CRISPR-Cas system

Hakim, Joseph A.; Koo, Hyunmin; van Elsas, Jan D.; Trevors, Jack T.; Bej, Asim K.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

CRISPR–Cas system: a new paradigm for bacterial stress response through genome rearrangement

Joseph A. Hakim[1], Hyunmin Koo[1], Jan D. van Elsas[2], Jack T. Trevors[3], and Asim K. Bej[1]

[1]Department of Biology, University of Alabama at Birmingham, Birmingham, Alabama, USA
[2]Biological Center, University of Groningen, Groningen, The Netherlands
[3]Laboratory of Microbiology, School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada

2.11.1 Introduction

Bacteria are the most abundant and diverse group of organisms inhabiting the Earth. As bacterial communities are integrated and sometimes interdependent within their environments, the privations or impositions accompanying both existing and new environments pose challenges to their survival (for reviews, see Cary *et al.*, 2010; Fierer and Jackson, 2006; Kostic *et al.*, 2013; Kristjánsson and Hreggvidsson, 1995; Schrenk *et al.*, 2010; Torsvik and Øvreås, 2002). Additionally, dynamic ecosystems with fluctuating ecophysico-chemical parameters cause stresses that can exceed the ranges of bacterial functions. Some of these parameters include pH, temperature, water, gases, depleted and/or limited nutrients leading to nutrient starvation, light–dark cycles, irradiation, cell envelope perturbation or disruption, and cohabiting bacteriophage populations (Buckling and Rainey, 2002; Burgess *et al.*, 1999; Gómez and Buckling, 2011; Middelboe *et al.*, 2001). To counter bacteriophage attack, some bacteria have restriction modification systems (Tock and Dryden, 2005; see Chapter 2.10). In addition, genome rearrangements or reprogramming cellular functions (i.e., evolution) conducive to their fitness in the ecosystem has been proposed (Gurung *et al.*, 2001; Rubin and Leff, 2007; Spano and Massa, 2006).

Some of the known genetic-based mechanisms (e.g., transformation, conjugation, transduction, and recombination) contributing to genome rearrangements are considered to be the means for bacterial cells evolving with the constraints of environmental stresses. However, recently discovered clustered regularly-interspaced short palindromic repeats (CRISPR) and the CRISPR-associated proteins (Cas) (Jansen *et al.*, 2002) are recognized as a unique strategy utilized by certain bacterial species.

Initial recognition of the CRISPR–Cas array occurred in 1987 through the work of Ishino *et al.* (1987) while characterizing the alkaline phosphatase isozyme gene (*iap*) in *Escherichia coli*. It was discovered that the 3′ end of the amplified segment of genome containing *iap* consisted of a series of homologous repeats of nucleotides, arranged in dyadic symmetry. Additionally, it was reported that these repeats were interspaced among sequences of 32 nucleotides that contained no homology within the genome (Ishino *et al.*, 1987). Although the functional significance of the repeat sequences was not apparent in their discovery, a predictive role was suggested for messenger RNA (mRNA) stability, due to an appearance of the internal post-transcriptional folding of the RNA. Jansen *et al.* (2002) offered an insight into the structure and function of the loci, and coined the terms "CRISPR" and "Cas", collectively referred to as the CRISPR–Cas system (Figure 2.11.1).

In subsequent research, the presence of "structural artifacts" was acknowledged, characterizing a species-conserved leader sequence upstream of the repeat region, and four closely located *cas* genes synchronous with the CRISPR (Jansen *et al.*, 2002). The amino acid sequence alignment of Cas1 proteins from various strains of *E. coli* showed significant homology, whereas Cas1 from various species of *Streptococcus* exhibited poor homology, depicting the species-specific conservativeness of this protein toward its target CRISPR sequences (this study) (Figure 2.11.2a,b).

Similar results were observed when the amino acid residues of Cas1 proteins from *Myxococcus xanthus* and *Pectobacterium atrosepticum* were aligned (data not shown). The comparison of the amino acid residues of the Cas2 protein of *E. coli* displayed a better homology as compared to various species of *Streptococcus* (this study) (Figure 2.11.3a,b). Evolution of interspecies diversity of the Cas1 and Cas2 proteins was further supported by
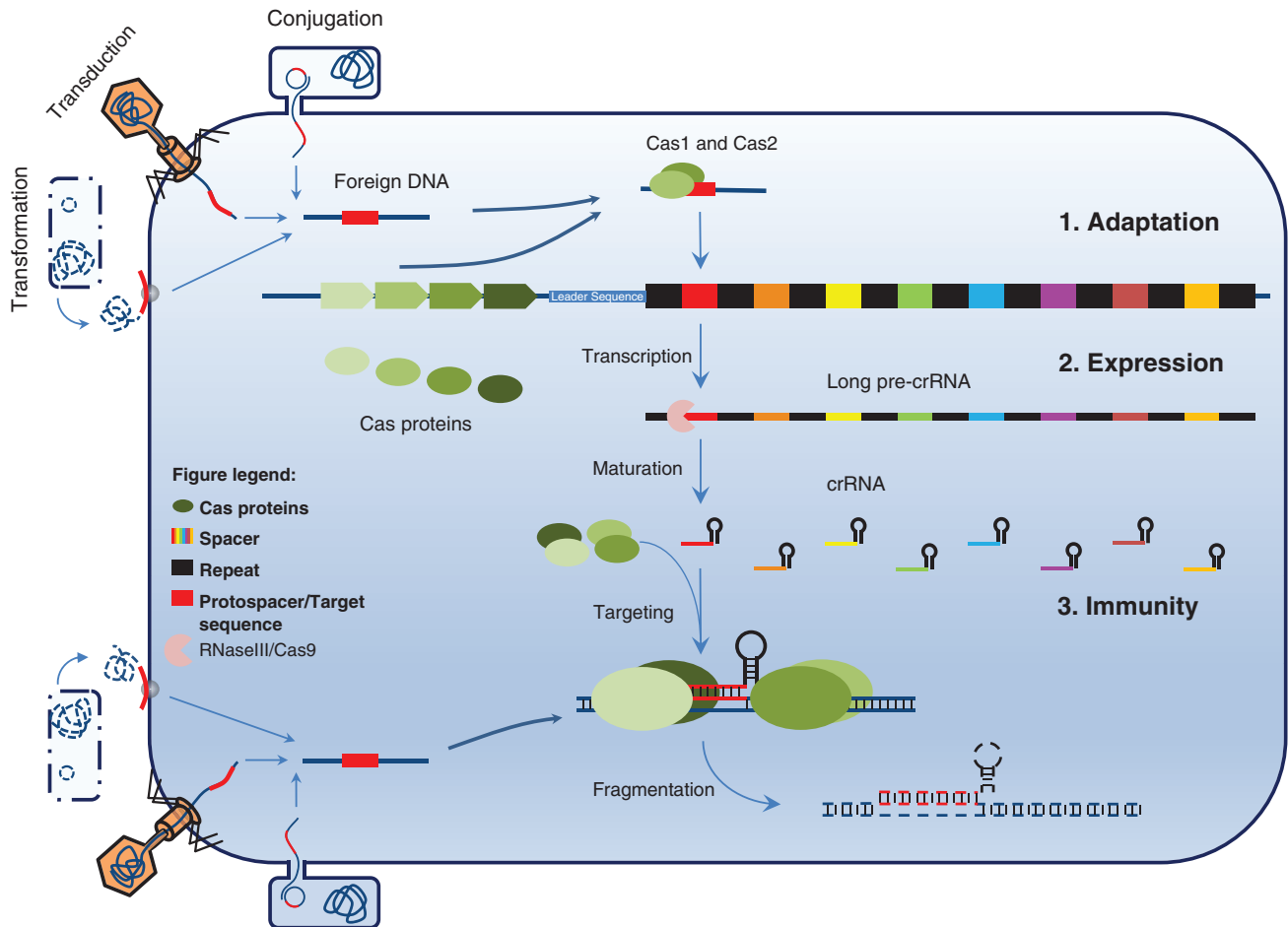
Figure 2.11.1 Schematic diagram for the CRISPR–Cas system in bacteria. Extraneous mobile genetic elements enter the recipient bacterial cells through various mechanisms (transformation, transduction, and/or conjugation). The process follows three steps: (i) Once inside the cell, Cas proteins target foreign stretch of nucleotides, excising a portion of the strand (protospacer) to be incorporated into the CRISPR array, at a location adjacent to the leader sequence in most bacteria. This process is hypothesized to depend on Cas1 and Cas2 proteins for acquisition of the protospacer. (ii) Transcription of the CRISPR array initiates at an upstream AT-rich region upon infection by foreign nucleotides. Long pre–CRISPR RNA (pre-crRNA) is refined into crRNA, and stabilized through dyadic symmetry of the repeat region. CRISPR–Cas Type I and III systems utilize Cas endonucleases, yet CRISPR–Cas II utilizes a trans-activating crRNA (tracrRNA), which complements the pre-crRNA and initiates cleavage through RNase III or Cas9. (iii) Spacer matching foreign nucleotide strand is incorporated into a Cas complex, which acts as a guide to locate nucleotides belonging to a repeated invasion by mobile genetic elements. In CRISPR–Cas Type III, this complex includes tracrRNA. This targeting initiates fragmentation of foreign genetic material.

the poor homologies obtained from the alignment of the amino acid residues of *E. coli*, *M. xanthus*, *P. atrosepticum*, and various *Streptococcus* species (this study) (Figures 2.11.4a,b).

Later, research on *Yersinia pestis*, *Streptococcus thermophilus*, and *Streptococcus vestibularis* determined the non-repetitive spacer region to be segments of exogenous DNA, from bacteriophage and plasmid origin, that was integrated into the host cell genome in a polar fashion, adjacent to the leader sequence (Bolotin *et al.*, 2005; Pourcel *et al.*, 2005). Recently, the polar addition scheme of spacer integration was questioned, as some archaea, namely *Sulfolobus solfataricus*, have been experimentally shown to integrate spacers at random locations internally, as opposed to locations near the leader sequence in the CRISPR array (Erdmann and Garrett, 2012). Nonetheless, there appears to be a loss of a previously embedded spacer in the array after integration of new spacer content, as in the CRISPR I array

in *Streptococcus agalactiae* (Lopez-Sanchez *et al.*, 2012), suggesting a dynamic nature of spacer content, a mechanism that assists in response to changing environments of mobile genetic elements.

Extrapolating from these research milestones, further genome-mapping efforts have revealed CRISPR–Cas to be present in many prokaryotes, with the current estimate at >40% for bacteria and, remarkably, >90% in archaeal species (Grissa *et al.*, 2007). Interestingly, there also exists an intraspecies incongruity of CRISPR arrays, as some strains may harbor the CRISPR–Cas system within their genome, and others may not. The absence of the CRISPR spacer repeats is typically accompanied by a lack of associated *cas* genes, leading to the assignment of functional interdependency of the artifacts in the interference of exogenous DNA (Touchon *et al.*, 2011). The number of spacer sequences within any particular strain

(a)

(b)

Figure 2.11.2 Multiple sequence alignment of the amino acid residues of the Cas1 protein from (a) *Escherichia coli*; and (b) *Streptococcus* species. The protein sequences were downloaded from the National Center for Biotechnology Information (NCBI) protein database (http://www.ncbi.nlm.nih.gov/protein/) and aligned with MAFFT alignment (MAFFT version 7, http://mafft.cbrc.jp/alignment/server/). The aligned protein sequences were managed and edited by using BioEdit (Hall, 1999). The identical sequences are shadowed in dark gray, and similar sequences are shadowed in light gray. The Cas1 protein sequences aligned well among different *E. coli* strains. This protein was poorly aligned among different *Streptococcus* species.

(a)

```
                                    10          20          30          40
                            . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
Cas2-Escherichia_coli_K-12      - - - - - - - - M S M L V V V T E N V P P R L R G R L A I W L L E V R A G V Y V
Cas2_Escherichia_coli_53638     M Q A I G V A E M S M L V V V T E N V P P R L R G R L A I W L L E V R A G V Y V
Cas2_Escherichia_coli_UMNK88    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Cas2_Escherichia_coli_O145:H28  - - - - - - - - M S M L V V V T E N V P P R L R G R L A I W L L E V R A G V Y V

                                    50          60          70          80
                            . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
Cas2-Escherichia_coli_K-12      G D V S A K I R E M I W E Q I A G L A E E G N V V M A W A T N T E T G F E F Q T
Cas2_Escherichia_coli_53638     G V V S A R I R E M I W E Q I S G L A E E G N V V M A W A T N T E S G F E F Q T
Cas2_Escherichia_coli_UMNK88    - - - - - - - - - M I W E Q I A G L A E E G N V V M A W A T N T E S G F E F Q T
Cas2_Escherichia_coli_O145:H28  G D V S A K I R E M I W E Q I S G L A E E G N V V M A W A T N T E S G F E F Q T

                                    90         100
                            . . . . | . . . . | . . . . | . . . . | . .
Cas2-Escherichia_coli_K-12      F G L N R R T P V D L D G L R L V S F L P V
Cas2_Escherichia_coli_53638     F G V N R R T P V D L D G L R L V S F L P V
Cas2_Escherichia_coli_UMNK88    F G V N R R T P V D L D G L R L V S F L P V
Cas2_Escherichia_coli_O145:H28  F G V N R R T P V D L D G L R L V S F L P F
```

(b)

```
                                        10          20          30          40
                                . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
Cas2_Streptococcus_agalactiae   M S Y R Y M R M I L M F D M P T E T A E E R K A Y R K F - - - R N F L L S E G -
Cas2_Streptococcus_pyogenes     M S Y R Y M R M I L M F D M P T D T A E E R K A Y R K F - - - R K F L L S E G -
Cas2_Streptococcus_dysgalactia  M S Y R Y M R M I L M F D M P T D T A E E R K V Y R K F - - - R K F L L S E G -
Cas2_Streptococcus_thermophilu  M S Y R Y M R M I L M F D M P T D T A E E R K A Y R K F - - - R K F L L S E G -
Cas2_Streptococcus_mutans       M S Y R Y M R M I L M F D M P T D T A E E R K A Y R K F - - - R K F L L S E G -
Cas2_Streptococcus_vestibulari  M R Y E A L R L L C F F D L P M E S K D E K R I Y R N F - - - R K E L I S N G -
Cas2_Streptococcus_sanguinis    M T - - - - - - - - F Y N L D S V S K E F A R K K T K F C L V I Y D I V S N K R

                                        50          60          70          80
                                . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
Cas2_Streptococcus_agalactiae   - F I M H Q F S V Y S K L L L N N T A N N A M I G R L K V N N P K K G N I T L L
Cas2_Streptococcus_pyogenes     - F I M H Q F S I Y S K L L L N N T A N N A M I G R L R E H N P N K G N I T L L
Cas2_Streptococcus_dysgalactia  - F I M H Q F S I Y S K L L L N N T A N N A M I G R L R E N N P H K G N I T L L
Cas2_Streptococcus_thermophilu  - F I M H Q F S V Y S K L L L N H T A N T A M V G R L K A N N P K K G N I T I L
Cas2_Streptococcus_mutans       - F I M H Q F S V Y S K L L L N N S A N T A M I A R L K E N N P K K G N I T M L
Cas2_Streptococcus_vestibulari  - F E M L Q F S V Y Y R T C P N R S F A N K F Y K K L K M S N L P A G N V R L L
Cas2_Streptococcus_sanguinis    R L K L A K L L E G Y G T R V Q R S C F E V D I E K L - - - - - - - - N F E L L

                                        90         100         110         120
                                . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
Cas2_Streptococcus_agalactiae   T V T E K Q F A R M - - - - - V Y L H G E R N T S V A N S D - - - - - - - - - S R
Cas2_Streptococcus_pyogenes     T V T E K Q F A R M - - - - - I Y L H G E R N N C I A N S D - - - - - - - - E R
Cas2_Streptococcus_dysgalactia  T V T E K Q F A R M - - - - - I Y L H G E R N N C I A N S D - - - - - - - - E R
Cas2_Streptococcus_thermophilu  T V T E K Q F A R M - - - - - I Y L Y G D K N T S I A N S E - - - - - - - - E R
Cas2_Streptococcus_mutans       T V T E K Q F A R M - - - - - I Y L N G E R D T S I T N S D - - - - - - - - S R
Cas2_Streptococcus_vestibulari  A V T E K Q F S E M - - - - T L I I G G K T K Q E E I V S D - - - - - - - - N K
Cas2_Streptococcus_sanguinis    I K D I R D F Y Q A D E G D N I I V Y V G H K E E T V V F N P Y D G A E L L E E

                                       130
                                . . . . | . . . . | . .
Cas2_Streptococcus_agalactiae   L V F L G D S Y D Q D -
Cas2_Streptococcus_pyogenes     L V F L G E A F D E S -
Cas2_Streptococcus_dysgalactia  L V F L G E A F D E S -
Cas2_Streptococcus_thermophilu  L V F L G D N Y C D E D
Cas2_Streptococcus_mutans       L V F L G E A F P D E T
Cas2_Streptococcus_vestibulari  L V V I - - - - - - - -
Cas2_Streptococcus_sanguinis    I L F F - - - - - - - -
```
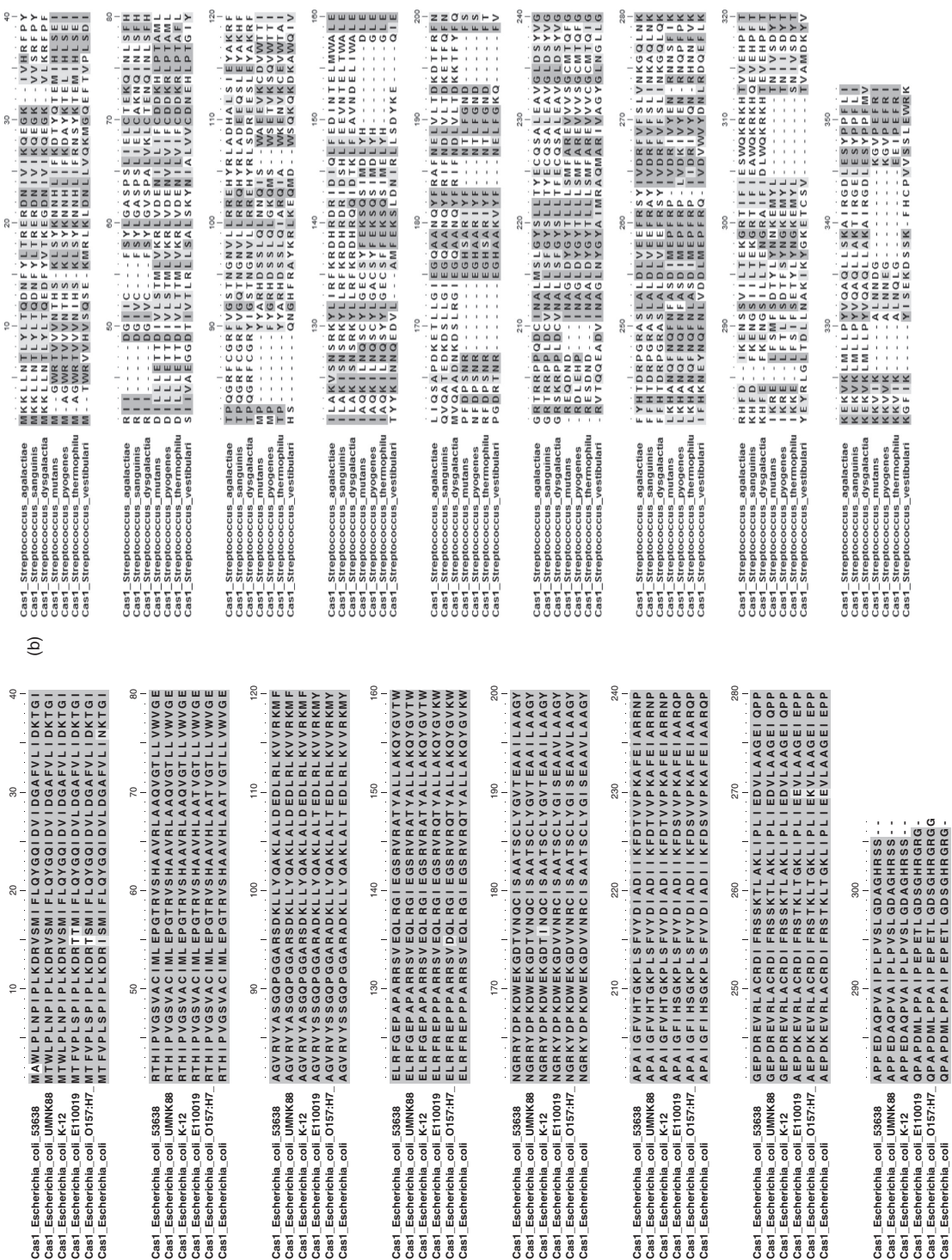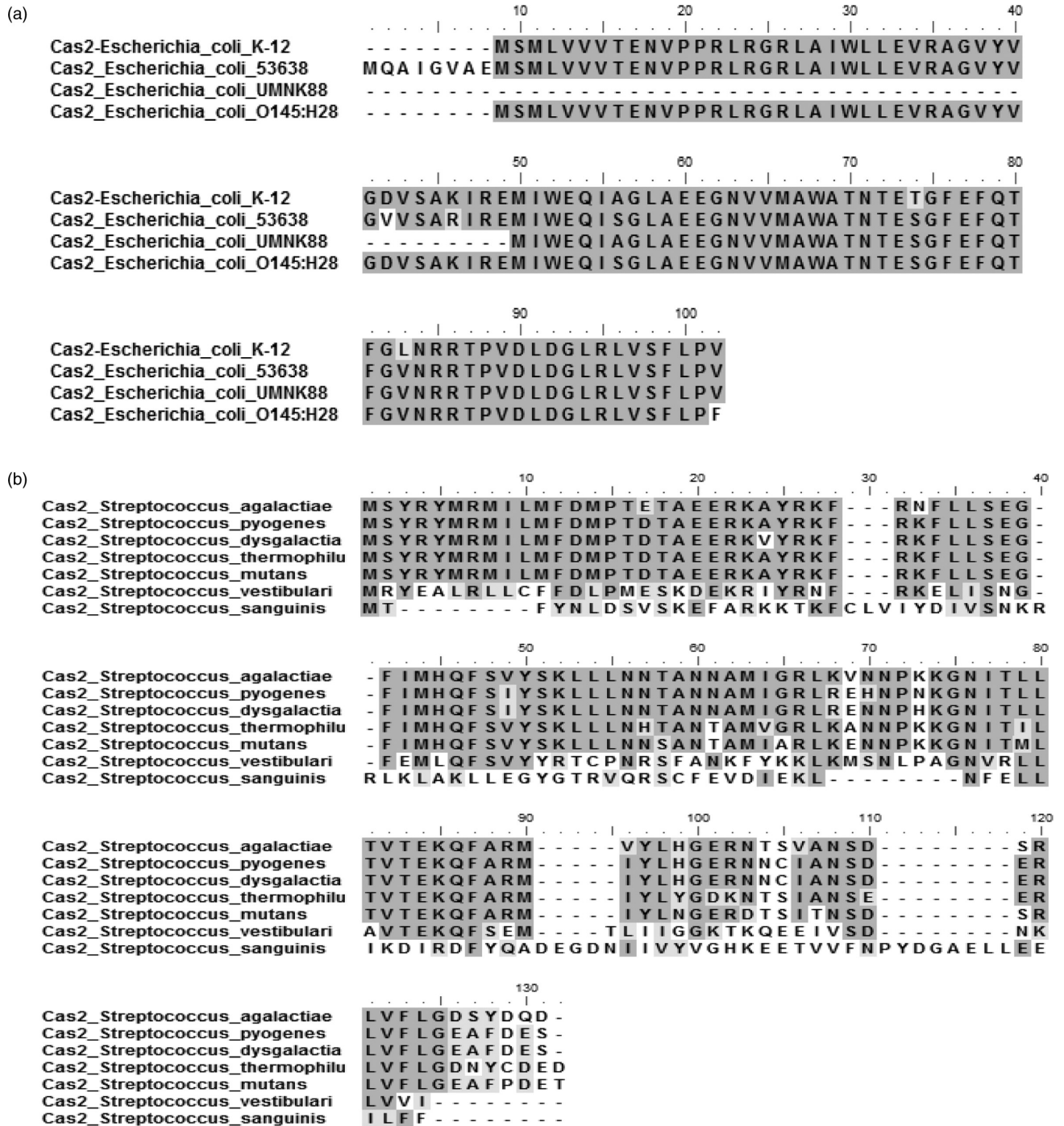
Figure 2.11.3 Multiple sequence alignment of the amino acid residues of the Cas2 protein sequences from (a) *Escherichia coli* strains; and (b) *Streptococcus* species. The Cas2 protein sequences were downloaded from the NCBI protein database (http://www.ncbi.nlm.nih.gov/protein/) and aligned with MAFFT alignment (MAFFT version 7, http://mafft.cbrc.jp/alignment/server/). The aligned protein sequences were managed and edited using BioEdit (Hall, 1999). The identical sequences are shadowed by dark gray, and similar sequences are shadowed by light gray. The alignment result showed that Cas2 protein sequences aligned well among various *E. coli* strains, whereas the *Streptococcus* species exhibited poor sequence homology.
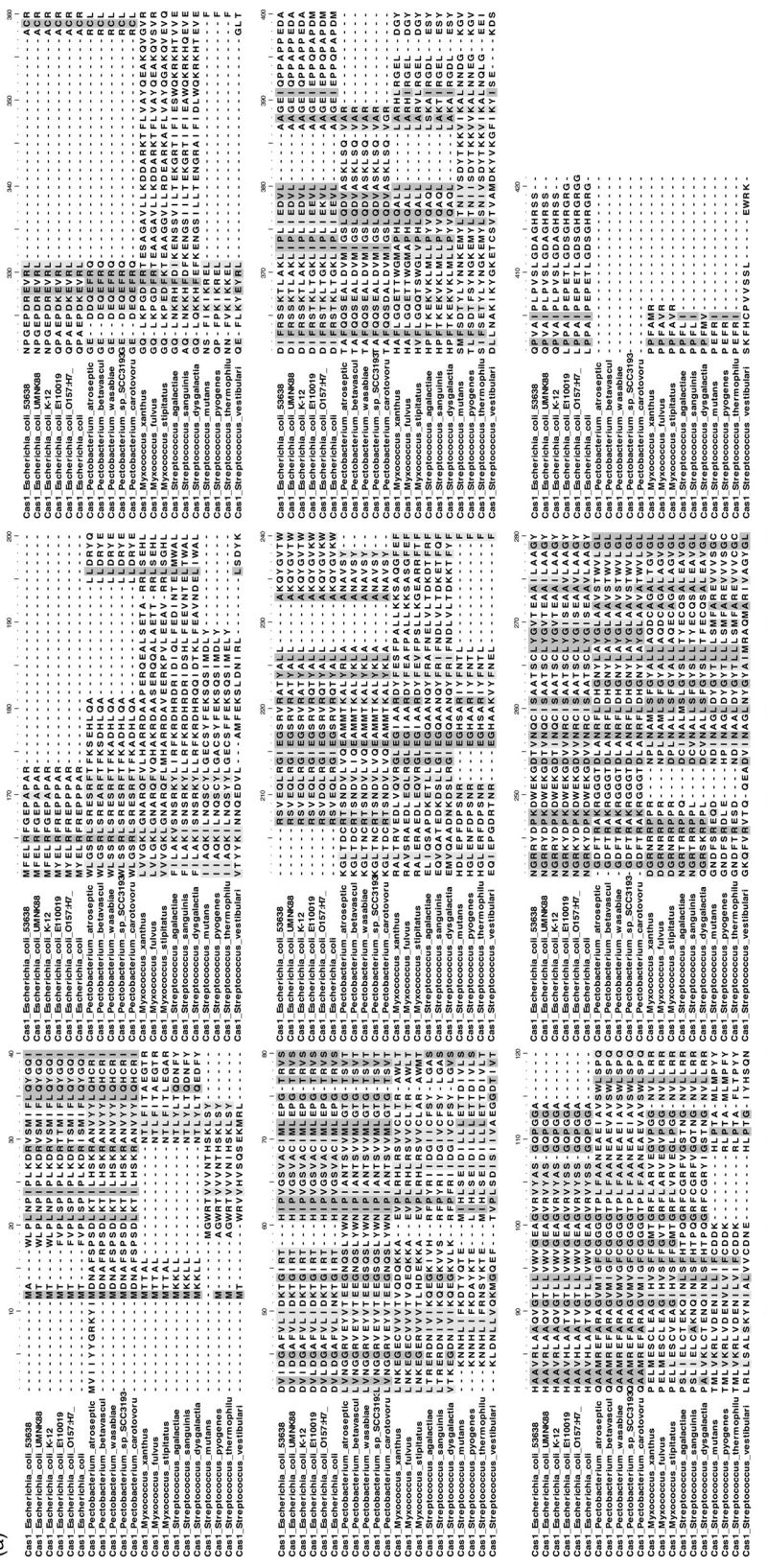
(a)

Figure 2.11.4 Multiple sequence alignment of amino acid residues of (a) Cas1 and (b) Cas2 from *Escherichia coli*, *Streptococcus*, *Pectobacterium*, and *Myxococcus* showing a diverse sequence structure and composition of these proteins. The protein sequences were downloaded from the NCBI protein database (http://www.ncbi.nlm.nih.gov/protein/) and aligned with MAFFT alignment (MAFFT version 7, http://mafft.cbrc.jp/alignment/server/). The aligned protein sequences were managed and edited using BioEdit (Hall, 1999). The identical sequences are shadowed in dark gray, and similar sequences are shadowed in light gray. (a) Cas1 and (b) Cas2 protein sequences did not align well within the different bacteria.

(b)

Cas2-Escherichia_coli_K-12
Cas2_Escherichia_coli_53638
Cas2_Escherichia_coli_UMNK88
Cas2_Escherichia_coli_O145:H28
Cas2_Pectobacterium_wasabiae
Cas2_Myxococcus_xanthus
Cas2_Myxococcus_sp._
Cas2_Myxococcus_stipitatus
Cas2_Myxococcus_fulvus
Cas2_Streptococcus_sanguinis
Cas2_Streptococcus_agalactiae
Cas2_Streptococcus_pyogenes
Cas2_Streptococcus_dysgalactia
Cas2_Streptococcus_thermophilu
Cas2_Streptococcus_mutans
Cas2_Streptococcus_vestibulari

(Multiple sequence alignment of Cas2 proteins, residues 1–120, shown across the page)

Figure 2.11.4 (*Continued*)

appears to be finite, as spacer quantity is maintained through removal of internal spacer sequences after integration of new spacers (Reyes *et al.*, 2008). This knowledge has been used as a means to subtype species on the basis of spacer count, coined "spoligotyping," as in "spacer typing" (Van der Zanden *et al.*, 2002). Intraspecies variation is further compounded as the nature, number, and arrangement of the Cas and other CRISPR-associated genes may vary across a single species (Pourcel and Drevet, 2013). Even the number of CRISPR–Cas systems in a genome may differ across the same species of bacterium (Carte *et al.*, 2014).

Although 12 subtypes have been recognized across bacterial species, it is generally agreed that three major types of CRISPR–Cas systems occur in bacteria, designated as Types I, II, and III – with Type I considered to be the most common (Makarova *et al.*, 2011a,b, 2013). Each type exemplifies a distinctive architecture of *cas* genes, varying in gene location relative to the CRISPR array, gene structure, and combinations of various *cas* genes (Makarova *et al.*, 2011b). Among all of the *cas* genes identified thus far, *cas1* and *cas2* have been recognized to be present in all CRISPR–Cas systems, with *cas1*–*cas6* considered the core genes of the CRISPR–Cas system (Haft *et al.*, 2005; Horvath and Barrangou, 2010). The target of the CRISPR–Cas mechanism is typically, but not always, double-stranded DNA, as it has been discovered that the Type III-B subtypes found in *Thermus thermophilus* target exogenous RNAs (Sorek *et al.*, 2013; Staals *et al.*, 2013).

2.11.2 Mechanism

2.11.2.1 The integration of spacer content

The CRISPR–Cas system functions as adaptive immunity through the complementation of a CRISPR RNA (crRNA) to an invading DNA element, typically of viral or plasmid origin. This causes the disruption of the foreign DNA through a series of nuclease-related events (Brouns *et al.*, 2008). Within the CRISPR array, crRNA is first transcribed as a long single-stranded product, and then processed by the Cas and CRISPR-related proteins, to produce short RNA sequences, which are 57 nucleotides in *E. coli* (Brouns *et al.*, 2008). This sequence is embedded into the Cas proteins (in which the specific Cas or CRISPR-related proteins will differ across bacterial species) and functions as a "guide" to locate complementary sequences on foreign DNA, perhaps introduced by a bacteriophage or a plasmid (Brouns *et al.*, 2008). The process from spacer acquisition to the RNA-mediated "immunity" or "protection" from the invading DNA occurs in three stages: (i) adaptation, (ii) expression, and (iii) interference (Figure 2.11.1) (Iranzo *et al.*, 2013; Van Der Oost *et al.*, 2009).

2.11.2.2 Adaptation

The acquisition of spacer content is dependent on invasion of foreign exogenous genetic elements, termed protospacers prior to integration, and is a fundamental requisite for the immunity of the CRISPR–Cas system (Bolotin *et al.*, 2005; Mojica *et al.*, 2005, 2009; Pourcel *et al.*, 2005). Of the three stages, evidence of the mechanism of initiation, as well as the underlining process of protospacer target and integration, has largely been inscrutable (Yosef *et al.*, 2012). This seems to be due to difficulties with recreating the adaptation stage under laboratory conditions, which contrasts the proficient and abundant integration of spacers under natural conditions (Andersson and Banfield, 2008; Tyson and Banfield, 2008; Weinberger *et al.*, 2012; Westra and Brouns, 2012). Some evidence has been forthcoming through investigation of *Pectobacterium atrosepticum* and *E. coli* (Datsenko *et al.*, 2012; Swarts *et al.*, 2012; Yosef *et al.*, 2012), strongly associating Cas1 and Cas2 as the hypothesized universal proteins responsible for spacer acquisition (Datsenko *et al.*, 2012; Yosef *et al.*, 2012). This is supported by evidence that both Cas1 and Cas2 proteins function in the absence of other CRISPR-related proteins to target and acquire spacers (Yosef *et al.*, 2012). In *E. coli*, Cas1 and Cas2 have been shown to form a complex that, once dissociated, will terminate spacer acquisition (Nuñez *et al.*, 2014). Additionally, a specialized 2 to 5 nt motif accompanying the foreign DNA has been observed to be a target for spacer acquisition: the protospacer adjacent motif (PAM), a recognizable sequence that has been observed to be a target across various bacterial CRISPR systems, and is hypothesized to be organism specific (Bolotin *et al.*, 2005; Mojica *et al.*, 2009, Shah *et al.*, 2013). This mode of adaptation has been termed the "naïve method", as the spacer acquisition occurs *de novo* (Fineran *et al.*, 2014).

Another type of adaptation has been observed within the K12 strain of *E. coli*, in which spacer acquisition was demonstrated to be associated with bacteriophage stress (see Chapter 2.10). As bacteria have evolved CRISPR–Cas systems of adaptive immunity, so too have phages and mobile elements to evade disruption. Specifically, mutations in the PAM sequence adjacent to protospacer regions allow for the avoidance of CRISPR-mediated immunity. As a mechanism, once a spacer sequence becomes ineffective as a crRNA for immunity through homologous complementation, the same sequence will act as a guide to partially complement stretches of genomic content on foreign mobile genetic elements (Fineran *et al.*, 2014). It has been observed that while utilizing CRISPR-associated proteins, namely Cas1, Cas2, and Cas3, alongside a guide RNA from the original spacer and an additional Cascade protein, the CRISPR machinery will recognize stretches of foreign DNA and select fragments to be incorporated as new spacer content (Datsenko *et al.*, 2012; Fineran *et al.*, 2014). This mechanism of adaptation has been designated the "priming" method of spacer acquisition (Fineran *et al.*, 2014), and it has recently begun to hold validity across other bacterial species, such as *Haloarcula hispanica* (Li *et al.*, 2014). As observed in *E. coli*, once the sequence has been acquired, insertion of the protospacer occurs through a staggered cleavage, followed by ligation, at the junction between the leader sequence and the first repeat, through the catalytic

potential of the metal-binding core of Cas1 in the Cas1–Cas2 complex (Arslan *et al.*, 2014). The use of bioinformatics tools revealed that occasionally the spacer sequences can be inserted internally within the CRISPR array, which was found to be unique as compared to the conventional process of the polar addition of spacers at the 5′ end of the CRISPR array (Erdmann and Garrett, 2012).

2.11.2.3 Expression

Although some mechanistic variation exists among the three types of CRISPR–Cas, there are common elements shared by all types, such as the transcription of the long pre-crRNA product, initiated at the upstream AT-rich region, followed by maturation of the transcript into short crRNA through cleavage internal to the repeat sequence (Horvath and Barrangou, 2010). Types I and III show a similarity in processing of the pre-crRNA product, as *cas* genes encoding endoribonucleases are responsible for cleavage, whereas Type II is dependent on a transactivating RNA (tracrRNA), which complements regions on the pre-crRNA and initiates cleavage through RNase III and Cas9 activity (Westra *et al.*, 2014).

2.11.2.4 Immunity

The CRISPR–Cas immunity occurs when crRNA is ushered by Cas proteins to complement stretches of nucleotides belonging to the mobile genetic elements of repeat invasions (Brouns *et al.*, 2008). As observed in the CRISPR–Cas Cse-subtype of *E. coli*, the matured crRNA is embedded into a complex of Cas proteins termed "CASCADE" (CRISPR-associated complex for antiviral defense), which is composed of CasA, CasB, CasC, CasD, and CasE (Brouns *et al.*, 2008; Jore *et al.*, 2011). After incorporation of the guide RNA into CASCADE, the complex is guided to the foreign element. Through the activity of an additional Cas protein, Cas3 (nuclease–helicase), foreign nucleotides will be fragmented (Figure 2.11.1) (Jore *et al.*, 2011).

2.11.3 Stress response of CRISPR

2.11.3.1 Bacteriophage stress

Interest has been placed on the CRISPR–Cas system, specifically addressing its function in adaptive immunity against predatory phages and harmful mobile genetic elements. Regarding stresses in the environment, bacteriophages have a prominent effect on shaping microbial communities in many niches, such as marine ecosystems (Mojica and Brussaard, 2014), temperate soil (Williamson *et al.*, 2013), and even the more extreme ecosystems such as Arctic soil (Allen *et al.*, 2010; see Chapter 10) and Antarctic lakes (Filippova *et al.*, 2013; Lauro *et al.*, 2010). Not surprisingly, bacteriophages imposing a lytic cycle will decrease microbial populations of susceptible species, with mortality estimates of about 50% in some environments (Fuhrman and Noble, 1995). Such an impact may affect the biotic ecosystem at a grander scale, especially when only select species are targeted. Decreases in select species may confer a shift in populations, thus leading to a change in nutrient cycling and metabolism, an occurrence that has been investigated in the context of the microbiome (Fuhrman and Noble, 1995; Koskella, 2013; Reyes *et al.*, 2012).

The effect of bacteriophages in causing bacterial stress is not restricted to population decline and species dynamics. Latent phages in the lysogenic life cycle, or prophages that integrate into a bacterial host as part of the life cycle, may confer a selective advantage of the host bacteria – a mechanism to increase persistence of the phage by enhancing the survivorship of the phage-harboring host (Bossi *et al.*, 2003). Additionally, phage-mediated gene trafficking, or transduction, may offer a fitness advantage to certain bacteria over others in the natural ecosystem by inserting DNA into the genome of the host that assists in metabolic function – again, conferring a fitness advantage to both the bacteria and the phage (Mann *et al.*, 2003). Although seemingly benign, these strategies may inadvertently cause stress to cohabiting bacteria through competition, creating shifts in the microbial populations and thereby affecting the ecosystem function.

The effect of CRISPR–Cas systems in response to lytic phages has been researched in various bacterial species, in which a fitness benefit was established (Brouns *et al.*, 2008; Marraffini and Sontheimer, 2008; Sorek *et al.*, 2008). However, CRISPR–Cas is not limited to protection against lytic phages, as it can interfere with the genome of lysogenic phages. In 2010, the range of CRISPR–Cas potential against invading genetic elements was investigated in *E. coli* in the presence of λ bacteriophage (Edgar and Qimron, 2010). A range of activity against the phage in each life cycle was observed. First, in the presence of a lysogenetic phage, CRISPR–Cas-mediated interference was shown to target segments of the phage genome, interrupting the cycle within the cell. Although segments of the phage genome responsible for lysogenization were not specifically targeted by CRISPRs, this is evidence in favor of broad CRISPR function in preserving the bacterial genome. CRISPR–Cas response to prophage integration was also investigated, as the CRISPR–Cas system was induced after the integration of the phage DNA segments. It was observed that the CRISPR harboring *E. coli* cleared prophage DNA in a small portion, showing that those bacteria harboring a spacer congruent with the integrated λcI857kan (kanamycin resistant temperature sensitive cI variant) prophage had a ≈500-fold increase in survival, even after the expected lytic cycle induction of the phage at a temperature increase (42 °C) (Edgar and Qimron, 2010). Yet largely, CRISPR–Cas-mediated apoptosis occurred through self-targeting, or the prophage cycle was too progressed for timely removal, leading to cell death of the bacteria affected by the prophages. In addition to the apparent function of CRISPR–Cas in response to bacteriophage stress, CRISPR–Cas systems are functional under many phage circumstances, an area that requires additional research.

2.11.3.2 Fruiting body and starvation in *Myxococcus xanthus*

When subjected to a nutrient-deprived environment, certain bacteria have been observed to undergo colony structure adjustments to cope with starvation stresses. One such adjustment, fruiting body formation, has recently been associated with the CRISPR–Cas system in *Myxococcus xanthus* (Viswanathan *et al.*, 2007a). During periods of long starvation, myxobacterial cells undergo fruiting body development through a process of cell–cell contact interaction (Kiskowski *et al.*, 2004; Sozinova *et al.*, 2005). The development from aggregation to sporulation, including rippling that accompanies and precedes aggregation (Shimkets and Kaiser, 1982), has been linked to the production of five intercellular signals, canonically named the A, B, C, D, and E factors (Boysen *et al.*, 2002; Downard *et al.*, 1993; Kim *et al.*, 1992). Crucial to the fruiting body process, the A factor signal is initiated by starvation, which induces the transcription of the FruA DNA-binding response regulator (Ellehauge *et al.*, 1998; Viswanathan *et al.*, 2007b). Additionally, cell–cell contact will initiate the C factor, encoded by *csgA* gene (Kaiser *et al.*, 2010), which will activate FruA to induce *devR* and *devS* genes (Jelsbak and Søgaard-Andersen, 2000). The *devRS* genes have been identified to be a part of a larger locus that includes *devT* gene – which stimulates FruA synthesis to positively regulate *devR* and *devS* genes (Boysen *et al.*, 2002).

Recent analysis of the *dev* genes of *M. xanthus* revealed an association between the *dev* locus and the CRISPR–Cas array. Specifically, genome mapping has placed the locus as nested within the arrangement of the *cas* genes of *M. xanthus*, and has determined *devT*, *devR*, and *devS* genes to be *cas8*, *cas7*, and *cas5*, respectively (Boysen *et al.*, 2002; Jelsbak and Søgaard-Andersen, 2000). The consequential transcription of the *dev* genes in response to starvation occurs alongside the *cas* genes (Boysen *et al.*, 2002), with only a portion of the CRISPR repeat segment transcribing with the *dev* locus. This CRISPR array portion contains a spacer corresponding to an integrase enzyme, necessary for the lysogenization of the Mx8 bacteriophage into the genome (Viswanathan *et al.*, 2007a). An interesting hypothesis proposed by Viswanathan *et al.* (2007a,b) was that, mechanistically, the Mx8 spacer is transcribed as a crRNA in defense against the Mx8 bacteriophage, as the *M. xanthus* may be particularly susceptible to lysogenization during starvation periods. Little evidence exists to verify the physiological importance of transcription of the Mx8 spacer. However, considering the co-transcription of the *dev* and *cas* genes, and the genetic architecture of the *dev* locus in relation to the CRISPR–Cas array, there may be an association of the two systems in response to environmental stresses, but the precise role of the system remains unknown.

2.11.3.3 DNA damage repair

Prior to adopting the Cas designation, the genes associated with the CRISPR system of exogenous DNA silencing were linked with DNA repair, an observation that came through the genomic analysis of thermophilic bacteria (Makarova *et al.*, 2002). Generally, extremophile bacteria inhabiting harsh niches will have a phenotype and underlying genotype that confer survival under the stress of DNA-damaging conditions (Stan-Lotter and Fendrihan, 2012). For example, desiccation, UV, and gamma radiation are common DNA-damaging agents in Antarctica, yet bacteria such as *Pseudomonas syringae* Lz4W can repair excessively damaged DNA and otherwise lethal amounts of genome fragmentation (Pavankumar *et al.*, 2010; Sinha and Häder, 2002; see Section 8).

A gene association exists between the Cas genes and DNA repair enzymes in the *E. coli* CRISPR–Cas type I-E (Babu *et al.*, 2011). Analysis of the Cas1 protein (YgbT in *E. coli*) revealed a multifaceted nuclease activity capable of processing substrates of branched DNA, such as 5′ flaps, Holiday junctions, and replication forks (Babu *et al.*, 2011). In the same research, co-purification of Cas1 revealed an association of Cas1 with RuvB, a helicase that functions in the RuvAB complex, which migrates Holiday junctions for the eventual RuvC-mediated double-stranded break of the target DNA (West, 1996). Additionally observed was an association of Cas1 with RecB and RecC, two enzymes that are crucial to the recombinational restoration of template imperfections in DNA (Kowalczykowski, 2000). RecB and RecC will form a complex with RecD, which inserts at the nicks induced by the RuvABC system, to initiate the homologous recombination of DNA (Kowalczykowski, 2000). RecBCD-mediated DNA repair was also identified, and examined, in the extremophile Antarctic bacteria *P. syringae* Lz4W (Pavankumar *et al.*, 2010). In *P. syringae*, deletion of the RecBCD led to increased susceptibility to DNA damage by UV radiation, as well as decreased viability when exposed to low temperatures (4 °C) (Pavankumar *et al.*, 2010; Regha *et al.*, 2005).

In addition to its observed function in the CRISPR–Cas system of immunity, the *E. coli* Cas1 (YgbT) interacts with RecB to repair DNA. Many CRISPR–Cas-containing bacteria possess the Cas4 gene (Haft *et al.*, 2005), and in certain species such as *M. xanthus*, the Cas1 and Cas4 gene sequences appear overlapping with one another, emphasizing the association of the two gene products (Zhang *et al.*, 2012). *E. coli* does not contain the Cas4 gene per se, yet there is a homology between the Cas4 and the RecB exonuclease, and an apparent association of the two proteins in DNA repair (Jansen *et al.*, 2002). Cas3 was discovered to contain functional domains consistent with helicases belonging to superfamily 2 of DNA and RNA metabolism (Jansen *et al.*, 2002). Although the CRISPR array of spacer repeats has not been verified to function in DNA repair, interestingly, the RecBCD system has previously been implicated in the interruption of foreign linear DNA, a characteristic of CRISPR–Cas immunity (Dillingham and Kowalczykowski, 2008). This is a noteworthy observation implicating a dual functionality of the *cas* genes in both immunity and host DNA repair.

2.11.3.4 Envelope stress as a trigger for CRISPR–Cas induction

Central to coping with environmental stress is the DnaK protein, which, in *E. coli*, is correlated with the heat stress response (see Section 13). Tilly *et al.* (1983) investigated the DnaK protein and its production during temperature increases in *E. coli*; and, in 1984, Bardwell and Craig (1984) reported their findings, classifying the protein as a heat shock protein (HSP) due to its previously described homology to known HSPs across many eukaryotic and prokaryotic genomes. Subsequent investigations elucidated the role and function of DnaK in folding and chaperoning of newly synthesized proteins in the cytosol (Pérez-Rodríguez *et al.*, 2007), for export across the cytoplasmic membrane via the Tat (twin-arginine translocation) pathway (Graubner *et al.*, 2007). Crucial to envelope integrity is the efficient transport of folded proteins and the avoidance of amassing misfolded or unfolded proteins in the periplasm (Raivio, 2005).

To test the association of the CRISPR–Cas system with the DnaK chaperone protein in the context of envelope stress (see Section 18), the involvement of CasE of *E. coli* in response to extracellular stress has been reported (Pérez-Rodríguez *et al.*, 2011). Plasmids containing *sstorA* fused to *gfp* (*sstorA-gfp*) were transformed into a DnaK-deficient (ΔDnaK) *E. coli* culture – which causes the accumulation of insoluble unfolded ssTorA product, conferring envelope stress (Deuerling *et al.*, 1999; Lim *et al.*, 2009). ssTorA is a bacterial Tat-specific *E. coli* trimethylamine *N*-oxide (TMAO) reductase (TorA) signal peptide, plus the first four residues of mature TorA (ssTorA), that helps reduce TMAO into trimethylamine (TMA) as part of the electron transport chain (Fisher *et al.*, 2006). It was discovered that endogenous sequences consisting of spacers homologous to *sstorA* silenced the plasmid and subdued expression of the misfolded protein (Pérez-Rodríguez *et al.*, 2011). Attention was placed on the BaeSR regulatory system, in which BaeS resides in the membrane and phosphorylates the BaeR transcription activator in the event of envelope stress (MacRitchie *et al.*, 2008). BaeR is a response regulator of the putative BaeSR two-component signal transduction system for regulation of *E. coli* multidrug transporter genes (Baranova and Nikaido, 2002). As BaeR was known to activate CasA in *E. coli* (Baranova and Nikaido, 2002), the binding of BaeR to CasA in the envelope stress due to inefficient protein translocation was inspected, and it was observed that BaeR binds internal to the CasA gene. This association with membrane-maintaining components has elucidated a potential role of CRISPR–Cas in membrane preservation.

An additional association of CRISPR–Cas to membrane integrity was investigated in the pathogen *Francisella novicida*, where membrane stress was offset by the CRISPR–Cas system, specifically through the function of Cas9 (Sampson *et al.*, 2014). *F. novicida* must cope with many antimicrobial agents residing within the target host organism and the macrophages that engulf the bacteria (Jones *et al.*, 2012). Unique to this bacterium is its ability to survive exposure to certain antimicrobials, namely those that are cationic peptides (Balagopal *et al.*, 2006). In their research, polymyxin B, a membrane-acting antimicrobial agent used to study the effect of cationic peptide antimicrobials (Mohapatra *et al.*, 2007; Wang *et al.*, 2006), was administered to *F. novicida* lacking the Cas9 gene. It was observed that at low doses (100 μg/mL as compared to the lethal dose of 800 μg/mL), the ΔCas9 bacteria displayed a susceptibility to the antimicrobial as compared to the wild-type strain (Sampson *et al.*, 2014). In addition, two small RNAs that associate with Cas9, tracrRNA and small Cajal body–specific RNA (scaRNA), were also investigated. It was determined that, upon deletion of the two RNA products, a similar susceptibility to that of Cas9 mutants was observed.

Besides the observed association of Cas9 to membrane stress in *F. novicida*, the Cas9 gene of *Streptococcus pyogenes* (SpCas9) is a crucial factor in the Type II CRISPR–Cas system, coding for an RNA-guided DNA endonuclease enzyme. This SpCas9 enzyme has been reported to perform efficient RNA-guided sequence-specific DNA cleavage. By modifying the SpCas9 and the CRISPR system to create a gene-editing tool, it has become possible to utilize a predetermined RNA guide strand to target a specific location in the genome of a target cell's DNA, effectively utilizing the Cas9 endonuclease property to nick the target and nontarget strand of a double-stranded DNA (Cong *et al.*, 2013). This could be used either to disrupt a gene for knockout study or to allow the cell's innate excision repair mechanism to incorporate a new strand of DNA, which may be introduced by the researcher to the system, allowing the DNA to integrate at the location of the site of cleavage (Jinek *et al.*, 2012; Mali *et al.*, 2013). Recently, the comparatively smaller Cas9 protein of the Type II CRISPR–Cas system of *Staphylococcus aureus* (SaCas9) was proposed as a more effective nuclease for the RNA-guided genome-editing strategy (Ran *et al.*, 2015). Homologous Cas9 genes are related, and the degree of their similarity is dependent on the CRISPR–Cas subtype: Type II-A, Type II-B, or Type II-C (Fonfara *et al.*, 2013). Conserved across these homologs are the amino acids of the NHN and RuvC domains, which nick the target and nontarget strand, respectively (Fonfara *et al.*, 2013; Jinek *et al.*, 2012). However, between the catalytic domains are a variability of amino acids (Chylinski *et al.*, 2014; Fonfara *et al.*, 2013). A study by Fonfara *et al.* (2013), using *S. pyogenes* (CRISPR–Cas Type II-A) as the model bacteria, demonstrated that only the Cas9 genes belonging to bacteria that possess a closely related Type II-A system can be interchanged while still maintaining CRISPR–Cas functionality. Particularly, it was shown that the key differences in the Cas9 protein functionality across the various Type II subtypes were in the maturing crRNA, as well as the stabilization in the tracrRNA and pre-crRNA duplex, which is the eventual complex that targets invading foreign genetic elements (Fonfara *et al.*, 2013). The striking diversity in the structure and gene-editing efficiency of the Cas9 between the two aforementioned bacteria reveals an

evolutionary divergence, potentially tailoring to the immunity requirements of the bacteria, as it is targeted by species-specific phage infection.

As the evidence implicating the CRISPR–Cas system in envelope stress amasses, speculation has arisen regarding the involvement of membrane cues to trigger the CRISPR–Cas system. The role of heat-stable nucleoid structuring (H-NS) protein in the suppressive function of CasA transcription in the *E. coli* CRISPR–Cas system has been reported (Pul *et al.*, 2010). The leucine response transcription factor, LeuO, was found to be a positive regulator of CasA (Westra *et al.*, 2010). The H-NS–LeuO regulation function was later observed in *Salmonella enterica* serovar Typhi (Medina-Aparicio *et al.*, 2011). An additional regulatory function of H-NS has been proposed indicating that the protein is sequestered from the CRISPR–Cas array in the event of exogenous genetic penetration, promoting the transcription of the CRISPR–Cas system (Navarre *et al.*, 2006, 2007). With this information pooled, Tracy Raivio (2011) summarized a potential mechanism for the initiation of CRISPR–Cas induction. It was hypothesized that a signal exists at the membrane, which is triggered through either the penetration of exogenous DNA or the production of atypical protein products through foreign DNA. This signal initiates the phosphorylation of BaeR by BaeS at the membrane, which then transfers to the CRISPR–Cas system for transcription initiation. Additional positive regulation of the CRISPR–Cas locus through LeuO will occur as H-NS binds with exogenous DNA. Although key components in the cascade are yet to be determined, this offers a deeper look into the CRISPR–Cas system, elucidating the induction of CRISPR–Cas in the event of membrane stress, and thereupon providing insight into the initiation of CRISPR–Cas as it pertains to cell immunity (Raivio, 2011).

2.11.3.5 Cas1 and Cas2 are toxin–antitoxin systems leading to dormancy

The overexpression of the toxin of a toxin–antitoxin (TA) system has been observed to induce cell dormancy (see Chapter 2.7), and eventual cell death in the advent of stress. It has also been hypothesized to assist in survivability of cells, as shown in *E. coli*, where the RelE and ChpAK toxins induce a static state of cell metabolism, allowing for metabolic halt until the antitoxin is expressed (Pedersen *et al.*, 2002). In a vegetative cell, the toxin and its associated antitoxin are co-expressed, forming a stable complex that subdues the activity of the toxin (Van Melderen and De Bast, 2009; Yamaguchi *et al.*, 2011). In the context of phage stress, the TA system, mainly the toxin, has been linked to abortive phage infection, neutralizing both phage and host cell transcripts, leading to cell suicide and phage elimination (Cook *et al.*, 2013; Fineran *et al.*, 2009).

Interestingly, Makarova *et al.* (2012) linked the CRISPR–Cas system of *E. coli* immunity to a TA system, specifically addressing the role of Cas1 and Cas2 genes as key genes for inducing dormancy and/or abortive phage infection in the event of

bacteriophage infection. They hypothesized that as a cell is invaded by a phage, cell dormancy is triggered to allow time for adequate CRISPR–Cas immunity, and/or to induce apoptosis to prevent integration and spread of the viral genetic material. Their assumption was based on the observation that Cas1 and Cas2 are universal across many CRISPR–Cas systems, and that Cas2 shares homology to the mRNA-cleaving VapD (Daines *et al.*, 2004; Kwon *et al.*, 2012), acting as a toxin that induces dormancy and/or death following infection. The antitoxin, Cas1, forms a complex with Cas2 in the nonstressed cell, and in the event of bacteriophage infection, Cas2 is presumably degraded, allowing for Cas1 (toxin) accumulation. However, the link of the *cas* genes to this area of stress coping requires more investigation, and it is highly theoretical (Westra *et al.*, 2014), as a sufficient link between the two systems is yet to be determined. This is an intriguing association in the realm of stress response and CRISPR–Cas.

Conclusion

As CRISPR–Cas is in the early stages of investigation, the scope of functionality has yet to be fully comprehended. As a means to adaptive immunity (i.e., evolution in progress), the CRISPR–Cas system has been observed to combat bacteriophage infection, in addition to other types of foreign genetic material in bacteria. This is an important function conferring survivorship to targeted bacterial cells, as phages have been observed to cohabit many ecological niches with bacteria, with the major observable effect being population decrease due to lysing as part of the bacteriophage life cycle (Ashelford *et al.*, 2003; Bergh *et al.*, 1989; Díaz-Muñoz and Koskella, 2014; Mojica and Brussaard, 2014; Proctor and Fuhrman, 1990; Williamson *et al.*, 2013). Alternate roles of CRISPR–Cas have been investigated, as components of the CRISPR–Cas pathway appear to be involved in other bacterial stress responses.

Acknowledgements

References

Allen B, Willner D, Oechel WC, Lipson D. 2010. Top-down control of microbial activity and biomass in an Arctic soil ecosystem. *Environmental Microbiology* 12(3): 642–648.

Andersson AF, Banfield JF. 2008. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* 320: 1047–1050.

Arslan Z, Hermanns V, Wurm R, Wagner R, Pul Ü. 2014. Detection and characterization of spacer integration intermediates in type IE CRISPR–Cas system. *Nucleic Acids Research* 42(12): 7884–7893.

Ashelford KE, Day MJ, Fry JC. 2003. Elevated abundance of bacteriophage infecting bacteria in soil. *Applied and Environmental Microbiology* 69(1): 285–289.

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. *Molecular Microbiology* 79(2): 484–502.

Balagopal A, MacFarlane AS, Mohapatra N, Soni S, Gunn JS, Schlesinger LS. 2006. Characterization of the receptor-ligand pathways important for entry and survival of *Francisella tularensis* in human macrophages. *Infection and Immunity* 74(9): 5114–5125.

Baranova N, Nikaido H. 2002. The *baeSR* two-component regulatory system activates transcription of the *yegMNOB* (*mdtABCD*) transporter gene cluster in *Escherichia coli* and increases its resistance to novobiocin and deoxycholate. *Journal of Bacteriology* 184(15): 4168–4176.

Bardwell JC, Craig EA. 1984. Major heat shock gene of *Drosophila* and the *Escherichia coli* heat-inducible *dnaK* gene are homologous. *Proceedings of the National Academy of Sciences USA* 81(3): 848–852.

Bergh Ø, BØrsheim KY, Bratbak G, Heldal M. 1989. High abundance of viruses found in aquatic environments. *Nature* 340(6233): 467–468.

Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. 2005. Clustered regularly inter-spaced short palindromic repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151(8): 2551–2561.

Bossi L, Fuentes JA, Mora G, Figueroa-Bossi N. 2003. Prophage contribution to bacterial population dynamics. *Journal of Bacteriology* 185(21): 6467–6471.

Boysen A, Ellehauge E, Julien B, Søgaard-Andersen L. 2002. The DevT protein stimulates synthesis of FruA, a signal transduction protein required for fruiting body morphogenesis in *Myxococcus xanthus*. *Journal of Bacteriology* 184(6): 1540–1546.

Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, *et al.* 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321(5891): 960–964.

Buckling A, Rainey PB. 2002. Antagonistic coevolution between a bacterium and a bacteriophage. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269(1494): 931–936.

Burgess JG, Jordan EM, Bregu M, Mearns-Spragg A, Boyd KG. 1999. Microbial antagonism: a neglected avenue of natural products research. *Progress in Industrial Microbiology* 35: 27–32.

Carte J, Christopher RT, Smith JT, Olson S, Barrangou R, *et al.* 2014. The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Molecular Microbiology* 93(1): 98–112.

Cary SC, McDonald IR, Barrett JE, Cowan DA. 2010. On the rocks: the microbiology of Antarctic Dry Valley soils. *Nature Reviews Microbiology* 8(2): 129–138.

Chylinski K, Makarova KS, Charpentier E, and Koonin EV. 2014. Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Research* 42(10): 6091–6105.

Cong L, Ran FA, Cox D, Lin S, Barretto R, *et al.* 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121): 819–823.

Cook GM, Robson JR, Frampton RA, McKenzie J, Przybilski R, *et al.* 2013. Ribonucleases in bacterial toxin–antitoxin systems. *Biochimica et Biophysica Acta (BBA) – Gene Regulatory Mechanisms* 1829(6): 523–531.

Daines DA, Jarisch J, Smith AL. 2004. Identification and characterization of a nontypeable *Haemophilus influenzae* putative toxin-antitoxin locus. *BMC Microbiology* 4(1): 30.

Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. 2012. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature Communications* 3: 945.

Deuerling E, Schulze-Specking A, Tomoyasu T, Mogk A, Bukau B. 1999. Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* 400(6745): 693–696.

Díaz-Muñoz SL, Koskella B. 2014. Bacteria–phage interactions in natural environments. *Advances in Applied Microbiology* 89: 135.

Díez-Villaseñor C, Marco Guzmán N, Almendros Romero C, García Martínez J, Martínez Mojica FJ. 2013. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas IE variants of *Escherichia coli*. *RNA Biology* 10(5): 792–802.

Dillingham MS, Kowalczykowski SC. 2008. RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiology and Molecular Biology Reviews* 72(4): 642–671.

Downard J, Ramaswamy SV, Kil KS. 1993. Identification of *esg*, a genetic locus involved in cell-cell signaling during *Myxococcus xanthus* development. *Journal of Bacteriology* 175(24): 7762–7770.

Edgar R, Qimron U. 2010. The *Escherichia coli* CRISPR system protects from λ lysogenization, lysogens, and prophage induction. *Journal of Bacteriology* 192(23): 6291–6294.

Ellehauge E, Nørregaard-Madsen M, Søgaard-Andersen L. 1998. The FruA signal transduction protein provides a checkpoint for the temporal co-ordination of inter-cellular signals in *Myxococcus xanthus* development. *Molecular Microbiology* 30: 807–817.

Erdmann S, Garrett RA. 2012. Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Molecular Microbiology* 85(6): 1044–1056.

Fierer N, Jackson RB. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America* 103(3): 626–631.

Filippova SN, Surgucheva NA, Kulikov EE, Sorokin VV, Akimov VN, *et al.* 2013. Detection of phage infection in the bacterial population of Lake Untersee (Antarctica). *Microbiology* 82(3): 383–386.

Fineran PC, Blower TR, Foulds IJ, Humphreys DP, Lilley KS, Salmond GP. 2009. The phage abortive infection system, ToxIN, functions as a protein–RNA toxin–antitoxin pair. *Proceedings of the National Academy of Sciences USA* 106(3): 894–899.

Fineran PC, Gerritzen MJ, Suárez-Diez M, Künne T, Boekhorst J, *et al.* 2014. Degenerate target sites mediate rapid primed CRISPR adaptation. *Proceedings of the National Academy of Sciences USA* 111(16): E1629–E1638.

Fisher AC, Kim W, Delisa MP. 2006. Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. *Protein Science*, 15(3): 449–458.

Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lécrivain AL, *et al.* 2013. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Research* gkt1074.

Fuhrman JA, Noble RT. 1995. Viruses and protists cause similar bacterial mortality in coastal seawater. *Limnology and Oceanography* 40(7): 1236–1242.

Gómez P, Buckling A. 2011. Bacteria-phage antagonistic coevolution in soil. *Science* 332(6025): 106–109.

Graubner W, Schierhorn A, Brüser T. 2007. DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone. *Journal of Biological Chemistry* 282(10): 7116–7124.

Grissa I, Vergnaud G, Pourcel C. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8(1): 172.

Gurung TB, Kagami M, Yoshida T, Urabe J. 2001. Relative importance of biotic and abiotic factors affecting bacterial abundance in Lake Biwa: an empirical analysis. *Limnology* 2(1): 19–28.

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 computational Biology* 1(6): e60.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). *Nucleic Acids Symposium Series* 41: 95–98.

Horvath P, Barrangou R. 2010. CRISPR/Cas, the immune system of bacteria and Archaea. *Science* 327: 167–170.

Iranzo J, Lobkovsky AE, Wolf YI, Koonin EV. 2013. Evolutionary dynamics of the prokaryotic adaptive immunity system CRISPR-Cas in an explicit ecological context. *Journal of Bacteriology* 195(17): 3834–3844.

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. *Journal of Bacteriology* 169: 5429–5433.

Jansen R, Embden JD, Gaastra W, Schouls LM. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology* 43(6): 1565–1575.

Jelsbak L, Søgaard-Andersen L. 2000. Pattern formation: fruiting body morphogenesis in *Myxococcus xanthus*. *Current Opinion in Microbiology* 3(6): 637–642.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, and Charpentier E. 2012. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096): 816–821.

Jones CL, Napier BA, Sampson TR, Llewellyn AC, Schroeder MR, Weiss DS. 2012. Subversion of host recognition and defense systems by *Francisella* spp. *Microbiology and Molecular Biology Reviews* 76(2): 383–404.

Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, *et al.* 2011. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nature Structural & Molecular Biology* 18(5): 529–536.

Kaiser D, Robinson M, Kroos L. 2010. *Myxobacteria*, polarity, and multicellular morphogenesis. *Cold Spring Harbor Perspectives in Biology* 2(8): a000380.

Kim SK, Kaiser D, Kuspa A. 1992. Control of cell density and pattern by intercellular signaling in *Myxococcus* development. *Annual Reviews in Microbiology* 46(1): 117–139.

Kiskowski MA, Jiang Y, Alber MS. 2004. Role of streams in *Myxobacteria* aggregate formation. *Physical Biology* 1(3): 173.

Koskella B. 2013. Phage-mediated selection on microbiota of a long-lived host. *Current Biology* 23(13): 1256–1260.

Kostic AD, Howitt MR, Garrett WS. 2013. Exploring host–microbiota interactions in animal models and humans. *Genes and Development* 27(7): 701–718.

Kowalczykowski SC. 2000. Initiation of genetic recombination and recombination-dependent replication. *Trends in Biochemical Sciences* 25(4): 156–165.

Kristjánsson JK, Hreggvidsson GO. 1995. Ecology and habitats of extremophiles. *World Journal of Microbiology and Biotechnology* 11(1): 17–25.

Kwon AR, Kim JH, Park SJ, Lee KY, Min YH, *et al.* 2012. Structural and biochemical characterization of HP0315 from *Helicobacter pylori* as a VapD protein with an endoribonuclease activity. *Nucleic Acids Research* 40(9): 4216–4228.

Lauro FM, DeMaere MZ, Yau S, Brown MV, Ng C, *et al.* 2010. An integrative study of a meromictic lake ecosystem in Antarctica. *The ISME Journal* 5(5): 879–895.

Li M, Wang R, Xiang H. 2014. *Haloarcula hispanica* CRISPR authenticates PAM of a target sequence to prime discriminative adaptation. *Nucleic Acids Research* 42(11): 7226–7235.

Lim HK, Mansell TJ, Linderman SW, Fisher AC, Dyson MR, DeLisa MP. 2009. Mining mammalian genomes for folding competent proteins using Tat-dependent genetic selection in *Escherichia coli*. *Protein Science* 18(12): 2537–2549.

Lopez-Sanchez MJ, Sauvage E, Da Cunha V, Clermont D, Ratsima Hariniaina E, *et al.* 2012. The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Molecular Microbiology* 85(6): 1057–1071.

MacRitchie DM, Buelow DR, Price NL, Raivio TL. 2008. Two-component signaling and gram negative envelope stress response systems. *In Bacterial Signal Transduction: Networks and Drug Targets* 631: 80–110.

Makarova KS, Anantharaman V, Aravind L, Koonin EV. 2012. Live virus-free or die: coupling of antivirus immunity and programmed suicide or dormancy in prokaryotes. *Biology Direct* 7: 40.

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 research* 30(2): 482–496.

Makarova KS, Aravind L, Wolf YI, Koonin EV. 2011a. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biology Direct* 6: 38.

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, *et al.* 2011b. Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology* 9: 467–477.

Makarova KS, Wolf YI, Koonin EV. 2013. The basic building blocks and evolution of CRISPR-cas systems. *Biochemical Society Transactions* 41: 1392–1400.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, *et al.* 2013. RNA-guided human genome engineering via Cas9. *Science* 339(6121): 823–826.

Mann NH, Cook A, Millard A, Bailey S, Clokie M. 2003. Marine ecosystems: bacterial photosynthesis genes in a virus. *Nature* 424(6950): 741.

Marraffini LA, Sontheimer EJ. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322(5909): 1843–1845.

Medina-Aparicio L, Rebollar-Flores JE, Gallego-Hernandez AL, Vazquez A, Olvera L, *et al.* 2011. The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *Journal of Bacteriology* 193(10): 2396–2407.

Middelboe M, Hagström A, Blackburn N, Sinn B, Fischer U, *et al.* 2001. Effects of bacteriophages on the population dynamics of four strains of pelagic marine bacteria. *Microbial Ecology* 42(3): 395–406.

Mohapatra NP, Soni S, Bell BL, Warren R, Ernst RK, Muszynski A, Carlson RW, Gunn JS. 2007. Identification of an orphan response regulator required for the virulence of *Francisella* spp. and transcription of pathogenicity island genes. *Infection and Immunity* 75(7): 3305–3314.

Mojica FJ, Díez-Villaseñor C, Garcia-Martinez J, Almendros C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defense system. *Microbiology* 155(3): 733–740.

Mojica FJ, Díez-Villaseñor C, Garcia-Martinez J, Soria E. 2005. Intervening sequences of regulatory spaced prokaryotic repeats derive from foreign genetic elements, *Journal of Molecular Evolution* 60(2): 174–182.

Mojica KD, Brussaard CP. 2014. Factors affecting virus dynamics and microbial host–virus interactions in marine environments. *FEMS Microbiology Ecology* 89(3): 495–515.

Navarre WW, McClelland M, Libby SJ, Fang FC. 2007. Silencing of xenogeneic DNA by H-NS—facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes and Development* 21(12): 1456–1471.

Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC. 2006. Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella. Science* 313(5784): 236–238.

Nuñez JK, Kranzusch PJ, Noeske J, Wright AV, Davies CW, Doudna JA. 2014. Cas1–Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity. *Nature Structural and Molecular Biology* 21(6): 528–534.

Pavankumar TL, Sinha AK, Ray MK. 2010. All three subunits of RecBCD enzyme are essential for DNA repair and low-temperature growth in the Antarctic *Pseudomonas syringae* Lz4W. *PloS One* 5(2): e9412.

Pedersen K, Christensen SK, Gerdes K. 2002. Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Molecular Microbiology* 45(2): 501–510.

Pérez-Rodríguez R, Fisher AC, Perlmutter JD, Hicks MG, Chanal A, *et al.* 2007. An essential role for the DnaK molecular chaperone in stabilizing over-expressed substrate proteins of the bacterial twin-arginine translocation pathway. *Journal of Molecular Biology* 367(3): 715–730.

Pérez-Rodríguez R, Haitjema C, Huang Q, Nam KH, Bernardis S, Ke A, DeLisa MP. 2011. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli. Molecular Microbiology* 79(3): 584–599.

Pourcel C, Drevet C. 2013. Occurrence, diversity of CRISPR-Cas systems and genotyping implications. In *CRISPR-Cas systems* (ed. Barrangou R, Van Der Oost J). Berlin: Springer, pp. 33–60.

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(3): 653–663.

Proctor LM, Fuhrman JA. 1990. Viral mortality of marine bacteria and cyanobacteria. *Nature* 343: 60–62.

Pul Ü, 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. *Molecular Microbiology* 75(6): 1495–1512.

Raivio TL. 2005. MicroReview: envelope stress responses and Gram-negative bacterial pathogenesis. *Molecular Microbiology* 56(5): 1119–1128.

Raivio T. 2011. Identifying your enemies – could envelope stress trigger microbial immunity? *Molecular Microbiology* 79(3): 557–561.

Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, *et al.* 2015. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*, 520(7546): 186–191.

Regha K, Satapathy AK, Ray MK. 2005. RecD plays an essential function during growth at low temperature in the Antarctic bacterium *Pseudomonas syringae* Lz4W. *Genetics* 170(4): 1473–1484.

Reyes A, Semenkovich NP, Whiteson K, Rohwer F, Gordon JI. 2012. Going viral: next-generation sequencing applied to phage populations in the human gut. *Nature Reviews Microbiology* 10(9): 607–617.

Reyes J, Francis AR, Tanaka MM. 2008. Models of deletion for visualizing bacterial variation: an application to tuberculosis spoligotypes. *BMC Bioinformatics* 9: 496–511.

Rubin MA, Leff LG. 2007. Nutrients and other abiotic factors affecting bacterial communities in an Ohio River (USA). *Microbial Ecology* 54(2): 374–383.

Sampson TR, Napier BA, Schroeder MR, Louwen R, Zhao J, *et al.* 2014. A CRISPR-Cas system enhances envelope integrity mediating antibiotic resistance and inflammasome evasion. *Proceedings of the National Academy of Sciences USA* 111(30): 11163–11168.

Schrenk MO, Huber JA, Edwards KJ. 2010. Microbial provinces in the subseafloor. *Annual Review of Marine Science* 2: 279–304.

Shah SA, Erdmann S, Mojica FJ, Garrett RA. 2013. Protospacer recognition motifs. *RNA Biology* 10(5): 891–899.

Shimkets LJ, Kaiser D. 1982. Induction of coordinated movement of *Myxococcus xanthus* cells. *Journal of Bacteriology* 152(1): 451–461.

Sinha RP, Häder DP. 2002. UV-induced DNA damage and repair: a review. *Photochemical and Photobiological Sciences* 1(4): 225–236.

Sorek R, Kunin V, Hugenholtz P. 2008. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews Microbiology* 6(3): 181–186.

Sorek R, Lawrence CM, Wiedenheft B. 2013. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annual Review of Biochemistry* 82: 237.

Sozinova O, Jiang Y, Kaiser D, Alber M. 2005. A three-dimensional model of myxobacterial aggregation by contact-mediated interactions. *Proceedings of the National Academy of Sciences of the United States of America* 102(32): 11308–11312.

Spano G, Massa S. 2006. Environmental stress response in wine lactic acid bacteria: Beyond *Bacillus subtilis. Critical Reviews in Microbiology* 32(2): 77–86.

Staals RH, Agari Y, Maki-Yonekura S, Zhu Y, Taylor DW, *et al.* 2013. Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of *Thermus thermophilus. Molecular Cell* 52(1): 135–145.

Stan-Lotter H, Fendrihan S. 2012. *Adaption of microbial life to environmental extremes*. New York: Springer.

Swarts DC, Mosterd C, Van Passel MW, Brouns SJ. 2012. CRISPR interference directs strand specific spacer acquisition. *PloS One* 7(4): e35888.

Tilly K, McKittrick N, Zylicz M, Georgopoulos C. 1983. The dnaK protein modulates the heat-shock response of *Escherichia coli*. *Cell* 34(2): 641–646.

Tock MR, Dryden DT. 2005. The biology of restriction and anti-restriction. *Current Opinion in Microbiology* 8: 466–472.

Torsvik V, Øvreås L. 2002. Microbial diversity and function in soil: from genes to ecosystems. *Current Opinion in Microbiology* 5(3): 240–245.

Touchon M, Charpentier S, Clermont O, Rocha EP, Erick Denamur E, Branger C. 2011. CRISPR Distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *Journal of Bacteriology* 193: 2460–2467.

Tyson GW, Banfield JF. 2008. Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environmental Microbiology* 10(1): 200–207.

Van Der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJ. 2009. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends in Biochemical Sciences* 34(8): 401–407.

Van der Zanden AG, Kremer K, Schouls LM, Caimi K, Cataldi A, *et al.* 2002. Improvement of differentiation and interpretability of spoligotyping for *Mycobacterium tuberculosis* complex isolates by introduction of new spacer oligonucleotides. *Journal of Clinical Microbiology* 40(12): 4628–4639.

Van Melderen L, De Bast MS. 2009. Bacterial toxin–antitoxin systems: more than selfish entities? *PLoS Genetics* 5(3): e1000437.

Viswanathan P, Murphy K, Julien B, Garza AG, Kroos L. 2007a. Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *Journal of Bacteriology* 189(10): 3738–3750.

Viswanathan P, Ueki T, Inouye S, Kroos L. 2007b. Combinatorial regulation of genes essential for *Myxococcus xanthus* development

involves a response regulator and a LysR-type regulator. *Proceedings of National Academy of Sciences of the United States of America* 104(19): 7969–7974.

Wang X, Ribeiro AA, Guan Z, Abraham SN, Raetz CR. 2006. Attenuated virulence of a *Francisella* mutant lacking the lipid A 4′-phosphatase. *Proceedings of National Academy of Sciences of the United States of America* 104(10): 4136–4141.

West SC. 1996. The RuvABC proteins and Holliday junction processing in *Escherichia coli*. *Journal of Bacteriology* 178(5): 1237.

Westra ER, Brouns SJ. 2012. The rise and fall of CRISPRs–dynamics of spacer acquisition and loss. *Molecular Microbiology* 85(6): 1021–1025.

Westra ER, Buckling A, Fineran PC. 2014. CRISPR-Cas systems: beyond adaptive immunity. *Nature Reviews Microbiology* 12(5): 317–326.

Westra ER, Pul Ü, 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. *Molecular Microbiology* 77(6): 1380–1393.

Weinberger AD, Sun CL, Pluciński MM, Denef VJ, Thomas BC, *et al.* 2012. Persisting viral sequences shape microbial CRISPR-based immunity. *PLoS Computational Biology* 8(4): e1002475.

Williamson KE, Corzo KA, Drissi CL, Buckingham JM, Thompson CP, Helton RR. 2013. Estimates of viral abundance in soils are strongly influenced by extraction and enumeration methods. *Biology and Fertility of Soils* 49(7): 857–869.

Yamaguchi Y, Park JH, Inouye M. 2011. Toxin-antitoxin systems in bacteria and archaea. *Annual Review of Genetics* 45: 61–79.

Yosef I, Goren MG, Qimron U. 2012. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Research* 40(12): 5569–5576.

Zhang J, Kasciukovic T, White MF. 2012. The CRISPR associated protein Cas4 Is a 5′ to 3′ DNA exonuclease with an iron-sulfur cluster. *PloS One* 7(10): e47232.