Viruses of the Archaea: a unifying view

David Prangishvili, Patrick Forterre* and Roger A. Garrett‡*

DNA viruses of the Archaea have highly diverse and often exceptionally complex morphotypes. Many have been isolated from geothermally heated hot environments, raising intriguing questions about their origins, and contradicting the widespread notion of limited biodiversity in extreme environments. Here, we provide a unifying view on archaeal viruses, and present them as a particular assemblage that is fundamentally different in morphotype and genome from the DNA viruses of the other two domains of life, the Bacteria and Eukarya.

Fusiform
An organism that is spindle-shaped: wider in the middle and tapering towards the ends.

Hyperthermophile
An organism that has an optimal growth temperature above 80°C.

Extreme halophile
An organism that requires extremely high levels of sodium chloride for growth.

**Molecular Biology of the Gene in Extremophiles Unit, Institut Pasteur, rue du Docteur Roux 25, F-75724 Paris Cedex 15, France.*
‡Danish Archaea Centre, Institute of Molecular Biology, Copenhagen University, Sølvgade 83H, DK-1307 Copenhagen K, Denmark.
Correspondence to D.P.
e-mail: prangish@pasteur.fr
doi:10.1038/nrmicro1527

The discovery of the Archaea was a significant breakthrough in the recent history of biology. Whereas cell-ultrastructure studies had initially suggested a division of living organisms into eukaryotes and prokaryotes, molecular sequence analyses — pioneered by Carl Woese in the 1970s — revealed the existence of three different classes of ribosomal RNAs and ribosomes in cellular organisms. This discovery led to the replacement of the prokaryote/eukaryote dichotomy by a trinity of domains, the Archaea, Bacteria and Eukarya[1]. Subsequently, ribosomal RNA sequence comparisons led to the division of the archaeal domain into two main kingdoms, the Crenarchaeota and the Euryarchaeota[2].

Over the past three decades we have accrued a broad knowledge of the biological diversity of the Archaea. This includes an outline of their physiology, biochemistry and molecular biology, and many insights into their evolutionary relationships with the Bacteria and the Eukarya[3]. Although Archaea resemble Bacteria in their cellular ultrastructure and genome organization, their DNA replication, transcription and translation machineries show many similarities to their eukaryotic counterparts. In addition, other features seem to have either arisen, or have been exclusively conserved, within the archaeal domain (for example, ether-linked membranes). In this context, archaeal viruses are particularly interesting.

The first two archaeal viruses that were isolated visually resembled bacteriophage T4 and other members of the family *Myoviridae*, with icosahedral heads, contractile helical tails and linear, double-stranded (ds) DNA genomes[4,5]. Subsequently, a few head-tail archaeal viruses were reported with non-contractile tails similar to lambdoid bacteriophages of the family *Siphoviridae*[6,7]. Based on these initial studies, it was inferred, albeit erroneously, that archaeal viruses constituted a variety of the ubiquitous head-tail bacteriophages.

Recently however, this view has changed radically. Electron-microscopy investigations of samples collected from natural environments that contain predominantly archaea, and the enrichment cultures derived therefrom, revealed that the head-tail phenotype is rare among the archaeal viruses (for reviews, see REFS 8,9). In fact, cultured archaeal viruses, which to date all have dsDNA genomes, exhibit a range of virion morphotypes, most of which have not been observed before for any dsDNA virus. There are exceptional forms, including fusiforms, droplet and bottle shapes, and linear and spherical virions, with more complex virions combining features of these different forms. Moreover, genome-sequence analyses have demonstrated that most of the archaeal viruses are unrelated to other known viruses and suggest that they might have different, and possibly multiple, evolutionary origins[10].

Below we present the archaeal viruses, grouped according to the gross features of their virion morphotypes, and describe their genomic properties and relationships with the host cell. Moreover, the evolutionary relationships of the archaeal viruses to those of the Bacteria and Eukarya are discussed. The taxonomic assignments used have either been approved, or are pending, at the International Committee on Taxonomy of Viruses (ICTV).

Fusiform viruses

Viruses with fusiform virions, single or two-tailed, are common in, and exclusive to, the Archaea[11–22]. They constitute a large fraction of the known archaeal viruses and are associated with a broad range of hosts representing the three main phenotypes of cultured archaeal species, the hyperthermophiles, extreme halophiles and anaerobic methane-producers from the Euryarchaeota and Crenarchaeota kingdoms (TABLE 1). Moreover, electron-microscopy analysis revealed an abundance of fusiform-virus-like particles in habitats where these archaea predominate, for example in hot, acidic springs[23–25] and in hypersaline waters[26,27].

The isolated fusiform viruses are diverse in their structural and genomic characteristics, and they

Positively supercoiled DNA
A DNA molecule in which the number of topological links between the two strands is superior to the number of turns.

Lysogen
A bacterium or archaeon that contains a viral genome integrated into the chromosome.

have been classified into the *Fuselloviridae* family, the proposed '*Bicaudaviridae*' family and the genus *Salterprovirus*, although some remain unclassified. Most have a circular genome and carry an integrase gene that can facilitate integration into host chromosomes[28]. The exceptions are the haloarchaeal salterproviruses *Haloarcula hispanica* virus 1 (His1) and *Haloarcula hispanica* virus (His2), which have linear genomes lacking integrase genes[11].

The* Fuselloviridae *family. Known members of the *Fuselloviridae* infect the hyperthermophilic crenarchaeon *Sulfolobus* (TABLE 1), and their replication and release leave the host cell intact. The virions, which are in the size range 55–60 × 80–100 nm, have short tails which are uniform in size and carry thin terminal fibres[15,16,18–20] (FIG. 1a). The best studied is the type species of the family, the *Sulfolobus* spindle-shaped virus 1 (SSV1). Its circular genome is positively supercoiled[29] and during infection it integrates into a tRNA gene in the host chromosome, producing a partitioned integrase gene while the tRNA gene remains intact[28]. UV irradiation or mytomycin treatment induce viral replication and temporarily inhibit the growth of lysogens without causing their lysis[15].

The transcription pattern of the SSV1 genome is relatively simple, with most transcripts being produced constitutively (for a review, see REF. 30). However, following UV irradiation, upregulation of some constitutive transcripts is observed, together with the appearance of a short RNA molecule that might facilitate the initiation of DNA replication.

The* 'Bicaudaviridae' *family. The sole member of this proposed family, the *Acidianus* two-tailed virus (ATV)[14,17], is also the only known virus of the acidophilic, hyperthermophilic archaea that is capable of host lysis. Its reproductive cycle has some unique features[14,17]. Virions are extruded from host cells as tail-less, fusiform particles, which then develop long tails at each pointed end at temperatures above 75°C, close to the temperature of the natural habitat of the host (FIG. 1b). This major, extracellular morphological development is independent of the host cells or any energy sources, and its molecular mechanism remains unclear. The tails consist of tubes, which terminate in an anchor-like structure, and contain a periodic filamentous structure (FIG. 2). One function of the elongated, flexible tails might be to enhance the probability of virion adsorption to a new host cell.

There is circumstantial evidence that the newly discovered process of extracellular tail development of ATV might be shared by other fusiform viruses of the hyperthermophilic archaea. In growth cultures of uncharacterized *Acidianus* species, fusiform particles have often been observed with one or two tails, which can differ in length[23]. This could reflect different stages of extracellular morphogenesis.

The* Salterprovirus *genus. The two distantly related viruses His1 and His2, which have been assigned to this genus, infect strains of the extremely halophilic genus *Haloarcula* of the Euryarchaeota[11,12]. Both viruses are lytic and the linear genomes show no sequence similarity to the circular genomes of the other fusiform viruses. The virions of these viruses are similar in size (44 × 67–77 nm) but, despite pronounced morphological similarities, their major structural proteins are not orthologous. Each of the viruses encodes a DNA polymerase which might be primed by proteins attached to the termini of their linear genomes.

Table 1 | **Archaeal viruses with exceptional morphologies**

Family and genus	Species	Archaeal kingdom	Genome details*	References
Fusiform viruses				
Fuselloviridae, Fusellovirus	*Sulfolobus* spindle-shaped virus 1 (SSV1)	Cr	c, 15.5	15, 16, 18
	Sulfolobus spindle-shaped virus 2 (SSV2)	Cr	c, 14.8	19
	Sulfolobus spindle-shaped virus - Yellowstone 1 (SSV-Y1)	Cr	c, 16.5	20
	Sulfolobus spindle-shaped virus - Kamchatka 1 (SSV-K1)	Cr	c, 17.4	20
Fuselloviridae, Salterprovirus	*Haloarcula hispanica* virus 1 (His1)	Eu	ln, 14.5	11, 12
	Haloarcula hispanica virus 2 (His2)	Eu	ln, 16.1	11, 12
'*Bicaudaviridae, Bicaudavirus*'‡	*Acidianus* two-tailed virus (ATV)	Cr	c, 62.7	14, 17
Unclassified	*Sulfolobus tengchongensis* spindle-shaped virus 1 (STSV1)	Cr	c, 75.3	22
	Pyrococcus abyssi virus 1 (PAV1)	Eu	c, 18.0	13
	Methanococcus voltae virus-like particle (?)	Eu	c, 23.0	21
Bottle- and droplet-shaped viruses				
'*Ampullaviridae, Ampullavirus*'‡	*Acidianus* bottle-shaped virus (ABV)	Cr	ln, 23.9	31
Guttaviridae, Guttavirus	*Sulfolobus neozealandicus* droplet-shaped virus (SNDV)	Cr	c, 20.0	32

*Genome details shown are the form of the genome and the size in kb. ‡Taxonomic proposal pending at the International Committee on Taxonomy of Viruses. c, covalently closed circular; Cr, Crenarchaeota; Eu, Euryarchaeota; ln, linear.

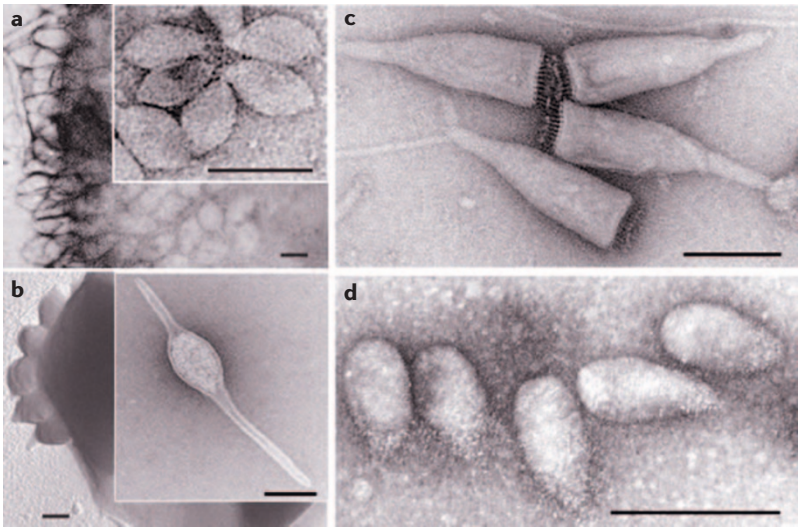


Figure 1 | **Electron micrographs of archaeal viruses with exceptional morphologies. a** | *Sulfolobus* spindle-shaped virus 1 (SSV1) (inset) and its extrusion from the host cell. **b** | The extracellularly developed *Acidianus* two tailed virus (ATV) (inset) and its extrusion from the host cell. **c** | *Acidianus* bottle-shaped virus (ABV). **d** | *Sulfolobus neozealandicus* droplet-shaped virus (SNDV). All images are negatively stained with uranyl acetate, except for part **b**, which was platinum-shadowed. Scale bars represent 100 nm. Parts **a** and **d** are courtesy of W. Zillig. Part **b** is reproduced from *Nature* REF. 14 © (2005) Macmillan Publishers Ltd. Part **c** is reproduced with permission from REF. 31 © (2005) American Society for Microbiology.

Unclassified fusiform viruses. The virion (107 × 230 nm) of the still unclassified *Sulfolobus tengchongensis* spindle-shaped virus 1 (STSV1) of the hyperthermophilic crenarchaeon *Sulfolobus* is the largest of the known fusiform viruses and has a short tail, the length of which varies in the range 0–133 nm[22]. Although variable tail length could result from the extracellular tail-development process described above for ATV, the genome shows minimal similarity in gene content to either ATV or other fusiform viruses. Exceptionally, it carries open reading frames (ORFs) encoding gene products implicated in DNA modification and has a complex pattern of DNA methylation. A putative origin of replication has been identified, which carries multiple AT-rich repeats.

Fusiform virus-like particles with short tails, named *Pyrococcus abyssi* virus 1 (PAV1), have also been reported to be produced by a member of the Euryarchaeota, the hyperthermophile '*Pyrococcus abyssi*'[13]. Although the putative virus resembles crenarchaeal fuselloviruses in size, morphology and genome structure, none of the annotated encoded proteins exhibit sequence similarity to proteins encoded by SSV1 and its relatives (C. Geslin, personal communication). Another fusiform virus of the Euryarchaeota might be represented by pleomorphic particles produced by the methanogen *Methanococcus voltae*[21], containing circular dsDNA which is also found integrated in the host chromosome. However, owing to the distortion of particles by purification procedures, the report of the morphotype of these particles is controversial and it is difficult to distinguish whether they are lemon-shaped or oblate.

Bottle-shaped and droplet-shaped viruses

The virions of two other archaeal viruses, the *Acidianus* bottle-shaped virus (ABV) and the *Sulfolobus neozealandicus* droplet-shaped virus (SNDV), have morphological features that are so unique that each of the viruses has been assigned to a new family (FIG. 1 c,d; TABLE 1).

The ABV virion has a complex form resembling a bottle (230-nm long, 4–75-nm wide). The virus infects the hyperthermophilic *Acidianus* genus and has been assigned to the new family '*Ampullaviridae*'[31]. The virion has no elements of icosahedral or helical symmetry (FIG. 1c) and differs in its basic architecture from any known virus. The envelope encases a cone-shaped core formed by a torroidally supercoiled nucleoprotein filament. A disc is present at the broader end, to which 20 (±2) short, thick filaments are attached (FIG. 1c). Their function remains unclear, while the virion seems to adsorb to the host cell through its narrow end. As for the genomes of salterproviruses, the ABV genome seems to be replicated by a virus-encoded DNA polymerase that is primed by a protein attached to the genomic termini. The virus genome also encodes a putative RNA molecule with notable secondary structural similarity to the RNA molecule that has been implicated in DNA packaging of the bacteriophage ϕ29 and its relatives (X. Peng, personal communication).

The virus SNDV of the hyperthermophilic genus *Sulfolobus* exhibits a complex droplet-shaped virion (90 × 180 nm) and is the sole member of the family *Guttaviridae*[32]. The droplet carries multiple long, thin fibres that are attached at its apex, and the surface seems to be helically ribbed (FIG. 1d). The circular dsDNA genome is extensively modified, probably by methylation, and has not been sequenced.

Linear viruses

Linear particles are the main virion type found in terrestrial hot environments (>80°C) where crenarchaea of the genera *Sulfolobus*, *Acidianus* and *Thermoproteus* predominate[23–25,31]. All linear viruses isolated from these environments to date infect members of these genera. They have dsDNA genomes, a property not previously observed for any linear virus, and therefore they have been classified into two new families: the stiff, rod-like *Rudiviridae*[33–35], and the flexible, filamentous *Lipothrixviridae*[36–40]. Originally, the discrimination of the families was based mainly on differences in virion structure and this was later supported by comparative genomics. Nevertheless, in contrast to members of the other archaeal viral families for which genome sequences are available, a substantial fraction of orthologous genes, including some encoding glycosyl transferases and transcriptional regulators, are shared by the rudiviruses and lipothrixviruses[10]. For example, of the 45 predicted genes of the rudivirus *Sulfolobus islandicus* rod-shaped virus 1 (SIRV1), nine share orthologues with the lipothrixvirus *Sulfolobus islandicus* filamentous virus (SIFV). These observations indicate that it is possible that the known linear archaeal viruses share a relatively recent common ancestor[10]. Therefore, we propose to classify the families

Rudiviridae and *Lipothrixviridae* into a new viral order, the '*Ligamenvirales*' (from the Latin *ligamen*, for string, thread).

Members of the '*Ligamenvirales*' have linear dsDNA genomes and show similarities in their relationships with their host cells. In contrast to archaeal viruses with circular genomes, no integrated viral genomes, or fragments thereof, have been detected in host chromosomes. Consistent with this observation, genomes of the ligamenviruses lack an integrase gene. It seems that these viruses persist stably in host cells and that their replication is not induced by stress factors such as UV irradiation or mytomycin C treatment. The only known exception is the lytic lipothrixvirus *Thermoproteus tenax* virus 1 (TTV1) of the neutrophilic hyperthermophile genus *Thermoproteus*[39].

The linear genomes of members of the '*Ligamenvirales*' have diverse terminal repeat structures ranging from the large (~1.5–2 kb) inverted terminal repeats (ITRs) of the rudiviruses to the shorter (~0.5–1 kb) ITRs of the lipothrixviruses. Many ITRs have conserved sequence motifs including, for example, short, regularly spaced, direct repeats in the rudiviruses and the lipothrixvirus SIFV[34]. Remarkably, the multiple repeat pattern at the termini of the genome of the lipothrixvirus *Acidianus* filamentous virus 1 (AFV1) resemble the telomeric ends of the linear chromosomes of eukaryotes[37].

The two DNA strands of the rudiviral genomes are covalently linked, 5′-3′, generating small terminal loops at each end[41], however there is no evidence so far for similar structures being present in the lipothrixviral genomes[36].

***The* Rudiviridae *family*.** Virions belonging to this family do not have an envelope and vary considerably in length (the average size is 23 × 610–900 nm), with the length proportional to the size of the linear dsDNA genomes for the three characterized rudiviruses[33,35] (FIG. 3a; TABLE 2). The virions contain a tube-like superhelix formed by linear dsDNA and copies of a single, glycosylated, basic DNA-binding protein. Plugs that are approximately 50 nm in length are located at each end of the tube-like structure and each carries three short tail fibres (FIG. 3a).

SIRV1 replicates producing head-to-head and tail-to-tail linked replicative intermediates, which are probably resolved by the Holliday junction resolvase that is common to all rudiviruses[34,42]. The formation of these intermediates is consistent with a self-priming mechanism of replication, similar to that proposed for some members of a group of the larger eukaryal nucleocytoplasmic DNA viruses (NCLDV)[43].

Consistent with the rudiviral–host relationships being relatively unsophisticated, *in vivo* studies have demonstrated a simple transcription pattern for the viral genomes, with few genes exhibiting temporal regulation[44]. A host-encoded transcription activator was shown to be involved in the regulation of viral transcription[45].

***The* Lipothrixviridae *family*.** The flexible, filamentous virions of members of this family are surrounded by envelopes containing lipids that have been acquired from the host. They show considerable diversity in their terminal structures, which, in combination with differing properties of their linear genomes and presumed differing replication mechanisms, has provided a basis for their classification into four genera.

The virion of the α lipothrixvirus TTV1 (400 nm × 38 nm) has a lipid bilayer envelope which encases a helical core consisting of DNA covered by multimers of two DNA-binding proteins[39]. The β lipothrixvirus SIFV (1,950 nm × 24 nm) differs in that the virion termini taper, ending in mop-like structures to which six tail fibres are attached[36]. This structure unfolds like a spider's legs, before attaching to receptors on the host cell membrane. For the γ lipothrixvirus AFV1 (900 × 24 nm), the virion termini carry exceptional claw-like structures (FIG. 3b), which clamp onto viral receptors located on host cell pili and maintain a firm contact[37]. The virion of the δ lipothrixvirus AFV2 (1100 × 24 nm) has a complex collar at the termini with two sets of attached filaments, resembling a bottle brush[38] (FIG. 3c).

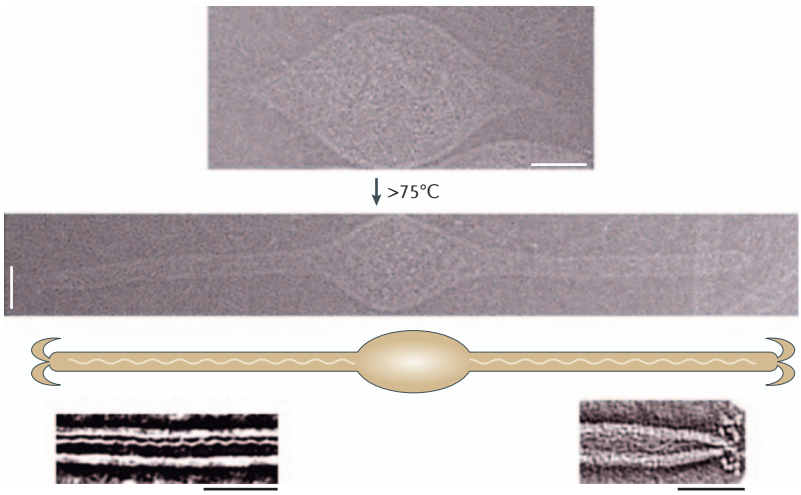


Figure 2 | **Extracellular tail development of *Acidianus* two-tailed virus (ATV).** Cryo-electron micrographs of tailless and two-tailed virions. The lower panels show sections through the three-dimensional reconstructions of different portions of the negatively stained tails obtained by electron tomography. Scale bars represent 50 nm. Modified with permission from REF. 17 © (2006) Elsevier.

Spherical viruses

The known species of archaeal spherical viruses exhibit virions of two main types (FIG. 3 d,e; TABLE 2). The spherical virion (diameter, 100 nm) of the *Pyrobaculum* spherical virus (PSV), a member of the proposed family '*Globuloviridae*', is enveloped and contains multimers of a 33-kDa protein[46] (FIG. 3d). The lipid-containing envelope encases a nucleoprotein core which has a superhelical arrangement. PSV infects species of the anaerobic and hyperthermophilic crenarchaeal genera *Pyrobaculum* and *Thermoproteus*. The linear genome carries 190-bp ITRs, and encodes a putative viroid-like RNA molecule of unknown function (X. Peng, personal communication). The *Thermoproteus tenax* spherical virus 1 (TTSV1), with similar morphological

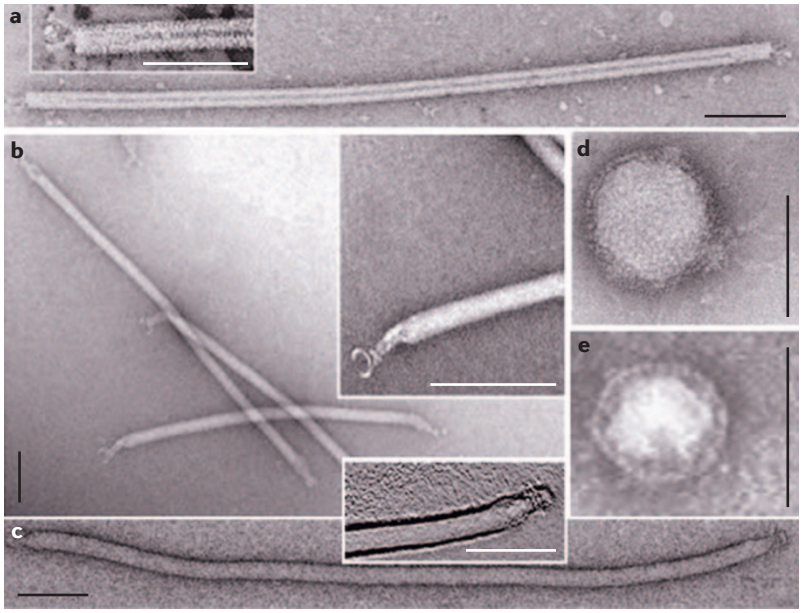


Figure 3 | **Electron micrographs of linear and spherical viruses of the archaea.** The figure shows: **a** | *Sulfolobus islandicus* rod-shaped virus 1 (SIRV1); **b** | *Acidianus filamentous* virus 1 (AFV1); and **c** | *Acidianus filamentous* virus 2 (AFV2) (with terminal structures shown in insets); **d** | *Pyrobaculum* spherical virus (PSV); **e** | *Haloarcula hispanica* virus (SH1). All negatively stained with uranyl acetate. Scale bars represent 100 nm in main images, and 50 nm in insets. Part **a** is courtesy of W. Zillig. Part **b** is reproduced with permission from REF.37 © (2003) Elsevier. Part **c** is reproduced with permission from REF 38 © (2005) American Society for Microbiology. Part **d** is reproduced with permission from REF.46 © (2004) Elsevier. Part **e** is reproduced with permission from REF.51 © (2005) Elsevier.

and genomic properties to PSV, should probably also be assigned to this family[47].

The two other known spherical viruses, *Sulfolobus* turreted icosahedral virus (STIV)[48,49] and *Haloarcula hispanica* virus (SH1) of the haloarchaeal genera *Haloarcula* and *Halorubrum*[50,51], reveal some morphological similarities. The virions of both are non-tailed icosahedra with an internal lipid layer (FIG. 3e). Therefore, they share an architectural principle with virions of the bacterial *Tectiviridae* family. Moreover, the crystal structure of the 37-kDa major capsid protein of STIV is closely similar to those of the major capsid proteins of the bacterial tectivirus PRD1 and the eukaryal phycodnavirus *Paramecium bursaria* Chlorella virus 1 (PBCV1), suggesting a common ancestry, although the protein sequences show no significant similarity[52]. Image reconstruction of the STIV virion revealed a unique virus architecture including complex, turret-like projections extending from each of the vertices[48].

SH1 is a lytic virus with a high burst-size of approximately 200 virus particles per cell, and the release mechanism is based on cell disruption[51]. By contrast, STIV is non-lytic and persists stably in the host cell. In line with differences in morphology and virus-host interactions, the genomes of the two viruses reveal no evidence of homologous gene content. Furthermore, although the STIV genome is circular, that of SH1 is linear, carrying 309-bp ITRs.

Mesophilic
An organism that grows best in a temperature range between 20 °C and 45 °C.

Head-tail viruses

In addition to the icosahedral viruses, archaea replicate another virion morphotype that is common to bacterial viruses. These are the head-tail phages, with non-enveloped virions carrying icosahedral heads and helical tails (FIG. 4). All are associated with the kingdom Euryarchaeota and they exclusively infect extreme halophiles or methanogens that are either mesophilic or moderately thermophilic. Sixteen head-tail phages have been reported, which exhibit diverse sizes of heads, tails and linear dsDNA genomes, and they have been assigned to the bacteriophage families *Myoviridae* and *Siphoviridae* (for reviews, see REFS 8,10,53). Most head-tail phages have not been studied beyond a basic description and are currently unavailable in laboratory collections. The best characterized are two pairs of closely related myoviruses, ΦH[54–57] and ΦCh1 (REFS 58–60), and HF1 and HF2 (REFS 61–64), which infect haloarchaea, and the siphovirus ΨM1 and its deletion mutant ΨM2, which infect *Methanothermobacter*[65–67] (TABLE 3). In addition to their phage-like virion structures, these viruses resemble dsDNA bacteriophages in their genome content (see below), and in possessing mosaic genomes that have undergone extensive genetic exchange[68,69].

Although so far viral diversity has been studied in only a few archaea-rich habitats, it is becoming increasingly clear that head-tail particles might be a rare virion morphotype in environments where archaea dominate. For example, in samples from high-temperature hydrothermal sites, many morphotypes are observed but rarely head-tail forms[23–25,70]. Moreover, in hypersaline waters where haloarchaea predominate, spindle-shaped, spherical and star-shaped virus-like particles are most common[26,27]. Therefore, the abundance of head-tail haloarchaeal viruses among those isolated almost certainly reflects a major bias in the approaches used for their detection and isolation. These were based mainly on their ability to induce host-cell lysis and produce plaques on lawns of a limited diversity of host cells[53].

The ecology of the archaeal viruses

Despite the widespread presence of archaea on our planet, specific screening for archaeal viruses so far has only been done in extreme hydrothermal and hypersaline environments. The available results suggest that the composition of viral communities reflects that of their hosts and is similar at different geographical locations with comparable environmental conditions. Therefore, overlapping subsets of known morphotypes of hyperthermophilic viruses have been observed at geothermal environments in Iceland, Eastern Russia (the Kamchatka peninsula), the Naples region of Italy and Yellowstone National Park in the USA, and in deep-sea hydrothermal vents[23–25,31,70].

Viruses from the same family which have been isolated from different geographical sites and infect closely related hosts often show strong sequence conservation. Therefore, approximately half of the genes of PSV and TTSV1, which have been assigned to the *'Globuloviridae'* with origins in the USA and Indonesia, respectively, are orthologous[10]. A comparable degree of similarity occurs for members of the

Table 2 | **Archaeal linear and spherical viruses**

Family and genus	Species	Archaeal kingdom	Genome details*	References
Linear viruses				
Rudiviridae, *Rudivirus*	*Sulfolobus islandicus* rod-shaped virus 1 (SIRV1)	Cr	ln, 32.3	33
	Sulfolobus islandicus rod-shaped virus 1 (SIRV2)	Cr	ln, 35.5	33
	Acidianus rod-shaped virus 1 (ARV1)	Cr	ln, 24.7	35
Lipothrixviridae, α-*lipothrixvirus*	*Thermoproteus tenax* virus 1 (TTV1)§	Cr	ln, 15.9	39, 40
Lipothrixviridae, β-*lipothrixvirus*	*Sulfolobus islandicus* filamentous virus (SIFV)	Cr	ln, 40.9	36
	Thermoproteus tenax virus 2 (TTV2)§	Cr	ND	39
	Thermoproteus tenax virus 3 (TT3)§	Cr	ND	39
Lipothrixviridae, γ-*lipothrixvirus*	*Acidianus* filamentous virus 1 (AFV1)	Cr	ln, 21.1	37
Lipothrixviridae, δ-*lipotrixvirus*‡	*Acidianus* filamentous virus 2 (AFV2)	Cr	ln, 31.8	38
Spherical viruses				
'*Globuloviridae*'‡, '*Globulovirus*'	*Pyrobaculum* spherical virus (PSV)	Cr	ln, 28.3	46
	Thermoproteus tenax spherical virus 1 (TTSV1)	Cr	ln, 20.9	47
Unclassified	*Haloarcula hispanica* virus (SH1)	Eu	ln, 30.9	50, 51
	Sulfolobus turreted icosahedral virus (STIV)	Cr	ln, 17.7	48

*Genome details show the form of the genome and the size in kb. ‡Currently not available in laboratory collections. §Taxonomic proposal pending at the International Committee on Taxonomy of Viruses. c, covalently closed circular; Cr, Crenarchaeota; Eu, Euryarchaeota; ln, linear; ND, not determined.

Fuselloviridae isolated on three different continents[20,71]. These genomes contain a fairly conserved region, or conserved sets of genes, and a more variable region[19,20]. Some of this genetic diversity might arise by intergenomic recombination between related species.

The high similarity in the gene content of viral species from the same family that are separated by large geographical distances is in contrast to the minimal similarity observed for species from different families that coexist in the same local microbial community. The most extreme example of this is provided by four viruses that were isolated from an acidic hot spring (pH2, 85°C) at Pozzuoli, near Naples, Italy. The rod-shaped ARV1, filamentous AFV2, bottle-shaped ABV and two-tailed ATV, representing four different viral families, coexist in the same habitat and can even replicate in the same host strains. Despite this, these viruses show no detectable signs of intergenomic genetic exchange.

Acidothermophilic aquatic environments, from which the most unusual viral morphotypes have been isolated, contain a concentration of viruses that is significantly lower than the concentration observed in other analysed ecosystems[72,73]. This could be caused by several factors including, for example, the limited stability of virions at high temperatures and low pH, or by the capacity of most known viruses from such environments to persist stably in the host cell rather than cause lysis. The latter strategy apparently reduces the possibility of direct exposure of a virus to the harsh environmental conditions. ATV, the only known lytic virus from an acidic hot spring, finalizes its replication cycle extracellularly under the environmental conditions that are favourable for host growth, and this might contribute to its survival strategy.

Ribbon-helix-helix (RHH). A structural motif consisting of four helices in an open array of two hairpins.

Helix-turn-helix (HTH). A structural motif common in DNA-binding proteins; typically the second helix fits into the DNA major groove.

Archaeal virus genomics

The genes of archaeal viruses generally yield few significant matches to sequences in the public sequence databases[10]. The exceptions are the several homologous matches between genes of the euryarchaeal head-tail viruses ΦCh1, ΨM1/2, HF1 and HF2, and those of the head-tail bacteriophages from the families *Myoviridae* and *Siphoviridae*[58,63,64,67]. These viruses carry homologous genes encoding, for example, proteins involved in capsid and tail formation and virion assembly, as well as transcriptional regulators, ATPases and nucleases.

For non-head-tail viruses, most predicted gene products lack recognizable functions and homologues in extant databases, other than in closely related archaeal viruses. Identified functions are confined to a few proteins involved in glycosylation (glycosyltransferases[74]), DNA replication, recombination and integration (a Holliday junction resolvase[42], DNA polymerase (REF. 11 and X. Peng, personal communication), RecB endonuclease and integrase[28]), nucleotide metabolism (a dUTPase[75] and thymidylate synthase), diverse transcriptional regulatory proteins, mainly of the ribbon-helix-helix (RHH) and helix-turn-helix (HTH) types, and functionally diverse ATPases. In a recent re-evaluation and comparative study of all the annotated genomes of crenarchaeal viruses, working close to the limits of statistical significance, several additional sequence matches were identified, in particular matches to small transcriptional regulatory proteins. The functional subgroups of, for example, AAA+ ATPases (ATPases associated with various cellular activities), and DNA replication enzymes, were classified in more detail[10]. Nevertheless, the current uniqueness of most gene products of these viruses

von Willebrand factor A motif
A structural motif that has been implicated in the formation of diverse types of specific protein–protein interaction and cell adhesion.

was verified, strengthening the idea that crenarchaeal viruses are unrelated to any other viruses[10]. These general properties and conclusions also extend to the non head-tail euryarchaeal viruses, including the fusiform viruses His1, His2 and PAV, and the spherical virus SH1, which have been sequenced and analysed more recently (REFS 11,50 and C. Geslin, personal communication).

Few homologous genes are shared between members of different virus families (except for the '*Ligamenviridae*') and there is no evidence so far for intergenomic recombination occurring between virus families. Nevertheless, comparative studies on fusellovirus genomes suggest that they carry one region in which the gene content varies[19,20,76], and evidence for intergenomic recombination has recently been observed among genomes of the lipothrixviruses (G. Vestergaard, personal communication).

Exceptionally, the bicaudavirus ATV carries four transposase genes, possibly present as insertion sequence (IS) elements, which are similar in sequence to the transposase genes in *Sulfolobus* spp. chromosomes and conjugative plasmids[17], suggesting that some intergenomic exchange has occurred (FIG. 5a). ATV is exceptional in that it encodes six large proteins, in the size range of 600–1,940 amino acids, which are rich in repeat structures, with some exhibiting extensive predicted coiled-coil regions, as well as regions of low complexity sequence, one of which carries a Von Willebrand factor A motif. At least some of these proteins seem to contribute to the extracellular development of the tail-like appendages[14,17].

Cellular defence or regulatory mechanisms

Almost all archaeal genomes contain large clusters of regularly spaced repeat elements (known as short regularly spaced repeats (SRSR) or clustered regularly interspaced short palindrome repeats (CRISPR)[77]) that can constitute more than 1% of the whole genome[78]. Much smaller clusters are found in about 50% of bacterial genomes[78]. For a given archaeon cluster, although the sequence of the repeats and size of the spacer sequences tend to be highly conserved, the spacer sequences themselves are generally unique[78]. Initially, evidence was presented for the clusters resembling centromere structures and being involved in chromosomal segregation or plasmid partitioning[77].

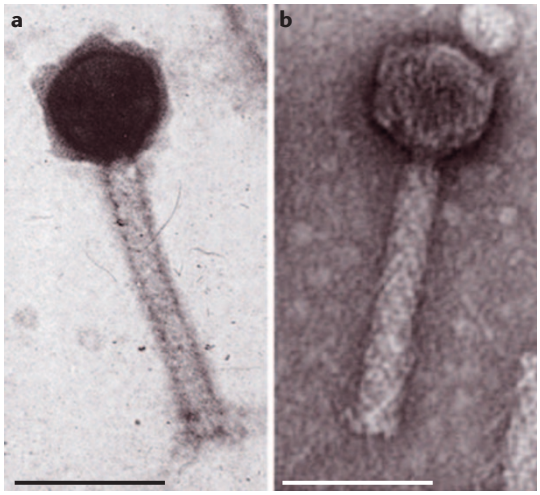


Figure 4 | **Electron micrographs of head-tail viruses of archaea and bacteria. a** | The haloarchaeal virus ϕH1. **b** | The bacteriophage P2. Both are negatively stained with uranyl acetate. Scale bars represent 100 nm. Part **a** is courtesy of W. Zillig. Part **b** is reproduced with permission from http://www.biochem.wisc.edu/inman/empics.

Although this hypothesis has not been refuted, recent evidence from the same authors demonstrated that a few of the spacer sequences present in these clusters are highly similar to sequences found in the genomes of archaeal rudiviruses or conjugative plasmids[79]. This result was reinforced on the completion of the ATV viral genome sequence, which revealed nine positive viral matches with spacer sequences in clusters of the *Sulfolobus solfataricus* P2 genome[78] (FIG. 5). As the repeat clusters produce larger transcripts that are probably processed into smaller fragments approaching the size of the spacer sequences, they can potentially target and inactivate viral messenger RNAs or genes[78,80,81]. Therefore, it has been suggested that the spacer sequences might provide the basis for a cellular defence mechanism against future invasion by related viruses or plasmids[79]. Bioinformatics evidence suggests that the group of genes, designated *cas* for CRISPR-associated, are physically associated and co-functional with some clusters[82]. They might be involved in adding new repeat-spacer units to the clusters by interacting with the flanking sequences that adjoin actively

Table 3 | **Archaeal head-tail viruses***

Family and genus	Species	Host	Archaeal kingdom	Genome details‡	References
Myoviridae, '*ϕH-like viruses*'	ϕH	*Halobacterium salinarum*	Eu	ln, 59.0	54–56
	ϕCh1	*Natrialba magadii*	Eu	ln, 58.5	58,60
Unassigned species in the family	HF1	*Haloferax volcanii* and *Halobacterium salinarum*	Eu	ln, 75.9	61,63
	HF2	*Halorubrum coriense*	Eu	ln, 77.7	61–64
Siphoviridae, '*ψM1-like viruses*'	ψM1/2	*Methanothermobacter marburgensis*	Eu	ln, 30.4/26.1	65–67

*Only those with sequenced genomes are listed. ‡Genome details show the form of the genome and the size in kb. Eu, Euryarchaeota; ln, linear.

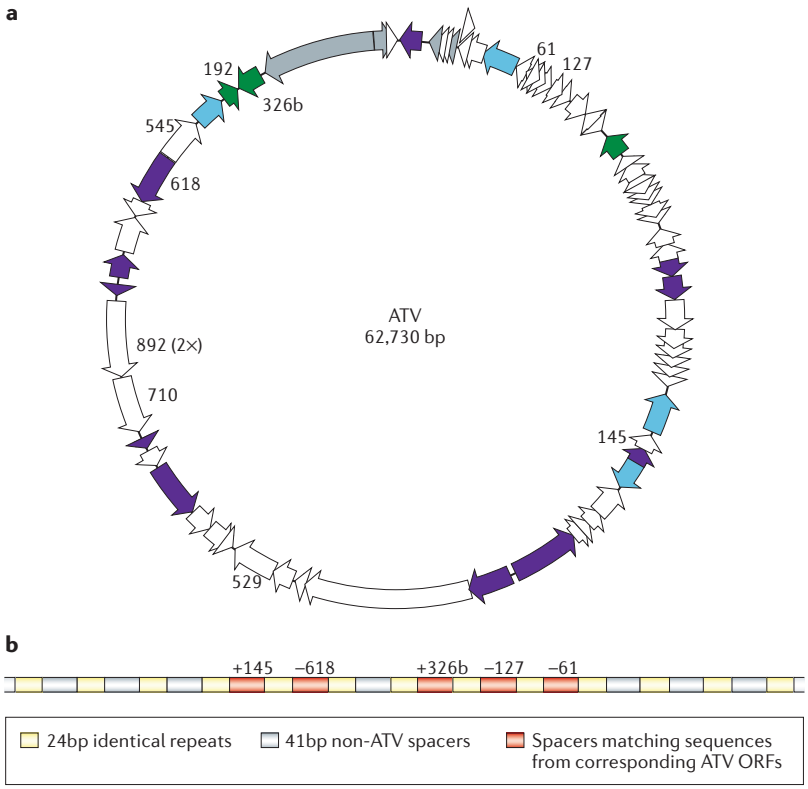


Figure 5 | **Genome structure of *Acidianus* two-tailed virus (ATV). a** | Map of the circular ATV genome with open reading frames (ORFs) represented by arrows and labelled according to the number of amino acids in the predicted protein. The ORFs are colour coded as follows: grey, have homologues in other crenarchaeal viral genomes; green, have homologues in archaeal conjugative plasmids; blue, transposases; purple, genes for structural proteins. For the numbered ORFs, fragments of identical, or near identical, sequences are present as CRISPR spacers in the chromosome of *Sulfolobus solfataricus* P2. **b** | Repeat-spacer elements from a section of the 96-repeat-spacer CRISPR cluster 'd' of *S. solfataricus* P2. The colour-coded elements are: yellow, 24-bp identical repeats; grey, ~41-bp non-ATV spacers; red, spacers matching sequences from corresponding ATV ORFs (+ and – indicate the same and opposite direction, respectively, relative to the ORF mRNA). Part **a** is modified with permission from REF. 17 © (2006) Elsevier. Part **b** is modified with permission from REF. 78.

growing clusters, possibly involving a retrotransposition mechanism[78,83]. The whole apparatus could be evolutionarily related to the eukaryotic small interfering RNA (siRNA) and micro RNA (miRNA) systems[83].

Unresolved questions include why the clusters are so extensive in archaea and why so many spacer matches to ATV are found? One possible explanation for the latter is that whereas most non-head-tail archaeal viruses enjoy a benign relationship with the host, and persist stably in the cell at a low copy number without causing its lysis, a high level of ATV replication results in host lysis and so perhaps the lytic ATV induces a stronger host response.

Some insight into how archaeal viruses overcome cellular defence mechanisms of the host is provided by the rudivirus SIRV1. When passed through diverse, closely related host strains of *Sulfolobus islandicus*, SIRV1 undergoes major genomic changes, including transpositions, duplications and deletions, all concentrated in six main regions of the viral genome and often resulting in an alteration in gene size[84]. Many of the changes are caused by insertions or deletions of 12 bp, although the mechanism by which this occurs is still not fully understood[84]. However, the changes seem to reflect the adaptation of the virus to the new host and could be a reaction to the RNAs being produced from spacers of the chromosomal repeat clusters.

Another mechanism of adaptation to cellular defences might be exemplified by the protein TPX, which is expressed from the α lipothrixvirus TTV1. The genome carries two low-complexity sequence regions consisting of the repetitive sequence CCNACN (where N is any nucleotide) one of which is located within the *tpx* gene and the other in a non-coding region. Frequent recombination seems to occur between these two regions, which generates multiple variants of the TPX protein in the cell[85].

Evolutionary considerations

Despite a modest number of isolated species, the morphological diversity of the dsDNA viruses of the Archaea strongly surpasses that of the numerous known species of dsDNA viruses of the Bacteria[86]. According to the recent report of the ICTV, 96% of the 452 known species of bacterial dsDNA viruses are non-enveloped head-tail bacteriophages, assigned to three families of the order *Caudovirales (Myoviridae*, *Siphoviridae* and *Podoviridae*), whereas only about 4% are non-enveloped tailless icosahedra or enveloped pleomorphic spheres[87]. As discussed above, in addition to the *Caudovirales* and tailless icosahedral viruses, archaea replicate a plethora of viruses with morphotypes that are not encountered among dsDNA viruses of either bacteria or eukarya (FIG. 6). The diversity of archaeal viruses in hot habitats is especially striking compared with the relative uniformity of the viral landscape in aquatic environments at moderate and low temperatures, which is dominated by head-tail viruses[88]. The origin and nature of this biodiversity raises intriguing evolutionary questions. Possibly, such diversity was once common in all environments but was later reduced by the successful expansion of bacteria and their phages in biotopes with moderate and low temperatures, whereas hot environments still remain a refuge for multiple unusual viral forms.

The occurrence of the *Caudovirales* among archaeal viruses was earlier interpreted as evidence for these viruses predating the divergence of the Archaea and Bacteria, as it was believed that viruses were unable to spread across domain boundaries owing to the pronounced differences in the molecular biology and lifestyles of the two domains[89]. However, it is remarkable that, in known cases, caudoviruses infect archaea carrying a high percentage of genes of bacterial origin, in particular certain mesophilic or moderately thermophilic haloarchaea or methanoarchaea[90,91]. Apparently, these bacterial genes have adapted successfully to an archaeal cellular context. Furthermore, extremely halophilic

Small interfering RNA
Non-coding RNAs (around 22 nucleotides long) derived from the processing of long double-stranded RNA during RNA interference. They direct the destruction of mRNA targets that have the same sequence.

Micro RNA
A short (21–22-nucleotide) RNA silencing trigger that is processed from short stem–loop precursors that are encoded in the genomes of metazoans and certain viruses.

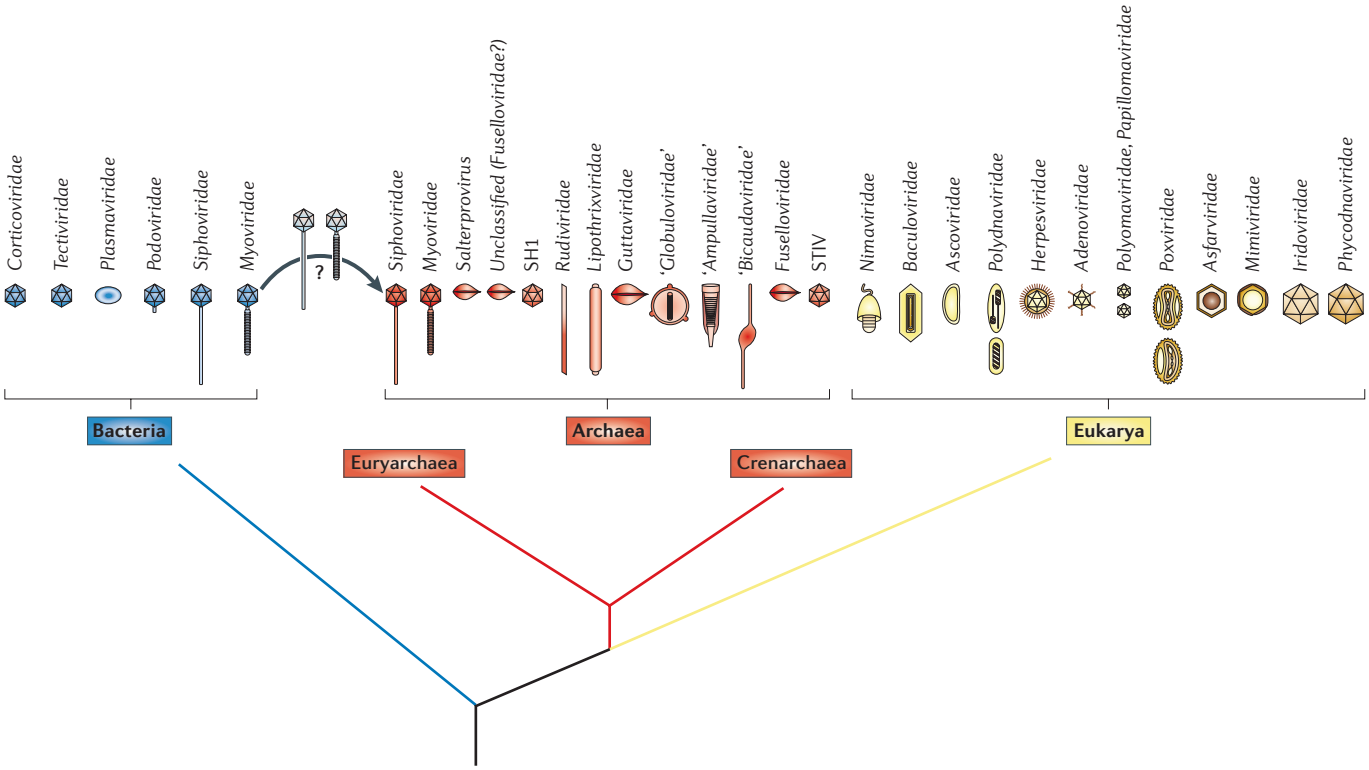


Figure 6 | **The families of dsDNA viruses associated with the three domains of life.** The virus families listed are approved by the International Committee on Taxonomy of viruses, and the schematic representation of virions (not drawn to scale) are presented as in REF. 87. Proposed families are shown in inverted commas. SH1, *Haloarcula hispanica* virus 1; STIV, *Sulfolobus* turreted icosahedral virus.

bacteria have been discovered, which could have facilitated the adaptation of bacteriophages to replicate in the haloarchaea[92]. Therefore, it now seems likely that the *Caudovirales* first entered archaea by interdomain spreading. Such an origin is also consistent with archaeal caudaviruses carrying many genes with homologues in either bacterial chromosomes or plasmids.

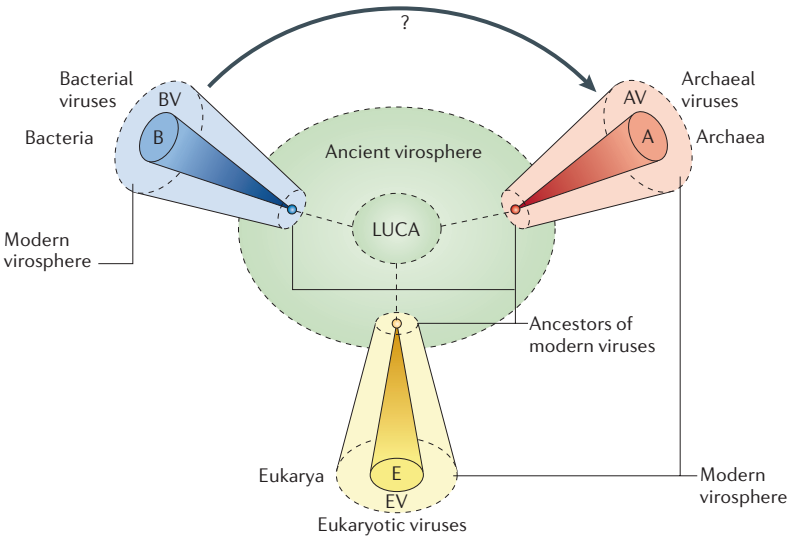
Assuming that archaeal *Caudovirales* do, indeed, originate from bacteria, then we are faced with the intriguing perspective that each of the three domains of life was originally characterized by a unique set of associated dsDNA viruses (FIG. 6). Could the association of a specific set of viruses with each domain mean that the old preconception of viruses originating from their hosts is correct? Probably not, because there are no apparent exclusive evolutionary links between the basic viral components, the capsid and replication apparatus found in one domain and the cellular components of that domain. On the contrary, recent structural studies have shown that some viruses present in different domains could have evolved from entities (or a common gene pool) that predated the divergence of the domains[93]. For instance, the icosahedral archaeal viruses STIV and SH1 share a basic architectural principle and/or a capsid protein with similar structures (demonstrated for STIV) with several bacterial (*Tectiviridae*) and eukaryotic (*Phycodnaviridae*, *Adenoviridae*) icosahedral viruses[52].

One possible explanation for the existence of three different 'virospheres', each associated with a specific domain, is that these virospheres were selected when the domains first arose. Therefore, the first evolving organisms of each separate domain could have already been infected by different subsets of viruses from the ancestral virosphere, which predated the last universal common ancestor, LUCA. Although the descendants of LUCA might have rapidly expanded on the planet, the lineages leading to the three ancestors and their associated viruses might well have originated later in distinct areas and remained geographically isolated for a long time period. Therefore, viruses might already have been specialized to co-evolve with cells of one domain and were unable to efficiently infect cells of another domain when the cells started to share the same environment. The mechanism and the timing of the speciation of the three domains remains unknown, however, a recent hypothesis suggests that DNA viruses could have played a direct role in their formation by introducing DNA genomes into ancient RNA cells (BOX 1). Such a model would be compatible with each domain being associated with its own viral landscape if one assumes that three ancestral RNA cells were infected by three different sets of DNA viruses at the time of the RNA–DNA transition (BOX 1).

The present lack of known archaeal RNA viruses also raises an intriguing evolutionary issue. The strategies recently used to detect and isolate archaeal viruses,

Last universal common ancestor
The progenitor from which all current life is thought to have evolved.

Box 1 | **The origins of viruses**



Three hypotheses have been proposed to explain the origin of viruses: (**a**) they originated in a pre-cellular world (the 'virus first' hypothesis); (**b**) they originated by a reduction from parasitic cells; or (**c**) they originated from fragments of cellular genetic material that escaped from cellular control and became parasites (the escape hypothesis). Originally, these hypotheses were proposed in the framework of the prokaryote/eukaryote dichotomy. Just as the erroneous concept of the prokaryotes became the paradigm for considering bacterial evolution, the escape hypothesis became the paradigm for explaining the origin of viruses.

In its classical version, the escape hypothesis maintained that bacteriophages originated from bacterial genomes and eukaryotic viruses from eukaryotic genomes. Amazingly, although most archaeal viruses are completely unrelated to bacterial viruses, they are still classified as 'bacteriophages' in this outdated framework. For example, archaeal viruses are illustrated under the heading "Families and genera infecting bacteria" in the latest edition of *Virus Taxonomy: Classification and Nomenclature of Viruses* (REF. 87). This occurs presumably because the archaeal domain is still not recognized by some biologists.

The problem of virus origin remained deadlocked until recently, when progress in the molecular description of viral proteins caused many scientists to realize that viruses form a world of their own, and that it is futile to continue to speculate on their origin in the framework of the discredited prokaryote/eukaryote dichotomy. The discovery that viruses which are associated with different domains can share similar, and apparently homologous, features strongly suggests that viruses are ancient and that they predated the last universal cellular ancestor (LUCA)[93,95].

The early hypotheses for viral origin have now been re-evaluated in the context of this new framework[95]. Currently, the main debate is between those who suggest a long period of acellular evolution (up to the actual emergence of archaea and bacteria) and those who favour an early appearance of cells. Those who suggest the former have revived the virus-first hypothesis, hypothesis (**a**) above. For instance, Koonin and Martin recently suggested that viruses emerged from an assemblage of self-replicating elements thriving in a hydrothermal vent, using inorganic compartments as their hosts[96]. For some of those who favour an early emergence of cellular organisms, viruses are considered to have originated from RNA–protein-based cells, either by reduction from parasitic RNA cells or from genetic material that escaped from the genomes of these cells (variants of hypotheses (**b**) and (**c**) above)[95]. A major question mark is the evolutionary relationship between DNA viruses and RNA viruses: did DNA viruses originate from RNA viruses or from primitive DNA cells, or both[95,97]? In one hypothesis, DNA itself is considered a viral invention, that is, DNA first appeared in viruses and was later transferred to cells[98]. In this model, three such independent transfers could have initiated the three modern DNA lineages and the modern virospheres associated with them[99].

The figure shows a model of the formation of the three modern virospheres, coincident with the formation of the three domains of life.

at least from hydrothermal environments[72], might have been expected to yield RNA viruses, although one cannot yet exclude a methodological problem. A possible explanation for their absence could be that they were already absent from the virosphere associated with the common archaeal ancestor. Given that RNA is much more labile than DNA at high temperature, this inference would concur with recent phylogenetic analyses based on the available sequences of the translation and transcription apparatus, which indicate that the archaeal ancestor was a hyperthermophile[94].

The discovery and exploration of the fascinating archaeal DNA viruses has paralleled, and contributed to, a recent upsurge of interest in the evolution of viruses in general (BOX 1). Although the known archaeal viruses reveal an exceptional degree of diversity with regard to both morphotypes and genomes, this might still be an underestimation as they have only been isolated from a limited number of host species and taxa (mainly from the orders *Sulfolobales* and *Halobacteriales*). Therefore, screening for virus–host systems in other phylogenetic taxa of the archaeal domain is an important research priority that promises to provide new insights into important questions concerning the nature and origin of the modern virosphere.

1. Woese, C. R. & Fox, G. E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl Acad. Sci. USA* **74**, 5088–5090 (1977).
2. Woese, C. R., Kandler, O. & Wheelis, M. L. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl Acad. Sci. USA* **87**, 4576–4579 (1990).
3. Garrett, R. A. & Klenk, H.-P. (eds) *Archaea. Evolution, Physiology and Molecular Biology* (Blackwell Publishing, Oxford, 2006).
This is a recent and comprehensive book that covers the ecology, physiology, molecular biology and evolution of the Archaea.
4. Torsvik, T. & Dundas, I. D. Bacteriophage of *Halobacterium salinarium*. *Nature* **248**, 680–681 (1974).
5. Wais, A. C., Kon, M., MacDonald, R. E. & Stollar, B. D. Salt-dependent bacteriophage infecting *Halobacterium cutirubrum* and *H. halobium*. *Nature* **256**, 314–315 (1975).
6. Pauling, C. Bacteriophages of *Halobacterium halobium*: isolated from fermented fish sauce and primary characterization. *Can. J. Microbiol.* **28**, 916–921 (1982).
7. Rohrmann, G. F., Cheney, R. & Pauling, C. Bacteriophages of *Halobacterium halobium:* virion DNAs and proteins. *Can. J. Microbiol.* **29**, 627–629 (1983).
8. Dyall-Smith, M., Tang, S. L. & Bath, C. Haloarchaeal viruses: how diverse are they? *Res. Microbiol.* **154**, 309–313 (2003).
9. Prangishvili, D. Evolutionary insights from studies on viruses of hyperthermophilic archaea. *Res. Microbiol.* **154**, 289–294 (2003).
10. Prangishvili, D., Garrett, R. A. & Koonin, E. V. Evolutionary genomics of archaeal viruses: Unique viral genomes in the third domain of life. *Virus Res.* **117**, 52–67 (2006).
Describes a detailed analysis of proteins encoded by archaeal viruses, with an emphasis on the comparative genomics of crenarchaeal viruses.
11. Bath, C., Cukalac, T., Porter, K. & Dyall-Smith, M. L. His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, *Salterprovirus*. *Virology* **350**, 228–239 (2006).
12. Bath, C. & Dyall-Smith, M. L. His1, an archaeal virus of the *Fuselloviridae* family that infects *Haloarcula hispanica*. *J. Virol.* **72**, 9392–9395 (1998).
13. Geslin, C. *et al.* PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, '*Pyrococcus abyssi*'. *J. Bacteriol.* **185**, 3888–3894 (2003).
14. Haring, M. *et al.* Virology: independent virus development outside a host. *Nature* **436**, 1101–1102 (2005).
15. Martin, A. *et al.* SAV1, a temperate u.v.-inducible DNA virus-like particle from the archaebacterium *Sulfolobus acidocaldarius* isolate B12. *EMBO J.* **3**, 2165–2168 (1984).
16. Palm, P. *et al.* Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae*. *Virology* **185**, 242–250 (1991).
17. Prangishvili, D. *et al.* Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle. *J. Mol. Biol.* **359**, 1203–1216 (2006).
18. Schleper, C., Kubo, K. & Zillig, W. The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. *Proc. Natl Acad. Sci. USA* **89**, 7645–7649 (1992).
19. Stedman, K. M. *et al.* Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. *Res. Microbiol.* **154**, 295–302 (2003).
20. Wiedenheft, B. *et al.* Comparative genomic analysis of hyperthermophilic archaeal *Fuselloviridae* viruses. *J. Virol.* **78**, 1954–1961 (2004).
21. Wood, A. G., Whitman, W. B. & Konisky, J. Isolation and characterization of an archaebacterial virus-like particle from *Methanococcus voltae* A3. *J. Bacteriol.* **171**, 93–98 (1989).
22. Xiang, X. *et al.* *Sulfolobus tengchongensis* spindle-shaped virus STSV1: virus–host interactions and genomic features. *J. Virol.* **79**, 8677–8686 (2005).
23. Rachel, R. *et al.* Remarkable morphological diversity of viruses and virus-like particles in hot terrestrial environments. *Arch. Virol.* **147**, 2419–2429 (2002).
24. Rice, G. *et al.* Viruses from extreme thermal environments. *Proc. Natl Acad. Sci. USA* **98**, 13341–13345 (2001).
25. Zillig, W. *et al.* Screening for *Sulfolobales*, their plasmids, and their viruses in Icelandic solfataras. *Syst. Appl. Microbiol.* **16**, 609–628 (1994).
References 23, 24 and 25 describe the diversity of viral morphotypes observed in samples from hot springs in the United States and Iceland.
26. Guixa-Boixareu, N., Calderon-Paz, J. I., Heldal, M., Bratbak, G., Pedros-Alio, C. Viral lysis and bacterivory as prokaryotic loss factors along a salinity gradient. *Aquat. Microb. Ecol.* **11**, 215–227 (1996).
27. Oren, A., Bratbak, G. & Heldal, M. Occurrence of virus-like particles in the Dead Sea. *Extremophiles* **1**, 143–149 (1997).
Describes virus-like particles in samples from the hypersaline Dead Sea.
28. Muskhelishvili, G., Palm, P. & Zillig, W. SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. *Mol. Gen. Genet.* **237**, 334–342 (1993).
29. Nadal, M., Mirambeau, G., Forterre, P., Reiter, W.-D. & Duguet, M. Positively supercoiled DNA in a virus-like-particle of an archaebacterium. *Nature* **321**, 256–258 (1986).
30. Zillig, W. *et al.* in *The Biochemistry of Archaea (Archaebacteria)* (ed. Kates, M. *et al.*) 367–391 (Elsevier Science Publishers B. V., 1993).
31. Haring, M., Rachel, R., Peng, X., Garrett, R. A. & Prangishvili, D. Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, *Acidianus* bottle-shaped virus, from a new family, the *Ampullaviridae*. *J. Virol.* **79**, 9904–9911 (2005).
Describes the viral diversity in a single acidic, hot spring in southern Italy.
32. Arnold, H. P., Ziese, U. & Zillig, W. SNDV, a novel virus of the extremely thermophilic and acidophilic archaeon *Sulfolobus*. *Virology* **272**, 409–416 (2000).
33. Prangishvili, D. *et al.* A novel virus family, the *Rudiviridae*: Structure, virus-host interactions and genome variability of the *Sulfolobus* viruses SIRV1 and SIRV2. *Genetics* **152**, 1387–1396 (1999).
34. Peng, X. *et al.* 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 (2001).
35. Vestergaard, G. *et al.* A novel rudivirus, ARV1, of the hyperthermophilic archaeal genus *Acidianus*. *Virology* **336**, 83–92 (2005).
36. Arnold, H. P. *et al.* A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus*. *Virology* **267**, 252–266 (2000).
37. Bettstetter, M., Peng, X., Garrett, R. A. & Prangishvili, D. AFV1, a novel virus infecting hyperthermophilic archaea of the genus *Acidianus*. *Virology* **315**, 68–79 (2003).
38. Haring, M. *et al.* Structure and genome organization of AFV2, a novel archaeal lipothrixvirus with unusual terminal and core structures. *J. Bacteriol.* **187**, 3855–3858 (2005).
39. Janekovic, D. *et al.* TTV1, TTV2, TTV3, a family of viruses of the extremely thermophilic, anaerobic sulfur-reducing archaebacterium *Thermoproteus tenax*. *Mol. Gen. Genet.* **192**, 39–45 (1983).
40. Neumann, H., Schwass, V., Eckerskorn, C. & Zillig, W. Identification and characterization of the genes encoding three structural proteins of the *Thermoproteus tenax* virus TTV1. *Mol. Gen. Genet.* **217**, 105–110 (1989).
41. Blum, H., Zillig, W., Mallok, S., Domdey, H. & Prangishvili, D. The genome of the archaeal virus SIRV1 has features in common with genomes of eukaryal viruses. *Virology* **281**, 6–9 (2001).
42. Birkenbihl, R. P., Neef, K., Prangishvili, D. & Kemper, B. Holliday junction resolving enzymes of archaeal viruses SIRV1 and SIRV2. *J. Mol. Biol.* **309**, 1067–1076 (2001).
43. Moss, B. in *Fields Virology* (eds Fields, B. N., Knipe, D. M. & Howley, P. M.) 2637–2671 (Lippincott-Raven, Philadelphia, 1996).
44. Kessler, A., Brinkman, A. B., van der Oost, J. & Prangishvili, D. Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon *Sulfolobus*. *J. Bacteriol.* **186**, 7745–7753 (2004).
45. Kessler, A. *et al.* A novel archaeal regulatory protein, Sta1, activates transcription from viral promoters. *Nucleic Acids Res.* 14 Sep 2006 (doi:10.1093/nar/gk1502).
Reports on the strategy of an archaeal virus to co-opt a host-cell regulator to promote the transcription of some of its genes.
46. Haring, M. *et al.* 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 (2004).
47. Ahn, D. G. *et al.* TTSV1, a new virus-like particle isolated from the hyperthermophilic crenarchaeote *Thermoproteus tenax*. *Virology* **351**, 280–290 (2006).
48. Rice, G. *et al.* 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. USA* **101**, 7716–7720 (2004).
49. Maaty, W. S. *et al.* 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 (2006).
50. Bamford, D. H. *et al.* Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon *Haloarcula hispanica*. *J. Virol.* **79**, 9097–9107 (2005).
51. Porter, K. *et al.* SH1: A novel, spherical halovirus isolated from an Australian hypersaline lake. *Virology* **335**, 22–33 (2005).
52. Khayat, R. *et al.* Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc. Natl Acad. Sci. USA* **102**, 18944–18949 (2005).
Shows that the major capsid protein of STIV is highly similar in structure to capsid proteins of some bacterial and eukaryal viruses.
53. Porter, K., Dyall-Smith, M. L. in *Methods in Microbiology — Extremophiles* (eds Reiney, F. A. & Oren, A.) 681–702 (Elsevier, London, 2006).
Reviews cultivation methods for the haloarchaea and their viruses.
54. Gropp, F., Grampp, B., Stolt, P., Palm, P. & Zillig, W. The immunity-conferring plasmid p phi HL from the *Halobacterium salinarium* phage phiH: nucleotide sequence and transcription. *Virology* **190**, 45–54 (1992).
55. Schnabel, H., Palm, P., Dick, K. & Grampp, B. Sequence analysis of the insertion element ISH1.8 and of associated structural changes in the genome of phage phiH of the archaebacterium *Halobacterium halobium*. *EMBO J.* **3**, 1717–1722 (1984).
56. Schnabel, H. *et al. Halobacterium halobium* phage phiH. *EMBO J.* **1**, 87–92 (1982).
57. Stolt, P. & Zillig, W. Transcription of the halophage phiH repressor gene is abolished by transcription from an inversely oriented lytic promoter. *FEBS Lett.* **344**, 125–128 (1994).
58. Klein, R. *et al. Natrialba magadii* virus phiCh1: first complete nucleotide sequence and functional organization of a virus infecting a haloalkaliphilic archaeon. *Mol. Microbiol.* **45**, 851–863 (2002).
59. Rossler, N., Klein, R., Scholz, H. & Witte, A. Inversion within the haloalkaliphilic virus phi Ch1 DNA results in differential expression of structural proteins. *Mol. Microbiol.* **52**, 413–426 (2004).
60. Witte, A. *et al.* Characterization of *Natronobacterium magadii* phage phi Ch1, a unique archaeal phage containing DNA and RNA. *Mol. Microbiol.* **23**, 603–616 (1997).
61. Nuttall, S. D. & Dyall-Smith, M. L. HF1 and HF2: novel bacteriophages of halophilic archaea. *Virology* **197**, 678–684 (1993).
62. Nuttall, S. D. & Dyall-Smith, M. L. Halophage HF2: genome organization and replication strategy. *J. Virol.* **69**, 2322–2327 (1995).
63. Tang, S. L., Nuttall, S. & Dyall-Smith, M. Haloviruses HF1 and HF2: evidence for a recent and large recombination event. *J. Bacteriol.* **186**, 2810–2817 (2004).
64. Tang, S. L. *et al.* HF2: a double-stranded DNA tailed haloarchaeal virus with a mosaic genome. *Mol. Microbiol.* **44**, 283–296 (2002).
65. Jordan, M., Meile, L. & Leisinger, T. Organization of *Methanobacterium thermoautotrophicum* bacteriophage psi M1 DNA. *Mol. Gen. Genet.* **220**, 161–164 (1989).
66. Meile, L., Jenal, U., Studer, D., Jordan, M., Leisinger, T. Characterization of psiM1, a virulent phage of *Methanobacterium thermoautotrophicum* Marburg. *Arch. Microbiol.* **152**, 105–110 (1989).
67. Pfister, P., Wasserfallen, A., Stettler, R. & Leisinger, T. Molecular analysis of *Methanobacterium* phage psiM2. *Mol. Microbiol.* **30**, 233–244 (1998).
68. Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E. & Hatfull, G. F. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl Acad. Sci. USA* **96**, 2192–2197 (1999).

69. Pedulla, M. L. *et al.* Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**, 171–182 (2003).
70. Geslin, C., Le Romancer, M., Gaillard, M., Erauso, G. & Prieur, D. Observation of virus-like particles in high temperature enrichment cultures from deep-sea hydrothermal vents. *Res. Microbiol.* **154**, 303–307 (2003).
71. Stedman, K. M., Clore, A. & Combet-Blac, Y. in *SGM Symposium 66: Prokaryotic Diversity — Mechanisms and Significance.* (eds Logan, N. M., Lappin-Scott, H. M. & Oyston, P. C. F.) 131–144 (Cambridge University Press, 2006).
72. Prangishvili, D. in *Methods in Microbiology —- Extremophiles* (eds Reiney, F. A. & Oren, A.) 331–348 (Elsevier, London, 2006).
Reviews the methods of isolating archaeal virus–host systems from hydrothermal environments.
73. Ortmann, A. C., Wiedenheft, B., Douglas, T. & Young, M. Hot crenarchaeal viruses reveal deep evolutionary connections. *Nature Rev. Microbiol.* **4**, 520–528 (2006).
74. Larson, E. T., Reiter, D., Young, M. & Lawrence, C. M. Structure of A197 from *Sulfolobus* turreted icosahedral virus: a crenarchaeal viral glycosyltransferase exhibiting the GT-A fold. *J. Virol.* **80**, 7636–7644 (2006).
75. Prangishvili, D. *et al.* Biochemical and phylogenetic characterization of the dUTPase from the archaeal virus SIRV. *J. Biol. Chem.* **273**, 6024–6029 (1998).
76. Snyder, J. C. *et al.* Effects of culturing on the population structure of a hyperthermophilic virus. *Microb. Ecol.* **48**, 561–566 (2004).
77. Mojica, F. J., Ferrer, C., Juez, G. & Rodriguez-Valera, F. Long stretches of short tandem repeats are present in the largest replicons of the archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. *Mol. Microbiol.* **17**, 85–93 (1995).
78. Lillestol, R. K., Redder, P., Garrett, R. A. & Brugger, K. A putative viral defence mechanism in archaeal cells. *Archaea* **2**, 59–72 (2006).
79. Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005).
80. Tang, T. H. *et al.* Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc. Natl Acad. Sci. USA* **99**, 7536–7541 (2002).
81. Tang, T. H. *et al.* Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* **55**, 469–481 (2005).
82. Jansen, R., van Embden, J. D., Gaastra, W. & Schouls, L. M. Identification of a novel family of sequence repeats among prokaryotes. *Omics* **6**, 23–33 (2002).
83. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1**, 7 (2006).
84. Peng, X., Kessler, A., Phan, H., Garrett, R. A. & Prangishvili, D. Multiple variants of the archaeal DNA rudivirus SIRV1 in a single host and a novel mechanism of genomic variation. *Mol. Microbiol.* **54**, 366–375 (2004).
85. Neumann, H. & Zillig, W. The TTV1-encoded viral protein TPX: primary structure of the gene and the protein. *Nucleic Acids Res.* **18**, 195 (1990).
86. Ackermann, H. W. Frequency of morphological phage descriptions in the year 2000. Brief review. *Arch. Virol.* **146**, 843–857 (2001).
87. Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U. & Ball, L. A. (eds) *Virus Taxonomy: Classification and Nomenclature of Viruses* (Elsevier, Amsterdam, 2005).
88. Wommack, K. E. & Colwell, R. R. Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* **64**, 69–114 (2000).
89. Zillig, W. *et al.* Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic Archaea. *FEMS Microbiol. Rev.* **18**, 225–236 (1996).
90. Koonin, E. V., Makarova, K. S. & Aravind, L. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55**, 709–742 (2001).
91. Wiezer, A, & Merkl, R. A comparative categorization of gene flux in diverse microbial species. *Genomics* **86**, 462–475 (2005).
92. Mongodin, E. F. *et al.* The genome of *Salinibacter ruber*: convergence and gene exchange among hyperhalophilic bacteria and archaea. *Proc. Natl Acad. Sci. USA* **102**, 18147–18152 (2005).
93. Bamford, D. H., Grimes, J. M. & Stuart, D. I. What does structure tell us about virus evolution? *Curr. Opin. Struct. Biol.* **15**, 655–663 (2005).
Discusses the potential for structural analyses of virion architecture and coat protein topology to provide insights into viral evolution.
94. Brochier, C., Forterre, P. & Gribaldo, S. An emerging phylogenetic core of Archaea: phylogenies of transcription and translation machineries converge following addition of new genome sequences. *BMC Evol. Biol.* **5**, 36 (2005).
95. Forterre, P. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res.* **117**, 5–16 (2006).
Presents a hypothesis that viruses played a crucial role in the invention of DNA and the development of its replication mechanisms, and in the formation of the three domains of life.
96. Koonin, E. V. & Martin, W. On the origin of genomes and cells within inorganic compartments. *Trends Genet.* **21**, 647–654 (2005).
97. Claverie, J. M. Viruses take center stage in cellular evolution. *Genome Biol.* **7**, 110 (2006).
98. Forterre, P. The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie* **87**, 793–803 (2005).
99. Forterre, P. Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. *Proc. Natl Acad. Sci. USA* **103**, 3669–3674 (2006).

Acknowledgements
We would like to thank all the members of our laboratories who have contributed to this work, and to thank colleagues who have made their unpublished results available to us.

Competing interests statement
The authors declare no competing financial interests.

DATABASES
The following terms in this article are linked online to:
Entrez Genome: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genome
AFV1 | His1 | His2 | PBCV1 | PSV | SH1 | SIFV | SIRV1 | STIV | STSV1 | TTSV1
Entrez Genome Project: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj
Methanococcus voltae | *Pyrococcus abyssi* | *Sulfolobus solfataricus* P2

FURTHER INFORMATION
David Prangishvili's homepage: http://www.pasteur.fr/recherche/unites/bmge-archaea/david%20Prangishvili/prangishvili%20personal%20page.htm
ICTV: http://www.ncbi.nlm.nih.gov/ICTVdb/index.htm
The *Sulfolobus* Database: http://www.sulfolobus.org/cbin/mutagen.pl
Access to this links box is available online.