Review

Viruses of archaea: Structural, functional, environmental and evolutionary genomics

Mart Krupovic[a,*], Virginija Cvirkaite-Krupovic[a], Jaime Iranzo[b], David Prangishvili[a], Eugene V. Koonin[b]

[a] *Department of Microbiology, Institut Pasteur, 25 rue du Dr. Roux, Paris 75015, Paris, France*
[b] *National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD, USA*

ABSTRACT

Viruses of archaea represent one of the most enigmatic parts of the virosphere. Most of the characterized archaeal viruses infect extremophilic hosts and display remarkable diversity of virion morphotypes, many of which have never been observed among viruses of bacteria or eukaryotes. The uniqueness of the virion morphologies is matched by the distinctiveness of the genomes of these viruses, with ∼75% of genes encoding unique proteins, refractory to functional annotation based on sequence analyses. In this review, we summarize the state-of-the-art knowledge on various aspects of archaeal virus genomics. First, we outline how structural and functional genomics efforts provided valuable insights into the functions of viral proteins and revealed intricate details of the archaeal virus-host interactions. We then highlight recent metagenomics studies, which provided a glimpse at the diversity of uncultivated viruses associated with the ubiquitous archaea in the oceans, including Thaumarchaeota, Marine Group II Euryarchaeota, and others. These findings, combined with the recent discovery that archaeal viruses mediate a rapid turnover of thaumarchaea in the deep sea ecosystems, illuminate the prominent role of these viruses in the biosphere. Finally, we discuss the origins and evolution of archaeal viruses and emphasize the evolutionary relationships between viruses and non-viral mobile genetic elements. Further exploration of the archaeal virus diversity as well as functional studies on diverse virus-host systems are bound to uncover novel, unexpected facets of the archaeal virome.

1. Archaea and their viruses

Archaea have been recognized as a third domain of life, in addition to bacteria and eukaryotes, 40 years ago (Woese and Fox, 1977). Although morphologically nearly indistinguishable from bacteria, at the molecular level, archaea present a mixture of features, some of which are closely related to those of eukaryotes, others are shared with bacteria, whereas some appear to be unique. Among the archaea-specific features, the most notable are ether-based lipid membranes, which in some hyperthermophiles, form monolayers rather than typical bilayers (Villanueva et al., 2014), and methanogenesis, the metabolic production of methane (Liu et al., 2012; Valentine, 2007). The eukaryotic-like traits include the information processing machineries responsible for DNA replication, transcription and translation (Forterre, 2013; Makarova and Koonin, 2013; Werner and Grohmann, 2011). Furthermore, members of the archaeal phylum Thaumarchaeota and most members of the Crenarchaeota also rely on the eukaryotic-like cell division machinery, the ESCRT complex, whereas archaea of the phylum

Euryarchaeota, similar to bacteria, employ the FtsZ-based apparatus for cell division (Lindås and Bernander, 2013; Makarova et al.,2010). Regardless of the type of cell division apparatus, many archaea contain certain bacterial-like components of the cell envelope, such as appendages related to type IV pili (Makarova et al., 2016), which could be either inherited from the common ancestor of bacteria and archaea or transferred horizontally between the two domains. Archaea and bacteria also share many defense systems against mobile genetic elements (MGE), such as viruses and plasmids. These include restriction-modification, abortive-infection and toxin-antitoxin systems as well as Argonaute-based innate immunity and CRISPR-Cas adaptive immunity (Koonin et al., 2017). Accordingly, it could be expected that bacteria and archaea would also share the MGE pool. This is indeed the case for insertion sequence (IS)-like transposons (Filée et al., 2007), self-synthesizing MGE named casposons (Krupovic et al., 2017), as well as a variety of conjugative and small, cryptic plasmids (Forterre et al., 2014; Wang et al., 2015). Unexpectedly, however, there is only a limited overlap in terms of common groups of viruses.

* Corresponding author.
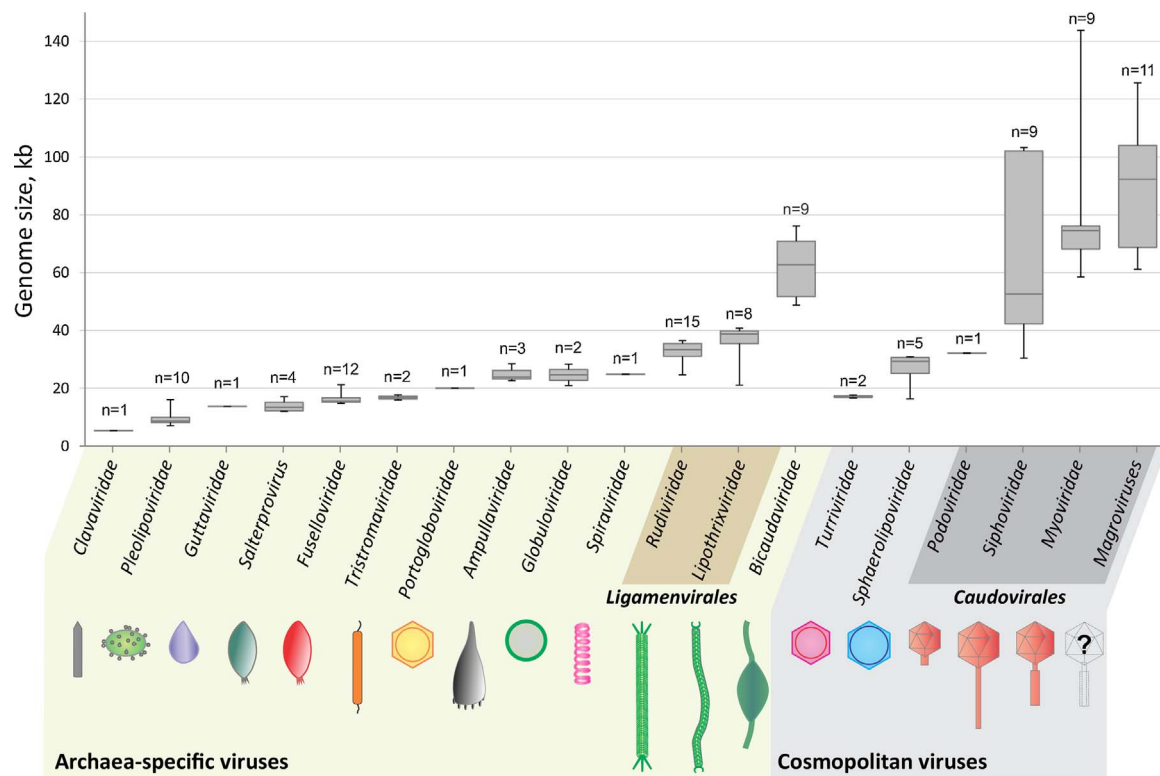 *E-mail address:* krupovic@pasteur.fr (M. Krupovic).

Fig. 1. Genome size distribution and morphotypes of archaeal viruses. In the box plot, each box represents the middle 50th percentile of the data set and is derived using the lower and upper quartile values. The median value is displayed by a horizontal line inside the box. Whiskers represent the maximum and minimum values. The number of genomes used for construction of the box plot is indicated for each group of archaeal viruses. The information on the genome sizes was collected from the GenBank records as well as from the published literature (for genomes assembled from metagenomic data). Archaea-specific and cosmopolitan fractions of the archaeal virosphere are indicated with different background colors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Archaeal viruses are currently formally classified into 17 families (with 5 families unified into 2 orders) and 1 unassigned genus (Fig. 1). In addition, several archaeal viruses have been isolated or discovered through metagenomics approaches but not yet classified due to the lack of appreciable similarity to classified species and thus likely represent new virus taxa. The vast majority of archaeal viruses have been thus far isolated from either hyperthermophiles or hyperhalophiles belonging to the phyla Crenarchaeota and Euryarchaeota, respectively; in addition, a handful of viruses have been reported to infect euryarchaeal metanogens. However, several metagenomic studies have hinted at the unexplored diversity of viruses infecting organisms of the third major archaeal phylum Thaumarchaeota (Danovaro et al., 2016; Roux et al., 2016a; Vik et al., 2017). Thaumarchaea are generally mesophilic and nearly ubiquitous in the environment, where they play important roles in nitrogen and carbon cycling (Offre et al., 2013). It has been demonstrated that, in the deep ocean, viruses more rapidly lyse archaea, mainly thaumarchaea, than bacteria, with the average abatement rate of 3.2% per day versus 1.6% per day, respectively. Furthermore, it has been estimated that virus-mediated turnover of archaea in surface deep-sea sediments accounts for up to one-third of the total microbial biomass killed, resulting in the release of ~0.3 to 0.5 gigatons of carbon per year globally (Danovaro et al., 2016). These recent findings greatly expand our appreciation of the diversity of archaeal viruses and illuminate their prominent role in the Biosphere (Danovaro et al., 2017).

2. Morphological diversity of archaeal viruses

Based on evolutionary relationship to bacterial and eukaryotic viruses, archaeal virosphere can be broadly divided into two major assemblages: (i) archaea-specific viruses and (ii) cosmopolitan archaeal viruses (Iranzo et al., 2016a). Morphological diversity of archaeal viruses has been extensively described in several recent reviews (Dellas

et al., 2014; Luk et al., 2014; Pietilä et al., 2014; Prangishvili, 2013; Prangishvili et al., 2017; Snyder et al., 2015), so we only briefly outline it below.

The archaea-specific virosphere includes viruses that are currently classified into 12 families and 1 unassigned genus (Fig. 1). These viruses display a combination of morphological and genomic features that is not observed among viruses infecting hosts in the other two cellular domains. Viruses with bottle-shaped (family *Ampullaviridae*), spindle-shaped (*Bicaudaviridae, Fuselloviridae* and genus *Salterprovirus*), coil-shaped (*Spiraviridae*) and droplet-shaped (*Guttaviridae*) virions are thus far exclusive to the Archaea (Arnold et al., 2000; Häring et al., 2005; Krupovic et al., 2014b; Mochizuki et al., 2012; Stedman, 2011; Stedman et al., 2015). Members of the families *Rudiviridae*, *Lipothrixviridae*, *Clavaviridae* and *Tristromaviridae* have filamentous virions, but unlike bacterial and eukaryotic filamentous viruses, that have single-stranded (ss) DNA and ssRNA genomes, respectively, all these archaeal viruses possess double-stranded (ds) DNA genomes (Bautista et al., 2017; Mochizuki et al., 2010; Prangishvili et al., 2013; Prangishvili and Krupovic, 2012; Ptchelkine et al., 2017; Rensen et al.,2016). Furthermore, virions of lipothrixviruses and tristromaviruses are enveloped (Kasson et al., 2017; Rensen et al., 2016), whereas all other known filamentous viruses do not contain a lipid membrane. Similarly, spherical virions of globuloviruses contain dsDNA genomes (Ahn et al., 2006; Häring et al., 2004), whereas eukaryotic viruses with similar morphology (e.g., paramyxoviruses) have ssRNA genomes. Members of the *Pleolipoviridae*, which have pleomorphic, membrane vesicle-like virions, are morphologically similar to bacterial viruses of the *Plasmaviridae* family but do not share a single gene in common (Pietilä et al., 2016), suggesting independent origins of the two virus groups. Finally, in the virions of portogloboviruses, the circular dsDNA genome appears to be condensed into a nucleoprotein filament that is wound up into a unique spherical coil, which is surrounded by a lipid membrane and

further encased by an outer icosahedral protein shell (Liu et al., 2017).

The cosmopolitan fraction of the archaeal virosphere falls into 5 officially recognized virus families (Fig. 1). Archaeal viruses with icosahedral capsids and long contractile, non-contractile, or short, stubby tails are classified into families *Myoviridae*, *Siphoviridae* and *Podoviridae*, respectively, which are further unified into the order *Caudovirales* (Senčilo and Roine, 2014). Archaeal members of the order *Caudovirales* are morphologically indistinguishable from tailed bacteriophages that comprise the bulk of this order and represent the dominant virus group in Bacteria (and generally, the most abundant type of viruses on earth). Indeed, bacterial and archaeal viruses of this group encode homologous proteins involved in virion structure, morphogenesis and maturation as well as genome packaging (Krupovic et al., 2010a). Notably, the key structural and morphogenetic proteins of tailed viruses of bacteria and archaea are also shared by eukaryotic herpesviruses (Pietilä et al., 2013; Prangishvili et al., 2017; Rixon and Schmid, 2014; Yu et al., 2017). Similarly, the family *Sphaerolipoviridae* includes both archaeal (genera *Alphasphaerolipovirus* and *Betasphaerolipovirus*) and bacterial (*Gammasphaerolipovirus*) viruses with tail-less icosahedral virions constructed from two paralogous single-jelly-roll (SJR) major capsid proteins (MCP) and carrying an internal membrane layer located underneath the protein capsid (Demina et al., 2017; Pawlowski et al., 2014). Archaeal viruses in the family *Turriviridae* are morphologically similar to sphaerolipoviruses, with the key difference being that the capsid is constructed from a single MCP which adopts a double-jelly-roll (DJR) fold (Happonen et al., 2010; Veesler et al., 2013). Viruses with DJR MCPs are found in bacteria (*Tectiviridae* and *Corticoviridae*) and are widespread in eukaryotes (big and giant viruses of the tentative order "Megavirales", adenoviruses, lavidaviruses, polintoviruses) (Abrescia et al., 2012; Krupovic and Bamford, 2008b; Krupovic and Koonin, 2015). With all these bacterial and eukaryotic viruses, turriviruses share at least 3 homologous proteins, namely, the major and minor capsid proteins, and the genome packaging ATPase of the FtsK/HerA superfamily. The similarities between the virion structures and shared gene contents strongly suggest that the cosmopolitan subset of archaeal viruses shares common ancestry with the corresponding bacterial and eukaryotic viruses.

Very recently, two new archaeal viruses were isolated, namely, *Metallosphaera* turreted icosahedral virus (MTIV) (Wagner et al., 2017) and *Methanosarcina* spherical virus (MetSV) (Weidenbach et al., 2017). Both viruses carry linear dsDNA genomes with inverted terminal repeats, but their genes generally show no similarity to those of other known viruses, with a notable exception of a DNA polymerase encoded by MetSV. Accordingly, MTIV and MetSV are likely to found two new archaeal virus families in the next future.

3. Genes of archaeal viruses

Functional annotation of archaeal virus proteins has previously shown that very few of these proteins, especially those encoded by crenarchaeal viruses, were homologous to any sequences in the public databases, be it proteins of other viruses or those of cellular organisms (Prangishvili et al., 2006). To investigate whether this conclusion still held after the massive expansion of sequence databases over the last decade, we performed a family-specific comparison of viral proteomes against the sequences available in the non-redundant protein database at NCBI (as of November 16th, 2017; Fig. 2, Table S1). With the exception of fuselloviruses and pleolipoviruses, which are frequently found integrated within the host genomes as proviruses (Fröls et al., 2007; Pietilä et al., 2016; Stedman et al., 2003; Wang et al., 2018) and hence give the false impression of encoding many "cellular" proteins (78% and 61%, respectively), the average fraction of archaeal virus proteins with no identifiable homologs at the selected E-value threshold of < 1e-5 was ~75% (~85% for crenarchaeal viruses). Notably, less than 10% of proteins encoded by members of the families *Ampullaviridae*, *Globuloviridae*, *Spiraviridae*, *Portogloboviridae* and *Tristromaviridae*

encode proteins with homologs in other viruses or cellular life forms (Fig. 2; Table S1). Even when the detection threshold is lowered to E < 1e-3, the fraction of identifiable homologs does not increase substantially (Table S1). Consequently, archaeal virus genomes remain a rich source of unknown genes, many of which could be responsible for unique mechanisms of virus-host interactions or possess unexpected properties of potential interest for biotechnological applications.

4. Genomes of archaeal viruses

All isolated archaeal viruses have DNA genomes, which can be either single-stranded or double-stranded, linear or circular (Prangishvili et al., 2017). Although putative RNA viruses were detected using metagenomics approaches in archaea-dominated hot springs of Yellowstone, the actual host of these viruses, archaeal or otherwise, remains to be identified (Bolduc et al., 2012; Bolduc et al., 2015). The genomes of archaeal viruses vary in size from 5.3 kb in clavavirus APBV1, one of the smallest known dsDNA genomes, to 143.8 kb in myovirus HGTV-1 (Fig. 1). Archaea-specific viruses typically have smaller genomes (median size of 23.9 kb, n = 69) compared to the cosmopolitan viruses (median size of 66.6 kb, n = 37), in particular when compared to the tailed viruses of the order *Caudovirales* (Fig. 1). The reasons for this disparity between archaea-specific and cosmopolitan archaeal viruses are not clear but might have to do with the life styles of the corresponding viruses (e.g., different genome replication strategies) as well as the inherent properties of the corresponding viral particles. In particular, virions of caudoviruses are known to encapsidate genomes of up to ~500 kb with packaging density approaching that of a crystal (500–550 mg/mL) and withstand internal pressures of ~25-60 atmospheres (Molineux and Panja, 2013; Rao and Feiss, 2015), i.e., 10- to 30-fold higher pressure than in a champagne bottle.

The vast majority of archaeal viruses have dsDNA genomes. Only two groups of viruses, members of the families *Spiraviridae* and *Pleolipoviridae*, have ssDNA genomes. The spiravirus ACV has the largest genome (24.9 kb) among currently known ssDNA viruses. The genome is condensed in a unique fashion: the circular ssDNA of the virus is covered with capsid proteins and the two halves of this circular nucleoprotein intertwine to form a rope-like structure which is further condensed into a spring-like coil (Mochizuki et al., 2012). Interestingly, closely related alphapleolipoviruses can have either ssDNA or dsDNA genomes indicating that there is considerable flexibility with regard to the replicative intermediates which are incorporated into the pleolipoviral particles (Pietilä et al., 2016; Senčilo et al.,2012). Archaeal viruses with linear genomes employ different solutions for protection and replication of the genome ends, including covalently closed hairpins, terminal inverted repeats and covalently-attached terminal proteins (see below).

5. Mechanisms of genome replication

With few exceptions, the mechanisms of genome replication in archaeal viruses were inferred from the recognizable genome replication-associated genes in the viral genomes. The rudivirus SIRV2 represents one of the few exceptions, where the replication mechanism has been actually investigated experimentally (Prangishvili et al., 2013). Similar to other rudiviruses, SIRV2 encodes several proteins involved in DNA replication and repair, most of which have been experimentally characterized *in vitro*, including a Holliday junction resolvase gp35 (Birkenbihl et al., 2001), a unique ssDNA-binding protein gp17 (Guo et al., 2015), an ssDNA annealing ATPase gp18 (Guo et al., 2015), a Cas4-like ssDNA nuclease gp19 (Gardner et al., 2011), a dUTPase gp23 (Prangishvili et al., 1998), and a Rep protein gp16 (Oke et al., 2011) homologous to the rolling-circle replication endonucleases (RCRE). Furthermore, one of the early-expressed virus-encoded DNA-binding proteins, gp1 (Peeters et al., 2017), interacts with the host-encoded proliferating cell nuclear antigen (PCNA) homolog and presumably
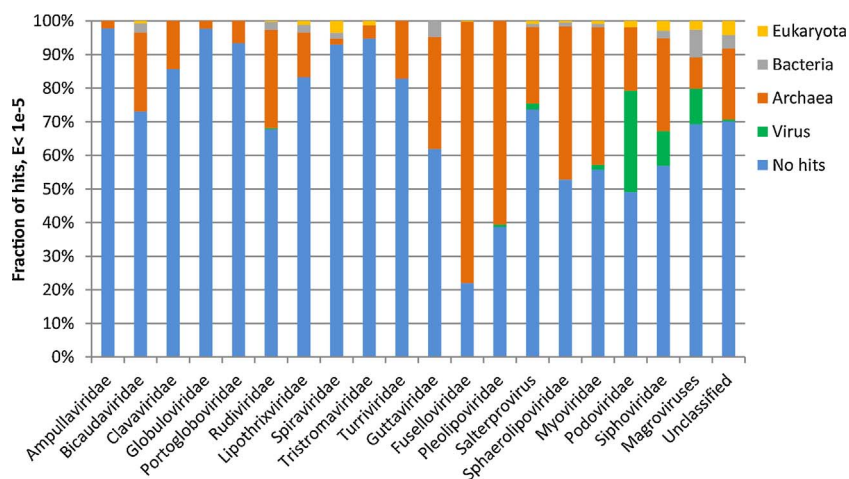
Fig. 2. Fraction of archaeal virus proteins with best-hits to proteins encoded by other viruses or cellular organisms. BLASTP searches were performed using two inclusion thresholds, E < 1e-05 and E < 1e-03 (Table S1). The plot shown in the figure is created using the former E value cut-off. The classification of the hits is based on the best-hit only. Self-hits were eliminated for the calculations. Hits to the *Sulfolobus acidocaldarius* genome, which contains a provirus closely related to turriviruses STIV and STIV2 (Anderson et al., 2017), were excluded as well. The "Unclassified" category includes Sulfolobales viruses YNP1 and YNP2 (Gudbergsdóttir et al., 2016), ANMV-1 (Paul et al., 2015), Hyperthermophilic archaeal virus 1 (Garrett et al., 2010), *Pyrococcus abyssi* virus 1 (Geslin et al., 2007) and *Thermococcus prieurii* virus 1 (Gorlas et al., 2012). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sequesters it for organization of a cellular replisome on the viral DNA (Gardner et al., 2014). Indeed, PCNA is known as a "molecular toolbelt" which interacts with multiple proteins involved in DNA replication and repair, including DNA polymerase, DNA ligase, replication factor C, etc (Pan et al., 2011). SIRV2 does not encode a DNA polymerase and instead relies on one of the four paralogous host DNA polymerases, namely Dpo1, which might be recruited for the SIRV2 genome replication via PCNA. A recent immuno-fluorescence study has shown that SIRV2 DNA synthesis is confined to a focus near the periphery of the infected cell to which PCNA and Dpo1 are sequestered (Martinez-Alvarez et al., 2017). The results of the 2D agarose gel and fluorescence microscopy analyses have spurred an exceedingly complex model of SIRV2 genome replication (Martinez-Alvarez et al., 2016). This model posits that SIRV2 employs a combination of strand-displacement, rolling-circle and strand-coupled genome replication mechanisms, which generate multimeric, highly branched 'brush-like' intermediates reaching a size of > 1200 kb (∼34 genome units). However, currently, it is not clear whether all these modes of replication are actually deployed during the course of the SIRV2 infection cycle. Unexpectedly, the gene encoding the viral Rep (gp16), the predicted key player of the proposed mode of replication, is one of the most poorly expressed SIRV2 genes as shown by several methods (Kessler et al., 2004; Okutan et al., 2013; Quax et al., 2013). Thus, although the major host- and virus-encoded components involved in SIRV2 genome replication have been elucidated, the exact mechanism of replication and orchestration of the molecular players remain obscure.

More generally, viruses with small to moderate-sized genomes (5–50 kb) typically carry genes for strategic components of the replication machinery, which enable them to hijack the replisome of the host. The exceptions to this general pattern are four groups of viruses − bottle-shaped ampullaviruses, spindle-shaped salterprovirus His1, pleolipovirus His2 and the recently discovered spherical MetSV − encoding protein-primed DNA polymerases (Bath et al., 2006; Peng et al., 2007; Weidenbach et al., 2017), which could be largely sufficient for genome replication, although none of these proteins has been studied experimentally. Global analysis of dsDNA virus genomes across all three domains of life has shown that helicases are the most common among the virus-encoded components of the DNA replication machinery (75% of the analyzed genomes) (Kazlauskas et al., 2016). Indeed, many viruses infecting euryarchaea encode replicative minichromosome-maintenance (MCM) helicases (Krupovic et al., 2010a). Phylogenetic analysis shows that the MCM helicases have been recruited by different viruses from their respective hosts on multiple independent occasions (Krupovic et al., 2010b). Interestingly, in methanogenic archaea of the order Methanococcales, the cellular *mcm* gene apparently underwent accelerated evolution in the course of the host-to-virus-to-host transfer loop (i.e., the gene was captured by a virus from the host, and following

the accelerated evolution in the viral genome, was reintegrated into the host genome to replace the original *mcm* gene). Some viruses infecting haloarchaea encode their own PCNA proteins, as in the case of the myovirus PhiCh1 (Klein et al., 2002), or homologs of the archaeal Orc1/Cdc6 replication initiators (Krupovic et al., 2010a; Liu et al., 2015; Pagaling et al., 2007). Notably, the spindle-shaped fuselloviruses and bicaudaviruses as well as the droplet-shaped guttaviruses encode AAA+ ATPases homologous to DnaA, the protein that triggers initiation of DNA replication in bacteria (Iranzo et al., 2016a; Koonin, 1992). However, given the broad range of functions associated with AAA+ ATPases, the involvement of this protein in viral genome replication should not be taken for granted. Finally, certain members of the families *Pleolipoviridae* and *Sphaerolipoviridae* encode diverse RCRE proteins (Senčilo et al., 2012; Wang et al., 2016). In the case of sphaerolipovirus SNJ1, the RCRE protein has been experimentally shown to be essential for viral genome replication and the corresponding gene was successfully used to develop shuttle vectors (Wang et al., 2016).

Haloarchaeal siphoviruses and myoviruses with large genomes (> 100 kb) encode nearly complete replisomes (Senčilo et al., 2013). For instance, haloarchaeal virus HVTV-1, with a genome of 102 kb, encodes its own DNA polymerase, DNA clamp and its loader, archaeo-eukaryotic primase, RNase HI and a putative replicative helicase (Kazlauskas et al., 2016). Remarkably, uncultivated members of the *Caudovirales*, dubbed Magroviruses and predicted to infect Marine Group II Euryarchaeota, encode even more complete ensembles of genome replication proteins, additionally including genes for RadA-like ATPase, ssDNA-binding protein, and ATP-dependent DNA ligases (Philosof et al., 2017). Overall, archaeal viruses appear to follow the general trend observed among dsDNA viruses, whereby viruses with larger genomes approach self-sufficiency for genome replication (Kazlauskas et al., 2016). Indeed, the tailed archaeal viruses appear to be minimally dependent on the host replication machinery. Sequence comparisons and phylogenetic analysis generally suggest that the replication proteins encoded by archaeal viruses share common ancestry with the archaeal counterparts but the exact routes of their evolution await detailed investigation (Kazlauskas et al., 2016; Krupovic et al., 2010b; Philosof et al., 2017; Senčilo et al., 2013). Interestingly, one clade of the archaeal family B DNA polymerases, namely PolB3 in halophilic and methanogenic archaea (Makarova et al., 2014), apparently has been recruited from archaeal members of the *Caudovirales* (Kazlauskas et al., 2016).

It should be noted, however, that for many archaeal viruses (families *Clavaviridae*, *Tristromaviridae*, *Portogloboviridae*, *Globuloviridae*, *Spiraviridae*, *Turriviridae*, *Sphaerolipoviridae*, *Lipothrixviridae*), replication proteins could not be predicted and the mechanisms of the genome replication remain enigmatic. For instance, linear dsDNA genomes of lipothrixviruses (Pina et al., 2014) and alphasphaerolipoviruses (Porter
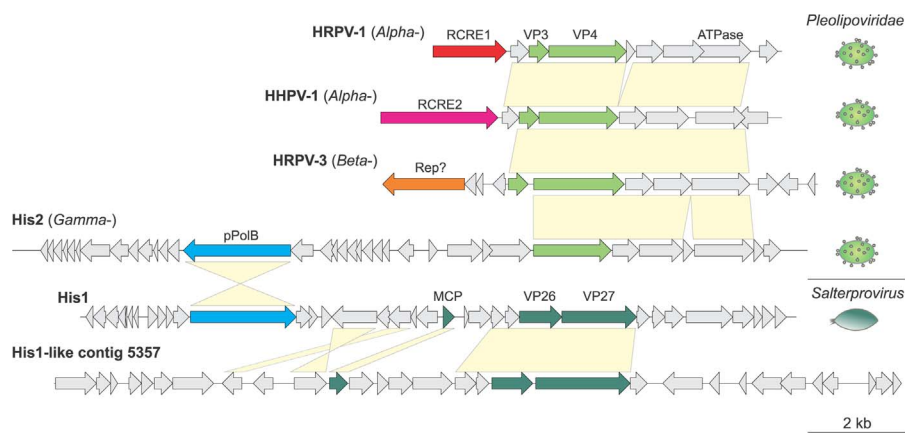
Fig. 3. Genome comparison of pleoloviruses and salterproviruses. The genome schematics are drawn roughly to scale (shown at the bottom of the figure). The genes are shown by block arrows indicating the direction of transcription. The genes encoding for genome replication-associated proteins and virion proteins are color coded: structural proteins of pleolipoviruses and salterproviruses are shown in light and dark green, respectively; two families of rolling circle replication initiation endonucleases (RCRE1 and RCRE2) are colored red and magenta, respectively; uncharacterized Rep is shown in orange; protein-primed family B DNA polymerase (pPolB), light blue. Genes shared between viruses are indicated by yellow shading. Morphologies of the corresponding viruses are shown on the right of the figure. MCP, major capsid protein. Genome accession numbers: HRPV-1, FJ685651; HHPV-1, GU321093; HRPV-3, JN882265; His2, AF191797; His1, AF191796; His1-like contig 5357, LFUF01004316 (BioProject PRJNA287316). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and Dyall-Smith, 2008) contain covalently-attached terminal proteins, a hallmark of genomes replicated by protein-primed family B DNA polymerases (Redrejo-Rodríguez and Salas, 2014). However, viruses of these groups do not encode recognizable DNA polymerases, implying that they might exercise novel mechanism(s) of genome replication. This was confirmed in the case of lipothrixvirus AFV1, for which genome replication has been suggested to start by a D-loop formation and proceed by strand displacement, whereas termination apparently relies on recombination through the formation of terminal loop-like structures (Pina et al., 2014). However, no proteins involved in this peculiar replication mechanism have been identified thus far. Notably, members of the same family in some cases encode non-orthologous or even non-homologous replication proteins. This pattern is most obvious in the family *Pleolipoviridae* (Fig. 3), where viruses of the genus *Alphapleolipovirus* encode two non-orthologous families of RCRE enzymes and presumably replicate by the rolling circle mechanism. In members of the *Betapleolipovirus* genus, the RCRE is replaced by a gene showing no resemblance to other known replication protein, whereas gamma-pleolipovirus His2 encodes a protein-primed DNA polymerase, with the closest homologue in a spindle-shaped salterprovirus His1 (Bath et al., 2006; Pietilä et al., 2016; Senčilo et al., 2012). Similarly, members of the genera *Alphasphaerolipovirus* (Demina et al., 2017; Jaakkola et al., 2012; Porter et al., 2005; Porter et al., 2013) and *Betasphaerolipovirus* (Zhang et al., 2012) have linear and circular genomes, respectively, and apparently replicate via distinct mechanisms (Pawlowski et al., 2014; Wang et al., 2016). Such modular evolution of viral genomes, where genome replication and virion structure modules have distinct histories, is a common theme in virus evolution (Koonin et al., 2015; Krupovic and Bamford, 2009) and is particularly prominent in different families of bacterial viruses (Krupovic and Bamford, 2007; Krupovic and Bamford, 2010; Weigel and Seitz, 2006).

6. Structural genomics of archaeal viruses

As mentioned above, most of the proteins encoded by archaeal viruses are refractory to informative bioinformatic analysis because of the lack of significant similarity to sequences in the public databases (Fig. 2), even when most sensitive of the available sequence analysis methods are employed. Given that tertiary protein structures typically outlast the conservation of the protein sequences (Chothia and Lesk, 1986), several groups have undertaken structural genomics projects in order to elucidate the functions of the enigmatic proteins encoded by hyperthermophilic archaeal viruses. As of this writing (November 11, 2017), high resolution structures are available for 43 proteins (59 structures) of hyperthermophilic archaeal viruses from the families *Fuselloviridae*, *Bicaudaviridae*, *Rudiviridae*, *Lipothrixviridae*, *Globuloviridae*, *Clavaviridae*, *Turriviridae* and one unclassified spindle-shaped virus

PAV1 (Table S2). The structures were determined using X-ray crystallography, NMR spectroscopy and, more recently, cryo-electron microscopy. More than a quarter of these proteins display unique folds with no structural homologs in public databases, providing limited information on the putative functions of the corresponding viral proteins (reviewed in (Dellas et al., 2014; Krupovic et al., 2012)). Many of these proteins could modulate specific aspects of virus-host interaction, in particular, inactivation of the host defense mechanisms, such as CRISPR-Cas. Nevertheless, in some cases, structural studies were crucial for inferring the possible function. For example, the X-ray structure of the protein ORF119 from the rudivirus SIRV1 revealed a fold characteristic of RCRE proteins (Oke et al., 2011), which is the signature replication protein in viruses with ssDNA genomes (Krupovic, 2013). The predicted nicking activity of this protein has been subsequently demonstrated *in vitro* (Oke et al., 2011), despite the fact that expression of a close homolog in SIRV2 (gp16) could not be detected *in vivo* (see above). Other notable examples when high resolution structures provided functional clues include the putative glycosyltransferase of turrivirus STIV (Larson et al., 2006) as well as the PD-(D/E)XK family nuclease of fusellovirus SSV-RH (Menon et al., 2010) and a novel-fold nuclease of lipothrixvirus AFV1 (Goulet et al., 2010a). In all these cases, sequence-based analyses were not effective for functional protein annotation.

One of the research directions which particularly benefitted from structural genomics efforts is the study of transcription regulation by archaeal viruses. Archaeal viruses frequently encode transcription factors with diverse DNA-binding motifs, including ribbon-helix-helix (RHH), winged helix-turn-helix (wHTH) and zinc-fingers (Peeters et al., 2013; Prangishvili et al., 2006), which due to their generally small size are amenable to crystallization and NMR. Structures of 10 such putative transcription factors have been determined thus far and some of them have been experimentally characterized revealing intricate patterns of transcriptional control during infection (Fusco et al., 2015b; Guilliere et al., 2009; Peixeiro et al., 2013; Selb et al., 2017). These studies have validated the conclusions initially reached by protein sequence analysis (Prangishvili et al., 2006), namely, that whereas the basal transcription machinery of archaea, in terms of the structure of promoters and subunit composition of the RNA polymerase, closely resembles the eukaryotic counterparts (Peeters et al., 2013; Werner and Grohmann, 2011), many of the transcription factors encoded by archaea and their viruses are bacterial-like.

Furthermore, structural studies have provided valuable insights into the evolution of archaeal viruses. High resolution structures of the major capsid protein (MCP) and the penton protein of STIV have shown that these proteins are homologous to the corresponding proteins of icosahedral viruses infecting bacteria and eukaryotes (Veesler et al., 2013). Thus, it has been proposed that STIV belongs to an ancient viral

lineage that could predate the divergence of cellular organisms into the three contemporary domains, Archaea, Bacteria and Eukarya (Krupovic and Bamford, 2011). By contrast, the MCP of rudiviruses displays a unique structural fold, which is only found in members of another archaeal virus family, the *Lipothrixviridae* (DiMaio et al., 2015; Goulet et al., 2009; Kasson et al., 2017; Szymczyna et al., 2009). The latter finding combined with the shared genomic content led to the unification of rudiviruses and lipothrixviruses into a new order of archaeal viruses, the *Ligamenvirales* (Prangishvili and Krupovic, 2012). Structures of two virion proteins were also resolved for the bicaudavirus ATV (Felisberto-Rodrigues et al., 2012; Goulet et al., 2010b); both proteins display unique folds observed neither in archaeal viruses from other families nor elsewhere in the virosphere (Krupovic and Koonin, 2017). Interestingly, a paralog of the major structural protein of ATV, ORF145, has been recently identified as a global transcriptional repressor of ATV genes. The protein does not bear canonical DNA-binding motifs; instead, it adopts a unique helical-bundle fold (Goulet et al., 2010b). The ORF145 protein forms a high-affinity complex with the host RNA polymerase by inserting into the DNA-binding channel, thereby counteracting the formation of transcription pre-initiation complexes and repressing transcription initiation as well as elongation (Sheppard et al., 2016).

7. Functional genomics of archaeal viruses

Studies on bacterial and eukaryotic viruses have benefited from the availability of well-established genetic tools developed for the respective hosts and, more generally, from the broad knowledge base on the host biology. This, unfortunately, has not been the case for most of the archaeal virus-host systems, although new genetic tools are being developed for an increasing number of archaea and their viruses (Iverson and Stedman, 2012; Iverson et al., 2017; Jaubert et al., 2013; Selb et al., 2017; Wang et al., 2016; Wirth et al., 2011). However, during the past few years, high-throughput functional genomics approaches have been adapted to study archaeal viruses, yielding valuable information on their biology. Below we summarize recent advances in the field obtained using various high-throughput "omics" methods, including DNA microarrays, whole-transcriptome sequencing (RNA-seq) and large-scale proteomic analyses. Most of this work has focused on four groups of *Sulfolobus*-infecting viruses, namely, rudiviruses (SIRV2), fuselloviruses (SSV1 and SSV2), turriviruses (STIV) and bicaudaviruses (STSV2).

Transcriptomics. Although structural genomics efforts shed some light on the functional potential and evolution of archaeal viruses, the details on virus-host interplay remained obscure. The first molecular insights into the course of events within the host cell during archaeal virus infection were provided by whole-genome transcriptomic analyses employing DNA microarray (Fröls et al., 2007; Fusco et al., 2015a; Okutan et al., 2013; Ortmann et al., 2008; Ren et al., 2013) and RNA-seq (Leon-Sobrino et al., 2016; Quax et al., 2013) technologies. For all studied viruses, i.e., STIV, SSV1, SSV2, STSV2 and SIRV2, a temporal control of viral gene expression has been observed, albeit the extent of this temporality was highly variable for different viruses. For example, upon UV induction, SSV1 lysogens show tight temporal transcriptional regulation of viral genes, resembling the regulation strategy employed by many bacterial and eukaryotic viruses (Fröls et al., 2007). Temporally regulated gene expression was also observed during the infection (rather than UV induction) of *Sulfolobus* cells with SSV2 (Ren et al., 2013). Notably, the transcriptional control during infection substantially differed from that observed during UV induction, with transcripts of the major structural genes being expressed early during infection but late in the case of UV induction (Fröls et al., 2007; Ren et al., 2013). By contrast, lytic viruses SIRV2 and STIV appeared to show little temporal regulation (Ortmann et al., 2008; Quax et al., 2013). The RNA-seq analysis of SIRV2 has shown that the expression of the majority of viral genes starts immediately after infection and subsequently steadily increases throughout the infection cycle (Quax et al., 2013).

Analysis of the host gene expression during virus infection provided valuable insights into virus-host interactions in archaea. Interestingly, the host response to infection varied greatly among the archaeal viruses, both in terms of the number of differentially expressed genes and the direction of regulation (i.e., up- versus downregulation). Only a handful of host genes were found to be differentially expressed during the induction of SSV1 lysogens (Fröls et al., 2007), whereas SSV2 induced strong host response, including transcriptional activation of the CRISPR loci and *cas* genes (Fusco et al., 2015a). Similarly, in the case of SIRV2 infection, 30% to 50% of the cellular genes showed changes in the expression level, including significant upregulation of transposases and antiviral defense genes, such as CRISPR-Cas and toxin-antitoxin systems (Quax et al., 2013). The type I CRISPR-Cas and various toxin-antitoxin systems were also upregulated during STSV2 infection (Leon-Sobrino et al., 2016). These findings indicate that, although most studied viruses induce an antiviral response, some, such as SSV1, are largely invisible to the host defense. Notably, the same genes that tend to be down-regulated in the case of UV induction of SSV1 appear to be upregulated during STIV infection (Fröls et al., 2007; Ortmann et al., 2008), suggesting that propagation of different archaeal viruses requires distinct − even if partially overlapping − subsets of host functions. For example, crenarchaeal cell division (*cdv*) genes that encode homologs of the eukaryotic ESCRT machinery (Moriscot et al., 2011; Samson et al., 2008) are downregulated during SIRV2 and STSV2 infections (Leon-Sobrino et al., 2016; Quax et al., 2013) but are upregulated in the case of STIV infection (Ortmann et al., 2008), suggesting an important role of the Cdv proteins in the infection cycle of the latter virus (see below).

Transcriptomics also provided clues as to which host proteins are likely to be involved in genome replication of SSV1, SSV2, STSV2 and STIV. In the case of both fuselloviruses and STIV, Orc1/Cdc6 and reverse gyrase genes were upregulated, whereas SSV2 additionally boosted expression of replicative MCM helicase, Dpo1, PCNA and topoisomerase VI (Fröls et al., 2007; Ortmann et al., 2008; Ren et al., 2013). By contrast, STSV2 infection leads to increased transcription of a replication initiation protein, WhiP (Winged-Helix Initiator Protein) (Leon-Sobrino et al., 2016), which is unrelated to the Orc/Cdc6 and is specifically associated with one of the three *ori* sites of *Sulfolobus* (Robinson and Bell, 2007). Furthermore, STSV2-infected cells showed increased transcription of Holliday junction resolvase, DNA topoisomerase I and several DNA repair enzymes, whereas expression of reverse gyrase, unlike for other viruses, was downregulated (Leon-Sobrino et al., 2016). Similarly, infection with different viruses resulted in distinct patterns of differential regulation of genes encoding proteins involved in transcription (Leon-Sobrino et al., 2016; Ortmann et al., 2008; Ren et al., 2013). Generally, transcriptomics studies have shown that each virus has a unique effect on the host gene expression although in many cases, it is impossible to distinguish whether the changes in expression levels are mediated by the virus as part of the host take-over strategy or by the host, in response to viral infection.

Proteomics. Large-scale proteomic analysis of infected cells using one- and two-dimensional differential gel electrophoresis (DIGE) coupled with protein identification by mass spectrometry and activity-based protein profiling represents another powerful approach that can provide important insights into virus-host interactions. Thus far, this methodology has been only employed to study the STIV infection in two *S. solfataricus* strains, P2 and P2-2-12, that substantially differ with respect to susceptibility to STIV (Maaty et al., 2012a; Maaty et al., 2012b). In the highly susceptible P2-2-12 strain, only 10 host proteins changed in abundance. By contrast, 71 host proteins representing 33 different cellular pathways were affected during the infection of the poorly susceptible strain P2 (Maaty et al., 2012a; Maaty et al., 2012b), shedding light on the basis of different susceptibility to infection of closely related *Sulfolobus* strains. Most notably, among the highly upregulated proteins were components of the CRISPR-Cas system as well

as Cdv proteins involved in cell division, again suggesting that the latter proteins play an important role in the infection cycle of STIV. The critical role of the archaeal ESCRT machinery during the assembly of STIV virions, in particular acquisition of the virion membrane located underneath the protein capsid, has bene recently confirmed in direct experiments (Snyder et al., 2013).

Notably, several high-throughput top-down and bottom-up proteomics studies have recently focused on the post-translational modifications (PTM) in *Sulfolobus* strains, revealing massive protein methylation and N-terminal acetylation as well as glycosylation of cellular surface proteins (Chu et al., 2016; Palmieri et al., 2013; Vorontsov et al., 2016); studies on the PTM changes in the course of infection with various *Sulfolobus* viruses might uncover new facets of archaeal virus biology.

8. Metagenomics of archaeal viruses

The advent of high-throughput sequencing and advanced bioinformatics has ushered archaeal virology into a new era of discoveries. Metagenomics has enabled researchers not only to probe the extent of genetic diversity of known archaeal virus groups from terrestrial hot springs (Bolduc et al., 2015; Gudbergsdóttir et al., 2016) and hypersaline environments (Adriaenssens et al., 2016; Emerson et al., 2013) but also provided a glimpse at unknown viruses, particularly those infecting oceanic archaea that are recalcitrant to cultivation (Danovaro et al., 2016; Nigro et al., 2017; Nishimura et al., 2017; Paul et al., 2015; Philosof et al., 2017; Roux et al., 2016a).

Although many of the virus genomes assembled from hot spring samples represented the established families *Rudiviridae* (Erdmann et al., 2014; Gudbergsdóttir et al., 2016; Servin-Garciduenas et al., 2013b), *Fuselloviridae* (Servin-Garciduenas et al., 2013a), *Lipothrixviridae* (Gudbergsdóttir et al., 2016), *Bicaudaviridae* (Gudbergsdóttir et al., 2016; Hochstein et al., 2016) and *Ampullaviridae* (Gudbergsdóttir et al., 2016), several viruses could not be assigned to existing families (Garrett et al., 2010; Gudbergsdóttir et al., 2016), indicating that our knowledge of the viral diversity in extreme geothermal environments remains incomplete. Indeed, metagenomic analysis of a single hot spring in the Yellowstone National Park revealed 110 viral groups, among which only 7 (6.3%) represented known archaeal viruses from acidic hot spring environments (Bolduc et al., 2015).

In hypersaline environments, the vast majority of complete or near-complete genome sequences were obtained for archaeal viruses of the order *Caudovirales* (Garcia-Heredia et al., 2012; Santos et al., 2007), although 3 potentially complete genomes were also assembled for novel members of the genus *Salterprovirus* (Adriaenssens et al., 2016). Around 50 complete genomes of head-tailed haloviruses have been obtained through sequencing of fosmid libraries into which DNA extracted from the viral assemblages was cloned (Garcia-Heredia et al., 2012; Martinez-Garcia et al., 2014; Santos et al., 2007). This approach has provided the first genomic insights into viruses of *Haloquadratum walsbyi*, a square-shaped archaeon that dominates hypersaline environments (Dyall-Smith et al., 2011), and nanohaloarchaea, a fast-evolving lineage of nano-sized halophilic archaea (Narasingarao et al., 2012). Further insights into functional genomics of these uncultivated haloviruses have been obtained using the metatranscriptomics approach, which revealed dynamic virus-host interactions in hypersaline settings (Santos et al., 2011). Comparison of all available hypersaline viromes has demonstrated worldwide distribution of the haloarchaeal members of the *Caudovirales*, with groups of related viruses detected in salt ponds on different continents (Roux et al., 2016b). The same conclusion has been also reached using large-scale virus isolation and cross-infection trials, which suggested that hypersaline environments worldwide effectively comprise a single habitat (Atanasova et al., 2012). By contrast, comparative genomic and phylogenomic analyses of *Sulfolobus islandicus*-infecting rudiviruses revealed a pronounced biogeographical signal, both at a global and local scale, with more closely related strains sharing more of their variable gene content (Bautista et al., 2017). A similar biogeographic pattern has been also observed in *S. islandicus* populations around the globe (Reno et al., 2009) as well as with *Sulfolobus*-infecting spindle-shaped fuselloviruses (Held and Whitaker, 2009). These studies suggest that gene flow is limited among spatially partitioned hyperthermophilic archaeal viruses, but not among archaeal viruses thriving in spatially open habitats.

Arguably, so far, the major contribution of metagenomics to the study of archaeal viruses was the characterization of the viral diversity associated with ubiquitous, environmentally important groups of archaea, many of which remain uncultivated. One of these is Marine Group II Euryarchaeota, which are particularly abundant in the oceanic surface waters (Karner et al., 2001). Over 50 complete and near-complete genomes (~90 kb in size; Fig. 1) have been assembled for viruses, dubbed Magroviruses, associated with these archaea (Nishimura et al., 2017; Philosof et al., 2017). Remarkably, magrovirus genomes are globally widespread in the marine environment, third only to the SAR11 phages and cyanophages, suggesting a prominent role for magroviruses in controlling the turnover of their hosts in global ocean (Philosof et al., 2017). Thaumarchaeota is another group of archaea that is ubiquitous in aquatic and terrestrial environments (Offre et al., 2013), but for which viruses have not been identified, except for a single provirus in the genome of a soil thaumarchaeon *Nitrososphaera viennensis* (Krupovic et al., 2011). Single-cell genomics and fosmid sequencing yielded the first two putative thaumarchaeal virus genomes from marine environments (Chow et al., 2015; Labonte et al., 2015). Uncultured viruses have been also described for nano-sized archaea known as ARMAN (Burstein et al., 2017; Paul et al., 2015). More recently, a dedicated bioinformatics tool, MArVD (for Metagenomic Archaeal Virus Detector), has been developed for mining the metagenomic datasets for archaeal virus genomes (Vik et al., 2017). Application of this tool to 10 viromes from the Eastern Tropical North Pacific oxygen minimum zone revealed 43 new putative archaeal virus genomes and large genome fragments (10–31 kb), which were suggested to represent 6 novel candidate archaeal virus genera. Co-occurrence analysis has suggested that these viruses are associated with archaea of the class Thermoplasmata (Belmar et al., 2011), another archaeal group for which viruses have not been isolated thus far.

Interestingly, all of the uncultivated viruses from mesophilic environments mentioned above are undisputable members of the order *Caudovirales*, as indicated by the conservation of signature genes, such as large terminase subunit, portal protein and HK97-like MCP. Whether any of these ecologically important, oceanic archaea are infected by archaea-specific viruses, remains unknown. Notably, however, electron microscopy studies have shown that ARMAN members are hosts to spindle-shaped and rod-shaped viruses, although their genomes have not yet been sequenced (Comolli et al., 2009). Furthermore, both morphotypes (as well as *Caudovirales*) have been recently reported in fluid samples recovered from boreholes 117–292 m deep into the ocean basement, where the archaeal fraction was represented by members of the Archaeoglobi, Thermococci, Bathyarchaeota and the marine benthic group E (Nigro et al., 2017).

Exploration of the diversity of uncultured archaeal viruses, especially those from mesophilic environments, is only commencing. Nevertheless, the obtained results have already significantly advanced our appreciation of the archaeal virosphere, revealing several interesting features of these viruses, including nearly complete genome replication machineries and diversity of chaperonins in magroviruses (Marine et al., 2017; Nishimura et al., 2017; Philosof et al., 2017); a diversity-generating retroelement, which uses mutagenic reverse transcription and retro-homing with a potential to generate up to 10^{18} variants of the tail fiber ligand-binding domain in ANMV-1 (Paul et al., 2015); and CRISPR arrays, albeit without associated *cas* genes, in certain uncultivated haloarchaeal viruses (Garcia-Heredia et al., 2012) as well as in *Acidianus* tailed spindle virus, which has been discovered by culture-independent methods and subsequently propagated in pure

culture in the laboratory settings (Hochstein et al., 2016). Development of further bioinformatics resources, such as MArVD (Vik et al., 2017), and establishment of a general framework for classification of 'metagenomic' viruses (Simmonds et al., 2017) will undoubtedly stimulate rapid progress in the field, which in combination with the culture-dependent efforts, will provide a more complete picture on the global impact of archaeal viruses in the Biosphere.

9. Evolutionary genomics of archaeal viruses

Due to horizontal gene transfer and high mutation rates in viral genomes as well as the lack of universal virus genes, standard phylogenetic methods have limited utility for studying deep evolutionary connections between distantly related virus groups. Indeed, evolution of viruses is more faithfully represented as a network rather than a tree (Iranzo et al., 2017). Thus, a bipartite network analysis, in which viral genomes are connected through shared gene families, has been recently applied to investigate the evolutionary relationships of different archaeal virus groups to each other as well as to viruses of bacteria and eukaryotes.

Place of archaeal viruses in the global virosphere. Global analysis of the dsDNA virus genomes under the framework of bipartite networks has uncovered a robust hierarchical modularity in the dsDNA virosphere (Iranzo et al., 2016b). This hierarchical organization rests on 3 classes of conserved genes: i) *hallmark genes*, which encode key proteins involved in genome replication and virion formation, and are shared by overlapping sets of diverse viruses, providing the broad connectivity in the virus world; ii) *connector genes* that are shared by multiple subgroups within a group, and iii) *signature genes* that are highly specific to sets of related viruses within a module. The network of dsDNA viruses included 19 modules that could be further grouped into 5 major supermodules (Iranzo et al., 2016b). The cosmopolitan subset of archaeal viruses was distributed between 2 supermodules containing bacterial and eukaryotic viruses (Fig. 4A). Specifically, the archaeal members of the *Caudovirales* were placed into a supermodule together with the tailed bacteriophages and eukaryotic herpesviruses, whereas turriviruses and sphaerolipoviruses were included into a supermodule containing bacterial viruses of the families *Tectiviridae* and *Corticoviridae* as well as eukaryotic viruses of the unofficial order "Megavirales", adenoviruses, virophages (*Lavidaviridae*), and various elements related to Polintons (Iranzo et al., 2016b), fully recapitulating the previous results of more focused comparative genomic analyses and structural studies on the corresponding viruses (Abrescia et al., 2012; Gil-Carton et al., 2015; Koonin and Krupovic, 2017; Krupovic et al., 2010a; Pawlowski et al., 2014; Pietilä et al., 2013; Rixon and Schmid, 2014; Strömsten et al., 2005; Veesler et al., 2013; Yu et al., 2017). By contrast, all archaea-specific virus groups formed a distinct supermodule, which was largely disconnected from the rest of the virosphere (Fig. 4A), suggesting independent origins of these viruses and the existence of stronger barriers to horizontal gene transfer between viruses of archaea and bacteria/eukaryotes.

Relationships among archaeal viruses. A detailed dissection of the archaeal virus network has revealed strong modularity (Iranzo et al., 2016a), with 11 distinct modules, whereas members of the families *Tristromaviridae* (Rensen et al., 2016) and *Clavaviridae* (Mochizuki et al., 2010), which do not share genes with other archaeal viruses, remained disconnected (Fig. 4B). Members of the *Caudovirales*, including the uncultivated Magroviruses, represent 4 modules, which are densely interconnected by several connector genes and four hallmark genes encoding the morphogenetic toolkit of these viruses, i.e. the HK97-like MCP, the large termnase subunit, portal protein, and capsid maturation protease. By contrast, the archaea-specific virosphere split into 6 modules, which were remarkably sparsely connected, sharing only a few hallmark genes, most of which were not implicated in the central virus functions, such as genome replication and virion assembly. Among the most widely distributed hallmark genes are those encoding

RHH domain-containing transcription factors and glycosyltransferases of the GT-B family (Fig. 4B), both of which are at the interface of virus-host interaction and could be independently acquired from the respective hosts by viruses within each module or spread between viruses horizontally. Consistent with this possibility, the RHH protein is the only connection between the cosmopolitan turriviruses and the archaea-specific viruses (Fig. 4B). Given that genome replication proteins – most of which are unidentifiable (see above) – and major virion proteins of archaea-specific viruses are mostly specific to particular modules, it has been suggested that most of the archaea-specific viral groups are evolutionarily distinct, i.e., do not share common viral ancestors (Iranzo et al., 2016a).

Links to non-viral mobile genetic elements. Although there was little connectivity between different modules of archaea-specific viruses, 5 of the 11 modules displayed strong links to various non-viral MGE. Notably, the MGE were primarily included into corresponding modules through shared genes encoding major genome replication proteins. In particular, bottle-shaped ampullaviruses are connected to casposons via pPolB (Krupovic et al., 2014a); betasphaerolipovirus SNJ1 shares the RCRE gene with euryarchaeal plasmids pZMX101 and pC2A (Wang et al., 2016); fuselloviruses are connected to small *Sulfolobus* plasmids via several genes, including the one for the DnaA-like ATPase (Wang et al., 2007); certain alphapleolipoviruses, such as HRPV-1 (Fig. 3), encode RCRE closely related to those encoded by several euryarchaeal plasmids (Gorlas et al., 2013). The most extensive gene sharing is observed between the spindle-shaped virus PAV1 (Geslin et al., 2007; Geslin et al., 2003) and a group of plasmids of Thermococcales (Fig. 5): nearly half of the PAV1 genome, carrying genes for several DNA-binding proteins and ABC ATPase, has been apparently recruited from plasmids (Krupovic et al., 2013), whereas the other half encoding major virion proteins has been inherited from spindle-shaped viruses of Thermococcales, such as TPV1 (Gorlas et al., 2012).

The paucity of connections in the archaeal virus network, the lack of connectivity with the global virosphere and the multiple, independent links between archaeal viruses and non-viral elements (Fig. 4) suggest that at least some of the archaea-specific virus groups could have evolved from non-viral MGE through acquisition of genetic determinants for virion formation from various sources. The latter hypothesis appears to be supported by the recent discovery of a haloarchaeal plasmid pR1SE which is transferred between cells within regularly-shaped membrane vesicles containing several plasmid-encoded proteins (Erdmann et al., 2017). Remarkably, akin to virus infection, the released vesicles are capable of 'infecting' plasmid-free haloarchaeal strains which then gain the ability to produce plasmid-containing vesicles. Although the actual origins of the morphogenetic modules remain unknown for most archaeal viruses, it has been recently shown that one of the major nucleocapsid proteins of a tristromavirus TTV1 has been exapted (i.e., refunctionalized) from a Cas4-like nuclease (Krupovic et al., 2015). A similar route of refunctionalization of ancestral cellular proteins might have played an important role in the evolution of the entire archaeal virosphere. This route of evolution seems all the more plausible because recruitment of bona fide cellular proteins to function as major structural components of virions is not unique to archaea but has been shown to recurrently occur throughout the history of the virosphere (Krupovic and Koonin, 2017). It is also of note that many of the structural proteins encoded by archaeal viruses have simple, often α-helical folds (DiMaio et al., 2015; Hochstein et al., 2015; Kasson et al., 2017; Krupovic et al., 2014b; Ptchelkine et al., 2017), which could have originated de novo at different stages of evolution. Thus, the history of the archaeal virosphere appears to combine descent from the ancient virosphere, in particular for the cosmopolitan viruses, and more recent emergence of novel virus groups from non-viral MGE. However, it cannot be ruled out that some of the archaea-specific virus groups date back to the last common cellular ancestor or even beyond but have been lost in the other cellular domains (Prangishvili, 2015). Furthermore, evolution of MGE is hardly a
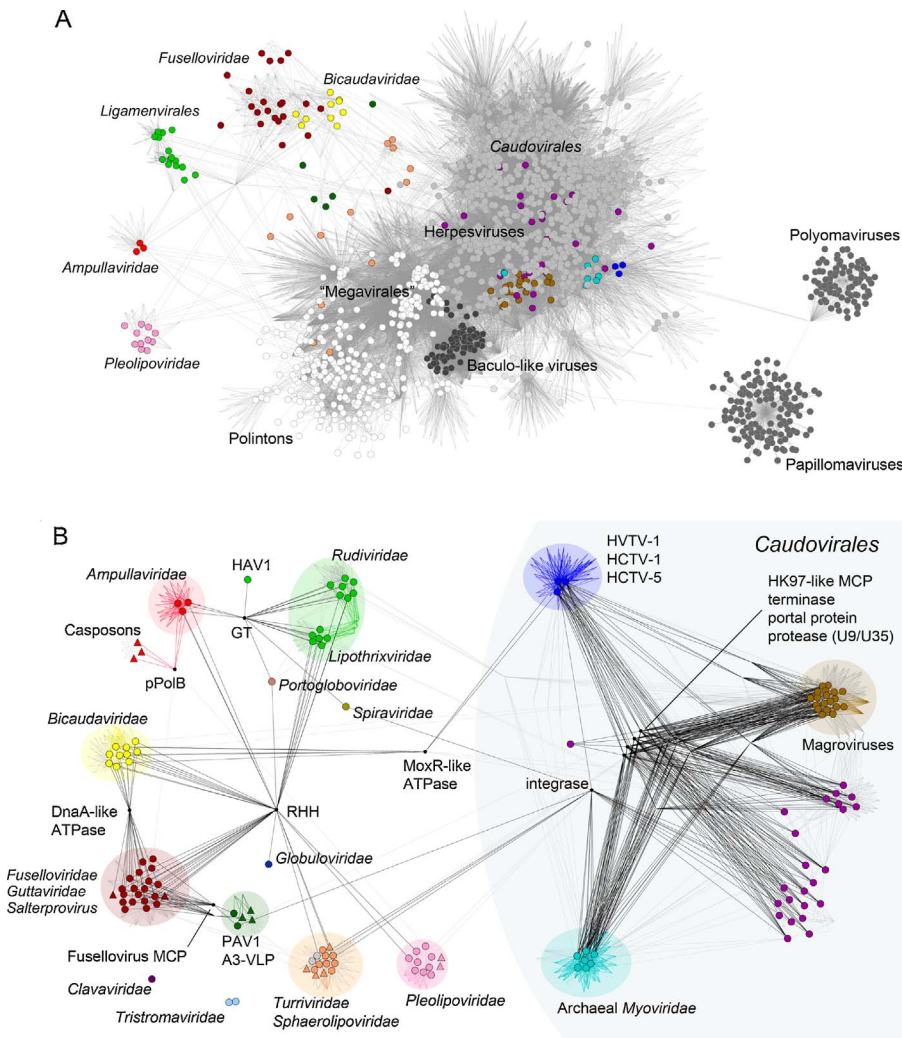
Fig. 4. Comparative genomics of archaeal viruses, represented as a network of genomes and shared genes.

A. The gene sharing network of the entire dsDNA virosphere highlights the distinct nature of the archaeal-specific viruses. Each viral genome is represented as a circle (colored circles correspond to archaeal viruses). Straight lines (edges) connect each virus with the gene families present in its genome. Consequently, edge junctions denote the presence of a gene family shared by multiple viruses. In this bipartite representation, closely related genomes are connected indirectly through a large number of shared gene families. The network has been projected onto a plane to show groups of similar genomes close to each other. Different shades of grey represent the four supermodules of the dsDNA virosphere, from lighter to darker: (i) the double-jelly roll fold MCP supermodule (including "Megavirales", *Adenoviridae* and polintons, among others), (ii) the HK97-like MCP supermodule (including *Caudovirales* and *Herpesvirales*), (iii) Papillomaviruses and Polyomaviruses, and (iv) Baculo-like viruses. The color scheme for archaeal genomes follows the classification in modules shown in panel B. The high density of edges in the main body of the network results from the widespread gene sharing among most bacterial and eukaryotic viral groups, as well as the subset of archaeal viruses that are related to tailed bacteriophages (*Caudovirales*). In contrast, the archaeal-specific portion of the network is characterized by the presence of largely isolated clusters of similar genomes (modules) that correspond to distinct taxa. The figure is modified from (Iranzo et al., 2016b).

B. Modular structure and gene sharing patterns among archaeal viruses. Groups of closely related genomes (modules) are highlighted with different colors. Black edges indicate the connections that involve broadly shared gene families (connector genes). Within each module, colored edges indicate the presence of "signature" genes, i.e. genes that are diagnostic of the members of the given module. The network also includes some capsid-less mobile genetic elements related to archaeal viruses and discussed in the text, such as casposons, which are represented as triangles. GT, glycosyltransferase of the GT-B superfamily; RHH, ribbon-helix-helix domain-containing protein; pPolB, protein-primed DNA polymerase B; MCP, major capsid protein; HAV1, Hyperthermophilic archaeal virus 1. The figure is updated from (Iranzo et al., 2016a). The original data set was supplemented with 26 genomes of Magroviruses (Philosof et al., 2017), 3 genomes of uncultivated His1-like viruses (Adriaenssens et al., 2016), 1 genome of an uncultivated member of the *Caudovirales* from the Oceanic basement (Nigro et al., 2017), and 1 representative genome of the family *Portogloboviridae* (Liu et al., 2017). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

one-way street; thus, it is also likely that some of the non-viral MGE have evolved from archaeal viruses by losing the encapsidation capacity and succumbing to vertical propagation within the host population. This might be the case with the TKV4-like proviruses and pTN3-like integrative plasmids of *Thermococcus*, which do not seem to produce viral particles despite encoding the homologs of the DJR MCP and the characteristic genome packaging ATPase (Gaudin et al., 2014; Krupovic and Bamford, 2008a). Notably, similar to pR1SE, pTN3 has been shown to be specifically incorporated into and transferred between *Thermococcus* cells within membrane vesicles, although the latter do not contain detectable plasmid-encoded proteins (Gaudin et al., 2014). Further exploration of the archaeal virus diversity, especially from
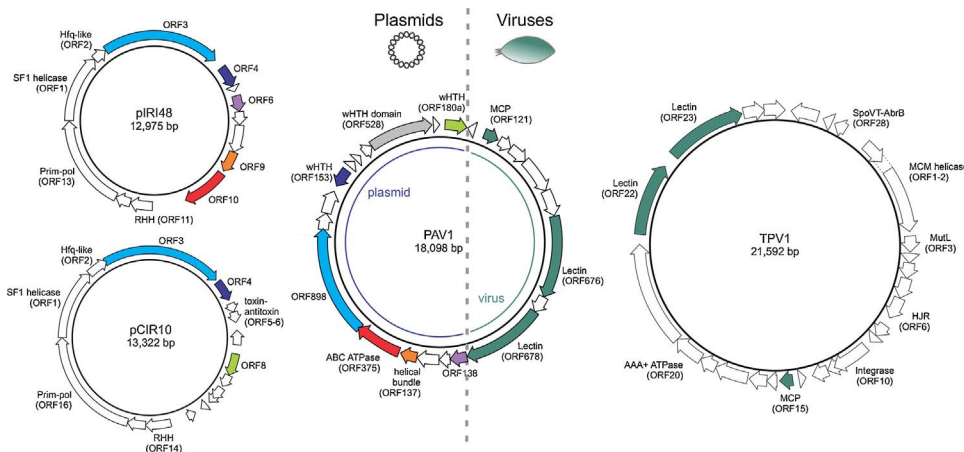

Fig. 5. Relationship between thermococcal plasmids and viruses. Homologous genes are colored similarly. PAV1 ORF528, which has homologues in haloarchaeal plasmids, is shaded grey. Note that viruses PAV1 and TPV1 share only genes encoding structural proteins (dark green arrows). Abbreviations, MCP, major capsid protein; SF1, superfamily 1; wHTH, winged helix-turn-helix; RHH, ribbon-helix-helix; HJR, Holliday junction resolvase; prim-pol, primase-polymerase. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

understudied environments and hosts, should provide crucial insights into the origins and evolution of this remarkable part of the global virosphere.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.virusres.2017.11.025.

References

Abrescia, N.G., Bamford, D.H., Grimes, J.M., Stuart, D.I., 2012. Structure unifies the viral universe. Annu. Rev. Biochem. 81, 795–822.

Adriaenssens, E.M., van Zyl, L.J., Cowan, D.A., Trindade, M.I., 2016. Metaviromics of namib desert salt pans: a novel lineage of haloarchaeal salterproviruses and a rich source of ssDNA viruses. Viruses 8 (1), E14.

Ahn, D.G., Kim, S.I., Rhee, J.K., Kim, K.P., Pan, J.G., Oh, J.W., 2006. TTSV1, a new virus-like particle isolated from the hyperthermophilic crenarchaeote Thermoproteus tenax. Virology 351 (2), 280–290.

Anderson, R.E., Kouris, A., Seward, C.H., Campbell, K.M., Whitaker, R.J., 2017. Structured populations of sulfolobus acidocaldarius with susceptibility to mobile genetic elements. Genom. Biol. Evol. 9 (6), 1699–1710.

Arnold, H.P., Ziese, U., Zillig, W., 2000. SNDV, a novel virus of the extremely thermophilic and acidophilic archaeon Sulfolobus. Virology 272 (2), 409–416.

Atanasova, N.S., Roine, E., Oren, A., Bamford, D.H., Oksanen, H.M., 2012. Global network of specific virus-host interactions in hypersaline environments. Environ. Microbiol. 14 (2), 426–440.

Bath, C., Cukalac, T., Porter, K., Dyall-Smith, M.L., 2006. His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, Salterprovirus. Virology 350 (1), 228–239.

Bautista, M.A., Black, J.A., Youngblut, N.D., Whitaker, R.J., 2017. Differentiation and structure in sulfolobus islandicus rod-Shaped virus populations. Viruses 9 (5).

Belmar, L., Molina, V., Ulloa, O., 2011. Abundance and phylogenetic identity of archaeoplankton in the permanent oxygen minimum zone of the eastern tropical South Pacific. FEMS Microbiol. Ecol. 78 (2), 314–326.

Birkenbihl, R.P., Neef, K., Prangishvili, D., Kemper, B., 2001. Holliday junction resolving enzymes of archaeal viruses SIRV1 and SIRV2. J. Mol. Biol. 309 (5), 1067–1076.

Bolduc, B., Shaughnessy, D.P., Wolf, Y.I., Koonin, E.V., Roberto, F.F., Young, M., 2012. Identification of novel positive-strand RNA viruses by metagenomic analysis of archaea-dominated Yellowstone hot springs. J. Virol. 86 (10), 5562–5573.

Bolduc, B., Wirth, J.F., Mazurie, A., Young, M.J., 2015. Viral assemblage composition in Yellowstone acidic hot springs assessed by network analysis. ISME J. 9 (10), 2162–2177.

Burstein, D., Harrington, L.B., Strutt, S.C., Probst, A.J., Anantharaman, K., Thomas, B.C., Doudna, J.A., Banfield, J.F., 2017. New CRISPR-Cas systems from uncultivated microbes. Nature 542 (7640), 237–241.

Chothia, C., Lesk, A.M., 1986. The relation between the divergence of sequence and structure in proteins. EMBO J. 5 (4), 823–826.

Chow, C.E., Winget, D.M., White 3rd, R.A., Hallam, S.J., Suttle, C.A., 2015. Combining genomic sequencing methods to explore viral diversity and reveal potential virus-host interactions. Front. Microbiol. 6, 265.

Chu, Y., Zhu, Y., Chen, Y., Li, W., Zhang, Z., Liu, D., Wang, T., Ma, J., Deng, H., Liu, Z.J., Ouyang, S., Huang, L., 2016. aKMT catalyzes extensive protein lysine methylation in the hyperthermophilic archaeon sulfolobus islandicus but is dispensable for the growth of the organism. Mol. Cell. Proteomics 15 (9), 2908–2923.

Comolli, L.R., Baker, B.J., Downing, K.H., Siegerist, C.E., Banfield, J.F., 2009. Three-dimensional analysis of the structure and ecology of a novel, ultra-small archaeon. ISME J. 3 (2), 159–167.

Danovaro, R., Dell'Anno, A., Corinaldesi, C., Rastelli, E., Cavicchioli, R., Krupovic, M., Noble, R.T., Nunoura, T., Prangishvili, D., 2016. Virus-mediated archaeal hecatomb in the deep seafloor. Sci. Adv. 2 (10), e1600492.

Danovaro, R., Rastelli, E., Corinaldesi, C., Tangherlini, M., Dell'Anno, A., 2017. Marine archaea and archaeal viruses under global change. F1000Res 6, 1241.

Dellas, N., Snyder, J.C., Bolduc, B., Young, M.J., 2014. Archaeal viruses: diversity, replication, and structure. Annu. Rev. Virol. 1 (1), 399–426.

Demina, T.A., Pietila, M.K., Svirskaite, J., Ravantti, J.J., Atanasova, N.S., Bamford, D.H., Oksanen, H.M., 2017. HCIV-1 and other tailless icosahedral internal membrane-Containing viruses of the family sphaerolipoviridae. Viruses 9 (2), E32.

DiMaio, F., Yu, X., Rensen, E., Krupovic, M., Prangishvili, D., Egelman, E.H., 2015. Virology. A virus that infects a hyperthermophile encapsidates A-form DNA. Science 348 (6237), 914–917.

Dyall-Smith, M.L., Pfeiffer, F., Klee, K., Palm, P., Gross, K., Schuster, S.C., Rampp, M., Oesterhelt, D., 2011. Haloquadratum walsbyi: limited diversity in a global pond. PLoS One 6 (6), e20968.

Emerson, J.B., Andrade, K., Thomas, B.C., Norman, A., Allen, E.E., Heidelberg, K.B., Banfield, J.F., 2013. Virus-host and CRISPR dynamics in archaea-dominated hypersaline lake Tyrrell, Victoria, Australia. Archaea 2013, 370871.

Erdmann, S., Le Moine Bauer, S., Garrett, R.A., 2014. Inter-viral conflicts that exploit host CRISPR immune systems of Sulfolobus. Mol. Microbiol. 91 (5), 900–917.

Erdmann, S., Tschitschko, B., Zhong, L., Raftery, M.J., Cavicchioli, R., 2017. A plasmid from an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells. Nat. Microbiol. 2 (10), 1446–1455.

Felisberto-Rodrigues, C., Blangy, S., Goulet, A., Vestergaard, G., Cambillau, C., Garrett, R.A., Ortiz-Lombardia, M., 2012. Crystal structure of ATV(ORF273), a new fold for a thermo- and acido-stable protein from the Acidianus two-tailed virus. PLoS One 7 (10), e45847.

Filée, J., Siguier, P., Chandler, M., 2007. Insertion sequence diversity in archaea. Microbiol. Mol. Biol. Rev. 71 (1), 121–157.

Forterre, P., Krupovic, M., Raymann, K., Soler, N., 2014. Plasmids from euryarchaeota. Microbiol. Spectr. 2 (6) (PLAS-0027-2014).

Forterre, P., 2013. Why are there so many diverse replication machineries? J. Mol. Biol. 425 (17), 4714–4726.

Fröls, S., Gordon, P.M., Panlilio, M.A., Schleper, C., Sensen, C.W., 2007. Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays. Virology 365 (1), 48–59.

Fusco, S., Liguori, R., Limauro, D., Bartolucci, S., She, Q., Contursi, P., 2015a. Transcriptome analysis of Sulfolobus solfataricus infected with two related fuselloviruses reveals novel insights into the regulation of CRISPR-Cas system. Biochimie 118, 322–332.

Fusco, S., She, Q., Fiorentino, G., Bartolucci, S., Contursi, P., 2015b. Unravelling the role of the F55 regulator in the transition from lysogeny to UV induction of sulfolobus spindle-Shaped virus 1. J. Virol. 89 (12), 6453–6461.

Garcia-Heredia, I., Martin-Cuadrado, A.B., Mojica, F.J., Santos, F., Mira, A., Anton, J., Rodriguez-Valera, F., 2012. Reconstructing viral genomes from the environment using fosmid clones: the case of haloviruses. PLoS One 7 (3), e33802.

Gardner, A.F., Prangishvili, D., Jack, W.E., 2011. Characterization of Sulfolobus islandicus rod-shaped virus 2 gp19, a single-strand specific endonuclease. Extremophiles 15 (5), 619–624.

Gardner, A.F., Bell, S.D., White, M.F., Prangishvili, D., Krupovic, M., 2014. Protein-protein interactions leading to recruitment of the host DNA sliding clamp by the hyperthermophilic Sulfolobus islandicus rod-shaped virus 2. J. Virol. 88 (12), 7105–7108.

Garrett, R.A., Prangishvili, D., Shah, S.A., Reuter, M., Stetter, K.O., Peng, X., 2010. Metagenomic analyses of novel viruses and plasmids from a cultured environmental sample of hyperthermophilic neutrophiles. Environ. Microbiol. 12 (11), 2918–2930.

Gaudin, M., Krupovic, M., Marguet, E., Gauliard, E., Cvirkaite-Krupovic, V., Le Cam, E., Oberto, J., Forterre, P., 2014. Extracellular membrane vesicles harbouring viral genomes. Environ. Microbiol. 16 (4), 1167–1175.

Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G., Prieur, D., 2003. PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, Pyrococcus abyssi. J. Bacteriol. 185 (13), 3888–3894.

Geslin, C., Gaillard, M., Flament, D., Rouault, K., Le Romancer, M., Prieur, D., Erauso, G., 2007. Analysis of the first genome of a hyperthermophilic marine virus-like particle, PAV1, isolated from Pyrococcus abyssi. J. Bacteriol. 189 (12), 4510–4519.

Gil-Carton, D., Jaakkola, S.T., Charro, D., Peralta, B., Castano-Diez, D., Oksanen, H.M., Bamford, D.H., Abrescia, N.G., 2015. Insight into the assembly of viruses with vertical single beta-barrel major capsid proteins. Structure 23 (10), 1866–1877.

Gorlas, A., Koonin, E.V., Bienvenu, N., Prieur, D., Geslin, C., 2012. TPV1, the first virus isolated from the hyperthermophilic genus Thermococcus. Environ. Microbiol. 14 (2), 503–516.

Gorlas, A., Krupovic, M., Forterre, P., Geslin, C., 2013. Living side by side with a virus: characterization of two novel plasmids from Thermococcus prieurii, a host for the spindle-shaped virus TPV1. Appl. Environ. Microbiol. 79 (12), 3822–3828.

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. U. S. A. 106 (50), 21155–21160.

Goulet, A., Pina, M., Redder, P., Prangishvili, D., Vera, L., Lichiere, J., Leulliot, N., van Tilbeurgh, H., Ortiz-Lombardia, M., Campanacci, V., Cambillau, C., 2010a. ORF157 from the archaeal virus Acidianus filamentous virus 1 defines a new class of nuclease. J. Virol. 84 (10), 5025–5031.

Goulet, A., Vestergaard, G., Felisberto-Rodrigues, C., Campanacci, V., Garrett, R.A., Cambillau, C., Ortiz-Lombardia, M., 2010b. Getting the best out of long-wavelength X-rays: de novo chlorine/sulfur SAD phasing of a structural protein from ATV. Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 3), 304–308.

Gudbergsdóttir, S.R., Menzel, P., Krogh, A., Young, M., Peng, X., 2016. Novel viral genomes identified from six metagenomes reveal wide distribution of archaeal viruses and high viral diversity in terrestrial hot springs. Environ. Microbiol. 18 (3), 863–874.

Guilliere, F., Peixeiro, N., Kessler, A., Raynal, B., Desnoues, N., Keller, J., Delepierre, M., Prangishvili, D., Sezonov, G., Guijarro, J.I., 2009. Structure, function, and targets of the transcriptional regulator SvtR from the hyperthermophilic archaeal virus SIRV1. J. Biol. Chem. 284 (33), 22222–22237.

Guo, Y., Kragelund, B.B., White, M.F., Peng, X., 2015. Functional characterization of a conserved archaeal viral operon revealing single-stranded DNA binding, annealing and nuclease activities. J. Mol. Biol. 427 (12), 2179–2191.

Häring, M., Peng, X., Brugger, K., Rachel, R., Stetter, K.O., Garrett, R.A., Prangishvili, D.,

2004. Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera *Pyrobaculum* and *Thermoproteus*: a novel virus family, the *Globuloviridae*. Virology 323 (2), 233–242.

Häring, M., Rachel, R., Peng, X., Garrett, R.A., Prangishvili, D., 2005. 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 (15), 9904–9911.

Happonen, L.J., Redder, P., Peng, X., Reigstad, L.J., Prangishvili, D., Butcher, S.J., 2010. Familial relationships in hyperthermo- and acidophilic archaeal viruses. J. Virol. 84 (9), 4747–4754.

Held, N.L., Whitaker, R.J., 2009. Viral biogeography revealed by signatures in Sulfolobus islandicus genomes. Environ. Microbiol. 11 (2), 457–466.

Hochstein, R., Bollschweiler, D., Engelhardt, H., Lawrence, C.M., Young, M., 2015. Large tailed spindle viruses of archaea: a new way of doing viral business. J. Virol. 89 (18), 9146–9149.

Hochstein, R.A., Amenabar, M.J., Munson-McGee, J.H., Boyd, E.S., Young, M.J., 2016. Acidianus tailed spindle virus: a new archaeal large tailed spindle virus discovered by culture-independent methods. J. Virol. 90 (7), 3458–3468.

Iranzo, J., Koonin, E.V., Prangishvili, D., Krupovic, M., 2016a. Bipartite network analysis of the archaeal virosphere: evolutionary connections between viruses and capsidless mobile elements. J. Virol. 90 (24), 11043–11055.

Iranzo, J., Krupovic, M., Koonin, E.V., 2016b. The double-stranded DNA virosphere as a modular hierarchical network of gene sharing. mBio 7 (4), e00978–16.

Iranzo, J., Krupovic, M., Koonin, E.V., 2017. A network perspective on the virus world. Commun. Integr. Biol. 10 (2), e1296614.

Iverson, E., Stedman, K., 2012. A genetic study of SSV1, the prototypical fusellovirus. Front. Microbiol. 3, 200.

Iverson, E.A., Goodman, D.A., Gorchels, M.E., Stedman, K.M., 2017. Extreme mutation tolerance: nearly half of the archaeal fusellovirus SSV1 genes are not required for virus function, including the minor capsid protein gene, vp3. J. Virol. 91 (10), e02406–16.

Jaakkola, S.T., Penttinen, R.K., Vilen, S.T., Jalasvuori, M., Ronnholm, G., Bamford, J.K., Bamford, D.H., Oksanen, H.M., 2012. Closely related archaeal Haloarcula hispanica icosahedral viruses HHIV-2 and SH1 have nonhomologous genes encoding host recognition functions. J. Virol. 86 (9), 4734–4742.

Jaubert, C., Danioux, C., Oberto, J., Cortez, D., Bize, A., Krupovic, M., She, Q., Forterre, P., Prangishvili, D., Sezonov, G., 2013. Genomics and genetics of Sulfolobus islandicus LAL14/1, a model hyperthermophilic archaeon. Open Biol. 3 (4), 130010.

Karner, M.B., DeLong, E.F., Karl, D.M., 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. Nature 409 (6819), 507–510.

Kasson, P., DiMaio, F., Yu, X., Lucas-Staat, S., Krupovic, M., Schouten, S., Prangishvili, D., Egelman, E.H., 2017. Model for a novel membrane envelope in a filamentous hyperthermophilic virus. eLife 6, e26268.

Kazlauskas, D., Krupovic, M., Venclovas, C., 2016. The logic of DNA replication in double-stranded DNA viruses: insights from global analysis of viral genomes. Nucleic Acids Res. 44 (10), 4551–4564.

Kessler, A., Brinkman, A.B., van der Oost, J., Prangishvili, D., 2004. Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon sulfolobus. J. Bacteriol. 186 (22), 7745–7753.

Klein, R., Baranyi, U., Rossler, N., Greineder, B., Scholz, H., Witte, A., 2002. Natrialba magadii virus phiCh1: first complete nucleotide sequence and functional organization of a virus infecting a haloalkaliphilic archaeon. Mol. Microbiol. 45 (3), 851–863.

Koonin, E.V., Krupovic, M., 2017. Polintons, virophages and transpovirons: a tangled web linking viruses, transposons and immunity. Curr. Opin. Virol. 25, 7–15.

Koonin, E.V., Dolja, V.V., Krupovic, M., 2015. Origins and evolution of viruses of eukaryotes: the ultimate modularity. Virology 479 (-480), 2–25.

Koonin, E.V., Makarova, K.S., Wolf, Y.I., 2017. Evolutionary Genomics of Defense Systems in Archaea and Bacteria. Annu Rev Microbiol.

Koonin, E.V., 1992. Archaebacterial virus SSV1 encodes a putative DnaA-like protein. Nucleic Acids Res. 20 (5), 1143.

Krupovic, M., Bamford, D.H., 2007. Putative prophages related to lytic tailless marine dsDNA phage PM2 are widespread in the genomes of aquatic bacteria. BMC Genom. 8, 236.

Krupovic, M., Bamford, D.H., 2008a. Archaeal proviruses TKV4 and MVV extend the PRD1-adenovirus lineage to the phylum Euryarchaeota. Virology 375 (1), 292–300.

Krupovic, M., Bamford, D.H., 2008b. Virus evolution: how far does the double beta-barrel viral lineage extend? Nat. Rev. Microbiol. 6 (12), 941–948.

Krupovic, M., Bamford, D.H., 2009. Does the evolution of viral polymerases reflect the origin and evolution of viruses? Nat. Rev. Microbiol. 7 (3), 250 (author reply 250).

Krupovic, M., Bamford, D.H., 2010. Order to the viral universe. J. Virol. 84 (24), 12476–12479.

Krupovic, M., Bamford, D.H., 2011. Double-stranded DNA viruses: 20 families and only five different architectural principles for virion assembly. Curr. Opin. Virol. 1 (2), 118–124.

Krupovic, M., Koonin, E.V., 2015. Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. Nat. Rev. Microbiol. 13 (2), 105–115.

Krupovic, M., Koonin, E.V., 2017. Multiple origins of viral capsid proteins from cellular ancestors. Proc. Natl. Acad. Sci. U. S. A. 114 (12), E2401–E2410.

Krupovic, M., Forterre, P., Bamford, D.H., 2010a. Comparative analysis of the mosaic genomes of tailed archaeal viruses and proviruses suggests common themes for virion architecture and assembly with tailed viruses of bacteria. J. Mol. Biol. 397 (1), 144–160.

Krupovic, M., Gribaldo, S., Bamford, D.H., Forterre, P., 2010b. The evolutionary history of archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of mobile genetic elements. Mol. Biol. Evol. 27 (12), 2716–2732.

Krupovic, M., Spang, A., Gribaldo, S., Forterre, P., Schleper, C., 2011. A thaumarchaeal

provirus testifies for an ancient association of tailed viruses with archaea. Biochem. Soc. Trans. 39 (1), 82–88.

Krupovic, M., White, M.F., Forterre, P., Prangishvili, D., 2012. Postcards from the edge: structural genomics of archaeal viruses. Adv. Virus Res. 82, 33–62.

Krupovic, M., Gonnet, M., Hania, W.B., Forterre, P., Erauso, G., 2013. Insights into dynamics of mobile genetic elements in hyperthermophilic environments from five new Thermococcus plasmids. PLoS One 8 (1), e49044.

Krupovic, M., Makarova, K.S., Forterre, P., Prangishvili, D., Koonin, E.V., 2014a. Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. BMC Biol. 12, 36.

Krupovic, M., Quemin, E.R., Bamford, D.H., Forterre, P., Prangishvili, D., 2014b. Unification of the globally distributed spindle-shaped viruses of the Archaea. J. Virol. 88 (4), 2354–2358.

Krupovic, M., Cvirkaite-Krupovic, V., Prangishvili, D., Koonin, E.V., 2015. Evolution of an archaeal virus nucleocapsid protein from the CRISPR-associated Cas4 nuclease. Biol. Direct 10, 65.

Krupovic, M., Béguin, P., Koonin, E.V., 2017. Casposons: mobile genetic elements that gave rise to the CRISPR-Cas adaptation machinery. Curr. Opin. Microbiol. 38, 36–43.

Krupovic, M., 2013. Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses. Curr. Opin. Virol. 3 (5), 578–586.

Labonte, J.M., Swan, B.K., Poulos, B., Luo, H., Koren, S., Hallam, S.J., Sullivan, M.B., Woyke, T., Wommack, K.E., Stepanauskas, R., 2015. Single-cell genomics-based analysis of virus-host interactions in marine surface bacterioplankton. ISME J. 9 (11), 2386–2399.

Larson, E.T., Reiter, D., Young, M., Lawrence, C.M., 2006. Structure of A197 from Sulfolobus turreted icosahedral virus: a crenarchaeal viral glycosyltransferase exhibiting the GT-A fold. J. Virol. 80 (15), 7636–7644.

Leon-Sobrino, C., Kot, W.P., Garrett, R.A., 2016. Transcriptome changes in STSV2-infected Sulfolobus islandicus REY15A undergoing continuous CRISPR spacer acquisition. Mol. Microbiol. 99 (4), 719–728.

Lindås, A.C., Bernander, R., 2013. The cell cycle of archaea. Nat. Rev. Microbiol. 11 (9), 627–638.

Liu, Y., Beer, L.L., Whitman, W.B., 2012. Methanogens: a window into ancient sulfur metabolism. Trends Microbiol. 20 (5), 251–258.

Liu, Y., Wang, J., Wang, Y., Zhang, Z., Oksanen, H.M., Bamford, D.H., Chen, X., 2015. Identification and characterization of SNJ2, the first temperate pleolipovirus integrating into the genome of the SNJ1-lysogenic archaeal strain. Mol. Microbiol. 98 (6), 1002–1020.

Liu, Y., Ishino, S., Ishino, Y., Pehau-Arnaudet, G., Krupovic, M., Prangishvili, D., 2017. A novel type of polyhedral viruses infecting hyperthermophilic archaea. J. Virol. 91 (13), e00589–17.

Luk, A.W., Williams, T.J., Erdmann, S., Papke, R.T., Cavicchioli, R., 2014. Viruses of haloarchaea. Life (Basel) 4 (4), 681–715.

Maaty, W.S., Selvig, K., Ryder, S., Tarlykov, P., Hilmer, J.K., Heinemann, J., Steffens, J., Snyder, J.C., Ortmann, A.C., Movahed, N., Spicka, K., Chetia, L., Grieco, P.A., Dratz, E.A., Douglas, T., Young, M.J., Bothner, B., 2012a. Proteomic analysis of sulfolobus solfataricus during sulfolobus turreted icosahedral virus infection. J. Proteome Res. 11 (2), 1420–1432.

Maaty, W.S., Steffens, J.D., Heinemann, J., Ortmann, A.C., Reeves, B.D., Biswas, S.K., Dratz, E.A., Grieco, P.A., Young, M.J., Bothner, B., 2012b. Global analysis of viral infection in an archaeal model system. Front. Microbiol. 3, 411.

Makarova, K.S., Koonin, E.V., 2013. Archaeology of eukaryotic DNA replication. Cold Spring Harb. Perspect. Biol. 5 (11), a012963.

Makarova, K.S., Yutin, N., Bell, S.D., Koonin, E.V., 2010. Evolution of diverse cell division and vesicle formation systems in Archaea. Nat. Rev. Microbiol. 8 (10), 731–741.

Makarova, K.S., Krupovic, M., Koonin, E.V., 2014. Evolution of replicative DNA polymerases in archaea and their contributions to the eukaryotic replication machinery. Front. Microbiol. 5, 354.

Makarova, K.S., Koonin, E.V., Albers, S.V., 2016. Diversity and evolution of type IV pili systems in archaea. Front. Microbiol. 7, 667.

Marine, R.L., Nasko, D.J., Wray, J., Polson, S.W., Wommack, K.E., 2017. Novel chaperonins are prevalent in the virioplankton and demonstrate links to viral biology and ecology. ISME J. 11 (11), 2479–2491.

Martinez-Alvarez, L., Bell, S.D., Peng, X., 2016. Multiple consecutive initiation of replication producing novel brush-like intermediates at the termini of linear viral dsDNA genomes with hairpin ends. Nucleic Acids Res. 44 (18), 8799–8809.

Martinez-Alvarez, L., Deng, L., Peng, X., 2017. Formation of a viral replication focus in sulfolobus cells infected by the rudivirus sulfolobus islandicus rod-shaped virus 2. J. Virol. 91 (13), e00486–17.

Martinez-Garcia, M., Santos, F., Moreno-Paz, M., Parro, V., Anton, J., 2014. Unveiling viral-host interactions in the 'microbial dark matter'. Nat. Commun. 5, 4542.

Menon, S.K., Eilers, B.J., Young, M.J., Lawrence, C.M., 2010. The crystal structure of D212 from sulfolobus spindle-shaped virus ragged hills reveals a new member of the PD-(D/E)XK nuclease superfamily. J. Virol. 84 (12), 5890–5897.

Mochizuki, T., Yoshida, T., Tanaka, R., Forterre, P., Sako, Y., Prangishvili, D., 2010. Diversity of viruses of the hyperthermophilic archaeal genus Aeropyrum, and isolation of the Aeropyrum pernix bacilliform virus 1, APBV1, the first representative of the family Clavaviridae. Virology 402 (2), 347–354.

Mochizuki, T., Krupovic, M., Pehau-Arnaudet, G., Sako, Y., Forterre, P., Prangishvili, D., 2012. Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. Proc. Natl. Acad. Sci. U. S. A. 109 (33), 13386–13391.

Molineux, I.J., Panja, D., 2013. Popping the cork: mechanisms of phage genome ejection. Nat. Rev. Microbiol. 11 (3), 194–204.

Moriscot, C., Gribaldo, S., Jault, J.M., Krupovic, M., Arnaud, J., Jamin, M., Schoehn, G., Forterre, P., Weissenhorn, W., Renesto, P., 2011. Crenarchaeal CdvA forms double-helical filaments containing DNA and interacts with ESCRT-III-like CdvB. PLoS One 6

(7), e21921.

Narasingarao, P., Podell, S., Ugalde, J.A., Brochier-Armanet, C., Emerson, J.B., Brocks, J.J., Heidelberg, K.B., Banfield, J.F., Allen, E.E., 2012. De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. ISME J. 6 (1), 81–93.

Nigro, O.D., Jungbluth, S.P., Lin, H.T., Hsieh, C.C., Miranda, J.A., Schvarcz, C.R., Rappe, M.S., Steward, G.F., 2017. Viruses in the oceanic basement. mBio 8 (2).

Nishimura, Y., Watai, H., Honda, T., Mihara, T., Omae, K., Roux, S., Blanc-Mathieu, R., Yamamoto, K., Hingamp, P., Sako, Y., Sullivan, M.B., Goto, S., Ogata, H., Yoshida, T., 2017. Environmental viral genomes shed new light on virus-Host interactions in the ocean. mSphere 2 (2).

Offre, P., Spang, A., Schleper, C., 2013. Archaea in biogeochemical cycles. Annu. Rev. Microbiol. 67, 437–457.

Oke, M., Kerou, M., Liu, H., Peng, X., Garrett, R.A., Prangishvili, D., Naismith, J.H., White, M.F., 2011. A dimeric Rep protein initiates replication of a linear archaeal virus genome: implications for the Rep mechanism and viral replication. J. Virol. 85 (2), 925–931.

Okutan, E., Deng, L., Mirlashari, S., Uldahl, K., Halim, M., Liu, C., Garrett, R.A., She, Q., Peng, X., 2013. Novel insights into gene regulation of the rudivirus SIRV2 infecting Sulfolobus cells. RNA Biol. 10 (5), 875–885.

Ortmann, A.C., Brumfield, S.K., Walther, J., McInnerney, K., Brouns, S.J., van de Werken, H.J., Bothner, B., Douglas, T., van de Oost, J., Young, M.J., 2008. Transcriptome analysis of infection of the archaeon Sulfolobus solfataricus with Sulfolobus turreted icosahedral virus. J. Virol. 82 (10), 4874–4883.

Pagaling, E., Haigh, R.D., Grant, W.D., Cowan, D.A., Jones, B.E., Ma, Y., Ventosa, A., Heaphy, S., 2007. Sequence analysis of an Archaeal virus isolated from a hypersaline lake in Inner Mongolia, China. BMC Genomics 8, 410.

Palmieri, G., Balestrieri, M., Peter-Katalinic, J., Pohlentz, G., Rossi, M., Fiume, I., Pocsfalvi, G., 2013. Surface-exposed glycoproteins of hyperthermophilic Sulfolobus solfataricus P2 show a common N-glycosylation profile. J. Proteome Res. 12 (6), 2779–2790.

Pan, M., Kelman, L.M., Kelman, Z., 2011. The archaeal PCNA proteins. Biochem. Soc. Trans. 39 (1), 20–24.

Paul, B.G., Bagby, S., Czornyj, E., Arambula, D., Handa, S., Sczyrba, A., Ghosh, P., Miller, J.F., Valentine, D.L., 2015. Targeted diversity generation by intraterrestrial archaea and archaeal viruses. Nat. Commun. 6, 6585.

Pawlowski, A., Rissanen, I., Bamford, J.K., Krupovic, M., Jalasvuori, M., 2014. Gammasphaerolipovirus, a newly proposed bacteriophage genus, unifies viruses of halophilic archaea and thermophilic bacteria within the novel family Sphaerolipoviridae. Arch. Virol. 159 (6), 1541–1554.

Peeters, E., Peixeiro, N., Sezonov, G., 2013. Cis-regulatory logic in archaeal transcription. Biochem. Soc. Trans. 41 (1), 326–331.

Peeters, E., Boon, M., Rollie, C., Willaert, R.G., Voet, M., White, M.F., Prangishvili, D., Lavigne, R., Quax, T.E.F., 2017. DNA-interacting characteristics of the archaeal rudiviral protein SIRV2_Gp1. Viruses 9 (7).

Peixeiro, N., Keller, J., Collinet, B., Leulliot, N., Campanacci, V., Cortez, D., Cambillau, C., Nitta, K., Vincentelli, R., Forterre, P., Prangishvili, D., Sezonov, G., van Tilbeurgh, H., 2013. Structure and function of AvtR, a novel transcriptional regulator from a hyperthermophilic archaeal lipothrixvirus. J. Virol. 87 (1), 124–136.

Peng, X., Basta, T., Haring, M., Garrett, R.A., Prangishvili, D., 2007. Genome of the Acidianus bottle-shaped virus and insights into the replication and packaging mechanisms. Virology 364 (1), 237–243.

Philosof, A., Yutin, N., Flores-Uribe, J., Sharon, I., Koonin, E.V., Béjà, O., 2017. Novel abundant oceanic viruses of uncultured Marine Group II Euryarchaeota identified by genome-centric metagenomics. Curr. Biol. 27 (9), 1362–1368.

Pietilä, M.K., Laurinmaki, P., Russell, D.A., Ko, C.C., Jacobs-Sera, D., Hendrix, R.W., Bamford, D.H., Butcher, S.J., 2013. Structure of the archaeal head-tailed virus HSTV-1 completes the HK97 fold story. Proc. Natl. Acad. Sci. U. S. A. 110 (26), 10604–10609.

Pietilä, M.K., Demina, T.A., Atanasova, N.S., Oksanen, H.M., Bamford, D.H., 2014. Archaeal viruses and bacteriophages: comparisons and contrasts. Trends Microbiol. 22 (5), 334–344.

Pietilä, M.K., Roine, E., Sencilo, A., Bamford, D.H., Oksanen, H.M., 2016. *Pleolipoviridae*, a newly proposed family comprising archaeal pleomorphic viruses with single-stranded or double-stranded DNA genomes. Arch. Virol. 161 (1), 249–256.

Pina, M., Basta, T., Quax, T.E., Joubert, A., Baconnais, S., Cortez, D., Lambert, S., Le Cam, E., Bell, S.D., Forterre, P., Prangishvili, D., 2014. Unique genome replication mechanism of the archaeal virus AFV1. Mol. Microbiol. 92 (6), 1313–1325.

Porter, K., Dyall-Smith, M.L., 2008. Transfection of haloarchaea by the DNAs of spindle and round haloviruses and the use of transposon mutagenesis to identify non-essential regions. Mol. Microbiol. 70 (5), 1236–1245.

Porter, K., Kukkaro, P., Bamford, J.K., Bath, C., Kivela, H.M., Dyall-Smith, M.L., Bamford, D.H., 2005. SH1: A novel, spherical halovirus isolated from an Australian hypersaline lake. Virology 335 (1), 22–33.

Porter, K., Tang, S.L., Chen, C.P., Chiang, P.W., Hong, M.J., Dyall-Smith, M., 2013. PH1: an archaeovirus of Haloarcula hispanica related to SH1 and HHIV-2. Archaea 2013, 456318.

Prangishvili, D., Krupovic, M., 2012. A new proposed taxon for double-stranded DNA viruses, the order Ligamenvirales. Arch. Virol. 157 (4), 791–795.

Prangishvili, D., Klenk, H.P., Jakobs, G., Schmiechen, A., Hanselmann, C., Holz, I., Zillig, W., 1998. Biochemical and phylogenetic characterization of the dUTPase from the archaeal virus SIRV. J. Biol. Chem. 273 (11), 6024–6029.

Prangishvili, D., Garrett, R.A., Koonin, E.V., 2006. Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life. Virus Res. 117 (1), 52–67.

Prangishvili, D., Koonin, E.V., Krupovic, M., 2013. Genomics and biology of Rudiviruses, a model for the study of virus-host interactions in Archaea. Biochem. Soc. Trans. 41 (1), 443–450.

Prangishvili, D., Bamford, D.H., Forterre, P., Iranzo, J., Koonin, E.V., Krupovic, M., 2017. The enigmatic archaeal virosphere. Nat. Rev. Microbiol. 15 (12), 724–739.

Prangishvili, D., 2013. The wonderful world of archaeal viruses. Annu. Rev. Microbiol. 67, 565–585.

Prangishvili, D., 2015. Archaeal viruses: living fossils of the ancient virosphere? Ann. N. Y. Acad. Sci. 1341, 35–40.

Ptchelkine, D., Gillum, A., Mochizuki, T., Lucas-Staat, S., Liu, Y., Krupovic, M., Phillips, S., Prangishvili, D., Huiskonen, J., 2017. Unique architecture of thermophilic archaeal virus APBV1 and its genome packaging. Nat. Commun. 8, 1436.

Quax, T.E., Voet, M., Sismeiro, O., Dillies, M.A., Jagla, B., Coppee, J.Y., Sezonov, G., Forterre, P., van der Oost, J., Lavigne, R., Prangishvili, D., 2013. Massive activation of archaeal defense genes during viral infection. J. Virol. 87 (15), 8419–8428.

Rao, V.B., Feiss, M., 2015. Mechanisms of DNA packaging by large double-Stranded DNA viruses. Annu. Rev.Virol. 2 (1), 351–378.

Redrejo-Rodríguez, M., Salas, M., 2014. Multiple roles of genome-attached bacteriophage terminal proteins. Virology 468-470, 322–329.

Ren, Y., She, Q., Huang, L., 2013. Transcriptomic analysis of the SSV2 infection of Sulfolobus solfataricus with and without the integrative plasmid pSSVi. Virology 441 (2), 126–134.

Reno, M.L., Held, N.L., Fields, C.J., Burke, P.V., Whitaker, R.J., 2009. Biogeography of the Sulfolobus islandicus pan-genome. Proc. Natl. Acad. Sci. U. S. A. 106 (21), 8605–8610.

Rensen, E.I., Mochizuki, T., Quemin, E., Schouten, S., Krupovic, M., Prangishvili, D., 2016. A virus of hyperthermophilic archaea with a unique architecture among DNA viruses. Proc. Natl. Acad. Sci. U. S. A. 113 (9), 2478–2483.

Rixon, F.J., Schmid, M.F., 2014. Structural similarities in DNA packaging and delivery apparatuses in Herpesvirus and dsDNA bacteriophages. Curr. Opin. Virol. 5, 105–110.

Robinson, N.P., Bell, S.D., 2007. Extrachromosomal element capture and the evolution of multiple replication origins in archaeal chromosomes. Proc. Natl. Acad. Sci. U. S. A. 104 (14), 5806–5811.

Roux, S., Brum, J.R., Dutilh, B.E., Sunagawa, S., Duhaime, M.B., Loy, A., Poulos, B.T., Solonenko, N., Lara, E., Poulain J., Pesant, S., Kandels-Lewis, S., Dimier, C., Picheral, M., Searson, S., Cruaud, C., Alberti, A., Duarte, C.M., Gasol, J.M., Vaque, D., Bork, P., Acinas, S.G., Wincker, P., Sullivan, M.B., 2016a. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. Nature 537 (7622), 689–693.

Roux, S., Enault, F., Ravet, V., Colombet, J., Bettarel, Y., Auguet, J.C., Bouvier, T., Lucas-Staat, S., Vellet, A., Prangishvili, D., Forterre, P., Debroas, D., Sime-Ngando, T., 2016b. Analysis of metagenomic data reveals common features of halophilic viral communities across continents. Environ. Microbiol. 18 (3), 889–903.

Samson, R.Y., Obita, T., Freund, S.M., Williams, R.L., Bell, S.D., 2008. A role for the ESCRT system in cell division in archaea. Science 322 (5908), 1710–1713.

Santos, F., Meyerdierks, A., Pena, A., Rossello-Mora, R., Amann, R., Anton, J., 2007. Metagenomic approach to the study of halophages: the environmental halophage 1. Environ. Microbiol. 9 (7), 1711–1723.

Santos, F., Moreno-Paz, M., Meseguer, I., Lopez, C., Rossello-Mora, R., Parro, V., Anton, J., 2011. Metatranscriptomic analysis of extremely halophilic viral communities. ISME J. 5 (10), 1621–1633.

Selb, R., Derntl, C., Klein, R., Alte, B., Hofbauer, C., Kaufmann, M., Beraha, J., Schoner, L., Witte, S., 2017. The viral gene ORF79 encodes a repressor regulating induction of the lytic life cycle in the haloalkaliphilic virus phiCh1. J. Virol. 91 (9), e00206–17.

Sencilo, A., Roine, E., 2014. A glimpse of the genomic diversity of haloarchaeal tailed viruses. Front. Microbiol. 5, 84.

Sencilo, A., Paulin, L., Kellner, S., Helm, M., Roine, E., 2012. Related haloarchaeal pleomorphic viruses contain different genome types. Nucleic Acids Res. 40 (12), 5523–5534.

Sencilo, A., Jacobs-Sera, D., Russell, D.A., Ko, C.C., Bowman, C.A., Atanasova, N.S., Osterlund, E., Oksanen, H.M., Bamford, D.H., Hatfull, G.F., Roine, E., Hendrix, R.W., 2013. Snapshot of haloarchaeal tailed virus genomes. RNA Biol. 10 (5), 803–816.

Servin-Garcidueñas, L.E., Peng, X., Garrett, R.A., Martinez-Romero, E., 2013a. Genome sequence of a novel archaeal fusellovirus assembled from the metagenome of a mexican hot spring. Genome Announc. 1 (2), e0016413.

Servin-Garcidueñas, L.E., Peng, X., Garrett, R.A., Martinez-Romero, E., 2013b. Genome sequence of a novel archaeal rudivirus recovered from a mexican hot spring. Genome Announc. 1 (1), e00040–12.

Sheppard, C., Blombach, F., Belsom, A., Schulz, S., Daviter, T., Smollett, K., Mahieu, E., Erdmann, S., Tinnefeld, P., Garrett, R., Grohmann, D., Rappsilber, J., Werner, F., 2016. Repression of RNA polymerase by the archaeo-viral regulator ORF145/RIP. Nat. Commun. 7, 13595.

Simmonds, P., Adams, M.J., Benko, M., Breitbart, M., Brister, J.R., Carstens, E.B., Davison, A.J., Delwart, E., Gorbalenya, A.E., Harrach, B., Hull, R., King, A.M., Koonin, E.V., Krupovic, M., Kuhn, J.H., Lefkowitz, E.J., Nibert, M.L., Orton, R., Roossinck, M.J., Sabanadzovic, S., Sullivan, M.B., Suttle, C.A., Tesh, R.B., van der Vlugt, R.A., Varsani, A., Zerbini, F.M., 2017. Consensus statement: virus taxonomy in the age of metagenomics. Nat. Rev. Microbiol. 15 (3), 161–168.

Snyder, J.C., Samson, R.Y., Brumfield, S.K., Bell, S.D., Young, M.J., 2013. Functional interplay between a virus and the ESCRT machinery in archaea. Proc. Natl. Acad. Sci. U. S. A. 110 (26), 10783–10787.

Snyder, J.C., Bolduc, B., Young, M.J., 2015. 40 Years of archaeal virology: expanding viral diversity. Virology 479–480, 369–378.

Stedman, K.M., She, Q., Phan, H., Arnold, H.P., Holz, I., Garrett, R.A., Zillig, W., 2003. Relationships between fuselloviruses infecting the extremely thermophilic archaeon Sulfolobus: SSV1 and SSV2. Res. Microbiol. 154 (4), 295–302.

Stedman, K.M., DeYoung, M., Saha, M., Sherman, M.B., Morais, M.C., 2015. Structural insights into the architecture of the hyperthermophilic Fusellovirus SSV1. Virology

474, 105–109.

Stedman, K., 2011. Fusellovirus. In: Tidona, C., Darai, G. (Eds.), The Springer Index of Viruses, 2nd edThe Springer Index of Viruses, 2nd ed. Springer-Verlag, New York, pp. 561–566.

Strömsten, N.J., Bamford, D.H., Bamford, J.K., 2005. In vitro DNA packaging of PRD1: a common mechanism for internal-membrane viruses. J. Mol. Biol. 348 (3), 617–629.

Szymczyna, B.R., Taurog, R.E., Young, M.J., Snyder, J.C., Johnson, J.E., Williamson, J.R., 2009. Synergy of NMR, computation, and X-ray crystallography for structural biology. Structure 17 (4), 499–507.

Valentine, D.L., 2007. Adaptations to energy stress dictate the ecology and evolution of the Archaea. Nat. Rev. Microbiol. 5 (4), 316–323.

Veesler, D., Ng, T.S., Sendamarai, A.K., Eilers, B.J., Lawrence, C.M., Lok, S.M., Young, M.J., Johnson, J.E., Fu, C.Y., 2013. Atomic structure of the 75 MDa extremophile Sulfolobus turreted icosahedral virus determined by CryoEM and X-ray crystallography. Proc. Natl. Acad. Sci. U. S. A. 110 (14), 5504–5509.

Vik, D.R., Roux, S., Brum, J.R., Bolduc, B., Emerson, J.B., Padilla, C.C., Stewart, F.J., Sullivan, M.B., 2017. Putative archaeal viruses from the mesopelagic ocean. Peer J. 5, e3428.

Villanueva, L., Damste, J.S., Schouten, S., 2014. A re-evaluation of the archaeal membrane lipid biosynthetic pathway. Nat. Rev. Microbiol. 12 (6), 438–448.

Vorontsov, E.A., Rensen, E., Prangishvili, D., Krupovic, M., Chamot-Rooke, J., 2016. Abundant lysine methylation and N-terminal acetylation in sulfolobus islandicus revealed by bottom-up and top-down proteomics. Mol. Cell. Proteomics 15 (11), 3388–3404.

Wagner, C., Reddy, V., Asturias, F., Khoshouei, M., Johnson, J.E., Manrique, P., Munson-McGee, J., Baumeister, W., Lawrence, C.M., Young, M.J., 2017. Isolation and Characterization of Metallosphaera turreted icosahedral virus (MTIV), a founding member of a new family of archaeal viruses. J. Virol. 91 e00925-17.

Wang, Y., Duan, Z., Zhu, H., Guo, X., Wang, Z., Zhou, J., She, Q., Huang, L., 2007. A novel Sulfolobus non-conjugative extrachromosomal genetic element capable of integration into the host genome and spreading in the presence of a fusellovirus. Virology 363 (1), 124–133.

Wang, H., Peng, N., Shah, S.A., Huang, L., She, Q., 2015. Archaeal extrachromosomal genetic elements. Microbiol. Mol. Biol. Rev. 79 (1), 117–152.

Wang, Y., Sima, L., Lv, J., Huang, S., Liu, Y., Wang, J., Krupovic, M., Chen, X., 2016. Identification, characterization, and application of the replicon region of the halophilic temperate sphaerolipovirus SNJ1. J. Bacteriol. 198 (14), 1952–1964.

Wang, J., Liu, Y., Liu, Y., Du, K., Xu, S., Wang, Y., Krupovic, M., Chen, X., 2018. A novel family of tyrosine integrases encoded by the temperate pleolipovirus SNJ2. Nucleic Acids Res (in press).

Weidenbach, K., Nickel, L., Neve, H., Alkhnbashi, O.S., Kunzel, S., Kupczok, A., Bauersachs, T., Cassidy, L., Tholey, A., Backofen, R., Schmitz, R.A., 2017. Methanosarcina spherical virus, a novel archaeal lytic virus targeting methanosarcina strains. J. Virol. 91 (22), e00955–17.

Weigel, C., Seitz, H., 2006. Bacteriophage replication modules. FEMS Microbiol. Rev. 30 (3), 321–381.

Werner, F., Grohmann, D., 2011. Evolution of multisubunit RNA polymerases in the three domains of life. Nat. Rev. Microbiol. 9 (2), 85–98.

Wirth, J.F., Snyder, J.C., Hochstein, R.A., Ortmann, A.C., Willits, D.A., Douglas, T., Young, M.J., 2011. Development of a genetic system for the archaeal virus Sulfolobus turreted icosahedral virus (STIV). Virology 415 (1), 6–11.

Woese, C.R., Fox, G.E., 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. U. S. A. 74 (11), 5088–5090.

Yu, X., Jih, J., Jiang, J., Zhou, Z.H., 2017. Atomic structure of the human cytomegalovirus capsid with its securing tegument layer of pp150. Science 356 (6345), eaam6892.

Zhang, Z., Liu, Y., Wang, S., Yang, D., Cheng, Y., Hu, J., Chen, J., Mei, Y., Shen, P., Bamford, D.H., Chen, X., 2012. Temperate membrane-containing halophilic archaeal virus SNJ1 has a circular dsDNA genome identical to that of plasmid pHH205. Virology 434 (2), 233–241.