



Contents lists available at ScienceDirect

Biochimie

journal homepage: [www.elsevier.com/locate/biochi](http://www.elsevier.com/locate/biochi)

Research paper

## Transcriptome analysis of *Sulfolobus solfataricus* infected with two related fuselloviruses reveals novel insights into the regulation of CRISPR–Cas system

Salvatore Fusco<sup>a</sup>, Rossana Liguori<sup>b</sup>, Danila Limauro<sup>a</sup>, Simonetta Bartolucci<sup>a</sup>,  
Qunxin She<sup>c,\*</sup>, Patrizia Contursi<sup>a,\*</sup>

<sup>a</sup> Dipartimento di Biologia, Università degli Studi di Napoli Federico II, Complesso Universitario Monte S. Angelo, Via Cinthia, 80126 Napoli, Italy

<sup>b</sup> Dipartimento di Scienze Chimiche, Università degli Studi di Napoli Federico II, Complesso Universitario Monte S. Angelo, Via Cinthia, 80126 Napoli, Italy

<sup>c</sup> Danish Archaea Centre, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark

### ARTICLE INFO

#### Article history:

Received 27 February 2015

Accepted 8 April 2015

Available online xxx

#### Keywords:

CRISPR–Cas system

*Sulfolobus solfataricus*

Fuselloviridae

*Sulfolobus* spindle-shaped virus

Transcriptome analysis

### ABSTRACT

Fuselloviruses SSV1 and SSV2 are model systems to investigate virus–host relationships in stably infected cells thanks to their temperate nature. Although they are very similar in morphology, genome organization and gene synteny, their replication is induced by different stimuli, i.e.: by UV-light exposure (for SSV1) and by the growth progression of the host (for SSV2). In this study, we have analysed global gene expression in SSV1- and SSV2-lysogens of *Sulfolobus solfataricus* P2 in the absence of any stimuli. Additionally, the interplay among SSV1, SSV2 and the host has been investigated in a double-infected strain to explore both virus–host and virus–virus interactions. Whereas SSV1 did not induce major changes of the host gene expression, SSV2 elicited a strong host response, which includes the transcriptional activation of CRISPR loci and *cas* genes. As a consequence, a significant decrease of the SSV2 copy number has been observed, which in turn led to provirus-capture into the host chromosome. Results of this study have revealed novel aspects of the host–viral interaction in the frame of the CRISPR–response.

© 2015 Elsevier B.V. and Société Française de Biochimie et Biologie Moléculaire (SFBBM). All rights reserved.

### 1. Introduction

The majority of organisms is susceptible to viral infection and accordingly it has been predicted that viruses inhabit several niches worldwide [1,2]. This makes viruses, especially those infecting prokaryotes (i.e. *Archaea* and *Bacteria*), the most predominant biological entity on Earth [3]. Soon after the discovery of the archaeal domain, many viruses and virus-like particles have been isolated from extremely hot, low pH or hypersaline niches [4,5]. Intriguingly, these viruses exhibit unique morphologies compared to those of bacteriophages, and fall into ten new virus families in the current classification [4,6,7]. For instance, spindle-shaped viruses are exceptional on their own since this morphotype is a hallmark of viruses infecting archaeal

microorganisms [8] and it has never been described for bacteriophages or eukaryal viruses.

Spindle-shaped viruses infecting organisms of the genus *Sulfolobus* (SSVs) belong to the family *Fuselloviridae*, which comprises so far ten members (SSV1, SSV2, SSV3, SSV4, SSV5, SSV6, SSV7, SSV8, SSV9 and ASV1) [5,9]. *Sulfolobus* spindle-shaped virus 1 (SSV1) is a model virus for investigating virus–host relationship, since it is the only UV-inducible archaeal virus isolated to date [9–11]. Analysis of SSV1 transcription has laid the basis for understanding how archaeal gene expression is regulated [12,13]. Moreover, its genome has served as a backbone for the construction of vectors for genetic manipulation and gene expression in *Sulfolobus* [9]. SSV1 is a temperate virus that, upon infection, establishes a stable coexistence with the host by keeping its copy number low and constant throughout the growth of infected cells. It is worth noting that, unlike lambda-lysogens, in which only proviruses exist, SSV1-lysogens carry both a provirus (an integrated viral genome in the host chromosome) and a few episomal copies of the viral DNA [14,15]. Consequently, viral progeny is

\* Corresponding authors.

E-mail addresses: [qunxin@bio.ku.dk](mailto:qunxin@bio.ku.dk) (Q. She), [contursi@unina.it](mailto:contursi@unina.it) (P. Contursi).

<sup>1</sup> These authors contributed equally to this work.

constitutively produced at low level even in the absence of the inducing stimulus (i.e. UV-light exposure). Therefore, the lysogenic state of SSV1 is better defined as a carrier state [11,16]. The transcriptional map of SSV1 highlighted the chronological expression of viral genes as consequence of the UV exposure [10] and helped speculating about the function of their products. However, so far, the general lack of sequence similarity with other proteins in public databases has hindered functional studies of these viral genes. In fact, only a limited number of proteins encoded by SSV1 or other fuselloviruses have been functionally and/or structurally characterized [11,17–24].

On the other hand, although SSV2 resembles SSV1 in the shape of the viral particle as well as in genome organization and gene synteny, its replication is not induced by an external stimulus as for SSV1. Rather, the SSV2 copy number is kept constant and low (around 1–3 copies per cell) until the native host *Sulfolobus islandicus* REY15/4 enters into the stationary growth phase, when a steep increase of the copy number occurs (25–50 copies per cell) [25]. This effect has not been observed for the permissive host *Sulfolobus solfataricus*, in which SSV2 shows a copy number comparable to that of the induced state of the natural host. This led to the hypothesis that a transcription factor encoded by the natural host *S. islandicus* was involved in the regulation of the viral replication induction and this factor is presumably absent in the foreign host *S. solfataricus* [25].

Recently, a transcriptome analysis has been carried out to define the gene expression pattern of SSV2 in the early stages of infection of *S. solfataricus* cells (up to 9 h post infection) [26]. This allowed detecting seven non-overlapping transcripts that were termed after the SSV1 counterparts. Although, genome transcription occurred (as for SSV1) in a temporal fashion, early genes were not adjacently located and displayed a distributive pattern of expression [26]. Since the study by Ren et al., is limited to a short post-infection time window, genes responsible for the maintenance of the SSV2 carrier state have not been identified.

Only a few global gene expression analyses of hosts infected with archaeal viruses have been reported [10,26–29]. Worth-mentioning are those performed on the lytic *Sulfolobus islandicus* rod-shaped virus 2 (SIRV2), which exhibits a temporal pattern of gene expression [28] as well as on the *Sulfolobus* turreted icosahedral virus (STIV), which does not show any temporal regulation [27]. Viral–host interactions in *Crenarchaea* have been successfully investigated at gene expression level both for lytic and temperate viruses, using as host *S. solfataricus* P2. Since it has been shown that several *S. solfataricus* strains isolated in Italy are susceptible to infection by all fuselloviruses [30], this host is a suitable crenarchaeal model to study fusellovirus–host relationships. Moreover, the strain P2 harbours the CRISPR-Cas antiviral defense system, which is one of the main players of the virus–host arms race. In brief, to achieve viral immunity, DNA sequences of invading genetic elements (i.e. protospacers) are integrated in the array of a CRISPR locus in a process known as adaptation. Transcription of these loci produces crRNAs that form ribonucleoprotein complexes with Cas proteins. Foreign nucleic acids are detected and degraded by these complexes in the interference stage [31].

The genome of *S. solfataricus* P2 harbours six CRISPR loci (from A to F) and *cas* gene cassettes encoding for up to seven different interference complexes, i.e. of subtypes I-A (Cascade complex), III-B (CMR complex) and III-A [31]. Whereas ribonucleoprotein complexes of subtypes IA and III-A exert DNA recognition and degradation, those belonging to the subtype III-B cleave RNA in a sequence-specific manner. These features make *S. solfataricus* P2 an exceptional model to investigate the modulation of the CRISPR-Cas system activity in response to the infection by different viruses. Herein we report a whole-transcriptome analysis that

highlights gene expression remodelling in *S. solfataricus* upon the establishment of SSV1 and SSV2 into the host cells. Previous studies have been carried out to investigate on variation of gene expression in the immediate aftermath of SSV2 infection (up to 9 h later) [26] or upon UV-stimulus in SSV1 lysogens [10]. We show that once SSV1 and SSV2 established a carrier state, the host response elicited by the two viruses is very different. Noteworthy, the up-regulation of CRISPR-Cas system occurs only in cells infected with SSV2.

## 2. Materials and methods

### 2.1. Strains, media and growth conditions

SSV1-, SSV2- and SSV1/SSV2-infected strains of *S. solfataricus* P2 were generated, as described elsewhere [11,25], using as host the uracil auxotrophic mutant InF1 [32]. Cultures were grown aerobically in TYSU, i.e. a glycine-buffered Brock's basal salt solution supplemented with 0.1% tryptone, 0.05% yeast extract, 0.2% sucrose and 0.002% uracil (w/v); the pH was adjusted to 3.2 with H<sub>2</sub>SO<sub>4</sub> [32]. Incubation was conducted in 250-ml Erlenmeyer flasks at 75 °C, with a shaking rate of 150 rpm in an Innova 3100 Water bath shaker (New Brunswick Scientific Corp).

Aliquots from frozen cultures of the uninfected InF1 and the infected strains (SSV1-InF1, SSV2-InF1 and SSV1/SSV2-InF1) were revitalized by inoculating in TYSU medium. Cell growth was spectrophotometrically monitored at 600 nm (OD<sub>600</sub>) throughout the cultivation by means of a Variant Cary<sup>®</sup> 50 Bio UV/Visible Spectrophotometer (McKinley Scientific). Once reached 0.4 OD<sub>600</sub>, cultures were diluted to 0.05 OD<sub>600</sub>, incubated back to 75 °C and samples collected at 0.4 OD<sub>600</sub> (early exponential phase) and 1.2 OD<sub>600</sub> (late exponential phase). Cellular pellets were obtained by centrifugation at 3000 × g for 10 min using the Centrifuge 5810R (Eppendorf) and treated for total DNA and RNA preparations.

To isolate single clones from the SSV2-InF1 strain, serial dilutions of this culture were plated on TYSU-Gelrite and incubated at 75 °C. Isolated colonies appeared on the plate surface after 7–10 days (about 100 colonies per plate). Several colonies were inoculated in liquid medium (about 10 per plate), let to grow until 0.4 OD<sub>600</sub> and culture supernatant tested by plaque assay for viral titre determination, using the uninfected InF1 strain as lawn and the supernatant of the SSV2-InF1 culture as a control. Single clones showing a lower viral titre were streaked on plates three times in the attempt to isolate SSV2-cured cells.

### 2.2. RNA extraction, cDNA synthesis and labelling

Total RNA samples were prepared using the TRIzol reagent (Sigma Aldrich<sup>®</sup>) and carried-over DNA was digested using Turbo<sup>™</sup> DNase (Ambion<sup>®</sup>), according to the manufacturer's instructions. The enzyme was thermal inactivated at 70 °C for 10 min after the addition of 5 mM EDTA (Ethylene Diamine Tetraacetic Acid). DNA-free RNA samples were purified by phenolic extraction and ethanol precipitation. RNA pellets were dissolved in nuclease-free water and both concentration and integrity were checked through: i) electrophoresis on denaturing, formaldehyde-containing 2.0% agarose gel and ii) determination of the 260nm/280 nm adsorption ratio using a Nanodrop 1000 Spectrophotometer (Thermo Scientific). Only samples showing ratios between 2.1 and 1.9 were used for cDNA synthesis.

cDNA labelled with 5-(3-aminoallyl)-dUTP (aa-dUTP) was generated using the Amersham CyScribe Post Labeling Kit (GE Healthcare) according to the manufacturer's directions, with few modifications as follows: 4 µl of random hexamers primers, 1 µl

random nonamers primers, 10–15 µg total RNA and nuclease-free water were mixed in a final volume of 11 µl. The mixture was denatured at 70 °C for 5 min and chilled at room temperature for 10 min. Four microlitres of 5 × CyScribe buffer, 1 µl dNTP mix, 1 µl aa-dUTP, 2 µl DTT 0.1 M and 1 µl CyScribe reverse transcriptase, were added to the reaction mixture and incubated at 42 °C for 90 min. Afterwards, the RNA template was degraded by adding 2 µl of 2.5 M NaOH and incubating at 37 °C for 15 min. The allyl-dUTP cDNA sample was purified using the Illustra Cyscribe GFX Purification kit (GE Healthcare) after neutralizing the reaction mixture with the addition of 10 µl of 2 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid). The concentration of the purified cDNA was spectrophotometrically determined.

The CyDye labelling of the amino allyl-modified cDNA was achieved using the Amersham CyScribe Post Labeling Kit (GE Healthcare) and purification of the labelled cDNA was carried out using the MinElute® PCR purification kit (Qiagen), following the manufacturer's instructions. Reference (ref) and experimental (exp) cDNA samples were labelled with cyanine-3 (Cy-3) and cyanine-5 (Cy-5), respectively. The ratio between the CyDye-labelled and the total cDNAs was used to monitor the labelling efficiency. Only reference and experimental cDNA samples showing similar labelling efficiency were co-hybridized on the same slide. For each hybridization reference and experimental samples are indicated in the text (Section 3, Results and discussion).

### 2.3. Microarray hybridization and data analysis

Customized microarray slides harbouring about 7000 spots were designed by the *Sulfolobus* genome chips consortium and manufactured by Ocimum Biosolutions (Hyderabad). A single array includes probes, spotted in duplicate, for 3042 *S. solfataricus* P2 genes, for several crenarchaeal viruses, three plasmids as well as for human and *Arabidopsis thaliana* sequences as negative controls [26–28]. The microarray slide was first dipped into a coupling jar containing 25 ml of prewarmed prehybridization solution (1% BSA; 5 × SSC; 0.1% SDS) and then incubated at 42 °C for 40 min under shaking (Thermo-Electron Corporation). The prehybridized slide was washed three times with distilled water and once with 100% isopropanol. Finally, the slide was dried by centrifugation and a LifterSlip coverslip was applied onto the array-area.

The hybridization solution, i.e. 60% deionized formamide, 7 × SSC (1.0 M NaCl, 0.1 M sodium citrate), 2% SDS, 0.2 µg/µl herring sperm DNA and 0.2 µg/µl tRNA was incubated at 95 °C for 2 min and chilled on ice for 1 min, before adding equal amounts of reference and experimental cDNAs (50 pmol CyDye for each). Subsequently, the hybridization mixture was carefully injected between the array surface and the coverslip. The slide was sealed in a hybridization chamber and incubated at 42 °C for 18 h. Washing of the hybridized slide was performed with: i) a prewarmed solution A (2 × SSC; 0.1% SDS), at 42 °C for 5 min under gentle shaking; ii) a prewarmed solution B (0.1 × SSC; 0.1% SDS) for 20 min at 42 °C and iii) a solution C (0.1 × SSC) for five times at room temperature. Finally, the slide was dried by centrifugation and immediately scanned using the Array Worf<sup>®</sup> (Applied Precision). Data analysis was conducted by ImaGene<sup>®</sup> v. 9.0 (Bio-Discovery) using default settings, and included the following steps: data import, background adjustment, normalization, summarization and quality assessment.

Up to four independent experiments were carried out for each reference/experimental cDNA couple. Moreover, since each probe is present in duplicate on a slide, the fold-change variation for each gene is the average among the collected data points. In particular, the software ImaGene<sup>®</sup> gives a log<sub>2</sub> ratio value for each analysed spot:

$$[\log_2(\text{FCy-5 exp})] - [\log_2(\text{FCy-3 ref})] = \log_2(\text{FCy5 exp/FCy3 ref})$$

where “F” indicates the normalized fluorescence intensity of the Cy-3 and Cy-5 detected in a given spot. The variation of the gene expression fold (VGEF) is calculated through the following equation:

$$\text{VGEF} = 2^{\log_2(\text{FCy-5 exp/FCy-3 ref})}$$

Only genes with a VGEF of ≥2 as well as a *p* value of <0.05 were regarded as differentially expressed.

### 2.4. PCR analysis of CRISPR loci leader-proximal regions and of the SSV2 integration site

A fundamental step in the CRISPR-Cas system is the adaptation, which consists of the Cas proteins-mediate incorporation of new spacers that occurs immediately downstream the leader sequence of a given CRISPR locus. With the purpose of detecting new spacers integration at the leader-proximal regions of all CRISPR loci (A–F) in the *S. solfataricus* genome, six primer couples were used (Table S1). In particular, these oligonucleotides were designed to amplify the first 5–8 repeat-spacer units located immediately downstream the leader sequence [33]. Thus, integration of new spacers will produce longer PCR products than the control sample (uninfected InF1).

Total DNA samples were prepared using the DNeasy tissue kit (Qiagen), following the manufacturer's instructions and their concentration was spectrophotometrically measured using a Nanodrop 2000 Spectrophotometer (Thermo-Scientific). PCR master mixes were prepared as follows: 1 × Taq Buffer, 2.5 mM MgCl<sub>2</sub>, 0.2 mM dNTP mix, 0.6 µM primer-fw, 0.6 µM primer-rv and 0.05 U/µl of Taq DNA Polymerase (Thermo Scientific). The thermal cycling protocol was as follows: an initial denaturation step of 5 min at 95 °C, followed by 35 cycles of 1 min at 95 °C, 1 min at 50 °C, and 1 min at 72 °C. A final step at 72 °C has been carried out for 10 min at the end of the 35th cycle. A negative control (no template) was also included in the analysis. PCR products were analysed in a 1% agarose gel.

In parallel, to check the occupancy of the SSV2-integration site in the *S. solfataricus* genome, a PCR analysis was carried out using the primers SSV2attApO1, SSV2attApO2 and SSV2attApO3 [25], whose sequences are listed in Table S1. The reactions were set up (as described above) except the annealing temperature was of 55 °C.

### 2.5. Semi-quantitative PCR analysis of the SSV2 DNA content

Cell pellets of the initial population InF1-SSV2 and of the isolated clones showing a lower SSV2 titre (clone 1, 1Q and 1Q<sub>1</sub>) were treated for total DNA extraction using the DNeasy tissue kit (Qiagen). To follow the variation of the SSV2 DNA content among the different clones, a semi-quantitative PCR analysis was performed. Two primer couples were designed using Primer3 software (available at the website: <http://bioinfo.ut.ee/primer3-0.4.0/>), in order to amplify: (i) a 108-bp region of the host single-copy gene *orc1* and (ii) a 238-bp region of the SSV2 single-copy gene *vp3* (Table S1). A master mix was prepared and PCRs were carried out as described elsewhere [34], using the following cycling protocol: an initial denaturation step of 5 min at 95 °C, followed by 30 cycles of 40 s at 95 °C, 40 s at 62 °C, and 1 min at 72 °C. For each reaction, tubes were taken from the thermocycler at the 20<sup>th</sup>, 25<sup>th</sup> and 30<sup>th</sup> cycle of amplification. PCR products were analysed by agarose gel

electrophoresis in  $1 \times$  TAE buffer pH 8.0 (40 mM Tris, 20 mM acetic acid and 1 mM EDTA).

### 3. Results and discussion

#### 3.1. Viral gene expression in the infected strains

Previous studies were carried out in the immediate aftermath of the viral infection [26], on lytic viruses [27–29] or upon induction of the viral replication [10], and very little is known about the viral/host gene expression in the carrier state. Two fuselloviruses, SSV1 and SSV2, were chosen as model in this study since they are not lytic and are therefore suitable for transcriptomic analysis in stably infected populations. Furthermore, the double-infected SSV1/SSV2-InF1 strain is a ternary system, which is suitable for analysing the mutual effect of the two viruses on each other as well as on the host gene expression.

To evaluate the effect of the viral infection on the host gene expression, *S. solfataricus* P2 (InF1) was infected to generate virus harbouring strains, i.e. SSV1-InF1, SSV2-InF1 and SSV1/SSV2-InF1. SSV1 and SSV2 show similar genome organization and gene synteny and, accordingly, SSV2 promoters were named after their SSV1 counterparts [9,26]. The only significant difference is the absence, in the SSV2 genome, of the SSV1 region involved in regulating the switch from the lysogenic to the UV-induced state. Microarray experiments were performed for all strains at two different growth phases, i.e. at 0.4 OD<sub>600nm</sub> (early-exponential) and 1.2 OD<sub>600nm</sub> (late-exponential), in order to define which viral genes were expressed. With this aim, microarray slides were co-hybridized with the total cDNA samples from:

- 1) uninfected InF1 (ref) vs either SSV1-, SSV2-, or SSV1/SSV2-InF1 (exp) collected at 0.4 OD<sub>600nm</sub>
- 2) uninfected InF1(ref) vs either SSV1-, SSV2-, or SSV1/SSV2-InF1 (exp) collected at 1.2 OD<sub>600nm</sub>

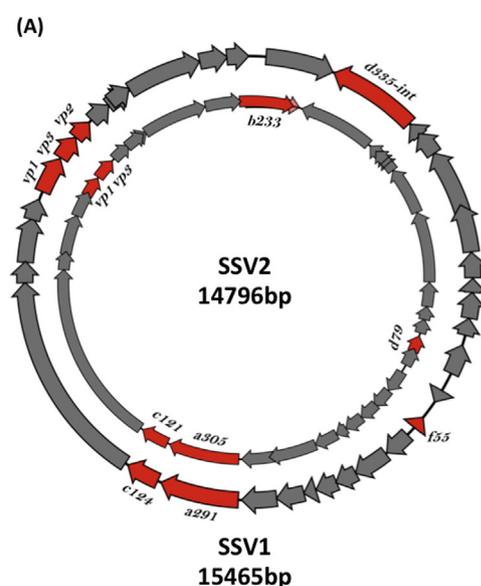
cDNAs were prepared from the uninfected and infected strains and used as reference and experimental samples, respectively.

SSV1 genes expressed at both growth phases analysed were those encoding: i) the structural proteins VP1, VP2 and VP3, ii) the integrase D335, iii) the transcription repressor F55 and iv) A291 and C124 (Fig. 1), for which virion docking/release and structural functions have been proposed, respectively [10,26]. Interestingly, most of viral genes expressed by the SSV2-InF1 strain were homologous to those expressed from SSV1, i.e.: VP1, VP3, A305 (SSV1-A291) and C121 (SSV1-C124) (Fig. 1). In addition, DnaA-like protein B233 and D79 were found expressed. A transcriptional activity was also detected by a probe matching the 3' non-coding region of *b233* gene, thus indicating that the mRNA of this gene carries a long untranslated tail like the homolog on the pSSVx genome [9,35] (Fig. 1). In the double-infected strain (SSV1/SSV2-InF1), viral genes expressed were those encoding: i) VP1 and VP3 from both SSV1 and SSV2, ii) the homologous proteins A291/A305 and C124/C121, iii) F55 from SSV1 and iv) B233 and D79 from SSV2 (Fig. 1). Interestingly, gene expression patterns of SSV1 and SSV2 are nearly identical, despite the fact that replication induction for these two viruses is triggered by different stimuli, i.e. the UV-light exposure and the host physiological/metabolic state, respectively. Therefore, this analysis defined a minimal set of genes required in the carrier state for the replication and packaging of both fuselloviruses (Fig. 1).

In order to evaluate variation of the expression levels of viral genes during the host growth, the following co-hybridizations were carried out:

- 1) SSV1-InF1 (ref) collected at 0.4 OD<sub>600nm</sub> vs SSV1-InF1 (exp) collected at 1.2 OD<sub>600nm</sub>
- 2) SSV2-InF1 (ref) collected at 0.4 OD<sub>600nm</sub> vs SSV2-InF1 (exp) collected at 1.2 OD<sub>600nm</sub>
- 3) SSV1/SSV2-InF1 (ref) collected at 0.4 OD<sub>600nm</sub> vs SSV1/SSV2-InF1 (exp) collected at 1.2 OD<sub>600nm</sub>

cDNAs from infected strains collected at 0.4 OD<sub>600nm</sub> and 1.2 OD<sub>600nm</sub> were used as reference and experimental samples, respectively. This analysis revealed a constitutive expression of all



(B)

Viral genes expressed				
SSV1-InF1	SSV2-InF1	SSV1/SSV2-InF1		Function
<i>SSV1-vp1</i>	<i>SSV2-vp1</i>	<i>SSV1-vp1</i>	<i>SSV2-vp1</i>	Structural
<i>SSV1-vp2</i>	<sup>a</sup> H.N.D.	<sup>c</sup> N.E.	<sup>a</sup> H.N.D.	Structural
<i>SSV1-vp3</i>	<i>SSV2-vp3</i>	<i>SSV1-vp3</i>	<i>SSV2-vp3</i>	Structural
<i>d335</i>	<sup>b</sup> H.N.E.	<sup>c</sup> N.E.	<sup>c</sup> N.E.	Integrase
<i>f55</i>	<sup>a</sup> H.N.D.	<i>f55</i>	<sup>a</sup> H.N.D.	Transcription repressor
<i>a291</i>	<i>a305</i>	<i>a291</i>	<i>a305</i>	Putative viral docking/release
<i>c124</i>	<i>c121</i>	<i>c124</i>	<i>c121</i>	Putative structural
<sup>b</sup> H.N.E.	<i>b233</i>	<sup>b</sup> H.N.E.	<i>b233</i>	DnaA-like protein
<sup>b</sup> H.N.E.	<i>d79</i>	<sup>b</sup> H.N.E.	<i>d79</i>	Unknown

<sup>a</sup> Homologue not detected, <sup>b</sup> Homologue not expressed, <sup>c</sup> No expression

**Fig. 1.** Schematic representation of the SSV1 and SSV2 genomes. (A) Red-filled arrows are viral genes expressed in the carrier state for both SSV1 (outer) and the SSV2 (inner) genomes. Clockwise- and anticlockwise-oriented arrow heads represent ORFs encoded by the plus and minus strand, respectively. (B) List of viral genes expressed in all strains analysed.

**Table 1**  
Number of up- and down-regulated genes and their functional distribution.

Functional group <sup>a</sup>	Gene expression							
	Up-regulated genes				Down-regulated genes			
	InF1	SSV1-InF1	SSV2-InF1	SSV1/SSV2-InF1	InF1	SSV1-InF1	SSV2-InF1	SSV1/SSV2-InF1
Amino acids biosynthesis	2	1	2	1	2	0	2	0
Cellular envelope and membrane	1	0	1	0	0	0	1	0
Cellular process	2	1	4	2	2	1	8	2
Central intermediary metabolism	0	2	0	2	0	0	2	0
Cofactor biosynthesis	0	0	0	0	0	0	2	1
Energy metabolism	10	1	19	0	1	0	12	3
Hypothetical protein <sup>b</sup>	16	3	10	1	6	4	16	3
IS elements	1	2	6	0	1	0	0	0
Lipid metabolism	2	0	3	1	0	0	0	0
Nucleotide metabolism	1	0	0	0	0	0	7	1
Protease and protein modification	0	0	0	0	1	0	2	0
Replication and repair	1	0	1	0	0	0	3	2
RNA	2	1	1	1	0	1	0	0
Transcription and regulation	3	1	11	2	5	2	7	2
Translation	0	0	0	0	7	4	14	0
Transport	17	4	7	1	1	1	5	2
Uncategorized/helicases	1	1	3	0	1	0	2	2
Cas genes	0	0	8	0	0	0	0	0
SSV2 genes <sup>c</sup>	0	0	0	0	0	0	1	1
STIV infection differentially regulated (Ortman et al., 2008)	1	0	1	0	0	0	1	0
Total	60	17	77	11	27	13	85	19

<sup>a</sup> Putative functions are derived from the website of the *Sulfolobus solfataricus* P2 complete sequencing project (<http://www-archbac.u-psud.fr/projects/sulfolobus/>).

<sup>b</sup> Hypothetical proteins with no putative function.

<sup>c</sup> The SSV2 down-regulated gene is *a305*.

the above-listed SSV1 and SSV2 genes, except for the SSV2 ORF *a305*, which was down-regulated (Table 1).

### 3.2. Gene expression analysis in the uninfected InF1 strain

To evaluate the effect of viral infection on host gene expression, we first identified up- and down-regulated genes during the growth of the uninfected InF1 strain. With this aim, microarray slides were co-hybridized with total cDNA samples as follows:

InF1 collected at 0.4 OD<sub>600nm</sub> (ref) vs InF1 collected at 1.2 OD<sub>600nm</sub> (exp)

A total of 60 up-regulated and 27 down-regulated genes were detected (Table 1). Among the up-regulated ones, expression of 38 genes increased by 2.0–4.9 folds, 13 showed a 5.0–9.9 fold-change and only for 5 genes expression was up-regulated between 10.0 and 33.0 folds. On the other hand, all down-regulated genes showed a fold of repression comprised in the range 2.0–4.9, except for *sso2574* whose transcription was down-regulated by 5.9 folds.

A remarkable number of the up-regulated *S. solfataricus* genes encode proteins belonging to three functional categories, i.e. hypothetical protein, transport and energy metabolism (Table 1). Transcriptional activation of genes involved in the transport of small metabolites is not surprising and it is likely due to the shortage of nutrients that occurs in a late stage of the growth. Moreover, it is interesting that a homolog of the gene encoding the bacitracin resistance protein (*sso1860*) is up-regulated, hinting to defence mechanisms acting against xenobiotic peptides produced by other member cells to face overpopulation. On the other hand, down-regulated genes mostly cluster into the categories of transcription and regulation as well as translation (Table 1). In particular, down-regulation of many genes encoding for ribosomal proteins is consistent with the reduction of protein synthesis and, in turn, with the decrease of cell growth rate. A

complete list of the differentially regulated genes, grouped in functional categories, is reported in the Table S2.

### 3.3. SSV1 infection has a slight effect on the host gene expression

In order to detect remodelling of the host gene expression in the SSV1-InF1 lysogenic strain, a microarray analysis was carried out by co-hybridizing cDNA samples as follows:

SSV1-InF1 collected at 0.4 OD<sub>600nm</sub> (ref) vs SSV1-InF1 collected at 1.2 OD<sub>600nm</sub> (exp)

To reveal host genes whose expression is altered as a consequence of the SSV1 infection, we excluded from the analysis those that were up- and down-regulated in the uninfected InF1 strain as a consequence of the physiological changes occurring in the passage from the early to the late exponential growth phase. Among 30 differentially expressed host genes, 17 were up- and 13 down-regulated throughout the growth of the SSV1-InF1 strain (Table 1). Changes in the expression level tended to be higher for up-regulated genes. The majority of the down-regulated ones decreased by 2–2.5-folds, whereas 5 genes were up-regulated with an increase from 2.8 to 4.3-fold (Table S3). Although these genes are scattered over several functional categories, a similar trend to that of the uninfected strain was observed. Indeed, genes encoding proteins involved in the transport and translation are up- and down-regulated, respectively.

These data show that the infection by SSV1 has only a limited effect on the host gene expression, suggesting a harmonic coexistence with the host in the carrier state. Accordingly, SSV1 is a temperate virus that self-regulates its gene expression in the carrier state through the pleiotropic effect of the transcription regulator F55 [11,16], which represses the expression of the UV-inducible as well as of the early genes.

### 3.4. Remodelling of the host gene expression induced by SSV2 infection includes the CRISPR-Cas system activation

To investigate the variation of host gene expression during the growth of the stably infected SSV2-InF1 strain, a microarray analysis was carried out by co-hybridizing cDNA samples as follows:

SSV2-InF1 collected at 0.4 OD<sub>600nm</sub> (ref) vs SSV2-InF1 collected at 1.2 OD<sub>600nm</sub> (exp)

As already described for the SSV1-InF1 strain, all variations of gene expression reported herein for the SSV2-InF1 strain have to be considered merely due to the SSV2 infection. A total of 162 host genes were differentially expressed throughout the growth of the SSV2-InF1 strain, with 77 and 85 genes that were found to be up- and down-regulated, respectively (Table 1). By comparing the total number of differentially regulated host genes in SSV1-InF1 (30 genes) to that of SSV2-InF1 (162 genes), it is evident that the remodelling occurring as consequence of the SSV2 infection is remarkable (Table 1). Expression variation is in general higher for up-regulated genes (between 2.0 and 7.9 folds) than for those down-regulated (between 2.0 and 6.0 folds). Moreover, expression of the majority of the latter decreased by 3.0-folds or less, whereas 34 up-regulated genes were differentially regulated of 4.0-folds or greater (Table S4). Intriguingly, although SSV2 is a not-lytic virus, its effect on the host gene expression is comparable to the lytic STIV and SIRV2 [27,28]. Indeed, 5% of the *S. solfataricus* genes were directly affected by the SSV2 infection, among which, those encoding for proteins belonging to the functional categories of energy metabolism as well as transcription and regulation tended to be up-regulated (Table 1). Furthermore, transcriptional induction of IS elements (Tab. S4) has already been reported after SIRV2 infection [29], as consequence of UV-light exposure [10] and heat shock [36], thus highlighting their involvement in the *S. solfataricus* stress response. On the other hand, down-regulated genes are mostly associated with categories of translation (Table 1 and Table S4). Moreover, repression of stress response genes encoding for the proteasome subunit (*sso0278*) and Bcp3 (*sso0225*) resembles the response of *S. solfataricus* to the infection by SIRV2, which is

able in this way to circumvent host defences [28]. Expression remodelling of genes involved in informational processing (i.e. transcription and translation) has been also observed after infection with STIV [27] and SIRV2 [28]. Accordingly, like the latter, SSV2 might take the control over the host informational machinery to replicate its own genome.

One of the most interesting outcomes from this analysis was the expression variation observed for a group of genes involved in CRISPR-Cas antiviral system, i.e.: i) the type IA interference cassette located downstream the locus C, which includes *sso1439* (*cas3''*), *sso1441* (*cas5*), *sso1442* (*cas7*) and *sso1443* (*cas5*); ii) *sso1424* (small subunit of CASCADE, CRISPR-associated Complex for Antiviral Defence) and *sso1425* (*cas3*) belonging to the type IIIA interference cassette localized downstream the former one; iii) *sso1389* (*csx1* putative transcription factor) localized upstream the locus A and iv) *sso1997* (*cas7* type IA) lying in the proximity of the locus F (Fig. 2, Table 2). In parallel, with the exception of the locus E, all CRISPR loci were found to be up-regulated during the growth of the SSV2-InF1 strain, with the highest expression induction observed for loci A, C and D (Table 2). Noteworthy, proteins encoded by the aforementioned up-regulated type IA interference cassette constitute the CASCADE effector complex of *S. solfataricus*, which is involved in the interference step of the defence system [31]. Interestingly, loci A, B, D and F of *S. solfataricus* P2 strain contain spacers matching fuselloviral genomes with the highest density observed at the leader-proximal region of the locus F [37]. Although this latter is a non-extending locus, due to the lack of a leader sequence, it can still be useful in challenging viral infection since it generates crRNAs. Accordingly, the aCASCADE (archaeal CASCADE) co-purified with crRNAs from all CRISPR loci [38] and probably exert its activity with all the crRNAs of *S. solfataricus*, including those from locus F [39]. Moreover, the expression of this locus was confirmed in the strain SSV2-InF1 by RT-PCR (data not shown).

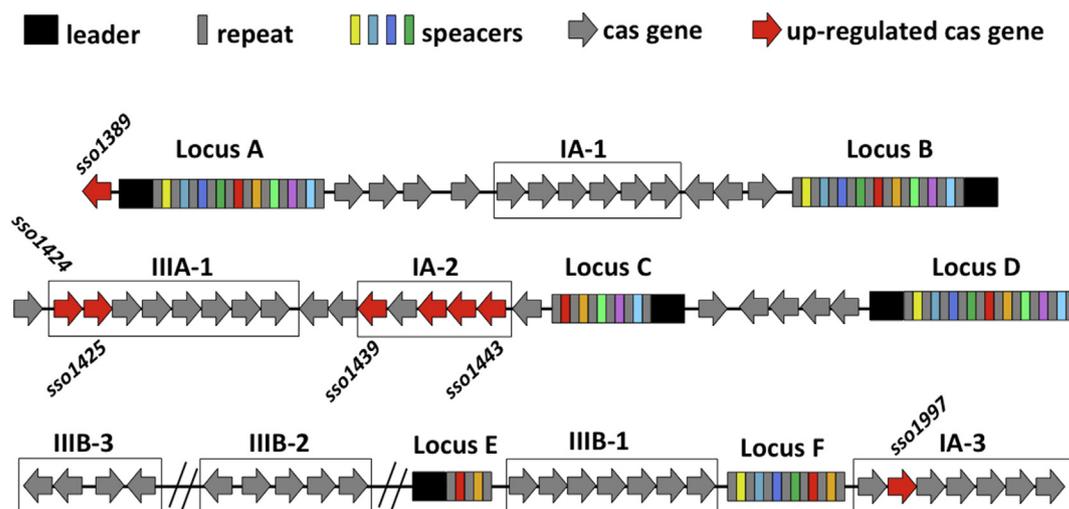


Fig. 2. CRISPR-Cas loci of *S. solfataricus* P2. The six CRISPR-Cas loci of the *S. solfataricus* P2 genome are schematized, where red-filled arrows are *cas* genes up-regulated in the SSV2-InF1 strain. Gene cassettes are black-framed and labelled.

### 3.5. The presence of SSV1 quenches the effect of SSV2 infection on the host gene expression

In order to evaluate variation of the host genes transcription caused by the co-presence of SSV1 and SSV2, hybridizations were carried out using cDNA samples as follows:

SSV1/SSV2-InF1 collected at 0.4 OD<sub>600nm</sub> (ref) vs SSV1/SSV2-InF1 collected at 1.2 OD<sub>600nm</sub> (exp)

As done before, data from the uninfected strain were used to filter those of the double-infected one, so that gene expression variation reported here is merely due to the presence of both viruses. As concerns viral genes, no significant variation was observed with the exception of the SSV2 gene *a305* that showed a down-regulation trend, similarly to that observed for the single-infected strain SSV2-InF1 (Table S5). On the other hand, among 30 differentially regulated host genes, 11 were up- and 19 down-

**Table 2**  
CRISPR loci and *cas* genes of *Sulfolobus solfataricus* P2 whose expression is affected by SSV2 infection.

Gene ID	Function
<i>Cas</i> gene product	
<i>sso1389</i>	<i>csx1</i> (CRISPR-Cas system-associated protein)
<i>sso1424</i>	<i>csa5</i> (CRISPR-Cas system-associated protein)
<i>sso1425</i>	<i>csm3</i> (CRISPR-associated RAMP)
<i>sso1439</i>	<i>cas''</i> (nuclease subunits of Cas3)
<i>sso1441</i>	<i>cas5</i> (CASCADE complex core)
<i>sso1442</i>	<i>cas7</i> (CRISPR-associated protein, Csa2)
<i>sso1443</i>	<i>csa5</i> (CRISPR-associated protein)
<i>sso1997</i>	CRISPR-associated auto-regulator, DevR family homolog
CRISPR locus	Probe position <sup>a</sup>
SSOLCTR-A	1708–1767
SSOLCTR-A	3373–3432
SSOLCTR-B	STIV
SSOLCTR-B	278–343
SSOLCTR-B	1493–1557
SSOLCTR-C	pNOB8
SSOLCTR-C	10–71
SSOLCTR-C	1916–1980
SSOLCTR-D	ATV
SSOLCTR-D	105–174
SSOLCTR-D	483–551
SSOLCTR-D	5568–5630
SSOLCTR-F	3689–3753
SSOLCTR-F	4350–4419

<sup>a</sup> Spacer position in each locus. STIV, pNOB8 and ATV are probes matching spacers against these genetic elements.

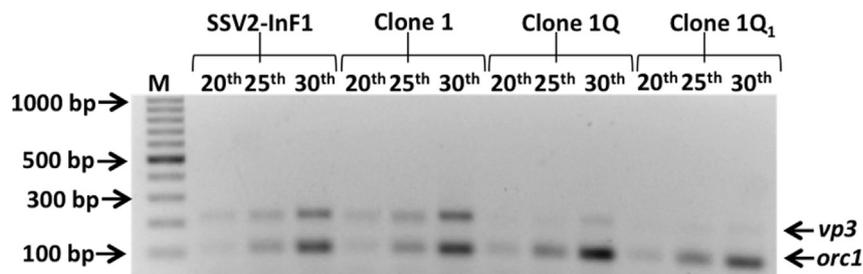
regulated during the growth of the double-infected strain (Table 1). Therefore, the presence of SSV1 in the same host cells of SSV2 quenched the gene expression variation from 162 (being differentially regulated in the single-infected SSV2-InF1 strain) to only 30 genes. Surprisingly, genes and clusters of the CRISPR-Cas system were not induced in double-infected cells. Instead, the

expression pattern of the viral genes is overall identical to that displayed by the single-infected SSV1- and SSV2-InF1 strains. The two fuselloviruses show syntenic genomes, with the only significant difference in the region responsible for the UV-inducibility of SSV1, which is lacking in the SSV2 genome [9]. Therefore, it is tempting to speculate that the SSV1 F55 transcription factor, encoded by this region, interferes with the remodelling of the host gene expression induced by SSV2.

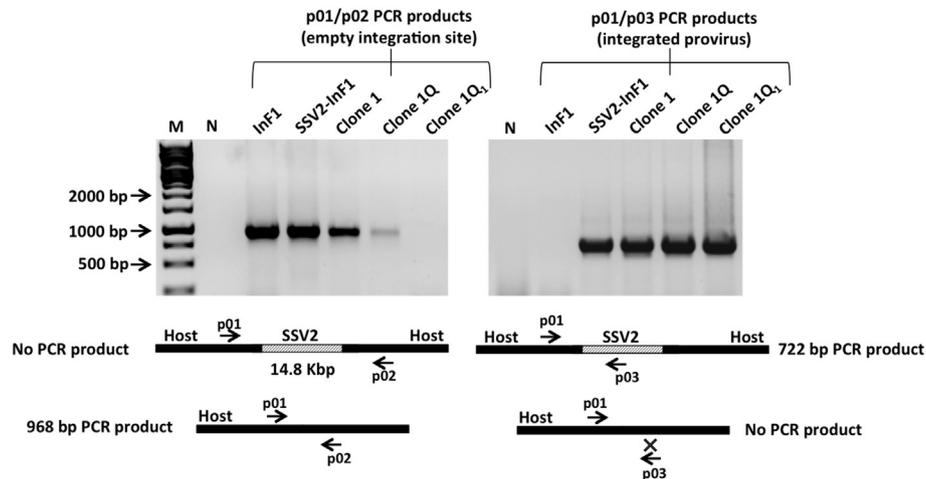
### 3.6. Isolation of single clones from the CRISPR-Cas responsive SSV2-InF1 strain

To study if the activation of the CRISPR-Cas system enabled the cells to get rid of SSV2, we performed a screening of the SSV2-InF1 strain to isolate cured cells. Therefore, clones showing a decreased viral titre, compared to the SSV2-InF1 initial population, were further characterized at DNA level to evaluate SSV2 content through semi-quantitative PCR. Results are shown for clones 1, 1Q and 1Q<sub>1</sub> that are representative of the first, second and third round of selection, respectively. The PCR products were analysed for each sample on agarose gel at the 20<sup>th</sup>, 25<sup>th</sup> and 30<sup>th</sup> cycle of amplification and the initial population SSV2-InF1 was used as reference (Fig. 3). Although the intensity of the *orc1* signal is nearly identical when the same amplification cycle is considered, the signal relative to the viral gene *vp3* progressively decreases until it becomes hardly visible for the clone 1Q<sub>1</sub> (Fig. 3). Therefore, the low viral titre observed for this latter clone (two orders of magnitude less than the SSV2-InF1 initial population), is mirrored by a comparable drop of the intracellular SSV2 DNA content (Fig. 3).

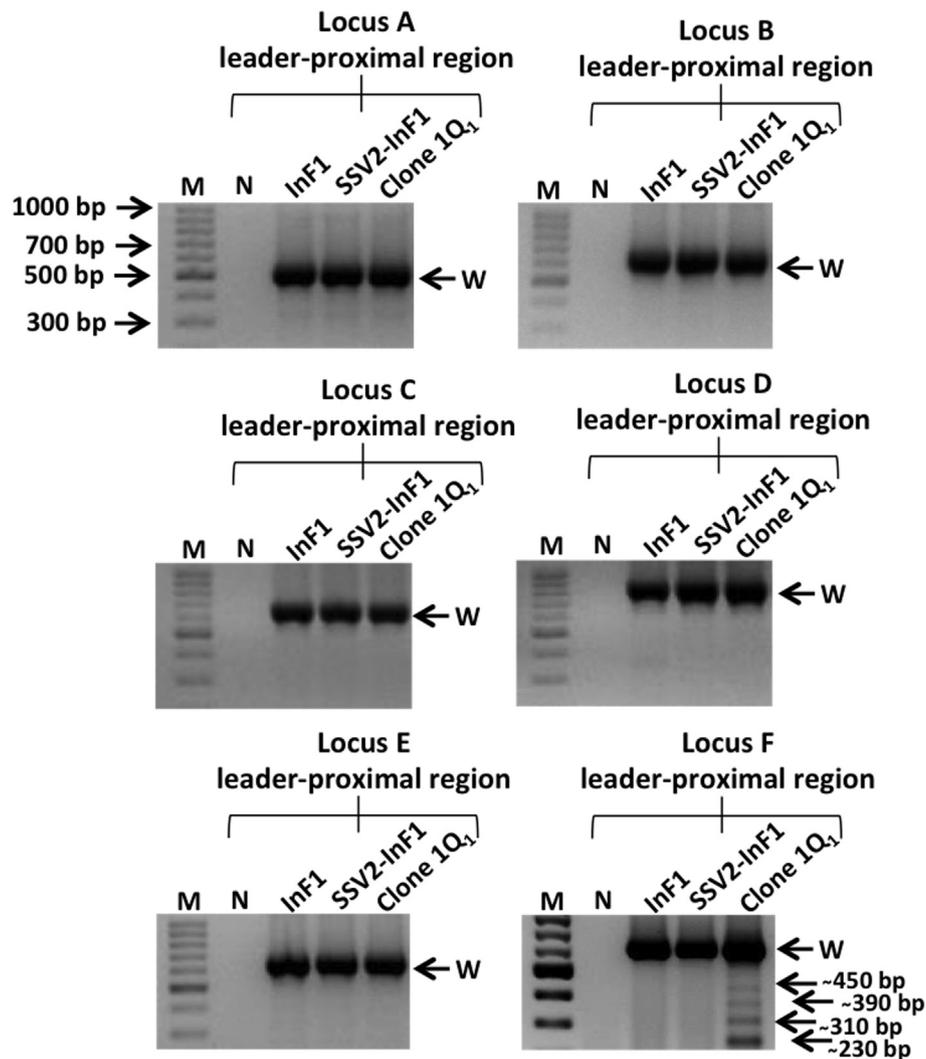
In parallel, a PCR analysis of the SSV2 integration site was carried out for the same clones using the primers SSV2attApO1, SSV2attApO2 and SSV2attApO3 [25]. The primer couple pO1–pO2 allows the amplification of an approximately 968 bp fragment if the integration site is empty (no provirus). On the other hand, the couple ApO1–ApO3 leads to the amplification of a 722 bp product only when the integration site is occupied by SSV2 (Fig. 4). Whereas



**Fig. 3.** SSV2 DNA detection by semi-quantitative PCR. Black-straight arrows point out to molecular size markers as well as to host (*orc1*) and viral (*vp3*) PCR products. Total DNA samples were prepared from the initial population (SSV2-InF1) as well as from the isolated clones (1, 1Q and 1Q<sub>1</sub>). The decreased intensity of the *vp3* signals indicates that the SSV2 copy number dropped throughout the isolation procedure.



**Fig. 4.** PCR analysis of the SSV2 integration site occupancy. Black-straight arrows point out to molecular size markers. To check the occupancy of the SSV2 integration site, PCR products were amplified using the p01/p02 (968 bp, free integration site) and p01/p03 (722 bp, integrated provirus) primer couples from the SSV2-InfF1 as well as from the isolated clones (1, 1Q and 1Q<sub>1</sub>). As shown in the right panel, all infected clones carry the provirus. However, provirus-free cells still persist in the cultures (left panel) although their amount decreases throughout the isolation procedure. Eventually, no empty integration sites were detected in the 1Q<sub>1</sub> clone.



**Fig. 5.** Leader-proximal region amplification of the *S. solfataricus* P2 CRISPR array. PCRs were performed to amplify leader-proximal regions of all six CRISPR loci (A–F) from the virus-free strain (InfF1), SSV2-infected initial population (SSV2-InfF1) and the 1Q<sub>1</sub> clone. The wild-type PCR product (W) has been amplified for all loci, thus indicating the lack of spacer acquisition (no adaptation). Intriguingly, shorter PCR products were amplified from the locus F of the clone 1Q<sub>1</sub>. Sequencing of these fragments confirmed loss of repeat/spacer units from this locus (see Fig. 6). M and N stay for marker and negative control, respectively.

amplification of the product pO1–pO2 for the infected strain SSV2-InF1 as well as for the clone 1 and 1Q indicates that a fraction of the cell population does not carry the provirus, the absence of such amplicon in the 1Q<sub>1</sub> clone suggests that virtually all cells carry the integrated SSV2 in the attachment site. The amplification of the product pO1–pO3 further confirms the presence of the provirus (Fig. 4). Since the excision of the provirus strictly depends on the presence of active integrase/excision enzymes, encoded by the episomal copies [40], degradation of the latter by the host CRISPR-Cas system leads to the entrapment of the provirus (Fig. 4), as already hypothesized elsewhere [41].

### 3.7. SSV2 infection causes deletions at the leader-proximal region of the locus F

A fundamental step of the CRISPR-Cas system is the adaptation, which consists in the Cas proteins-mediate incorporation of new spacers from foreign DNAs. This occurs immediately downstream the leader sequence of a given CRISPR locus upon viral infection. In order to detect the integration of new spacers, leader-proximal regions of all *S. solfataricus* loci were amplified (Fig. 5). In particular, primer couples were chosen as described elsewhere [33], so that integration of new spacers would have produced longer PCR products when compared with a control sample (uninfected InF1).

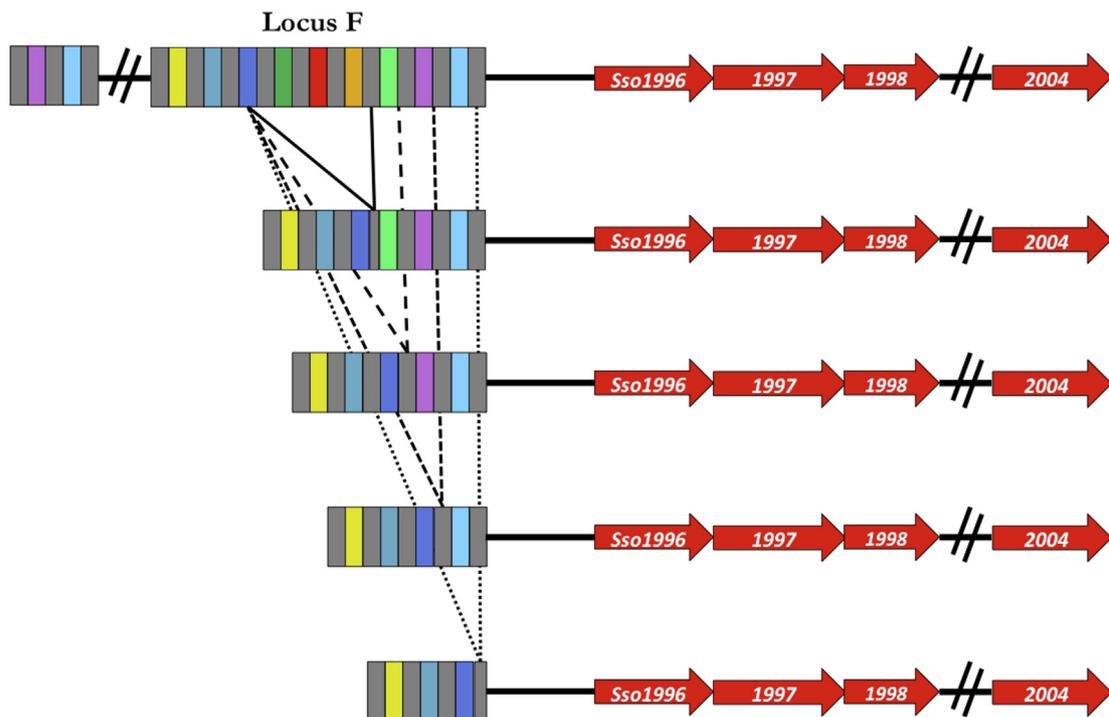
Our PCR analysis detected no differences in the length of the amplified products (loci A–E) among the initial population SSV2-InF1 and the isolated clones (Fig. 5), thus confirming that no adaptation occurred upon infection with SSV2 at these loci. Nevertheless, smaller PCR products (between 230 and 450 bp) were found when the leader-proximal region of the CRISPR locus F was amplified from the 1Q<sub>1</sub> DNA (Fig. 5). In order to shed light on the identity of these amplicons, PCR products were run on an agarose gel and single bands were cut off, purified and sequenced. Surprisingly, this analysis showed that progressive deletions

occurred at the leader-proximal region of the locus F and up to six repeat-spacer units were found to be lacking (Fig. 6).

In this regard, it is noteworthy that Stern and co-workers have hypothesized that the incorporation of self-targeting spacers might lead to autoimmunity and select for the loss of CRISPR functions [42]. Moreover, it has been demonstrated that when *S. solfataricus* is transformed with a plasmid that carries a gene essential for the host and which is also a target of the CRISPR system, surviving transformants show mutations that eliminate the plasmid-targeting spacer [32]. What would happen if the invading DNA, such as a virus, could not be destroyed because it managed to integrate into the host genome, thus becoming 'self'? In the hypothesis that the host is challenging the SSV2 infection (as shown by microarray data), it is likely that the CRISPR-Cas system targets also the integrated viral genome (i.e., the provirus). This event casts a huge fitness cost on cells harbouring self-targeting spacers [43]. As a consequence, the selective pressure drives the inactivation of the CRISPR-Cas system to ensure the survival of SSV2-lysogens. Our results indicate that this is achieved through deletion of spacer units responsible for the recognition of the SSV2 genome (Figs. 5 and 6).

## 4. Conclusions

Analysis of both viral and host transcriptomes upon SSV1 and SSV2 infection in *S. solfataricus* has revealed new insights into archaeal host–virus interactions. Previous studies were carried out to detect variation of viral and host gene expression in the early phases of the SSV2 infection [26], upon UV-irradiation for SSV1 [10] or on lytic viruses [27–29]. Therefore, the dissection of the viral transcription patterns during the carrier state has never been performed. Herein, by exploiting stably infected cells in the absence of any stimuli, we demonstrate not only that the host response towards the two viruses is extremely different, but that it is also



**Fig. 6.** Schematic illustration of repeat/spacer units loss from the locus F. The array is reported as grey rectangles (repeats) interspersed by colourful rectangles (spacers). A short intergenic region separates the array from the opposite-oriented *cas* genes cassette (IA3, see also Fig. 2). The four different loci F with deletions are schematised to show the loss of 2–6 repeat/spacer units in the clone 1Q<sub>1</sub>.

influenced by the co-presence of both fuselloviruses (Table 1). In particular, whereas *S. solfataricus* establishes a harmonious coexistence with SSV1, the reaction against the infection by SSV2 is remarkable and includes the transcriptional activation of CRISPR loci and *cas* genes (Table 2). This result is consistent with the isolation of single clones showing a low SSV2 copy number (Fig. 3). Our study highlights an interesting aspect of the host–virus interaction in the frame of the CRISPR response, i.e. host cells containing an integrated provirus (Fig. 4) are forced to develop a survival strategy in order to avoid self-attack by the CRISPR–Cas system. In fact, since the provirus is identical in sequence to the episomal copies, it can be targeted by the CRISPR–Cas system as well, thus casting a fitness cost on cells that are actively challenging the SSV2 infection. Our data show, for the first time, that *S. solfataricus* cells developed a specific strategy to safeguard host genome integrity, i.e. throughout deletion of self-targeting spacers (Figs. 5 and 6). How deletion of specific spacers at CRISPR loci occurs is matter of further investigation and it is expected to reveal completely novel molecular components and/or mechanisms of the CRISPR–Cas system.

Another relevant aspect highlighted by this study is the absence of the CRISPR–Cas response in the double-infected SSV1/SSV2–InF1 strain (Table 1 and Table S5). The two viral genomes are overall identical with the only significant difference being the absence in the SSV2 genome of a region responsible for the UV-induction in SSV1 (Fig. 1). Therefore, it is tempting to speculate that transcription factors and/or other molecular components encoded by this SSV1 region are responsible for silencing the CRISPR–Cas response in the double-infected strain. A combination of biochemical and genetic approaches is needed to move toward a better understanding of the mutual influence of the two fuselloviruses in this virus–host system.

### Conflict of interest

No conflict of interest.

### Acknowledgements

We thank the *Sulfolobus* Gene Chip Consortium for constructing the microarrays used for our experiments. This study was supported by grants from Regione Campania, legge 5 (CUP number E69D15000210002) as well as from the Danish Independent Research Council (DFF-0602-02196B, DFF1323-00330 and 11-106683), the Ministero dell'Istruzione, dell'Università e della Ricerca Scientifica (Progetti di Ricerca di Interesse Nazionale E61J1000020001) and the Carlsberg Foundation (2013\_01\_0060). Dr. Donata Cafasso is also acknowledged for performing DNA sequencing of this work.

### Appendix A. Supplementary material

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.biochi.2015.04.006>.

### References

- [1] C.A. Suttle, Marine viruses – major players in the global ecosystem, *Nat. Rev. Microbiol.* 5 (2007) 801–812.
- [2] A.M. Comeau, G.F. Hatfull, H.M. Krisch, D. Lindell, N.H. Mann, D. Prangishvili, Exploring the prokaryotic virosphere, *Res. Microbiol.* 159 (2008) 306–313.
- [3] M.K. Pietilä, T.A. Demina, N.S. Atanasova, H.M. Oksanen, D.H. Bamford, Archaeal viruses and bacteriophages: comparisons and contrasts, *Trends Microbiol.* 22 (2007) 334–344.
- [4] D. Prangishvili, The wonderful world of archaeal viruses, *Annu. Rev. Microbiol.* 8 (2013) 565–585.
- [5] H. Wang, N. Peng, S.A. Shah, L. Huang, Q. She, Archaeal extrachromosomal genetic elements, *Microbiol. Mol. Biol. Rev.* 79 (2015) 117–152.
- [6] M. Pina, A. Bize, P. Forterre, D. Prangishvili, The archeoviruses, *FEMS Microbiol. Rev.* 35 (2011) 1035–1054.
- [7] X. Peng, R.A. Garrett, Q. She, Archaeal viruses – novel, diverse and enigmatic, *Sci. China Life Sci.* 55 (2012) 422–433.
- [8] M. Krupovic, E.R. Quemin, D.H. Bamford, P. Forterre, D. Prangishvili, Unification of the globally distributed spindle-shaped viruses of the *Archaea*, *J. Virol.* 88 (2013) 2354–2358.
- [9] P. Contursi, S. Fusco, R. Cannio, Q. She, Molecular biology of fuselloviruses and their satellites, *Extremophiles* 3 (2014) 473–489.
- [10] S. Fröls, P.M. Gordon, M.A. Panlilio, C. Schleper, C.W. Sensen, Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays, *Virology* 365 (2007) 48–59.
- [11] S. Fusco, Q. She, S. Bartolucci, P. Contursi,  $T_{lys}$ , 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 (2013) 5926–5936.
- [12] W.D. Reiter, P. Palm, W. Zillig, Analysis of transcription in the archaeobacterium *Sulfolobus* indicates that archaeobacterial promoters are homologous to eukaryotic pol II promoters, *Nucleic Acids Res.* 16 (1988) 1–19.
- [13] W.D. Reiter, P. Palm, W. Zillig, Transcription termination in the archaeobacterium *Sulfolobus*: signal structures and linkage to transcription initiation, *Nucleic Acids Res.* 16 (1988) 2445–2459.
- [14] M. Nadal, G. Mirambeau, P. Forterre, W.D. Reiter, M. Duguet, Positively supercoiled DNA in a virus-like particle of an archaeobacterium, *Nature* 321 (1986) 256–258.
- [15] C. Schleper, K. Kubo, W. Zillig, 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 (1992) 7645–7649.
- [16] S. Fusco, Q. She, G. Fiorentino, S. Bartolucci, P. Contursi, Unravelling the role of the F55 regulator in the transition from lysogeny to UV induction of *Sulfolobus* spindle-shaped virus, *J. Virol.* 89 (2015) [in press].
- [17] P. Kraft, D. Kümmel, A. Oeckinghaus, G.H. Gauss, B. Wiedenheft, M. Young, C.M. Lawrence, 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 (2004) 7438–7442.
- [18] P. Kraft, A. Oeckinghaus, D. Kümmel, G.H. Gauss, J. Gilmore, B. Wiedenheft, M. Young, C.M. Lawrence, Crystal structure of F-93 from *Sulfolobus* spindle-shaped virus 1, a winged-helix DNA binding protein, *J. Virol.* 78 (2004) 11544–11550.
- [19] S.K. Menon, W.S. Maaty, G.J. Corn, S.C. Kwok, B.J. Eilers, P. Kraft, E. Gillitzer, M.J. Young, B. Bothner, C.M. Lawrence, Cysteine usage in *Sulfolobus* spindle-shaped virus 1 and extension to hyperthermophilic viruses in general, *Virology* 376 (2008) 270–278.
- [20] F. Guillière, N. Peixeiro, A. Kessler, B. Raynal, N. Desnoues, J. Keller, M. Delepierre, D. Prangishvili, G. Sezonov, J.I. Guijarro, Structure, function, and targets of the transcriptional regulator SvtR from the hyperthermophilic archaeal virus SIRV1, *J. Biol. Chem.* 284 (2009) 22222–22237.
- [21] P. Contursi, K. D'Ambrosio, L. Pirone, E. Pedone, T. Aucelli, Q. She, G. De Simone, S. Bartolucci, C68 from the *Sulfolobus islandicus* plasmid–virus pSSVx is a novel member of the AbrB-like transcription factor family, *Biochem. J.* 435 (2011) 157–166.
- [22] C. Schlenker, A. Goel, B.P. Tripet, S. Menon, T. Willi, M. Dlakić, M.J. Young, C.M. Lawrence, V. Copié, Structural studies of E73 from a hyperthermophilic archaeal virus identify the “RH3” domain, an elaborated ribbon–helix–helix motif involved in DNA recognition, *Biochemistry* 51 (2012) 2899–2910.
- [23] P. Contursi, S. Fusco, D. Limauro, G. Fiorentino, Host and viral transcriptional regulators in *Sulfolobus*: an overview, *Extremophiles* 17 (2013) 881–895.
- [24] P. Contursi, B. Farina, L. Pirone, S. Fusco, L. Russo, S. Bartolucci, R. Fattorusso, E. Pedone, 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.* 42 (2014) 5993–6011.
- [25] P. Contursi, S. Jensen, T. Aucelli, M. Rossi, S. Bartolucci, Q. She, Characterization of the *Sulfolobus* host–SSV2 virus interaction, *Extremophiles* 10 (2006) 615–627.
- [26] Y. Ren, Q. She, L. Huang, Transcriptomic analysis of the SSV2 infection of *Sulfolobus solfataricus* with and without the integrative plasmid pSSVi, *Virology* 441 (2013) 126–134.
- [27] A.C. Ortmann, S.K. Brumfield, J. Walther, K. McInerney, S.J. Brouns, H.J. van de Werken, B. Bothner, T. Douglas, J. van de Oost, M.J. Young, Transcriptome analysis of infection of the archaeon *Sulfolobus solfataricus* with *Sulfolobus* turreted icosahedral virus, *J. Virol.* 82 (2008) 4874–4883.
- [28] E. Okutan, L. Deng, S. Mirlashari, K. Uldahl, M. Halim, C. Liu, R.A. Garrett, Q. She, X. Peng, Novel insights into gene regulation of the rudivirus SIRV2 infecting *Sulfolobus* cells, *RNA Biol.* 10 (2013) 875–885.
- [29] T.E. Quax, M. Voet, O. Sismeiro, M.A. Dillies, B. Jagla, J.Y. Coppée, G. Sezonov, P. Forterre, J. van der Oost, R. Lavigne, D. Prangishvili, Massive activation of archaeal defense genes during viral infection, *J. Virol.* 87 (2013) 8419–8428.
- [30] R.M. Ceballos, C.D. Marceau, J.O. Marceau, S. Morris, A.J. Clore, K.M. Stedman, Differential virus host-ranges of the *Fuselloviridae* of hyperthermophilic *Archaea*: implications for evolution in extreme environments, *Front. Microbiol.* 3 (2012) 295.

- [31] J. Zhang, M.F. White, Hot and crispy: CRISPR-Cas systems in the hyperthermophile *Sulfolobus solfataricus*, *Biochem. Soc. Trans.* 41 (2013) 1422–1426.
- [32] S. Gudbergdottir, L. Deng, Z. Chen, J.V.K. Jensen, L.R. Jensen, Q. She, R.A. Garrett, Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers, *Mol. Microbiol.* 79 (2011) 35–49.
- [33] S. Erdmann, R.A. Garrett, Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms, *Mol. Microbiol.* 85 (2012) 1044–1056.
- [34] S. Fusco, M. Aulitto, S. Bartolucci, P. Contursi, A standardized protocol for the UV induction of *Sulfolobus* spindle-shaped virus 1, *Extremophiles* 19 (2015) 539–546.
- [35] P. Contursi, R. Cannio, Q. She, Transcription termination in the plasmid/virus hybrid pSSVx from *Sulfolobus islandicus*, *Extremophiles* 14 (2010) 453–463.
- [36] C.R. Cooper, A.J. Daugherty, S. Tachdjian, P.H. Blum, R.M. Kelly, Role of vapBC toxin–antitoxin loci in the thermal stress response of *Sulfolobus solfataricus*, *Biochem. Soc. Trans.* 37 (2009) 123–126.
- [37] R.K. Lillestøl, S.A. Shah, K. Brügger, P. Redder, H. Phan, J. Christiansen, R.A. Garrett, CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties, *Mol. Microbiol.* 72 (2009) 259–272.
- [38] N.G. Lintner, M. Kerou, S.K. Brumfield, S. Graham, H. Liu, J.H. Naismith, M. Sdano, N. Peng, Q. She, V. Copié, M.J. Young, M.F. White, C.M. Lawrence, Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE), *J. Biol. Chem.* 286 (2011) 21643–21656.
- [39] A. Manica, C. Schleper, CRISPR-mediated defense mechanisms in the hyperthermophilic archaeal genus *Sulfolobus*, *RNA Biol.* 10 (2013) 671–678.
- [40] Q. She, X. Peng, W. Zillig, R.A. Garrett, Gene capture events in archaeal chromosomes, *Nature* 409 (2001) 478.
- [41] R.A. Garrett, S.A. Shah, G. Vestergaard, L. Deng, S. Gudbergdottir, C.S. Kenchappa, S. Erdmann, Q. She, CRISPR-based immune systems of the *Sulfolobales*: complexity and diversity, *Biochem. Soc. Trans.* 39 (2011) 51–57.
- [42] A. Stern, L. Keren, O. Wurtzel, G. Amitai, R. Sorek, Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* 26 (2010) 335–340.
- [43] M. Dyal-Smith, Dangerous weapons: a cautionary tale of CRISPR defence, *Mol. Microbiol.* 79 (2011) 3–6.