Molecular Microbiology (2014) 91(5), 900–917 ■
doi:10.1111/mmi.12503
First published online 17 January 2014

Inter-viral conflicts that exploit host CRISPR immune systems of *Sulfolobus*

Susanne Erdmann, Sven Le Moine Bauer† and Roger A. Garrett*
Archaea Centre, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 N Copenhagen, Denmark.

Summary

Infection of *Sulfolobus islandicus* REY15A with mixtures of different *Sulfolobus* viruses, including STSV2, did not induce spacer acquisition by the host CRISPR immune system. However, coinfection with the tailed fusiform viruses SMV1 and STSV2 generated hyperactive spacer acquisition in both CRISPR loci, exclusively from STSV2, with the resultant loss of STSV2 but not SMV1. SMV1 was shown to activate adaptation while itself being resistant to CRISPR-mediated adaptation and DNA interference. Exceptionally, a single clone S-1 isolated from an SMV1 + STSV2-infected culture, that carried STSV2-specific spacers and had lost STSV2 but not SMV1, acquired spacers from SMV1. This effect was also reproducible on reinfecting wild-type host cells with a variant SMV1 isolated from the S-1 culture. The SMV1 variant lacked a virion protein ORF114 that was shown to bind DNA. This study also provided evidence for: (i) limits on the maximum sizes of CRISPR loci; (ii) spacer uptake strongly retarding growth of infected cultures; (iii) protospacer selection being essentially random and non-directional, and (iv) the reversible uptake of spacers from STSV2 and SMV1. A hypothesis is presented to explain the interactive conflicts between SMV1 and the host CRISPR immune system.

Introduction

Clustered regularly interspaced short palindromic repeats (CRISPR) provide the basis for diverse adaptive immune systems common to most archaea and many bacteria. The spacer regions derive from invading genetic elements and are inserted into CRISPR loci commonly, but not invariably, adjacent to the CRISPR leader. Generally transcripts are initiated within the leader and are processed within repeats to yield small crRNAs, carrying most or all of the individual spacer sequences, which then act as guide RNAs for diverse interference complexes that target and cleave DNA or RNA (reviewed in Garrett *et al.*, 2011; Makarova *et al.*, 2011; Terns and Terns, 2011; Wiedenheft *et al.*, 2012; Barrangou and van der Oost, 2013; Zhang *et al.*, 2013).

Although considerable progress has been made in determining the different mechanisms of CRISPR RNA processing and maturation, and in elucidating the structures and targeting modes of some of the different interference complexes, our understanding of the molecular mechanisms involved in the adaptation process remains minimal. Adaptation involves the selection of protospacers on invading genetic elements and their insertion into CRISPR loci and it appears to be the most conserved stage of the adaptive immune response. It was first observed in the laboratory for the bacterial-specific subtype II-A system of *Streptococcus thermophilus* (Barrangou *et al.*, 2007; Deveau *et al.*, 2008). Three proteins Cas1, Cas2 and generally Cas4 have been implicated in this process, although in the streamlined subtype I-E and I-F systems of *Escherichia coli* lack Cas4, as do subtype II-A and some subtype III-A systems (Garrett *et al.*, 2011; Makarova *et al.*, 2011). In the *E. coli* systems, the interference endonuclease Cas3 may replace Cas4, a supposition that is supported by the coexpression of Cas1, Cas2 and Cas3 in these systems, and by the interaction of Cas1 with a Cas2-Cas3 hybrid protein in the subtype I-F system of *Pectobacterium atrosepticum* (Richter *et al.*, 2012). Cas1, the largest and most conserved protein, exhibits DNA endonuclease activity (Wiedenheft *et al.*, 2009; Babu *et al.*, 2011) and its mutation in the *E. coli* subtype I-E system can inhibit spacer acquisition (Yosef *et al.*, 2012). Cas2 protein from *Bacillus halodurans* also exhibits dsDNA endonuclease activity (Nam *et al.*, 2012) while another Cas2 protein, from *Sulfolobus solfataricus* and other archaea, shows low specificity ssRNA endonuclease activity (Beloglazova *et al.*, 2008). Cas4 of *Sulfolobus solfataricus* carries 5′- to 3′-DNA exonuclease activity that may generate recombigenic 3′-overlaps for CRISPR spacer insertion (Zhang *et al.*, 2012).

Accepted 20 December, 2013. *For correspondence. E-mail garrett@bio.ku.dk; Tel. (+45) 3532 2010; Fax (+45) 3532 2128. †Present address: Centre for Geobiology, Allégt. 41, University of Bergen, NO 5020 Bergen, Norway.

© 2013 The Authors. *Molecular Microbiology* published by John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

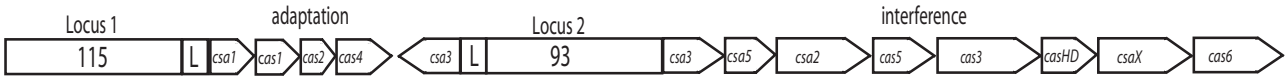

Fig. 1. Schematic representation of the CRISPR loci 1 and 2 of *S. islandicus* and the associated *cas* genes. The gene clusters implicated in spacer acquisition (adaptation) and in DNA interference are marked. The number of spacer-repeat units is given for each CRISPR locus and L denotes the leader regions.

Leaders were originally implicated, indirectly, in spacer acquisition because, for *S. solfataricus* strains, CRISPR loci lacking leaders did not accrue new spacers (Lillestøl *et al.*, 2006; 2009). Moreover, neighbour-joining trees provided support for the coevolution of Cas1 proteins, leaders and CRISPR repeats, for members of the Sulfolobales, consistent with their cofunctionality in spacer acquisition (Shah *et al.*, 2009; Shah and Garrett, 2011). More recently, experimental evidence has demonstrated that at least part of the 43 bp of the leader adjacent to the first CRISPR repeat is essential for acquisition activity in an *E. coli* subtype I-E system (Yosef *et al.*, 2012; Díez-Villaseñor *et al.*, 2013).

In subtype I-A systems of the Sulfolobales initial protospacer selection is highly specific, occurring one base pair after the PAM sequence. However, recognition at the other end of the protospacer shows no detectable sequence specificity and, moreover, multiple recognition sites can occur over up to six base pairs for a given protospacer region in different copies of the same genetic element (Erdmann and Garrett, 2012). These, and earlier results on *Sulfolobus* species (Lillestøl *et al.*, 2009) contrast with the recent finding of a potential motif downstream from the protospacer in *E. coli* (Yosef *et al.*, 2013). The *S. solfataricus* spacer acquisition results led to the proposal that a ruler mechanism operates for protospacer excision measured from the PAM sequence and this hypothesis was reinforced by the observation that many protospacers contain internal PAM sequences that can also be recognized, independently, during the putative protospacer excision from other intracellular copies of the same genetic element (Erdmann and Garrett, 2012). Experimental evidence for a second ruler mechanism operating during spacer insertion into CRISPR loci was provided for the *E. coli* subtype I-E system whereby initial cleavage occurs at the leader-repeat junction with a secondary cut occurring at the other end of the first repeat (Díez-Villaseñor *et al.*, 2013). These authors propose, further, that the occurrence of a dual ruler strategy during the two main acquisition steps, protospacer excision and spacer insertion, could ensure maintenance of the regular periodicity within CRISPR loci.

Members of the Sulfolobales are hosts for a wide range of viruses which exhibit greater morphological diversity than those of bacteria (Prangishvili *et al.*, 2006a; Pina *et al.*, 2011) and, moreover, single environmental isolates often carry different viruses propagating stably within cells (Prangishvili and Garrett, 2005). In addition, individual species generally carry large and multiple CRISPR loci and mixtures of type I and different type III interference systems, some of which remain to be characterized functionally (Garrett *et al.*, 2011; Zhang *et al.*, 2012; Deng *et al.*, 2013). The CRISPR spacers reflect this diversity by showing multiple predicted sequence matches to characterized viruses of the Sulfolobales exhibiting both linear and circular DNA genomes (Lillestøl *et al.*, 2006; 2009; Shah *et al.*, 2009; 2011; Shah and Garrett, 2011).

Earlier, we showed that on infecting *S. solfataricus* P2 cells with an environmental virus mixture, sampled from Yellowstone National Park, USA, that contained a single-tailed fusiform *Sulfolobus* virus and a minor conjugative plasmid component, later labelled *Sulfolobus* monocaudavirus SMV1 and pMGB1 respectively (Erdmann *et al.*, 2013), hyperactive spacer acquisition occurred exclusively from the plasmid (Erdmann and Garrett, 2012). Moreover the acquisition occurred in different CRISPR loci by two distinct mechanisms; in two loci, spacers were inserted adjacent to the leaders and in a third locus, spacers were incorporated throughout a single CRISPR array (Erdmann and Garrett, 2012). Here we examine further these unusual effects of SMV1 on different *Sulfolobus* viruses and the exceptional influence of SMV1 on CRISPR immune systems of *S. islandicus* REY15A for which a range of genetic tools are now available (Deng *et al.*, 2009; Peng *et al.*, 2009).

Results

Infection of an S. islandicus *culture with the tailed-fusiform virus SMV1 leads to cell death*

S. islandicus REY15A carries two CRISPR loci containing 115 and 93 spacer-repeat units, a single gene cassette encoding proteins Cas1, Cas2 and Cas4, implicated in spacer acquisition, a single gene for the CRISPR RNA processing enzyme Cas6, and one *cas* gene cassette for a subtype I-A DNA interference system, as well as two *cmr* gene cassettes encoding subtype III-B interference complexes (Guo *et al.*, 2011; Deng *et al.*, 2013). The CRISPR loci are coupled to the subtype I-A *cas* genes and are illustrated schematically in Fig. 1.

S. islandicus was tested initially for infectivity with an enriched virus mixture from Yellowstone National Park containing the tailed-fusiform virus SMV1 as a major com-

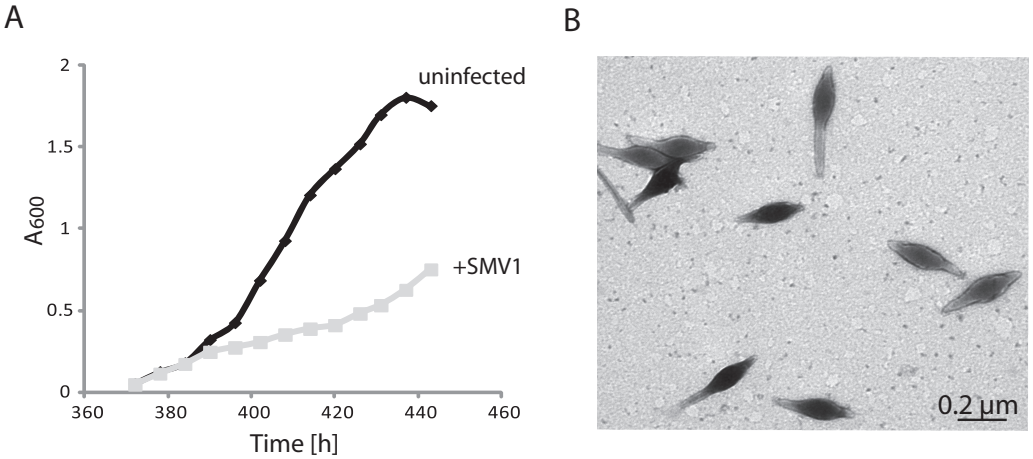

Fig. 2. *S. islandicus* infected with the SMV1 isolate from Yellowstone National Park, USA.
A. Growth curves of uninfected and SMV1-infected cultures measured over 18 days p.i.
B. Electron micrographs of virus-like particles isolated 15 days p.i. from the infected culture.

ponent, a rod-shaped virus, and the conjugative plasmid pMGB1 as a very minor component (Erdmann and Garrett, 2012; Erdmann *et al.*, 2013). Growth retardation of the culture occurred 15–17 days post infection (p.i.) (Fig. 2A) and electron micrographs of sample supernatants taken 15 days p.i. revealed the exclusive presence of single-tailed fusiform particles similar in size and form to SMV1 (Fig. 2B). The particles were purified by density gradient centrifugation and the dsDNA genome was sequenced. The sequence matched precisely to DNA contig sequences of the virus propagated earlier in *S. solfataricus* P2 (Erdmann and Garrett, 2012) and, further, yielded a circular genome.

We demonstrated by PCR amplification that SMV1, but not pMGB1, was present in the culture over 30 days p.i. (Fig. 3A). This occurred despite the presence of a match between spacer 87 of CRISPR locus 1 and an IS element in SMV1 (positions 25475–25515 with a single mismatch at position 25477). A high number of mismatches are tolerated by the CRISPR DNA interference machinery in both *S. solfataricus* and *S. islandicus* (Gudbergsdottir *et al.*, 2011; Manica *et al.*, 2011; 2013).

In earlier work, activation of spacer acquisition coincided with the onset of growth retardation of an SMV1 + pMGB1-infected *S. solfataricus* culture (Erdmann and Garrett, 2012). Therefore, CRISPR loci 1 and 2 of SMV1-infected *S. islandicus* were screened for the presence of new spacers over a 30 day period p.i. In both CRISPR loci, regions covering part of the leader and the first four repeat-spacer units were amplified by PCR (Table S1) but no larger products were formed consistent with a lack of *de novo* spacers. Instead, the main PCR products from the two leader-adjacent regions decreased dramatically in yield and they were no longer detectable at 30 days p.i. for both loci (Fig. 3B). This correlates approximately with the absence of PCR-amplified products from a core protein of *S. islandicus* (YP_005647721.1) 25 days p.i. (Table S1; Fig. 3A). We inferred that the infected host cells were being destroyed.

Properties of SMV1

Virions of SMV1 are fusiform (averaging 200 × 70 nm) with a single tail varying in length from 20 to 250 nm and

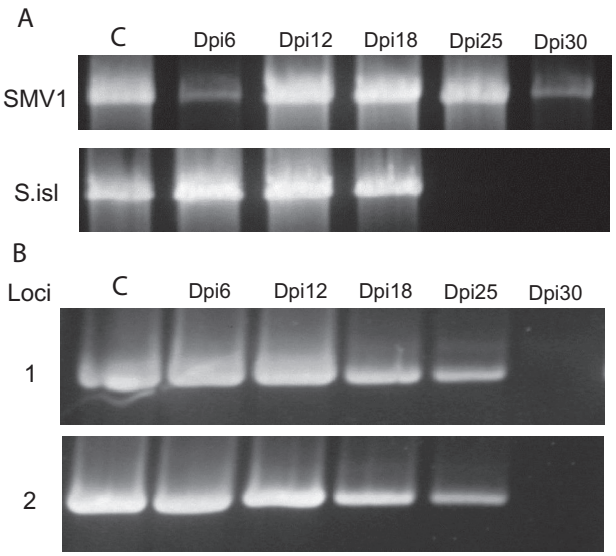

Fig. 3. Infection of *S. islandicus* with purified SMV1.
A. PCR-amplified products from SMV1 and a core gene (YP_005647721.1) of *S. islandicus*, S.isl. The control samples (C) are purified SMV1 (upper lane) and the core gene of uninfected *S. islandicus* (lower lane). Dpi, days post infection.
B. PCR-amplified products from leader proximal regions of CRISPR loci 1 and 2 for SMV1-infected *S. islandicus* and an uninfected control (C). PCR priming sites are listed in Table S1.

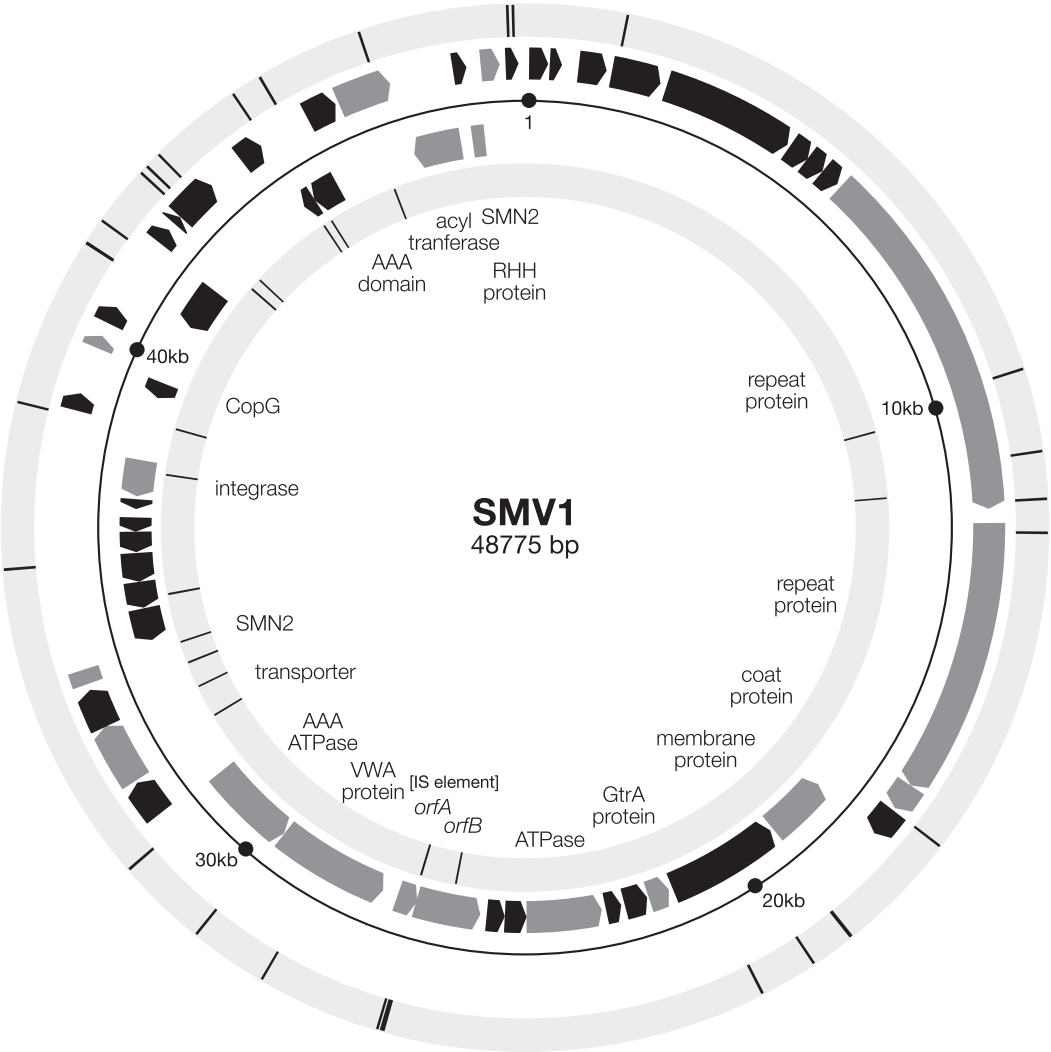

Fig. 4. Circular genome map of SMV1. Locations of the 45 *de novo* spacers which matched to SMV1 DNA are shown in the inner and outer concentric circles. Genes with predicted functions are shaded grey and are labelled. Full details of the SMV1 genomic properties are presented separately in Table S2. The nucleotide sequence is available in the European Nucleotide Archive, Accession Number HG322870.

a nose-like structure on the opposite pole (Fig. 2B), which in rare cases extended to generate a short second tail. On infection growth retardation occurs concurrently with virus replication, but no evidence was found for cell lysis upon virion release. No clear plaques were seen on Gelrite plates, nor was cell debris visible in infected cultures. Seven protein bands were resolved from the purified virions by gel electrophoresis with a major component of about 18 kDa (Fig. S1A).

The dsDNA sequence of the purified virus was determined by high throughput sequencing and there was sequence identity with the few contigs generated in the earlier study, covering about 74% of the genome which established that it was the same virus (Erdmann and Garrett, 2012). Restriction enzyme digestion of the viral DNA using a range of enzymes revealed normal digestion patterns with no evidence for base-specific modifications (Fig. S1B). Moreover, no evidence was found for integration in the host genome. The circular genome of 48 775 bp carries 51 annotated protein genes, 25 of which are arranged in nine operons. Sequence comparisons with public sequence databases established that 14 of the predicted gene products are homologous with proteins encoded by the *Acidianus* two-tailed virus ATV (Häring *et al.*, 2005; Prangishvili *et al.*, 2006b; Scheele *et al.*, 2011) with most of the remaining gene products yielding no significant matches in public sequence databases. The genome map is presented in Fig. 4 and details of the ATV homologues and predicted protein functions, and their e-values, are given in Table S2. The genome sequence is available in the European Nucleotide Archive, Accession Number HG322870.

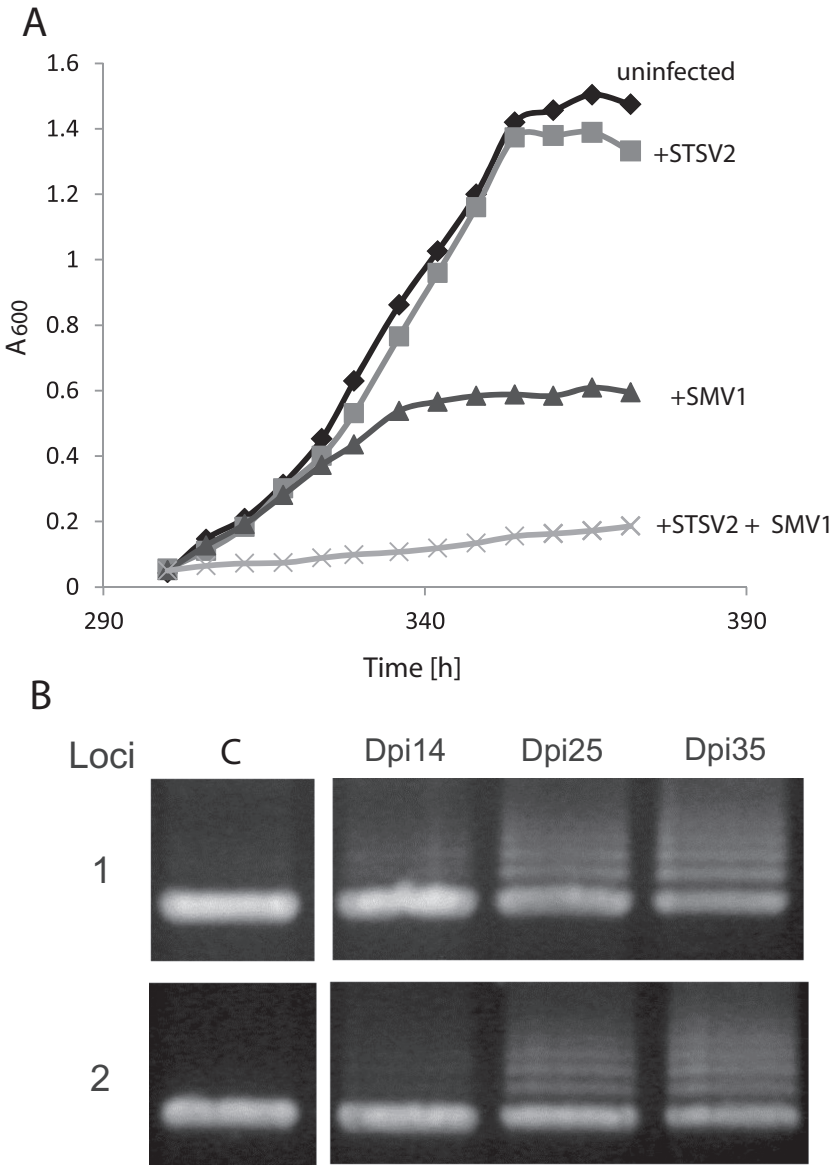

Fig. 5. Infection of *S. islandicus* with SMV1 and STSV2.
A. Growth curves of *S. islandicus* uninfected and infected with SMV1 alone, STSV2 alone, and a mixture of SMV1 + STSV2 over the period 12–16 days p.i.
B. PCR products amplified from leader proximal regions of CRISPR loci 1 and 2 after infection with a mixture of SMV1 + STSV2, and from an uninfected *S. islandicus* control (C). Dpi, days post infection.

Spacer acquisition on infection with SMV1 + STSV2

Earlier spacer acquisition in *S. solfataricus* was observed only after infection with a mixture of SMV1 and a conjugative plasmid pMGB1 and not after infection with single *Sulfolobus* viruses, including STSV2 (Erdmann and Garrett, 2012; S.E., S.L.B., R.A.G., unpublished work). STSV2 is a large single-tailed fusiform virus isolated from the Tengchong region, China, with a circular genome of 76 107 bp that is modified (Erdmann *et al.*, 2014). No evidence was found for integration in the *S. islandicus* genome. STSV2 infection only caused weak growth retardation of *S. islandicus* consistent with the putative viral release by a budding mechanism without causing cell lysis (Erdmann *et al.*, 2014). Therefore, in order to test whether SMV1 coinfection could induce spacer acquisition from STSV2, the following *S. islandicus* cultures were prepared: STSV2-infected, SMV1-infected, and SMV1 + STSV2-infected, together with an uninfected control culture. Growth retardation was only detected for the SMV1-infected and SMV1 + STSV2-infected cultures, both at 12 to 14 days p.i. (Fig. 5A). Leader proximal regions of CRISPR loci 1 and 2 were then analysed for *de novo* spacers by PCR amplification. Larger PCR fragments, indicative of spacer uptake, were produced from both loci only for the SMV1 + STSV2-infected culture (Fig. 5B). These products were isolated and cloned, and then sequencing of a fraction of the clones yielded 215 *de novo* spacers sequences with up to 5 new spacers in locus 1 and up to 8 new spacers in locus 2. They all

Table 1. *De novo* spacers matching STSV2 identified in 27 unique single clones of *S. islandicus* infected with SMV1 + STSV2.

Clone	Locus 1	Locus 2
Dpi 12		
S-1	1 (–)	1 (+)
Dpi 20		
S-2		1 (+)
S-3		1 (+)
S-4	1 (+), 1 (–)	
S-5		2 (+)
S-6		2 (–)
S-7/S-8	2 (+)	1 (–)
S-9	2 (+)	1 (+)
S-10	1 (+), 2 (–)	2 (+)
Dpi 30		
S-11		1 (–)
S-12		3 (+)
S-13	2 (+)	1 (–)
S-14/S-15	1 (+), 2 (–)	1 (+)
S-16	1 (+)	2 (+), 1 (–)
S-17	1 (–)	3 (+), 1 (–)
S-18	1 (+), 1 (–)	2 (+), 2 (–)
S-19	3 (+)	3 (+)
S-20	1 (+), 1 (–)	3 (+), 1 (–)
S-21	2 (–), 1 (+)	3 (+)
S-22	1 (+), 1 (–)	4 (+), 1 (–)
S-23	1 (+), 3 (–)	7 (+)
S-24	2 (–)	2 (+), 4 (–)
S-25	3 (+), 3 (–)	2 (+)
S-26	3 (+), 2 (–)	2 (+), 3 (–)
S-27	3 (+), 2 (–)	5 (+), 3 (–)
S-29	2 (+), 3 (–)	4 (+), 4 (–)

(+) matching the forward and (–) matching the reverse strand on the STSV2 genome. Spacer sequences are given in Table S3.
Dpi, days post infection when the clones were isolated.

matched exclusively to STSV2 and no sequence matches were observed for SMV1.

Single colonies from the culture infected with SMV1 + STSV2 were then isolated in order to determine the numbers of spacers and their distributions between the two CRISPR loci. In total 29 clones were analysed, 27 of which yielded unique *de novo* spacer patterns (Table 1 and Table S3). S-1, isolated 12 days p.i., exhibited single new spacers in each CRISPR locus, nine clones (S-2 to S-10) were selected 20 days p.i., four of which carried new spacers in both loci. A further eighteen clones (S-11 to S-29) were analysed 30 days p.i and all except two carried new spacers in both loci. On average there were 2.4 new spacers per clone 20 days p.i and 7 new spacers per clone 30 days p.i consistent with spacer acquisition being an ongoing process over the 30 day period (Table 1). In total 141 new, unique, spacer sequences were obtained from the single clones.

Next, in order to show that no undetected SMV1-matching spacers were inserted internally in the CRISPR arrays, both loci were screened for indels using PCR primers annealing at about 750 bp intervals along each locus (Fig. 1). DNA samples extracted from the SMV1 + STSV2-infected wild-type culture at 10, 15, 25 and 40 days p.i. served as PCR templates. The product sizes were consistent with neither spacer-repeat insertions nor deletions having occurred in either CRISPR locus (data not shown). Corresponding analyses of PCR products of both CRISPR loci from each of the 27 unique clones (Table 1) also revealed no significant size changes, in either locus, except for three clones. In S-10, one new spacer matching STSV2 was inserted between spacers 12 and 13 of locus 2; in S-35, the three spacers adjacent to the leader were replaced by two *de novo* spacers each matching STSV2, and in S-17 there was a large deletion of 2954 bp at the leader distal end of locus 1. In summary, while both loci appear to be quite stable on viral infection and during spacer acquisition for most of the infected culture, a minority of cells undergo diverse structural changes in their CRISPR loci.

Properties of de novo *CRISPR spacers from STSV2*

In total 347 new spacers were sequenced of which 29 were duplicated. The 318 unique sequences all matched the STSV2 genome. Spacer lengths varied from 38 to 46 bp, with single outliers of 35 bp and 48 bp. They are superimposed on the two DNA strands of the STSV2 genome in concentric circles (Fig. S2). The additional outermost and innermost circles show the distributions of CCN PAM sequences present in the genome. Several of the STSV2 protospacers overlap and many share the same end adjoining the PAM sequence but exhibit different lengths. The protospacers are distributed almost equally between the two DNA strands (52.7% and 47.3%), and fairly evenly throughout the circular STSV2 genome (Fig. S2), consistent with earlier results obtained for *S. solfataricus* on infection with a mixture of SMV1 + pMGB1 (Erdmann and Garrett, 2012; Erdmann *et al.*, 2013). Moreover, they appear to be distributed proportionally between protein genes and intergenic regions. These constitute an estimated 89% and 11% of the genome, respectively, and carry 85.7% and 14.3% of the newly acquired protospacers respectively (Table 2). There is also an equal distribution between the coding and mRNA complementary strands of the protein genes (43% and 42.7% respectively). With very few exceptions the protospacers exhibit perfect sequence matches to the *de novo* spacers.

We also examined the location and strand-specificity of multiple spacers that were acquired per CRISPR locus in single clones. The strand-specificity results are shown in Table 1 where (+) following a *de novo* spacer indicates a DNA forward strand match while (–) denotes a reverse strand match. There is no evidence for a biased strand directionality for successively acquired spacers in a given

Table 2. Analysis of the protospacer locations in the STSV2 genome with respect to DNA strand, and coding versus non coding regions, where the percentage of intergenic DNA was estimated at 11%.

Protospacer locations	Number of protospacers (%)
Forward strand	56.4
Reverse strand	43.6
Within protein genes	85.7
Coding sequence	43.0
mRNA complement	42.7
Intergenic	14.3
PAM sequences	
CCN	95.9
CCA	39.3 (35.5)
CCT	38.4 (35.5)
CCG	11.0 (11.3)
CCC	7.2 (17.7)
Inverted PAM	2.2
No PAM	1.9

A total of 347 *de novo* spacers were sequenced and analysed, 318 of which exhibited unique sequences. Percentage values are presented for the degree of conservation of the CCN PAM sequence. Percentage numbers in brackets for the four PAM sequences indicate the total percentages present in the genome. Inverted PAM indicates that the PAM sequence is located at the end of the protospacer that becomes leader distal rather than leader proximal in the CRISPR locus. No PAM denotes that no CCN sequence was present.

CRISPR locus of each clone, nor was there any evidence for a close genomic linkage of successively acquired spacers.

The PAM sequence CCN was conserved for 95.9% of the protospacers, consistent with earlier analyses (Shah *et al.*, 2013). Moreover, analysis of the –1 position revealed a strong and equal bias to A and T (39.3% and 38.4% respectively) relative to G (11%) and C (7.3%) which correlates approximately with the genomic bias of the 4816 predicted PAM sequences (A-35.5%, T-35.5%, C-17.7% and G-11.3%) (Table 2). This suggests that the identity of the –1 position is not important for spacer selection. Furthermore, alignments of the downstream regions of protospacers on the STSV2 genome did not reveal any further conserved sequence motifs in agreement with earlier studies on the Sulfolobales (Lillestøl *et al.*, 2009). Of the remaining protospacers, seven appear to match spacers that have been inverted on insertion into the CRISPR loci, with the PAM sequence located at the leader distal end of the *de novo* spacer rather than at the usual leader proximal end. These seven protospacers are presented together with the downstream PAM sequences in Table 3. The remaining six protospacers did not exhibit detectable PAM sequences.

Resistance of single S. islandicus *clones, carrying* de novo *STSV2 spacers, to infection by STSV2*

Next, single clones were selected and examined further to establish whether the *de novo* STSV2 spacers had produced resistance to viral infection. Five single clones S-1, S-2, S-4, S-7 and S-23 carrying *de novo* STSV2 spacers were chosen (Table 1 and Table S3). PCR amplification of STSV2 *orf145* and SMV1 (positions 33510–35507) demonstrated that none of these clones were still infected, except S-1 that contained SMV1 (Table 4). Cultures of these clones, and of the wild-type strain, were then challenged separately with STSV2, SMV1 and SMV1 + STSV2 and we tested again for the presence of viral DNA by PCR amplification. STSV2 was not detectable in any of the clone cultures indicating that each clone was resistant to the virus. In contrast, SMV1 was present in all of the clones

Table 3. Spacer-matching protospacer sequences in the STSV2 genome that are inverted, with the PAM sequence located at the end of the protospacer that becomes leader distal (bold) rather than leader proximal (underline) in the CRISPR locus.

CRISPR locus	STSV2 genome location	Protospacer sequence (bordered by triplets)
1	1458–1501	ATTAAAACAATTAATTGATGAGCGTGATGATTTACCAGTCATGAAAA**TGG** **ACC**
1	1874–1915	TGCCGCAGGCCCATCGCCGCATTTTACATTATTATTATATTATAC**GGG** **CCC**
2	26745–26786	CTTCATCGATGAGTTAAGTAGAGACTACATCATAAAGCCTGTAGC**TGG** **ACC**
2	43235–43277	AAATCACTAACACAATATGGTATGGATCATTGCTGTTAAGTTTATT**TGG** **ACC**
1	61216-61177 (reverse)	GTAACAATTACTTTAATGCAATTTTGCAGGTATTCGCTACGTT**AGG** **TCC**
1	65640–65679	ATCAACTATCAGGATGGGTGCCACGAATATTATGATAGAAATA**AGG** **TCC**
1	69890–69928 (reverse)	CTCAATAAACGCTCTCAAGCCATCTAGGCTCTTCAACACGAG**TGG** **ACC**

The bold PAM (CCN) sequence is shown on the right hand side on the complementary strand.

Table 4. Single clones of *S. islandicus* carrying STSV2-derived spacers (Table 1) after challenging with SMV1 and STSV2.

		STSV2 *de novo* spacers		Infected by		Sensitive on reinfection to	
Days p.i.	Single clone	Locus 1	Locus 2	SMV1	STSV2	SMV1	STSV2
12	S-1 (29)	1	1	+	–	+	–
20	S-2 (9)	0	1	–	–	+	–
20	S-4 (17)	2	0	–	–	+	–
20	S-7 (7)	2	1	–	–	+	–
30	S-23 (8)	4	7	–	–	–	–

Days p.i – indicates when cells from the culture were plated for isolation of single clones. Infected by – denotes that the virus was still present after isolation of single clones. Sensitive to – indicates the susceptibility of single clones to infection with SMV1 or STSV2. Sequences and DNA strand locations of the STSV2 protospacers are listed in Table 1 and Table S3.

S-1, S-2, S-4 and S-7 except S-23 (Table 4). PCR results are illustrated for cultures of the representative clones S-7 and S-23 over 30 days p.i and for the control wild-type strain (Fig. S3A).

Growth curves were prepared for the five clones over a 25 day period and compared with growth curves of uninfected and infected wild-type cultures. Each uninfected clone produced closely similar growth curves but they grew more slowly than the uninfected wild-type culture as exemplified for clone S-2 (Fig. 6A). Strong growth retardation was detected for the wild-type culture at the onset of spacer acquisition suggesting that the process involves a strong cost for the cells. Total recovery of the growth rate for the wild-type culture was seen at about 30–35 days p.i. indicating that several generations are required for full recovery. Single clones isolated from this culture showed recovery of the growth rate at about 20 days after isolation as illustrated for clone S-2 (Fig. 6B). However, when treated with SMV1 + STSV2, the growth retardation of each single clone was much less than observed for the wild-type *S. islandicus* strain infected with SMV1 + STSV2 (Fig. 6B).

In summary, the results are consistent with all the clones, except for S-23, being resistant to re-infection with STSV2 but not SMV1. Exceptionally S-23 was insensitive to SMV1 infection, probably due to the development of non CRISPR-based resistance during the 30 day initial exposure to the virus. The observed absence of STSV2 was also reinforced by the lack of any additional *de novo* STSV2 spacers being detectable by PCR in any of the five single clones after about 30 days p.i., a time point at which STSV2 spacer uptake generally occurred in wild-type cultures. This result is also illustrated from clones S-7 and S-23 (Fig. S3B).

S. islandicus *infection with different virus mixtures did not produce spacer acquisition*

Our inability to induce spacer acquisition in *S. islandicus* with any single genetic element, including STSV2, raised the possibility that coinfecting genetic elements are required to activate spacer uptake. Therefore, new viral isolates were purified from enrichment cultures that we had collected from terrestial hot-springs on Iceland and in Italy. They included a *Sulfolobus islandicus* filamentus virus 2 (SIFV2) and a *Sulfolobus islandicus* rod-shaped virus (SIRV3), both isolated from Iceland and a *Sulfolobus* short rod-shaped virus-like particle denoted SSRV from Naples,

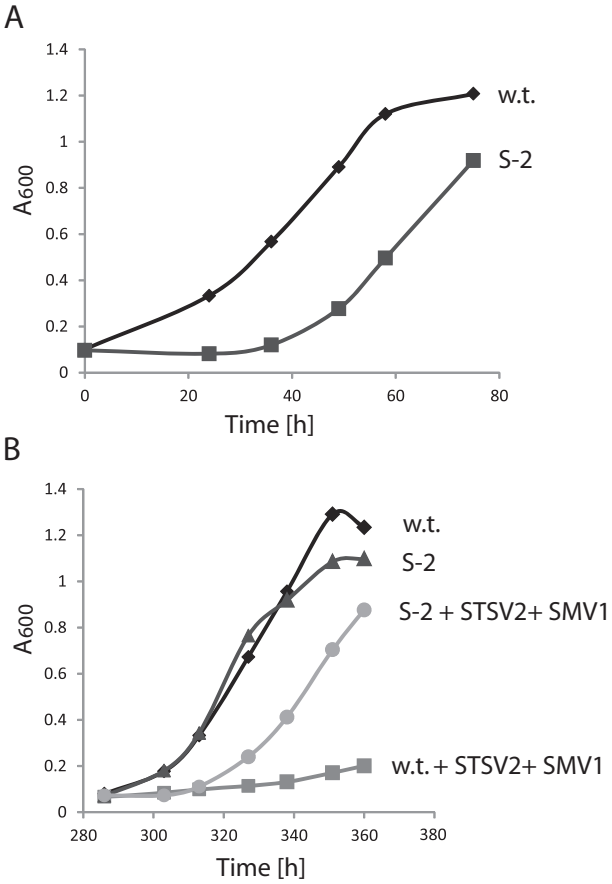

Fig. 6. Growth curves of cultures of wild-type *S. islandicus* and clone S-2: (A) alone and (B) uninfected and infected with SMV1 + STSV2, and initiated 12 days p.i.

Italy. SIFV2 yielded partial genome sequences that matched closely to the genome of the lipothrixvirus SIFV (Arnold *et al.*, 2000) and, also, partial sequences of SIRV3 matched closely to the rudiviruses SIRV1 and SIRV2 (Peng *et al.*, 2001). Electron micrographs of the isolated virus particles are presented in Fig. S4.

The *S. islandicus* wild-type culture infected with STSV2 was additionally challenged with (a) SIFV2 and (b) SIRV3 + SSRV, with an uninfected culture as a control. Growth retardation was observed for each culture carrying virus mixtures 3 days p.i. but growth recovered about 17 days p.i. (Fig. 7A). Mixed virus cultures were maintained over 45 days and tested every few days for evidence of spacer acquisition by PCR amplification but none was detected (Fig. 7B). Infection by STSV2 and SIFV2 was established by PCR amplification (Table S1), and infection by SIRV3 and SSRV was confirmed by electron microscopy (data not shown). We conclude that infection with a virus mixture could not alone explain the uptake of *de novo* STSV2 spacers observed in the presence of SMV1.

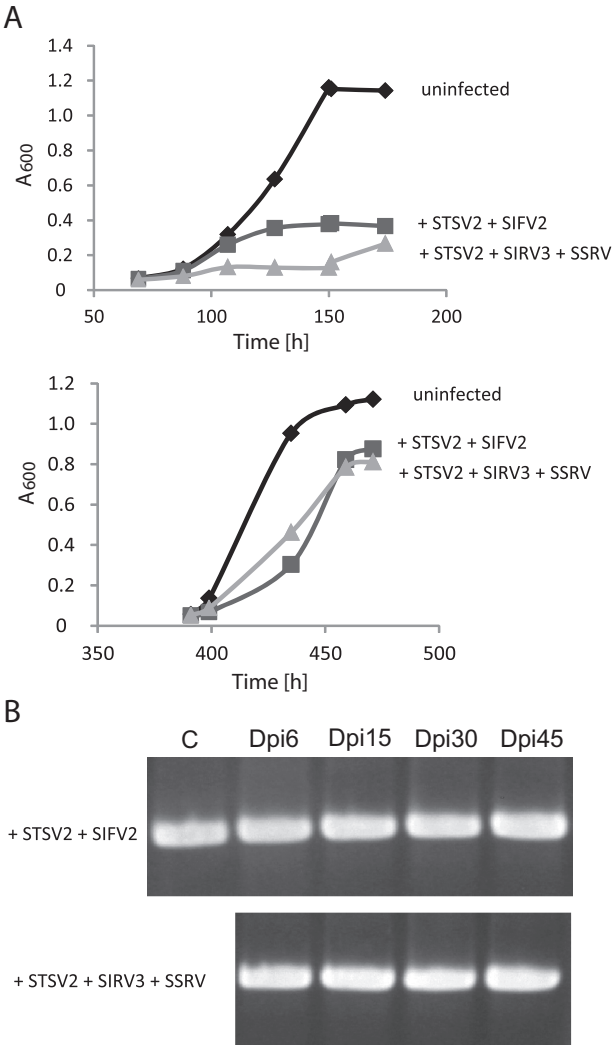

Fig. 7. Effect of multiple viral infections on growth rates and the CRISPR loci of wild-type *S. islandicus*.
A. Growth curves of uninfected *S. islandicus* cultures and cultures infected with STSV2 + SIFV2 or with STSV2 + SIRV3 + SSRV that were measured 4–7 days, and 17–20 days p.i. Electron micrographs of SIFV2, SIRV3 and SSRV virus particles are presented in Fig. S4.
B. PCR-amplified products from the leader proximal regions of CRISPR locus 1 of *S. islandicus* infected with each virus mixture, and from an uninfected *S. islandicus* control (C). Samples were analysed over 45 days p.i. Dpi, days post infection.

Clone S-1 rapidly acquired spacers from SMV1 after storage at –80°C

Clone S-1 originated from a culture that was actively acquiring *de novo* STSV2 spacers 12 days p.i. It was still infected with SMV1 but had lost STSV2 and carried a single *de novo* STSV2 spacers in each CRISPR locus (Table S3). Initially, it was selected for further study to determine whether it is able to acquire *de novo* spacers from different viruses. When cultured from a glycerol stock at –80°C, growth of S-1 was strongly retarded and, consistent with the preceding results, it could not be reinfected with STSV2. Unexpectedly, PCR amplification results of the leader adjacent regions indicated that spacer uptake had occurred exceptionally rapidly, already 2 days p.i. (Fig. 8A).

The *de novo* spacers in both CRISPR loci were sequenced and, surprisingly, they all matched to SMV1. Forty three unique *de novo* spacers were sequenced and they are superimposed on the genome map of the SMV1 virus (Fig. 4). This indicated that either SMV1 had lost its ability to protect itself against spacer acquisition or that mutations had occurred in the host enabling it to circumvent the underlying viral protection mechanism. We concluded that the cold-shock treatment had rendered SMV1 susceptible to spacer acquisition.

Differential loss of de novo *SMV1 and STSV2 spacers*

Viral infection of the SMV1-infected S-1 culture was monitored over a 27 day period and we also examined the *de novo* spacer composition of both CRISPR loci. The results showed that SMV1 was propagating stably despite the presence of *de novo* SMV1 spacers (Fig. 8). However, a significant difference was observed in the CRISPR loci at about 21 days after reactivation of the culture in that both *de novo* SMV1 spacers were lost. Furthermore, about 27 days after reactivation both *de novo* STSV2 spacers were also lost and the CRISPR loci had reverted to the wild-type composition (Fig. 8A). This difference was confirmed for several single clones isolated from this culture

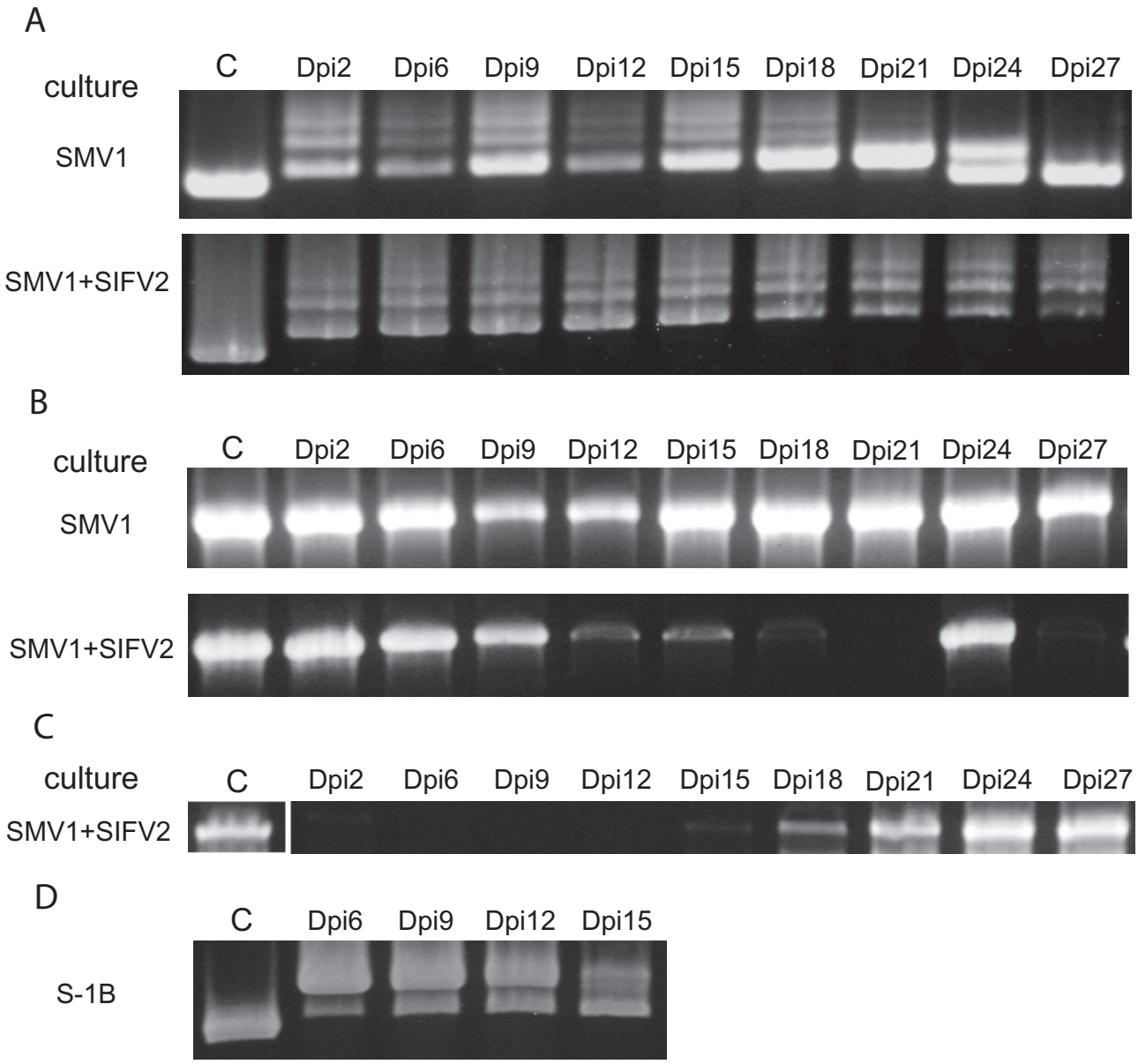

Fig. 8. PCR analyses of CRISPR loci and viral DNA in SMV1-infected clone S-1, the latter culture challenged with SIFV2, and the SMV1-reinfected subclone S-1B.
A. PCR-amplified products from the leader proximal region of CRISPR locus 1 of SMV1-infected clone S-1 before, and after, infection with SIFV2 compared with the PCR product from the uninfected *S. islandicus* control (C).
B. PCR-amplified products from SMV1 in cultures with and without SIFV2 in (A) and from an isolated SMV1 control (C).
C. PCR-amplified products from SIFV2 in the S-1 culture coinfected with SIFV2 and from a purified SIFV2 control (C).
D. PCR-amplified products from the leader proximal region of CRISPR locus 1 for the S-1B subclone infected with a mixture of SMV1v + SMV1 isolated from clone S-1 and an uninfected control sample (C). In each experiment template DNA concentrations were adjusted to be the same for each sample. Dpi, days post infection.

(Table 5) and it was concluded that all the *de novo* spacers were reversibly lost in the SMV1-infected S-1 culture. Moreover, *de novo* spacer uptake and loss coincided with strongly retarded growth over the whole 27 day period and these results were reproducible in two independent experiments.

Infection with SIFV2 influenced SMV1 propagation

Next we examined whether SMV1 could induce spacer acquisition from the coinfecting filamentous virus SIFV2 by challenging the SMV1-infected S-1 culture with SIFV2 and testing for spacer acquisition over 27 days. We compared

Table 5. Characterization of subclones isolated from S-1 cultures that carried single *de novo* STSV2 spacers in each CRISPR locus (Table 4).

Parent culture	Single clone	STSV2 infected	SMV1 infected	SIFV2 infected	Locus 1 *de novo* spacers		Locus 2 *de novo* spacers		STSV2 sensitive	SMV1 sensitive
					STSV2	SMV1	STSV2	SMV1		
S-1	S-1A	–	–	–	0	0	0	0	+	–
S-1 + SIFV2	S-1B	–	–	–	1	2	1	2	–	+
	S-1C	–	–	+	1	0	1	5	–	–

The cultures were either uninfected or infected with SIFV2. Each subclone was tested for the presence of STSV2, SMV1 and SIFV2 and the *de novo* spacers of each subclone were PCR amplified and sequenced. 0 indicates that no *de novo* spacers were present. Each subclone was also tested for sensitivity to reinfection by STSV2, and by the SMV1 mixture isolated from S-1. '+' and '–' denote positive and negative results respectively.

the CRISPR loci of the control SMV1-infected S-1 culture and after challenging it with SIFV2. The latter culture was shown by PCR amplification to be infected by both SMV1 and SIFV2 (Fig. 8B and C) but, unexpectedly, after the initial spacer acquisition, the CRISPR loci remained unchanged over the whole 27 day period. This contrasted with the loss of all *de novo* spacers observed for the SMV1-infected S-1 control culture (Fig. 8A). Sequencing of the *de novo* spacers in the SIFV2-infected culture at 3 and 21 days p.i. revealed only SMV1-matching spacers and none from SIFV2. The results suggested that acquisition of *de novo* SMV1 spacers occurred prior to SIFV2 infection but that coinfection with SIFV2 influenced propagation of SMV1 because, in contrast to the control culture lacking SIFV2, no loss of *de novo* spacers occurred (Fig. 8A). Furthermore, subclones carrying *de novo* SMV1 spacers could be isolated from the SIFV2-infected culture (Table 5).

Complex relationship between de novo *spacer composition and SMV1, STSV2 and SIFV2 sensitivity of different S-1 subclones*

Next we tested whether the *de novo* SMV1 spacers generated resistance to SMV1 infection. Single clones were isolated from the SMV1-infected S-1 cultures and leader proximal regions of the CRISPR loci of several clones were sequenced and shown to have reverted to wild-type spacer compositions (Fig. 8A). A single subclone S-1A was then selected for further characterization. Two subclones, S-1B and S-1C, with *de novo* SMV1 spacers, were selected from the SMV1+SIFV2-infected S-1 culture.

PCR amplification demonstrated further that none of the three subclones contained SMV1, but S-1C was still infected with SIFV2 (Table 5). The three subclones were then tested for sensitivity to reinfection with both STSV2 and SMV1 isolated from the S-1 culture. Only S-1A was reinfected by STSV2, consistent with the loss of *de novo* STSV2 spacers, but unexpectedly it was resistant to SMV1 (Table 5). We infer that cells had developed SMV1 resistance by a CRISPR-independent mechanism, as was observed for clone S-23 isolated from an STSV2 + SMV1-infected wild-type culture (Table 4).

The resistance of S-1B to STSV2 reinfection was consistent with the presence of *de novo* STSV2 spacers but it was sensitive to SMV1 infection despite the presence of *de novo* SMV1 spacers. In contrast, S-1C was resistant to both STSV2 and SMV1 infections, consistent with SIFV2 affecting SMV1 propagation (Table 5). Furthermore, subclone S-1B differed from the subclone S-1C in that the *de novo* SMV1 spacers were also gradually lost (Fig. 8D). We concluded that the CRISPR loci of S-1C were stable because the subclone was not infected by SMV1.

A variant SMV1, SMV1v, with an altered response to the CRISPR immune system is deficient in a virion protein

We investigated further the S-1 clone carrying *de novo* SMV1 spacers with a view to understand how SMV1 avoids the CRISPR immune system. PCR amplification of continuous contigs along both host CRISPR loci revealed no size changes (data not shown). Therefore, we analysed virus particles released from S-1 cells for altered properties. No morphological changes were observed in electron micrographs and no evidence of viral DNA modification was observed with multiple restriction enzyme digests which all produced digestion patterns identical to those of wild-type SMV1 DNA (Fig. S1B). However, when SMV1 was isolated from S-1 and purified by standard CsCl density gradient centrifugation, exceptionally two closely migrating virion bands were resolved. Both products were analysed for protein content by SDS-PAGE and virions from the lower band lacked a small protein component that was present in both the upper band and wild-type SMV1 virions (Fig. 9A). The latter virus was named SMV1 variant, SMV1v, and the deficient protein was purified from the wild-type SMV1 and subjected to Edman N-terminal sequencing. It carried the N-terminal sequence VDEYF which constituted a truncated version

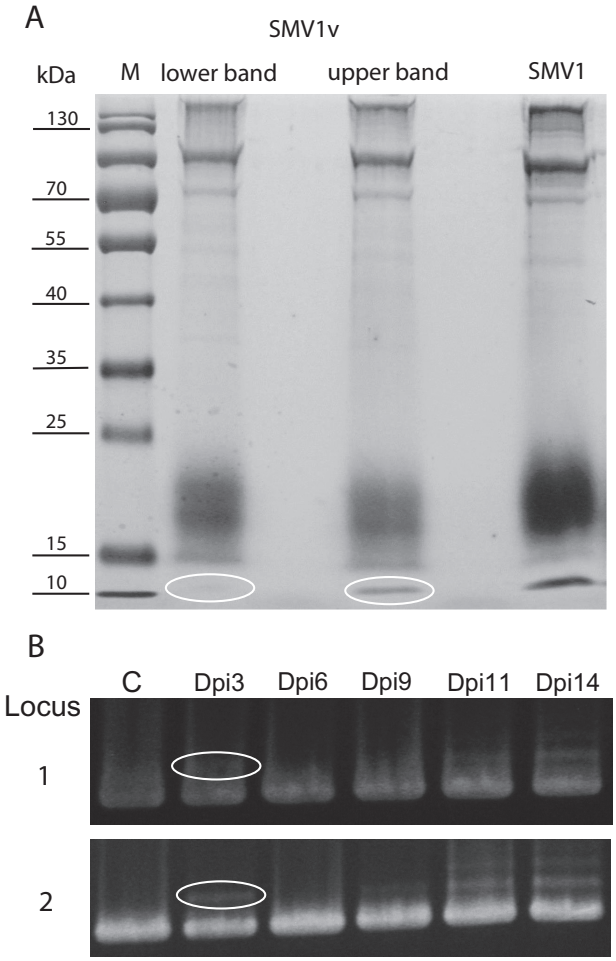

Fig. 9. Isolation of the SMV1v from SMV1-infected clone S-1.
A. 11% SDS-PAGE of CsCl gradient-purified virus particles. Two closely migrating SMV1 bands were resolved in the gradient. The lower band lacked the small encircled protein component that was present in the upper band and wild-type SMV1.
B. PCR amplified products from the leader proximal region of CRISPR locus 1 of wild-type *S. islandicus* infected with the SMV1v + SMV1 mixture and STSV2. Larger PCR products that were visible for both CRISPR loci at 3 days p.i. are encircled. Control (C) – uninfected *S. islandicus*.

of ORF153, a putative protein showing significant sequence similarity to the virion proteins ORF131 of ATV and ORF145 of STSV2 (Table S2). The protein contains 114 amino acids, lacking the N-terminal 39 amino acids of ORF153, and it exhibits the potential start codon GUG.

Evidence supporting that SMV1v DNA is susceptible to spacer acquisition

Next, we tested the hypothesis that SMV1v DNA is sensitive to spacer acquisition and challenged wild-type *S. islandicus* with an approximately equal mixture of SMV1 and SMV1v isolated from clone S-1. An additional culture infected with the SMV1 virus mixture + STSV2 was tested as a control. Infection produced immediate growth retardation in both cultures, much earlier than normally observed during the initial SMV1 infections at 12 days p.i. (Fig. 2A). Moreover, the *de novo* spacer uptake had occurred in both CRISPR loci 3 days p.i., as judged from the formation of larger PCR products but the reaction ceased at about 6 days p.i. after which larger PCR product yields decreased (Fig. 9B). These larger PCR products were cloned and sequenced and the *de novo* spacer sequences matched DNA from either SMV1 or the host *S. islandicus*, with 10 of the 12 unique spacer sequences matching the host chromosome.

Furthermore, the culture challenged with the SMV1 + SMV1v mixture + STSV2 produced a second growth retardation at 12 days p.i and uptake of multiple *de novo* spacers per CRISPR locus was observed exclusively from STSV2 (Fig. 9B) as seen earlier for SMV1 + STSV2-infected cells (Fig. 2A). We infer that the initial *de novo* SMV1 spacers are produced from SMV1v while the subsequent uptake of STSV2 spacers is induced by SMV1.

Next *S. islandicus* was challenged with the SMV1v + SMV1 mixture and the rudivirus SIRV3 (Fig. S4). Spacer acquisition was observed from both SMV1 and SIRV3 DNA about 3 days p.i. again consistent with the SMV1v inducing spacer acquisition immediately after infection (data not shown). This also established that SMV1 can facilitate spacer acquisition from a second virus carrying a linear genome that did not activate spacer acquisition when infected with other viruses (Fig. 7A and B).

ORF114 binds DNA cooperatively

ORF114 was absent from the SMV1v virions that had apparently lost the ability to avoid spacer acquisition. The secondary structure prediction for ORF114 revealed a helix–turn–helix motif and the overall protein prediction was similar to the structure determined for the homologous ATV virion protein ORF131 and the virion protein ORF145 of STSV2 (Goulet *et al.*, 2010; Erdmann *et al.*, 2014). To examine the properties of the protein further, the gene was cloned into vectors producing either a C-terminal His6-tag or an N-terminal GST-tag and it was expressed recombinantly in *E. coli*. The purified protein was shown to be contaminated with *E. coli* DNA, that was difficult to remove, consistent with DNA-binding activity. DNA-binding assays were performed with ORF114-His6 on a 125 bp DNA-substrate and the resulting binding curve was consistent with cooperative binding of multiple copies of ORF114 per DNA fragment (Fig. 10A). We infer that the high protein : DNA molar ratio at the plateau reflects firstly that multiple protein copies are bound per DNA molecule and secondly that many protein molecules

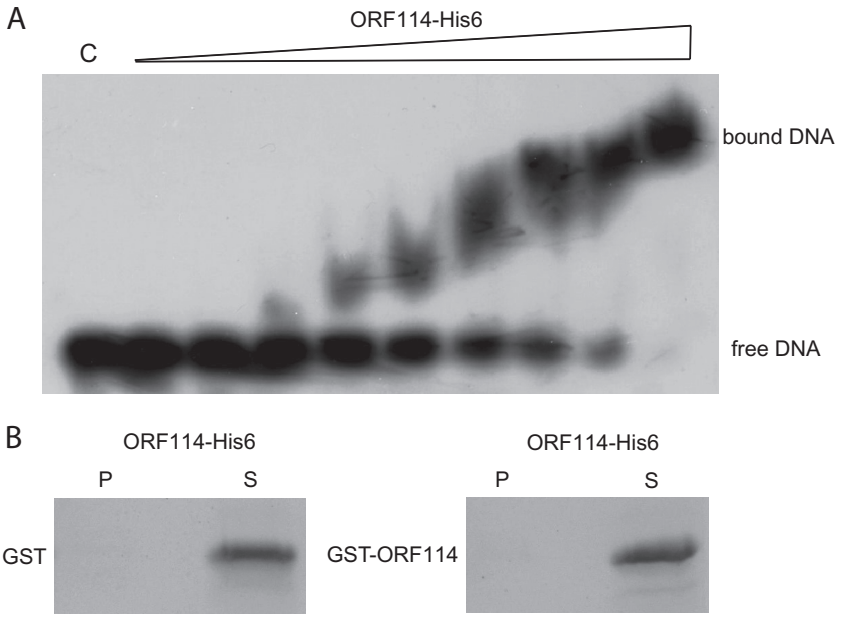

Fig. 10. Biochemical characterization of ORF114.
A. DNA binding assay with increasing concentrations of ORF114 (6.25–200 pmol) and 0.35 pmol of ^{32}P-labelled 125 bp dsDNA substrate in tracks from left to right, 0 (control sample C), 18:1, 36:1, 72:1, 90:1, 144:1, 216:1, 288:1 and 567:1 protein : DNA molar ratios.
B. Pulldown assay of GST-ORF114 with ORF114-His$_6$ with GST serving as control. P denotes the pellet (beads with immobilized protein) and S indicates supernatant (free protein after incubation with beads).

were not displaced from *E. coli* DNA fragments during preparation.

Next, triplicate pull-down assays were performed with GST-ORF114 immobilized on Glutathione Sepharose and free ORF114-His6 showing no evidence of interaction. This suggests that ORF114 does not oligomerize like the homologue ORF145 in STSV2 (Erdmann *et al.*, 2014) and it supports the observation of cooperative DNA binding (Fig. 10A).

SMV1 DNA replication is induced by cold-shock of infected cultures

Earlier we demonstrated that when a *S. solfataricus* culture was infected with a mixture of SMV1 and pMGB1 and subjected to cold-shock by freezing at –80°C, *de novo* spacer uptake from pMGB1 was specifically activated for one of six CRISPR loci (locus A) (Erdmann *et al.*, 2013). Moreover, in this study clone S-1 showed unexpected spacer acquisition from SMV1 on infection with SMV1 after storage at –80°C. Therefore, we investigated whether cold stress influenced viral replication by sampling an SMV1 + STSV2-infected culture 2 days p.i., storing the cells overnight in 15% glycerol at –80°C and reactivating the culture after 12 h. The original culture, and the cold-shocked culture, were then sampled each day, and DNA was extracted and diluted to the same concentration for each sample. The results showed that STSV2 DNA was present at a high level in both cultures consistent with the virus actively replicating. In contrast, SMV1 DNA was present at a very low concentration in the untreated culture but it increased strongly in yield after cold-shock, consistent with the cold-shock activating SMV1 replication (Fig. S5).

Discussion

SMV1 activates spacer acquisition indirectly from coinfecting genetic elements

Evidence is presented demonstrating that when the single-tailed fusiform viruses SMV1 and STSV2 coinfect *S. islandicus* REY15A, hyperactive spacer acquisition occurs exclusively from STSV2 in both CRISPR loci. In contrast, neither STSV2 alone, nor multiple infections with diverse linear DNA *Sulfolobus* viruses, nor mixtures of both STSV2 and linear DNA viruses, stimulated detectable spacer acquisition. Spacer uptake was also observed earlier from a conjugative plasmid pMGB1 that was coinfected with SMV1 in *S. solfataricus* P2 (Erdmann and Garrett, 2012) but neither study showed spacer acquisition from SMV1. This indicated that SMV1 was protected against spacer acquisition and therefore we infer that the spacer uptake from co-infecting viruses is a secondary effect induced by, but not directed by, SMV1.

SMV1 infection of *S. islandicus* alone resulted in cell death (Fig. 3A) and, given that SMV1 does not cause cell lysis, and that no coinfecting genetic elements were present, this may have been caused by spacer uptake from, and DNA interference of, chromosomal DNA. This interpretation is strengthened by the observation that there was a preferential uptake of spacers from the *S. islandicus* chromosome only three days after infection of the wild-type strain with SMV1v.

Spacer acquisition is associated with strong growth retardation of the culture

Culture growth was strongly retarded during spacer acquisition on SMV1-infection of *S. islandicus*, and also of *S. solfataricus* (Erdmann and Garrett, 2012). The growth rate recovered about 20 days after acquisition was initiated. Moreover, single clones isolated from a culture undergoing spacer acquisition, that had lost the virus, still showed retarded growth for about 12 days after isolation (Fig. 6), indicating that growth retardation was not only caused by viral infection. Spacer acquisition and viral replication seems to produce a high cost for the cells. However, the adaptation process could be facilitated by toxin activity which would cause a decrease in cell growth and could provide a window for spacer uptake to occur prior to cell division, because several generations were required in our study for reestablishing growth rates after termination of spacer acquisition. Toxin-antitoxin gene pairs are often interwoven among *cas* gene cassettes and it has also been proposed that some Cas proteins may exhibit toxin activities (Garrett *et al*., 2011; Makarova *et al*., 2012).

Essentially random protospacer selection

318 unique *de novo* spacer sequences were determined and superimposed on the STSV2 genome. Their distribution is essentially random with respect to the DNA strand and to protein coding versus non coding regions, following the pattern described recently, for spacer acquisition from the *Sulfolobus* conjugative plasmid pMGB1 (Erdmann and Garrett, 2012; Erdmann *et al*., 2013), and in agreement with earlier *in silico* predicted distributions of protospacers localized on linear and circular genomes of diverse genetic elements of the Sulfolobales (Shah *et al*., 2009; 2011; Shah and Garrett, 2011). This is in marked contrast to the recent demonstration of biased protospacer selection from a bacteriophage by a bacteria-specific subtype II-A CRISPR system of *S. thermophilus* (Paez-Espino *et al*., 2013). Another significant difference from bacterial systems is that no evidence was found for the same DNA strand directionality of *de novo* spacers taken up consecutively in a given CRISPR locus, as reported for a genetically manipulated *E. coli* subtype I-E spacer acquisition (Datsenko *et al*., 2012; Swarts *et al*., 2012).

Limits on CRISPR locus sizes

Hyperactive spacer acquisition occurred in both CRISPR loci of *S. islandicus*, with a maximum of 5 new spacers in locus 1 and 8 in locus 2 and these maxima were not exceeded despite spacer uptake continuing over many days. This suggests that there is a mechanism regulating the maximum length of a CRISPR locus. In support of this, maximum limits on the numbers of acquired *de novo* pMGB1 spacers were observed for loci C, D and E of *S. solfataricus* (Erdmann and Garrett, 2012).

All de novo *STSV2 spacers generated viral resistance*

Experiments with several isolated single clones carrying different *de novo* STSV2 spacers provided evidence for the degree of viral resistance not being influenced by the number, or sequence, of the *de novo* STSV2 spacers present. This differs from the earlier observation of increasing numbers of *de novo* spacers providing increased bacteriophage resistance for the type II-A CRISPR system of *S. thermophilus* (Barrangou *et al*., 2007; Deveau *et al*., 2008).

The 27 unique single clones from *S. islandicus* carried a wide range of numbers of *de novo* spacers extending from one (S-2, S-3, S-11) and two (S-1, S-4, S-5, S-6, S-9) up to ten (S-26), eleven (S-23) and twelve (S-27) (Table 1). All were completely resistant to infection with STSV2. This suggests that the virus can only escape DNA interference by undergoing protospacer mutations or deletions such that multiple *de novo* spacers will give the host added protection over longer periods.

Evidence for reversible uptake of CRISPR spacers

The progressive loss of *de novo* STSV2 and SMV1 spacers observed for clone S-1 and subclone S-1B (Fig. 8A) reinforces that *de novo* spacers are potentially less stable than older spacers since a deletion of internal CRISPR spacers was shown to be a rare event. This loss of *de novo* spacers was only observed when SMV1v was present but not for single clones infected only with SMV1 or for single clones that were resistant to SMV1 infection. Therefore, we infer that loss of *de novo* spacers is induced, directly or indirectly, by the SMV1 variant. It was demonstrated earlier that plasmids carrying protospacers matching host spacers of *S. solfataricus* P2 or *S. islandicus* REY15A could survive DNA interference when matching spacers in the CRISPR loci were deleted (Gudbergsdottir *et al*., 2011). Thus, spacer loss observed in *S. islandicus* could result from a reversible spacer acquisition mechanism, possibly induced by the SMV1 variant.

Inter-viral conflicts between SMV1 and SIFV2

Contrasting spacer acquisition results were obtained for *S. islandicus* cultures infected with SMV1 alone or with SMV1 + STSV2. Clone S-1 carrying *de novo* STSV2 spacers lost these spacers. However, when the same culture was challenged with SIFV2, no spacer loss occurred. Given that spacer loss only occurred in the presence of SMV1v, the result indicates that this variant

was unable to reinfect SIFV2-infected cells. Moreover, a subclone S-1C isolated from this SIFV2-infected culture that was still infected with SIFV2, was shown to be insensitive to SMV1 infection. Therefore we infer that intracellular SIFV2 can inhibit coinfection by SMV1.

Conflicting interactions between SMV1 and the CRISPR immune system: a working hypothesis

SMV1 activates spacer acquisition: a link between viral DNA replication and spacer acquisition. The activation mechanism for spacer acquisition remains unclear. There was no evidence for the operation of a matching spacer 'priming' mechanism (Table 1 and Table S3) as proposed for a bacterial type I-E system (Datsenko *et al*., 2012), unless it can occur indirectly, where the CRISPR spacers matching SMV1 can act as 'primers' for spacer uptake from a second genetic element. Nor was there evidence for the successive uptake of *de novo* spacers, deriving from one DNA strand, in a given CRISPR locus (Table 1). However, our experimental results did suggest a link between spacer acquisition and viral DNA replication. On coinfection with STSV2 + SMV1, STSV2 replicated immediately on infection, whereas SMV1 replication was only detectable 12–16 days later, or 2 to 3 days later after exposure to freezing at –80°C (Fig. S5). Consistent with this inference, SMV1 replication was observed immediately after infecting an *S. solfataricus* P2 mutant lacking both CRISPR loci A to D and *cas* genes required for spacer acquisition (Erdmann and Garrett, 2012).

STSV2 alone did not activate spacer acquisition but, unlike SMV1, it exhibits a modified genome (Erdmann *et al*., 2014) which might enable the viral DNA to avoid being recognized as an invading genetic element. Thus, although SMV1 initially remains dormant in cells carrying active CRISPR systems for self-protection, host cell ageing, cold-shock, or other stress factors, seem to stimulate SMV1 replication which results in activation of the CRISPR immune system and spacer acquisition.

SMV1 resistance to spacer acquisition: viral DNA is protected by a protein. No spacer acquisition from SMV1 was detected after coinfection of *S. solfataricus* with SMV1 + pMGB1 (Erdmann and Garrett, 2012) nor after coinfection of *S. islandicus* with SMV1 + STSV2 (this study) which indicates that SMV1 is protected against protospacer selection. Exceptionally, SMV1 DNA from clone S-1 was sensitive to spacer acquisition. Analysis of the infecting virions revealed two distinct particles one lacking ORF114 which was subsequently shown to bind DNA cooperatively. Assuming that SMV1 DNA carries some bound ORF114 during replication, and that ORF114 binds cooperatively, it will always preferentially assemble on the SMV1 genome and could, thereby, physically protect SMV1 DNA from protospacer selection. Thus the SMV1v virion could result from the rapid viral DNA replication that follows cold shock of host cells at –80°C and results in a deficient production of ORF114.

SMV1 avoids CRISPR-directed DNA interference: indirectly, and directly? On infecting *S. islandicus*, both SMV1 and the SMV1v + SMV1 mixture avoided DNA interference despite a spacer with a single mismatch encoded by *S. islandicus* and *de novo* SMV1 spacers present in subclone S-1B, each with an associated cognate -CCN- PAM sequence. SMV1 also infected *S. solfataricus* P2 despite the presence of eight perfectly matching spacers with cognate PAM sequences (Erdmann and Garrett, 2012). Therefore, we infer that both SMV1 and SMV1v are able to avoid CRISPR-directed DNA interference although the mechanism remains unclear.

SMV1 appears to be dormant in cells with active CRISPR systems for 12–14 days p.i., although not in CRISPR-deficient mutants (Erdmann and Garrett, 2012), possibly to protect itself from DNA interference. Thus, as suggested for activation of spacer acquisition, DNA interference might also be activated by the onset of viral replication. Non replicating SMV1 could provide a limited period for viral protein biosynthesis. Although a role for ORF114 in avoiding DNA interference can be excluded, because SMV1v is able to infect wild-type *S. islandicus* or S-1 subclone S-1B carrying matching SMV1 spacers (Table 5), other viral proteins may be expressed which directly inhibit the interference complex or indirectly protect the viral DNA from the interference complex, possibly in a related way to anti-CRISPR viral proteins described for some *Pseudomonas* strains (Bondy-Denomy *et al*., 2012).

Perspectives

Phylogenetically diverse archaeal viruses have been shown to coinfect hosts in natural environments and under laboratory conditions (Prangishvili and Garrett, 2005). The experiments described here with the different viruses SMV1, STSV2, SIFV2 and SIRV3 infecting *Sulfolobus* show that interactions with the host CRISPR system can be very complex, and complicated further by the potential activation of other resistance mechanisms observed for some clones. This raises many unresolved questions including why does STSV2 not activate spacer acquisition although it is sensitive to DNA interference, and how does SIFV2 inhibit infection of *S. islandicus* with SMV1? Clearly, each virus has evolved its own mechanisms for virus-host interactions, and for coexisting intracellularly with other genetic elements. Continuous infections with different viruses produce diverse reactions from both virus and host as observed, for example, for

both the variant virus SMV1v and the different clones isolated from virus-infected *S. islandicus* cultures. This work yields insights into the conflicting interactions of SMV1 with the host CRISPR immune system of *S. islandicus* and we propose potential mechanisms for the viral resistance to spacer acquisition and DNA interference. Clearly further studies will be required to elucidate the complex mechanistic details of these diverse interactions.

Experimental procedures

Infection, isolation and sequencing of the SMV1 viral genome

S. islandicus REY15A was grown in *Sulfolobus* medium supplemented with 0.2% trypton, 0.1% yeast extract and 0.2% sucrose (TYS medium) (Zillig *et al.*, 1994). 10 ml of cells were harvested from a fresh culture by centrifuging (6000 g, 10 min) and resuspending in 1 ml of TYS medium. A virus mixture from Yellowstone National Park, USA, was isolated (Erdmann and Garrett, 2012) and 20 μl were added. After incubating for 2 h at 78°C, infected cells were transferred to 50 ml of pre-heated (78°C) TYS medium. The culture was then incubated for 3 to 6 days at 78°C before isolating and purifying the virus (Erdmann and Garrett, 2012). The viral genome was sequenced using a Hiseq 2000 sequenator (Beijing Genomics Institute, Shenzhen, China) to yield about a 200-fold coverage and the genome was assembled using the CLC genome package (Aalborg, Denmark) and Velvet (Zerbino and Birney, 2008). Gaps between assembled contigs where closed by PCR amplification and sequencing.

Isolation of SIFV2, SIRV3 and SSRV and infection of S. islandicus

SIFV2 and SSRV were isolated from a mud sample taken from the Gunnuhver geothermal area, Iceland and SIRV3 was isolated from a mud sample collected at Pozzuoli, Naples, Italy. Isolation was performed as described for SMV1 using *S. islandicus* REY15A as host strain (Erdmann and Garrett, 2012). Virus particles were purified by loading onto 0.45 g ml^{-1} CsCl and centrifuging at 38 000 rpm for 48 h in a SW41 rotor (Beckman, Fullerton, USA). Virus bands were extracted and CsCl was removed by dialysis against 10 mM Tris-HCl, pH 8. DNA was isolated from purified virus particles using DNeasy® Blood&Tissue Kit (Qiagen, Westberg, Germany) and digested with HindIII (Fermentas, St. Leon-Rot, Germany). Ends were filled in using Klenow Polymerase (Fermentas) according to the manufacturer's protocol. Blunt end fragments were cloned using CloneJET PCR Cloning Kit (Thermo Fisher Scientific) and sequenced (MWG Eurofins, Ebersberg, Germany). Contigs were used for primer design to verify infection by PCR. 10 ml of *S. islandicus* cells were harvested from a fresh culture by centrifuging (6000 *g*, 10 min) and resuspending in 1 ml of TYS medium. Three microlitres of each purified virus, SMV1 SIFV2 or SIRV3 at 50 PFU μl^{-1} and 22.5×10^6 particles μl^{-1} of STSV2 were added and, after incubating for 2 h at 80°C, infected cells were transferred to 50 ml of pre-heated TYS medium at 78°C. Infection with SSRV was achieved in the same manner using an environmental virus mixture containing SSRV that was isolated as described (Erdmann and Garrett, 2012) together with an additional virus-like particle that did not propagate in *S. islandicus* REY15A.

Growth curves and PCR amplification of CRISPR loci and viral genomes

Wild-type *S. islandicus* REY15A and isolated single clones were cultivated in 50 ml TYS medium at 78°C, both uninfected and after infection with different viruses. One millilitre was taken from each culture every 12 h and A_{600} values were measured. In addition, 2 ml samples were extracted every 24 h and cells were harvested by centrifugation (6000 *g*, 10 min) and DNA was isolated using DNeasy® Blood&Tissue Kit (Qiagen). Infection with SMV1, STSV2 and SIFV2 was monitored by PCR amplification using primers listed in Table S1. Infection with SSRV and SIRV3 was monitored by electron microscopy. Leader proximal regions of CRISPR loci 1 and 2, extending from the leader to spacer five, and approximately 750 bp regions covering the whole of each CRISPR locus, were amplified by PCR using the listed primers (Table S1). The template DNA concentration was measured using NanoDrop 1000 (Thermo Fisher Scientific) and each sample was adjusted to the same concentration.

Cloning and sequencing of CRISPR products

PCR products were separated on 1% agarose gels and bands larger than those produced from the uninfected control sample were excised from gels and purified with QIAquick Gel Extraction Kit (Qiagen). The PCR products were then cloned using InsTAclone™ PCR Cloning Kit (Thermo Fisher Scientific) following the manufacturer's protocol. Plasmid purification and sequencing were performed by GATC Biotech AG (Konstanz, Germany).

Transmission electron microscopy

Virus particles were adsorbed onto carbon-coated copper grids for 5 min and stained with 2% uranyl acetate. Images were recorded using a Tecnai G2 transmission electron microscope (FEI, Eindhoven, the Netherlands), with a CCD camera, at an acceleration voltage of 120 kV.

Identification of viral protein ORF114

Purified virus particles were dissolved in loading buffer, 25 mM Tris-HCl, pH 8.0, 12.5% glycerol, 50 mM DTT, 0.01% bromophenol blue, 2.5% SDS, heated at 95°C for 5 min, and electrophoresed in an 15% polyacrylamide gel with 0.1% SDS. The resolved protein bands were blotted onto a PVDF membrane (Millipore, Billerica, USA) and the membrane was stained with Ponceau red. The protein band corresponding to ORF114 was excised and identified by N-terminal Edmann sequencing (Alphalyse, Odense, Denmark).

Purification of ORF114 and DNA binding studies

The ORF114 gene was amplified by PCR from DNA isolated from purified SMV1 particles and cloned into pET28d

(Erdmann *et al*., 2013). The ORF114 forward primer, 5′-ACGGAATTCATGGTGGATGAGTATTTC-3′ carried an additional EcoRI restriction site. The reverse primer 5′-GCTCTCGAGACATTGATAAGCCATCTG-3′ contained an XhoI restriction site that was used for cloning. The GST fusion construct of ORF114 was generated using forward primer 5′-ACGGAATTCACATGGTGGATGAGTATTTC-3′ and reverse primer 5′-GCTCTCGAGATTATTGATAAGCCATCTG-3′. The PCR product was cleaved with EcoRI and XhoI and ligated into the linearized pGEX6P-2 vector. Expression and purification of ORF114 and pull-down experiments were performed as described (Erdmann *et al*., 2014). DNA binding experiments were performed for 30 min at 50°C in 50 mM NaH_2PO_4, pH 6.4 containing 150 mM KCl, 2.5 mM MgCl and 1% glycerol. The DNA substrate was generated by PCR using forward primer 5′-ATAAACAGCTGTCCTATCC-3′ and reverse primer 5′-TCTTCAAGAATGAGCAAAC-3′ from a genomic library of the *Acidianus* two-tailed virus (Häring *et al*., 2005). It was 5′ end-labelled using ^{32}P-ATP (Perkin Elmer, Waltham, USA) and T4 polynucleotide kinase (Thermo Fisher Scientific). Protein-bound and free DNA were separated in an non denaturing 9% polyacrylamide gel and DNA bands were visualized autoradiographically.

Acknowledgements

The research was supported by the Danish Natural Science Research Council and S.E. also received grants from Copenhagen University. Steffi Krause and Jörg Watzke helped with virus sampling on Iceland and we are very grateful to Xu Peng for providing environmental samples from Naples, Italy. Shiraz A. Shah contributed with spacer and PAM sequence analyses, Ling Deng facilitated the SMV1 genome assembly and Qunxin She providing the STSV2 virus. We are grateful to all, and Jan Christiansen, for their helpful advice and stimulating discussions.

References

Arnold, H.P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Utterback, T., *et al*. (2000) A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus*. *Virology* **267:** 252–266.

Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., *et al*. (2011) A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol Microbiol* **79:** 484–502.

Barrangou, R., and van der Oost, J. (eds) (2013) *CRISPR-Cas Systems*. Heidelberg: Springer Press, pp. 1–291.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., *et al*. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315:** 1709–1712.

Beloglazova, N., Brown, G., Zimmerman, M.D., Proudfoot, M., Makarova, K.S., Kudritska, M., *et al*. (2008) A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *J Biol Chem* **283:** 20361–20371.

Bondy-Denomy, J., Pawluk, A., Maxwell, K.L., and Davidson, A.R. (2012) Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493:** 429–432.

Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., and Semenova, E. (2012) Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun* **3:** 945.

Deng, L., Zhu, H., Chen, Z., Liang, Y.X., and She, Q. (2009) Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. *Extremophiles* **13:** 735–746.

Deng, L., Garrett, R.A., Shah, S.A., Peng, X., and She, Q. (2013) A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol Microbiol* **87:** 1088–1099.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., *et al*. (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* **190:** 1390–1400.

Díez-Villaseñor, C., Guzmán, N.M., Almendros, C., García-Martínez, J., and Mojica, F.J.M. (2013) CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli*. *RNA Biol* **10:** 792–802.

Erdmann, S., and Garrett, R.A. (2012) Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Mol. Microbiol*. 85: 1044–1056. Corrigendum. *Mol Microbiol* **86:** 757.

Erdmann, S., Shah, S.A., and Garrett, R.A. (2013) SMV1 virus-induced CRISPR spacer acquisition from the conjugative plasmid pMGB1 in *Sulfolobus solfataricus* P2. *Biochem Soc Trans* **41:** 1449–1458.

Erdmann, S., Chen, B., Huang, X., Deng, L., Liu, C., Shah, S.A., *et al*. (2014) A novel single-tailed fusiform *Sulfolobus* virus STSV2 infecting model *Sulfolobus* species. *Extremophiles* **18:** 51–60.

Garrett, R.A., Vestergaard, G., and Shah, S.A. (2011) Archaeal CRISPR-based immune systems: exchangeable functional modules. *Trends Microbiol* **19:** 549–556.

Goulet, A., Vestergaard, G., Felisberto-Rodrigues, C., Campanacci, V., Garrett, R.A., Cambillau, C., *et al*. (2010) Getting the best out of long-wavelength X-rays: *de novo* chlorine/sulfur SAD phasing of a structural protein from ATV. *Acta Crystallogr D Biol Crystallogr* **D66:** 304–308.

Gudbergsdottir, S., Deng, L., Chen, Z., Jensen, J.V., Jensen, L.R., She, Q., *et al*. (2011) 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:** 35–49.

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

Häring, M., Rachel, R., Peng, X., Garrett, R.A., and Prangishvili, D. (2005) Viral diversity in hot springs of Pozzuoli, Italy, including a unique bottle-shaped virus ABV from a new family the *Ampullaviridae*. *J Virol* **79:** 9904–9911.

Lillestøl, R.K., Redder, P., Garrett, R.A., and Brügger, K. (2006) A putative viral defence mechanism in archaeal cells. *Archaea* **2:** 59–72.

Lillestøl, R.K., Shah, S.A., Brügger, K., Redder, P., Phan, H., Christiansen, J., *et al*. (2009) CRISPR families of the

crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. *Mol Microbiol* **72:** 259–272.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., *et al*. (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9:** 467–477.

Makarova, K.S., Anatharaman, V., Aravind, L., and Koonin, E.V. (2012) Live virus-free or die: coupling of antivirus immunity and programmed suicide or dormancy in prokaryotes. *Biol Direct* **7:** 40.

Manica, A., Zebec, Z., Teichmann, D., and Schleper, C. (2011) *In vivo* activity of CRISPR-mediated virus defence in a hyperthermophilic archaeon. *Mol Microbiol* **80:** 481–491.

Manica, A., Zebec, Z., Steinkellner, J., and Schleper, C. (2013) Unexpectedly broad target recognition of the CRISPR-mediated virus defence system in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res* **41:** 10509–10517.

Nam, K.H., Ding, F., Haitjema, C., Huang, Q., Delisa, M.P., and Ke, A. (2012) Double-stranded endonuclease activity in *Bacillus halodurans* clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas2 protein. *J Biol Chem* **287:** 35943–35952.

Paez-Espino, D., Morovic, W., Sun, C.L., Thomas, B.C., Ueda, K., Stahl, B., *et al*. (2013) Strong bias in the bacterial CRISPR elements that confer immunity to phage. *Nat Commun* **4:** 1430.

Peng, N., Xia, Q., Chen, Z., Liang, Y.X., and She, Q. (2009) An upstream activation element exerting differential transcriptional activation on an archaeal promoter. *Mol Microbiol* **74:** 928–939.

Peng, X., Blum, H., She, Q., Domdey, H., Mallok, S., Brügger, K., *et al*. (2001) Sequences and replication of the archaeal rudiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses. *Virology* **291:** 226–234.

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

Prangishvili, D., and Garrett, R.A. (2005) Viruses of hyperthermophilic Crenarchaea. *Trends Microbiol* **13:** 535–542.

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

Prangishvili, D., Vestergaard, G., Häring, M., Aramayo, R., Basta, T., Rachel, R., and Garrett, R.A. (2006b) Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle. *J Mol Biol* **359:** 1203–1216.

Richter, C., Gristwood, T., Clulow, J.S., and Fineran, P.C. (2012) *In Vivo* Protein interactions and complex formation in the *Pectobacterium atrosepticum* subtype I-F CRISPR/Cas system. *PLoS ONE* **7:** e49549.

Scheele, U., Erdmann, E., Ungewickell, E.J., Felisberto-Rodrigues, C., Ortiz-Lombardía, M., and Garrett, R.A. (2011) Chaperone role for proteins p618 and p892 in the extra-cellular tail development of *Acidianus* two-tailed virus ATV. *J Virol* **85:** 4812–4821.

Shah, S.A., and Garrett, R.A. (2011) CRISPR/Cas and Cmr modules – mobility and evolution of an adaptive immune system. *Res Microbiol* **162:** 27–38.

Shah, S.A., Hansen, N.R., and Garrett, R.A. (2009) Distributions of CRISPR spacer matches in viruses and plasmids of crenarchaeal acidothermophiles and implications for their inhibitory mechanism. *Biochem Soc Trans* **37:** 23–28.

Shah, S.A., Vestergaard, G., and Garrett, R.A. (2011) *CRISPR/Cas and CRISPR/Cmr Immune Systems of Archaea. Regulatory RNAs in Prokaryotes*. Marchfelder, A., and Hess, W. (eds). Vienna: Springer Press, pp. 163–181.

Shah, S.A., Erdmann, S., Mojica, F.J.M., and Garrett, R.A. (2013) Protospacer motifs: mixed identities and functional diversity. *RNA Biol* **10:** 891–899.

Swarts, D.C., Mosterd, C., van Passel, M.W., and Brouns, S.J. (2012) CRISPR interference directs strand specific spacer acquisition. *PLoS ONE* **7:** e35888.

Terns, M.P., and Terns, R.M. (2011) CRISPR-based adaptive immune systems. *Curr Opin Microbiol* **14:** 321–327.

Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S.M., Ma, W., and Doudna, J.A. (2009) Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* **17:** 904–912.

Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482:** 331–338.

Yosef, I., Goren, M.G., and Qimron, U. (2012) Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res* **40:** 5569–5576.

Yosef, I., Shitrit, D., Goren, M.G., Burstein, D., Pupko, T., and Qimron, U. (2013) DNA motifs determining the efficiency of adaptation into *Escherichia coli* CRISPR array. *Proc Natl Acad Sci USA* **110:** 14396–14401.

Zerbino, D.R., and Birney, E. (2008) Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* **18:** 821–829.

Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brügger, K., Graham, S.J., *et al*. (2012) Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol Cell* **45:** 303–313.

Zhang, Y., Heidrich, N., Ampattu, B.J., Gunderson, C.W., Seifert, H.S., Schoen, C., *et al*. (2013) Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol Cell* **50:** 488–503.

Zillig, W., Kletzin, A., Schleper, C., Holz, I., Janekovic, D., Hain, J., *et al*. (1994) Screening for Sulfolobales, their plasmids and their viruses in Icelandic solfataras. *Syst Appl Microbiol* **16:** 109–128.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.