CRISPR-Cas System: The Powerful Modulator of Accessory Genomes in Prokaryotes

Anca Butiuc-Keul\textsuperscript{a, b}, Anca Farkas\textsuperscript{a, b}, Rahela Carpa\textsuperscript{a, b}, Dumitrana Iordache\textsuperscript{a}

\textsuperscript{a}Department of Molecular Biology and Biotechnology, Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania; \textsuperscript{b}Center of Systems Biology, Biodiversity and Bioresources, Babeş-Bolyai University, Cluj-Napoca, Romania

Keywords
Adaptation · \textit{cas} genes · CRISPR · Interference · Spacer acquisition

Abstract
Being frequently exposed to foreign nucleic acids, bacteria and archaea have developed an ingenious adaptive defense system, called CRISPR-Cas. The system is composed of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) array, together with CRISPR (\textit{cas})-associated genes. This system consists of a complex machinery that integrates fragments of foreign nucleic acids from viruses and mobile genetic elements (MGEs), into CRISPR arrays. The inserted segments (spacers) are transcribed and then used by cas proteins as guide RNAs for recognition and inactivation of the targets. Different types and families of CRISPR-Cas systems consist of distinct adaptation and effector modules with evolutionary trajectories, partially independent. The origin of the effector modules and the mechanism of spacer integration/deletion is far less clear. A review of the most recent data regarding the structure, ecology, and evolution of CRISPR-Cas systems and their role in the modulation of accessory genomes in prokaryotes is proposed in this article. The CRISPR-Cas system’s impact on the physiology and ecology of prokaryotes, modulation of horizontal gene transfer events, is also discussed here. This system gained popularity after it was proposed as a tool for plant and animal embryo editing, in cancer therapy, as antimicrobial against pathogenic bacteria, and even for combating the novel coronavirus – SARS-CoV-2; thus, the newest and promising applications are reviewed as well.

Introduction
The CRISPR-Cas system was observed for the first time in 1987 by a group of Japanese researchers [Ishino et al., 1987], even though they did not realize what this structure represents and how it works. It was only after 2005 that its role in immunity was foreseen, when it was observed that many of the inserted sequences in bacterial genomes were of viral or plasmidic origin [Mojica et al., 2005; Doudna and Charpentier, 2014]. In 2012, the system was proposed as a powerful gene-editing tool [Jinek et al., 2012] and since then, the CRISPR-Cas engineered system had many applications. It was already used in...
plant genome editing [Schaeffer and Nakata, 2015; Samanta et al., 2016], treatment of neurodegenerative diseases [Kolli et al., 2018], and especially in cancer therapy [Zhen et al., 2014; Dai et al., 2019; Zhang et al., 2020]. Recently, it was proposed for combating pathogenic bacteria [Bikard and Barrangou, 2017] and human viruses [Kolli et al., 2018]. Most recent outcomes on CRISPR usage are focused on the surveillance [Kellner et al., 2019] and therapy of retroviral infections, such as coronavirus disease 2019 (COVID-19) caused by the novel coronavirus SARS-CoV-2 [Li et al., 2020; Nguyen et al., 2020].

Despite the multitude of research articles about the CRISPR system, the detailed mechanism of action and the possible role of this immune system in the acquisition of foreign genetic material is still unknown. Thus, this article proposes a review of the most recent data regarding the structure, ecology, and evolution of CRISPR-Cas systems and their role in the modulation of accessory genomes in prokaryotes. The potential and applied impact of spacer acquisition and CRISPR array polymorphism are also discussed in this article as well as the main applications in biotechnology.

**CRISPR-Cas Structure**

Viruses were considered the key factors of life evolution, acting as a selective pressure on bacteria from different environments and facilitating genetic exchange. Approximately 20–40% of bacteria are killed every day by bacteriophage infection [Hampton et al., 2020]. Subsequently, it was shown that the acquisition of exogenous DNA via transduction, conjugation, and horizontal gene transfer (HGT) was followed by selection of an array of defence mechanisms in bacteria that allow the recognition of foreign DNA and distinguishing from self DNA [Horvath and Barrangou, 2010]. There are several different protection mechanisms in bacteria, and it is unclear whether clustering defence genes in islands offers a specific selective advantage [Doron et al., 2018] upon “genomic junkyards,” in which the defence genes frequently acquired via HGT are accumulated, because insertion in these regions is unlikely to be deleterious [Koonin, 2018].

Thus, most archaea and many bacteria developed several autoprotection strategies against foreign nucleic acids. Among them, CRISPR arrays together with associated *cas* genes form the CRISPR-Cas immune system [Rath et al., 2015; Høyland-Kroghsbo et al., 2017; Jackson et al., 2017; García-Martínez et al., 2018]. Information stored within CRISPR arrays is used to direct the sequence-specific neutralization of invading genetic elements, including viruses and plasmids [Jackson et al., 2017].

The first description of the system was in 1987 [Ishino et al., 1987], and since then it has been identified in 85% of archaea and 45% of bacteria [McGinn and Marraffini, 2019]. Their localization is different, CRISPR loci can be found in chromosomes and plasmids as well [Horvath and Barrangou, 2010]. The CRISPR-Cas locus generally consists of an operon of *cas* genes and a CRISPR array composed of a series of conserved, sequence-specific repeats (25–35 bp long), flanking unique inserts of similar length known as spacers (26–72 bp) (as shown in Fig. 1), acquired from extrachromosomal elements [Wright et al., 2016; Jackson et al., 2017; Koonin and Makarova, 2019]. Several sequences of CRISPR spacers show homology with extrachromosomal DNA [Mojica et al., 2005; Pourcel et al., 2005], such as sequences detected in viral genomes and plasmids, called protospacers [Horvath et al., 2008]. Spacers are key elements of adaptive immunity, their acquisition during previous infections [Pourcel et al., 2005; Makarova et al., 2006] ensuring the recognition and neutralization of invaders in case of subsequent attacks [Barrangou et al., 2007].

Partially palindromic sequences were found in most of the repeats, conferring stability and highly conserved secondary structures. Palindromic repeats are predicted to generate RNAs with stable hairpin structures [Kunin et al., 2007; Koonin and Makarova, 2019]. Usually, one CRISPR locus contains conserved repeat sequences, but there is a polymorphism in the sequence and the length of the repeats between different CRISPR loci [Kunin et al., 2007]. Prior to the first repeat of each CRISPR array, there is a DNA sequence rich in adenine and thymine, called...
the leader sequence [Grissa et al., 2007]. New spacers, derived from invading MGEs are integrated into the CRISPR array through the activity of Cas proteins [Jackson et al., 2017]. The spacer sequences of a given bacterial strain, therefore, reveal the history of past CRISPR-Cas-mediated interactions. The number of CRISPR loci and the length of each locus are variable. Several bacteria contain multiple CRISPR loci (e.g., 18 CRISPR loci in Methanocaldus sp. FS406-22), and some of these loci can be hundreds of repeat-spacer units. The number and length of CRISPR loci are not correlated with genome size; some of the smallest microbial genomes (e.g., Nanoarchaeum equitans) contain multiple CRISPR loci [Sorek et al., 2013].

Cas Proteins

CRISPR systems consist of CRISPR arrays (spacers and repeat sequences) and adjacent CRISPR associated cas genes, which code for proteins involved in the immune response [Brouns et al., 2008] and DNA repair [Makarova et al., 2002]. Diverse cas genes were discovered in different organisms, making their classification very difficult, as the number of annotated genomes increases. At first, four cas genes were identified in genomes containing the CRISPR-Cas system [Jansen et al., 2002], but accumulating genome sequences have led to the identification of 93 cas genes grouped in ~45 different gene families based on sequence similarity of the encoded proteins [Haft et al., 2005; Makarova et al., 2015]. Six of these cas genes (cas1–cas6) are widely distributed, cas1 and cas2 being considered as a hallmark in genomes that contain CRISPR loci [Haft et al., 2005]. The analysis of cas1 sequences suggests several distinct versions of CRISPR-Cas systems [Haft et al., 2005], each of them being defined by a unique composition and conserved arrangement of cas genes (Table 1).

The Cas proteins interact with specific sets of CRISPR loci [Makarova et al., 2006]. Different subtypes of Cas proteins were named after representative organisms that contained a single CRISPR-Cas locus (for example, Cas proteins from Escherichia coli were designated Cse1 [CRISPR system of E. coli gen1]; other subtypes included Aeropyrum [Csa], Desulfovibrio [Csd], Haloarcula [Csh], Mycobacterium [Csm], Neisseria [Csn], Thermotoga [Cst], and Yersinia [Csy]) [Haft et al., 2005]. These initial categories could not easily handle the relationships between homologous Cas proteins, the variability of cas operons, or organisms that contain multiple CRISPR loci. A new classification system based on the evolutionary relationships between conserved proteins and cas operon suggests the organization of the CRISPR-Cas system in two modules: the adaptation module that requires the proteins Cas1 and Cas2, which are involved in spacer acquisition, and the effector module, that is required for the processing of primary CRISPR transcripts (crRNA), interference and degradation of foreign nucleic acids. Thus, the CRISPR subtypes may differ in the Cascade proteins involved in the processing of crRNA, some CRISPR systems require multiple proteins, whereas other types require a single multifunctional protein. Multiple protein complexes were identified mostly in the class 1 CRISPR-Cas system, where the effector complexes of type I and type III are phylogenetically related ribonucleoproteins [Brouns et al., 2008; Makarova et al., 2015]. Unlike these, the class 2 system contains a single interference protein as it was observed in types II, V and VI [Shmakov et al., 2017]. Several other proteins containing at least one RNA recognition motif with unknown function (RAMPs) seem to be involved in the processing of pre-crRNA [Makarova et al., 2011].

Classification of CRISPR-Cas Systems

A simplified classification of these immune systems is almost impossible because most of the prokaryotes acquired multiple CRISPR loci by frequent HGT, and the CRISPR-Cas systems show a high diversity of Cas proteins. Nevertheless, 2 classes and six types with multiple subtypes of CRISPR-Cas systems were identified. Based on the distinct architecture of the effector modules, types I, II, and IV belong to class 1, and types II, V, and VI belong to class 2 [Koonin and Makarova, 2019]. Class 1 systems possess multisubunit crRNA-effector complexes, while in class 2 systems all functions of the effector complex are carried out by a single protein, such as Cas9 (Table 1). There is evidence for types IV and V, which belong to class 1 and class 2, respectively [Koonin and Makarova 2019]. There are several differences between the CRISPR-Cas systems in terms of expression, interference, and adaptation modules [Makarova et al., 2015]. Although most of the prokaryotes contain one type of CRISPR-Cas system, the coexistence of different types of systems was also reported. Class 1 CRISPR-Cas systems are considered evolutionary ancestral systems. The class 2 systems have evolved from class 1 systems via the insertion of transposable elements encoding various nucleases, and are now being used as tools for genome editing [Mohanraju et al., 2016].
The common feature of all type I systems is the presence of protein Cas3 [Sinkunas et al., 2011]. The number of cas genes is variable and defines six subtypes of the type I system (type I-A to type I-F). All type I systems encode the Cas1, Cas2, Cas3 proteins as a Cascade-like complex. Cascade complex is involved in target localization, spacer acquisition, and crRNA processing. The type I-A systems show a particular Cascade complex with Cas3 as part of the complex [Rath et al., 2015]. The Cas1, Cas2, Cas9 proteins, and sometimes Csn2 or Cas4 are encoded by type II CRISPR-Cas systems [Barrangou et al., 2007; Heler et al., 2015; Wei et al., 2015]. Cas9 is involved in adaptation, crRNA processing, and cleavage of the target DNA [Heler et al., 2015; Wei et al., 2015]. Type II systems were further divided in subtypes II-A, II-B and II-C. Type II-C lacks the gene for Csn2 or Cas4 [Chylinski et al., 2013; Koonin and Makarova, 2013]. The protein Cas10 was found in type III CRISPR-Cas systems, but their function is not clear yet. Most Cas proteins are part of the complexes similar to Cascade as Csm (in Type III-A) or Cmr (in

**Table 1. Major Cas proteins (other Cas proteins are reviewed by Makarova et al., 2011)**

<table>
<thead>
<tr>
<th>Process</th>
<th>Protein</th>
<th>Distribution type/subtype</th>
<th>Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptation</td>
<td>Cas1</td>
<td>Universal</td>
<td>DNase, not sequence specific, can bind RNA; present in all types; structure available for several cas 1 protein</td>
<td>Bhaya et al., 2011; Makarova et al., 2011</td>
</tr>
<tr>
<td></td>
<td>Cas2</td>
<td>Universal</td>
<td>Small RNase specific to U-rich regions; present in all types; structures available from <em>Thermus thermophiles</em> and <em>Sulfobolus solfatarius</em> and others</td>
<td>Bhaya et al., 2011; Makarova et al., 2011</td>
</tr>
<tr>
<td></td>
<td>Cas3</td>
<td>Type I signature (A, B, C, D, E, F)</td>
<td>DNA helicase and DNase activity; most proteins have a fusion with HD nuclease Single-stranded DNA nuclease (HD domain) and ATP-dependent helicase</td>
<td>Bhaya et al., 2011; Huo et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Cas4</td>
<td>(A, B, C, D), II (B), V</td>
<td>RecB-like nuclease with exonuclease activity homologous to RecB</td>
<td>Bhaya et al., 2011</td>
</tr>
<tr>
<td></td>
<td>Cas7</td>
<td>I (A, B, C, D, E, F)</td>
<td>RAMP protein, endoribonuclease involved in crRNA biogenesis; part of Cascade</td>
<td>Makarova et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Cas8</td>
<td>I (A, B, C)</td>
<td>Subunit of the interference module. Important in targeting of invading DNA by recognizing the protospacer adjacent motif sequence</td>
<td>Makarova et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Cas9</td>
<td>Type II signature (A, B)</td>
<td>Large multidomain protein with McrA-HNH nuclease domain and RuvC-like nuclease domain</td>
<td>Richter et al., 2012; Heler et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Cas10</td>
<td>I (D), III (A, B)</td>
<td>HD nuclease domain, palm domain, Zn ribbon; some homologies with Cascade elements</td>
<td>Richter et al., 2012; Sorek et al., 2013</td>
</tr>
<tr>
<td>Expression</td>
<td>Cas5</td>
<td>I (A, B, C, E), III, IV</td>
<td>RAMP protein endoribonuclease involved in crRNA biogenesis; part of Cascade</td>
<td>Makarova et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Cas6</td>
<td>I (A, B, C, D, E, F), III (A, B)</td>
<td>RAMP protein, an endoribonuclease involved in crRNA biogenesis; part of Cascade; structure available from <em>P. furiosus</em></td>
<td>Sorek et al., 2013</td>
</tr>
<tr>
<td></td>
<td>Cas7</td>
<td>I (A, B, C, D, E, F), III, IV</td>
<td>RAMP protein endoribonuclease involved in crRNA biogenesis; part of Cascade</td>
<td>Makarova et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Cas8</td>
<td>I (A, B, C)</td>
<td>Large protein with McrA-HNH-nuclease domain and RuvC-like nuclease; part of Cascade</td>
<td>Bhaya et al., 2011</td>
</tr>
<tr>
<td></td>
<td>Cas10</td>
<td>I (D), type III signature (A, B)</td>
<td>HD nuclease domain, palm domain, Zn ribbon; some homologies with Cascade elements</td>
<td>Richter et al., 2012; Sorek et al., 2013</td>
</tr>
<tr>
<td>Interference</td>
<td>Cas8</td>
<td>I (A, B), IV</td>
<td>Large protein with McrA-HNH-nuclease domain and RuvC-like nuclease; part of Cascade</td>
<td>Bhaya et al., 2011</td>
</tr>
<tr>
<td></td>
<td>Cas9</td>
<td>Type II signature (A, B)</td>
<td>Large multidomain protein with McrA-HNH nuclease domain and RuvC-like nuclease domain</td>
<td>Richter et al., 2012; Heler et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Cas13a</td>
<td>VI</td>
<td>RNA-guided RNase</td>
<td>Cox et al., 2017</td>
</tr>
<tr>
<td></td>
<td>Cas13b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cas13c</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cas13d</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Type III-B) complexes [Rouillon et al., 2013; Staals et al., 2013].

Another difference between CRISPR-Cas systems is the target DNA in the case of type I and II systems and DNA and/or RNA in the case of type III systems. Type II system has been found only in bacteria, while the type I and type III systems in bacteria and archaea as well [Makarova et al., 2011]. Type III systems showing dual DNA/RNA interference activity are composed of several different subunit complexes comprising multiple copies of the Cas7 and the small subunit Cas11 (Csm2 or Cmr5) [Staals et al., 2014]. Type III CRISPR-Cas system is the only system that utilizes three different nuclease activities such as specific DNA/RNA cleavage and nonspecific ssDNA cleavage to provide efficient protection against DNA and RNA invaders. The Cas7 protein might function as a switch to lower the DNAse activity of Cas10 after releasing the cleaved target RNA from the complex [Samai et al., 2015]. Type III system is flexible to mutations in the protospacer sequence [Maniv et al., 2016], being able to neutralize escape mutants from that particular system [Silas et al., 2017]. The structure and particular functions of type III system have been reviewed by Zhu et al. [2018]. Due to their particularities, different biotechnological applications based on type III systems have been developed, such as genetic manipulations, including genome engineering and gene silencing [Liu et al., 2018].

Mechanism of CRISPR-Cas Adaptive Immunity

The defence process mediated by CRISPR-Cas system can be divided into three stages: adaptation, expression, and interference (shown in Fig. 2). The first step occurs only when the cell encounters new foreign DNA for the first time. The second and third steps happen any time the cell gets infected. The sequences of invading DNA corresponding to the spacers, called protospacers, are recognized by the system due to a short conserved region within a few bases (2–5 bp) near them, known as Protospacer-Adjacent Motif (PAM) [Mojica et al., 2009; Shah et al., 2013; Gleditzsch et al., 2019]. Protospacer recognition is followed by the generation and integration of a new sequence of nucleic acids identical to the protospacer into the CRISPR array as a new spacer, followed by reparation of the CRISPR array by cellular repair proteins [Wang et al., 2015; Hille and Charpentier, 2016], and the duplication of the proximal repeat [Liu et al., 2017]. Spacer acquisition from RNA (transcripts of a DNA genome of a MGE via reverse transcription) is possible by a reverse transcriptase mostly fused to the Cas1 protein [Silas et al., 2016, 2017].

Many of the CRISPR-Cas systems have highly consistent system-specific spacer lengths, except the type III CRISPR with variable length of spacers. The PAM sequence within the pre-spacer substrate allows the correct integration of new spacers into the CRISPR array, to
avoid the alignment at the wrong end of the crRNA target binding site. This process is mediated by Cas1–Cas2 complex. The Cas1 protein is involved in the cleavage of the CRISPR array at the leader-repeat junction and joining of the incoming spacer in-between of the repeat strands, the nicking and ligation occur in a concerted manner, corresponding to a classical integrase reaction [Arslan et al., 2014]. Such a mechanism is consistent with the predicted integrase activity of the Cas1 protein [Makarova et al., 2006]. Some authors consider that the integration of new spacers consists in two cleavages at the insertion site. The first cut, mediated by Cas1 or Cas2, occurs at the CRISPR/leader boundary, while the second nick occurs at the leader-distal end of the CRISPR, being determined by the distance from the previous cleavage. Cas1 protein interacts with Cas2 protein to form a complex that acts as a spacer integrase. This heterohexameric complex (Cas1–Cas2) contains two separate DNA-binding regions, one that binds the incoming protospacer and one that binds the CRISPR array. Once loaded with the incoming spacer, the Cas1–Cas2 complex catalyzes two cleavage-ligation reactions, first at the leader end of the first repeat, and subsequently at the spacer end of the repeat [McGinn and Marraffini, 2019]. Acquisition process also requires non-Cas proteins involved in DNA repair by the interaction of Cas1 with key components of repair systems such as RecB, RecC, and RuvB [Babu et al., 2011]. This mechanism was confirmed by Díez-Villaseñor et al. [2013]. Details about the function of Cascade complex are reviewed by Jackson et al. [2017].

Another important aspect is the acquisition of spacers from foreign DNA, and not from host DNA, to avoid the mechanism of self-targeting, which is similar to autoimmunity in euarkytic adaptive immune systems [Stern et al., 2010]. Acquisition of spacers from MGEs, other than those from host CRISPRs is termed naïve CRISPR adaptation and requires pre-spacer substrates generated from foreign material [Fineran and Charpentier, 2012]. Further studies are required to clarify how the diverse CRISPR-Cas systems balance the requirement for naïve production of pre-spacers from MGEs against the risk of acquiring spacers from host DNA. It is already known that in type II systems, spacer acquisition appears biased towards MGEs, and nuclease-deficient Cas9 fails to discriminate between host and foreign DNA [Wei et al., 2015].

In the second stage, the spacer is transcribed into a precursor of CRISPR RNA (pre-crRNA) that is subsequently processed into mature crRNA. The expressed spacer sequence provided by the crRNA is thought to recognize and guide the complex to bind the specific protospacer target. At the expression-processing stage, the pre-crRNA is processed to generate mature crRNAs by a distinct complex of Cas proteins, a dedicated processing nuclease (Cas6), a single large Cas protein or an external RNase [Brouns et al., 2008; Sorek et al., 2013; Rath et al., 2015; Mohanraju et al., 2016; Barrangou and Horvath, 2017].

In the third stage, interference, the foreign nucleic acid is recognized and destroyed by crRNA and Cas proteins. The complex of crRNA-Cas proteins locates the corresponding protospacer and triggers degradation of the target by specific Cas nucleases [Rath et al., 2015]. If there is no match between the CRISPR spacer and the foreign DNA, then the DNA sequence is not neutralized by the spacer transcript. In the case of a phage, it can reproduce inside the cell, leading to bacteria lysis and death [Han et al., 2013]. Different aspects of CRISPR-Cas biology were previously reviewed in numerous articles [Sorek et al., 2013; Charpentier et al., 2015; Mohanraju et al., 2016; Wright et al., 2016; Barrangou and Horvath, 2017; Jackson et al., 2017; Komor et al., 2017; Nishimasu and Nureki, 2017].

**Origin and Evolution of CRISPR-Cas Systems**

In the last decade, CRISPR-Cas systems were studied in detail, the structure and some of the biological functions of the system being deciphered [Sorek et al., 2013; Barrangou and Marraffini, 2014; Wright et al., 2016; Barrangou and Horvath, 2017]. Moreover, the Cas9, Cas12, and Cas13 endonucleases were successfully employed as genome-editing tools [Doudna and Charpentier, 2014; Hsu et al., 2014; Komor et al., 2017; Wu et al., 2018]. Being a defence mechanism, rapid evolution of the sequence of some of the cas genes in the effector module [Takeuchi et al., 2012] and diversification of the organization of the CRISPR-Cas loci were observed [Makarova et al., 2015; Koonin et al., 2017]. Despite the unclear origin of the effector modules, most of them were acquired by several classes of MGEs [Makarova et al., 2015; Mohanraju et al., 2016; Koonin and Makarova, 2019] and the components of CRISPR-Cas systems were recruited by various MGEs as well. Transposable elements similar to Tn7 encode different variants of CRISPR-Cas [Faure et al., 2019]. The adaptation process and the structure and function of the effector modules of the CRISPR-Cas systems are not completely elucidated due to their high diversity and autonomy. Usually, the adaptation module remains fixed,
while the effector module is changed in different variants of the same CRISPR-Cas type or subtype [Hudaiberdiev et al., 2017; Koonin and Makarova, 2019]. Moreover, in numerous bacteria and archaea, stand-alone adaptation and effector modules have been found. Thus, the Cas proteins give a functional separation of the modules of the CRISPR-Cas system.

CRISPR array is usually located in the direct vicinity of cas genes [Haft et al., 2005], but in some species such as *Listeria monocytogenes* [Mandin et al., 2007], *Aggregatibacter actinomycetemcomitans* [Jorth and Whiteley, 2012], and *Enterococcus faecalis* [Hullahalli et al., 2015], multiple CRISPR arrays may be distan to the cas genes. Such isolated CRISPRs, called orphan CRISPR arrays, are considered as remnants of previous functional CRISPR-Cas systems [Makarova et al., 2015], being involved in bacterial autoimmunity [Mandin et al., 2007]. Analysis of the CRISPR arrays in the orphan CRISPRs of *A. actinomycetemcomitans* showed that spacer sequences were antisense to bacterial self-coding genes [Jorth and Whiteley, 2012], suggesting the relationship between orphan CRISPRs and the regulation of other gene expression [Stern et al., 2010]. Even if all CRISPR loci are expressed, not all of these produced crRNA trigger interference, as it was shown in *Haloferax volcanii* [Maier et al., 2013].

Bioinformatic analysis of cas genes and genomic cas1 homologues showed that these proteins are not always encoded within the CRISPR-Cas loci [Makarova et al., 2013]. *Cas1* genes are found within 12–18 kb regions of genomic DNA flanked by terminal inverted repeats, very similar to transposable elements [Krupovic et al., 2014]. These transposable elements share two universal genes encoding Cas1, a family B DNA polymerase, and additional genes, mostly encoding diverse nucleases and DNA-binding proteins. Moreover, the mechanism of spacer integration catalyzed by Cas1 is very similar to transposon integration by integrase or casposon integration by casposase. Casposons are a newly identified group of transposons that use Cas1-like transposases for transposition. This led to the hypothesis that Cas1 could be assimilated to a casposase [Krupovic et al., 2014; Krupovic and Koonin, 2016]. Several casposons encode some virus capsid proteins, being similar to politons and are predicted to form virions [Yutin et al., 2018].

Four distinct casposon families integrated mostly into archaea and some bacteria were identified [Krupovic et al., 2017]. Family 1 of casposons were found exclusively in the archaeal phylum *Thaumarchaeota* and encoded PolBs of the protein-primed variety, being very similar to the corresponding proteins from archael viruses. The casposons from three other families encode divergent PolBs different from either protein-primed or RNA-primed PolBs [Makarova et al., 2015]. There is no evidence of the transposition of casposons, but some of the casposons are regarded as active transposons due to their mobility [Krupovic et al., 2016]. Thus, Krupovic et al. [2017], supposed that CRISPR repeats evolved directly from the preexisting casposon target site. Rollie et al. [2015] showed that Cas1, in the absence of Cas2, displays intrinsic sequence specificity to the sequence flanking the integration site. Similarly, casposase alone is sufficient to direct the specific integration of casposons into the target site [Béguin et al., 2016]. Other data strongly suggests that the leader sequence also evolved directly from the target site employed by the ancestral casposon. It was also shown that Cas1 evolved as the core enzyme of casposons and was eventually coopted to form the basis of the CRISPR-Cas immune system [McGinn and Marraffini, 2019].

Another key step sustaining the evolution of the CRISPR-Cas adaptation machinery from casposons was the formation of a pseudo-symmetrical heterohexameric complex by recruitment of Cas2. There is a significant difference between DNA of casposons and protospacers. The two ends of the linear casposon DNA designated for integration are identical and can be both specifically recognized by the casposase; thus, the two termini are poised for integration at a fixed distance, but the two ends of a protospacer are not identical. The protospacer end containing the PAM-complementary sequence is recognized by one of the four Cas1 subunits, whereas the accommodation of the other protospacer terminus is dictated by the quaternary structure of the Cas1-Cas2 complex [Wang et al., 2015]. The persistence length of dsDNA is 35–50 nm (100–150 bp) [Mitchell et al., 2017], which precludes bending of the short 30–70 bp protospacers such that both ends of the molecule would be held by the homodimer. Thus, the transformation of the casposon into CRISPR-Cas1 requires the recruitment of Cas2 enzyme that is homologous to mRNA interference, which could have already been present in the casposon that gave rise to CRISPR-Cas adaptation machinery [Koonin and Krupovic, 2015].

**Role of the CRISPR-Cas System in HGT**

**CRISPR-Cas Mediation of HGT**

Protection against chemical (e.g., antibiotics, disinfectants, or heavy metals) and biological agents (e.g., foreign DNA structures, such as viruses) is achieved via two ap-
presents that CRISPR-Cas systems are associated with small genomes and higher GC content. This suggests that CRISPR-Cas inhibits the acquisition of foreign DNA, since MGEs usually have lower GC content than their bacterial hosts. Another analysis of 300 *P. aeruginosa* genomes led to the conclusion that CRISPR-Cas restricts HGT, and the evolutionary mechanisms that ensure its maintenance or drive its loss are key to the ability of this pathogen to adapt to new niches and stressors [Wheatley and MacLean, 2020].

However, the relationship between CRISPR-Cas systems and HGT has not yet been elucidated. Despite the earlier lines of evidence that suggested the impairment of HGT by the CRISPR-Cas system, Gophna et al. [2015] assumed a more subtle relationship between the magnitude of HGT and the activity of CRISPR-Cas at the evolutionary scale. They tested the hypothesis that the activity of the CRISPR-Cas system would be negatively correlated with the magnitude of the HGT phenomenon. To assess the activity of the system, they considered the length of the CRISPR array: a large number of spacer and repeat sequences denoted a strong activity, but no statistically significant dependence was found between the number of recent HGT events and the activity of the CRISPR-Cas system. Most likely, the massive limitation of HGT would mean too high cost for the fitness of the CRISPR-Cas system. Most likely, the massive limitation of HGT would mean too high cost for the fitness of the CRISPR-Cas system. Gophna et al. [2015] tested the hypothesis that the activity of the CRISPR-Cas system controls the acquisition of resistance genes. They found that the CRISPR-Cas system prevents the acquisition of these plasmids in approximately 85% of the isolates in vivo (from the gastrointestinal tract of mice), but it has a much lower effect on the plasmids under different in vitro conditions (planktonic and biofilm). One study on an impressive collection of *Pseudomonas aeruginosa* isolates [van Belkum et al., 2015] showed that genomes have types I-F and I-E of the CRISPR-Cas system are shorter by about 300 kb than those without the system. This difference, which consists of approximately 22,036 genes, was called the accessory genome and was found to be DNA targeted by CRISPR-Cas systems. In *P. aeruginosa*, active CRISPR-Cas systems are associated with smaller genomes and higher GC content. This suggests that CRISPR-Cas inhibits the acquisition of foreign DNA, since MGEs usually have lower GC content than their bacterial hosts. Another analysis of 300 *P. aeruginosa* genomes led to the conclusion that CRISPR-Cas restricts HGT, and the evolutionary mechanisms that ensure its maintenance or drive its loss are key to the ability of this pathogen to adapt to new niches and stressors [Wheatley and MacLean, 2020].

However, the relationship between CRISPR-Cas systems and HGT has not yet been elucidated. Despite the earlier lines of evidence that suggested the impairment of HGT by the CRISPR-Cas system, Gophna et al. [2015] assumed a more subtle relationship between the magnitude of HGT and the activity of CRISPR-Cas at the evolutionary scale. They tested the hypothesis that the activity of the CRISPR-Cas system would be negatively correlated with the magnitude of the HGT phenomenon. To assess the activity of the system, they considered the length of the CRISPR array: a large number of spacer and repeat sequences denoted a strong activity, but no statistically significant dependence was found between the number of recent HGT events and the activity of the CRISPR-Cas system. Most likely, the massive limitation of HGT would mean too high cost for the fitness of the CRISPR-Cas system, and it would no longer be maintained [Mccann and Marraffini, 2019]. Therefore, the events in which the CRISPR-Cas system controls the acquisition of resistance genes occur rather on a population scale and not on an evolutionary scale. On the other hand, selection for mobility of CRISPR-Cas components by MGEs contributes to the evolution of CRISPR-Cas systems and of bacteria as well.

### HGT of the CRISPR-Cas System

Recent outcomes also revealed new directions with respect to the relationship between CRISPR-Cas systems and HGT. In archaea, besides spacers matching MGEs, such as plasmids or viruses, other spacers were found to match non-self chromosomal genes, providing evidence for the extensive inter-genus and inter-species HGT [Brodt et al., 2011]. Located on chromosomes, plasmids, or genomic islands, CRISPR-Cas systems as well are disseminated horizontally by conjugation or transduction [Watson et al., 2018; Varble et al., 2019]. Despite the ability of CRISPR-Cas systems to limit HGT through conjugation, transformation, and transduction, their main role in viral/bacteriophage resistance was demonstrated to promote HGT. A
complex dynamics between CRISPR-Cas-mediated phage resistance and transduction leads to the mobilization of a chromosomal CRISPR-Cas system containing phage-targeting spacers and enhances generalized transduction [Watson et al., 2018]. Protopacers can recombine with spacers in either chromosomal or plasmid-borne CRISPR loci, leading to either the transfer of CRISPR-adjacent genes or the propagation of acquired immunity to other bacteria in the population, respectively [Varble et al., 2019]. A comparative genomic and phylogenetic analysis of CRISPR-Cas variants associated with distinct families of transposable elements generated the hypothesis that they contribute to the propagation of these MGEs by facilitating transposition into specific sites. As it was previously shown, three distinct groups of Tn7-like transposons that encode different variants of type I CRISPR-Cas systems were identified, and a new mechanism of RNA-guided transposition was described for DNA transposons [Peters et al., 2017]. More recently, CRISPR-Cas systems were detected predominantly on MGEs such as genomic islands, plasmids, and transposon-like elements. Phylogenetic analysis of Cas proteins indicated that the CRISPR-Cas systems were acquired by HGT [McDonald et al., 2019]. The new paradigm is the mechanism by which CRISPR-Cas systems associated with transposons recognize protospacers and, instead of destroying them, transpose adjacent to them [Strecker et al., 2019]. Bacterial Tn7-like transposons have coopted nuclease-deficient CRISPR-Cas systems to catalyze RNA-guided integration of MGEs into the genome. Integration of donor DNA occurs highly specifically [Klompe et al., 2019]. The precision of RNA-guided DNA insertion with CRISPR-associated transposases and the discovery of a fully programmable integrase lays the foundation for genomic manipulations that obviate the requirements for double-strand breaks and homology-directed repair [Klompe et al., 2019; Strecker et al., 2019]. Having an important role in the ecology and evolution of microbial populations [Brockhurst et al., 2019], due to the variation induced by accessory genes such as MGEs, HGT is a risky evolutionary strategy. Recruitment of CRISPR-Cas systems by MGEs is a common phenomenon under which defence systems are repeatedly recruited for offence activities [Faure et al., 2019].

**CRISPR-Cas Regulation of Virulence Gene Expression in Bacteria**

During their infection cycle, pathogenic bacteria encounter different stress conditions and the presence of foreign DNA elements; thus, it is expected to rely on efficient defence systems such as CRISPR-Cas. Virulence, as a specific stress response of pathogens, results in the expression of virulence genes modulated by CRISPR-Cas system [Louwen et al., 2014]. Most pathogens involved in nosocomial infections possess biofilm-forming abilities, which play major roles in virulence and drug resistance. An increased ability to form biofilms was observed in CRISPR-Cas-positive strains of *E. faecalis* [Bourgogne et al., 2008] and *P. aeruginosa* [Zegans et al., 2009]. CRISPR-Cas gene expression is induced by quorum sensing in *Seratia marcescens* [Patterson et al., 2016], *P. aeruginosa* [Hoyland-Kroghsbo et al., 2017] and *Clostridium difficile* [Maikova et al., 2018]. In *Acinetobacter baumannii*, specific genes involved in biofilm formation appear almost exclusively in strains enriched in CRISPR-Cas systems [Mangas et al., 2019]. Deletion of the cas3 gene in *Streptococcus mutans* affects biofilm formation [Tang et al., 2019], leading to diminished growth under oxidative stress, enhanced growth under low pH, and reduced survival under heat shock and DNA-damaging conditions [Serbanescu et al., 2015]. In *P. aeruginosa*, the cas3 gene is involved in the upregulation of virulence factors allowing the bacteria to evade recognition by the host immune system [Li et al., 2016]. In *Legionella pneumophila*, the cas2 gene is required for intracellular infection [Gunderson and Cianciotto, 2013]. Downregulation of biofilm formation ability and reduction of intracellular invasion by deletion of cas3 were also demonstrated in *Salmonella enterica* [Cui et al., 2020]. Cas9 gene is important for the virulence of *Campylobacter jejuni* [Louwen et al., 2013; Shabbir et al., 2018], *Francisella novicida* and *Neisseria meningitidis* [Sampson et al., 2013; Heidrich et al., 2019], affecting adhesion, invasion and intracellular survival. In *Streptococcus pyogenes*, deletion of the cas9 gene was linked to a reduction in the abundance of virulence determinants and the expression of several virulence regulatory proteins. The mutant strain has reduced adherence to epithelial cells and other altered parameters such as growth in human whole blood ex vivo, and virulence in a murine necrotizing skin infection model, compared to the wild-type parent [Gao et al., 2019]. Although the presence of cas genes was found to correlate with the upregulation of certain virulence factors, the complex role of the CRISPR-Cas systems in major pathogens remains to be discovered. A positive correlation between the virulence of *P. aeruginosa* and the presence of CRISPR-Cas systems was observed, but CRISPR-Cas activity is neither necessary nor sufficient for increased virulence. The findings of Vasquez-Rifo et al. [2019] suggest that bacterial adaptive immunity and virulence are indirectly associated.
with the effects of physiological, ecological, and evolutionary factors.

It was demonstrated that the CRISPR-Cas immune system prevented the virulence acquisition of *Streptococcus pneumoniae* [Bikard et al., 2012]. In *E. coli*, the negative correlation between the number of CRISPR units and the presence of pathogenicity traits was associated with the repeat content [García-Gutiérrez et al., 2015]. Investigating the impact of CRISPR-Cas systems in the *Bacillus cereus* group, they were found to be barriers to HGT, being selectively inactivated during bacterial evolution, to allow the acquisition of MGEs for adaptation to diverse environments. Introduction of a functional CRISPR-Cas system into a strain lacking the system resulted in reduced adaptability to various stresses and decreased pathogenicity [Zheng et al., 2020].

**Anti-CRISPR Defence**

As in any other prey-predator relationship, bacteria and bacteriophages constantly exert pressure on each other, which leads to a co-dependent evolution: bacteria find new ways to survive, and phages develop new ways to avoid or counteract bacterial defence mechanisms. The bacterial defence against bacteriophages is based on innate immunity [Seed, 2015; Trasanidou et al., 2019] found in all bacteria, and on the acquired immunity, represented by the CRISPR-Cas system. The innate immunity has low efficiency, while the CRISPR-Cas system is a considerable hindrance to phage invasion. Therefore, phages have developed mechanisms to escape this defence by encoding proteins that inactivate the system, which have been called anti-CRISPR proteins [Bondy-Denomy et al., 2013; Pawluk et al., 2014; Bondy-Denomy et al., 2015].

The first genes for anti-CRISPR proteins (*Acr*) have been identified in phages that infect *P. aeruginosa* with the type I-F CRISPR system [Bondy-Denomy et al., 2013]. Since then, numerous studies focused on finding proteins which also inactivate other types. To date, there are known proteins capable to inactivate several CRISPR types: I-C, I-D, I-E, I-F, II-A, II-C, V-A, and VI-B. For more details regarding the protein structure, mechanism of action, and bacterial species where they have been identified, see Trasanidou et al. [2019]. Recently, it has been discovered that a bacteriophage is able to form a nuclease-like structure, where it can hide from the DNA-targeting CRISPR immunity, but remains susceptible to RNA-targeting CRISPR-Cas systems [Malone et al., 2020; Mendoza et al., 2020]. One of those RNA-targeting systems is the type III CRISPR-Cas, for which a mechanism by which bacteriophages can subvert it has very recently been uncovered [Athukoralage et al., 2020]. The latest studies are focused on finding proteins that inactivate CRISPR-Cas9 systems used in genetic editing. CRISPR-Cas9 represents the most powerful and specific gene editing instrument currently available, even though one possible risk of this tool might be the nonspecific action of the system, which could occur over time. In these conditions, finding an off-switch for the system is as useful as it can be [Pawluk et al., 2016b; Rauch et al., 2017; Dong et al., 2017].

In addition to being located on phages, *Acr* genes have been found in other mobile DNA structures, such as plasmids and transposons, which also carry out HGT. It was shown that these genes are located in the genome of bacterial cells holding CRISPR-Cas systems, and most likely they come from MGEs, being kept in the bacterial genome because a temporary inactivation of the system can confer an advantage on the acquisition of beneficial genes [Pawluk et al., 2016a].

**Biotechnological Applications**

Besides the CRISPR-Cas-based research on genetic editing and cancer therapy, which started years ago, the latest research focuses on using the system to combat pathogens, such as bacteria or viruses, and many other biotechnological applications [Carroll and Zhou, 2017]. Self-targeting by CRISPR-Cas was previously found to be lethal in bacteria. By analyzing CRISPRs from 330 organisms, Stern et al. [2010] found that self-targeting occurs in 18% of all CRISPR-bearing organisms. Therefore, the use of CRISPR-Cas systems to combat bacterial pathogens was suggested by transferring the CRISPR loci targeting chromosomal regions associated with virulence or antibiotic resistance [Bikard et al., 2012]. The RNA-programmable genome editing using the Cas9 endonuclease proposed by Jinek et al. [2012] has been widely adopted. Stern et al. [2010], observed the abundance of degraded repeats near self-targeting spacers and no conservation across species; thus, they suggested that self-targeting is a form of autoimmunity rather than a regulatory mechanism that incurs an autoimmune fitness cost and could explain the abundance of degraded CRISPR systems across prokaryotes [Stern et al., 2010]. Therefore, CRISPR-Cas systems have been successfully repurposed to target pathogenic bacteria. Removal of individual bacterial
strains from mixed populations of *E. coli* opened new avenues for the development of “smart” antibiotics that prevent resistance and differentiate between pathogenic and beneficial microorganisms [Gomaa et al., 2014]. Such effective tools are the CRISPR-Cas3 and CRISPR-Cas9 constructs, which can be designed to target specific virulence genes and thus to attack pathogens, or to target antibiotic resistance genes and thus to fight drug resistance. This strategy is efficient in the case of chromosomal genes but not in the case of plasmid-encoded virulence and resistance, but research never stops. A series of CRISPR-Cas13a-based antibacterial nucleocapsids capable to kill resistant pathogens by recognizing antimicrobial resistance genes have recently been developed. CapsidCas13a constructs generated by packaging programmed CRISPR-Cas13a from *Leptotrichia shahii* into a bacteriophage capsid demonstrated sequence-specific killing by RNase activity. The systems were proposed as both therapeutic agents against bacterial infections and diagnostic agents to detect bacterial resistance genes [Kiga et al., 2020]. An interesting perspective of programmed CRISPR-Cas systems with discriminatory power between pathogens and commensal bacteria could be the modification of microbial communities (i.e., the human microbiome associated with different diseases).

A possible resistance could also be developed here, in time, under the selective pressure of using CRISPR as antimicrobials, due to the existence of Acr genes that we mentioned in the previous section. Development of the self-targeting gene editing technology allowed to target different positions of the virulence genes. In *S. mutans*, precise modification of glucosyltransferases resulted in decreased exopolysaccharide synthesis and a weakened ability to form biofilms [Gong et al., 2018]. CRISPR-Cas technology can be applied to create antimicrobials with the spectrum of activity chosen by design. RNA-guided nucleases targeting specific DNA sequences were delivered efficiently, improving survival in a *Galleria mellonella* infection model [Citorik et al., 2014]. The phage-based approach was also used to deliver a CRISPR-Cas9 system targeting the chromosomal kanamycin resistance gene in *Staphylococcus aureus*. As a result, a strong growth inhibition of resistant bacteria due to chromosome cleavage and subsequent cell death were obtained. In an in vivo mouse skin infection model, the treatment also led to a significant reduction of antibiotic-resistant bacteria [Bikard et al., 2014].

A new CRISPR-Cas9-mediated genome editing method and a base editing system were recently developed, enabling the cytidine base editor to inactivate a target gene by generating a premature stop codon. The highly efficient genetic manipulation in *Pseudomonas* species is estimated to accelerate a wide variety of investigations of bacterial physiology, drug target exploration, and metabolic engineering [Chen et al., 2018].

In addition, CRISPR interference (CRISPRi), a genetic perturbation technique, allows the sequence-specific repression of gene expression in prokaryotic and eukaryotic cells [Qi et al., 2013]. The CRISPRi mechanism was applied in *E. coli* for inhibiting bacterial biofilm by silencing the luxS gene encoding a synthase involved in the initial stage of biofilm formation [Zuberi et al., 2017]. The identification of rules for specific targeting of transcriptional repressors (CRISPRi) and activators (CRISPRa) to endogenous genes via endonuclease-deficient Cas9 laid new foundations to enable modulation of gene expression [Gilbert et al., 2014]. CRISPR-Cas systems have been successfully repurposed to target virulence factors and antibiotic resistance genes in bacteria and constitute an appealing option for programmable and sequence-specific antimicrobials [Bikard and Barrangou, 2017]. Genome editing became the major outcome of self-targeting using CRISPR-based technologies in eukaryotes. Targeted gene regulation is now used to influence the expression of genes in human cells for cell and gene therapy, genetic reprogramming, and regenerative medicine.

CRISPR-Cas system has been proposed for treating, besides bacterial infections, also viral infections. CRISPR-Cas9 has been engineered for combating human viruses, such as Epstein-Barr virus, human papillomavirus, and hepatitis B virus, all of this being DNA viruses. Since Cas9 cannot interact directly with RNA, CRISPR-Cas9 is not efficient for curing retroviral infections [Li et al., 2020]. Considering the current situation regarding the COVID-19 pandemic, caused by the retrovirus SARS-CoV-2, the CRISPR research is focusing on engineering a system able to target RNA, which can be used as a treatment. Recent studies showed that Cas13 proteins, Cas13a [Li et al., 2020] and Cas13d [Nguyen et al., 2020], are able to target RNA. Nguyen et al. were able to design efficient guide RNAs, which target the coding regions of the novel coronavirus. This guide RNAs are also safe since they do not affect the human transcriptome. A study available online as a preprint, while still under review, claimed that they managed to engineer a system able to target over 90% of all coronaviruses, which could be quickly used in other rising strains that could cause a pandemic [Abbott et al., 2020]. Another promising application for CRISPR-Cas systems in the context of the ongoing pandemic, is for the surveillance of the disease. The most recent research papers proposed
some diagnostic tools for SARS-CoV-2, alongside other RNA viruses [Metsky et al., 2020; Ding et al., 2020; Curti et al., 2020]. These tools are based on the SHERLOCK method (specific high-sensitivity enzymatic reporter unlocking) developed by Kellner et al. in 2019. This method is based on the capacity of Cas12 and Cas13 proteins to cleave unspecific RNA in their proximity, if the associated crRNA finds its target, which can lead to a color or florescent signal emission, when the DNA or RNA target is present. The specificity of the test depends on the crRNA design. Several crRNAs have been designed to target specific sequences in the SARS-CoV-2 [Metsky et al., 2020; Ding et al., 2020; Curti et al., 2020], other human coronaviruses [Metsky et al., 2020], and HIV [Ding et al., 2020]. These diagnostic tools seem to be rapid, highly specific, comparable in sensitivity to the qPCR method, and possibly portable.

**Conclusion**

Despite the numerous studies, the CRISPR-Cas system is still of great interest, several aspects are far less clear than others, such as the origin of the effector modules and the mechanism of spacer integration/deletion. Although the evolution of adaptation and effector modules seem to be partially independent, and the origin of the adaptation module in casposons is accepted, little is known about the effector module. The classification of CRISPR-Cas systems in 2 classes is based mostly on the organization of the effector module. The class 1 systems have multisubunit effectors with origin in the nucleases encoded by different MGEs. The class 2 systems, in which the effector consists of a single, large protein, have an uncertain origin. The contribution of MGEs to the recruitment of minimalist variants of CRISPR-Cas systems remains to be elucidated. Moreover, phylogenomic studies revealed a continuous diversification of the CRISPR-Cas systems and their adaptation modules as well, which needs more attention as well.

The role of the CRISPR-Cas system in the modulation of genotypes, physiology, and ecology of prokaryotes, the implication of HGT in the acquisition of CRISPR-Cas modules, and the role of the CRISPR-Cas system in the limitation of HGT, to allow the acquisition of advantageous genes are topics of great interest, as well as the development of new applications in the field of diagnostic and treatment of infectious diseases.

**Conflict of Interest Statement**

The authors have no conflicts of interest to declare.

**Funding Sources**

This work was supported by Babeş-Bolyai University, Cluj-Napoca, Romania. This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

**Author Contributions**


**References**


Crispr-Cas: The Immunomodulatory System in Prokaryotes


McDonald ND, Regmi A, Morreale DP, Borowski JD, Boyd EF. CRISPR-Cas systems are present predominantly on mobile genetic elements in Vibrio species. BMC Genomics. 2019; 20:105.


CRISPR-Cas: The Immunomodulatory System in Prokaryotes


Shmakov SA, Sitnik V, Makarova KS, Wolf YI, Severinov KV, Koonin EV. The CRISPR spacer space is dominated by sequences from species-specific mobiles. MBio. 2017;8.


Watson BNJ, Staals RHJ, Fineran PC. CRISPR-cas-mediated phage resistance enhances horizontal gene transfer by transduction. MBio. 2018;9(1).

Wei Y, Terns RM, Terns MP. Cas9 function and host genome sampling in Type II – A CRISPR-Cas adaptation. Genes Dev. 2015;29:536–61.

DOI: 10.1159/000516643
Wheatley RM, MacLean RC. CRISPR-Cas systems restrict horizontal gene transfer in Pseudomonas aeruginosa. ISME J. 2020.


