ARTICLE

doi:10.1038/nature09523

The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA

Josiane E. Garneau[1], Marie-Ève Dupuis[1], Manuela Villion[1], Dennis A. Romero[2], Rodolphe Barrangou[2], Patrick Boyaval[3], Christophe Fremaux[3], Philippe Horvath[3], Alfonso H. Magadán[1] & Sylvain Moineau[1]

Bacteria and Archaea have developed several defence strategies against foreign nucleic acids such as viral genomes and plasmids. Among them, clustered regularly interspaced short palindromic repeats (CRISPR) loci together with *cas* (CRISPR-associated) genes form the CRISPR/Cas immune system, which involves partially palindromic repeats separated by short stretches of DNA called spacers, acquired from extrachromosomal elements. It was recently demonstrated that these variable loci can incorporate spacers from infecting bacteriophages and then provide immunity against subsequent bacteriophage infections in a sequence-specific manner. Here we show that the *Streptococcus thermophilus* CRISPR1/Cas system can also naturally acquire spacers from a self-replicating plasmid containing an antibiotic-resistance gene, leading to plasmid loss. Acquired spacers that match antibiotic-resistance genes provide a novel means to naturally select bacteria that cannot uptake and disseminate such genes. We also provide *in vivo* evidence that the CRISPR1/Cas system specifically cleaves plasmid and bacteriophage double-stranded DNA within the proto-spacer, at specific sites. Our data show that the CRISPR/Cas immune system is remarkably adapted to cleave invading DNA rapidly and has the potential for exploitation to generate safer microbial strains.

Clustered regularly interspaced short palindromic repeats (CRISPR) loci were discovered in 1987 in *Escherichia coli*[1]. However, the interest in these genetic elements increased in the early 2000s, as they were identified, along with many CRISPR-associated (Cas) proteins, in several prokaryotes[2,3]. Recently, it was shown that the short spacers (21–72 base pairs (bp)) between these repeats originated from extrachromosomal DNA[4–7]. Most importantly, it was experimentally demonstrated that those short spacers can provide resistance against bacteriophage infection and plasmid transformation[8–10].

The CRISPR/Cas immune systems act in at least two general steps: (1) the adaptation stage, where new spacers derived from foreign DNA (proto-spacers) are generally acquired at the leader end of the CRISPR locus[11,12]; and (2) the interference stage, where the CRISPR/Cas system targets either invading DNA[10] or RNA[13]. The mechanistic details of spacer acquisition are still unknown, but a clearer picture is emerging for the interference stage, which starts with the transcription of the CRISPR locus from a promoter located within the leader sequence[14,15]. The full-length RNA is subsequently cleaved by a protein or protein complex, generating short CRISPR RNAs (crRNAs)[16–19]. In *Pyrococcus*, Cas proteins use the crRNAs to target foreign RNA by complementarity in a ruler-anchored manner[13]. However, the *in vivo* mechanism of plasmid and viral interference has yet to be determined.

We previously showed that when bacteriophage-sensitive *Streptococcus thermophilus* cells are infected by virulent bacteriophages, a subset of cells (frequency of $<10^{-6}$) naturally diversify into bacteriophage-insensitive mutants through the acquisition of novel spacers derived from the invading bacteriophage genome into CRISPR1 and/or CRISPR3 (refs 9, 11, 20). Here we investigate the *in vivo* activity of the CRISPR/Cas system in *S. thermophilus* against both bacteriophage and plasmid DNA.

CRISPR/Cas affects plasmid stability

In silico analyses previously indicated that spacers may also be derived from *S. thermophilus* plasmids[11]. Because the CRISPR/Cas system was recently demonstrated to interfere with plasmid transfer in *Staphylococcus*[10], we put forward the hypothesis that the CRISPR/Cas system is responsible for the scarcity of plasmids in wild-type strains of *S. thermophilus*[21,22]. To test this hypothesis, we first introduced by electroporation the vector pNT1 (ref. 23) into the plasmid-free *S. thermophilus* strain DGCC7710. This vector, which is derived from a native *S. thermophilus* plasmid, replicates via a rolling-circle mode and carries an added chloramphenicol-resistance (*cat*) gene as a selection marker. A representative transformant was grown in liquid medium for about 60 generations in the absence of chloramphenicol and aliquots were screened for antibiotic-sensitive colonies.

A total of 54 colonies (out of 900 tested) became sensitive to chloramphenicol and concomitantly lost pNT1. Thirty of these fifty-four colonies acquired a new spacer-repeat unit in CRISPR1. No spacer was integrated in the three other CRISPR loci of *S. thermophilus* DGCC7710 (ref. 8). Sequence analysis of CRISPR1 in the aforementioned 30 colonies identified 14 different spacers (S43–S56), all of which were homologous to pNT1 sequences (Table 1 and Fig. 1). Plasmid stability assays were also performed using two isogenic DGCC7710 strains in which *cas5* (*csn1*-like) or *cas7* genes associated to CRISPR1 were inactivated before the introduction of pNT1. Plasmid pNT1 was highly stable in the DGCC7710::p*cas5*$^{-}$ mutant as no chloramphenicol-sensitive colonies could be isolated, out of 1,800 screened. Of 170 randomly selected chloramphenicol-resistant colonies, none had acquired a new spacer in CRISPR1. On the other hand, chloramphenicol-sensitive colonies were readily obtained with the strain DGCC7710::p*cas7*$^{-}$, but none of the 200 colonies tested

[1]Département de biochimie, de microbiologie et de bio-informatique, Faculté des sciences et de génie, Groupe de recherche en écologie buccale, Faculté de médecine dentaire, Félix d'Hérelle Reference Center for Bacterial Viruses, Université Laval, Quebec City, Quebec G1V 0A6, Canada. [2]Danisco USA Inc., 3329 Agriculture Drive, Madison, Wisconsin 53716, USA. [3]Danisco France SAS, Boîte Postale 10, F-86220 Dangé-Saint-Romain, France.

Table 1 | Plasmid-interfering mutants obtained during plasmid stability assay.

PIM	Spacer acquired* (5′ to 3′)	PAM†	Gene‡	Position on pNT1	Strand	Transformation with pNT1§ (CFU per μg DNA)	Number of times isolated‖
DGCC7710$_{pNT1}^{+S43}$	GATCAAATAACTAATAAATACCCAGTACTT	TGCAGAAG	-	29	−	-	1
DGCC7710$_{pNT1}^{+S44}$	GACCCCCTTTTAAGTGCCGAGTGCCAAAT	TGAGAAA	*dso*	63	+	-	1
DGCC7710$_{pNT1}^{+S45}$	TATACTTGGGTTAATTATACCGTATGGCAA	AAGAAA	-	436	+	4.4 ± 3.5	1
DGCC7710$_{pNT1}^{+S46}$	TTTCCCAATCTTCTGGAATTGAATCGGGAT	AGAGTAG	*rep*	528	−	3.7 ± 2.6 × 10^2	2
DGCC7710$_{pNT1}^{+S47}$	CATGATCTGCAATAATATTGCAGACCTCGT	CTAGAAT	*rep*	917	−	-	1
DGCC7710$_{pNT1}^{+S48}$	GATGATCTGCAATAATATTGCAGACCTCGT	CTAGAAT	*rep*	918	−	-	2
DGCC7710$_{pNT1}^{+S49}$	CGATGATCTGCAATAATATTGCAGACCTCGT	CTAGAAT	*rep*	919	−	-	1
DGCC7710$_{pNT1}^{+S50}$	AATTTAGTTCCGTCAGTAGATTATGAAACT	GGAGAAG	*rep*	1074	+	-	1
DGCC7710$_{pNT1}^{+S51}$	AAAAGCAATGAGTTACATGGTTGCAAGAAT	GCAGAAA	*mob*	1488	+	-	4
DGCC7710$_{pNT1}^{+S52}$	GCCCCAGCTTACTATCAAGGAGCTTTCACG	GCATAAA	*sso*	1999	−	-	1
DGCC7710$_{pNT1}^{+S53}$	CGCCACAGGTTACTTGCTGTCAAGGAGACC	ATGGAAT	-	2066	+	-	6
DGCC7710$_{pNT1}^{+S54}$	TCGTTTGTTGAACTAATGGGTGCTTTAGTT	GAAGAAT	-	2246	−	-	4
DGCC7710$_{pNT1}^{+S55}$	AGAGTTTTATGATTTATACCTTTCTGATGT	AGAGAAA	*cat*	2717	+	-	3
DGCC7710$_{pNT1}^{+S56}$	TTCTTCAACTAACGGGGCAGGTTAGTGACA	TTAGAAA	-	3114	−	-	2
DGCC7710	-					1.4 ± 0.6 × 10^3	

CFU, colony-forming units; PAM, proto-spacer adjacent motif; PIM, plasmid-interfering mutant.
* Nucleotide mismatch with pNT1 sequence is underlined.
† Nucleotide mismatches with the CRISPR1 proto-spacer adjacent motif (NNAGAAW) are underlined.
‡ The '-' symbol indicates an intergenic region.
§ The '-' symbol indicates that pNT1 could not be electroporated into the plasmid-interfering mutants (<1 CFU per μg DNA), $n \geq 2$.
‖ The number of times that each plasmid-interfering mutant was isolated in the assay.

had acquired a new spacer in CRISPR1, indicating that plasmid loss was probably the result of other mechanisms responsible for plasmid instability[24,25]. These data are consistent with previous findings that *cas5* is involved in the interference stage[9] and its inactivation favours plasmid stability, whereas *cas7* is linked to the spacer acquisition stage[9]. Taken together, these observations indicate that the CRISPR/Cas system causes plasmid loss in *S. thermophilus*.

CRISPR/Cas targets antibiotic-resistance genes

To assess whether the acquired spacers from pNT1 cause plasmid interference in these plasmid-interfering mutants, we tested their propensity for pNT1 reintroduction. We were unable to obtain transformants of the plasmid-interfering mutant that carried a spacer perfectly matching a proto-spacer associated with the consensus proto-spacer adjacent motif NNAGAAW (Table 1). We previously showed that the newly added spacer must be identical to the proto-spacer in the bacteriophage genome to confer bacteriophage resistance, and that a proto-spacer adjacent motif located downstream of the proto-spacer is required for the resistance phenotype[4,9,20,26]. Here, approximately half of the proto-spacer adjacent motif sequences contained one or two nucleotide mismatches. This tolerance for proto-spacer adjacent motif degeneracy could be due to the lower selective pressure for plasmids as compared to bacteriophages.

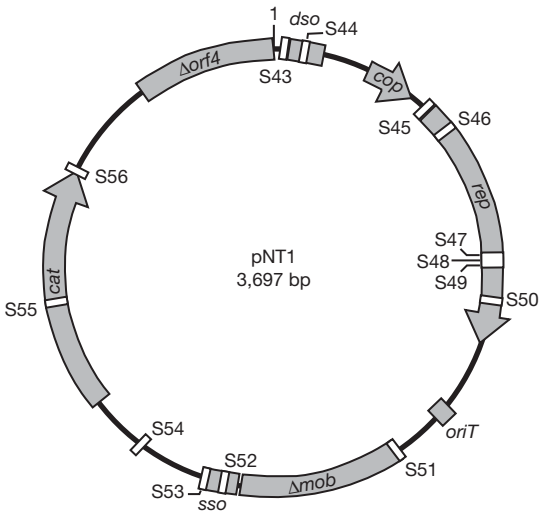

Figure 1 | pNT1 proto-spacers. The numbers outside and inside the map correspond to proto-spacers from the positive and negative strand, respectively.

Some plasmid-interfering mutants targeting proto-spacers associated with non-consensus motifs (NNNAGAAG, NNATAAA, NNGGAAT or NNAGAAG) were also refractory to pNT1 reintroduction. Notably, a plasmid carrying the degenerate motif NNATAAA downstream of proto-spacer S52 could not be transformed into the corresponding plasmid-interfering mutants, whereas bacteriophages carrying the same proto-spacer adjacent motif could infect the matching bacteriophage-insensitive mutants[20] (Supplementary Table 1). It is also worth mentioning that one plasmid-interfering mutant contained a spacer (S47) that matched the last 29 nucleotides (out of 30) of the corresponding proto-spacer in pNT1, indicating that sequence identity at the 5′ end of the spacer might be less important than in the middle or at the 3′ end.

Two plasmid-interfering mutants (DGCC7710$_{pNT1}^{+S55}$ and DGCC7710$_{pNT1}^{+S56}$, also named PIM S55 and PIM S56 for simplicity) carried a spacer targeting the *cat* gene (S55) or its downstream region (S56) (Table 1 and Fig. 1). These spacers rendered the cells untransformable with pNT1 and also with another plasmid (pTRK687) carrying the same *cat* gene. Thus, we conclude that the CRISPR/Cas system provides a simple and natural means to develop a bacterial strain that is refractory to the acquisition of plasmids that carry antibiotic-resistant genes.

Motif degeneracy influences plasmid interference

Two other plasmid-interfering mutants (PIM S45 and PIM S46) carrying a spacer associated with a non-consensus proto-spacer adjacent motif could be re-transformed with pNT1 but at lower frequencies (Table 1). Unexpectedly, the re-transformed PIM S46 visibly carried a linear form of pNT1 (Fig. 2a). *S. thermophilus* strain PIM S46 also contained the usual circular form of the vector, albeit at much lower concentration than the linear form (Fig. 2b), and could still grow in a medium containing chloramphenicol. However, 74–100% of the cells lost pNT1 within 14 generations after growth in a non-selective medium. Under the same conditions, 0–10% of the wild-type colonies lost the plasmid. Therefore, although the NNAGTAG motif was initially permissive for pNT1, the CRISPR/Cas machinery still eliminated the circular and linear plasmid forms within a few generations. The reason for the relative stability of the linear form of the plasmid, under selective pressure, is still unclear.

To assess whether the observed plasmid linearization was the result of CRISPR/Cas activity, the *cas5* and *cas7* genes of PIM S46 were also inactivated and the isogenic strains transformed with pNT1. Plasmid content analysis revealed the presence of only the circular form in PIM S46::p*cas5*$^-$, confirming the involvement of *cas5* in plasmid

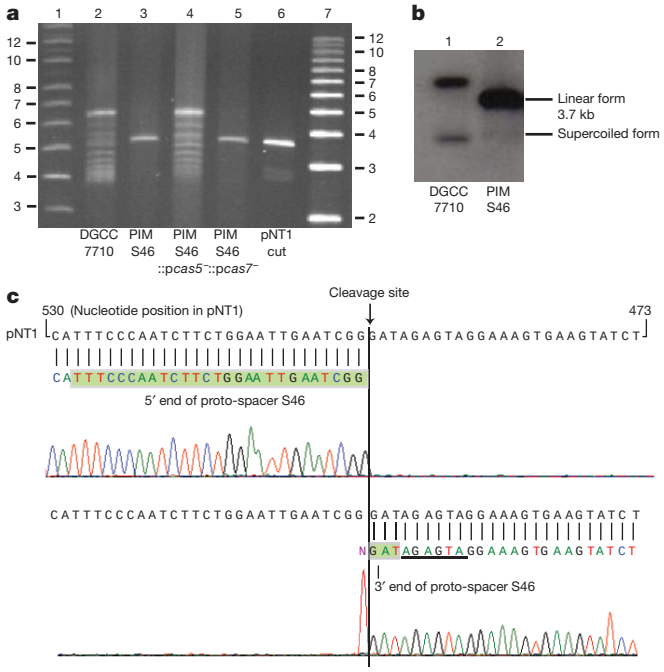

Figure 2 | The CRISPR1/Cas system in *S. thermophilus* targets incoming plasmid DNA. **a**, Plasmid pNT1 in *S. thermophilus* strains. Twenty nanograms of plasmid DNA were loaded per lane. pNT1 plasmid was extracted from wild-type strain DGCC7710 (lane 2), from plasmid-interfering mutant (PIM) strain S46 (lane 3), PIM S46::p*cas5*$^{-}$ (lane 4) and PIM S46::p*cas7*$^{-}$ (lane 5). pNT1 from lane 2 was linearized with EcoRV (lane 6). Lanes 1 and 7: supercoiled and 1-kb DNA ladders (Invitrogen), respectively. **b**, Southern hybridization of plasmid pNT1 in *S. thermophilus* strains. Lane 1: 1 ng of native pNT1 plasmid extracted from wild-type strain DGCC7710. Lane 2: 10 ng of pNT1 extracted from PIM S46. The DIG-labelled probe 6 (Supplementary Table 2) was used. **c**, Direct sequencing electropherograms from primers NT17215_3 (upper part) and NT17225_2 (lower part) (see Methods). The non-templated addition of adenine (T in the reverse complement sequence shown here) at the extremity of the primer NT17225_2 sequence is a sequencing artefact due to the polymerase[35].

interference. On the other hand, pNT1 was still linear in mutant PIM S46::p*cas7*$^{-}$ and in PIM S46 (Fig. 2a).

We next investigated the terminal ends of the pNT1 DNA molecules by directly sequencing the linear plasmid extracted from the PIM S46 strain. We unambiguously determined that pNT1 had been cut once (blunt), within the S46 proto-spacer, after the 27th nucleotide, 3 bases upstream of the proto-spacer adjacent motif (Table 2 and Fig. 2c). We conclude that double-stranded plasmid DNA is targeted and cleaved within the proto-spacer by the CRISPR/Cas machinery.

CRISPR/Cas system cuts viral DNA in the proto-spacers

To test whether the CRISPR/Cas system cleaves other invading DNA within the proto-spacer, we analysed the fate of bacteriophage DNA during the infection of bacteriophage-insensitive mutants containing bacteriophage-derived spacers (Fig. 3a). We used the virulent streptococcal bacteriophage 2972 (ref. 27), its host DGCC7710, and three previously described bacteriophage-insensitive mutants derived from DGCC7710 following challenge with virulent bacteriophages 2972 and/or 858, namely DGCC7710$_{\Phi2972}^{+S4}$, DGCC7710$_{\Phi2972}^{+S7}$ and DGCC7710$_{\Phi2972}^{+S4}{}_{\Phi858}^{+S32}$ (bacteriophage-insensitive mutant (BIM) S4, BIM S7 and BIM S4/S32, respectively)[20]. These bacterial strains were infected with bacteriophage 2972 and total DNA was extracted from bacteriophage-infected cells sampled over time. Southern hybridizations of the DNA throughout the time course of infection of the bacteriophage-insensitive mutants with probes targeting regions close to the proto-spacers revealed that the bacteriophage DNA is rapidly cleaved by the CRISPR/Cas machinery (Fig. 3b, c). The cleavage site in the bacteriophage genome seemed to be within or in the vicinity of the proto-spacer during the infection of bacteriophage-insensitive mutants, whereas no cleavage of bacteriophage DNA was observed in the sensitive strain DGCC7710 (Fig. 3c). The cleavage patterns were similar in the bacteriophage-infected BIMs, even though the S4 and S7 spacers are derived from different DNA strands and transcriptional modules of the bacteriophage genome[28] (Fig. 3a). We also conducted the same experiments with isogenic derivatives of *S. thermophilus* BIM S4, in which the *cas5* and *cas7* genes were independently inactivated. The inactivation of *cas5* restored bacteriophage sensitivity (efficiency of plaquing of 1) and no bacteriophage DNA cleavage occurred, confirming the involvement of *cas5* in interference. In contrast, the inactivation of *cas7* did not affect bacteriophage resistance (efficiency of plaquing of 10^{-6}) and bacteriophage DNA was cleaved (Fig. 3c).

We also investigated the fate of bacteriophage DNA during the infection of *S. thermophilus* BIM S4/S32, which contains two new spacers that target bacteriophage 2972. The bacteriophage genome was cleaved within each proto-spacer in this infected bacteriophage-insensitive mutant (Fig. 3c). The 5.2-kb and 5.8-kb bands corresponded to a cleavage at the S4 or S32 proto-spacers, respectively, whereas the 1.4-kb band resulted from cleavage at both sites. This is consistent with the previous observation that resistance increases with the number of spacers acquired[20]. Our results show that the *in vivo* target of the *S. thermophilus* CRISPR1/Cas system is DNA for both plasmid (Fig. 2) and bacteriophage (Fig. 3).

We then wanted to determine whether the bacteriophage DNA cleavage site was identical to that of the plasmid. Contrary to the method used for the determination of the cleavage site of plasmid pNT1, the low amount of cleaved bacteriophage DNA in the infected bacteriophage-insensitive mutants rendered direct sequencing of the extremities impossible. An alternative sequencing method was used (see Supplementary Methods and Supplementary Fig. 1), as the cleaved 1.4-kb fragment (from *S. thermophilus* BIM S4/S32) was subjected to inverse PCR to obtain a product that joins the ends of the S4- and S32-cleaved proto-spacers. Sequence analysis of the amplicon revealed that the bacteriophage fragment had indeed been cleaved within S4 and S32. For the other cleavage sites, the extracted bacteriophage DNA was digested *in vitro* by restriction endonucleases producing blunt ends

Table 2 | Cleavage sites within the proto-spacers in bacteriophage 2972 or plasmid pNT1

BIM or PIM	Spacer (size)	Homology (position)	Strand	Proto-spacer* (5′ to 3′)	PAM†
DGCC7710$_{\Phi2972}^{+S4}$	S4 (30 nt)	Bacteriophage 2972 (31582)	+	CTCAGTCGTT↓ACTGGTGAACCAGTTTC↓AAT	TGAGAAA
DGCC7710$_{\Phi2972}^{+S4}{}_{\Phi858}^{+S32}$	S4 (30 nt)	Bacteriophage 2972 (31582)	+	CTCAGTCGTT↓ACTGGTGAACCAGTTTC↓AAT	TGAGAAA
DGCC7710$_{\Phi2972}^{+S4}{}_{\Phi858}^{+S32}$	S32 (30 nt)	Bacteriophage 2972 (33044)	+	ATTGTCTATTA↓CGACAACATGGAAGAT↓GAT	GTAGAAA
DGCC7710$_{\Phi2972}^{+S40}$	S40 (29 nt)	Bacteriophage 2972 (31583)	+	TCAGTCGTT↓ACTGGTGAACCAGTTTC↓AAT	TGAGAAA
DGCC7710$_{\Phi2972}^{+S7}$	S7 (30 nt)	Bacteriophage 2972 (10299)	−	AAGCAAGTTGATATATTTCTCTTTCTT↓TAT	TAAGAAA
DGCC7710$_{\Phi2972}^{+S41}$	S41 (30 nt)	Bacteriophage 2972 (31518)	−	TTCCCTTCGATAATGGCAAGACCGAAA↓CGT	TCAGAAA
DGCC7710$_{\Phi2972}^{+S42}$	S42 (30 nt)	Bacteriophage 2972 (31084)	−	ATATTCATATTCCCTGCTCATGTTTGA↓TAG	CAAGAAT
DGCC7710$_{pNT1}^{+S46}$	S46 (30 nt)	Plasmid pNT1 (528)	−	TTTCCCAATCTTCTGGAATTGAATCGG↓GAT	AGAGTAG

BIM, bacteriophage-insensitive mutant; nt, nucleotide; PAM, proto-spacer adjacent motif; PIM, plasmid-interfering mutant.
* ↓ Indicates the cleavage site
† The mismatches with the consensus proto-spacer adjacent motif (NNAGAAW) are underlined.

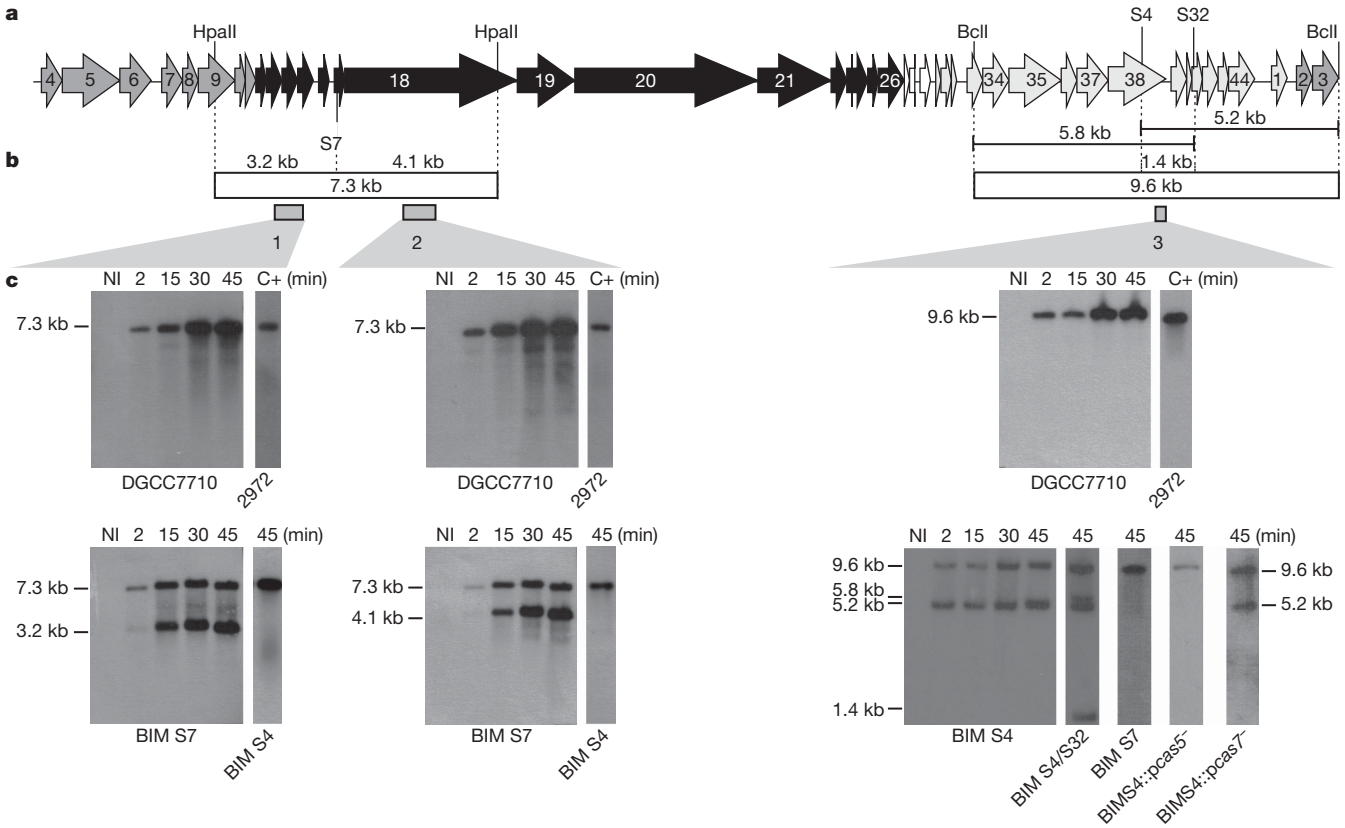

Figure 3 | The CRISPR1/Cas system targets bacteriophage DNA. **a**, Bacteriophage 2972 genome and position of proto-spacers. Light grey, dark grey and black arrows indicate early, middle and late transcription module, respectively[28]. Proto-spacers above and below the genome indicate positive and negative strands, respectively. **b**, Restriction fragments detected by hybridization and position of probes (1, 2 and 3). **c**, Southern blots of bacteriophage-infected strains DGCC7710, BIM S4, BIM S7, BIM S4/S32, BIM S4::p*cas5*$^{-}$ and BIM S4::p*cas7*$^{-}$ over 45 min (conducted at 2 min, 15 min, 30 min and 45 min). NI, non-infected strain. C+, positive control, 10 ng of digested DNA from bacteriophage 2972. One microgram of total DNA from DGCC7710 and 10 μg of bacteriophage-insensitive mutant per lane.

and cutting upstream and downstream of each proto-spacer (see Supplementary Figs 1 and 2). Fragments comprising the cleaved proto-spacer at one extremity and a blunt site at the other were ligated and inverse PCRs were performed as described previously. PCR products were obtained for every fragment and sequenced.

Analysis of the amplicons revealed that all of the cleavage sites were within the proto-spacers (S4, S7 and S32). In every cleaved proto-spacer analysed, one cleavage site was located after the 27th nucleotide, 3 bases upstream of the proto-spacer adjacent motif, as observed for the linearized plasmid in PIM S46 (Table 2). Curiously, a second cleavage site was detected within the proto-spacers but only for the spacers targeting the positive strand of the bacteriophage genome (see S32 and S4 in Table 2). This second site was located 19 or 20 nucleotides upstream of the proto-spacer adjacent motif. Notably, the plasmid proto-spacer S46 described above was on the negative strand and was cut only once (Table 2). To confirm these results, three additional bacteriophage-infected BIMs that have acquired different spacers (S40, S41, or S42) were analysed similarly. The single (negative strand) and double (positive strand) cleavage sites within the proto-spacer were confirmed in these three bacteriophage-insensitive mutants (Table 2). Of note, the two cleavage sites within the S40 proto-spacer (a natural 5′-end truncated version of the S4 proto-spacer with 29 nucleotides instead of 30) were at the same position, indicating that the CRISPR1/Cas system in *S. thermophilus* probably acts in a 3′-end ruler-anchored manner. Ligation and amplification of each cleaved fragment could be obtained and no nucleotide was missing in the S7, S41, S42 and S46 proto-spacer sequences, confirming the blunt-end cleavage activity of *S. thermophilus* CRISPR1/Cas system.

Overall, we have established that the *S. thermophilus* CRISPR1/Cas system cleaves both bacteriophage and plasmid DNA *in vivo*. This endonuclease activity, which seems to require *cas*5, is proto-spacer specific and orientation dependent. In that regard, the *S. thermophilus* CRISPR1/Cas interference mechanism differs from that recently described in *Pyrococcus furiosus*, whereby RNA is the target[13] but corroborates the previous work on *Staphylococcus* which identified DNA as the target[10]. Furthermore, we show that the DNA cleavage activity is responsible for both bacteriophage resistance in bacteriophage-insensitive mutants and plasmid instability in plasmid-interfering mutants. The antagonism between plasmid maintenance and CRISPR/Cas activity probably explains the natural scarcity of plasmids in *S. thermophilus*. The CRISPR/Cas immune system can be used to naturally generate safer and more robust organisms with increased bacteriophage resistance and that are able to interfere with the dissemination of plasmids that carry antibiotic-resistance markers.

METHODS SUMMARY

Microbial conditions. *S. thermophilus* strains were grown at 37 °C or 42 °C in LM17 medium (ref. 20). Strains were infected with CsCl-purified bacteriophage 2972 at a multiplicity of infection of 5 (ref. 29). *cas5* and *cas7* genes were inactivated as described[9]. The identity of *S. thermophilus* strains was confirmed by sequencing the CRISPR loci[8] and by bacteriophage assays[20].

Southern hybridization. DNA was extracted from bacterial cells as described[30]. DNA probes for Southern blot hybridizations[29] were prepared with the PCR DIG labelling mix (Roche) (Supplementary Table 2). Hybridization and detection (CDP-star) were performed as recommended by Roche.

Plasmid stability. *S. thermophilus* DGCC7710 (or its *cas5*$^{-}$ and *cas7*$^{-}$ derivatives) transformed with pNT1 (ref. 23) was used to inoculate 10 ml of LM17. One-hundred microlitres of the previous culture was inoculated into 10 ml of fresh LM17 medium every morning (42 °C) and night (37 °C) for 5 days, for a total of 9 inoculations. For each culture, 100 colonies were screened for chloramphenicol sensitivity. The CRISPR1 of the chloramphenicol-sensitive clones was analysed.

Plasmid-interfering mutants with new spacers were electroporated[31] with ~600 ng of pNT1 or pTRK687 (ref. 32). PIM S55 and PIM S56 were also transformed with pLS1 (ref. 33). Plasmid content of re-transformed PIM S45 and PIM S46 was verified by extraction (Qiagen) and extremities of the linearized plasmid from PIM S46 were sequenced.

Cleavage site of bacteriophage DNA. Total DNA from bacteriophage-infected bacteriophage-insensitive mutants was extracted[30] and digested using endonucleases generating blunt ends (Supplementary Fig. 2). After migration, agarose bands corresponding to both sides of the cleaved proto-spacers were recovered, DNA eluted and used for overnight intramolecular ligation at 16 °C, with T4 DNA ligase (Invitrogen). Ligation product was used as a template for inverse PCR[34] (Supplementary Figs 1 and 2 and Supplementary Table 3) and amplicons were sequenced (primers in Supplementary Table 3).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 April; accepted 20 September 2010.

1. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* **169,** 5429–5433 (1987).
2. Jansen, R., van Embden, J. D. A., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43,** 1565–1575 (2002).
3. Mojica, F. J., Díez-Villaseñor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* **36,** 244–246 (2000).
4. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151,** 2551–2561 (2005).
5. Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60,** 174–182 (2005).
6. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151,** 653–663 (2005).
7. Andersson, A. F. & Banfield, J. F. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320,** 1047–1050 (2008).
8. Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* **327,** 167–170 (2010).
9. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712 (2007).
10. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322,** 1843–1845 (2008).
11. Horvath, P. *et al.* Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* **190,** 1401–1412 (2008).
12. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Rev. Genet.* **11,** 181–190 (2010).
13. Hale, C. R. *et al.* RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139,** 945–956 (2009).
14. Lillestøl, R. K. *et al.* CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. *Mol. Microbiol.* **72,** 259–272 (2009).
15. Pul, U. *et al.* Identification and characterization of *E. coli* CRISPR-*cas* promoters and their silencing by H-NS. *Mol. Microbiol.* **75,** 1495–1512 (2010).
16. Tang, T. H. *et al.* Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc. Natl Acad. Sci. USA* **99,** 7536–7541 (2002).
17. Carte, J., Wang, R. Y., Li, H., Terns, R. M. & Terns, M. P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22,** 3489–3496 (2008).
18. Hale, C., Kleppe, K., Terns, R. M. & Terns, M. P. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* **14,** 2572–2579 (2008).
19. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321,** 960–964 (2008).
20. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190,** 1390–1400 (2008).
21. Turgeon, N. & Moineau, S. Isolation and characterization of a *Streptococcus thermophilus* plasmid closely related to the pMV158 family. *Plasmid* **45,** 171–183 (2001).
22. Somkuti, G. A. & Steinberg, D. H. Distribution and analysis of plasmids in *Streptococcus thermophilus*. *J. Ind. Microbiol.* **1,** 157–163 (1986).
23. Vaillancourt, K. *et al.* Role of *galK* and *galM* in galactose metabolism by *Streptococcus thermophilus*. *Appl. Environ. Microbiol.* **74,** 1264–1267 (2008).
24. Girard, S. L. & Moineau, S. Analysis of two theta-replicating plasmids of *Streptococcus thermophilus*. *Plasmid* **58,** 174–181 (2007).
25. Kiewiet, R., Kok, J., Seegers, J. F., Venema, G. & Bron, S. The mode of replication is a major factor in segregational plasmid instability in *Lactococcus lactis*. *Appl. Environ. Microbiol.* **59,** 358–364 (1993).
26. Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155,** 733–740 (2009).
27. Lévesque, C. *et al.* Genomic organization and molecular analysis of virulent bacteriophage 2972 infecting an exopolysaccharide-producing *Streptococcus thermophilus* strain. *Appl. Environ. Microbiol.* **71,** 4057–4068 (2005).
28. Duplessis, M., Russell, W. M., Romero, D. A. & Moineau, S. Global gene expression analysis of two *Streptococcus thermophilus* bacteriophages using DNA microarray. *Virology* **340,** 192–208 (2005).
29. Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 2001).
30. Fortier, L.-C. & Moineau, S. Morphological and genetic diversity of temperate phages in *Clostridium difficile*. *Appl. Environ. Microbiol.* **73,** 7358–7366 (2007).
31. Buckley, N. D., Vadeboncoeur, C., LeBlanc, D. J., Lee, L. N. & Frenette, M. An effective strategy, applicable to *Streptococcus salivarius* and related bacteria, to enhance or confer electroporation competence. *Appl. Environ. Microbiol.* **65,** 3800–3804 (1999).
32. Sturino, J. M. & Klaenhammer, T. R. Expression of antisense RNA targeted against *Streptococcus thermophilus* bacteriophages. *Appl. Environ. Microbiol.* **68,** 588–596 (2002).
33. Nieto, C., Fernández de Palencia, P., López, P. & Espinosa, M. Construction of a tightly regulated plasmid vector for *Streptococcus pneumoniae*: controlled expression of the green fluorescent protein. *Plasmid* **43,** 205–213 (2000).
34. Ochman, H., Gerber, A. S. & Hartl, D. L. Genetic applications of an inverse polymerase chain reaction. *Genetics* **120,** 621–623 (1988).
35. Clark, J. M. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **16,** 9677–9686 (1988).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. R. Klaenhammer for pTRK687, B. Conway for editorial assistance, D. Tremblay, A. Fleury and B. Stahl for technical assistance. M.-E.D. is the recipient of a Novalait/FQRNT graduate scholarship and A.H.M. of a Clarín/FICYT postdoctoral scholarship. S.M. acknowledges funding from NSERC (Discovery program) and CIHR (Team Grant–Emerging: Novel Alternatives to Antibiotics).

Author Contributions S.M. conceived and headed the project. All the authors participated in the design of the study. J.E.G., M.-E.D., M.V. and A.H.M. performed the experiments. J.E.G. and S.M. wrote the main parts of the manuscript. All of the authors analysed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.M. (Sylvain.Moineau@bcm.ulaval.ca).

METHODS

Bacterial, bacteriophage and culture conditions. *Streptococcus thermophilus* strain DGCC7710 (ref. 9) and bacteriophage-insensitive mutants (BIMs) or plasmid-interfering mutants (PIMs) were grown in M17 broth supplemented with 0.5% lactose (LM17) at 37 °C or 42 °C. The identity of *S. thermophilus* strains was confirmed by sequencing the CRISPR loci[8]. Bacteriophage sensitivity of the isolated bacteriophage-insensitive mutants was estimated as reported previously[36]. Bacteriophages were propagated in LM17 supplemented with 10 mM $CaCl_2$. Purified bacteriophage preparations (10^{11}–10^{12} PFU ml^{-1}) were obtained by ultracentrifugation using a discontinuous CsCl gradient[29]. *cas5* and *cas7* genes were inactivated as described previously[9] in three *S. thermophilus* strains: DGCC7710, BIM S4 and PIM S46. All $cas5^-$ and $cas7^-$ derivatives were grown under erythromycin selective pressure. Efficiency of plaquing was determined as reported[20].

Bacteriophage infection. Bacterial strains were incubated at 42 °C in 100 ml of LM17 to an optical density of 0.5 at 600 nm. Cultures were concentrated by centrifugation and pellets were re-suspended in 10 ml of pre-incubated LM17 medium containing 10 mM $CaCl_2$. After removing a 1-ml uninfected sample, each bacterial culture was infected with purified bacteriophage 2972 at a multiplicity of infection (MOI) of 5, and incubated at 42 °C. Samples were taken after 2, 15, 30 and 45 min, centrifuged for 15 s at 16,000*g* and pellets were flash frozen, and stored at −80 °C until DNA extraction. The latent period of bacteriophage 2972 was previously determined to be 34 min on *S. thermophilus* DGCC7710. The maximum burst of bacteriophages is 40 min after the beginning of the infection[28].

DNA extraction and Southern hybridization. Total DNA extractions were performed as described elsewhere[30], with these modifications: 25% (w/v) sucrose and 60 mg ml^{-1} lysozyme were used as well as only one phenol-chloroform extraction. Dissolved DNA solutions were treated with 1 μg of RNase, incubated for 15 min at 37 °C, and their concentration determined using the Nanodrop 2000 spectrophotometer (Thermo Scientific).

Ten micrograms of bacteriophage-insensitive mutant DNA and 1 μg of wild-type DNA were digested with either HpaII or BclI (Roche). As controls, 10 μg of bacterial DNA and 10 ng of bacteriophage DNA were digested and were used for probe hybridizations. In addition, 500 ng of bacteriophage DNA was used as a digestion and transfer control. DNA fragments were separated in a 0.8% agarose gel in 1× TAE buffer, stained with EZ-Vision DNA dye (Amresco) and photographed under ultraviolet light. DNA was transferred onto positively charged nylon membranes (Roche) by capillary blotting[29]. Membranes were stained for 45 s with methylene blue, rinsed with distilled water, and dried in ambient air. DNA probes were prepared with PCR DIG labelling mix (Roche) (Supplementary Table 2). Pre-hybridization, hybridization, washes and detection (CDP-star) were performed as recommended by Roche.

Plasmid stability and plasmid-interfering mutant analysis. *S. thermophilus* DGCC7710 (or the $cas5^-$ and $cas7^-$ isogenic derivatives) transformed with pNT1 plasmid[23] (GenBank accession number HQ010044) was picked from a LM17 plate containing 5 μg ml^{-1} of chloramphenicol. Plasmid and CRISPR contents of this transformant were confirmed by sequencing and used to inoculate 10 ml of LM17. One-hundred microlitres of the previous culture was inoculated into 10 ml of fresh LM17 medium every morning and night for 5 days, for a total of 9 inoculations. Cultures were grown at 37 °C for overnight incubations and at 42 °C for 8 h during the day. Finally, these cultures were serially diluted and plated on LM17. For each culture, 100 colonies were screened for the absence of the antibiotic resistance phenotype. Plasmid stability of pNT1 was analysed by comparing the number of colonies appearing on selective agar (5 μg ml^{-1} of chloramphenicol) versus the number of colonies on LM17 agar. The CRISPR1 of chloramphenicol-sensitive clones was verified: the 5′ end of the CRISPR1 of the plasmid-interfering mutants or bacteriophage-insensitive mutants was amplified by PCR with the primers yc70 (5′-TGCTGAGACAACCTAGTCTCTC-3′)[4] and RDS7revBamHI (5′-GGATCCG GATCCGTTGAGGCCTTGTTC-3′), and sequenced using the same primers. Fourteen different plasmid-interfering mutants having acquired a new spacer from pNT1 were electroporated as described elsewhere[31] with ~600 ng of pNT1 or pTRK687 (ref. 32) vectors. All plasmid-interfering mutants (except for PIM S55 and PIM S56) could be transformed with the control vector pTRK687, which is carrying a chloramphenicol-resistance gene as selection marker. In the case of PIM S55 and PIM S56, they could be transformed with pLS1, which is carrying an erythromycin-resistance gene as a selection marker[33].

The proto-spacer region of pNT1 isolated from the pNT1-transformed PIM S45 was sequenced using the primers pNT1_104 (5′-GTGCCTTGAACCTTAGA GCCACAA-3′) and NT17215_3 (5′-GTTCAGAGTATGGACTGCCG-3′). Chloramphenicol-resistant clones were also checked for spacer acquisition. Their CRISPR1 was amplified by PCR using the primers yc70 and RDS7revBamHI cited above.

Plasmid extraction and plasmid cleavage site determination. The plasmid content of *S. thermophilus* strains was isolated using QIAquick Spin Miniprep columns (Qiagen). The QIAquick protocol was modified by treating the cells with P1 buffer containing sucrose (25%) and lysozyme (60 mg ml^{-1}), and incubating at 37 °C for 15 to 30 min. Extremities of the linearized plasmid were directly sequenced with the primers NT17225_2 (5′-TACGTACTTGTGTTACTA TTG-3′) and NT17215_3 (5′-CGGCAGTCCATACTCTGAAC-3′).

Bacteriophage cleavage-site determination. Each bacteriophage-insensitive mutant was independently infected for 45 min with bacteriophage 2972 at a multiplicity of infection of 5. Total DNA from bacteriophage-infected bacteriophage-insensitive mutant was extracted[30] and digested using endonucleases cutting upstream and downstream of each proto-spacer and generating blunt ends (Roche and NEB) (Supplementary Fig. 2). DNA was migrated on a 0.8% agarose gel. Bands corresponding to the left and right fragments of the cleaved proto-spacer were extracted from the gel (band sizes estimated according to cleavage within the proto-spacer). DNA was extracted using the QIAEX II Gel Extraction Kit (Qiagen). Each fragment comprised the cleaved proto-spacer at one extremity and a blunt site at the other. Total eluted product (50 μl) of each fragment was used for overnight intramolecular ligation at 16 °C, with T4 DNA ligase (Invitrogen). In the case of BIM S4/S32, the fragment delimited by the cleaved proto-spacers S4 and S32 could be directly used for intramolecular ligation. Five microlitres of the ligation product were used as a template for PCR (100 μl reaction volume), with Taq Polymerase (Roche) (Supplementary Table 3). The cycling conditions were: 94 °C/45 s, 55 °C/45 s, 72 °C/1 min for 35 cycles with a final extension of 72 °C/10 min[34]. PCR products were sequenced using primers listed in Supplementary Table 3.

Sequences. All sequencing reactions were performed by the 'Plateforme de Séquençage et de Génotypage des Génomes' service from the CHUL-CHUQ Research Center. The sequences from both strands were analysed using BioEdit software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) or Staden package (http://staden.sourceforge.net/staden_home.html).

36. Moineau, S. *et al.* Characterization of lactococcal bacteriophages from Québec cheese plants. *Can. J. Microbiol.* **38,** 875–882 (1992).