

A new proposed taxon for double-stranded DNA viruses, the order “Ligamenvirales”

David Prangishvili · Mart Krupovic

Received: 20 September 2011 / Accepted: 9 December 2011 / Published online: 22 January 2012
© Springer-Verlag 2012

Abstract Linear viruses with double-stranded DNA genomes are classified into two families, *Lipothrixviridae* and *Rudiviridae*. The members of these two families, all of which infect hyperthermophilic members of the domain Archaea, differ significantly in the complexity of their virions as well as in their mechanisms of genome replication. However, recent structural and genomic studies have revealed a robust evolutionary link between members of the two families. To acknowledge this relationship we propose to unify the two families into the new taxonomic order “Ligamenvirales”.

Known double-stranded DNA viruses infecting hyperthermophilic members of the Archaea, the third domain of life next to the domains Bacteria and Eukarya, form a unique group in the viral world. Due to the unique properties of these viruses, eight novel families have been introduced by the International Committee on Taxonomy of Viruses (ICTV) for their classification: *Lipothrixviridae*, *Rudiviridae*, *Globuloviridae*, *Guttaviridae*, *Fuselloviridae*, *Bicaudaviridae*, *Clavaviridae*, and *Ampullaviridae* (for reviews, see refs. 16, 17). Virion morphologies of members of the latter four families are unique: viruses similar to the spindle-shaped fuselloviruses, droplet-shaped guttaviruses, two-tailed bicaudaviruses or bottle-shaped ampullaviruses have never been observed among bacteria or eukarya. Unique for the Archaea are also linear viruses with double-stranded (ds) DNA genomes, which have been assigned to the families *Lipothrixviridae* and *Rudiviridae*.

The members of the family *Lipothrixviridae* (from the Greek *lipos*, “fat” and *thrix*, “thread”) were the first viruses of hyperthermophilic archaea to be described [10]. Their flexible, linear virions consist of the outer envelope, inner helical core and linear dsDNA genome (Fig. 1A). Following the characterization of lipothrixviruses, an additional group of linear viruses that differed considerably from the members of the family *Lipothrixviridae* in their virion characteristics was discovered. Unlike lipothrixviruses, these viruses did not contain a lipid envelope and were stiff, rather than flexible (Fig. 1B). Due to these differences, the latter group of viruses was classified into the novel family *Rudiviridae* (from the Latin *rudis*, “small rod”) [19]. At present, the family *Lipothrixviridae* is comprised of nine virus species [2, 4, 8, 10, 22], while the family *Rudiviridae* includes 4 virus species [19, 23, 24]. All of the members of these families were isolated from extreme geothermal environments in Europe, North America and Asia. Their known host range is restricted to hyperthermophilic archaea from the genera *Sulfolobus*, *Acidianus*, *Stygiolobus* and *Thermoproteus* (all members of the phylum *Crenarchaeota*).

Besides the apparent differences in virion organization and complexity, members of the two families also differ in the mechanisms they exploit for genome replication. Although genomes of both rudiviruses and lipothrixviruses consist of linear dsDNA molecules, only those of rudiviruses have hairpin ends produced by two terminally linked DNA strands [5]. A rolling hairpin model for the genome replication of rudiviruses was recently proposed based on the structural and biochemical characterization of the REP protein, which is absolutely conserved in all members of the family [13]. The replication appears to be initiated by introducing a nick on one of the genome strands, 11

D. Prangishvili (✉) · M. Krupovic
Department of Microbiology, Institut Pasteur,
25 rue du Dr. Roux, 75015 Paris, France
e-mail: david.prangishvili@pasteur.fr

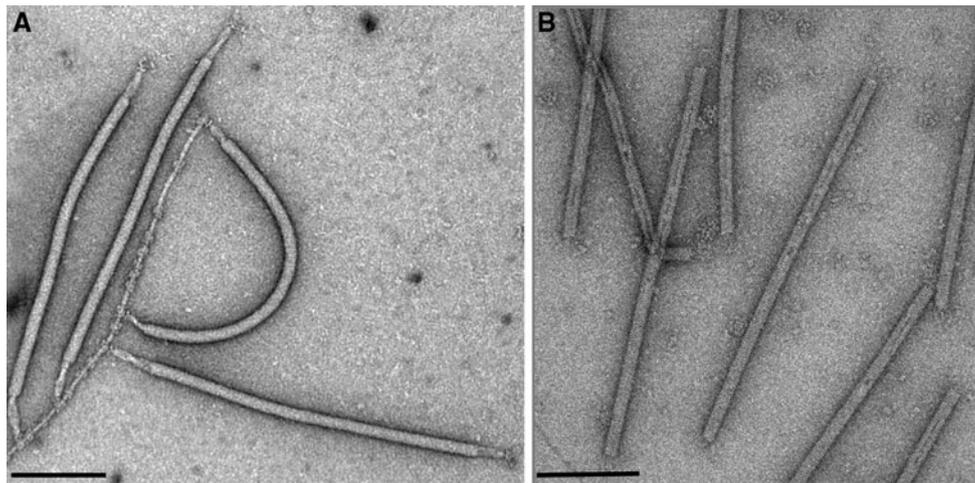


Fig. 1 Negative-contrast electron micrographs of the representatives of the families *Lipothrixviridae* and *Rudiviridae*. (A) Acidianus filamentous virus 1, AFV1; (B) Sulfolobus islandicus rod-shaped virus 2, SIRV2. Scale bars, 200 nm

nucleotides away from the terminus. A protein implicated in this process displays remarkable structural similarity to the rolling-circle replication initiation proteins and was shown to introduce a nick at the predicted replication initiation site, forming a covalent adduct with the newly created 5' end, with the concomitant release of the 3'-hydroxyl terminus, which is suitable for priming DNA replication [13]. The replication proceeds by strand displacement and results in the generation of head-to-head concatemers [15]. A four-way junction between the concatemers is likely to be subsequently resolved by the viral Holliday-junction-resolving enzyme (Hjr) encoded by all members of the family [3, 6]. The two key enzymes in rudiviral genome replication, REP and Hjr, do not have

homologues in any of the isolated lipothrixviruses. Moreover, the genome topology of viruses in the family *Lipothrixviridae*, with a protein covalently attached to the termini of their linear genomes [16], apparently excludes the rolling hairpin model. Unusual traits of genome replication of the lipothrixvirus AFV1 strongly suggest a novel replication mechanism, distinct from that operating in rudiviruses [16].

Despite the pronounced differences in virion morphology and genome replication mechanisms, accumulation and thorough comparative analysis of complete genome sequences from both rudiviruses and lipothrixviruses has revealed an unexpected evolutionary link between the two families [12, 17, 18]. It has become apparent that, at the

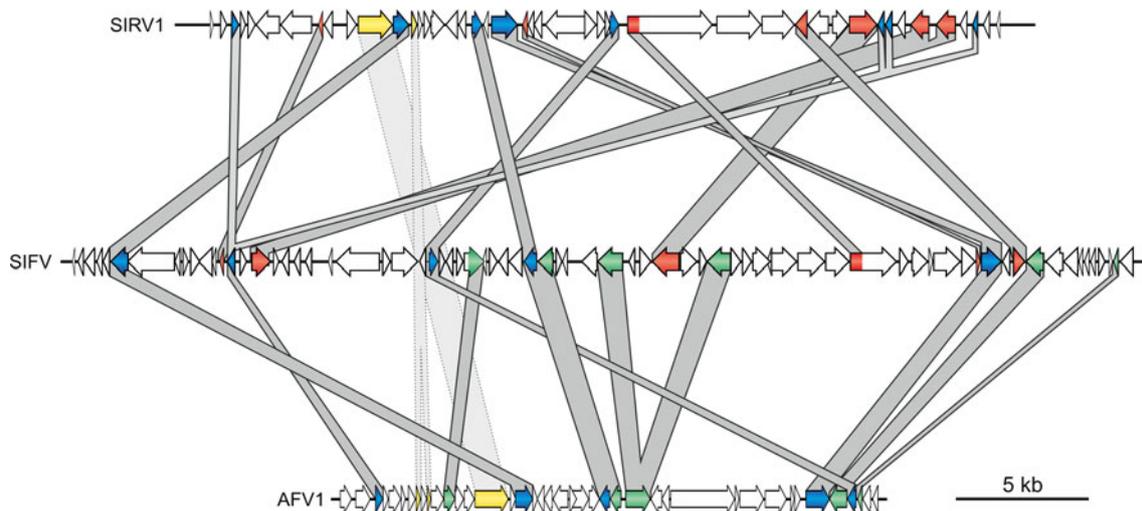


Fig. 2 Genomic relationship between linear archaeal viruses of the families *Rudiviridae* and *Lipothrixviridae*. Genes shared by Sulfolobus islandicus rod-shaped virus 1 (SIRV1; *Rudiviridae*) and lipothrixviruses Sulfolobus islandicus filamentous virus (SIFV) and

Acidianus filamentous virus 1 (AFV1) are shaded blue. Genes restricted to virus pairs SIRV1-SIFV, SIRV1-AFV1 and SIFV-AFV1 are shown in red, yellow and green, respectively. The figure is adapted from ref. [12] (color figure online)

genome level, some lipothrixviruses are no more similar to other members of the family *Lipothrixviridae* than they are to the rudiviruses (Fig. 2). An extreme example is provided by the lipothrixvirus SIFV, which shares the same number of homologous genes (ten) with the lipothrixvirus AFV1 and the rudivirus SIRV1. The two sets of genes (SIFV–AFV1 and SIFV–SIRV1) are not identical, but do overlap (Fig. 2; [12]). The overlapping set of genes includes genes encoding glycosyl transferases, transcription factors and small genes of unknown function. Such a relatively close genetic relationship between viruses belonging to different families is unusual in the archaeal virosphere, where viruses from different families typically share only a few (if any) homologous genes [18]. Rudiviruses and lipothrixviruses do not encode recognizable integrases and are not known to lysogenize their hosts, a phenomenon which could potentially favour gene shuffling between unrelated viruses [12]. Consequently, there is no reason to

believe that horizontal gene exchange, which could explain the common gene content of rudiviruses and lipothrixviruses, is more vigorous between these linear viruses than between other hyperthermophilic archaeal viruses. We therefore conclude that the genomic relationship between the two groups of viruses most likely reflects their common ancestry.

Even more convincing evidence for the evolutionary link between rudiviruses and lipothrixviruses has come from structural analysis of their major capsid proteins (MCPs), which turned out to be highly similar despite their very low sequence similarity (Fig. 3A). The filamentous, enveloped virions of lipothrixviruses are composed of two MCPs, MCP1 and MCP2. The structures of both proteins from the virus AFV1 have been determined by X-ray crystallography [7], revealing the same unique four-helix-bundle fold at the C-termini of both proteins (Fig. 3B). However, due to the differences in their N-terminal

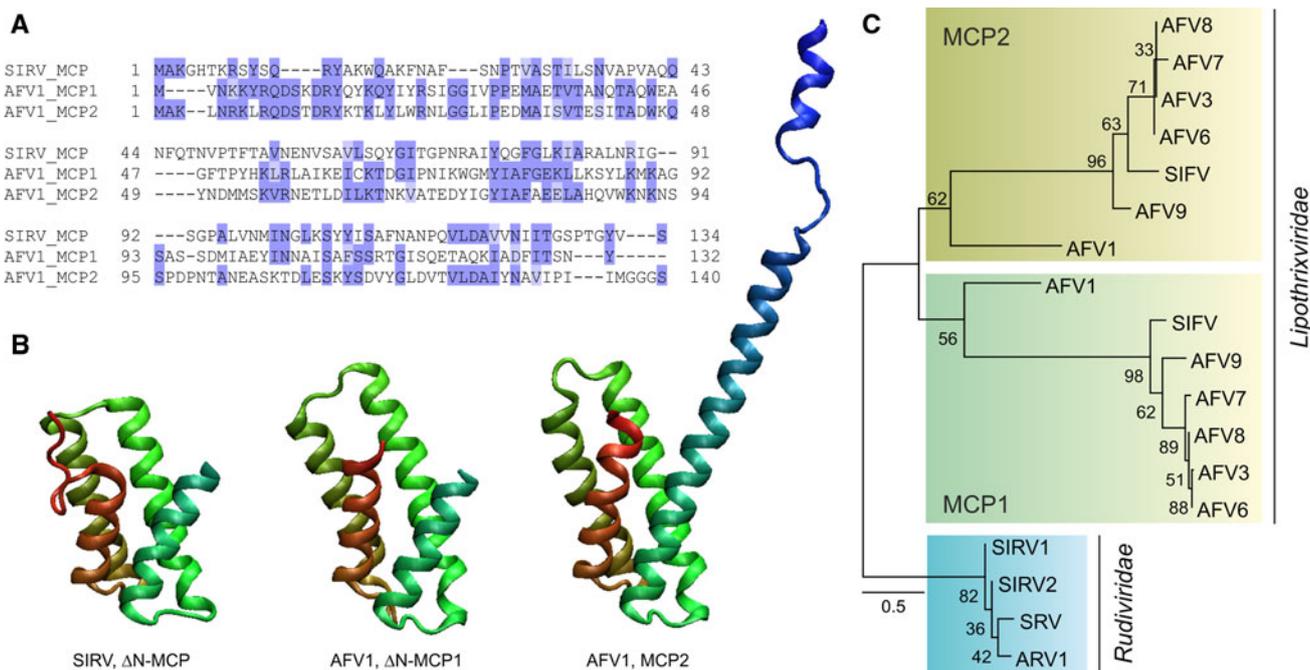


Fig. 3 Major capsid proteins (MCP) of linear dsDNA viruses infecting archaea. **A**. Sequence alignment of the MCP of the rudivirus SIRV (ORF134) with the two MCPs of the lipothrixvirus AFV1. The alignment is coloured according to sequence conservation (BLOSUM62 matrix). **B**. Comparison of the N-terminally truncated MCP of SIRV with the two MCPs of AFV1. The structures are colored using a rainbow colour gradient from the N-terminus (blue) to the C-terminus (red). **C**. Maximum-likelihood phylogeny of the MCPs of rudiviruses and lipothrixviruses. For phylogenetic analysis, homologous MCP sequences were collected using PSI-BLAST [1] and aligned using PROMALS3D [14]. MCP sequences of lipothrixviruses AFV2 and TTV1 could not be retrieved using sequence-similarity-based approaches and are therefore not included in this analysis. A maximum-likelihood phylogeny was constructed in MEGA5 [21] using the JTT matrix model (+I, +G [4

categories]). All positions containing gaps were eliminated, and the final dataset consisted of 120 positions. Numbers at the branch points represent bootstrap values (100 replicates). The scale bar represents the number of substitutions per site. The tree was rooted on the branch between rudiviral and lipothrixviral MCPs. GenBank accession numbers: *Rudiviridae* MCPs (SIRV1 ORF134, NP_666607; SIRV2 ORF134, NP_666560; SRV ORF134, CAQ58456; ARV gp24, YP_001542641), *Lipothrixviridae* MCP1 (AFV1 ORF132, YP_003749; AFV3 gp34, YP_001604376; AFV6 gp35, YP_001604193; AFV7 gp28, YP_001604252; AFV8 gp30, YP_001604311; AFV9 gp32, YP_001798550; SIFV ORF35, NP_445700) and MCP2 (AFV1 ORF140, YP_003750; AFV3 gp35, YP_001604377; AFV6 gp36, YP_001604194; AFV7 gp29, YP_001604253; AFV8 gp31, YP_001604312; AFV9 gp33, YP_001798551; SIFV ORF36, NP_445701) (color figure online)

regions, the two AFV1 MCPs display a distinct hydrophobicity profile, which allowed the topological model to be proposed. According to this model, the basic MCP1 forms a core around which the genomic dsDNA is wrapped, whereas the MCP2 interacts with the genome with its basic N-terminal region, and the hydrophilic C-terminal domain is embedded in the lipid envelope [7]. The rod-shaped, non-enveloped virions of rudiviruses are composed of one major protein: the highly basic MCP accounts for 99% of virion proteins and associates with the genomic DNA to form the helical body of the virion. Strikingly, the N-terminally truncated MCP from the rudivirus SIRV was found to have the same four-helix bundle topology as the MCPs of the lipothrixvirus AFV1 [20] (Fig. 3). This fact is remarkable considering that there is only 17% sequence identity between the MCP1 of AFV1 and MCP of SIRV (Fig. 3). In maximum-likelihood phylogenetic reconstruction, the MCPs of rudiviruses and lipothrixviruses are robustly segregated into separate clusters (Fig. 3C), suggesting that these MCP genes have evolved within their respective viral genomes for an appreciable period of time without being transferred between members of the two families.

Phylogenetic analysis of the MCP sequences supports the current classification of flexible enveloped and rigid non-enveloped linear viruses of archaea into two distinct families. However, the structural similarity between the MCPs of rudiviruses and lipothrixviruses (Fig. 3B) unequivocally points towards the common ancestry of these viruses. Indeed, based on the structural data, it has been envisioned that lipothrixviruses might have evolved from a “simpler” non-enveloped rudivirus-like ancestor [7]. One can suggest a sequence of evolutionary events in which the gene for the single MCP of such putative ancestor has been duplicated and evolved so as to facilitate interactions with a hydrophobic envelope, producing the more complex, lipothrixvirus-like virions.

Importantly, the four-helix bundle topology of the rudiviral and lipothrixviral MCPs has not been observed previously in the MCPs of other known dsDNA viruses [11]. Indeed, when rudiviral and lipothrixviral MCPs were compared to all currently available protein structures using the DALI server [9], the AFV1 MCP2 and SIRV MCP were reciprocally found to be the closest structural relatives, with highly significant Z scores (13.3 and 13.4, respectively). No structural virion proteins from other viruses were identified as similar. Notably, the arrangement of the four helices in the MCPs of rudiviruses and lipothrixviruses is different from that found in the capsid protein of the ssRNA-genome-containing tobacco mosaic virus [7] (Fig. 4), suggesting an independent origin for these archaeal and plant virus capsid proteins. Consequently, structural information not only illuminates the

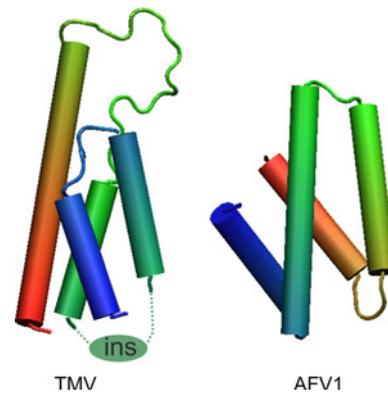


Fig. 4 Different arrangement of α -helices in the four-helix-bundle major capsid proteins of tobacco mosaic virus (TMV; PDB ID:1EI7) and *Acidianus filamentous virus 1* (AFV1; PDB ID:3FBL). The structures are colored using a rainbow color gradient from the N-terminus (blue) to the C-terminus (red). The insertion (ins) between the second and the third α -helices in the TMV protein was omitted for more convenient comparison. The figure is modified from ref. [11] (color figure online)

relationship between the two groups of linear archaeal viruses but also highlights their distinctiveness from any other known virus group.

In summary, recent structural and genomic studies have led to the accumulation of compelling evidence that points to a common ancestry of rudiviruses and lipothrixviruses. Accordingly, in order to acknowledge the evolutionary relationship between linear viruses of the two families, we propose to unify them in the new taxonomic order “Ligamenvirales” (from the Latin *ligamen*, for “string”, “thread”). The proposal is under consideration by the Executive Committee of the International Committee on Taxonomy of Viruses.

References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Bettstetter M, Peng X, Garrett RA, Prangishvili D (2003) AFV1, a novel virus infecting hyperthermophilic archaea of the genus *Acidianus*. *Virology* 315:68–79
- Birkenbihl RP, Neef K, Prangishvili D, Kemper B (2001) Holiday junction resolving enzymes of archaeal viruses SIRV1 and SIRV2. *J Mol Biol* 309:1067–1076
- Bize A, Peng X, Prokofeva M, McLellan K, Lucas S, Forterre P, Garrett RA, Bonch-Osmolovskaya EA, Prangishvili D (2008) Viruses in acidic geothermal environments of the Kamchatka Peninsula. *Res Microbiol* 159:358–366
- Blum H, Zillig W, Mallok S, Domdey H, Prangishvili D (2001) The genome of the archaeal virus SIRV1 has features in common with genomes of eukaryal viruses. *Virology* 281:6–9
- Gardner AF, Guan C, Jack WE (2011) Biochemical characterization of a structure-specific resolving enzyme from *Sulfolobus islandicus* rod-shaped virus 2. *PLoS one* 6:e23668

7. Goulet A, Blangy S, Redder P, Prangishvili D, Felisberto-Rodrigues C, Forterre P, Campanacci V, Cambillau C (2009) *Acidianus* filamentous virus 1 coat proteins display a helical fold spanning the filamentous archaeal viruses lineage. *Proc Natl Acad Sci USA* 106:21155–21160
8. Haring M, Vestergaard G, Brügger K, Rachel R, Garrett RA, Prangishvili D (2005) Structure and genome organization of AFV2, a novel archaeal lipothrixvirus with unusual terminal and core structures. *J Bacteriol* 187:3855–3858
9. Holm L, Kääriäinen S, Rosenström P, Schenkel A (2008) Searching protein structure databases with DaliLite v. 3. *Bioinformatics* 24:2780–2781
10. Janekovic S, Wunder I, Holz W, Zillig W, Gierl A, Neumann H (1983) TTV1, TTV2 and TTV3, a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaeobacterium *Thermoproteus tenax*. *Mol Gen Genet* 192:39–45
11. Krupovic M, Bamford DH (2011) Double-stranded DNA viruses: 20 families and only five different architectural principles for virion assembly. *Curr Opin Virol* 1(2):118–124
12. Krupovic M, Prangishvili D, Hendrix RW, Bamford DH (2011) Genomics of bacterial and archaeal viruses: dynamics within the prokaryotic virosphere. *Microbiol Mol Biol Rev* 75:610–635
13. Oke M, Kerou M, Liu H, Peng X, Garrett RA, Prangishvili D, Naismith JH, White MF (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
14. Pei J, Tang M, Grishin NV (2008) PROMALS3D web server for accurate multiple protein sequence and structure alignments. *Nucleic Acids Res* 36:W30–W34
15. Peng X, Blum H, She Q, Mallok S, Brügger K, Garrett RA, Zillig W, Prangishvili D (2001) Sequences and replication of genomes of the archaeal ruidiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses. *Virology* 291:226–234
16. Pina M, Bize A, Forterre P, Prangishvili D (2011) Archeoviruses. *FEMS Microbiol Rev* 35:1035–1054
17. Prangishvili D, Forterre P, Garrett RA (2006) Viruses of the Archaea: a unifying view. *Nat Rev Microbiol* 4:837–848
18. Prangishvili D, Garrett RA, Koonin EV (2006) Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life. *Virus Res* 117:52–67
19. Prangishvili D, Arnold HP, Gotz D, Ziese U, Holz I, Kristjansson JK, Zillig W (1999) A novel virus family, the Ruidiviridae: Structure, virus-host interactions and genome variability of the sulfobolus viruses SIRV1 and SIRV2. *Genetics* 152:1387–1396
20. Szymczyna BR, Taurog RE, Young MJ, Snyder JC, Johnson JE, Williamson JR (2009) Synergy of NMR, computation, and X-ray crystallography for structural biology. *Structure* 17:499–507
21. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
22. Vestergaard G, Aramayo R, Basta T et al (2008) Structure of the *Acidianus* filamentous virus 3 and comparative genomics of related archaeal lipothrixviruses. *J Virol* 82:371–381
23. Vestergaard G, Haring M, Peng X, Rachel R, Garrett RA, Prangishvili D (2005) A novel ruidivirus, ARV1, of the hyperthermophilic archaeal genus *Acidianus*. *Virology* 336:83–92
24. Vestergaard G, Shah SA, Bize A, Reitberger W, Reuter M, Phan H, Briegel A, Rachel R, Garrett RA, Prangishvili D (2008) *Stygiolobus* rod-shaped virus and the interplay of crenarchaeal ruidiviruses with the CRISPR antiviral system. *J Bacteriol* 190:6837–6845