Identification, Characterization, and Application of the Replicon Region of the Halophilic Temperate Sphaerolipovirus SNJ1

Yuchen Wang, Linshan Sima, Jie Lv, Suiyuan Huang, Ying Liu, Jiao Wang, Mart Krupovic, Xiangdong Chen

Identification, characterization and application of the replicon region of the halophilic temperate sphaerolipovirus SNJ1

Yuchen Wang[1], Linshan Sima[1], Jie Lv[1, 3], Suiyuan Huang[1], Ying Liu[1], Jiao Wang[1], Mart Krupovic[4], Xiangdong Chen[1,2,*]

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

[2]Hubei Provincial Cooperative Innovation Center of Industrial Fermentation, Wuhan 430072, China

[3]Henan Academy of Medical and Pharmaceutical Sciences, Zhengzhou University, Zhengzhou 450052, China

[4]Unit of Molecular Biology of the Gene in Extremophiles, Department of Microbiology, Institut Pasteur, 25 rue du Dr Roux, 75015 Paris, France

* Corresponding author

Xiangdong Chen, Wuhan University, China

Address: Wuhan University, Luojia Hill, 430072, Wuhan, Hubei Province, P.R.China

Telephone number: (+86) 027 68754533

Fax: (+86) 27 68754533.

E-mail address: xdchen@whu.edu.cn

Key words: Haloarchaea; temperate virus; replication; *Natrinema* ; expression shuttle vector

Running title: Halophilic temperate sphaerolipovirus SNJ1 replicon

ABSTRACT

Temperate haloarchaeal virus SNJ1 displays lytic and lysogenic life cycles. During the lysogenic cycle, the virus resides in its host, *Natrinema* sp. J7-1, in the form of an extrachromosomal circular plasmid, pHH205. In this study, a 3.9-kb region containing seven predicted genes organized in two operons was identified as the minimal replicon of SNJ1. Only RepA, encoded by ORF11-12, was found to be essential for replication and its expression increased during the lytic cycle. Sequence analysis suggested that RepA is a distant homolog of HUH endonucleases, a superfamily which includes rolling circle replication initiation proteins from various viruses and plasmids. In addition to RepA, two genetic elements located within both termini of the 3.9-kb replicon were also required for SNJ1 replication. SNJ1 genome and SNJ1 replicon-based shuttle vectors were present at 1−3 copies per chromosome. However, deletion of ORF4 significantly increased SNJ1 copy number, suggesting that the product of ORF4 is a negative regulator of SNJ1 abundance. Shuttle vectors based on the SNJ1 replicon were constructed and validated for stable expression of heterologous proteins, both in J7 derivatives and *Natrinema pallidum* JCM8980, suggesting their broad applicability as genetic tools for *Natrinema* species.

IMPORTANCE

Archaeal viruses exhibit striking morphological diversity and unique gene content. In this study, the minimal replicon of the temperate haloarchaeal virus SNJ1 was identified. A number of ORFs and genetic elements controlling virus genome replication, maintenance, and copy number were characterized. In addition, based on the replicon, a novel expression shuttle vector has been constructed and validated for protein expression and purification in *Natrinema* sp. CJ7 and *Natrinema pallidum* JCM8980. This study provided not only mechanistic and functional insights into SNJ1 replication but

47 also led to the development of useful genetic tools to investigate SNJ1 and other viruses infecting
48 *Natrinema* species as well as their hosts.

INTRODUCTION

Haloarchaea are the dominant microbes in hypersaline environments, such as salt lakes and salterns. Viruses infecting haloarchaea outnumber their hosts by several orders of magnitude (1). It has been speculated that in such harsh environment with no other predators, interactions between viruses and their hosts constitute the main evolutionary driving force (2). About 90 haloarchaeal viruses have been discovered so far, a relatively small number compared to the ~ 6200 reported bacteriophages (3, 4). Given that characterization of bacterial and eukaryotic viruses has yielded remarkable insights into the physiology and cell biology of their hosts, studies on haloarchaeal viruses are expected to be equally revealing. However, thus far research has been limited by the scarcity of suitable and genetically tractable virus-host models. Consequently, most of the current knowledge on viral entry, transcription, genome replication, assembly, and release derives from bioinformatics analyses of viral genome sequences. Therefore, there is a growing interest in developing genetic tools for understanding these viruses and their hosts.

Recently, *Sphaerolipoviridae*, a new family of viruses that infect extremophiles has been established by the International Committee on Taxonomy of Viruses (5). These are icosahedral, tailless viruses, with an internal lipid membrane sandwiched between the capsid and the dsDNA genome. Currently, *Sphaerolipoviridae* comprises three genera: *Alphasphaerolipovirus*, which encompasses *Haloarcula hispanica* viruses SH1 (6), HHIV-2 (7), and PH1 (8); *Betasphaerolipovirus*, which includes the *Natrinema* virus SNJ1 (9); and *Gammasphaerolipovirus*, which contains *Thermus*-infecting bacteriophages P23-77 (10) and IN93 (11). Each genus presents distinct features. SH1, PH1, and HHIV-2 are lytic viruses with a linear dsDNA. It is thought that their replication relies on a protein-primed mechanism, because they present inverted terminal repeats and proteins attached to the termini of their genomes (12, 13). In addition, all three viruses are highly similar to one another in terms

of genomic structure, protein homology, and life cycle. By contrast, SNJ1 has a temperate life cycle and a circular dsDNA genome (9). It is not clear how the genome of SNJ1 is replicated, because none of the putative open reading frames (ORFs) show similarity to genes encoding known replication proteins. P23-77 and IN93 contain circular dsDNA genomes and exhibit lytic and lysogenic life cycles, respectively. Viruses of this genus infect bacteria and are thought to employ the theta-type mode of genome replication, similar to that of *Thermus* plasmids (14, 15). Despite differences in genomic design and replication mechanisms, the architectural similarities between *Sphaerolipoviridae* suggest a common ancestor. Consequently, they represent an ideal experimental model to study the evolution of morphological and genetic properties of archaeal and bacterial viruses.

Betasphaerolipovirus SNJ1 was identified in *Natrinema* sp. J7-1 (16, 17), a derivative of *Natrinema* sp. J7. Upon lysogeny, unlike most other temperate viruses (18), SNJ1 does not integrate into the host genome but instead resides as an extrachromosomal circular plasmid, called pHH205. Upon mitomycin C (MMC) induction, large amounts of SNJ1 virions can be produced (9, 19, 20). It is not known how SNJ1 switches between lysogenic and lytic cycles, although its dependence on MMC is very similar to that displayed by bacterial phages such as the lambda phage (21, 22). Similarly, it remains of interest to determine how pHH205 is partitioned and sorted into daughter cells during the lysogenic cycle, because little is known about DNA segregation in archaea. Interestingly, we found that SNJ1 could not infect J7-1 that already contained the SNJ1 proviral genome, a phenomenon known as superinfection exclusion or immunity. Although relatively widespread among bacterial viruses, superinfection immunity has only been observed among archaea in *Sulfolobus islandicus* rod-shaped virus SIRV2 (23), the exact mechanism remaining unclear. Recently, SNJ1 has been reported to promote the replication of another temperate virus, SNJ2 (24). SNJ2, a member of the proposed family *Pleolipoviridae*, integrates its proviral genome into the chromosome of J7. In total, 17 SNJ2-like proviruses have been found in

genomes of 10 genera of haloarchaea, suggesting that *Pleolipoviridae* and related proviruses are ubiquitous in haloarchaea. Efficient production of SNJ2 virions could only be achieved in J7 strains co-infected with SNJ1, indicating that SNJ1 replication or expression of some of its proteins promoted replication of SNJ2. The mechanism behind this virus-virus interaction is poorly understood. Collectively, these observations suggest that SNJ1 could serve as a model virus to investigate the basic aspects of viral life cycle, virus-host, and virus-virus interactions as well as the fundamental cell biology of *Natrinema* species. Its existence as a dsDNA plasmid also suggests its possible use as a genetic tool.

In this study, we determined the minimal replication region of SNJ1 and identified an essential protein for its replication. We also determined regions and ORFs that were important for maintenance and copy number control of SNJ1. Furthermore, we developed several cloning and expression shuttle vectors based on the SNJ1 replicon. These vectors were verified for cloning and expression purposes in J7. Notably, we found that these vectors could also be transformed into and express target genes in *Natrinema pallidum* JCM8980, indicating that they might be widely applicable to *Natrinema* species. Construction of shuttle vectors should facilitate future studies on SNJ1 and other haloarchaeal virus life cycles, as well as promote research on the biology of *Natrinema* species.

111 **MATERIALS AND METHODS**

112

113 **Strains, media and growth conditions**

114 All the strains used in this study are listed in Table 1 and Table S1. J7 and other haloarchaeal strains
115 were cultured on Halo-2 or 18% modified growth medium (MGM) as previously described (25, 26). The
116 growth curves of *Natrinema* CJ7 with or without 5 μg/mL mevinolin (Mev) antibiotic in Halo-2 and
117 18% MGM medium at 45℃were measured respectively as previously described (9). *Escherichia coli*
118 DH5α and JM110 were used for plasmid DNA construction. Routinely, they were cultured in
119 Luria-Bertani medium at 37℃. 1.5% agar was added to the medium to prepare the solid medium. 5
120 μg/mL Mev or 100 μg /mL ampicillin was added to the medium for colony selection.

121

122 **Standard molecular manipulations**

123 Standard *E. coli* molecular manipulations were performed as described (27). The SNJ1 proviral genome
124 was extracted from J7-1 as previously reported (16). Total RNA from J7-1 was extracted using TRIZOL
125 reagent (Invitrogen). For reverse transcription PCR (RT-PCR), residual genomic DNA was removed
126 using gDNA Eraser (Takara) and cDNA was synthesized from 5 μg total RNA using random primers
127 (RT reagent kit; Takara). The resulting cDNA was used as a template for transcription analysis with
128 specific primers designed according to each ORF's sequence. Co-transcription analysis of ORF5−ORF6
129 and ORF7−ORF11-12 was performed with primers ORF5-fwd and ORF6-rev, and ORF7-up and
130 ORF11-12-down, respectively (Table S2).

131

132 **Transfection and transformation**

133 CJ7 transformation was performed at room temperature using the PEG method (28, 29) with the

following modifications. Briefly, CJ7 cells were grown in Halo-2 medium at 45°C to an optical density (OD600) of 0.4–0.6. Generation time was adjusted to be at least 6 h at 37°C and the saline concentration in transformation solutions was changed to 0.8 M. CJ7 cells transfected with pHH205 were induced by MMC as previously described (9) and incubated for another 12 h. The supernatant was collected, mixed with CJ7 cells, and then poured on top of an agar plate. CJ7 transformed with shuttle vectors were plated on 18% MGM containing 5 μg/mL Mev and incubated at 45°C for 7 days protected from light. Transformation efficiency was calculated by counting the number of transformants obtained with 1 μg plasmid DNA.

Plasmid construction

The basic vector, pUC-mev, was created by ligating the 2.3 kb *Eco*RI-*Sac*I Mev resistance fragment amplified from pNBK07 (Table S1), into digested pUC19. pYC-S was generated by inserting the *Sac*I-digested full-length SNJ1 (S fragment) into pUC-mev. A 6.4 kb *Sac*I-*Hin*dIII portion (SHS fragment) digested from the S fragment was ligated into pUC-mev to generate pYC-SHS. Fragments B, C, D, E, F, and G were amplified using the primers listed in Table S2. A series of vectors (pYC-BC, pYC-CD, pYC-DE, pYC-EF, and pYC-FG) was constructed by three-fragment ligation. Their specific locations and the enzymes used for digestion are listed in Fig. 2A.

Portions covering different lengths of the SHS fragment were generated by PCR and named according to their specific locations (the first G base of the *Sac*I site was regarded as position 1). Vectors listed in Fig. 2B were constructed by ligating two truncated parts of the SHS fragment into pUC-mev. For example, pYC1 was generated by ligating *Sac*I-*Sca*I 575−3414 and *Sca*I-*Bam*HI 3414−6482 fragments into *Sac*I-*Bam*HI pUC-mev. The enzymes used for digestion are shown in Table S1 and the names of the new corresponding vectors are listed in Fig. 2B.

157 A fragment encompassing the *amyH* gene and its 200-bp promoter (Apro-amyH) was amplified from the
158 genome of *Haloarcula hispanica* DSM4426 (30, 31). In order to add three restriction sites to the
159 multiple cloning site of pYC13, we designed forward primers bearing restriction sequences for: *Sma*I,
160 *Bam*HI, *Sna*BI, and *Mfe*I; and reverse primers containing sites for *Mfe*I, *Age*I, and *Xba*I. An Apro-amyH
161 fragment containing multiple restriction sites was amplified by PCR, digested by *Sma*I and *Xba*I, and
162 ligated into pYC13 to generate pYC13–A-amyH. pYC13-A-amyH was then digested with *Mfe*I to
163 remove Apro-amyH and re-ligated to create pYC-J. The promoter of *hsp70* (H pro) amplified from the
164 genome of *Haloferax volcanii* DS52 (32) was ligated into pYC-J to generate pYCJ-H. The digested
165 *Hin*dIII-*Nde*I region was amplified with primers bearing a histidine (his)-tag sequence at the *Hin*dIII end.
166 The shuttle expression vector pYCJ-HH containing a his-tag sequence was constructed by replacing the
167 *Hin*dIII-*Nde*I portion of pYCJ-H with the newly cloned fragment, as shown in Fig. S1. All constructs
168 were confirmed by sequencing (Beijing Genomics Institute, China).
169
170 **Mutagenesis of start codons of ORFs located within the minimal replicon**
171 Mutations were introduced at different locations into the 1−3414 bp fragment using overlap PCR;
172 primers are listed in Table S2. These fragments and segment 3414−6482 bp were cloned into pUC-mev
173 to create pYC-SHS-5M, pYC-SHS-6M, pYC-SHS-7M, pYC-SHS-8M, pYC-SHS-9M, pYC-SHS-10M,
174 and pYC-SHS-11-12M. The start codons of ORF5 and ORF7 were both mutated to ATT. In order to
175 keep the amino acid sequence consistent, the start codons of the overlapping ORFs (ORF6, ORF8,
176 ORF9, ORF10, and ORF11) were mutated to ACG, ACG, TTG, TTG, and ACG, respectively. Finally
177 all vectors were transformed into CJ7 and the corresponding transformation efficiency was calculated.
178
179 **Western blot analysis**

Western blot analysis was performed to determine the expression of RepA in J7-1. An exponentially growing J7-1 culture was induced by MMC as previously described (9) and then incubated in fresh Halo-2 medium for 0, 12, 24, and 36 h. Cells were collected and resuspended in distilled water for complete cell lysis. Total protein concentration was measured by the Bradford assay after a short ultrasonication step. Samples with equal amounts of protein were separated by 12% SDS-PAGE and probed with an anti-ORF11-12 antibody (1:500 dilution) as previously described (9). The secondary anti-rabbit antibody was peroxidase-conjugated (1: 1000 dilution). The membranes were revealed by chemiluminescence.

Sequence analyses

The non-redundant database of protein sequences at the NCBI was searched using the PSI-BLAST (33). Protein sequences were aligned with Promals3D (34). The alignment was visualized using Jalview (35). Profile-against-profile searches were performed using HHpred (36) against different protein databases, including PFAM, PDB, CDD, and COG.

Stability and maintenance of the shuttle vectors

To assess the structural stability of vectors bearing different fragments of SNJ1, they were isolated from CJ7 transformants, back-transformed into *E. coli* for amplification, extracted, and subjected to restriction enzyme digestion. Plasmid DNA of the constructs extracted from *E. coli* was used as positive control. Constructs containing different fragments of SNJ1 were tested for plasmid maintenance in CJ7. Briefly, selected clones were cultured in 5 mL 18% MGM supplemented with 5 µg/mL Mev to late exponential phase. The culture was then diluted 1:100 in 5 mL Halo-2 medium and incubated for another 24 h without antibiotic. The dilution process mentioned above was repeated 12 times (the generation time of

CJ7 was about 2.9 h). After each dilution, the cultures were spread onto 18% MGM plates. Fifty random clones were selected, reseeded onto fresh plates with or without Mev, and the clones counted. The survival frequency on 18% MGM with 5 μg/mL Mev was considered to be an indicator of vector maintenance.

Relative copy number determination

Real-time PCR was conducted to estimate the relative copy number of the constructs and pHH205 in CJ7. Total DNA from CJ7 cells containing the different vectors was prepared as described (37). Briefly, cells were harvested from 1 mL aliquots, suspended in 100 μL 18% (w/v) NaCl solution, and lysed after addition of 900 μL distilled water. The lysate was diluted 1:100 in distilled water and used as template. Relative copy numbers of the single copy gene *rad*A (37), located on the chromosome of the host strain, and *repA*, located in the minimal replication region, were determined. The specific primers used for Real-time PCR are listed in Table S2; each experiment was performed in triplicate as previously described (38).

Amylase activity assay

Amylase activity of CJ7/pYC13-A-amyH and JCM8980/pYC13-A-amyH was qualitatively determined as previously described (28). Briefly, 12 random transformants were inoculated into 5 mL 18% MGM. When cultures grew to log phase, 2 μL of the suspension was spotted onto 18% MGM plates supplemented with 2% (w/v) soluble starch and 5 μg/mL Mev. After at least 5 days of incubation at 45°C, iodine solution (0.3% I_2, 0.6% KI) was spread gently on the plate. Haloes around the colonies indicated starch hydrolysis. A strain containing the pYC-J vector and wild type CJ7 were used as negative controls.

Amylase specific activity was measured as previously described by Fuwa (39) with one modification, the 0.5% starch solution was prepared in 18% salt solution (pH7.5) containing 25g NaCI, 50g $MgCl_2 \cdot 6H_2O$, 5g K_2SO_4, and 0.26g $CaCl_2 \cdot 2H_2O$ per litter (25). One unit of amylase activity was defined as the amount of amylase required for the hydrolysis of 1 mg of starch in 1 h. To facilitate comparison, the same amounts of CJ7/pYC13-A-amyH and CJ7/pYCJ-HH-amyH cultures were used for amylase measurements. Briefly, fresh cultures were adjusted to the same biomass and diluted 1:10, 1:100, 1:1000, 1:10 000, and 1:100 000. Cultures (2 μL) at the different dilutions were reseeded onto 2% starch solid plate medium and amylase activity was assessed as described above. The sizes of haloes around the colonies indicated promoter activity.

Protein purification

In order to test whether the his-tag on pYCJ-HH could be used to purify exogenous proteins, 3 mL of fresh CJ7/pYCJ-HH-amyH culture was seeded in 300 mL 18% MGM containing 5 μg/mL Mev at 45°C and cultured for at least 5 days. The cell extract was prepared as previously described (16) and applied onto a 3 mL Ni^+-charged HisTrap affinity column. The purification procedure was performed as described previously (25). Fractions were eluted by imidazole gradient, separated by 12% SDS-PAGE, and probed with anti-his-tag antibodies. CJ7 harboring the empty pYCJ-HH vector was used as negative control.

Journal of Bacteriology

RESULTS

Proviral genome pHH205 can be transfected into CJ7 to produce SNJ1 virions

We first tested whether pHH205 could be transformed into CJ7. CJ7 cells were transformed with pHH205, treated with MMC, incubated for 12 h, and finally the supernatant was collected and tested for plaque formation. As shown in Fig. 1A, the CJ7 culture transformed with pHH205 was able to form plaques, whereas the one transformed with an equal volume of water could not (Fig. 1B), suggesting that successful transformation of pHH205 was achieved in CJ7.

As direct genetic manipulation of pHH205 in CJ7 was not feasible, we constructed an *E. coli-Natrinema* sp. CJ7 shuttle vector, pYC-S, which contained the full-length pHH205 proviral genome and the Mev resistance marker (mevR). Given that CJ7 was sensitive to Mev in 18% MGM but not in Halo-2 medium (Fig. S2), the former was used to select for Mev-resistant transformants. As shown in Fig. 2A, pYC-S could be successfully transformed into CJ7 with an efficiency of about 3.3×10^2 cfu/μg DNA. Noticeably, upon MMC treatment, infectious SNJ1 particles could be induced from pYC-S-transformed CJ7 cells, suggesting that the exogenous fragment inserted into the SNJ1 genome did not affect packaging and release of the virus.

Identification of the minimal replication region of SNJ1

Having established that the shuttle vector pYC-S could be stably maintained in CJ7, we systematically deleted different regions of SNJ1 to determine the portion required for replication of the proviral genome. As shown in Fig. 2A, only pYC-SHS, which lacked the genes encoding for the viral structural proteins identified by Zhang (9), could be efficiently transformed into CJ7. Several transformants were also obtained when vectors pYC-BC and pYC-CD were introduced into CJ7. However, transformation

Journal of Bacteriology

efficiencies were low and their sequences were altered following introduction into CJ7 cells, as observed by enzymatic digestion (data not shown). Thus, we concluded that pYC-SHS contained all the necessary elements for SNJ1 replication.

To further delimit the region required for SNJ1 replication, a series of shorter elements derived from the SHS fragment were ligated into the pUC19-mev vector. As shown in Fig. 2B, pYC1, pYC17, pYC18, and pYC19, which contain different deletions in the 1−575-bp region, were successfully transformed into CJ7 cells. In contrast, pYC2, which contained a deletion in the 575−861-bp region, was not, suggesting that this region was critical for SNJ1 replication. Although a deletion at 575−861 bp may block the transcription and expression of ORF5 and ORF6 in pYC2, we found that mutations in the start codons of ORF5 and ORF6 did not prevent the transformation of pYC-SHS-5M and pYC-SHS-6M, respectively, into CJ7 (Fig. 3C). An AT-rich stretch, usually adjacent to the origin of replication and necessary for DNA unwinding, was found within the 575−861-bp region; however, additional tests are required to confirm its function. On the other end of the SHS fragment, deletion of the 4481−6482-bp region did not affect transformation into CJ7 cells (pYC7 and pYC13). Thus, ORF13−ORF16 are not essential for autonomous replication. Instead, a deletion of the 3881−4481-bp region abolished stable replication of the construct. Taken together, fragment 575−4481 bp, which contains seven putative ORFs, an AT-rich region (576−861 bp), and an intergenic region (3881−4481 bp), seem to represent the minimal regions necessary for SNJ1 replication.

All the fragments mentioned above (excluding SHS and S) were amplified by PCR and sequenced. A cytosine located at position 5078 in the published genome of SNJ1 (GI: AY048850) was not present in the newly obtained sequence. Following the correction, ORF11 and ORF12, originally annotated within the flanking sequences of the missing nucleotide (17), were found to correspond to a single contiguous ORF, hereinafter renamed ORF11-12. In the following experiments, we found that ORF11-12 encoded a

291 protein essential for SNJ1 replication, suggesting that the additional cytosine in the original sequence
292 was likely a sequencing error.
293
294 **Identification of the ORFs necessary for SNJ1 replication**
295 Having defined the minimal replication region of SNJ1 (Fig. 2B, 575−4481bp), we used RT-PCR to
296 determine transcription and co-transcription of the seven predicted ORFs within this region. As shown
297 in Fig. 3A, specific mRNAs of the seven ORFs in addition to co-transcripts of ORF5−ORF6 and
298 ORF7−ORF11-12 were detected in J7-1, suggesting that these predicted ORFs formed two distinct
299 operons (Fig. 3B).
300 To determine which of the seven ORFs were essential for SNJ1 replication, the start codon of each of
301 them was mutated in pYC-SHS, the most stable CJ7 vector. Sequencing of pYC-SHS-5M,
302 pYC-SHS-6M, pYC-SHS-7M, pYC-SHS-8M, pYC-SHS-9M, pYC-SHS-10M, and pYC-SHS-11-12M
303 indicated that mutations were successfully introduced into the vectors. Only pYC-SHS-11-12M
304 completely prevented plasmid transformation into CJ7 cells (Fig. 3C). Transformation efficiencies of
305 pYC-SHS-7M, pYC-SHS-9M, and pYC-SHS-10M were unchanged, whereas those of the remaining
306 plasmids were significantly reduced (< 10^2 cfu/μg DNA). Overall, we conclude that only ORF11-12 is
307 indispensable for SNJ1 replication, while ORF5, ORF6, and ORF8 promote replication efficiency of
308 SNJ1. It remains to be determined whether the mRNAs of the identified ORFs are indeed functional.
309 Similarly, we cannot exclude the presence of other functional ORFs located in the replication region.
310 To further confirm the indispensability of ORF11-12 in SNJ1 replication, expression of ORF11-12
311 protein in J7-1 cells was assessed at different time points (0, 12, 24, and 36 h) after MMC induction
312 (lanes 2, 3, 4, and 5 in Fig. 4, respectively). ORF11-12 protein was detected in J7-1 before and after
313 MMC induction, but not in CJ7 cells (lacking the provirus), confirming that ORF11-12 encodes a

functional protein. The level of ORF11-12 protein increased with incubation time following induction. At the same time, OD600 of the culture first increased and then dropped, whereas the titer of SNJ1 increased gradually. Collectively, these results indicate that ORF11-12 encodes an essential protein required for the replication of SNJ1 and its expression in J7-1 increases along with the accumulation of extracellular virions. Thus, ORF11-12 was designated as the *repA* gene of SNJ1.

Sequence analysis of the predicted proteins encoded within the minimal replicon

In light of the new results, we decided to repeat the sequence analysis of the predicted proteins encoded within the SHS fragment (Fig. 2B) encompassing the minimal replicon of SNJ1 as well as several other ORFs (ORFs 4 to 16). Ten of the twelve putative proteins could be functionally annotated. Interestingly, sensitive HHpred searches showed that eight of the ten gene products (gp) carry various DNA-binding domains (Fig. S3), suggesting their involvement in transcription regulation. Specifically, gp4, gp5 and gp15 contain AbrB-like swapped-hairpin domain (Fig. S3A-C), which is found in various bacterial and archaeal transcriptional regulators, including those encoded by some archaeal plasmids (40) and within toxin-antitoxin systems (41, 42). Gp7 and gp10 contain helix-turn-helix DNA-binding domains (Fig. S3D and E, respectively), most closely related to those found in bacterial sigma-70-like sigma factors (43), whereas gp13 was found to contain an N-terminal winged helix-turn-helix domain followed by an additional C-terminal helix-turn-helix domain (Fig. S3F). Furthermore, gp8 and gp9 possess distinct Zn-finger motifs (Fig. S3G and H, respectively).

Interestingly, gp14 encodes a VapC-like toxin component of the VapBC toxin-antitoxin system (Fig. S3I), widespread in archaea (44) and also present in some archaeal plasmids (45). In the VapBC pair, VapC is the PIN-domain containing ribonuclease, whereas VapB is typically a transcription factor with the AbrB-like swapped-hairpin domain (46). As mentioned above, gp15 contains the AbrB-like

337 DNA-binding domain and is encoded immediately upstream of ORF14, strongly suggesting that the two
338 proteins, gp14 and gp15, constitute a functional toxin-antitoxin system of SNJ1.
339 HHpred analysis of the SNJ1 RepA sequence revealed only a partial match to the IS91-like Zn-binding
340 motif (Fig. S3J). PSI-BLASTP searches seeded with the sequence of gp11-12 (RepA), after the first
341 iteration resulted in significant matches restricted to hypothetical proteins encoded by halophilic archaea,
342 namely *Natronococcus jeotgali* (GI:495694890) and *Halococcus thailandensis* JCM 13552
343 (GI:445806310). However, following additional iterations, significant hits to Rep proteins encoded by
344 plasmid pC2A of *Methanosarcina acetivorans* (47) and plasmid pZMX101 of *Halorubrum*
345 *saccharovorum* (48) were obtained. The relationship among these proteins was further investigated by
346 constructing a multiple sequence alignment (Fig. 5). Plasmid pC2A has been previously suggested to
347 replicate via the rolling-circle mechanism (47). Consistently, analysis of the conserved motifs shared by
348 RepA of SNJ1 and those of other plasmids, allowed delineation of the three motifs characteristic of the
349 rolling circle replication initiation proteins belonging to the HUH endonuclease superfamily (49, 50).
350
351 **Stability and maintenance of the shuttle vectors**
352 In order to assess the structural stability of the shuttle vectors, plasmid DNA extracted from three
353 independent transformants was back-transformed into *E. coli* (post-transformed plasmids) and
354 re-isolated for enzyme digestion analysis. As shown in Fig. 6A, pre- (lanes 1, 2, and 3) and post-
355 transformed (lanes 4, 5, and 6) plasmids showed similar digestion profiles, indicating that structural
356 stability of all constructs was maintained in CJ7.
357 The proviral genome of SNJ1, pHH205, was stably maintained in J7-1, suggesting a faithful partition
358 into daughter cells at division. To determine which region(s) were critical for its ongoing maintenance,
359 stability of the shuttle vectors was evaluated in the absence of Mev. As shown in Fig. 6B, pYC-SHS

could be stably maintained in CJ7 cells for 12 days (about 100 generations), suggesting that it contained all the necessary elements for correct partitioning. All other constructs showed different degrees of plasmid loss. pYC1 was completely lost after 2 days, suggesting that the 1−575-bp region was extremely important for maintenance of the plasmid. pYC17, pYC18, and pYC19 were constructed to further narrow down the region required for stable maintenance of pYC-SHS. As shown in Fig. 6B, pYC18 and pYC19 containing a deletion in the 1−294-bp and 1−140-bp regions, respectively, were relatively stable in CJ7. However, pYC17, which lacked the 1−401-bp region, was highly unstable, implying that the 294−401 nucleotide sequence was essential for the stable maintenance of the vectors. Deletion at the other end of the SHS fragment also significantly affected plasmid stability. Vector pYC7, which contains a deletion in the 5160−6482-bp region (ORFs 14−16), started to disappear from CJ7 cells after 6 days. After 12 days, most of the cells had lost the plasmid. When the fragment was shortened from 5510 to 4481 bp (pYC13), the vector was lost completely within 9 days. Overall, these results suggest that the 1−575-bp region is critical for the maintenance of SNJ1, whereas the 4481−6482-bp region affects maintenance to a certain extent.

To test which ORFs were critical for the maintenance of the vector, the stability of plasmids containing start codon mutations in ORF4−ORF13 (except for ORF11-12, which was essential for replication) were determined. As shown in Fig. 6C, mutation in the start codon of ORF4 resulted in a complete loss of plasmid pYC-SHS-4M within 2 days, suggesting that the product of ORF4 was critical for the maintenance of the shuttle vectors. Surprisingly, pYC-SHS-13M, containing a mutation in the start codon of ORF13 was stably maintained after 12 days without antibiotic (Fig. 6C). Mutations in the start codons of ORF5, ORF6, and ORF8 seemed to affect maintenance of these plasmids, but also transformation efficiency. Instead, mutations in the start codons of ORF7, ORF9, and ORF10 did not seem to affect plasmid stability suggesting that, like ORF13, the products of these ORFs did not play a

role in the maintenance of the plasmid, at least not under the tested conditions. Noticeably, pYC7 was relatively unstable, indicating that, although not crucial, ORF14, ORF15, and ORF16 may also affect plasmid maintenance.

Copy number of the SNJ1 proviral genome and shuttle vectors

Temperate viruses existing as plasmids during the lysogenic cycle are usually maintained in low copy number in their hosts (51). As shown in Fig. 7A, pHH205 was maintained at 1−3 copies per chromosome in J7-1 cells determined by Real-time PCR. Except for pYC1, pYC17, and pYC20, most of the shuttle vectors were also present at 1−3 copies per chromosome. However, deletion of the 1−575 bp portion in the stable replication fragment (1−6428, pYC-SHS) increased the number of pHH205 to about 30 copies per chromosome, suggesting that this region was crucial for controlling the copy number of SNJ1. Since this region contained only one ORF, ORF4, the copy number of pYC-SHS-4M was determined. As shown in Fig. 7B, pYC-SHS-4M yielded a similar copy number as pYC1. Interestingly, we found that mutating the start codon of the other ORFs located in the minimal replication region did not affect copy number. These results suggest that only ORF4 negatively regulates the copy number of the SNJ1 proviral genome and shuttle vectors.

Construction of expression vectors for J7 using the stable replicon of SNJ1

As shown in Fig. S1, we constructed the cloning vector pYC-J and expression vectors pYCJ-HH, based on the stable replicon of SNJ1 (1−4481 fragment). Similar to pHH205 and its parent shuttle vector pYC13, Real-time PCR revealed that pYC-J and pYCJ-HH were maintained at 1−3 copies per CJ7 chromosome (data not shown). To determine the applicability of these shuttle vectors, the *amyH* amylase gene was cloned from *Haloarcula hispanica* DSM 4426 and inserted into pYC13 to generate

406 pYCJ-A-amyH. After transforming pYCJ-A-amyH into CJ7, 12 randomly selected colonies were
407 transferred onto 18% MGM plates supplemented with 2% (w/v) soluble starch and 5 µg/mL Mev. As
408 shown in Fig. 8, haloes were detected around the colonies after exposure to iodine solution, suggesting
409 that AmyH was well expressed. On the contrary, CJ7 without the plasmid could not grow on the Mev
410 plate (N1), and those harboring the empty pYC13 vector showed no enzymatic activity (N2), confirming
411 that starch consumption was indeed due to expression of AmyH.
412 Vector pYCJ-HH contained a *Haloferax volcanii* DS52 *hsp70* promoter (H pro), which showed higher
413 activity than A pro (Y. Huang, unpublished data), followed by a multiple cloning site and a His-tag for
414 protein expression and purification (Fig. S4A and B). Single enzyme digestion of pYCJ-HH by eight
415 different restriction enzymes showed that the cleavage sites within the multiple cloning site could easily
416 generate sticky ends (Fig. S4C). The *amyH* gene was inserted into pYCJ-HH and amylase was purified
417 from 300 mL of fresh CJ7/pYCJ-HH-amyH culture. As shown in Fig. S4D, the 48 kDa His-tagged
418 amylase zymogen protein was successfully expressed and purified from CJ7 by affinity purification, and
419 then confirmed using an anti-His-tag antibody (Fig. S4E). Therefore, pYCJ-HH could be used for
420 protein purification in *Natrinema* species.
421 A pro (30) was the only promoter that had already been used for efficient expression of exogenous
422 proteins in J7 (28). In order to compare the activities of H pro and A pro in CJ7, amylase activity of
423 CJ7/pYC13-A-amyH and CJ7/pYCJ-HH-amyH cultures was assessed. As shown in Fig. S5A, H pro
424 activity during different growth periods was higher than that of A pro. Moreover, haloes around
425 CJ7/pYCJ-HH-amyH colonies were substantially bigger than those around CJ7/pYC13-A-amyH
426 colonies at the same dilution (Fig. S5B), confirming higher H pro activity.
427
428 **pYCJ-HH host range**

Finally, we tested whether pYCJ-HH could be transformed into haloarchaeal species other than J7, such as the model organisms in *Natrinema* sp. (Table 1). While all strains were successfully transfected with pHH205, only those sensitive to 5 μg/mL Mev, were transformed with pYCJ-HH. This was achieved only in *Natrinema pallidum* JCM8980, with an efficiency of 10^3 cfu/μg DNA, somewhat lower than that in CJ7. After a 12 h recovery period, the supernatant of pHH205-transfected JCM8980 was induced to form plaques on CJ7, suggesting that SNJ1 could replicate and be released from JCM8980. However, as previously described (9), SNJ1 virions could not adsorb to or enter JCM8980 cells.

Having established that pHH205 could be transformed into JCM8980, we tested whether the shuttle expression vector could be used to express proteins in this species. As shown in Fig. 8B, plasmid pYCJ-HH-amyH was successfully transformed into JCM8980 and amylase was expressed to significant levels. These results suggest that pYCJ-HH might be a valuable genetic tool for studies in several *Natrinema* species.

DISCUSSION

Archaeal viruses present strikingly diverse virion morphology and unique gene content (52). Currently, little is known about the mechanisms controlling their life cycles, such as entry, genome replication, virion assembly, and release (53). In this study, we identified the replication region of *Betasphaerolipovirus* SNJ1, and constructed several SNJ1-based *E. coli-Natrinema* sp. J7 shuttle vectors. A number of ORFs and genetic elements were found to be involved in replication, maintenance, and copy number control of SNJ1. Regardless of their exact function, their identification will enable the study of the viral life cycle of temperate haloarchaeal viruses such as SNJ1. Based on the replication region of SNJ1, a set of expression vectors was validated for protein expression and purification in J7 and JCM8980 cells. These vectors will undoubtedly facilitate studies of SNJ1 and other viruses infecting *Natrinema* species, as well as their hosts.

The replicon of SNJ1

Although many archaeal viruses and plasmids have been discovered over the last few decades, genome replication has been studied in only a handful of them (54-57). In this study, a systemic deletion analysis of SNJ1 proviral genome, pHH205, allowed us to define the minimal region required for virus replication. This region is a 3.9-kb fragment (575−4481 bp), containing seven ORFs organized into two operons. One ORF, *repA*, and two genetic elements, encompassing regions 576−861 bp and 3881−4481 bp, were found to be essential for SNJ1 replication, with other ORFs significantly affecting replication. Their functions and roles in replication are discussed in the following section.

The region containing *repA* was originally predicted to harbor two ORFs, ORF11 and ORF12. However, after correcting a previous sequencing error, we found that the two ORFs were in fact a single gene, *repA*. We confirmed that *repA* encoded a functional protein and found that its expression in J7-1 cells

increased along with the accumulation of extracellular viruses (Fig. 4). Accordingly, RepA is probably responsible for genome replication during the lytic cycle of SNJ1. Specifically, RepA shows significant similarity to Rep proteins encoded by small archaeal plasmids, such as pC2A (5.4 kb) and pZMX101 (3.9 kb). All the homologs contain conserved signature motifs characteristic of rolling-circle replication proteins (49, 50). Interestingly, whereas pC2A was indeed found to replicate via the rolling-circle mechanism (47, 58), pZMX101 was suggested to employ a theta-like replication strategy (48). This discrepancy has been pointed out previously (59) but remains unresolved. Studies on the mechanism of SNJ1 genome replication and the exact function of RepA in this process are currently under way.

Two genetic elements have been identified as critical for SNJ1 replication in this study, one located in an AT-rich region (576−861 bp) and the other in an intergenic region (3881−4481 bp). Deletion of the former eliminates part of ORF5 and could potentially affect the transcription of ORF6. However, we found that substitutions in the start codons of ORF5 and ORF6 did not prevent the plasmid from replicating in J7 cells (Fig. 3C), suggesting that gene products of ORF5 and ORF6 were not essential for SNJ1 replication. The GC content of this region is 49.9%, significantly lower than the average for SNJ1 (61.6%), which makes it a likely part of the SNJ1 origin of replication. Additional evidence is required to prove this hypothesis. Although it remains unclear how the second genetic element contributes to the replication of SNJ1, we found a number of repeated sequences —CGTCGACG— within this region. In fact, this palindromic sequence occurs up to 13 times in the 1−6428 bp fragment of SNJ1. Whether these repeats are important for SNJ1 replication remains to be seen and is currently under investigation.

ORF7, ORF9, and ORF10 within the minimum replication region of SNJ1 encode putative DNA-binding proteins. Changing the start codon of these ORFs did not affect the transformation efficiency of shuttle vectors (Fig. 3C), indicating that they were not required for replication. Although ORF5, ORF6, and ORF8 were also not strictly required, mutations in their start codons significantly

488 reduced transformation efficiency. Currently, we do not know how these ORFs affect SNJ1 replication.
489 Our preliminary data indicates that the mutated shuttle vectors become rather unstable during culturing.
490 Perhaps these ORFs function as accessory factors for replication and in their absence replication
491 proceeds less efficiently. Notably, ORF5 and ORF8 encode putative DNA-binding proteins, which might
492 act as transcriptional regulators of the cognate *repA* operon.
493
494 **Maintenance and copy number control of SNJ1**
495 The genome of temperate prokaryotic viruses can exist as a plasmid or integrate into the host
496 chromosome during the lysogenic cycle. Integration enables the viral genome to be replicated and
497 partitioned passively during the host cell cycle. Instead, when existing as plasmids (e.g., pHH205 for
498 SNJ1), the proviruses have to coordinate their replication and partitioning with the host cell cycle.
499 Otherwise, they will be quickly lost from the host cells. DNA segregation has been well studied in
500 bacteria (60). The most widespread and well understood bacterial DNA segregation system is comprised
501 of three components: a DNA centromere, a ParA ATPase, and a centromere-binding protein, ParB (61,
502 62). Similar segregation systems have been reported in *Sulfolobus* species, which also contain ParA-like
503 ATPase and site-specific DNA-binding proteins that attach to centromere-like sequences (63-65).
504 Inspection of the ORFs within the stable replicated fragment of SNJ1 (pYC-SHS) did not reveal any
505 ParA- or ParB-like ORFs, nor did we find candidates similar to the identified archaeal
506 centromere-binding proteins. However, we found that the stability of SNJ1 could be dramatically altered
507 following localized deletions and mutations in some ORFs, suggesting that these regions and ORFs have
508 important roles in the segregation of SNJ1. Among these factors, ORF4 is perhaps the most important
509 one, because its deletion resulted in the loss of shuttle vectors within 2 days in the absence of antibiotic
510 selection (Fig. 6C). How ORF4 regulates the stability of SNJ1 is not clear, but sequence analysis

511 suggests that it belongs to the SpoVT/AbrB family of transcription regulators (66). It is possible that
512 ORF4 regulates the expression of proteins required for partitioning of SNJ1. In its absence, these
513 proteins are not expressed, resulting in plasmid loss. However, we cannot exclude the possibility that
514 ORF4 itself is a component of the partitioning system, given that it is predicted to bind dsDNA. In
515 addition, deletion of the 1−294 bp (pYC18 and pYC19) and 4481−5510 bp (pYC7 and pYC13) regions
516 from the stable pYC-SHS vector also significantly reduces stability of the plasmid, suggesting that these
517 regions may contain sequences that are important for plasmid partitioning. Alternatively, the proviral
518 genome of SNJ1 might not be actively partitioned during cell division. It is worth mentioning that the
519 VapBC toxin-antitoxin system encoded by ORF14 and ORF15 might be important for ensuring retention
520 of SNJ1 by a post-segregational killing mechanism (41, 67). Accordingly, cells that did not acquire the
521 toxin-antitoxin-encoding proviral genome and do not produce the labile VapB antitoxin component are
522 killed by the VapC toxin.
523 Temperate prokaryotic viruses also need to control their copy number during the lysogenic cycle,
524 otherwise they risk becoming a burden for their hosts. Integrated proviruses avoid this problem, but
525 proviral genomes existing as plasmids depend on a tight control of their copy number. Bacterial
526 plasmids maintain their copy number by self-encoded negative control systems (68). How archaeal
527 plasmids control their copy number is poorly understood. In this study, we found that ORF4 was critical
528 for SNJ1 copy number control (Fig. 7B). It is possible that ORF4 regulates plasmid copy number by
529 controlling the level of RepA, or that of other proteins or RNAs. It will be interesting to determine how
530 ORF4 controls both the stability and copy number of the SNJ1 genome.
531
532 ***Natrinema* species shuttle vectors based on the SNJ1 proviral genome**
533 Establishing a genetic system for halophilic archaea is not straightforward. Currently, efficient genetic

systems are available for only two haloarchaeal model organisms: *Halobacterium salinarum* and *Haloferax volcanii* (69). Replication origins used to construct most of the replicative shuttle vectors were obtained from indigenous haloarchaeal plasmids, such as pGRB1 (70), pHH1 (71), and pNRC100 (72) in *Halobacterium salinarum*, and pHK2 (73, 74), pHV2 (75, 76), and pHV1/4 (77) in *Haloferax volcanii*. Viral replicons represent another important source. For example, shuttle vectors pUBP1 and pUBP2, which were constructed on the basis of the ΦH replicon, were artificially introduced into *Halobacterium salinarum* and *Haloferax volcanii* spheroplasts using PEG (78). The genomic DNA of several other haloarchaeoviruses, such as His2 and SH1, has also been successfully transformed into their hosts (13). However, no replicative shuttle vector is available for J7, limiting our ability to study J7 and its viruses. Here, we used the replicon region of SNJ1, to construct a set of shuttle vectors for protein expression and purification.

Recently, an exogenous gene was successfully expressed upon integration in the genome of two J7 auxotrophic mutants (28). This system is complementary to the vectors developed in our study. Many genetic experiments such as complementation tests will become possible by a combination of these two genetic tools. We expect that research on SNJ1 and other viruses infecting *Natrinema* species, as well as studies of *Natrinema* species themselves, will rapidly lead to the future utilization of these vectors.

ACKNOWLEDGMENTS

We thank Dr. Shishen Du (University of Kansas Medical Center, USA) for many helpful discussions and for revising the manuscript. We also thank Dr. David Prangishvili (Pasteur Institute, France) and Dr. Yoshizumi Ishino (Kyushu University, Japan) for giving valuable suggestions for the manuscript.

FUNDING INFORMATION

This work was supported by grants from the National Basic Research Program of China (973 Program) (no. 2011CB808800), the National Natural Science Foundation of China (no. 31570174), the National Found for Fostering Talents of Basic Sciences (J1103513) and Research (Innovative) Fund of Laboratory Wuhan University.

REFERENCES

1. **Atanasova N, Oksanen H, Bamford D.** 2015. Haloviruses of archaea, bacteria, and eukaryotes. Current Opinion in Microbiology **25:**40-48.

2. **Guixaboixereu N, Calderonpaz J, Heldal M, Bratbak G, Pedrosalio C, Pedrosalio C.** 1996. Viral lysis and bacterivory as prokaryotic loss factors along a salinity gradient. Aquatic Microbial Ecology **11:** 213-227.

3. **Williams T, Erdmann S, Cavicchioli R.** 2014. Viruses of Haloarchaea. Life **4:**681-715.

4. **Pina M, Bize A, Forterre P, Prangishvili D.** 2011. The archeoviruses. FEMS Microbiology Reviews **35:**1035-1054.

5. **Alice Pawlowski IR, Jaana K, H. Bamford, Mart Krupovic, Matti Jalasvuori.** 2014. Gammasphaerolipovirus, a newly proposed bacteriophage genus, unifies viruses of halophilic archaea and thermophilic bacteria within the novel family Sphaerolipoviridae. Archives of Virology **159:**1541.

6. **Dyall-Smith M, Tang S-L, Bath C.** 2003. Haloarchaeal viruses: how diverse are they? Research in microbiology **154:**309-313.

7. **Jaakkola ST, Penttinen RK, Vilén ST, Jalasvuori M, Rönnholm G, Bamford JK, Bamford DH, Oksanen HM.** 2012. Closely related archaeal Haloarcula hispanica icosahedral viruses HHIV-2 and SH1 have nonhomologous genes encoding host recognition functions. Journal of

580 virology **86:**4734-4742.

581 8. **Porter K, Tang S, Chiang P, Hong M, Dyallsmith M.** 2013. PH1: An Archaeovirus of
582 Haloarcula hispanica Related to SH1 and HHIV-2, vol 2013.

583 9. **Zhang ZQ, Liu Y, Wang S, Yang D, Cheng YC, Hu JN, Chen J, Mei YJ, Shen P, Bamford**
584 **DH, Chen XD.** 2012. Temperate membrane-containing halophilic archaeal virus SNJ1 has a
585 circular dsDNA genome identical to that of plasmid pHH205. Virology **434:**233-241.

586 10. **Jalasvuori M, Jaatinen ST, Laurinavičius S, Ahola-Iivarinen E, Kalkkinen N, Bamford DH,**
587 **Bamford JK.** 2009. The closest relatives of icosahedral viruses of thermophilic bacteria are
588 among viruses and plasmids of the halophilic archaea. Journal of virology **83:**9388-9397.

589 11. **Matsushita I, Yanase H.** 2009. The genomic structure of Thermus bacteriophage ϕIN93. Journal
590 of biochemistry **146:**775-785.

591 12. **Bamford DH, Ravantti JJ, Rönnholm G, Laurinavičius S, Kukkaro P, Dyall-Smith M,**
592 **Somerharju P, Kalkkinen N, Bamford JK.** 2005. Constituents of SH1, a novel lipid-containing
593 virus infecting the halophilic euryarchaeon Haloarcula hispanica. Journal of virology
594 **79:**9097-9107.

595 13. **Porter K, Dyall‐Smith ML.** 2008. Transfection of haloarchaea by the DNAs of spindle and
596 round haloviruses and the use of transposon mutagenesis to identify non‐essential regions.
597 Molecular microbiology **70:**1236-1245.

598 14. **de Grado M, Lasa I, Berenguer J.** 1998. Characterization of a plasmid replicative origin from
599 an extreme thermophile. FEMS microbiology letters **165:**51-57.

600 15. **Matsushita I, Yanase H.** 2011. Characterization of the protein gp1 from bacteriophage φIN93.
601 Biological Letters **48:**47-56.

602 16. **Shen P, Chen Y.** 1993. Plasmid from Halobacterium halobium and its restriction map. Yi chuan

xue bao **21:**409-416.

17. **Ye X, Ou J, Ni L, Shi W, Shen P.** 2003. Characterization of a novel plasmid from extremely halophilic Archaea: nucleotide sequence and function analysis. FEMS microbiology letters **221:**53-57.

18. **Krupovic M, Prangishvili D, Hendrix R, Bamford D.** 2011. Genomics of Bacterial and Archaeal Viruses: Dynamics within the Prokaryotic Virosphere, Microbiology and Molecular Biology Reviews **75(4):**610-635.

19. **Luo W, Shen P.** 1995. Studies on protoplast fusion of halobacterium halobium by killed-parents method. Wuhan University Natural Science edition**:**751-756.

20. **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. Canadian journal of microbiology **53:**1106-1110.

21. **Casjens S.** 2003. Prophages and bacterial genomics: what have we learned so far? Molecular microbiology **49:**277-300.

22. **Casjens SR, Hendrix RW.** 2015. Bacteriophage lambda: Early pioneer and still relevant. Virology **479:**310-330.

23. **Quemin ER, Lucas S, Daum B, Quax TE, Kühlbrandt W, Forterre P, Albers S-V, Prangishvili D, Krupovic M.** 2013. First insights into the entry process of hyperthermophilic archaeal viruses. Journal of virology **87:**13379-13385.

24. **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. Molecular microbiology **98(6)**:1002-20.

25. **Dyall-Smith M.** 2008. The Halohandbook: protocols for haloarchaeal genetics. Haloarchaeal

Genetics Laboratory, Melbourne:2008.

26. **Nuttall SD, Dyall-Smith ML.** 1993. HF1 and HF2: novel bacteriophages of halophilic archaea. Virology **197:**678-684.

27. **Green MR, Sambrook J.** 2012. Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press New York vol 1.

28. **Lv J, Wang S, Wang Y, Huang Y, Chen X.** 2015. Isolation and molecular identification of auxotrophic mutants to develop a genetic manipulation system for the haloarchaeon Natrinema sp. J7-2. Archaea **2015:**483194.

29. **Cline SW, Lam WL, Charlebois RL, Schalkwyk LC, Doolittle WF.** 1989. Transformation methods for halophilic archaebacteria. Canadian Journal of Microbiology **35:**148-152.

30. **Zeng C, Zhao Y-Z, Cui C-Z, Zhang H, Zhu J-Y, Tang X-F, Shen P, Huang Y-P, Chen X-D.** 2009. Characterization of the Haloarcula hispanica amyH gene promoter, an archaeal promoter that confers promoter activity in Escherichia coli. Gene **442:**1-7.

31. **Lv J, Wang S, Zeng C, Huang Y, Chen X.** 2013. Construction of a shuttle expression vector with a promoter functioning in both halophilic Archaea and Bacteria. FEMS microbiology letters **349:**9-15.

32. **Chen W, Yang G, He Y, Zhang S, Chen H, Shen P, Chen X, Huang Y-P.** 2015. Nucleotides Flanking the Start Codon in hsp70 mRNAs with Very Short 5'-UTRs Greatly Affect Gene Expression in Haloarchaea. PloS one **10:**e0138473.

33. **Altschul S, Madden T, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman D.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25:**3389-3402.

34. **Pei J, Grishin N.** 2014. PROMALS3D: multiple protein sequence alignment enhanced with

evolutionary and three-dimensional structural information, vol 1079.

35. **Waterhouse A, Procter J, Martin D, Clamp M, Barton G.** 2009. Jalview Version 2 – a multiple sequence alignment editor and analysis workbench. Bioinformatics **25:**1189-1191.

36. **Soding J, Biegert A, Lupas A.** 2005. The HHpred interactive server for protein homology detection and structure prediction, Nucleic Acids Research **33**:W244–W248.

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

38. **Zhang H, Cui P, Lin L, Shen P, Tang B, Huang Y-P.** 2009. Transcriptional analysis of the hsp70 gene in a haloarchaeon Natrinema sp. J7 under heat and cold stress. Extremophiles **13:**669-678.

39. **Fuwa H.** 1954. A new method for microdetermination of amylase activity by the use of amylose as the substrate. The Journal of Biochemistry **41:**583-603.

40. **Bartolucci S, De Simone G, She Q, Aucelli T, Pedone E, Pirone L, Dambrosio K, Contursi P.** 2011. C68 from the Sulfolobus islandicus plasmid–virus pSSVx is a novel member of the shAbrB-like transcription factor family. Biochemical Journal **435(1)**:157-66

41. **Unterholzner S, Poppenberger B, Rozhon W.** 2013. Toxin-antitoxin systems: Biology, identification, and application, Mob Genet Elements **3(5)**: e26219.

42. **Wessner F, Lacoux C, Goeders N, Dherouel A, Matos R, Serror P, Van Melderen L, Repoila F.** 2015. Regulatory crosstalk between type I and type II toxin-antitoxin systems in the human pathogen Enterococcus faecalis. RNA Biology **12(10)**:1099-108.

43. **Campbell E, Muzzin O, Chlenov M, Sun J, Olson C, Weinman O, Tresterzedlitz M, Darst S.** 2002. Structure of the Bacterial RNA Polymerase Promoter Specificity σ Subunit. Molecular Cell **9:**527.

672 44. **Arcus V, Mckenzie J, Robson J, Cook G.** 2011. The PIN-domain ribonucleases and the
673 prokaryotic VapBC toxin–antitoxin array, Protein Engineering, Design & Selection
674 **24(1-2)**:33-40.
675 45. **Krupovic M, Gonnet M, Hania W, Forterre P, Erauso G.** 2013. Insights into dynamics of
676 mobile genetic elements in hyperthermophilic environments from five new Thermococcus
677 plasmids. PLOS ONE **8:**e49044.
678 46. **Cooper C, Daugherty A, Tachdjian S, Blum P, Kelly R.** 2009. Role of vapBC toxin-antitoxin
679 loci in the thermal stress response of Sulfolobus solfataricus. Biochemical Society Transactions
680 **37:**123-126.
681 47. **Metcalf W, Zhang J, Apolinario E, Sowers K, Wolfe R.** 1997. A genetic system for Archaea of
682 the genus Methanosarcina: Liposome-mediated transformation and construction of shuttle
683 vectors. Proceedings of the National Academy of Sciences of the United States of America
684 **94:**2626.
685 48. **Zhou L, Zhou M, Sun C, Xiang H, Tan H.** 2007. Genetic analysis of a novel plasmid
686 pZMX101 from Halorubrum saccharovorum: determination of the minimal replicon and
687 comparison with the related haloarchaeal plasmid pSCM201. Fems Microbiology Letters
688 **270:**104-108.
689 49. **Ilyina T, Koonin E.** 1992. Conserved sequence motifs in the initiator proteins for rolling circle
690 DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria.
691 Nucleic Acids Research **20:**3279-3285.
692 50. **Chandler M, La Cruz F, Dyda F, Hickman A, Moncalian G, Tonhoang B.** 2013. Breaking and
693 joining single-stranded DNA: the HUH endonuclease superfamily. Nature Reviews
694 Microbiology **11:**525.

695 51. **Sternberg N, Austin S.** 1981. The maintenance of the P1 plasmid prophage. Plasmid **5:**20-31.

696 52. **Prangishvili D.** 2013. The wonderful world of archaeal viruses. Annual review of microbiology
697 **67:**565-585.

698 53. **Dellas N, Snyder JC, Bolduc B, Young MJ.** 2014. Archaeal viruses: diversity, replication, and
699 structure. Annual Review of Virology **1:**399-426.

700 54. **Gardner A, Bell S, White M, Prangishvili D, Krupovic M.** 2014. Protein-Protein Interactions
701 Leading to Recruitment of the Host DNA Sliding Clamp by the Hyperthermophilic Sulfolobus
702 islandicus Rod-Shaped Virus 2. Journal of Virology **88:**7105-7108.

703 55. **Oke M, Kerou M, Liu H, Peng X, Garrett R, Prangishvili D, Naismith J, White M.** 2011. A
704 Dimeric Rep Protein Initiates Replication of a Linear Archaeal Virus Genome: Implications for
705 the Rep Mechanism and Viral Replication. Journal of Virology **85:**925-931.

706 56. **Pina M, Basta T, Quax T, Joubert A, Baconnais S, Cortez D, Lambert S, Cam E, Bell S,**
707 **Forterre P, Prangishvili D.** 2014. Unique genome replication mechanism of the archaeal virus
708 AFV1. Molecular Microbiology **92:**1313-1325.

709 57. **Mayrhoferiro M, Ladurner A, Meissner C, Derntl C, Reiter M, Haider F, Dimmel K,**
710 **Rossler N, Klein R, Baranyi U, Scholz H, Witte A.** 2013. Utilization of Virus ϕCh1 Elements
711 To Establish a Shuttle Vector System for Halo(alkali)philic Archaea via Transformation of
712 Natrialba magadii. Applied and Environmental Microbiology **79:**2741-2748.

713 58. **Sowers K, Gunsalus R.** 1988. Plasmid DNA from the acetotrophic methanogen Methanosarcina
714 acetivorans. Journal of Bacteriology **170:**4979.

715 59. **Forterre P, Krupovic M, Raymann K, Soler N.** 2015. Plasmids from Euryarchaeota,
716 Microbiology Spectrum **2(6)**.

717 60. **Gerdes K, Howard M, Szardenings F.** 2010. Pushing and pulling in prokaryotic DNA

segregation. Cell **141:**927-942.

61. **Gerdes K, Møller‐Jensen J, Jensen RB.** 2000. Plasmid and chromosome partitioning: surprises from phylogeny. Molecular microbiology **37:**455-466.

62. **Larsen RA, Cusumano C, Fujioka A, Lim-Fong G, Patterson P, Pogliano J.** 2007. Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in Bacillus thuringiensis. Genes & development **21:**1340-1352.

63. **She Q, Phan H, Garrett RA, Albers S-V, Stedman KM, Zillig W.** 1998. Genetic profile of pNOB8 from Sulfolobus: the first conjugative plasmid from an archaeon. Extremophiles **2:**417-425.

64. **Peng X, Brügger K, Shen B, Chen L, She Q, Garrett RA.** 2003. Genus-specific protein binding to the large clusters of DNA repeats (short regularly spaced repeats) present in Sulfolobus genomes. Journal of bacteriology **185:**2410-2417.

65. **Schumacher MA, Tonthat NK, Lee J, Rodriguez-Castañeda FA, Chinnam NB, Kalliomaa-Sanford AK, Ng IW, Barge MT, Shaw PL, Barillà D.** 2015. Structures of archaeal DNA segregation machinery reveal bacterial and eukaryotic linkages. Science **349:**1120-1124.

66. **Hsu C-H, Wang AH-J.** 2011. The DNA-recognition fold of Sso7c4 suggests a new member of SpoVT-AbrB superfamily from archaea. Nucleic acids research **39(15)**:6764-74.

67. **Gerdes K, Rasmussen P, Molin S.** 1986. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. Proceedings of the National Academy of Sciences of the United States of America **83:**3116.

68. **Del Solar G, Espinosa M.** 2000. Plasmid copy number control: an ever‐growing story. Molecular microbiology **37:**492-500.

69. **Berquist BR, Müller JA, DasSarma S.** 2006. 27 Genetic Systems for Halophilic Archaea.

Methods in microbiology **35:**649-680.

70. **Krebs MP, Hauss T, Heyn MP, RajBhandary UL, Khorana HG.** 1991. Expression of the bacterioopsin gene in Halobacterium halobium using a multicopy plasmid. Proceedings of the National Academy of Sciences **88:**859-863.

71. **Pfeifer F, Ghahraman P.** 1993. Plasmid pHH1 of Halobacterium salinarium: characterization of the replicon region, the gas vesicle gene cluster and insertion elements. Molecular Genetics and Genomics **238(1-2)**:193-200.

72. **Dassarma S.** 1993. Minimal replication origin of the 200-kilobase Halobacterium plasmid pNRC100. Journal of Bacteriology **175:**4584-4596.

73. **Holmes M, Pfeifer F, Dyall-Smith M.** 1994. Improved shuttle vectors for Haloferax volcanii including a dual-resistance plasmid. Gene **146:**117-121.

74. **Holmes M, Dyall-Smith M.** 1990. A plasmid vector with a selectable marker for halophilic archaebacteria. Journal of bacteriology **172:**756-761.

75. **Allers T, Ngo H-P, Mevarech M, Lloyd RG.** 2004. Development of additional selectable markers for the halophilic archaeon Haloferax volcanii based on the leuB and trpA genes. Applied and environmental microbiology **70:**943-953.

76. **Charlebois RL, Lam WL, Cline SW, Doolittle WF.** 1987. Characterization of pHV2 from Halobacterium volcanii and its use in demonstrating transformation of an archaebacterium. Proceedings of the National Academy of Sciences **84:**8530-8534.

77. **Norais C, Hawkins M, Hartman AL, Eisen JA, Myllykallio H, Allers T.** 2007. Genetic and physical mapping of DNA replication origins in Haloferax volcanii **3(5)**:e77

78. **Blaseio U, Pfeifer F.** 1990. Transformation of Halobacterium halobium: development of vectors and investigation of gas vesicle synthesis. Proceedings of the National Academy of Sciences

764 **87:**6772-6776.

765 79. **Chen S, Wang C, Xu J, Yang Z.** 2013. Molecular characterization of pHRDV1, a new virus-like
766 mobile genetic element closely related to pleomorphic viruses in haloarchaea. Extremophiles
767 **18:**195-206.

768 80. **Woods W, Dyallsmith M.** 1997. Construction and analysis of a recombination - deficient (radA)
769 mutant of Haloferax volcanii. Molecular Microbiology **23:**791.

770

771 **Table 1.** Host range of pYCJ-HH/pHH205.

Strain	Description	Source
Natrinema versifome JCM10478[T]	NO replicability of pYCJ-HH or pHH205	CGMCC[a]
Natrinema pellirubrum JCM10476 [T]	NO replicability of pYCJ-HH or pHH205	Henglin Cui, Universiy of Jiangsu, China
Natrinema pallidum JCM8980 [T]	Replicability of pYCJ-HH and pHH205, immunity to SNJ1 infection	Henglin Cui, Universiy of Jiangsu, China
Natrinema ejinorense JCM13890 [T]	NO replicability of pYCJ-HH or pHH205	CGMCC [a]
Natrinema altunense CGMCC 1.3731 [T]	NO replicability of pYCJ-HH or pHH205	Hua Xiang, Institute of Microbiology, Chinese Academy of Sciences, China
Halobacterium sp. AB91037	NO replicability of pYCJ-HH or pHH205	CTTCC[b]
Haloarcula hispanica 33960[T]	NO replicability of pYCJ-HH or pHH205	Kamekura. M. Noda Institute for Scientific Research, Japan
Halobiforma lacisalsi 1.3738[T]	NO replicability of pYCJ-HH or pHH205	CGMCC [a]
Haloarcula marismortui 1.1784[T]	NO replicability of pYCJ-HH or pHH205	CGMCC [a]
Halorubrum sp. T3	NO replicability of pYCJ-HH or pHH205	(79)
Halobacterium salinarum NRC-1	NO replicability of pYCJ-HH or pHH205	CCTCC[b]
Haloferax volcanii DS52	NO replicability of pYCJ-HH or pHH205	(80)

772 [a] China general microbiological culture collection center.

773 [b] China Center for Type Culture Collection.

774

775 **FIGURES LEGENDS**

776

777 **Figure 1. pHH205-transformed CJ7 produce SNJ1 virions.** A CJ7 culture transformed with either

778 water (A), as a control, or SNJ1 DNA (B), was induced by mitomycin C. After incubation for 12 h, the
779 supernatant from each culture was collected, mixed with CJ7 cells, and poured on agar plates. Plaques
780 on a lawn of CJ7 cells indicate production of infectious SNJ1 virions.
781
782 **Figure 2. Analysis of the SNJ1 minimal replicon.** Schematic representations of pYC vectors
783 containing different deletions of the full-length S (A) and SHS fragments (B) of SNJ1, together with
784 their transformation efficiencies. The relative location and direction of transcription are indicated by
785 horizontal arrows with numbers corresponding to ORFs. Genes encoding structural proteins [9] are
786 indicated by gray filled arrows. Restriction sites are marked. Solid lines indicate different fragments,
787 numbers denote start/end positions, and letters the arbitrary location used for cloning. The minimal
788 replicon of SNJ1 is marked by two vertical arrows.
789
790 **Figure 3. Transcriptional analysis and determination of essential genes within the SNJ1 minimal**
791 **replicon.** Total RNA from J7-1 cells was used to analyze the transcription (A) and co-transcription (B)
792 of the seven predicted ORFs within the SNJ1 minimal replication region by RT-PCR. Lanes: M, 2k plus
793 II marker; 1, distilled water negative control; 2, total RNA negative control; 3, J7-1 DNA positive
794 control; 4, cDNA. (C) Schematic diagram of the SHS minimal replication region with mutations in the
795 start codons of predicted ORFs. Horizontal arrows, ORFs; vertical arrows, minimal replicon. Specific
796 mutation sites are marked by numbers and solid vertical lines. Transformation efficiencies of the
797 constructs are presented to the right.
798
799 **Figure 4. Western blot analysis of RepA in J7-1 cell extracts.**
800 Samples from a J7-1 culture were collected 0, 12, 24, and 48 h after mitomycin C induction.

Immunoblots were probed using an anti-ORF11-12 antibody. Lane 1, crude extract from CJ7 culture (negative control); lanes 2−5: samples of J7-1 culture collected 0, 12, 24, and 48 h post MMC induction, respectively; lane 6, purified RepA protein (positive control). Molecular weight markers are indicated on the left. The samples' OD600 values and titers of SNJ1 virus in the corresponding supernatants are listed underneath.

Figure 5. Multiple sequence alignment of SNJ1 RepA with homologs encoded by haloarchaeal and methanoarchaeal plasmids. Sequences are indicated with GenBank identifiers followed by species names. Conserved Cys residues which could be involved in Zn binding are indicated with green starts under the alignment, whereas active site His residues and the catalytic Tyr are indicated with blue and red stars, respectively.

Figure 6. Stability and maintenance of shuttle vectors. (A) Structural stability of shuttle vectors in CJ7 cultures subjected to antibiotic selection. Pre-transformed plasmid (extracted from *E. coli* before being transformed into CJ7, lanes 1, 2, and 3) and back-transformed plasmid (extracted from *E. coli* following transformation into CJ7, lanes 4, 5, and 6) were digested with restriction enzymes to compare stability. M, λ *Hind*III marker; lanes 1 and 4, undigested plasmids; lanes 2 and 5, plasmids digested with *Nde*I and *Hind*III; lanes 3 and 6, plasmids digested with *Sca*I. (B) Maintenance of shuttle vectors in CJ7 cultures without antibiotic selection. CJ7 cultures (50 µL), were diluted in 5 mL Halo-2 medium and grown for 24 h. After each dilution (12 in total), aliquots were spread on Halo-2 medium, 50 random colonies were selected, and spotted onto 18% MGM plates containing 5 µg/mL Mev. Maintenance was measured by calculating the percentage of Mev-resistant colonies. (C) Maintenance of shuttle vectors containing mutations in the start codon of different ORFs in CJ7 cultures without antibiotic selection.

Experiments were performed as described above.

Figure 7. Relative copy numbers of pHH205 and shuttle vectors in CJ7.

Real-time PCR was performed to determine the relative copy numbers of pHH205 and constructs containing different fragments of SNJ1 (A), and constructs containing mutations in the start codon of different ORFs (B). Primers based on *repA* and *radA* were used to determine the copy number of constructs and chromosomes, respectively. Plasmid copy number was calculated based on the assumption that each chromosome had only one copy of *radA*.

Figure 8. Amylase expression in CJ7 and JCM8980 cells.

pYCJ-A-amyH was transformed into CJ7 (A) and JCM8980 (B) cells. Twelve random transformants were transferred to 5 mL 18% MGM with antibiotic, grown to log phase, and 2 μL were spotted onto 18% MGM plates supplemented with 2% (w/v) soluble starch and 5 μg/mL antibiotic. After five days, iodine solution was added to the plates, and haloes formed immediately around the selected tansformants indicating amylase expression. N1, CJ7 or JCM8980 cultures (negative control); N2, CJ7 or JCM8980 transformed with pYC-J (negative control); 1-12, transformants of CJ7 or JCM8980 harboring pYCJ-A-amyH.

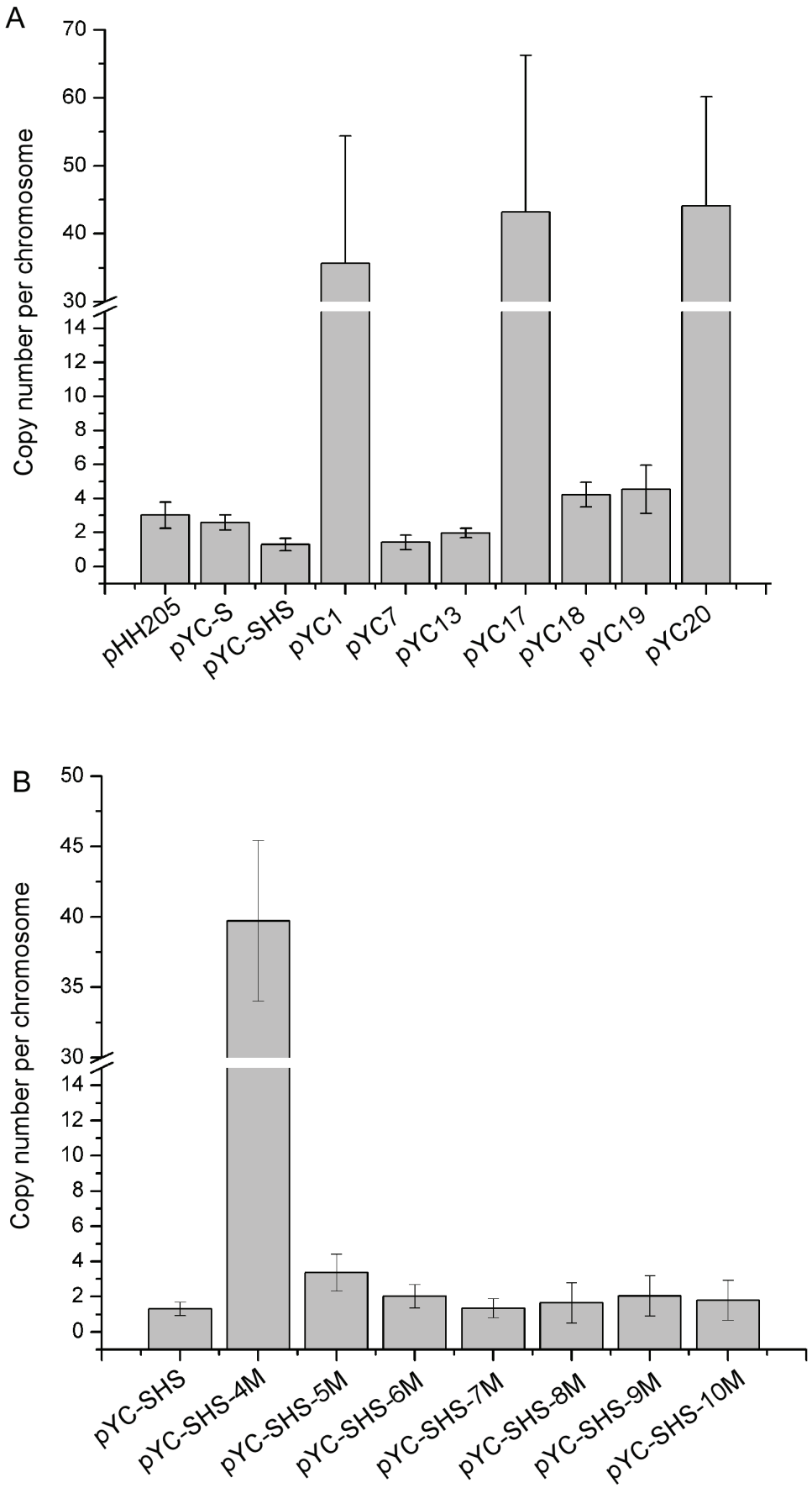
Journal of Bacteriology

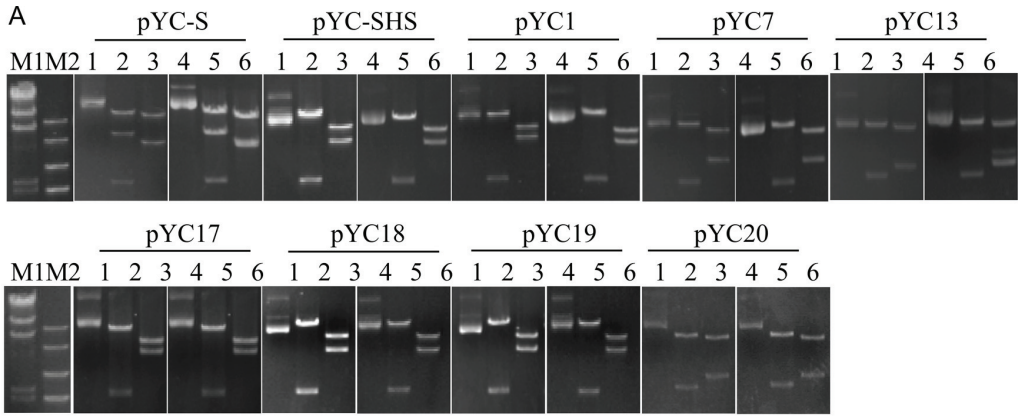
JB

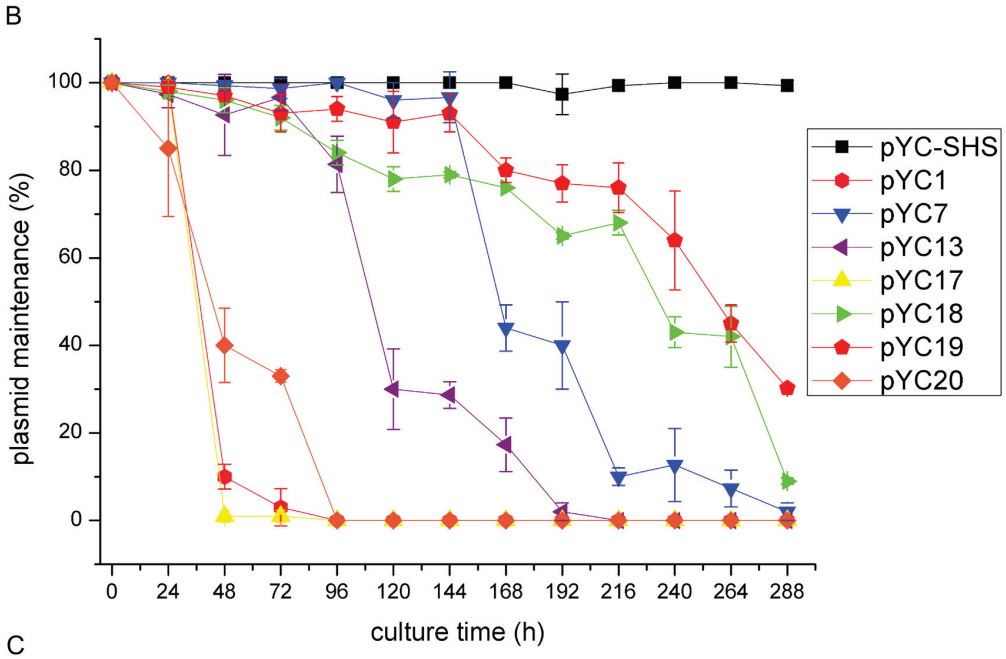
Journal of Bacteriology

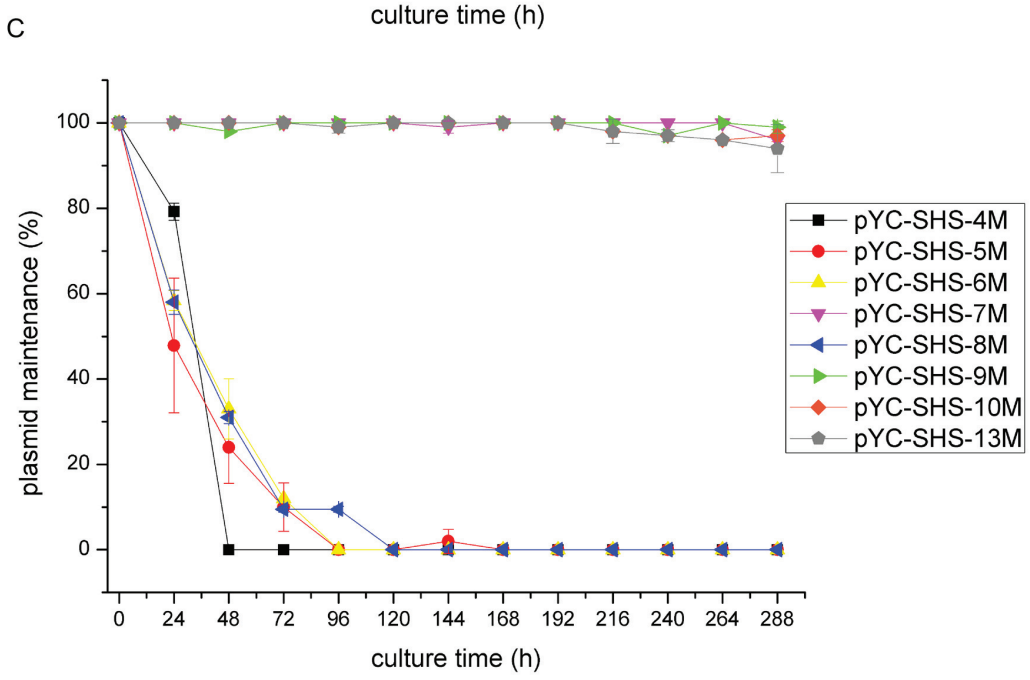
A

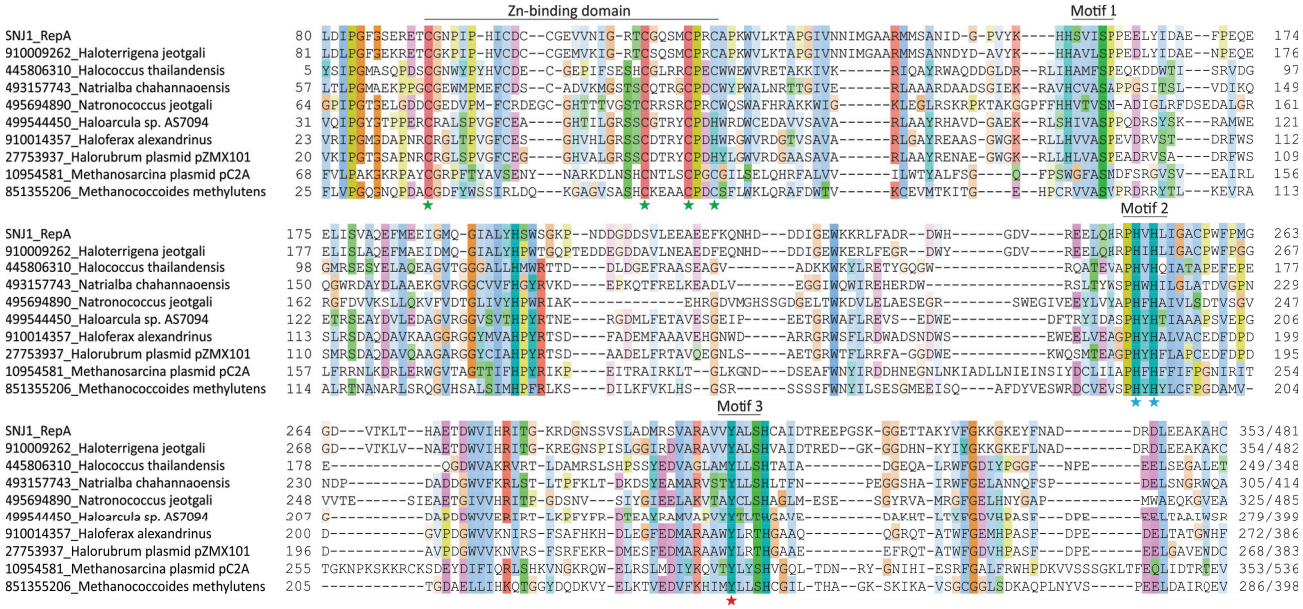
B

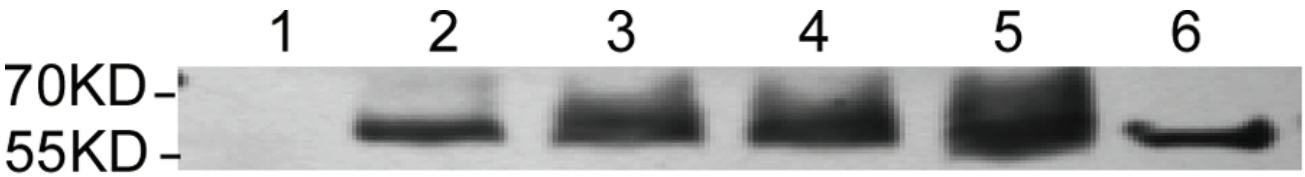






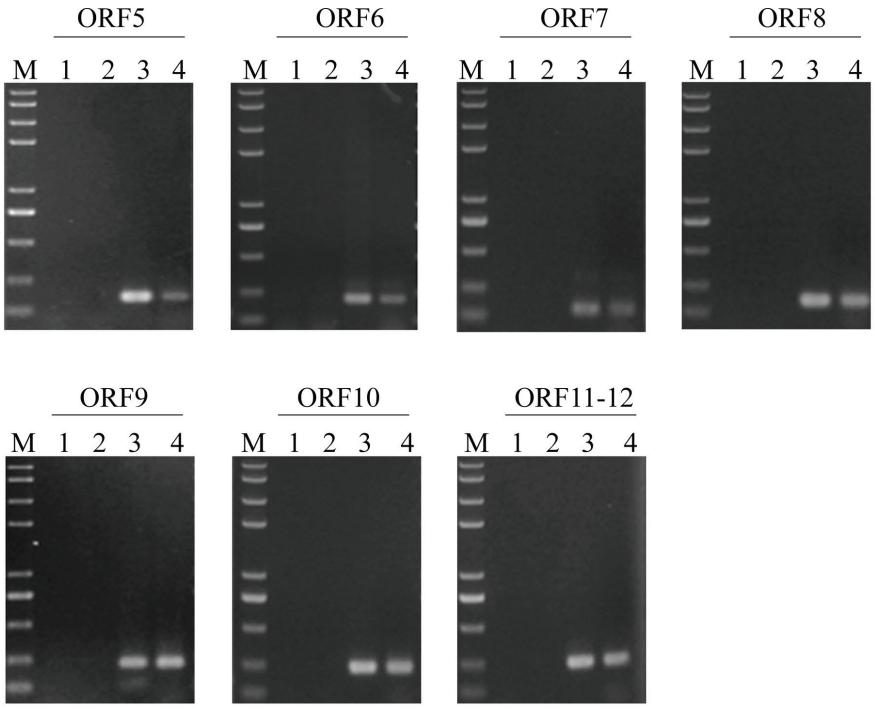




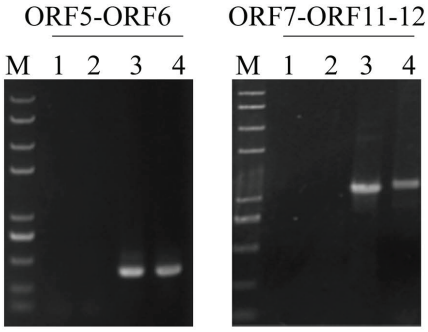


OD600	0.541	0.901	0.774	0.743
pfu/mL	1.8×10^5	1×10^8	1.6×10^9	5×10^9

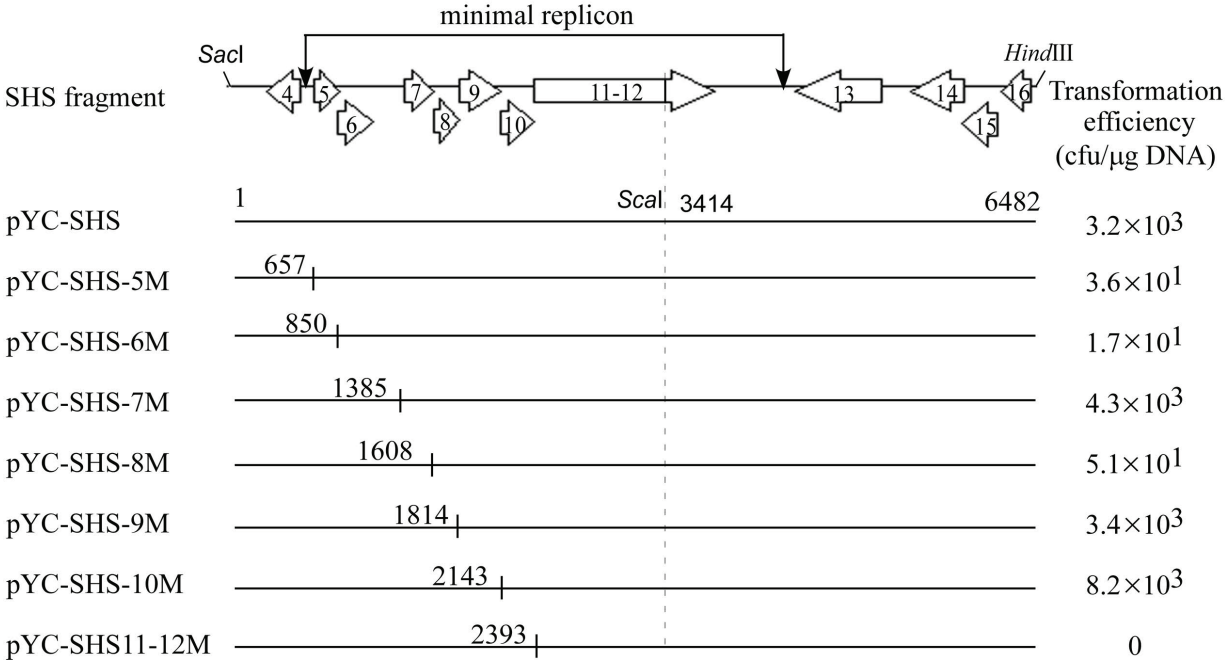
A

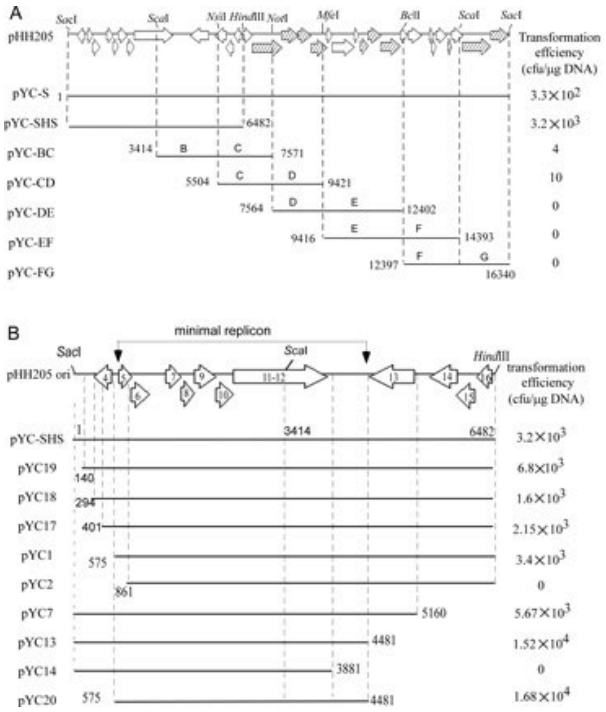


B



C





Journal of Bacteriology

JB

Journal of Bacteriology