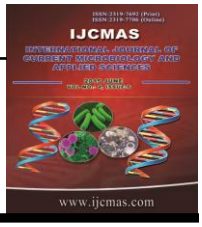
Int.J.Curr.Microbiol.App.Sci (2015) 4(6): 187-200

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 4 Number 6 (2015) pp. 187-200**
http://www.ijcmas.com



Review Article

CRISPR Systems: RNA-Guided Defence Mechanisms in Bacteria and Archaea

Manish Dhawan[1]*, Manish Sharma[2] and Rajveer Singh Grewal[3]

[1]Department of Microbiology, Punjab Agricultural University, Ludhiana, India
[2]Department of Microbiology, Guru Nanak Dev University, Amritsar, India
[3]School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, India
**Corresponding author*

Keywords

Cas proteins, CRISPR, crRNAs, Repeats and Spacers

A B S T R A C T

Bioinformatics, genetics and biochemical studies have revealed that many prokaryotes use an RNA-based adaptive immune system to target and destroy genetic parasites. Prokaryotic Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) systems provide a sophisticated adaptive immune system that offers protection against foreign nucleic acid. CRISPR-Cas system consists of CRISPR loci contains array of identical repeats interspaced with spacers carrying unique sequences and CRISPR-associated (Cas) proteins exhibit helicases, nucleases, polymerases and RNA-binding functions. In CRISPR-based immune systems, short sequence tags from invading genetic elements (Protospacers) are actively incorporated into the host's CRISPR locus to be transcribed and processed into library of short CRISPR-derived RNAs (crRNAs). These crRNAs packaged into a large surveillance complex which guides destruction of foreign genetic material. CRISPR-Cas systems have been grouped into three types (Types I–III), along with subtypes (e.g., Type I-E), on the basis of the Cas genes they possess and their mode of action. CRISPR system proves useful in genome editing, development of industrially important phage resistant strains, reduction of antibiotic resistance (horizontal gene flow). Discovery of fundamental mechanisms of CRISPR systems raised many new questions such as how and when target sequences are identified during a phage infection.

Introduction

Viruses are responsible for 4 to 50% destruction of bacterial and archaeal communities (Breitbart and Rohwer, 2005; Rohwer and Thurber, 2009) and the selective pressures imposed by these rapidly evolving viruses (parasites) has driven the diversification of microbial defence systems. Previously our understanding of antiviral immune systems in bacteria has focused on encounter blocks, adsorption resistance, penetration blocks, abortive infection, restriction modification, phage growth limitation systems (Hyman and Abedon, 2010; Labrie *et al.*, 2010). Now, our understanding has been expanded to include the CRISPR systems of Clustered Regularly Interspaced Short Palindromic Repeats. These CRISPR systems are essential

components of nucleic-acid based adaptive immune systems that are widespread in bacteria and archaea. Prokaryotic CRISPR systems are highly adaptive and heritable resistance mechanisms that incorporate sequences derived from the foreign element into a small-RNA-based repertoire that provides protection against foreign DNA. In response to invading nucleic acid, bacteria and archaea integrate short fragments of invaded nucleic acid into the host chromosome at one end of a repetitive element known as CRISPR loci. These sequences further transcribe and precisely process into small RNAs which guide a multifunctional protein complex of Cas proteins. This complex further recognizes and cleaves incoming foreign genetic material. These repetitive loci maintaining a genetic record of prior encounters with foreign DNA or RNA hence serve as molecular 'vaccination cards' (Bondy-Denomy and Davidson, 2014).

Brief history

CRISPRs were originally identified in the *Escherichia coli* genome in 1987, when they were described as an unusual sequence elements consisting of a series of 29-nucleotide repeats separated by unique 32-nucleotide 'spacer' sequences (Ishino *et al.*, 1987). Repetitive sequences with a similar repeat–spacer–repeat pattern were later identified in phylogenetically diverse bacterial and archaeal genomes but the function of these repeats remained unknown. Jansen *et al.* (2002) coined the term CRISPR and reported that CRISPRs co-localize with specific Cas (CRISPR-associated proteins) genes. Pourcel *et al.* (2005) reported that many spacer sequences were identical to viral and plasmid sequences. These observations suggested that the spacers were derived from invading genomes and indicated the possible role of CRISPR in microbial immunity. Makarova *et al.* (2006) led to the hypothesis that CRISPRs provide a genetic memory of infection and detection of short CRISPR-derived RNA transcripts suggested that there may be functional similarities between CRISPR-based immunity and RNA-interference. Barrangou *et al.* (2007) provided the first experimental evidence that CRISPR-Cas system functions as an inheritable and adaptive immune system in prokaryotes. Marraffini *et al.* (2008) also reported the ability of CRISPRs to prevent plasmid transfer in staphylococcus

Novel features of CRISPR-Cas system

A functional CRISPR-Cas system has two distinguishable components required for activity: CRISPR locus/arrayand Cas genes. CRISPR locus/array located on the genome (either chromosome or plasmid) and is a series of direct repeats (approximately 20–50 base pairs) separated by unique hypervariable spacers sequences of a similar length acquired from virus or plasmid DNA (Marraffini *et al.*, 2010) (Figure 1). The repeat sequences within a CRISPR locus are conserved, but repeats in different CRISPR loci can vary in both sequence and length. In addition, the number of repeat–spacer units in a CRISPR locus varies widely within and among organisms (Kunin *et al.*, 2007). Repeats are partially palindromic in nature and transcripts from these regions may form stable and highly conserved RNA secondary structures (Marrafini *et al.*, 2008). Majority of genomes contain a single CRISPR-Cas locus but some bacterial genomes contain as many as 13–15 CRISPR loci. Many archaeal genomes harbor numerous CRISPR loci (Grissa *et al.*, 2007).

Bioinformatics analyses have shown that CRISPR loci are flanked by a large number of extremely diverse *cas* genes (Haft *et al.*,

2005). Diverse group of *cas* genes located in the vicinity of CRISPR locus and encode proteins (generally called Cas proteins) which are required for the multi step defense against invasive genetic elements. Initial homology comparisons by Jansen and colleagues delineated four core CRISPR-associated gene families, *cas*1–4 (Jansen *et al.,* 2002). These gene families were further extended to include *cas5* and *cas6* (Bolotin *et al.,* 2005). Haft *et al.* (2005) reported eight subtypes of Cas proteins based on the phylogenetics of highly conserved Cas1 protein and the operonic organization of *cas* genes. The *cas1* gene is a common component of all CRISPR systems. Phylogenetic analyses of *cas*1 sequences suggested several versions of the CRISPR systems (Kunin *et al.,* 2007). Additional evidences for the classification of distinct CRISPR types such as neighborhood analysis of CRISPR loci has identified conserved arrangements of four to ten *cas*genes which can be found in association with CRISPR loci harbouring specific repeat sequences (Kunin *et al.,* 2007). *Cas1* to *Cas10* genes have been identified till now (Table 1). Makarova *et al.* (2011) divided these distinct immune systems into three major CRISPR types on the basis of gene conservation and locus organization. Based on phylogeny, sequence and locus organization there are three types (Type I, Type II, and Type III) of CRISPR-Cas systems (Figure 2).

Based on computational analyses Cas proteins were predicted to contain identifiable domains characteristic of helicases, nucleases, polymerases, and RNA-binding proteins. Cas proteins led to the initial speculation that they may be part of a novel DNA repair system (Makarova *et al.,* 2002). Sequence analysis of genomes containing multiple CRISPR loci uncovered an additional structural feature directly adjacent to the short repeats (Figure 1B) (Jansen *et al.,* 2002). This region of conservation between CRISPR loci termed as leader sequence and extends several hundred base pairs. This region lacks coding potential and generally found on one side of the CRISPR in a fixed orientation. Generally leaders are 80% identical within a genome but dissimilar among the species.

Mechanism of CRISPR-Mediated Defense

The mechanism of CRISPR-Cas system (RNA-guided Defense) divided into three phases which are CRISPR spacer acquisition, CRISPR locus expression (transcription and processing), and CRISPR interference. In following sections we will discuss detailed mechanism of CRISPR-Cas system.

CRISPR spacer acquisition

Escherichia coli (Type I), *Streptococcus thermophilus* (Type II) and *Pyrococcus furiosus* (Type III) are the best model organisms to study the mechanism of CRISPR acquisition. To acquire resistance, new spacer information must be incorporated into the CRISPR locus. Barrangou *et al.* (2007) isolated bacteriophage insensitive mutants (BIMs) of *S. thermophiles* and reported that all of the resistant mutants had gained from one to four novel spacers with sequence identity to the invading genome. They also reported that the degree of resistance in the mutants can be correlated with the number of acquired sequences. Evolutionary analyses predicted that the insertions occurred proximal to the leader. The presence of the leader sequence next to the CRISPR locus proved essential for the acquisition of the spacers (Karginov and Hannon, 2010). Various studies on the Cas-proteins suggested that spacer acquisition is driven

by recognition of phage sequences by subtype-specific proteins in different species (Mojica *et al.,* 2009). The process of spacer acquisition can be divided into three distinct steps involving (a) recognition of the invasive nucleic acid and scanning foreign DNA for potential Protospacer Associated Motifs (PAMs) that identify protospacers (b) the generation of a new repeat spacer by processing of the nucleic acid and (c) the integration of the new CRISPR repeat spacer unit at the leader end of the CRISPR locus (Figure 3). Presently, the precise mechanisms by which information is transferred from phage or plasmids into CRISPR loci are obscure.

CRISPR locus expression

CRISPR locus transcribes into primary transcript or pre-crRNA. This process has been examined in *Escherichia coli*, *Xanthomonasoryzae*, *Thermus thermophilus*, *Pyrococcus furiosus* and *Sulfolobus*. In the effector phase transcription of the CRISPR repeats initiates in or near the leader sequence and generates a long pre-crRNA precursor that can span the entire locus (Lillestol *et al.,* 2006; 2009). The pre-crRNA is then endonucleolytically processed into fragments which are corresponding to the interval between repeats. This produces mature products and a laddering pattern of intermediates (Brouns *et al.,* 2008).

In *S. acidocaldarius* CRISPR derived small RNAs appear as products from 35 to 52 nt, presumably generated by endonucleolytic cleavage of long precursors. In Type I, II and III systems Cas6e/Cas6f, Cas9 and Cas6 acts as endonuclease respectively to make mature crRNA from pre-crRNA. The structures of *T. thermophilus* cse3 and *Pyrococcus furiosus* cas6 explain their common endonucleolytic function. These proteins act as endonuclease which produces the small prokaryotic silencing RNAs (Figure 4). Detailed analyses in *S. thermophilus* (Bolotin *et al.,* 2005), in bacteria and archaea (Makarova *et al.,* 2006; Shah *et al.,* 2009) showed that spacers encode crRNAs corresponding to both the coding and template strands of the phage. Similar conclusions can be reached by examination of spacers arising in experimentally induced phage-resistant mutants of *S. thermophilus* (Barrangou *et al.,* 2007).

CRISPR interference

Third stage of CRISPR-mediated immunity is target interference. In this stage crRNAs associate with Cas proteins to form large CRISPR-associated ribonucleoprotein complexes that can recognize invading nucleic acids. Foreign nucleic acids are identified by base-pairing interactions between the crRNA spacer sequence and a complementary sequence from the intruder. RNA-guided DNA-targeting system is indeed a pathway for DNA silencing has recently been demonstrated in *S. thermophilus*. DNA sequencing and Southern blots indicated that both strands of the target DNA are cleaved within the region that is complementary to the crRNA spacer sequence. This mechanism efficiently eliminates foreign DNA sequences which have been specified by the spacer region of the crRNA. This avoids targeting the complementary DNA sequences in the CRISPR region of the host chromosome. The mechanism for distinguishing self from non-self is built into the crRNA. The spacer sequence of each crRNA is flanked by a portion of the adjacent CRISPR repeat sequence. Any complementarity beyond the spacer into the adjacent repeat region signals self which prevents the destruction of the host chromosome (Garneau *et al.,* 2010). In

Type II system crRNA forms hybrid with tracrRNA (trans activating RNA). The crRNA: tracrRNA hybrid further triggers the Cas9 protein for making double stranded break (Figure 5). All CRISPR systems do not target DNA. In vitro experiments using enzymes from the Type III-B CRISPR system of *P. furiosus* have shown that this system cleaves target RNA rather than DNA. All DNA targeting systems encode a complementary DNA sequence for each crRNA in the CRISPR locus and therefore require a mechanism for distinguishing self (CRISPR locus) from non-self (invading DNA). CRISPR systems that target RNA may be uniquely capable of defending against viruses that have RNA-based genomes. Adaptation of the CRISPR in response to a challenge by an RNA-based virus will probably require the invading RNA to be reverse-transcribed into DNA before it can be integrated into the CRISPR locus. There are diverse mechanisms of CRISPR RNA biogenesis for the interference. On the basis of these mechanisms there are three major CRISPR-Cas systems (Type I, II, and III system) (Wiedenheft *et al.*, 2012).

Type I CRISPR-Cas system

A large multidomain Cas3 protein with distinct DNA nuclease and helicase activities is the signature protein in Type I system (Sinkunas *et al.*, 2011). There are also multiple Cas proteins that form CASCADE-like complexes which are involved in the interference step. Many proteins such as Cas5, Cas6 and Cas7 are in distinct RAMP (Repeat-associated mysterious proteins) super-families. Among the three systems, Type I is the most diverse with six different subtypes (Type I-A through Type II-F) (Makarova *et al.*, 2011). Recent experiments on *E. coli* and *Pseudomonas aeruginosa* have suggested the mechanism of Type I CRISPR-Cas system (Haurwitz *et al.*, 2010; Wiedenheft *et al.*, 2011). For the DNA interference, CASCADE associates with processed crRNA to form a ribonucleoprotein complex that drives the formation of R-loops in invasive double-stranded DNA (dsDNA) through seed sequence-driven base pairing (Figure 5). CASCADE target dsDNA by sequence-specific hybridization between crRNA and the target DNA over a 7-8 bp sequence (the seed sequence) at the 5' end of the spacer (Figure 5) (Wiedenheft *et al.*, 2011). In Type I CRISPR systems, Cas6e and Cas6f recognize the major groove of the crRNA stem-loop primarily through electrostatic interactions using a β-hairpin and α-helix, respectively. Cleavage occurs at the double-stranded-single stranded junction, leaving 8-nt 5'-handle on mature crRNAs (Figure 6) (Wiedenheft *et al.*, 2012).

Type II CRISPR-Cas System

A large multifunctional protein Cas9 is the signature protein in Type II CRISPR-Cas system. Cas9 protein has ability to generate crRNA, target phage and plasmid DNA for degradation (Garneau *et al.*, 2010). In Type II CRISPR systems tracrRNA hybridizes to the pre-crRNA repeat to form duplex RNAs which are substrates for endonucleolytic cleavage by host RNase III. This activity may also require Cas9. Subsequent trimming by an unidentified nuclease removes repeat sequences from the 5' end (Figure 6) (Wiedenheft *et al.*, 2012). Bhaya *et al.* (2011) reported that Cas9 protein may contain two nuclease domains one at the N-terminus (RuvC-like nuclease) and an HNH (McrA-like) nuclease domain in the middle section. They further reported that HNH nuclease domain might be involved in target cleavage based on its endonuclease activity. Type II is the simplest of the three CRISPR-Cas types with only four genes that compose

the operon (this includes *cas9*, *cas1*, *cas2*, and either *cas4* or *csn2*).

Type III CRISPR-Cas System

This system has a number of recognizable features including RAMP as the signature protein Cas 10 which is involved in the processing of crRNA. This protein targets cleavage of DNA which is functionally analogous to Type I CASCADE. Cas6 may also acts as signature protein in Type III system which is involved in crRNA processing. Type III systems distinguished further in two types (Type IIIA and IIIB) on the basis of target nucleic acid (DNA or RNA) for the interference step. In *Pyrococcus furiosus*, a Type IIIA system targets mRNA for CRISPR interference (Hale *et al.*, 2009). Whereas Type IIIB system found in *Staphylococcus epidermidis* targets DNA for interference (Marraffini and Sontheimer, 2008). Type IIIB CRISPR systems recognize single-stranded RNA, upstream of the scissile phosphate, on a face of the protein opposite that of the previously identified active site residues (Carte *et al.*, 2008; Wang *et al.*, 2011). The remainder of the repeat substrate probably wraps around the protein to allow cleavage 8 nucleotides upstream of the repeat-spacer junction. Further 3' trimming produces mature crRNA of two discrete lengths (Figure 6). The N-terminal domain of all Cas6 family proteins adopts a ferredoxin-like fold. Moreover, C-terminal domain of Cas6 and cas6e also adopts a ferredoxin-like fold but the C-terminal domain of Cas6f is structurally distinct (wiedenheft *et al.*, 2009, 2012).The distribution of CRISPR-Cas systems is a notable feature. It is because Type I systems found in both bacteria and archaea. Type II systems are exclusively present in bacteria. Type III systems usually found in archaea, although it is also found in bacteria (Makarova *et al.*, 2011; Terns and Terns 2011).

Applications of CRISPR systems

The sequence diversity of CRISPR loci, even within closely related strains, has been used for high-resolution genotyping. Spoligotyping (spacer oligotyping), has been used successfully for the analysis of human pathogens, including *Mycobacterium tuberculosis* (Groenen *et al.*, 1993), *Corynebacterium diphtheria* (Mokrousov *et al.*, 2007) and *Salmonella enterica* (Liu *et al.*, 2011). CRISPR mediated genetic silencing provides new opportunities for making the spoligotyping more creative and applicable.

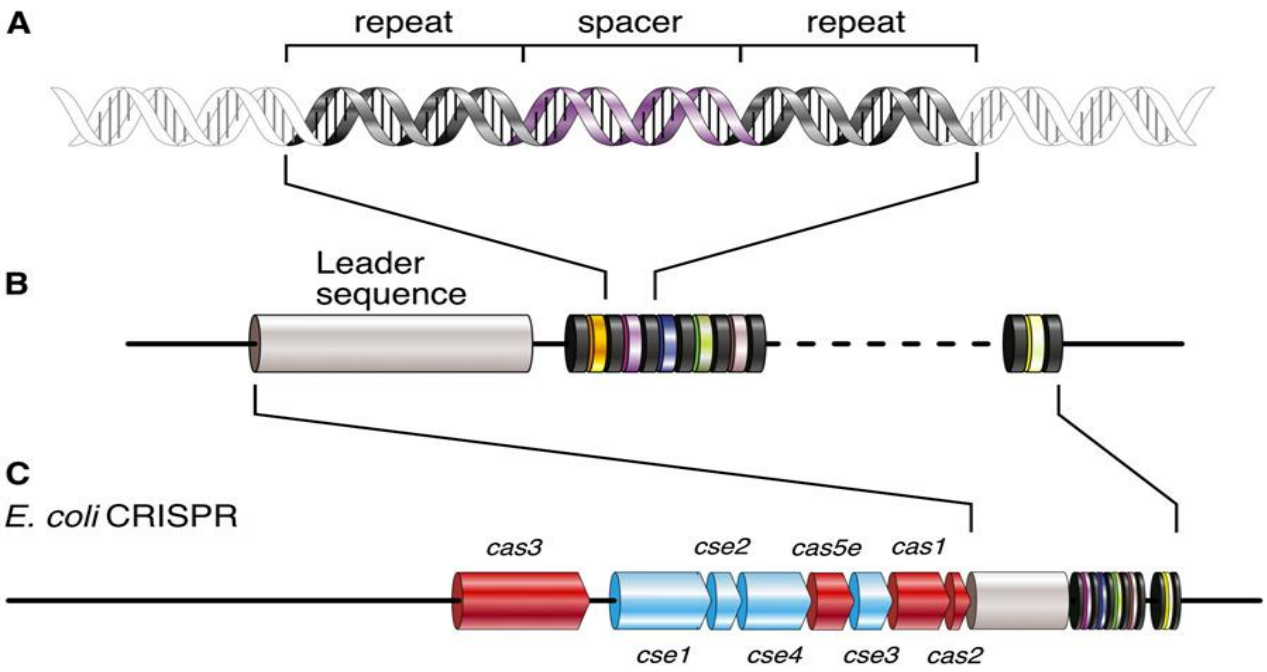
Laboratory strains of bacteria are grown in high-density bioreactors for many different applications in the food and pharmaceutical industry. These strains are becoming increasingly important in the production of biofuels. Some industrially important bacteria found as susceptible to phages. CRISPR systems offer a natural mechanism for adapting economically important bacteria for resistance against multiple phages (Barrangou *et al.*, 2007).The biochemical activities of various Cas proteins may have useful applications in molecular biology. These enzymes can be used in the same way that DNA restriction enzymes have revolutionized cloning and DNA manipulation. A wide range of CRISPR-specific endoribonucleases that recognize small RNA motifs with high affinity expand the number of tools available for manipulating nucleic acids. AcrRNA-guided ribonucleoprotein complex in *P. furiosus* cleaves target RNAs (Hale *et al.*, 2009). So, Site-specific cleavage of target RNA molecules could have a range of uses, from generating homogeneous termini after *in vitro* transcription to targeting a specific intracellular messenger RNA for inactivation in similar way to RNAi.

Table.1 Major Cas proteins and their role in different CRISPR systems

Protein	Distribution	COG	Process	Function
Cas1	Universal	COG1518	Spacer acquisition	DNAse, not sequence specific, can bind RNA; present in all Types; structure available for several Cas 1 proteins
Cas2	Universal	COG1343, COG3512	Spacer acquisition	Small RNAse specific to U-rich regions; present in all Types; structure available from *Thermus thermophilus* and *Sulfolobus solfataricus*and others
Cas3	**Type I signature**	COG1203, COG2254	Target interference	DNA helicase; most proteins have a fusion to HD nuclease
Cas4	Type I,II	COG1468	Spacer acquisition	RecB-like nuclease with exonuclease activity homologous to RecB
Cas5	Type I	COG1688,RAMP	crRNA expression	RAMP protein, endoribonuclease involved in crRNA biogenesis; part of CASCADE
Cas6	Type I, III	COG1583, COG5551,RAMP	crRNA expression	RAMP protein, endoribonuclease involved in crRNA biogenesis; part of CASCADE; structure available from *P. furiosus*
Cas7	Type I	COG1857,COG3649,RAMP	crRNA expression	RAMP protein, endoribonuclease involved in crRNA biogenesis; part of CASCADE
Cas8	Type I	Not determined	crRNA expression	Large protein with McrA/HNH-nuclease domain and RuvC-like nuclease; part of CASCADE
Cas9	Type II signature	COG3513	Target interference	Large multidomain protein with McrA-HNH nuclease domain and RuvC-like nuclease domain; necessary for interference and target cleavage
Cas10	Type III signature	COG1353	crRNA expression and interfernece	HD nuclease domain, palm domain, Zn ribbon; some homologies with CASCADE elements

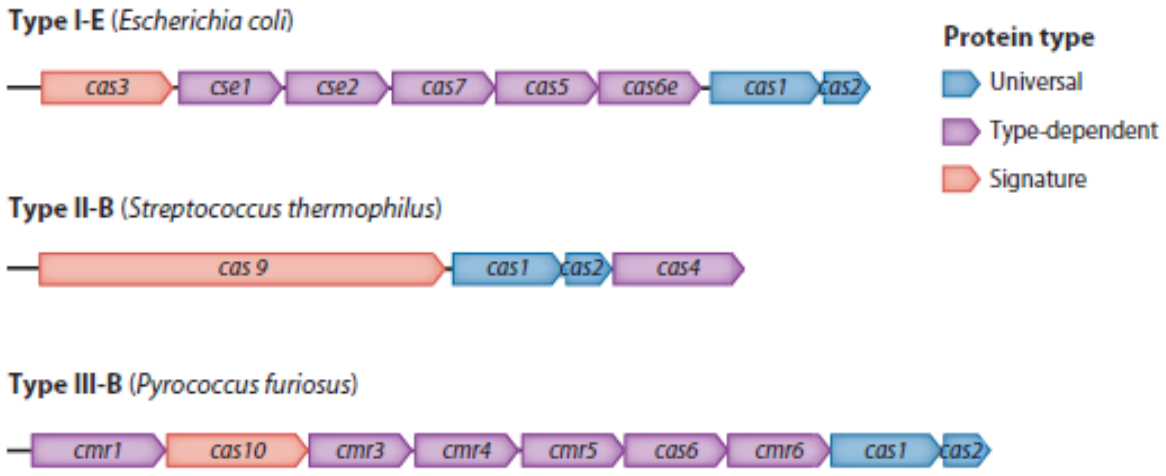
Source: Bhaya *et al.*, 2011

Figure.1 Components of a typical CRISPR-Cas system



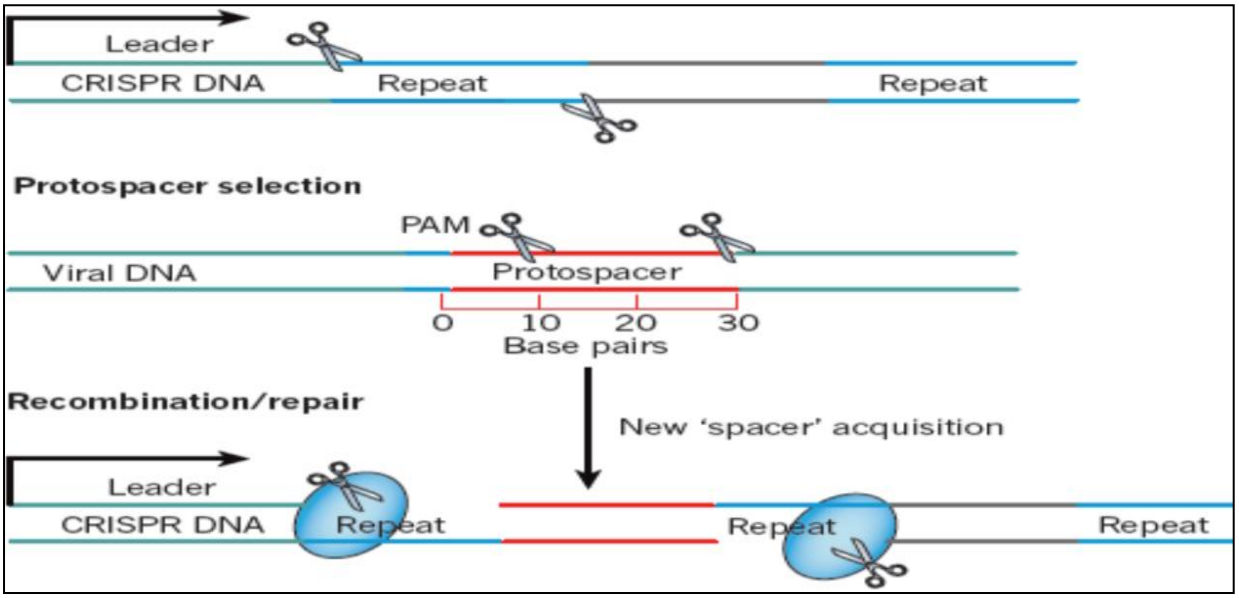
Source: Karginov and Hannon, 2010

Figure.2 Genetic constitution of various CRISPR-Cas systems



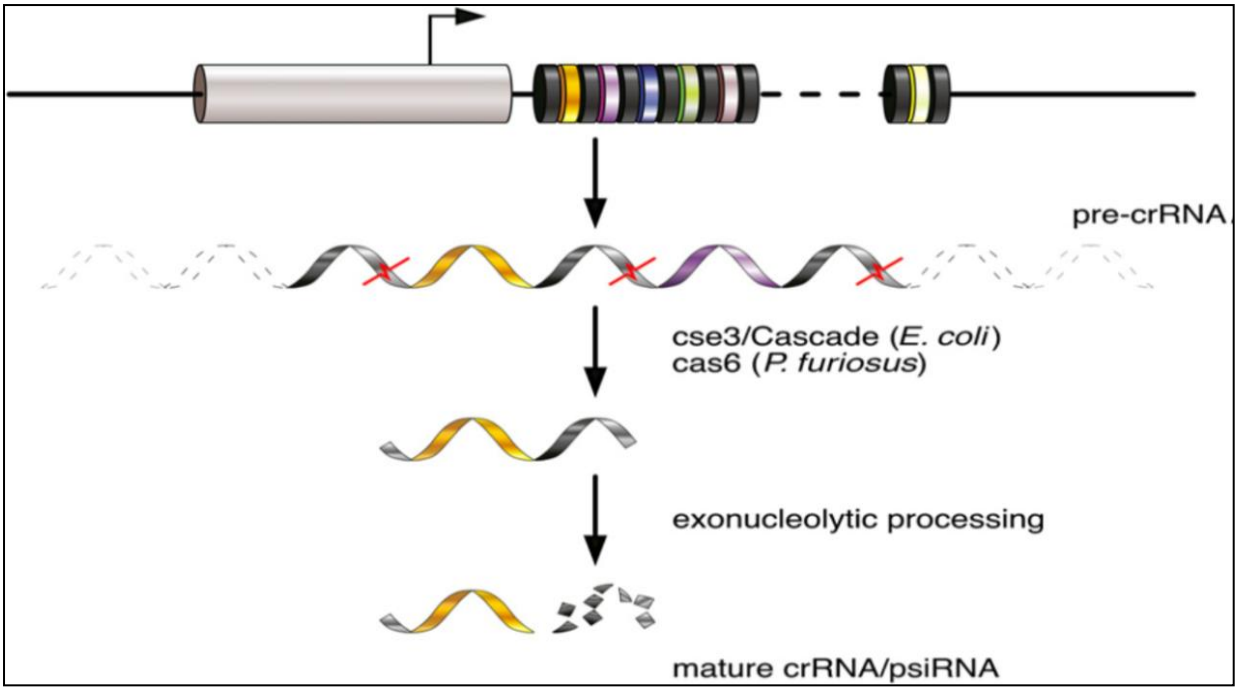
Source: Bhaya *et al.*, 2011

Figure.3 First step (Spacer Acquisition) of CRISPR mediated immune response



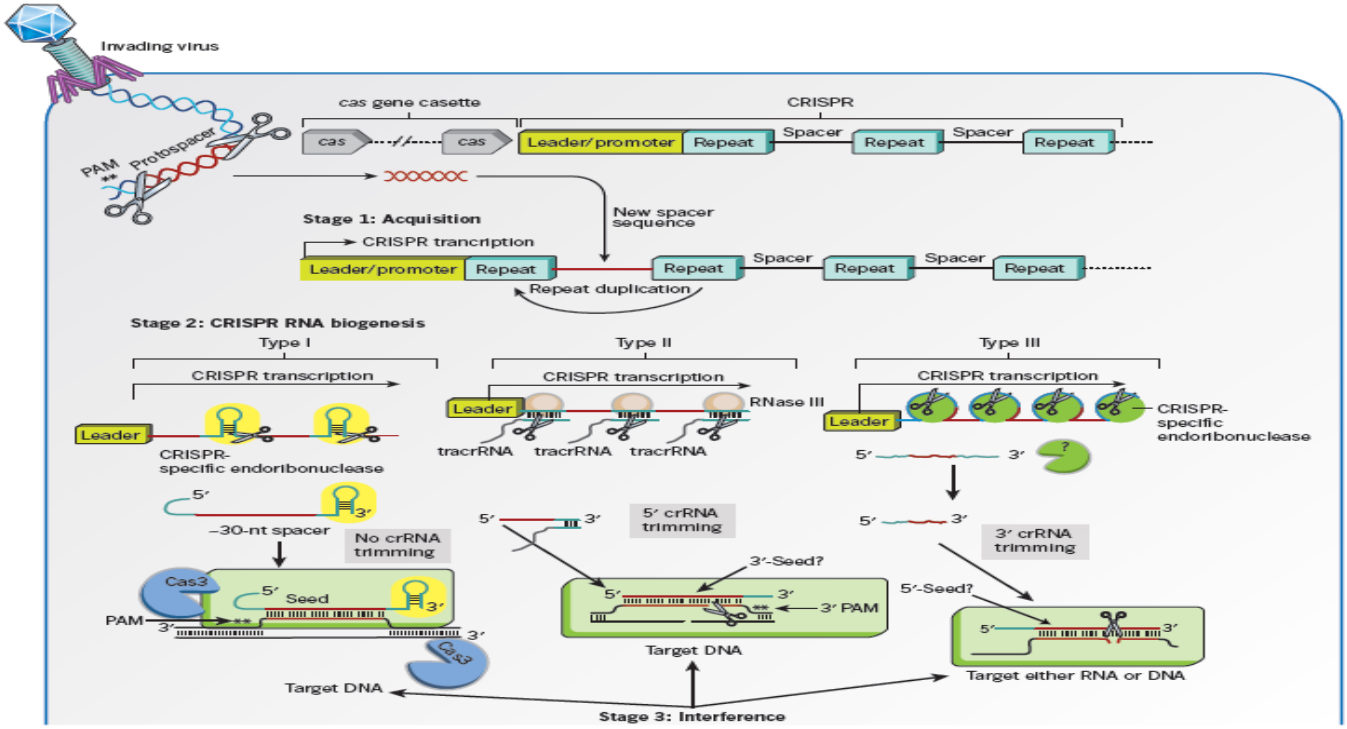
Source: Wiedenheft *et al.*, 2012

Figure.4 Second step (Locus Expression) of CRISPR mediated immune response



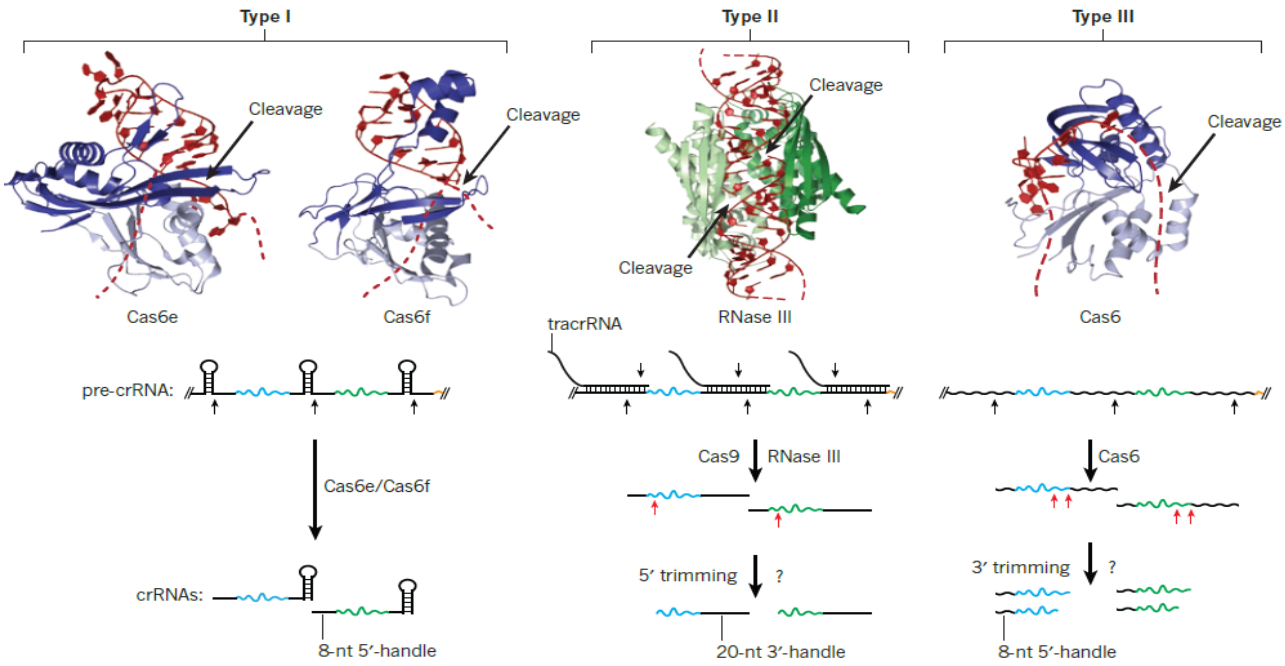
Source: Karginov and Hannon, 2010

Figure.5 CRISPR Interference and the overview of CRISPR mediated immune response



Source: Wiedenheft *et al.*, 2012

Figure.6 Difference in mechanisms of three different types of CRISPR-Cas systems



Source: Wiedenheft *et al.*, 2011

CRISPRs also provide a new mechanism for limiting the spread of antibiotic resistance or the transfer of virulence factors by blocking horizontal gene transfer (Garneau *et al.*, 2010; Marraffini and Sontheimer, 2008). CRISPRs playa regulatory mechanism that alters biofilm formation in *P. aeruginosa* (Cady *et al.*, 2011; Zegans *et al.*, 2009). The clinical relevance of CRISPRs has to be discovered and this new gene-regulation system provides new opportunities for creative implementation in wide areas.

Limitations

Major limitation of CRISPR-Cas system is the significant off-target effects. Non-target DNA resembling the guide RNA can become cut, activated, or deactivated. Delivery of the CRISPR system in host is an enormous challenge so, delivery of the Guide RNAs should be specific and these should target the specific DNA sequences. There must be development in these systems for more specificity. A number of Cas proteins have been biochemically or structurally characterized but still lack a functional assignment within the CRISPR mechanism.

Future prospective of CRISPR biology

New questions have been raised by the discovery of fundamental mechanisms of CRISPR-based adaptive immunity. These highlighted the areas with the greatest potential for future research such as how and when target sequences are identified during a phage infection. Moreover, plasmid transformations are still unclear. Further the mechanisms such as why DNA or RNA target sequences are chosen, and their fates when they are bound to a crRNA-targeting complex are not fully understood (Wiedenheft *et al.*, 2012). Further, the mechanisms have to be fully understood by which foreign sequences are selected and integrated into CRISPR loci. The prevalence and diversity of CRISPR-Cas systems throughout bacterial and archaeal communities ensures that new findings and applications in this field will be forthcoming in the years ahead.

Conclusion

The CRISPR system is an effective prokaryotic adaptive immune system against foreign genetic elements. It is rightly described as an adaptive immune system. CRISPR–Cas systems provide powerful means for bacteria and archaea to destroy potentially harmful foreign DNA. Common occurrence of these systems within bacterial genomes suggests the positive influence that these systems must have on evolutionary fitness. CRISPR-Cas system seems to fit more firmly with a Lamarckian paradigm, in essence because increases in fitness do not rely on random mutations but on a much more specific acquisition of genetic information from environmental sources (Karginov and Hannon, 2010). It is difficult to predict net biological outcome of a CRISPR–Cas system within a given organism in a given environment. In addition, some systems may be inactive (i.e., suppressed, repressed, or defective) or be performing alternative roles. Future studies must focus on in vivo characterization of more systems operating in diverse species for the improvement of understanding about the complex biological outcomes of CRISPR–Cas systems. The recent discovery of anti-CRISPR genes in *P. aeruginosa* suggests that many such systems may exist. The identification and characterization of more CRISPR–Cas suppression mechanisms will be crucial for assessing the general impact of CRISPR–Cas systems (Bondy-Denomy and Davidson, 2014).

Acknowledgements

The authors acknowledge Dr. Shammi Kapoor (Professor), Punjab Agricultural University, Ludhiana for the valuable guidance and time.

References

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 315: 1709–1712.

Bhaya, D., Davison, M., Barrangou, R. 2011. CRISPR-Cas Systems in Bacteria and Archaea:Versatile Small RNAsfor Adaptive Defenseand Regulation. *Ann. Rev. Genetics*, 45: 273–297.

Bolotin, A., Quinquis, B., Sorokin, A., Ehrlich, S.D. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, 151: 2551–2561.

Bondy-Denomy, J., Davidson, A.R. 2014. To acquire or resist: the complex biological effects of CRISPR–Cas systems. *Trends Microbiol.,* 22(4): 218–225.

Breitbart, M., Rohwer, F. 2005. Here a virus, there a virus, everywhere the same virus? *Trends Microbiol.,* 13: 278–284.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., Van der Oost, J. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science,* 321: 960–964.

Cady, K.C., White, A.S., Hammond, J.H., Abendroth, M.D., Karthikeyan, R. S. 2011. Prevalence, conservationand functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology,* 157: 430–37.

Carte, J., Wang, R., Li, H., Terns, R.M., Terns, M.P. 2008. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.*, 22: 3489–3496.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 468: 67–71.

Grissa, I., Vergnaud, G., Pourcel, C. 2007. CRISPR Finder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.,* 35: W52–57.

Groenen, P.M., Bunschoten, A.E., van Soolingen, D., van Embden, J.D. 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: 1057–1065.

Haft, D.H., Selengut, J., Mongodin, E.F., Nelson, K.E. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computat. Biol.,* 1**:** 60.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R. 2009. RNA-guided RNAcleavage by a CRISPR RNA–Cas protein complex. *Cell,* 139: 945–956.

Haurwitz, R.E., Jinek, M., Wiedenheft, B., Zhou, K., Doudna, J.A. 2010. Sequence- and structure-specific

RNA processing by a CRISPR endonuclease. *Science* 329:1355–58

Hyman, P., Abedon, S.T. 2010. Bacteriophage host range and bacterial resistance. *Adv. Appl. Microbiol.,* 70: 217–48.

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, J.D., Gaastra, W., Schouls, L.M. 2002. Identification of a novel family of sequence repeats among prokaryotes. *OMICS,* 6: 23–33.

Karginov, F.V., Hannon, G.J. 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, S.J., Samson, J.E., Moineau, S. 2010. Bacteriophage resistance mechanisms. *Nature Rev. Microbiol.,* 8: 317–27.

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

Lillestol, R.K., Shah, S.A., Brugger, K., Redder, P., Phan, H., Christiansen, J., Garrett, R.A. 2009. CRISPR families of thecrenarchaeal genus Sulfolobus:bidirectional transcription and dynamic properties. *Mol. Microbiol.,* 72: 259–272.

Liu, F., Barrangou, R., Gerner-Smidt, P., Ribot, E.M., Knabel, S.J. 2011. 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: 1946–1956.

Makarova, K.S., Aravind, L., Grishin, N.V., Rogozin, I.B., Koonin, E.V. 2002. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.,* 30: 482–496.

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

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E. 2011. Evolution and classification of the CRISPR/Cas systems. *Nature Rev. Microbiol.,* 9: 467–77.

Marraffini, L.A, Sontheimer, E.J. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science,* 322: 1843–45.

Marraffini, L.A., Sontheimer, E.J. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Rev. Genetics,* 11: 181–190.

Mojica, F.J., 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.

Mokrousov, I., Limeschenko, E., Vyazovaya, A., Narvskaya, O. 2007. *Corynebacterium diphtheria* spoligo

typing based on combined use of two CRISPR loci. *Biotechnol. J.*, 2: 901–906.

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.

Rohwer, F., Thurber, R.V. 2009. Viruses manipulate the marine environment. *Nature*, 459: 207–212.

Shah, S.A., Hansen, N.R., Garrett, R.A. 2009. Distribution of CRISPR spacer matches in viruses and plasmids of crenarchaealacidothermophiles and implications for their inhibitory mechanism. *Biochem. Soc. Transact.*, 37: 23–28.

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.

Terns, M.P., Terns, R.M. 2011. CRISPR-based adaptive immune systems. *Curr. Opin. Immunol.*, 14: 321–27.

Wang, R., Preamplume, G., Terns, M.P., Terns, R.M., Li, H. 2011. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage. *Structure*, 19: 257–264.

Wiedenheft, B., Sternberg, S.H., Doudna, J.A. 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, 482: 331–38.

Wiedenheft, B., van Duijn E., Bultema, J., Waghmare, S., Zhou, K. 2011. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Nat. Acad. Sci. USA.*, 108: 10092–10097.

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

Zegans, M. E., Wagner, J. C., Cady, K. C., Murphy, D. M., Hammond, J. H., and O'Toole, G. A. (2009). Interaction between bacteriophageDMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *Journal of Bacterioogy*. 191:210–219.