STUDIA UNIVERSITATIS BABEŞ-BOLYAI BIOLOGIA, 67, 1, 2022 (p. 131-142)
doi:10.24193/subbbiol.2022.1.07

Revealing the CRISPR array in bacteria living in our organism

Matei-Ştefan Dobrescu[1]✉, Dumitrana Iordache[1,2,3] and Anca Butiuc-Keul[1,2]

*[1]Department of Molecular Biology, Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania; [2]Center of Systems Biology, Biodiversity and Bioresources, Babeş-Bolyai University Cluj-Napoca, Romania; [3]Doctoral School of Integrative Biology, Babeş-Bolyai University, Cluj-Napoca, Romania; ✉**Corresponding author, E-mail: matei.dobrescu@stud.ubbcluj.ro.***

Article history: Received: 31 March 2022; Revised: 27 April 2022; Accepted: 9 June 2022; Available online: 30 June 2022.

Abstract. CRISPR (clustered regularly interspaced short palindromic repeats) is an immune system used by bacteria to defend themselves from different types of pathogens. It was discovered that this immune system can modify itself in specific regions called spacers due to previous interaction with foreign genetic material from phages and plasmids. Through our research, we have identified in different bacterial isolates CRISPR arrays belonging to the subtypes I-E (present in 42 samples) and I-F (present in 9 samples). The number of spacers in CRISPR arrays was also estimated based on the array length as a possible connection with the systems activity. Our results yielded arrays as small as 200 bp and as large as 1400 bp.

Keywords: CRISPR array; clinical isolates; pathogens; bacteria

Introduction

The end of the XXth century represented a reference point regarding genetics and molecular biology. The discovery of the CRISPR-Cas system (Clustered Regularly Interspaced Short Palindromic Repeats) meant a new opportunity for research, an area never explored before, that can bring new therapies in the medical domain, as well in associated sciences and areas.

CRISPR-Cas system consists of an immune system used by bacteria to protect themselves from pathogens. Over the years, scientists have discovered the inner workings of this immune system and harvested the power it holds: targeted genetic manipulation. The CRISPR-Cas system has been discovered in around 90% of Archaea and 60% of Bacteria (Palmer and Gilmore, 2010).

CRISPR arrays are a family of diverse DNA sequences with very similar structures. Each locus comprises repeated sequences of 20–40 bp, called repeats, and sequences containing the genetic information, called spacers, ranging between 25–70 bp (Kunin *et al.,* 2007). Besides the CRISPR array, there are many associated proteins. These proteins have been named Cas proteins (CRISPR Associated Proteins). In the beginning, only 4 Cas proteins have been identified (Jansen *et al.,* 2002). However, over the years, at least 45 families of Cas proteins have been discovered. Out of these proteins, Cas 1 and Cas 2 are universally conserved (Haft *et al.,* 2005).

There are different classes of CRISPR loci. The main method for classifying them is by the Cas cluster (Makarova *et al.,* 2011a). Thus, there is CRISPR Class I, defined by the use of multiproteic effectors. In this CRISPR class, there is a multitude of different proteins involved in the immune process (Barrangou, 2015). The second class is defined by the now-famous protein Cas 9. The Cas 9 protein can cut DNA strands, resulting in DNA sequences with straight ends and aids in the insertion of new spacers (Garneau *et al.,* 2010).

The immune response mediated by CRISPR-Cas works in three different stages. The first stage is adaptation, which consists of acquiring new spacer sequences, creating a genetic memory that will be used in further stages (Devashish *et al.,* 2015). In this stage, the Cas 1 and Cas 2 proteins play a crucial role in the integration of new genetic data (Yosef *et al.,* 2012). These proteins create a complex that will bind to the CRISPR DNA due to the Cas 1 protein (Nuñez *et al.,* 2014). Vital to this stage is the existence of PAM (Protospacer Adjacent Motif), which helps differentiation between self and non-self sequences (Mojica *et al.,* 2009). After the non-self sequence has been identified, the protein cascade or Csy complex in junction with host cell proteins (polymerases, ligases) will begin working towards incorporating the new spacer sequence (Devashish *et al*., 2015).

The second stage is the expression. The result of this stage is the formation of an RNA molecule called pre-crRNA (Devashish *et al*., 2015). This RNA molecule needs to mature; thus, maturation is essential to ensure the proper functionality of the CRISPR-Cas system. This stage requires the synthesis of specific proteins that modify the pre-crRNA molecule in the final crRNA (Gesner *et al.,* 2015). This initial transcript will be processed by Cas proteins (Carte *et al*., 2014), Cas 6 being the main protein realizing the transcript

processing (Nam *et al.,* 2014). This protein recognizes stem-loops present in pre-crRNA repeat sequences. The cutting takes place at the base of these loops, ensuring that cleaving only occurs in repeat sequences, not affecting the spacers sequences. After this process is finished, the mature crRNA is obtained (Xue and Sashital, 2019).

The last stage is interference. The PAM sequence plays a significant role in the identification of pathogens. After the pathogen is identified, the mature crRNA is guided with the help of the protein cascade or Csy complex. The interaction between these complexes and the foreign DNA destabilizes double-stranded DNA (dsDNA) (Xue and Sashital, 2019). This destabilization enables the crRNA to invade, leading to the formation of crRNA-DNA heteroduplex and the creation of the R-loop. The creation of the R-loop will then trigger the recruitment of the Cas 3 protein that will degrade the target DNA (Guo *et al*., 2017).

The Enterobacteriaceae family contains a high diversity of bacteria, responsible for a variety of diseases in humans. (Fritz *et al.,* 2005). In a previous study, 228 genomes from 38 species of bacteria were analysed. Out of these genomes, 38,6% (*Escherichia coli, Cronobacter, Citrobacter, Erwinia, Klebsiella, Proteus, Salmonella, Shigella*) contain the I-E CRISPR subtype, 14% have the I-F subtype (*Enterobacter, Erwinia, E. coli, Klebsiella, Pectobacterium, Providencia, Rahnella, Serratia, Yersinia*) and 4% have subtypes I-E and I-F (*Dickeya, Erwinia, Pectobacterium*). 0.44% of the genomes, the equivalent of one isolate, identified as *Serratia*, contains three different subtypes (Medina-Aparicio *et al*., 2018).

The study of the CRISPR-Cas system in these bacteria is essential for genotyping and developing of possible therapies for hard-to-treat infectious diseases in humans. This study investigated the presence of CRISPR loci in clinical isolates of *Pseudomonas aeruginosa* and *Escherichia coli*. The subtype of the CRISPR array and the number of spacers offer an integrative view of the activity of the CRISPR-Cas system in a particular environment exposed to permanent selective pressure exerted by antibiotics and bacteriophages.

Materials and methods

Bacterial isolates

The bacterial isolates have been collected in hospital units from patients with: urinary infections, infections resulting from wounds or burns, pneumonia, ocular infections and circulatory system infections. Patients were not involved in this study. Different bacteria have been isolated and identified in the hospital laboratory according to specific protocols, including *Citrobacter, Enterobacter,*

Hafnia, Klebsiella, Proteus, Providencia, Salmonella, Yersinia and *E. coli*. (Farkas *et al.,* 2019). Thus, 80 different isolates were analyzed for the presence of CRISPR arrays.

Molecular identification of CRISPR array

CRISPR arrays were identified by PCR amplification with different primer pairs designed in our laboratory. PCR amplification was performed with bacterial suspensions as templates. Bacterial pure cultures were suspended in sterile water to a concentration of approximately 106 cells/ml (Crăciunaș *et al,* 2010). PCR reaction mix contained in 25 µL total volume: 12.5 µL DreamTaq Green PCR master mix (2x) (Thermo Fisher Scientific, USA), 10.25 µL nuclease-free water (Lonza, Switzerland), 25 pmol each primer, and 2 µL bacterial suspension. As a negative control, we used 2 µL of sterile water in the PCR mixture. The PCR programs are shown in Table 1. PCRs were performed using a thermocycler TProfessional Trio (Analytik Jena, Germany), Mastercycler Nexus (Eppendorf AG, Germany) or Gradient Palm-Cycler (Corbett Life Science, Australia). The amplicons were separated in 1.5% w/v agarose (Cleaver Scientific, United Kingdom) gel in 1×TBE buffer (Lonza, Switzerland) and stained with 0.5 µg/ml ethidium bromide (Thermo Fisher Scientific, USA). Data acquisition was performed using the BDA Digital Compact System and BioDocAnalyze Software (Analytik Jena, Germany).

Table 1. PCR (Polymerase Chain Reaction) program used for the identification of CRISPR in clinical samples. EC represents the primers used to amplify CRISPR structures of Type I-E. PA represents the primers used to amplify CRISPR structures of Type I-F.

Step	Temperature	Time	Number of cycles
Initial denaturation	94ºC	4 min	
Denaturation	94ºC	40 s	x 35
Annealing	EC: 55ºC PA: 57ºC	20 sec	
Elongation	72ºC	2 min	
Final Elongation	72ºC	7 min	

Data analysis

Gel analysis and estimation of the number of spacers was performed further. Several numbers of base pairs were removed from the length of bands, flanking the CRISPR structure. These flanking elements are usually the genes encoding Cas proteins. In Table 2, the number of bp removed for each

primer pairs are shown. After the subtractions, the resulted number is divided to 58, the equivalent of the average size of the spacer-repeat pair, to obtain the number of spacers in CRISPR arrays in each bacterial isolate. Bands having ≤ 100 bp cannot be considered because they are too small to contain CRISPR arrays.

Table 2. Number of base pairs subtracted in the case of each primer pairs. Only primer pairs that give PCR amplification of CRISPR arrays are shown.

Primer	Size of flanking regions in bp	Size of a spacer/repeat/spacer+repeat
EC2-A	100 bp	30bp /28bp /58bp
EC2-B	94 bp	
EC3	71 bp	
PA1	138 bp	
PA2	138 bp	

Results

The amplification using the primer EC1 resulted in bands being 100 bp or smaller than that. Thus, these results were excluded when calculating the number of CRISPR spacers. Two primer pairs EC2-A and EC2-B were used to target different variants of the same CRISPR locus and to identify as many CRISPR arrays as possible. The amplifications using EC3 primers yielded a high amount of amplicons that could be used further in the study.

Regarding the PA primers, we used two sets of primers: PA1 and PA2. Amplification with the primer pair PA1 showed 1400 bp fragments, meaning CRISPR arrays with high number of spacers. CRISPR arrays are identified in 34 bacterial isolates by PCR amplification with different primer pairs (Table 3). In Table 4, the estimated number of spacers per CRISPR array is shown.

The highest number of spacers was found in samples EM1, EM3, EM13, EM14 and EM26, having 1400 bp, 23 spacers in average. These spacer sequences were found by using the PA1 primer pairs. The lowest number of spacer sequences were identified in the following samples: EM2, EM3, EM7, EM13, EM14, EM42, EM44, EM45, EM47, EM48, EM56, EM61, EM64, EM65, EM67 and EM81. The spacer sequences with 200 bp, the equivalent of 1 spacer, were identified using the EC3 primer pairs.

Figure 1 shows the distribution of isolates identified by each set of primers. By using the primer pair EC3, CRISPR arrays were identified in 22 isolates. The lowest yield can be observed in the case of EC2-B, two isolates that contain CRISPR array. This was expected because this set of primers was

designed to identify extra spacers that EC2-A set of primers could not amplify. The EC2-A primer pair showed positive results in 18 isolates, and the PA-C primer pair showed positive results in 9 isolates.

The most common length is 200 bp being the equivalent of 1.2 spacers. The downfall of these samples is that there might be a possibility that the CRISPR loci in this bacteria might be inactive. It is worth mentioning that bands over 800 bp have a high number of spacers which is correlated with an intense activity of the system, and they were identified in 11 isolates.

Table 3. CRISPR array identified in different bacterial isolates.

Sample	Size in bp			
	EC2-A	EC2-B	EC3	PA1
EM1				1400
EM2			200	
EM3			200	1400
EM4		600		
EM7			200	900
EM8				900
EM13			200	1400
EM14			200	1400
EM18	800			
EM19	800			
EM25	400		250	900
EM26	400		250	1400
EM28	500			900
EM37			400	
EM38	400	400	250	
EM39			250	
EM42	450		200	
EM44	500		200	
EM45	450		200	
EM47	450		200	
EM48			200	
EM49	450			

Sample	Size in bp			
	EC2-A	EC2-B	EC3	PA1
EM53	400			
EM56	400		200	
EM59	400			
EM61			200	
EM64			200	
EM65			200	
EM67	400		200	
EM81	400		200	

Table 4. Number of CRISPR spacer sequences identified in bacterial isolates.

Sample	Number of spacers			
	EC2-A	EC2-B	EC3	PA1
EM1				23
EM2			1	
EM3			1	23
EM4		8		
EM7			1	14
EM8				14
EM13			1	23
EM14			1	23
EM18	11			
EM19	11			
EM25	4		2	14
EM26	4		2	23
EM28	6			14
EM37			5	
EM38	4	4	2	
EM39			2	
EM42	5		1	
EM44	6		1	
EM45	5		1	
EM47	5		1	

Sample	Number of spacers			
	EC2-A	EC2-B	EC3	PA1
EM48			1	
EM49	5			
EM53	4			
EM56	4		1	
EM59	4			
EM61			1	
EM64			1	
EM65			1	
EM67	4		1	
EM81	4		1	

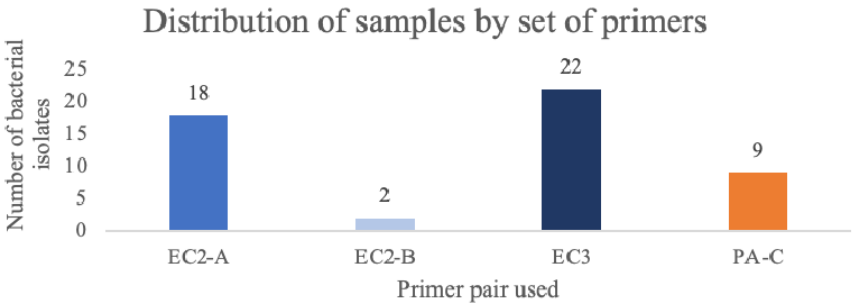

Figure 1. Bacterial isolates bearing CRISPR arrays revealed by different primer pairs.

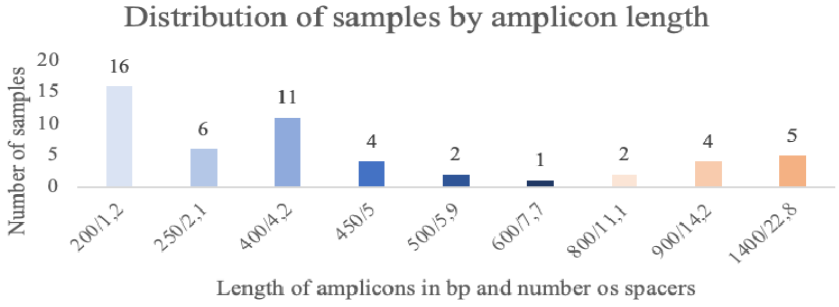

Figure 2. Distribution of samples by amplicon length.

The distribution of CRISPR subsystems between the bacteria in the clinical samples is shown in Table 5.

Table 5. Distribution of CRISPR subtypes between bacteria.

Genus	CRISPR subtype
Citrobacter	I-E
Enterobacter	I-F
Klebsiella	I-E and I-F
Proteus	I-E
Providencia	I-F
Salmonella	I-E
Yersinia	I-F
E. coli	I-E and I-F

Discussion

The CRISPR-Cas system is divided into many types. This classification has been made based on the Cas proteins that take part in the CRISPR machinery (Makarova et al., 2011b). One of the most frequent subtypes of CRISPR is the I-E subtype. This subtype was first discovered in *E. coli* and it was shown that the spacers contained are derived from bacteriophages, mobile genetic elements and plasmids (Kiro et al., 2013).

The I-F subtype was discovered in *P. aeruginosa*. An interesting discovery regarding this subtype is that most of the spacer sequences are 100% identical to mobile genetic elements that can insert themselves in the bacterial chromosome (Cady et al., 2012). Another significant discovery regarding the I-F subtype is that a slight mismatch of nucleotides between the PAM sequence and the protospacer of this subtype can lead to the inactivation of the whole system, while in I-E subtype, only the CRISPR loci would be inactive, the Cas proteins were still active and transcribed (Semenova et al., 2011).

Type I of CRISPR arrays is characterized by the presence of a protein cascade that fulfils the role of the CRISPR-Cas system. At the core of this protein cascade is the Cas 3 protein that acts as a helicase and has DNA-ase activity (Sinkunas *et al.,* 2011). Besides the Cas 3 protein, many other proteins create the protein cascade in the case of subtype I-E (Brouns et al., 2008). Finally, the protein cascade is replaced by the Csy surveillance complex in the case of subtype I-F (Xue and Sashital, 2019).

Previous studies (Mlaga *et al*., 2021; Horvath *et al*., 2008) have identified a positive correlation between the number and diversity of spacers and the activity of the CRISPR-Cas system. Thus, looking at our results, we can

identify that most CRISPR arrays in isolates amplified with the PA-1 primer set have an intense activity. The isolates having CRISPR arrays identified with EC2-A, EC2-B and EC3 primer sets have, on average, reduced activity, compared to isolates with CRISPR arrays identified with primer set PA-1.

Most bacteria contain a single subsystem of CRISPR, as identified by Medina-Aparicio et al., 2018. However, it can be noticed that some genus can have more than one subtype. This is possible because of the ability of bacteria to transfer genes horizontally.

Horizontal gene transfer (HGT) represents one of the most important tool used by bacteria to enrich their genome, especially when talking about antibiotic resistance genes (ARGs). However, active CRISPR arrays may inhibit HGT (Wheatley and MacLean, 2021). When looking at the isolates having a high number of spacers in CRISPR array we can hypothesize that this system is active. A previous study (Farkas *et al.,* 2019) showed that these isolates bear a high number of ARGs encoding the resistance to different antibiotics. Assuming that CRISPR-Cas system inhibits the HGT of ARGs, most probably in these isolates, the CRISPR-Cas system is not active at present, ensuring the accumulation of ARGs. This inactivation is most probably recent; the unuseful spacers sequences were not deleted yet. If the selection pressure of antibiotics persists (as happens in the clinical environment), some of the spacers will probably be lost. Another hypothesis is that the system is active, but the spacers sequences are accumulated by previous bacteriophage infection (Wheatley and MacLean, 2021).

Conclusions

In the Enterobacteriaceae family, CRISPR is present under two subtypes: I-E and I-F.

When looking at the length of CRISPR array and number of spacers, we identified the largest CRISPR array in isolates containing I-F subtype. We have identified spacer sequences having 1400 bp, equivalent to 23 spacers. This indicates that the I-F subsystem is active.

The I-E subtype has smaller spacer sequences. However, there are some isolates that presented larger CRISPR arrays with higher number of spacers (600 bp – 8 spacers). This might indicate that the I-E subtype might not be as active as the isolates containing the I-F subtype.

The subtype of the CRISPR array and the number of spacers offer an integrative view of the activity of the CRISPR-Cas system in a particular environment exposed to permanent selective pressure exerted by antibiotics and bacteriophages.

References

Barrangou, R. (2015). Diversity of CRISPR-Cas immune systems and molecular machines. *Genome Biol*. 16, 247, doi:10.1186/s13059-015-0816-9.

Brouns S. J. J, Jore M. M., Lundgren M., Westra E. R., Slijkhuis R. J. H., Snijders A. P. L., Dickman M. J., Makarova K. S., Koonin E. V., van der Oost J. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science*, 321(15891): 960-964.

Cady, K.C.; Bondy-Denomy, J.; Heussler, G.E.; Davidson, A.R.; O'toole G.A. (2012). The CRISPR/Cas Adaptive Immune System of *Pseudomonas aeruginosa* Mediates Resistance to Naturally Occurring and Engineered Phages. *J. Bacteriol.* 194, 5728–5738, doi:10.1128/jb.01184-12.

Carte J., Christopher R. T., Smith J. T., Olson S., Barrangou R., Moineau S., Glover C.V.C. III, Graveley B. R., Terns R. M., Terns, M. P. (2014). The three major types of CRISPR- Cas systems function independently in CRISPR RNA biogenesis in Streptococcus thermophilus. *Molecular Microbiology*. 93(1): 98-112.

Crăciunaş C, Butiuc-Keul A, Florta M, Brad A, Sigarteu M (2010) Application of molecular techniques to the study of *Pseudomonas aeruginosa* clinical isolate in Cluj-Napoca, Romania. *Ann Univ Oradea Biology* 243-247.

Devashish R., Lina A., Archana R., Magnus L. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie*. 117: 119-128.

Farkas, A.; Tarco, E.; Butiuc-Keul, A. (2019). Antibiotic resistance profiling of pathogenic Enterobacteriaceae from Cluj-Napoca, Romania. *Germs*. 9, 17–27, doi:10.18683/germs.2019.1153.

Fritz H. K., Kurt A. B., Johannes E., Rolf M. Z. (2005) Medical Microbiology. ch. 4, Thieme, pp. 278, 279.

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

Gesner, E.M.; Schellenberg, M.J.; Garside, E.L.; George, M.M.; Macmillan, A.M. (2015) Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol*. 18, 688–692, doi:10.1038/nsmb.2042.

Guo TW, Bartesaghi A, Yang H, Falconieri V, Rao P, Merk A, Eng ET, Raczkowski AM, Fox T, Earl LA, Patel DJ, Subramaniam S. (2017). Cryo-EM Structures Reveal Mechanism and Inhibition of DNA Targeting by a CRISPR-Cas Surveillance Complex. Cell. 171:414–426.

Haft D. H., Selengut J., Mongodin E. F., Nelson K.E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLOS Computational Biology*, 1(6).

Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, Moineau S, Boyaval P, Fremaux C, Barrangou R. (2008). Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol.* Feb;190(4):1401-12. doi: 10.1128/JB.01415-07.

Jansen R., Embden J. D., Gaastra W., Schouls L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*. 43(6): 1565- 1575.

Kiro, R.; Goren, M.G.; Yosef, I.; Qimron, U. (2013). CRISPR adaptation in *Escherichia coli* subtype I-E system. *Biochem. Soc. Trans*. 41, 1412–1415, doi:10.1042/bst20130109.

Kunin, V.; Sorek, R.; Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol*. 8, R61–7, doi:10.1186/gb-2007-8-4-r61.

Makarova, K.S.; Aravind, L.; Wolf, Y.I.; Koonin, E.V. (2011) Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct*. 6, 1–27, doi:10.1186/1745-6150-6-38.

Makarova, K.S.; Haft, D.H.; Barrangou, R.; Brouns, S.J.J.; Charpentier, E.; Horvath, P.; Moineau, S.; Mojica, F.J.M.; Wolf, Y.I.; Yakunin, A.F.; et al. (2011). Evolution and classification of the CRISPR–Cas systems. *Nat. Rev. Genet*. 9, 467–477, doi:10.1038/nrmicro2577.

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

Mlaga, K.D.;Garcia, V.; Colson, P.; Ruimy, R.; Rolain, J.-M.; Diene, S.M. (2021) Extensive Comparativ Genomic Analysis of *Enterococcus faecalis* and *Enterococcus faecium* Reveals a Direct Association between the Absence of CRISPR–Cas Systems, the Presence of Anti-Endonuclease (ardA) and the Acquisition of Vancomycin Resistance in *E. faecium*. *Microorganisms*. 9, 1118. https://doi.org/10.3390/microorganisms.

Mojica F. J. M., Díez-Villaseñor C., García-Martínez J., Almendros C. (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*. 155(3): 773-740.

Nam K. H., Haitjema C., Liu X., Ding F., Wang H., DeLisa M. P., Ke A. (2014). Cas5d Protein Processes Pre-crRNA and Assembles into a Cascade-like Interference Complex in Subtype I-C/Dvulg CRISPR-Cas System. *Structure*. 20(9): 1574-1584.

Palmer K. L., Gilmore M. S. (2010) Multidrug-resistant enterococci lack CRISPR-cas. mBio. vol. 1(4).

Semenova, E.; Jore, M.M.; Datsenko, K.A.; Semenova, A.; Westra, E.R.; Wanner, B.; Van Der Oost, J.; Brouns, S.J.J.; Severinov, K. (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci.* 108, 10098–10103, doi:10.1073/pnas.1104144108.

Sinkunas T., Gasiunas G., Fremaux C., Barrangou R., Horvath P., Siksnys V., (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO Journal*. 30(7): 1335-13342.

Wheatley, R.M., MacLean, R.C. (2021). CRISPR-Cas systems restrict horizontal gene transfer in *Pseudomonas aeruginosa*. *ISME J.* 15, 1420–1433. https://doi.org/10.1038/s41396-020-00860-3.

Xue C, Sashital DG. (2019). Mechanisms of Type I-E and I-F CRISPR-Cas Systems in Enterobacteriaceae. *EcoSal Plus*. 8(2):10.1128/ecosalplus.ESP-0008-2018. doi:10.1128/ecosalplus.ESP-0008-2018.

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