Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

A short overview of the CRISPR-Cas adaptation stage

Cas Mosterd[1,2], Geneviève M. Rousseau[1,2] and Sylvain Moineau[1,2,3]*

[1] Département de biochimie, de microbiologie et de bio-informatique, Faculté des sciences et de génie, Université Laval, Québec, Canada, G1V 0A6

[2] Groupe de recherche en écologie buccale, Faculté de médecine dentaire, Université Laval, Québec, Canada, G1V 0A6

[3] Félix d'Hérelle Reference Center for Bacterial Viruses, Faculté de médecine dentaire, Université Laval, Québec, Canada, G1V 0A6

Correspondence should be addressed to:

Tel: +1 418 656 3712; Email: Sylvain.Moineau@bcm.ulaval.ca

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

Abstract

CRISPR research began over 30 years ago with the incidental discovery of an unusual nucleotide arrangement in the *Escherichia coli* genome. It took twenty years to find the main function of CRISPR-Cas systems as an adaptive defence mechanism against invading nucleic acids and our knowledge of their biology has been steadily increased ever since. In parallel, the number of applications derived from CRISPR-Cas systems has risen spectacularly. The CRISPR-based genome editing tool is arguably the most exciting application in both basic and applied research. Lately, CRISPR-Cas research has partially shifted to the least understood aspect of its biology: the ability of CRISPR-Cas systems to acquire new immunities during the so-called adaptation step. To date, the most efficient natural system to readily acquire new spacers is the type II-A system of the gram-positive dairy bacterium *Streptococcus thermophilus*. The discovery of additional systems able to acquire new spacers will hopefully draw more attention to this step of CRISPR-Cas biology. This review focuses on the breakthroughs that have helped to unravel the adaptation phase and on questions that remain to be answered.

Keywords: CRISPR-Cas, bacteriophage, plasmid, anti-viral, defence, resistance

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

Introduction

Clustered regularly interspaced short palindromic repeats (CRISPRs) were first described in bacteria in 1987, during the analysis of the nucleotide sequence of the *iap* gene responsible for the isozyme conversion of alkaline phosphatase (and its flanking region) in *Escherichia coli*. A series of five repeated palindromic nucleotide sequences flanking the 3' end of the *iap* gene were separated by variable sequences of uniform length, with no homology to each other (Ishino et al. 1987). Six years later, another research group described a similar nucleotide sequence pattern in Archaea (Mojica et al. 1993). During these early days, multiple groups studying these mysterious repeats used different names and acronyms, resulting in confusion within the field. In 2002, the term CRISPR was introduced and the first CRISPR-associated (*cas*) genes were described (Jansen et al. 2002). These genes were found to flank CRISPR loci and were often absent in CRISPR-negative strains. The importance of the variable sequences, named spacers, between the repeats remained a mystery until 2005. Homologies were discovered between some spacers within a few CRISPR loci, and virus and plasmid sequences (Bolotin et al. 2005; Mojica et al. 2005; Pourcel et al. 2005). Several hypotheses were proposed as the mode of action, one of them included that CRISPRs may play a role in the defence against foreign genetic material (Mojica et al. 2005; Pourcel et al. 2005). In 2006, bioinformatics analyses of various CRISPR loci and *cas* genes suggested the cooperation between CRISPRs and Cas proteins in defence against plasmids and phages, by using a mechanism comparable to eukaryotic RNA interference systems (Makarova et al. 2006).

The first experimental demonstration of the function of CRISPR-Cas was published in 2007 in the dairy bacterial species *Streptococcus thermophilus*. Upon exposure to virulent (also known as lytic) streptococcal phages, the vast majority of the phage-infected cells died but a small number of cells survived phage infection. Analysis of the CRISPR array of the

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

surviving bacterial cells, revealed that a new spacer was integrated into the CRISPR locus along with the addition of a repeat (Barrangou et al. 2007). The Eureka moment soon followed when it was realised that the new spacers originated from the genome of the virulent phage used in the assay. Later, the exact same sequence present in the phage genome was termed "protospacer" to differentiate from the bacterial spacer found in the array (Deveau et al. 2008). Integrating a novel spacer identical to a protospacer resulted in resistance towards the phage carrying this sequence. The removal of the spacer sequence restored the phage-sensitive phenotype (Barrangou et al. 2007). Interference activity against plasmids was demonstrated in 2008 (Marraffini and Sontheimer 2008). It was then shown in *E. coli*, that the CRISPR locus is transcribed and later processed by Cas proteins to serve as small guide RNAs that allow a complex of Cas proteins and CRISPR-RNA (crRNA) to interfere with an invasion by recognised foreign DNA (Brouns 2008). In 2010, the CRISPR-Cas system of *S. thermophilus* was shown to precisely cleave target phage and plasmid DNA within the protospacer sequence *in vivo*. This provided a clear explanation for the phage resistance and plasmids interference (Garneau et al. 2010). Although multiple Cas proteins are involved in the interference step in *E. coli*, Cas9 is the only Cas protein required to cut the target DNA in *S. thermophilus* (12, 13) and *Streptococcus pyogenes* (Jinek et al. 2012). This paved the way for the development of the CRISPR-Cas9-based genome editing tool (14).

General principle of CRISPR-Cas function

CRISPR-Cas systems are divided into two classes, as well as in six types and several subtypes (see below). The differences between the CRISPR-Cas systems are based on the composition of the *cas* genes (Koonin et al. 2017). Differences in repeat sequences and number of spacer-repeat units also occur between CRISPR-Cas systems of the same subtype and within the

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

same species, respectively (Horvath et al. 2008; Díez-Villaseñor et al. 2010). Although on a molecular level, the mechanisms used by CRISPR-Cas systems inside the cell differ depending on the type of system, the general principle consists of three steps that all systems appear to follow.

The first step is the adaptation phase, during which a novel spacer is acquired from an invading nucleic acid sequence. In some types of CRISPR-Cas systems, adaptation occurs only when a sequence motif flanking the protospacer, named the protospacer adjacent motif (PAM), is recognised and the protospacer is inserted into the CRISPR locus (Deveau et al. 2008; Horvath et al. 2008). The duplication of a repeat sequence indicates that the expansion of the CRISPR locus by one novel repeat-spacer unit is completed. Cas1 and Cas2 are conserved within almost all CRISPR-Cas types and form a complex involved in the adaptation step (Nuñez et al. 2014). The adaptation phase is followed by the expression phase, during which the CRISPR locus is transcribed into one long strand of pre-crRNA while additional *cas* genes are expressed. Some of these expressed Cas proteins are involved in the processing of pre-crRNA into mature short crRNAs (Charpentier et al. 2015). These crRNAs then form a complex with one or more Cas proteins. Finally, the third and final step is called the interference phase. Upon the subsequent invasion by nucleic acids harbouring a protospacer, the sequence is recognised and cleaved by the crRNA-Cas complexes (Brouns 2008). Whereas the Cas1 and Cas2 proteins of the adaptation phase are ubiquitous among the different system types, there is more diversity in the Cas proteins involved in interference (Koonin et al. 2017). Proteins involved in the interference phase can also sometimes play a role in the initial adaptation step (Heler et al. 2015; Wei et al. 2015b).

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

Classification

CRISPR-Cas systems have been found in approximately 40% of the bacterial genomes (Makarova et al. 2019) and their classification is being re-evaluated and updated on a regular basis (Makarova et al. 2011, 2015, 2018, 2019; Koonin et al. 2017). As indicated above, two classes have been defined. Systems that belong to Class 1 use complexes made up of multiple Cas proteins that mediate interference, whereas Class 2 systems use one single Cas multidomain effector protein to cut nucleic acids. An overview of the different classes, types and subtypes is presented in Figure 1.

Class 1 CRISPR-Cas systems comprise types I, III and IV, with each type having its unique signature gene. In type I, the gene is *cas3*, which encodes a protein with both helicase and nuclease domains (Brouns 2008; Sinkunas et al. 2011). For type III, the signature gene is *cas10*, a gene encoding a multidomain protein with nuclease activity involved in the interference step (Ramia et al. 2014). The *cas10* gene is missing in subtype III-E (Makarova et al. 2019). The proposed signature gene for type IV is *csf1* (Makarova et al. 2015, 2018; Koonin et al. 2017), although it is missing in subtype IV-C (Makarova et al. 2019). Type I and III systems both make use of effector complexes consisting of several Cas proteins and crRNA. The effector complex of type I systems is the Cascade complex; while for type III systems, it is the Cmr complex. These effector complexes differ in protein composition and share little sequence identity. Nevertheless, the architecture and function of these two complexes are similar since they evolved from a common ancestor (Makarova et al. 2011; Jackson and Wiedenheft 2015). A major functional difference between Cascade complexes and Cmr complexes is that Cascade complexes target DNA (Brouns 2008; Sinkunas et al. 2011), while Cmr complexes target both RNA (Hale et al. 2009) and DNA (Samai et al. 2015). Type IV is different from the two other types in that it does not contain *cas1* and *cas2*

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

genes. Moreover, its *cas* genes are usually not found close to a CRISPR locus, but the associated proteins are nevertheless predicted to form an effector complex (Makarova et al. 2015, 2018; Koonin et al. 2017).

The class 2 CRISPR-Cas systems include types II, V and VI. Although class 1 CRISPR-Cas systems are more common in bacteria, the class 2 has received significantly more attention due to its application in genome editing. Type II CRISPR-Cas systems, the first of the class 2 to be described, are the simplest in terms of the number of genes. The signature gene *cas9* encodes a protein that fulfils the roles of the multiple proteins found in the class 1 effector complexes. Cas9 contains two nuclease domains, RuvC and HNH, and is the only Cas protein required to cleave the target DNA (Jinek et al. 2012; Gasiunas et al. 2012). Besides Cas1 and Cas2, Cas9 has also been shown to play a role in the adaptation step (Heler et al. 2015; Wei et al. 2015b). In addition to the *cas* genes and CRISPR locus, trans-activating small RNAs (tracrRNAs) and the host RNaseIII are also essential components of type II systems required for the maturation of pre-crRNA into crRNAs (Deltcheva et al. 2011; Chylinski et al. 2013). Type V and VI were later additions to the class 2. As in other systems, Cas1 and Cas2 are involved during adaptation in type V and VI systems (Makarova et al. 2015). Type V was proposed after the discovery of the *cas12* gene and it was predicted that Cas12 (also known as Cpf1) is the only protein involved in the interference stage (Strecker et al. 2019; Swarts and Jinek 2019). Cas12 contains only a RuvC nuclease domain, thereby lacking the HNH domain (Zetsche et al. 2015; Swarts and Jinek 2019). No tracrRNA is needed for the maturation of the crRNAs (Zetsche et al. 2015; Swarts and Jinek 2019). The exceptions are subtypes V-F1 and V-G which do require tracrRNA. The *cas12* gene is not found in subtype V-U (Makarova et al. 2019). The discovery of *cas13* gave rise to type VI and Cas13 encodes a tracrRNA-independent nuclease, which targets RNA (Shmakov et al. 2015; Abudayyeh et al. 2016).

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

Adaptation

In recent years, steady progress has been made in our understanding of CRISPR-Cas systems with the interference phase attracting the most attention. Many molecular details remain poorly understood in the spacer acquisition process. During this phase, a stretch of DNA (prespacer) is first captured by Cas proteins (Wilkinson et al. 2019). Then, the Cas proteins and the bound prespacer are directed to the CRISPR locus, a repeat sequence is duplicated and the prespacer is integrated between the two repeats as a novel spacer. The *cas1* and *cas2* genes are essential and specific to the adaptation phase and both are relatively well conserved in almost all CRISPR-Cas systems (Brouns 2008). Other universal elements of the acquisition process are the leader sequence (discussed below) and the first repeat of the CRISPR locus (Amitai and Sorek 2016), suggesting that some processes are likely conserved in all types (Makarova et al. 2015).

Cas1 is a metal-dependent endonuclease that cleaves single-stranded DNA (Babu et al. 2011), double-stranded DNA (dsDNA) (Wiedenheft et al. 2009), and single-stranded RNA (ssRNA) (Kim et al. 2013). Like Cas1, Cas2 is also a metal-dependent endonuclease. The Cas2 protein was first described as having activity against ssRNA (Beloglazova et al. 2008). However, depending on the CRISPR-Cas system, Cas2 may have activity against dsDNA (Nam et al. 2012). Cas1 and Cas2 form a complex consisting of a Cas2 dimer flanked by two Cas1 dimers ($Cas1_4$-$Cas2_2$). Within the complex only the nuclease activity of Cas1 is essential for adaptation. Cas1 has a strong binding affinity for the CRISPR locus, as it is involved in the integration of new spacers. In the absence of Cas2, Cas1 is incapable of binding to the CRISPR locus, making both proteins essential during adaptation (Nuñez et al. 2014). Cas2

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

functions as an adaptor protein, bridging the Cas1 proteins as well as binding and stabilising the prespacer DNA (Wang et al. 2015).

In addition, a CRISPR repeat and a leader sequence are required for the integration of a novel spacer. The AT-rich leader sequence is typically located upstream of the CRISPR locus, which often contains the promoter that directs transcription of the CRISPR locus into pre-crRNA (Pougach et al. 2010; Pul et al. 2010). Repressor proteins silence transcription in certain species (Pougach et al. 2010; Pul et al. 2010; Medina-Aparicio et al. 2011), whereas in other species expression of the CRISPR operon can be constitutive and up-regulated during phage infection (Agari et al. 2010; Young et al. 2012; Quax et al. 2013). In addition to its function in transcription, the leader sequence has another role within the integration process as the integration of novel spacers usually occurs at the leader end of the CRISPR locus (Barrangou et al. 2007; Yosef et al. 2012; Díez-Villaseñor et al. 2013). Indeed, the leader sequence and the flanking repeat sequence harbour recognition signals with the ability to direct the Cas1-Cas2 complex to this position. These recognition sites are located within 10 base pairs of the integration site in both the leader and repeat sequences(Wei et al. 2015b).

The method by which sequences are integrated into the CRISPR locus suggests the probable acquisition of spacers that also target the bacterial chromosome (self-DNA) instead of foreign DNA. However, there appears to be a bias towards acquiring new spacers from foreign DNA compared to self-DNA. The number of PAM sequences present on bacterial self-DNA is similar to that of foreign DNA and therefore, the reason cannot be related to PAM prevalence. It has been shown that one of the sources of protospacers for the Cas1-Cas2 complex are degradation products of the bacterial RecBCD complex, which processes double stranded DNA breaks (DSBs), although this does not apply to all the CRISPR-Cas system types. The

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

large majority of these DSBs occur at replication forks during DNA replication and more are found on plasmids compared to the bacterial chromosome. The RecBCD complex unwinds and degrades the DNA, starting at the DSB until it reaches a crossover hotspot instigator (Chi) site. Bacterial genomic DNA is rich in Chi sites compared to plasmid and phage DNA, and therefore is indirectly protected from spacer acquisition. Because phage DNA enters the cell as linear DNA, contains few Chi sites and is highly replicative, it is presumably an easier target for the RecBCD complex than chromosomal DNA. This may explain why the CRISPR-Cas system preferentially acquires spacers from foreign DNA sources rather than its own (Levy et al. 2015). In type I systems, the degradation products from the activity of Cas3 during primed adaptation (see below) have also been demonstrated to function as prespacers (Künne et al. 2016; Staals et al. 2016). While the RecBCD complex is found only in Gram-negative bacteria, Gram-positive bacteria possess a highly similar mechanism in the form of the AddAB repair machinery. As with the RecBCD complex, the function of the AddAB machinery has been demonstrated to have a large impact on spacer acquisition (Modell et al. 2017).

Depending on the type of CRISPR-Cas system, additional factors are required to process the prespacers into spacers ready for integration. In type I-E systems, Cas2 is fused to a DnaQ domain, which degrades prespacers at the 3'-end to generate suitable spacers for integration (Drabavicius et al. 2018; Kim et al. 2020). A similar role has been proposed for Cas4, present in types I, II, and V systems. Cas4 cuts the 3' overhangs of prespacer DNA until a bound $Cas1_4$-$Cas2_2$ complex is encountered to generate spacers of specific length and a correct PAM. As such, Cas4 prevents the integration of non-functional spacers with an inappropriate length or PAM (Kieper et al. 2018; Lee et al. 2018, 2019; Rollie et al. 2018; Shiimori et al.

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

2018; Almendros et al. 2019). The function of Cas4 and DnaQ can also be performed by non-CRISPR nucleases in other systems (Yoganand et al. 2019).

Spacer acquisition under laboratory conditions

The large diversity of spacers found in strains possessing CRISPR loci indicates that spacer acquisition clearly occurs in numerous natural ecosystems. Surprisingly, spacer acquisition in bacterial strains has rarely been observed under laboratory conditions. The first case of spacer acquisition was observed with the dairy bacterium *S. thermophilus* and correlated with phage resistance (8) and plasmid interference (12). Spacer acquisition was also demonstrated in *Streptococcus mutans* following phage infection, although in this species, other phage resistance mechanisms appear to be at play (van der Ploeg 2009). Strains of *S. thermophilus* and *S. mutans* possess type II-A CRISPR-Cas systems. Spacer acquisition against plasmids in a type II-C system has also been described in *Riemerella anatipestifer* (He et al. 2018). The type II-C system of *Campylobacter jejuni* was also able to acquire novel spacers into its CRISPR array, but only after infection by a phage expressing a Cas4-like protein (Hooton and Connerton 2015).

Spacer acquisition in type I CRISPR-Cas systems has also been described (51-54). In the *E. coli* K12 type I-E system, spacer acquisition required the upregulation of *cas* genes or the deletion of a repressor. With other CRISPR-Cas systems, a fitness cost was associated with constitutive expression of the *cas* genes and to a lesser extent, during the acquisition of novel spacers, which may explain the presence of these repressors (Vale et al. 2015). The first evidence of spacer acquisition through priming was also observed in *E. coli*, although both primed and naïve spacer acquisition were observed (Datsenko et al. 2012; Swarts et al. 2012). Whereas naïve spacer acquisition refers to spacer acquisition from a virus not previously encountered, priming refers to spacer acquisition following the previous acquisition of a

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

253 spacer from the same foreign DNA. Spacer acquisition was also observed in the CRISPR
254 locus of the type I-B system of the archeon *Haloarcula hispanica*, although pre-existing
255 spacers partially matching the infecting viral genome were required for adaptation (51). These
256 partially matching spacers did not provide resistance against the virus, but still increase spacer
257 acquisition events during new encounters with the same virus. Similar results were obtained
258 with the type I-F system of *Pectobacterium atrosepticum* (Richter et al. 2014). Spacer
259 acquisition was also observed in the archaea *Sulfolobus solfataricus* and *S. islandicus*,
260 although only during viral infection combined with the transformation of a plasmid (Erdmann
261 and Garrett 2012).
262
263 Growth conditions have been reported to be a major influence on the capacity of CRISPR-Cas
264 to acquire novel spacers. *Pseudomonas aeruginosa* strains harbouring an active CRISPR-Cas
265 predominantly became phage-resistant through surface receptor modification under nutrient-
266 rich growth conditions, whereas under low nutrient conditions derivatives became phage-
267 resistant due to the acquisition of new spacers (Westra et al. 2015). CRISPR-Cas appears to
268 provide a fitness advantage during growth in low-nutrient conditions (Westra et al. 2015;
269 Chabas et al. 2016). The fitness disadvantage of surface receptor modification was further
270 increased in a co-culture of multiple bacterial species, increasing the prevalence of CRISPR
271 resistance (Alseth et al. 2019). A possible explanation for this is that the type IV pilus that
272 functions as phage receptor in *P. aeruginosa* has a role in the competition between species
273 (Westra et al. 2015; Alseth et al. 2019).
274
275 While the acquisition of a single new spacer is the norm, the acquisition of multiple spacers
276 has also been observed under laboratory conditions, particularly on solid media. Depending
277 on the capacity of the first acquired spacer to protect the bacterium against an invading virus,

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

a second spacer can be acquired to increase the overall phage resistance. It has been proposed that the acquisition of the second spacer is driven by priming during a single infection (Nussenzweig et al. 2019; Pyenson and Marraffini 2020) or by the re-infection of the virus (Pyenson and Marraffini 2020).

To identify systems that acquire novel spacers, tactics or factors that promote spacer acquisition can be employed. In the priming phenomenon described above the mismatches between the interference complex (crRNA) and the incoming phage genome do not allow for a strong interaction but still favour the Cas adaptation module to acquire additional spacers (Datsenko et al. 2012; Swarts et al. 2012; Jackson et al. 2019). Priming has been mainly associated with type I CRISPR-Cas systems. Evidence for primed acquisition has not been demonstrated in type II systems, nonetheless, it has been noted that upon acquisition of a second spacer, this spacer is typically acquired from the vicinity of the first spacer (Nussenzweig et al. 2019). Bioinformatics analysis revealed that acquisition of newer spacers does not appear to be random but in the proximity of presumed older spacers (Nicholson et al. 2019). In *S. thermophilus*, it was demonstrated that prior phage exposure to UV light increased the number of spacer acquisition events. Specifically, the UV light generated defective phage particles that retained the capacity to inject their genome into a bacterial host but lost the capacity to generate infectious virions (Hynes et al. 2014). Presumably, it allows more time to select and integrate protospacers within the CRISPR locus, compared to an infection by an infective virulent phage where the adaptation module must be expressed and a spacer acquired before the virus lyses the cell (Hynes et al. 2014). A similar observation was made when the invading phage DNA is cleaved by restriction enzymes (82). Cleavage of phage DNA by restriction enzymes neutralises infection and provides prespacers for the CRISPR-Cas system (Dupuis et al. 2013).

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

Type II-A CRISPR-Cas system of *S. thermophilus* as a model for spacer acquisition

As indicated previously, the first observation of spacer acquisition under laboratory conditions was in *S. thermophilus*, specifically with the dairy industrial strain DGCC7710 (Barrangou et al. 2007). This unique strain is used worldwide in the production of yoghurt due to its production of exopolysaccharides, which leads to the desired organoleptic properties of the final fermented milk product. The dairy industry relies on DGCC7710 strains that are naturally resistant to phages, therefore this resulted in CRISPR research with this specific bacterial strain.

In the initial experimental studies on CRISPR-Cas systems (8, 9, 12), *S. thermophilus* DGCC7710 was challenged with the virulent phages 2972 or 858 or both, and the CRISPR loci of the resulting bacteriophage-insensitive mutants (BIMs) were analysed (8, 9). Novel spacers were acquired from the phage genome in response to phage exposure and these new spacers came from the phage genome. These BIMs were not only resistant to the phages used in the assay, but also to other phages that carry the identical protospacer sequence. Furthermore, an overall stronger phage resistance phenotype was observed when multiple spacers were acquired (Barrangou et al. 2007). Interestingly, some phages were also found to evade CRISPR immunity (CRISPR escaping mutants, or CEMs). Genomic analysis of CEMs revealed point mutations in protospacers or the flanking PAMs, which allowed the circumvention of the interference complexes. Subsequently, when exposed to CEMs, former BIMs were also able to establish renewed phage resistance by acquiring additional spacers (9, 59). This adaptive immunity is thus comparable to adding a new vaccine to your

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

327 immunisation record and the spacer content of a bacterial strain can be seen as a molecular
328 archive of past infections by foreign nucleic acids.
329
330 Despite its adaptability, any given bacterial strain can also carry multiple CRISPR-Cas
331 systems and *S. thermophilus* is no exception. Four of these systems have been reported in this
332 food-grade bacterium. CRISPR1 and CRISPR3 are both type II-A systems and have been
333 experimentally shown to actively acquire novel spacers, with the large majority of the
334 acquisition events taking place in the CRISPR1 locus (Horvath et al. 2008; Magadán et al.
335 2012). The two systems have the capacity to cooperate and increase the overall phage
336 resistance phenotype within the same cell (Magadán et al. 2012). In addition, the CRISPR-
337 Cas systems of *S. thermophilus* are compatible with other bacterial defence mechanisms,
338 including restriction-modification systems (61). Furthermore, a type III-A system (CRISPR2)
339 and a type I-E system (CRISPR4) can be found in some *S. thermophilus* strains (Magadán et
340 al. 2012). No *in vivo* activity has been observed in the streptococcal type I-E system, but the
341 isolated Cascade complex has been shown to interfere with targeted plasmid DNA (Sinkunas
342 et al. 2013).
343
344 The two type II-A systems of *S. thermophilus* consist of four *cas* genes. In addition to the type
345 II system signature gene *cas9* and the conserved *cas1* and *cas2* genes, a fourth gene called
346 *csn2* distinguishes type II-A systems from other type II systems (Makarova et al. 2015). The
347 *csn2* gene is essential for adaptation, although its precise role is unknown (Barrangou et al.
348 2007). Although the two type II-A systems have the same gene composition, there is a
349 significant variation at the amino acid levels between the *cas* gene products. The percent
350 identity is as low as 24.4% for Cas9, 32.5% for Cas1, 38.7% for Cas2, and 21.76% for Csn2

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

(Figure 2). All four proteins and distinct tracrRNAs are needed for spacer acquisition, with Cas9 providing PAM specificity (Heler et al. 2015; Ka et al. 2018).

Structural analysis of the Cas proteins of the CRISPR3 type II-A system of *S. thermophilus* revealed a complex consisting of $Cas1_8$-$Cas2_4$-$Csn2_8$. Although not detected in the complex, Cas9 was hypothesised to interact with Csn2 through DNA (Wilkinson et al. 2019). Csn2 has been shown to bind double-stranded DNA ends (Ellinger et al. 2012; Lee et al. 2012; Nam et al. 2012). It was therefore proposed that the $Cas1_8$-$Cas2_4$-$Csn2_8$ complex binds the free ends of prespacers generated by the AddAB repair machinery through Csn2. Then, the complex slides down the DNA until it encounters Cas9 that is bound to a PAM, and then, the spacer may be cut from the DNA (Wilkinson et al. 2019). This process is highlighted in Figure 3. In the interference stage, Cas9 is the sole protein required in both CRISPR1 and CRISPR3 as it cuts the target DNA within the protospacer, three nucleotides upstream of the PAM (12, 61).

Spacer acquisition events have typically been observed at the 5' end of the CRISPR locus, flanking the leader sequence. However, ectopic spacer acquisition has also been observed in *S. thermophilus* (60). This phenomenon was first described in *S. pyogenes* and refers to the acquisition of spacers at positions other than the 5'-end (McGinn and Marraffini 2016). The leader sequence in strains that acquired ectopic spacers was compared to strains that had not acquired ectopic spacers. A deletion of one nucleotide was discovered within the leader sequence directly upstream of the CRISPR locus in the strains that acquired ectopic spacers (Achigar et al. 2017). This region of the leader sequence is where the Cas proteins involved in adaptation interact with the CRISPR locus, likely explaining the reduced ability of these strains to integrate spacers at this position (Wei et al. 2015a). Furthermore, an AG motif was found at the 5' end of the spacer upstream of the newly acquired ectopic spacers. This AG

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

motif may play a role in guiding the adaptation module to this specific position (Achigar et al. 2017).

Anti-CRISPR

Phages have evolved several ways to cope with a wide range of anti-viral defence systems. This on-going evolutionary battle has been called an arms race as every evolutionary step, from either bacteria or phages, drives a new evolutionary step for the other (Safari et al. 2020). While escaping CRISPR immunity can be facilitated by a single mutation or a deletion within the protospacer or PAM sequence as mentioned previously (Deveau et al. 2008), it is more challenging to replicate in a diverse population of cells globally carrying a large number of different spacers. It has been hypothesised that this limitation led to the evolution of a more general method to disable the CRISPR immunity in the form of anti-CRISPR proteins (ACRs) (Van Houte et al. 2016).

The first anti-CRISPR genes were discovered in temperate phages infecting *P. aeruginosa* (94). While integrated into the *P. aeruginosa* genome as a prophage, the phage expressed proteins that interfered with the activity of the host CRISPR-Cas system. Since these phage proteins did not interfere with the transcription of *cas* genes and the CRISPR locus, the anti-CRISPR activity was found to take place after the formation of the crRNA-Cascade complex (Bondy-Denomy et al. 2013). The activity of these anti-CRISPR proteins was against type I-F systems only, but soon after, additional anti-CRISPR proteins that exhibited activity against type I-E systems were also discovered (Pawluk et al. 2014). At the sequence level, a signature motif that allowed for the identification of anti-CRISPR genes in other phages and archaeal viruses were found. This helix-turn-helix motif was found in proteins for which the genes were present directly downstream from anti-CRISPR genes. The anti-CRISPR associated

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

401 (*aca*) genes, which code for transcription regulators, helped to identify new anti-CRISPR
402 genes for other systems, including type II using bioinformatic approaches.
403
404 The first type II anti-CRISPR genes were found in prophages of *Neisseria meningitides*. They
405 interfered with the nuclease activity of the type II-C Cas9 in *N. meningitides* but not with the
406 type II-A Cas9 of *S. pyogenes* (Pawluk et al. 2016). Anti-CRISPR genes interfering with type
407 II-A CRISPR-Cas systems were found using other approaches. One such approach was to
408 screen for prophage sequences that carry a sequence identical to one of the spacers present in
409 the cell. This would typically result in self-targeting in the case of an active CRISPR-Cas
410 system, but would be prevented by the expression of an ACR protein. Using this strategy,
411 several new *acr* genes were discovered in prophages of *Listeria monocytogenes,* which were
412 capable of inhibiting SpCas9 (Rauch et al. 2017). The first anti-CRISPR genes discovered in
413 virulent phages were found in *S. thermophilus*. These ACR-containing phages were
414 discovered because phage-resistant mutants were extremely rare following some phage-
415 bacteria interactions while some phages were able to infect BIMs carrying spacers targeting
416 their genomes (Hynes et al. 2017, 2018). Additional anti-CRISPR proteins targeting type II
417 systems were discovered by the transformation of metagenomic libraries, after which the
418 ACR activity against *S. pyogenes* Cas9 was confirmed *in vitro* and *in vivo* (Uribe et al. 2019).
419
420 Several mechanisms for anti-CRISPR proteins have also been discovered, revealing that
421 different ACR proteins target different Cas proteins. For example, binding of the ACR to Cas
422 proteins prevented binding or cleaving of nucleic acids (Bondy-Denomy et al. 2015; Dong et
423 al. 2017; Harrington et al. 2017; Shin et al. 2017; Yang and Patel 2017; Hardouin and Goulet
424 2020; Marino et al. 2020; Davidson et al. 2020). A nomenclature for ACRs has been proposed
425 and an online spreadsheet is available to avoid names being used twice in this fast growing

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

field (Bondy-Denomy et al. 2018). The number of anti-CRISPR genes that has been discovered has expanded dramatically and so has the number of targeted system types. To date, over 50 families of anti-CRISPR proteins have been described that inhibit various types of CRISPR-Cas systems, and the search for new anti-CRISPR proteins still continues (Stanley and Maxwell 2018; Davidson et al. 2020).

Conclusion

A few decades have passed since the incidental discovery of a CRISPR locus. Several research groups have uncovered many of its mysteries and it has turned out to be one of the most fascinating discoveries in microbiology. Currently, six CRISPR-Cas system types have been described as well as several subtypes (Makarova et al. 2019) and many additional variants likely await discovery. The development of the ever-growing CRISPR-Cas tools for genome editing was due, in part, to a better understanding of the crRNA synthesis and interference steps. This CRISPR frenzy as a genome editing tool resulted in more research into these two steps of the CRISPR-Cas biology. As such, studies on the adaptation phase have been limited for the most part. The experimental demonstration of spacer adaptation remains elusive for several CRISPR-Cas subtypes as it still appears to be a rare event under laboratory conditions.

While a very efficient crRNA synthesis or interference stage is needed to readily protect the cell against invasion of foreign nucleic acids, a spacer acquisition process in overdrive would be detrimental to the cell as self-targeting would likely occur, thereby killing the cells. The presence of *acr* genes in the bacterial genome (as prophages) or on plasmids are also likely responsible for these apparent experimental frustrations associated with spacer acquisition assays. The use of ACR-containing virulent phages in these experiments may also have been

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

playing a role. The recent observation of spacer acquisition in a bacterial community (Alseth et al. 2019) opens new avenues and could draw more attention to the adaptation phase.

Although CRISPR-Cas is considered a defence mechanism in a cellular context, other functions have been proposed and described (Faure et al. 2019). Under certain circumstances, CRISPR-Cas has been associated with gene regulation, genome evolution, DNA repair and cooperation with other anti-viral systems (Westra et al. 2014). As the number of publications on CRISPR-Cas biology increases each year, novel functions and applications will likely be discovered in the future. It will be interesting to see if Cas proteins involved in the adaptation phase also play other functions. Cas1 was found to increase the resistance of *Mycobacterium smegmatis* towards multiple anti-tuberculosis drugs (Wei et al. 2019) and has also been associated with DNA repair (Babu et al. 2011). In *Legionella pneumophila*, Cas2 has been shown to be an essential factor to facilitate the infection of amoebae (Gunderson and Cianciotto 2013; Gunderson et al. 2015). New functions for these versatile proteins are likely waiting to be discovered.

Acknowledgments

We would like to Michi Waygoo and Amanda Toperoff for editorial assistance. S.M. acknowledges funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Program. S.M. holds a Tier 1 Canada Research Chair in Bacteriophages.

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E.S., Koonin, E. V., and Zhang, F. 2016. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science **353**(6299). doi:10.1126/science.aaf5573.

Achigar, R., Magadán, A.H., Tremblay, D.M., Julia Pianzzola, M., and Moineau, S. 2017. Phage-host interactions in *Streptococcus thermophilus*: genome analysis of phages isolated in Uruguay and ectopic spacer acquisition in CRISPR array. Sci. Rep. **7**: 43438. doi:10.1038/srep43438.

Agari, Y., Sakamoto, K., Tamakoshi, M., Oshima, T., Kuramitsu, S., and Shinkai, A. 2010. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. J. Mol. Biol. **395**(2): 270-281. doi:10.1016/j.jmb.2009.10.057.

Almendros, C., Nobrega, F.L., McKenzie, R.E., and Brouns, S.J.J. 2019. Cas4-Cas1 fusions drive efficient PAM selection and control CRISPR adaptation. Nucleic Acids Res. **47**(10): 5223-5230. doi:10.1093/nar/gkz217.

Alseth, E.O., Pursey, E., Luján, A.M., McLeod, I., Rollie, C., and Westra, E.R. 2019. Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. Nature **574**(7779): 549-552. doi:10.1038/s41586-019-1662-9.

Amitai, G., and Sorek, R. 2016. CRISPR-Cas adaptation: Insights into the mechanism of action. Nat. Rev. Microbiol. **14**(2): 67–76. doi:10.1038/nrmicro.2015.14.

Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Pogoutse, O., Brown, G., Binkowski, A., Phanse, S., Joachimiak, A., Koonin, E. V., Savchenko, A., Emili, A., Greenblatt, J., Edwards, A.M., and Yakunin, A.F. 2011. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. Mol. Microbiol. **79**(2): 484–502. doi:10.1111/j.1365-2958.2010.07465.x.

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**(5819): 1709–1712. doi:10.1126/science.1138140.

Beloglazova, N., Brown, G., Zimmerman, M.D., Proudfoot, M., Makarova, K.S., Kudritska, M., Kochinyan, S., Wang, S., Chruszcz, M., Minor, W., Koonin, E. V., Edwards, A.M., Savchenko, A., and Yakunin, A.F. 2008. A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. J. Biol. Chem. **283**(29): 20361–20371. doi:10.1074/jbc.M803225200.

Bolotin, A., Quinquis, B., Sorokin, A., and Dusko Ehrlich, S. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology **151**(8): 2551–2561. doi:10.1099/mic.0.28048-0.

Bondy-Denomy, J., Davidson, A.R., Doudna, J.A., Fineran, P.C., Maxwell, K.L., Moineau, S., Peng, X., Sontheimer, E.J., and Wiedenheft, B. 2018. A Unified Resource for Tracking Anti-CRISPR Names. Cris. J. **1**(5). doi:10.1089/crispr.2018.0043.

Bondy-Denomy, J., Garcia, B., Strum, S., Du, M., Rollins, M.F., Hidalgo-Reyes, Y., Wiedenheft, B., Maxwell, K.L., and Davidson, A.R. 2015. Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. Nature **526**(7571): 136–139. NIH Public Access. doi:10.1038/nature15254.

Bondy-Denomy, J., Pawluk, A., Maxwell, K.L., and Davidson, A.R. 2013. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature **493**(7432): 429–432. doi:10.1038/nature11723.

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

519 Brouns. 2008. Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. Science **321**(5891):
520 960-964. doi:10.1126/science.1229223.
521 Chabas, H., Van Houte, S., Høyland-Kroghsbo, N.M., Buckling, A., and Westra, E.R. 2016.
522 Immigration of susceptible hosts triggers the evolution of alternative parasite defence
523 strategies. Proc. R. Soc. B Biol. Sci. **283**(1837): 20160721. doi:10.1098/rspb.2016.0721.
524 Charpentier, E., Richter, H., van der Oost, J., and White, M.F. 2015. Biogenesis pathways of
525 RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. FEMS Microbiol.
526 Rev. **39**(3): 428-441. doi:10.1093/femsre/fuv023.
527 Chylinski, K., Le Rhun, A., and Charpentier, E. 2013. The tracrRNA and Cas9 families of type II
528 CRISPR-Cas immunity systems. RNA Biol. **10**(5): 726–737. doi:10.4161/rna.24321.
529 Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., and Semenova, E.
530 2012. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial
531 immunity system. Nat. Commun. **3**(945). doi:10.1038/ncomms1937.
532 Davidson, A.R., Lu, W.-T., Stanley, S.Y., Wang, J., Mejdani, M., Trost, C.N., Hicks, B.T., Lee, J.,
533 and Sontheimer, E.J. 2020. Anti-CRISPRs: Protein inhibitors of CRISPR-Cas systems.
534 Annu. Rev. Biochem. doi:10.1146/annurev-biochem-011420-111224.
535 Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R.,
536 Vogel, J., and Charpentier, E. 2011. CRISPR RNA maturation by trans-encoded small RNA
537 and host factor RNase III. Nature **471**(7340): 602–607. doi:10.1038/nature09886.
538 Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A.,
539 Horvath, P., and Moineau, S. 2008. Phage response to CRISPR-encoded resistance in
540 *Streptococcus thermophilus*. J. Bacteriol. **190**(4): 1390–1400. doi:10.1128/JB.01412-07.
541 Díez-Villaseñor, C., Almendros, C., García-Martínez, J., and Mojica, F.J.M. 2010. Diversity of
542 CRISPR loci in *Escherichia coli*. Microbiology **156**(5): 1351–1361.
543 doi:10.1099/mic.0.036046-0.
544 Díez-Villaseñor, C., Guzmán, N.M., Almendros, C., García-Martínez, J., and Mojica, F.J.M.
545 2013. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition
546 specificities among CRISPR-Cas I-E variants of *Escherichia coli*. RNA Biol. **10**(5): 792–802.
547 doi:10.4161/rna.24023.
548 Dong, D., Guo, M., Wang, S., Zhu, Y., Wang, S., Xiong, Z., Yang, J., Xu, Z., and Huang, Z. 2017.
549 Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. Nature
550 **546**(7658): 436–439. doi:10.1038/nature22377.
551 Drabavicius, G., Sinkunas, T., Silanskas, A., Gasiunas, G., Venclovas, Č., and Siksnys, V. 2018.
552 DnaQ exonuclease-like domain of Cas2 promotes spacer integration in a type I-E
553 CRISPR-Cas system. EMBO Rep. **19**(7):e45543. doi:10.15252/embr.201745543.
554 Dupuis, M.È., Villion, M., Magadán, A.H., and Moineau, S. 2013. CRISPR-Cas and restriction-
555 modification systems are compatible and increase phage resistance. Nat. Commun.
556 **4**(2087). doi:10.1038/ncomms3087.
557 Ellinger, P., Arslan, Z., Wurm, R., Tschapek, B., MacKenzie, C., Pfeffer, K., Panjikar, S.,
558 Wagner, R., Schmitt, L., Gohlke, H., Pul, Ü., and Smits, S.H.J. 2012. The crystal structure
559 of the CRISPR-associated protein Csn2 from *Streptococcus agalactiae*. J. Struct. Biol.
560 **178**(3):350-362. doi:10.1016/j.jsb.2012.04.006.
561 Erdmann, S., and Garrett, R.A. 2012. Selective and hyperactive uptake of foreign DNA by
562 adaptive immune systems of an archaeon via two distinct mechanisms. Mol. Microbiol.
563 **85**(6): 1044–1056. doi:10.1111/j.1365-2958.2012.08171.x.
564 Faure, G., Makarova, K.S., and Koonin, E. V. 2019. CRISPR–Cas: complex functional networks
565 and multiple roles beyond adaptive immunity. J. Mol. Bio **431**(1): 3-20.

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

doi:10.1016/j.jmb.2018.08.030.

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**(7320): 67–71. doi:10.1038/nature09523.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. **109**(39): E2579-86. doi:10.1073/pnas.1208507109.

Gunderson, F.F., and Cianciotto, N.P. 2013. The CRISPR-associated gene cas2 of *Legionella pneumophila* is required for intracellular infection of amoebae. MBio. **4**(2): e00074-13. doi:10.1128/mBio.00074-13.

Gunderson, F.F., Mallama, C.A., Fairbairn, S.G., and Cianciotto, N.P. 2015. Nuclease activity of *Legionella pneumophila* Cas2 promotes intracellular infection of amoebal host cells. Infect. Immun. **83**(3): 1008-1018. doi:10.1128/IAI.03102-14.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and Terns, M.P. 2009. RNA-Guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell **139**(5): 945–956. doi:10.1016/j.cell.2009.07.040.

Hardouin, P., and Goulet, A. 2020. Diversity of molecular mechanisms used by anti-CRISPR proteins: the tip of an iceberg? Biochem. Soc. Trans. **48**(2): 507-516. doi:10.1042/bst20190638.

Harrington, L.B., Doxzen, K.W., Ma, E., Liu, J.J., Knott, G.J., Edraki, A., Garcia, B., Amrani, N., Chen, J.S., Cofsky, J.C., Kranzusch, P.J., Sontheimer, E.J., Davidson, A.R., Maxwell, K.L., and Doudna, J.A. 2017. A broad-spectrum inhibitor of CRISPR-Cas9. Cell **170**(6): 1224-1233.e15. doi:10.1016/j.cell.2017.07.037.

He, Y., Wang, M., Liu, M., Huang, L., Liu, C., Zhang, X., Yi, H., Cheng, A., Zhu, D., Yang, Q., Wu, Y., Zhao, X., Chen, S., Jia, R., Zhang, S., Liu, Y., Yu, Y., and Zhang, L. 2018. Cas1 and Cas2 from the Type II-C CRISPR-Cas system of *Riemerella anatipestifer* are required for spacer acquisition. Front. Cell. Infect. Microbiol. **8**: 195. doi:10.3389/fcimb.2018.00195.

Heler, R., Samai, P., Modell, J.W., Weiner, C., Goldberg, G.W., Bikard, D., and Marraffini, L.A. 2015. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. Nature **519**(7542): 199–202. doi:10.1038/nature14245.

Hooton, S.P.T., and Connerton, I.F. 2015. *Campylobacter jejuni* acquire new host-derived CRISPR spacers when in association with bacteriophages harboring a CRISPR-like Cas4 protein. Front. Microbiol. **5**: 744. doi:10.3389/fmicb.2014.00744.

Horvath, P., Romero, D.A., Coûté-Monvoisin, A.C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., and Barrangou, R. 2008. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. J. Bacteriol. **190**(4): 1401–1412. doi:10.1128/JB.01415-07.

Van Houte, S., Ekroth, A.K.E., Broniewski, J.M., Chabas, H., Ashby, B., Bondy-Denomy, J., Gandon, S., Boots, M., Paterson, S., Buckling, A., and Westra, E.R. 2016. The diversity-generating benefits of a prokaryotic adaptive immune system. Nature **532**(7599): 385–388. doi:10.1038/nature17436.

Hynes, A.P., Rousseau, Lemay, M.-L., Horvath, P., Romero, D.A., Fremaux, C., and Moineau, S. 2017. An anti-CRISPR from a virulent streptococcal phage inhibits *Streptococcus pyogenes* Cas9. Nat. Microbiol. **2**(10): 1374-1380. doi:10.1038/s41564-017-0004-7.

Hynes, A.P., Rousseau, G.M., Agudelo, D., Goulet, A., Amigues, B., Loehr, J., Romero, D.A., Fremaux, C., Horvath, P., Doyon, Y., Cambillau, C., and Moineau, S. 2018. Widespread

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

anti-CRISPR proteins in virulent bacteriophages inhibit a range of Cas9 proteins. Nat. Commun. **9**(1): 2919. doi:10.1038/s41467-018-05092-w.

Hynes, A.P., Villion, M., and Moineau, S. 2014. Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. Nat. Commun. **5**: 4399. doi:10.1038/ncomms5399.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. J. Bacteriol. **169**(12): 5429–5433. doi:10.1128/jb.169.12.5429-5433.1987.

Jackson, R.N., and Wiedenheft, B. 2015. A conserved structural chassis for mounting versatile CRISPR RNA-guided immune responses. Mol. Cell. **58**(5): 722-728. doi:10.1016/j.molcel.2015.05.023.

Jackson, S.A., Birkholz, N., Malone, L.M., and Fineran, P.C. 2019. Imprecise spacer acquisition generates CRISPR-Cas immune diversity through primed adaptation. Cell Host Microbe **25**(2): 250-264.e4.. doi:10.1016/j.chom.2018.12.014.

Jansen, R., Embden, J.D.A. van, Gaastra, W., and Schouls, L.M. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. **43**(6): 1565–1575. doi:10.1046/j.1365-2958.2002.02839.x.

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**(6096): 816–821. doi:10.1126/science.1225829.

Ka, D., Jang, D.M., Han, B.W., and Bae, E. 2018. Molecular organization of the type II-A CRISPR adaptation module and its interaction with Cas9 via Csn2. Nucleic Acids Res. **46**(18): 9805–9815. doi:10.1093/nar/gky702.

Kieper, S.N., Almendros, C., Behler, J., McKenzie, R.E., Nobrega, F.L., Haagsma, A.C., Vink, J.N.A., Hess, W.R., and Brouns, S.J.J. 2018. Cas4 facilitates PAM-compatible spacer selection during CRISPR adaptation. Cell Rep. **22**(13): 3377-3384. doi:10.1016/j.celrep.2018.02.103.

Kim, S., Loeff, L., Colombo, S., Jergic, S., Brouns, S.J.J., and Joo, C. 2020. Selective loading and processing of prespacers for precise CRISPR adaptation. Nature. **579**(7797): 141-145. doi:10.1038/s41586-020-2018-1.

Kim, T.Y., Shin, M., Huynh Thi Yen, L., and Kim, J.S. 2013. Crystal structure of Cas1 from *Archaeoglobus fulgidus* and characterization of its nucleolytic activity. Biochem. Biophys. Res. Commun. **441**(4): 720–725. doi:10.1016/j.bbrc.2013.10.122.

Koonin, E. V., Makarova, K.S., and Zhang, F. 2017. Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. **37**: 67-78. doi:10.1016/j.mib.2017.05.008.

Künne, T., Kieper, S.N., Bannenberg, J.W., Vogel, A.I.M., Miellet, W.R., Klein, M., Depken, M., Suarez-Diez, M., and Brouns, S.J.J. 2016. Cas3-derived target DNA degradation fragments fuel primed CRISPR adaptation. Mol. Cell **63**(5): 852–864. doi:10.1016/j.molcel.2016.07.011.

Lee, H., Dhingra, Y., and Sashital, D.G. 2019. The Cas4-Cas1-Cas2 complex mediates precise prespacer processing during CRISPR adaptation. Elife. **8**: e44248. doi:10.7554/eLife.44248.

Lee, H., Zhou, Y., Taylor, D.W., and Sashital, D.G. 2018. Cas4-dependent prespacer processing ensures high-fidelity programming of CRISPR arrays. Mol. Cell. **70**(1): 48–59.e5. doi:10.1016/j.molcel.2018.03.003.

Lee, K.H., Lee, S.G., Eun Lee, K., Jeon, H., Robinson, H., and Oh, B.H. 2012. Identification, structural, and biochemical characterization of a group of large Csn2 proteins involved

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

in CRISPR-mediated bacterial immunity. Proteins **80**(11):2573-2582. doi:10.1002/prot.24138.

Levy, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U., and Sorek, R. 2015. CRISPR adaptation biases explain preference for acquisition of foreign DNA. Nature **520**(7548): 505–510. doi:10.1038/nature14302.

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**(7): e40913. doi:10.1371/journal.pone.0040913.

Makarova, K.S., Aravind, L., Wolf, Y.I., and Koonin, E. V. 2011. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. Biol. Direct **6**: 38. doi:10.1186/1745-6150-6-38.

Makarova, K.S., Grishin, N. V., Shabalina, S.A., Wolf, Y.I., and Koonin, E. V. 2006. A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol. Direct **1**: 7. doi:10.1186/1745-6150-1-7.

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., Horvath, P., Moineau, S., Mojica, F.J.M., Terns, R.M., Terns, M.P., White, M.F., Yakunin, A.F., Garrett, R.A., Van Der Oost, J., Backofen, R., and Koonin, E. V. 2015. An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. **13**(11): 722–736. doi:10.1038/nrmicro3569.

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., Moineau, S., Mojica, F.J.M., Scott, D., Shah, S.A., Siksnys, V., Terns, M.P., Venclovas, Č., White, M.F., Yakunin, A.F., Yan, W., Zhang, F., Garrett, R.A., Backofen, R., van der Oost, J., Barrangou, R., and Koonin, E. V. 2019. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol. **18**(2):67-83. doi:10.1038/s41579-019-0299-x.

Makarova, K.S., Wolf, Y.I., and Koonin, E. V. 2018. Classification and nomenclature of CRISPR-Cas systems: where from here? Cris. J. **1**(5): 325–336. doi:10.1089/crispr.2018.0033.

Marino, N.D., Pinilla-Redondo, R., Csörgő, B., and Bondy-Denomy, J. 2020. Anti-CRISPR protein applications: natural brakes for CRISPR-Cas technologies. Nat. Methods **17**(5): 471-479. doi:10.1038/s41592-020-0771-6.

Marraffini, L.A., and Sontheimer, E.J. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science **322**(5909): 1843–1845. doi:10.1126/science.1165771.

McGinn, J., and Marraffini, L.A. 2016. CRISPR-Cas systems optimize their immune response by specifying the site of spacer integration. Mol. Cell **64**(3): 616–623. doi:10.1016/j.molcel.2016.08.038.

Medina-Aparicio, L., Rebollar-Flores, J.E., Gallego-Hernández, A.L., Vázquez, A., Olvera, L., Gutiérrez-Ríos, R.M., Calva, E., and Hernández-Lucas, I. 2011. The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. J. Bacteriol. **193**(10): 2396-407. doi:10.1128/JB.01480-10.

Modell, J.W., Jiang, W., and Marraffini, L.A. 2017. CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. Nature **544**(7648): 101–104. doi:10.1038/nature21719.

Mojica, F.J., Juez, G., and Rodríguez-Valera, F. 1993. Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. Mol.

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

Microbiol. **9**(3): 613–621. doi:10.1111/j.1365-2958.1993.tb01721.x.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. **60**(2): 174–182. doi:10.1007/s00239-004-0046-3.

Nam, K.H., Ding, F., Haitjema, C., Huang, Q., DeLisa, M.P., and Ke, A. 2012. Double-stranded endonuclease activity in *Bacillus halodurans* clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas2 protein. J. Biol. Chem. **287**(43): 35943–35952. doi:10.1074/jbc.M112.382598.

Nicholson, T.J., Jackson, S.A., Croft, B.I., Staals, R.H.J., Fineran, P.C., and Brown, C.M. 2019. Bioinformatic evidence of widespread priming in type I and II CRISPR-Cas systems. RNA Biol. **16**(4):566-576. doi:10.1080/15476286.2018.1509662.

Nuñez, J.K., Kranzusch, P.J., Noeske, J., Wright, A. V., Davies, C.W., and Doudna, J.A. 2014. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. Nat. Struct. Mol. Biol. **21**(6): 528–534. doi:10.1038/nsmb.2820.

Nussenzweig, P.M., McGinn, J., and Marraffini, L.A. 2019. Cas9 cleavage of viral genomes primes the acquisition of new immunological memories. Cell Host Microbe. **26**(4):515-526.e6. doi:10.1016/j.chom.2019.09.002.

Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M., Sontheimer, E.J., Maxwell, K.L., and Davidson, A.R. 2016. Naturally occurring off-switches for CRISPR-Cas9. Cell **167**(7): 1829-1838.e9. doi:10.1016/j.cell.2016.11.017.

Pawluk, A., Bondy-Denomy, J., Cheung, V.H.W., Maxwell, K.L., and Davidson, A.R. 2014. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. MBio **5**(2): e00896. doi:10.1128/mBio.00896-14.

van der Ploeg, J.R. 2009. Analysis of CRISPR in *Streptocccus mutans* suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. Microbiology **155**(6): 1966–1976. doi:10.1099/mic.0.027508-0.

Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K.A., Djordjevic, M., Wanner, B.L., and Severinov, K. 2010. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. Mol. Microbiol. **77**(6): 1367–1379. doi:10.1111/j.1365-2958.2010.07265.x.

Pourcel, C., Salvignol, G., and Vergnaud, G. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology **151**(3): 653–663. doi:10.1099/mic.0.27437-0.

Pul, Ü., Wurm, R., Arslan, Z., Geißen, R., Hofmann, N., and Wagner, R. 2010. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. Mol. Microbiol. **75**(6): 1495–1512. doi:10.1111/j.1365-2958.2010.07073.x.

Pyenson, N.C., and Marraffini, L.A. 2020. Co-evolution within structured bacterial communities results in multiple expansion of CRISPR loci and enhanced immunity. Elife **9**: e53078. doi:10.7554/eLife.53078.

Quax, T.E.F., Voet, M., Sismeiro, O., Dillies, M.-A., Jagla, B., Coppee, J.-Y., Sezonov, G., Forterre, P., van der Oost, J., Lavigne, R., and Prangishvili, D. 2013. Massive activation of archaeal defense genes during viral infection. J. Virol. **87**(15): 8419-8428. doi:10.1128/jvi.01020-13.

Ramia, N.F., Tang, L., Cocozaki, A.I., and Li, H. 2014. *Staphylococcus epidermidis* Csm1 is a 3'-5' exonuclease. Nucleic Acids Res. **42**(2): 1129-1138. doi:10.1093/nar/gkt914.

Rauch, B.J., Silvis, M.R., Hultquist, J.F., Waters, C.S., McGregor, M.J., Krogan, N.J., and Bondy-Denomy, J. 2017. Inhibition of CRISPR-Cas9 with bacteriophage proteins. Cell **168**(1–2):

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

150-158.e10. doi:10.1016/j.cell.2016.12.009.

Richter, C., Dy, R.L., McKenzie, R.E., Watson, B.N.J., Taylor, C., Chang, J.T., McNeil, M.B., Staals, R.H.J., and Fineran, P.C. 2014. Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. Nucleic Acids Res. **42**(13): 8516–8526. doi:10.1093/nar/gku527.

Rollie, C., Graham, S., Rouillon, C., and White, M.F. 2018. NAR breakthrough article: Prespacer processing and specific integration in a type I-A CRISPR system. Nucleic Acids Res. **46**(3):1007-1020. doi:10.1093/nar/gkx1232.

Safari, F., Sharifi, M., Farajnia, S., Akbari, B., Karimi Baba Ahmadi, M., Negahdaripour, M., and Ghasemi, Y. 2020. The interaction of phages and bacteria: the co-evolutionary arms race. Crit Rev Biotechnol. **40**(2): 119-137. doi:10.1080/07388551.2019.1674774.

Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., and Marraffini, L.A. 2015. Co-transcriptional DNA and RNA cleavage during type III CRISPR-cas immunity. Cell **161**(5): 1164–1174. doi:10.1016/j.cell.2015.04.027.

Shiimori, M., Garrett, S.C., Graveley, B.R., and Terns, M.P. 2018. Cas4 nucleases define the PAM, length, and orientation of DNA fragments integrated at CRISPR loci. Mol. Cell. **70**(5): 814-824.e6. doi:10.1016/j.molcel.2018.05.002.

Shin, J., Jiang, F., Liu, J.J., Bray, N.L., Rauch, B.J., Baik, S.H., Nogales, E., Bondy-Denomy, J., Corn, J.E., and Doudna, J.A. 2017. Disabling Cas9 by an anti-CRISPR DNA mimic. Sci. Adv. **3**(7): e1701620.. doi:10.1126/sciadv.1701620.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., and Koonin, E. V. 2015. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. Mol. Cell **60**(3): 385–397. doi:10.1016/j.molcel.2015.10.008.

Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. 2011. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. EMBO J. **30**(7): 1335–1342. doi:10.1038/emboj.2011.41.

Sinkunas, T., Gasiunas, G., Waghmare, S.P., Dickman, M.J., Barrangou, R., Horvath, P., and Siksnys, V. 2013. In vitro reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. EMBO J. **32**(3): 385-394. doi:10.1038/emboj.2012.352.

Staals, R.H.J., Jackson, S.A., Biswas, A., Brouns, S.J.J., Brown, C.M., and Fineran, P.C. 2016. Interference dominates and amplifies spacer acquisition in a native CRISPR-Cas system. Nat. Commun. **7**: 12853. doi:10.1038/ncomms12853.

Stanley, S.Y., and Maxwell, K.L. 2018. Phage-encoded anti-CRISPR defenses. Annu. Rev. Genet. **52**(1): 445–464. doi:10.1146/annurev-genet-120417-031321.

Strecker, J., Jones, S., Koopal, B., Schmid-Burgk, J., Zetsche, B., Gao, L., Makarova, K.S., Koonin, E. V., and Zhang, F. 2019. Engineering of CRISPR-Cas12b for human genome editing. Nat. Commun. **10**(1):212. doi:10.1038/s41467-018-08224-4.

Swarts, D.C., and Jinek, M. 2019. Mechanistic insights into the cis- and trans-acting DNase activities of Cas12a. Mol. Cell. **73**(3): 589-600.e4. doi:10.1016/j.molcel.2018.11.021.

Swarts, D.C., Mosterd, C., van Passel, M.W.J., and Brouns, S.J.J. 2012. CRISPR interference directs strand specific spacer acquisition. PLoS One **7**(4): e35888. doi:10.1371/journal.pone.0035888.

Uribe, R. V., van der Helm, E., Misiakou, M.-A., Lee, S.-W., Kol, S., and Sommer, M.O.A. 2019. Discovery and characterization of Cas9 inhibitors disseminated across seven bacterial phyla. Cell Host Microbe **26**(5): 702. doi:10.1016/j.chom.2019.01.003.

Vale, P.F., Lafforgue, G., Gatchitch, F., Gardan, R., Moineau, S., Gandon, S. 2015. Costs of

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

CRISPR-Cas-mediated resistance in *Streptococcus thermophilus*. Proc. Biol. Sci. **282**(1812): 20151270. doi:10.1098/rspb.2015.1270.

Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M., and Wang, Y. 2015. Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. Cell **163**(4): 840–853. doi:10.1016/j.cell.2015.10.008.

Wei, J., Lu, N., Li, Z., Wu, X., Jiang, T., Xu, L., Yang, C., and Guo, S. 2019. The *Mycobacterium tuberculosis* CRISPR-Associated Cas1 involves persistence and tolerance to anti-tubercular drugs. Biomed Res. Int. **2019**: 7861695. doi:10.1155/2019/7861695.

Wei, Y., Chesne, M.T., Terns, R.M., and Terns, M.P. 2015a. Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in *Streptococcus thermophilus*. Nucleic Acids Res. **43**(3): 1749–1758. doi:10.1093/nar/gku1407.

Wei, Y., Terns, R.M., and Terns, M.P. 2015b. Cas9 function and host genome sampling in type II-A CRISPR–cas adaptation. Genes Dev. **29**(4): 356–361. doi:10.1101/gad.257550.114.

Westra, E.R., Buckling, A., and Fineran, P.C. 2014. CRISPR-Cas systems: Beyond adaptive immunity. Nat. Rev. Microbiol. **12**(5): 317–326. doi:10.1038/nrmicro3241.

Westra, E.R., Van houte, S., Oyesiku-Blakemore, S., Makin, B., Broniewski, J.M., Best, A., Bondy-Denomy, J., Davidson, A., Boots, M., and Buckling, A. 2015. Parasite exposure drives selective evolution of constitutive versus inducible defense. Curr. Biol. **25**(8): 1043-1049. doi:10.1016/j.cub.2015.01.065.

Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S.M., Ma, W., and Doudna, J.A. 2009. Structural Basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. Structure **17**(6): 904–912. doi:10.1016/j.str.2009.03.019.

Wilkinson, M., Drabavicius, G., Silanskas, A., Gasiunas, G., Siksnys, V., and Wigley, D.B. 2019. Structure of the DNA-Bound Spacer Capture Complex of a Type II CRISPR-Cas System. Mol. Cell. **75**(1): 90-101.e5. doi:10.1016/j.molcel.2019.04.020.

Yang, H., and Patel, D.J. 2017. Inhibition mechanism of an anti-CRISPR suppressor AcrIIA4 targeting SpyCas9. Mol. Cell **67**(1): 117-127.e5. doi:10.1016/j.molcel.2017.05.024.

Yoganand, K.N., Muralidharan, M., Nimkar, S., and Anand, B. 2019. Fidelity of prespacer capture and processing is governed by the PAM-mediated interactions of Cas1-2 adaptation complex in CRISPR-Cas type I-E system. J. Biol. Chem. **294**(52): 20039-20053. doi:10.1074/jbc.RA119.009438.

Yosef, I., Goren, M.G., and Qimron, U. 2012. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. Nucleic Acids Res. **40**(12): 5569–5576. doi:10.1093/nar/gks216.

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**(5): e38077. doi:10.1371/journal.pone.0038077.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., Van Der Oost, J., Regev, A., Koonin, E. V., and Zhang, F. 2015. Cpf1 Is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell **163**(3): 759–771. doi:10.1016/j.cell.2015.09.038.

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

Figure 1: Classification of CRISPR-Cas systems. In this figure an example is given for each CRISPR system type. The signature genes are highlighted in orange. The other genes are blue. The Type III-B and Type VI-A systems do not possess their own *cas1* and *cas2* genes, but use those of other systems and these genes are therefore presented in white with dotted lines. The CRISPR locus is illustrated in the form of red diamonds (repeats) and green squares (spacers).

Figure 2: Comparison of Cas9 from *Streptococcus thermophilus* and *Streptococcus pyogenes*. The percentage identity of the Cas proteins from CR1 of *S. thermophilus* DGCC7710 and *S. pyogenes* M1 compared to the Cas proteins of C3 of *S. thermophilus* DGCC7710 are shown in the figure. Within Cas9, the RuvC domains are highlighted in yellow, the recognition lobe in marine blue, the HNH domains in orange and the PAM recognition domain in purple.

Figure 3: Spacer acquisition in CRISPR3 (Type II-A) of *Streptococcus thermophilus*. The CRISPR locus of *S. thermophilus* DGCC7710CRISPR3 consists of 13 repeats (red diamonds) and 12 spacers (green squares). The acquisition complex consists of 4 Cas1 dimers (blue), 2 Cas2 dimers (yellow) and 2 Csn2 tetramers (purple) and likely slides down the viral DNA until it encounters Cas9 (orange) bound to a PAM sequence (black). The flanking protospacer (white square) is somehow cut from the viral DNA and is ready for integration into the CRISPR locus below as a novel spacer.

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

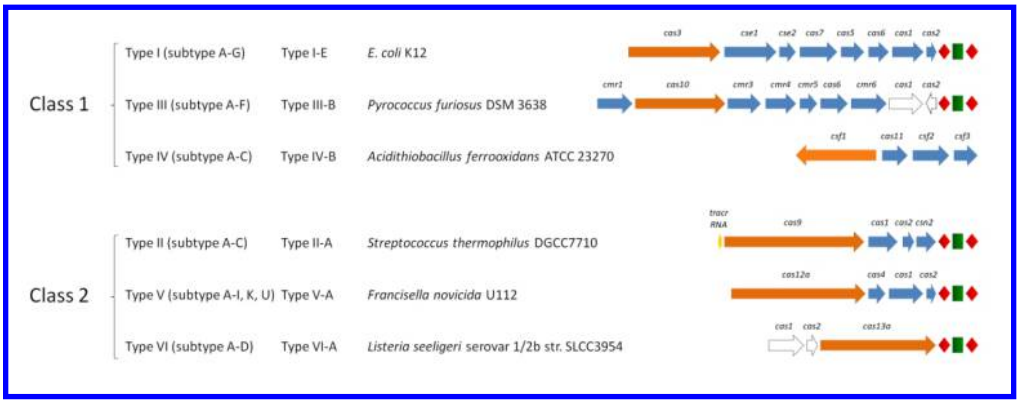


Figure 1: Classification of CRISPR-Cas systems. In this figure an example is given for each CRISPR system type. The signature genes are highlighted in orange. The other genes are blue. The Type III-B and Type VI-A systems do not possess their own cas1 and cas2 genes, but use those of other systems and these genes are therefore presented in white with dotted lines. The CRISPR locus is illustrated in the form of red diamonds (repeats) and green squares (spacers).

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

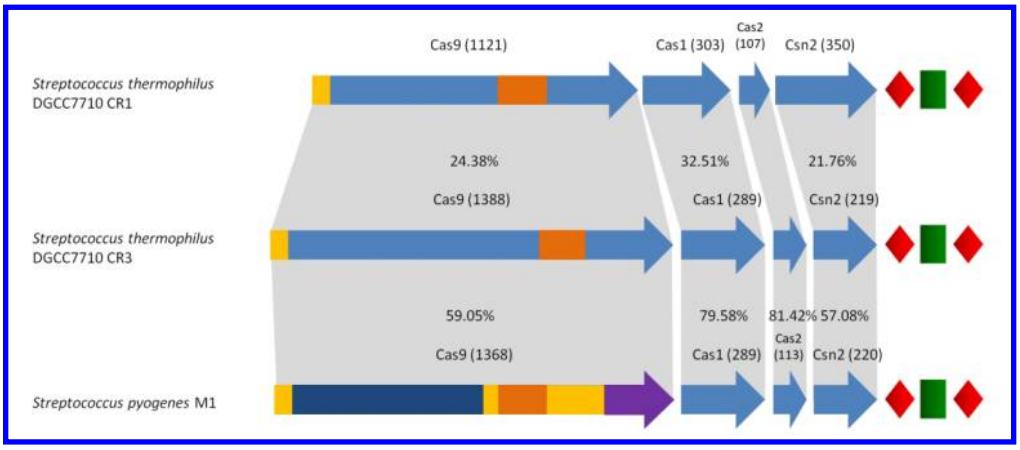


Figure 2: Comparison of Cas9 from Streptococcus thermophilus and Streptococcus pyogenes. The percentage identity of the Cas proteins from CR1 of S. thermophilus DGCC7710 and S. pyogenes M1 compared to the Cas proteins of C3 of S. thermophilus DGCC7710 are shown in the figure. Within Cas9, the RuvC domains are highlighted in yellow, the recognition lobe in marine blue, the HNH domains in orange and the PAM recognition domain in purple.

Can. J. Microbiol. Downloaded from nrc-prod.literatumonline.com by 23.89.151.97 on 07/09/20
For personal use only. This Just-IN manuscript is the accepted manuscript prior to copy editing and page composition. It may differ from the final official version of record.

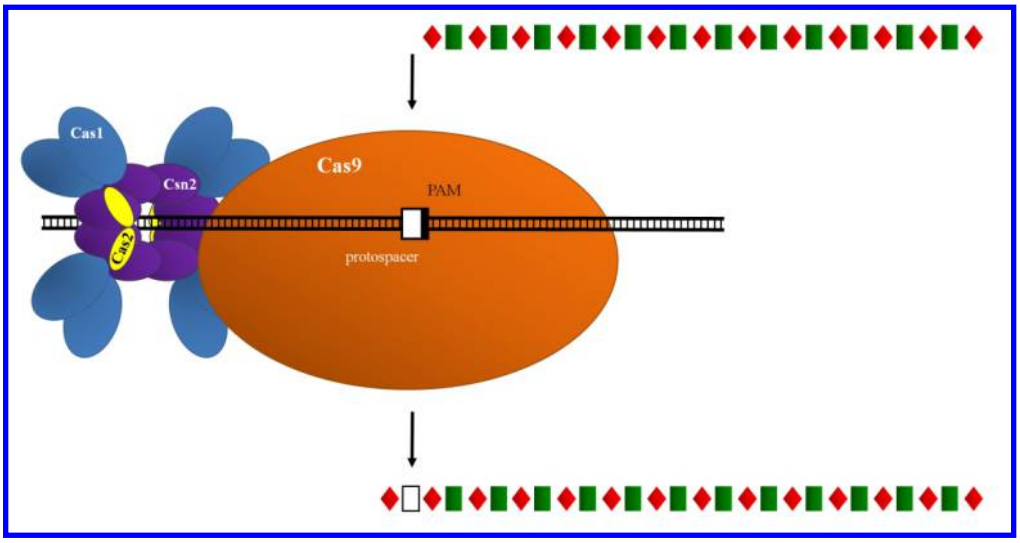


Figure 3: Spacer acquisition in CRISPR3 (Type II-A) of Streptococcus thermophilus. The CRISPR locus of S. thermophilus DGCC7710 CRISPR3 consists of 13 repeats (red diamonds) and 12 spacers (green squares). The acquisition complex consists of 4 Cas1 dimers (blue), 2 Cas2 dimers (yellow) and 2 Csn2 tetramers (purple) and likely slides down the viral DNA until it encounters Cas9 (orange) bound to a PAM sequence (black). The flanking protospacer (white square) is somehow cut from the viral DNA and is ready for integration into the CRISPR locus below as a novel spacer.