ORF4 of the temperate archaeal virus SNJ1 governs the lysis-lysogeny switch and superinfection immunity

Beibei Chen, Zhao Chen, Yuchen Wang, Han Gong, Linshan Sima, Jiao Wang, Shushan Ouyang, Wenqiang Gan, Mart Krupovic, Xiangdong Chen, et al.

ORF4 of the temperate archaeal virus SNJ1 governs the lysis-lysogeny switch and superinfection immunity

Beibei Chen[1], Zhao Chen[1], Yuchen Wang[1], Han Gong[1], Linshan Sima[1], Jiao Wang[1], Shushan Ouyang[1], Wenqiang Gan[1], Mart Krupovic[2], Xiangdong Chen[1]*, Shishen Du[1]*

[1] *State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan 430072, China*

[2] *Department of Microbiology, INSTITUT PASTEUR,25 rue du Dr Roux, 75015 Paris, FRANCE*

* Correspondence

E-mail: ssdu@whu.edu.cn; xdchen@whu.edu.cn

Abstract

Recent environmental and metagenomic studies have considerably increased the repertoire of archaeal viruses and suggested that they play important roles in nutrient cycling in the biosphere. However, very little is known about how they regulate their life cycles and interact with their hosts. Here, we report that the life cycle of the temperate haloarchaeal virus SNJ1 is controlled by the product ORF4, a small protein belonging to the antitoxin MazE superfamily. We show that ORF4 controls the lysis-lysogeny switch of SNJ1 and mediates superinfection immunity by repression of genomic DNA replication of the superinfecting viruses. Bioinformatic analysis shows that ORF4 is highly conserved in two SNJ1-like proviruses, suggesting that the mechanisms for lysis-lysogeny switch and superinfection immunity are conserved in this group of viruses. As lysis-lysogeny switch and superinfection immunity of archaeal viruses are poorly studied, we suggest that SNJ1 could serve as a model system to study these processes.

Importance

Archaeal viruses are important parts of the virosphere. Understanding how they regulate their life cycles and interact with host cells provide crucial insights into their biological functions and the evolutionary histories of viruses. However, mechanistic studies of the life cycle of archaeal viruses are scarce due to a lack of genetic tools and demanding cultivation conditions. Here, we discover that the temperate haloarchaeal virus SNJ1, which infects *Natrinema* sp. J7, employs a lysis-lysogeny

switch and establishes superinfection immunity like bacteriophages. We show that its ORF4 is critical for both processes and acts as a repressor of the replication of SNJ1.These results establish ORF4 as a master regulator of SNJ1 life cycle and provides novel insights on the regulation of life cycles by temperate archaeal viruses and on their interactions with host cells.

Introduction

The study of archaeal viruses is still in its infancy compared to the widely reported bacteriophages (about 8,500 species) (1). Only about 120 species of archaeal viruses have been sequenced and investigated to date; however, they display remarkable diversity in virion morphologies and gene contents (2-4). Temperate life cycles are common among viruses of bacteria and archaea, whereby viral genomes are stably maintained within the cell either as extrachromosomal episomes or are integrated into the host chromosomes. This mode of propagation provides ample opportunities for the virus and its host to co-evolve (5, 6). Therefore, research on the viral life cycle and the interaction between temperate viruses and their hosts can provide crucial insights into the evolutionary histories of viruses and novel modes of virus-host interactions (7, 8). However, due to the lack of suitable genetic tools and demanding cultivation conditions, only a handful of archaeal temperate viruses have been studied experimentally so far (9-17).

Unlike temperate archaeal viruses, temperate bacteriophages are ubiquitous and the regulation of their life cycle has been extensively studied (18-20). A logical parallel may be drawn from these studies: temperate viruses have a dual life cycle, which can terminate either in host cell lysis or lysogeny. If the lytic pathway is chosen, the virus actively replicates its genome, produces progeny virions and lyses the host cell. By contrast, if the lysogenic pathway ensues, the viral genome will be present in a dormant state either as an extrachromosomal element or as a provirus integrated into the host chromosome, and the expression of the genes responsible for the lytic

66 pathway is repressed. Temperate viruses often confer immunity to their host cells
67 against subsequent infection by the same and closely related viruses, a phenomenon
68 known as superinfection immunity (21, 22). When host cell fitness decreases, such as
69 upon exposure to DNA damaging agents, the provirus switches to the lytic pathway
70 (23). The switch between the lytic and lysogenic states of temperate bacteriophages is
71 very sophisticated and is normally controlled by a master regulator which senses the
72 fitness of the host cell and represses the expression of the lytic pathway genes. The cI
73 protein of the *Escherichia coli* phage λ is a prototypical master regulator which
74 represses the expression of the lytic pathway genes and is cleaved under conditions of
75 DNA damage (23-26). Among archaeal viruses, the switch between lytic and
76 temperate life cycles has been explored in the case of the spindle-shaped virus SSV1
77 infecting hyperthermophilic archaea of the genus *Sulfolobus* (15-17, 27). It has been
78 suggested that SSV1 protein F55 binds to host protein RadA and this complex
79 represses the genes responsible for active virus replication; dissociation of the
80 F55-RadA complex upon DNA damage leads to transcriptional derepression and
81 active replication of SSV1 (17). However, whether such a mechanism is conserved in
82 other archaeal viruses, especially in viruses infecting distantly related archaea, such as
83 halophiles, remains unknown.

84 The temperate haloarchaeal virus SNJ1 isolated from *Natrinema* sp. J7-1 resides in
85 the cytoplasm as a circular plasmid called pHH205 (9, 28). A small amount of virions
86 [with a titer of about 10^6 plaque-forming unit (pfu) mL^{-1}] is released during
87 cultivation, presumably due to spontaneous induction in some cells (Supplementary

Fig. 1). Upon treatment of J7-1 with the DNA damaging agent mitomycin C (MMC), SNJ1 is triggered to undergo the lytic life cycle and up to 10^{10} pfu mL^{-1} of SNJ1 virions are produced. Released SNJ1 viruses can infect a strain cured from SNJ1, named *Natrinema* sp. CJ7 (10), and forms turbid plaques on lawns of CJ7. Here we investigated how SNJ1 controls its life cycle and identified its ORF4 as a critical regulator.

Results

The discovery of clear plaque mutants of SNJ1

Our lab previously used SNJ1's proviral genome, plasmid pHH205, to construct a series of *E. coli-Natrinema* shuttle vectors (29). One of these vectors, pYC-S, was obtained by insertion of the *E. coli* vector pUC19-mev at the *Sac*I site of the SNJ1 genome (Supplementary Fig. 2). Since pYC-S contains the whole genome of SNJ1, we tested whether infectious viruses could be produced from a SNJ1-cured strain carrying pYC-S (CJ7/pYC-S) upon MMC treatment using a double-layer plaque assay. As shown in Fig. 1a, many plaques formed on the lawn of CJ7 using supernatants from MMC treated CJ7/pYC-S culture, suggesting that SNJ1 progeny viruses were generated. However, unlike the homogenous turbid plaques formed by wild type SNJ1 viruses, viruses produced using pYC-S formed two kinds of plaques, one was similar to the turbid plaques formed by wild type SNJ1, while the other was much bigger and clearer (Fig. 1a). This suggested that viruses produced using pYC-S were a mixture of

viruses. Moreover, viruses purified from the turbid or clear plaques maintained their plaque morphotypes on the lawns of CJ7 (Fig. 1a bottom). In liquid culture of CJ7, addition of the turbid-plaque forming viruses slowed down the growth of CJ7 similarly to wild type SNJ1, while clear-plaque forming viruses completely blocked its growth (Fig. 1b). Consistent with this, we were unable to obtain stable lysogen of the clear-plaque SNJ1 mutant viruses, whereas lysogens of the turbid-plaque SNJ1 mutant viruses were obtained at the same frequency as wild type SNJ1, suggesting that the clear-plaque SNJ1 mutant viruses could only undergo the lytic pathway. Since disruption of the lysis-lysogeny switch of temperate viruses frequently resulted in the appearance of clear plaques, we suspected that insertion of the pUC19-mev fragment at the *Sac*I site of SNJ1 disrupted the lysis-lysogeny life cycle of SNJ1.

ORF4 is critical for the lysis-lysogeny switch of SNJ1

To determine how insertion of foreign DNA disrupted the life cycle of SNJ1, we first used PCR to check the genomic DNA (gDNA) integrity of the mutant viruses generated from pYC-S. Interestingly, we found that a primer pair designed to amplify the pUC19-mev region of pYC-S always generated products smaller than its expected size using mutant SNJ1 viruses as templates. This suggested that parts of the pUC19-mev fragment were lost in the SNJ1 mutant viruses. To confirm this, the pUC19-mev region and its flanking sequences from 20 clear or turbid-plaque viruses were amplified by PCR (HJ-test-F/R primers listed in Supplementary Table 2) and sequenced. As shown in Fig. 2, all of the sequenced viral gDNA had varying degrees of deletion of the pUC19-mev fragment; some also contained deletions in the flanking

sequences from SNJ1. The total length of pYC-S (21,425 bp) was 30% longer than the length of wild type SNJ1 genome (16,492 bp, Supplementary Fig. 2) and it had been reported that phage packaging systems never pack exceed about 10% over the parental gDNA (30, 31). Hence, it was unsurprising that all the gDNA of mutant viruses generated from pYC-S contained deletions. Examination of the length of the mutant viruses genomes showed that they ranged from 16,010 bp to 18,054 bp (97% to 109% of the genome length), suggesting that the capsid of SNJ1 could tolerate a maximum of approximately 1.5 kb of foreign DNA.

Strikingly, the *orf4* gene of SNJ1 (nucleotide sequence 499-293) was disrupted in all clear-plaque SNJ1 viral genomes, whereas it was intact in every turbid-plaque mutant virus (Fig. 2), suggesting that disruption of *orf4* resulted in the appearance of clear plaque forming viruses. To test this, the plaque morphologies of ORF4-disrupted viruses were determined using viruses induced from CJ7/pYC-S-4M (start codon mutation of *orf4*, SNJ1$^{orf4\ mut}$) and CJ7/pYC-S Δ1-575 (deletion mutation, SNJ1$^{\Delta orf4}$). As shown in Fig. 3a, SNJ1$^{orf4\ mut}$ and SNJ1$^{\Delta orf4}$ viruses formed only clear plaques on lawns of CJ7, suggesting that ORF4 controlled the lysis-lysogeny switch of SNJ1. To further test this, we checked whether expression of ORF4 in trans would restore turbid plaque formation of the clear-plaque SNJ1 mutant viruses. For this test, we used a plasmid containing *orf4* and its co-transcribed downstream gene *orf3* as well as flanking sequences of this operon (1-656) to preserve its promoter and elements critical for its expression. As shown in Fig. 3b, plaque formation of the SNJ1$^{orf4\ mut}$ and SNJ1$^{\Delta orf4}$ viruses was significantly inhibited on lawns of CJ7 expressing *orf4*

153 (with plasmid pFJ6-1-656), although some turbid plaques could form at the 10^0 to 10^{-3}
154 dilutions. On lawns of CJ7 containing the empty vector, SNJ1$^{orf4\ mut}$ and SNJ1$^{\Delta orf4}$
155 viruses formed clear plaques up to the 10^{-7} dilutions. Notably, the complementation of
156 ORF4 in CJ7 also blocked the plaque formation of SNJ1 completely. These results
157 demonstrated that ORF4 was critical for the lysis-lysogeny switch of SNJ1 and
158 suggested that it also mediated superinfection immunity of the temperate virus (see
159 below).

160 We previously found that inactivation of ORF4 affected the stability and copy number
161 of SNJ1 based *E. coli-Natrinema* shuttle vectors (29), suggesting that it was likely
162 also critical for the maintenance of the SNJ1 lysogenic state. Consistent with previous
163 observations, we found that in the presence of ORF4, the shuttle vector pYC-SHS,
164 which contained the replication and regulatory elements of SNJ1, was stably
165 maintained at about 2 copies per chromosome. However, in the absence of ORF4
166 (pYC-SHS-4M), the copy number of this plasmid increased to 20 copies per
167 chromosome and was extremely unstable (Fig. 4). Expression of ORF4 in trans
168 (pFJ6-1-656) restored stability and decreased the copy number of pYC-SHS-4M,
169 confirming the role of ORF4 in plasmid stability and copy number control. More
170 strikingly, we found that the copy number of pYC-SHS (with *orf4*) increased from 2
171 copies to 25 copies per chromosome upon MMC treatment, whereas the copy number
172 of pYC-SHS-4M (without *orf4*) was only slightly affected by MMC (35 copies
173 comparing to 20 copies per chromosome). Since pYC-SHS contained the gene
174 clusters required for SNJ1 replication and regulation, the mechanism responsible for

the induction of SNJ1 during MMC treatment of J7-1 was likely the same. Thus, these results suggested that ORF4 acted as a repressor of replication of SNJ1 to maintain it in the lysogenic state. Since the lack of ORF4 and MMC induction both lead to activation of SNJ1 lytic cycle, we speculated that ORF4 may be inactivated upon DNA damage, resulting in the active replication of SNJ1.

ORF4 mediates superinfection immunity of SNJ1

The above finding (Fig. 3b) that the presence of ORF4 in CJ7 cells inhibited plaque formation of SNJ1 viruses suggested that SNJ1 had immunity against a second infection by itself and ORF4 played a critical role. To confirm this, we infected CJ7 or strain J7-1 with SNJ1 viruses. As shown in Fig. 5a, while SNJ1 formed plaques on CJ7, J7-1 was completely immune to the infection of SNJ1, confirming that SNJ1 establishes superinfection immunity. Moreover, the presence of a plasmid carrying the *orf3-4* operon in CJ7 completely blocked infection by SNJ1 and inactivation of ORF4, but not ORF3, by a frame shift mutation abolished the ability to block SNJ1 infection. This confirmed ORF4 is the critical regulator of superinfection immunity.

orf4 was predicted to encode a 7.7 kDa protein possessing 68 residues. Western blot of a GFP fusion of ORF4 confirmed that it encoded for a protein, although the fusion protein was not functional in superinfection immunity (Supplementary Fig. 3). Bioinformatic analysis of ORF4 revealed that it belonged to the MazE antitoxin superfamily and showed homology with the *Bacillus subtilis* sporulation regulatory protein SpoVT and the C68 protein of the hybrid virus-plasmid pSSVx (Fig. 5b).

196 These proteins usually form dimers through a swapped hairpin domain and act as
197 transcriptional regulators, which is a distinctive structural characteristic of this family
198 (32-35). Protein secondary-structure prediction suggested that ORF4 contained 5
199 β-strands and one α-helix. To determine which part of ORF4 was important for
200 superinfection immunity, it was truncated and expressed under the promoter of heat
201 shock protein 70 from *Haloferax volcanii* D52 in plasmid pFJ6. Remarkably, a
202 fragment containing only the N-terminal 33 amino acids was sufficient to block SNJ1
203 infection (Fig. 5c). Further truncation of the 33 amino acids abolished its ability to
204 restrict SNJ1 infection, demonstrating that the N-terminal 33 amino acids of ORF4
205 were necessary and sufficient for superinfection immunity.

206 **ORF4 represses DNA replication of SNJ1**

207 Generally, there are five stages in a viral life cycle: adsorption, gDNA ejection,
208 macromolecular synthesis, packaging and release. ORF4 may inhibit any step of the
209 SNJ1 viral life cycle to prevent its replication. We first checked whether adsorption of
210 SNJ1 to the host cells was blocked by ORF4. To do this, SNJ1 was incubated with
211 cells with or without SNJ1 (CJ7, J7-1, respectively) or cells with or without ORF4
212 expression (CJ7-F/pFJ6-MCS and CJ7-F/pFJ6-Hpro-*orf4*) at multiplicity of infection
213 (MOI) of 0.1, 1 and 10 for 1 h. After incubation, the titer of unabsorbed viruses was
214 quantified by double-layer plaque assay and compared to the titer of the viruses
215 before the adsorption assay. The difference of the titers before and after the adsorption
216 assay was considered as the adsorption efficiency. As shown in Fig. 6a, there was no
217 obvious difference in adsorption efficiency in the presence or in the absence of SNJ1,

218 nor with or without ORF4 expression at low or high MOI. Therefore, we concluded
219 that ORF4 did not act by preventing viral adsorption.

220 Next, we tested whether viral gDNA ejection was inhibited by ORF4. The relative
221 copy number of viral gDNA in the host cells with or without ORF4 expression were
222 quantified 1 h post infection by qPCR. The single-copy gene *radA* located on the
223 chromosome was used as a reference, while *orf14* of SNJ1 was used for viral gDNA.
224 As shown in Fig. 6b, at a MOI of 0.1, copies of viral gDNA in cells without ORF4
225 (0.0065 in CJ7, 0.0068 in CJ7/pFJ6-MCS) was about 3-fold higher than cells with
226 ORF4 (0.0020 in CJ7-F/pFJ6-Hpro-*orf4*). The difference increased to about 6-fold at
227 a MOI of 1 (0.0691, 0.0601 *vs* 0.0106) and 9-fold at a MOI of 10 (0.8877, 0.8100 *vs*
228 0.1080). These results suggested that ORF4 may prevent gDNA from ejecting into the
229 host cells. However, we could not exclude the possibility that the ejected gDNA had
230 been degraded in the presence of ORF4 given that this was examined 1 hour post
231 infection. It was also possible that ejection of gDNA was not affected, but replication
232 of the ejected gDNA was suppressed by ORF4 (see below).

233 To test whether replication of SNJ1 was repressed by ORF4, we quantified the copy
234 number of gDNA of SNJ1 in the infected cells every hour post infection for 9 h. For a
235 better comparison, we infected cells without ORF4 (CJ7 and CJ7/pFJ6-MCS) at a
236 MOI of 0.1, while cells with ORF4 (CJ7-F/pFJ6-Hpro-*orf4*) at a MOI of 0.5 to ensure
237 similar gDNA copy number of SNJ1 at 1 h post infection (0.0046, 0.0039 and 0.0038
238 copies per chromosome, respectively). As shown in Fig. 6c, in the absence of ORF4
239 (strain CJ7 and CJ7/pFJ6-MCS), the viral gDNA duplicated rapidly over time and

reached up to 2.5280 or 2.0998 copies per chromosome 9 h post infection. However, the viral gDNA was barely replicated in the strain CJ7-F/pFJ6-Hpro-*orf4* (0.0582 copies per chromosome). These results showed that gDNA replication in the host cell was significantly repressed in the presence of ORF4.

To ensure that ORF4 repressed replication of the SNJ1 gDNA instead of degrading it, we repeated the experiment by southern blot analysis using probes that specifically recognized SNJ1. Note that restriction digestion of the gDNA was not carried out in this assay to preserve the integrity of the sample and as a consequence multiple bands of SNJ1 gDNA could show up as the DNA molecule could be supercoiled. Consistent with the above results, gDNA bands were detected in strain CJ7 and CJ7/pFJ6-MCS 1 h post infection and accumulated extensively as time increased in the absence of ORF4 (Fig. 6d). However, only a weak signal of SNJ1 genome was observed in the ORF4 expressing strain CJ7-F/pFJ6-Hpro-*orf4* even 9 h post infection. More importantly, SNJ1 genome remained intact during the whole cultivation process, suggesting that ORF4 did not prevent replication by degrading viral DNA. In conclusion, these results suggested that ORF4 precluded SNJ1 infection through inhibition of viral gDNA replication.

ORF4 is highly conserved in SNJ1-like proviruses

Because ORF4 played such an important role in the SNJ1 life cycle, its presence in a genome may indicate the existence of a SNJ1-like provirus. Thus, we searched for ORF4-like proteins using a protein-protein BLAST algorithm. We found that an ORF

in *Natrinema versiforme* strain BOL5-4 plasmid pNVE19 and an ORF in *Haloterrigena jeotgali* strain A29 plasmid unnamed5 shared high homologies to ORF4 (94% and 83% identity, respectively). When the genomes of these two plasmids were aligned with the genome of SNJ1, we found that they were highly similar, indicating that they could be SNJ1-like proviruses (Fig. 7). First, their genomes were about the same size and encoded roughly similar number of ORFs. BOL5-4 plasmid pNVE19 was 18,925 bp and contained 28 putative genes (GenBank Acc. No. NZ_CP040333), while A29 plasmid unnamed5 was 17,189 bp and contained 31 putative genes (GenBank Acc. No. CP031302)(36). Second, most of the ORFs on these two plasmids were homologous to ORFs of SNJ1 and arranged into clusters similarly as SNJ1. Most importantly, genes homologous to the ORFs important for SNJ1 replication shared identity over 40% and homologs of the core structural elements for *Sphaerolipoviridae* were also present in these two plasmids (37). These included homologs of the two major capsid proteins and the putative packaging ATPase of SNJ1 (colored blue and pink, respectively). Based on these features, these two plasmids were likely SNJ1-like proviruses and ORF4-like protein may play a critical role in the regulation of their life cycles.

Discussion

Overall, our results showed that ORF4 is a master regulator of the life cycle of SNJ1 (Fig. 8). It not only regulates the lysis-lysogeny switch of SNJ1, but also mediates superinfection immunity, two phenomena which are presumably widely present but

283 have rarely been studied for archaeal viruses. Sequence analysis of ORF4 showed that
284 it contains the SpoVT/AbrB-like domain and shares homology with proteins
285 belonging to the antitoxin MazE superfamily. Proteins of the SpoVT/AbrB family
286 normally act as transcriptional regulators (38-40). MazE is the antidote to the toxin
287 MazF, and MazE-MazF in *E. coli* is a regulated prokaryotic chromosomal addiction
288 module. MazE is a labile protein that is degraded by ClpAP serine protease (41, 42).
289 We suspect that ORF4 works as a transcriptional repressor of the genes responsible
290 for the lytic pathway.

291 The expression level of ORF4 in the host cell appears to be critical for determining
292 the route of the SNJ1 infection. In the absence of ORF4, SNJ1 formed clear plaques,
293 whereas in its presence, SNJ1 formed turbid plaques, and with its ectopic expression,
294 SNJ1 infection was blocked. This suggests that the lysis-lysogeny switch is controlled
295 by a ratio of ORF4 with a yet to be discovered SNJ1-encoded protein(s). The
296 observations that MMC treatment and disruption of ORF4 both resulted in the active
297 replication of SNJ1 and increased the copy number of SNJ1-based plasmids indicate
298 that DNA damage in the host cell may lead to the inactivation of ORF4. Presumably,
299 ORF4 is cleaved and/or degraded by host proteases in a manner similar to the
300 cleavage/degradation of the cI repressor of λ phage (23). It is also possible that ORF4
301 dissociates from its target sites like the F55 repressor of SSV1 (43). Future studies to
302 test the ability of ORF4 to bind DNA and regulate gene expression as well as its
303 inactivation upon DNA damage in the host cell will undoubtedly provide critical
304 missing information on the mechanism of the lysis-lysogeny switch of SNJ1.

Superinfection immunity has been well documented for many bacteriophages and eukaryotic viruses (44-46). However, only a couple of cases have been reported for archaeal viruses. It was found that the adsorption of *Sulfolobus islandicus* rod-shaped virus 2 (SIRV2) to its host cells was significantly reduced by a previous infection, indicating that SIRV2 established superinfection exclusion (47). However, the exact mechanism underlying this phenomenon has not yet been elucidated. In addition, hyperthermophilic archaeal viruses SPV1 and SPV2 carry mini-CRISPR arrays containing spacers against each other and, consistently, seem to restrict each others replication. Thus, it has been suggested that virus-borne mini-CRISPR arrays might represents a distinct mechanism of heterotypic superinfection exclusion (48). Here we present evidences that lysogenic SNJ1 confers immunity against subsequent infections and identify ORF4 as a critical factor in this process by repressing SNJ1 genome replication. However, we cannot exclude the possibility that it may also block gDNA ejection into the host cells because our results showed that at early stage of infection (1 h post infection) gDNA in cells without ORF4 were remarkably higher than cells with ORF4. Strikingly, we found that only the N-terminal 33 amino acids of ORF4 are necessary and sufficient for blocking SNJ1 infection. It is likely that the N-terminal 33 amino acids of ORF4 is sufficient to bind DNA and repress transcription of the lytic pathway genes. However, biochemical characterization of ORF4 is necessary to confirm this and reveal its mechanism of action.

In conclusion, we showed that the temperate haloarchaeal virus SNJ1 employs a lysis-lysogeny switch and establishes superinfection immunity. We also identified

ORF4 as a key factor for both of these processes and found that it acts as a repressor of SNJ1 replication. We suggest that SNJ1 could be used as a model system to understand the regulation of archaeal viral life cycle and virus-host interactions. Using the SNJ1 mutant viruses and *E. coli-Natrinema* shuttle vectors generated in this study, many other viral proteins of SNJ1 and host factors necessary for the regulation of SNJ1 life cycle could be discovered. Characterization and elucidation of how these proteins and ORF4 work should greatly increase our understanding of the enigmatic archaeal virosphere.

Materials and Methods

Strains, culture conditions and transformation methods

All strains, plasmids and primers used in this study were listed in Supplementary Table 1-3. Different derivatives of *Narinema* sp. J7 were cultured in Halo-2 medium at 45℃ as described before (10). Halo-2 medium contained 250 g NaCl, 30 g $MgCl_2 \cdot 6H_2O$, 2.5 g Lactalbumin hydrolysate (Difco Laboratories) and 2 g Bacto yeast extract (Difco Laboratories) per liter of water. Casamino Acids medium (Hv-Ca) or 18% modified growth medium (MGM) were prepared as described before (29, 49). Hv-Ca contained 144 g NaCl, 18 g $MgCl_2 \cdot 6H_2O$, 21 g $MgSO_4 \cdot 7H_2O$, 4.2 g KCl, 5 g Amicase (Sigma), 0.5 g $CaCl_2$ and 30 mL 1M Tris-HCl (pH = 7.5) per liter of water. 18% MGM contained 144 g NaCl, 18 g $MgCl_2 \cdot 6H_2O$, 21 g $MgSO_4 \cdot 7H_2O$, 4.2 g KCl, 5 g Peptone (Difco Laboratories), 3 g Bacto yeast extract (Difco Laboratories), 0.5 g $CaCl_2$ and 30 mL 1M Tris-HCl (pH = 7.5) per liter of water.

Transformation of *Narinema* sp. CJ7 and CJ7-F was performed using the modified polyethylene glycol method as described previously (50, 51). CJ7 transformed with pYC vectors were plated on 18% MGM plates with 5 μg mL^{-1} mevinolin (Mev), while CJ7-F transformed with pFJ6 vectors were selected on Hv-Ca medium. *E. coli* were cultured in Luria-Bertani medium at 37℃, supplemented with ampicillin (0.1 mg mL^{-1}) when necessary. *E. coli* were transformed following the $CaCl_2$ method (52). Solid and soft agar media were prepared by adding 12 g L^{-1} and 5 g L^{-1} agar respectively.

Plasmid construction

To construct pYC-S-4M, overlap extension PCR was performed by primer pairs pYC-S-*Afl*Ⅱ-F/4M-R and 4M-F/pYC-S-*Nco*Ⅰ-R using pYC-S as a template. PCR products were ligated into *Afl*Ⅱ-*Nco*Ⅰ digested pYC-S by Hieff Clone™ Plus One Step Cloning Kit (Yeasen, China). To construct pYC-S Δ1-575, the *Sac*Ⅰ-*Nco*Ⅰ fragment on the pYC-S was replaced with the *Sac*Ⅰ-*Nco*Ⅰ fragment on the pYC-1 by T4 DNA Ligase (Takara). To construct *E. coli*-*Narinema* sp. J7 shuttle vectors containing full-length SNJ1, SNJ1 proviral genomes pHH205 was extracted from strain J7-1 using TIANpure Mini Plasmid Kit (TIANGEN). Purified pHH205 was digested with FastDigest *Sac*I restriction enzyme (ThermoFisher Scientific) and ligated with *Sac*I digested pUC-mev.

Derivatives of pFJ6 plasmids were all constructed by the same approach. For pFJ6-1-656, a DNA fragment containing nucleotide 1-656 of SNJ1 viral genome was

amplified with primer pairs *Sph*Ⅰ-1-F/Mun Ⅰ-656-ORF4-R from J7-1. The product was digested with *Mun*Ⅰ-*Sph*Ⅰ and ligated into *Mun*Ⅰ-*Sph*Ⅰ digested pFJ6-MCS by T4 DNA Ligase (Takara). Frame-shift mutations were introduced by overlap extension PCR using primer pairs *Sph*Ⅰ-1-F/(ORF3-FSM-R, ORF4 FSM-R) and (ORF3-FSM-F, ORF4 FSM-F)/*Mun*Ⅰ-656-ORF4-R. For pFJ6-Hpro, target genes were amplified by *Mun*I-Hpro-F/*Not*I-Hpro-R from pYCJ-H , the product was digested with *Mun*Ⅰ-*Not*Ⅰ and ligated into pFJ6-MCS. For pFJ6-Hpro-*orf4*, pFJ6-Hpro-*orf4*-4FSM and pFJ6-Hpro-$orf4^{x}$, target genes were amplified by *Not*Ⅰ-ORF4-499-F/(*Sph*Ⅰ-ORF4-293-R, *Sph*Ⅰ-ORF4-150-R, *Sph*Ⅰ-ORF4-99-R, *Sph*Ⅰ-ORF4-72-R, *Sph*Ⅰ-ORF4-48-R, *Sph*Ⅰ-ORF4-24-R) and (*Not*Ⅰ-ORF4-FSM-F, *Not*Ⅰ-ORF4-400-F)/*Sph*Ⅰ-ORF4-293-R respectively from J7-1, then the products were digested with *Not*Ⅰ-*Sph*Ⅰ and ligated into pFJ6-Hpro.

For pFJ6-PphaRP-GFP, PCR was performed by primer pairs ORF4-GFP-1-F/ORF4-GFP-3-R using plasmid pRF as a template(53). PCR products were ligated into *Afl*Ⅱ-*Sph*Ⅰ digested pFJ6-MCS by Hieff Clone™ Plus One Step Cloning Kit (Yeasen, China). For pFJ6-PphaRP-ORF4-GFP-1 and pFJ6-PphaRP-ORF4-GFP-2, target genes were amplified by primer pairs ORF4-GFP-1-F/ORF4-GFP-1-R, ORF4-GFP-2-F/(ORF4-GFP-2-R, ORF4-GFP-4-R) and (ORF4-GFP-3-F, ORF4-GFP-4-F)/ORF4-GFP-3-R respectively, and PCR products were ligated into *Afl*Ⅱ-*Sph*Ⅰ digested pFJ6-MCS as well.

All of the plasmids were transformed into *E. coli* DH5α and JM110 successively to get demethylated plasmids. Finally, the plasmids were transformed into *Natrinema* sp.

CJ7 or CJ7-F accordingly.

SNJ1 induction, propagation and infection procedures

SNJ1 was induced from J7-1 strain by MMC (Roche). J7-1 was cultivated in Halo-2 medium at 45℃ for about 24 h to late-exponential phase (OD_{600} ≈ 0.6), and treated with MMC (1 μg mL^{-1}) for 30 min at 37°C with aeration. Cells were collected by centrifugation (10000 rpm, 5 min) and resuspended in the same volume of Halo-2 medium. After 24 h of cultivation at 45℃, 200 rpm, the culture was centrifuged at 10000 rpm for 20 min, and the supernatant was collected and passed through a filter (0.22 μm) to remove cell debris. The titer of the virus stock was measured by double-layer plaque assay as follows.

SNJ1 titers were calculated as pfu mL^{-1} as described previously(54), with minor modifications. 100μL of SNJ1 virus stock dilutions and 400 μL of CJ7 culture (late-exponential phase, OD_{600} ≈ 0.6) were mixed with 4 mL of soft melted Halo-2 medium (top layer), then poured onto Halo-2 solid plates (bottom layer) immediately. The plates were incubated at 37°C for 48 h and counted. If necessary, CJ7 cells at early-exponential growth phase (OD_{600} ≈ 0.3) were collected and infected with SNJ1 (MOI = 10) to promote virus titer.

The mutant virus stocks were obtained from CJ7/pYC-S, CJ7/pYC-S-4M and CJ7/pYC-S Δ1-575 by MMC treated as same as above. Late exponential phase culture of CJ7 was infected with mutant viruses by double-layer plaque assay, then the plates were incubated at 37°C for 48 h and photographed.

Analysis of the lytic ability of different viruses to CJ7

CJ7 was cultured in Halo-2 medium at 45℃ to early-exponential growth phase ($OD_{600} \approx 0.2$). The same amount of viruses (SNJ1, two turbid and clear-plaque viruses from CJ7/pYC-S culture) was added into CJ7 culture and incubated at 45°C for an hour. Cells were collected by centrifugation (10000 rpm, 5 min) and resuspended in the same volume of Halo-2 medium. The cultures were then cultivated at 45°C, 200 rpm and the growth curve was monitored for 49 h post infection by measuring OD_{600}. CJ7 without virus infection was set as a control.

Superinfection immunity assay

Superinfection immunity assay was performed by the double-layer plaque assay with minor modifications. First, 400 μL of late-exponential phase cultures of indicated strains were mixed with soft melted Halo-2 agar (4 mL) and poured onto Halo-2 agar plates immediately. 10 minutes later, when the soft agar was solidified, 2 μL of the ten-fold serial dilutions of SNJ1 virus was spotted onto the plates carefully. The plates were incubated at 37°C for 48 h and photographed.

Western blot assay

CJ7/pFJ6-PphaRP-GFP and CJ7/pFJ6-PphaRP-ORF4-GFP were cultivated in Halo-2 medium to late-exponential phase and crude extracts were taken and loaded for SDS-PAGE. Anti-GFP antibodies (TransGen Biotech) were used for Western blot assay of GFP and ORF4-GFP fusion protein.

Adsorption assay

SNJ1 was incubated with 20 mL of early-exponential phase (OD_{600} ≈ 0.3) of cultures indicated strains at different MOI (0.1, 1 and 10) for about 1 h at 45℃. After absorption, the culture was centrifuged (10000 rpm, 5 min), supernatant was collected and titer of unbound viruses in supernatant was measured by double-layer method. Adsorption efficiency was determined by comparing the titer before and after virus adsorption. Three independent experiments were performed, and error bars indicated standard deviations.

Quantification of viral gDNA

The relative copy number of viral gDNA intercellular 1 h post infection was detected by qPCR method using cells collected from adsorption assay as templates. The templates from different cultures were prepared as described previously (29, 55). Briefly, 1 mL aliquot of the cell culture was harvested by centrifugation, and washed twice by Halo-2 medium to eliminate the free viruses in supernatant. Cells were resuspended in 1 mL of 18% (wt/vol) NaCl solution. To avoid the negative effect of high salt concentration on qPCR, 10 μL of the cell suspension was added to 490 μL of distilled water, which also resulted in rapid cell lysis. A single-copy gene *radA* located on the chromosome of host strain was used as a reference, while *orf14* of SNJ1 gnome was used as target gene. The specific primers used for qPCR were listed in Supplementary Table 3. For the reactions, 20 μL mixtures were prepared containing 5 μL of template, 10 μL of iTaq Universal SYBR Green Supermix (Bio-Rad, USA), 1 μL of primer pairs (10 μM) and 4 μL of distilled water. Amplification was performed according to the manufacturer's instructions. Finally, the qPCR data were analyzed by

456 $2^{-\Delta C_T}$ method (56). Three independent experiments were performed, and error bars
457 indicated standard deviations.

458 **SNJ1 replication assay**

459 SNJ1 was incubated with 100 mL of early-exponential phase ($OD_{600} \approx 0.3$) cultures of
460 CJ7 or CJ7-F/pFJ6-MCS at a MOI of 0.1 and CJ7-F/pFJ6-Hpro-*orf4* at a MOI of 0.5
461 for 1 h at 45℃. After 1 h of absorption, the culture was centrifuged (10000 rpm, 20
462 min) to end the infection procedure. Cell pellets were washed twice with fresh Halo-2
463 medium to remove the free viruses, resuspended in the same volume of Halo-2
464 medium, and incubated at 45℃ for 9 h. Samples were taken every hour post infection.
465 qPCR analyses were performed using the same primer pairs and procedure as in
466 **Quantification of viral gDNA**. The qPCR data were analyzed by $2^{-\Delta C_T}$ method (56)

467 **Genomic integrity of SNJ1 by Southern blot**

468 CJ7, CJ7-F/pFJ6-MCS and CJ7-F/pFJ6-Hpro-orf4 cells were infected by SNJ1 as
469 described in **SNJ1 replication assay** except that SNJ1 was incubated at a MOI of 5
470 with all strains. Multiple aliquots of 8 mL of culture were collected every two hours
471 after 1 h of infection. To remove the unbound viruses, cells were collected by
472 centrifugation (10000 rpm, 5 min) and washed twice by the same volume of Halo-2
473 medium. The total genomic DNA of infected cells was extracted according to the
474 online protocol (http://www.haloarchaea.com/resources/halohandbook) with
475 modifications. Briefly, cell pellets were resuspended gently in 2 mL of distilled water
476 to completely lyse the cells. 2 mL of phenol-chloroform solution was mixed the cell

lysate to extract proteins (and probably some carbohydrate). Finally, the genomic DNA was precipitated by adding 2 volumes of ethanol and dissolved in 150 μl of TE buffer. 5 μL of DNA preparations were electrophoresed on 0.8 % agarose gel. The DNA bands were then transferred onto positively charged nylon membranes with alkaline transfer buffer according to the online protocol (Molecular Cloning chapter 6 protocol 8, http://www.molecularcloning.com). The primers ORF7-F and ORF11-R were designed to synthesize a probe corresponding to the region from 1,483 to 2,537 bp of the SNJ1 genome (Supplementary Table 3). Southern blot was performed as described before (57).

The DNA probe preparation, hybridization and detection were performed using DIG High Prime DNA Labeling and Detection Starter kit I (Roche) according to the manufacturer's instructions.

Determination of plasmid copy number and stability

CJ7/pYC-SHS and CJ7/pYC-SHS-4M were cultured to late-exponential phase (24 h) in 18% MGM + Mev medium and treated with MMC (1 μg mL^{-1}) for 30 min, while cultures without MMC were set as controls. Cells were collected by centrifugation and washed twice in same volume of 18% MGM to remove MMC. Cell pellets were resuspended in 18% MGM + Mev medium and cultured for 24 h. Samples were taken for qPCR analysis using primer pairs vector-F/vector-R, orf14-F/orf14-R and radA-F/radA-R to determine the relative plasmid copy number to chromosome. The qPCR data were analyzed by $2^{-\Delta C_T}$ method (56). Three independent experiments were

performed, and error bars indicated standard deviations.

For plasmid stability assay of pYC-SHS and pYC-SHS-4M during passage: 100 μL of stationary phase culture of CJ7/pYC-SHS and CJ7/pYC-SHS-4M in Halo-2 medium was inoculated into 5 mL of Halo-2 medium every day. Samples were taken and measured by qPCR using primer pairs vector-F/vector-R, orf14-F/orf14-R and radA-F/radA-R. Three independent experiments were performed, and error bars indicated standard deviations. The qPCR data were analyzed by $2^{-\Delta C_T}$ method (56)

Bioinformatic and statistical analysis.

Homologous proteins were searched by using BlastP and SyntTax (58), and conserved domains were detected by CD-search at National Center for Biotechnology Information (NCBI). Clustal Omega and Constraint-based Multiple Alignment Tool at NCBI (COBALT) was used for multiple protein alignments (https://www.ebi.ac.uk/Tools/msa/clustalo, https://www.ncbi.nlm.nih.gov/tools/cobalt/). The secondary-structure of ORF4 was predicted by PSIPRED in MPI Bioinformatics Toolkit (https://toolkit.tuebingen.mpg.de) (59, 60). The global alignment of two proteins were performed by EMBOSS Needle toll at EMBI-EBI (https://www.ebi.ac.uk/Tools/psa/emboss_needle/).

References

1. Dion MB, Oechslin F, Moineau S. 2020. Phage diversity, genomics and phylogeny. Nat Rev Microbiol doi:10.1038/s41579-019-0311-5.
2. Krupovic M, Cvirkaite-Krupovic V, Iranzo J, Prangishvili D, Koonin EV. 2018. Viruses of archaea: Structural, functional, environmental and evolutionary genomics. Virus Res 244:181-193.
3. Prangishvili D, Bamford DH, Forterre P, Iranzo J, Koonin EV, Krupovic M. 2017. The enigmatic archaeal virosphere. Nat Rev Microbiol 15:724-739.
4. Zhang J, Zheng X, Wang H, Jiang H, Dong H, Huang L. 2020. Novel Sulfolobus Fuselloviruses with Extensive Genomic Variations. J Virol 94.
5. Harrison E, Brockhurst MA. 2017. Ecological and evolutionary benefits of temperate phage: What does or doesn't kill you makes you stronger. Bioessays 39.
6. Gandon S. 2016. Why be temperate: Lessons from bacteriophage λ. Trends Microbiol 24:356-365.
7. Koskella B, Brockhurst MA. 2014. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiol Rev 38:916-31.
8. Nasir A, Kim KM, Caetano-Anolles G. 2017. Long-term evolution of viruses: A Janus-faced balance. Bioessays 39.
9. Zhang Z, Liu Y, Wang S, Yang D, Cheng Y, Hu J, Chen J, Mei Y, Shen P, Bamford DH, Chen X. 2012. Temperate membrane-containing halophilic archaeal virus SNJ1 has a circular dsDNA genome identical to that of plasmid pHH205. Virology 434:233-41.
10. Liu Y, Wang J, Liu Y, Wang Y, Zhang Z, Oksanen HM, Bamford DH, Chen X. 2015. Identification and characterization of SNJ2, the first temperate pleolipovirus integrating into the genome of the SNJ1-lysogenic archaeal strain. Mol Microbiol 98:1002-20.
11. Witte A, Baranyi U, Klein R, Sulzner M, Luo C, Wanner G, Kruger DH, Lubitz W. 1997. Characterization of *Natronobacterium magadii* phage φCh1, a unique archaeal phage containing DNA and RNA. Mol Microbiol 23:603-616.
12. Schnabel H, Zillig W, Pfaffle M, Schnabel R, Michel H, Delius H. 1982. *Halobacterium halobium* phage PHI-H. EMBO J 1:87-92.
13. Schleper C, Kubo K, Zillig W. 1992. The particle SSV1 form the extremely thermophilic archaeon *Sulfolobus* is a virus - demonstration of infectivity and of transfection with viral - DNA. Proc Natl Acad Sci U S A 89:7645-7649.
14. Sheppard C, Blombach F, Belsom A, Schulz S, Daviter T, Smollett K, Mahieu E, Erdmann S, Tinnefeld P, Garrett R, Grohmann D, Rappsilber J, Werner F. 2016. Repression of RNA polymerase by the archaeo-viral regulator ORF145/RIP. Nat Commun 7:13.
15. Fusco S, She Q, Bartolucci S, Contursi P. 2013. T(lys), a newly identified Sulfolobus spindle-shaped virus 1 transcript expressed in the lysogenic state, encodes a DNA-binding protein interacting at the promoters of the early genes. J Virol 87:5926-36.
16. Fusco S, She Q, Fiorentino G, Bartolucci S, Contursi P. 2015. Unravelling the Role of the F55 Regulator in the Transition from Lysogeny to UV Induction of Sulfolobus Spindle-Shaped Virus 1. J Virol 89:6453-61.
17. Fusco S, Aulitto M, Iacobucci I, Crocamo G, Pucci P, Bartolucci S, Monti M, Contursi P. 2020. The interaction between the F55 virus-encoded transcription regulator and the RadA host

recombinase reveals a common strategy in Archaea and Bacteria to sense the UV-induced damage to the host DNA. Biochim Biophys Acta Gene Regul Mech doi:10.1016/j.bbagrm.2020.194493:194493.

18. Kim M, Ryu S. 2013. Antirepression system associated with the life cycle switch in the temperate *Podoviridae* phage SPC32H. J Virol 87:11775-11786.

19. Erez Z, Steinberger-Levy I, Shamir M, Doron S, Stokar-Avihail A, Peleg Y, Melamed S, Leavitt A, Savidor A, Albeck S, Amitai G, Sorek R. 2017. Communication between viruses guides lysis-lysogeny decisions. Nature 541:488-493.

20. Ranquet C, Toussaint A, de Jong H, Maenhaut-Michel G, Geiselmann J. 2005. Control of bacteriophage Mu lysogenic repression. J Mol Biol 353:186-195.

21. Mavrich TN, Hatfull GF. 2019. Evolution of superinfection immunity in Cluster A mycobacteriophages. MBio 10.

22. Bondy-Denomy J, Qian J, Westra ER, Buckling A, Guttman DS, Davidson AR, Maxwell KL. 2016. Prophages mediate defense against phage infection through diverse mechanisms. ISME J 10:2854-2866.

23. Sauer RT, Ross MJ, Ptashne M. 1982. Cleavage of the lambda and P22 repressors by recA protein. J Biol Chem 257:4458-62.

24. Oppenheim AB, Kobiler O, Stavans J, Court DL, Adhya S. 2005. Switches in bacteriophage lambda development. Annu Rev Genet 39:409-29.

25. Johnson AD, Poteete AR, Lauer G, Sauer RT, Ackers GK, Ptashne M. 1981. λ Repressor and cro--components of an efficient molecular switch. Nature 294:217-23.

26. Galkin VE, Yu X, Bielnicki J, Ndjonka D, Bell CE, Egelman EH. 2009. Cleavage of bacteriophage lambda cI repressor involves the RecA C-terminal domain. J Mol Biol 385:779-87.

27. Fusco S, Liguori R, Limauro D, Bartolucci S, She Q, Contursi P. 2015. Transcriptome analysis of Sulfolobus solfataricus infected with two related fuselloviruses reveals novel insights into the regulation of CRISPR-Cas system. Biochimie 118:322-32.

28. Ye XC, Ou JH, Ni L, Shi WL, Shen P. 2003. Characterization of a novel plasmid from extremely halophilic Archaea: nucleotide sequence and function analysis. FEMS Microbiol Lett 221:53-57.

29. Wang Y, Sima L, Lv J, Huang S, Liu Y, Wang J, Krupovic M, Chen X. 2016. Identification, characterization, and application of the replicon region of the halophilic temperate sphaerolipovirus SNJ1. J Bacteriol 198:1952-1964.

30. Casjens SR, Gilcrease EB, Winn-Stapley DA, Schicklmaier P, Schmieger H, Pedulla ML, Ford ME, Houtz JM, Hatfull GF, Hendrix RW. 2005. The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. J Bacteriol 187:1091-1104.

31. Casjens SR, Gilcrease EB. 2009. Determining DNA Packaging Strategy by Analysis of the Termini of the Chromosomes in Tailed-Bacteriophage Virions, p 91-111. *In* Clokie MRJ, Kropinski AM (ed), Bacteriophages: Methods and Protocols, Volume 2 Molecular and Applied Aspects doi:10.1007/978-1-60327-565-1_7. Humana Press, Totowa, NJ.

32. Contursi P, D'Ambrosio K, Pirone L, Pedone E, Aucelli T, She Q, De Simone G, Bartolucci S. 2011. C68 from the *Sulfolobus islandicus* plasmid-virus pSSVx is a novel member of the AbrB-like transcription factor family. Biochem J 435:157-66.

33. Kamada K, Hanaoka F, Burley SK. 2003. Crystal structure of the MazE/MazF complex: Molecular bases of antidote-toxin recognition. Mol Cell 11:875-884.
34. Coles M, Djuranovic S, Soding J, Frickey T, Koretke K, Truffault V, Martin J, Lupas AN. 2005. AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi β barrels. Structure 13:919-928.
35. Asen I, Djuranovic S, Lupas AN, Zeth K. 2009. Crystal structure of SpoVT, the final modulator of gene expression during spore development in *Bacillus subtilis*. J Mol Biol 386:962-975.
36. Xiong L, Liu S, Chen S, Xiao Y, Zhu B, Gao Y, Zhang Y, Chen B, Luo J, Deng Z, Chen X, Wang L, Chen S. 2019. A new type of DNA phosphorothioation-based antiviral system in archaea. Nat Commun 10:1688.
37. Demina TA, Pietila MK, Svirskaite J, Ravantti JJ, Atanasova NS, Bamford DH, Oksanen HM. 2017. HCIV-1 and other tailless icosahedral internal membrane-containing viruses of the family *Sphaerolipoviridae*. Viruses 9.
38. Ramirez-Peralta A, Stewart KA, Thomas SK, Setlow B, Chen Z, Li YQ, Setlow P. 2012. Effects of the SpoVT regulatory protein on the germination and germination protein levels of spores of *Bacillus subtilis*. J Bacteriol 194:3417-25.
39. Bagyan I, Hobot J, Cutting S. 1996. A compartmentalized regulator of developmental gene expression in *Bacillus subtilis*. J Bacteriol 178:4500-7.
40. Chumsakul O, Takahashi H, Oshima T, Hishimoto T, Kanaya S, Ogasawara N, Ishikawa S. 2011. Genome-wide binding profiles of the *Bacillus subtilis* transition state regulator AbrB and its homolog Abh reveals their interactive role in transcriptional regulation. Nucleic Acids Res 39:414-28.
41. Aizenman E, Engelberg-Kulka H, Glaser G. 1996. An *Escherichia coli* chromosomal "addiction module" regulated by guanosine 3',5'-bispyrophosphate: a model for programmed bacterial cell death. Proc Natl Acad Sci U S A 93:6059-63.
42. Nikolic N. 2019. Autoregulation of bacterial gene expression: lessons from the MazEF toxin-antitoxin system. Curr Genet 65:133-138.
43. Fusco S, Aulitto M, Iacobucci I, Crocamo G, Pucci P, Bartolucci S, Monti M, Contursi P. 2020. The interaction between the F55 virus-encoded transcription regulator and the RadA host recombinase reveals a common strategy in Archaea and Bacteria to sense the UV-induced damage to the host DNA. Biochim Biophys Acta Gene Regul Mech 1863:194493.
44. Zhou X, Sun K, Zhou X, Jackson AO, Li Z. 2019. The Matrix Protein of a Plant Rhabdovirus Mediates Superinfection Exclusion by Inhibiting Viral Transcription. J Virol 93.
45. Wagenaar TR, Moss B. 2009. Expression of the A56 and K2 proteins is sufficient to inhibit vaccinia virus entry and cell fusion. J Virol 83:1546-54.
46. Lu MJ, Henning U. 1994. Superinfection exclusion by T-even-type coliphages. Trends Microbiol 2:137-9.
47. Quemin ER, Lucas S, Daum B, Quax TE, Kuhlbrandt W, Forterre P, Albers SV, Prangishvili D, Krupovic M. 2013. First insights into the entry process of hyperthermophilic archaeal viruses. J Virol 87:13379-85.
48. Medvedeva S, Liu Y, Koonin EV, Severinov K, Prangishvili D, Krupovic M. 2019. Virus-borne mini-CRISPR arrays are involved in interviral conflicts. Nat Commun 10:5204.
49. Wang J, Liu Y, Liu Y, Du K, Xu S, Wang Y, Krupovic M, Chen X. 2018. A novel family of

tyrosine integrases encoded by the temperate pleolipovirus SNJ2. Nucleic Acids Res 46:2521-2536.

50. Charlebois RL, Lam WL, Cline SW, Doolittle WF. 1987. Characterization of pHV2 from *Halobacterium volcanii* and its use in demonstrating transformation of an archaebacterium. Proc Natl Acad Sci U S A 84:8530-4.

51. Cline SW, Lam WL, Charlebois RL, Schalkwyk LC, Doolittle WF. 1989. Transformation methods for halophilic archaebacteria. Can J Microbiol 35:148-52.

52. Huff JP, Grant BJ, Penning CA, Sullivan KF. 1990. Optimization of routine transformation of *Escherichia coli* with plasmid DNA. Biotechniques 9:570-2, 574, 576-7.

53. Cai S, Cai L, Zhao D, Liu G, Han J, Zhou J, Xiang H. 2015. A novel DNA-binding protein, PhaR, plays a central role in the regulation of polyhydroxyalkanoate accumulation and granule formation in the haloarchaeon *Haloferax mediterranei*. Appl Environ Microbiol 81:373-85.

54. Mei Y, Chen J, Sun D, Chen D, Yang Y, Shen P, Chen X. 2007. Induction and preliminary characterization of a novel halophage SNJ1 from lysogenic *Natrinema* sp. F5. Can J Microbiol 53:1106-10.

55. Breuert S, Allers T, Spohn G, Soppa J. 2006. Regulated polyploidy in halophilic archaea. PLoS One 1:e92.

56. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25:402-8.

57. Wang Y, Chen B, Cao M, Sima L, Prangishvili D, Chen X, Krupovic M. 2018. Rolling-circle replication initiation protein of haloarchaeal sphaerolipovirus SNJ1 is homologous to bacterial transposases of the IS*91* family insertion sequences. J Gen Virol doi:10.1099/jgv.0.001009.

58. Oberto BB. 2013. SyntTax: a web server linking synteny to prokaryotic taxonomy. BMC Bioinf 14:4.

59. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, Gabler F, Soding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI Bioinformatics Toolkit with a new HHpred server at its core. J Mol Biol 430:2237-2243.

60. Jones DT. 1999. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292:195-202.

Acknowledgments

We thank members of the Chen's lab and Du's lab for comments and advice in preparing the manuscripts. We thank Prof. Hua Xiang (Institute of Microbiology, Chinese Academy of Sciences) for kindly providing the pRF plasmid. This study was supported by grants from National Natural Science Foundation of China [31570174], National Found for Fostering Talents of Basic Sciences [J1103513]; Research

(Innovative) Fund of Laboratory Wuhan University to Dr. Xiangdong Chen.

Author Contributions

B.B.C, X.D.C and S.D. designed the research; B.B.C., Z.C, Y.C.W, H.G, L.S.S, J.W, S.S.O and W.Q.G performed the research; B.B.C, M.K., X.D.C and S.D analyzed data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Correspondence

ssdu@whu.edu.cn; xdchen@whu.edu.cn

Figure legends

Fig 1. Discovery of clear-plaque mutants of SNJ1. a. Viruses obtained from CJ7/pYC-S culture formed both clear and turbid plaques on lawns of CJ7. Late exponential phase culture of CJ7 was infected with SNJ1 viruses or viruses generated using CJ7/pYC-S. SNJ1 formed turbid plaques, while viruses generated using pYC-S formed two kinds of plaques. Arrows represented turbid plaques and triangles

represented clear plaques. Viruses purified from the turbid or clear plaques were propagated and used to infect CJ7 as above, the morphotypes of the plaques were maintained for either virus. **b**. The ability of different viruses to lyse cells. The same amount of viruses was added into CJ7 culture ($OD_{600} \approx 0.1$). Growth of the culture was monitored for 49 h post infection by measuring OD_{600}. SNJ1-Ca/Cb and SNJ1-Ta/Tb represented two clear-plaque viruses and two turbid-plaque viruses isolated form lawns of CJ7 infected with viruses from CJ7/pYC-S culture, respectively. CJ7 without virus infection was used as a control. The experiment was repeated twice, and error bars indicated standard deviations.

Fig 2. Schematic diagrams of the genomic deletion of 20 turbid and clear plaque viral genomes. a. Scheme of SNJ1 genome from 1 bp to 1,134 bp and locations of genomic deletion of turbid/clear plaque viruses of SNJ1. The putative ORFs of SNJ1 were indicated by numbered arrows, while the end site of genomic deletion in turbid/clear plaque viruses were marked by T1-T20 and C1-C20. Red dotted arrow represented the region of genomic deletion. Detailed deletion location was shown in **b** (turbid plaque viruses) and **c** (clear plaque viruses). The pUC19-mev fragment and flanking sequences of 20 turbid and clear plaque viruses were amplified by HJ-F/R primer pair and sequenced. Black solid lines represented genomes packaged in viral particles, while black dotted lines represented deletions. The putative ORFs were indicated by numbered arrows and the *Sac*I site, where foreigner DNA was inserted, was indicated. The numbers on the right side represented the start (x) and end (y)

location of the genomic deletion. (z) represented the genome size of mutant SNJ1 viruses and the smaller genomes were noted in bold, compared to the wild-type SNJ1 (16,492 bp).

Fig 3. Identification of *orf4* as a critical factor for SNJ1 lysis-lysogeny switch. a. *orf4* disrupted viruses formed only clear plaques on lawns of CJ7. Late exponential phase culture of CJ7 was infected with SNJ1$^{orf4\ mut}$ and SNJ1$^{\Delta orf4}$ viruses (generated from CJ7/pYC-S-4M and CJ7/pYC-S Δ1-575 culture respectively), and only clear plaques were observed. **b**. Expression of ORF4 in trans restored turbid plaque formation of clear-plaque SNJ1 mutant viruses and inhibited plaque formation dramatically. Ten-fold serial dilutions of SNJ1 viruses or its clear plaque mutants were spotted onto the lawns of CJ7-F/pFJ6-MCS, or CJ7-F/pFJ6-1-656 (with *orf4*). Plates were incubated at 37°C for 48 h and photographed. The maximum dilution for observed plaques was highlighted by asterisk.

Fig 4. ORF4 is critical for stability and copy number control of SNJ1-based plasmids. a. The relative copy number of SNJ1-based plasmid pYC-SHS (with *orf4*) and pYC-SHS-4M (without *orf4*) in CJ7 cells with or without MMC treatment. WT represents wild type pYC-SHS plasmid, Δ*orf4* represents pYC-SHS-4M (start codon mutation of *orf4*) and Δ*orf4*+*orf4* represents pFJ6-1-656 complemented pYC-SHS-4M. These plasmids were transformed to CJ7 strain and cultured to late-exponential phase (24 h) in 18% MGM + Mev medium and treated with MMC (1

745 μg mL^{-1}) for 30 min, while cultures without MMC treated were set as control. Cells
746 were collected by centrifugation and washed twice in same volume of 18% MGM to
747 remove MMC. Cell pellets were resuspended in 18% MGM + Mev medium and
748 cultured for 24 h. Samples were taken for qPCR analysis using primer pairs
749 vector-F/vector-R, orf14-F/orf14-R and radA-F/radA-R to determine the relative
750 plasmid copy number to chromosome. Three independent experiments were
751 performed, and error bars indicated standard deviations. **b**. Plasmid stability of
752 pYC-SHS and pYC-SHS-4M during passage. 100 μL of stationary phase culture of
753 CJ7/pYC-SHS, CJ7/pYC-SHS-4M and CJ7/pYC-SHS-4M+pFJ6-1-656 was
754 inoculated into 5 mL of Halo-2 medium every day. Samples were taken and measured
755 by qPCR using primer pairs vector-F/vector-R, orf14-F/orf14-R and radA-F/radA-R
756 for five days. Three independent experiments were performed, and error bars
757 indicated standard deviations.
758
759 **Fig. 5. ORF4 mediates superinfection immunity of SNJ1. a**. Expression of ORF4
760 confers resistance to SNJ1 infection in CJ7. To test the immunity to against SNJ1,
761 ten-fold serial dilutions of SNJ1 virus stocks were spotted onto lawns of CJ7, J7-1 and
762 CJ7-F strains carrying pFJ6 plasmids with or without *orf3* and *orf4*. The maximum
763 dilution for observed plaques was highlighted by asterisk. **b**. Sequence alignment of
764 ORF4 with other MazE/SpoVT family members. C68 protein from hybrid
765 virus-plasmid pSSVx (residues 1-68, PDB code 3O27)(32), the N-terminal domains
766 of MazE (residues 1-53, PDB code 1UB4)(33), AbrB (residues 1-51, PDB code

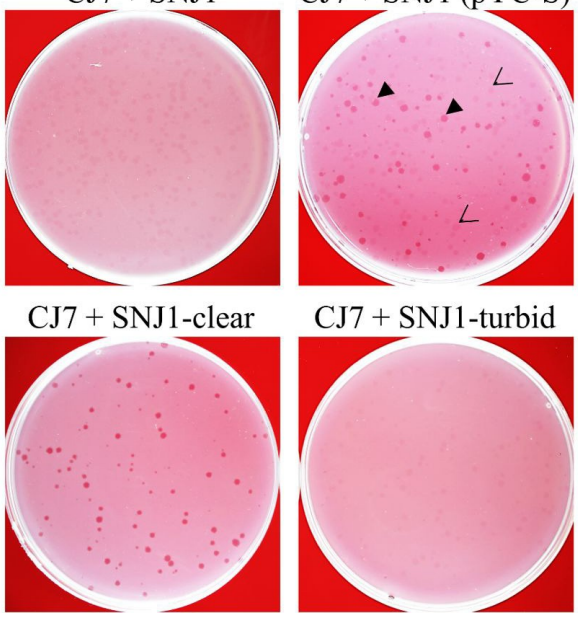
1YSF)(34), and SpoVT (residues 1-55, PDB code 2W1T)(35) are aligned with ORF4 by Clustal Omega. The secondary-structure of ORF4 was predicted by PSIPRED in MPI Bioinformatics Toolkit. ORF4 contains 5 β-strands (light-grey arrows) and one α-helix (dark-grey cylinder). The same color is used for other proteins. **c**. The N-terminal 33 amino acids of SNJ1 were necessary and sufficient for immunity to against SNJ1. Ten-fold serial dilutions of SNJ1 were spotted onto lawns of CJ7-F harboring plasmid pFJ6-MCS or its derivatives carrying different portion of ORF4, superscript denotes the amino acid residues of ORF4. Hpro stands for the promoter of heat-shock protein 70 from *Haloferax volcanii* DS52. Plates were incubated at 37°C for 48 h and photographed.

Fig. 6. ORF4 blocks SNJ1 infection by inhibiting viral gDNA replication. a. ORF4 did not affect SNJ1 adsorption to host cells CJ7. SNJ1 was incubated with early-exponential phase cultures of CJ7, J7-1, CJ7-F/pFJ6-MCS (- *orf4*) and CJ7-F/pFJ6-Hpro-*orf4* (+ *orf4*) at different MOI (0.1, 1 and 10) for about 1 h at 45°C. After absorption, the culture was centrifuged, and the titer of unbound viruses in supernatant was measured by double-layer method. Adsorption efficiency was determined by comparing the titer before and after virus adsorption. Three independent experiments were performed, and error bars indicated standard deviations. **b**. Relative copy number of viral gDNA intercellular 1 h post infection. Indicated strains were infected with SNJ1 as in (**a**). After 1 h of adsorption, the cells were centrifuged and the relative copy number of viral gDNA intercellular were measured

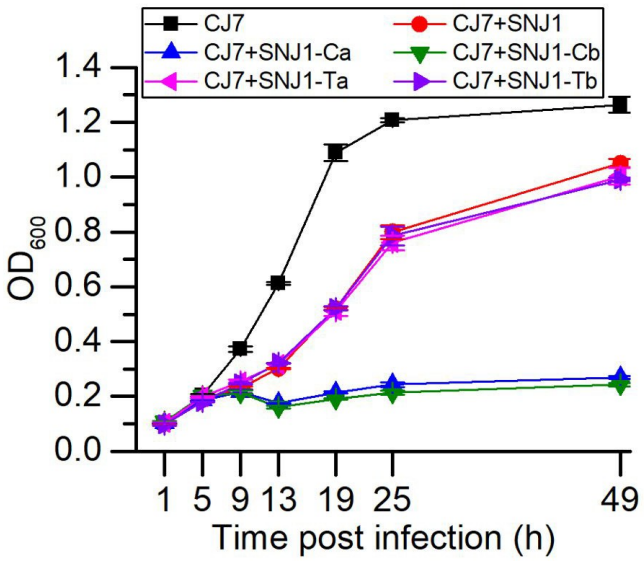
789 by qPCR using primer pairs orf14-F/orf14-R and radA-F/radA-R, which represented
790 proviral genome and host chromosome, respectively. The bars represent the means
791 and standard deviations of three independent experiments. Significance testing against
792 CJ7-F/pFJ6-Hpro-*orf4* was performed using a one-sample t-test (***: $P < 0.001$, **: P
793 < 0.01). **c**. ORF4 repressed SNJ1 genome replication. SNJ1 was incubated with CJ7
794 or CJ7-F/pFJ6-MCS at a MOI of 0.1 and CJ7-F/pFJ6-Hpro-*orf4* at a MOI of 0.5 for 1
795 h. The cells were collected and washed in fresh Halo-2 medium twice to eliminate
796 free viruses. Cells were then cultivated for 8 h and samples were taken every hour
797 post infection. qPCR analyses were performed using the same primer pairs as in (**b**).
798 The bars represent the means and standard deviations of three independent
799 experiments. Significance testing against CJ7-F/pFJ6-Hpro-*orf4* was performed using
800 a one-sample t-test (***: $P < 0.001$, **: $P < 0.01$, *: $P < 0.05$). **d**. Genomic integrity of
801 provirus was not affected by ORF4. Experiments were performed as in (**c**) expect that
802 SNJ1 was incubated at a MOI of 5 with all strains and samples for southern blot
803 analysis were taken every two hours after 1 h of infection. The DNA samples were
804 electrophoresed on agarose gels and transferred onto positively charged nylon
805 membranes with alkaline transfer buffer. A specific probe recognized nucleotide
806 1,483-2,537 bp of SNJ1 was used for southern blot analysis. The DNA marker were
807 shown on the left. C: CJ7 strain; -: CJ7-F/pFJ6-MCS strain; +: CJ7-F/pFJ6-Hpro-*orf4*
808 strain.
809
810 **Fig. 7. Alignment of SNJ1 with two SNJ1-like plasmids.** Genomic alignment of

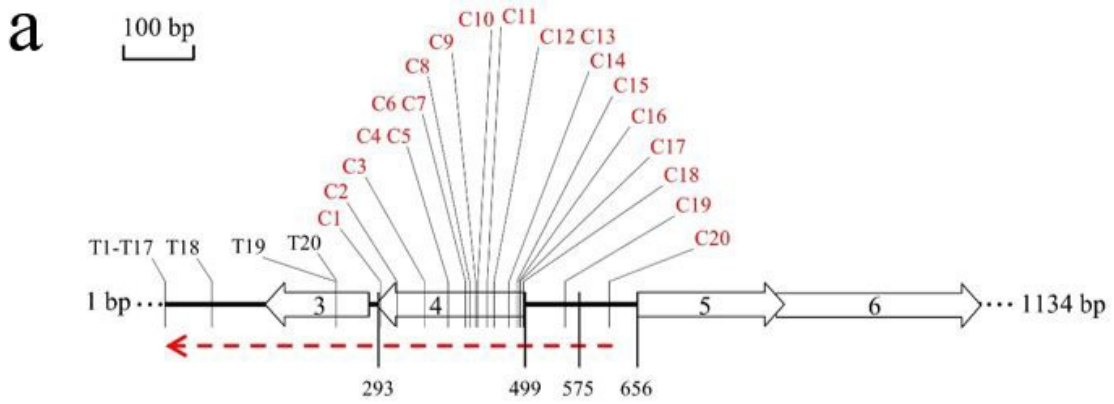
811 SNJ1, *Natrinema versiforme* strain BOL5-4 plasmid pNVE19 (GenBank Acc. No.
812 NZ_CP040333, start from 8,613 bp) and *Haloterrigena jeotgali* strain A29 plasmid
813 unnamed5 (GenBank Acc. No. CP031302, start from 9,455 bp). ORFs of SNJ1 was
814 noted in the arrows. Proteins predicted to be regulators or related to DNA replication
815 were indicated by red or yellow arrows. Capsid proteins defined by mass
816 spectrometric analysis previously(9) were colored orange. Characteristic conserved
817 proteins in family *Sphaerolipoviridae*, including the packaging ATPase, small major
818 capsid protein and large major capsid protein, were marked with pink, light blue and
819 dark blue color (bottom legend). Homologous ORFs in the other two plasmids are
820 shown in the same color, and the percentages of protein identities were shown.
821
822 **Fig. 8. Schematic view of the life cycle of SNJ1.** i: SNJ1 proviral genome resides in
823 J7-1 cytoplasm as plasmid pHH205 with a relative copy number of 1 to 3. ORF4
824 represses expression of the lytic pathway genes presumably by binding to the viral
825 DNA as a dimer, thus maintains SNJ1 in the lysogenic state. ii-v: Upon MMC
826 treatment, ORF4 was inactivated by an unknown mechanism, resulting in expression
827 of the lytic pathway genes and replication of viral genome by the RepA protein. Host
828 cells are lysed and assembled progeny viruses are released into supernatant. vi-x:
829 Released SNJ1 viruses infect CJ7, either enter the lysogenic state as a plasmid or
830 replicate actively, in a process controlled by ORF4. xi: A lysogen of SNJ1 (J7-1) was
831 immune to superinfection of SNJ1 because ORF4 represses replication of the ejected
832 gDNA.

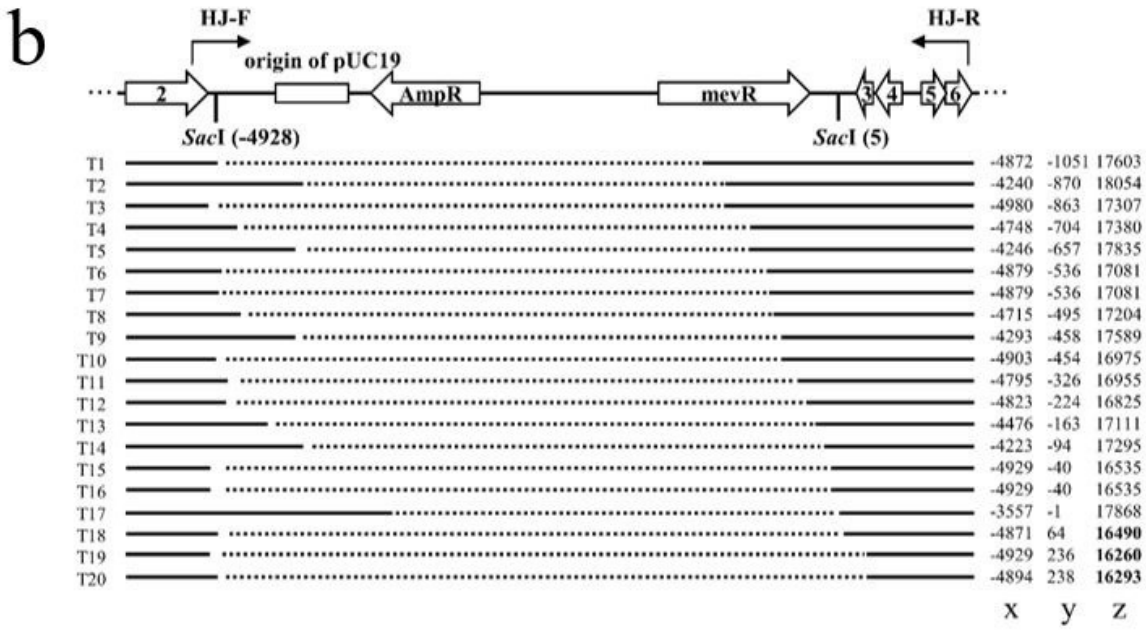
a


CJ7 + SNJ1
CJ7 + SNJ1 (pYC-S)
CJ7 + SNJ1-clear
CJ7 + SNJ1-turbid

b

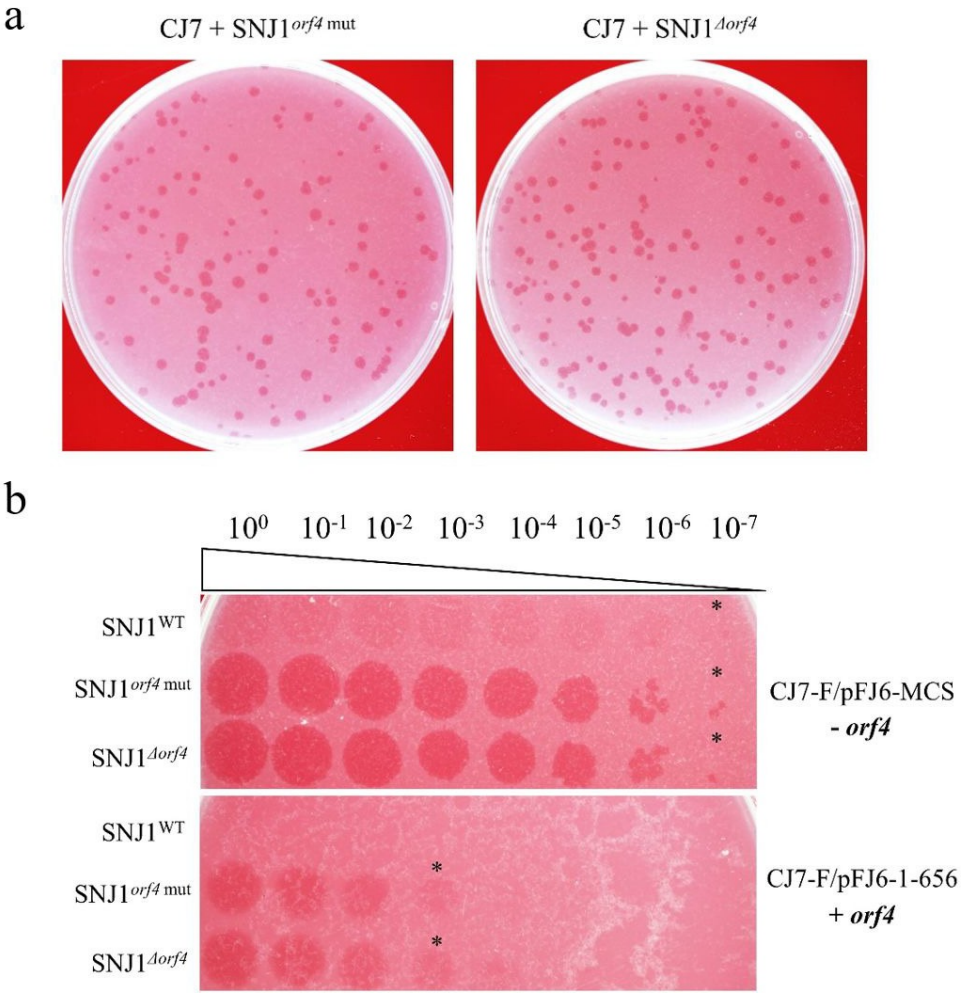

CJ7
CJ7+SNJ1
CJ7+SNJ1-Ca
CJ7+SNJ1-Cb
CJ7+SNJ1-Ta
CJ7+SNJ1-Tb
OD600
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1 5 9 13 19 25 49
Time post infection (h)



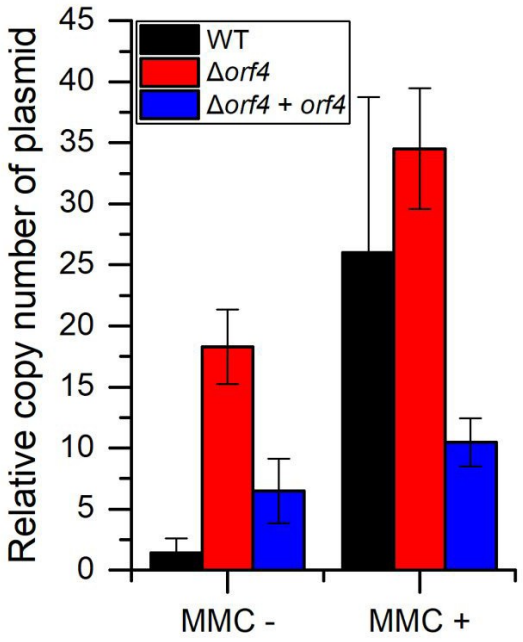


T1	-4872	-1051	17603
T2	-4240	-870	18054
T3	-4980	-863	17307
T4	-4748	-704	17380
T5	-4246	-657	17835
T6	-4879	-536	17081
T7	-4879	-536	17081
T8	-4715	-495	17204
T9	-4293	-458	17589
T10	-4903	-454	16975
T11	-4795	-326	16955
T12	-4823	-224	16825
T13	-4476	-163	17111
T14	-4223	-94	17295
T15	-4929	-40	16535
T16	-4929	-40	16535
T17	-3557	-1	17868
T18	-4871	64	**16490**
T19	-4929	236	**16260**
T20	-4894	238	**16293**
	x	y	z

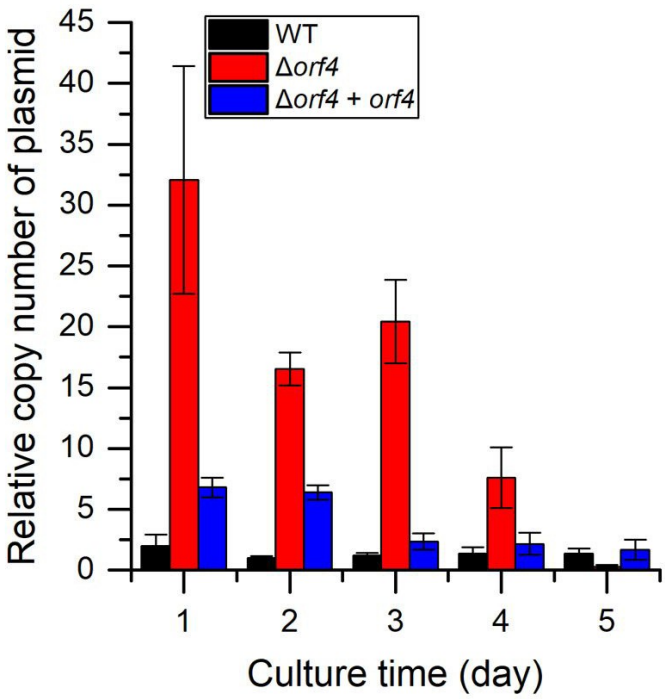
C1	-4882	299	**16244**
C2	-4902	322	**16201**
C3	-4547	360	16518
C4	-4821	393	**16211**
C5	-4144	393	16888
C6	-4392	417	16616
C7	-4392	417	16616
C8	-4911	423	**16091**
C9	-3825	431	17169
C10	-4571	435	**16419**
C11	-4570	448	**16407**
C12	-4489	457	**16479**
C13	-4791	457	**16177**
C14	-4495	478	**16452**
C15	-4862	489	**16074**
C16	-4557	492	**16376**
C17	-4784	493	**16148**
C18	-4917	498	**16010**
C19	-4715	556	**16154**
C20	-4369	617	**16439**
	x	y	z

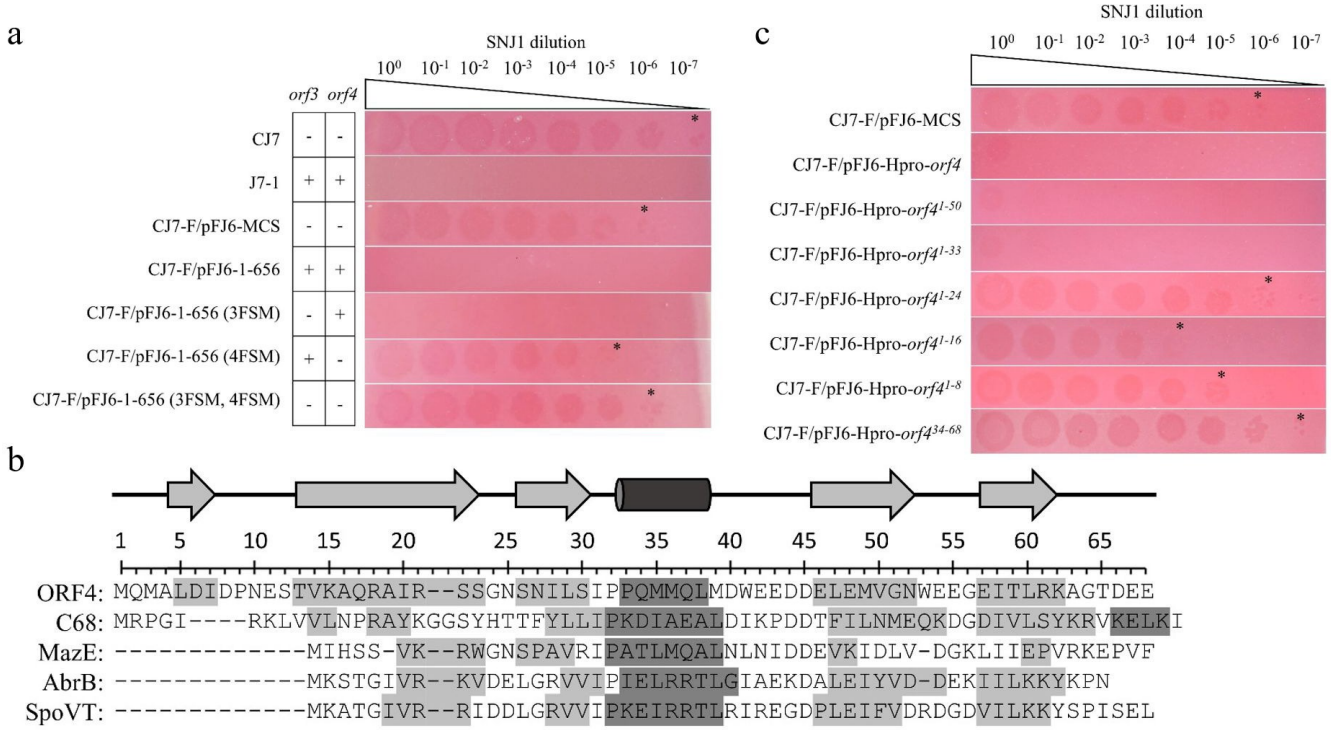

a
CJ7 + SNJ1orf4 mut
CJ7 + SNJ1Δorf4
b
10^0 10^-1 10^-2 10^-3 10^-4 10^-5 10^-6 10^-7
SNJ1WT
SNJ1orf4 mut
SNJ1Δorf4
CJ7-F/pFJ6-MCS
- orf4
SNJ1WT
SNJ1orf4 mut
SNJ1Δorf4
CJ7-F/pFJ6-1-656
+ orf4

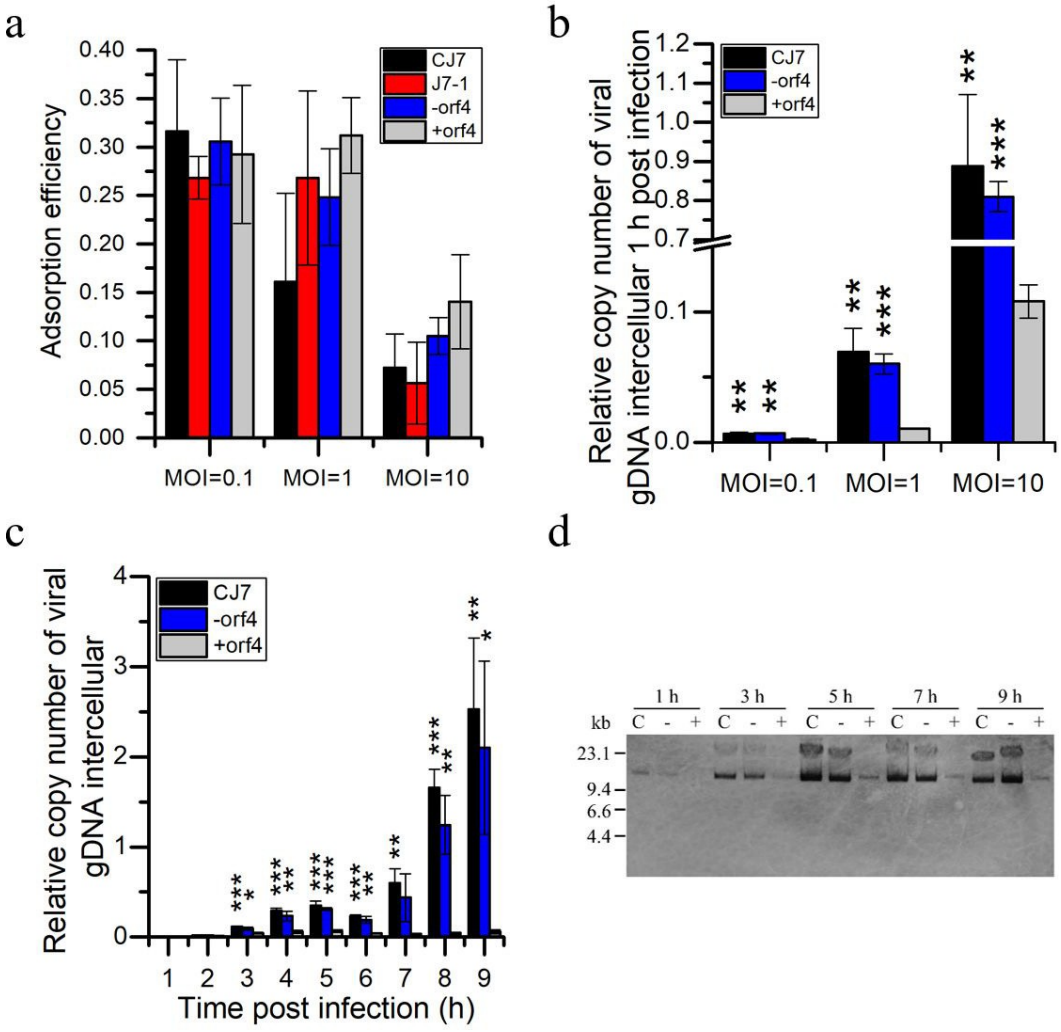
a

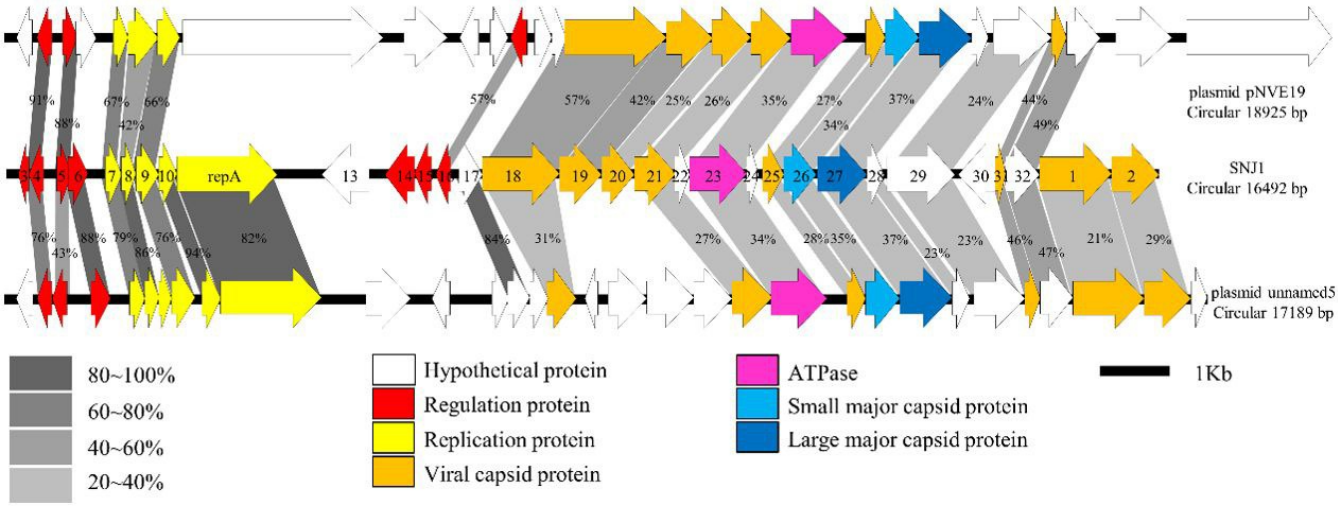

Relative copy number of plasmid
45
40
35
30
25
20
15
10
5
0
WT
Δorf4
Δorf4 + orf4
MMC -
MMC +

b


Relative copy number of plasmid
45
40
35
30
25
20
15
10
5
0
WT
Δorf4
Δorf4 + orf4
1
2
3
4
5
Culture time (day)


a
SNJ1 dilution
10^0 10^-1 10^-2 10^-3 10^-4 10^-5 10^-6 10^-7
orf3 orf4
CJ7 - -
J7-1 + +
CJ7-F/pFJ6-MCS - -
CJ7-F/pFJ6-1-656 + +
CJ7-F/pFJ6-1-656 (3FSM) - +
CJ7-F/pFJ6-1-656 (4FSM) + -
CJ7-F/pFJ6-1-656 (3FSM, 4FSM) - -
c
SNJ1 dilution
10^0 10^-1 10^-2 10^-3 10^-4 10^-5 10^-6 10^-7
CJ7-F/pFJ6-MCS
CJ7-F/pFJ6-Hpro-orf4
CJ7-F/pFJ6-Hpro-orf4^1-50
CJ7-F/pFJ6-Hpro-orf4^1-33
CJ7-F/pFJ6-Hpro-orf4^1-24
CJ7-F/pFJ6-Hpro-orf4^1-16
CJ7-F/pFJ6-Hpro-orf4^1-8
CJ7-F/pFJ6-Hpro-orf4^34-68
b
1 5 10 15 20 25 30 35 40 45 50 55 60 65
ORF4: MQMALDIDPNESTVKAQRAIR--SSGNSNILSIPPQMMQLMDWEEDDELEMVGNWEEGEITLRKAGTDEE
C68: MRPGI----RKLVVLNPRAYKGGSYHTTFYLLIPKDIAEALDIKPDDTFILNMEQKDGDIVLSYKRVKELKI
MazE: ----------MIHSS-VK--RWGNSPAVRIPATLMQALNLNIDDEVKIDLV-DGKLIIEPVRKEPVF
AbrB: ----------MKSTGIVR--KVDELGRVVIPIELRRTLGIAEKDALEIYVD-DEKIILKKYKPN
SpoVT: ----------MKATGIVR--RIDDLGRVVIPKEIRRTLRIREGDPLEIFVDRDGDVILKKYSPISEL


a
Adsorption efficiency
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
CJ7
J7-1
-orf4
+orf4
MOI=0.1
MOI=1
MOI=10
b
Relative copy number of viral gDNA intercellular 1 h post infection
1.2
1.1
1.0
0.9
0.8
0.7
0.1
0.0
CJ7
-orf4
+orf4
MOI=0.1
MOI=1
MOI=10
c
Relative copy number of viral gDNA intercellular
4
3
2
1
0
CJ7
-orf4
+orf4
1 2 3 4 5 6 7 8 9
Time post infection (h)
d
1 h
3 h
5 h
7 h
9 h
kb
C - + C - + C - + C - + C - +
23.1
9.4
6.6
4.4


plasmid pNVE19
Circular 18925 bp
SNJ1
Circular 16492 bp
plasmid unnamed5
Circular 17189 bp
repA
80~100%
60~80%
40~60%
20~40%
Hypothetical protein
Regulation protein
Replication protein
Viral capsid protein
ATPase
Small major capsid protein
Large major capsid protein
1Kb