The CRISPR conundrum: evolve and maybe die, or survive and risk stagnation

Jesús García-Martínez[1], Rafael D. Maldonado[1], Noemí M. Guzmán[1] and Francisco J. M. Mojica[1,2,]*

[1] Departamento de Fisiología, Genética y Microbiología. Universidad de Alicante, Campus de San Vicente, 03690 San Vicente del Raspeig (Alicante), Spain.
[2] I.M.E.M. Ramón Margalef. Universidad de Alicante, Campus de San Vicente, 03690 San Vicente del Raspeig (Alicante), Spain.
* Corresponding Author:
Francisco J. M. Mojica, Tel: +34 965909761. Fax +34 965909569; E-mail: fmojica@ua.es

ABSTRACT CRISPR-Cas represents a prokaryotic defense mechanism against invading genetic elements. Although there is a diversity of CRISPR-Cas systems, they all share similar, essential traits. In general, a CRISPR-Cas system consists of one or more groups of DNA repeats named CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats), regularly separated by unique sequences referred to as spacers, and a set of functionally associated *cas* (CRISPR associated) genes typically located next to one of the repeat arrays. The origin of spacers is in many cases unknown but, when ascertained, they usually match foreign genetic molecules. The proteins encoded by some of the *cas* genes are in charge of the incorporation of new spacers upon entry of a genetic element. Other Cas proteins participate in generating CRISPR-spacer RNAs and perform the task of destroying nucleic acid molecules carrying sequences similar to the spacer. In this way, CRISPR-Cas provides protection against genetic intruders that could substantially affect the cell viability, thus acting as an adaptive immune system. However, this defensive action also hampers the acquisition of potentially beneficial, horizontally transferred genes, undermining evolution. Here we cover how the model bacterium *Escherichia coli* deals with CRISPR-Cas to tackle this major dilemma, evolution versus survival.

INTRODUCTION

The prokaryotic world has been historically the main source of tools for genetic engineering and molecular biology in general. CRISPR-Cas is a recent example of how the study of prokaryotes has revolutionized life sciences. Besides becoming the most important tool for genomic editing to date [1], the discovery of this immune system has marked an important milestone in the history of Microbiology.

Cas proteins, CRISPR loci and CRISPR RNAs are the core functional parts of an adaptive and heritable resistance system against foreign DNA. They enable the cell to keep memory of infections by exogenous elements and fight against the invader. There is a significant diversity of genes associated with CRISPR, presumably reflecting the selective pressure viruses exert on the evolution of the system. Classification of CRISPR-Cas systems has been proven a chal-

lenging task [2–4], and new variants are emerging as sequencing data increases and functional studies on these systems are performed. Distinct CRISPR-Cas systems can coexist in a genome [4–8]. Moreover, the number of CRISPR loci pertaining to the same type varies among organisms, and both the identity and number of spacers within each array greatly changes even among genomes of closely related strains [9].

In this paper, we present an overview of the CRISPR-Cas systems outlining their discovery, classification and functional role, and we discuss about the evolutionary importance of these systems in the model organism *Escherichia coli*. The chromosome of *E. coli* strains may harbor up to two CRISPR-Cas systems involving as much as two repeat arrays each [6]. Equivalent arrays show a considerable intraspecific polymorphism in terms of spacer number and sequence. Fundamental knowledge about the CRISPR-

Cas mechanism has been generated from the analysis of these two systems in *E. coli* [10–13] and related Enterobacteriaceae [14].

DISCOVERY OF CRISPR LOCI AS DNA-MEMORY STORES

The serendipitous finding by Nakata and collaborators in 1987 [15] of five direct repeats next to the *iap* gene in *E. coli* was the first report of a CRISPR locus. Subsequently, in 1989 [16] the Nakata's team documented another array of repeats at approximately 20 kb from the first one. Soon after, Hermans *et al.* [17] found direct repeats in the unrelated, Gram-positive *Mycobacterium tuberculosis* complex, launching the use of the repeat loci for strain typing based on their particular spacer content [18]. Archaea first CRISPRs were discovered in 1993 [19], and the earliest functional studies on these sequences were performed in 1995 [20]. By the end of the 1990's, similar direct repeats were found in other prokaryotes and denominations given to these sequences started to multiply: DR, direct repeats [17]; TREPs, tandem repeats [20]; SRSR, short regularly spaced repeats [21]; DVR, direct variant repeats [22]; LCTR, large clusters of tandem repeats [23]; SPIDR, spacers interspersed direct repeats [24]. To avoid confusion, an agreement was made on naming the repeated sequences as CRISPR [25]. This acronym appeared published for the first time in 2002 [26]. By then, the biological relevance of these sequences was recognized, since they were distributed among many different, distantly related prokaryotes, representing a widespread family of repeats [21]. However, even though protein coding genes commonly associated to CRISPR arrays were discovered [26], unraveling their function was still pending. These Cas proteins, some of them related to helicases or nucleases, could play a role on DNA metabolism or expression[26].

Nevertheless, the definitive hint for the biological function of CRISPR-Cas came from the spacers rather than from the Cas or CRISPR units. In 2005, three independent studies found that some spacers matched sequences from transmissible genetic elements [27–29]. Notably, a comprehensive survey of the literature published on viruses and plasmids carrying spacer homologs, pointed to a relationship between immunity to these carriers and the presence of the cognate spacer in a potential host [27]. Therefore, it was suggested that the spacers represent a memory of past infections, and this information might be used to guide a defense mechanism. This fundamental breakthrough in the understanding of the CRISPR role in nature came hand in hand with the advent of increasing amounts of sequence data generated from viral, plasmid and complete genome sequences of prokaryotic strains which allowed researchers to cross-compare them. The existence of an adaptive, immunity-like system in Bacteria and Archaea was such an innovative idea that the three research groups undergone difficulties in publishing their results [30]. Historical perspectives of the initial moments of this discovery have been published elsewhere [25, 30–34] showing interesting insights into the way modern science works and how scientific discoveries are made.

In 2007, the function of CRISPR-Cas as a specific immune system was experimentally proven in *Streptococcus thermophilus* [35]: phage resistance was endowed after the incorporation of small fragments of the foreign genetic material as spacers into the CRISPR loci of the bacterium. Moreover, Cas proteins were shown to be involved in this immunity. One year later it was demonstrated that transcripts derived from CRISPR arrays in *E. coli* were processed by Cas proteins and that the resulting small RNAs (crRNAs) are necessary to achieve immunity [12].

CRISPR-CAS MECHANISM

Despite the diversification of CRISPR-Cas systems and their wide distribution in distantly related bacteria and archaea [4], the fundamental mechanism of this immune system is quite conserved, following three basic steps: adaptation, expression and interference.

Adaptation, or spacer acquisition, requires the integration of fragments of nucleic acids from invader molecules [36–38]. In addition to Cas, non-Cas proteins are involved in this stage [39]. Fragments of foreign nucleic acids selected for integration, named proto-spacers [40], are usually flanked by short conserved sequences, the proto-spacer adjacent motif (PAM) [41]. New spacers are preferentially integrated in a polarized manner [29], next to the terminal CRISPR unit downstream to an AT-rich region called leader [12, 26, 42]. The PAM sequence is needed for most, but not all systems to recognize foreign targets, and its absence in the own CRISPR array avoids self-targeting [43]. Most CRISPR-Cas systems acquire spacers directly from DNA donors but a few systems are able to gain new spacers derived from RNA precursors after retrotranscription [44].

The transcription of a CRISPR array from the leader generates a multi-spacer RNA (pre-crRNA) which is processed to single-spacer crRNAs with the participation of Cas proteins [12] and, in some systems, of non-Cas ribonucleases as well as a trans-activating crRNA (tracrRNA) that partially hybridizes with the pre-crRNA [45]. After processing, each mature crRNA (or crRNA/tracrRNA duplex) remains assembled with Cas proteins in a CRISPR ribonucleoprotein (crRNP) complex [46, 47]. This completes the second step of the CRISPR mechanism.

During the interference stage, the crRNP complex recognizes and directs cleavage of spacer-complementary sequences resulting in the elimination of molecules that carry potential targets [48]. Specific PAMs are crucial for efficient interference by many CRISPR-Cas systems [48–50]. In this case, upon PAM recognition by a protein of the crRNP complex, double-strand pairing is disrupted at the target DNA, leading to a R-loop conformation through progressive hybridization (starting from the PAM) with the spacer sequence in the crRNA [46]. The R-loop is the substrate for cleavage by Cas endonucleases [51]. Some CRISPR-Cas systems target RNA instead of, or in addition to, DNA [52–54].

CRISPR-Cas SYSTEMS CLASSIFICATION

Cas proteins are categorized in three functional modules [55]. The suite of proteins for the acquisition module is quite uniform. Regular members are Cas1 and Cas2 [36, 56], which have nuclease activities and form a multi-protein complex [57]. The Cas1-Cas2 adaptation complex appears to be assisted by Cas4 when present, and might be included in this module [58, 59]. In contrast to the acquisition proteins, the effector module (that is, proteins involved in pre-crRNA processing, target recognition and cleavage) is highly variable [3, 4, 60]. There is a third module of ancillary Cas proteins, involved in regulatory and other unknown roles [3, 60].

Due to the fast evolution and wide diversification of the CRISPR-Cas systems, a multiple criteria approach has been used for classification: signature *cas* genes specific for some types, sequence similarity between common Cas proteins, the phylogeny of Cas1 (the most conserved Cas protein) and gene configuration in the loci [3, 4]. The application of these criteria resulted in the current classification principle of two classes (1 and 2) and six types (from I to VI) [3]. Several subtypes (designated by letters, from 'A' forward) have been proposed based on signature genes and characteristic genomic arrangements [3, 4]. Moreover, at least in the case of *E. coli*, subtype variants showing substantial differences in *cas* sequence and PAM preference have been recognized within the species [61, 62]. This classification system also involves a systematic naming for Cas proteins that, in some cases, has changed over time to adapt to new discoveries [2–4].

Class 1 systems rely on multi-protein effector complexes [3]. They include Type I and Type III systems (distinguishable by the presence of Cas3 and Cas10, respectively) as well as the uncommon Type IV, devoid of an adaptation module. Class 2 is defined by the presence of a single-protein effector, namely Cas9, Cas12 or Cas13, depending on the particular type of system (Type II, Type V and Type VI, respectively) [3, 63]. In spite of the need for tracrRNAs by Type II systems [45], not being involved in Class 1 systems [3], most applications of CRISPR technology in heterologous hosts are based on Type II components. This is mainly because, in contrast to Class 1, a single protein is required for interference and the target is cleaved just once at precise sites [31].

HOW PROKARYOTES BYPASS THE GENETIC BARRIER DICTATED BY CRISPR: THE CASE OF *Escherichia coli*

Once the biological function of CRISPR-Cas was revealed, the potential drawbacks that fully efficient CRISPR-mediated interference could pose to prokaryotic evolution became evident [64]. Horizontal Gene Transfer (HGT) is one of the main forces driving genetic change in Bacteria and Archaea [65, 66]. However, the uptake of foreign nucleic acids might be constrained by functional CRISPR-Cas systems. To cope with such situation, prokaryotes either lack these systems or place them under stringent control [67, 68]. This is exemplified by the case of *E. coli*, a paradigm of genome plasticity [69, 70] in spite of being in pos-

session of CRISPR-Cas systems [6]: a subtype I-E system is present in the majority of strains and a complete I-F system exists in a reduced number of isolates. Still, cells harboring *cas* genes of the two subtypes are extremely rare [6, 61]. Unexpectedly, the *E. coli* I-F system is constitutively expressed under normal laboratory growth conditions [71, 72]. Therefore, in principle, it is permanently acting against gene transfer. However, the PAMs of the spacers present in the I-F arrays of *E. coli* differ from the proto-spacer adjacent motifs that elicit the most efficient interference [71]. Such a relaxed interference could provide the opportunity for beneficial foreign DNA to be acquired, while at the same time still limiting exchange of unwanted genetic material. Remarkably, when I-F *cas* are absent in the *E. coli* genome, an array with a limited number of I-F repeats is invariably present, allegedly as a remnant of an ancient complete I-F system [6, 61, 73]. Most strikingly, the vast majority of spacers in these orphan arrays match *cas* I-F genes [6, 73], playing a crucial role in preventing the barrier effect of their cognate genes [73]. Strains harboring these arrays use them as a constitutively expressed anti-*cas* mechanism that avoids the establishment of a fully equipped, immunity-prone CRISPR-Cas I-F system: intrusive DNA containing *cas* I-F genes is degraded through the action of the encoded Cas proteins guided by the resident crRNAs [73]. This anti-*cas* mechanism strongly supports the hypothesis that CRISPR-Cas immunity could be annoying for the carrier cell.

Opposite to I-F, expression of the *E. coli* I-E system is precisely regulated. H-NS protein is the main repressor of the system and its silencing effect can be lessened by the transcription factor LeuO [74–76]. The cAMP receptor protein (CRP) also contributes to CRISPR inhibition, acting as a competitor of LeuO for binding to the regulatory regions in the CRISPR-*cas* locus [77]. However, activity of the I-E system of *E. coli* has not been detected under the multiple laboratory growth conditions so far tested (our unpublished results), and the natural circumstances upon which such silencing is relieved remain to be clearly elucidated. In this regard, quorum sensing autoinducers of the N-Acyl-homoserine-lactone (AHL) class appear to activate CRISPR-Cas systems in Gram-negative bacteria such as *Pseudomonas aeruginosa* [78] and *Serratia* sp. [79] at elevated cell densities, when the risk of infection by bacteriophages is the highest [80]. Although this sort of induction has not been detected in pure cultures of *E. coli*, the presence in this species of AHL receptors [81, 82] raises the possibility that their CRISPR-Cas systems might be regulated through an interspecific crosstalk, by signals secreted by other members of the microbial community. Overall, these findings illustrate the complexity of I-E CRISPR-Cas regulation in *E. coli*. Moreover, its diverging spacer count and identity within the species is an indication that CRISPR activity, at least at the adaptation stage, is turned on at a different pace depending on the particular group of strains.

Related to this, a notable case is that of pathogenic strains. When compared to non-pathogens (i.e., commensals), they gain a selective advantage via the acquisition through HGT of virulence factors, allowing them the ability

to colonize more varied ecological niches within their hosts [70, 83, 84]. Inquiringly, a recent work from our group [85] established a negative correlation between pathogenicity and I-E CRISPR repeat count in *E. coli*: commensal strains tend to have more repeats than pathogenic isolates. This observation is compatible with the hypothesis that the activity of CRISPR-Cas I-E is kept limited when environmental adaptation needs to take precedence over protection, to minimize the negative effects of an evolutionary constraint. Another related question is why *E. coli* strains have lost either the I-E or the I-F *cas* genes, depending on their particular environment [6, 85]. Indeed, most extraintestinal pathogens pertaining to diverse phylogroups retain a I-F CRISPR-Cas system while the majority of commensals and enteric pathogens harbor a I-E system [85]. The preference for one or the other CRISPR-Cas subtype is suggestive of functional differences between the two systems. In this sense, previous works have reported that whereas spacers within I-E arrays of *E. coli* target viruses and plasmids alike, most I-F spacers matching known sequences have a plasmid origin [6, 72, 85]. Being plasmids the primary vectors for antibiotic resistance genes [86], this bias of I-F toward targeting plasmids is in agreement with the observation that the carrier strains are particularly susceptible to antibiotics [72]. Even though the reason for this apparent specialization is unknown, it highlights the inconvenience of an indiscriminate interference and the burden of carrying multiple CRISPR-Cas systems.

In summary, the analysis of the different CRISPR-Cas settings found in *E. coli* strengthen the idea that these systems, despite conferring protection, could severely hamper prokaryote evolution, hinting at how detrimental they could become if left unrestricted. Therefore, avoiding *cas* genes and limiting CRISPR-Cas activity when present appears to be a necessary evil for a prokaryote, where a delicate balance should be reached between the two extremes, those of promiscuity or chastity in terms of genetic exchange.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

The authors have no conflict of interests to declare.

COPYRIGHT

REFERENCES

1. 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. doi: 10.1126/science.1225829

2. Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, Moineau S, Mojica FJM, Wolf YI, Yakunin AF, van der Oost J, and Koonin E V. (**2011**). Evolution and classification of the CRISPR–Cas systems. **Nat Rev Microbiol** 9(6): 467–477. doi: 10.1038/nrmicro2577

3. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJJ, Charpentier E, Haft DH, Horvath P, Moineau S, Mojica FJM, Terns RM, Terns MP, White MF, Yakunin AF, Garrett RA, van der Oost J, Backofen R, and Koonin E V. (**2015**). An updated evolutionary classification of CRISPR–Cas systems. **Nat Rev Microbiol** 13(11): 722–736. doi: 10.1038/nrmicro3569

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

5. Haft DH, Selengut J, Mongodin EF, and Nelson KE (**2005**). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/cas subtypes exist in prokaryotic genomes. **PLoS Comput Biol** 1(6): 0474–0483. doi: 10.1371 /journal.pcbi.0010060

6. Díez-Villaseñor C, Almendros C, García-Martínez J, and Mojica FJM (**2010**). Diversity of CRISPR loci in *Escherichia coli*. **Microbiology** 156(5): 1351–1361. doi: 10.1099/mic.0.036046-0

7. Rousseau C, Gonnet M, Le Romancer M, and Nicolas J (**2009**). CRISPI: A CRISPR interactive database. **Bioinformatics** 25(24): 3317–3318. doi: 10.1093/bioinformatics/btp586

8. Zhang Q, and Ye Y (**2017**). Not all predicted CRISPR–Cas systems are equal: isolated *cas* genes and classes of CRISPR like elements. **BMC Bioinformatics** 18(1): 92. doi: 10.1186/s12859-017-1512-4

9. Grissa I, Vergnaud G, and Pourcel C (**2007**). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. **BMC Bioinformatics** 8: 1–10. doi: 10.1186/1471-2105-8-172

10. Redding S, Sternberg SH, Marshall M, Gibb B, Bhat P, Guegler CK, Wiedenheft B, Doudna JA, and Greene EC (**2015**). Surveillance and processing of foreign DNA by the *Escherichia coli* CRISPR-Cas system. **Cell** 163(4): 854–865. doi: 10.1016/j.cell.2015.10.003

11. van Erp PBG, Jackson RN, Carter J, Golden SM, Bailey S, and Wiedenheft B (**2015**). Mechanism of CRISPR-RNA guided recognition of DNA targets in *Escherichia coli*. **Nucleic Acids Res** 43(17): 8381–8391. doi: 10.1093/nar/gkv793

12. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin E V, and van der Oost J (**2008**). Small CRISPR RNAs guide antiviral defense in prokaryotes. **Science** 321(5891): 960–964. doi: 10.1126/science.1159689

13. Levy A, Goren MG, Yosef I, Auster O, Manor M, Amitai G, Edgar R,

Qimron U, and Sorek R (**2015**). CRISPR adaptation biases explain preference for acquisition of foreign DNA. **Nature** 520(7548): 505–510. doi: 10.1038/nature14302

14. Medina-Aparicio L, Dávila S, Rebollar-Flores JE, Calva E, and Hernández-Lucas I (**2018**). The CRISPR-Cas system in Enterobacteriaceae. **Pathog Dis** 76(1): fty002. doi: 10.1093/femspd/fty002

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

16. Nakata A, Amemura M, and Makino K (**1989**). Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. **J Bacteriol** 171(6): 3553–3556. doi: 10.1128/jb.171.6.3553-3556.1989

17. Hermans PWM, van Soolingen D, Bik EM, De Haas PEW, Dale JW, and van Embden JDA (**1991**). Insertion element IS*987* from *Mycobacterium bovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. **Infect Immun** 59(8): 2695–2705. PMID:1649798

18. Groenen PMA, Bunschoten AE, van Soolingen D, and van Embden JDA (**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(5): 1057–1065. doi: 10.1111/j.1365-2958.1993.tb00976.x

19. Mojica FJM, Juez G, and Rodríguez-Valera F (**1993**). Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified *Pst*I sites. **Mol Microbiol** 9(3): 613–621. doi: 10.1111/j.1365-2958.1993.tb01721.x

20. Mojica FJM, Ferrer C, Juez G, and Rodríguez-Valera F (**1995**). Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. **Mol Microbiol** 17(1): 85–93. doi: 10.1111/j.1365-2958.1995.mmi_17010085.x

21. Mojica FJM, Díez-Villaseñor C, Soria E, and Juez G (**2000**). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. **Mol Microbiol** 36(1): 244–246. doi: 10.1046/j.1365-2958.2000.01838.x

22. van Embden JDA, van Gorkom T, Kremer K, Jansen R, van der Zeijst BAM, and Schouls LM (**2000**). Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. **J Bacteriol** 182(9): 2393–2401. doi: 10.1128/JB.182.9.2393-2401.2000

23. She Q et al. (**2001**). The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. **Proc Natl Acad Sci** 98(14): 7835–7840. doi: 10.1073/pnas.141222098

24. Jansen R, van Embden JD, Gaastra W, and Schouls LM (**2002**). Identification of a novel family of sequence repeats among prokaryotes. **Omics** 6(1): 23–33. doi: 10.1089/15362310252780816

25. Mojica FJM, and Garrett RA (**2013**). Discovery and seminal developments in the CRISPR field. In: Barrangou R, van der Oost J, editors. CRISPR-Cas systems: RNA-mediated adaptive immunity in Bacteria and Archaea. **Springer, Berlin, Heidelberg**; pp 1–31.

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

27. Mojica FJM, Díez-Villaseñor C, García-Martínez J, and Soria E (**2005**). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. **J Mol Evol** 60(2): 174–182. doi:

10.1007/s00239-004-0046-3

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

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

30. Lander ES (**2016**). The heroes of CRISPR. **Cell** 164(1–2): 18–28. doi: 10.1016/j.cell.2015.12.041

31. Mojica FJM, and Montoliu L (**2016**). On the origin of CRISPR-Cas technology: From prokaryotes to mammals. **Trends Microbiol** 24(10): 811–820. doi: 10.1016/j.tim.2016.06.005

32. Mojica FJM, and Rodríguez-Valera F (**2016**). The discovery of CRISPR in archaea and bacteria. **FEBS J** 283: 3162–3169. doi: 10.1111/febs.13766

33. Barrangou R, and Horvath P (**2017**). A decade of discovery: CRISPR functions and applications. **Nat Microbiol** 2: 17092. doi: 10.1038/nmicrobiol.2017.92

34. Morange M (**2015**). What history tells us XXXVII. CRISPR-Cas: The discovery of an immune system in prokaryotes. **J Biosci** 40(2): 221–223. doi: 10.1007/s12038-015-9532-6

35. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, and Horvath P (**2007**). CRISPR provides acquired resistance against viruses in prokaryotes. **Science** 315(5819): 1709–12. doi: 10.1126/science.1138140

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

37. Jackson SA, McKenzie RE, Fagerlund RD, Kieper SN, Fineran PC, and Brouns SJJ (**2017**). CRISPR-Cas: Adapting to change. **Science** 356(6333): eaal5056. doi: 10.1126/science.aal5056

38. Sternberg SH, Richter H, Charpentier E, and Qimron U (**2016**). Adaptation in CRISPR-Cas systems. **Mol Cell** 61(6): 797–808. doi: 10.1016/j.molcel.2016.01.030

39. Killelea T, and Bolt EL (**2017**). CRISPR-Cas adaptive immunity and the three Rs. **Biosci Rep** 37(4): BSR20160297. doi: 10.1042/BSR20160297

40. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, Romero DA, Horvath P, and Moineau S (**2008**). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. **J Bacteriol** 190(4): 1390–1400. doi: 10.1128/JB.01412-07

41. Mojica FJM, Díez-Villaseñor C, García-Martínez J, and Almendros C (**2009**). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. **Microbiology** 155: 733–740. doi: 10.1099/mic.0.023960-0

42. Lillestøl RK, Redder P, Garrett RA, and Brügger K (**2006**). A putative viral defence mechanism in archaeal cells. **Archaea** 2(1): 59–72. doi: 10.1155/2006/542818

43. Westra ER, Semenova E, Datsenko KA, Jackson RN, Wiedenheft B, Severinov K, and Brouns SJJ (**2013**). Type I-E CRISPR-Cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. **PLoS Genet** 9(9). doi: 10.1371/journal.pgen.1003742

44. Silas S, Mohr G, Sidote DJ, Markham LM, Sanchez-Amat A, Bhaya D, Lambowitz AM, and Fire AZ (**2016**). Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. **Science** 351(6276): aad4234. doi: 10.1126/science.aad4234

45. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, and Charpentier E (**2011**). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. **Nature** 471(7340): 602–607. doi: 10.1038/nature09886

46. Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, Wiedenheft B, Pul Ü, Wurm R, Wagner R, Beijer MR, Barendregt A, Zhou K, Snijders APL, Dickman MJ, Doudna JA, Boekema EJ, Heck AJR, van der Oost J, and Brouns SJJ (**2011**). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. **Nat Struct Mol Biol** 18(5): 529–536. doi: 10.1038/nsmb.2019

47. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns RM, and Terns MP (**2009**). RNA-Guided RNA cleavage by a CRISPR RNA-Cas Protein Complex. **Cell** 139(5): 945–956. doi: 10.1016/j.cell.2009.07.040

48. Garneau JE, Dupuis M-È, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, and Moineau S (**2010**). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. **Nature** 468(7320): 67–71. doi: 10.1038/nature09523

49. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin E V, and Zhang F (**2015**). Cpf1 Is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. **Cell** 163(3): 759–771. doi: 10.1016/j.cell.2015.09.038

50. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, van der Oost J, Brouns SJJ, and Severinov K (**2011**). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. **Proc Natl Acad Sci** 108(25): 10098–10103. doi: 10.1073/pnas.1104144108

51. Sternberg SH, Redding S, Jinek M, Greene EC, and Doudna JA (**2014**). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. **Nature** 507(7490): 62–67. doi: 10.1038/nature13011

52. Hatoum-Aslan A, Palmer KL, Gilmore MS, and Marraffini LA (**2013**). Type III CRISPR-Cas systems and the roles of CRISPR-Cas in bacterial virulence. In: Barrangou R, van der Oost J, editors. CRISPR-Cas systems: RNA-mediated adaptive immunity in Bacteria and Archaea. **Springer, Berlin, Heidelberg**; pp 201–219.

53. Abudayyeh OO, Gootenberg JS, Essletzbichler P, Han S, Joung J, Belanto JJ, Verdine V, Cox DBT, Kellner MJ, Regev A, Lander ES, Voytas DF, Ting AY, and Zhang F (**2017**). RNA targeting with CRISPR–Cas13. **Nature** 550(7675): 280–284. doi: 10.1038/nature24049

54. Rousseau BA, Hou Z, Gramelspacher MJ, and Zhang Y (**2018**). Programmable RNA cleavage and recognition by a natural CRISPR-Cas9 system from *Neisseria meningitidis*. **Mol Cell** 69(5): 906–914.e4. doi: 10.1016/j.molcel.2018.01.025

55. Makarova KS, Wolf YI, and Koonin EV (**2013**). The basic building blocks and evolution of CRISPR–Cas systems. **Biochem Soc Trans** 41(6): 1392–1400. doi: 10.1042/BST20130038

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

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

58. Kieper SN, Almendros C, Behler J, McKenzie RE, Nobrega FL, Haagsma AC, Vink JNA, Hess WR, and Brouns SJJ (**2018**). Cas4 facilitates PAM-compatible spacer selection during CRISPR adaptation. **Cell Rep** 22(13): 3377–3384. doi: 10.1016/j.celrep.2018.02.103

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

60. Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin E V., and van der Oost J (**2016**). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. **Science** 353(6299): aad5147. doi: 10.1126/science.aad5147

61. Almendros C, Mojica FJM, Díez-Villaseñor C, Guzmán NM, and García-Martínez J (**2014**). CRISPR-Cas functional module exchange in *Escherichia coli*. **MBio** 5(1): 1–10. doi: 10.1128/mBio.00767-13

62. Díez-Villaseñor C, Guzmán NM, Almendros C, García-Martínez J, and Mojica FJM (**2013**). CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli*. **RNA Biol** 10: 792–802. doi: 10.4161/rna.24023

63. Shmakov S, Smargon A, Scott D, Cox D, Pyzocha N, Yan W, Abudayyeh OO, Gootenberg JS, Makarova KS, Wolf YI, Severinov K, Zhang F, and Koonin E V. (**2017**). Diversity and evolution of class 2 CRISPR–Cas systems. **Nat Rev Microbiol** 15(3): 169–182. doi: 10.1038/nrmicro.2016.184

64. Marraffini LA, and Sontheimer EJ (**2008**). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. **Science** 322(5909): 1843–5. doi: 10.1126/science.1165771

65. McDaniel LD, Young E, Delaney J, Ruhnau F, Ritchie KB, and Paul JH (**2010**). High frequency of horizontal gene transfer in the oceans. **Science** 330(6000): 50. doi: 10.1126/science.1192243

66. Thomas CM, and Nielsen KM (**2005**). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. **Nat Rev Microbiol** 3(9): 711–721. doi: 10.1038/nrmicro1234

67. Patterson AG, Yevstigneyeva MS, and Fineran PC (**2017**). Regulation of CRISPR–Cas adaptive immune systems. **Curr Opin Microbiol** 37: 1–7. doi: 10.1016/j.mib.2017.02.004

68. Leon LM, Mendoza SD, and Bondy-Denomy J (**2018**). How bacteria control the CRISPR-Cas arsenal. **Curr Opin Microbiol** 42: 87–95. doi: 10.1016/j.mib.2017.11.005

69. Porse A, Gumpert H, Kubicek-Sutherland JZ, Karami N, Adlerberth I, Wold AE, Andersson DI, and Sommer MOA (**2017**). Genome dynamics of *Escherichia coli* during antibiotic treatment: Transfer, loss, and persistence of genetic elements *in situ* of the infant gut. **Front Cell Infect Microbiol** 7: 126. doi: 10.3389/fcimb.2017.00126

70. Dobrindt U (**2005**). (Patho-)Genomics of *Escherichia coli*. **Int J Med Microbiol** 295(6–7): 357–371. doi: 10.1016/j.ijmm.2005.07.009

71. Almendros C, Guzmán NM, Díez-Villaseñor C, García-Martínez J, and Mojica FJM (**2012**). Target motifs affecting natural immunity by a constitutive CRISPR-Cas system in *Escherichia coli*. **PLoS One** 7(11): e50797. doi: 10.1371/journal.pone.0050797

72. Aydin S, Personne Y, Newire E, Laverick R, Russell O, Roberts AP, and Enne VI (**2017**). Presence of Type I-F CRISPR/Cas systems is associated with antimicrobial susceptibility in *Escherichia coli*. **J Antimicrob Chemother** 72(8): 2213–2218. doi: 10.1093/jac/dkx137

73. Almendros C, Guzmán NM, García-Martínez J, and Mojica FJM (**2016**). Anti-cas spacers in orphan CRISPR arrays prevent uptake of active CRISPR–Cas I-F systems. **Nat Microbiol** 1(8): 16081. doi: 10.1038/nmicrobiol.2016.81

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

75. Westra ER, Pul Ü, Heidrich N, Jore MM, Lundgren M, Stratmann T, Wurm R, Raine A, Mescher M, Van Heereveld L, Mastop M, Wagner EGH, Schnetz K, Van Der Oost J, Wagner R, and Brouns SJJ (**2010**). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. **Mol**

Microbiol 77(6): 1380–1393. doi: 10.1111/j.1365-2958.2010.07315.x

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

77. Yang CD, Chen YH, Huang HY, Huang H Da, and Tseng CP (**2014**). CRP represses the CRISPR/Cas system in *Escherichia coli*: Evidence that endogenous CRISPR spacers impede phage P1 replication. **Mol Microbiol** 92(5): 1072–1091. doi: 10.1111/mmi.12614

78. Høyland-Kroghsbo NM, Paczkowski J, Mukherjee S, Broniewski J, Westra E, Bondy-Denomy J, and Bassler BL (**2017**). Quorum sensing controls the *Pseudomonas aeruginosa* CRISPR-Cas adaptive immune system. **Proc Natl Acad Sci** 114(1): 131–135. doi: 10.1073/pnas.1617415113

79. Patterson AG, Jackson SA, Taylor C, Evans GB, Salmond GPC, Przybilski R, Staals RHJ, and Fineran PC (**2016**). Quorum sensing controls adaptive immunity through the regulation of multiple CRISPR-Cas systems. **Mol Cell** 64(6): 1102–1108. doi: 10.1016/j.molcel.2016.11.012

80. Knowles B et al. (**2016**). Lytic to temperate switching of viral communities. **Nature** 531(7595): 466–470. doi: 10.1038/nature17193

81. Lee J, Maeda T, Hong SH, and Wood TK (**2009**). Reconfiguring the quorum-sensing regulator SdiA of *Escherichia coli* to control biofilm formation via indole and N-acylhomoserine lactones. **Appl Environ Microbiol** 75(6): 1703–1716. doi: 10.1128/AEM.02081-08

82. Sperandio V (**2010**). SdiA sensing of acyl-homoserine lactones by enterohemorrhagic *E. coli* (EHEC) serotype O157. **Gut Microbes** 1(6): 432–435. doi: 10.4161/gmic.1.6.14177

83. Kaper JB, Nataro JP, and Mobley HLT (**2004**). Pathogenic *Escherichia coli*. **Nat Rev Microbiol** 2(2): 123–140. doi: 10.1038/nrmicro818

84. Diard M, and Hardt W-D (**2017**). Evolution of bacterial virulence. **FEMS Microbiol Rev** 41(5): 679–697. doi: 10.1093/femsre/fux023

85. García-Gutiérrez E, Almendros C, Mojica FJM, Guzmán NM, and García-Martínez J (**2015**). CRISPR content correlates with the pathogenic potential of *Escherichia coli*. **PLoS One** 10(7): e0131935. doi: 10.1371/journal.pone.0131935

86. Blair JMA, Webber MA, Baylay AJ, Ogbolu DO, and Piddock LJ V (**2015**). Molecular mechanisms of antibiotic resistance. **Nat Rev Microbiol** 13(1): 42–51. doi: 10.1038/nrmicro3380