9504–9513 *Nucleic Acids Research, 2014, Vol. 42, No. 14* *Published online 24 July 2014*
doi: 10.1093/nar/gku628

CRISPR-Cas: an efficient tool for genome engineering of virulent bacteriophages

Bruno Martel[1,2] and Sylvain Moineau[1,2,*]

[1]Département de biochimie, microbiologie et bio-informatique, Faculté des sciences et de génie, Université Laval, Quebec city, Quebec, G1V 0A6, Canada and [2]Groupe de recherche en écologie buccale, Faculté de médecine dentaire, Université Laval, Quebec City, Quebec, G1V 0A6, Canada

Received May 21, 2014; Revised June 25, 2014; Accepted June 27, 2014

ABSTRACT

Bacteriophages are now widely recognized as major players in a wide variety of ecosystems. Novel genes are often identified in newly isolated phages as well as in environmental metavirome studies. Most of these novel viral genes have unknown functions but appear to be coding for small, non-structural proteins. To understand their biological role, very efficient genetic tools are required to modify them, especially in the genome of virulent phages. We first show that specific point mutations and large deletions can be engineered in the genome of the virulent phage 2972 using the *Streptococcus thermophilus* CRISPR-Cas Type II-A system as a selective pressure to increase recombination efficiencies. Of significance, all the plaques tested contained recombinant phages with the desired mutation. Furthermore, we show that the CRISPR-Cas engineering system can be used to efficiently introduce a functional methyltransferase gene into a virulent phage genome. Finally, synthetic CRISPR bacteriophage insensitive mutants were constructed by cloning a spacer-repeat unit in a low-copy vector illustrating the possibility to target multiple regions of the phage genome. Taken together, this data shows that the CRISPR-Cas system is an efficient and adaptable tool for editing the otherwise intractable genomes of virulent phages and to better understand phage-host interactions.

INTRODUCTION

Bacteriophages (or phages) are recognized as the most prevalent biological entities on Earth and are believed to outnumber their bacterial hosts 10-fold (1). However, phages are still not well understood, as roughly half of their predicted genes have no functional assignments. Typically, after isolation and purification, phages are characterized by electron microscopy followed by genome sequencing. Significant effort has been devoted to studying structural proteins (capsid, tail) and their assembly pathways, which are responsible for phage morphology (2). Comparative genome analysis also revealed that late-expressed genes coding for these structural proteins are conserved between closely related phages (2). Moreover, the organization of these late-expressed genes is often similar among phages sharing the same morphology. Conversely, early-expressed phage genes coding for small, non-structural proteins are highly diverse, even between related viral isolates. These genes are often located next to each other, have no predicted function and represent most of the novel genes found in phage genomes. Metavirome studies and subsequent bioinformatic analyses have also revealed several novel phage genes flanking genes coding for known structural proteins (3).

The ability to construct phage mutants is central to studying the function of these proteins and determining which genes are essential. In contrast to the range of tools available for targeted mutagenesis of bacterial genomes, far fewer methods have been implemented for strictly virulent phages (4). It is relatively easy to inactivate (or add) a gene to the genome of a prophage through homologous recombination, effectively treating it as a chromosomal locus. Such an efficient system does not exist for virulent phages, because, in the absence of selection, high frequency recombination events are required. Historically, mutations in virulent phages were constructed by general mutagenesis using UV irradiation or chemical compounds (5). A homologous recombination approach can be used to make targeted mutations, but the frequency is low and screening is time-consuming (5).

The Bacteriophage Recombineering of Electroporated DNA (BRED) technique was recently developed to generate specific mutations, insertions, deletions and gene replacement in lytic mycobacteriophage (4,6). BRED exploits a phage recombinase to induce homologous recombination

*To whom correspondence should be addressed. Tel: +1 418 656 3712; Fax: +1 418 656 2861; Email: Sylvain.Moineau@bcm.ulaval.ca
Present address:
Sylvain Moineau, Groupe de recherche en écologie buccale, Faculté de médecine dentaire, Université Laval, Quebec City, Quebec, G1V 0A6, Canada.

(5,7–10) between a phage genome and a template, which must be electroporated into the cell (5). While also efficient in coliphages (11), this system must be optimized for each bacterial species and is difficult to use in strains refractory to electroporation. Depending on the type of mutation sought, BRED requires the screening of several lysis plaques (4).

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and their associated genes (*cas*) form an adaptive immune system that protects microbial cells from DNA invasion (12). A CRISPR locus is composed of direct repeats separated by short stretches of variable DNA called spacers. Overall, these systems possess extraordinarily diverse architectures for both the *cas* genes and the spacer content. Still, CRISPR-Cas systems are currently classified into three major types (I, II and III), divided into several subtypes (13). To provide immunity, the type II system follows three general steps: spacer acquisition, biogenesis of small RNAs and interference. In the acquisition step, a bacterial cell (BIM, Bacteriophage-Insensitive Mutant) acquires a new repeat-spacer unit in its CRISPR locus following a phage challenge (12). The new spacer in the CRISPR array is acquired from the invading DNA through the involvement of spacer acquisition motif located in the vicinity of a (proto)spacer in the phage genome (14). The CRISPR locus is then transcribed and processed with the help of *trans*-acting RNA (tracrRNA) and the host RNase III to produce smaller RNAs (crRNAs) (14). In the interference step, crRNAs and Cas9 proteins guide and cleave the invading DNA in a sequence-specific manner to ensure cell defence (15–19). Of interest here, phage mutants can bypass the interference activity (CEM, CRISPR-Escape Mutants) by point mutation in the protospacer (PS) sequence or adjacent motif (20).

The type II-A CRISPR-Cas system of *Streptococcus pyogenes* was recently optimized by the design of a guided RNA (gRNA) containing a crRNA fused with the tracrRNA (18). The addition of the *cas9* gene generated a programmable and adaptable tool for cleaving a specific DNA target. If a mutated template is provided in *trans*, the break can be repaired by homologous recombination, resulting in precise genome editing (18). The CRISPR-Cas9 engineering system has now been adapted and widely used with high efficiency for editing the genome of cells from humans (21–23), monkeys (24), livestock (25–27), plants (28–30), rabbits (31), rodents (32–34), amphibians (35,36), fishes (37), insects (38,39), nematodes (40), yeasts (41), bacteria (42) and latent eukaryotic virus (43). Very recently, Kiro *et al.* demonstrated that the type I-E CRISPR-Cas system of *Escherichia coli* can be used to generate a scarless deletion of two genes from the lytic coliphage T7 (44). Using this type I-E system, ~40% of the phages analysed carried the desired deletions.

Here, we show that the type II CRISPR-Cas system of *Streptococcus thermophilus* DGCC7710 can be exploited to edit at very high efficiencies the genome of virulent phages through point mutations, small and large deletions, as well as complete gene replacement.

MATERIALS AND METHODS

Bacterial strains, phages and culture conditions

The bacterial strains, phages and plasmids used in this study are listed in Supplementary Table S1. The bacterial strain *S. thermophilus* DGCC7710 and its derivatives were grown at 37°C or 42°C without agitation in M17 broth supplemented with 0.5% lactose (LM17) (45). The Gram-positive cloning host strain *Lactococcus lactis* MG1363 was grown at 30°C in M17 supplemented with 0.5% glucose (GM17). Chloramphenicol was added to media at a final concentration of 5 μg/ml for growth and selection of *L. lactis* and *S. thermophilus* containing pNT1 and pNZ123 plasmids and their derivatives. Virulent phage 2972 and its CEMs were routinely amplified on *S. thermophilus* DGCC7710 or *S. thermophilus* DGCC7710 BIM-derivatives, respectively, at 42°C in LM17 supplemented with 10 mM $CaCl_2$. Efficiencies of plaquing (EOPs) of phage 2972 and CEM derivatives were used to measure the effect of the CRISPR-Cas system as well as the efficiency of recombination. The EOPs were determined by dividing the phage titer obtained on the resistant strain by the phage titer obtained on the wild-type strain.

Isolation and characterization of CEMs of phage 2792

Streptococcus thermophilus DGCC7710 BIMs S90 to S95 were selected from our BIM collection (16) based on a single newly added spacer at the 5' end of the CRISPR1 locus targeting a phage gene of interest (Figure 1). CEM plaques were obtained at a frequency of $\sim 10^{-6}$ following the challenge of a BIM with a lysate of phage 2972 (10^9 pfu/ml). Ninety-four well-separated plaques from each BIM were picked with sterile tips and transferred into a flat-bottom 96-well plate containing 250 μl of sterile phage buffer (50 mM TRIS-hydrochloride pH 7.5, 100 mM NaCl, 8 mM MgSO4). After diffusion for 30 min at room temperature, 5 μl of each phage sample was then used as template to amplify the phage protospacer region by polymerase chain reaction (PCR) using a CFX96 C1000 Touch thermal cycler (BioRad). Primers used for the PCR reactions are listed in Supplementary Table S2. PCR products were sequenced using the ABI 3730xl DNA analyser (Applied Biosystems) at the Plateforme de Séquençage et de Génotypage des Génomes at the CHUL/CHUQ Center. Sequences were aligned with 2972 sequence and analysed using BioEdit (46). CEMs of interest following sequencing were then plaque purified three times (47).

Plasmid constructions and DNA manipulations

Plasmid constructs were obtained using the same general cloning strategy (Supplementary Figure S1) using XhoI and EcoRI restriction enzymes (Roche Applied Sciences) as indicated by the manufacturer. Each template was cloned into pNZ123 and confirmed by sequencing (49). Vectors were dephosphorylated with Antarctic phosphatase (New England Biolabs) and the fragments were ligated with T4 DNA ligase (Invitrogen) as specified by the manufacturer. Ligation products were electroporated in *L. lactis* MG1363 and,

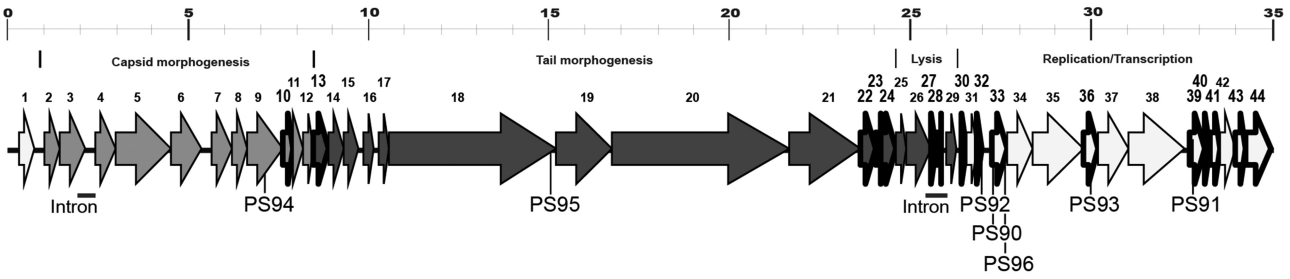

Figure 1. Genomic organization of virulent wild-type *S. thermophilus* phage 2972. Arrows represent *orfs*. Color shows transcription profile based on (48). Light gray indicates early gene expression (0–7 min post-infection), middle gray indicates middle gene expression (7–22 min post-infection), dark gray indicates late gene expression (12–27 min post-infection). *Orfs* without assigned function have bold outlines and are bold numbered. The relative locations of the protospacers targeted in this study are indicated by ≪PS≫ number.

ultimately, in *S. thermophilus* DGCC7710 and BIM derivatives as described elsewhere (50).

The S91R, oligo39ed and oligo39null templates were constructed by oligonucleotides annealing (Supplementary Table S2). One hundred pmol of each oligonucleotide was mixed in Taq PCR buffer (Feldan). Annealing conditions were: denaturation at 95°C for 5 min followed by cooling at the rate of −0.1°C/s to a temperature of 25°C and incubation for 10 min at 25°C. Oligonucleotide templates were purified using Qiaquick PCR Purification Kit (Qiagen).

RS91R template was constructed by annealing and elongating oligonucleotides XhoI-S91-for/EcoRI-S91-rev (Supplementary Table S2). Once annealed, Taq DNA polymerase (Feldan) and deoxyribonucleotides were added to the annealing mix and heated for 20 min at 72°C for elongation in a thermal cycler. RS91R template was purified as indicated above. RS91R and S91 templates were cloned into pNT1 (51), and oligo39ed and oligo39null were cloned into pNZ123. Phage 2972 *orf33* editing templates were constructed by PCR amplification using primers orf33ed-500bp-for/rev (pORF33ed-500bp) and orf33ed-1kbp-for/rev (pORF33ed-1kbp) with phage 2972 CEMS90 genomic DNA as template. Final templates were cloned into pNZ123.

LlaDCHIAed and orf39ed templates were designed and obtained through Integrated DNA Technologies (IDT) as a cloned fragment in the vector pIDTSmart, which we first transformed into *E. coli* MC1061. Plasmid DNA was extracted as mentioned above and editing template was recovered by a XhoI/EcoRI restriction, cloned and transformed as aforementioned.

Phage resistance efficacy of a repeat-spacer unit transcribed from a plasmid

pRS91R and pS91R were separately transformed into the wild-type, phage-sensitive strain *S. thermophilus* DGCC7710. The EOPs of wild-type phage 2972 were determined in the recombinant strain and compared with the wild-type strain.

Effect of phage genome editing on CRISPR-Cas system efficacy

Editing plasmids were separately transformed into *S. thermophilus* DGCC7710 BIM S90 (pORF33ed-500bp, pORF33ed-1kbp and pORF33-LlaDCHIAed) and BIM S91 (pORF39ed, pOLIGO39ed and pOLIGO39null). The EOPs of wild-type phage 2972, CEMS90 and CEMS91 were determined in *S. thermophilus* BIM S90 and BIM S91 with and without editing plasmids to verify if recombination between phage DNA and editing template increased the numbers of escaping phage mutants. The EOPs were calculated by dividing the phage titer obtained on a BIM (with and without editing plasmid) by the phage titer on the phage-sensitive wild-type strain.

Editing *orf39* from phage 2972

The CEM plaques obtained by challenging *S. thermophilus* BIM S91 containing editing plasmids (pORF39ed, pOLIGO39ed or pOLIGO39null) with phage 2972 or phage CEMS90 were transferred into phage buffer for diffusion. The protospacer region PS91 and its Protospacer-Adjacent-Motif located at the 3' end of a protospacer (PAM) were amplified by PCR and sequenced to check for mutations in CEMs. The sequences were aligned with the wild-type sequence and analyzed with BioEdit.

Editing *orf33* from phage 2972

The CEM plaques obtained by challenging *S. thermophilus* DGCC7710 BIM S90 containing editing plasmids (pORF33ed-500bp, pORF33ed-1kbp or pORF33-LlaDCHIAed) with wild-type phage 2972 were transferred into phage buffer for diffusion. The protospacer region PS90 and its PAM were amplified by PCR. The PCR products were migrated on a 2% agarose gel to confirm whether the CEMs contained deletions (from the pORF33ed-500bp or the pORF33ed-1kbp) or insertions (from pORF33-LlaDCHIAed). The CEMs from pORF33ed-500bp and pORF33ed-1kbp editing were purified on BIM S90 and the CEMs from pORF33-LlaDCHIAed were purified on BIM S96. The EOPs of CEMs obtained on BIM S90 containing pORF33-LlaDCHIAed were tested on *S. thermophilus* SMQ1107 to confirm whether the LlaDHCIA methyltransferase gene was acquired and was functional during phage infection. Phage CEM-LlaDCHIA DNA (amplified on DGCC7710) was extracted from a 10 ml lysate using QIAGEN Lambda DNA mini preparations (QIAGEN). Finally, DNA was digested separately with Sau3A and DpnI

(Roche Applied Sciences) and DNA fragments were separated by electrophoresis.

RESULTS

Determination of mutation pattern in protospacer and PAM

Deveau *et al.* previously reported that point mutation and deletion in the PS (30 nucleotides) or in the PAM (seven nucleotides, NNAGAAW, where W stands for A or T) allowed *S. thermophilus* phages to escape type II CRISPR-Cas interference (20). Therefore, the *S. thermophilus* DGCC7710's endogenous CRISPR1-Cas system could be directly used for the selection of random mutations within these 37-bp regions in any given streptococcal phage genome. However in the type I-E system of *E. coli*, only an 8-bp seed region of the PS flanking the PAM appears to be the region critical to the interference activity (52). Similarly, a 7-bp seed sequence in addition to the PAM was identified for the type II-A CRISPR3-Cas locus of *S. thermophilus* (53), which closely resembles the CRISPR-Cas system of *S. pyogenes* (17). However, this minimal sequence is unknown for the most active type II-A CRISPR1-Cas system of *S. thermophilus* DGCC7710. Accordingly, we predicted a theoretical region based on an *in silico* base-pairing model analysis of the crRNA and protospacer complex for selection of random mutations (17,18). This region would include 20 nucleotides from the 3'end of the PS and seven from the PAM (54).

To confirm this region, we selected *S. thermophilus* DGCC7710 derivatives (BIM S91 to S95) that are resistant to the virulent phage 2972 through the acquisition of a new spacer in their CRISPR1 locus (Supplementary Table S1) (12,16,19,20). These BIMs have acquired spacers that target either early-, middle- or late-expressed genes of phage 2972, with or without a known function (Figure 1). Next, high-titer preparations of virulent phage 2972 were used to challenge the selected *S. thermophilus* BIMs and numerous escaping phages (CEMs) were isolated at frequencies ranging from 10^{-5} to 10^{-7} (Table 1).

A total of 400 phage plaques from these assays were purified and their targeted PS regions were amplified by PCR and sequenced. No escaping mutations were detected in 18 of them (data not shown). Of the remaining 382 CEM phages, single nucleotide mutations were obtained in 380 of them and found, as expected, in the PS or in the PAM (Figure 2). Two out of the 382 CEMs had double mutations in the PS and/or PAM (Supplementary Table S3). Of note, two positions in the PAM (NNA**G**A**A**W) represent over 40% of the CEM mutations detected. As we obtained no escape phages with mutations in PAM positions other than the core AGAA, the other positions (**NN**AGAA**W**) may be only important for acquisition of new spacers and not for interference. While the seed sequence in the PS for CRISPR1-Cas appears to be 9 bp long, three CEMs with a mutation at position 13 of the PS also prevented the *S. thermophilus* CRISPR1-Cas interference activity.

This mutation frequency pattern in the PS and PAM provides critical information for targeting a protospacer region to isolate a CEM with a specific mutation at higher frequencies. For example, the G within the PAM of the CRISPR1-Cas system (A**G**AA) is the most frequently mutated CEM nucleotide allowing evasion of the CRISPR-Cas system. Thus, a nonsense mutation could readily be obtained by substituting a G for a T in the PAM, leading to an ochre stop codon (UAA), depending on the translation frame.

Isolation and characterization of phage 2972 CEMs with inactivated *orfs*

Previous analysis of the phage 2972 genome predicted 44 *orfs* but biological studies reduced that number to 40 due to the excision of two group I introns during phage infection (45). Fourteen of the 40 *orfs* have unknown function (Supplementary Table S4). Our goal was to select for CRISPR escaping mutants in PS regions of these genes with unknown functions using the CRISPR1-Cas. We targeted specific PS, which according to the now-identified seed sequence would be more likely to generate inactivating mutations such as the aforementioned UAA stop codon. Obtaining an inactivated gene would indicate that the phage gene is nonessential for replication under laboratory conditions. As a proof-of-concept, we selected four BIMs that acquired a spacer in their CRISPR1 locus from genes coding for unknown functions in phage 2972 genome. These genes are *orf32* (BIM S92), *orf33* (BIM S90), *orf36* (BIM S93) and *orf39* (BIM S91). All four genes are transcribed during infection (48,55).

S. thermophilus BIM S90 to S93 were challenged with virulent phage 2972 and the resulting CEMs were analyzed by sequencing a PCR product of the PS region. Three CEMs were of particular interest (Table 2). CEMS90, obtained on BIM S90, had an in-frame deletion of 306 bp (out of 471 bp) indicating that *orf33* could be non-essential for 2972. CEMS91 and CEMS92 escaped interference through nonsense mutations within the PAM sequence at the fourth codon of *orf39* and in the middle of *orf32,* respectively (Table 2), indicating that these genes are also likely dispensable. These results indicate that the CRISPR-Cas system can be used to isolate phage mutants with nonsense mutations or large deletions, but many phages must be screened to identify any given specific mutation (Table 2).

In contrast, nonsense mutations were not found in any of the CEMs obtained by challenging BIM S93 with phage 2972, despite the possibility of generating two stop codons (UAA) with a C to T nucleotide substitution at position 28 of PS93 or G to T nucleotide substitution in the PAM. Many synonymous and non-synonymous mutations were found, but none leading to inactivation showing that *orf36* is probably essential for phage 2972 propagation.

Construction of synthetic BIMs

After infection of *S. thermophilus* DGCC7710 with virulent phage 2972, CRISPR BIMs are relatively easy to obtain naturally (12,15,16,20) but isolating a BIM with the spacer targeting the desired phage sequence can be tedious (56,57). Thus, we constructed a plasmid-based system to provide phage resistance without the need to screen hundreds of colonies for the needed acquired spacer. We cloned a repeat-spacer-repeat unit (R-S-R) containing repeats from CRISPR1 into the vector pNT1. The expression of the R-S-R unit is under the control of the P_{mob} promoter found

Table 1. EOPs of phage 2972 and CEM derivatives on various *S. thermophilus* strains

Phage	Strains	EOPs
2972	BIM S90	$1.8 \pm 0.4 \times 10^{-6}$
	BIM S91	$3.2 \pm 0.3 \times 10^{-5}$
	BIM S92	$5.2 \pm 0.6 \times 10^{-7}$
	BIM S93	$7.7 \pm 0.4 \times 10^{-5}$
	BIM S94	$2.8 \pm 0.0 \times 10^{-6}$
	BIM S95	$2.4 \pm 0.2 \times 10^{-6}$
	SMQ1276	$1.3 \pm 0.1 \times 10^{-5}$
	DGCC7710-pRS91R	$1.5 \pm 0.1 \times 10^{-5}$

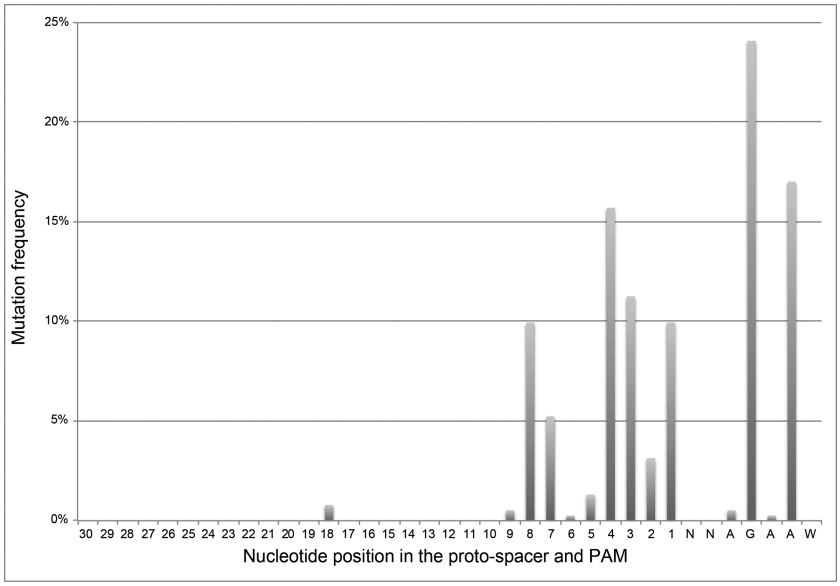

Figure 2. Mutation frequency of protospacer and PAM in CEM phages. We analyzed the protospacers and PAMs of 400 CEM using PCR and sequencing. The 380 CEMs containing a unique mutation were compiled for this chart. Isolated CEMs were obtained on different BIMs with spacers targeting regions of *orf9* (BIM S95), *orf18* (BIM S94), *orf32* (BIM S92), *orf36* (BIM S93) and *orf39* (BIM S91) of phage 2972.

Table 2. Isolated CEMs of interest and their specificities

Phages	Frequency[a]	Mutated *Orf*[b]	Position[b]	Effects	EOP on host BIM
CEMS90	1.1%	*orf33*	27006–27311	Deletion of 102 aa	1.3 ± 0.1
CEMS91	8.2%	*orf39*	32475	Glu4Stop	$6.9 \pm 0.9 \times 10^{-1}$
CEMS92	9.5%	*orf32*	26642	Glu25Stop	$9.8 \pm 0.5 \times 10^{-1}$
CEMS93–1	17.2%	*orf36*	29584	Phe49: TTC→TTT Rare to frequent codon[c]	$6.4 \pm 0.6 \times 10^{-1}$
CEMS93–2	21.8%	*orf36*	29593	Glu52: GAA→GAG Frequent to rare codon[c]	$8.6 \pm 0.5 \times 10^{-1}$

[a]Compared with CEMs obtained on the same BIM.
[b]Based on the GenBank sequence of phage 2972.
[c]According to the codon usage of *S. thermophilus* LMD-9 determined in the Codon Usage Database.

on the vector (Supplementary Figure S1d). The spacer naturally acquired by BIM S91 was used as a proof-of-concept and cloned into pNT1 (51). The resulting recombinant plasmid was introduced into *S. thermophilus* DGCC7710 and the level of resistance against phage 2972 (as determined by efficiency of plaquing) provided by the plasmid was similar to that observed for the natural BIM S91 (Table 1).

The processing of the crRNA (17) in *S. thermophilus* DGCC7710 was recently determined (56,57) and the results suggested that the repeat next to the 5' end of the spacer was not essential for interference. Therefore, the minimal construct S91R (spacer-repeat unit) was inserted into the pNT1 vector, as for pRS91R, and the resulting plasmid introduced into *S. thermophilus* DGCC7710. The presence of pS91R provided a phage resistance phenotype but it was less effective than BIM S91 or pRS91R (Table 1). Taken together, our data show the possibility of using a plasmid-based sys-

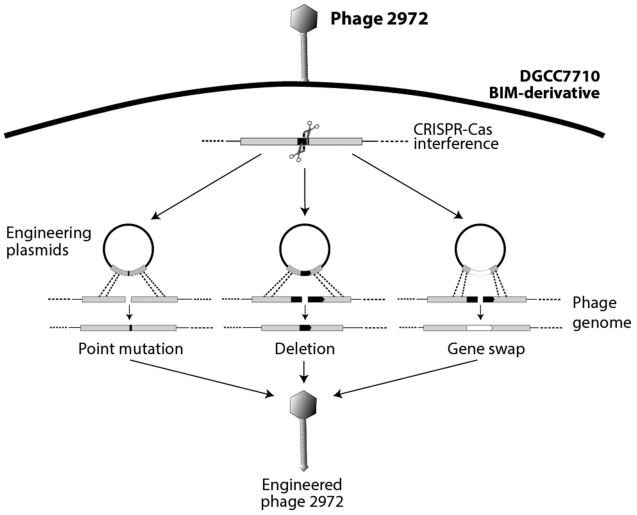

Figure 3. General schematic view of genome editing for phage 2972 through different strategies. Gray boxes indicate homologous regions between editing template (on plasmid shown by circles) and phage 2972 genome. Scissors show the Cas9 DNA double-stranded break of phage genome DNA. Dotted lines show possible recombination events. White and black arrows represent open reading frames. Point mutation is showed by a small black rectangle.

tem to target a phage region and generate a phage resistance phenotype.

Genome editing of virulent streptococcal phage genomes

Since the CRISPR-Cas system can be used to select CEMs, it may be possible to directly edit a phage genome and obtain the CEM of interest without the need for large-scale screening. Indeed, direct mutagenesis of a phage genome could be achieved by challenging a BIM host, harboring a selected targeting spacer in its CRISPR array as well as an editing plasmid, with a virulent phage and selecting for recombinant phage in the resulting lysate (42,44). The editing plasmid must contain a recombination template with the desired mutation that also prevent CRISPR-Cas interference as well as include sequences homologous to the phage genome to be engineered (42,44). Templates were constructed, accordingly, cloned in a high copy vector and transformed into BIMs of DGCC7710 (Supplementary Figure S1). These BIMs provide selective pressure for recombinant phages by targeting a region of the phage genome. Mutation type can go to point mutation, deletion or gene exchange (Figure 3).

Phage genome editing: proof-of-concept. The plasmid pORF39ed was constructed and used as a proof-of-concept by mimicking a nonsense mutation in *orf39* as observed in the natural CEMS91. pORF39ed was introduced into *S. thermophilus* BIM S91 and challenged with phage 2972 to compare the number of CRISPR escaping phages obtained on BIM S91 with or without the editing template. We obtained a stunning increase of almost four orders of magnitude in the phage EOP on the strain carrying pORF39ed (Table 3). Sequencing of the *orf39* region from 10 phage plaques isolated on BIM S91-pORF39ed confirmed the presence, in the phage genome, of the mutation provided by pORF39ed (Supplementary Figure S2a).

While the above mutated template (339 bp) cloned into pORF39ed was synthesized through a commercial service, we also constructed a shorter editing template (72 bp) at lower cost using an oligonucleotide annealing approach (Supplementary Table S2 and Supplementary Figure S1a). The resulting pOLIGO39ed was introduced into *S. thermophilus* BIM S91 and the resulting recombinant strain infected with phage 2972. The EOP also significantly increased but by two logs (as compared to four logs for the longer mutated template), presumably because of reduced recombination frequency due to the smaller homologous regions. Again, all the escaping phages recovered on *S. thermophilus* BIM S91-pOLIGO39ed contained the specific CEMS91 mutation (Supplementary Figure S2a), indicating that a DNA template as small as 72 bp can be used to introduce point mutations in the genome of a virulent phage.

Moreover, we recreated the 306 bp deletion in *orf33* observed in the natural CEMS90 (Table 2). We used the genomic DNA of CEMS90 to construct the orf33ed-500bp and orf33–1kbp, ∼250 bp and 500 bp of homologous sequences each side of the deletion respectively, templates by PCR amplification (Supplementary Table S1 and Supplementary Figure S1b). We cloned each template into the vector pNZ123 to construct pORF33ed-500bp and pORF33ed-1kbp (Supplementary Figure S1b) and introduced them, separately, into *S. thermophilus* BIM S90. Infection assays with phage 2972 led in both cases to a two-log increase in EOP (Table 3). Sequence analysis of escaping phages confirmed the expected genome deletion in all of them, indicating that the presence of these templates favoured homologous recombination, although here the template size had no discernable impact (Table 3). No deletions were found in CEM phages isolated from *S. thermophilus* BIM S90 containing the empty pNZ123 vector.

Phage genome editing: small frameshift deletion. We constructed pOLIGO39null, which contains a 72-bp template (cloned into pNZ123) from which the first AG nucleotides of the PS91 PAM were removed to obtain a -2 frameshift, inactivating *orf39* (Supplementary Table S2). This small deletion had not been previously obtained with natural CEMs (Supplementary Table S3). We challenged BIM S91-pOLIGO39null with virulent phage 2972 (Table 3). Plaques were smaller than the one obtained on BIM S91-OLIGO39ed. Sequencing results of the PS region of CEMs showed the 2 bp deletion (Supplementary Figure S2a), illustrating the adaptability of this genome engineering tool.

Phage genome editing: gene exchange. We then designed a gene exchange strategy in the genome of virulent phage 2972 to illustrate the versatility of the CRISPR-Cas system as an engineering tool. We opted to replace an existing gene rather than simply incorporate a new one, as the length of phage genomes is limited by the capsid size and the maximum genome length is currently unknown for phage 2972 (58). We elected to introduce a methyltransferase gene of the type II restriction/modification (R/M) system LlaDCHI (59) from *L. lactis* since this system is functional in *S. thermophilus* DGCC7710 and does not interfere with the

Table 3. EOPs of phage 2972 on BIMs and strains used for genome editing

Phages	Strains	Plasmids	EOPs	Δ $\log_{10}$[a]
2972	BIM S91	pNZ123	$3.2 \pm 0.3 \times 10^{-5}$	-
		pORF39ed	$1.8 \pm 0.1 \times 10^{-1}$	3.8
		pOLIGO39ed	$3.8 \pm 0.8 \times 10^{-3}$	2.1
		pOLIGO39null	$9.6 \pm 0.4 \times 10^{-5}$	0.5
CEMS91		pNZ123	$6.9 \pm 0.9 \times 10^{-1}$	4.3
2972	BIM S90	pNZ123	$1.8 \pm 0.4 \times 10^{-6}$	-
		pORF33ed-500bp	$1.1 \pm 0.3 \times 10^{-3}$	2.8
		pORF33ed-1kp	$1.9 \pm 0.6 \times 10^{-3}$	3.0
		pORF33-LlaDCHIAed	$5.3 \pm 0.4 \times 10^{-5}$	1.5
CEMS90		pNZ123	1.3 ± 0.1	5.9
2972	SMQ1107	pSRQ707	$3.6 \pm 0.8 \times 10^{-6}$	-
CEM-LlaDCHIA		pSRQ707	$9.6 \pm 0.7 \times 10^{-1}$	-

[a]Increase of EOP calculated by dividing EOP of phage 2972 on BIM containing editing plasmid by EOP of phage 2972 obtained on BIM containing empty vector and expressed in a base 10 logarithmic scale.

DGCC7710 CRISPR-Cas systems (19). Therefore, the presence of a functional methyltransferase gene in the phage genome would provide a marker for selection. Our strategy was to replace *orf33* in the phage 2972 genome with *LlaDCHIA* since *orf33* is a non-essential gene (see above). We calculated that this gene swapping would increase the genome length of phage 2972 by only 1%.

The ORF33-LlaDCHIAed template was designed to contain *LlaDCHIA*, as well as 220 bp (5' end) and 227 bp (3'end) sequences homologous to phage 2972. The resulting plasmid pORF33-LlaDCHIAed was transformed into *S. thermophilus* BIM S90 containing a spacer targeting a PS in *orf33* (Supplementary Figure S1c) and the recombinant strain was infected with the wild-type 2972 (Table 3). Genome sequencing of the resulting escape phages confirmed that *orf33* had been exchanged with the plasmid template, leading to escaping phages with a methyltransferase gene.

Phage plaques were purified using *S. thermophilus* BIM S96, which has a spacer targeting a PS located near the 3'-end of *orf33* (Figure 1), to avoid amplification of CEMs containing a CRISPR escape mutation into PS90. Purified CEM-LlaDCHIA plaques were subsequently amplified on *S. thermophilus* DGCC7710 for DNA methylation assays (Supplementary Figure S2c). The EOPs of phages 2972 and CEM-LlaDCHIA were then determined on *S. thermophilus* SMQ-1107, a DGCC7710 derivative containing the LlaDCHI R/M system (Table 3). As expected, the EOP of phage 2972 on SMQ-1107 was of 10^{-6}. Conversely, the EOP of CEM-LlaDCHIA was almost 1, indicating total resistance to the LlaDCHI R/M system, likely through *in vivo* methylation. To confirm methylation, the genomic DNA of phage CEM-LlaDCHIA propagated on the wild-type *S. thermophilus* DGCC7710 was extracted and digested with DpnI, which cleaves methylated G^{m}ATC sites (the methylation site of LlaDCHIA methyltransferase). As expected, DpnI cut the genomic DNA of CEM-LlaDCHIA but not the DNA of the wild-type phage 2972 propagated on the same host (Supplementary Figure S2d). Both genomes of wild-type and CEM-LlaDCHIA phages were successfully cut by the Sau3A restriction enzyme, which cleaves GATC sites regardless of methylation. These data indicate that new and functional genes can be added to the genome of virulent phage using the CRISPR-Cas system.

DISCUSSION

The CRISPR1-Cas system (Type II-A) of *S. thermophilus* DGCC7710 can be used as a tool to select CEMs containing mutations in a specific region (PS or PAM) of a virulent phage. We determined the locations of natural mutations in almost 400 naturally obtained CEMs generated due to CRISPR interference. Most mutations were found in a 9-bp region at the 3'-end of the PS and at two positions within the PAM (NNA**G**A**A**W) (Figure 2), which is similar to those obtained for CRISPR3 of *S. thermophilus* (53) and for *S. pyogenes* Cas9 (18). A perfect match between the 18th nucleotide of the PS and the crRNA is also needed for interference. Of note, recent off-target studies of *S. pyogenes* Cas9 protein and several single-guided RNAs have shown a longer and discontinuous seed sequence (60,61). Cas9 can tolerate some mutations and they vary depending on the targeted genomic region and this may explain some of the escape mutations for CRISPR1-Cas system.

The genome (34 704 bp) of phage 2972 contains 233 PAM sequences (NNAGAAW) including at least one in each of its 40 *orfs*, indicating that every gene can be targeted by the *S. thermophilus* CRISPR1-Cas system. To target a gene, you need either a natural BIM that has acquired a spacer targeting the phage gene of interest or, as we demonstrated, an engineered BIM obtained by cloning a repeat-spacer-repeat into a low-copy plasmid. The engineered approach shows the feasibility of generating BIMs at low cost and avoiding the time-consuming procedure of randomly isolating a natural BIM containing the spacer of interest (56). Selection of a spacer for engineering a BIM is based on *in silico* analysis to target a mutation of interest, specifically using the base pairing of the CRISPR1-Cas interference complex (Cas9 and crRNA::tracrRNA duplex) with the phage 2972 genome, as well as the seed sequence proposed here.

This approach can be used to isolate phages with the desired mutations, such as the presence of a stop codon to determine if a given phage gene is essential for replication. The isolation of CEMS93–1 and CEMS93–2 (Table 2) showed that synonymous and non-synonymous mutations can also

be obtained to study codon optimization. However, several phages need to be screen to obtain the desired mutated phages.

To reduce the screening of phage plaques, we successfully developed a strategy for precise genome engineering of the virulent phage 2972 genome using the wild-type CRISPR1-Cas locus of *S. thermophilus* DGCC7710. By providing an editing template on a high-copy number plasmid, we were able to generate phage 2972 CEMs with specific point mutations and deletions. More importantly, point mutation editing results showed a 1000-fold increase of CEMs in the presence of the 339 bp template ORF39ed (Table 3), indicating a high recombination frequency. In fact, 100% of the tested CEMs were recombinant phages. Even with as small a length of template as 72 bp (OLIGO39ed), a 100-fold increase of escaping phages was still observed, and again all primary plaques tested were recombinant CEM phages.

We successfully inactivated *orf39* of phage 2972 with a 2-bp frameshift deletion by recombination with the OLIGO39null template. However, the recombination efficiency was lower compared to the results obtained with the nonsense mutation templates (ORF39ed/OLIGO39ed), but still eight plaques out of 10 were from recombinant phages. We hypothesize that this *orf39* mutation affects phage 2972 fitness as evidenced, in part, by smaller plaques than the ones obtained with nonsense mutation. A large deletion was also generated using 449 bp and 997 bp editing templates and, again, a 600- to 1000-fold increase in recombinant escaping phages was obtained.

One of the most exciting applications of the CRISPR-Cas system presented here is the addition of a new gene into the genome of the virulent phage 2972. We showed that *orf33* of phage 2972 is dispensable and can be replaced by a larger bacterial gene, the lactococcal methyltransferase *LlaDCHI*A. Phage assays showed that the introduced methyltransferase was fully functional, as demonstrated by *in vivo* methylation leading to immunity to restriction by the LlaDCHI R/M system.

This work expounds an easy and rapid method to inactivate or modify any phage gene of interest, even in strictly lytic phages, at very high efficiencies. The procedure is likely amenable to modification for very high throughput. Such a CRISPR-based tool will allow us to study the roles of various novel phage genes *in vivo* during replication as well as their interactions with hosts by fusion of tags to proteins of interest. Moreover, we show the possibility of swapping a non-essential gene for a new functional gene in a phage genome. We believed that this approach can be easily adapted to engineer any virulent phage infecting other bacterial species. As interest in the use of virulent phages for biocontrol and therapy purposes increases, we offer a new tool to modify a phage of interest to maximize its efficacy.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank A. P. Hynes for discussions and Barb Conway for editorial assistance.

FUNDING

B.M. is the recipient of a Novalait/FQRNT graduate scholarship. S.M. acknowledges funding from NSERC of Canada (Discovery program). S.M. holds a Tier 1 Canada Research Chair in Bacteriophages.
Conflict of interest statement. None declared.

REFERENCES

1. Breitbart,M. and Rohwer,F. (2005) Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.*, **13**, 278–284.
2. San Martín,C. (2013) In Mateu,M.G. (ed.), *Structure and Physics of Viruses*. Springer, Netherlands, pp. 329–360.
3. Rosario,K. and Breitbart,M. (2011) Exploring the viral world through metagenomics. *Curr. Opin. Virol.*, **1**, 289–297.
4. Marinelli,L.J., Piuri,M., Swigoňová,Z., Balachandran,A., Oldfield,L.M., van Kessel,J.C. and Hatfull,G.F. (2008) BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. *PloS One*, **3**, e3957.
5. Marinelli,L.J., Hatfull,G.F. and Piuri,M. (2012) Recombineering: a powerful tool for modification of bacteriophage genomes. *Bacteriophage*, **2**, 5–14.
6. da Silva,J.L., Piuri,M., Broussard,G., Marinelli,L.J., Bastos,G.M., Hirata,R.D., Hatfull,G.F. and Hirata,M.H. (2013) Application of BRED technology to construct recombinant D29 reporter phage expressing EGFP. *FEMS Microbiol. Lett.*, **344**, 166–172.
7. Court,D.L., Sawitzke,J.A. and Thomason,L.C. (2002) Genetic engineering using homologous recombination. *Annu. Rev. Genet.*, **36**, 361–388.
8. Murphy,K.C. (2012) Phage recombinases and their applications. *Adv. Virus Res.*, **83**, 367–414.
9. Oppenheim,A.B., Rattray,A.J., Bubunenko,M., Thomason,L.C. and Court,D.L. (2004) *In vivo* recombineering of bacteriophage lambda by PCR fragments and single-strand oligonucleotides. *Virology*, **319**, 185–189.
10. Sharan,S.K., Thomason,L.C., Kuznetsov,S.G. and Court,D.L. (2009) Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.*, **4**, 206–223.
11. Fehér,T., Karcagi,I., Blattner,F.R. and Pósfai,G. (2012) Bacteriophage recombineering in the lytic state using the lambda red recombinases. *Microb. Biotechnol.*, **5**, 466–476.
12. Barrangou,R., Fremaux,C., Deveau,H., Richards,M., Boyaval,P., Moineau,S., Romero,D.A. and Horvath,P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**, 1709–1712.
13. Makarova,K.S., Haft,D.H., Barrangou,R., Brouns,S.J., Charpentier,E., Horvath,P., Moineau,S., Mojica,F.J., Wolf,Y.I., Yakunin,A.F. *et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.*, **9**, 467–477.
14. Almendros,C., Guzmán,N.M., Díez-Villaseñor,C., García-Martínez,J. and Mojica,F.J. (2012) Target motifs affecting natural immunity by a constitutive CRISPR-Cas system in *Escherichia coli*. *PloS One*, **7**, e50797.
15. Garneau,J.E., Dupuis,M.È., Villion,M., Romero,D.A., Barrangou,R., Boyaval,P., Fremaux,C., Horvath,P., Magadán,A.H. and Moineau,S. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, **468**, 67–71.
16. Magadán,A.H., Dupuis,M.È., Villion,M. and Moineau,S. (2012) Cleavage of Phage DNA by the *Streptococcus thermophilus* CRISPR3-Cas System. *PloS One*, **7**, e40913.
17. Deltcheva,E., Chylinski,K., Sharma,C.M., Gonzales,K., Chao,Y., Pirzada,Z.A., Eckert,M.R., Vogel,J. and Charpentier,E. (2011) CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature*, **471**, 602–607.
18. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
19. Dupuis,M.È., Villion,M., Magadán,A.H. and Moineau,S. (2013) CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nat. Commun.*, **4**,2087, 1–7.

20. Deveau,H., Barrangou,R., Garneau,J.E., Labonté,J., Fremaux,C., Boyaval,P., Romero,D.A., Horvath,P. and Moineau,S. (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.*, **190**, 1390–1400.
21. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
22. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
23. Hsu,P.D., Lander,E.S. and Zhang,F. (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, **157**, 1262–1278.
24. Niu,Y., Shen,B., Cui,Y., Chen,Y., Wang,J., Wang,L., Kang,Y., Zhao,X., Si,W., Li,W. *et al.* (2014) Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell*, **156**, 836–843.
25. Fan,Z., Lee,S., Park,H., Lucibello,K., Meng,Q., Polejaeva,I., White,K., Bunch,T. and Wang,Z. (2013) 107 genetic inactivation of the *sry* gene in Argali wild and Romney domestic sheep with CRISPR/Cas systems for producing sex-reversed female animals. *Reprod. Fertil. Dev.*, **26**, 167.
26. Hu,S., Wang,Z. and Polejaeva,I. (2013) 40 knockout of goat nucleoporin 155 (*NUP155*) gene using CRISPR/Cas9 systems. *Reprod. Fertil. Dev.*, **26**, 134.
27. Tan,W., Carlson,D.F., Lancto,C.A., Garbe,J.R., Webster,D.A., Hackett,P.B. and Fahrenkrug,S.C. (2013) Efficient nonmeiotic allele introgression in livestock using custom endonucleases. *Proc. Natl Acad. Sci. U.S.A.*, **110**, 16526–16531.
28. Jiang,W., Zhou,H., Bi,H., Fromm,M., Yang,B. and Weeks,D.P. (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Res.*, **41**, e188.
29. Li,J.F., Norville,J.E., Aach,J., McCormack,M., Zhang,D., Bush,J., Church,G.M. and Sheen,J. (2013) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.*, **31**, 688–691.
30. Shan,Q., Wang,Y., Li,J., Zhang,Y., Chen,K., Liang,Z., Zhang,K., Liu,J., Xi,J.J., Qiu,J.L. *et al.* (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.*, **31**, 686–688.
31. Yang,D., Xu,J., Zhu,T., Fan,J., Lai,L., Zhang,J. and Chen,Y.E. (2014) Effective gene targeting in rabbits using RNA-guided Cas9 nucleases. *J. Mol. Cell Biol.*, **6**, 681–683.
32. Li,D., Qiu,Z., Shao,Y., Chen,Y., Guan,Y., Liu,M., Li,Y., Gao,N., Wang,L., Lu,X. *et al.* (2013) Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat. Biotechnol.*, **31**, 681–683.
33. Li,W., Teng,F., Li,T. and Zhou,Q. (2013) Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat. Biotechnol.*, **31**, 684–686.
34. Wang,H., Yang,H., Shivalila,C.S., Dawlaty,M.M., Cheng,A.W., Zhang,F. and Jaenisch,R. (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, **153**, 910–918.
35. Blitz,I.L., Biesinger,J., Xie,X. and Cho,K.W. (2013) Biallelic genome modification in F0 *Xenopus tropicalis* embryos using the CRISPR/Cas system. *Genesis*, **51**, 827–834.
36. Nakayama,T., Fish,M.B., Fisher,M., Oomen-Hajagos,J., Thomsen,G.H. and Grainger,R.M. (2013) Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in *Xenopus tropicalis*. *Genesis*, **51**, 835–843.
37. Hwang,W.Y., Fu,Y., Reyon,D., Maeder,M.L., Tsai,S.Q., Sander,J.D., Peterson,R.T., Yeh,J.R. and Joung,J.K. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.*, **31**, 227–229.
38. Gratz,S.J., Cummings,A.M., Nguyen,J.N., Hamm,D.C., Donohue,L.K., Harrison,M.M., Wildonger,J. and O'Connor-Giles,K.M. (2013) Genome engineering of drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics*, **194**, 1029–1035.
39. Wang,Y., Li,Z., Xu,J., Zeng,B., Ling,L., You,L., Chen,Y., Huang,Y. and Tan,A. (2013) The CRISPR/Cas system mediates efficient genome engineering in *Bombyx mori*. *Cell Res.*, **23**, 1414–1416.
40. Friedland,A.E., Tzur,Y.B., Esvelt,K.M., Colaiácovo,M.P., Church,G.M. and Calarco,J.A. (2013) Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods*, **10**, 741–743.
41. DiCarlo,J.E., Norville,J.E., Mali,P., Rios,X., Aach,J. and Church,G.M. (2013) Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.*, **41**, 4336–4343.
42. Jiang,W., Bikard,D., Cox,D., Zhang,F. and Marraffini,L.A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.*, **31**, 233–239.
43. Ebina,H., Misawa,N., Kanemura,Y. and Koyanagi,Y. (2013) Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci. Rep.*, **3**, 2510.
44. Kiro,R., Shitrit,D. and Qimron,U. (2014) Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. *RNA Biol.*, **11**, 42–44.
45. Lévesque,C., Duplessis,M., Labonté,J., Labrie,S., Fremaux,C., Tremblay,D. and Moineau,S. (2005) Genomic organization and molecular analysis of virulent bacteriophage 2972 infecting an exopolysaccharide-producing *Streptococcus thermophilus* strain. *Appl. Environ. Microbiol.*, **71**, 4057–4068.
46. Hall,T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, **41**, 95–98.
47. Moineau,S., Pandian,S. and Klaenhammer,T.R. (1994) Evolution of a lytic bacteriophage via DNA acquisition from the *Lactococcus lactis* chromosome. *Appl. Environ. Microbiol.*, **60**, 1832–1841.
48. Duplessis,M., Russell,W.M., Romero,D.A. and Moineau,S. (2005) Global gene expression analysis of two *Streptococcus thermophilus* bacteriophages using DNA microarray. *Virology*, **340**, 192–208.
49. De Vos,W.M. (1987) Gene cloning and expression in lactic streptococci. *FEMS Microbiol. Lett.*, **46**, 281–295.
50. Moineau,S., Walker,S.A., Holler,B.J., Vedamuthu,E.R. and Vandenbergh,P.A. (1995) Expression of a *Lactococcus lactis* phage resistance mechanism by *Streptococcus thermophilus*. *Appl. Environ. Microbiol.*, **61**, 2461–2466.
51. Vaillancourt,K., Bédard,N., Bart,C., Tessier,M., Robitaille,G., Turgeon,N., Frenette,M., Moineau,S. and Vadeboncoeur,C. (2008) Role of *galK* and *galM* in galactose metabolism by *Streptococcus thermophilus*. *Appl. Environ. Microbiol.*, **74**, 1264–1267.
52. Semenova,E., Jore,M.M., Datsenko,K.A., Semenova,A., Westra,E.R., Wanner,B., van der Oost,J., Brouns,S.J. and Severinov,K. (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl Acad. Sci. U.S.A.*, **108**, 10098–10103.
53. Sapranauskas,R., Gasiunas,G., Fremaux,C., Barrangou,R., Horvath,P. and Siksnys,V. (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.*, **39**, 9275–9282.
54. Carte,J., Christopher,R.T., Smith,J.T., Olson,S., Barrangou,R., Moineau,S., Glover,C.V. III, Graveley,B.R., Terns,R.M. and Terns,M.P. (2014) The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Mol. Microbiol.*, **93**, 98-112 .
55. Young,J.C., Dill,B.D., Pan,C., Hettich,R.L., Banfield,J.F., Shah,M., Fremaux,C., Horvath,P., Barrangou,R. and Verberkmoes,N.C. (2012) Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PloS One*, **7**, e38077.
56. Paez-Espino,D., Morovic,W., Sun,C.L., Thomas,B.C., Ueda,K., Stahl,B., Barrangou,R. and Banfield,J.F. (2013) Strong bias in the bacterial CRISPR elements that confer immunity to phage. *Nat. Commun.*, **4**,1430, 1–7.
57. Sun,C.L., Barrangou,R., Thomas,B.C., Horvath,P., Fremaux,C. and Banfield,J.F. (2013) Phage mutations in response to CRISPR diversification in a bacterial population. *Environ. Microbiol.*, **15**, 463–470.
58. Black,L.W. (1989) DNA packaging in dsDNA bacteriophages. *Annu. Rev. Microbiol.*, **43**, 267–292.
59. Moineau,S., Walker,S.A., Vedamuthu,E.R. and Vandenbergh,P.A. (1995) Cloning and sequencing of LlaDCHI [corrected] restriction/modification genes from *Lactococcus lactis* and relatedness of this system to the *Streptococcus pneumoniae* DpnII system. *Appl. Environ. Microbiol.*, **61**, 2193–2202.

60. Fu,Y., Foden,J.A., Khayter,C., Maeder,M.L., Reyon,D., Joung,J.K. and Sander,J.D. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.*, **31**, 822–826.

61. Hsu,P.D., Scott,D.A., Weinstein,J.A., Ran,F.A., Konermann,S., Agarwala,V., Li,Y., Fine,E.J., Wu,X., Shalem,O. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.*, **31**, 827–832.