Application Of CRISPR-Cas in Pathogen Detection

Yunfei Wu*

Haide Institute, Ocean University of China, Qingdao, China

* Corresponding Author Email: wuyunfei@stu.ouc.edu.cn

Abstract. Infectious diseases have had a significant influence on the market economy and human civilization in recent years. Traditional nucleic acid tests typically include a lot of procedures, are complicated, take a long time to complete, and have significant space, staff, and equipment needs. More and more study is being done in this area as a result of the identification of clustered regularly interspaced short palindromic repeats and related(CRISPR-Cas) proteins in prokaryotes. The CRISPR-Cas system has been extensively used to detect viruses that cause infectious diseases due of its high specificity and sensibility. The history of the CRISPR-Cas system, including its revised categorization, is discussed in this paper. Besides, a summary of the use of various CRISPR-Cas system types in pathogen detection is also provided in the paper, along with an introduction to the CRISPR-Cas system's basic principles employing CRISPR-Cas9 as an example. As a biosciences frontier hot spot, research on CRISPR-Cas systems is developing quickly and technology based on the CRISPR-Cas system for detecting pathogens is extremely useful in clinical diagnosis and biological research.

Keywords: CRISPR-Cas, Pathogen Detection, Infectious Disease.

1. Introduction

Infectious diseases have been a huge hidden danger of economic loss and social disruption to human society, with Zika, Ebola, Middle East Respiratory Syndrome virus and the ongoing SARS-CoV-2 outbreak already posing a significant risk to humanity. Therefore, the key to monitoring and controlling infectious disease outbreaks is the quick, accurate, and affordable detection and identification of pathogens [1]. As stated by WHO standards, the ideal pathogen detection method should be sensitive, specific, rapid, low-cost, easy to use and require no large equipment. Currently, traditional pathogen detection techniques include microbial culture methods, molecular-level detection methods, serological analysis and other adaptive immune response tests [2]. Microbial culture techniques are the gold standard for identifying and detecting pathogens and were among the first diagnostic methods for infectious disorders [2]. Microbial culture methods, in which samples are inoculated into one or more culture flasks or tubes, are used to detect and identify bacteria or other culturable microorganisms. The detection of microorganisms in a patient's blood is clinically important for the diagnosis, treatment and prognosis of infectious diseases. However, conventional testing requires multiple operational steps such as media preparation, sterilization, plate pouring, multi-stage dilution, inoculation, incubation, counting, isolation and identification. The test operation is complex and the technical requirements of the operator can be high in order to obtain accurate test results. Molecular biological techniques have the advantages of strong specificity, high sensitivity and can be effectively applied to the rapid detection of pathogenic bacteria, while the disadvantage is that it is time consuming and also requires a high level of competence in terms of experimental equipment and testing personnel. The serological test is a general term for the visible reaction of antigenic antibodies in vitro and is therefore also known as the antigen-antibody reaction. It can be used to detect an unknown antigen (the bacteria to be tested) with a known antibody (bacterial antiserum) or to detect the corresponding bacterial antibody and its potency in the patient's serum with a known antigen (known pathogen), and is an important tool for clinical diagnosis, laboratory studies and bacteriological identification. Serological analysis is effective, quick, sensitive and suited for testing a large number of samples, but it frequently yields false positives, reduces the specificity of the test, takes longer to prepare the antiserum and only one virus can be tested at a time using this method.

In contrast to the methods mentioned above, CRISPR-Cas has shown great potential for pathogen detection. Prokaryotes have a special adaptive immune system to protect them from the invasion of foreign nucleic acids called clustered regularly interspaced short palindromic repeats and associated (CRISPR-Cas) proteins. CRISPR technology has a variety of technical uses, including the ability to identify pathogens, eradicate disease-causing bacteria, revive certain species and fix genetic flaws that lead to illness. At present, the CRISPR-Cas systems are divided into two major categories: the first category and the second category. The first category contains types Ⅰ, Ⅲ and Ⅳ CRISPR-Cas systems. The second category contains types Ⅱ, Ⅴ and Ⅵ CRISPR-Cas systems. Cas9, Cas12 and Cas13, which are often used for nucleic acid detection, belong to types Ⅱ, Ⅴ and Ⅵ, respectively. Researchers have also created a range of pathogen detection techniques over time using various CRISPR-Cas systems, including SHERLOCKM, HOLMES, TOETOLD, DECTECTR, etc. Based on its high sensitivity and specificity, the study of the CRISPR system is expanding quickly, nucleic acid detection is using an increasing number of Cas proteins that have been discovered.

2. Development, principles and classification of CRISPR-Cas system

2.1. The development of CRISPR-Cas systems

CRISPR is an adaptive immune defence system that has evolved over time in bacteria. When a foreign virus invades, bacteria can capture fragments of its genetic material and integrate them into their own CRISPR sequence, which in turn can be combined with Casase to precisely cut the virus' genetic sequence and effectively defend against the virus. In 1987 Ishino and colleagues at Osaka University in Japan discovered a set of 29nt repeats downstream of E. coli iap gene during their study of it, which was the first discovery of a short repeat sequence in the CRISPR-Cas system [3]. Mojica et al. in 2000 found that these short repeats were widely present in 20 species of bacteria or archaea [4]. Jansen et al. in 2002 also identified a number of genes associated with this structure near the repeats and named them Cas genes [5]. In 2005, Mojica et al., who have continued to study CRISPR, found after systematic analysis that the spacer sequence may be plasmid or viral in origin [6]. In addition, it was found that plasmids or viruses cannot infect certain bacteria or archaea that have spacer sequences with the same genetic sequence as plasmids or viruses, so it was hypothesized that CRISPR is an acquired immunity mechanism for prokaryotes. In the following year, 2012, Siksnys and Charpentier/Doudna independently demonstrated in vitro that Cas9 can be cleaved in vitro and that individual RNAs, i.e. complexes of crRNA and tracrRNA, can be equally active in vitro [7,8]. In 2013, Feng Zhang's lab achieved gene editing with CRISPR-Cas9 in eukaryotic cells for the first time [9]. After this, method for editing DNA using the CRISPR-Cas9 system has received widespread attention and research, and various applications based on it have been rapidly developed.

2.2. The updated classification of CRISPR-Cas systems

In 2015, a thorough examination and research of all recognized protein families used by CRISPR-Cas systems was carried out by Makarova et al. [10]. Makarova et al. argued that the classification of CRISPR-Cas systems should be done from a broader and more fundamental perspective. In his classification method, those using multi-protein effector complexes to achieve gene editing should be classified in one category, while those with only a few protein subunits exercising function in total should be classified in another. This downward classification led to five types and 16 subtypes, which have become the classification method now applied. The first major class of CRISPR-Cas systems includes the most commonplace and the most diverse type Ⅰ, type Ⅲ and the relatively rare type Ⅱ, where effector modules such as type Ⅰ and type Ⅲ CRISPR-Cas are multi-protein effector complexes consisting of several subunits of Cas protein, and the connection between the Cas protein effector complex and the target cut nuclease involved in CRISPR pre-crRNA processing is what distinguishes the two kinds, i.e. if the effector nuclease is a portion of the effector complex [11-13]. In the type Ⅰ CRISPR-Cas system, the effector nuclease is an HD nuclease that makes up one

domain of the Cas3 protein. Neither helicase nor the nuclease is an original component of the effector complex, but is recruited to the effector complex after successful binding to the target sequence. However, in the type Ⅲ CRISPR-Cas system, the effecting HD nuclease is a structural area of the Cas10 protein, which in the constituent effector complex [14-16]. The second major class of CRISPR-Cas systems includes the commonplace type Ⅱ (represented in type Ⅱ systems by the CRISPR-Cas9 protein system) and the less common types Ⅴ and Ⅵ. Only a single protein effector is required in type Ⅱ systems, which contain a single, multi-structural domain binding protein for crRNA, and this binding protein contains all the components needed to perform cleavage of the nucleic acid. Although Class Ⅱ CRISPR-Cas systems currently represent only 20% of the overall CRISPR-Cas systems, Class Ⅱ systems are simpler than those in Class Ⅰ and are preferred for biotechnology applications.

2.3. The principles of CRISPR-Cas systems

For instance, the CRISPR-Cas9 system's operating process entails three phases.. The CRISPR-Cas9 system in bacteria begins to work when a virus or plasmid infects the bacteria and injects genetic material into the bacteria. The first step is the acquisition and synthesis of the spacer sequence, where a small segment of DNA called the proto-spacer is intercepted from the exogenous gene and integrated into the bacterial CRISPR sequence. The process relies on the formation of a dimer of the conserved adaptive Cas1 and Cas2 proteins to perform the spacer sequence shearing. This dimer recognises the sequence with the NGG feature at the 3' end of the target gene, where N stands a base. In the CRISPR-Cas system, the NGG sequence is known as the protospacer adjacentmotif (PAM), which varies between CRISPR-Cas system, and in this case NGG is the common PAM in the CRISPR-Cas9 system. The second phase in the CRISPR-Cas9 system is the transcription of the CRISPR sequence to produce a CRISPR precursor RNA (pre-crRNA), which is then cleaved into a mature, d small-molecular crRNA with the assistance of the Cas9 protein. The Cas9 protein, a type Ⅲ RNase, and a stretch of trans-activating crRNA (tracrRNA) work together in the CRISPR-Cas9 system to shear the precursor RNA in this phase. The third step is interference with the target gene. Cas9 proteins carry tracrRNA and mature crRNA to find the pre-spacer sequence of the target gene complementary to the crRNA and bind to it to shear the target gene.

3. Application of CRISPR-Cas systems in pathogen detection

The CRISPR technology creates a new revolution in gene editing because of its high efficiency, ease of use and low cost. As the research on CRISPR system deepens, various CRISPR-based application technologies are gradually developed and perfected, and they are making a big impact in many fields, and the following is an example of the application of each type of CRISPR-Cas in pathogen detection.

3.1. Application of CRISPR-Cas9 system

Based on CRISPR-Cas9, Keith et al. were the first to create a low-cost test paper assay for RNA virus detection and typing with the use of nucleic acid-dependent amplification technology and a developed Toehold molecular switch [17]. This technique is mainly designed to target only the US ZIKA virus but not the gRNA of African ZIKA virus by using the recognition property of Cas9 protein for PAM, to accurately detect ZIKA virus and distinguish different subtypes of the virus simultaneously with a detection accuracy of 1 fm [17]. About the Toehold molecular switch, the technique operates by introducing a Toehold molecule, which, by binding to an mRNA molecule encoding a specific protein, coordinates the translation of the target mRNA molecule. At the same time, Toehold RNA also contains a sequence that binds to another trigger mRNA and acts as a trigger so that the Toehold switch is activated and the stalled protein translation process can proceed if this target mRNA sequence is found. Any gene, such as a fluorescent reporter molecule, can serve as the

trigger mRNA sequence. Researchers can visually assess whether the desired mRNA sequence is present using this fluorescent signal. In addition, Vigilant designed sgRNA targeting SARS-Cov2 N gene based on CRISPR-Cas9 [18]. Its sensitivity is 96.4% and its specificity is 100%. Furthermore, its detection limit may go as low as 2.5 copies per microliter. After isothermal amplification pre-assembly, the process may be finished in 35 minutes and is therefore practical for usage in places with limited resources.

3.2. Application of CRISPR-Cas13 system

Cas13a protein is the second protein with similar effects identified after Cas9 protein with non-targeted cleavage ssRNA activity and was the first protein to be used in the nucleic acid detection toll developed in the CRISPR-Cas13 based system. LbuCas13a protein's ability to detect target RNA was originally established by East-Seletsky et al. in 2016 [19]. However due to its poor detection sensitivity, it did not match the criteria for clinical diagnosis. The SHERLOCKv1 indirect detection method for dsDNA was created in 2017 by Gootenberg et al [20]. It combines RPA, T7 RNA polymerase and Cas13a to identify the target RNA produced by transcriptional amplification by T7 RNA polymerase. The system can specifically detect RNA of Zika virus (ZIKV) and dengue virus (DENV). At the heart of the SHERLOCK diagnostic technology is the use of Cas13a's incidental cutting properties. The sensitivity of the detection is increased when the viral RNA is isothermally amplified by RT-RPA to produce a DNA molecule containing a T7 promoter, which can then be translated by T7 RNA polymerase to produce a high number of viral RNA molecules. Designing crRNA that matches the specific region of viral RNA so that the complex of Cas13a protein and crRNA targets and cleaves the viral RNA target site. Due of Cas13a's incidental cleavage characteristics, the presence of viral RNA is characterized by the release of a corresponding fluorescent signal from the misfiring RNA reporter molecule. Because of the technical constraints of SHERLOCKv1, the researchers then further developed SHERLOCKv2, which generates multi-color fluorescent signals and enables simultaneous detection of four nucleic acids. When RPA is used in combination with LFD, SHERLOCKv2 achieves single molecule detection in two hours with a sensitivity of up to 2 amol/L [21]. Qin et al, on the other hand, examined Ebola viral RNA using microfluidic technology integrated with Cas13, 20 PFU/ml of purified viral RNA could be detected within five minutes by using a specially designed fluorometer [22]. Besides, CRISPR-Cas13 has been applied in the detection of novel coronaviruses. CARMEN is a microarray chip-based high-throughput detection system [23]. Its high-throughput detection capabilities has a sensitivity of 99.7% in the monitoring of 400 novel coronavirus samples, which might be very useful in large-scale infectious disorders. The CRISPR-Cas13a-based ERASE system, on the other hand, is simple, fast, inexpensive, and convenient. The approach employs the nucleoprotein N gene of the novel coronavirus as the target sequence and band brightness as the identification standard, which is appropriate for nucleic acid detection in large populations and makes it easier to identify the infection's origin [24].

3.3. Application of CRISPR-Cas12 system

The above-discussed nucleic acid detection techniques based on CRISPR-Cas9 frequently require the labor-intensive combination with nucleic acid amplification technology currently in use for target identification. As a result of Cas13a's propensity for targeting RNA instead of DNA, nucleic acid detection methods built on the CRISPR-Cas13 system can only directly detect the target DNA. However, with the discovery of Cas12 effector proteins, direct DNA detection has also become possible [25]. Using PCR technology intergrated into the CRISPR-Cas12a system, Li et al. [26] created the one-hour low-cost multipurpose highly efficient system (HOLMES) by using the collateral cleavage capabilities of Cas12a. With a sensitivity of 1-10 amol/L, they were able to identify the pseudorabies virus and the Japanese encephalitis virus using fluorescent reporter/quencher-labeled ssDNA. However, this system separates PCR amplification of the target sequence and Cas12a detection into a two-step reaction with numerous operational steps.

Streamlining the HOLMES system's operational procedures, Li et al. developed a second-generation HOLMES system based on the first-generation HOLMES system by using Cas12b protein and coupling it to RT-LAMP amplification technology [27]. The system enables a one-step reaction for target amplification and signal detection, making the operational steps easier. The CRISPR-Cas12a technique has been used to identify novel coronaviruses due of the incidence of novel coronavirus-2019. When the sample contains very little of the target gene, the probability of Cas12a matching the gRNA complex to the target sequence to be detected is very low. PCR is a common amplification tool for this purpose, but requires a special PCR instrument for thermal amplification reactions. Another signal amplification technique called RPA allows the signal to be amplified at a constant temperature without the need for complex warming and cooling processes. Professor Doudna's innovative combination of Cas12a's targeted cleavage of single-stranded DNA and RPA has created the method of DETECTR. Joung et al., on the other hand, introduced in their report the SHERLOCK technique, which uses a magnetic bead method to simplify the nucleic acid and reduces detection time using a lateral flow system [28]. Combining the ring-mediated isothermal amplification technique and the CRISPR-Cas12b system, this technique allows the detection of novel coronavirus in 1 hour with a sensitivity of 93.1% and a specificity of 98.5%.

4. Conclusion

The use of Cas9, Cas12 and Cas13 in CRISPR-Cas for pathogen identification is described in this paper, ideas like SHERLOCK and Toehold are also briefly discussed. The aforementioned uses of the CRISPR-Cas system in pathogen identification show that, in addition to being a powerful tool for gene editing, this technology has a lot of promise for use and advancement. At present, the application of CRISPR-Cas in pathogen detection is mainly characterized by limitations in target sequence, dependence on targeted amplification techniques and off-target effects. However, as further research on this technology it is believed that in the future, solutions to these limitations will be developed one by one. Traditional nucleic acid tests generally have many steps, are complex, take a long time to perform and have high requirements for space, personnel and equipment. Therefore, the development of pathogen detection technologies that are sensitive, specific, rapid, low-cost, easy to use and do not require large equipment is of great significance for clinical diagnosis and biological research. As a frontier hot spot in the biosciences, more and more researches are performed on the basis of CRISPR-Cas systems. The use of CRISPR-Cas in pathogen identification is expected to advance as more studies are undertaken and additional Cas proteins are discovered.

References

[1] Li JW, Wang YX, Wang B, Lou J, Ni P, Jin YF, Chen SY, Duan GC, Zhang RG. Application of CRISPR-Cas systems in the Nucleic Acid Detection of Infectious Diseases [J]. Diagnostics, 2022, 12(10): 2455.

[2] Gao HD, Shang ZF, Siew YC, Ma DL. Recent advances in the use of the CRISPR-Cas system for the detection of infectious pathogens [J]. Journal of Zhejiang University-SCIENCE B, 2022, 23(11):881-899.

[3] Ishino Y, Shinagawa H, Makino K, et al. Nucleotide sequence of the iap gene, responsible for alkaline phosphataseisozyme conversion in Escherichia coli, and identification of the gene product [J]. Journal of Bacteriology, 1987, 168(12):5429-5433.

[4] Mojica FJ, Diez-Villasenor C Soria E, et al. biological significance of a family of regularly spaced repeats in the genomes of archaea, bacteria and mitochondria [J] Mol Micro-biol, 2000, 36(1):244-246.

[5] Jansen R, Embden JD, Gaastra W, et al. Identification of a novel family of sequence repeats among prokaryotes [J]. OMICS, 2002, 6(1):23-33.

[6] Mojica FJ, Diez-Villasensor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements [J]. J Mol Evol, 2005, 60:174-182.

[7] Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage foradaptive immunity in bacteria [J]. Proc Natl Acad Sci U S A, 2012, 109:E2579-2586.

[8] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna, JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity [J]. Science, 2012, 337:816-821.

[9] Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo, JE. RNA-guided human genome engineering via Cas9 [J]. Science, 2013. 339(6121): 823-826.

[10] Makarova KS, Wolf YI, Alkhnbashi OS, et al. Anupdated evolutionary classification of CRISPR-Cas systems [J]. Nat Rev Microbiol, 2015, 13(11):722-736.

[11] Beloglazova N, Petit P, Flick R, et al. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference [J]. EMBO J, 2011, 30:4616-4627.

[12] Xiao Y, Luo M, Dolan AE, et al. Structure basis for RNA-guided DNA deg-radation by Cascade and Cas3 [J]. Science, 2018, 361.

[13] Westra ER, van Erp PB, Kunne T, et al. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3 [J]. Mol cell, 2012, 46:595-605.

[14] Jung TY, A Y, Park KH, et al. Crystal structure of the Csm1 subunit of the Csm complex and its single-stranded DNA-specific nuclease activity [J]. Structure, 2015, 23:782-790

[15] Zhang J, Graham S, Tello A, et al. Mutiple nucleic acid cleavage modes in divergent type Ⅲ CRISPR systems [J]. Nucleic Acids Res, 2016, 44:1789-1799.

[16] Liu TY, Lavarone AT, Doudna JA. RNA and DNA targeting by a reconstituted Thermus thermophilus type Ⅲ-A CRISPR-Cas system [J]. PLoS One, 2017, 12:e0170552.

[17] Feng WY, Meng LN, Chen C, He LZ, Qin GS, Shi DS. Research progress in CRISPR-Cas system classification and pathogen detection [J]. Biotechnology, 2022, 32(02):258-267.

[18] Marsic T, Ali Z, Tehseen M, Mahas A, Hamdan S, Mahfouz M. Vigilant: An Engineered VirD2-Cas9 Complex for Lateral Flow Assay-Based Detection of SARS-CoV2-2 [J]. Nano Lett, 2021, 21:3596-3603.

[19] East-Seletsky A, O'Connell MR, Knight SC, et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection [J]. Nature, 2016, 538(7624):270-263.

[20] Gootenberg JS, Abudayyeh OO, Lee JW, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2 [J]. Science, 2017, 356(6336):438-442.

[21] Gootenberg JS, Abudayyeh OO, Kellner MJ, et al. multi-plexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6 [J]. Science, 2018, 360(6387):439-444.

[22] Qin PW, Park M, Alfson KJ, et al. Rapid and fully microfluidic ebola virus detection with CRISPR-Cas13a [J]. ACS Sens, 2019, 4(4):1048-1054.

[23] Ackerman CM, Myhrvold C, Thakku SG, Freije CA, Metsky HC, Yang DK, Ye SH, Boehm SK, Kosoko-Thoroddsen TF, Kehe J, et al. Massively mutiplexed nucleic acid detection with Cas13 [J]. Nature, 2020, 582:277-282.

[24] Li H, Dong X, Wang Y, Yang L, Cai L, Zhang X, Kou Z, He L, Sun S, Li T, et al. Sensitive and Easy-Read CRISPR Strip for COVID-19 Rapid Point-of-Care Testing. The CRISPR Journal, 2021, 4(3): 392-399.

[25] Gao HD, Shang ZF, Zeng YY, Ma DL. Research progress of CRISPR-Cas system for detection of infectious disease pathogens [J]. Journal of Zhejiang University-Science B, 2022, 23(11):881-899.

[26] Ackerman C, Myhrvold C, Boehm C, et al. Programmable inhibition and detection of RNA viruses using Cas13 [J]. Molecule Cell, 2019, 76(5):826-837.

[27] Li SY, Cheng QX, Wang JM, et al. CRISPR-Cas12a-assisted nucleic acid detection [J]. Cell Discovery, 2018, 4(1):1255-1266.

[28] Joung J, Ladha A, Satio M, et al. Detection of SARS-CoV-2 with SHERLOCK one-pot testing [J]. N Engl J Med, 2020, 383(15):1492-1494.