



Communication

Ectopic Spacer Acquisition in *Streptococcus thermophilus* CRISPR3 Array

Rodrigo Achigar ¹, Martina Scarrone ², Geneviève M. Rousseau ^{2,3}, Cécile Philippe ^{2,3}, Felipe Machado ¹, Valentina Duvós ¹, María Pía Campot ¹, Moïra B. Dion ^{2,3}, Yuyu Shao ^{2,3}, María Julia Pianzzola ⁴ and Sylvain Moineau ^{2,3,5,*}

- ¹ Laboratorio de Biotecnología, Facultad de Ingeniería, Universidad ORT Uruguay, Montevideo 11100, Uruguay; rodrigoachigar@gmail.com (R.A.); felipe@enteria.uy (F.M.); valentina@enteria.uy (V.D.); pia@enteria.uy (M.P.C.)
- ² Département de Biochimie, De Microbiologie et de Bio-Informatique, Faculté des Sciences et de Génie, Université Laval, Québec, QC G1V 0A6, Canada; martina.scarrone96@gmail.com (M.S.); genevieve.rousseau@greb.ulaval.ca (G.M.R.); cecilephilippe@gmail.com (C.P.); moira.dion.1@ulaval.ca (M.B.D.); yuyu.shao.1@ulaval.ca (Y.S.)
- ³ Groupe de Recherche en Écologie Buccale, Faculté de Médecine Dentaire, Université Laval, Québec, QC G1V 0A6, Canada
- ⁴ Departamento de Biociencias (DEPBIO), Facultad de Química, Universidad de la República, Montevideo 11800, Uruguay; mpianzzo@gmail.com
- ⁵ Félix d'Hérelle Reference Center for Bacterial Viruses, Faculté de Médecine Dentaire, Université Laval, Québec, QC G1V 0A6, Canada
- * Correspondence: Sylvain.Moineau@bcm.ulaval.ca; Tel.: +1-418-656-3712



Citation: Achigar, R.; Scarrone, M.; Rousseau, G.M.; Philippe, C.; Machado, F.; Duvós, V.; Campot, M.P.; Dion, M.B.; Shao, Y.; Pianzzola, M.J.; et al. Ectopic Spacer Acquisition in *Streptococcus thermophilus* CRISPR3 Array. *Microorganisms* **2021**, *9*, 512. <https://doi.org/10.3390/microorganisms9030512>

Academic Editor: Jennifer Mahony

Received: 3 February 2021

Accepted: 24 February 2021

Published: 1 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: *Streptococcus thermophilus* relies heavily on two type II-A CRISPR-Cas systems, CRISPR1 and CRISPR3, to resist siphophage infections. One hallmark of these systems is the integration of a new spacer at the 5' end of the CRISPR arrays following phage infection. However, we have previously shown that ectopic acquisition of spacers can occur within the CRISPR1 array. Here, we present evidence of the acquisition of new spacers within the array of CRISPR3 of *S. thermophilus*. The analysis of randomly selected bacteriophage-insensitive mutants of the strain Uy01 obtained after phage infection, as well as the comparison with other *S. thermophilus* strains with similar CRISPR3 content, showed that a specific spacer within the array could be responsible for misguiding the adaptation complex. These results also indicate that while the vast majority of new spacers are added at the 5' end of the CRISPR array, ectopic spacer acquisition is a common feature of both CRISPR1 and CRISPR3 systems in *S. thermophilus*, and it can still provide phage resistance. Ectopic spacer acquisition also appears to have occurred naturally in some strains of *Streptococcus pyogenes*, suggesting that it is a general phenomenon, at least in type II-A systems.

Keywords: *Streptococcus thermophilus*; *Streptococcus pyogenes*; CRISPR-Cas; adaptation; phages

1. Introduction

Streptococcus thermophilus is a lactic acid bacterium used extensively for the manufacture of several fermented dairy products, such as yogurt and several cheeses [1–4]. Siphophage infections of these Gram-positive bacteria are the leading cause of milk fermentation failures worldwide [5–8]. One important strategy for controlling virulent phages in industrial settings is to select and use natural bacteriophage-insensitive mutant (BIM) strains as starter cultures. It is now well-documented that *S. thermophilus* strains rely on the CRISPR-Cas system (a prokaryotic adaptive immune system) to protect itself against phage attacks [9–15]. This system is composed of a clustered regularly interspaced short palindromic repeats (CRISPR) array and its associated *cas* genes [16–18]. These *cas* genes encode Cas proteins, some of which are used by the bacteria to acquire new immunities by integrating short DNA sequences, called spacers, from invading DNA, such as phage genomes, at

the leader 5' end of the CRISPR array. Then, the CRISPR array is transcribed and processed into small interfering RNAs (called crRNAs) [19,20]. These crRNAs form ribonucleoprotein complexes with Cas proteins and destroy invading DNA through base-pair recognition and cleavage [21–23].

CRISPR-Cas systems are currently divided into two classes, six types, and several subtypes [24]. For most of them, the integration of new spacers is mostly driven by two proteins, Cas1 and Cas2, which form an integrase-like complex [25]. During the adaptation process, the first repeat of the CRISPR array is also duplicated upstream of the newly acquired spacer [26]. *S. thermophilus* strains have been shown to possess up to four different CRISPR-Cas systems: two distinct type II-A systems, one type I-E, and one type III-A [22]. So far, only the two type II-A systems, CRISPR1 (CR1) and CRISPR3 (CR3), appear to be active in spacer acquisition [27,28], with the large majority of the acquisition events taking place in the CR1 array [29], which is the most predominant system in this bacterial species.

Type II-A CRISPR-Cas systems are composed of four genes coding for Cas9, Cas1, Cas2, and Csn2 [30]. All four are required for the adaptation step *in vivo*, which implies probable interactions between them [31]. For type II CRISPR-Cas systems, the signature gene is *cas9*, which encodes a multidomain protein that combines the functions of the crRNA–effector complex, targets DNA cleavage, and contributes to the selection of new spacers during the adaptation stage [30]. Cas9 recognizes short motifs called protospacer adjacent motifs (PAMs) within the invading DNA. When a PAM is recognized, the adjacent protospacer sequence can be integrated into the CRISPR array as a new spacer. Csn2 also appears to be critical in the adaptation stage as a *csn2* gene insertion mutant was found to be incapable of acquiring new spacers in response to phage infection [3,11]. Of note, the two type II-A systems of *S. thermophilus* use a different PAM [21,22,27].

Cryoelectron microscopy analyses of the Cas1-Cas2-Csn2 complex from the *S. thermophilus* CR3 system showed a large multi-subunit complex (Cas1₈-Cas2₄-Csn2₈) with a channel occupied by (approx. 30 bp) double-stranded DNA, also suggesting a protective role for the complex. Spacer adaptation complexes may have quite different architectures, but the speculated model for spacer capture suggest that: (1) the Cas1-Cas2-Csn2 complex engages free DNA ends from invading dsDNA phage genome and encircles it within the complex, (2) the Cas1-Cas2-Csn2 complex slide on the DNA until it encounters Cas9 that is bound to a PAM, (3) DNA is cleaved, releasing Cas9 and the Cas1-Cas2-Csn2 complex, encapsulating the DNA as a new spacer ready for integration [31].

One of the hallmarks of the acquisition of new immunities is that novel spacers are typically integrated at the leader-proximal region of the CRISPR array [32]. This polarity is guided by the leader sequence upstream of the CRISPR array [29]. It is likely that sequences within the leader elements of CRISPR loci are important given that novel spacers are introduced adjacent to the leader in several systems. It was reported that the integrity of the 3' end of the leader sequence (called leader anchoring sequence or LAS) is crucial for the polarized acquisition of new spacers [33]. It has been also shown that Cas1 contains a DNA-binding region that binds this leader DNA [34–36]. This sequence (most notably the 5'-GAG-3' at the 3' end) is highly conserved across type II systems. In the absence of an appropriate LAS, other short nucleotide sequences within the *Streptococcus pyogenes* type II-A associated CRISPR array were shown to guide ectopic (acquired at positions other than the 5' end) spacer integration in the heterologous host *Staphylococcus aureus* [33]. Ectopic spacer acquisition was also recently observed in the CRISPR array associated with the type II-A system of *Streptococcus mutans* [37].

In a previous study, we observed ectopic spacer acquisition in the CR1 array of *S. thermophilus* BIMs, obtained after a phage-sensitive host was exposed to virulent phages of the *Siphoviridae* family [38]. In this study, we investigated whether ectopic spacer acquisition could also occur following a phage infection in the second active type II-A CRISPR-Cas system of *S. thermophilus*, namely CR3, as well as determining if the presence of a LAS sequence was needed.

2. Materials and Methods

2.1. Phage and Bacterial Strains

Virulent siphophage 53 (*cos*-type) was previously isolated from a failed mozzarella production in Uruguay and its complete genome is available (accession no. KT717084) [38]. Phage-sensitive *S. thermophilus* strains (Uy series) were obtained from a local starter culture supplier and grown in LM17 medium at 42 °C. Phages and bacterial strains were stored, as frozen stocks, in LM17 supplemented with 15% (v/v) glycerol. Phage 53 was amplified in LM17 supplemented with 10 mM CaCl₂ (LM17-CaCl₂). Briefly, 0.1 mL of a fresh bacterial culture (OD₆₀₀ = 0.6) was inoculated in 10 mL of broth and incubated at 42 °C for three hours. Then, 0.1 mL of phage lysate was added and incubated at 42 °C until complete lysis was observed. The lysate was filtered (0.45 µm filters) and stored at 4 °C until used. Phage titer was determined using methods described elsewhere [39].

2.2. *S. thermophilus* and *S. pyogenes* CRISPR Loci Analysis

The CR3 loci of 50 strains of the Uy collection were amplified by PCR (NEB Phusion High-Fidelity DNA Polymerase) and sequenced using primers CR3-fwd (5-CTGAGATTAA TAGTGCGATTACG-3) and CR3-rev (5-GCTGGATATTCGTATAACATGTC-3) [27]. Bioinformatics analyses to identify the CR3 spacer content of 45 Uy strains were firstly performed with SnapGene (version 4.1.9). In addition, 39 CR3 loci were retrieved from 64 complete *S. thermophilus* genomes available in GenBank as of October 2020. Similarly, type II-A CRISPR loci were searched from *S. pyogenes* genomes. A set of 213 complete and circular genomes was available from GenBank as of October 2020. All type II-A loci (116 for *S. pyogenes* and 84 for *S. thermophilus*) were identified with CRISPRDetect (version 2.2.3, <http://crispr.otago.ac.nz/CRISPRDetect>, accessed on 1 March 2021) using default parameters [40]. The putative CRISPRs were also manually checked. CRISPRDetect gff output files were then used with the Python script CRISPRStudio (version 1) [41] to extract, align and cluster the spacer sequences to generate a SVG file for CRISPR loci representation. Output SVG files were edited manually using Illustrator 2020 or Inkscape 1.0.0.

2.3. Bacteriophage Insensitive Mutants (BIMs)

BIMs were obtained by infecting the phage-sensitive strain *S. thermophilus* Uy01 with the virulent phage 53. Briefly, approximately 5×10^8 CFU of *S. thermophilus* were mixed with 1×10^8 PFU of phages in 4 mL of soft LM17-CaCl₂ (0.75% agar) and poured on a LM17-CaCl₂ agar plate (1.5% agar). Plates were incubated at least 48 h at 42 °C. Individual colonies were recovered, streaked and re-streaked for purity. Individual colonies were then inoculated and incubated overnight in LM17 broth. The cultures were then tested for phage resistance as described elsewhere [10].

3. Results

3.1. Evidence of Ectopic Acquisition Events in CRISPR3

The analysis of the spacer content of the CR3 array was performed on 64 complete public *S. thermophilus* genomes available at the time of the study in GenBank along with 50 strains of the Uy collection. A total of 39 (60.9%) CR3 loci were retrieved from publicly available *S. thermophilus* genomes and 45 (90%) CR3 loci were detected by PCR from the Uy strain collection. Analysis of these arrays showed that they contained a minimum of 5 spacers and a maximum of 44, with a median of 15 spacers per strain. Moreover, the same spacers appeared to have been acquired by some of these wild-type strains or were derived from a common ancestor (Figure 1a). These data also suggested that spacers were either deleted or acquired at specific positions within the CR3 array (Figure 1b). For example, when comparing *S. thermophilus* strain Uy23 with the strain Uy44 and the reference strains LMD-9 and Uy44, it would appear that two spacers were either deleted in the CR3 array of the Uy44/LMD-9 strains or were acquired by the strain Uy23. Similar events of either deleted or ectopically acquired spacers appear to have occurred within the CR3 arrays of other strains (Figure 1b). In some strains, new spacers appear to have

also been subsequently acquired at the 5' end of the array, as with the strain APC151 when compared to strains Uy07 and Uy33.

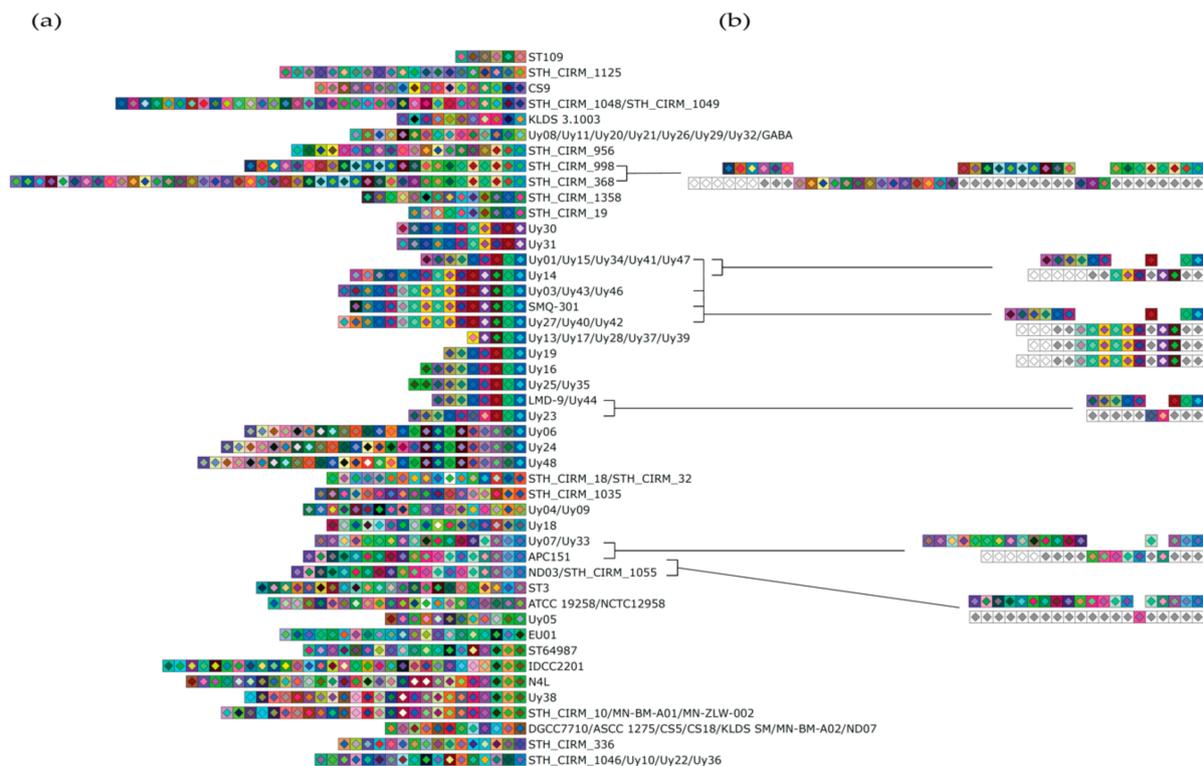


Figure 1. Spacer representation of CRISPR3 loci of *S. thermophilus* genomes using CRISPRStudio. (a) CRISPR3 loci of complete public *S. thermophilus* genomes and of strains of the Uy collection. (b) Selected examples of probable ectopic spacer acquisition or insertion/deletion events in CRISPR3 loci. The order of the spacers is the same as in panel a. White square harboring a light gray diamond represents similar spacer. White square harboring a white diamond represents unique spacer on the 5' end of the array.

We next sought to identify similar potential and natural ectopic acquisition events in other type II-A systems, outside *S. thermophilus*. The type II-A system of *S. pyogenes* system has been shown to acquire spacer ectopically when expressed in a heterologous host [33]. Thus, we explored if natural ectopic spacer acquisition could also be inferred from publicly available genomic sequences of *S. pyogenes*. Moreover, the Cas9 (1368 aa) of *S. pyogenes* is related to *S. thermophilus* CR3 Cas9 (1409 aa, 57% ID). Interestingly, the *S. pyogenes* system uses the PAM 5'-NGG-3', while the CR3 system of *S. thermophilus* uses the PAM 5'-NGGNG-3' [22,27].

From the 213 *S. pyogenes* genomes found in GenBank, 116 (54.5%) CRISPR arrays with at least three repeats were detected. Analysis of these arrays showed that they contained up to a maximum of 13 spacers, with a median of 4 spacers per strain. Of note, five strains (2.3%) with only two repeats and one spacer were also detected. Therefore, the number of spacers is generally smaller in *S. pyogenes* than in *S. thermophilus*.

Further analyses suggested that some of the strains acquired the same spacers or that the strains were derived from a common ancestor (Figure 2a). Moreover, specific spacers could have also been either deleted or acquired within the CRISPR array (Figure 2b). For example, when comparing *S. pyogenes* strain NCTC12044 with the strain NCTC12059, two spacers appear to have been deleted in the CRISPR array of strain NCTC12059 or acquired by the strain NCTC12044. The latter strain may also have acquired two unique spacers at the 5' end, compared to the reference strain.

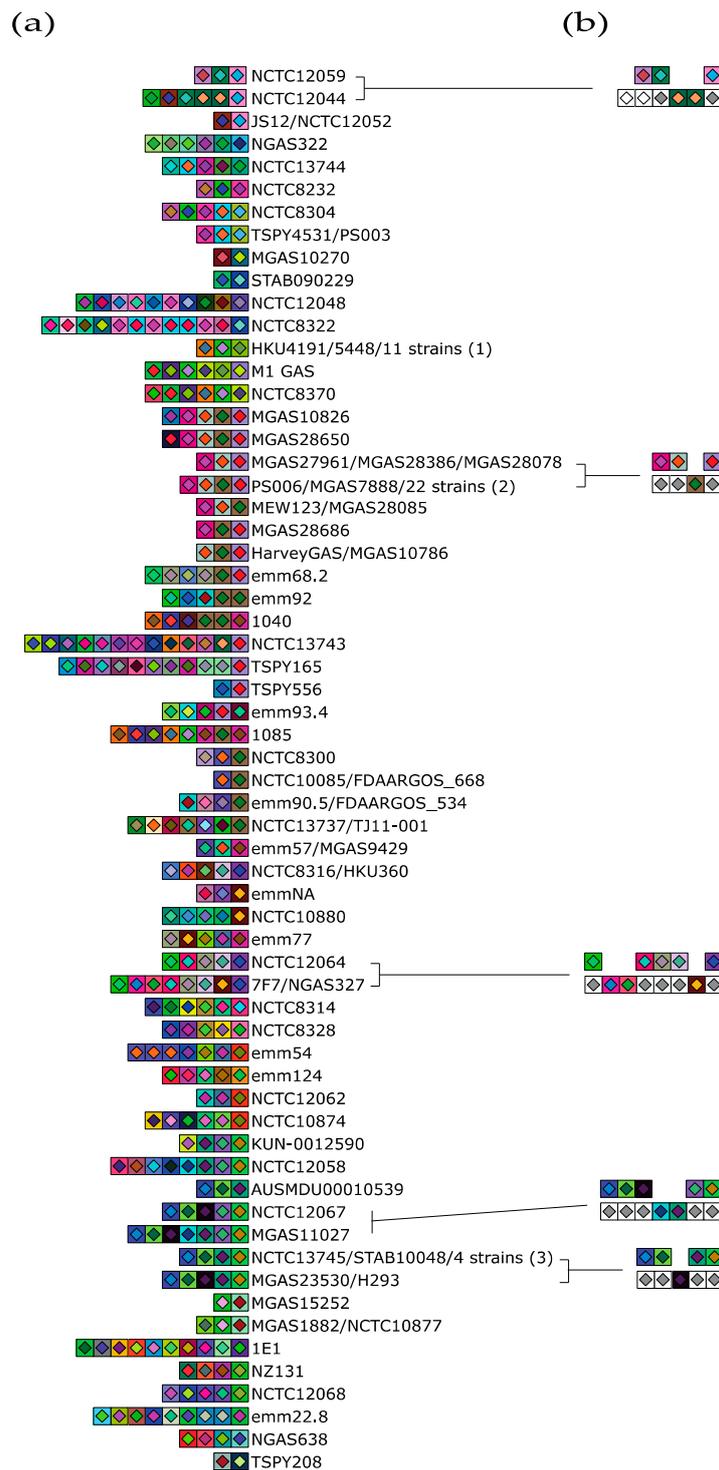


Figure 2. Spacer representation of CRISPR type II-A loci of *S. pyogenes* genomes using CRISPRStudio. (a) Type II-A CRISPR loci from public *S. pyogenes* genomes and containing at least two spacers. Some strains have identical CRISPR arrays and it includes Group 1: Strains S119, MGAS2221, emm1, FDAARGOS_149, HKU488, 5448, 10-85, MTB314, M1 476, MGAS5005, A20. Group 2: MGAS7914, MGAS8347, MGAS11052 MGAS11108, MGAS11115, MGAS28191 MGAS28271, MGAS28278, MGAS28330, MGAS28360, MGAS28533, MGAS28669, MGAS28746, MGAS29064, MGAS29284, MGAS29326, MGAS29409, M28PF1, STAB09014, STAB10015, MGAS6180, NIH35. Group 3: STAB09023, MGAS27061, JMUB1235, KUN-0014944. (b) Examples of possible ectopic acquisition or insertion/deletion events in *S. pyogenes* CRISPR arrays. The order of the spacers is the same as in panel a. White square harboring a light gray diamond represents similar spacer. White square harboring a white diamond represents unique spacer on the 5' end of the array.

3.2. Evaluation of Spacer Acquisition

To test if ectopic acquisition can indeed occur in the CR3 array of *S. thermophilus*, bacteriophage insensitive mutants (BIMs) were generated by infecting the phage-sensitive *S. thermophilus* strain Uy01 (which has nine spacers in its CR3 array) with the virulent phage 53 for a prolonged period of time. Forty-six randomly selected BIMs were analyzed for new spacer acquisition. All of them acquired new spacers in the CR3 array. Only one spacer was acquired per BIM, except for BIM1, which acquired two. A total of 25 different spacers (out of 47) were acquired by these 46 BIMs (Supplementary Table S1), indicating that some BIMs acquired the same spacer. We observed clear ectopic spacer acquisition events in the CR3 of 5 BIMs (10.9%), representing three CRISPR cluster (C) types (C2, C3 and C4; Figure 3). Spacer acquisitions with concomitant spacer deletions were also observed in four BIMs (8.7%) (see C4 and C5, Figure 3). It is unclear if additional ectopic spacer acquisitions also occurred in the three BIMs grouped in C5 (Figure 3). The remaining 38 BIMs acquired spacers at the leader 5' end of the CRISPR array (CR1, Figure 3). Interestingly, over a third (16) of all the acquisition events involved protospacers from *orf14* (the longest gene in the phage genome which codes for the tape measure protein), including six of the eight ectopic events (Figure 4). All the protospacers were flanked by the previously identified PAM for CR3 (5'-NGGNG-3') [4,8].

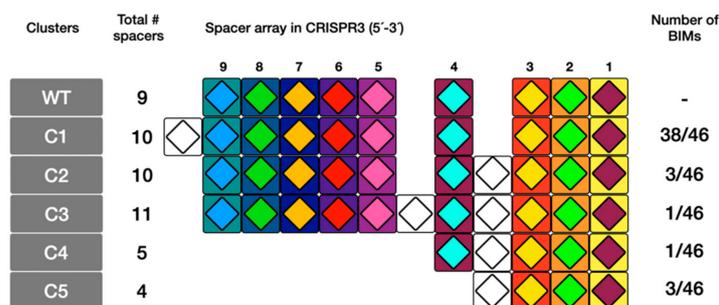


Figure 3. Ectopic spacer acquisition events observed for Uy01 bacteriophage insensitive mutants (BIMs). Spacers are aligned and their direction is shown from 5' to 3', with respect to the leader sequence. The position of each spacer is relative to the leader end of the reference wild-type (WT) strain Uy01, with spacer 1 being the most distal. Each spacer is represented by a combination of colors based on their respective nucleotide sequence. The white spacers indicate that various spacers were acquired at these positions. The letter C followed by a number indicates a BIM cluster.

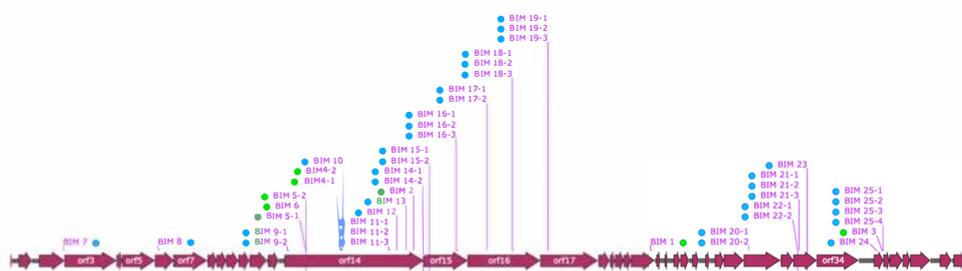


Figure 4. Graphic representation of phage 53 genome (accession no. KT717084). The arrows indicate the different open reading frames. The blue dots show the position of the protospacers acquired at the 5' end of the CR3 array by the BIMs obtained in this work. The green dots represent the position of protospacers acquired ectopically. The BIM number represents the different protospacers, and a hyphen followed by a number indicates that the protospacer was incorporated by more than one BIM.

Most ectopic acquisition events occurred between spacers 3 and 4. A single ectopic spacer acquisition also occurred between spacers 4 and 5 in BIM1. Interestingly, loss of spacer content also occurred in three of these BIMs, namely BIMs 3, 4, and 6 (Figure 3). The newly acquired spacers were mapped in the phage genome and perfectly matched

protospacers (Table 1), except for one new spacer that could not be matched to the phage genome. The BIMs were tested against phage 53 and all showed a phage-resistant phenotype, including those that had acquired the spacer in the middle of the array.

Table 1. Nucleotide sequence and position of newly acquired ectopic spacers in CRISPR3 of the BIMs obtained after a challenge with phage 53.

BIM	Sequence of New Spacer (5'-3')	Position in CRISPR3	Protospacer Position in Phage 53 Genome	PAM NGGNG
BIM1	TGATAGTAAAATATTGTCATCATTGAATAC CATTACAGACACAGGAGAAGGCGGCTATTA	6 4	None * 22837-22866, <i>orf22</i>	None TGGTG
BIM2	TTATGCAAACGGTGGCCTAGTCCACAAGAA	4	14360-14389, <i>orf14</i>	CGGCG
BIM3	AGTTGATGGTAAAACGGTGGAAATGACCATA	4	31076-31105, <i>orf34</i>	TGGCG
BIM4	AAACGTCAAAAAAGCTGGTAGTAAAGTCAA	4	11789-11818, <i>orf14</i>	TGGCG
BIM5	TGTTTCAGTATCGTTCGACTTCATTCCCCAAA	4	10537-10508, <i>orf14</i>	CGGCG
BIM6	TGTTTCAGTATCGTTCGACTTCATTCCCCAAA	4	10537-10508, <i>orf14</i>	CGGCG

* Did not match any region in phage 53 genome and no significant results were obtained from BLAST analysis.

Because the region flanking spacers 3 and 4 appears to be a hot spot for ectopic spacer acquisition in CR3, we investigated the presence of a motif that could mimic the leader sequence, as previously observed in ectopic events in CR1 of *S. thermophilus* [38]. In both cases, spacers did not contain the GAG motif found in the LAS sequence in the leader sequence (Figure 5). Only the adenine at position -2 matching the GAG motif was found in five spacers, including spacers 3 and 4.

```

Leader  -TACGAAATCTGGTAGAAAGATATCCTACGAGGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 1 --ATAGTATTAATTTTCATTGAAAAATAATTGTGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 2 --AGAACTTTATCAAGATAAACTACTTTAAAGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 3 --CTCATATTCGTTAGTTGCTTTTGTCAATAAGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 4 ACAGAGTACAATATTGTCCTCATTGGAGACACGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 5 --TGTGTTGGGAAACCGCAGTAGCCATGATTAAGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 6 --CCTTGTGTCGTGGCTCTCCATACGCCATATAGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 7 --TCAATGAGTGGTATCCAAGACGAAAACCTTAGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 8 --GGTGAAAAGGTTCACTGTACGAGTACTTAGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Spacer 9 --TATGCAAGTAAAGGAATATGCTTTTATATAAGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC

```

Figure 5. Nucleotide alignment of the leader sequence and spacers of *S. thermophilus* Uy01. The nucleotide sequence of the repeat is in red and the leader and spacers are in black. The underlined portion of the sequence is the putative LAS sequence of the *S. thermophilus* CRISPR3 array with the GAG motif highlighted in yellow.

4. Discussion

Integration of new spacers at the 5' end of the CRISPR loci is one of the hallmark features of the CRISPR-Cas systems in *S. thermophilus* [42] and other bacterial species [3], although ectopic spacer acquisition in CR1 was previously described in *S. thermophilus* [38]. Our results now show that ectopic spacer acquisition can also occur in the CR3 system of *S. thermophilus*. Spacer analysis of CRISPR type II-A loci of *S. pyogenes* genomes hinted at ectopic spacer acquisition in this species as well. BIMs were generated by infecting the phage-sensitive *S. thermophilus* strain Uy01 with the virulent phage 53. All of them acquired spacers in CR3. We also observed that many of the newly acquired spacers came from protospacers found in the *orf14* (coding for the tape measure protein) of phage 53, without any obvious reason, other than being the longest gene. Of interest, there are 478 CR3 PAMs in the genome of phage 53, and 68 of them (14%) are in *orf14*. The absence of spacer acquisition in the CR1 of this strain is unknown, but it could be related to the presence of genes coding for anti-CRISPR proteins in the phage or bacterial genome [43–45] or a defect in the spacer acquisition machinery.

Recent work showed the importance of a seven-nucleotide sequence at the 3' end of the leader sequence, called the “leader-anchoring sequence” (LAS), in the acquisition of

new spacer [42]. When this (or similar) LAS sequence is found in specific spacers, this may lead to ectopic spacer acquisition events, especially if they have a GAG motif at the 3' end [16]. In a previous study, we observed that ectopic integration occurred when the last G of the leader sequence was missing, and that new spacers could be added at five different positions within the array [38]. We noticed that in four out of the five adjacent spacers to the newly acquired one, an adenine and a guanine were found at position -2 and -1 (3' end), respectively. In agreement with previous studies [17–19], our data suggested that the LAS may be limited to only a few nucleotides, including the adenine at position -2 [21]. In the present study, the LAS sequence of the leader did contain the GAG motif, while spacers 3 and 4 did not have the GAG motif at the 3' end (just the adenine at position -2). Similar results were observed in *S. pyogenes* [16]. In that case, while the LAS sequence (conserved between the two species) was critical for the integration into the first repeat, one of the spacers may have led the addition of new spacers into its downstream repeat while not fully matching the GAG motif of the leader sequence (only the G at the position -3). Therefore, it could be possible that other factors guide ectopic spacer acquisition. Further experimental studies are needed to address the above.

Nevertheless, these results represent the first study showing ectopic spacer acquisition in the CR3 of *S. thermophilus*. Our analyses also suggest that a similar phenomenon may be naturally occurring in *S. pyogenes*, supporting the previously observed ectopic spacer acquisition in this type II-A system when expressed in *S. aureus* [33]. It should be noted that the acquisition of novel spacers at the 5' end of the CRISPR array is still the preferred location of new immunities. However, ectopic spacer acquisition also occurs at various frequencies. While CRISPR arrays still represent molecular archives of past nucleic acids invasion, the chronology of these events may not always be correlated with the spacer position within a given array.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2076-2607/9/3/512/s1>, Table S1: Nucleotide sequences and positions of newly acquired spacers in CRISPR3 of all the BIMs obtained after a challenge with the virulent *cos*-type phage 53.

Author Contributions: R.A. and S.M. conceived of this research. R.A., F.M., V.D. and M.P.C. performed the microbiological experiments and data analysis. G.M.R., C.P., M.B.D. and Y.S. performed bioinformatics analysis and generated figures. M.J.P. and S.M. critically reviewed the results and participated in the discussion. R.A., M.S., G.M.R. and S.M. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: R.A. is supported by the Uruguayan National Research and Innovation Agency (ANII). R.A. was also supported by the ELAP program of the Department of Foreign Affairs and International Trade of Canada. S.M. acknowledges funding from NSERC of Canada (Discovery program). S.M. holds a Tier 1 Canada Research Chair in Bacteriophages.

Data Availability Statement: The data presented in this study are available in the article's figures and tables and in the Supplementary Materials. The analyzed genome sequences are available on GenBank (NCBI). Additional information are available on request from the corresponding author.

Acknowledgments: We thank Carlos Sanguinetti for allowing the use of the facilities in the Biotechnology Research and Innovation Center (CBI+I) for part of this study. We are also grateful to Amanda Toperoff and Michi Waygood for editorial assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hols, P.; Hancy, F.; Fontaine, L.; Grossiord, B.; Prozzi, D.; Leblond-Bourget, N.; Decaris, B.; Bolotin, A.; Delorme, C.; Ehrlich, S.D.; et al. New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. *FEMS Microbiol. Rev.* **2005**, *29*, 435–463. [[CrossRef](#)]
2. Delorme, C. Safety assessment of dairy microorganisms: *Streptococcus thermophilus*. *Int. J. Food Microbiol.* **2008**, *126*, 274–277. [[CrossRef](#)] [[PubMed](#)]
3. Iyer, R.; Tomar, S.K.; Uma Maheswari, T.; Singh, R. *Streptococcus thermophilus* strains: Multifunctional lactic acid bacteria. *Int. Dairy J.* **2010**, *20*, 133–141. [[CrossRef](#)]

4. Uriot, O.; Denis, S.; Junjua, M.; Roussel, Y.; Dary-Mourot, A.; Blanquet-Diot, S. *Streptococcus thermophilus*: From yogurt starter to a new promising probiotic candidate? *J. Funct. Foods* **2017**, *37*, 74–89. [[CrossRef](#)]
5. Garneau, J.E.; Moineau, S. Bacteriophages of lactic acid bacteria and their impact on milk fermentations. *Microb. Cell Fact.* **2011**, *10*, S20. [[CrossRef](#)]
6. Mahony, J.; Cambillau, C.; van Sinderen, D. Host recognition by lactic acid bacterial phages. *FEMS Microbiol. Rev.* **2017**, *41*, S16–S26. [[CrossRef](#)] [[PubMed](#)]
7. Lavelle, K.; Martinez, I.; Neve, H.; Lugli, G.A.; Franz, C.M.A.P.; Ventura, M.; Bello, F.D.; van Sinderen, D.; Mahony, J. Biodiversity of *Streptococcus thermophilus* phages in global dairy fermentations. *Viruses* **2018**, *10*, 577. [[CrossRef](#)]
8. Romero, D.A.; Magill, D.; Millen, A.; Horvath, P.; Fremaux, C. Dairy lactococcal and streptococcal phage–host interactions: An industrial perspective in an evolving phage landscape. *FEMS Microbiol. Rev.* **2020**, *44*, 909–932. [[CrossRef](#)]
9. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D.A.; Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **2007**, *315*, 1709–1712. [[CrossRef](#)]
10. Deveau, H.; Barrangou, R.; Garneau, J.E.; Labonté, J.; Fremaux, C.; Boyaval, P.; Romero, D.A.; Horvath, P.; Moineau, S. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **2008**, *190*, 1390–1400. [[CrossRef](#)]
11. Mills, S.; Griffin, C.; Coffey, A.; Meijer, W.C.; Hafkamp, B.; Ross, R.P. CRISPR analysis of bacteriophage-insensitive mutants (BIMs) of industrial *Streptococcus thermophilus*—implications for starter design. *J. Appl. Microbiol.* **2010**, *108*, 945–955. [[CrossRef](#)] [[PubMed](#)]
12. Paez-Espino, D.; Sharon, I.; Morovic, W.; Stahl, B.; Thomas, B.C.; Barrangou, R.; Banfield, J.F. CRISPR immunity drives rapid phage genome evolution in *Streptococcus thermophilus*. *mBio* **2015**, *6*, e00262-15. [[CrossRef](#)]
13. Hao, M.; Cui, Y.; Qu, X. Analysis of CRISPR-Cas system in *Streptococcus thermophilus* and its application. *Front. Microbiol.* **2018**, *9*, 257. [[CrossRef](#)] [[PubMed](#)]
14. Naumenko, O.; Skrypkina, I.; Zhukova, Y.; Vakulenko, M.; Kigel, N. Selection and analysis of bacteriophage-insensitive mutants of *Streptococcus thermophilus* isolated in Ukraine. *Int. J. Dairy Technol.* **2019**, *72*, 515–523. [[CrossRef](#)]
15. Cañez, C.; Selle, K.; Goh, Y.J.; Barrangou, R. Outcomes and characterization of chromosomal self-targeting by native CRISPR-Cas systems in *Streptococcus thermophilus*. *FEMS Microbiol. Lett.* **2019**, *366*, fnz105. [[CrossRef](#)] [[PubMed](#)]
16. Bolotin, A.; Quinquis, B.; Sorokin, A.; Dusko Ehrlich, S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **2005**, *151*, 2551–2561. [[CrossRef](#)] [[PubMed](#)]
17. 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* **2005**, *151*, 653–663. [[CrossRef](#)] [[PubMed](#)]
18. Hille, F.; Richter, H.; Wong, S.P.; Bratovič, M.; Ressel, S.; Charpentier, E. The biology of CRISPR-Cas: Backward and forward. *Cell* **2018**, *172*, 1239–1259. [[CrossRef](#)] [[PubMed](#)]
19. Brouns, S.J.J.; Jore, M.M.; Lundgren, M.; Westra, E.R.; Slijkhuis, R.J.H.; Snijders, A.P.L.; Dickman, M.J.; Makarova, K.S.; Koonin, E.V.; van der Oost, J. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **2008**, *321*, 960–964. [[CrossRef](#)]
20. Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pírzada, Z.A.; Eckert, M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **2011**, *471*, 602–607. [[CrossRef](#)]
21. Garneau, J.E.; Dupuis, M.È.; Villion, M.; Romero, D.A.; Barrangou, R.; Boyaval, P.; Fremaux, C.; Horvath, P.; Magadán, A.H.; Moineau, S. The CRISPR/cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **2010**, *468*, 67–71. [[CrossRef](#)] [[PubMed](#)]
22. Magadán, A.H.; Dupuis, M.È.; Villion, M.; Moineau, S. Cleavage of phage DNA by the *Streptococcus thermophilus* CRISPR3-Cas system. *PLoS ONE* **2012**, *7*, 40913. [[CrossRef](#)]
23. Gasiunas, G.; Sinkunas, T.; Siksnys, V. Molecular mechanisms of CRISPR-mediated microbial immunity. *Cell. Mol. Life Sci.* **2014**, *71*, 449–465. [[CrossRef](#)] [[PubMed](#)]
24. Makarova, K.S.; Wolf, Y.I.; Iranzo, J.; Shmakov, S.A.; Alkhnbashi, O.S.; Brouns, S.J.J.; Charpentier, E.; Cheng, D.; Haft, D.H.; Horvath, P.; et al. Evolutionary classification of CRISPR–Cas systems: A burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **2020**, *18*, 67–83. [[CrossRef](#)]
25. Nuñez, J.K.; Kranzusch, P.J.; Noeske, J.; Wright, A.V.; Davies, C.W.; Doudna, J.A. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* **2014**, *21*, 528–534. [[CrossRef](#)]
26. Amitai, G.; Sorek, R. CRISPR-Cas adaptation: Insights into the mechanism of action. *Nat. Rev. Microbiol.* **2016**, *14*, 67–76. [[CrossRef](#)]
27. Horvath, P.; Romero, D.A.; Coûté-Monvoisin, A.C.; Richards, M.; Deveau, H.; Moineau, S.; Boyaval, P.; Fremaux, C.; Barrangou, R. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* **2008**, *190*, 1401–1412. [[CrossRef](#)] [[PubMed](#)]
28. Carte, J.; Christopher, R.T.; Smith, J.T.; Olson, S.; Barrangou, R.; Moineau, S.; Glover 3rd, C.V.C.; Graveley, B.R.; Terns, R.M.; Terns, M.P. The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Mol. Microbiol.* **2014**, *93*, 98–112. [[CrossRef](#)]
29. Mosterd, C.; Rousseau, G.M.; Moineau, S. A short overview of the CRISPR-Cas adaptation stage. *Can. J. Microbiol.* **2020**, *12*, 1–12. [[CrossRef](#)]
30. Makarova, K.S.; Wolf, Y.I.; Alkhnbashi, O.S.; Costa, F.; Shah, S.A.; Saunders, S.J.; Barrangou, R.; Brouns, S.J.J.; Charpentier, E.; Haft, D.H.; et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **2015**, *13*, 722–736. [[CrossRef](#)] [[PubMed](#)]

31. Wilkinson, M.; Drabavicius, G.; Silanskas, A.; Gasiunas, G.; Siksnys, V.; Wigley, D.B. Structure of the DNA-Bound Spacer Capture Complex of a Type II CRISPR-Cas System. *Mol. Cell* **2019**, *75*, 90–101. [[CrossRef](#)] [[PubMed](#)]
32. McGinn, J.; Marraffini, L.A. Molecular mechanisms of CRISPR–Cas spacer acquisition. *Nat. Rev. Microbiol.* **2019**, *17*, 7–12. [[CrossRef](#)] [[PubMed](#)]
33. McGinn, J.; Marraffini, L.A. CRISPR-Cas systems optimize their immune response by specifying the site of spacer integration. *Mol. Cell* **2016**, *64*, 616–623. [[CrossRef](#)]
34. Wright, A.V.; Doudna, J.A. Protecting genome integrity during CRISPR immune adaptation. *Nat. Struct. Mol. Biol.* **2016**, *23*, 876–883. [[CrossRef](#)]
35. Wright, A.V.; Liu, J.J.; Knott, G.J.; Doxzen, K.W.; Nogales, E.; Doudna, J.A. Structures of the CRISPR genome integration complex. *Science* **2017**, *357*, 1113–1118. [[CrossRef](#)]
36. Xiao, Y.; Ng, S.; Hyun Nam, K.; Ke, A. How type II CRISPR-Cas establish immunity through Cas1-Cas2-mediated spacer integration. *Nature* **2017**, *550*, 137–141. [[CrossRef](#)]
37. Mosterd, C.; Moineau, S. Characterization of a Type II-A CRISPR-Cas system in *Streptococcus mutans*. *mSphere* **2020**, *5*, e00235-20. [[CrossRef](#)]
38. Achigar, R.; Magadán, A.H.; Tremblay, D.M.; Julia Pianzola, M.; Moineau, S. Phage-host interactions in *Streptococcus thermophilus*: Genome analysis of phages isolated in Uruguay and ectopic spacer acquisition in CRISPR array. *Sci. Rep.* **2017**, *7*, 43438. [[CrossRef](#)]
39. Hynes, A.P.; Villion, M.; Moineau, S. Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. *Nat. Commun.* **2014**, *5*, 4399. [[CrossRef](#)] [[PubMed](#)]
40. Biswas, A.; Staals, R.H.J.; Morales, S.E.; Fineran, P.C.; Brown, C.M. CRISPRDetect: A flexible algorithm to define CRISPR arrays. *BMC Genomics* **2016**, *17*, 356. [[CrossRef](#)]
41. Dion, M.B.; Labrie, S.J.; Shah, S.A.; Moineau, S. Crisprstudio: A user-friendly software for rapid crispr array visualization. *Viruses* **2018**, *10*, 602. [[CrossRef](#)]
42. Wei, Y.; Chesne, M.T.; Terns, R.M.; Terns, M.P. Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in *Streptococcus thermophilus*. *Nucleic Acids Res.* **2015**, *43*, 1749–1758. [[CrossRef](#)] [[PubMed](#)]
43. Hynes, A.P.; Rousseau, G.M.; Lemay, M.-L.; Horvath, P.; Romero, D.A.; Fremaux, C.; Moineau, S. An anti-CRISPR from a virulent streptococcal phage inhibits *Streptococcus pyogenes* Cas9. *Nat. Microbiol.* **2017**, *2*, 1374–1380. [[CrossRef](#)]
44. Landsberger, M.; Gandon, S.; Meaden, S.; Rollie, C.; Chevallereau, A.; Chabas, H.; Buckling, A.; Westra, E.R.; van Houte, S. Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. *Cell* **2018**, *174*, 908–916. [[CrossRef](#)] [[PubMed](#)]
45. Stanley, S.Y.; Maxwell, K.L. Phage-Encoded Anti-CRISPR Defenses. *Annu. Rev. Genet.* **2018**, *52*, 445–464. [[CrossRef](#)] [[PubMed](#)]