

Review

# Genomics, Transcriptomics, and Proteomics of SSV1 and Related Fusellovirus: A Minireview

Martina Aulitto <sup>1,2</sup>, Laura Martinez-Alvarez <sup>3,†</sup>, Salvatore Fusco <sup>4,†</sup>, Qunxin She <sup>5</sup>, Simonetta Bartolucci <sup>1</sup>, Xu Peng <sup>3,\*</sup> and Patrizia Contursi <sup>1,6,7,\*</sup>

<sup>1</sup> Dipartimento di Biologia, University of Naples Federico II, 80126 Naples, Italy

<sup>2</sup> Lawrence Berkeley National Laboratory, Biological Systems and Engineering Division, Berkeley, CA 94720, USA

<sup>3</sup> Archaea Centre, Department of Biology, University of Copenhagen, DK-1165 Copenhagen, Denmark

<sup>4</sup> Biochemistry and Industrial Biotechnology Laboratory, Department of Biotechnology, University of Verona, 37134 Verona, Italy

<sup>5</sup> CRISPR and Archaea Biology Research Center, State Key Laboratory of Microbial Technology, Microbial Technology Institute, Shandong University, Qingdao 250100, China

<sup>6</sup> BAT Center—Interuniversity Center for Studies on Bioinspired Agro-Environmental Technology, University of Naples Federico II, 80055 Naples, Italy

<sup>7</sup> Task Force on Microbiome Studies, University of Naples Federico II, 80126 Naples, Italy

\* Correspondence: peng@bio.ku.dk (X.P.); contursi@unina.it (P.C.)

† These authors contributed equally to this work.

**Abstract:** *Saccharolobus* spindle-shaped virus 1 (SSV1) was one of the first viruses identified in the archaeal kingdom. Originally isolated from a Japanese species of *Saccharolobus* back in 1984, it has been extensively used as a model system for genomic, transcriptomic, and proteomic studies, as well as to unveil the molecular mechanisms governing the host–virus interaction. The purpose of this mini review is to supply a compendium of four decades of research on the SSV1 virus.

**Keywords:** *Saccharolobus*; SSV1; genomic; transcriptomic; proteomic



**Citation:** Aulitto, M.; Martinez-Alvarez, L.; Fusco, S.; She, Q.; Bartolucci, S.; Peng, X.; Contursi, P. Genomics, Transcriptomics, and Proteomics of SSV1 and Related Fusellovirus: A Minireview. *Viruses* **2022**, *14*, 2082. <https://doi.org/10.3390/v14102082>

Academic Editor: Terje Dokland

Received: 20 July 2022

Accepted: 15 September 2022

Published: 20 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



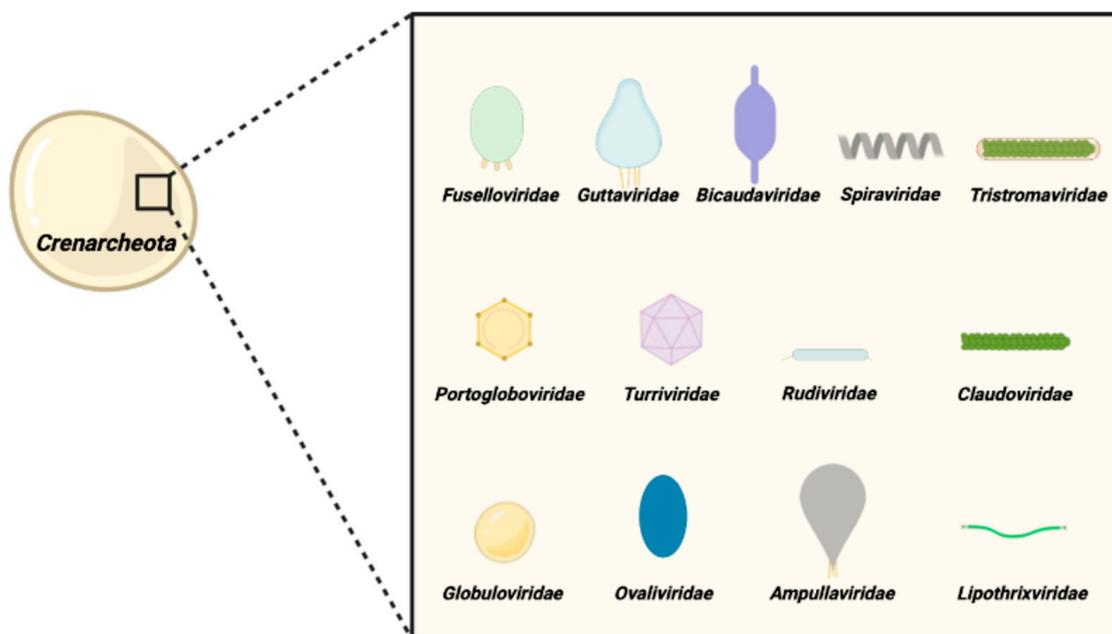
**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With some exceptions, archaeal viruses are genomically and morphologically unique since they are evolutionarily unrelated to viruses infecting bacteria and eukarya, thus representing one of the least understood part of the virosphere [1–3]. Indeed, most of them harbor many orphan open reading frames (ORFs) with little or no similarity to known sequences in the public databases [4]. Most of the known archaeal viruses infect members of the phylum *Crenarchaeota*, and those unique to hyperthermophilic *Archaea* belong to 13 families, i.e., *Turrioviridae*, *Portogloboviridae*, *Rudiviridae*, *Clavaviridae*, *Lipothrixviridae*, *Tristromaviridae*, *Globuloviridae*, *Ovaliviridae*, *Guttaviridae*, *Spiraviridae*, *Ampullaviridae*, *Fuselloviridae*, and *Bicaudaviridae* [5]. Some morphologies are thus far exclusive to archaeal viruses, including the bottle (*Ampullaviridae*), the droplet (*Guttaviridae*), and the spindle or lemon shape (*Fuselloviridae*). Among these latter ones, SSVs (*Saccharolobus-ex Sulfolobus*-spindle virus) were the first discovered and one of the best studied archaeal viral families (Figure 1) [6].

SSVs are abundant in environments characterized by extreme chemo-physical conditions such as high temperature (>70 °C) and low pH (<4), where their hosts, i.e., mainly *Sulfolobus* spp. (recently renamed *Saccharolobus* [7]) and its close relatives, thrive [8]. SSVs contain circular double-stranded DNA genomes that vary in size between 14 and 17 kbp and are found as episomal DNA, integrated in the host genome as proviruses, or packaged as fully assembled virions [6,8]. Their size is approximately 90 nm by 60 nm with short tail fibers extending from the end of the major axis. A total of 50 SSV representatives have been deposited into the NCBI (20 June 2022, Supplementary File S1); nevertheless, the

majority of them have been poorly characterized, apart from SSV2, SSV8, SSV9, and SSV22, for which sporadic biochemical and physiological studies are reported [9–14].



**Figure 1.** Representation of viruses infecting Crenarcheota. The schematic morphologies of virions are shown above the corresponding family names.

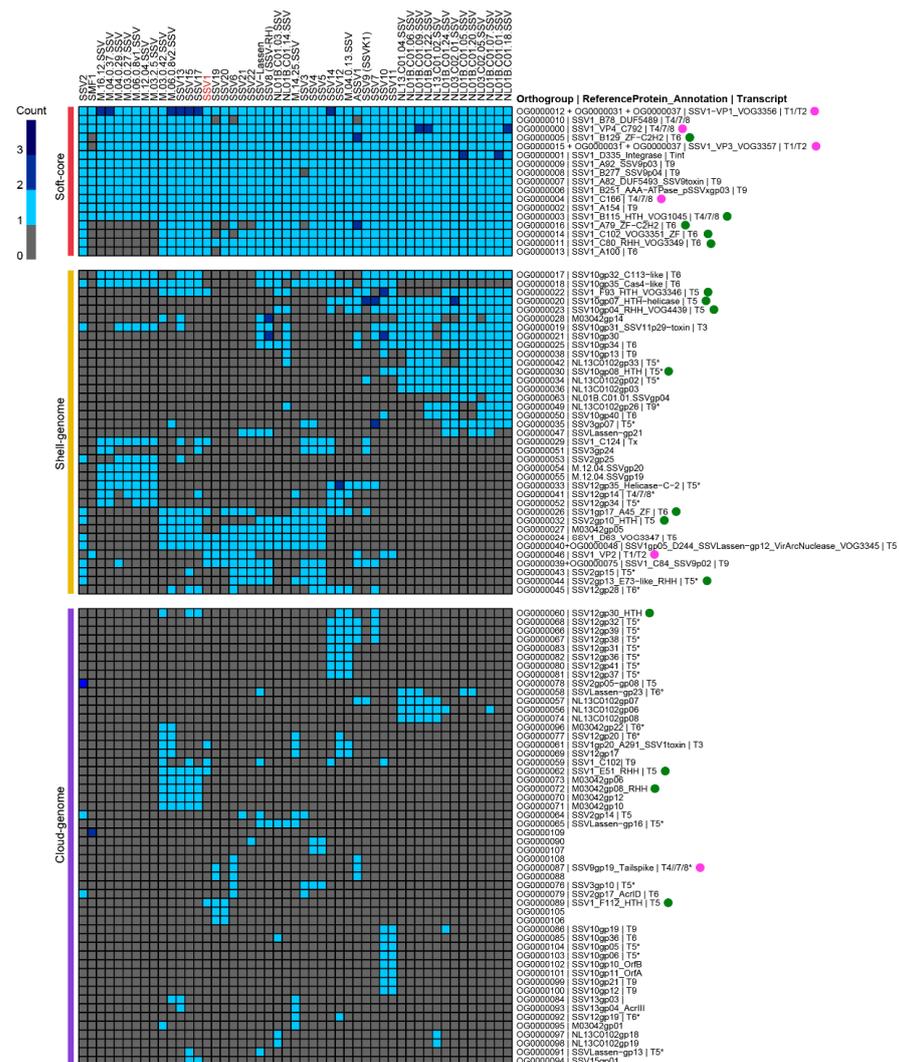
Among SSVs, *Saccharolobus* spindle-shaped virus 1 (SSV1) is the best-understood member of the *Fuselloviridae* and represents one of the model archaeal viruses for physiological, genetic, transcriptional, and virus–host interaction studies [13,15–18]. SSV1 was isolated from its natural host (*Saccharolobus shibatae* strain B12) at a geothermal spring in Beppu (Japan). Initially, SSV1 was named SAV1 because of the erroneous species annotation of its host that was at recognized as *Saccharolobus acidocaldarius* B12 [17,19–21]. SSV1 can replicate in few other hosts, including a strain of *S. solfataricus* isolated from the solfataric field of Pisciarelli near Naples (Italy) [12]. Upon infection, a copy of the SSV1 genome (i.e., a double-stranded DNA molecule of 15,465 bp) integrates into the host chromosome at an arginyl-tRNA gene by means of a virus-encoded integrase (D355) [22]. Besides the integrated form (pro-virus), about five copies per cell of SSV1 DNA are present in host cells as episomal DNA [23]. As reported by initial transcription analyses, upon UV irradiation, a well-coordinated temporal expression of the SSV1 transcripts was revealed, similarly to what has been reported for many bacteriophages and eukaryotic viruses [24,25]. Sequence analyses of SSV1 transcriptional start sites (TSSs) led to the discovery of archaeal promoter elements such as the TATA-box-like hexanucleotide sequence as well as the nearby TGA consensus trinucleotide sequence [24,26]. These two conserved motifs resemble those recognized by the eukaryotic RNA polymerase II [26], thus implying a similar functional recognition mechanism of the archaeal RNA polymerase [26,27]. Most proteins encoded by SSV1 have little sequence similarities with other characterized proteins in public databases; consequently, their function has been challenging to infer through homology analysis. Notorious exceptions are some of the SSV1 structural proteins (VP1, VP2, and VP3) and the integrase [22,28].

The purpose of this mini-review is to provide a concise overview of all the work performed on SSV1 including genomic, transcriptomic, and proteomic studies; virus–host interaction; and co-evolution.

## 2. Genomic Analysis

### 2.1. The Fusellovirus Core Genome

Previous efforts to identify the core genes of fuselloviruses used 9–11 genomes and identified a set of 12–13 genes conserved in members of this family [16,29,30]. However, there are 50 fusellovirus sequences currently available in the NCBI, and we set up to determine if the earlier defined set of core genes accurately reflected gene conservation of the expanded family. The genomic sequences of the *Fuselloviridae* available in the NCBI (June 2022) were downloaded and clustered to 98% nucleotide identity and used for further analysis (49 genomes). A total of 1595 ORFs were retrieved from these genomes, and OrthoFinder v2.5.4 was used to define clusters of homologous proteins [31,32] (see Supplementary Files S1–S4). This resulted in 109 clusters of two or more protein sequences (orthogroups) covering 1528 (96%) of the ORFs analyzed, plus 67 singletons (proteins with no homologs in other fuselloviruses). A heatmap of the distribution of orthogroups in the fuselloviruses is shown in Figure 2. With the result of our analysis, we revisited the definition of the *Fuselloviridae* core genome. We propose 17 genes as the soft-core of the fuselloviruses, which are present in over 80% of the genomes (Figure 2, red line). The shell genome (yellow line) includes genes present in >20% of the genomes, and the variable genome in Figure 2 (purple line) contains genes with more than two occurrences in the dataset. For simplicity, singletons are not shown.



**Figure 2.** The pangenome of the *Fuselloviridae*. Presence and absence heatmap of clusters of homologous proteins in 49 fuselloviruses. Genes present in > 80% of the strains were labeled as part

of the soft-core genome (red line), genes present in >20% of the genomes are part of the shell-genome (yellow line), and the rest are considered the cloud-genome (purple line). Singletons are not shown. Pink circles indicate virion proteins, and green circles indicate putative transcriptional regulators. Asterisks (\*) indicate that location of the reference gene within the corresponding genome is putative.

## 2.2. Location of Core Genes

The fuselloviruses have a similar genome organization where genes are clustered into 10–12 transcriptional units and homologous genes are encoded usually in the same transcript across different genomes [11,24,25,33]. Transcripts  $T_{4/7/8}$ ,  $T_9$ , and  $T_{1/2}$  contain most of the core-genome ORFs, while the shell and variable portions of the genome are identifiable in transcripts  $T_5$  and  $T_6$ . The latter are highly flexible regions in the fuselloviruses and contain most of the non-essential genes identified in previous functional genomics studies [11,34,35]. Among the genes in the soft-core genome, nine are present in all genomes and constitute the absolute-core (VP1, VP4, D355-integrase, A82, A92, B251-AAA ATPase, C166, A154, and B115), while four genes are present in over 90% of the genomes (VP3, B129, B277, and B78). The four remaining soft-core genes (A79, C102, C80, and A100), despite being missing from a small subgroup of viruses isolated primarily from Russia, were thus far considered as part of the absolute-core [29,30].

## 2.3. Functional Annotation of Core Genes

The 13 formerly defined core genes [14,29,30] were kept in our analysis as part of the soft-core, and we incorporated four additional genes (A92, VP3, C102, and A100). Core genes VP1, VP3, VP4, and C166 play a structural role [28,36–38], the integrase D355 serves virus integration into the genome [22], and proteins B251 an ATPase with similarity to the bacterial DnaA [39] and A154 have been implicated in replication and packaging of the genome. Indeed, B251 and A154 are also encoded by the satellite virus pSSV $\times$ , which replicates during infection of an SSV2 helper [40]. Although homologs of the tail-spike VP4 are present in all the SSV sequences analyzed, six genomes do not have a full-length version of the protein and it has been questioned if this truncated version is able to complement the structural function of the full-length VP4 [30]. Three of these genomes (SSV NL101B.C01.09, NL101B.C01.18, and NL101B.C01.22) harbor two VP4 homologs that seem to be the result of the full protein split into its N- and C-terminal domains. For the remaining three (SSV6, SSV19, and ASSV1), it was recently reported that the short VP4 homolog serves as an adaptor protein for the tail, while another protein (belonging to OG0000087) has the same genomic location as VP4 and serves as the tail-spike [38]. A truncated VP4 and the OG0000087 tail-spike are associated with viruses that show a pleomorphic morphology [30]. Since genome organization seems conserved in fuselloviruses and given that transcripts  $T_{4/7/8}$  and  $T_{1/2}$  appear to encode for structural proteins, we predict B78 to be also a virion protein as it has two transmembrane helices (similarly to VP1, VP3, VP4, and C166, which also contain transmembrane helices). Intriguingly, homologs of this protein are missing from SSV6, SSV19, and ASSV1, which have a different tail structure [38]. Interestingly, the absolute core gene B115, which is a putative transcriptional regulator, is encoded in  $T_{4/7/8}$  and may be involved in the regulation of structural gene expression. The core protein A82 was recently demonstrated to act as a toxin that kills uninfected cells [41]. The corresponding gene lies in the same region of all SSV genomes, downstream of the genes encoding capsid proteins VP1 and VP3, in the proximity of the core proteins A92 and B277 and of the non-core protein C84. The role of these accompanying proteins in the infection cycle remains to be elucidated and may not be related to the replicative advantage conferred by the toxin [41]. Interestingly, the work of DeWerff et al. [41] identified additional toxins (OG0000019 represented by SSV11p29 and OG0000061 represented by A291), which are not part of the core genome of the *Fuselloviridae*.

#### 2.4. The *Fusellovirus* Pangenome Contains Multiple Transcriptional Regulators

It is noteworthy that around one-third of the genes in the core genome (5/17) are putative transcriptional regulators (A79, C80, C102, B115, and B129), and this holds true also for several genes of the shell- and cloud-genomes (Figure 2, green circles). These transcriptional factors have not been experimentally characterized, although a couple of studies address the role of regulators that do not belong to the core genes [33,42–47]. C80 is homologous to the host transcriptional regulator aCcr-1 (SSO\_RS11690 in *S. solfataricus* P2), a global cell cycle regulator in the Sulfolobales. Homologs of aCcr-1 are widespread in viruses infecting Sulfolobales and are thought to benefit virus replication by driving the cell into the S-phase of the cycle where DNA synthesis occurs. This ensures the presence of proteins and resources required for viral genome replication, which depends on the cellular machinery [47]. However, while the C80 homologs of SSV19–22 and ASSV1 have high sequence similarity to aCcr-1, the other members of the C80 cluster are divergent, cluster separately to aCcr-1 and consist of two different RHH-domains linked together, implying a functional difference [47]. The prevailing presence of transcriptional factors in fuselloviruses suggests that transcriptional regulation of viral gene expression is critical to their infection, as illustrated by F55 (see below), a protein unique to SSV1 involved in the control of early gene expression after UV induction of viral replication [33,45].

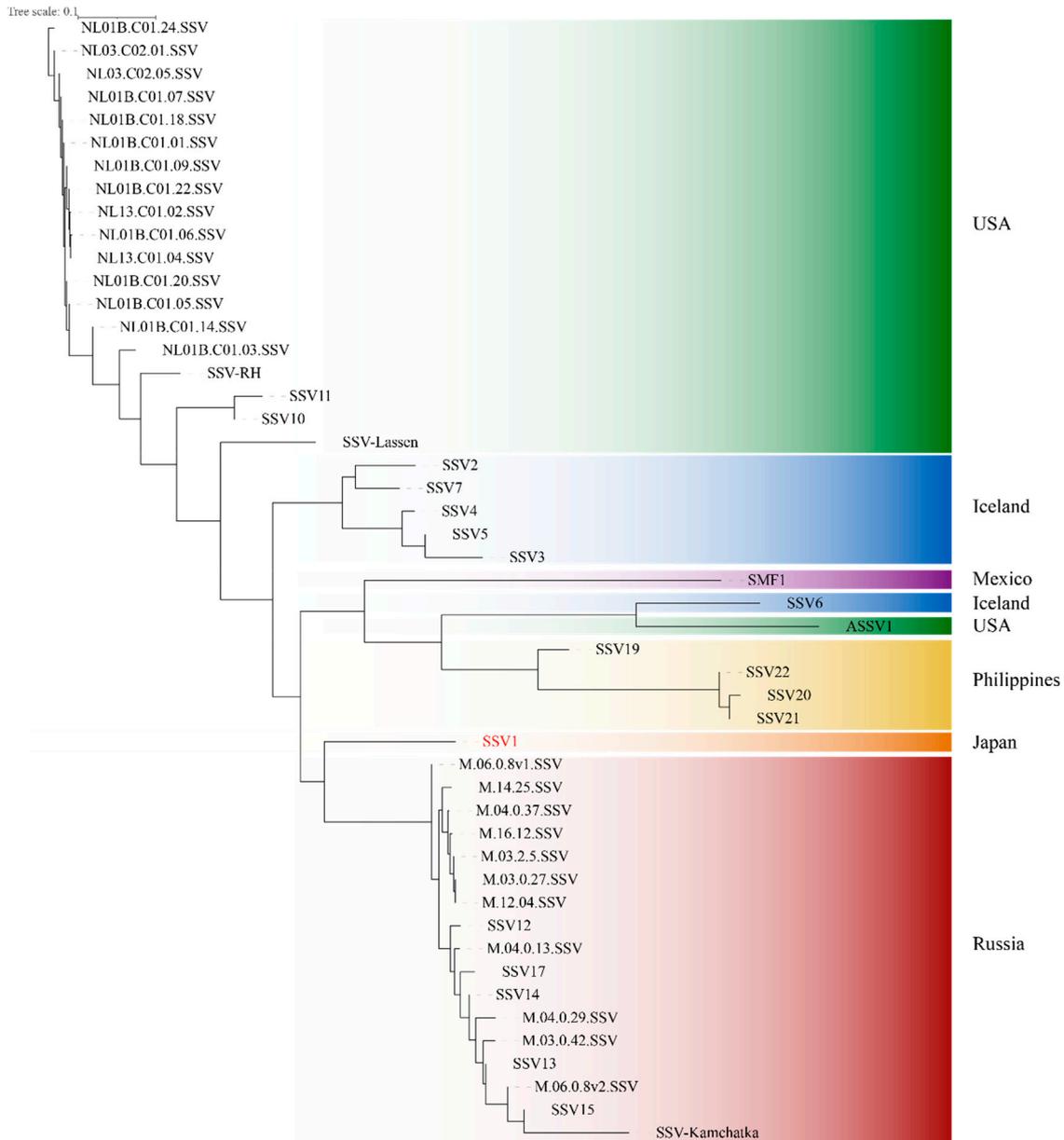
Table 1 summarizes the 10 putative transcriptional factors in SSV1. The remaining core proteins C166, B78, and A100 have no annotation and their functions remain to be investigated.

**Table 1.** Putative transcriptional regulators in SSV1.

Protein	Annotation	Core-Gene	Distribution in Crenarchaeal Viruses	Reference
F93	winged helix-turn-helix (wHTH)	N	<i>Fuselloviridae</i> , <i>Portogloboviridae</i> , <i>Guttaviridae</i>	[48]
F112	winged helix-turn-helix (wHTH)	N	unique to SSV1	[43]
B115	winged helix-turn-helix (wHTH)	Y	<i>Fuselloviridae</i>	
F55	ribbon-helix-helix (RHH)	N	unique to SSV1	[33,45]
C80	ribbon-helix-helix (RHH)	Y	<i>Lipothrixviridae</i> , <i>Bicaudaviridae</i> , <i>Rudiviridae</i> , <i>Spiraviridae</i>	
E51	ribbon-helix-helix (RHH)	N	<i>Fuselloviridae</i>	
A79	zinc-finger	Y	<i>Fuselloviridae</i> , <i>Rudiviridae</i> , <i>Portogloboviridae</i> , <i>Ovaloviridae</i>	
B129	zinc-finger	Y	<i>Fuselloviridae</i>	[44]
A45	zinc-finger	N	<i>Fuselloviridae</i> , <i>Lipothrixviridae</i>	
C102	zinc-finger, putative membrane protein	Y	<i>Fuselloviridae</i> , <i>Guttaviridae</i>	
A100	Pur-alpha repeat	Y	<i>Fuselloviridae</i> , <i>Rudiviridae</i> , <i>Portogloboviridae</i>	

The genome of SSV1 contains a high number of singletons in comparison to the other members of the family (7/35, F55, A132, B49, E96, E178, F92, E54). This is among the highest number of singletons (only surpassed by SMF1, which has 10), as the majority of fuselloviral members (70% of the genomes) possess 0–1 singletons. Most of these unique proteins are encoded in transcript T5, which, together with T6, contains the bulk of the variable gene pool of fuselloviruses. Overall, the core genome of the *Fuselloviridae* appears well defined, and the incorporation of new genomes did not dramatically change the set of core genes. Genome organization is also conserved and maintained in transcriptional units of genes participating in similar processes (e.g., virion structure and assembly). Nevertheless,

the role of most of the viral proteins remains to be elucidated. Finally, a phylogenetic tree of the *Fuselloviridae* family is shown in Figure 3. Viruses cluster according to their geographic location, a pattern that has been observed previously for the *Fuselloviridae* and the *Rudiviridae* [29] and for genes of other viruses [49].



**Figure 3. Phylogeny of the *Fuselloviridae*.** The phylogeny of the 49 fusellovirus sequences was obtained with OrthoFinder, which uses the predicted orthogroups to infer a tree using STAG and roots the tree with STRIDE [31,50,51]. Viruses cluster according to their geographic location, as shown by the color code.

### 3. Transcriptomic Analysis

#### 3.1. SSV1 Genes Regulation

The analysis of SSV1 transcripts dates back to 1987, when Zillig and coworkers performed a pioneering study on transcription expression in *Archaea* using SSV1 as a model system [24,28]. At that time, it was known that SSV1 was a UV-inducible virus able to infect *S. shibatae* and *S. solfataricus*, exhibiting a unique life cycle that did not imply cell lysis [20]. However, the molecular mechanisms underpinning the switch from lysogeny (carrier state)

to the induction state were not known for a long time, as well as those regulating the maintenance of the lysogeny and the reversion to the steady state in the aftermath of UV induction. Moreover, the complex network regulating the host–virus interaction had not yet been elucidated.

### 3.2. SSV1 Genes Expressed in the Carrier State

The first studies on transcriptional mapping of SSV1 highlighted the presence of 11 transcripts that started from 7 mapped promoter sites and covered almost the whole genome, which includes 35 open reading frames (ORFs) [17,24]. Constitutive messengers expressed once the lysogenic state has been established are T1/T2, T3, and Tx, encoding *vp1-2-3*, *a291* (a putative toxin) and *c124* genes [23,25] (Figure 4). A transcriptional activity has been detected also for *d335*, encoding the SSV1 integrase. Then, the carrier state of SSV1 is sustained by the expression of a few genes, among which only those encoding the structural capsid proteins (VP1, VP2, and VP3) [34,52] and the integrase have been functionally characterized [22]. Moreover, a recent study has revealed that most of the lysogenic genes (*vp1*, *vp2*, *d355*, and *a291*) are apparently essential for SSV1 infectivity, except for *c124* and *vp3* [16]. Subsequently, a transcriptomic analysis revealed an additional mRNA (named T<sub>lys</sub>) in uninduced cells. T<sub>lys</sub> encodes a 55-amino-acid protein (F55) that is involved in the repression of the early UV-inducible genes (see below) in the absence of UV irradiation [33]. Genetic functional analysis on *f55* has shown that this gene is not required to produce infectious virus particles. Rather, it seems that  $\Delta F55$  mutants exhibit a constant expression of early gene products because of the absence of the repressive action of F55 on the early promoters [16].

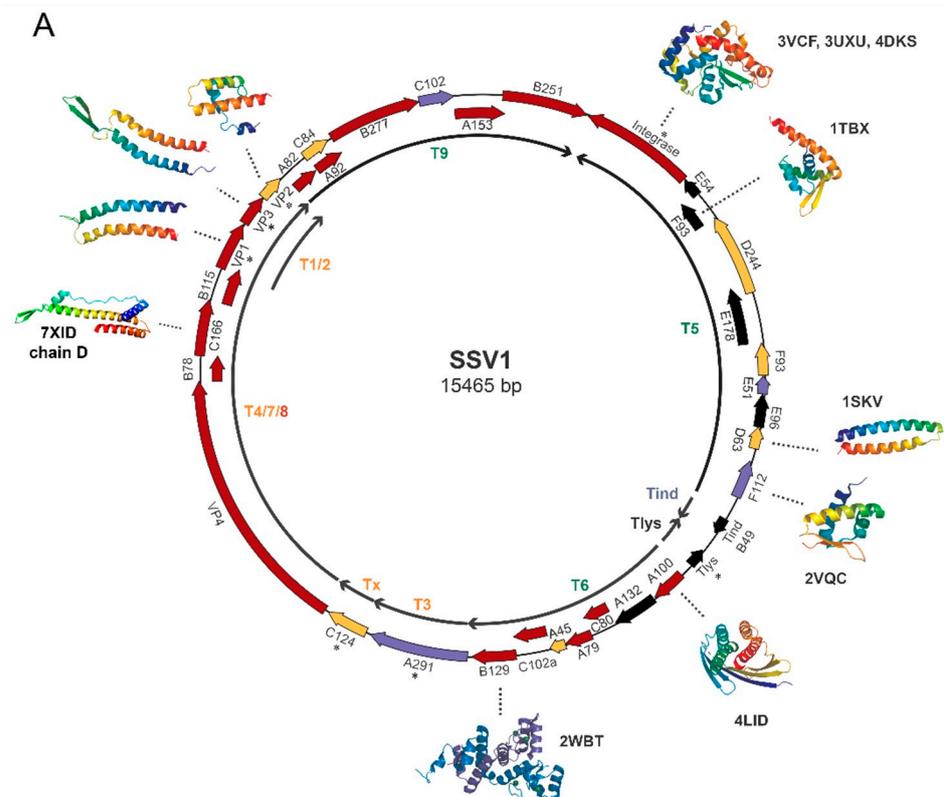
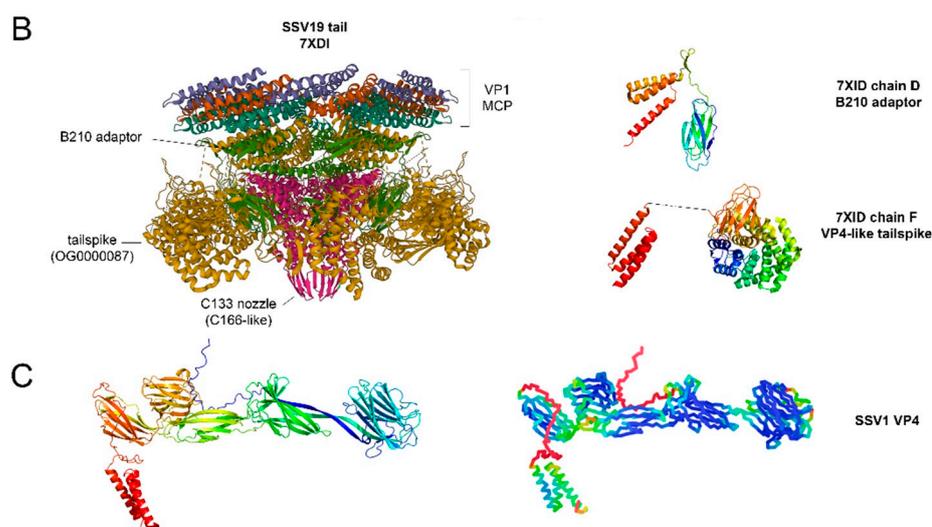


Figure 4. Cont.



**Figure 4. Genomic map of SSV1.** (A) ORFs in the genome of SSV1 are depicted as arrows blocks, and coloring corresponds to their classification as part of the soft-core (red), cloud (yellow), or variable (purple) genome. Proteins expressed during lysogeny are marked with an asterisk (\*). Inner arrows represent the transcripts of SSV1, and the font color indicates their classification according to their transcriptional profile into UV-induced (blue), early (green), late (orange), and late-extended (red) transcripts. Tlys is transcribed in the lysogenic state. Crystallographic structures available for SSV1 open-reading frames are depicted for the integrase D355, F93, D63, F112, A100, B129, and C166 (from the SSV19 homolog C131). The structures of VP1, VP2, VP3 (A), and VP4 (C) were modelled using AlphaFold2 [53]. (B) Cryo-electron micrograph of the SSV19 tail (7XID, [38]). The atomic models of the B210 adaptor and the VP4-like tail-spike are shown to the right. (C) Model structure of SSV1 VP4. Left: atomic model of SSV1 VP4. Right: model colored according to the confidence measure of the prediction (pLDDT) on a scale from 0 to 100, where blue indicates high confidence and red indicates low confidence.

### 3.3. SSV1 Genes Expressed upon UV Induction

The study of the SSV1 life cycle in the aftermath of UV induction revealed the very early expression of  $T_{ind}$  transcript (Figure 4) whose role is still murky [16,25]. Initially, it was hypothesized that this transcript triggered the initiation of replication, thus pointing to the  $T_{ind}$  surrounding region as the possible replication origin of SSV1 [24]. Subsequently, the identification of a 49 aa long protein (B49) within  $T_{ind}$  messenger suggested an alternative role in switching on the viral transcription, either directly or indirectly [19,33]. Indeed, secondary structure prediction indicates that B49 might bear a helix-turn-helix motif, thus indicating a possible role as transcription factor. However, expression trials of B49 gene in *E. coli* and in *S. solfataricus* were unsuccessful (unpublished data), thus casting doubts about the true translation of  $T_{ind}$  into a functional B49 protein. Although direct evidence of the  $T_{ind}$  role is still lacking, the fact that its expression extensively increased in the timeframe following the UV induction and preceding the onset of SSV1 replication suggested that it plays a role in UV-induced replication either as a transcript and/or as a protein product [24,54]. Finally, mutational studies demonstrated that deletions of  $T_{ind}$  regions are indeed tolerated, thus showing that this transcript is dispensable for the infectivity of virus particles in the absence of UV stimulus [16] (Figure 4). Upon UV induction, the very early expression of  $T_{ind}$  transcript is followed by the transcription of the early ( $T_5$ ,  $T_6$ , and  $T_9$ ), late ( $T_{1/2}$ ,  $T_3$ ,  $T_x$ , and  $T_{4/7}$ ), and late-extended ( $T_{4/7/8}$ ) RNAs [24,25]. This cascade of events led, in turn, to the induction of the SSV1 genome replication and eventually to a steep increase in the viral titer. The two subsets of  $T_1/T_2$  and  $T_4/T_7/T_8$  transcripts originated from a common promoter located upstream of  $T_1$  and  $T_4$ , respectively, and terminator read-through resulted in the formation of more than one RNA species [17,26] (Figure 4).

### 3.4. Promoters and Terminators of SSV1

Mapping of the transcriptional start sites (TSSs) of the SSV1 transcripts has allowed the discovery of archaeal promoter elements such as the TATA-box-like hexanucleotide sequence (box A, TTTAAA) and the nearby TGA consensus trinucleotide sequence (later identified as the BRE, B recognition element sequence) located about at  $-23$  and  $-28$  nucleotides upstream of the TSS [24,26]. Moreover, transcription termination signals of SSV1 messengers were associated to the presence of a TTTTTYT conserved motif. All these consensus sequences resembled those found in eukaryotic promoters and terminators recognized by RNA polymerase II. This evidence mirrored a similar functional recognition mechanism of the archaeal RNA pol as confirmed by its homology to the eukaryotic RNA polymerase [26,27].

### 3.5. SSV1 Transcription Regulation

The transcriptional regulation of SSV1 genes upon irradiation is quite complex and implies a combination of different mechanisms including mRNA degradation [24], early transcriptional termination/antitermination processes [26], up- and downregulation [20], and possible cooperation of some of these. For instance, terminator read-through took place for terminators of transcripts T7 and T4 that were included in a polycistronic T8 messenger upon UV irradiation. Likewise, T2 transcript was included in T1 upon UV exposure [26]. A similar mechanism is employed by pSSV<sub>x</sub>, thus suggesting that viruses can discard the regulative role of transcriptional sequences under conditions that trigger virus replication [10]. Whilst most of the mechanisms featuring the viral transcriptional response to UV irradiation are known, the host and/or viral molecular components involved in this process as well as the chronological transcriptional activation have not been elucidated yet. Conversely, a fine-tuned transcriptional regulation circuit that ruled the maintenance of the SSV1 carrier state has been recently unveiled [33,54]. The discovery of the T<sub>lys</sub> transcript led to the evidence that its encoded transcriptional factor F55 is the main viral player in the maintenance of the lysogeny of SSV1 and in regulating the transition from the carrier to the UV-inducible state. Intracellular concentration of F55 was regulated by transcriptional activation/repression as well as by a post-transcriptional mechanism of RNA degradation, as observed in the aftermath of UV irradiation [33,54]. A detailed description of F55's role is included in its proteomic paragraph. It remains to be established whether other regulation circuits might be involved in the transcriptional activation of the UV-inducible promoters.

## 4. Proteomic Analysis

As for most of the spindle-shaped archaeal hyperthermophilic fuselloviruses, for a long time, putative genes of SSV1 did not match with any genes in public databases [55,56]. Consequently, it has been challenging to infer, through homology analysis, the role of SSV1 proteins in its life cycle, although some structural and bioinformatics investigations have provided reasonable functional predictions for several other SSV1 proteins [44]. The function of some SSV1 structural proteins (VP1, VP2, and VP3) was discovered in the past [24]. Conversely, all the constituents of SSV1 virus particles (virions) remained only partially described until 2015, when Quemain et al. carried out a thorough biochemical characterization of SSV1 virions [36]. Specifically, they identified a fourth (besides VP1, VP2, and VP3) virus-encoded structural protein, i.e., VP4 (formerly known as C792). Thus, the bulk of SSV1 virions is made up of multiple copies of VP1, VP2, and VP3. Terminal virion fibers involved in virion aggregation [36] and in host recognition and attachment [52] are instead constituted by oligomers of VP4. Moreover, by means of a glycoprotein-specific stain, it was shown that VP1, VP3, and VP4 are glycosylated at multiple sites. As for other proteins from *Sulfolobales* [57], SSV1 structural proteins are glycosylated at asparagine residues within the consensus motif N-X-S/T (where X is any amino acid except proline). Both VP1 and VP3 possess two such motifs, whilst VP4 contains 20 consensus motifs that could potentially undergo glycosylation [36].

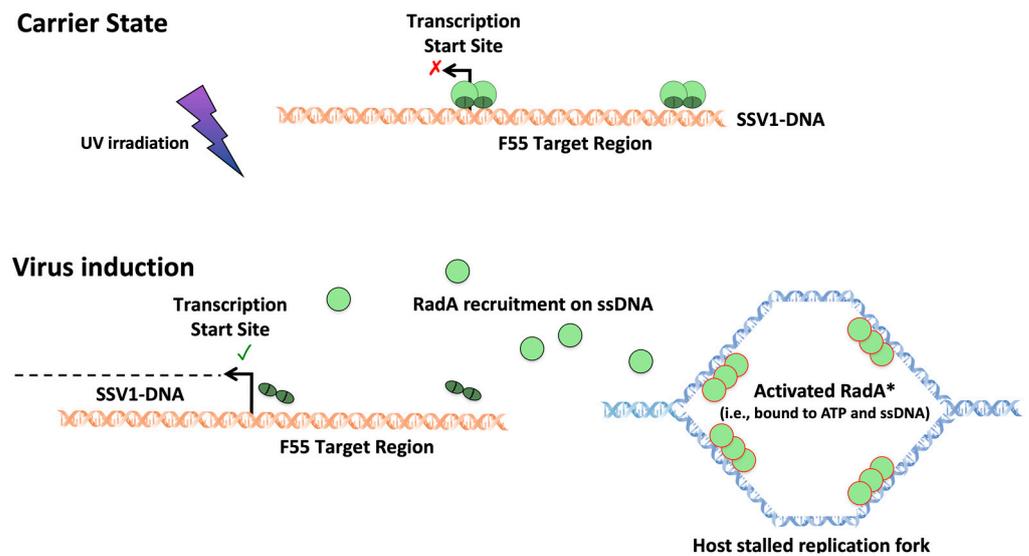
Furthermore, the same authors discovered that a DNA-binding protein of cellular origin (Sso7d) is part of the virion structure. Since the host-encoded Sso7d is a member of the Sul7d protein family, which are responsible for chromosome organization in *Saccharolobus*, it has been speculated that Sso7d could act during virion packaging to organize and condense the viral genome. Whereas VP1 and VP3 paralogous have homologs in all known spindle-shaped viruses, VP2 has been found to be encoded only by 10 fuselloviruses for which complete genome sequences are available (Figure 2, shell genome, OG0000046). Moreover, VP2 is not specific to fuselloviruses; indeed, homologs of this protein are encoded by unrelated archaeal and bacterial viruses. In accordance with the scarce conservation among fuselloviruses, deletion of the *vp2* gene did not affect virion assembly or infectivity, thus suggesting that the protein VP2 is dispensable at least under laboratory growth conditions. In line with its scattered phyletic distribution and dispensability, it has been hypothesized that the *vp2* gene was acquired relatively late in the history of fuselloviruses from a different group of viruses [36].

So far, among the 35 predicted SSV1-encoded proteins, structural data are available only for very few proteins; the last SSV1 protein structure (A100) was released in 2014. Atomic resolution structures are available for five putative proteins with unknown functions (i.e., B129, D63, F93, F112, and A100), as well as for the integrase (D335) [55,58,59]. Nevertheless, the lack of functional characterization makes it difficult to assign a defined function to these proteins. For instance, a reasonable prediction of D63 function was limited by the fact that its secondary structural motifs are widespread in many proteins covering a broad range of functions. Only the similarity of its three-dimensional structure with the bacterial adaptor protein repressor of primer (ROP) suggested a possible role of D63 in the regulation of the SSV1 episomal copy number [55,60]. Conversely, structural analyses of the integrase have shown a similar core fold to both bacterial and eukaryal recombinases, except for the lack of the typical helix corresponding to  $\alpha I$  of Cre. Therefore, this enzyme was unequivocally assigned to the type I tyrosine recombinase family. Moreover, functional analyses have shown that it possesses an N-terminal domain responsible for protein dimerization and a C-terminal domain capable of DNA cleavage and ligation [59].

Except for few examples, most of SSV1 proteins remain uncharacterized and/or lack an assigned function. In the attempt to address this issue, Iverson et al. designed a genetic approach that included both specific and random mutagenesis to introduce mutations in all SSV1 ORFs [16]. Although this approach did not clarify the exact role of SSV1 proteins, it helped to understand at least which proteins were essential for the infectivity of the SSV1 virions. In particular, by means of long inverse PCR (LIPCR) and transposon mutagenesis, the authors constructed 78 variants of the SSV1 genomes harboring mutations in each of the 35 ORFs. All these virus mutants were tested for their ability to infect the permissive host *S. solfataricus*. By doing so, the authors have shown that nearly half of the SSV1 genes (16 out of 35 ORFs) were dispensable for the SSV1 virions infectivity and that most of essential genes are highly conserved within the *Fuselloviridae* family and constitute the so-called "fusellovirus core" [34]. Whereas almost the entire T5 early transcript was shown to be dispensable (7 out of 10 ORFs), many proteins encoded by the T6 early transcript were essential, which is in line with the abundance of well-conserved fusellovirus core genes in this transcript. Most of core genes were intolerant to insertions and/or deletions, thus indicating an essential role of the encoded proteins for the SSV1 infection cycle. A surprising exception was represented by the minor capsid gene *vp3* that, despite belonging to the fusellovirus core genes, appears to be dispensable for SSV1 infectivity. However, mutations of this gene resulted in an abnormal morphology with virions longer and thinner than the wild type [16]. Even though deletion of the integrase gene led to a virus mutant (SSV1- $\Delta$ d335) that could not infect the strain used in the study by Iverson and colleagues, this core gene was dispensable at least in some other hosts (i.e., *S. solfataricus* strain P2 and G $\theta$ ) [61]. These apparently contradictory results might have been due to the presence of a host-encoded integrase that, in some *Saccharolobus* strains, could be active on SSV1 DNA [16,34]. This study also showed that *b49* and *f55* can be deleted from SSV1 without

loss of infectivity. However, the spot-on-lawn or halo assays performed in this study did not allow for the quantification of virion number in the culture supernatant; therefore, it cannot be completely excluded that the deletion of *f55* caused an alteration of SSV1 titer during lysogeny (carrier state) without interfering with the infectivity of SSV1 virions.

Even though SSV1 has been the most studied fusellovirus in the past decades, the molecular mechanism involved in the transition from carrier state to transcription/replication induction remained unclarified for long time. It was only in 2013 that the discovery of an additional and previously unidentified viral transcript (i.e.,  $T_{lys}$ ) allowed the shedding of light on how this virus regulates its infection cycle [33]. This initial in vitro study demonstrated that  $T_{lys}$  encodes for a 55-amino-acid protein (F55). The recombinant F55 was shown to specifically bind the SSV1 DNA at an 11 nt binding sequence (consensus sequence: 5'-ATAGATAGAGT-3'), which is present in several copies in the promoter region of  $T_5$ ,  $T_6$ , and  $T_{ind}$ , as well of its own transcript (i.e.,  $T_{lys}$ ). In order to elucidate the regulatory role of F55, more analyses were carried out to measure the level of F55 protein (before and after UV light exposure) in infected cells of *S. solfataricus* to confirm its interaction in vivo with the promoters of  $T_5$ ,  $T_6$ ,  $T_{ind}$ , and  $T_{lys}$  [54]. Upon UV irradiation, it was shown that F55 first dissociates from target sequences in the  $T_{ind}$  promoter (2 h post treatment) and later from those in the  $T_5$  and  $T_6$  promoters (4 h post-treatment). The progressive release of this transcription regulator from its binding sites is in agreement with the differential (in vitro measured) affinity of F55 for these sequences ( $K_d = T_5 \simeq T_6 < T_{ind}$ ) [33]. Given that  $T_{ind}$  and  $T_5/T_6$  are transcribed after 2 h and 4 h after UV irradiation, respectively, the release of F55 from the promoter regions of these transcripts might be the molecular event that activates their transcription [54]. Nonetheless, it was still not clear how SSV1 could “sense” the host cell damage caused by the environmental stress (i.e., the UV irradiation). For this reason, a variant of the electrophoretic mobility shift assay (EMSA) coupled to Western blot and mass spectrometry analyses was used to unravel the protein-protein interaction network (interactome) between SSV1 and its host *S. solfataricus* [45]. Intriguingly, among F55 interactors, the enzyme RadA (the archaeal homolog of the bacterial RecA) was identified, which is essential for genome stability. Specifically, RadA mediates non-specific DNA repair by binding single-stranded DNA (ssDNA) at collapsed replication forks and catalyzing homologous base pairing and strand exchange [62]. The interaction of a viral regulator with a host protein involved in DNA repair it is not unusual. For instance, it is used by the well-characterized bacteriophage  $\lambda$  to regulate the switch from lysogenic to lytic development. Interestingly, it was shown that in SSV1-infected cells, RadA binds to F55 to generate a stable tripartite RadA–F55–dsDNA complex that can be destabilized by the formation of stalled DNA replication forks (Figure 5). Because of the RadA recruitment at ssDNA regions, the interaction of F55 with DNA at its target sites was destabilized, thus leading to the dissociation of this transcription repressor from promoters of  $T_{ind}$ ,  $T_5$ , and  $T_6$  that, in turn, activates their transcription. Even though it has not been unequivocally demonstrated that the interaction between F55 and its operators is destabilized because of RadA detachment, it was speculated that the interaction between F55 and RadA represents the molecular sensor that SSV1 uses to control host genome integrity and, in turn, to regulate its own infection cycle [45].



**Figure 5. SSV1 carrier-to-induction state transition stimulated by host DNA damage.** In the carrier state, multimers of F55 are bound to all the target sites on the SSV1 genome, thus repressing transcription (represented as red cross) of the viral early genes. The interaction of F55 to the virus DNA is stabilized by the co-presence of the host encoded RadA protein. UV irradiation damages the host DNA, which in turn causes the accumulation of stalled replication forks (ssDNA) in regions of the host genome where DNA contains unrepaired lesions or lesions that are undergoing repair. This event fosters the recruitment of RadA to ssDNA regions (indicated as activated RadA\*), where it forms nucleoprotein filaments. In turn, this event results in the progressive release of RadA from the virus DNA and, later on, in the dissociation of F55 from its target sites with the consequent activation of the temporary coordinated transcription pattern of SSV1. The model is adapted from Fusco et al., 2022 [45].

## 5. Host–Virus Interaction

SSV1 was first identified as a plasmid replicating in *S. shibatae* B12 [63]. Then, the viral nature of SSV1 was demonstrated when spindle-shaped virus-like particles present in the supernatant of the culture were found to infect *S. solfataricus*. The conditions to obtain the highest virus titer under lab conditions have been studied, and a protocol for the UV induction of SSV1 replication based on the combination of several parameters is available [15]. Since the physiology of the SSV1 life cycle was identical in *S. shibatae* and *S. solfataricus*, the latter has been widely used as a host model for molecular, genetic, and physiological studies on SSV1 [63]. It integrates into the host genome at the arginyl-tRNA gene, which is maintained intact after integration. Moreover, the integrase that promotes site-specific integration has been biochemically characterized [20,61]. As a unique feature of the host–virus interaction of an UV-inducible virus, the release of SSV1 viral particles does not cause lysis of host membrane, and cells recover their growth rate as well as their carrier state within a few hours [20].

### 5.1. Transcriptome Analysis of *S. Solfataricus*

A transcriptomic analysis has been performed on *S. solfataricus* lysogens to study the variation of host/virus gene expression upon UV irradiation [23]. It was observed that the host response included down- or upregulation of several genes in comparison to the uninfected cells, suggesting a major sensitivity of SSV1 lysogens to UV irradiation. Only a few genes exhibited a differential transcriptional variation specifically related to the presence of the SSV1 virus; however, the strength and speed of the general UV response was significantly different in uninfected and infected cells, indicating that the presence of the virus and/or of some viral components increased the host sensitive to the UV-induced DNA damage. Among the very few genes that showed a differential response, worth

noting are those encoding the two subunits of topoisomerase VI that might be involved in DNA supercoiling during DNA replication [25]. Interestingly, the temporal concurrence of the upregulation of chromosomal genes and of the early genes of SSV1 indicates a potential virus–host co-regulation. Differently from what was seen for infected cells upon exposure to UV irradiation, the host metabolism and growth is not visibly impaired by the presence of the SSV1 virus in the absence of UV stimulus. Indeed, about 30 genes were found to be up- or downregulated in stable SSV1 lysogens, confirming that SSV1 had only a limited effect on the host gene expression in the carrier state [23]. The absence of a strong host response hinted to a harmonic host–virus co-existence in the absence of UV irradiation. Accordingly, SSV1 is a temperate virus able to self-regulate its own replication through the pleiotropic effect of F55, which represses the expression of the UV-inducible as well as of the early genes (see above) [33,54]. Whereas SSV1 did not induce major changes on the host gene expression, SSV2, which is closely related to SSV1 in terms of morphology, genome organization, and gene synteny, elicited a strong host response (162 up- and downregulated genes vs. 30 SSV1 genes), which includes transcriptional activation of CRISPR loci and *cas* genes. This response has been attributed to the difference in life cycles between SSV1 and SSV2, since only the latter exhibits a strong physiological induction during the host growth. This led to the activation of the CRISPR response to control the SSV2 induction and to safeguard host genome integrity through deletion of self-targeting spacers [23]. Interestingly, it was observed that in lysogens containing both SSV1 and SSV2, the presence of the former quenched the effect of SSV2 infection on the host gene expression, and specifically, no differential expression has been observed for genes and clusters of the CRISPR-Cas system. It is plausible that transcription factors and/or other molecular components encoded by SSV1 are involved in silencing the CRISPR-Cas response in the SSV1-SSV2-infected strain [23]. The harmonic coexistence between SSV1 and its host is possible thanks also to the virus egress strategy that apparently do not cause cell lysis [20,64]. The morphogenesis of the SSV1 virus particles has been extensively studied in the natural host *S. shibatae*, demonstrating that SSV1 assembly and egress are concomitant and occur at the cellular cytoplasmic membrane via a budding-like process followed by the maturation into characteristic spindle-shaped virions. The proposed model explains why SSV1 is encased in lipid-containing envelopes similarly to the enveloped eukaryotic viruses that exploit the same egress strategy. Less understood is how SSV1 gains access to the *Saccharolobus* host cells. Previous studies indicated that the S-layer in *S. solfataricus* P1 is required for SSV1 adsorption since S-layer depleted cells are less susceptible to SSV1 infection [65]. To sum up, the coexistence of the virus with the natural and *S. solfataricus* lysogens is characterized by (i) maintenance of low copy number in the absence of UV irradiation, (ii) negligible effect on the host metabolism, and (iii) the non-lytic egress of SSV1 [64,66]. All together, these features allowed for a close co-evolution of this virus with its natural host.

### 5.2. Regulation of CRISPR-Cas System during the Infection

Compared to other SSV representatives, SSV1 seems to have the narrowest host range, infecting only three *S. solfataricus* strains and a single *Saccharobus* sp. isolated from Lassen Volcanic National Park [52]. Records of previous infection of *S. tokodaii* have been revealed by the analysis of the CRISPR spacers containing a 41 bp long sequence matching SSV1. However, *S. tokodaii* turned out to be resistant to the infection of all the SSVs viruses, including SSV1. This suggests that this host has developed efficient strategies to circumvent the SSVs virus entry/infection. Similarly, *S. acidocaldarius* is also insensitive to SSV infection, although no CRISPR matches toward these viruses have been found [12]. It is now widely accepted that CRISPR-Cas defense systems might have, to a large extent, shaped the evolution of the viral genomes. Nevertheless, the above-mentioned pieces of evidence suggest that the infectivity of SSVs toward their potential host is related not only to the CRISPR system but also to complex not yet elucidated mechanisms, including regulation of adsorption, other aspects of DNA entry, genome replication, transcription, translation,

assembly, or virus release. At present, there is no explanation for the narrow host range of SSV1. In general, SSV infectivity and *Saccharolobus* susceptibility are independent of the geographic regions from which the hosts and viruses were isolated [12]. It is tempting to speculate that the UV-inducible nature of SSV1, which is unique among all the known crenarchaeal viruses, along with the lack of evolutionary connection with other SSVs, have influenced the virus physiology and in turn the “coevolutionary arms race” of the host–virus interaction [67].

## 6. Conclusions

Viruses represent one of the major agents of evolution by virtue of their capacity to operate exchange of genetic material (horizontal gene transfer, HGT) and to move actively among biomes [68]. In particular, archaeal virology and metaviromics are opening the way to answer questions regarding the origin of viruses and their position in the global virosphere [68]. Furthermore, the study of virus–host model systems, such as SSV1 and *Saccharolobus* spp., is functional to shedding light on the evolution and on the regulatory mechanisms of viruses and their relative hosts [9,10,24,28,69–72]. Comprehensive studies on SSV1 have definitely contributed to elucidating fundamental molecular mechanisms in *Saccharolobus* hosts as well as to set up valuable genetic tools [21,73–75]. For instance, these systems have been successfully used to unveil the role of SSV1 proteins in the infection cycle [16] and to express heterologous enzymes [18,75]. Finally, a wide variety of biotechnological applications such as medicine, agriculture, pharmacology, cosmetics, electronics, nanotechnology, and environmental safeguarding locate extremophilic microorganisms [76–86] and their viruses as valuable tools [2,87,88].

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/v14102082/s1>, File S1: Genome accession numbers of the members of the Fuselloviridae family (15 January 2022); File S2: Presence-absence matrix of groups of orthologous proteins in 49 members of the Fuselloviridae; File S3: Accession numbers of proteins of the different orthologous groups of the Fuselloviridae; File S4: Annotation of fuselloviral capsid proteins VP1 and VP3 using the prokaryotic viral ortholog groups (pVOGs) database.

**Author Contributions:** M.A., L.M.-A., S.F. and P.C. drafted the manuscript; Q.S., S.B. and X.P. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (grant no. 31771380 to Q.S.).

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Prangishvili, D.; Bamford, D.H.; Forterre, P.; Iranzo, J.; Koonin, E.V.; Krupovic, M. The enigmatic archaeal virosphere. *Nat. Rev. Microbiol.* **2017**, *15*, 724–739. [[CrossRef](#)] [[PubMed](#)]
2. Gil, J.F.; Mesa, V.; Estrada-Ortiz, N.; Lopez-Obando, M.; Gómez, A.; Plácido, J. Viruses in extreme environments, current overview, and biotechnological potential. *Viruses* **2021**, *13*, 81. [[CrossRef](#)] [[PubMed](#)]
3. Maaty, W.S.A.; Ortmann, A.C.; Dlakić, M.; Schulstad, K.; Hilmer, J.K.; Liepold, L.; Weidenheft, B.; Khayat, R.; Douglas, T.; Young, M.J.; et al. Characterization of the Archaeal Thermophile *Sulfolobus* Turreted Icosahedral Virus Validates an Evolutionary Link among Double-Stranded DNA Viruses from All Domains of Life. *J. Virol.* **2006**, *80*, 7625–7635. [[CrossRef](#)] [[PubMed](#)]
4. Krupovic, M.; Cvirkaite-Krupovic, V.; Iranzo, J.; Prangishvili, D.; Koonin, E.V. Viruses of archaea: Structural, functional, environmental and evolutionary genomics. *Virus Res.* **2018**, *244*, 181–193. [[CrossRef](#)] [[PubMed](#)]
5. Baquero, D.P.; Contursi, P.; Piochi, M.; Bartolucci, S.; Liu, Y.; Cvirkaite-Krupovic, V.; Prangishvili, D.; Krupovic, M. New virus isolates from Italian hydrothermal environments underscore the biogeographic pattern in archaeal virus communities. *ISME J.* **2020**, *14*, 1821–1833. [[CrossRef](#)]
6. Contursi, P.; Fusco, S.; Cannio, R.; She, Q. Molecular biology of fuselloviruses and their satellites. *Extremophiles* **2014**, *18*, 473–489. [[CrossRef](#)]

7. Sakai, H.D.; Kurosawa, N. *Saccharolobus caldissimus* gen. nov., sp. nov., a facultatively anaerobic iron-reducing hyperthermophilic archaeon isolated from an acidic terrestrial hot spring, and reclassification of *Sulfolobus solfataricus* as *saccharolobus solfataricus* comb. nov. and. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 1271–1278. [[CrossRef](#)]
8. Wiedenheft, B.; Stedman, K.; Roberto, F.; Willits, D.; Gleske, A.-K.; Zoeller, L.; Snyder, J.; Douglas, T.; Young, M. Comparative Genomic Analysis of Hyperthermophilic Archaeal Fuselloviridae Viruses. *J. Virol.* **2004**, *78*, 1954–1961. [[CrossRef](#)]
9. Contursi, P.; Jensen, S.; Aucelli, T.; Rossi, M.; Bartolucci, S.; She, Q. Characterization of the *Sulfolobus* host-SSV2 virus interaction. *Extremophiles* **2006**, *10*, 615–627. [[CrossRef](#)]
10. Aucelli, T.; Contursi, P.; Girfoglio, M.; Rossi, M.; Cannio, R. 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.* **2006**, *34*, e114. [[CrossRef](#)]
11. Goodman, D.A.; Stedman, K.M. Comparative genetic and genomic analysis of the novel fusellovirus *Sulfolobus* spindle-shaped virus 10. *Virus Evol.* **2018**, *4*, vey022. [[CrossRef](#)] [[PubMed](#)]
12. Ceballos, R.M.; Marceau, C.D.; Marceau, J.O.; Morris, S.; Clore, A.J.; Stedman, K.M. Differential virus host-ranges of the Fuselloviridae of hyperthermophilic Archaea: Implications for evolution in extreme environments. *Front. Microbiol.* **2012**, *3*, 295. [[CrossRef](#)] [[PubMed](#)]
13. Ceballos, R.M.; Drummond, C.G.; Stacy, C.L.; Padilla-Crespo, E.; Stedman, K.M. Host-Dependent Differences in Replication Strategy of the *Sulfolobus* Spindle-Shaped Virus Strain SSV9 (a.k.a., SSVK1): Infection Profiles in Hosts of the Family Sulfolobaceae. *Front. Microbiol.* **2020**, *11*, 1218. [[CrossRef](#)] [[PubMed](#)]
14. Zhang, J.; Zheng, X.; Wang, H.; Jiang, H.; Dong, H.; Huang, L. Novel *Sulfolobus* Fuselloviruses with Extensive Genomic Variations. *J. Virol.* **2020**, *94*, 1624–1643. [[CrossRef](#)]
15. Fusco, S.; Aulitto, M.; Bartolucci, S.; Contursi, P. A standardized protocol for the UV induction of *Sulfolobus* spindle-shaped virus 1. *Extremophiles* **2015**, *19*, 539–546. [[CrossRef](#)]
16. Iverson, E.A.; Goodman, D.A.; Gorchels, M.E.; Stedman, K.M. Extreme Mutation Tolerance: Nearly Half of the Archaeal Fusellovirus *Sulfolobus* Spindle-Shaped Virus 1 Genes Are Not Required for Virus Function, Including the Minor Capsid Protein Gene vp3. *J. Virol.* **2017**, *91*, e02406–e2416. [[CrossRef](#)]
17. Reiter, W.D.; Palm, P.; Zillig, W. Analysis of transcription in the archaeobacterium *Sulfolobus* indicates that archaeobacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res.* **1988**, *16*, 1–19. [[CrossRef](#)]
18. Jonuscheit, M.; Martusewitsch, E.; Stedman, K.M.; Schleper, C. A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. *Mol. Microbiol.* **2003**, *48*, 1241–1252. [[CrossRef](#)]
19. Palm, P.; Schleper, C.; Grampp, B.; Yeats, S.; McWilliam, P.; Reiter, W.D.; Zillig, W. Complete nucleotide sequence of the virus SSV1 of the archaeobacterium *Sulfolobus shibatae*. *Virology* **1991**, *185*, 242–250. [[CrossRef](#)]
20. Schleper, C.; Kubo, K.; Zillig, W. The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: Demonstration of infectivity and of transfection with viral DNA. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7645–7649. [[CrossRef](#)]
21. Iverson, E.; Stedman, K. A genetic study of SSV1, the prototypical fusellovirus. *Front. Microbiol.* **2012**, *3*, 200. [[CrossRef](#)] [[PubMed](#)]
22. Muskhelishvili, G.; Palm, P.; Zillig, W. SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. *MGG Mol. Gen. Genet.* **1993**, *237*, 334–342. [[CrossRef](#)] [[PubMed](#)]
23. Fusco, S.; Liguori, R.; Limauro, D.; Bartolucci, S.; She, Q.; Contursi, P. Transcriptome analysis of *Sulfolobus solfataricus* infected with two related fuselloviruses reveals novel insights into the regulation of CRISPR-Cas system. *Biochimie* **2015**, *118*, 322–332. [[CrossRef](#)]
24. Reiter, W.D.; Palm, P.; Yeats, S.; Zillig, W. Gene expression in archaeobacteria: Physical mapping of constitutive and UV-inducible transcripts from the *Sulfolobus* virus-like particle SSV1. *MGG Mol. Gen. Genet.* **1987**, *209*, 270–275. [[CrossRef](#)]
25. Fröls, S.; Gordon, P.M.K.; Panlilio, M.A.; Schleper, C.; Sensen, C.W. Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays. *Virology* **2007**, *365*, 48–59. [[CrossRef](#)]
26. Reiter, W.D.; Palm, P.; Zillig, W. Transcription termination in the archaeobacterium *Sulfolobus*: Signal structures and linkage to transcription initiation. *Nucleic Acids Res.* **1988**, *16*, 2445–2460. [[CrossRef](#)] [[PubMed](#)]
27. Best, A.A.; Olsen, G.J. Similar subunit architecture of archaeal and eukaryal RNA polymerases. *FEMS Microbiol. Lett.* **2001**, *195*, 85–90. [[CrossRef](#)] [[PubMed](#)]
28. Reiter, W.-D.; Palm, P.; Henschen, A.; Lottspeich, F.; Zillig, W.; Grampp, B. Identification and characterization of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. *Mol. Gen. Genet. MGG* **1987**, *206*, 144–153. [[CrossRef](#)]
29. Held, N.L.; Whitaker, R.J. Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. *Environ. Microbiol.* **2009**, *11*, 457–466. [[CrossRef](#)]
30. Redder, P.; Peng, X.; Brügger, K.; Shah, S.A.; Roesch, F.; Greve, B.; She, Q.; Schleper, C.; Forterre, P.; Garrett, R.A.; et al. Four newly isolated fuselloviruses from extreme geothermal environments reveal unusual morphologies and a possible interviral recombination mechanism. *Environ. Microbiol.* **2009**, *11*, 2849–2862. [[CrossRef](#)]
31. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 238. [[CrossRef](#)] [[PubMed](#)]
32. Aulitto, M.; Martinez-Alvarez, L.; Fiorentino, G.; Limauro, D.; Peng, X.; Contursi, P. A Comparative Analysis of *Weizmannia coagulans* Genomes Unravels the Genetic Potential for Biotechnological Applications. *Int. J. Mol. Sci.* **2022**, *23*, 3135. [[CrossRef](#)]

33. Fusco, S.; She, Q.; Bartolucci, S.; Contursi, P. Tlys, 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.* **2013**, *87*, 5926–5936. [[CrossRef](#)] [[PubMed](#)]
34. Iverson, E.A.; Goodman, D.A.; Gorchels, M.E.; Stedman, K.M. Genetic analysis of the major capsid protein of the archaeal fusellovirus SSV1: Mutational flexibility and conformational change. *Genes* **2017**, *8*, 373. [[CrossRef](#)] [[PubMed](#)]
35. Zhang, J.; Huang, L. Construction of a Fusellovirus with a Minimal Set of Genes. *ACS Synth. Biol.* **2021**, *10*, 2617–2627. [[CrossRef](#)] [[PubMed](#)]
36. Quemin, E.R.J.; Pietilä, M.K.; Oksanen, H.M.; Forterre, P.; Rijpstra, W.I.C.; Schouten, S.; Bamford, D.H.; Prangishvili, D.; Krupovic, M. Sulfolobus Spindle-Shaped Virus 1 Contains Glycosylated Capsid Proteins, a Cellular Chromatin Protein, and Host-Derived Lipids. *J. Virol.* **2015**, *89*, 11681–11691. [[CrossRef](#)]
37. Wang, F.; Cvirkaite-Krupovic, V.; Vos, M.; Beltran, L.C.; Kreutzberger, M.A.B.; Winter, J.M.; Su, Z.; Liu, J.; Schouten, S.; Krupovic, M.; et al. Spindle-shaped archaeal viruses evolved from rod-shaped ancestors to package a larger genome. *Cell* **2022**, *185*, 1297–1307.e11. [[CrossRef](#)]
38. Han, Z.; Yuan, W.; Xiao, H.; Wang, L.; Zhang, J.; Peng, Y.; Cheng, L.; Liu, H.; Huang, L. Structural insights into a spindle-shaped archaeal virus with a sevenfold symmetrical tail. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2119439119. [[CrossRef](#)]
39. Koonin, E.V. Archaeobacterial virus SSV1 encodes a putative DNAA-like protein. *Nucleic Acids Res.* **1992**, *20*, 1143. [[CrossRef](#)]
40. Arnold, H.P.; She, Q.; Phan, H.; Stedman, K.; Prangishvili, D.; Holz, I.; Kristjansson, J.K.; Garrett, R.; Zillig, W. The genetic element pSSVx of the extremely thermophilic crenarchaeon Sulfolobus is a hybrid between a plasmid and a virus. *Mol. Microbiol.* **1999**, *34*, 217–226. [[CrossRef](#)]
41. Dewerff, S.J.; Zhang, C.; Schneider, J.; Whitaker, R.J. Intraspecific antagonism through viral toxin encoded by chronic Sulfolobus spindle-shaped virus. *Philos. Trans. R. Soc. B Biol. Sci.* **2022**, *377*, 1842. [[CrossRef](#)]
42. Kraft, P.; Oeckinghaus, A.; Kümmel, D.; Gauss, G.H.; Gilmore, J.; Wiedenheft, B.; Young, M.; Lawrence, C.M. Crystal Structure of F-93 from Sulfolobus Spindle-Shaped Virus 1, a Winged-Helix DNA Binding Protein. *J. Virol.* **2004**, *78*, 11544–11550. [[CrossRef](#)] [[PubMed](#)]
43. Menon, S.K.; Maaty, W.S.; Corn, G.J.; Kwok, S.C.; Eilers, B.J.; Kraft, P.; Gillitzer, E.; Young, M.J.; Bothner, B.; Lawrence, C.M. Cysteine usage in Sulfolobus spindle-shaped virus 1 and extension to hyperthermophilic viruses in general. *Virology* **2008**, *376*, 270–278. [[CrossRef](#)] [[PubMed](#)]
44. Lawrence, C.M.; Menon, S.; Eilers, B.J.; Bothner, B.; Khayat, R.; Douglas, T.; Young, M.J. Structural and functional studies of archaeal viruses. *J. Biol. Chem.* **2009**, *284*, 12599–12603. [[CrossRef](#)] [[PubMed](#)]
45. Fusco, S.; Aulitto, M.; Iacobucci, I.; Crocamo, G.; Pucci, P.; Bartolucci, S.; Monti, M.; Contursi, P. The interaction between the F55 virus-encoded transcription regulator and the RadA host recombinase reveals a common strategy in Archaea and Bacteria to sense the UV-induced damage to the host DNA. *Biochim. Biophys. Acta Gene Regul. Mech.* **2020**, *1863*, 194493. [[CrossRef](#)]
46. Yang, Y.; Junfeng, L.; Xiaofei, F.; Zhou, F.; Zhang, S.; Zhang, X.; Huang, Q.; Krupovic, M.; She, Q.; Ni, J.; et al. A novel RHH family transcription factor aCcr1 and its viral homologs dictate cell cycle progression in archaea. *bioRxiv* **2022**. [[CrossRef](#)]
47. Li, X.; Lozano-Madueño, C.; Martínez-Alvarez, L.; Peng, X. A clade of RHH proteins ubiquitous in Sulfolobales and their viruses regulates cell cycle progression. *bioRxiv* **2022**. [[CrossRef](#)]
48. Michelin, M.; Gomes, D.G.; Romani, A.; de Maria, L.T.M.P.; Teixeira, J.A. Nanocellulose production: Exploring the enzymatic route and residues of pulp and paper industry. *Molecules* **2020**, *25*, 3411. [[CrossRef](#)]
49. Bhoobalan-Chitty, Y.; Johansen, T.B.; Di Cianni, N.; Peng, X. Inhibition of Type III CRISPR-Cas Immunity by an Archaeal Virus-Encoded Anti-CRISPR Protein. *Cell* **2019**, *179*, 448–458.e11. [[CrossRef](#)]
50. Emms, D.M.; Kelly, S. Affiliations STAG: Species Tree Inference from All Genes. *bioRxiv* **2018**. [[CrossRef](#)]
51. Emms, D.M.; Kelly, S. STRIDE: Species tree root inference from gene duplication events. *Mol. Biol. Evol.* **2017**, *34*, 3267–3278. [[CrossRef](#)] [[PubMed](#)]
52. Stedman, K.M.; DeYoung, M.; Saha, M.; Sherman, M.B.; Morais, M.C. Structural insights into the architecture of the hyperthermophilic fusellovirus SSV1. *Virology* **2015**, *474*, 105–109. [[CrossRef](#)] [[PubMed](#)]
53. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]
54. Fusco, S.; She, Q.; Fiorentino, G.; Bartolucci, S.; Contursi, P. Unravelling the Role of the F55 Regulator in the Transition from Lysogeny to UV Induction of Sulfolobus Spindle-Shaped Virus 1. *J. Virol.* **2015**, *89*, 6453–6461. [[CrossRef](#)] [[PubMed](#)]
55. Contursi, P.; Fusco, S.; Limauro, D.; Fiorentino, G. Host and viral transcriptional regulators in Sulfolobus: An overview. *Extremophiles* **2013**, *17*, 881–895. [[CrossRef](#)]
56. Makarova, K.S.; Wolf, Y.I.; Koonin, E.V. Towards functional characterization of archaeal genomic dark matter. *Biochem. Soc. Trans.* **2019**, *47*, 389–398. [[CrossRef](#)]
57. Jarrell, K.F.; Ding, Y.; Meyer, B.H.; Albers, S.-V.; Kaminski, L.; Eichler, J. N-Linked Glycosylation in Archaea: A Structural, Functional, and Genetic Analysis. *Microbiol. Mol. Biol. Rev.* **2014**, *78*, 304–341. [[CrossRef](#)]
58. Fulton, J.; Bothner, B.; Lawrence, M.; Johnson, J.E.; Douglas, T.; Young, M. Genetics, biochemistry and structure of the archaeal virus STIV. *Biochem. Soc. Trans.* **2009**, *37*, 114–117. [[CrossRef](#)]

59. Zhan, Z.; Ouyang, S.; Liang, W.; Zhang, Z.; Liu, Z.-J.; Huang, L. Biological Crystallography Structural and functional characterization of the C-terminal catalytic domain of SSV1 integrase. *Res. Pap. Acta Cryst.* **2012**, *68*, 659–670. [[CrossRef](#)]
60. Eberle, W.; Pastore, A.; Sander, C.; Rösch, P. The structure of ColE1 rop in solution. *J. Biomol. NMR* **1991**, *1*, 71–82. [[CrossRef](#)]
61. Clore, A.J.; Stedman, K.M. The SSV1 viral integrase is not essential. *Virology* **2007**, *361*, 103–111. [[CrossRef](#)] [[PubMed](#)]
62. Patel, M.; Jiang, Q.; Woodgate, R.; Cox, M.M.; Goodman, M.F. A New Model for SOS-induced Mutagenesis: How RecA Protein Activates DNA Polymerase V. *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 171–184. [[CrossRef](#)] [[PubMed](#)]
63. Martin, A.; Yeats, S.; Janekovic, D.; Reiter, W.-D.; Aicher, W.; Zillig, W. SAV 1, a temperate u.v.-inducible DNA virus-like particle from the archaeobacterium *Sulfolobus acidocaldarius* isolate B12. *EMBO J.* **1984**, *3*, 2165–2168. [[CrossRef](#)] [[PubMed](#)]
64. Quemina, E.R.J.; Chlanda, P.; Sachse, M.; Forterre, P.; Prangishvili, D.; Krupovic, M. Eukaryotic-like virus budding in archaea. *MBio* **2016**, *7*, e01439-16. [[CrossRef](#)] [[PubMed](#)]
65. Zink, I.A.; Pfeifer, K.; Wimmer, E.; Sleytr, U.B.; Schuster, B.; Schleper, C. CRISPR-mediated gene silencing reveals involvement of the archaeal S-layer in cell division and virus infection. *Nat. Commun.* **2019**, *10*, 4797. [[CrossRef](#)]
66. Rowland, E.F.; Bautista, M.A.; Zhang, C.; Whitaker, R.J. Surface resistance to SSVs and SIRVs in pilin deletions of *Sulfolobus islandicus*. *Mol. Microbiol.* **2020**, *113*, 718–727. [[CrossRef](#)]
67. Pauly, M.D.; Bautista, M.A.; Black, J.A.; Whitaker, R.J. Diversified local CRISPR-Cas immunity to viruses of *Sulfolobus islandicus*. *Philos. Trans. R. Soc. B Biol. Sci.* **2019**, *374*, 20180093. [[CrossRef](#)]
68. Koonin, E.V.; Senkevich, T.G.; Dolja, V.V. The ancient virus world and evolution of cells. *Biol. Direct* **2006**, *1*, 29. [[CrossRef](#)]
69. Contursi, P.; Farina, B.; Pirone, L.; Fusco, S.; Russo, L.; Bartolucci, S.; Fattorusso, R.; Pedone, E. Structural and functional studies of Stf76 from the *Sulfolobus islandicus* plasmid-virus pSSVx: A novel peculiar member of the winged helix-turn-helix transcription factor family. *Nucleic Acids Res.* **2014**, *42*, 5993–6011. [[CrossRef](#)]
70. Contursi, P.; D’Ambrosio, K.; Pirone, L.; Pedone, E.; Aucelli, T.; She, Q.; De Simone, G.; Bartolucci, S. C68 from the *Sulfolobus islandicus* plasmid-virus pSSVx is a novel member of the AbrB-like transcription factor family. *Biochem. J.* **2011**, *435*, 157–166. [[CrossRef](#)]
71. Prato, S.; Vitale, R.M.; Contursi, P.; Lipps, G.; Saviano, M.; Rossi, M.; Bartolucci, S. Molecular modeling and functional characterization of the monomeric primase-polymerase domain from the *Sulfolobus solfataricus* plasmid pIT3. *FEBS J.* **2008**, *275*, 4389–4402. [[CrossRef](#)] [[PubMed](#)]
72. Contursi, P.; Cannio, R.; Prato, S.; She, Q.; Rossi, M.; Bartolucci, S. Transcriptional analysis of the genetic element pSSVx: Differential and temporal regulation of gene expression reveals correlation between transcription and replication. *J. Bacteriol.* **2007**, *189*, 6339–6350. [[CrossRef](#)]
73. Stedman, K.M.; Schleper, C.; Rumpf, E.; Zillig, W. Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*: Construction and testing of viral shuttle vectors. *Genetics* **1999**, *152*, 1397–1405. [[CrossRef](#)]
74. Cannio, R.; Contursi, P.; Rossi, M.; Bartolucci, S. An autonomously replicating transforming vector for *Sulfolobus solfataricus*. *J. Bacteriol.* **1998**, *180*, 3237–3240. [[CrossRef](#)]
75. Contursi, P.; Cannio, R.; Prato, S.; Fiorentino, G.; Rossi, M.; Bartolucci, S. Development of a genetic system for hyperthermophilic Archaea: Expression of a moderate thermophilic bacterial alcohol dehydrogenase gene in *Sulfolobus solfataricus*. *FEMS Microbiol. Lett.* **2003**, *218*, 115–120. [[CrossRef](#)] [[PubMed](#)]
76. Aulitto, M.; Gallo, G.; Puopolo, R.; Mormone, A.; Limauro, D.; Contursi, P.; Piochi, M.; Bartolucci, S.; Fiorentino, G. Genomic Insight of *Alicyclobacillus mali* FL18 Isolated From an Arsenic-Rich Hot Spring. *Front. Microbiol.* **2021**, *12*, 639697. [[CrossRef](#)] [[PubMed](#)]
77. Aulitto, M.; Tom, L.M.; Ceja-Navarro, J.A.; Simmons, B.A.; Singer, S.W. Whole-Genome Sequence of *Brevibacillus borstelensis* SDM, Isolated from a Sorghum-Adapted Microbial Community. *Microbiol. Resour. Announc.* **2020**, *9*, 8–9. [[CrossRef](#)]
78. Aulitto, M.; Strazzulli, A.; Sansone, F.; Cozzolino, F.; Monti, M.; Moracci, M.; Fiorentino, G.; Limauro, D.; Bartolucci, S.; Contursi, P. Prebiotic properties of *Bacillus coagulans* MA-13: Production of galactoside hydrolyzing enzymes and characterization of the transglycosylation properties of a GH42  $\beta$ -galactosidase. *Microb. Cell Fact.* **2021**, *20*, 71. [[CrossRef](#)]
79. Aulitto, M.; Fusco, S.; Bartolucci, S.; Franzén, C.J.; Contursi, P. *Bacillus coagulans* MA-13: A promising thermophilic and cellulolytic strain for the production of lactic acid from lignocellulosic hydrolysate. *Biotechnol. Biofuels* **2017**, *10*, 210. [[CrossRef](#)]
80. Aulitto, M.; Fusco, S.; Nickel, D.B.; Bartolucci, S.; Contursi, P.; Franzén, C.J. Seed culture pre-adaptation of *Bacillus coagulans* MA-13 improves lactic acid production in simultaneous saccharification and fermentation. *Biotechnol. Biofuels* **2019**, *12*, 45. [[CrossRef](#)]
81. Aulitto, M.; Fusco, S.; Franzén, C.J.; Strazzulli, A.; Moracci, M.; Bartolucci, S.; Contursi, P. Draft Genome Sequence of *Bacillus coagulans* MA-13, a Thermophilic Lactic Acid Producer from Lignocellulose. *Microbiol. Resour. Announc.* **2019**, *8*, 341–360. [[CrossRef](#)] [[PubMed](#)]
82. Gallo, G.; Aulitto, M.; Contursi, P.; Limauro, D.; Bartolucci, S.; Fiorentino, G. Bioprospecting of Extremophilic Microorganisms to Address Environmental Pollution. *J. Vis. Exp.* **2021**, *178*, e63453. [[CrossRef](#)] [[PubMed](#)]
83. Gaglione, R.; Pirone, L.; Farina, B.; Fusco, S.; Smaldone, G.; Aulitto, M.; Dell’Olmo, E.; Roschetto, E.; Del Gatto, A.; Fattorusso, R.; et al. Insights into the anticancer properties of the first antimicrobial peptide from Archaea. *Biochim. Biophys. Acta - Gen. Subj.* **2017**, *1861*, 2155–2164. [[CrossRef](#)]
84. Aulitto, M.; Fusco, S.; Limauro, D.; Fiorentino, G.; Bartolucci, S.; Contursi, P. Galactomannan degradation by thermophilic enzymes: A hot topic for biotechnological applications. *World J. Microbiol. Biotechnol.* **2019**, *35*, 32. [[CrossRef](#)] [[PubMed](#)]

85. Aulitto, M.; Fusco, F.A.; Fiorentino, G.; Bartolucci, S.; Contursi, P.; Limauro, D. A thermophilic enzymatic cocktail for galactomannans degradation. *Enzyme Microb. Technol.* **2018**, *111*, 7–11. [[CrossRef](#)] [[PubMed](#)]
86. Ing, N.; Deng, K.; Chen, Y.; Aulitto, M.; Gin, J.W.; Pham, T.L.M.; Petzold, C.J.; Singer, S.W.; Bowen, B.; Sale, K.L.; et al. A multiplexed nanostructure-initiator mass spectrometry (NIMS) assay for simultaneously detecting glycosyl hydrolase and lignin modifying enzyme activities. *Sci. Rep.* **2021**, *11*, 11803. [[CrossRef](#)] [[PubMed](#)]
87. Varanda, C.; Félix, M.D.R.; Campos, M.D.; Materatski, P. An overview of the application of viruses to biotechnology. *Viruses* **2021**, *13*, 13–18. [[CrossRef](#)]
88. Michán, C.; Michán-Doña, A. Viruses: Friends or foes. *Microb. Biotechnol.* **2022**, *15*, 88–90. [[CrossRef](#)]