteriol. **184:**4891–4905.
159. **Lefkowitz, E. J., C. Wang, and C. Upton.** 2006. Poxviruses: past, present and future. Virus Res. **117:**105–118.
160. **Leplae, R., A. Hebrant, S. J. Wodak, and A. Toussaint.** 2004. ACLAME: a classification of mobile genetic elements. Nucleic Acids Res. **32:**D45–D49.
161. **Licis, N., Z. Balklava, and J. Van Duin.** 2000. Forced retroevolution of an RNA bacteriophage. Virology **271:**298–306.
162. **Lima-Mendez, G., A. Toussaint, and R. Leplae.** 2007. Analysis of the phage sequence space: the benefit of structured information. Virology **365:**241–249.
163. **Lin, N. T., R. Y. Chang, S. J. Lee, and Y. H. Tseng.** 2001. Plasmids carrying cloned fragments of RF DNA from the filamentous phage (phi)Lf can be integrated into the host chromosome via site-specific integration and homologous recombination. Mol. Genet. Genomics **266:**425–435.
164. **Lindell, D., J. D. Jaffe, Z. I. Johnson, G. M. Church, and S. W. Chisholm.** 2005. Photosynthesis genes in marine viruses yield proteins during host infection. Nature **438:**86–89.
165. **Lindell, D., et al.** 2004. Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. Proc. Natl. Acad. Sci. U. S. A. **101:**11013–11018.
166. **Lipps, G.** 2006. Plasmids and viruses of the thermoacidophilic crenarchaeote *Sulfolobus*. Extremophiles **10:**17–28.
167. **Lipps, G.** 2011. Structure and function of the primase domain of the replication protein from the archaeal plasmid pRN1. Biochem. Soc. Trans. **39:**104–106.
168. **López-Bueno, A., et al.** 2009. High diversity of the viral community from an Antarctic lake. Science **326:**858–861.
169. **Luo, Y., P. Pfister, T. Leisinger, and A. Wasserfallen.** 2001. The genome of archaeal prophage psiM100 encodes the lytic enzyme responsible for autolysis of *Methanothermobacter wolfeii*. J. Bacteriol. **183:**5788–5792.
170. **Maaty, W. S., et al.** 2006. Characterization of the archaeal thermophile *Sulfolobus* turreted icosahedral virus validates an evolutionary link among double-stranded DNA viruses from all domains of life. J. Virol. **80:**7625–7635.

171. **Maniloff, J., and K. Dybvig.** 2006. Mycoplasma phages, p. 636–652. *In* R. Calendar (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.
172. **Mann, N. H., A. Cook, A. Millard, S. Bailey, and M. Clokie.** 2003. Marine ecosystems: bacterial photosynthesis genes in a virus. Nature **424:**741.
173. **Männistö, R. H., A. M. Grahn, D. H. Bamford, and J. K. Bamford.** 2003. Transcription of bacteriophage PM2 involves phage-encoded regulators of heterologous origin. J. Bacteriol. **185:**3278–3287.
174. **Männistö, R. H., H. M. Kivelä, L. Paulin, D. H. Bamford, and J. K. Bamford.** 1999. The complete genome sequence of PM2, the first lipid-containing bacterial virus to be isolated. Virology **262:**355–363.
175. **Marvin, D. A.** 1998. Filamentous phage structure, infection and assembly. Curr. Opin. Struct. Biol. **8:**150–158.
176. **McDaniel, L. D., M. delaRosa, and J. H. Paul.** 2006. Temperate and lytic cyanophages from the Gulf of Mexico. J. Mar. Biol. Assoc. U.K. **86:**517–527.
177. **McLeod, S. M., H. H. Kimsey, B. M. Davis, and M. K. Waldor.** 2005. CTXphi and *Vibrio cholerae*: exploring a newly recognized type of phage-host cell relationship. Mol. Microbiol. **57:**347–356.
178. **McLeod, S. M., and M. K. Waldor.** 2004. Characterization of XerC- and XerD-dependent CTX phage integration in *Vibrio cholerae*. Mol. Microbiol. **54:**935–947.
179. **Melcher, U., Y. Sha, F. Ye, and J. Fletcher.** 1999. Mechanisms of spiroplasma genome variation associated with SpV1-like viral DNA inferred from sequence comparisons. Microb. Comp. Genomics **4:**29–46.
180. **Millard, A., M. R. Clokie, D. A. Shub, and N. H. Mann.** 2004. Genetic organization of the *psbAD* region in phages infecting marine *Synechococcus* strains. Proc. Natl. Acad. Sci. U. S. A. **101:**11007–11012.
181. **Millard, A. D., K. Zwirglmaier, M. J. Downey, N. H. Mann, and D. J. Scanlan.** 2009. Comparative genomics of marine cyanomyoviruses reveals the widespread occurrence of *Synechococcus* host genes localized to a hyperplastic region: implications for mechanisms of cyanophage evolution. Environ. Microbiol. **11:**2370–2387.
182. **Mindich, L.** 1996. Heterologous recombination in the segmented dsRNA genome of bacteriophage ϕ6. Semin. Virol. **7:**389–397.
183. **Mindich, L.** 2006. Phages with segmented double-stranded RNA genomes, p. 197–207. *In* R. Calendar (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.
184. **Mochizuki, T., et al.** 2010. Diversity of viruses of the hyperthermophilic archaeal genus *Aeropyrum*, and isolation of the *Aeropyrum pernix* bacilliform virus 1, APBV1, the first representative of the family *Clavaviridae*. Virology **402:**347–354.
185. **Moyer, K. E., H. H. Kimsey, and M. K. Waldor.** 2001. Evidence for a rolling-circle mechanism of phage DNA synthesis from both replicative and integrated forms of CTXphi. Mol. Microbiol. **41:**311–323.
186. **Muskhelishvili, G., P. Palm, and W. Zillig.** 1993. SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. Mol. Gen. Genet. **237:**334–342.
187. **Nolan, J. M., V. Petrov, C. Bertrand, H. M. Krisch, and J. D. Karam.** 2006. Genetic diversity among five T4-like bacteriophages. Virol. J. **3:**30.
188. **Oke, M., et al.** 2010. The Scottish Structural Proteomics Facility: targets, methods and outputs. J. Struct. Funct. Genomics **11:**167–180.
189. **Oke, M., et al.** 2011. A dimeric Rep protein initiates replication of a linear archaeal virus genome: implications for the Rep mechanism and viral replication. J. Virol. **85:**925–931.
190. **O'Keefe, K. J., et al.** 2010. Geographic differences in sexual reassortment in RNA phage. Evolution **64:**3010–3023.
191. **Olsen, R. H., and P. Shipley.** 1973. Host range and properties of the *Pseudomonas aeruginosa* R factor R1822. J. Bacteriol. **113:**772–780.
192. **Olsthoorn, R. C., and J. van Duin.** 1996. Evolutionary reconstruction of a hairpin deleted from the genome of an RNA virus. Proc. Natl. Acad. Sci. U. S. A. **93:**12256–12261.
193. **Onodera, S., Y. Sun, and L. Mindich.** 2001. Reverse genetics and recombination in Phi8, a dsRNA bacteriophage. Virology **286:**113–118.
194. **Osborn, A. M., and D. Boltner.** 2002. When phage, plasmids, and transposons collide: genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. Plasmid **48:**202–212.
195. **Pagaling, E., et al.** 2007. Sequence analysis of an archaeal virus isolated from a hypersaline lake in Inner Mongolia, China. BMC Genomics **8:**410.
196. **Palasingam, K., and P. N. Shaklee.** 1992. Reversion of Q beta RNA phage mutants by homologous RNA recombination. J. Virol. **66:**2435–2442.
197. **Paul, J. H.** 2008. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? ISME J. **2:**579–589.
198. **Pavesi, A.** 2006. Origin and evolution of overlapping genes in the family *Microviridae*. J. Gen. Virol. **87:**1013–1017.
199. **Pedulla, M. L., et al.** 2003. Origins of highly mosaic mycobacteriophage genomes. Cell **113:**171–182.
200. **Peeters, B. P., R. M. Peters, J. G. Schoenmakers, and R. N. Konings.** 1985. Nucleotide sequence and genetic organization of the genome of the N-specific filamentous bacteriophage IKe. Comparison with the genome of the F-specific filamentous phages M13, fd and f1. J. Mol. Biol. **181:**27–39.
201. **Peng, X.** 2008. Evidence for the horizontal transfer of an integrase gene from a fusellovirus to a pRN-like plasmid within a single strain of *Sulfolobus* and the implications for plasmid survival. Microbiology **154:**383–391.
202. **Peng, X., et al.** 2001. Sequences and replication of genomes of the archaeal rudiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses. Virology **291:**226–234.
203. **Pfister, P., A. Wasserfallen, R. Stettler, and T. Leisinger.** 1998. Molecular analysis of Methanobacterium phage psiM2. Mol. Microbiol. **30:**233–244.
204. **Pietilä, M. K., E. Roine, L. Paulin, N. Kalkkinen, and D. H. Bamford.** 2009. An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. Mol. Microbiol. **72:**307–319.
205. **Pina, M., A. Bize, P. Forterre, and D. Prangishvili.** 2011. The archeoviruses. FEMS Microbiol. Rev. **35:**1035–1054.
206. **Poddar, S. K., S. P. Cadden, J. Das, and J. Maniloff.** 1985. Heterogeneous progeny viruses are produced by a budding enveloped phage. Intervirology **23:**208–221.
207. **Pope, W. H., et al.** 2011. Expanding the diversity of mycobacteriophages: insights into genome architecture and evolution. PLoS One **6:**e16329.
208. **Poranen, M. M., and A. Domanska.** 2008. Virus entry to bacterial cells, p. 365–370. *In* B. W. Mahy and M. H. Van Regenmortel (ed.), Encyclopedia of virology, 3rd ed. Academic Press, San Diego, CA.
209. **Poranen, M. M., R. Tuma, and D. H. Bamford.** 2005. Assembly of double-stranded RNA bacteriophages. Adv. Virus Res. **64:**15–43.
210. **Porter, K., and M. L. Dyall-Smith.** 2008. Transfection of haloarchaea by the DNAs of spindle and round haloviruses and the use of transposon mutagenesis to identify non-essential regions. Mol. Microbiol. **70:**1236–1245.
211. **Porter, K., et al.** 2005. SH1: a novel, spherical halovirus isolated from an Australian hypersaline lake. Virology **335:**22–33.
212. **Porter, K., B. E. Russ, and M. L. Dyall-Smith.** 2007. Virus-host interactions in salt lakes. Curr. Opin. Microbiol. **10:**418–424.
213. **Prangishvili, D., P. Forterre, and R. A. Garrett.** 2006. Viruses of the Archaea: a unifying view. Nat. Rev. Microbiol. **4:**837–848.
214. **Prangishvili, D., R. A. Garrett, and E. V. Koonin.** 2006. Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life. Virus Res. **117:**52–67.
215. Reference deleted.
216. **Proux, C., et al.** 2002. The dilemma of phage taxonomy illustrated by comparative genomics of Sfi21-like *Siphoviridae* in lactic acid bacteria. J. Bacteriol. **184:**6026–6036.
217. **Qiao, J., X. Qiao, Y. Sun, and L. Mindich.** 2010. Role of host protein glutaredoxin 3 in the control of transcription during bacteriophage Phi2954 infection. Proc. Natl. Acad. Sci. U. S. A. **107:**6000–6004.
218. **Qiao, X., J. Qiao, and L. Mindich.** 1997. An in vitro system for the investigation of heterologous RNA recombination. Virology **227:**103–110.
219. **Qiao, X., Y. Sun, J. Qiao, F. Di Sanzo, and L. Mindich.** 2010. Characterization of Phi2954, a newly isolated bacteriophage containing three dsRNA genomic segments. BMC Microbiol. **10:**55.
220. **Qiao, X., Y. Sun, J. Qiao, and L. Mindich.** 2008. The role of host protein YajQ in the temporal control of transcription in bacteriophage Phi6. Proc. Natl. Acad. Sci. U. S. A. **105:**15956–15960.
221. **Qiao, X., Y. Sun, J. Qiao, and L. Mindich.** 2009. Temporal control of message stability in the life cycle of double-stranded RNA bacteriophage phi8. J. Virol. **83:**633–639.
222. **Quax, T. E., M. Krupovic, S. Lucas, P. Forterre, and D. Prangishvili.** 2010. The *Sulfolobus* rod-shaped virus 2 encodes a prominent structural component of the unique virion release system in Archaea. Virology **404:**1–4.
223. **Rakonjac, J., N. J. Bennett, J. Spagnuolo, D. Gagic, and M. Russel.** 2011. Filamentous bacteriophage: biology, phage display and nanotechnology applications. Curr. Issues Mol. Biol. **13:**51–76.
224. **Rao, V. B., and M. Feiss.** 2008. The bacteriophage DNA packaging motor. Annu. Rev. Genet. **42:**647–681.
225. **Raoult, D., et al.** 2004. The 1.2-megabase genome sequence of Mimivirus. Science **306:**1344–1350.
226. **Rappé, M. S., and S. J. Giovannoni.** 2003. The uncultured microbial majority. Annu. Rev. Microbiol. **57:**369–394.
227. **Ravantti, J. J., A. Gaidelytė, D. H. Bamford, and J. K. Bamford.** 2003. Comparative analysis of bacterial viruses Bam35, infecting a gram-positive host, and PRD1, infecting gram-negative hosts, demonstrates a viral lineage. Virology **313:**401–414.
228. **Ravin, N. V.** 2011. N15: the linear phage-plasmid. Plasmid **65:**102–109.
229. **Ravin, V., et al.** 2000. Genomic sequence and analysis of the atypical temperate bacteriophage N15. J. Mol. Biol. **299:**53–73.
230. **Read, T. D., et al.** 2003. Genome sequence of *Chlamydophila caviae* (*Chlamydia psittaci* GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae. Nucleic Acids Res. **31:**2134–2147.
231. **Redder, P., et al.** 2009. Four newly isolated fuselloviruses from extreme geothermal environments reveal unusual morphologies and a possible interviral recombination mechanism. Environ. Microbiol. **11:**2849–2862.
232. **Reigstad, L. J., S. L. Jorgensen, and C. Schleper.** 2010. Diversity and abundance of *Korarchaeota* in terrestrial hot springs of Iceland and Kamchatka. ISME J. **4:**346–356.
233. **Renaudin, J., and J. M. Bové.** 1994. SpV1 and SpV4, spiroplasma viruses

with circular, single-stranded DNA genomes, and their contribution to the molecular biology of spiroplasmas. Adv. Virus Res. **44:**429–463.
234. **Rice, G., et al.** 2004. The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. Proc. Natl. Acad. Sci. U. S. A. **101:**7716–7720.
235. **Rohwer, F., and R. V. Thurber.** 2009. Viruses manipulate the marine environment. Nature **459:**207–212.
236. **Roine, E., et al.** 2010. New, closely related haloarchaeal viral elements with different nucleic acid types. J. Virol. **84:**3682–3689.
237. **Rokyta, D. R., C. L. Burch, S. B. Caudle, and H. A. Wichman.** 2006. Horizontal gene transfer and the evolution of microvirid coliphage genomes. J. Bacteriol. **188:**1134–1142.
238. **Rossmann, M. G., and J. E. Johnson.** 1989. Icosahedral RNA virus structure. Annu. Rev. Biochem. **58:**533–573.
239. **Rubin, E. J., W. Lin, J. J. Mekalanos, and M. K. Waldor.** 1998. Replication and integration of a *Vibrio cholerae* cryptic plasmid linked to the CTX prophage. Mol. Microbiol. **28:**1247–1254.
240. **Ruokoranta, T. M., A. M. Grahn, J. J. Ravantti, M. M. Poranen, and D. H. Bamford.** 2006. Complete genome sequence of the broad host range single-stranded RNA phage PRR1 places it in the Levivirus genus with characteristics shared with allolevivirus. J. Virol. **80:**9326–9330.
241. **Russel, M., and P. Model.** 2006. Filamentous phage, p. 146–160. *In* R. Calendar (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.
242. **Sait, M., et al.** 2011. Identification, sequencing and molecular analysis of Chp4, a novel chlamydiaphage of *Chlamydophila abortus* belonging to the *Microviridae*. J. Gen. Virol. **92:**1733–1737.
243. **Salas, M.** 1991. Protein-priming of DNA replication. Annu. Rev. Biochem. **60:**39–71.
244. **Saren, A. M., et al.** 2005. A snapshot of viral evolution from genome analysis of the *Tectiviridae* family. J. Mol. Biol. **350:**427–440.
245. **Scaltriti, E., et al.** 2011. Lactococcal phage p2 ORF35-Sak3 is an ATPase involved in DNA recombination and AbiK mechanism. Mol. Microbiol. **80:**102–116.
246. **Schleper, C., and G. W. Nicol.** 2010. Ammonia-oxidising archaea—physiology, ecology and evolution. Adv. Microb. Physiol. **57:**1–41.
247. **Sha, Y., U. Melcher, R. E. Davis, and J. Fletcher.** 2000. Common elements of Spiroplasma plectroviruses revealed by nucleotide sequence of SVTS2. Virus Genes **20:**47–56.
248. **Sharon, I., et al.** 2009. Photosystem I gene cassettes are present in marine virus genomes. Nature **461:**258–262.
249. **Sharon, I., et al.** 2011. Comparative metagenomics of microbial traits within oceanic viral communities. ISME J. **5:**1178–1190.
250. **She, Q., B. Shen, and L. Chen.** 2004. Archaeal integrases and mechanisms of gene capture. Biochem. Soc. Trans. **32:**222–226.
251. **Silander, O. K., et al.** 2005. Widespread genetic exchange among terrestrial bacteriophages. Proc. Natl. Acad. Sci. U. S. A. **102:**19009–19014.
252. **Sime-Ngando, T., et al.** 2011. Diversity of virus-host systems in hypersaline Lake Retba, Senegal. Environ. Microbiol. **13:**1956–1972.
253. **Soler, N., M. Gaudin, E. Marguet, and P. Forterre.** 2011. Plasmids, viruses and virus-like membrane vesicles from *Thermococcales*. Biochem. Soc. Trans. **39:**36–44.
254. **Soler, N., et al.** 2010. Two novel families of plasmids from hyperthermophilic archaea encoding new families of replication proteins. Nucleic Acids Res. **38:**5088–5104.
255. **Sozhamannan, S., et al.** 2008. Molecular characterization of a variant of *Bacillus anthracis*-specific phage AP50 with improved bacteriolytic activity. Appl. Environ. Microbiol. **74:**6792–6796.
256. **Spang, A., et al.** 2010. Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum *Thaumarchaeota*. Trends Microbiol. **18:**331–340.
257. **Srinivasiah, S., et al.** 2008. Phages across the biosphere: contrasts of viruses in soil and aquatic environments. Res. Microbiol. **159:**349–357.
258. **Stassen, A. P., E. F. Schoenmakers, M. Yu, J. G. Schoenmakers, and R. N. Konings.** 1992. Nucleotide sequence of the genome of the filamentous bacteriophage I2-2: module evolution of the filamentous phage genome. J. Mol. Evol. **34:**141–152.
259. **Steven, A. C., et al.** 2005. Structure, assembly, and antigenicity of hepatitis B virus capsid proteins. Adv. Virus Res. **64:**125–164.
260. **Steven, A. C., J. B. Heymann, N. Cheng, B. L. Trus, and J. F. Conway.** 2005. Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. Curr. Opin. Struct. Biol. **15:**227–236.
261. **Stewart, C. R., et al.** 2009. The genome of *Bacillus subtilis* bacteriophage SPO1. J. Mol. Biol. **388:**48–70.
262. **Strömsten, N. J., D. H. Bamford, and J. K. Bamford.** 2005. In vitro DNA packaging of PRD1: a common mechanism for internal-membrane viruses. J. Mol. Biol. **348:**617–629.
263. **Strömsten, N. J., S. D. Benson, R. M. Burnett, D. H. Bamford, and J. K. Bamford.** 2003. The *Bacillus thuringiensis* linear double-stranded DNA phage Bam35, which is highly similar to the *Bacillus cereus* linear plasmid pBClin15, has a prophage state. J. Bacteriol. **185:**6985–6989.
264. **Sullivan, M. B., et al.** 2010. Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. Environ. Microbiol. **12:**3035–3056.
265. **Sullivan, M. B., et al.** 2006. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. PLoS Biol. **4:**e234.
266. **Sun, Y., X. Qiao, J. Qiao, S. Onodera, and L. Mindich.** 2003. Unique properties of the inner core of bacteriophage phi8, a virus with a segmented dsRNA genome. Virology **308:**354–361.
267. **Suttle, C. A.** 2007. Marine viruses—major players in the global ecosystem. Nat. Rev. Microbiol. **5:**801–812.
268. **Szymczyna, B. R., et al.** 2009. Synergy of NMR, computation, and X-ray crystallography for structural biology. Structure **17:**499–507.
269. **Tadmor, A. D., E. A. Ottesen, J. R. Leadbetter, and R. Phillips.** 2011. Probing individual environmental bacteria for viruses by using microfluidic digital PCR. Science **333:**58–62.
270. **Tang, S. L., S. Nuttall, and M. Dyall-Smith.** 2004. Haloviruses HF1 and HF2: evidence for a recent and large recombination event. J. Bacteriol. **186:**2810–2817.
271. **Toussaint, A., and C. Merlin.** 2002. Mobile elements as a combination of functional modules. Plasmid **47:**26–35.
272. **Tu, A. H., L. L. Voelker, X. Shen, and K. Dybvig.** 2001. Complete nucleotide sequence of the mycoplasma virus P1 genome. Plasmid **45:**122–126.
273. **Tucker, K. P., R. Parsons, E. M. Symonds, and M. Breitbart.** 2011. Diversity and distribution of single-stranded DNA phages in the North Atlantic Ocean. ISME J. **5:**822–830.
274. **Val, M. E., et al.** 2005. The single-stranded genome of phage CTX is the form used for integration into the genome of *Vibrio cholerae*. Mol. Cell **19:**559–566.
275. **Valentine, R. C., P. M. Silverman, K. A. Ippen, and H. Mobach.** 1969. The F-pilus of *Escherichia coli*. Adv. Microb. Physiol. **3:**1–52.
276. **van Duin, J., and N. Tsareva.** 2006. Single-stranded RNA phages, p. 175–196. *In* R. Calendar (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.
277. **Vestergaard, G., et al.** 2008. Structure of the acidianus filamentous virus 3 and comparative genomics of related archaeal lipothrixviruses. J. Virol. **82:**371–381.
278. **Vestergaard, G., et al.** 2008. *Stygiolobus* rod-shaped virus and the interplay of crenarchaeal rudiviruses with the CRISPR antiviral system. J. Bacteriol. **190:**6837–6845.
279. **Waldor, M. K., and J. J. Mekalanos.** 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science **272:**1910–1914.
280. **Waldor, M. K., E. J. Rubin, G. D. Pearson, H. Kimsey, and J. J. Mekalanos.** 1997. Regulation, replication, and integration functions of the *Vibrio cholerae* CTXphi are encoded by region RS2. Mol. Microbiol. **24:**917–926.
281. **Walkinshaw, M. D., et al.** 2002. Structure of Ocr from bacteriophage T7, a protein that mimics B-form DNA. Mol. Cell **9:**187–194.
282. **Wang, C., et al.** 2011. A reverse transcriptase-related protein mediates phage resistance and polymerizes untemplated DNA in vitro. Nucleic Acids Res. **39:**7620–7629.
283. **Wang, Y., et al.** 2007. A novel *Sulfolobus* non-conjugative extrachromosomal genetic element capable of integration into the host genome and spreading in the presence of a fusellovirus. Virology **363:**124–133.
284. **Ward, D. E., et al.** 2002. Characterization of plasmid pRT1 from *Pyrococcus* sp. strain JT1. J. Bacteriol. **184:**2561–2566.
285. **Webb, J. S., M. Lau, and S. Kjelleberg.** 2004. Bacteriophage and phenotypic variation in *Pseudomonas aeruginosa* biofilm development. J. Bacteriol. **186:**8066–8073.
286. **Weigel, C., and H. Seitz.** 2006. Bacteriophage replication modules. FEMS Microbiol. Rev. **30:**321–381.
287. **Weigele, P. R., et al.** 2007. Genomic and structural analysis of Syn9, a cyanophage infecting marine *Prochlorococcus* and *Synechococcus*. Environ. Microbiol. **9:**1675–1695.
288. **Wikoff, W. R., et al.** 2000. Topologically linked protein rings in the bacteriophage HK97 capsid. Science **289:**2129–2133.
289. **Willetts, N., and R. Skurray.** 1980. The conjugation system of F-like plasmids. Annu. Rev. Genet. **14:**41–76.
290. **Wirth, J. F., et al.** 2011. Development of a genetic system for the archaeal virus *Sulfolobus* turreted icosahedral virus (STIV). Virology **415:**6–11.
291. **Woese, C. R., O. Kandler, and M. L. Wheelis.** 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.
292. **Wommack, K. E., and R. R. Colwell.** 2000. Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. **64:**69–114.
293. **Yamada, T., et al.** 2010. A jumbo phage infecting the phytopathogen *Ralstonia solanacearum* defines a new lineage of the *Myoviridae* family. Virology **398:**135–147.
294. **Yu, M. X., M. R. Slater, and H. W. Ackermann.** 2006. Isolation and characterization of *Thermus* bacteriophages. Arch. Virol. **151:**663–679.
295. **Zlotnick, A., et al.** 1998. Shared motifs of the capsid proteins of hepadnaviruses and retroviruses suggest a common evolutionary origin. FEBS Lett. **431:**301–304.