Extremophiles (2014) 18:473–489
DOI 10.1007/s00792-014-0634-0

REVIEW

Molecular biology of fuselloviruses and their satellites

Patrizia Contursi · Salvatore Fusco · Raffaele Cannio · Qunxin She

Received: 18 September 2013 / Accepted: 31 January 2014 / Published online: 23 February 2014

Abstract Fuselloviruses, also known as *Sulfolobus* Spindle-shaped viruses (SSVs), are "lemon"- or "spindle"-shaped double-stranded DNA viruses. Among them, SSV1, SSV2 and the satellite viruses pSSVx and pSSVi have been investigated at the structural, genetic, transcriptomic, proteomic and biochemical levels, thus becoming models for dissecting DNA replication/gene expression in *Archaea*. Important progress has been made including elucidation of temporal genome expression during virus infection and induction of replication, SSV1 lysogeny maintenance as well as differentially expression of pSSVx replicase. Future researches focusing on these model systems would yield insightful knowledge of life cycle and DNA replication of fuselloviruses.

Keywords Fuselloviridae · *Sulfolobus* · Satellite viruses · Transcription regulation

Communicated by S. Albers.

P. Contursi (✉) · S. Fusco
Dipartimento di Biologia, Università degli Studi di Napoli Federico II, Complesso Universitario Monte S. Angelo, Via Cinthia, Naples, Italy
e-mail: contursi@unina.it

R. Cannio
Unity of Rome, c/o Department of Chemistry, Institute of Biomolecular Chemistry (ICB), Sapienza University of Rome, Rome, Italy

Q. She (✉)
Department of Biology, Danish Archaea Centre, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen N, Denmark
e-mail: qunxin@bio.ku.dk

Introduction

Prokaryotes are the most abundant organisms on Earth but they are outnumbered tenfold by the viruses infecting them. Thus, prokaryotic viruses are the most predominant biological entities on Earth (Bamford 2003; Bamford et al. 2005; Le Romancer et al. 2007; Krupovic et al. 2011; Forterre and Prangishvili 2013). Shortly after the recognition of the *Archaea* (formerly *Archaebacteria*) (Woese et al. 1978; Albers et al. 2013), research focused on isolation of viruses and virus-like particles from extremely hot, low pH or hypersaline niches (Guixa-Boixareu et al. 1996; Oren et al. 1997; Rice et al. 2001; Rachel et al. 2002; Häring et al. 2005; Porter et al. 2007; Bize et al. 2008; Redder et al. 2009; Sime-Ngando et al. 2011; Snyder and Young 2011; Pietilä et al. 2013; Porter et al. 2013). Intriguingly, archaeal viruses exhibited completely novel and unique morphologies (Prangishvili and Garrett 2005; Lawrence et al. 2009; Pina et al. 2011; Ackermann and Prangishvili 2012; Peng et al. 2012; Prangishvili 2013) and appear to be adapted to the environments from which their hosts thrive (Snyder et al. 2003; Ortmann et al. 2006; Lawrence et al. 2009; Atanasova et al. 2012; Prangishvili 2013; Bartolucci et al. 2013), thus pointing to a process of host–virus coevolution within each biogeographic context (Held and Whitaker 2009). For example, for the viruses that infect hyperthermophilic organisms, their virions are also extremely thermostable and viral infection occurs most effectively at the optimal growth temperatures of their hosts (Schleper et al. 1992; Zillig et al. 1996; Ceballos et al. 2012; Quemin et al. 2013; Fu and Johnson 2012). This has led to the hypothesis that archaeal viruses are very ancient and may have preceded the separation of the three Domains of life (Prangishvili 2003; Forterre 2006; Ortmann et al. 2006) such that each Domain had already been

accommodated with a fraction of the existing diverse virus population when it was formed (Prangishvili et al. 2006b; Koonin et al. 2006).

Compared to bacteriophages that mainly show the typical head-to-tail morphotypes, archaeal viruses show a remarkable diversity in morphology, which imposed the establishment of a number of new virus families (Pina et al. 2011; Peng et al. 2012; Prangishvili 2013). To date, all known archaeal viruses contain double-stranded DNA (dsDNA) as genetic material except for two recent reports of single-stranded DNA viruses isolated from a haloarchaeaon (Pietilä et al. 2009) and from a crenarchaeon (Mochizuki et al. 2012). It is also worth mentioning that putative archaeal RNA viruses have been detected in a metagenomic study but such biological entities still remain to be obtained from an archaeal hosts (Bolduc et al. 2012; Stedman et al. 2013).

The majority of known archaeal viruses infect organisms belonging to *Creanarchaea*, including, but not limited to, members of the genera of *Sulfolobus*, *Acidianus*, *Thermoproteus*, *Aeropyrum* and *Pyrobaculum* (Prangishvili et al. 2006a, Prangishvili 2013). This attributes at least partly to the major effects of isolating archaeal viruses from thermophilic environments in which these archaeal organisms flourish. Archaeal virus research was initiated by Wolfram Zillig, Wolf-Dieter Reiter and colleagues (Reiter et al. 1987a, b, 1988a, b, 1989; Zillig et al. 1994, 1996; Albers et al. 2013) and followed by David Prangishvili, Roger Garrett, Mark Young and colleagues (Prangishvili et al. 2001; Prangishvili and Garrett 2004, 2005; Wiedenheft et al. 2004; Pina et al. 2011). In parallel, there were focused researches on developing *Sulfolobus* model organisms for studying novel biological principles in *Archaea*. In particular genome sequencing of *Sulfolobus solfataricus* P2 (She et al. 2001b) and *Sulfolobus islandicus* REY15A, HVE10/4 and LAL14/1 (Guo et al. 2011; Jaubert et al. 2013) as well as the development of microarrays containing probes of all host genes and of several viral genomes has allowed archaeal host–virus interaction to be investigated at genome scale using DNA microarrays (Fröls et al. 2007; Ortmann et al. 2008; Okutan et al. 2013; Ren et al. 2013) and RNA sequencing (Quax et al. 2013). These developments render *Sulfolobus* host–virus systems good models for studying molecular biology of archaeal viruses. Nevertheless, gene and protein sequences from crenarchaeal viruses are often orphans, i.e., they do not show detectable sequence similarity to proteins present in the public databases. As a result, functional annotation has not been performed for these gene products so far (Prangishvili et al. 2006a, b; Menon et al. 2010). For this reason, the comprehension of the fundamental viral processes in archaea, such as uptake to host cells, transcriptional regulation, genome replication virus assembly and release, is still at a stage of infancy.

Nearly all crenarchaeal viruses fall into eight families including *Fuselloviridae*, *Lipothrixviridae*, *Rudiviridae*, *Guttaviridae*, *Globuloviridae*, *Bicaudaviridae*, *Ampullaviridae* and *Clavaviridae* (Prangishvili et al. 2006a; Mochizuki et al. 2010; Pina et al. 2011; Peng et al. 2012) and two other families whose approval is still pending at the International Committee on Taxonomy of Viruses (ICTV) (Prangishvili 2013). Research on *Fuselloviridae* has been centered on a few model viruses. SSV1 (*Sulfolobus* spindle-shaped virus 1) is the first fusellovirus to be characterized (Martin et al. 1984), representing the prototype of this family for physiological, genetic and transcriptional studies (Reiter et al. 1987a, b; Reiter et al. 1988a, b, 1989; Schleper et al. 1992; Stedman et al. 1999; Fröls et al. 2007; Iverson and Stedman 2012). The second studied fusellovirus is SSV2, which shows distinct life cycles in the natural versus foreign hosts (Stedman et al. 2003; Contursi et al. 2006). Also investigated are the interactions of SSV2 with virus satellites, which are plasmid–virus hybrids that form virions by hijacking the viral packaging machinery of a helper virus (Arnold et al. 1999; Wang et al. 2007; Ren et al. 2013). Although SSV8 (aka SSV RH) has not been studied physiologically, a few proteins encoded in this viral genome have been characterized to reveal functional diversity of viral proteins. Interestingly, integration of SSV9 (aka SSV K1) occurs not only at tRNA locations, thus representing an exception to the evolutionary conserved process of integrating viruses and phages (Wiedenheft et al. 2004; Prangishvili 2013).

In this review, we focus on the current knowledge of the molecular and physiological features of the *Fuselloviridae* and how the acquired knowledge significantly expanded our understanding of archaeal biology.

An overview of fuselloviral morphology, diversity and evolution

To date, *Fuselloviridae* comprise nine members derived from four different geographic locations. The founding member SSV1 was isolated from a hot spring in Beppu, Japan (Martin et al. 1984). Subsequently, five fuselloviruses (SSV2, SSV4, SSV5, SSV6, and SSV7) were obtained from different geothermal environments in Iceland (Stedman et al. 2003; Redder et al. 2009). Moreover, Mark Young and colleagues have isolated several new archaeal viruses including SSV9 from a hot spring sample from Kamchatka (Russia) and SSV8 from the Yellowstone National Park (USA; Wiedenheft et al. 2004). The only non-*Sulfolobus* fusellovirus isolated thus far is ASV1 (*Acidianus* spindle virus 1), which was obtained from

Table 1 Features of all known fuselloviruses

Virus name	Sampling site	Genome size (bp)	ORFs number	NCBI number	Reference
SSV1	Japan	15,465	35	NC_001338	Palm et al. 1991
SSV2	Iceland	14,796	35	NC_005265	Stedman et al. 2003
SSV4	Iceland	15,135	34	EU030938	Redder et al. 2009
SSV5	Iceland	15,330	34	EU030939	Redder et al. 2009
SSV6	Iceland	15,684	33	NC_013587	Redder et al. 2009
SSV7	Iceland	17,602	33	NC_013588	Redder et al. 2009
SSV8	USA	16,473	37	NC_005360	Wiedenheft et al. 2004
SSV9	Kamchatka	17,385	31	NC_005361	Wiedenheft et al. 2004
ASV1	USA	24,186	38	NC_013585	Redder et al. 2009

Acidianus brierleyi, a thermoacidophilic archaeon isolated from a hot spring in the Yellowstone National Park, USA (Redder et al. 2009) (Table 1).

Whereas all the known *S. solfataricus* strains and some icelandic species such as *S. islandicus* HVE10/4 can propagate efficiently most of the SSV viruses, SSVs infection is not effective on *S. islandicus* RENH1, *S. acidocaldarius* and *S. tokodaii* species. On the other hand, SSVs display significant variations in their host ranges with SSV1 exhibiting the narrowest and SSV8 exhibiting the broadest host range, respectively. Therefore, SSV infectivity and *Sulfolobus* susceptibility are not related to the geographical context from which the hosts and viruses were isolated (Held and Whitaker 2009; Ceballos et al. 2012).

These nine genomes provide geographically distinct isolates that have been used in comparative genomic and morphological analyses (Wiedenheft et al. 2004; Redder et al. 2009).

The genome of *Fuselloviradae* contains dsDNA ranging from ca. 15–22 kb in size that is encased by 2–3 structural proteins (VP1, VP2 and VP3) to form spindle-shaped viral particles (Fig. 1a). In particular, for SSV1 it has been experimentally demonstrated that the capsid is formed by the interaction of VP1, VP2 and VP3 proteins (Reiter et al. 1987a). Nevertheless, most of fuselloviral genomes contain only genes encoding for VP1 and VP3 (Redder et al. 2009). Under electron microscope, virions of most known fuselloviruses appear as 55−60 × 80−100 lemon-shaped particles (Stedman et al. 2003; Wiedenheft et al. 2004; Redder et al. 2009). However, two fuselloviruses SSV6 and ASV1 show pleomorphic and malleable morphologies, varying from a pear-like shape to an elongated cigar shape (Fig. 1b, c). A close examination of these virions revealed filaments attached to one of the pointed ends (Martin et al. 1984; Palm et al. 1991; Schleper et al. 1992; Wiedenheft et al. 2004; Redder et al. 2009), which are likely to be implicated in anchoring the viruses to the host membrane upon infection as well as to cause virus clustering into rosette formations as seen in cultures of almost all the isolates.

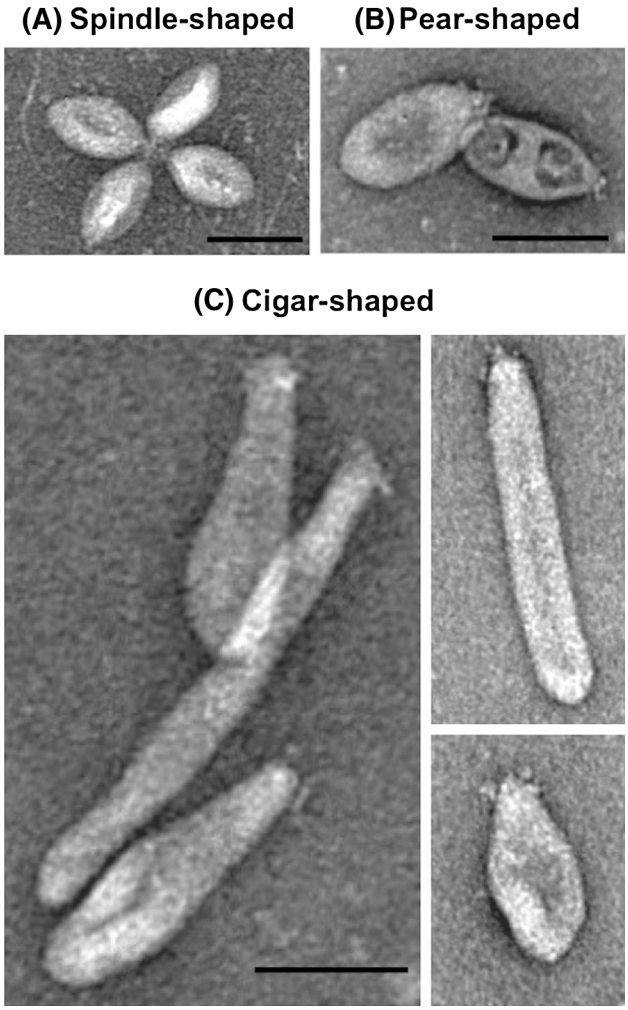


Fig. 1 Electron microscopy images of *Fuselloviridae* virions. **a** Exemplification of the typical lemon-/spindle-shaped morphology, which is displayed by SSV1, SSV2, SSV4, SSV5, SSV7, SSV8 and SSV9. **b**, **c** Viral particles of ASV1 and SSV6 exhibiting a wide range of morphotypes varying from the pear shape to the elongated cigar shape. *Scale bars* 100 nm. Reproduced with permission from Redder et al. 2009

Two different variants in the number and in the structure of these filaments have been discovered: multiple, thin, quite sticky filaments, readily linking virions are present in all

the *Fuselloviridae* with the exception of SSV6 and ASV1. The latter two, indeed, carry 3–4 thick and slightly curved short filaments, forming a crown around the virion tip. It seems that the difference in the numbers and structure of the terminal ends in the two morphotypes correlates with the presence of two different variants of the longest ORF (open reading frame) present on the fuselloviral genomes, i.e., the SSV1-C792 or the SSV6-B1232 module, respectively (Redder et al. 2009). Despite the pleomorphism in virions, all fuselloviral genomes show a high similarity in gene synteny indicating a clear relationship among the viral isolates. As is typical of many viral genomes, the predicted ORFs are tightly arranged on the genome with very small intervening non-coding sequences. Sequence alignments of most of the ORFs show little or no sequence similarity in the public databases. However, most of the fuselloviral ORFs are homologoues to each other and are also collinearly organized (Wiedenheft et al. 2004). A very illustrative picture of the fuselloviral genomes comparison is reported in the paper of Redder et al. 2009. Thirteen core genes are conserved among all members, which are also considered as the minimal genetic signature of this virus family and may represent viral functions common to all fuselloviruses, despite their different geographic isolation (Wiedenheft et al. 2004; Redder et al. 2009; Held and Whitaker 2009; Krupovic et al. 2012). The core set of ORFs includes *vp1* and *vp3* coding for the capsid proteins, three putative transcriptional regulators, six orphan proteins and a type I tyrosine recombinase that facilitates provirus formation (Argos et al. 1986; Serre et al. 2002). This common set of 13 ORFs shared by all the nine isolates is not contiguous. Nevertheless, it is noteworthy that conserved ORFs are mainly confined to one half of the SSV virus genomes and are arranged in the same orientation. Conversely, the other half of the genomes is more divergent and ORFs unique to any one fuselloviral isolates may reflect their evolutionary history, geographic isolation, requirements for replication in their specific hosts or adaptations to features of their thermal environments (Wiedenheft et al. 2004; Redder et al. 2009).

In the less conserved half of the genomes, a clustering of ORFs is evident containing cysteine codons that are unusual for hyperthermophiles (Stedman et al. 2003). Further analyses have indicated that there is a general abundance of disulfide bonds in the intracellular proteins of hyperthermophilic viruses but a decreased cysteine content in their membrane proteins, which are often late gene products of hyperthermophilic viruses (Menon et al. 2008). Therefore, it has been reasoned that such a genomic organization for fuselloviruses more likely reflects the clustering of late genes of these viruses rather than a fusion event of two genomes with distinct histories (Menon et al. 2008).

Apparently there are evolutionary constraints to maintain gene synteny for fuselloviruses, which can be exemplified with the comparison between SSV1 and SSV2. Although isolated from hot springs of geographically distant locations, these two viruses show collinear genomic organization, with the majority of their ORFs being homologous to each other (Stedman et al. 2003). Nevertheless, SSV1 and SSV2 exhibit interesting differences in physiology, life cycle and relationship with the relative hosts (see below), which reflect evolutionary adaptation to their hosts and/or to the environment.

A mechanism has been proposed for the evolution of fuselloviruses invoking homologous recombination between two integrated viruses. The hypothesis is based on the fact that closely related fuselloviruses possess multiple highly similar or even identical regions. For instance, SSV4 and SSV5 share an almost identical 7.9 kb DNA region. If co-integration of two viral genomes into the same site generates concatemers of different viruses, their excision can occur at multiple positions leading to diverse virus variants (Redder et al. 2009).

SSV1: the prototype of the *Fuselloviridae*

SSV1–host interaction

SSV1 was first isolated as a plasmid present in *S. shibatae* B12 (Yeats et al. 1982). Then virus-like particles (VLP) were observed and their production was found to be UV-inducible (Martin et al. 1984). Subsequently, it was found that the VLP infected *S. solfataricus* (Schleper et al. 1992), thereby demonstrating the viral nature for SSV1.

SSV1 episomal DNA is stably maintained in host cells in three different forms: the DNA is either positively or negatively supercoiled or relaxed double stranded (Nadal et al. 1986). Whereas the integrated form might be used as a template for transcription, the positively supercoiled is likely encased in the viral particle, while the negatively supercoiled might be used for genome replication (Snyder et al. 2003). SSV1 is the only known member of the *Fuselloviridae* family that shows an UV-inducible gene expression and genome replication. Cells are not lysed upon virus particle release and recover their growth rate as well as their lysogenic state within few hours (Schleper et al. 1992).

The complete genome of SSV1 was sequenced (Palm et al. 1991) and the integrase that furthers site-specific integration has been biochemically characterized (Muskhelishvili et al. 1993; Serre et al. 2002; Zhan et al. 2012). The integration occurred site specifically at an arginyl-tRNA gene and an intact host gene is maintained after integration (Reiter et al. 1989; Muskhelishvili et al. 1993;

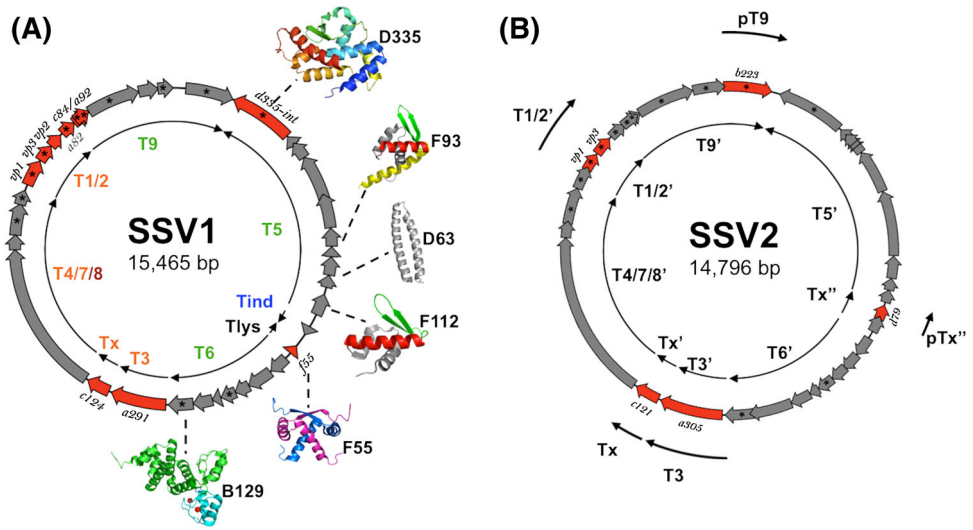


Fig. 2 Genomic map of SSV1 and SSV2. ORFs labeled with the *asterisk* belong to the core set of 13 genes conserved in all *Fuselloviridae* (Redder et al. 2009) ORFs in *red* are those expressed in the lysogenic state of both SSV1 and SSV2. *Inner black arrows* represent the identified transcripts of the fuselloviruses. The transcriptional map has been experimentally determined for SSV1 (Reiter et al. 1987b; Fröls et al. 2007) and deduced from microarray data for SSV2 (Ren et al. 2013). **a** The three-dimensional structures of SSV1 proteins (determined for D63, F93, F112, B129 and D355-int C-terminal domain and predicted for F55) are shown. **b** The transcriptional map of SSV2 in stable *S. solfataricus* transfectants deduced from the expression pattern of the ORFs is indicated by *the outer black arrows*. pT_9 and pT_x'' stand for partial T_9 and T_x'' transcripts

Serre et al. 2002). Conversely, since the attachment site is located within the *int* gene, integration results in the partitioning of the gene that inactivates the expression of the encoded integrase. This mechanism accounts for generating gene capture in archaeal genomes (She et al. 2001a). Interestingly, the integrase gene has been shown to be dispensable for virus replication as well as for its spreading into *Sulfolobus* cultures, thus demonstrating that the viral integration is an optional step in the replication of SSV1 and this is probably true for all fuselloviruses (Clore and Stedman 2007). Although Clore and Stedman showed that the wild-type SSV1 outcompeted the virus without an integrase gene, a positive pressure for the maintenance of these genes has been hypothesized, since integrase-encoding sequences are widespread in fuselloviruses genomes as well as in other integrative genetic elements (She et al. 2001a, 2002, 2004; Peng et al. 2000; Peng 2008; Cortez et al. 2009). Obviously, integrase genes have played an important role in horizontal gene transfer and genome evolution.

SSV1 gene expression: insights into promoters and terminators of archaeal transcription

The regulation of gene expression on SSV1 was studied immediately after the isolation of this genetic element. First, constitutive and UV-inducible transcripts were identified, including an UV-inducible transcript (T_{ind}), and nine constitutive messengers (T_1–T_9) (Reiter et al. 1987a, 1988a, Zillig et al. 1988). More detailed analysis using microarrays showed that the expression of these viral transcripts is temporally regulated such that they fall into four distinct classes: (1) UV-inducible (T_{ind}), (2) early (T_5, T_6 and T_9), (3) late (T_3, $T_{4/7}$), and (4) late-extended ($T_{4/7/8}$) messengers. Moreover, an additional late monocistronic transcript was detected and mapped (T_x) (Fig. 2a). This fashion of regulation is reminiscent of that adopted by many bacteriophages and eukaryotic viruses (Fröls et al. 2007).

Then, transcriptional start sites (TSSs) were identified for all the SSV1 transcripts (Reiter et al. 1988a), which facilitated the identification of promoter elements of SSV1. Two conserved sequence motifs, including a TATA-box-like hexanucleotide sequence (box A, TTTAAA) that is centered about 26 nucleotides upstream of the TSS and a trinucleotide sequence (box B; consensus sequence TGA) that is localized nearby the TSS, were identified (Reiter et al. 1988a). Strikingly, these elements resemble those of the eukaryotic basal gene promoters recognized by the eukaryotic RNA polymerase II (Reiter et al. 1988a) and this is consistent with the results of the similar subunit composition of the eukaryotic and archaeal RNA polymerases unraveled from another study (Huet et al. 1983).

Analyzing transcription termination signals of these SSV1 transcripts has identified a conserved motif with the consensus TTTTTYT (Reiter et al. 1988b). Once again these sequences resemble the pyrimidine-rich elements that are implicated in transcription termination by eukaryotic

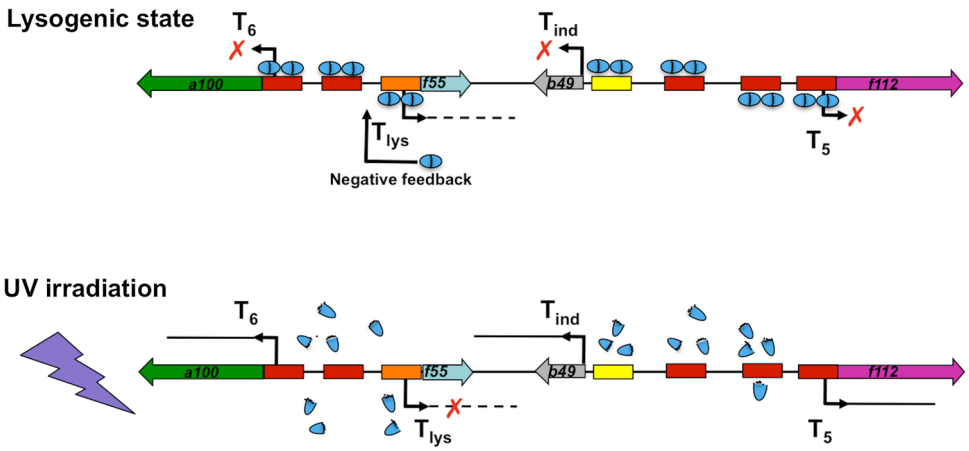


Fig. 3 Model of the F55 interaction at its binding sites as modified from Fusco et al. 2013. In the lysogenic state, F55 (*cyan ovals*) binds as dimers to the target sequences in the promoters of T_5, T_6, T_{ind} (*red and yellow boxes*) as well as to its own promoter (*orange box*). As shown by *red crosses*, transcription of T_5, T_6 and T_{ind} is locked, while the expression of its own gene is progressively turned off following a negative feedback control. Upon UV irradiation, F55 might be degraded and/or inactivated and consequently transcription of the early T_{ind}, T_5 and T_6 transcripts is unlocked

RNA polymerase I (Grummt et al. 1985; Reiter et al. 1988b).

More recently, our research has gained some insights into the molecular components and the mechanisms underpinning the maintenance of the SSV1 lysogeny. Transcription analysis of an SSV1-harboring *S. solfataricus* strain, in the absence of the UV stimulus (lysogenic state), showed the expression of a novel SSV1 transcript, named T_{lys}. This messenger is expressed from a genomic region located nearby the UV-inducible T_{ind} (Fig. 2a), which is in turn totally repressed under the same conditions. T_{lys} encodes for a 6.3-kDa protein, F55, which is able to interact specifically at operator sequences in the promoters of the early transcripts (T_{ind}, T_5 and T_6) as well as of its own messenger. F55 is predicted to bear the ribbon–helix–helix fold typical of negative transcription regulators. Therefore, by binding to its target sequences, F55 might exert down-regulation of the early genes in the lysogenic state. By analogy with the cI repressor of lambda phage, F55 is possibly degraded and/or inactivated upon UV irradiation, thus unlocking the transcriptional circuit of the early genes. The hypothesis of SSV1 induction and the molecular components involved in the SSV1 life cycle are illustrated in Fig. 3 (Fusco et al. 2013).

Structural and functional analyses of SSV1 proteins

There are 34 predicted ORFs on the SSV1 genome, encoding protein products that range from 6 to 86 KDa and about 75 % of these ORFs have not been reliably identified by bioinformatic approaches. Function is only known for the viral structural proteins (VP1, VP2 and VP3) (Reiter et al. 1987a) and for the integrase D335, whose structural analysis of the C-terminal catalytic domain has revealed that it possesses a core fold similar to those of type I tyrosine recombinases of both bacterial and eukaryal origin. In vitro studies showed that this enzyme is capable of transferring a phosphodiester bond from host to viral attachment sites (attA and attP, respectively) as well as of carrying out the reverse reaction (Muskhelishvili et al. 1993; Serre et al. 2002; Eilers et al. 2012; Zhan et al. 2012). VP2 is thought to function as a small, packaged, DNA-binding protein (Stedman et al. 2003; Wiedenheft et al. 2004). Two additional components that copurify with the viral particles have also been identified, C792 and D244 (Menon et al. 2008). C792 is a predicted membrane protein serving probably a structural role by generating filaments at the tail end of virions implicated in host receptor (Redder et al. 2009), while D244 is a soluble protein of unknown function (Menon et al. 2010) whose homolog from the SSV8 virus has been studied at structural level revealing a possible role in DNA replication, repair, or recombination (see below).

B251 exhibits limited similarity to DnaA (Koonin 1992). Furthermore, molecular modeling allowed structure prediction for four SSV1 proteins. While E51, C80 and F55 might adopt the ribbon–helix–helix fold, A45 and A79 carry a C2H2 zinc finger-like motifs (Prangishvili et al. 2006b). However, with the exception of F55, their DNA-binding activity remains to be tested and therefore their role in the virus life cycle is murky.

The general lack of insights into protein functions from sequence analyses stimulated the X-ray crystallographic approach to determine structures of SSV1 proteins. Indeed, structural analysis of viral proteins provides an important alternative approach to obtain a deeper comprehension of

the life cycle of SSV1 and crenarchaeal viruses in general (Lawrence et al. 2009; Krupovic et al. 2012). Beside the C-terminal domain of the integrase, 3D structures are available for 4 SSV1 ORFs, all encoding putative transcriptional factors. In the structure of D63 (Kraft et al. 2004a), the helix-turn-helix fold resembles that of the "repressor of primer" (ROP), an adaptor protein that regulates colE1 plasmid copy number in *E. coli* (Helmer-Citterich et al. 1988). F93 shows a fold similar to proteins belonging to the MarR (Alekshun et al. 2001; Di Fiore et al. 2009; Fiorentino et al. 2011) and SlyA (Wu et al. 2003) subfamilies of winged-helix transcription regulators (Kraft et al. 2004b). The winged-helix protein F112 shows structural homologies with DP2 (Gibrat et al. 1996), a regulator that plays a central role in the eukaryotic cell cycle (Menon et al. 2008). Finally, B129 encodes a C2H2 Zinc finger transcription factor (Lawrence et al. 2009).

However, functions of these SSV1 regulators as well as of other gene products of unknown function remain to be dissected by genetic analysis. Stedman and colleagues have made first attempts to reveal essential genes of the SSV1 virus (Stedman et al. 1999).

In a subsequent study, three ORFs were deleted from the SSV1 genome, i.e., the universally conserved *b129*, the well-conserved *d244* and the poorly conserved *vp2*. The inactivation of the ORF encoding for the predicted transcriptional regulator B129 resulted in a loss of infectivity, while deletion of the one encoding the predicted DNA-binding protein VP2 yielded viable viruses that were indistinguishable from the wild-type. Interestingly, a new phenotype was observed for a *d244*-defective mutant, i.e., it was able to slow down the growth of the host in liquid culture (Iverson and Stedman 2012).

SSV2 and its satellites

Life cycle

SSV2 and its hosts have been subjects of physiological characterization for the study of fusellovirus–hosts interactions. SSV2 was isolated from *S. islandicus* REY 15/4 together with pSSVx, a virus satellite (Arnold et al. 1999) and as for SSV1, this virus also replicates in *S. solfataricus* P2 and infected cells carry a provirus in the chromosomes site specifically integrated at a $tRNA^{Gly}$ gene (Contursi et al. 2006).

Differently from SSV1, which exhibits a dramatic UV-inducible virus production, SSV2 apparently does not and shows only a modest induction of 2-fold (Stedman et al. 2003). Instead, SSV2 exhibits a physiological induction of virus replication dependent on the growth (Contursi et al. 2006). Thus, both viruses are interesting models for studying molecular biology of archaeal viruses.

SSV2 is very similar to SSV1 in virion morphology and in gene synteny and in sequence, among a total of 34 ORFs of SSV2, 26 show significant similarity to SSV1 ORFs (Stedman et al. 2003). SSV2 carries a direct repeat of 62 base pairs, with the core sequence repeated 4.5 times, which could be regarded as a hallmark of the SSV2 genome as it is absent from all other known fuselloviruses (Stedman et al. 2003, Wiedenheft et al. 2004). In comparison, the T_{ind} and T_{lys} transcripts are the hallmarks of SSV1 (Fig. 2a). Together, these differences may account for the lack of ultraviolet induction of SSV2 and for an elevated level of virion production detected in the lysogenic state of SSV2-infected *S. solfataricus* cells (see below).

Another striking difference is that SSV2 shows a broader host range compared to SSV1. Among the Icelandic *Sulfolobus* species susceptible to SSV2 infection (Ceballos et al. 2012) there is *S. islandicus* REY15A (Contursi et al. 2006), a strain isolated from the same enrichment culture of its natural host (Arnold et al. 1999). *S. islandicus* 15A and *S. solfataricus* P2 have been chosen to study the virus–host interaction in closely related and foreign systems, respectively.

The life cycle of SSV2 in the natural host is characterized by induction of virus replication at a later growth stage. It has been hypothesized that the induction is triggered by an unidentified signal molecule or by a transcriptional factor that exhibits growth phase regulation. Viral DNA replication upon induction shows the following features: (1) a steep increase of the SSV2 copy number (about 50 folds) within 4 h; (2) a concurrent inhibition of the host growth at a late exponential growth phase; (3) a consistent packaging and extrusion of the viral particles and (4) the reversibility of the replication induction and of the growth inhibition (Contursi et al. 2006).

The induction of SSV2 replication also occurs for the infected *S. islandicus* REY15A. However, the viral replication is not induced in SSV2-infected *S. solfataricus* cells, suggesting that the SSV2 physiological induction has resulted from mutual interactions that only exist for SSV2 and certain *S. islandicus* strains.

The interplay between SSV2 and *S. solfataricus* during the process of the infection has been studied through global analysis of the gene expression of the viral and host genomes in primary infected cells (Ren et al. 2013). In this case, the activation of SSV2 genes follows a chronological scheme based on a distributive pattern with all the genes transcribed within 7.5 h post-infection (Fig. 2b). More investigations are required to illustrate the involved mechanisms.

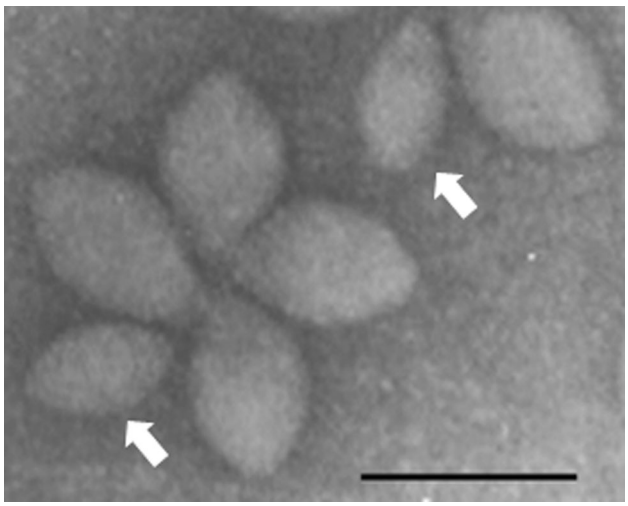


Fig. 4 Electron microscopy image of SSV2 and pSSVi viral particles. *White arrows* point to the smaller pSSVi-containing virions. *Scale bar* 100 nm. Reproduced with permission from Wang et al. 2007

Two hybrid virus/plasmid genetic elements, denoted as pSSVx (Arnold et al. 1999) and pSSVi (Wang et al. 2007), are packaged into small spindle-shaped virion particles (60 nm x 40 nm) in the presence of SSV1 or SSV2 as helpers (Fig. 4), thus representing the only two satellite-helper virus systems known in the archaeal domain. Interesting insights have been gained into interactions between these genetic elements as well as between them and their hosts.

Interplays between SSV2 and virus satellites

Archaeal virus satellites were first identified by analyzing virus-like particles present in an *S. islandicus* REY15/4 culture in which spindle virions of two distinct sizes were observed: the larger and the smaller ones contained the genome of SSV2 and pSSVx, respectively. In this system, SSV2 is an ordinary virus and acts as a helper to pSSVx such that the latter hijacks the virus packaging machinery for its own virion formation (Arnold et al. 1999; Stedman et al. 2003).

pSSVx has seven ORFs, 3 of which are viral-like while the remaining 4 are plasmid-like. The viral-like genes *orf154* and *orf288* exhibit high sequence similarity to two of the 13 core genes well conserved among all known SSV viruses (Redder et al. 2009) and the third gene, *orf-c68*, is only present in virus satellites and integrated elements (Contursi et al. 2007; Wang et al. 2007). Strikingly, pSSVx devotes ca. half of its coding capacity to code for *orf892*, a homolog of the putative replicase of pRN1 and for three additional plasmidic ORFs, ORF51 (a putative copy number control protein), ORF91 (containing a putative zinc-binding motif) and ORF76 (a putative leucine-zipper protein), the latter three belonging to the category of DNA-binding proteins/transcriptional regulators (Fig. 5a). Taken together, this suggests that pSSVx replicates its genome in a fashion similar to plasmids of the pRN family (Keeling et al. 1996, 1998; Peng et al. 2000; Kletzin et al. 1999).

Putative replicases encoded in the pRN family have only been characterized for *S. islandicus* pRN1 plasmid (ORF904). The protein is a multiple domain enzyme: the N-terminal domain shows primase and DNA polymerase activities and the C terminus contains the winged-helix DNA-binding domain displaying helicase/ATPase activity (Lipps et al. 2003). The latter is necessary for DNA unwinding at the replication origin. To date, a detailed mechanism remains to be illustrated for the DNA replication of pSSVx and other plasmids of this family.

The second virus satellite, the pSSVi (Fig. 5b), has been identified in the genome of a laboratory strain of *S. solfataricus* P2 where it was integrated into the $tRNA^{Arg}$ gene. Integration is catalyzed by an integrase encoded by its own genome and the integrated form is stably maintained and propagated over several generations (Wang et al. 2007). This condition is perturbed by the co-presence of SSV2 (or SSV1), leading to the active replication of the pSSVi episomal form and to its packaging into spindle-shaped virus particles (Ren et al. 2013). Unlike pSSVx, its putative replicase gene is clearly different from those encoded by pRN plasmids (Wang et al. 2007). Characterization of the encoded enzyme shows that it is a DNA helicase belonging to the superfamily 3 helicase which interacts with the host primase in vitro, suggesting it could function as the replication initiator of pSSVi (Guo and Huang 2010). Interestingly, among the integrated elements identified in archaeal genomes, pSA2 and pST1 exhibit a genetic organization similar to that of pSSVi, suggesting they could also represent integrated virus satellites in *Sulfolobus* genomes (She et al. 2002, 2004; Cortez et al. 2009).

To date, mechanisms of virions satellite formation in *Archaea* are mainly inferred from bioinformatic analyses of their genomes. Initially, three viral-like ORFs of pSSVx, ORF-c68, ORF154 and ORF288, were supposed to be implicated in the virion formation. However, the absence of any homologues of the pSSVx ORF154 and ORF288 in the pSSVi genome suggests that these two ORFs are not essential for the formation of satellite virions (Wang et al. 2007). This raises the possibility that the packaging machinery of a fusellovirus recognizes a sequence motif to start virus packaging. Interestingly, the sequence motif 5′-AAGGGAAANAGNA-3′ is present in the genomes of pSSVx, pSSVi and SSV2 (989–1001, 2412–2424, 1406–1418 bp on their linear map, respectively), which is

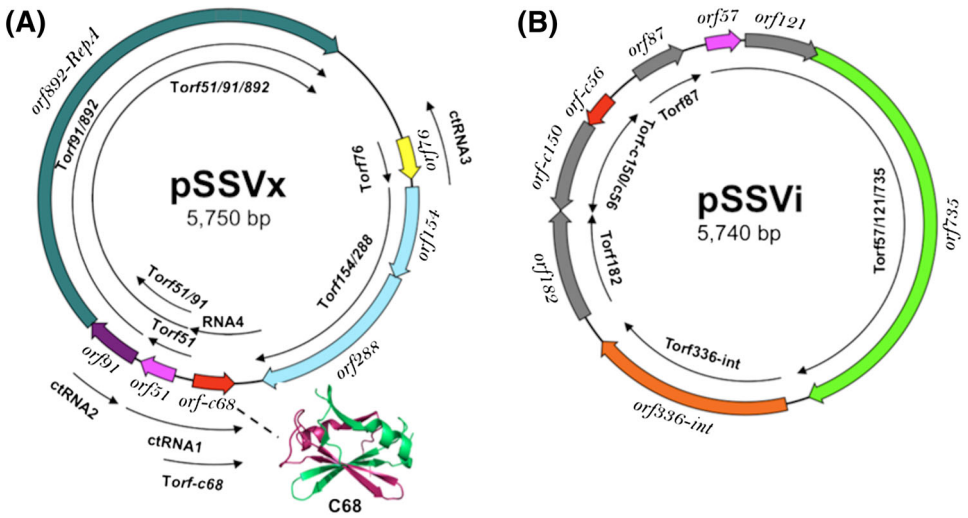


Fig. 5 Genomic map of the satellites pSSVx and pSSVi. ORFs homologues between the two genomes are filled in the same color. The pSSVx transcriptional map has been experimentally determined (by Northern blot and Primer extension) (Contursi et al. 2007, 2010) while that of pSSVi has been deduced from microarray data (Ren et al. 2013). **a** The biggest ORF (*dark green*) encodes for a Primase/Replicase which is homologue to that of the pRN family. The ORFs filled in *light blue* encode for proteins involved in DNA prepackaging and are homologues to two of the 13 ORFs of the *Fuselloviridae* core set. *orf91* and *orf51* encode for putative transcriptional factors, while *orf-c68* and *orf76* encode for experimentally determined transcriptional factors (Contursi et al. 2011 and manuscript submitted). **b** *orf735*, filled in green, encodes for the pSSVi replicase, the putative integrase is encoded by *orf336*, while ORFs in *gray* are orphans

absent from all known pRN plasmids that replicate autonomously.

An additional insight into the formation of satellites viral particles has been gained from another study of *Sulfolobus* genetic elements in which pXZ1, a plasmid that carries an SSV1-type integrase gene and co-exists with SSV4. This plasmid does not have the capability of forming virions but it carries an integrase gene (*int*) that is highly similar to the *int* gene of SSV4; a stretch of 500-bp DNA sequence is identical for the two *int* genes except for one single mismatch at the integration attachment sites, which guides the integration of the plasmid and virus into different tRNA sites in the host genome (Peng 2008).

Together, the genetic elements discussed above, including pSSVx, pSSVi, pXZ1, SSV2 and SSV4, provide very suitable materials to unravel the mechanism of virion packaging for fuselloviruses and their satellites.

Tripartite interactions among SSV2, pSSVi and their host

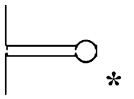
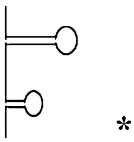
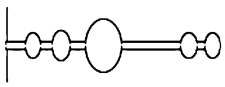
Interactions between pSSVi, SSV2 and the host *S. solfataricus* were investigated by transcriptomic analysis through microarrays. In the pSSVi-containing *S. solfataricus* P2 cells, the integrated form of pSSVi is propagated passively over generations. All pSSVi genes were expressed constitutively at a low level except for the integrase gene, which is partitioned in the host genome. Upon infection with SSV2 (or SSV1), pSSVi replication is activated, producing multiple copies of the episomal DNA on which an intact *int* gene is yielded. Under this condition, expression is elevated strongly for all pSSVi genes. Interestingly, it appears that there is a mutual interaction between the two genetic elements such that the replication of pSSVi facilitates that of its helper virus since the copy number of the latter also increases (Ren et al. 2013).

Notably, elevated expression of pSSVi genes correlates with an enhanced level of the expression from SSV2 early genes. Also the expression level of some host genes is elevated, including that encoding the RNA polymerase subunit F, RpoF. Therefore, it is plausible to suggest that pSSVi stimulates transcription of SSV2 early genes by regulating the expression level of RpoF and/or of other host factors (Ren et al. 2013).

pSSVx: a model for studying gene regulation circuits in *Archaea*

Similar to SSV2, pSSVx exhibits inducible DNA replication in the natural host at a late growth phase. We have demonstrated that the pSSVx genes are differentially expressed during its life cycle, providing a good model for studying regulation of gene expression in *Archaea* (Contursi et al. 2007, 2010). Three putative transcriptional factors seem to be implicated in the regulation of gene expression at the *rep* locus including ORF-c68, the virus satellite-specific factor, and ORF51 and ORF91, the plasmidic DNA-binding proteins (Fig. 5).

Table 2 Features of the pSSVx transcripts as experimentally determined by Northern blot, Primer extension and RT-PCR

Transcript	Transcription start site	Transcription termination sites	Encoded protein	Expressed in the plasmidic phase	Expressed in the viral phase	Promoter feature	Termination signal
T_{orf76}	1095	1325/1405	Transcription factor	Y	N	Canonical	
$T_{orf154/288}$	1367	3091/3150	Proteins involved in DNA packaging	Y	Y	Canonical	ttttttcatttt
T_{orfc68}	3027	2676/2710	Swapped-hairpin-like transcription factor	Y	Y	Canonical	ttatttttccgttattt
T_{orf51}	3150	3343	Putative CopG protein	Y	Y	Atypical	
$T_{orf51/91}$	3150	3550/3612	Putative CopG protein/Putative Zinc-motif protein	Y	Y	Atypical	
$T_{orf91/892}$	3338	669/709	Putative Zinc-motif protein/ RepA	N	Y	Atypical	Ttttttttctttt *
$T_{orf51/91/892}$	3150	669/709	Putative CopG protein/Putative Zinc-motif protein/RepA	N	Y	Atypical	ttttttttctttt
ctRNA1	3341	2710/2676	X	Y	N	Canonical	ttatttttccgttattt
ctRNA2	3557	3370/3324	X	Y	N	Canonical	ttttttactccatttt
ctRNA3	1383	2023/1056	X	Y	Y	Atypical	tttcgcctttt
RNA4	2686	3091/3150	X	Y	Y	Canonical	ttttttcatttt

The hypothesized DNA-binding sites for the CopG protein are the promoters Pr1 and/or Pr2, whereas ORF-c68 factor binds to its own regulatory sequence (Pr3)

The termination signals marked with an asterisk are predicted

A systematic study on initiation and termination of gene transcription for the pSSVx genome has been conducted. TSSs have been mapped for all the transcripts based on which archaeal canonical promoters as well as atypical promoter sequences have been identified (Table 2). The overlapping arrangement of genes increases the number of proteins that may be produced from a small genome but requires more complex regulatory mechanisms for the modulation of gene expression (see below).

Analyzing the expression of pSSVx genes in the natural host has revealed that they exhibit a temporal regulation. At an early growth phase (plasmidic phase) pSSVx is strictly controlled to a low copy number and transcripts that are specifically detected in this phase, i.e., ORF76, ctRNA1 and ctRNA2, are likely to be important for the stringent control of plasmid replication (Table 2). When pSSVx replication is induced, the expression from the putative replicase operon is strongly elevated and this occurs concurrently with the highest copy number attained by pSSVx (viral phase) (Fig. 6). Furthermore, the expression of three additional transcripts (*Torf51*, *Torf154/orf288* and RNA4) appears to be proportional to pSSVx copy number throughout the entire growth of the host (Contursi et al. 2007, 2010).

Northern analysis of *Torf892-RepA* expression indicates that *Torf51/orf91/orf892-RepA* and *Torf91-orf892-RepA*, the two *rep* transcripts, are only expressed at a later growth phase when virus replication is induced. Since ctRNA1 and ctRNA2 are oppositely transcribed to *orf51* and *orf91* and are highly expressed at an early growth phase of host cells, they are likely to be implicated in the inhibition of the replicase expression at this stage, possibly causing premature termination of *rep* transcripts. Indeed, as revealed from our transcriptional data, transcription terminates at the 3′ ends of *Torf51* and *Torf51-orf91* messengers when ctRNA1 and ctRNA2 are expressed at a consistent high

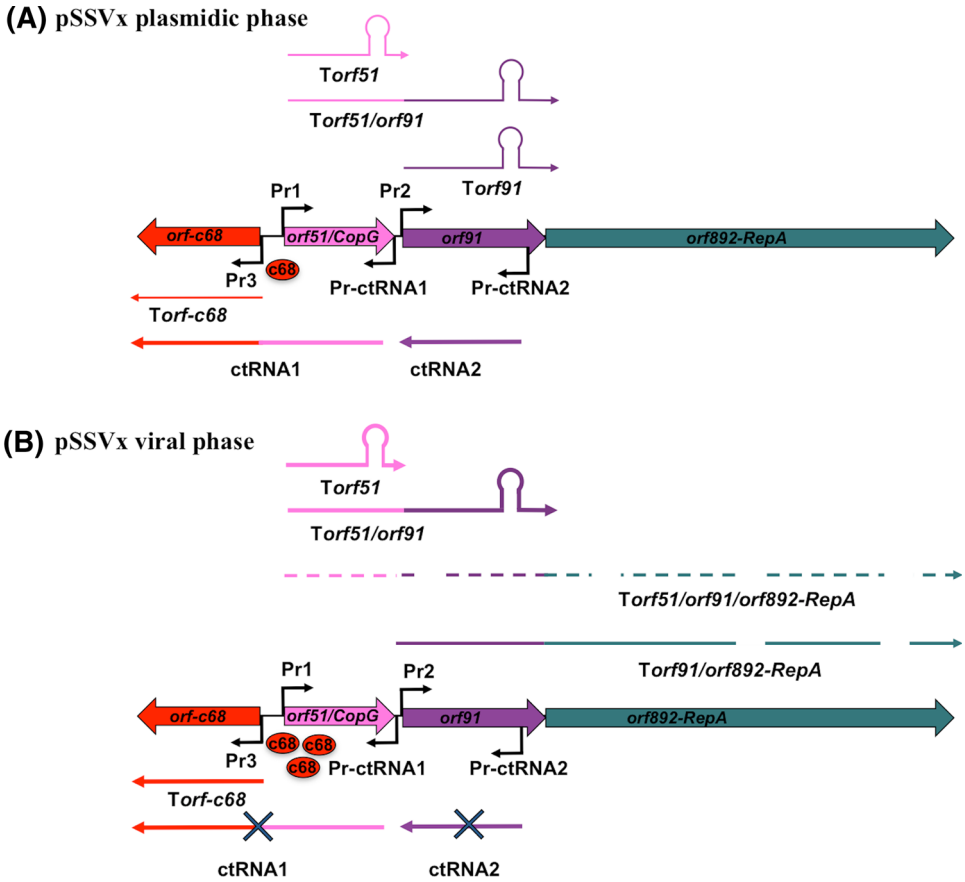


Fig. 6 Schematic representation of the transcription regulation of the pSSVx at the *rep* locus. The transcription of the ORF892-RepA can start at both the promoters Pr1 and Pr2 but with differential initiation frequency as indicated by *solid* (highly transcribed messenger RNAs) and *dashed* (low abundance transcripts) *arrows* in the viral phase of pSSVx replication. The increase in the relative abundance of the other transcripts at the *rep* locus in the passage from the plasmidic to the viral phase is represented by the *thickness* of the *lines*. Differential transcription termination is indicated by the relative length of the transcripts and stem-and-loop structures. ORF892-RepA transcripts are *sketched* as interrupted lines to outline nuclease susceptibility and consequent fast turnover (**b**, viral phase)

level during the plasmidic phase (Fig. 6a). Conversely, at a late growth phase levels of ctRNA1 and ctRNA2 are undetectable (Contursi et al. 2007, 2010) (Fig. 6b). Therefore, the relative abundance of the mRNAs and relatives ctRNAs might move the balance either towards premature termination or to anti-termination. Similarly ctRNA3 may regulate the expression of *orf76*. Regulation via anti-sense RNAs is a common mechanism among bacterial plasmids and anti-sense RNAs are frequently found in archaeal transcriptomes (Siemering et al. 1994; del Solar and Espinosa 2000, Brantl 2002a, b; Tang et al. 2002, 2005; Dodd et al. 2005; Kwong et al. 2006; Straub et al. 2009; Wurtzel et al. 2009).

Besides the shift of the equilibrium from premature termination (more properly a bacterial mode) to anti-termination/read-through (more typically a viral mode) for the *Torf51*, *Torf51/orf91*, *Torf51/orf91/orf892-RepA* messengers, another level of regulation is effective in controlling the expression of the RepA protein at post-transcriptional level and consists in the degradation of the two *rep* messengers (*Torf91/orf892-RepA* and *Torf51/orf91/orf892-RepA*) during the viral phase of pSSVx (Contursi et al. 2007). The degradation of specific transcripts is a typical viral tactic for controlling gene expression at the post-transcriptional and protein synthesis levels, when a compensation is necessary to correct the "sloppiness" in the transcription regulation. This type of mechanism is indeed employed not only by the well known bacterial T4 and lambda phages (Mosig and Hall 1994) but also by the archaeal and closely related fuselloviruses, SSV1 (Reiter et al. 1987b).

The regulatory copy number control systems are usually negative feedback loops that often involve a constitutively expressed regulatory molecule, the abundance of which is determined by gene dosage (directly proportional to plasmid copy number), which negatively affects *rep* gene expression at transcriptional and/or translational levels. The apparent candidate as a "sensor" of the copy number fluctuation of pSSVx is the putative copy number control protein (copG homolog) encoded by *orf51*. The relative abundance of *Torf51* strictly parallels all the variations of the plasmid copy number during the pSSVx life cycle. CopGs generally bind their own promoters (Pr1 in the case of pSSVx) repressing both their own and the transcriptionally coupled expression of the *rep* gene. Biochemical characterization of ORF-c68 has revealed that it specifically binds to a target site upstream of the TATA-box and the BRE element in its own promoter, suggesting that it

functions as a transcriptional activator (Fig. 6a, b). This agrees well with the simultaneous increase of its own transcript as well as of the protein. The biological function and contribution of ORF-c68 to the pSSVx life cycle are still unclear thus far. The messenger *Torf-c68* accumulates even after the plasmid copy number has reached its plateau value, namely up to the very final stage of the pSSVx life cycle. It has been hypothesized that the acquisition of *orf-c68* (as well as of the *orf154* and *orf288*) by the pSSVx genome has been crucial for acquiring the viral nature and the responsiveness to viral stimuli of the helper (Contursi et al. 2011, 2013).

Biochemical characterization of *Sulfolobus* spindle-shaped virus Ragged Hills proteins

Structural and biochemical analyses of *Fuselloviridae* conserved ORFs are functional to unravel the archaeal viral biology. Two proteins encoded by the fusellovirus SSV8 have been characterized, i.e., E73 and D212.

The protein E73 represents an interesting variation on the theme of the Ribbon–Helix–Helix transcription regulators. Indeed, it possesses an additional α-helix located at its C terminal, thus giving rise to the RH3 motif. Likewise most of the RHH proteins, also E73 is an homodimer strengthened by interactions between the N-terminal β-strands of each monomer. The third α-helix stabilizes the quaternary structure by wrapping the helix α2 of the nearby subunit and therefore contributing to the formation of a tightly intertwined dimer (Schlenker et al. 2012; Contursi et al. 2013). E73 shares significant homology with other proteins encoded by other SSV genomes like SSV2 79a, SSV4 73, SSV5 GP23 and SSV6 GP17 and with E51 of SSV1, thus suggesting that this peculiar fold suits with the mechanisms and molecular components required for transcription regulation in the *Fuselloviridae* family.

Orthologs of D212 are present in many *Fuselloviridae* (Wiedenheft et al. 2004). The overall structure of D212 resembles that of type II restriction endonuclease. Furthermore, D212 bears the PD-(D/E)XK catalytic motif that is conserved in this nuclease superfamily. Since, other classes of nucleases involved in DNA replication, repair and recombination also possess the same association of this fold with the PD-(D/E)XK catalytic residues, similar roles also for the SSV8 protein have been hypothesized. A number of functions related to the life cycle and/or viral evolution mechanisms of the *Fuselloviridae* might require the catalytic activity of such nuclease. For instance, D212 might: (1) recognize DNA branch point of the SSV8 genome with a specific geometric structure and specific sequences or (2) participate to the correction of replication defects or other physical damages to the viral genome or (3) catalyze the excision of a covalent closed circular DNA virus from a concatemer as suggested by Redder et al. 2009 (Menon et al. 2010). Although, none of these in vivo roles have been proven, a hint about its function comes from a study performed on the SSV1 homolog (D212). Indeed, the host growth slowed down in cells infected with the mutated virus SSV-Δd244. This effect is probably due to defects in SSV-Δd244 replication or in the resolution of its replication intermediates that might lead to accumulation of aberrant DNA and in turn to a slowing down of the host growth (Iverson and Stedman 2012).

SSV9 integrates not only at tRNA locations

SSV9 was isolated from a hot (75 °C) acidic (pH 4.0) pool in the Geyser Valley region of a National park in the Kamchatka peninsula (Russia) (Wiedenheft et al. 2004).

Analysis of potential *attP* and *attA* sites for directing integration of the viral genome, pointed out to 4 putative *attA* sites on the *S. solfataricus* genome, one at an aspartic acid tRNA, the other two at the glutamic acid tRNA genes (E1 and E2) and the fourth at a non-tRNA site. Differently from SSV8, SSV1 and SSV2 that integrate at single site, SSV9 is able to target all the four predicted *attA* sites. Nevertheless, an exact copy of the tRNA gene is reconstituted only upon integration at the glutamic acid gene, whereas integration at the remaining locations is not as precise (Wiedenheft et al. 2004). This is the first report of an integrating virus that target a non-tRNA site (Campbell 1992; Reiter et al. 1989) and the biological significance of multiple integration events needs to be further investigated.

Shuttle vectors based on *Fuselloviridae* genomes

Genetic analysis is fundamental in dissecting the molecular biology and the physiology of any organisms. Effective and more sophisticated "hot" expression systems are indispensable not only for better understanding gene function and protein sorting in vivo but also for the production of fully active and best performing enzymes.

Despite the extensive characterization of extrachromosomal elements in *Sulfolobus (*Zillig et al. 1996; Prato et al. 2006; Snyder et al. 2003; Lipps 2006; Fu and Johnson 2012; Prangishvili 2013), only few of them have been used as template for the setting up of suitable genetic tools (Atomi et al. 2012; Zheng et al. 2012; Zhang et al. 2013). The putative replication origin of the fusellovirus SSV1 has been employed for the construction of the first shuttle vector for *S. solfataricus*, the pEXSs, in which the hygromycin phosphotransferase gene was the selective marker (Cannio et al. 1998, 2001; Contursi et al. 2004). Although it was able to replicate efficiently both in *S. solfataricus* and *E. coli*, the low copy number in the hyperthermophilic host hindered its applicability for heterologous expression

or genetic manipulations (Fiorentino et al. 2009; Contursi et al. 2003).

Later on, shuttle vectors based on the whole SSV1 genome were developed in a first instance for studying viral ORFs essentiality and subsequently used successfully for heterologous gene expression and complementation (Stedman et al. 1999; Jonuscheit et al. 2003; Albers et al. 2006; Iverson and Stedman 2012). The vector pKMSD48 contains the pBluescript II SK + *E. coli* plasmid site specifically inserted into e178, an ORF encoded by the polycistronic messenger T_5. Despite its bigger size (about 18 kbp), if compared to the wild-type SSV1 (15.5 Kbp), pKMSD48 is efficiently packaged into viable virions and able to spread through *Sulfolobus* cultures without undergoing to recombination or rearrangement events (Stedman et al. 1999). A similar shuttle vector (pJM03) has been assembled by fusing the SSV1 moiety with exogenous sequences within the non-essential ORF e51, i.e., the *E. coli* plasmid pUC18, the *lacS* expression cassette and the *pyrE/F* genes as reporter and selection marker, respectively (Jonuscheit et al. 2003). Its efficient spreading as well as the successful genetic complementation showed by this vector, prompted its use for heterologous expression at high level of hyperthermophilic proteins (Albers et al. 2006).

Another suitable shuttle vector for *S. solfataricus* was obtained by fusing site specifically the pSSVx chromosome with an *E. coli* plasmid replicon. The resulting recombinant vector was able to propagate in *E. coli* under ampicillin selection and at high copy number in *S. solfataricus* with no recombination/integration events occurring. The stable maintenance of this vector was achieved by inserting the *lacS* cassette as selection marker. The resulting recombinant vector relied on the co-presence of the SSV2 for its replication, packaging and spreading (Aucelli et al. 2006). Moreover, it was successfully used for homologous gene transfer and overexpression of the *sso1354* gene encoding one of the three putative endo-β-1-4-glucanases from *S. solfataricus* (Limauro et al. 2001; Girfoglio et al. 2012).

Concluding remarks

Fuselloviridae provide unique models for studying molecular virology and molecular biology in *Archaea*. To date, SSV1 and SSV2 are still the only archaeal viruses exhibiting inducible virus production. These unique features have facilitated their research ever since their discovery. Early studies on constitutive and UV-inducible transcription represent a milestone in archaeal transcription study including definition of archaeal basal promoter elements as well as revealing signals of archaeal transcription termination (Reiter et al. 1988a, b). Then, an extensive study of the pSSVx transcription has revealed complex pre- and post-transcriptional mechanisms to regulate the expression of its replicase during the physiological induction of pSSVx replication in the presence of SSV2 (Contursi et al. 2007, 2010). Transcriptomic analysis of viral and host genomes using whole-genome microarrays has unraveled cascade regulation of gene expression. However, very limited insights have been gained into the regulation of these processes and we still know very little about the life cycle of fuselloviruses and their regulation.

One of the first outstanding questions for the fuselloviral research is what are the main regulators leading to the inducible virus production? Identification of the SSV1 lysogenic regulator represents an important step towards the mechanism of the induction process (Fusco et al. 2013) but detailed mechanisms remain to be investigated. Crystal structures have been solved for several putative SSV1 transcriptional factors. A combinatorial approach of biochemical and genetic assays will be more powerful to study functions of these putative transcriptional factors such that their target sites will be revealed from biochemical analysis whereas insights into their functions will be gained from sophisticated genetic analyses involving gene knockout and mutagenesis.

Both viruses exhibit a cascade regulation of gene expression during virus induction, a process that very likely involves host transcriptional factors. Fortunately, tools for genetic study are not only available for SSV1 and SSV2/pSSVx (Aucelli et al. 2006; Iverson and Stedman 2012) but also for their hosts (She et al. 2009; Wagner et al. 2009). Therefore, SSV1 and SSV2/pSSVx provide excellent model systems for these investigations.

There are several important themes in fuselloviral research to be investigated including mechanisms of viral DNA replication, cell receptors for virus infection, viral and host components that participate in virion assembly and cellular extrusion. Mechanism of virus satellites formation is a very exciting theme as it reveals how archaeal plasmids could have exploited the fuselloviruses to enhance their possibilities for survival. Finally, the arm race between archaeal hosts and their viruses involves a novel genetic entity termed CRISPR systems, i.e., "Clustered regularly interspaced palindromic repeats" (CRISPR) and CRISPR-associated proteins (Garrett et al. 2011; Held and Whitaker 2009). Apparently fuselloviruses have developed the capability of circumventing the CRISPR systems of their hosts because they infect these archaeal organisms despite of the fact that these hosts carry many significant matches to fuselloviral genomes identified in their CRISPR loci (Redder et al. 2009). Clearly, focused researches are required to provide important insights into these research themes.

Acknowledgments Researches in the authors' laboratories are supported by the grant "Programma F.A.R.O, IV° tornata", founded by Università Federico II di Napoli and "Compagnia di San Paolo" (Naples Laboratory) and grants (DFF–1323-00330 and FTP/11-106683) from Danish Council of Independent Research (Copenhagen laboratory). We thank Peter Redder and Li Huang for providing pictures of fuselloviruses and Prof. Simonetta Bartolucci for valuable scientific discussions throughout the preparation of the manuscript.

References

Ackermann HW, Prangishvili D (2012) Prokaryote viruses studied by electron microscopy. Arch Virol 157:1843–1849

Albers SV, Jonuscheit M et al (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Appl Environ Microbiol 72:102–111

Albers SV, Forterre P, Prangishvili D, Schleper C (2013) The legacy of Carl Woese and Wolfram Zillig: from phylogeny to landmark discoveries. Nat Rev Microbiol 11:713–719

Alekshun MN, Levy SB, Mealy TR, Seaton BA, Head JF (2001) The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 A resolution. Nat Struct Biol 8:710–714

Argos P, Landy A et al (1986) The integrase family of site-specific recombinases: regional similarities and global diversity. EMBO J 5:433–440

Arnold HP, She Q et al (1999) The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. Mol Microbiol 34:217–226

Atanasova NS, Roine E, Oren A, Bamford DH, Oksanen HM (2012) Global network of specific virus-host interactions in hypersaline environments. Environ Microbiol 14:426–440

Atomi H, Imanaka T, Fukui T (2012) Overview of the genetic tools in the *Archaea*. Front Microbiol 3:337

Aucelli T, Contursi P, Girfoglio M, Rossi M, Cannio R (2006) A spreadable, non-integrative and high copy number shuttle vector for *Sulfolobus solfataricus* based on the genetic element pSSVx from *Sulfolobus islandicus*. Nucleic Acids Res 34:e114

Bamford DH (2003) Do viruses form lineages across different domains of life? Res Microbiol 154:231–236

Bamford DH, Grimes JM, Stuart DI (2005) What does structure tell us about virus evolution? Curr Opin Struct Biol 15:655–663

Bartolucci S, Contursi P, Fiorentino G, Limauro D, Pedone E (2013) Responding to toxic compounds: a genomic and functional overview of Archaea. Front Biosci 18:165–189

Bize A, Peng X et al (2008) Viruses in acidic geothermal environments of the Kamchatka Peninsula. Res Microbiol 159:358–366

Bolduc B, Shaughnessy DP, Wolf YI, Koonin EV, Roberto FF, Young M (2012) Identification of novel positive-strand RNA viruses by metagenomic analysis of archaea-dominated Yellowstone hot springs. J Virol 86:5562–5573

Brantl S (2002a) Antisense RNAs in plasmids: control of replication and maintenance. Plasmid 48:165–173

Brantl S (2002b) Antisense-RNA regulation and RNA interference. Biochim Biophys Acta 1575:15–25

Campbell AM (1992) Chromosomal insertion sites for phages and plasmids. J Bacteriol 174:7495

Cannio R, Contursi P, Rossi M, Bartolucci S (1998) An autonomously replicating transforming vector for *Sulfolobus solfataricus*. J Bacteriol 180:3237–3240

Cannio R, Contursi P, Rossi M, Bartolucci S (2001) Thermoadaptation of a mesophilichygromycin B phosphotransferase by directed evolution in hyperthermophilic Archaea: selection of a stable genetic marker for DNA transfer into *Sulfolobus solfataricus*. Extremophiles 5:153–159

Ceballos RM, Marceau CD, Marceau JO, Morris S, Clore AJ, Stedman KM (2012) Differential virus host-ranges of the *Fuselloviridae* of hyperthermophilic Archaea: implications for evolution in extreme environments. Front Microbiol 3:295

Clore AJ, Stedman KM (2007) The SSV1 viral integrase is not essential. Virology 361:103–111

Contursi P, Cannio R, Prato S, Fiorentino G, Rossi M, Bartolucci S (2003) Development of a genetic system for hyperthermophilic Archaea: expression of a moderate thermophilic bacterial alcohol dehydrogenase gene in *Sulfolobus solfataricus*. FEMS Microbiol Lett 218:115–120

Contursi P, Pisani FM, Grigoriev A, Cannio R, Bartolucci S, Rossi M (2004) Identification and autonomous replication capability of a chromosomal replication origin from the archaeon *Sulfolobus solfataricus*. Extremophiles 8:385–391

Contursi P, Jensen S, Aucelli T, Rossi M, Bartolucci S, She Q (2006) Characterization of the *Sulfolobus* host-SSV2 virus interaction. Extremophiles 10:615–627

Contursi P, Cannio R, Prato S, She Q, Rossi M, Bartolucci S (2007) Transcriptional analysis of the genetic element pSSVx: differential and temporal regulation of gene expression reveals correlation between transcription and replication. J Bacteriol 189:6339–6350

Contursi P, Cannio R, She Q (2010) Transcription termination in the plasmid/virus hybrid pSSVx from *Sulfolobus islandicus*. Extremophiles 14:453–463

Contursi P, D'Ambrosio K et al (2011) C68 from the *Sulfolobus islandicus* plasmid-virus pSSVx is a novel member of the AbrB-like transcription factor family. Biochem J 435:157–166

Contursi P, Fusco S, Limauro D, Fiorentino G (2013) Host and viral transcriptional regulators in *Sulfolobus*: an overview. Extremophiles 17:881–895

Cortez D, Forterre P, Gribaldo S (2009) A hidden reservoir of integrative elements is the major source of recently acquired foreign genes and ORFans in archaeal and bacterial genomes. Genome Biol 10:R65

del Solar G, Espinosa M (2000) Plasmid copy number control: an ever-growing story. Mol Microbiol 37:492–500

Di Fiore A, Fiorentino G, Vitale RM, Ronca R, Amodeo P, Pedone C, Bartolucci S, De Simone G (2009) Structural analysis of BldR from *Sulfolobus solfataricus* provides insights into the molecular basis of transcriptional activation in archaea by MarR family proteins. J Mol Biol 388:559–569

Dodd IB, Shearwin KE, Egan JB (2005) Revisited gene regulation in bacteriophage lambda. Curr Opin Genet Dev 15:145–152

Eilers BJ, Young MJ, Lawrence CM (2012) The structure of an archaeal viral integrase reveals an evolutionarily conserved catalytic core yet supports a mechanism of DNA cleavage in trans. J Virol 86:8309–8313

Fiorentino G, Ronca R, Bartolucci S (2009) A novel *E. coli* biosensor for detecting aromatic aldehydes based on a responsive inducible archaeal promoter fused to the green fluorescent protein. Appl Microbiol Biotechnol 82:67–77

Fiorentino G, Del Giudice I, Bartolucci S, Durante L, Martino L, Del Vecchio P (2011) Identification and physicochemical characterization of BldR2 from *Sulfolobus solfataricus*, a novel archaeal member of the MarR transcription factor family. Biochemistry 50:6607–6621

Forterre P (2006) The origin of viruses and their possible roles in major evolutionary transitions. Virus Res 117:5–16

Forterre P, Prangishvili D (2013) The major role of viruses in cellular evolution: facts and hypotheses. Curr Opin Virol 3:558–565

Fröls S, Gordon PM, Panlilioc MA, Schleper C, Sensen CW (2007) Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays. Virology 365:48–59

Fu CY, Johnson JE (2012) Structure and cell biology of archaeal virus STIV. Curr Opin Virol 2:122–127

Fusco S, She Q, Bartolucci S, Contursi P (2013) T(lys), a newly identified *Sulfolobus* spindle-shaped virus 1 transcript expressed in the lysogenic state, encodes a DNA-binding protein interacting at the promoters of the early genes. J Virol 87:5926–5936

Garrett RA, Shah SA, Vestergaard G, Deng L, Gudbergsdottir S, Kenchappa CS, Erdmann S, She Q (2011) CRISPR-based immune systems of the Sulfolobales: complexity and diversity. Biochem Soc Trans 39:51–57

Gibrat JF, Madej T, Bryant SH (1996) Surprising similarities in structure comparison. Curr Opin Struct Biol 6:377–385

Girfoglio M, Rossi M, Cannio R (2012) Cellulose degradation by *Sulfolobus solfataricus* requires a cell-anchored endo-beta-1-4-glucanase. J Bacteriol 194:5091–5100

Grummt I, Maier U, Ohrlein A, Hassouna N, Bachellerie JP (1985) Transcription of mouse rDNA terminates downstream of the 3′ end of 28 s rRNA and involves interaction of factors with repeated sequences in the 3′ spacer. Cell 43:801–810

Guixa-Boixareu N, Calderon-Paz JI, Heldal M, Bratbak G, Pedros-Alio C (1996) Viral lysis and bacterivory as prokaryotic loss factors along a salinity gradient. Aquat Microb Ecol 11:215–227

Guo X, Huang L (2010) A superfamily 3 DNA helicase encoded by plasmid pSSVi from the hyperthermophilic archaeon *Sulfolobus solfataricus* unwinds DNA as a higher-order oligomer and interacts with host primase. J Bacteriol 192:1853–1864

Guo L, Brugger K, Liu C et al (2011) Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus-host interaction studies. J Bacteriol 193:1672–1680

Häring M, Rachel R, Peng X, Garrett RA, 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 79:9904–9911

Held NL, Whitaker RJ (2009) Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. Environ Microbiol 11:457–466

Helmer-Citterich M, Anceschi MM, Banner DW, Cesareni G (1988) Control of ColE1 replication: low affinity specific binding of Rop (Rom) to RNAI and RNAII. EMBO J 7:557–566

Huet J, Schnabel R, Sentenac A, Zillig W (1983) Archaebacteria and eukaryotes possess DNA dependent. RNA polymerases of a common type. EMBO J 2:1291–1294

Iverson E, Stedman K (2012) A genetic study of SSV1, the prototypical fusellovirus. Front Microbiol 3:1–7

Jaubert C, Danioux C, Oberto J, Cortez D, Bize A, Krupovic M, She Q, Forterre P, Prangishvili D, G Sezonov (2013) Genomics and genetics of *Sulfolobus islandicus* LAL14/1, a model hyperthermophilic archaeon. Open Biol 3(4):130010. doi:10.1098/rsob.130010

Jonuscheit M, Martusewitsch E, Stedman KM, Schleper C (2003) A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. Mol Microbiol 48:1241–1252

Keeling PJ, Klenk HP et al (1996) Complete nucleotide sequence of the *Sulfolobus islandicus* multicopy plasmid pRN1. Plasmid 35:141–144

Keeling PJ, Klenk HP et al (1998) *Sulfolobus islandicus* plasmids pRN1 and pRN2 share distant but common evolutionary ancestry. Extremophiles 2:391–393

Kletzin A, Lieke A, Urich T, Charlebois RL, Sensen CW (1999) Molecular analysis of pDL10 from *Acidianus ambivalens* reveals a family of related plasmids from extremely thermophilic and acidophilic archaea. Genetics 152:1307–1314

Koonin EV (1992) Archaebacterial virus SSV1 encodes a putative DnaA-like protein. Nucleic Acids Res 11:1143

Koonin EV, Senkevich TG, Dolja VV (2006) The ancient Virus World and evolution of cells. Biol Direct 19:1–29

Kraft P, Kümmel D, Oeckinghaus A, Gauss GH, Wiedenheft B, Young M, Lawrence CM (2004a) Structure of D-63 from *Sufolobus* spindle-shaped virus 1: surface properties of the dimeric four-helix bundle suggest an adaptor protein function. J Virol 78:7438–7442

Kraft P, Oeckinghaus A et al (2004b) Crystal structure of F-93 from *Sulfolobus* spindle-shaped virus 1, a winged-helix DNA binding protein. J Virol 78:11544–11550

Krupovic M, Prangishvili D, Hendrix RW, Bamford DH (2011) Genomics of bacterial and archaeal viruses: dynamics within the prokaryotic virosphere. Microbiol Mol Biol Rev 75:610–613

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

Kwong SM, Skurray RA, Firth N (2006) Replication control of staphylococcal multiresistance plasmid pSK41: an antisense RNA mediates dual-level regulation of Rep expression. J Bacteriol 188:4404–4412

Lawrence CM, Menon S, Eilers BJ, Bothner B, Khayat R, Douglas T, Young MJ (2009) Structural and functional studies of archaeal viruses. J Biol Chem 284:12599–12603

Le Romancer M, Gaillard M, Geslin C, Prieur D (2007) Viruses in extreme environments. Extremophiles 6:17–31

Limauro D, Cannio R, Fiorentino G, Rossi M, Bartolucci S (2001) Identification and molecular characterization of an endoglucanase gene, celS, from the extremely thermophilic archaeon *Sulfolobus solfataricus*. Extremophiles 5:213–219

Lipps G (2006) Plasmids and viruses of the thermoacidophilic crenarchaeote *Sulfolobus*. Extremophiles 10:17–28

Lipps G, Rother S, Hart C, Krauss G (2003) A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. EMBO J 22:2516–2525

Martin A, Yeats S, Janekovic D, Reiter WD, Aicher W, Zillig W (1984) SAV 1, a temperate UV.-inducible DNA virus-like particle from the archaebacterium *Sulfolobus acidocaldarius* isolate B12. EMBO J 3:2165–2168

Menon SK, Maaty WS et al (2008) Cysteine usage in *Sulfolobus* spindle-shaped virus 1 and extension to hyperthermophilic viruses in general. Virology 376:270–278

Menon SK, Eilers BJ, Young MJ, Lawrence CM (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 84:5890–5897

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 402:347–354

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 USA 109:13386–13391

Mosig G, Hall DH (1994) Regulation of gene expression. In: Karam JD (ed) Molecular biology of bacteriophage T4. ASM Press, Washington, DC, pp 127–193

Muskhelishvili G, Palm P, Zillig W (1993) SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. Mol Gen Genet 237:334–342

Nadal M, Mirambeau G, Forterre P, Reiter W-D, Duguet M (1986) Positively supercoiled DNA in a virus-like particle of an archaebacterium. Nature 321:256–258

Okutan E, Deng L, Mirlashari S, Uldahl K, Halim M, Liu C, Garrett RA, She Q, Peng X (2013) Novel insights into gene regulation of

the rudivirus SIRV2 infecting *Sulfolobus* cells. RNA Biol 10:875–885

Oren A, Bratbak G, Hendal M (1997) Occurrence of virus-like particles in the Dead Sea. Extremophiles 1:143–149

Ortmann AC, Wiedenheft B, Douglas T, Young M (2006) Hot crenarchaeal viruses reveal deep evolutionary connections. Nat Rev Microbiol 4:520–528

Ortmann AC, Brumfield SK et al (2008) Transcriptome analysis of infection of the archaeon *Sulfolobus solfataricus* with *Sulfolobus* turreted icosahedral virus. J Virol 82:4874–4883

Palm P, Schleper C, Grampp B, Yeats S, McWilliam P, Reiter WD, Zillig W (1991) Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae*. Virology 185:242–250

Peng X (2008) Evidence for the horizontal transfer of an integrase gene from a fusellovirus to a pRN-like plasmid within a single strain of *Sulfolobus* and the implications for plasmid survival. Microbiology 154:383–391

Peng X, Holz I, Zillig W, Garrett RA, She Q (2000) Evolution of the family of pRN plasmids and their integrase-mediated insertion into the chromosome of the crenarchaeon *Sulfolobus solfataricus*. J Mol Biol 303:449–454

Peng X, Garrett RA, She Q (2012) Archaeal viruses–novel, diverse and enigmatic. Sci China Life Sci 55:422–433

Pietilä MK, Roine E, Paulin L, Kalkkinen N, Bamford DH (2009) An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. Mol Microbiol 72:307–319

Pietilä MK, Laurinmäki P, Russell DA, Ko CC, Jacobs-Sera D, Butcher SJ, Bamford DH, Hendrix RW (2013) Insights into head-tailed viruses infecting extremely halophilic archaea. J Virol 87:3248–3260

Pina M, Bize A, Forterre P, Prangishvili D (2011) The archeoviruses. FEMS Microbiol Rev 35:1035–1054

Porter K, Russ BE, Dyall-Smith ML (2007) Virus–host interactions in salt lakes. Curr Opin Microbiol 10:418–424

Porter K, Tang SL, Chen CP, Chiang PW, Hong MJ, Dyall-Smith M (2013) PH1: an archaeovirus of Haloarcula hispanica related to SH1 and HHIV-2. Archaea 2013:456318

Prangishvili D (2003) Evolutionary insights from studies on viruses of hyperthermophilic archaea. Res Microbiol 154:289–294

Prangishvili D (2013) The wonderful world of archaeal viruses. Annu Rev Microbiol 8:565–585

Prangishvili D, Garrett RA (2004) Exceptionally diverse morphotypes and genomes of crenarchaeal hyperthermophilic viruses. Biochem Soc Trans 32:204–208

Prangishvili D, Garrett RA (2005) Viruses of hyperthermophilic Crenarchaea. Trends Microbiol 13:535–542

Prangishvili D, Stedman K, Zillig W (2001) Viruses of the extremely thermophilic archaeon *Sulfolobus*. Trends Microbiol 9:39–43

Prangishvili D, Forterre P, Garrett RA (2006a) Viruses of the Archaea: a unifying view. Nat Rev Microbiol 4:837–848

Prangishvili D, Garrett RA, Koonin EV (2006b) Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life. Virus Res 117:52–67

Prato S, Cannio R, Klenk HP, Contursi P, Rossi M, Bartolucci S (2006) pIT3, a cryptic plasmid isolated from the hyperthermophilic crenarchaeon *Sulfolobus solfataricus* IT3. Plasmid 56:35–45

Quax TE, Voet M, Sismeiro O, Dillies MA, Jagla B, Coppée JY, Sezonov G, Forterre P, van der Oost J, Lavigne R, Prangishvili D (2013) Massive activation of archaeal defense genes during viral infection. J Virol 87:8419–8428

Quemin ER, Lucas S, Daum B, Quax TE, Kühlbrandt W, Forterre P, Albers SV, Prangishvili D, Krupovic M (2013) First insights into the entry process of hyperthermophilic archaeal viruses. J Virol 87:13379–13385

Rachel R, Bettstetter M, Hedlund BP, Häring M, Kessler A, Stetter KO, Prangishvili D (2002) Remarkable morphological diversity of viruses and virus-like particles in hot terrestrial environments. Arch Virol 147:2419–2429

Redder P, Peng X, Brügger K, Shah SA, Roesch F, Greve B, She Q, Schleper C, Forterre P, Garrett RA, Prangishvili D (2009) Four newly isolated fuselloviruses from extreme geothermal environments reveal unusual morphologies and a possible interviral recombination mechanism. Environ Microbiol 11:2849–2862 John Wiley & Sons Ltd

Reiter WD, Palm P et al (1987a) Identification and characterization of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. Mol Gen Genet 206:144–153

Reiter WD, Palm P, Yeats S, Zillig W (1987b) Gene expression in archaebacteria: physical mapping of constitutive and UV-inducible transcripts from the *Sulfolobus* virus-like particle SSV1. Mol Gen Genet 209:270–275

Reiter WD, Palm P, Zillig W (1988a) Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. Nucleic Acids Res 16:1–19

Reiter WD, Palm P, Zillig W (1988b) Transcription termination in the archaebacterium *Sulfolobus*: signal structures and linkage to transcription initiation. Nucleic Acids Res 16:2445–2459

Reiter WD, Palm P, Yeats S (1989) Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. Nucleic Acids Res 17:1907–1914

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

Rice G, Stedman K et al (2001) Viruses from extreme thermal environments. Proc Natl Acad Sci USA 98:13341–13345

Schlenker C, Goel A, Tripet BP, Menon S, Willi T, Dlakić M, Young MJ, Lawrence CM, Copié V (2012) Structural studies of E73 from a hyperthermophilic archaeal virus identify the "RH3" domain, an elaborated ribbon-helix-helix motif involved in DNA recognition. Biochemistry 51:2899–2910

Schleper C, Kubo K, Zillig W (1992) The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. Proc Natl Acad Sci USA 89:7645–7649

Serre MC, Letzelter C, Garel JR, Duguet M (2002) Cleavage properties of an archaeal site-specific recombinase, the SSV1 integrase. J Biol Chem 277:16758–16767

She Q, Peng X, Zillig W, Garrett RA (2001a) Gene capture in archaeal chromosomes. Nature 409:478

She Q, Singh RK et al (2001b) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci USA 98:7835–7840

She Q, Brügger K, Chen L (2002) Archaeal integrative genetic elements and their impact on genome evolution. Res Microbiol 153:325–332

She Q, Shen B, Chen L (2004) Archaeal integrases and mechanisms of gene capture. Biochem Soc Trans 32:222–226

She Q, Zhang C, Deng L, Peng N, Chen Z, Liang YX (2009) Genetic analyses in the hyperthermophilic archaeon *Sulfolobus islandicus*. Biochem Soc Trans 37:92–96

Siemering KR, Praszkier J, Pittard AJ (1994) Mechanism of binding of the antisense and target RNAs involved in the regulation of IncB plasmid replication. J Bacteriol 176:2677–2688

Sime-Ngando T, Lucas S, Robin A, Tucker KP, Colombet J, Bettarel Y, Desmond E, Gribaldo S, Forterre P, Breitbart M, Prangishvili D (2011) Diversity of virus-host systems in hypersaline Lake Retba, Senegal. Environ Microbiol 13: 1956–1972

Snyder JC, Young MJ (2011) Advances in understanding archaea-virus interactions in controlled and natural environments. Curr Opin Microbiol 14:497–503

Snyder JC, Stedman K, Rice G, Wiedenheft B, Spuhler J, Young MJ (2003) Viruses of hyperthermophilic Archaea. Res Microbiol 154:474–482

Stedman KM, Schleper C, Rumpf E, Zillig W (1999) Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*: construction and testing of viral shuttle vectors. Genetics 152:1397–1405

Stedman KM, She Q et al (2003) Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: sSV1 and SSV2. Res Microbiol 154:295–302

Stedman KM, Kosmicki NR, Diemer GS (2013) Codon usage frequency of RNA virus genomes from high-temperature acidic-environment metagenomes. J Virol 87:1919

Straub J, Brenneis M, Jellen-Ritter A, Heyer R, Soppa J, Marchfelder A (2009) Small RNAs in haloarchaea: identification, differential expression and biological function. RNA Biol 6:281–292

Tang TH, Bachellerie JP, Rozhdestvensky T, Bortolin ML, Huber H, Drungowski M, Elge T, Brosius J, Huttenhofer A (2002) Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. Proc Natl Acad Sci USA 99:7536–7541

Tang TH, Polacek N, Zywicki M et al (2005) Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. Mol Microbiol 55:469–481

Wagner M, Berkner S, Ajon M, Driessen AJM, Albers SV (2009) Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. Biochem Soc Trans 37: 97–101

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: 124–133 Elsevier Publishers

Wiedenheft B, Stedman K et al (2004) Comparative genomic analysis of hyperthermophilic archaeal Fuselloviridae viruses. J Virol 78:1954–1961

Woese CR, Magrum LJ, Fox GE (1978) Archaebacteria. J Mol Evol 11:245–251

Wu RY, Zhang RG et al (2003) Crystal structure of *Enterococcus faecalis* SlyA-like transcriptional factor. J Biol Chem 278: 20240–20244

Wurtzel O, Sapra R, Chen F, Zhu Y, Simmons BA, Sorek R (2009) A single-base resolution map of an archaeal transcriptome. Genome Res 20:133–141

Yeats S, McWilliam P, Zillig W (1982) A plasmid in the archaebacterium *Sulfolobus acidocaldarius*. EMBO J 1:1035–1038

Zhan Z, Ouyang S, Liang W, Zhang Z, Liu ZJ, Huang L (2012) Structural and functional characterization of the C-terminal catalytic domain of SSV1 integrase. Acta Crystallogr D Biol Crystallogr 68:659–670

Zhang C, Cooper TE, Krause DJ, Whitaker RJ (2013) Augmenting the genetic toolbox for *Sulfolobus islandicus* with a stringent positive selectable marker for agmatine prototrophy. Appl Environ Microbiol 79:5539–5549

Zheng T, Huang Q, Zhang C, Ni J, She Q, Shen Y (2012) Development of a simvastatin selection marker for a hyperthermophilicacidophile, *Sulfolobus islandicus*. Appl Environ Microbiol 78:568–574

Zillig W, Palm P, Reiter WD, Gropp F, Puhler G, Klenk HP (1988) Comparative evaluation of gene expression in archaebacteria. Eur J Biochem 173:473–482

Zillig W, Kletzin A et al (1994) Screening for *Sulfolobales*, Their Plasmids and Their Viruses in Icelandic Solfataras. Syst Appl Microbiol 16:609–628

Zillig W, Prangishvilli D et al (1996) Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic *Archaea*. FEMS Microbiol Rev 18:225–236