Proceedings of the 3rd International Conference on Biological Engineering and Medical Science
DOI: 10.54254/2753-8818/23/20231012

Research progress of high precision single base editing technology with a focus on prime editing

Jiaxin Zhang*

School of Life Science and Engineering, Southwest Jiaotong University, Chengdu, 610031, China

jzhang93@student.gsu.edu

Abstract. Gene editing is a technique in genetic engineering that aims to change specific genes in the genome of a certain organism mainly by modifying specific genes, such as by adding or deleting them. It has a profound meaning in the treatment of human genetic ailments and in the modification of diseased genes in plants and animals, and is also one of the hottest research areas in the life sciences in recent years. This paper focuses on the development of CRISPR/Cas9-based single-base editing technologies after 2019, especially after the advent of lead editing. Researchers in various countries have applied it to rice and mice for experiments to ameliorate the accuracy of single-base editing and decrease off-target effects, and have tested the accuracy and conversion rate of prime editing (PE) by whole-genome sequencing and other methods, and have continued to innovate on this basis to pursue more efficient editing results. This paper summarizes the problems that still need to be solved for PE, so that it can serve as a reference in the future research process and better develop prime editor and gene editing.

Keywords: CRISPR/Cas9, single-base editing, prime editing, CBE, ABE.

1. Introduction

The revolution in molecular biology that began in the mid-20th century has led to breakthroughs in deciphering the genetic code of genes, and scientists have sought to further modify genes to fundamentally cure diseases, and as theories have accumulated, the application of gene editing has become increasingly apparent. "Gene editing" describes the precise modification of gene sequences made possible by the use of enzymes to cut DNA strands, remove damaged DNA, or insert new DNA. It has the ability to "cut" and "repair" genes. To achieve genetic therapy, it is possible to change and modify the DNA sequence in a way that interferes with the operation of poisonous or suppressive genes (or restores the function of necessary genes).

As science and technology continue to advance, gene editing technologies are becoming more and more sophisticated, and editing the genome is becoming easier and easier, especially the third-generation technology Clustered regularly interspaced short palindromic repeats (CRISPR)/ CRISPR-associated systems (Cas), which is now considered the simplest and most efficient means of gene editing. The CRISPR/Cas system is an immune system found in prokaryotes (including bacteria and arhcea). Some prokaryotes, after being intruded by a virus, have the ability to remember a part of the virus gene in a space, which is called CRISPR, in their own genome. When the same virus attacks prokaryotes the second time, it is able to identify the virus on the previously remembered clip basis

DOI: 10.54254/2753-8818/23/20231012

and incapacitate it by cutting away its DNA segment. This system contains both the CRISPR motif and the Cas gene. The CRISPR sequence of this defensive system expresses RNAs that are sequence-identical to the invasive genome in response to the invasion of an exogenous gene, and the Cas then cleaves the exogenous genomic DNA at the sequence recognition, achieving the protective goal. After the emergence of CRISPR technology represented by Cas9, gene editing has become extremely easy and the field of gene editing has begun to flourish. Among the many directions of CRISPR applications, one of the most imaginative is single-base editing. A precise and effective method for achieving single base replacement is essential since single base variation is the genetic cause of around two-thirds of human genetic illnesses and is the genetic basis for variation in many significant features in plants and animals.

By creating specific DSBs in the DNA and then depending on the cell's own repair processes, gene editing tools including ZFN, TALEN, and CRISPR/Cas finish the editing process. Such approaches can effectively disrupt gene expression. However, their lack of control over the editing outcome, off-target effects, and reliance on DNA double-strand breaks can lead to unanticipated disruptions in cells after gene editing. Many genomic mutations occur in a single base, and to make gene editing more precise, base gene editing techniques were created to target these single base errors without leading DSBs in the DNA.

In 2016, David. R. Liu's team first introduced the single-base editing technology on the basis of CRISPR/Cas9 and proposed the Cytosine Base Editor (CBE), a kind of single-base editor that enables to transform cytosine-guanine base pairs to adenine-thymine base pairs [1]. Then in 2017, David's team developed another single-base editing tool called Adenine Base Editor (ABE), which can make the transformation of C-G→A-T base pairs [2]. Then, Chinese scientists Hui Yang and Caixia Gao reported the first findings of severe off-target effects of single-base editing systems in different species by using different methods. The two studies confirmed in mice and rice, respectively, that the single-base editing system has severe off-target effects and also induces a large number of DNA mutations [3,4]. Among the known human pathogenic mutations, point mutations, i.e., single base mutations, cause the highest proportion of diseases [5], and according to ClinVar statistics [6], base mutations are responsible for around 58% of human genetic disorders. Therefore, effective and precise single base mutation repair is crucial for the understanding and management of genetic illnesses.

In the following part of this paper, it will look at the following aspects, the definition of PE, the applications of PE, the development, limitations and future outlook of prime editing, and finally a summary will be made.

2. Prime editing

In October 2019, a paper published in Nature by David Liu's team indicated that they had developed a Prime Editor (PE), a gene editor capable of searching and replacing bases. All 12 base conversions, including cytosine to thymine, cytosine to adenine, cytosine to guanine, thymine to cytosine, thymine to guanine, thymine to adenine, guanine to adenine, guanine to cytosine, guanine to thymine, adenine to guanine, adenine to cytosine, adenine to thymine can be efficiently implemented without relying on DSB (DNA double-strand break) and highly DNA template of the same origin, it could also achieve accurate insertion of more than 44 bases or accurate deletion of more than 80 bases at the same time [7].

The major benefit of PE over David Liu's earlier CBE and ABE involved terrific site editing and DNA clip insertion and deletion (less than 100 bp). Compared with traditional Homologous Directed Recombination (HDR), PE is more efficient and safer because it can edit without DSB. Prime editing has better performance at all known gene editing sites, i.e., fewer off-target effects, compared to Cas9, and much less HDR compared to Cas9 initiation, and fewer by-products, and is more efficient or similar to it, with the advantages of prime editing complementing the disadvantages of base editing [7].

The first core of the Prime editor is that it is not simply a CRISPR/Cas9 construct, its protein is a composite of a modified Cas9 protein and reverse transcriptase, and the second core is that it is also not a normal guide RNA, but uses a prime editing guide RNA (pegRNA). A special feature of pegRNA is that it embodies a spacer region specifying the goal region, a single guide RNA (sgRNA) backbone and

DOI: 10.54254/2753-8818/23/20231012

a 3' stretch encoding the desired edit. pegRNA causes the new fusion protein to precisely cleave the specified DNA site, after which the template is applied to synthesize the new DNA strand, and the newly created sequence will be automatically incorporated into the genome by cells. The original sequence becomes redundant and is removed by the repair mechanism. Thus, the gene editing function has been "taken over". For instance, PE1 first copies genetic data from elongation of pegRNA to the target genomic region using reverse transcriptase (RT) coupled to RNA programmable nickase and pegRNA [7]. PE3 further increases editing efficiency by marking non-edited strands to promote their replacement, which may be done in human HEK293T cells up to 20–50% and produce 1–10% insertion and deletion mutations (indels) [7]. PE2 employs modified reverse transcriptase to improve editing efficiency [7]. HEK293T cells are often used to construct models of early cancer development because have some characteristics that can represent renal epithelial cells and are constructed by co-culturing models of normal and cancer cells [8].

After this, research on lead editing is still in full swing. David R.Liu et al. systematically analyzed and compared the risk of Cas9 non-dependent DNA off-target for multiple CBE systems based on bacterial resistance screening methods and found that multiple mutants of APOBEC1 such as YE1, YE2, EE, YEE, R33A and R33A/R34A had the lowest risk of Cas9 non-dependent DNA off-targeting, and found that YE1 was the best version with unchanged base conversion efficiency and reduced both DNA and RNA off-targeting [9]. Subsequently, the researchers also optimized the CBE system by protein engineering and other strategies to effectively reduce the occurrence of Cas9 non-dependent DNA off-target risk [9]. Members of the APOBEC family play an important role in innate antiviral immunity, such as limiting the expression of HIV-1, HBV and HPV 5-7. However, these enzymes can also deaminate cytosine in the host genome and produce cytosine to thymine conversion and cytosine to guanine conversion in the TCW motif, i.e., APOBEC mutagenesis. Nowadays, an increasing number of researchers have concentrated on the role of APOBEC family members and APOBEC mutagenesis in tumors. mutational patterns of APOBEC are widespread in human cancers and are associated with cancer progression and malignant progression.

Also in 2020, Feng Yang's team described the screening of 23 rationally designed CBE variants and found that mutant residues in the anticipated DNA binding sites have the potential to significantly lower the Cas9-independent off-target effect. In addition, one of the CBE variants, CBE variant-YE1-BE3-FNLS was obtained, which not only maintained high target editing efficiency but also could produce extremely low off-target effects [10]. The outcomes of this research are generally consistent with the findings published by David Liu's team, both of which reported that YE1 could reduce off-target effects on both DNA and RNA while maintaining high editing efficiency, meanwhile causing narrower editing window and reducing the proportion of indels produced [9,10]. David Liu's team's method based on bacterial resistance screening is only applicable to CBE, while the Hui Yang team's method based on GOTI (genome-wide off-target analysis by two-cell embryo injection) is unrestricted. Because it is possible to visually compare the edited and unedited cells even without knowledge of genes, it has a high sensitivity for the detection of probable off-target variations, so this method can not only detect off-target effects of single-base editing, but also be used for safety testing and improvement of other fusion protein-based gene editing tools [10,11].

In 2021, Caixia Gao's team’s research provides an in-depth and systematic assessment of the off-target impacts of the PE system at both the plant cell and individual plant levels [12]. High genome-wide editing specificity and no off-target effects are produced by the prime editing mechanism. This study is the first comprehensive analysis at the genome-wide phase of the off-target effects of the PE system, confirming the high specificity of the prime editing so that providing confidence for further applications of this novel gene editor.

David Liu's team has further improved the Prime Editor. Their paper mentions that although prime editing can cut DNA with precision, little is known about the cellular factors that affect editing efficiency [13]. In the study, they conducted a large-scale screening using CRISPR interference (CRISPRi)-based technology and found that mismatch repair of DNA affects prime editing, producing unintended insertion/deletion mutations. In response, they developed more advanced prime editing

DOI: 10.54254/2753-8818/23/20231012

systems that transiently express proteins that inhibit DNA mismatch repair to improve editing efficiency. PE4 and PE5 were invented as new prime editors, and they stimulated a gene named MLH1 dominant negative (MLH1dn) in a short time, so PE4 (PE2+MLH1dn) and PE5 (PE3+MLH1dn) were generated [13]. Not only did their presence allow for a significant increase in gene editing efficiency, but the rate of Indel production was also reduced. Now PE5 can achieve 30-60% editing efficiency in Primary T Cells. It should be noted that PE5 still generates a lot of Indel because it needs to cut the opposite single strand. In contrast, PE4, which does not cut the opposite single chain, has an Indel rate of less than 1%, but the editing efficiency will be sacrificed. It was previously found that the editing efficiency of PE2/PE3 in HEK293T cells is usually more than it is in HeLa, K562 and U2OS cells [7] because MLH1 gene has promoter in HEK293T cells has been highly methylated and suppressed its expression, while MLH1 is normally expressed in several other cells. Since mismatch repair (MMR) is a very important pathway for DNA repair, especially for suppressing microsatellite instability (congenital deletion of a single copy of MMR greatly increases the incidence of colorectal cancer and many other cancers), and for repairing single base mutations caused by misreplication (about 100-fold less likely to occur). Therefore, transient expression of MLH1dn) gene is risky. However, the study found no detectable effect on microsatellite instability if expressed only transiently for a few days.

The researchers also found that the MMR pathway is not very sensitive to all single-base mutations and is less likely to detect guanine to cytosine single-base mutations, so using the original PE2/PE3 to achieve this edit is generally more efficient [13]. Taking advantage of this, the researchers could selectively introduce some silent or benign mutation into the vicinity of the target editing site to generate a relatively insensitive mutation in MMR, thus achieving high editing efficiency even with the original PE2/PE3. Finally, they further optimized the Cas9 protein, entry signal, linkage composition, and codon, together with the previously optimized epegRNA to improve the editing productiveness of the PE4/PE5, named PE4max/PE5max [13].

Prime editing has been widely used since its introduction, yet there is a bottleneck that has not yet been broken—it can only delete genes of limited length. If a sequence exceeds 100 bases, it would probably reduce how effective this technique is at editing. Tingting Jiang et al., Junhong Choi and so on break this bottleneck. Two independent teams have each developed optimized prime editing systems that can precisely excise up to 10,000 bases at once [14,15]. This greatly expands the application scenario of prime editing.

The first study came from Professor Wen Xue's group. Unlike the modified Cas9 protein used in the prime edit, they used a fully functional Cas9 protein, which was then fused with reverse transcriptase [14]. Such a fusion protein can cut two DNA strands at a time. At the same time, they also added an additional pegRNA (two in total) that targets the complementary sequence on the DNA. In turn, the two new DNA segments synthesized by reverse transcriptase from complementary sticky ends. When these two ends are combined, the large segment of DNA that was in the middle can be removed [14]. This technique can delete up to nearly 10,000 bases. The second study came from the group of Professor Jay Shendure. Unlike the first study, the researchers used the same special Cas9 protein that was originally used in the prime edit, which cuts through DNA one at a time, but with two pegRNAs, it was also able to delete intermediate sequences [15]. In human kidney cells, the researchers did a series of deletion tests ranging from 20 bases to 10,000 bases [15]. Although the editing efficiency is not yet very high, the positive results of these experiments validate the feasibility of this technique. More critically, the success of editing does not depend on the length of the deleted sequence, but more on the sequence selected. This also further demonstrates the feasibility of performing large DNA deletions.

David Liu's team has recently deepened their research on prime editing by creating an optimized dual AAV delivery system that achieves unprecedentedly efficient in vivo prime editing, achieving very high editing efficiencies of 42% in the brain and 46% in the liver of experimental animals (rat), increasing the efficiency of in prime editing by approximately 10-fold and achieving a breakthrough in the exploration of genetic diseases in humans [16].

DOI: 10.54254/2753-8818/23/20231012

3. Limitations, challenges and prospects

Although current improved methods reduce the potential off-target effects, it is difficult to pursue higher editing efficiency and wider editing windows. Existing base editors can only catalyze some types of bases, which limits their wide application over a longer period of time and to a large extent.

Each step of the procedure must be completed flawlessly to attain prime editing. Scientists will first produce the prime editor, which will then go to work in the nucleus to form a complex with the pegRNA and finally complete the ligation to the target DNA. After the successful completion of the steps above, the prime editor can break the specific DNA strand, which is usually exposed, and begin using pegRNA to reverse transcription. Afterwards, a cellular DNA repair process is required to copy the processed DNA segment to the unedited strand, allowing the editing to continue. Attempting to increase the effectiveness of prime editing may necessitate concentrating on the following issues.

Single-base editing leads to a large number of unpredictable off-target mutations. Since its introduction, there has been debate concerning the real off-target rate of CRISPR/Cas9 and its variants, and the risk of off-target has never been adequately and effectively detected, among which, there are even fewer studies on the off-target effect after single-base editing. Previous protocols for detecting off-target have limitations in detecting off-target mutations, especially single nucleotide mutations. Once off-target effects occur in the clinic, they may cause a variety of side effects, including cancer.

Single-base editing has limited applicability with a single edit window. If there are numerous cytosines or adenines in the edit window or close by, this can result in unwanted base conversions. If a base editor is used to correct single-nucleotide polymorphism (SNPs), any other nucleotide within the same edit window should be avoided except for the target nucleotide to be edited. Therefore, precise implementation of single-base editing is of clinical importance.

Due to the preferential differences in the action sites of these base editors, there are no studies that have systematically compared the efficacy of these editors, while the human disease-associated single-base mutation sites on which different base editors can act have not been consolidated and summarized in the domain of gene editing.

Up to now, the advance of base gene editing technology is continuously updated, and single-base editing technology provides more references for the progression direction of gene editing tools and their wide application in various fields, providing new thoughts for further development and utilization of base gene editing tools for human beings.

CRISPR technology has become a must-have gene editing tool in every biology laboratory and is also being actively translated into the clinical medical field, forming a market with great potential. In the medical field, safety always comes first, which requires CRISPR to achieve accurate and efficient gene editing, and it is in recent years that new CRISPR technologies like base editing have emerged and made great progress in a short period of time, providing strong and continuous momentum for gene editing as a major direction in biomedicine.

4. Conclusion

The precise manipulation of the genome in a programmable manner has transformed life science research. Compared to other gene editing technologies, prime editing is more flexible and produces products with higher purity. Prime editing has great potential to drive many innovative studies, improve the efficiency and capability of gene editing, and demonstrate its applications and unlimited possibilities in basic research and therapeutics.

New base editors are full of opportunities and challenges. How to significantly enhance the base editors' editing effectiveness, purity of products and specificity of base editors, balance the editing range and editing window size, and develop safer and more efficient editing tools with wider applicability are the urgent technical problems to be solved in this field.

In the end, any science and technology need to serve all human beings, and people all look forward to the clinical use of technologies like gene editing to treat diseases that were once untreatable. But it should not be forgotten that any therapy used in the clinic must be very clear about its risks. With more accurate and sensitive detection technology, it will be possible to develop safer gene editing tools in the

future. It is hoped that the above work will help to develop industry standards as soon as possible, so that this field can develop in an orderly and healthy way.

References

[1] Komor, A., Kim, Y., Packer, M. et al. "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage." Nature 533, 420-44 (2016).
[2] Gaudelli, N., Komor, A., Rees, H. et al. "Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage." Nature 551, 464–471 (2017).
[3] Zuo, E., Sun, Y., Wei, W. et al. "Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos." Science 364(6437), 289-292 (2019).
[4] Jin, S., Zong,Y., Gao,Q. et al. "Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice." Science 364(6437), 292-295 (2019).
[5] Chen, Y., Chen, YX., Ping, Y. "Application of base editing systems in the treatment of genetic diseases." Journal of Pharmaceutical Sciences, 55(07):1562-1572 (2020).
[6] Landrum, J., Lee, M., Benson, M. et al. "ClinVar: public archive of interpretations of clinically relevant variants." Nucleic Acids Res 44(D1): D862-8 (2016).
[7] Anzalone, A.V., Randolph, P.B., Davis, J.R. et al. "Search-and-replace genome editing without double-strand breaks or donor DNA." Nature 576, 149–157 (2019).
[8] Chung J., et al. "Molecular changes in appearance of a cancer cell among normal HEK293T cells." Journal of Bioscience and Bioengineering 123, 3 (2017).
[9] Doman, J.L., Raguram, A., Newby, G.A. et al. "Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors." Nat Biotechnol 38, 620–628 (2020).
[10] Zuo, E., Sun, Y., Yuan, T. et al. "A rationally engineered cytosine base editor retains high on-target activity while reducing both DNA and RNA off-target effects." Nature Methods, 17(6), 600-604 (2020).
[11] Zuo, E., Sun, Y., Wei, W. et al. "GOTI, a method to identify genome-wide off-target effects of genome editing in mouse embryos." Nat Protoc 15, 3009–3029 (2020).
[12] Jin, S., Lin, Q., Luo, Y. et al. "Genome-wide specificity of prime editors in plants." Nat Biotechnol 39, 1292–1299 (2021).
[13] Chen, P., Hussmann, J. et al. "Enhanced prime editing systems by manipulating cellular determinants of editing outcomes." Cell. 184(22):5635-5652.e29 (2021).
[14] Jiang, T., Zhang, XO., Weng, Z. et al. "Deletion and replacement of long genomic sequences using prime editing." Nat Biotechnol 40, 227–234 (2022).
[15] Choi, J., Chen, W., Suiter, C.C. et al. "Precise genomic deletions using paired prime editing." Nat Biotechnol 40, 218–226 (2022).
[16] Davis, J.R., Banskota, S., Levy, J.M. et al. "Efficient prime editing in mouse brain, liver and heart with dual AAVs." Nat Biotechnol (2023).