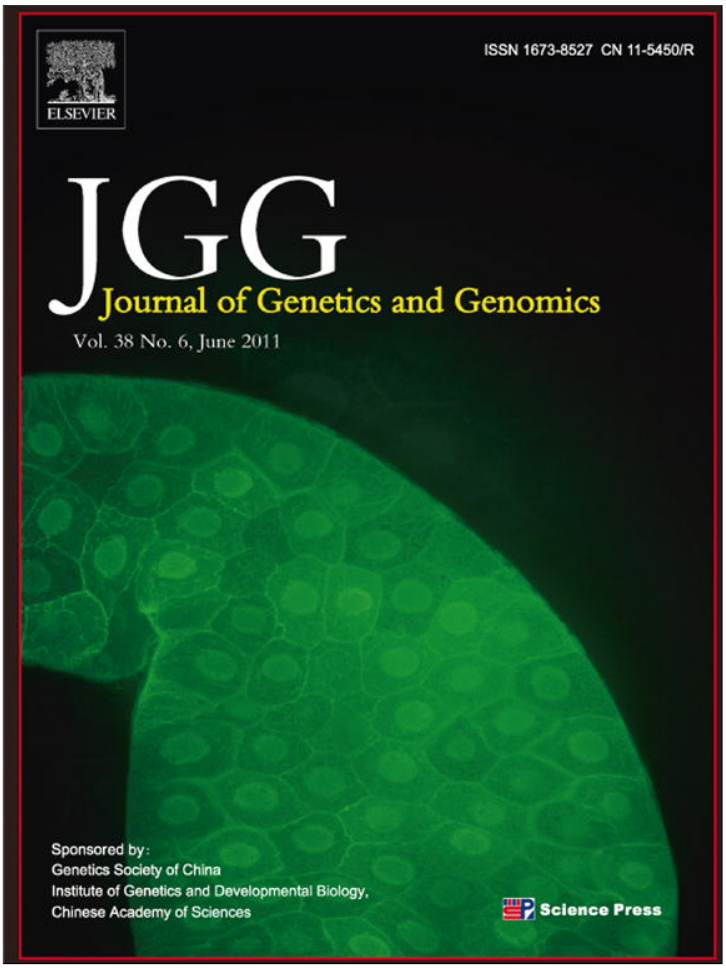

Available online at www.sciencedirect.com
ScienceDirect
Journal of Genetics and Genomics 38 (2011) 261−269
JOURNAL OF GENETICS AND GENOMICS
www.jgenetgenomics.org

Development of *pyrF*-based gene knockout systems for genome-wide manipulation of the archaea *Haloferax mediterranei* and *Haloarcula hispanica*

Hailong Liu, Jing Han, Xiaoqing Liu, Jian Zhou, Hua Xiang*

State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

Received 8 April 2011; revised 3 May 2011; accepted 3 May 2011

Abstract

The haloarchaea *Haloferax mediterranei* and *Haloarcula hispanica* are both polyhydroxyalkanoate producers in the domain Archaea, and they are becoming increasingly attractive for research and biotechnology due to their unique genetic and metabolic features. To accelerate their genome-level genetic and metabolic analyses, we have developed specific and highly efficient gene knockout systems for these two haloarchaea. These gene knockout systems consist of a suicide plasmid vector with the *pyrF* gene as the selection marker and a uracil auxotrophic haloarchaeon ($\triangle pyrF$) as the host. For in-frame deletion of a target gene, the suicide plasmid carrying the flanking region of the target gene was transferred into the corresponding $\triangle pyrF$ host. After positive selection of the single-crossover integration recombinants (pop-in) on AS-168SY medium without uracil and counterselection of the double-crossover *pyrF*-excised recombinants (pop-out) with 5-fluoroorotic acid (5-FOA), the target gene knockout mutants were confirmed by PCR and Southern blot analysis. We have demonstrated the effectiveness of these systems by knocking out the *crtB* gene which encodes a phytoene synthase in these haloarchaea. In conclusion, these well-developed knockout systems would greatly accelerate the functional genomic research of these halophilic archaea.

Keywords: Knockout system; *pyrF*; Pop-in/pop-out method; Haloarchaea

1. Introduction

Since Archaea were identified as the third domain of life by Woese in 1990 (Woese et al., 1990), their ability to thrive in harsh habitats has endeared them to microbiologists. Specifically, the mechanism for adaptation to extreme temperatures, pH and salinity, and their unique metabolic pathways, such as the methylaspartate cycle in haloarchaea or methanogenesis in methanogens, have gained extensive attention (Kennedy et al., 2001; Vieille and Zeikus, 2001; Angelov and Liebl, 2006; Trivedi et al., 2006; Falb et al., 2008; Liu and Whitman, 2008; Cardenas et al., 2010; Khomyakova et al., 2011). Moreover, the Bacteria and Eukaryota chimera features of Archaea are also very important in understanding the origin and evolution of cellular processes (Bell and Jackson, 1998).

In the domain Archaea, haloarchaea are becoming increasingly attractive for research and biotechnology due to their versatile metabolic abilities, ease of culture, and genetic tractability. So far, 14 haloarchaeal genomes have been sequenced: *Halobacterium* sp. NRC-1, *Haloarcula marismortui* ATCC 43049, *Natronomonas pharaonis* DSM 2160, *Haloquadratum walsbyi* DSM 16790, *Halorhabdus utahensis* DSM 12940, *Halorubrum lacusprofundi* ATCC 49239, *Haloterrigena turkmenica* DSM 5511, *Halogeometricum borinquense* DSM 11551, *Halomicrobium mukohataei* DSM 12286, *Natrialba magadii* ATCC 43099, *Haloferax volcanii* DS2, *Halalkalicoccus jeotgali* B3^T, *Haloferax mediterranei* CGMCC 1.2087 (our laboratory, unpublished data) and *Haloarcula hispanica* CGMCC 1.2049 (our laboratory, unpublished data). With the continuous increase in the number of genome sequences

* Corresponding author. Tel/fax: +86 10 6480 7472.
E-mail address: xiangh@sun.im.ac.cn (H. Xiang).

doi:10.1016/j.jgg.2011.05.003

available for haloarchaea, it is urgent to utilize their powerful genetics to support research on their biochemistry and functional genomics. Though the genome sequences of haloarchaea are increasingly available, the major model haloarchaea include only two species, *Halobacterium salinarum* and *Hfx. volcanii* (Leigh et al., 2011), chiefly because most haloarchaea harbor restriction/modification systems that significantly reduce the efficiency of transformation. The extremely halophilic archaea *Hfx. mediterranei* and *Har. hispanica* are obligate halophiles that were first isolated from solar salterns in Spain (Rodriguez-Valera et al., 1980; Juez et al., 1986). Fortunately, they both possess nearly all of the advantages of the two model organisms, such as stable genome organization, ease of culture, and a sequenced genome. Particularly, these two strains can also accumulate large amounts of polyhydroxyalkanoate (PHA) when their carbon source is available in excess (Han et al., 2007; Lu et al., 2008).

Poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) (PHBV) is a popular PHA produced by microbes with desired biodegradable, biocompatible, and thermoplastic features. It can be employed to substitute for conventional petrochemical plastics. However, using bacteria to produce PHBV is still costly because several related carbon sources, such as propionate, which is expensive and highly toxic, are usually required as precursors (Steinbuchel and Lutke-Eversloh, 2003). In contrast, haloarchaea can produce PHBV from inexpensive renewable sources like glucose or starch. Although many haloarchaea, belonging to at least 12 genera, are able to synthesize short-chain-length (SCL)-PHAs (Han et al., 2010), the well-studied model haloarchaea, *Hbt. salinarum* and *Hfx. volcanii*, cannot. Especially, apart from the PHA synthase genes (*phaEC*) (Han et al., 2007; Lu et al., 2008) and one PHA precursor supplying gene (*phaB*) (Han et al., 2009) being identified in *Hfx. mediterranei* and *Har. hispanica*, little is known about the PHA metabolic pathway in halophilic archaea. Thus, genome-wide analyses of the *Hfx. mediterranei* and *Har. hispanica* would provide considerable insights into the PHA metabolism and its regulation in the domain Archaea. If a highly efficient method for mutant construction is developed, a genome-wide analysis of gene function is possible (Hammelmann and Soppa, 2008). It would not only accelerate the genome-wide identification of the genes responsible for PHA biosynthesis, but may also be used to genetically engineer the haloarchaea to be superior PHA producers. Therefore, highly efficient gene knockout systems are becoming important and urgent for more haloarchaeal species, such as *Hfx. mediterranei* and *Har. hispanica*.

In this study, we report the development of the knockout systems for target gene in-frame deletion in *Hfx. mediterranei* and *Har. hispanica*, respectively. These systems are based on selection against the orotidine-5′-phosphate decarboxylase gene (*pyrF*) with or without a toxic uracil analog 5-fluoroorotic acid (5-FOA), a method first described for *Saccharomyces cerevisiae* (Boeke et al., 1984). Although similar systems have been established in two model haloarchaea, *Hbt. salinarum* and *Hfx. volcanii* (Peck et al., 2000; Bitan-Banin et al., 2003), this is the first report for such a system developed in haloarchaeal PHA producers. We found that the *pyrF* gene can conveniently serve as both a selectable genetic marker and a counterselectable genetic marker for efficient generation of gene knockouts in both *Hfx. mediterranei* and *Har. hispanica*. We have also demonstrated that these knockout systems are highly effective *via* the deletion of the *crtB* gene, which encodes phytoene synthase, in their respective genomes.

2. Materials and methods

2.1. Strains, plasmids, and primers

The strains and plasmids used in this study are listed in Table 1, and the oligonucleotides are listed in Table 2.

Table 1
Strains and plasmids used in this study.

Strain or plasmid	Description	Source or reference
Strains		
Escherichia coli JM109	*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi*	(Sambrook et al., 1989)
Haloferax mediterranei	Wild-type strain CGMCC 1.2087 = ATCC 33500 (Hm)	CGMCC
DF50	*Hfx. mediterranei* △*pyrF*	This study
DF50in	*Hfx. mediterranei* △*pyrF* containing pHFX△*crtB* (pop-in)	This study
DF50△*crtB*	*Hfx. mediterranei* △*pyrF* △*crtB* (pop-out)	This study
Haloarcula hispanica	Wild-type strain CGMCC 1.2049 = ATCC 33960 (Hh)	CGMCC
DF60	*Har. hispanica* △*pyrF*	This study
DF60in	*Har. hispanica* △*pyrF* containing pHAR△*crtB* (pop-in)	This study
DF60△*crtB*	*Har. hispanica* △*pyrF* △*crtB* (pop-out)	This study
Plasmids		
pUBP	6.6 kb; derivative of pUBP2 obtained by removing the pHH9 *ori* region	(Han et al., 2007)
pUBPHm△*pyrF*	7.7 kb; integration vector for deletion of $pyrF_{Hm}$	This study
pUBPHh△*pyrF*	7.7 kb; integration vector for deletion of $pyrF_{Hh}$	This study
pHFX	4.0 kb; integration vector containing $pyrF_{Hm}$ and its native promoter	This study
pHAR	4.0 kb; integration vector containing $pyrF_{Hh}$ and its native promoter	This study
pHFX△*crtB*	5.0 kb; pHFX containing DF50△*crtB* for deletion of DF50 *crtB*	This study
pHAR△*crtB*	5.0 kb; pHAR containing DF60△*crtB* for deletion of DF60 *crtB*	This study

Table 2
Oligonucleotides used in this study.

Primer	Sequence (5′→3′)[a]
HmpyrFRTF	TTCGTCCTCTGTCGCACCTCT
HmpyrFRTR	CTGGTTCAGCCGCTCTTTG
HmpyrFF	GGCGAATTCGTTTTCTTGTGTGCGGTT
HmpyrFR	GTTGGTACCTTACCGGTACTGGTTCAG
HmpyrFF1	GCTGGTACCCGTCAACATGGCGTACATCC
HmpyrFR1	GCGGGATCCAGTGCGACGACCGTATGTAAG
HmpyrFF2	GCGGGATCCATCACAGACTGGCACTGGACT
HmpyrFR2	ATACTGCAGGCGTCTGGATGCGTCTCCT
HmcrtBF1	ATACTGCAGGTGCGGCGTGTGGGTACTTC
HmcrtBR1	AATGGATCCCCGCGTACCGTCCCTGTCAC
HmcrtBF2	GGTGGATCCGTTTGGTCGTTGATTTTTGA
HmcrtBR2	ACAGGTACCGGCTTCCCGTGTTTTTCATC
HhpyrFRTF	AGTTCAACCGCCGCATCAT
HhpyrFRTR	GGAGAAACGGCTCCAACGAG
HhpyrFF	TATGAATTCGAGCGGGCTTCTACCTGC
HhpyrFR	GGCGGTACCTTAGCGGAACTGATTCAG
HhpyrFF1	ATACTGCAGCGACTCGGCTCGGCAATA
HhpyrFR1	AATGGATCCCGGCTGGCAGCGATACAA
HhpyrFF2	GCGGGATCCAAACAGCTCAAACAGCGACTG
HhpyrFR2	GCTGGTACCCCAGCATTCCGAGTATCCA
HhcrtBF1	ATTGGTACCAAGCGTGGTCGAGCAGGTCC
HhcrtBR1	TTAGGATCCCGTCGCGCACTTGTCGGCAG
HhcrtBF2	AATGGATCCCAGCCACGCACCGGGTATCG
HhcrtBR2	ATACTGCAGCGCCGACGGTCCGCCACTCC

[a] Restriction endonuclease sites are underlined.

Escherichia coli JM109 was grown in Luria-Bertani (LB) medium (Sambrook et al., 1989), and *Hfx. mediterranei* and *Har. hispanica* strains were generally cultivated at 37 °C in a nutrient-rich medium (AS-168) (Han et al., 2007). When needed, ampicillin was added to a final concentration of 100 μg/mL for *E. coli*, mevinolin to 3 μg/mL for *Hfx. mediterranei* transformants, and 5 μg/mL for *Har. hispanica* (Lu et al., 2008). AS-168SY medium was similar to AS-168, except that yeast extract was omitted. When required, AS-168SY medium was supplemented with uracil (Sangon, China) to a concentration of 50 μg/mL, and 5-FOA (Sangon, China) to 250 μg/mL and 150 μg/mL, for *Hfx. mediterranei* and *Har. hispanica*, respectively. The haloarchaeal plasmids were first constructed in *E. coli* JM109 and then transformed into *Hfx. mediterranei* or *Har. hispanica*. Transformation of the haloarchaeal strains was carried out as described previously (Cline et al., 1989).

2.2. Reverse transcription polymerase chain reaction (RT-PCR)

Total RNA of *Hfx. mediterranei* and *Har. hispanica* was extracted using TRIzol Reagent (Invitrogen, USA) from the cells grown in AS-168 medium at stationary phase. Primer pairs (HmpyrFRTF/HmpyrFRTR and HhpyrFRTF/HhpyrFRTR) were designed specifically for the $pyrF_{Hm}$ and $pyrF_{Hh}$ genes, respectively (Table 2). RNA samples were treated with RNase-free RQ1 DNase (Promega) (Han et al., 2009) to eliminate any DNA contamination, and was confirmed to be DNA-free using a control PCR reaction with treated RNA as the template. The DNA-free RNA samples were then used as templates for RT-PCR with a OneStep RT-PCR kit (Qiagen, USA) according to the manufacturer's instructions.

2.3. Cloning and analysis of pyrF genes from Hfx. mediterranei and Har. hispanica

To clone the full-length *pyrF* genes and the flanking sequences, genomic DNA from *Hfx. mediterranei* and *Har. hispanica* was prepared as described previously (DasSarma and Fleischmann, 1995). The 897 bp (*Hfx. mediterranei*) or 954 bp (*Har. hispanica*) DNA fragment of the *pyrF* gene was amplified by PCR using the primers HmpyrFF/HmpyrFR or HhpyrFF/HhpyrFR. DNA fragments were cloned into the pGEM-T vector (Promega, USA) and confirmed by sequencing. DNA or protein database searches were performed at the National Center for Biotechnology Information Blast website (http://www.ncbi.nlm.nih.gov/BLAST/). To predict the encoded amino acid sequences, the DNASTAR software (Burland, 2000) was used. Multiple sequence alignments were constructed with the GeneDoc program (http://www.nrbsc.org/gfx/genedoc).

2.4. Construction of plasmids pUBPHm△pyrF and pUBPHh△pyrF for knockout of the pyrF gene in Hfx. mediterranei or Har. hispanica

A 602-bp fragment containing the upstream flanking sequence of the *Hfx. mediterranei pyrF* gene was amplified by PCR with the primers HmpyrFF1 and HmpyrFR1 (Table 2). A 567-bp DNA fragment containing the downstream flanking region of *Hfx. mediterranei pyrF* was amplified by PCR with the primers HmpyrFF2 and HmpyrFR2 (Table 2). The two PCR products were sequenced and cloned into pUBP to generate pUBPHm△*pyrF* (Table 1). Similarly, the 585-bp and 548-bp flanking sequences of the *Har. hispanica pyrF* gene were amplified by PCR with the primers HhpyrFF1/HhpyrFR1 or HhpyrFF2/HhpyrFR2, respectively (Table 2) and cloned into pUBP to generate the plasmid pUBPHh△*pyrF* (Table 1). The resulting plasmids were then transformed into *Hfx. mediterranei* or *Har. hispanica*, respectively, to knock out the *pyrF* gene by double-crossover homologous recombination. The *pyrF* deletion mutants, named *Hfx. mediterranei* DF50 or *Har. hispanica* DF60, were screened by PCR as previously described (Han et al., 2007).

2.5. Construction of pyrF-based suicide plasmid vector pHFX and pHAR

The cloned and sequenced *pyrF* genes of *Hfx. mediterranei* and *Har. hispanica*, with their native promoters (99 bp upstream of the start codon ATG for *Hfx. mediterranei* and 120 bp for *Har. hispanica*), were digested from pGEM-T by *Eco*R I and *Kpn* I and then ligated with an *Eco*R I-*Kpn* I fragment of pUBP containing the *E. coli* replication origin and ampicillin resistance gene. The generated plasmids were

named pHFX for *Hfx. mediterranei* and pHAR for *Har. hispanica*. The plasmids pHFX and pHAR will serve as the *pyrF*-based knockout vectors in the newly developed knockout systems.

2.6. Construction of plasmids pHFX △ *crtB* and pHAR △ *crtB* for knockout of the *crtB* gene in strain DF50 or DF60

For construction of the plasmid pHFX △ *crtB* (Table 1), the 504-bp and 528-bp flanking regions of *Hfx. mediterranei crtB* gene were amplified by PCR with the primer pairs HmcrtBF1/HmcrtBR1 or HmcrtBF2/HmcrtBR2 (Table 2), respectively, and cloned into pHFX (Table 1). Similarly, the 517-bp and 562-bp flanking sequences of the *Har. hispanica crtB* gene were amplified by PCR using the primers HhcrtBF1/HhcrtBR1 or HhcrtBF2/HhcrtBR2 (Table 2), and cloned into pHAR to generate the plasmid pHAR △ *crtB* (Table 1). The resulting plasmids were then transformed into DF50 or DF60, respectively, to knock out the *crtB* gene by homologous recombination. The *crtB* deletion mutants, named DF50 △ *crtB* or DF60 △ *crtB*, were obtained through positive selection (pop-in) and counterselection (pop-out) of *pyrF* gene on the plasmid pHFX △ *crtB* or pHAR △ *crtB*, respectively.

2.7. Southern blot and PCR analysis of recombinant strains

Southern blot analysis was performed as described previously (Sambrook et al., 1989). Briefly, the genomic DNA was prepared as described above. Hybridization probe labeling and color detection with NBT/BCIP were performed using a DIG-High Prime DNA Labeling and Detection Starter Kit I (Roche, USA). For PCR analysis, genomic DNA template was obtained by transferring a colony into a mixture of 500 μL of sterile water and 100 μL phenol−chloroform, vortexing the sample for 10 s and centrifuging at 12,000 × *g* for 5 min. The extracted DNA samples were then subjected to PCR analysis.

3. Results and discussion

3.1. Identification and analysis of the *pyrF* genes in *Hfx. mediterranei* and *Har. hispanica*

The complete gene sets for *de novo* synthesis of pyrimidines from carbamoyl phosphate and ribose 5-phosphate are usually present in haloarchaeal genomes (Falb et al., 2008). Moreover, *pyrF* (equivalent to *ura3*), encoding orotidine-5′-phosphate decarboxylase (Peck et al., 2000), and *pyrE2*, encoding orotate phosphoribosyltransferase (Bitan-Banin et al., 2003), have been developed into selectable markers in haloarchaeal genetic manipulation systems. Though the wild-type haloarchaeal strains are usually sensitive to 5-FOA, the mutants, which were devoid of *ura3* or *pyrF*, were resistant to 5-FOA (Peck et al., 2000; Bitan-Banin et al., 2003). Notably, genome-wide analysis has revealed only one copy of putative orotidine-5′-phosphate decarboxylase gene (*pyrF*) in *Hfx. mediterranei* and *Har. hispanica*, respectively. The deduced *Hfx. mediterranei* and *Har. hispanica* PyrF proteins consist of 265 and 277 amino acids, respectively, with high homology to the previously characterized Ura3 of *Hbt. salinarum* (Peck et al., 2000). Alignment of the two sequences with *Hbt. salinarum* Ura3 revealed that the homology extends over the full length of the protein (Fig. 1A), indicating that the *pyrF* genes encode the putative orotidine-5′-phosphate decarboxylases in these two strains. To determine whether the *pyrF* gene was dynamically transcribed, RT-PCR was performed. As shown in Fig. 1B, a 279-bp PCR fragment for $pyrF_{Hh}$ mRNA and a 381-bp PCR fragment for $pyrF_{Hm}$ mRNA were amplified by the primer pairs HhpyrFRTF/HhpyrFRTR and HmpyrFRTF/HmpyrFRTR, respectively. The RT-PCR results suggested that the *pyrF* genes were actively expressed and would function well in both *Hfx. mediterranei* and *Har. hispanica*, and it

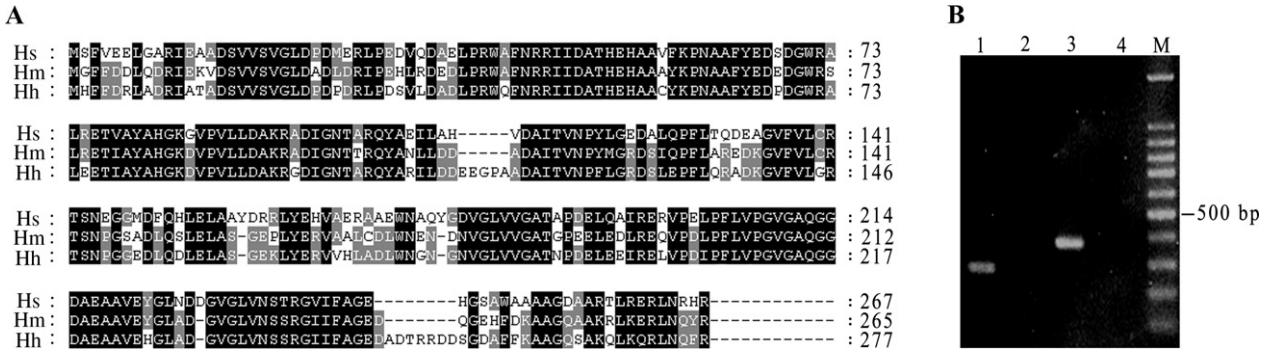

Fig. 1. Characterization of *pyrF* genes in *Hfx. mediterranei* and *Har. hispanica*. **A:** multiple alignments of *Hbt. salinarum* Ura3 (Hs), *Hfx. mediterranei* PyrF (Hm), and *Har. hispanica* PyrF (Hh) proteins. Amino acids are given in standard one-letter abbreviations, and the numbers indicate the positions of the amino acids within the respective proteins. The conserved residues are darkly shaded, and the residues identical in two of the three are lightly shaded. GenBank accession numbers for *Hbt. salinarum* Ura3 is AF187997. **B:** RT-PCR analysis of *pyrF* genes from strains cultured in AS-168 medium. Lane 1, *Har. hispanica* $pyrF_{Hh}$ mRNA (279 bp) with cDNA as PCR template; lane 2, negative control with *Har. hispanica* treated RNA as PCR template; lane 3, *Hfx. mediterranei* $pyrF_{Hm}$ mRNA (381 bp) with cDNA as PCR template; lane 4, negative control with *Hfx. mediterranei* treated RNA as PCR template; lane M, 100 bp DNA ladder.

would be possible to develop them into selection markers in gene knockout systems.

3.2. Development of the pyrF-based gene knockout systems for Hfx. mediterranei and Har. hispanica

The *pyrF*-based gene knockout system should include a uracil auxotrophic host ($\triangle pyrF$) and a non-replicative plasmid vector carrying the *pyrF* gene, which can act as both positive selection and counterselection marker for generation of knockouts. To obtain the suitable *Hfx. mediterranei* or *Har. hispanica* host strains, DF50 and DF60, respectively, the *pyrF* gene was first knocked out with a traditional vector, pUBP, without the haloarchaeal replicon, as previously described (Lu et al., 2008). As shown in Fig. 2A and C, $pyrF_{Hm}$ and $pyrF_{Hh}$ were the only genes in their locations, so the possible polar effects to other genes could be ignored when deleting them. A 1169-bp fragment, with the 816-bp $pyrF_{Hm}$ coding region deleted, was cloned into pUBP to generate pUBPHm$\triangle pyrF$ (Fig. 2A), and a 1133-bp fragment, with the 809-bp $pyrF_{Hh}$ coding region deleted, was cloned into pUBP to generate pUBPHh$\triangle pyrF$ (Fig. 2C). These plasmids, pUBPHm$\triangle pyrF$ and pUBPHh$\triangle pyrF$, which contain the mevinolin resistance genes and the *pyrF* flanking sequences, were introduced into the wild-type strains *Hfx. mediterranei* and *Har. hispanica*, respectively. The plasmid-harboring recombinants of single-crossover were selected on mevinolin plates, and single-crossover recombinants were then resuspended and subcultured at least three times (>60 generations) in medium with uracil, but not mevinolin, to allow double-crossover recombination to occur. The cell suspensions were then grown on plates without mevinolin; subsequently, the resulting colonies were replicated on mevinolin plates and mevinolin-free plates. Colonies that could only grow on mevinolin-free plates and not on mevinolin plates may have lost the mevinolin resistance gene as a result of double-crossover. Primer pairs HmpyrFF1/HmpyrFR2 or HhpyrFF1/HhpyrFR2 were used for identification of the single-crossover and double-crossover mutants by PCR analysis. The identified *pyrF*-deficient mutant was further confirmed by Southern blot analysis using either the 567-bp or 585-bp flanking sequence mentioned above as a probe (Fig. 2A and C). Southern blot analysis of strains DF50 and DF60 confirmed the expected decrease in the size of the *pyrF*-coding sequences relative to the wild-type strain *Hfx. mediterranei* or *Har. hispanica* (Fig. 2B and D). As expected, both *Hfx. mediterranei* DF50 and *Har. hispanica* DF60 are uracil auxotrophic, which grew almost as well as the wild-type strains on AS-168SY medium containing uracil, and thus would be sufficient for routine application as the desired *pyrF*-deficient hosts.

To obtain an appropriate concentration of 5-FOA for counterselection, a series of concentration gradients was investigated for the wild-type strains in AS-168. We determined that 250 μg/mL 5-FOA was sufficient for inhibiting the growth of *Hfx. mediterranei*, whereas 150 μg/mL 5-FOA was sufficient for *Har. hispanica*. Consistently, when plated onto AS-168SY medium containing 50 μg/mL uracil plus either 250 μg/mL or 150 μg/mL of 5-FOA for DF50 or DF60, it was obvious that the DF50 and DF60 strains were able to grow with few inhibitory effects, whereas the wild-type *Hfx. mediterranei* and *Har. hispanica* were unable to grow. We concluded that the uracil auxotrophic DF50 and DF60 were resistant to 5-FOA, whereas the uracil prototrophic *Hfx. mediterranei* and *Har. hispanica* were 5-FOA-sensitive. Thus, the AS-168SY medium without uracil could be used for positive selection of the single-crossover recombinants

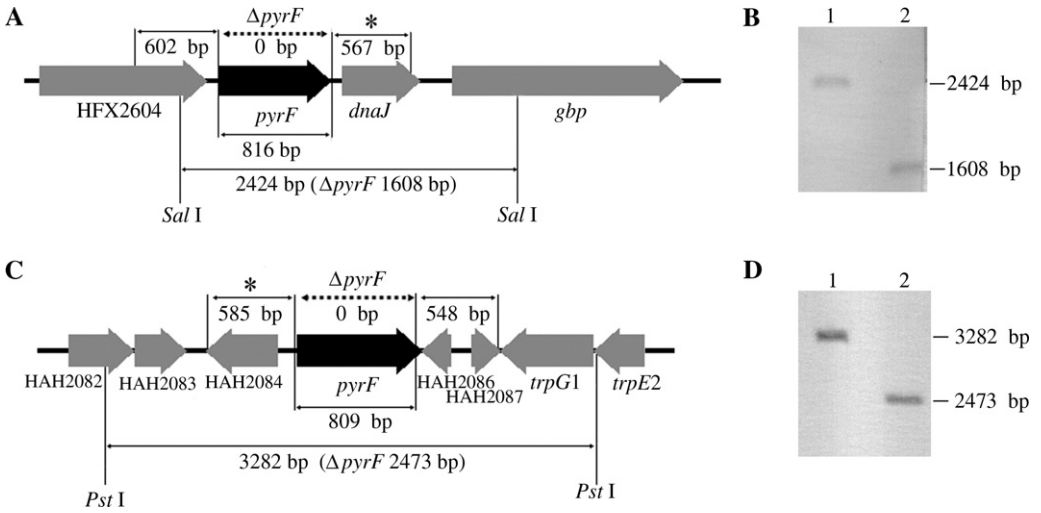

Fig. 2. Schematic diagram of *pyrF* regions and Southern blot analysis of mutant strains. **A:** *Hfx. mediterranei pyrF* region. The upstream 602 bp and downstream 567 bp of *pyrF* to be cloned into pUBP and the deleted 816 bp region are indicated. The extracted genomic DNA was digested with *Sal* I. Asterisk indicates the 567-bp hybridization probe. The size of hybridization fragments is also indicated: 2424 bp for *Hfx. mediterranei* and 1608 bp for DF50. **B:** Southern blot analysis of DF50 mutant. Lane 1, *Hfx. mediterranei*, 2424 bp; lane 2, DF50, 1608 bp. **C:** *Har. hispanica pyrF* region. The upstream 585 bp and downstream 548 bp of *pyrF* to be cloned into pUBP and the deleted 809-bp region are indicated. The extracted genomic DNA was digested with *Pst* I. Asterisk indicates the 585-bp hybridization probe. The size of hybridization fragments is also indicated: 3282 bp for *Har. hispanica* and 2473 bp for DF60. **D:** Southern blot analysis of DF60 mutant. Lane 1, *Har. hispanica*, 3282 bp; lane 2, DF60, 2473 bp.

(*pyrF*-integrative plasmid pop-in), and the medium containing uracil and 5-FOA could be used for counterselection of the double-crossover recombinants (*pyrF*-integrative plasmid pop-out), with the strains DF50 and DF60 as the suitable hosts.

In correspondence to the above *pyrF*-deleted hosts, the integrative plasmid vectors pHFX and pHAR (Table 1) carrying the *pyrF* marker genes were constructed as described in the Materials and methods (Fig. 3). These *pyrF*-carrying plasmid vectors and corresponding *pyrF*-deleted strains, pHFX/DF50 and pHAR/DF60, constituted the new gene knockout systems for *Hfx. mediterranei* and *Har. hispanica*, respectively. Lacking the haloarchaeal replication origins, the pHFX and pHAR vectors could not replicate in *Hfx. mediterranei* and *Har. hispanica,* so they could only integrate into the genome by homologous recombination with the flanking regions of the target gene cloned into these vectors. It is noteworthy that, both pHFX and pHAR could replicate in *E. coli*, which facilitated the gene cloning and manipulation. As mentioned above, after integration into DF50 or DF60 at the target region, the *pyrF* genes in pHFX and pHAR could then serve as positive selection markers, conferring uracil prototroph phenotypes to DF50 or DF60, and could also serve as counterselection markers with excision of the integrated plasmids under the pressure of 5-FOA.

3.3. Deletion of crtB genes using the pyrF-based gene knockout systems

The synthetic pathway leading to carotenogenesis begins with the formation of a phytoene from geranylgeranyl pyrophosphate (Sieiro et al., 2003). This step is well conserved in all carotenogenic organisms, including haloarchaea, and is catalyzed by a phytoene synthase, encoded by the *crtB* gene. Knockout of *crtB* will disrupt the synthesis of carotenoids, such as beta-carotene and bacterioruberin; thus the strains will lose pigment-producing ability, exhibiting a white phenotype. To determine the efficiency of our newly developed *pyrF*-based gene knockout systems, we performed knockouts of the *crtB* genes with the procedure as described in Fig. 4.

A 1032-bp fragment, deleting the 1029-bp $crtB_{Hm}$ coding region, was cloned into pHFX to generate pHFX△*crtB* (Fig. 5A), and a 1079-bp fragment, deleting the 1006-bp $crtB_{Hh}$ coding region, was cloned into pHAR to generate pHAR△*crtB* (Fig. 5C). The pHFX△*crtB* and pHAR△*crtB* were transformed into DF50 and DF60, respectively, and the transformants were plated onto AS-168SY medium to positively select the single-crossover recombinants (pop-in). One clone for each recombinant, in which pHFX△*crtB* or pHAR△*crtB* had integrated into the *crtB* flanking region, was designated DF50in or DF60in. Excision of the integrated plasmids in DF50in or DF60in was performed by propagating for approximately 20−30 generations in AS-168SY medium containing uracil, followed by counterselection for the *pyrF*-excised recombinants with 5-FOA on AS-168SY medium supplemented with uracil. Excision of pHFX△*crtB* or pHAR△*crtB* by homologous recombination events may result in either reconstitution of the wild-type allele or deletion of *crtB* genes (pop-out). The strains with *crtB* gene deleted were designated as DF50△*crtB* or DF60△*crtB*. Primer pairs HmcrtBF1/HmcrtBR2 or HhcrtBF1/HhcrtBR2 were used for identification of the single-crossover and double-crossover mutants by PCR analysis (data not shown). The PCR-identified mutant was further confirmed by Southern blot analysis using the 528-bp or 517-bp flanking sequence (Fig. 5A and C) mentioned above as the probe. Southern blot analysis of DF50△*crtB* and DF60△*crtB* further confirmed the deletion of the *crtB* coding region from the host strains DF50 and DF60, respectively (Fig. 5B and D). As counterselection was applied in this *pyrF*-based system, the time to get the double-crossover recombinants from single-crossover recombinants was largely shortened, generally from more than two weeks to only about one week.

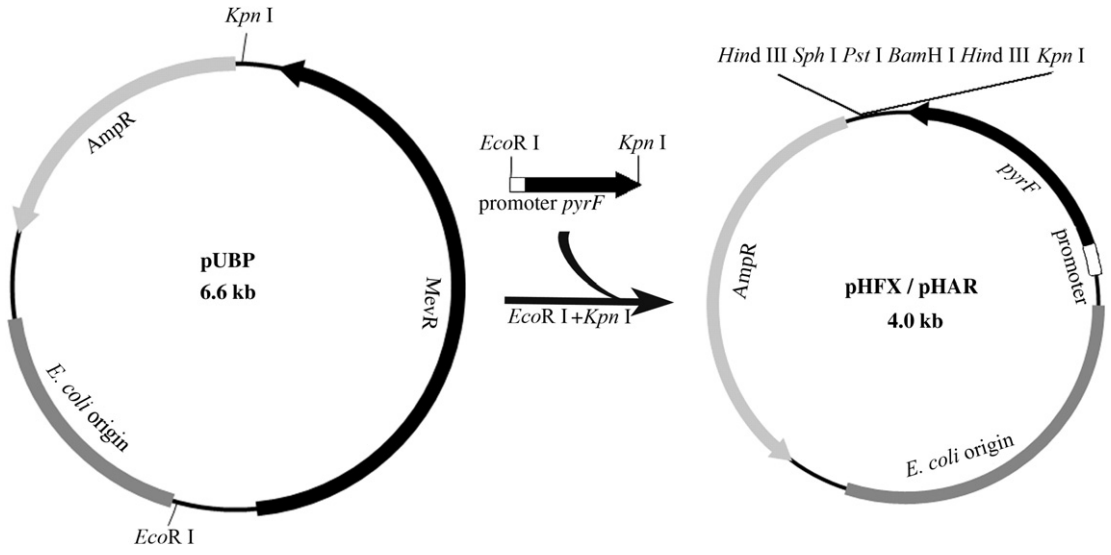

Fig. 3. Construction of pHFX and pHAR, the *pyrF*-based suicide vectors for *Hfx. mediterranei* and *Har. hispanica*, respectively. Only relevant restriction sites are shown.

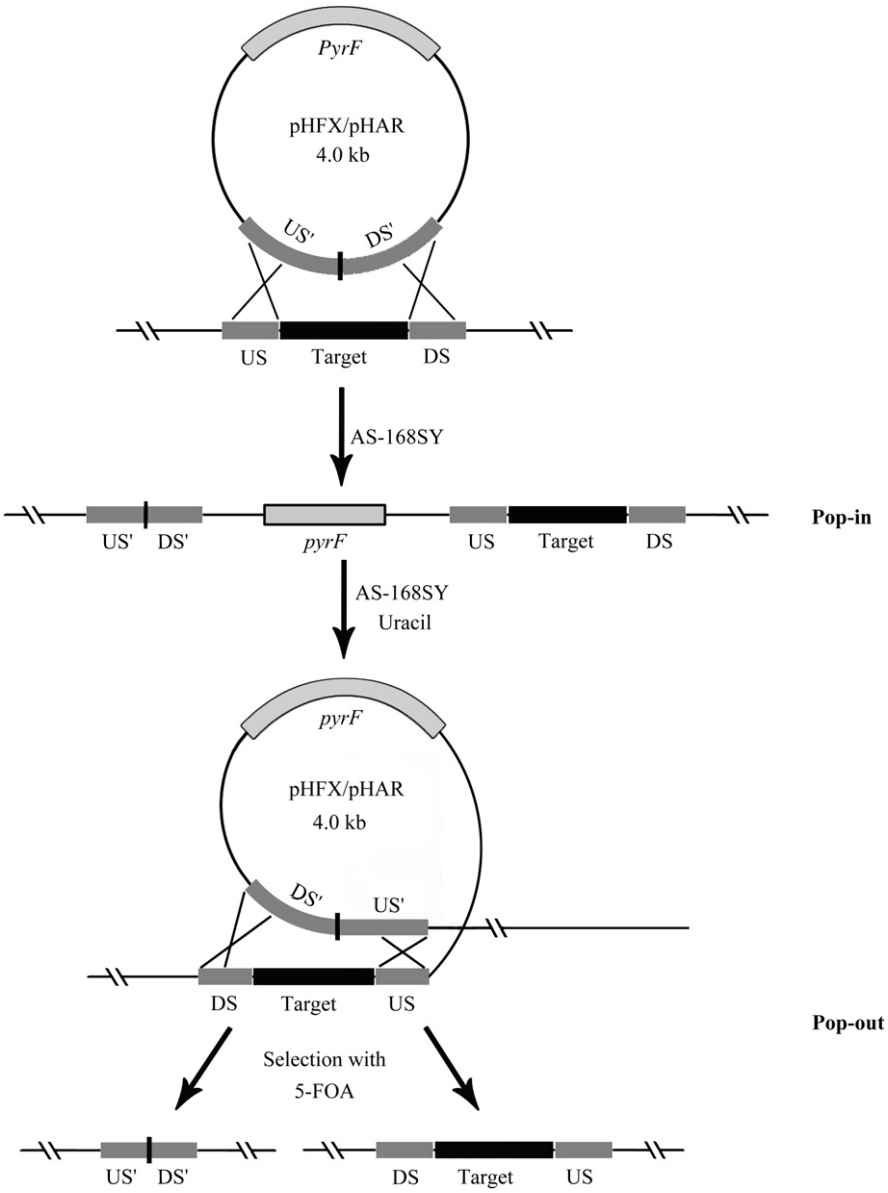

Fig. 4. Schematic diagram of the *pyrF*-based knockout system using the pop-in/pop-out method. US/US', upstream of the target gene to be deleted; DS/DS', downstream of the target gene to be deleted. It should be noted that US/US' and DS/DS' did not represent the real sequence transformed during homologous recombination.

Notably, the DF50 △ *crtB* and DF60 △ *crtB* mutants showed white phenotypes when compared to the pink phenotypes of DF50 and DF60, while they were cultured in AS-168SY medium containing uracil (Fig. 6). This result also demonstrated that the *crtB* gene indeed encodes the functional phytoene synthase, CrtB, and participates in the carotenoid biosynthesis pathway in both *Hfx. mediterranei* and *Har. hispanica*.

4. Conclusion

As those *pyr*-based systems were established in the two model haloarchaea, *Hbt. salinarum* and *Hfx. volcanii* (Peck et al., 2000; Bitan-Banin et al., 2003), the newly developed *pyrF*-based gene knockout systems for *Hfx. mediterranei* and *Har. hispanica* are much more efficient than the mevinolin-resistance gene-based systems (e.g. pUBP). There are at least two reasons for this highly efficiency. First, as the mevinolin resistance gene, which encodes the 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase, has homologs in both *Hfx. mediterranei* and *Har. hispanica*, homologous recombination between the HMG-CoA reductase gene of pUBP and the chromosomal alleles of the host strains occurs frequently, especially for *Hfx. mediterranei*. Thus, the false positive integrations were usually generated with the mevinolin-resistance gene-based systems. Second and most importantly, as counterselection of the *pyrF* gene was performed to obtain double-crossover recombinants with the *pyrF*-based systems, the experimental time was largely shortened. Thus, these newly developed gene knockout

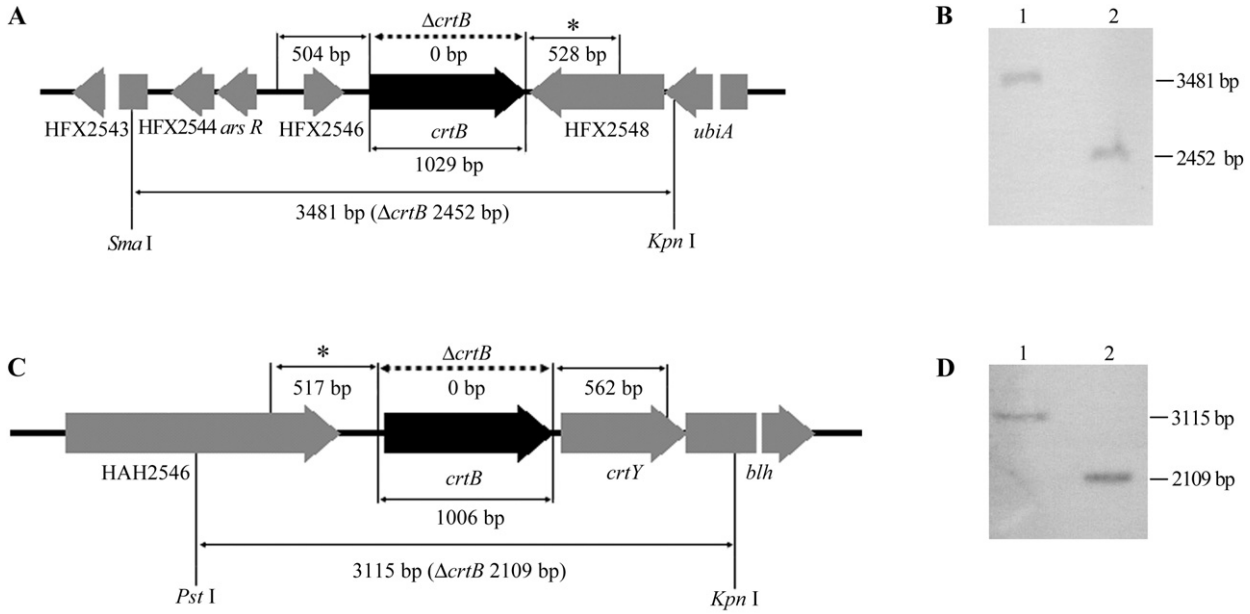

Fig. 5. Schematic diagram of *crtB* region and Southern blot analysis of mutant strains. **A:** *Hfx. mediterranei crtB* region. The upstream 504 bp and downstream 528 bp of *crtB* to be cloned into pHFX and the deleted 1029-bp region are indicated. The extracted genomic DNAs were digested with *Sma* I and *Kpn* I. Asterisk indicates 528-bp hybridization probe. The sizes of the hybridization fragments are also indicated: 3481 bp for DF50 and 2452 bp for DF50 △ *crtB*. **B:** Southern blot analysis of DF50 △ *crtB* mutant. Lane 1, DF50, 3481 bp; lane 2, DF50 △ *crtB*, 2452 bp. **C:** *Har. hispanica crtB* region. The upstream 517 bp and downstream 562 bp of *crtB* to be cloned into pHAR and the deleted 1006-bp region are indicated. The extracted genomic DNAs were digested with *Pst* I and *Kpn* I. Asterisk indicates the 517-bp hybridization probe. The sizes of hybridization fragments are also indicated: 3115 bp for DF60 and 2109 bp for DF60 △ *crtB*. **D:** Southern blot analysis of DF60 △ *crtB* mutant. Lane 1, DF60, 3115 bp; lane 2, DF60 △ *crtB*, 2109 bp.

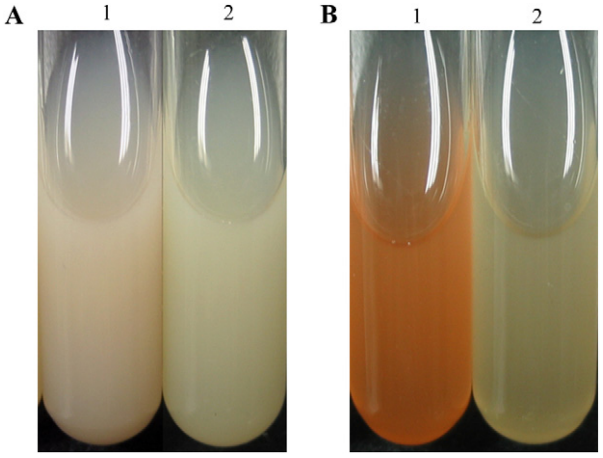

Fig. 6. Phenotypic comparisons. **A:** the color of the cell suspension was compared between DF50 (Tube 1) and DF50 △ *crtB* (Tube 2). **B:** the color of the cell suspension was compared between DF60 (Tube 1) and DF60 △ *crtB* (Tube 2).

systems would promote high-throughout generation of in-frame gene deletions in their respective genomes, and would accelerate researches on functional genomics and biotechnology of *Hfx. mediterranei* and *Har. hispanica*.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 30830004, 30925001), the National 863 Program of China (No. 2009AA09Z401), and the Chinese Academy of Sciences (No. KSCX2-EW-G-2).

References

Angelov, A., Liebl, W., 2006. Insights into extreme thermoacidophily based on genome analysis of *Picrophilus torridus* and other thermoacidophilic archaea. J. Biotechnol. 126, 3−10.

Bell, S.D., Jackson, S.P., 1998. Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. Trends Microbiol. 6, 222−228.

Bitan-Banin, G., Ortenberg, R., Mevarech, M., 2003. Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the *pyrE* gene. J. Bacteriol. 185, 772−778.

Boeke, J.D., LaCroute, F., Fink, G.R., 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol. Gen. Genet. 197, 345−346.

Burland, T.G., 2000. DNASTAR's Lasergene sequence analysis software. Methods Mol. Biol. 132, 71−91.

Cardenas, J.P., Valdes, J., Quatrini, R., Duarte, F., Holmes, D.S., 2010. Lessons from the genomes of extremely acidophilic bacteria and archaea with special emphasis on bioleaching microorganisms. Appl. Microbiol. Biotechnol. 88, 605−620.

Cline, S.W., Lam, W.L., Charlebois, R.L., Schalkwyk, L.C., Doolittle, W.F., 1989. Transformation methods for halophilic archaebacteria. Can. J. Microbiol. 35, 148−152.

DasSarma, S., Fleischmann, E.M., 1995. Halophiles. Cold Spring Harbor Laboratory Press, Plainview, N.Y.

Falb, M., Muller, K., Konigsmaier, L., Oberwinkler, T., Horn, P., von Gronau, S., Gonzalez, O., Pfeiffer, F., Bornberg-Bauer, E., Oesterhelt, D., 2008. Metabolism of halophilic archaea. Extremophiles 12, 177−196.

Hammelmann, M., Soppa, J., 2008. Optimized generation of vectors for the construction of *Haloferax volcanii* deletion mutants. J. Microbiol. Methods 75, 201−204.

Han, J., Lu, Q., Zhou, L., Zhou, J., Xiang, H., 2007. Molecular characterization of the $phaEC_{Hm}$ genes, required for biosynthesis of poly (3-hydroxybutyrate) in the extremely halophilic Archaeon *Haloarcula marismortui*. Appl. Environ. Microbiol. 73, 6058−6065.

Han, J., Lu, Q., Zhou, L., Liu, H., Xiang, H., 2009. Identification of the polyhydroxyalkanoate (PHA)-specific acetoacetyl coenzyme A reductase among multiple FabG paralogs in *Haloarcula hispanica* and reconstruction of the PHA biosynthetic pathway in *Haloferax volcanii*. Appl. Environ. Microbiol. 75, 6168−6175.

Han, J., Hou, J., Liu, H., Cai, S., Feng, B., Zhou, J., Xiang, H., 2010. Wide distribution among halophilic archaea of a novel polyhydroxyalkanoate synthase subtype with homology to bacterial type III synthases. Appl. Environ. Microb. 76, 7811−7819.

Juez, G., Rodriguezvalera, F., Ventosa, A., Kushner, D.J., 1986. *Haloarcula hispanica* spec. nov. and *Haloferax gibbonsii* spec. nov., two new species of extremely halophilic archaebacteria. Syst. Appl. Microbiol. 8, 75−79.

Kennedy, S.P., Ng, W.V., Salzberg, S.L., Hood, L., DasSarma, S., 2001. Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment through computational analysis of its genome sequence. Genome Res. 11, 1641−1650.

Khomyakova, M., Bukmez, O., Thomas, L.K., Erb, T.J., Berg, I.A., 2011. A methylaspartate cycle in haloarchaea. Science 331, 334−337.

Leigh, J.A., Albers, S.V., Atomi, H., Allers, T., 2011. Model organisms for genetics in the domain archaea: methanogens, halophiles, *thermococcales* and *sulfolobales*. FEMS Microbiol. Rev. doi:10.1111/j.1574-6976.2011.00265.x.

Liu, Y., Whitman, W.B., 2008. Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. Ann. N.Y. Acad. Sci. 1125, 171−189.

Lu, Q., Han, J., Zhou, L., Zhou, J., Xiang, H., 2008. Genetic and biochemical characterization of the poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) synthase in *Haloferax mediterranei*. J. Bacteriol. 190, 4173−4180.

Peck, R.F., DasSarma, S., Krebs, M.P., 2000. Homologous gene knockout in the archaeon *Halobacterium salinarum* with *ura3* as a counterselectable marker. Mol. Microbiol. 35, 667−676.

Rodriguez-Valera, F., Ruiz-Berraquero, F., Ramos-Cormenzana, A., 1980. Isolation of extremely halophilic bacteria able to grow in defined inorganic media with single carbon sources. J. Gen. Microbiol. 119, 535−538.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sieiro, C., Poza, M., de Miguel, T., Villa, T.G., 2003. Genetic basis of microbial carotenogenesis. Int. Microbiol. 6, 11−16.

Steinbuchel, A., Lutke-Eversloh, T., 2003. Metabolic engineering and pathway construction for biotechnological production of relevant polyhydroxyalkanoates in microorganisms. Biochem. Eng. J. 16, 81−96.

Trivedi, S., Gehlot, H.S., Rao, S.R., 2006. Protein thermostability in Archaea and Eubacteria. Genet. Mol. Res. 5, 816−827.

Vieille, C., Zeikus, G.J., 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol. Mol. Biol. Rev. 65, 1−43.

Woese, C.R., Kandler, O., Wheelis, M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA 87, 4576−4579.