CRISPR: New Horizons in Phage Resistance and Strain Identification

Rodolphe Barrangou[1] and Philippe Horvath[2]

[1]Danisco USA, Inc., Madison, Wisconsin 53716; email: rodolphe.barrangou@danisco.com

[2]Danisco France SAS, Dangé-Saint-Romain F-86220, France

Annu. Rev. Food Sci. Technol. 2012. 3:143–62

First published online as a Review in Advance on December 20, 2011

The *Annual Review of Food Science and Technology* is online at food.annualreviews.org

This article's doi:
10.1146/annurev-food-022811-101134

1941-1413/12/0410-0143$20.00

Keywords

bacteriophage, *cas* gene, interference, typing, immunity, starter culture

Abstract

Bacteria have been widely used as starter cultures in the food industry, notably for the fermentation of milk into dairy products such as cheese and yogurt. Lactic acid bacteria used in food manufacturing, such as lactobacilli, lactococci, streptococci, *Leuconostoc*, pediococci, and bifidobacteria, are selectively formulated based on functional characteristics that provide idiosyncratic flavor and texture attributes, as well as their ability to withstand processing and manufacturing conditions. Unfortunately, given frequent viral exposure in industrial environments, starter culture selection and development rely on defense systems that provide resistance against bacteriophage predation, including restriction-modification, abortive infection, and recently discovered CRISPRs (clustered regularly interspaced short palindromic repeats). CRISPRs, together with CRISPR-associated genes (*cas*), form the CRISPR/Cas immune system, which provides adaptive immunity against phages and invasive genetic elements. The immunization process is based on the incorporation of short DNA sequences from virulent phages into the CRISPR locus. Subsequently, CRISPR transcripts are processed into small interfering RNAs that guide a multifunctional protein complex to recognize and cleave matching foreign DNA. Hypervariable CRISPR loci provide insights into the phage and host population dynamics, and new avenues for enhanced phage resistance and genetic typing and tagging of industrial strains.

INTRODUCTION

Bacteriophage: virus that infects bacteria; also known as a phage

Clustered regularly interspaced short palindromic repeats (CRISPRs): a novel DNA repeat family

CRISPR-associated proteins (Cas): diverse family of proteins directing CRISPR-encoded immunity

Spacer: hypervariable sequence between DNA repeats of CRISPR loci, derived from invasive nucleic acid

Interference: process by which invasive nucleic acid is targeted by crRNA-guided CASCADE

Historically, a diverse number of domesticated microbes have been exploited by humans to ferment natural resources into a wide array of palatable food products and beverages. Over time, food processing operations have improved together with technological advances, notably in microbiology. In the recent past, advances in bacterial genomics have provided insights into the genetic basis for desirable functional traits of industrial microbes, notably in the lactic acid bacteria widely used in the fermentation of dairy products such as yogurt and cheese. A tremendous amount of knowledge has been developed on the physiology of gram-positive bacteria that efficiently ferment carbohydrates into lactic acid while generating desirable organoleptic compounds through the catabolism of lipids and proteins into fatty acids and peptides. Although considerable attention is dedicated to the biochemical capabilities that provide desirable sensory attributes, such as flavor and texture, the sustainable use of bacterial starter cultures relies on their ability to overcome phage exposure in industrial settings.

Notwithstanding dramatic improvements in hygiene and industrial operations throughout the twentieth century, the dairy industry remains victimized by predatory bacteriophages that often contaminate raw milk (Brüssow 2001, Labrie et al. 2010). Their ability to survive pasteurization and their resistance to multiple sanitation measures have forced the industry to rely on bacteria with naturally occurring phage resistance systems for the formulation of starter cultures. Indeed, most current mesophilic lactococci and thermophilic streptococci used in culture rotation strategies exhibit phage resistance phenotypes, such as prevention of adsorption, blocking of phage DNA injection, restriction-modification, abortive infection, and the novel CRISPRs (clustered regularly interspaced short palindromic repeats). Recent advances in microbiology, molecular biology, and genomics of bacteria and phages have set the stage for the selection and development of sustainable industrial starter cultures that carry effective and enhanced phage resistance systems (Sturino & Klaenhammer 2006).

CRISPRs together with CRISPR-associated proteins (Cas), form the CRISPR/Cas system, which provides immunity against bacteriophages (Deveau et al. 2010, Horvath & Barrangou 2010, Karginov & Hannon 2010, Marraffini & Sontheimer 2010, Terns & Terns 2011, van der Oost et al. 2009). Although this novel DNA-repeat family, which is present in the genomes of many bacteria and most archaea (Mojica et al. 1995), was originally discovered in the intergenic region adjacent to the alkaline phosphatase (*iap*) gene in *Escherichia coli* K12 (Ishino et al. 1987), the CRISPR acronym was coined in 2002 (Jansen et al. 2002a,b), and its function remained elusive until recently. Although several putative roles were initially proposed based on in silico functional predictions for Cas proteins, the observation in 2005 that spacer sequences showed homology to extrachromosomal elements prompted the hypothesis that CRISPRs may provide immunity against invasive genetic elements (Bolotin et al. 2005, Mojica et al. 2005, Pourcel et al. 2005). Until then, the limited availability of phage and plasmid sequence information in public databases hindered the interpretation of the origin of CRISPR spacers. Subsequently, it was proposed in 2006 that CRISPR-encoded immunity may be mediated through RNA interference (Makarova et al. 2006). It was shown shortly thereafter that biologically, CRISPR loci incorporate novel spacers from viruses during the natural generation of phage-resistant mutants in bacteria and that there is a direct genetic link between spacer content and phage resistance based on sequence homology (Barrangou et al. 2007). Further, it was established that spacer acquisition and resistance rely on *cas* genes. Multiple genetic and biochemical studies in the following years revealed that CRISPR/Cas systems provide immunity against plasmids (Marraffini & Sontheimer 2008) and phages (Deveau et al. 2008, van der Ploeg 2009), and that the Cas machinery mediates small-RNA interference (Al-Attar et al. 2011, Brouns et al. 2008). A series of metagenomic surveys established

that CRISPR-mediated immunity plays a key role in host/virus population dynamics in natural communities and that CRISPR sequences provide historical and geographical insights (Andersson & Banfield 2008, Heidelberg et al. 2009, Held & Whitaker 2009, Shah & Garrett 2011, Tyson & Banfield 2008).

Interest in CRISPR/Cas systems has grown exponentially and prompted studies in multiple disciplines across phylogenetic branches. Recent advances in understanding the genetic and biochemical underpinnings of the mechanism of action of the CRISPR/Cas immune system have set the stage for implementation and industrial exploitation of CRISPR-encoded phage resistance, and the use of its hypervariable and adaptive nature for the typing of bacterial strains.

CRISPR database: Web site offering several tools to identify and analyze CRISPR loci

CRISPR/Cas SYSTEMS

Elements and Features of CRISPR/Cas Systems

CRISPR sequences constitute a novel DNA repeat family, which occurs in the genomes of many bacteria and most archaea. CRISPR loci consist of noncontiguous direct repeat sequences that are separated by unique sequences called spacers (Jansen et al. 2002b). These spacer sequences typically derive from foreign genetic elements, such as viruses and plasmids (Horvath & Barrangou 2010, Marraffini & Sontheimer 2010, Terns & Terns 2011, van der Oost et al. 2009). CRISPR repeats and spacers usually vary between 21–48 and 21–72 bp, respectively. The number of spacers within a particular locus can vary widely and reach several hundred (587 spacers have been identified in *Haliangium ochraceum* DSM 14365). Although repeat sequence can vary across loci, GTTTg/c and GAAAC sequences are common at the 5′ and 3′ end of the repeat, respectively (Deveau et al. 2010, Godde & Bickerton 2006, Jansen et al. 2002b, Kunin et al. 2007), providing partially palindromic sequences. Chromosomes may carry multiple CRISPR loci from distinct families, as documented in public CRISPR databases, notably CRISPRdb (Grissa et al. 2007) and CRISPI (Rousseau et al. 2009). Often, CRISPR-associated genes (*cas*) are present adjacent to the CRISPR repeat/spacer array (Haft et al. 2005, Jansen et al. 2002a). *Cas* genes encode a polymorphic family of proteins that contain functional domains involved in interaction with various nucleic acids, including nucleases, polymerases, helicases, and a variety of nucleotide-binding proteins (Haft et al. 2005; Makarova et al. 2006, 2011). Although initial functional predictions focused on involvement in chromosome partitioning and DNA repair (Makarova et al. 2002), Cas homologies with RNA interference machinery led to the hypothesis that CRISPR/Cas might be a defense system analogous to eukaryotic RNA interference (RNAi) (Makarova et al. 2006).

Those proteins, in combination with nucleic acids, form nucleoproteic complexes that carry out essential functions throughout the CRISPR/Cas immune system mechanism of action. Although polymorphism is observed in terms of *cas* gene number, sequence, type, and function across CRISPR/Cas systems, there are two universal genes, namely *cas1* and *cas2*, as well as core and signature genes that are idiosyncratic across the three CRISPR/Cas types (Makarova et al. 2011). In addition to the repeat/spacer array and *cas* genes, CRISPR/Cas system elements also include a leader sequence, which serves as a promoter for the transcription of the repeated elements (**Figure 1**, overview of the *Streptococcus thermophilus* CRISPR1/Cas system).

Classification and Occurrence

CRISPR/Cas systems were recently classified into three types that can be further divided into 10 subtypes (Makarova et al. 2011), based on a polythetic system that notably includes *cas* content. For each system, signature genes were established to specify each category. For particular systems, current knowledge is available about idiosyncratic genetic and biochemical processes, notably for

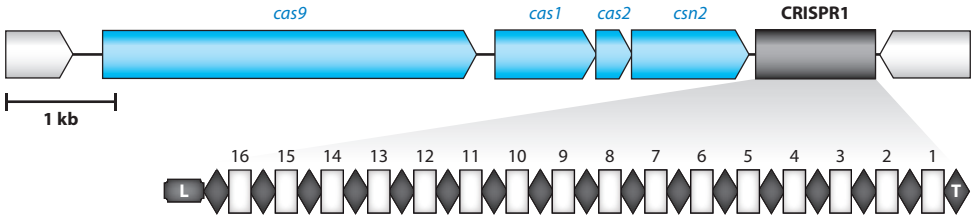


Figure 1

Overview of the *Streptococcus thermophilus* CRISPR1/Cas system locus in the LMD-9 genome. The *cas* genes are represented in blue, including universal markers *cas1* and *cas2*, the Type-II signature gene *cas9*, and *csn2*. The repeat-spacer array is represented in black (*top*). CRISPR repeats and spacers are represented as black diamonds and gray rectangles, respectively. Abbreviations: L, leader; T, terminal repeat. Numbers indicate the order in which spacers have been acquired over time.

S. thermophilus, *Pseudomonas aeruginosa*, *E. coli*, *Staphylococcus epidermidis*, and *Pyrococcus furiosus*, all of which have been investigated in a series of CRISPR studies. The occurrence of CRISPR/Cas systems in the genomes of bacteria commonly used in food cultures is highlighted and summarized in **Table 1**. Historically, attempts to classify CRISPR/Cas systems have focused on the content and sequences of *cas* genes and the proteins they encode. Unfortunately, given their sequence diversity, this has proven difficult (Haft et al. 2005; Makarova et al. 2006, 2011). Initial classifications relied on six core genes (Haft et al. 2005, Jansen et al. 2002a). A new classification system based on multiple criteria, including evolutionary relationships of conserved proteins and *cas* operon organization (Makarova et al. 2011), covers two partially independent subsystems: an information processing module based on universally present core proteins Cas1 and Cas2, involved in new spacer acquisition; and an executive subsystem responsible for processing primary CRISPR transcripts [CRISPR RNA (crRNA)] and interference with invasive nucleic acids. Based on these criteria, the new classification has established three distinct types based on locus organization and content. The Type I CRISPR/Cas systems are defined by the ubiquitous presence of a large signature gene, *cas3*, that encodes a helicase/nuclease, likely involved in R-loop-dependent target DNA cleavage in the interference step (Cady & O'Toole 2011, Jore et al. 2011a, Sinkunas et al. 2011). In addition, several genes (*cas5*, *cas6*, *cas7*) predicted to encode proteins of the interference complex are present in different combinations. With six different subtypes (Type I-A through Type I-F), this is the most diverse CRISPR/Cas system type. The Type II CRISPR/Cas systems typically contain the *cas9* signature gene, which encodes a large multifunctional protein with the ability to generate crRNA and targets phage and plasmid DNA for degradation (Garneau et al. 2010, Deltcheva et al. 2011). Cas9 contains a RuvC-like nuclease domain and an McrA-like HNH nuclease domain. The functional *S. thermophilus* CRISPR model is a Type II system that has been shown to provide defense against bacteriophage and plasmid DNA (Barrangou et al. 2007, Deveau et al. 2008, Garneau et al. 2010), in which *trans*-encoded small CRISPR RNA (tracrRNA) is involved in the processing of pre-CRISPR RNA (pre-crRNA) into crRNA (Deltcheva et al. 2011). The Type III CRISPR/Cas systems have a number of distinguishing features, notably the signature *cas10* gene, which encodes a repeat-associated mysterious protein (RAMP) involved in the processing of crRNA and target DNA cleavage (Anantharaman et al. 2010). Also, the *cas6* signature gene encodes an endoribonuclease involved in crRNA processing (Carte et al. 2008, Wang et al. 2011). So far, two distinct Type III subtypes have been distinguished: Type III-B, which includes the *P. furiosus* system targeting mRNA (Hale et al. 2008, Hale et al. 2009), and Type III-A, which includes the *S. epidermidis* system targeting DNA (Marraffini & Sontheimer 2008, 2010).

CRISPR RNA (crRNA): small noncoding ribonucleic acid mediating CRISPR-encoded immunity

***Trans*-encoded small CRISPR RNA (tracrRNA):** required for crRNA maturation in Type II systems

Pre-CRISPR RNA (pre-crRNA): full-length transcript produced from CRISPR repeat-spacer arrays

Repeat-associated mysterious protein (RAMP): diverse protein family involved in crRNA processing

Table 1 CRISPR/Cas systems in food cultures. Occurrence of CRISPR/Cas systems in the genomes of bacteria widely used in food cultures. Accession numbers are provided for the genome sequences. Each CRISPR/Cas system is characterized by a typical CRISPR repeat sequence, and the number of repeats in each CRISPR repeat-spacer array is provided

Genus, species, and strain name	Accession number(s) for chromosomal sequences	CRISPR/Cas systems	Number of repeats	Typical repeat sequence (5′-3′)
Arthrobacter arilaitensis RE117	FQ311875	0		
Bifidobacterium adolescentis ATCC 15703	AP009256	1	86+3	GTCGCTCTCCTTACGGAGAGCGTGGATTGAAAT
Bifidobacterium animalis subsp. *lactis* AD011 (+ BB-12, CNCM I-2494, DSM 10140, V9)	CP001213 (+ CP001853, CP002915, CP001606, CP001892)	1	20+2	ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
Bifidobacterium animalis subsp. *lactis* Bl-04	CP001515	1	23+2	ATCTCCGAAGTCTCGGCTTCGGAGCTTCATTGAGGG
Bifidobacterium bifidum PRL2010	CP001840	0		
Bifidobacterium bifidum S17	CP002220	1	45	GTTTCAGATGCCTGTCAGATCAATGACTTTGACCAC
Bifidobacterium breve ACS-071-V-Sch8b	CP002743	1	16+12+37+2	GTCGCTCCCCATCCGGGGAGCGTGGATTGAAAT
Bifidobacterium breve UCC2003	CP000303	1	5+43+2	GTCGCTCCCCATCCGGGGAGCGTGGATTGAAAT
Bifidobacterium dentium Bd1	CP001750	2	18 81	CAAGTTTATCAAGAAGGGTAGAAGCTAATTCCCAGT GTCGCTCTCCTCACGGAGAGCGTGGATTGAAAT
Bifidobacterium longum DJO10A	CP000605	1	43	CAAGCTTATCAAGAAGGGTGAATGCTAATTCCCAGC
Bifidobacterium longum NCC2705	AE014295	0		
Bifidobacterium longum subsp. *infantis* 157F (+ ATCC 15697, JCM 1222)	AP010890 (+ CP001095, AP010889)	0		
Bifidobacterium longum subsp. *longum* BBMN68	CP002286	1	10+13+10	GTTTGCCCCGCATGCGCGGGGATGATCCG
Bifidobacterium longum subsp. *longum* JCM 1217 (+ JDM301)	AP010888 (+ CP002010)	0		
Bifidobacterium longum subsp. *longum* KACC 91563	CP002794	1	33	CAAGCTTATCAAGAAGGGTGAATGCTAATTCCCAGC
Carnobacterium sp. 17-4	CP002563	0		
Lactobacillus acidophilus 30SC	CP002559	0		
Lactobacillus acidophilus NCFM	CP000033	1[a]	34	ATTTTCTCCACGTATGTGGAGGTGATCCT
Lactobacillus amylovorus GRL 1112 (+ GRL 1118)	CP002338 (+ CP002609)	0		
Lactobacillus brevis ATCC 367	CP000416	1[a]	6+5+1	GTATTCCCCACACATGTGGGGGTGATCCT

(*Continued*)

Table 1 (*Continued*)

Genus, species, and strain name	Accession number(s) for chromosomal sequences	CRISPR/Cas systems	Number of repeats	Typical repeat sequence (5′-3′)
Lactobacillus buchneri NRRL B-30929	CP002652	1	26	GTTTTAGAAGGATGTTAAATCAATAAGGTTAAACCC
Lactobacillus casei ATCC 334	CP000423	1	22	GTTTTCCCCGCACATGCGGGGGTGATCCC
Lactobacillus casei BD-II (+ BL23, LC2W)	CP002618 (+ FM177140, CP002616)	1	22	GTCTCAGGTAGATGTCGAATCAATCAGTTCAAGAGC
Lactobacillus casei Zhang	CP001084	1	17	GTCTCAGGTAGATGTCGAATCAATCAGTTCAAGAGC
Lactobacillus crispatus ST1	FN692037	1	17+16+8	GTATTCTCCACGTATGTGGAGGTGATCCT
Lactobacillus delbrueckii subsp. *bulgaricus* 2038	CP000156	1	20	GTTTTAGAAGGTTGTCTATTCAATAAGGTTTAACCC
Lactobacillus delbrueckii subsp. *bulgaricus* ATCC 11842	CR954253	1	41	GTATTCCCCACGCAAGTGGGGGTGATCC
Lactobacillus delbrueckii subsp. *bulgaricus* ATCC BAA-365	CP000412	1	21	GTATTCCCCACGCAAGTGGGGGTGATCC
Lactobacillus delbrueckii subsp. *bulgaricus* ND02	CP002341	1	66	GTATTCCCCACGCAAGTGGGGGTGATCC
Lactobacillus fermentum CECT 5716	CP002033	1	20+24	GTATTCCCCATGTATGTGGGGGTGATCCT
Lactobacillus fermentum IFO3956	AP008937	2	21+24 4	GTATTCCCCATGTATGTGGGGGTGATCCT GTCGCACTCCTTCGCGGGGGTGCGTGGATTGAAAT
Lactobacillus gasseri ATCC 33323	CP000413	0		
Lactobacillus helveticus DPC 4571	CP000517	1 + 1[a]	23 2	GTCGCACTCCTTGTGAGTGCGTGGATTGAAAT GTATTCTCCACGTATGTGGAGGTGATCCT
Lactobacillus helveticus H10	CP002429	1	42	GTCGCACTCCTTGTGAGTGCGTGGATTGAAAT
Lactobacillus johnsonii FI9785 (+ NCC 533)	FN298497 (+ AE017198)	0		
Lactobacillus kefiranofaciens ZW3	CP002764	1	5+4	GTGTTCTCCACGTATGTGGAGGTGATCCT
Lactobacillus plantarum JDM1 (+ ST-III, WCFS1)	CP001617 (+ CP002222, AL935263)	0		
Lactobacillus reuteri DSM 20016 (+ JCM 1112, SD2112)	CP000705 (+ AP007281, CP002844)	0		
Lactobacillus rhamnosus ATCC 53103 (+ GG)	AP011548 (+ FM179322)	1	25	GTCTCAGGTAGATGTCAGATCAATCAGTTCAAGAGC
Lactobacillus rhamnosus Lc 705	FM179323	0		

Lactobacillus sakei subsp. *sakei* 23K	CR936503	0		
Lactobacillus salivarius CECT 5713 (+ UCC118)	CP002034 (+CP000233)	1	30	GTTTCAGAAGTATGTTAAATCAATAAGGTTAAGACC
Lactococcus lactis subsp. *cremoris* MG1363 (+ NZ9000, SK11)	AM406671 (+ CP002094, CP000425)	0		
Lactococcus lactis subsp. *lactis* CV56 (+ IL1403, KF147)	CP002365 (+ AE005176, CP001834)	0		
Leuconostoc citreum KM20	DQ489736	0		
Leuconostoc gasicomitatum LMG 18811	FN822744	0		
Leuconostoc kimchii IMSNU11154	CP001758	0		
Leuconostoc mesenteroides subsp. *mesenteroides* ATCC 8293	CP000414	0		
Oenococcus oeni PSU-1	CP000411	0		
Pediococcus pentosaceus ATCC 25745	CP000422	0		
Propionibacterium freudenreichii subsp. *shermanii* CIRM-BIA1	FN806773	1	34	ATTGCCCCTCCTTCTGGAGGGGCCCTTCATTGAGGC
Staphylococcus carnosus subsp. *carnosus* TM300	AM295250	0		
Streptococcus thermophilus CNRZ1066	CP000024	1 + 1[a]	42	GTTTTTGTACTCTCAAGATTTAAGTAACTGTACAAC
Streptococcus thermophilus JIM 8232	FR875178	2 + 1[a]	43 18 3	GTTTTTGTACTCTCAAGATTTAAGTAACTGTACAAC GATATAAACCTAATTACCTCGAGAGGGGACGGAAAC GTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Streptococcus thermophilus LMD-9	CP000419	3	17 4 8	GTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC GATATAAACCTAATTACCTCGAGAGGGGACGGAAAC GTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Streptococcus thermophilus LMG18311	CP000023	2	34 5	GTTTTTGTACTCTCAAGATTTAAGTAACTGTACAAC GATATAAACCTAATTACCTCGAGAGGGGACGGAAAC
Streptococcus thermophilus ND03	CP002340	3	37 2 21	GTTTTTGTACTCTCAAGATTTAAGTAACTGTACAAC GATATAAACCTAATTACCTCGAGAGGGGACGGAAAC GTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAAC
Weissella koreensis KACC 15510	CP002899	0		

[a]Incomplete CRISPR/Cas system, where either the CRISPR array or the *cas* operon is missing.

Although Type I systems occur in multiple phylogenetic clades of bacteria and archaea, Type II is exclusively present in bacteria and Type III systems appear more commonly in archaea (Makarova et al. 2011). Intriguingly, there are multiple examples of species that contain more than one CRISPR/Cas type, which is consistent with their propensity for horizontal gene transfer via plasmids, transposons, and insertion sequences (Godde & Bickerton 2006, Heidelberg et al. 2009, Horvath et al. 2009, Portillo & Gonzalez 2009, Yang et al. 2011). Starting in the late 1990s and through the 2000s, as bacterial and archaeal genome sequencing/annotation projects greatly expanded, it became obvious that CRISPR sequences were widespread in bacteria and archaea. Within organisms widely used in food operations, a wide array of CRISPR loci have been identified in the genomes of streptococci and many *Lactobacillus* species (**Table 1**). Occurrence and distribution of various CRISPR/Cas types do not correlate with phylogeny, which is consistent with their documented propensity for horizontal gene transfer (Haft et al. 2005, Godde & Bickerton 2006). In addition to *Firmicutes*, CRISPR loci have also been identified in the genomes of *Actinobacteria* used in food manufacturing, notably *Bifidobacterium* (Horvath et al. 2009, Ventura et al. 2009) (**Table 1**). As genome sequencing becomes more readily accessible and complete and draft genomes for food bacteria are made available, the occurrence and distribution of CRISPR/Cas systems will become available. Indeed, although such loci are yet to be identified in the genomes of *Lactococcus*, *Leuconostoc*, *Oenococcus*, and *Pediococcus* (**Table 1**), more sequencing is required to assess whether their absence is inherent to low and/or biased sampling. Further, the current ratio of CRISPR occurrence in bacteria may be underestimated because of the sampling bias in phylogeny and origin of genomes sequenced to date.

CRISPR Mechanism of Action

The CRISPR/Cas immune system encompasses a multistep process by which DNA-encoded spacers lead to sequence-specific interference of invasive nucleic acids, mediated by a ribonucleoproteic complex. There are two main processes characteristic of CRISPR-based immunity, namely spacer acquisition (immunization) and crRNA-mediated interference (immunity). Thus far, most of the molecular basis for the CRISPR/Cas mechanism of action has focused on the interference step, with notable focus on the biochemical details of crRNA biogenesis and target nucleic acid interaction.

In the CRISPR RNA (crRNA) biogenesis phase (**Figure 2**, top), the CRISPR repeat-spacer array is transcribed and processed into small units that include partial repeat and spacer sequences. First, a primary transcript covering the full length repeat-spacer array, called pre-crRNA, is transcribed from the CRISPR locus, as shown in *E. coli* (Brouns et al. 2008, Pougach et al. 2010, Pul et al. 2010), *Xanthomonas oryzae* (Semenova et al. 2009), *Thermus thermophilus* (Agari et al. 2010), *P. furiosus* (Hale et al. 2008), and *Sulfolobus* (Lillestøl et al. 2009, Gudbergsdottir et al. 2011). Reports have established that CRISPR loci are constitutively transcribed at low levels (Pougach et al. 2010, Phok et al. 2011) and can be specifically induced by stress and exposure to viruses using a complex regulatory mechanism (Agari et al. 2010; Lillestøl et al. 2006, 2009; Lintner et al. 2011a; Perez-Rodriguez et al. 2011; Pul et al. 2010; Shinkai et al. 2007). Specific endoribonucleases then cleave the pre-crRNA into small crRNAs that contain a single spacer flanked by partial CRISPR repeats (Hale et al. 2008, Lillestøl et al. 2009). In most cases, cleavage occurs 8 bp from the 5′ end of the spacer, at the base of the hairpin formed by the palindromic CRISPR repeats (Brouns et al. 2008, Carte et al. 2008, Haurwitz et al. 2010, Jore et al. 2011a, Wiedenheft et al. 2011). The partially palindromic nature of CRISPR repeats provides a somewhat conserved secondary structure in the pre-crRNA (Kunin et al. 2007, Makarova et al. 2006). The enzymatic machinery involved in crRNA biogenesis varies across systems and involves a multimeric

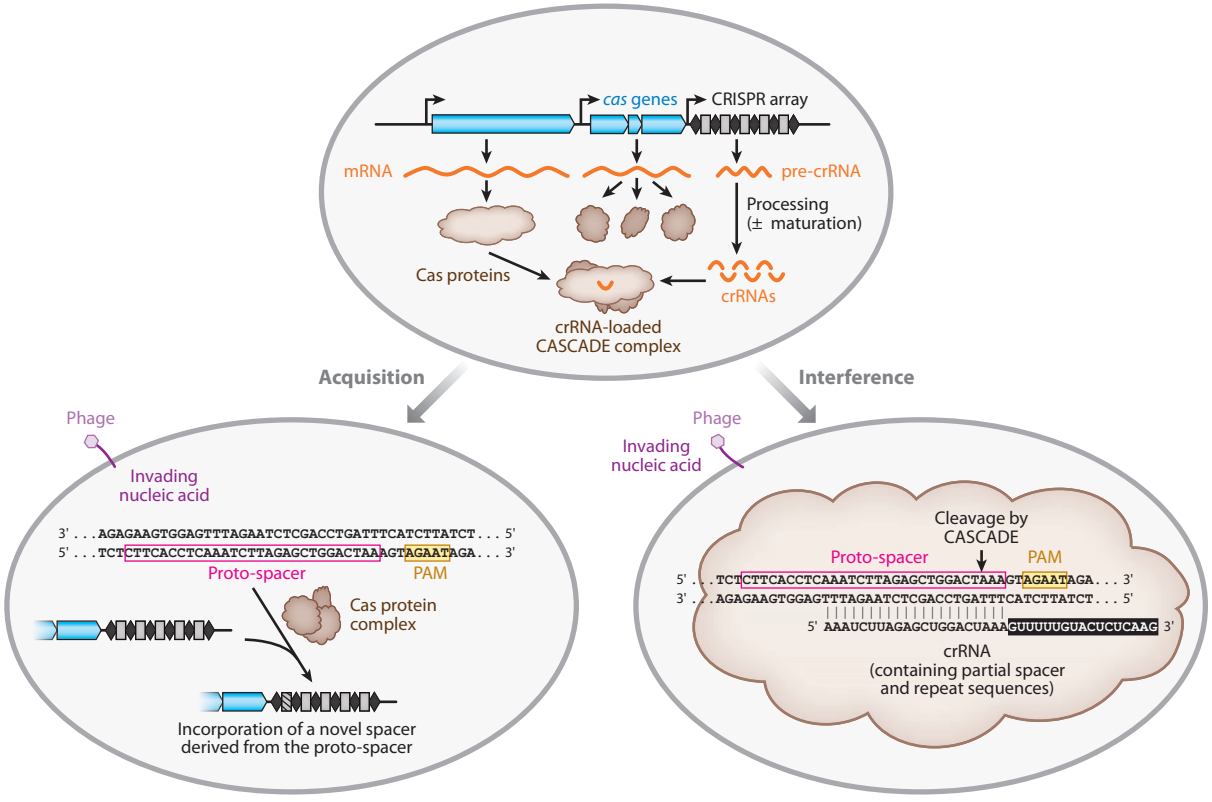


Figure 2

CRISPR mechanism of action. (*Top*) Biogenesis of crRNA and Cas proteins. *Cas* genes are transcribed into mRNA, subsequently translated into Cas proteins that form the CASCADE (CRISPR-associated complex for antiviral defense) complex. The CRISPR repeat-spacer array is transcribed into a full-length pre-crRNA, which is subsequently processed to generate crRNAs that contain partial repeat and spacer sequences. (*Bottom left*) Spacer acquisition. Phage proto-spacer sequences, adjacent to PAMs (proto-spacer associated motifs), are recognized by Cas proteins and subsequently used to generate a novel repeat-spacer unit that is inserted at the leader end of the repeat-spacer array. (*Bottom right*) Interference. The crRNA-loaded CASCADE ribonucleoproteic complex recognizes homologous invading sequences, and the dsDNA is cleaved within the proto-spacer in the vicinity of the PAM.

interference complex in *E. coli* (Brouns et al. 2008), Csy4 in *P. aeruginosa* (Haurwitz et al. 2010, Cady & O'Toole 2011, Wiedenheft et al. 2011), and Cas6 in *P. furiosus* (Carte et al. 2008).

In the spacer acquisition process (**Figure 2**, bottom left), small fragments of invasive nucleic acid are incorporated into the host genome between CRISPR repeats at the leader end of the locus (Barrangou et al. 2007, Deveau et al. 2008). The polarized integration of novel repeat-spacer units at the leader end provides a chronological time line of spacer acquisition events. The sequence on the viral genome that corresponds to a spacer is termed proto-spacer (Deveau et al. 2008). In most cases, a very short stretch of nucleotides is conserved in the immediate vicinity of the proto-spacer, the CRISPR motif, or the proto-spacer-associated motif (PAM) (Deveau et al. 2008, Horvath et al. 2008, Mojica et al. 2009). The ability to acquire novel spacers in vivo has been experimentally documented in *S. thermophilus* (Barrangou et al. 2007, Deveau et al. 2008, Garneau et al. 2010) and *Streptococcus mutans* (van der Ploeg 2009). In both species, one or more novel spacers were incorporated at the leader end of the locus, following exposure to lytic phages or plasmid transformation. The steps involved in spacer acquisition likely include recognition of invasive nucleic

Proto-spacer: sequence in invasive nucleic acid acquired as a spacer into a CRISPR array

Proto-spacer-associated motif (PAM): conserved sequence adjacent to proto-spacers

CRISPR-associated complex for antiviral defense (CASCADE): protein complex guiding crRNA interference

acid, scanning for PAMs that identify potential proto-spacers, and manufacturing and integrating a new repeat-spacer unit. PAMs, which typically are 2–5 nt long and occur within 1–4 bp of proto-spacers, have been identified in multiple CRISPR/Cas systems, notably *S. thermophilus* (AGAAW, GGNG), *S. mutans* (NGG, NAA, TTC) (van der Ploeg 2009), *Xanthomonas* (TTC) (Semenova et al. 2009), and *Sulfolobus* (CC) (Lillestøl et al. 2009). Cas1, a homodimeric, metal-dependent DNase that can process double-stranded DNA (Haurwitz et al. 2010, Wiedenheft et al. 2009) and resolve Holliday junctions (Babu et al. 2011) has been linked to spacer acquisition (Wiedenheft et al. 2009), which is consistent with the coevolutionary pattern and inherent functional linkage observed between Cas1 and specific CRISPR repeat sequences (Horvath et al. 2009). Cas2, which carries an RNA recognition domain and exhibits endoribonucleic activity, has also been implicated in spacer acquisition (Beloglazova et al. 2008, Ebihara et al. 2006). It was also shown that in Type II systems, *csn2* is necessary for spacer acquisition following exposure to phages (Barrangou et al. 2007) and plasmids (Garneau et al. 2010).

In the interference process (**Figure 2**, bottom right), the crRNAs (also called prokaryotic silencing RNAs), together with specific Cas proteins, form a CRISPR ribonucleoprotein complex (crRNP) that guides homologous base-pairing between crRNA and target nucleic acid proto-spacer (Brouns et al. 2008, Perez-Rodriguez et al. 2011) and subsequent cleavage. A multiprotein complex called the CASCADE (CRISPR-associated complex for antiviral defense) complex recognizes and cleaves incoming DNA (or RNA) through complementarity (Brouns et al. 2008, Jore et al. 2011b). Recently, two studies showed that the ribonucleoprotein complex facilitates target recognition by enthalpically driving sequence-specific hybridization between crRNA and the target sequence over a 7–8 bp seed sequence at the 5′ end of the spacer (Semenova et al. 2011, Wiedenheft et al. 2011). The involvement of a CASCADE complex in crRNA-mediated interference has been documented in bacterial Type I and archaeal Type III systems (Brouns et al. 2008, Lintner et al. 2011b). Additionally, several RAMPs with RNA recognition motifs and ferredoxin-fold domains have been implicated in pre-crRNA processing (Brouns et al. 2008, Carte et al. 2008, Haurwitz et al. 2010). A graphical display of the CRISPR/Cas mechanism of action is shown in **Figure 2**. Although initial results suggested that perfect identity was required between spacer and proto-spacer sequences (Barrangou et al. 2007, Deveau et al. 2008), follow-up experiments have shown that perfect matches are most critical in the direct vicinity of the seed sequence and cleavage site (Garneau et al. 2010, Semenova et al. 2009, Wiedenheft et al. 2011).

Although most evidence points to CRISPR/Cas targeting foreign DNA (Marraffini & Sontheimer 2008, Garneau et al. 2010, Manica et al. 2011), it was also shown that a system can target RNA (Hale et al. 2009, Garrett et al. 2011). Interestingly, the CRISPR/Cas mechanism of action shares similarities with the eukaryotic RNAi system, although there are notable genetic and biochemical differences.

APPLICATIONS OF CRISPR/Cas SYSTEMS

Even though the CRISPR field is still in its infancy and most published research has focused primarily on the genetic and biochemical underpinnings of the interference mechanism of action, several studies have provided insights into a series of applications that cover multiple areas of industrial and academic interest, notably enhancing phage resistance in industrial cultures, controlling the dissemination of undesirable genes via plasmid transfer, controlling RNA transcript levels from within, strain genotyping, natural genetic tagging, and investigating microbial population dynamics. The tremendous interest in and focus on eukaryotic RNA interference, together with the ever-expanding repertoire of roles of small noncoding RNAs in eukaryotes, have established that small RNAs can be multifunctional and versatile (Jinek & Doudna 2009). Likewise,

there is growing evidence supporting the diverse regulatory roles that small noncoding RNAs play in bacteria (Gottesman & Storz 2011, Waters & Storz 2009), and recent advances in sequencing technologies have provided new insights into the diversity and high transcript levels of small RNAs, including crRNAs (Deltcheva et al. 2011, Phok et al. 2011).

Phage Resistance

Notwithstanding the historical accumulation of knowledge regarding phage resistance, and the plethora of strategies devised to strategically formulate starter cultures with enhanced phage resistance mechanisms, bacteriophages still remain an issue in industrial dairy fermentation processes. The option to enhance current phage-resistance strategies by harnessing the natural ability of the CRISPR/Cas immune system to generate increased phage resistance barriers provides a novel experimental framework to iteratively build up phage resistance in both resistance levels and spectrum (**Figure 3**). This would theoretically extend the lifespan of valuable industrial strains for perennial use in starter cultures.

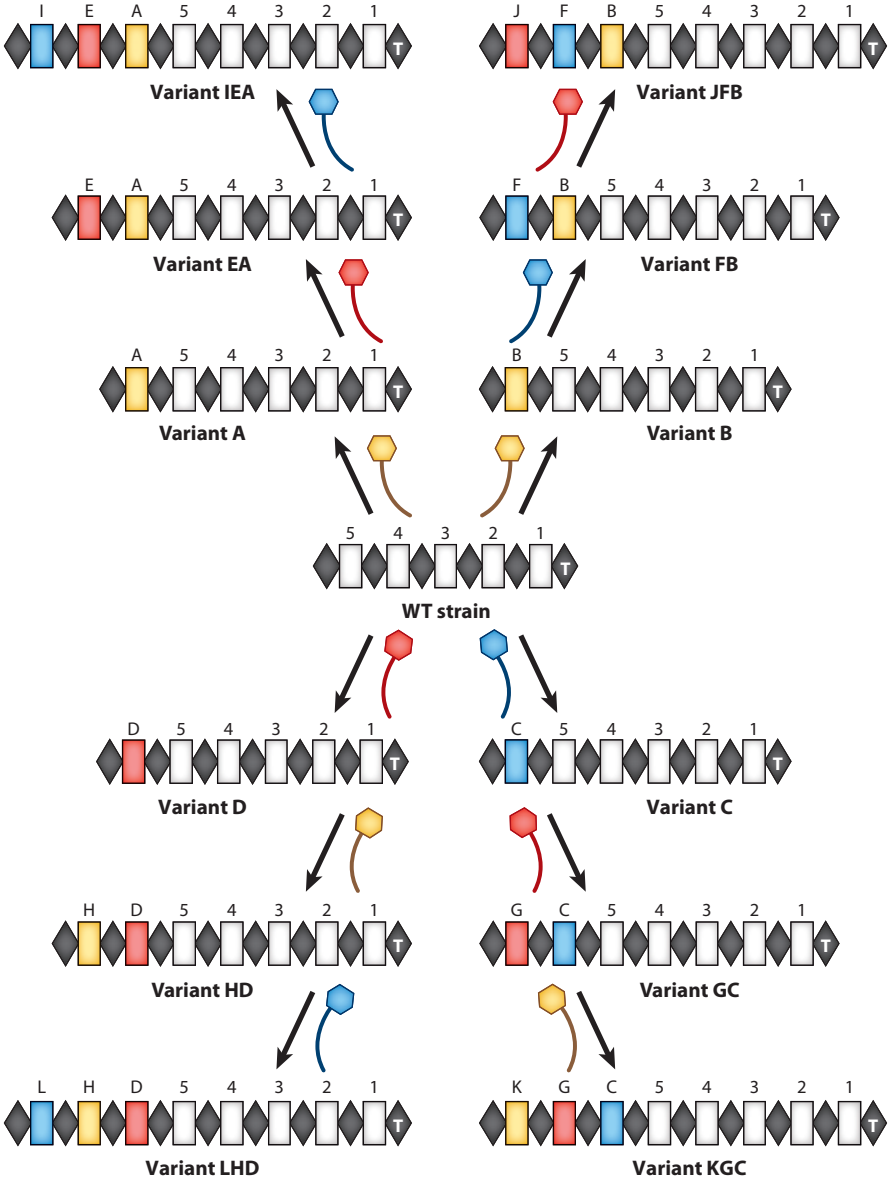
Further, the unique spacer combination resulting from multiple consecutive rounds of CRISPR mutant selection can be used as a natural genetic tagging system for proprietary strains. This natural alternative could also be replicated artificially in vitro through genetic engineering approaches that would generate synthetic spacer sequences.

From a practical standpoint, CRISPR-mediated phage resistance is best implemented iteratively, through multiple rounds of selecting natural CRISPR mutants following exposure to industrially relevant phages. Phages used in different rounds should be selected from a diverse collection based on different host ranges and/or genotypes. CRISPR variants can be readily screened by polymerase chain reaction, selecting mutants that have novel spacers, which preferably have sequence homology to highly conserved phage sequences. Iterative acquisition of CRISPR spacers provides variants with increased phage resistance in terms of reduced phage sensitivity levels (low efficiencies of phage plaquing) and expanded phage resistance spectra (broad spectrum of phage resistance across diverse phage types) (see **Figure 3**). Quantitatively, CRISPR provides high levels of resistance (up to six orders of magnitude of reduction in the efficiency of plaquing for each spacer). Subsequently, multiple CRISPR variants isolated from independent rounds of iterative phage challenges can be selectively combined in culture rotation schemes for increased robustness and optimal industrial lifespan (see **Figure 3**).

Although several examples in the literature document the propensity for mutations, deletions, and loss of *cas* genes in various CRISPR loci (Godde & Bickerton 2006, Horvath et al. 2009), there are multiple lines of evidence indicating that recently acquired CRISPR spacers are stable in long-term experiments. Absent ancillary mutations elsewhere in the genome, an advantage of CRISPR-based phage resistance over other alternatives is the conservation of functional properties inherent to isogenic strains.

Plasmid Interference

A recent study showed that the CRISPR/Cas system can target plasmids that contain antimicrobial resistance markers (Garneau et al. 2010). It was shown that CRISPR spacers can specifically target sequences from antibiotic resistance genes. Thus, CRISPR-encoded immunity provides a natural means to generate strains refractory to the uptake of plasmids that carry undesirable genes. This can be leveraged to generate strains that would reduce uptake and limit the dissemination of undesirable genetic elements, such as prophages, antibiotic resistance markers, and pathogenicity islands (Edgar & Qimron 2010, Nozawa et al. 2011, Palmer & Gilmore 2010, Shimomura et al. 2011).



Strain	Phage		
	Yellow	Blue	Red
WT	S	S	S
A	R	S	S
B	R	S	S
C	S	R	S
D	S	S	R
EA	R	S	R
FB	R	R	S
GC	S	R	R
HD	R	S	R
IEA	R	R	R
JFB	R	R	R
KGC	R	R	R
LHD	R	R	R

Figure 3

Iterative buildup of phage resistance. (*Center*) Wild-type (WT) CRISPR repeat-spacer array (parental strain), with repeats represented as black diamonds and spacers represented as gray rectangles. Four distinct paths represent independent phage challenges, whereby each phage is represented by a unique color: yellow, red, or blue. Each arrow represents a unique phage challenge cycle, generating four isogenic variants, each with a unique set of three novel spacers acquired iteratively that ultimately provide resistance to all three phages used. (*Right inset*) Phage resistance (R, *red*) and sensitivity (S, *blue*) profiles of each intermediate and final isogenic strain. Combinations of different isogenic variants in culture systems provide increased robustness against phage predation. Abbreviation: T, terminal repeat.

Today, capabilities in genetic engineering provide the technology to clone and transfer functional CRISPR/Cas systems into a wide array of hosts (Sapranauskas et al. 2011). Further, the spectrum of interference targets can include mobile genetic elements. Given the negative correlation between CRISPRs and antibiotic resistance markers in multidrug-resistant enterococci, CRISPRs likely mitigate the dissemination of mobile genetic elements, such as plasmids and prophages, which account for up to 25% of pathogenic *Enterococcus faecalis* (Palmer & Gilmore 2010). Likewise, a negative correlation has been reported between the occurrence of prophages and the number of CRISPR spacers in *Streptococcus pyogenes* (Nozawa et al. 2011). Thus, there is potential to leverage CRISPRs to reduce the spread of antimicrobial resistance genes and virulence factors in pathogens such as methicillin-resistant *Staphylococcus aureus* (Marraffini & Sontheimer 2010).

Intriguingly, within industrial cultures, CRISPR-dependent plasmid interference might be responsible for the relative scarcity of plasmids in dairy *S. thermophilus*, as compared with *Lactococcus lactis.* Although both species have similar genome sizes and environments, the former relies on chromosomally encoded CRISPR resistance, whereas the latter depends on plasmid-encoded restriction-modification and abortive-infection systems (Deveau et al. 2010, Labrie et al. 2010). Further studies should investigate the orchestrated deployment of different defense systems, including CRISPRs, abortive-infection, toxin-antitoxin, and restriction-modification and regulatory systems (Labrie et al. 2010). Likewise, future studies should assess the impact of CRISPRs on phage ecology (Angly et al. 2006, Edwards & Rohwer 2005).

RNA Regulation

Although the primary target of the CRISPR/Cas system is DNA in most studied systems, it was shown that some systems can target RNA (Hale et al. 2009, Garrett et al. 2011). Accordingly, there is potential to leverage CRISPR loci targeting RNA to regulate or silence transcript levels within the cell. A recent report showed that a CRISPR spacer that matches the histidyl-tRNA synthetase sequence attenuated intracellular histidyl-tRNA pools (Aklujkar & Lovley 2010). Likewise, in silico analyses have revealed several examples of self-targeting CRISPR spacers (Stern et al. 2010).

Recent studies investigating transcriptional control of CRISPR/Cas systems in *E. coli* and *Salmonella* have revealed a complex repression scheme involving histone-like nucleoid-structuring protein (H-NS), LysR-type transcription factor LeuO, and LRP (Diez-Villasenor et al. 2010, Medina-Aparicio et al. 2011, Pul et al. 2010, Westra et al. 2010), implicating a global regulatory network consistent with roles beyond immunity.

Self-targeting spacers appear to be involved in lysogeny-dependent inhibition of biofilm formation in *P. aeruginosa* (Zegans et al. 2009, Cady & O'Toole 2011) and swarming of myxobacteria (Viswanathan et al. 2007). As our knowledge of CRISPRs expands, we anticipate that new functions of this versatile system will be uncovered.

Typing

Given the dynamic nature of spacer acquisition, CRISPRs provide a genetic basis for strain typing and epidemiological studies. Indeed, spacers provide a hypervariable time-resolved recording of past acquisition events that has been used to investigate *Mycobacterium tuberculosis* (Abadia et al. 2010, 2011; Borile et al. 2011; Brudey et al. 2006; Groenen et al. 1993; Zhang et al. 2010), *Yersinia pestis* (Cui et al. 2008, Pourcel et al. 2005), *Corynebacterium diphtheriae* (Mokrousov et al. 2007, 2009), *P. aeruginosa* (Cady et al. 2011), *E. coli* (Diez-Villasenor et al. 2010), *Legionella* (D'auria et al. 2010), *S. pyogenes* (Hoe et al. 1999, McShan et al. 2008), *Salmonella* (Fricke et al. 2011; Liu et al. 2011a,b), and industrially relevant organisms, such as lactobacilli and streptococci

(Horvath et al. 2008, 2009; Guinane et al. 2011). Surprisingly, although the first documented use of CRISPR sequences for spoligotyping of *M. tuberculosis* has been ongoing for nearly 20 years (Groenen et al. 1993), the origin of spacer polymorphism remained obscure until recently. Given the hypervariable nature of CRISPR loci, when active they provide the ability to segregate nearly identical strains over time within clonal populations (Andersson & Banfield 2008) and to track sublineages within monomorphic populations, such as *S. thermophilus* (Horvath et al. 2008) (see **Figure 3**), *Erwinia amylovora* (Rezzonico et al. 2011), and *S. enterica* (Fricke et al. 2011). The degree of spacer polymorphism in terms of both number of unique spacers and spacer arrangements usually correlates with the activity level of a given locus; thus, in cases where there are multiple CRISPR loci in the genome, it is important to choose a genetically polymorphic locus (Horvath et al. 2008). In contrast, in cases where CRISPR loci appear inactive and/or absent selective pressure by invasive genetic elements, such as phages, CRISPR content provides limited insights into phylogeny, as shown in *E. coli* (Touchon & Rocha 2010, Touchon et al. 2011).

CRISPR spacer hypervariability in space and time can be exploited to resolve population-level genotypes in complex environmental samples (Andersson & Banfield 2008, Heidelberg et al. 2009, Held & Whitaker 2009, Held et al. 2010, Pride et al. 2011, Sorokin et al. 2010, Tyson & Banfield 2008). These studies have shown that CRISPR spacers can be used to monitor complex dynamic systems over time and provide insights into ancestral relationships as well as allow the detection of events, such as blooms, selective sweeps, and bottlenecks. Over time, it is likely that active and hypervariable CRISPR loci will be increasingly leveraged in metagenomic studies (Anderson et al. 2011, Andersson & Banfield 2008, Tyson & Banfield 2008). In addition to resolving host population dynamics, spacers provide insights into the coevolutionary dynamics between host and viruses (Andersson & Banfield 2008, Garrett et al. 2010, Heidelberg et al. 2009). Indeed, studies have shown that viruses specifically mutate their genomes in proto-spacer and/or PAM regions in direct response to CRISPR spacer acquisition (Andersson & Banfield 2008, Deveau et al. 2008, Garneau et al. 2010). Recently, CRISPR-escape strategies were expanded with the observation that phages can encode H-NS proteins that repress *cas* genes in bacteria (Skennerton et al. 2011). These genetic patterns provide important insights into genome evolution of both the host and phage populations (Anderson et al. 2011, Touchon & Rocha 2010, Touchon et al. 2011, Tyson & Banfield 2008) and set the stage for mathematical modeling of their evolutionary interplay (He & Deem 2010, Levin 2010, Vale & Little 2010).

SUMMARY POINTS

1. The CRISPR/Cas system provides DNA-encoded and RNA-mediated immunity against phages and plasmids by sequence-specific cleavage of invasive nucleic acids.
2. This immune system has been identified in most archaea and many bacteria, and three different types are distributed across distant phylogenetic phyla.
3. CRISPR/Cas systems providing resistance against phages can be used to iteratively build up resistance levels to generate robust industrial starter cultures.
4. CRISPR/Cas systems providing interference against plasmids can be used to limit the uptake and dissemination of undesirable genetic elements, such as antibiotic resistance genes.
5. The hypervariable nature of CRISPR spacer sequences can be used for high-resolution genotyping of bacteria and archaea for epidemiological studies and metagenomic surveys.

FUTURE ISSUES

1. Notwithstanding current knowledge in *E. coli*, *P. aeruginosa*, *P. furiosus*, *S. epidermidis*, and *S. thermophilus*, there remains a need for new genetically tractable systems for analyses of the mechanism of action across CRISPR/Cas systems.
2. Mechanistic aspects remain unclear, and very little is known about the elements involved in CRISPR spacer selection and acquisition.
3. Given the potential practical applications for phage resistance, plasmid interference, typing, and RNA level control, CRISPRs should be leveraged industrially for the development of safer and more robust strains. This will require an understanding of elements that drive CRISPR activity.
4. Genetically, the speed and rate at which CRISPR loci evolve should be investigated, notably regarding CRISPR locus size control and spacer turnover.
5. Genomically, as the amount and diversity of available bacterial and archaeal sequences expand, global studies need to assess the occurrence and origin of the CRISPR/Cas system across phylogenetic phyla.
6. The availability of genetic engineering tools will allow the transfer of active CRISPR/Cas systems into a wide array of hosts for heterologous exploitation, notably for the enhancement of phage resistance and control of the dissemination and uptake of undesirable genetic material.

DISCLOSURE STATEMENT

Both authors are currently employed by Danisco and are co-inventors on several patent applications related to CRISPR.

ACKNOWLEDGMENTS

We thank our Danisco colleagues and academic collaborators for their support of our work and many contributions to the ever-expanding CRISPR field.

LITERATURE CITED

Abadia E, Zhang J, dos Vultos T, Ritacco V, Kremer K, et al. 2010. Resolving lineage assignation on *Mycobacterium tuberculosis* clinical isolates classified by spoligotyping with a new high-throughput 3R SNPs based method. *Infect. Genet. Evol.* 10:1066–74

Abadia E, Zhang J, Ritacco V, Kremer K, Tuimy R, et al. 2011. The use of microbead-based spoligotyping for *Mycobacterium tuberculosis* complex to evaluate the quality of the conventional method: providing guidelines for quality assurance when working on membranes. *BMC Infect. Dis.* 11:110

Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. 2010. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J. Mol. Biol.* 395:270–81

Aklujkar M, Lovley DR. 2010. Interference with histidyl-tRNA synthetase by a CRISPR spacer sequence as a factor in the evolution of *Pelobacter carbinolicus*. *BMC Evol. Biol.* 10:230

Al-Attar S, Westra ER, van der Oost J, Brouns SJ. 2011. Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol. Chem.* 392:277–89

Anantharaman V, Iyer LM, Aravind L. 2010. Presence of a classical RRM-fold palm domain in Thg1-type 3′-5′ nucleic acid polymerases and the origin of the GGDEF and CRISPR polymerase domains. *Biol. Direct.* 5:43

Anderson RE, Brazelton WJ, Baross JA. 2011. Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol. Ecol.* 77:120–33

Andersson AF, Banfield JF. 2008. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* 320:1047–50

Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, et al. 2006. The marine viromes of four oceanic regions. *PLoS Biol.* 4:e368

Babu M, Beloglazova N, Flick R, Graham C, Skarina T, et al. 2011. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.* 79:484–502

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–12

Beloglazova N, Brown G, Zimmerman MD, Proudfoot M, Makarova KS, et al. 2008. A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *J. Biol. Chem.* 283:20361–71

Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151:2551–61

Borile C, Labarre M, Franz S, Sola C, Refregier G. 2011. Using affinity propagation for identifying subspecies among clonal organisms: lessons from *M. tuberculosis*. *BMC Bioinforma.* 12:224

Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, et al. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321:960–64

Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, et al. 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol.* 6:23

Brüssow H. 2001. Phages of dairy bacteria. *Annu. Rev. Microbiol.* 55:283–303

Cady KC, O'Toole GA. 2011. Non-identity targeting of *Yersinia*-subtype CRISPR-prophage interaction requires the Csy and Cas3 proteins. *J. Bacteriol.* 193:3433–45

Cady KC, White AS, Hammond JH, Abendroth MD, Karthikeyan RS, et al. 2011. Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* 157:430–37

Carte J, Wang R, Li H, Terns RM, Terns MP. 2008. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* 22:3489–96

Cui Y, Li Y, Gorge O, Platonov ME, Yan Y, et al. 2008. Insight into microevolution of *Yersinia pestis* by clustered regularly interspaced short palindromic repeats. *PLoS ONE* 3:e2652

D'auria G, Jimenez-Hernandez N, Peris-Bondia F, Moya A, Latorre A. 2010. *Legionella pneumophila* pangenome reveals strain-specific virulence factors. *BMC Genomics* 11:181

Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, et al. 2011. CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature* 471:602–7

Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, et al. 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1390–400

Deveau H, Garneau JE, Moineau S. 2010. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.* 64:475–93

Diez-Villasenor C, Almendros C, Garcia-Martinez J, Mojica FJ. 2010. Diversity of CRISPR loci in *Escherichia coli*. *Microbiology* 156:1351–61

Ebihara A, Yao M, Masui R, Tanaka I, Yokoyama S, Kuramitsu S. 2006. Crystal structure of hypothetical protein TTHB192 from *Thermus thermophilus* HB8 reveals a new protein family with an RNA recognition motif-like domain. *Protein Sci.* 15:1494–99

Edgar R, Qimron U. 2010. The *Escherichia coli* CRISPR system protects from lambda lysogenization, lysogens, and prophage induction. *J. Bacteriol.* 192:6291–94

Edwards RA, Rohwer F. 2005. Viral metagenomics. *Nat. Rev. Microbiol.* 3:504–10

Fricke WF, Mammel MK, McDermott PF, Tartera C, White DG, et al. 2011. Comparative genomics of 28 *Salmonella enterica* isolates: evidence for CRISPR-mediated adaptive sublineage evolution. *J. Bacteriol.* 193:3556–68

Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, et al. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468:67–71

Garrett RA, Prangishvili D, Shah SA, Reuter M, Stetter KO, Peng X. 2010. Metagenomic analyses of novel viruses and plasmids from a cultured environmental sample of hyperthermophilic neutrophiles. *Environ. Microbiol.* 12:2918–30

Garrett RA, Shah SA, Vestergaard G, Deng L, Gudbergsdottir S, et al. 2011. CRISPR-based immune systems of the *Sulfolobales*: complexity and diversity. *Biochem. Soc. Trans.* 39:51–57

Godde JS, Bickerton A. 2006. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* 62:718–29

Gottesman S, Storz G. 2011. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb. Perspect. Biol.* 3:a003798

Grissa I, Vergnaud G, Pourcel C. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinforma.* 23:172

Groenen PM, Bunschoten AE, van Soolingen D, van Embden JD. 1993. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol. Microbiol.* 10:1057–65

Gudbergsdottir S, Deng L, Chen Z, Jensen JV, Jensen LR, et al. 2011. Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Mol. Microbiol.* 79:35–49

Guinane C, Kent RM, Norberg S, Hill C, Fitzgerald GF, et al. 2011. Host specific diversity in *Lactobacillus johnsonii* as evidenced by a major chromosomal inversion and phage resistance mechanisms. *PLoS ONE* 6:e18740

Haft DH, Selengut J, Mongodin EF, Nelson KE. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* 1:e60

Hale C, Kleppe K, Terns RM, Terns MP. 2008. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* 14:2572–79

Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, et al. 2009. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139:945–56

Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. 2010. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* 329:1355–58

He J, Deem MW. 2010. Heterogeneous diversity of spacers within CRISPR (clustered regularly interspaced short palindromic repeats). *Phys. Rev. Lett.* 105:128102

Heidelberg JF, Nelson WC, Schoenfeld T, Bhaya D. 2009. Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS ONE* 4:e4169

Held NL, Herrera A, Cadillo-Quiroz H, Whitaker RJ. 2010. CRISPR associated diversity within a population of *Sulfolobus islandicus*. *PLoS ONE* 5:e12988

Held NL, Whitaker RJ. 2009. Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. *Environ. Microbiol.* 11:457–66

Hoe N, Nakashima K, Grigsby D, Pan X, Dou SJ, et al. 1999. Rapid molecular genetic subtyping of serotype M1 group A *Streptococcus* strains. *Emerg. Infect. Dis.* 5:254–63

Horvath P, Barrangou R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327:167–70

Horvath P, Coûté-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, Barrangou R. 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int. J. Food Microbiol.* 131:62–70

Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, et al. 2008. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1401–12

Ishino Y, Shinagawa H, Makino K, Amemura M, 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.* 169:5429–33

Jansen R, van Embden JD, Gaastra W, Schouls LM. 2002a. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43:1565–75

Jansen R, van Embden JD, Gaastra W, Schouls LM. 2002b. Identification of a novel family of sequence repeats among prokaryotes. *OMICS* 6:23–33

Jinek M, Doudna JA. 2009. A three-dimensional view of the molecular machinery of RNA interference. *Nature* 457:405–12

Jore MM, Brouns SJ, van der Oost J. 2011a. RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. *Cold Spring Harb. Perspect. Biol.* doi: 10.1101/cshperspect.a003657

Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, et al. 2011b. Structural basis for CRISPR RNA-guided DNA recognition by CASCADE. *Nat. Struct. Mol. Biol.* 18:529–36

Karginov FV, Hannon GJ. 2010. The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol. Cell* 37:7–19

Kunin V, Sorek R, Hugenholtz P. 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* 8:R61

Labrie SJ, Samson JE, Moineau S. 2010. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* 8:317–27

Levin BR. 2010. Nasty viruses, costly plasmids, population dynamics, and the conditions for establishing and maintaining CRISPR-mediated adaptive immunity in bacteria. *PLoS Genet.* 6:e1001171

Lillestøl RK, Redder P, Garrett RA, Brugger K. 2006. A putative viral defence mechanism in archaeal cells. *Archaea* 2:59–72

Lillestøl RK, Shah SA, Brugger K, Redder P, Phan H, et al. 2009. CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. *Mol. Microbiol.* 72:259–72

Lintner NG, Frankel KA, Tsutakawa SE, Alsbury DL, Copie V, et al. 2011a. The structure of the CRISPR-associated protein Csa3 provides insight into the regulation of the CRISPR/Cas system. *J. Mol. Biol.* 405:939–55

Lintner NG, Kerou M, Brumfield SK, Graham S, Liu H, et al. 2011b. Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). *J. Biol. Chem.* 286:12643–56

Liu F, Barrangou R, Gerner-Smidt P, Ribot EM, Knabel SJ, et al. 2011a. Novel virulence gene and clustered regularly interspaced short palindromic repeat (CRISPR) multilocus sequence typing scheme for subtyping of the major serovars of *Salmonella enterica* subsp. *enterica*. *Appl. Environ. Microbiol.* 77:1946–56

Liu F, Kariyawasam S, Jayarao B, Barrangou R, Gerner-Smidt P, et al. 2011b. Subtyping *Salmonella* serovar Enteritidis isolates from different sources using sequence typing based on virulence genes and CRISPRs. *App. Environ. Microbiol.* 77:4520–26

Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV. 2002. A DNA repair system specific for thermophilic archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* 30:482–96

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, et al. 2011. Evolution and classification of the CRISPR/Cas systems. *Nat. Rev. Microbiol.* 9:467–77

Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 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. Dir.* 1:7

Manica A, Zebec Z, Teichmann D, Schleper C. 2011. In vivo activity of CRISPR-mediated virus defence in a hyperthermophilic archaeon. *Mol. Microbiol.* 80:481–91

Marraffini LA, Sontheimer EJ. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322:1843–45

Marraffini LA, Sontheimer EJ. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11:181–90

McShan WM, Ferretti JJ, Karasawa T, Suvorov AN, Lin S, et al. 2008. Genome sequence of a nephritogenic and highly transformable M49 strain of *Streptococcus pyogenes*. *J. Bacteriol.* 190:7773–85

Medina-Aparicio L, Rebolar-Flores JE, Gallego-Hernandez AL, Vasquez A, Olvera VL, et al. 2011. The CRISPR/Cas immune system is an operon regulated by LeuO, H-NA and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* 193:2396–407

Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155:733–40

Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60:174–82

Mojica FJ, Ferrer C, Juez G, Rodriguez-Valera F. 1995. Long stretches of short tandem repeats are present in the largest replicons of the archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. *Mol. Microbiol.* 17:85–93

Mokrousov I, Limeschenko E, Vyazovaya A, Narvskaya O. 2007. *Corynebacterium diphtheriae* spoligotyping based on combined use of two CRISPR loci. *Biotechnol. J.* 2:901–6

Mokrousov I, Vyazovaya A, Kolodkina V, Limeschenko E, Titov L, Narvskaya O. 2009. Novel macroarray-based method of *Corynebacterium diphtheriae* genotyping: evaluation in a field study in Belarus. *Eur. J. Clin. Microbiol. Infect. Dis.* 28:701–3

Nozawa T, Furukawa N, Aikawa C, Watanabe T, Haobam B, et al. 2011. CRISPR inhibition of prophage acquisition in *Streptococcus pyogenes*. *PLoS ONE* 6:e19543

Palmer KL, Gilmore MS. 2010. Multidrug-resistant enterococci lack CRISPR-*cas*. *MBio* 1:e00227-10

Perez-Rodriguez R, Haitjema C, Huang Q, Nam KH, Bernardis S, et al. 2011. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol. Microbiol.* 79:584–99

Phok K, Moisan A, Rinaldi D, Brucato N, Carpousis AJ, et al. 2011. Identification of CRISPR and riboswitch related RNAs among novel non-coding RNAs of the euryarchaeon *Pyrococcus abyssi*. *BMC Genomics* 12:312

Portillo MC, Gonzalez JM. 2009. CRISPR elements in the thermococcales: evidence for associated horizontal gene transfer in *Pyrococcus furiosus*. *J. Appl. Genet.* 50:421–30

Pougach K, Semenova E, Bogdanova E, Datsenko KA, Djordjevic M, et al. 2010. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol. Microbiol.* 77:1367–79

Pourcel C, Salvignol G, 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* 151:653–63

Pride DT, Sun CL, Salzman J, Rao N, Loomer P, et al. 2011. Analysis of streptococcal CRISPRs from human saliva reveals substantial sequence diversity within and between subjects over time. *Genome Res.* 21:126–36

Pul U, Wurm R, Arslan Z, Geissen R, Hofmann N, Wagner R. 2010. Identification and characterization of *E. coli* CRISPR-*cas* promoters and their silencing by H-NS. *Mol. Microbiol.* 75:1495–512

Rezzonico F, Smits THM, Duffy B. 2011. Diversity, evolution and functionality of clustered regularly interspaced short palindromic repeat (CRISPR) regions in the fire blight pathogen *Erwinia amylovora*. *Appl. Environ. Microbiol.* 77:3819–29

Rousseau C, Gonnet M, Le Romancer M, Nicolas J. 2009. CRISPI: a CRISPR interactive database. *Bioinformatics* 25:3317–18

Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. 2011. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* 39:9275–82

Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, et al. 2011. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* 108:10098–103

Semenova E, Nagornykh M, Pyatnitskiy M, Artamonova II, Severinov K. 2009. Analysis of CRISPR system function in plant pathogen *Xanthomonas oryzae*. *FEMS Microbiol. Lett.* 296:110–16

Shah SA, Garrett RA. 2011. CRISPR/Cas and Cmr modules, mobility and evolution of adaptive immune systems. *Res. Microbiol.* 162:27–38

Shimomura Y, Okumura K, Yamagata Murayama S, Yagi J, Ubukata K, et al. 2011. Complete genome sequencing and analysis of a Lancefield group G *Streptococcus dysgalactiae* subsp. *equisimilis* strain causing streptococcal toxic shock syndrome (STSS). *BMC Genomics* 12:17

Shinkai A, Kira S, Nakagawa N, Kashihara A, Kuramisty S, Yohohama S. 2007. Transcription activation mediated by a cyclic AMP receptor protein from *Thermus thermophilus* HB8. *J. Bacteriol.* 189:3891–901

Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. 2011. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* 30:1335–42

Skennerton CT, Angly FE, Breitbart M, Bragg L, He S, et al. 2011. Phage encoded H-NS: a potential Achilles heel in the bacterial defense system. *PLoS ONE* 6:e20095

Sorokin VA, Gelfand MS, Artamonova II. 2010. Evolutionary dynamics of clustered irregularly interspaced short palindromic repeat systems in the ocean metagenome. *Appl. Environ. Microbiol.* 76:2136–44

Stern A, Keren L, Wurtzel O, Amitai G, Sorek R. 2010. Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* 26:335–40

Sturino JM, Klaenhammer TR. 2006. Engineered bacteriophage-defence systems in bioprocessing. *Nat. Rev. Microbiol.* 4:395–404

Terns MP, Terns RM. 2011. CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* 14:321–27

Touchon M, Charpentier S, Clermont O, Rocha EPC, Denamur E, Branger C. 2011. CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *J. Bacteriol.* 193:2460–67

Touchon M, Rocha EP. 2010. The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS ONE* 5:e11126

Tyson GW, Banfield JF. 2008. Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environ. Microbiol.* 10:200–7

Vale PF, Little TJ. 2010. CRISPR-mediated phage resistance and the ghost of coevolution past. *Proc. Biol. Sci.* 277:2097–103

van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJ. 2009. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* 34:401–7

van der Ploeg JR. 2009. Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. *Microbiology* 155:1966–76

Ventura M, Turroni F, Lima-Mendez G, Foroni E, Zomer A, et al. 2009. Comparative analyses of prophage-like elements present in bifidobacterial genomes. *Appl. Environ. Microbiol.* 75:6929–36

Viswanathan P, Murphy K, Julien B, Garza AG, Kroos L. 2007. Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J. Bacteriol.* 189:3738–50

Wang R, Preamplume G, Terns MP, Terns RM, Li H. 2011. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage. *Structure* 19:257–64

Waters LS, Storz G. 2009. Regulatory RNAs in bacteria. *Cell* 136:615–28

Westra ER, Pul U, Heidrich N, Jore MM, Lundgren M, et al. 2010. H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* 77:1380–93

Wiedenheft B, van Duijn E, Bultema J, Waghmare S, Zhou K, et al. 2011. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* 108:10092–97

Wiedenheft B, Zhou K, Jinek M, Coyle SM, Ma W, Doudna JA. 2009. Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* 17:904–12

Yang Y, Kurokawa T, Takahama Y, Nindita Y, Mochizuki S, et al. 2011. pSLA2-M of *Streptomyces rochei* is a composite linear plasmid characterized by self-defense genes and homology with pSLA2-L. *Biosci. Biotechnol. Biochem.* 75:110054-1–7

Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH, O'Toole GA. 2009. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* 191:210–19

Zhang J, Abadia E, Refregier G, Tafaj S, Boschiroli ML, et al. 2010. *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of "spoligotyping" with new spacers and a microbead-based hybridization assay. *J. Med. Microbiol.* 59:285–94