AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENETIC DIVERSITY AND EVOLUTION

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*

Eric A. Iverson,* David A. Goodman, Madeline E. Gorchels,*
Kenneth M. Stedman

Center for Life in Extreme Environments and Biology Department, Portland State University, Portland, Oregon, USA

ABSTRACT Viruses infecting the *Archaea* harbor a tremendous amount of genetic diversity. This is especially true for the spindle-shaped viruses of the family *Fuselloviridae*, where >90% of the viral genes do not have detectable homologs in public databases. This significantly limits our ability to elucidate the role of viral proteins in the infection cycle. To address this, we have developed genetic techniques to study the well-characterized fusellovirus Sulfolobus spindle-shaped virus 1 (SSV1), which infects *Sulfolobus solfataricus* in volcanic hot springs at 80°C and pH 3. Here, we present a new comparative genome analysis and a thorough genetic analysis of SSV1 using both specific and random mutagenesis and thereby generate mutations in all open reading frames. We demonstrate that almost half of the SSV1 genes are not essential for infectivity, and the requirement for a particular gene correlates well with its degree of conservation within the *Fuselloviridae*. The major capsid gene *vp1* is essential for SSV1 infectivity. However, the universally conserved minor capsid gene *vp3* could be deleted without a loss in infectivity and results in virions with abnormal morphology.

IMPORTANCE Most of the putative genes in the spindle-shaped archaeal hyperthermophile fuselloviruses have no sequences that are clearly similar to characterized genes. In order to determine which of these SSV genes are important for function, we disrupted all of the putative genes in the prototypical fusellovirus, SSV1. Surprisingly, about half of the genes could be disrupted without destroying virus function. Even deletions of one of the known structural protein genes that is present in all known fuselloviruses, *vp3*, allows the production of infectious viruses. However, viruses lacking *vp3* have abnormal shapes, indicating that the *vp3* gene is important for virus structure. Identification of essential genes will allow focused research on minimal SSV genomes and further understanding of the structure of these unique, ubiquitous, and extremely stable archaeal viruses.

KEYWORDS *Archaea*, morphogenesis, mutational studies, virus assembly

The *Archaea*, particularly hyperthermophilic *Archaea* that grow optimally above 80°C, are infected by some of the most structurally and genetically diverse viruses known (reviewed by Prangishvili [1]). Of the numerous morphologically unique viruses, those possessing spindle-shaped architectures are widespread within the *Archaea* and include some of the best-characterized archaeal viruses. Spindle-shaped archaeal viruses belong to two viral families, the *Fuselloviridae* and *Bicaudaviridae*, which have few similarities other than overall capsid morphology (2). Fuselloviruses have been isolated

Received 14 December 2016 **Accepted** 27 January 2017

Accepted manuscript posted online 1 February 2017

Citation Iverson EA, Goodman DA, Gorchels ME, Stedman KM. 2017. 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 91:e02406-16. https://doi.org/10.1128/JVI.02406-16.

Editor Anne E. Simon, University of Maryland

Address correspondence to Kenneth M. Stedman, kstedman@pdx.edu.

* Present address: Eric A. Iverson, Catholic Memorial School, West Roxbury, Massachusetts, USA; Madeline E. Gorchels, Woods Hole Marine Biological Laboratory, Woods Hole, Massachusetts, USA.

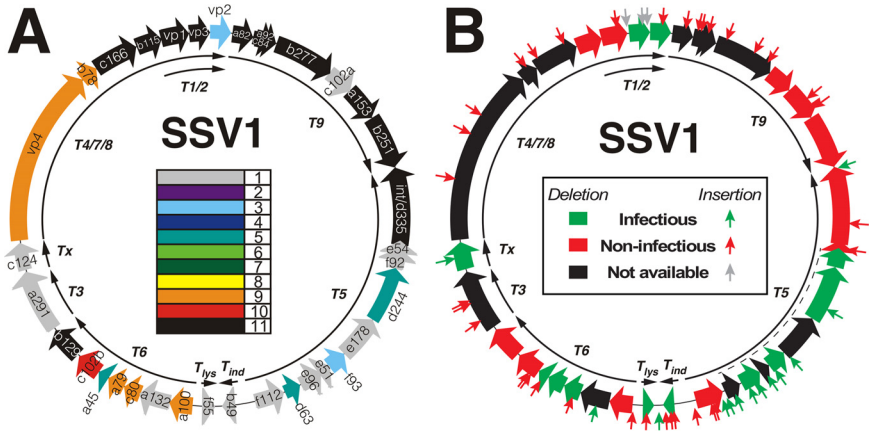

FIG 1 Annotated SSV1 genome and mutation summary. (A) Conserved ORFs in SSV1. Open reading frames (ORFs) are displayed as block arrows and labeled as described in Palm et al. (7). The conservation of each SSV1 ORF relative to 10 other fusellovirus genomes is indicated by the color code in the middle of the map. Color coding is from black, present in all 11 genomes, to gray, present only in the SSV1 genome. Conserved ORFs were identified by BLAST-P (34) using an E value of <0.001 to identify putative homologs. Core SSV1 ORFs, black, are labeled in white text, and all others are in black text. Mapped viral transcripts (9, 10, 25) are shown as thin black arrows on the interior of the map and are labeled. (B) SSV1 mutants. The ORF fill color indicates if deletion of the ORF in the genome allows production of infectious virus (green) or not (red) when transformed into *S. solfataricus* strain S441. ORFs for which deletions are not available are filled in black. The dotted line near the T5 transcript indicates the region deleted in the infectious EAI582 mutant. Each of the small arrows perpendicular to the map denotes the location of an EZ-Tn*5* transposon insertion. The color of the arrow indicates whether the virus genome containing that insertion results in the production of infectious virus (green) or not (red) when transformed into *S. solfataricus* strain S441. Gray arrows in the *vp1*/*vp3* gene region indicate insertions that apparently are removed via recombination, allowing the production of infectious virus (see the text). Transcripts are as described for panel A. All noninfectious mutants labeled in red have been tested independently at least 5 times.

worldwide from volcanic hot spring environments (70 to 80°C, pH ~2 to 4) in which their hosts, *Sulfolobus solfataricus* and its close relatives, thrive (1, 3–6).

Sulfolobus spindle-shaped virus 1 (SSV1) and its host, *S. shibatae*, originally were isolated from a hot spring in Beppu, Japan (3). SSV1 encapsidates a circular, positively supercoiled 15.5-kbp double-stranded DNA (dsDNA) genome (7, 8). Transcription of the SSV1 genome following UV irradiation, which also leads to virion production, proceeds via early and late transcripts (9, 10) (Fig. 1). The transcripts in the SSV1 genome carry 35 open reading frames (ORFs), most of which cannot be assigned functions due to undetectable homology with sequences in public databases (11). This situation is not unique among viruses; however, it seems especially pronounced among crenarchaeal viruses and hinders our understanding of these viruses and their life cycles (12–14). All fuselloviruses encode a tyrosine family recombinase or integrase that facilitates virus genome integration into a host tRNA gene. However, the viral integrase gene does not appear to be required for virus infectivity (15). To date, only the SSV1 integrase and structural proteins (VP1, VP2, VP3, and VP4) have been assigned functions, although a combination of structural studies and bioinformatics have provided predictions for the roles of several others (16).

Fusellovirus capsids are comprised of the major capsid protein VP1, host-derived lipids, and smaller amounts of the minor capsid proteins VP3 and VP4 (17, 18). The VP1 and VP3 proteins are highly conserved within the *Fuselloviridae* and are clearly homologous to each other (11, 17). SSV1 and three other fuselloviruses, SSV6, Acidianus spindle-shaped virus 1 (ASV1), and Sulfolobus Mexican fusellovirus 1 (SMF1), also encode a small and extremely basic structural protein (VP2) that is thought to bind to viral DNA within the capsid (19). VP2 has been found in SSV1 virions (17, 18). A fourth capsid protein, VP4 (formerly SSV1-C792), has been identified via mass spectrometry of purified virions and likely forms the tail fiber of the virion (18, 20). Recently it was shown that purified SSV1 virions contain host-derived glycerol dibiphytanyl glycerol tetraether

(GDGT) lipids (18). The structure of the SSV1 virion has been solved via cryoelectron microscopy, providing insight into the architecture and assembly of this diverse family of viruses (21).

A number of putative transcription factors have been identified in the SSV1 genome, none of which have been experimentally characterized. Atomic resolution structures of SSV1-B129, SSV1-F112, SSV1-F93, and SSV8-E73 (a homologue of SSV1-E51) reveal DNA binding domains characteristic of transcriptional regulators (16, 22–24). The products of four other well-conserved SSV1 ORFs (SSV1-A45, SSV1-A79, SSV1-C80, and SSV1-B115) are predicted to encode ribbon-helix-helix (RHH), helix-turn-helix (HTH), or zinc finger (ZNF) DNA binding domains and may also be involved in transcription regulation (12). The protein product of ORF *f55* is predicted to possess an RHH DNA binding domain and was shown to bind weakly to a number of viral promoters *in vitro* that presumably control the expression of early gene products. Based on these data, SSV1-F55 is hypothesized to maintain lysogeny by repressing early viral promoters (25, 26).

SSV1-B251 and SSV1-A153 are highly conserved and are the only fusellovirus proteins encoded by the satellite nucleic acid pSSVx, whose genome can be packaged in smaller virus capsids upon infection by a bona fide helper SSV. This result led to the prediction that the SSV1-B251 and SSV1-A153 ORFs are involved in replication or packaging (27). SSV1-B251 possesses NTP binding motifs and is predicted to be homologous to the bacterial *dnaA* gene (28). The structure of SSV8-D212, a homologue of SSV1-D244, has a predicted nuclease fold, although activity was not demonstrated biochemically (20). SSV1-D244 was identified in one study by mass spectrometry of purified SSV1 virions, but it was not detected in a later analysis (18, 20). The structure of SSV1-D63 displays a four-helix bundle that is characteristic of a large number of proteins, making a functional prediction difficult (29).

SSV1 is one of only two archaeal viruses amenable to genetic study (30, 31). The ability of SSV1 to tolerate large insertions of foreign DNA allowed the construction of SSV1-based shuttle vectors and provided tools for studying the viral genome itself (30, 32). Partial digestion of SSV1 DNA followed by the insertion of an *Escherichia coli* plasmid provided the first data on which genes are required for virus infectivity (30). These data indicated that a majority of ORFs in SSV1 are essential for virus function.

SSV1 shuttle vectors were essential for subsequent work in which specific ORFs were deleted from the viral genome. Using long inverse PCR (LIPCR), the SSV1 *integrase* gene was deleted and subsequently shown not to be essential for production of infectious virus in *S. solfataricus* P2 (15). More recently, in a structure-guided study, the *vp2* gene and ORFs *b129* and *d244* were deleted (11). Deletion of *b129*, a putative transcriptional regulator, resulted in loss of infectivity, whereas removal of the putative DNA binding protein gene *vp2* and predicted nuclease gene *d244* resulted in production of infectious virus. Homologues of SSV1-B129 have been identified in the genomes of all 11 isolated fuselloviruses, indicating that it is important for virus infectivity (Fig. 1A). SSV1-VP2 is much less well conserved, which may explain why it can be deleted without a loss of infectivity (Fig. 1). However, host-encoded chromatin proteins that have been found in purified virions may be able to functionally complement this lack of the *vp2* gene (18). Interestingly, cells that were infected with SSV1-Δ*d244* exhibited a retarded growth phenotype compared to cells infected with wild-type SSV1 or SSV1-Δ*vp2* (11).

To further characterize the genetic requirements for SSV1 function, LIPCR and transposon mutagenesis were employed to construct 78 mutant SSV1 genomes harboring mutations in each of the 35 ORFs. SSV1 appears to be much more tolerant of mutagenesis than previously thought, as half of the genes could be mutated without abrogating infectivity. Almost the entirety of the T5 early transcript appears to be expendable, while the T6 early transcript is much less so. This correlates with the abundance of well-conserved fusellovirus genes in the T6 transcript. The genes of the newly defined fusellovirus core appear to be essential with the surprising exception of the minor capsid gene *vp3*, which was shown not to be essential for SSV1 infectivity.

(Parts of this research, namely, the mutagenesis results and modified versions of Fig. 1 to 4A, were previously published as part of Eric A. Iverson's Ph.D. dissertation [33].)

TABLE 1 Fusellovirus genomes used in this study

Virus	Genome size (bp)	No. of ORFs	GenBank accession no.	Reference or source
SSV1	15,465	35	NC_001338	7
SSV2	14,795	35	NC_005265	4
SSV3	15,230	32	KY579375	35
SSV4	15,135	34	NC_009986	36
SSV5	15,330	34	NC_011217	5
SSV6	15,684	33	NC_013587	5
SSV7	17,602	33	NC_013588	5
SSV8[a]	16,473	37	NC_005360	6
SSV9[b]	17,385	31	NC_005361	6
SSVL	16,271	36	KY563228	Personal communication
ASV1	24,186	38	NC_013585	5

[a]SSV8 formerly referred to as Sulfolobus spindle-shaped virus Ragged Hills.
[b]SSV9 formerly referred to as Sulfolobus spindle-shaped virus Kamchatka.

RESULTS

Identification of fusellovirus core genes. In order to understand gene conservation in SSV1 and provide context for our mutagenesis results, all SSV1 ORFs were compared to all other full-length circular SSV genomes from virion-containing cultures using BLAST-P (34). The SSVs used were SSV2, SSV3, SSV4, SSV5, SSV6, SSV7, SSV8, SSV9, ASV, and SSVL (citations are listed in Table 1). Putative homologs were identified as having a BLAST-P E score of $< 10^{-3}$ using the SSV1 ORF as a query. This analysis identified a set of 12 genes/ORFs, the fusellovirus core, that are conserved in all *Fuselloviridae* (black arrows with white lettering in Fig. 1A) (Table 2).

Summary of SSV1 mutagenesis. To better understand the function of SSV1 genetic elements, both specific and random mutagenesis of SSV1 were performed. All 35 ORFs in the SSV1 genome were mutated via insertion and/or deletion (Table 3), and all 78 of these mutants were tested for their ability to infect *S. solfataricus* S441, a host permissive and susceptible to infection by all SSVs tested to date (38). Insertion mutants in intergenic regions were also isolated and characterized. Infectivity of SSV1 mutants was assayed by transformation of purified mutant DNA into uninfected *Sulfolobus* cells and then spotting these cultures onto an uninfected lawn. Spotting a productively infected culture on a host lawn results in the appearance of a ring of growth inhibition, or halo, around the spot after 48 to 72 h of incubation at 75°C (11, 38). Transformation of many SSV1 mutants into uninfected *Sulfolobus* did not result in cultures that inhibited growth of uninfected *Sulfolobus* on lawns, indicating that the mutation generated a defective virus genome (red arrows and ORFs in Fig. 1B). To lower the frequency of false negatives, a mutant was only characterized as noninfectious after a minimum of 5 independent negative halo assay results. The most frequently used positive control, mutant REC262, produced a halo in 97% of transformations.

Mutants in the mostly conserved ORFs in the T6 transcript. The ~2-kbp T6 transcript (9) encodes seven SSV1 ORFs (*a100*, *a132*, *c80*, *a79*, *a45*, *c102b*, and *b129*),

TABLE 2 Fusellovirus core genes

SSV1 gene	Annotation	Reference
vp1	Major capsid protein	17
vp3	Minor capsid protein	17
integrase	Viral integrase	37
a153	None (conserved in SSVx)	
a82	None	
a92	None	
b115	Putative HTH transcription regulator	12
b129	Putative C_2H_2 ZNF transcription regulator	16
b251	ATPase, DnaA homologue, lon-like protease (conserved in pSSVx)	28
b277	None	
c84	None	
c166	None	

TABLE 3 SSV1 mutants

Plasmid	Description	Infectious in s441 (+ or −)	Reference or source
pAJC96	pAJC97 background with integrase (ORF *d335*) deleted (also called *SSV1-Δint*)	−	15
pAJC97	SSV1 shuttle vector (TOPO PCR blunt II inserted at bp 3173) (ORF *e178*)	+	15
REC228[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 4680 (ORF *f112*)	+	This work
REC229[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 11265 (*vp4*)	−	This work
REC230[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 8277 (ORF *a291*)	−	This work
REC231[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 11227 (*vp4*)	−	This work
EAI232	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 8662 (ORF *a291*)	−	This work
EAI239	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5218 (ORF *b49*)	−	This work
EAI240	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 4792 (ORF *f112*)	−	This work
EAI241	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 100 (ORF *a100*)	−	This work
EAI242	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 10159 (*vp4*)	−	This work
REC243[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 14046 (ORF *c84*/*a82*)	−	This work
REC244[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 4788 (ORF *f112*)	−	This work
REC245[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 12718 (*vp1*)	−	This work
EAI247	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 13992 (ORF *c84*/*a82*)	−	This work
EAI248	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 953	−	This work
EAI249	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 6375 (ORF *a132*)	+	This work
EAI250	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 11641 (ORF *b78*)	−	This work
EAI251	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 14677 (ORF *b277*)	−	This work
EAI253	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 3889 (ORF *e51*)	+	This work
EAI254	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 8633 (ORF *a291*)	−	This work
EAI255	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 7509 (ORF *b129*)	−	This work
EAI256	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 2776 (ORF *d*244)	+	This work
EAI257	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 4249 (ORF *e96*)	+	This work
EAI258	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 8998 (ORF *c124*)	+	This work
EAI260	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 80 (ORF *a153*)	−	This work
EAI261	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 14209 (ORF *b277*)	−	This work
REC262[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 4209 (ORF *e96*)	+	This work
EAI266	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 1988 (ORF *e54*)	−	This work
EAI267	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 1717 (*integrase*)	−	This work
EAI271	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 6018 (ORF *a100*)	−	This work
EAI278	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 14834 (ORF *b277*)	−	This work
EAI281	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 13709 (ORF *c84*/*a82*)	−	This work
EAI282	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 11807 (ORF *c166*)	−	This work
EAI283	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 3572 (ORF *e178*)	+	This work
EAI286	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 11407 (*vp4*)	−	This work
EAI296	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5783	−	This work
EAI297	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 12170 (ORF *c166*)	−	This work
EAI305	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 7359 (ORF *c102b*)	−	This work
EAI319	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 7387 (ORF *c102b*)	−	This work
REC322[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5573 (ORF *f55*)	−	This work
REC324[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 4394 (ORF *d63*)	+	This work
REC325[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 967 (*integrase*)	+	This work
EAI446	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 13211 (*vp3*)	+[b]	This work
EAI452	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 13003	+[b]	This work
EAI453	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5451	+	This work
EAI469	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 13491 (*vp2*)	−	This work
EAI476	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 13191 (*vp3*)	+[b]	This work
EAI477	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5247 (*b49*)	−	This work
EAI486	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5024	+	This work
EAI492	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 3837 (*f92*)	+	This work
EAI202	pAJC97 background with ORF *b129* deleted	−	This work
EAI201	pAJC97 background with ORF *d244* deleted	−	This work
EAI205	pAJC97 background with ORF *b49* deleted	+	This work
EAI206	pAJC97 background with ORF *b251* deleted	−	This work
EAI214	pAJC97 background with ORF *b115* deleted	−	This work
EAI216	pAJC97 background with ORF *e96* deleted	+	This work
EAI233	pAJC97 background with ORF *a100* deleted	−	This work
EAI327	pAJC97 background with ORF *f112* deleted	−	This work
EAI390	pAJC97 background with ORF *c124* deleted	+	This work
EAI394	pAJC97 background with ORF *a79* deleted	+	This work
EAI398	pAJC97 background with ORF *a45* deleted	+	This work
EAI400	pAJC97 background with ORF *f55* deleted	+	This work
EAI407	pAJC97 background with ORF *f92* deleted	+	This work

(Continued on following page)

TABLE 3 (Continued)

Plasmid	Description	Infectious in s441 (+ or −)	Reference or source
EAI413	pAJC97 background with ORF *f93* deleted	+	This work
EAI420	pAJC97 background with *vp3* deleted	+	This work
EAI421	pAJC97 background with ORF *c102a* deleted	−	This work
EAI422	pAJC97 background with C terminus of ORF *b129* deleted	−	This work
EAI430	pAJC97 background with N terminus of *b129* deleted	−	This work
EAI435	pAJC97 background with ORF *e54* deleted	−	This work
EAI439	pAJC97 background with ORF *c80* deleted	+	This work
EAI496	pAJC97 background with ORF *c102b* deleted	−	This work
EAI499	pAJC97 background with ORF *b129* deleted	−	This work
JAH572[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5016	−	This work
JAH573[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5264 (*b49*)	−	This work
JAH576[a]	SSV1::Tn*5* mutant, EZ-Tn*5* inserted at bp 5681	−	This work
EAI580	EAI283 background with ORF *d335* deleted	−	This work
EAI582	EAI228 background with ORFs *d63* through *f92* deleted	+	This work

[a]Plasmids with the REC prefix refer to SSV1 mutants that were isolated in the Recombinant DNA Techniques Laboratory course at Portland State University. Plasmids with the JAH prefix refer to SSV1 mutants isolated by Jordan Hartunians.
[b]Tn*5* insertions in these mutants apparently were removed from *Sulfolobus* via homologous recombination.

which, with the exception of ORF *a132*, are well conserved within the *Fuselloviridae* (Fig. 1A). ORFs *c80*, *a79*, and *a45* could be deleted, and ORF *a132* tolerated transposon insertion without a loss of infectivity in any case. In contrast, the highly conserved ORFs *c102b* and *a100* appeared to be essential for infectivity. Infectious virus was not produced when the transposon was inserted into ORFs *a45* and *c80*, probably due to polar effects on essential ORFs *c102b* and *b129*. ORF *b129* was previously determined to be essential for SSV1 infectivity (11). The structure of the SSV1-B129 protein has both N-terminal and C-terminal DNA binding domains, each of which has been shown to bind DNA (16 and personal communication). Mutants lacking ORF *b129* amino acids (aa) 2 to 74 or 75 to 129 appear not to be capable of producing infectious virus, implying that the full-length SSV1-B129 protein is required for infectivity.

Mutants in the monocistronic transcripts T3 and Tx. The monocistronic transcripts T3 and Tx are both expressed early after UV irradiation (10). The T3 transcript encodes SSV1 ORF *a291*, and Tx encodes SSV1 ORF *c124* (Fig. 1A). None of three separate insertion mutants in ORF *a291* resulted in the production of infectious virus (Fig. 1B). Conversely, ORF *c124* tolerated both insertion in and deletion of the entire ORF without loss of infectivity. Interestingly, the ORF *c124* insertion mutant produced infectious virus in only 3 of 12 independent transformations, which was not observed for other insertion mutants in ORFs or for the ORF *c124* deletion mutant.

Fusellovirus core ORFs are intolerant of mutagenesis. The completely conserved fusellovirus core gene ORFs *c166*, *b115*, *a82*, *c84*, *a92*, *b277*, *a153*, and *b251* did not tolerate deletions and/or insertions (Fig. 1). The *vp4* gene and ORF *b78*, both proposed to encode the SSV1 tail filament (5, 18), likewise appeared to be essential for infectivity. The only two ORFs in this region that are not well conserved in fuselloviruses are *c102a* and the structural gene *vp2* (Fig. 1A). ORF *c102a* apparently is essential for infectivity. In contrast, the *vp2* gene was previously shown to be nonessential (11). Unexpectedly, SSV1 lost infectivity when a transposon was inserted into the *vp2* gene (Fig. 1B).

The T5 transcript region and ORFs therein are almost entirely dispensable. The T5 transcript encodes a number of mostly nonconserved ORFs (Fig. 1A). This part of the SSV1 genome is extremely tolerant of mutation (Fig. 1B). Ten ORFs occupy this region, seven of which were shown to be nonessential (Fig. 1B). ORF *f112* appeared to be essential, as it could not be deleted without abrogating infectivity. However, ORF *f112* tolerated transposon insertion at amino acid 111 of the ORF but did not tolerate either of two insertions in the middle of the protein-coding sequence (Fig. 1B).

As the large stretch of nonessential ORFs occupying the T5 transcript indicated that the majority of this region was not required for SSV1 infectivity, we deleted a 2.4-kb region encompassing ORFs *f92*, *d244*, *e178*, *f93*, *e51*, *e96*, and *d63* (Fig. 1B, dotted line). Somewhat surprisingly, SSV1 harboring this deletion remained infectious.

TABLE 4 *Sulfolobus* and *E. coli* strains used

Strain	Description/genotype	Reference or source
S. solfataricus P1	*S. solfataricus* isolate with complete genome sequence	DSM 1616 (39, 40)
S. solfataricus P2	*S. solfataricus* isolate with complete genome sequence	DSM 1617 (41, 39)
S. solfataricus S441	*S. solfataricus* isolate, SSV1 host	38
S. solfataricus GΘ	*S. solfataricus* MT4 derivative	42
E. coli EC100D pir$^+$	F$^-$ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80d*lacZ*ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ(*ara, leu*)*7697 galU galK* λ$^-$ *rpsL* (*Str^R*) *nupG pir*$^+$(*DHFR*)	Epicentre, Inc.
E. coli EC100D pir-116	F$^-$ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80d*lacZ*ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ(*ara, leu*)*7697 galU galK* λ$^-$ *rpsL* (*Str^R*) *nupG pir-116*(*DHFR*)	Epicentre, Inc.
E. coli NovaBlue	*endA1 hsdR17* (rK12$^-$ mK12$^+$) *supE44 thi-1 recA1 gyrA96 relA1 lac* [F′ *proA*$^+$*B*$^+$ *lacI^q*ZΔM15::Tn*10*] (TetR)	Millipore, Inc.

Integrase and e54 deletion mutants have a variable host range. The *integrase* gene and ORF *e54* are adjacent at the distal end of the T5 transcript (Fig. 1A). An SSV1 mutant lacking the integrase gene, pAJC96, also called *SSV1-Δint* (Table 3), was previously shown to be capable of infecting *S. solfataricus* strain P2 (15). We were unable to successfully infect *Sulfolobus* sp. strain S441 with *SSV1-Δint* but were able to productively infect *Sulfolobus* sp. strain GΘ (Table 4) with *SSV1-Δint*. Similarly, SSV1 lacking ORF *e54* was unable to infect *Sulfolobus* sp. strain S441 but was able to infect *Sulfolobus* sp. strain GΘ. To confirm this result, a new *integrase* deletion was constructed in the genetic background of a Tn*5* insertion in ORF *e178*, EAI283, also called *SSV1-Δint2*. This *SSV1-Δint2* mutant has the same host range as the original *SSV1-Δint* virus. A mutant with a transposon inserted in ORF *e54* did not productively infect *Sulfolobus* sp. strain S441 or GΘ. An insertion in the integrase gene likewise inhibited SSV1 infectivity, although a mutant with a Tn*5* insertion near the C terminus of the integrase was capable of producing virus in *Sulfolobus* sp. strain S441 (Fig. 1B).

ORFs *b49* and *f55* and the predicted origin of replication. The 1.3-kbp region between the T5 and T6 transcripts harbors two nonconserved ORFs (*b49* and *f55*), the putative origin of replication (42), and several promoters (Fig. 1). ORF *b49* is located on the UV-inducible transcript T_{ind}, while ORF *f55* is located on the recently discovered transcript T_{lys} (9, 10, 25). Both of these ORFs could be deleted from SSV1 without a loss of infectivity, although insertion was not tolerated in either ORF (Fig. 1B). Three independent ORF *b49* insertion mutants failed to yield infectious virus. Likewise, an ORF *f55* insertion mutant was unable to produce infectious virus. Transposon insertions in the intergenic region surrounding these ORFs were also analyzed (Fig. 1B and Table 3). Insertion of the transposon between ORFs *b49* and *f55* did not inhibit virus infectivity, whereas insertions upstream of ORF *f112* and upstream of ORF *a100* eliminated infectivity.

The major capsid protein gene, *vp1*, is essential, but the highly conserved minor capsid gene *vp3* is not. The SSV1 structural genes *vp1* and *vp3* are highly conserved (Fig. 1A), similar to each other (Fig. 2), and hypothesized to be essential for SSV1 infectivity. As expected, an SSV1 *vp1* deletion and an insertion in the middle of the ORF both failed to yield infectious virus (Fig. 1B). In contrast, SSV1 remained infectious following deletion of the universally conserved SSV1 *vp3* gene (Fig. 1B). The SSV1-Δ*vp3*

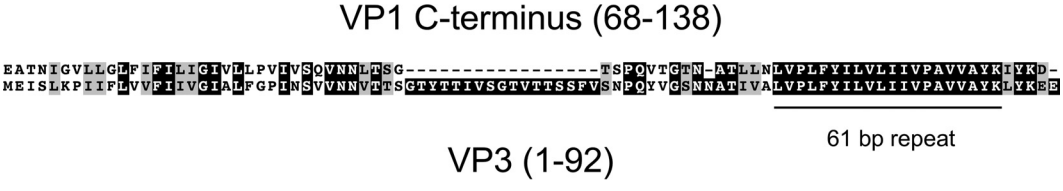

FIG 2 VP1 and VP3 sequence comparison. Pairwise alignment of the C terminus of SSV1-VP1 (aa 68 to 138) and SSV1-VP3 amino acid sequences. Identical amino acids are highlighted in black, chemically similar amino acids are highlighted in gray, and nonconserved amino acids are not highlighted. Gaps in the alignment are indicated by dashes. The alignment was performed with CLUSTAL-W (59). The 61-bp repeated sequence is indicated with a line under the amino acid sequences.

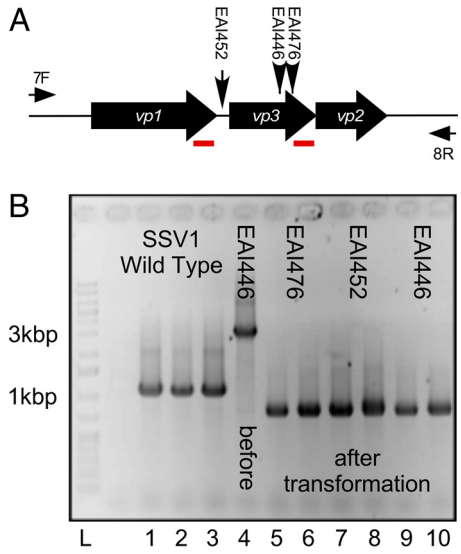

FIG 3 Insertion mutants in *vp3*. (A) Overview of *vp3* insertion mutants and their analyses. SSV1 structural genes *vp1*, *vp3*, and *vp2* are indicated as block arrows and labeled in white. The locations of EZ-Tn*5* insertion mutants EAI446, EAI452, and EAI476 in the SSV1 *vp1*/*vp3* structural gene region are indicated with vertical arrows. PCR primer (univ_7F and univ_8R) annealing sites are indicated by thin horizontal arrows and labeled 7F and 8R, respectively. Red underlined regions indicate 61-bp direct repeats in the *vp1* and *vp3* genes. (B) Analysis of *vp3* spontaneous deletions. PCR with primers univ_7F and univ_8R was performed on DNA purified from transformed *Sulfolobus* sp. strain S441 (except lane 4, where DNA was purified from transformed *E. coli*). Templates: lanes 1 to 3, wild-type SSV1; lane 4, EAI446 DNA purified from *E. coli* used to transform S441; lanes 5 to 10, DNA from *Sulfolobus* species transformed with EAI476 (lanes 5 and 6), EAI452 (lanes 7 and 8), and EAI446 (lanes 9 and 10). Lane L, GeneRuler 1-kb plus DNA ladder (Fisher). Relevant molecular masses are indicated beside the gel.

mutant was also shown to be infectious in *Sulfolobus solfataricus* strains GΘ, P1, and P2 (Table 4), suggesting this is not a strain-specific phenomenon.

Deletion of the *vp3* gene occurs when transposon insertions are present. Two plasmids with transposon insertions in the SSV1 *vp3* gene and one with an insertion in the small intergenic region between SSV1 *vp1* and SSV1 *vp3* (Fig. 3A) generated infectious virus when transformed into *Sulfolobus* sp. strain S441. Following transformation, viral DNA was purified from *Sulfolobus* and screened via PCR using primers that amplified the region encompassing the SSV1 *vp1*, *vp3*, and *vp2* structural genes (Fig. 3A). Oddly, PCR products generated from DNA isolated from *Sulfolobus* transformed with the insertion mutants were shorter than those from wild-type SSV1, indicating that a deletion had occurred within this region instead of an insertion (Fig. 3B). DNA sequencing of each PCR product showed that almost the entire *vp3* gene was missing. There is an identical 61-bp sequence (7) in the C termini of both the *vp1* and *vp3* genes (Fig. 2 and 3). Apparently recombination between the 61-bp direct repeats occurred in all of these insertion mutants (Fig. 3). This also resulted in a deletion of the final 15 bp from the *vp1* gene, including the stop codon. The new *vp1* stop codon is supplied by the native *vp3* stop codon (Fig. 2 and 3).

SSV1 virions lacking *vp3* are abnormal. Virions from mutants lacking the *vp3* minor capsid protein gene (SSV1-Δ*vp3*), either constructed by LIPCR or generated by recombination, were examined by transmission electron microscopy. While the overall shape of SSV1-Δ*vp3* virions is that of a spindle, they appeared to be longer and thinner than the wild type (Fig. 4). Length, width, and aspect ratio were measured for 240 wild-type virions stained with uranyl acetate or phosphotungstate. Only 5.4% of wild-type virions were more than 2 standard deviations from the mean of length, width, or aspect ratio. Similarly, 246 SSV1-Δ*vp3* virions were measured and 99% were more than 2 standard deviations from the means of the wild-type measurements. Moreover, there was much more variability in the shapes of the SSV1-Δ*vp3* virions than the wild type (Fig. 5).

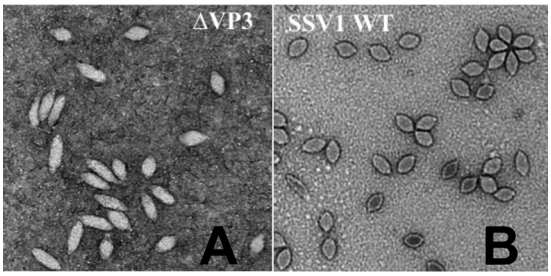

FIG 4 SSV1 virion electron micrographs. (A) Typical transmission electron micrographs of SSV1-Δ*vp3* (EAI420) virions. (B) Typical transmission electron micrographs of SSV1 wild-type virions. SSV1 wild-type particles are ~76 by 40 nm. Samples were negatively stained with uranyl acetate and imaged on a Tecnai F-20 TEM (FEI Inc.) at a 200-keV accelerating voltage. Images were obtained with a Gatan Ultrascan camera.

DISCUSSION

The SSV1 genome is extremely malleable. Overall, SSV1 appears to be extremely tolerant of mutagenesis, allowing a variety of insertions and deletions throughout the genome without loss of function (Table 3; Fig. 1B). The extent to which a particular ORF was conserved within the *Fuselloviridae* generally dictated whether or not the ORF was essential for SSV1 infectivity. Of the 12 ORFs that are unique to SSV1, 8 could be mutated without a loss of infectivity. Conversely, none of the 12 universally conserved genes in the fusellovirus core could be mutated and yet generate infectious virus, with the notable exception of the *vp3* gene (Fig. 1). However, two mutants that were initially judged to be noninfectious after 5 negative results, EAI258, a Tn*5* insertion in SSV1 ORF *c124*, and EAI453, a Tn*5* insertion between ORFs *b49* and *f55*, were found later to be capable of producing infectious virus (Table 3). Thus, although positive results are definitive, negative results must be considered not to be completely conclusive. Verification of negative results using complementation is ongoing. With a few exceptions, ORFs that could be deleted from the virus genome without loss of infectivity also tolerated insertion of the 2-kb EZ-Tn*5* transposon. Many dispensable ORFs encode short, <100-aa, putative proteins. It is not clear whether these ORFs encode proteins or possibly noncoding RNAs. Only products of *vp1*, *vp2*, *vp3*, *vp4*, and ORF *d244* have been reported in purified virions, and no proteomic studies of infected cells have been published (18, 22).

There are fewer essential genes in SSV1 than expected. We have shown that 16 of the 35 SSV1 ORFs can be disrupted without a loss of infectivity (Fig. 1B). This is a significantly higher number of nonessential genes than previously thought. Insertion of the bacterial plasmid pBluescript into the SSV1 genome following partial endonuclease restriction demonstrated that insertions of pBluescript into SSV1 ORFs *e178* and *e51* were tolerated, whereas insertions into ORFs *e96*, *b129*, *vp4*, and *d335* (*integrase*) all failed to produce infectious virus (30). These data are in agreement with this work, with the exception of *e96*, which was shown here to be nonessential by both deletion and insertional mutagenesis (Fig. 1B). Polar effects caused by the insertion of pBluescript could explain this discrepancy, although this seems unlikely, as we have not only isolated a number of infectious insertion mutants in this region but also have shown that this entire quadrant of the SSV1 genome can be deleted without loss of infectivity (Fig. 1B). Alternatively, previous results (30) could have been false negatives.

ORFs encoded by the T5 transcript are almost entirely dispensable. The T5 transcript is expressed early in the SSV1 transcription cycle after UV irradiation (10) and encodes some of the least conserved fusellovirus ORFs (Fig. 1A). Only three of the 10 ORFs in T5 appear to be essential for infectivity, and two of these (*integrase* and *e54*, see below) appear to be essential only in specific hosts (Fig. 1B). Moreover, a deletion mutant lacking ORFs *f92*, *d244*, *e178*, *f93*, *e51*, *e96*, and *d63* (~2.4 kbp) is infectious. This result suggests that SSV1 devotes at least 15% of its genome to ORFs that are seemingly superfluous.

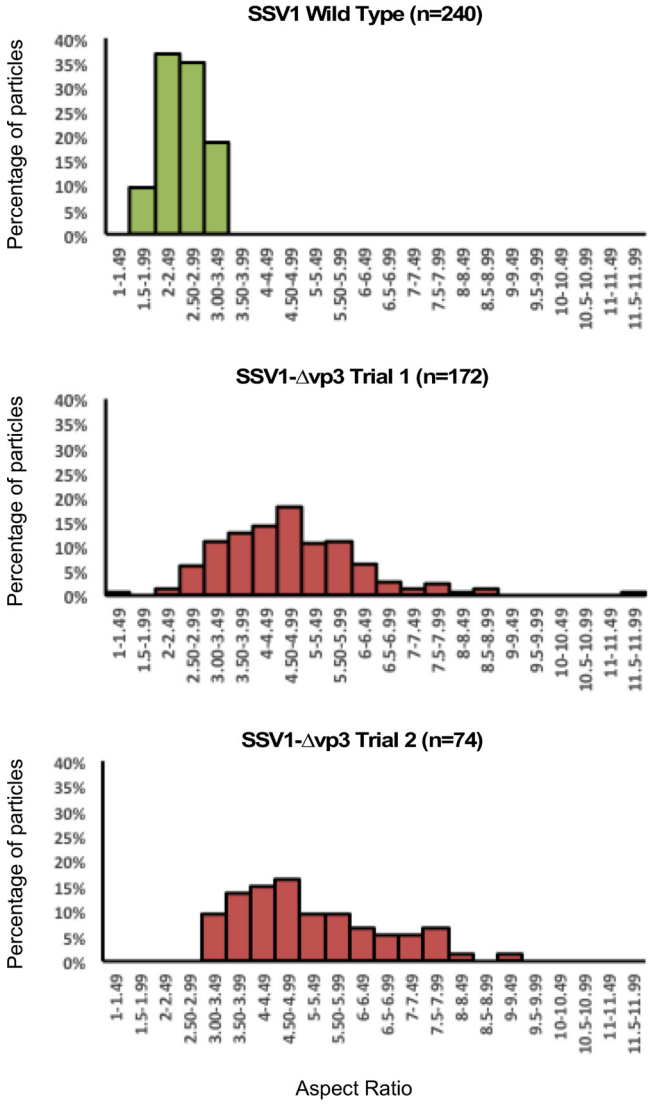

FIG 5 Aspect ratios of SSV1 and SSV1-Δ*vp3* virions. Aspect ratios, defined as the ratio of the long to short axis, of SSV1 wild-type virions ($n = 240$) and two independent preparations of SSV1-Δ*vp3* virions ($n = 172$ and $n = 74$). Aspect ratios are calculated from measurements of negative-stained TEM images using Image J. Histograms are plotted in bins of 0.5.

Mycobacteriophage genomes contain a set of well-conserved structural/assembly genes, but much of their genomes are composed of small ORFs (~500 bp) of unknown function and whose presence varies considerably from isolate to isolate (43). Roughly 2/3 of the nonstructural/assembly genes were not essential for mycobacteriophage function (44), very reminiscent of SSV1 (Fig. 1B). One hypothesis is that these genes were required for growth in an ancestral host or environment but are superfluous under current conditions. Alternatively, some viruses encode genes to compete with cooccurring viruses and/or for protection against host defense systems (45). It is unknown if any fusellovirus genes are involved in any of these putative roles, although many fuselloviruses (not including SSV1) encode a putative Cas4 homologue (46).

The SSV1 integrase gene may be essential in some hosts. The *integrase* gene was the first gene deleted from SSV1 (15). The virus lacking the integrase gene was infectious in *S. solfataricus* strain P2, although the mutant (SSV1-Δ*int*) was quickly outcompeted by wild-type virus (15). However, SSV1-Δ*int* could not infect the *Sulfolobus solfataricus* strain used in this study (S441). The SSV1-Δ*int* virus was able to infect *Sulfolobus solfataricus* strain GΘ, indicating that the SSV1-Δ*int* mutant has a variable

host range. *Sulfolobus* genomes, including that of strain P2, are known to carry a number of integrase-like genes, one of which may be active on a virus lacking an integrase gene (41, 47). Thus, SSV1, contrary to Clore and Stedman (15), does appear to require an *integrase* gene for infectivity in some host strains. The reason for this requirement remains to be determined.

Deletion of nonconserved ORF *e54*, which lies just upstream of the integrase start codon, resulted in the same phenotype as the SSV1-Δ*int* mutant. It is possible that deletion of *e54* somehow disrupts expression of the integrase gene, effectively resulting in a double mutant. The *integrase* gene occupies the 3′ end of the T5 transcript and is believed to be transcribed mainly via the T5 promoter. There is evidence that the integrase gene is transcribed from its own promoter. *integrase* mRNA was found in greater abundance than T5 mRNA following UV induction, and a similar phenomenon was observed for the SSV2 *integrase* gene during an analysis of the SSV2 transcription cycle (10, 48). Although an obvious promoter has not been identified upstream of the SSV1 *integrase* gene, several fuselloviruses (SSV2, SSV3, SSV4, and SSV9) encode putative promoters upstream of the integrase gene (49). Thus, based on the low conservation of the *e54* ORF and the evidence for an *integrase* promoter in this region, it seems likely that deletion of *e54* simultaneously disrupts expression of the *integrase* gene and explains the identical phenotype exhibited by both mutants. However, nonconserved ORF *f92*, which occupies the region of the SSV1 genome where such a promoter would be found, can be deleted without loss of infectivity (Fig. 1).

SSV1 ORF *a291* may be a cryptic conserved gene. The monocistronic transcripts T3 and Tx carry ORFs *a291* and *c124*, respectively, both of which reportedly have homologues in SSV2 (SSV2 ORFs *305* and *126*) (50) but were not identified using our BLAST-P analysis. ORF *c124* was shown to be nonessential for SSV1 infectivity, while ORF *a291* apparently is essential (Fig. 1B). Interestingly, most fusellovirus genomes encode similarly sized ORFs in an identical position just upstream of the putative tail fiber gene (SSV1 *vp4*), but with little to no detectable overall sequence similarity (5, 6).

Pairwise alignment of SSV1-A291 amino acid sequences and the syntenic SSV2 ORF SSV2-305 revealed that the N-terminal 20 amino acids of the two proteins are highly similar, while the remaining protein exhibits little to no similarity (33). The remaining 10 fusellovirus genomes (Table 1) were reexamined, and nine additional homologues of the N terminus of SSV1-A291 were identified (33). Furthermore, each of these ORFs is preceded by a putative promoter that shares significant similarity with the T3 promoter in SSV1, indicating that transcription of these genes is probably conserved (33). Because at least part of the SSV1 ORF *a291* is well conserved in the *Fuselloviridae*, it is not surprising that it appears essential for SSV1 infectivity. Whether only the N terminus of SSV1-A291 is required is not known.

The fusellovirus core is smaller than expected and is intolerant of mutagenesis. The fusellovirus core is the set of genes/ORFs that are carried by all known fuselloviruses. By analyzing 11 fusellovirus genomes, the core was reduced to 12 genes/ORFs (Fig. 1A) (5, 6), almost all of which appear to be essential for SSV1 infectivity (Fig. 1B). This is not surprising and reinforces the idea that these are critical to the viral life cycle. The only nonessential core gene identified was SSV1 *vp3*, which is discussed below. Core fusellovirus ORFs are clustered in one-half of the genome, with the exception of SSV1 ORF *b129*, and are upregulated during the middle to late portion of the SSV1 transcription cycle (10). Their timing of transcription and coexpression with known structural genes hints that most of the core genes have roles in virus replication, assembly, and packaging, although there is no experimental evidence to support this. We were not able to isolate any functional insertion mutants within this entire half of the genome, even in the poorly conserved and nonessential *vp2* gene (11). This insertion could cause a polar effect on the T9 transcript (Fig. 1) but has yet to be tested.

Insertions in ORFs *b49* and *f55* probably disrupt the SSV1 origin of replication. Unlike other known fuselloviruses, transcription of SSV1 is strongly induced by UV irradiation and is highly temporally regulated (9, 10). Following UV irradiation, the transcript T_{ind} is immediately upregulated and is swiftly followed by upregulation of the

two flanking transcripts, T5 and T6 (9, 10). ORF *b49* is the only ORF encoded by T_{ind} and possesses no homology to sequences in public databases, including other members of the *Fuselloviridae* (Fig. 1A). Due to the abundance of T_{ind} immediately following UV irradiation, it seems likely that the B49 protein plays a role in activation of viral transcription, either directly or indirectly. The *b49* ORF can be deleted and apparently is not essential for SSV1 infectivity (Fig. 1B). This agrees well with transcriptomic data from non-UV-induced SSV1-infected cells, where T_{ind} was not detected and presumably not required for infection (25). The effect of the ORF *b49* deletion on the SSV1 response to UV irradiation is unknown and could provide insight into this mechanism.

SSV1 encodes a second monocistronic transcript in this region, T_{lys}, that apparently is expressed constitutively (25, 26). T_{lys} encodes a 55-amino-acid protein (F55) that is hypothesized to repress transcription from early viral promoters, maintaining low virus expression in the absence of UV irradiation (25). Our results show that ORF *f55* is not required for the production of infectious virus. Absence of the F55 protein should result in a loss of repression of early viral promoters that could lead to constant expression of early gene products throughout the infection. However, we have not observed this to date, and plaque morphology appears to be similar to that of the wild type (data not shown).

This area of the SSV1 genome, near T_{ind} and T_{lys}, contains an abundance of promoters as well as the putative origin of replication and was intolerant of transposon insertion (Fig. 1B). The only functional insertion mutant in this region contained a transposon between the *f55* and *b49* ORFs, a significant distance from any of the known regulatory elements (Fig. 1B). All other insertions in this region fall within one of the two ORFs or were located adjacent to a promoter. Since ORFs *b49* and *f55* are not essential, it is unclear why insertions within these ORFs do not produce functional virus. These insertions may disrupt transcription of the T5 or T6 promoter (25, 51), but it is more likely that they disrupt replication. Unpublished data from the Steven D. Bell laboratory has mapped the origin of replication to this region (personal communication). Moreover, GC and purine skew analyses also indicated that the origin is within this area and appears to be well conserved in other fuselloviruses (49). Alternatively, this region could carry essential noncoding RNAs; however, none have been identified to date (10, 50).

The *vp3* minor capsid gene is not essential, but mutants have abnormal morphology. The only nonessential core gene was *vp3*, which is surprising considering its high degree of conservation and presence as a minor structural protein within the virion (17, 18, 20). Because the VP3 and proteolytically processed VP1 proteins are highly similar (Fig. 2), we hypothesize that VP1 partially complements the SSV1-Δ*vp3* mutant. SSV1-Δ*vp3* virions are highly abnormal relative to wild-type SSV1 (Fig. 4 and 5). Elongated particles are often observed in the virions of SSV6, SSV9, and ASV1; however, each of these viruses encodes a VP3 homologue (5, 6). Nonetheless, the dispensability of a seemingly critical gene is unexpected, and the consequences of its loss in regard to virion stability, infectivity, and structure remain to be investigated. Many *Sulfolobus* genomes encode cryptic fusellovirus genomes/genes (52), thus it is possible that the SSV1-Δ*vp3* mutant was able to remain infectious via complementation from a host-derived gene product. The sequence of *S. solfataricus* GΘ is unavailable, but the *S. solfataricus* P1 and P2 genomes have been sequenced and do not contain any obvious *vp3* homologues, making this scenario unlikely. The abnormal morphology of SSV1-Δ*vp3* virions is strikingly similar to intermediates in SSV1 budding after UV irradiation (53), indicating that VP3 plays a role in virion maturation.

Insertions in the *vp3* minor capsid protein gene are removed. Following the finding that SSV1-Δ*vp3* was infectious, several mutants harboring transposon insertions in the *vp1-vp3* gene region were isolated and assayed for infectivity (Fig. 1B and 3). Unlike other insertion mutants in the fusellovirus core, insertions in *vp3* and the *vp1*/*vp3* intergenic space appeared to be tolerated and did not appear to interfere with the production of infectious virus. However, viral DNA isolated from infected cultures

contained a deletion instead of an insertion. The deletion appears to have been facilitated by homologous recombination between two identical 61-bp sequences at the C termini of the *vp1* and *vp3* genes and results in the nearly complete deletion of the *vp3* gene (Fig. 3A). Full-length viral DNA harboring the transposon in this region could not be recovered from infected *Sulfolobus* cells, indicating that the transposon disrupts SSV1 infectivity and must be eliminated to produce infectious virus; presumably a partially functional VP3 protein or *vp3* gene is deleterious. These were the only mutations for which any modification other than LIPCR-mediated deletion or Tn*5* insertion was observed.

Our data suggest that this recombination event occurs in *Sulfolobus* following transformation of transposon-containing mutant DNA and not in *E. coli* prior to transformation. The same recombination likely occurs in wild-type SSV1, resulting in sporadic loss of the nonessential *vp3* gene. However, we have not been able to identify or isolate a spontaneous SSV1-Δ*vp3* mutant, suggesting that deletion mutants are outcompeted by wild-type virus containing VP3. Thus, an insertion in the *vp3* gene is more detrimental to SSV1 than a deletion, but a deletion is less fit than the wild type. SSV1 is the only fusellovirus that possesses such a long stretch of 100% identical bases within its *vp1* and *vp3* genes. Thus, the phenotype of a *vp3* deletion in another fusellovirus would be very interesting.

MATERIALS AND METHODS

Culture conditions. Infected and uninfected *Sulfolobus* strains were grown in yeast-sucrose liquid media (YS) and on Gelrite plates at 75°C as described previously (11). *E. coli* strains were grown on LB media both on agar plates and in liquid cultures with appropriate antibiotics (54). Strains used are listed in Table 4.

Purification of DNA. Plasmid DNA for LIPCR was purified from transformed *E. coli* using alkaline lysis (55). Plasmid DNA for transformation into *Sulfolobus* was purified from *E. coli* using the GeneJet plasmid purification kit by following the manufacturer's protocols (Thermo-Fisher).

SSV1 DNA used in transposon mutagenesis was purified from previously infected *S. solfataricus* strain S441 (38). Briefly, 50 ml of SSV1-infected cells was grown for 72 h at 75°C. SSV1 DNA was purified via alkaline lysis followed by 3 phenol-chloroform-isoamyl alcohol (25:24:1) extractions (56). Solvent was removed by passing DNA through a GeneJet plasmid purification column by following the manufacturer's protocol. SSV1 DNA was analyzed by UV absorption spectroscopy (absorption at 260/280 nm of ~1.8) and endonuclease digestion followed by agarose gel electrophoresis (54).

LIPCR. LIPCR (15) was used to delete SSV1 ORFs or portions thereof. For deletions, primers were designed to overlap the start and stop codons of the ORF to be deleted. Due to primer design considerations (e.g., incompatible melting temperatures [T_m], unfavorable secondary structures, primer dimers, etc.), most primers include portions of the 5′/3′ ends of the ORFs (Table 5). The optimal concentration of template DNA for LIPCR was determined empirically for each set of primers. Template DNA purified from *E. coli* with an initial concentration of approximately 200 ng/ml was initially diluted in 30 μl of H_2O with 0.01 μg RNase A and further diluted 1:10, 1:50, and 1:100 in double-distilled water (ddH_2O). LIPCR was performed as described in Iverson and Stedman (11) using Phusion DNA polymerase. The template for LIPCR reactions was pAJC97 (an SSV1 shuttle vector containing an *E. coli* plasmid in ORF *e178*), except for EAI580 and EAI582, which were constructed using a Tn*5* insertion in SSV1 as the template. Annealing temperatures for each primer pair were estimated using NEB T_m prediction software (http://tmcalculator.neb.com/#!/) and were experimentally optimized. LIPCR products were purified, phosphorylated, ligated, and transformed into chemically competent NovaBlue *E. coli* as in Iverson and Stedman (11).

Transposon mutagenesis. The EZ-Tn*5* <R6Kγori/KAN-2> insertion kit (Epicentre) was used to perform transposon mutagenesis on purified SSV1 DNA. A molar ratio of 30:1 SSV1 DNA to EZ-Tn*5* transposon was found to yield significantly more plasmid constructs containing the entire SSV1 genome than the manufacturer's recommended equimolar ratio (data not shown). This was the only deviation from the manufacturer's protocol (Epicentre). One microliter of the SSV1–EZ-Tn*5* reaction was electroporated into 50 μl of Transformax EC100D pir^+ electrocompetent *E. coli* (Epicentre) (Table 4) and plated on LB-agar plates with kanamycin.

Isolation and identification of transposon and deletion mutants. Plasmid DNA was purified via alkaline lysis (described above) from cultures from single colonies of *E. coli* following transformation with LIPCR products or transposon insertion reactions. Plasmid DNA was analyzed by restriction endonuclease digestion. All mutations were confirmed by DNA sequencing.

Transformation of *Sulfolobus*. Electrocompetent *Sulfolobus* species cells were prepared from mid-logarithmic cultures (optical density at 600 nm [OD_{600}] of 0.15 to 0.25) by washing with decreasing volumes of 20 mM sucrose essentially as in Schleper et al. (56). The final concentration of cells is ~10^{10} cells/ml. One hundred microliters of washed cells was added to a chilled 0.1-cm-gap-length cuvette (Bulldog Bio), and 2 μl of SSV1 DNA (100 to 500 ng/μl) was added to the cells. Cells were transformed by electroporation (Gene Pulser II; Bio-Rad) using the following conditions: 1.5 kV, 400 Ω, 25 μF.

TABLE 5 Primers used

LIPCR mutation	Primer sequence[a] (T_m, °C)		No. of amino acids remaining in ORF/% of ORF remaining[b]
	Forward	Reverse	
ΔC102a	**CTG AAT GGC TAA AAA GAA CGG** (62)	**TGT TAT TAT TTC TGT TAC TGA GAC C** (57)	12/11
ΔB251	**ATT GAC GAC GTA ACA AGA TAG** (56)	**TAA CCA TAA CCA TTC ATT ACT C** (55)	23/9
ΔE54	**ACT CAT TTG TCC ACC TTG** (56)	**CAA CCA TAA** TAC TGT GAG G (53)	13/24
ΔF92	**CTC TGA GTT GAA TAT CAT** TTT CC (58)	**ATA GCA TGG CTA GAA TAC AAG G** (59)	16/17
ΔF93	**CAT** ATC CTC CTC ACT CCT CAG (60)	**ATC AAA AAG AGG TGA** ACT AGA TGG (61)	5/5
ΔE96	**CAG ATC GAA TAT AGG AAC TTG C** (59)	**TAA** ATG ATA GAG AAG AGG AAA GAT AGG (60)	10/10
ΔF112	CAT CTT GTA TGA ATT TAG AGT TTG TGC (63)	**AAG GCA AAG CAG TGA** ACT GAC (63)	14/13
ΔB49	**GTA GAA GCA ATA AAT GAT TTG** (53)	**CTC AGA TTT TGC ACA TCC** (56)	15/30
ΔF55	**TTT CCT CGG CAT** ACG CTA TC (64)	**TAA** ATG CCC TAC TAT ACT CTA TCT CTC TC (60)	4/7
ΔA100	**TTT GAC TTC TGA** GGA GG (53)	**TAA CTC TTC TTC TTT TCG GG** (58)	15/15
ΔC80	**TTA GCG AGG TAT GTA GAA AAT GTT TAG** ATG C (67)	**TGT GTA ACA TCT AGG TAA TTT GAT GTA TTC** (62)	23/29
ΔA79	**GTT GAG TGA** ATA ATG TAT CAA TGT CTA C (60)	**GCA TCT AAA CAT** TTT CTA CAT ACC TC (60)	6/8
ΔA45	**GTG AGA GGA CAA TGA** ACC (55)	**TTG ATA CAT** TAT TCA CTC AAC C (56)	7/15
ΔC102b	**GGA TGA** CGG AGT CAG ACG TTG (67)	**GTT CAT** TGT CCT CTC ACC CTG AAG (67)	3/3
B129 (ΔN-terminus)	**AAA GCG ATT TCA CAG TTT GTC** (61)	**TGA CTC CGT CAT** CCT CTA AC (59)	64/50
B129 (ΔC-terminus)	**AGT TAG** GCT CTT TTT AAA GTC TAC C (58)	**TGA AAT CGC TTT ACT CGC** (59)	74/57
ΔC124	**AGA AGA TAG** CCC TTT TTA AAG CC (62)	**GAA AAT AGA ACC TAC AAC TGT AAA CAG** (59)	14/11
ΔB115	**TGG AGG GGT TTA AAA ACG TAA G** (63)	**TCA T**TC CGA CCC CCT AAT TAA C (65)	33/29
ΔVP1	**TGA** GGG ATG GAA ATC AGT TTA AAG (64)	**CAA** ACT CCT TAG GAG TCT CAT CC (62)	2/1
ΔVP3	**TGA** TAT GAA GTG GGT GCA AAA GG (67)	**CAT** CCC TCA CAC CTC AGT CTT TG (67)	2/2
ΔD335 (Integrase)	**CAT** TTC GCC TCA CAG TAT TAT GG (64)	**GTC TGA** CAT TAC CCG TAT CAC (58)	2/1
ΔD63-F92	**TTC TTT ACT CAT** TGT TTT TCA CCT TAG (62)	**ATA GCA TGG CTA GAA TAC AAG G** (59)	NA
Other primers			
Univ_7F[c]	ATTCAGATTCTGWATWCAGAA		
Univ_8R[d]	TCSCCTAACGCACTCATC		

[a]All sequences are 5′→3′. Bases in boldface denote bases in the SSV1 ORF.
[b]Does not include the stop codon. Determined after deletion with LICPR.
[c]Amplifies structural gene region of all SSVs with univ_8R (57).
[d]Amplifies structural gene region of all SSVs with univ_7F (57).

Immediately following electroporation, cells were resuspended in 1 ml of 70°C YS, transferred to a 1.5-ml tube, and incubated for 1 h in a 70°C incubator. Following incubation, cells were transferred to 50 ml of preheated YS in a long-neck Erlenmeyer flask and grown with shaking at 70°C. Cultures appeared turbid after ~24 to 36 h.

Confirmation of infectious SSV DNA. Spot-on-lawn or halo assays were performed in duplicate 48 and 72 h after transformation of *Sulfolobus* with SSV DNA. Halo assays were performed as in Iverson and Stedman (11) by spotting 1 to 5 μl of transformed cultures on an indicator lawn of uninfected *Sulfolobus* on a Gelrite plate, followed by incubation at 70°C for 48 to 72 h. Positive controls included transformation with wild-type SSV1 DNA and known functional mutants. Negative controls were uninfected *Sulfolobus* cultures. Transformed cultures that inhibited host growth (halo producers) were further analyzed to confirm the identity of the viral DNA. Viral DNA purified from infected cells was amplified with PCR using primers that flank the mutated region of the viral DNA. Control PCRs used the mutant DNA used for transformation and wild-type SSV1 DNA.

Transmission electron microscopy. For transmission electron microscopy, samples were prepared on 400-mesh carbon-Formvar-coated copper grids (Ted Pella). Grids were placed, carbon-Formvar down, on a 5-μl droplet of culture or culture supernatant for 2 min. Samples were removed from the grid by wicking. Grids were then stained for 60 s on 5 μl of either 2% uranyl acetate stain (pH ~3) or 2% sodium phosphotungstate tribasic hydrate stain (pH ~6). Phosphotungstate stain was made freshly every week to ensure that the solution did not disassociate. Grids were allowed to dry in air overnight and were examined within 48 h of staining. Images were obtained at 8,500× to 34,000× magnification on an FEI Tecnai F20 transmission electron microscope (TEM). Grids were analyzed by examining randomly selected grid squares. Images were obtained with a BM UltraScan camera and stored in digital micrograph 3 and TIFF formats.

Particle analysis. The length and width of images of stained virus particles were measured in ImageJ (58). Normal particle width, length, and aspect ratio were determined using the means of measurements of wild-type SSV1 particles ($n = 240$). Any particles whose width, length, or aspect ratio was more than two standard deviations from the means were classified as abnormal.

Accession number(s). Sequences for SSVL and SSV3 have been deposited in GenBank under accession numbers KY563228 and KY579375, respectively.

ACKNOWLEDGMENTS

We thank undergraduate students in the Recombinant DNA Techniques Laboratory at Portland State University (spring terms 2012 and 2013) for initial mapping and activity screening of REC mutants (228-231, 243-245, 262, and 322-325). We also thank Jordan Hartunians for mapping and screening of mutants JAH572, JAH573, and JAH576. We thank C. Martin Lawrence for sharing data on SSV1-B129 and Stephen D. Bell for sharing unpublished data on the SSV1 origin of replication. We thank Valerie P. Huang for isolation of some of the SSV1 EZ-Tn*5* insertion mutants.

This work was funded by grants from the National Science Foundation (MCB0702020, MCB1243963, and DMR1263339) and the National Institutes of Health (UL1GM118964) and received support from Portland State University.

REFERENCES

1. Prangishvili D. 2013. The wonderful world of archaeal viruses. Annu Rev Microbiol 67:565–585. https://doi.org/10.1146/annurev-micro-092412-155633.
2. Krupovic M, Quemin ERJ, Bamford DH, Forterre P, Prangishvili D. 2014. Unification of the globally distributed spindle-shaped viruses of the Archaea. J Virol 88:2354–2358. https://doi.org/10.1128/JVI.02941-13.
3. Yeats S, McWilliam P, Zillig W. 1982. A plasmid in the archaebacterium *Sulfolobus acidocaldarius*. EMBO J 1:1035–1038.
4. Stedman KM, She QX, Phan H, Arnold HP, Holz I, Garrett RA, Zillig W. 2003. Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. Res Microbiol 154:295–302. https://doi.org/10.1016/S0923-2508(03)00074-3.
5. Redder P, Peng X, Brugger K, Shah SA, Roesch F, Greve B, She QX, 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. https://doi.org/10.1111/j.1462-2920.2009.02009.x.
6. Wiedenheft B, Stedman K, Roberto F, Willits D, Gleske AK, Zoeller L, Snyder J, Douglas T, Young M. 2004. Comparative genomic analysis of hyperthermophilic archaeal *Fuselloviridae* viruses. J Virol 78:1954–1961. https://doi.org/10.1128/JVI.78.4.1954-1961.2004.
7. 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. https://doi.org/10.1016/0042-6822(91)90771-3.
8. Nadal M, Mirambeau G, Forterre P, Reiter WD, Duguet M. 1986. Positively supercoiled DNA in a virus-like particle of an archaebacterium. Nature 321:256–258. https://doi.org/10.1038/321256a0.
9. Reiter WD, Palm P, Yeats S, Zillig W. 1987. 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. https://doi.org/10.1007/BF00329653.
10. Fröls S, Gordon PMK, Panlilio 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. https://doi.org/10.1016/j.virol.2007.03.033.
11. Iverson E, Stedman K. 2012. A genetic study of SSV1 the prototypical fusellovirus. Front Microbiol 3:200.
12. Prangishvili D, Forterre P, Garrett RA. 2006. Viruses of the Archaea: a unifying view. Nat Rev Microbiol 4:837–848. https://doi.org/10.1038/nrmicro1527.
13. Iranzo J, Koonin EV, Prangishvili D, Krupovic M. 2016. Bipartite network analysis of the archaeal virosphere: Evolutionary connections between viruses and capsidless mobile elements. J Virol 90:11043–11055. https://doi.org/10.1128/JVI.01622-16.
14. Iranzo J, Krupovic M, Koonin EV. 2016. The double-stranded DNA viro-

sphere as a modular hierarchical network of gene sharing. mBio 7:e00978–16.
15. Clore AJ, Stedman KM. 2007. The SSV1 viral integrase is not essential. Virology 361:103–111. https://doi.org/10.1016/j.virol.2006.11.003.
16. 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. https://doi.org/10.1074/jbc.R800078200.
17. Reiter WD, Palm P, Henschen A, Lottspeich F, Zillig W, Grampp B. 1987. Identification and characterization of the genes encoding 3-structural proteins of the *Sulfolobus* virus-like particle SSV1. Mol Gen Genet 206: 144–153. https://doi.org/10.1007/BF00326550.
18. Quemin ERJ, Pietila MK, Oksanen HM, Forterre P, Rijpstra WIC, Schouten S, Bamford DH, Prangishvili D, Krupovic M. 2015. Sulfolobus spindle-shaped virus 1 contains glycosylated capsid proteins, a cellular chromatin protein, and host-derived lipids. J Virol 89:11681–11691. https://doi.org/10.1128/JVI.02270-15.
19. Servin-Garcidueñas LE, Peng X, Garrett RA, Martinez-Romero E. 2013. Genome sequence of a novel archaeal fusellovirus assembled from the metagenome of a Mexican hot spring. Genome Announc 1:e0016413.
20. 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. https://doi.org/10.1128/JVI.01663-09.
21. Stedman KM, DeYoung M, Saha M, Sherman MB, Morais MC. 2015. Structural insights into the architecture of the hyperthermophilic fusellovirus SSV1. Virology 474:105–109. https://doi.org/10.1016/j.virol.2014.10.014.
22. Menon SK, Maaty WS, Corn GJ, Kwok SC, Eilers BJ, Kraft P, Gillitzer E, Young MJ, Bothner B, Lawrence CM. 2008. Cysteine usage in Sulfolobus spindle-shaped virus 1 and extension to hyperthermophilic viruses in general. Virology 376:270–278. https://doi.org/10.1016/j.virol.2008.03.026.
23. Kraft P, Oeckinghaus A, Kümmel D, Gauss GH, Gilmore J, Wiedenheft B, Young M, Lawrence CM. 2004. Crystal structure of F-93 from Sulfolobus spindle-shaped virus 1, a winged-helix DNA binding protein. J Virol 78:11544–11550. https://doi.org/10.1128/JVI.78.21.11544-11550.2004.
24. Schlenker C, Menon S, Lawrence CM, Copié V. 2009. (1)H, (13)C, (15)N backbone and side chain NMR resonance assignments for E73 from Sulfolobus spindle-shaped virus ragged hills, a hyperthermophilic crenarchaeal virus from Yellowstone National Park. Biomol NMR Assign 3:219–222. https://doi.org/10.1007/s12104-009-9179-z.
25. Fusco S, She QX, 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. https://doi.org/10.1128/JVI.00458-13.
26. Fusco S, She QX, Fiorentino G, Bartolucci S, Contursi P. 2015. Unravelling the role of the F55 regulator in the transition from lysogeny to UV induction of Sulfolobus spindle-shaped virus 1. J Virol 89:6453–6461. https://doi.org/10.1128/JVI.00363-15.
27. Arnold HP, She Q, Phan H, Stedman K, Prangishvili D, Holz I, Kristjansson JK, Garrett R, Zillig W. 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. https://doi.org/10.1046/j.1365-2958.1999.01573.x.
28. Koonin EV. 1992. Archaebacterial virus SSV1 encodes a putative DnaA-like protein. Nucleic Acids Res 20:1143. https://doi.org/10.1093/nar/20.5.1143.
29. Kraft P, Kümmel D, Oeckinghaus A, Gauss GH, Wiedenheft B, Young M, Lawrence CM. 2004. Structure of D-63 from Sulfolobus spindle-shaped virus 1: surface properties of the dimeric four-helix bundle suggest an adaptor protein function. J Virol 78:7438–7442. https://doi.org/10.1128/JVI.78.14.7438-7442.2004.
30. 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.
31. Wirth JF, Snyder JC, Hochstein RA, Ortmann AC, Willits DA, Douglas T, Young MJ. 2011. Development of a genetic system for the archaeal virus Sulfolobus turreted icosahedral virus (STIV). Virology 415:6–11. https://doi.org/10.1016/j.virol.2011.03.023.
32. 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. https://doi.org/10.1046/j.1365-2958.2003.03509.x.
33. Iverson EA. 2015. A genetic and biochemical analysis of *Sulfolobus* spindle-shaped virus 1. Ph.D. Thesis, Portland State University, Portland, OR.
34. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.
35. Stedman K, Clore A, Combet-Blanc Y. 2006. Biogeographical diversity of archaeal viruses, p 131–143. *In* Logan NA, Lappin-Scott HM, Oyston PCF (ed), Prokaryotic diversity: mechanisms and significance. Cambridge University Press, Cambridge, United Kingdom.
36. 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.
37. Muskhelishvili G, Palm P, Zillig W. 1993. SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. Mol Gen Genet 237: 334–342.
38. 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.
39. Zillig W, Stetter KO, Wunderl S, Schulz W, Priess H, Scholz I. 1980. The *Sulfolobus-Caldariella* group–taxonomy on the basis of the structure of DNA-dependent RNA-polymerases. Arch Microbiol 125:259–269.
40. Liu G, She Q, Garrett RA. 2016. Diverse CRISPR-Cas responses and dramatic cellular DNA changes and cell death in pKEF9-conjugated *Sulfolobus* species. Nucleic Acids Res 44:4233–4242.
41. She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CCY, Clausen IG, Curtis BA, De Moors A, Erauso G, Fletcher C, Gordon PMK, Heikamp-de Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Van der Oost J. 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci U S A 98: 7835–7840. https://doi.org/10.1073/pnas.141222098.
42. Cannio R, Contursi P, Rossi M, Bartolucci S. 1998. An autonomously replicating transforming vector for *Sulfolobus solfataricus*. J Bacteriol 180:3237–3240.
43. Hendrix RW, Smith MCM, Burns RN, Ford ME, Hatfull GF. 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc Natl Acad Sci U S A 96:2192–2197. https://doi.org/10.1073/pnas.96.5.2192.
44. Marinelli LJ, Piuri M, Swigonová Z, Balachandran A, Oldfield LM, van Kessel JC, Hatfull GF. 2008. BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. PLoS One 3:e3957. https://doi.org/10.1371/journal.pone.0003957.
45. Hatfull GF. 2015. Dark matter of the biosphere: the amazing world of bacteriophage diversity. J Virol 89:8107–8110. https://doi.org/10.1128/JVI.01340-15.
46. Krupovic M, Cvirkaite-Krupovic V, Prangishvili D, Koonin EV. 2015. Evolution of an archaeal virus nucleocapsid protein from the CRISPR-associated Cas4 nuclease. Biol Direct 10:65. https://doi.org/10.1186/s13062-015-0093-2.
47. She Q, Peng X, Zillig W, Garrett RA. 2001. Gene capture in archaeal chromosomes. Nature 409:478. https://doi.org/10.1038/35054138.
48. 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. https://doi.org/10.1016/j.virol.2013.03.012.
49. Clore AJ. 2008. The family *Fuselloviridae*: diversity and replication of a hyperthermophilic virus infecting the archaeon genus *Sulfolobus*. Ph.D. thesis, Portland State University, Portland, OR.
50. Fusco S, Liguori R, Limauro D, Bartolucci S, She QX, Contursi P. 2015. Transcriptome analysis of *Sulfolobus solfataricus* infected with two related fusselloviruses reveals novel insights into the regulation of CRISPR-Cas system. Biochimie 118:322–332. https://doi.org/10.1016/j.biochi.2015.04.006.
51. Qureshi SA. 2007. Protein-DNA interactions at the Sulfolobus spindle-shaped virus-1 (SSV1) T5 and T6 gene promoters. Can J Microbiol 53:1076–1083. https://doi.org/10.1139/W07-065.
52. Held NL, Whitaker RJ. 2009. Viral biogeography revealed by signatures in

Sulfolobus islandicus genomes. Environ Microbiol 11:457–466. https://doi.org/10.1111/j.1462-2920.2008.01784.x.
53. Quemin ERJ, Chlanda P, Sachse M, Forterre P, Prangishvili D, Krupovic M. 2016. Eukaryotic-like virus budding in Archaea. mBio 7: e01439–16.
54. Green MR, Sambrook J. 2012. Molecular cloning: a laboratory manual, 4th ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
55. Birnboim HC, Doly J. 1979. Rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523. https://doi.org/10.1093/nar/7.6.1513.
56. 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 U S A 89: 7645–7649. https://doi.org/10.1073/pnas.89.16.7645.
57. Snyder JC, Spuhler J, Wiedenheft B, Roberto FF, Douglas T, Young MJ. 2004. Effects of culturing on the population structure of a hyperthermophilic virus. Microb Ecol 48:561–566. https://doi.org/10.1007/s00248-004-0246-9.
58. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675. https://doi.org/10.1038/nmeth.2089.
59. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948.