The Challenge Facing CRISPR/Cas9 System: Off-Target Effects and Their Optimization

Haotian Sun
Department of Biological Sciences, University of Bristol, Bristol, United Kingdom
wi21257@bristol.ac.uk

Abstract. The field of genome editing has undergone a profound revolution with the emergence of CRISPR-Cas9 technology, which enables precise modifications to the genetic code. However, a critical concern is the possibility for the appearance of off-target, where the modifications induced by Cas9 nuclease are at non-intended targets. The mismatches of the seed sequence with the single guide RNA (sgRNA) and the inappropriate length of it could induce off-target effects. Moreover, the inflammatory response triggered by virus-mediated delivery methods may also be responsible for off-target. For the expected events, a large number of deletions of the sequences of the targeted sequence induced by CRISPR-Cas9 may bring the uncertainty of safety. The random changes of genes induced by CRISPR-Cas9 may affect future generations through gene drives, which still lacks long-term studies on long-lived organisms. The off-target effects could be detected by some sequencing methods such as GUIDE-seq, and ITR-seq; however, the suitable ways are situation-dependent. Eventually, the specificity of the technique may be improved through the predictions implemented by Deep Learning, enhancing fidelity and lowering the binding affinity of the Cas9, binding of dual sgRNAs, and delivery method of RNP complexes. This review introduced the causes, effects, and detection methods of the off-target and further discussed of several ways that can mitigate the off-target effects.

Keywords: CRISPR/Cas9 system, off-target effect, optimization.

1. Introduction

The introduction of CRISPR-Cas9 technology has been influential and significance in genome editing, essentially altering the landscape of molecular biology and biomedical research. Derived from a prokaryotic immune system, the CRISPR-Cas9 system is considered an unparalleled apparatus in the manipulation of genetic information with remarkable efficiency.

The CRISPR-Cas9 system comprises two distinct components: CRISPR region, and the Cas9. By definition, besides the discontinuous repeats, there are also incorporated spacers originating from external genetic entities. During an infection, prokaryotic cells possess the capability to integrate concise DNA fragments from external sources through the cas1-cas2 complex, referred to as protospacers, into the CRISPR array, where they function as spacers. This process is called adaptation. When the cells encounter the same or similar foreign invaders, the spacers will be used for the recognition and targeting of the foreigner’s DNA for destruction. The process of targeting will involve the family of Cas proteins which are nuclease, serving as cutting tools. Furthermore, the sgRNA is a combined RNA that is the combination of tracrRNA and crRNA. The sgRNA will locate the cas9 to the intended position. In this process, the PAM targeted by cas9 is adjacent to the target sequence, which is a critical sequence for the initiation of gene editing. Then, the interaction between the sgRNA-Cas9 complex and the intended gene causes the initiation of a DSB, which triggers the cell-repairing pathways. The repairing machinery will implement either error-prone end joining or homologous recombination depending on demand [1].

Owing to its heightened efficiency, cost-effectiveness, and simplified design in contrast to the initial two generations of gene-editing methodologies, this technique has garnered extensive utilization for both the investigation of diseases and the advancement of therapeutic interventions for genetic disorders. For example, the gene editing in hematopoietic stem and progenitor cells (HSPCs) resulted in the acquisition of HIV-1 resistance among patients [1]. Moreover, CRISPR-Cas9 allowed
the researchers to establish the organism models of various diseases such as haemophilia and diabetes, which provided scientists insights for relevant therapies [1].

Despite the broad prospect of CRISPR, there are some challenges have arisen. One crucial concern is that the cas9 guided by the sgRNA may unintendedly bind and cleave the non-target sequence, which may cause some insertions and deletions (indels), unpredicted outcomes, severe situations such as loss of chromosome, the development of the oncogenic cells due to the activation of p53 [1]. Therefore, due to the potentially detrimental influences resulting from off-target effects, minimizing the off-get effects and improving precise targeting are significant. In the review, the causes of the off-target effects will be introduced in the first place. Then, the catastrophic events induced by off-target effects will be discussed, followed by the efficient ways to detect the off-target effects. Eventually, other strategies that help to mitigate the off-target effects will be analyzed and discussed respectively.

2. Off-target effects of the CRISPR/Cas9 system

2.1. Factors affecting off-target of CRISPR/Cas9 systems

2.1.1 Effects of sgRNA

Within the framework of the CRISPR-Cas9 system, the design of sgRNAs is orchestrated to establish a state of complementarity with the aimed sequence, thereby facilitating the precise guidance of the Cas9 enzyme to the anticipated loci. However, the presence of genomic regions bearing sequences closely resembling the target site introduces the possibility of off-target effects. In such instances, the sgRNA may inadvertently guide Cas9 to these non-target sites, leading to unintended cleavage events and consequential off-target effects. Furthermore, the precision of the CRISPR/Cas9 system is chiefly dictated by the arrangement of the sgRNA, specifically highlighting the significance of the seed sequence spanning about 10 nucleotides adjoining the PAM-proximal region. This seed sequence plays an essential role in the recognition of target sites and binding to the matching sequence. When the sgRNA sequence identifies partial mismatches outside the seed sequence, rather than at the intended on-target site, the editing of the off-target sites will occur [2]. Moreover, the length of the sgRNA could potentially contribute to the emergence of off-target. To elaborate further, in the editing of stem cells, sporadic occurrences of off-target effects were noted with the utilization of 17-nucleotide (nt) sgRNAs, whereas such effects were absent when employing 20-nt sgRNAs. The truncated sgRNAs (17 or 18 bp) in human cells could increase the specificity of targeting by more than 500-fold [2].

2.1.2 Effects of delivery methods

The delivery method used to introduce Cas9 and sgRNA into cells can influence the specificity of the CRISPR-Cas9 system. Different delivery methods may result in varying off-target profiles. To be specific, Adenoviruses (AVs) have gained widespread application in clinical trials as vectors for gene delivery. One remarkable aspect of AVs is their capacity to transduce liver cells and neurons while maintaining the critical characteristic of not integrating into the host cell genomes. Nevertheless, the predominant challenge associated with AV-based delivery pertains to the robust induction of innate immune responses in host cells, which triggers a pronounced inflammatory reaction in tissues, culminating in the subsequent clearance of AV vectors [3]. The induced inflammatory response will release several immune components such as cytosines that may interfere with the interaction of Cas9 with sgRNA, decreasing the specificity of on-target. Furthermore, Lentiviruses (LVs) are considered another vector for the delivery of genes. However, the DNA form of this viral genome will be integrated into the host cell’s genome through the activity of viral integrase. The random integrations can introduce genetic changes at sites that are not intended targets, causing off-target effects [4].

2.2. The unexpected events arisen by off-target effects

In 2017, a critical apprehension about the accuracy of the CRISPR-Cas9 system was triggered by a study [5]. It investigated the consequences of CRISPR-Cas9 by sequencing seventeen targeted sites
of 632 founder mice. It was found that one sgRNA can result in up to 600 deletions of kilobase pairs (kb). Only one sgRNA was injected to explore the deletions in the targeted sites, which is useful to control the other confounding factors such as variable sgRNA efficiencies, and different cellular responses to the repeated injections. Moreover, off-target effects may also affect gene drives. The gene drive is a genetic concept, which means that the desired or unintended genetic variant can propagate through the germline of the population faster than the normal Mendelian inheritance, which is typically called super-Mendelian inheritance. One research utilized mosquito models created by the CRISPR-Cas9 technology to control the spread of Wolbachia. Based on their achievements, nearly 90 percent of the modified females of mosquitoes possessed Wolbachia after being released 11 weeks later, reducing the transmission of dengue [6]. In their models, the super-version of Mendelian inheritance a modified gene has more than 50 percent even 90 percent of inherited frequencies, which is much higher than the normal Mendelian inheritance (normally estimated to be 50%). However, it may also have potential to cause some problems. In the first place, the off-target effect will apply the modifications in the unintended gene sequence, which may affect or destroy the construction of the gene drives. In this case, the gene-drive models may not perform their functions at the expected efficiency. Furthermore, the unpredicted outcomes induced by CRISPR-Cas9 may also cause the models to target unintended regions, species, and populations. The untargeted modification of genes may induce new (unwanted) traits in the organisms, which may trigger the super-Mendelian inheritance, rendering the whole population or even the ecosystem to be disrupted. The viability of CRISPR-Cas9-induced super-Mendelian inheritance was proven in the female mouse models [7]. However, the proposed dangers and challengers still lack solid evidence. The bad (unintended) effects resulting from the influence of the off-target of CRISPR-Cas9 need primate models to be proven.

3. Current tools for the detection of off-targets effects of CRISPR-Cas9

Widely employed techniques, such as GUIDE-seq and CIRCLE-seq, have the capability to capture the off-target effects associated with CRISPR-Cas9 editing [8, 9]. To be specific, for GUIDE-seq, the induced DSBs by Cas9 are marked by the integration of double-stranded oligodeoxynucleotide (dsODN) containing DNA barcodes with the NHEJ repair machinery. The genome DNA is then extracted, and PCR is used to amplify both genomic DNA and the integrated dsODN. Then the amplified fragments will be sequenced in high throughput. Finally, the researcher can compare the sequenced fragment with the reference fragment to find the off-target sites with the assistance of DNA barcodes. For CIRCLE-seq, the cleaved fragments by Cas9 were circularized by ligases. The circularized fragments will then be amplified and sequenced, followed by the alignment to the reference sequence. Both methods can be either in vivo or in vitro. The need for high efficiency of dsODN transfection becomes the limitation of the GUIDE-seq. Compared with GUIDE-seq, the ability of the detection by CIRCLE-seq can cover that of GULDE-seq. Moreover, CIRCLE- seq can also find new off-target sites, which makes it a more sensitive method [9]. The demand for a large amount of genome DNA is the limitation of CRICLE-seq [9]. In addition, supplementary methodologies such as PolyA-seq, DIG-seq, and ITR-seq are harnessed for the identification of off-target effects. However, none of the methods can optimize all aspects of the off-target effect. Thus, the utilization of the methods of detection depends on different situations [8]. For example, if viral transduction is more efficient than plasmid transfection, ITR-seq could be the appropriate method. In cases where cells exhibit a facile and efficient transfection capability, GUIDE-seq emerges as an appropriate option owing to its heightened sensitivity and efficacy.
4. Tactics for mitigating the off-target effects of CRISPR-Cas9

4.1. Deep learning

The identification of off-target effects is not limited solely to robust techniques like GUIDE-seq and CIRCLE-seq but can also be anticipated using computational tools such as Deep learning methods. Deep learning neural networks analyze and alter incoming data to generate desired outputs by connecting numerous layers of artificial neurons. The phrase "deep learning" refers to networks that can include a lot of layers and are "deep." One study established a prediction model using Deep learning tools, testing the performance of the model on an off-target dataset called CRISPOR. The investigation revealed that among various approaches, the convolutional neural network (CNN) demonstrated superior performance on the dataset, attaining the highest level of efficacy. Notably, the deep feedforward neural network also demonstrated competitiveness, achieving an average AUC of 97.0% under the same conditions [10]. The value of more than 95.0% indicated that the model had excellent discriminative ability. Namely, it could accurately classify the off-target and on-target sites. The 5-fold stratification indicated the model possessed good generalization capabilities. Therefore, when the potential off-target effects could be well-predicted by the deep learning model, the gene editing by CRISPR-Cas9 could accurately target the gene sequence of interest.

4.2. Fidelity and Binding affinity

Devising a Cas9 variant characterized by elevated fidelity represents an additional strategy for mitigating the incidence of off-target effects. For example, enhanced specificity Cas9 (eSpCas9) is an engineered variant of Cas9, which aims to improve the specificity of targeting [11]. To be specific, one subvariant (eSpCas9-1.1) of it contained substitutions of amino acids in the Cas9, which reduced the interaction between the Cas9 with non-target sequencing, then reduced the possibilities of off-target. Moreover, Kleinstiver and his team [11] employed similar principles in the development of SpCas9-HF1, a high-fidelity Cas9 variant. Their aim encompassed the minimization of off-target editing, while simultaneously upholding resilient on-target functionality. By strategically modifying specific residues within Cas9 that conventionally engage with the phosphate backbone of the target DNA strand, a deliberate weakening of binding to off-target sites was achieved. The consequent SpCas9-HF1 variant demonstrated exceptional on-target efficacy while significantly mitigating off-target editing. Therefore, lowering binding affinity and maintaining high fidelity are both useful to improve on-target activity.

4.3. Dual binding

Enhancing targeting precision can be also achieved through a strategy that involves requiring two Cas proteins to bind adjacent regions in order to induce a DSB [12]. The dual-binding strategy involves the utilization of paired Cas9 nickases, coupled with two distinct single-guide RNAs that target complementary strands of the DNA target. The dual nickases induce double-strand breaks (DSBs) exclusively when both Cas9 proteins bind to closely adjacent target sequences. Consequently, the likelihood of off-target DSBs is diminished, as this scenario necessitates the occurrence of two off-target sites in close proximity within the genome. Through the implementation of these tactics, the potential for unintended off-target effects in gene editing is significantly curtailed.

4.4. Optimization of delivery methods

RNP (ribonucleoprotein) complexes are a delivery strategy that involves delivering Cas9 protein and sgRNA as pre-assembled complexes rather than expressing Cas9 from a plasmid inside the target cells. This approach has several advantages and can help reduce off-target effects. Namely, one of the main advantages of using RNP complexes is that Cas9 protein has a shorter half-life in cells compared to the expression of Cas9 from a plasmid. When Cas9 is delivered as an RNP complex, it acts quickly and efficiently, reducing the time available for potential off-target cleavage. After the
complexes interact with the target sequence, the RNP complexes will be degraded by the protease and nucleases in the cells [13].

5. Conclusion

The detection of off-target effects is a complex endeavour, as these events can occur at genomic loci with high similarity to the targeted site. Off-target effects can result from inconsistencies in the alignment of the seed sequence within the sgRNA, as well as the inappropriate length of the sgRNA. The unexpected events, especially the unpredicted modifications of genes that could induce the detrimental genes in the population through gene drives, still need long-term experiments to be demonstrated. Sophisticated sequencing methodologies, notably GUIDE-seq and CIRCLE-seq, have significantly enriched the capacity to pinpoint and delineate off-target cleavage sites. The use of such tools could be determined based on the different situations. The combinational use of the tools may make the researchers a comprehensive understanding of the off-target effects in the CRISPR-Cas9 experiments. Moreover, the employment of deep learning algorithms has exhibited promise in the prediction of unintended regions, thereby enhancing the efficacy of detection processes.

In response to the quandary of off-target effects, a range of strategies have been devised. Notably, Cas9 variants distinguished by heightened fidelity have exhibited a decrease in off-target activity while concurrently upholding their on-target efficiency. The use of paired Cas9 nickases, RNP complexes, and modified guide RNAs has also exhibited success in improving targeting precision and minimizing off-target effects. Additionally, employing bioinformatics tools and computational approaches in sgRNA design could show potential for optimizing specificity and reducing the likelihood of off-target events.

References


