Computational Molecular Biology 2025, Vol.15, No.2, 102-111
http://bioscipublisher.com/index.php/cmb

Review Article Open Access

Computational Prediction of Off-Target Effects in CRISPR Systems

Wenzhong Huang
Biomass Research Center, Hainan Institute of Tropical Agricultural Resouces, Sanya, 572025, Hainan, China
Corresponding author: wenzhong.huang@hitar.org

Computational Molecular Biology, 2025, Vol.15, No.2 doi: 10.5376/cmb.2025.15.0010
Received: 13 Feb., 2025
Accepted: 25 Mar., 2025
Published: 16 Apr., 2025

Preferred citation for this article:
Huang W.Z., 2025, Computational prediction of off-Target effects in CRISPR systems, Computational Molecular Biology, 15(2): 102-111 (doi: 10.5376/cmb.2025.15.0010)

Abstract CRISPR/Cas gene editing technology, with its advantages of simple operation, strong specificity and high efficiency, has become an important tool in life science research and molecular breeding. However, the off-target effect has always been a key issue restricting the further application of this technology, especially in clinical and agricultural genetic improvement, and its potential risks need to be addressed urgently. In recent years, methods based on computational prediction have gradually developed into important means for identifying and reducing off-target effects, providing theoretical support and practical guidance for CRISPR experimental design and safety assessment. This article systematically reviews the CRISPR system and the molecular mechanisms underlying its off-target effects, with a focus on three mainstream computational prediction strategies: sequence aligning methods, rule and machine learning-based methods, and deep learning frameworks. The article further explores the commonly used model evaluation indicators and experimental verification methods, and demonstrates the application process of off-target prediction through a case study of the human EMX1 gene. Finally, the contributions of computational prediction methods in enhancing editing specificity were summarized, the current limitations were analyzed, and the future directions for promoting the development of this field through multimodal data integration, algorithm optimization, and preclinical safety assessment were prospected. This article aims to provide a systematic reference for subsequent research on CRISPR-based security applications.

Keywords CRISPR; Off-target effect; Computational prediction; Machine learning; Functional genomics

1 Introduction

Since the CRISPR/Cas9 system was officially applied to gene editing in 2012, this technology has rapidly become one of the core tools in life science research (Guo et al., 2023). Compared with the traditional zinc finger nucleases (ZFNs) and transcription activator effector nucleases (TALENs), the CRISPR/Cas system is not only easy to construct, but also shows good applicability in a variety of biological systems (Naeem et al., 2020). However, with its wide application in basic research, medical treatment and agricultural improvement and other fields, off-target effects have gradually become a safety hazard that urgently needs attention. Off-target effect refers to the cleavage or regulation of Cas nucleases at unexpected targets, thereby causing non-specific alterations in the genome. This phenomenon may lead to incorrect knockout of functional genes, accumulation of potential mutations, and even cause unpredictable biological consequences. Therefore, how to effectively predict and reduce off-target effects is the key to whether CRISPR technology can further move towards precision and application.

At present, researchers have developed a variety of off-target detection methods, such as experimental techniques like guiding seq, Digenome-seq, and CIRCLE seq (Martin et al., 2016). However, these methods are usually costly, time-consuming, and difficult to promote in large-scale research. In contrast, computational prediction methods have become an important supplement to the study of off-target effects due to their characteristics of rapidity, low cost and high throughput. By establishing mathematical models and algorithmic tools, researchers can conduct a comprehensive assessment of potential off-target sites before gene editing experiments, thereby improving the rationality of the design and the success rate of the experiments (Zhang et al., 2020).

The existing computational prediction strategies can be roughly divided into three categories: methods based on sequence alignment, methods based on rules and machine learning, and prediction frameworks based on deep learning. Sequence alignment methods have become the earliest applied means due to their simplicity.

Representative tools such as BLAST and Bowtie can search for loci similar to the target sequence through genome-wide alignment. With the deepening of research, rule-based and machine learning-based methods have gradually emerged, such as MIT scoring and CFD models. These methods not only consider the location and quantity of base mismatches but also introduce statistical patterns and experimental data for modeling. In recent years, with the development of artificial intelligence, deep learning frameworks have been introduced into the field of off-target prediction. Representative models such as DeepCRISPR, R-CRISPR and DL-CRISPR can improve the accuracy and universality of prediction through automatic feature extraction and learning from large-scale training data (Sherkatghanad et al., 2023).

In addition to the progress of the predictive model itself, how to evaluate its performance and reliability is equally crucial. Common evaluation indicators include ROC curve, AUC value, accuracy rate and recall rate, etc. Experimental verification methods such as GUIDE-seq and high-throughput sequencing provide solid support for computational prediction. By combining computational and experimental methods, scholars have gradually established a relatively complete off-target assessment system.

2 Introduction to the CRISPR/Cas System and Off-Target Effect Mechanism

2.1 Overview of CRISPR/Cas system

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) system was initially discovered in bacteria and archaea as an acquired immune mechanism. Its core function is to resist the invasion of viruses and plasmids by identifying and cutting exogenous invading DNA. As research deepened, scientists transformed it into a programmable gene editing tool. Among them, the Cas9 protein, which is the most widely used, can be cleaved at specific positions in the genome under the guidance of single-guide RNA (sgRNA), thereby achieving knockout, insertion or modification of the target gene (Yuan, 2024). CRISPR/Cas9 has demonstrated highly efficient editing capabilities in crops such as rice, wheat, and corn, as well as in human cells and model animals. Therefore, it holds broad application prospects in molecular breeding, disease model construction, and gene therapy. In addition, new CRISPR-derived systems (such as Cas12a/Cpf1, Cas13, etc.) are constantly being discovered and utilized, providing researchers with more diverse means of gene manipulation.

2.2 Definition and generation mechanism of off-target effects

Off-target effect refers to the cleavage or regulation of the CRISPR/Cas system at non-target sites, resulting in unexpected alterations in the genome. Its generation mechanism mainly includes the following aspects: Base mismatch tolerance: The Cas9-sgRNA complex is not absolutely strict with the target sequence. Sometimes, a small amount of mismatch is allowed, especially in the non-PAM proximal region, thereby triggering non-specific cleavage. PAM sequence diversity: Although SpCas9 most commonly recognizes NGG PAM, it can also recognize approximate sequences such as NAG and NGA, which increases the number of potential off-target sites. Genomic complexity: In large-genome species, there are numerous fragments partially similar to the target sequence, further intensifying the risk of off-target. Chromatin state and accessibility: Open chromatin regions are more easily recognized by Cas9, and cleavage may occur even in the presence of base mismatches.

2.3 The importance of off-target effect prediction research

Off-target effects not only affect the accuracy of experiments but may also bring serious safety hazards in medical and agricultural applications. In human cell gene therapy, non-targeted cleavage may lead to the activation of oncogenes or the inactivation of tumor suppressor genes. In crop breeding, off-target effects may cause unexpected trait changes and affect product safety. Therefore, in order to reduce risks, it is very necessary to establish efficient and accurate calculation and prediction methods. Through computational prediction, researchers can identify potential off-target sites before experiments and optimize the design of sgRNA, thereby reducing non-specific effects at the source.

3 Sequence Alignment Type Prediction Methods

3.1 Principles and applications of BLAST algorithm

BLAST (Basic Local Alignment Search Tool) is one of the earliest tools used for off-target prediction. The basic principle is to locally compare the sgRNA sequence with the reference genome to search for fragments similar to

the target sequence (Sun, 2023). Due to its mature algorithm and extensive database, BLAST has certain advantages in the initial screening of potential off-target sites. However, BLAST was not specifically designed for CRISPR, and its results do not adequately consider the weights of PAM sequences and mismatch positions, which can easily lead to false positives or false negatives. Therefore, it is more suitable as an early exploration or rough screening tool.

3.2 Application of sequence alignment tools such as Bowtie in off-target prediction

Bowtie is a fast short sequence alignment tool that can achieve efficient matching in large-scale data (Guo and Zhen, 2020). Compared with BLAST, Bowtie has more advantages in terms of speed and memory utilization, and thus is widely used in high-throughput off-target prediction research. Researchers usually input sgRNA sequences as query sequences, search for similar sites in the genome through Bowtie, and screen potential off-target sites in combination with PAM sequences. In addition, tools such as BWA are also used for comparative prediction. Their advantage lies in the flexible setting of mismatch tolerance.

3.3 Advantages and limitations of sequence alignment method

The main advantages of sequence alignment methods lie in their simple principle, rapid calculation, and convenient operation, making them suitable for the preliminary screening of potential off-target sites. However, its limitations are also quite obvious: it is unable to accurately distinguish the impact of different mismatch positions on cutting efficiency; Epigenetic factors such as chromatin opening degree and DNA methylation were not considered; Lack of experimental data-driven, limited prediction accuracy (Zhang and Jiang, 2022). Therefore, sequence alignment methods are usually combined with other computational methods to establish a more comprehensive prediction system.

4 Prediction Methods Based on Rules and Machine Learning

4.1 Rule-based off-target scoring algorithms (such as MIT algorithms, etc.)

The MIT algorithm is one of the earlier proposed rule-based off-target prediction methods. Its core idea is to assign different weights based on the different influences of mismatch positions on cutting efficiency. It is generally believed that mismatches closer to the PAM end have a greater impact on cutting activity, while mismatches farther from the PAM have a higher tolerance (Chao and Fei, 2023). This method is concise and intuitive, and combines certain experimental rules, thus it was widely used in the early days. However, its weight parameters rely on limited experimental data and lack universality (Fan and Xu, 2021).

4.2 CFD off-target scoring model and optimization

The cutting frequency determination (CFD) model has been improved on the basis of the MIT algorithm, integrating more experimental data, especially the influence of different base mismatch types and positions. By statistically analyzing a large number of experimental results, the CFD model can assign more reasonable probability values to each mismatch combination, thereby predicting the off-target risk more accurately. Currently, CFD has been integrated into multiple CRISPR design platforms, such as CRISPOR and CRISPR-DO, and serves researchers widely.

4.3 Application of machine learning models in off-target prediction

With the accumulation of data and the development of algorithms, researchers began to introduce machine learning methods for off-target prediction (Anuradha et al., 2024). Common models include support vector machine (SVM), random forest (RF), and logistic regression, etc. These methods can learn patterns from a large amount of known off-target and non-off-target data and establish classifiers to determine whether new sequences are potential off-target sites (Figure 1) (Toufikuzzaman et al., 2024). Compared with traditional rule-based methods, machine learning models can consider multiple features simultaneously, such as the number of mismatches, distribution, GC content, PAM sequence, etc., thereby improving the accuracy of prediction. However, machine learning methods rely on high-quality training data and may have overfitting problems (Choubisa, 2024).

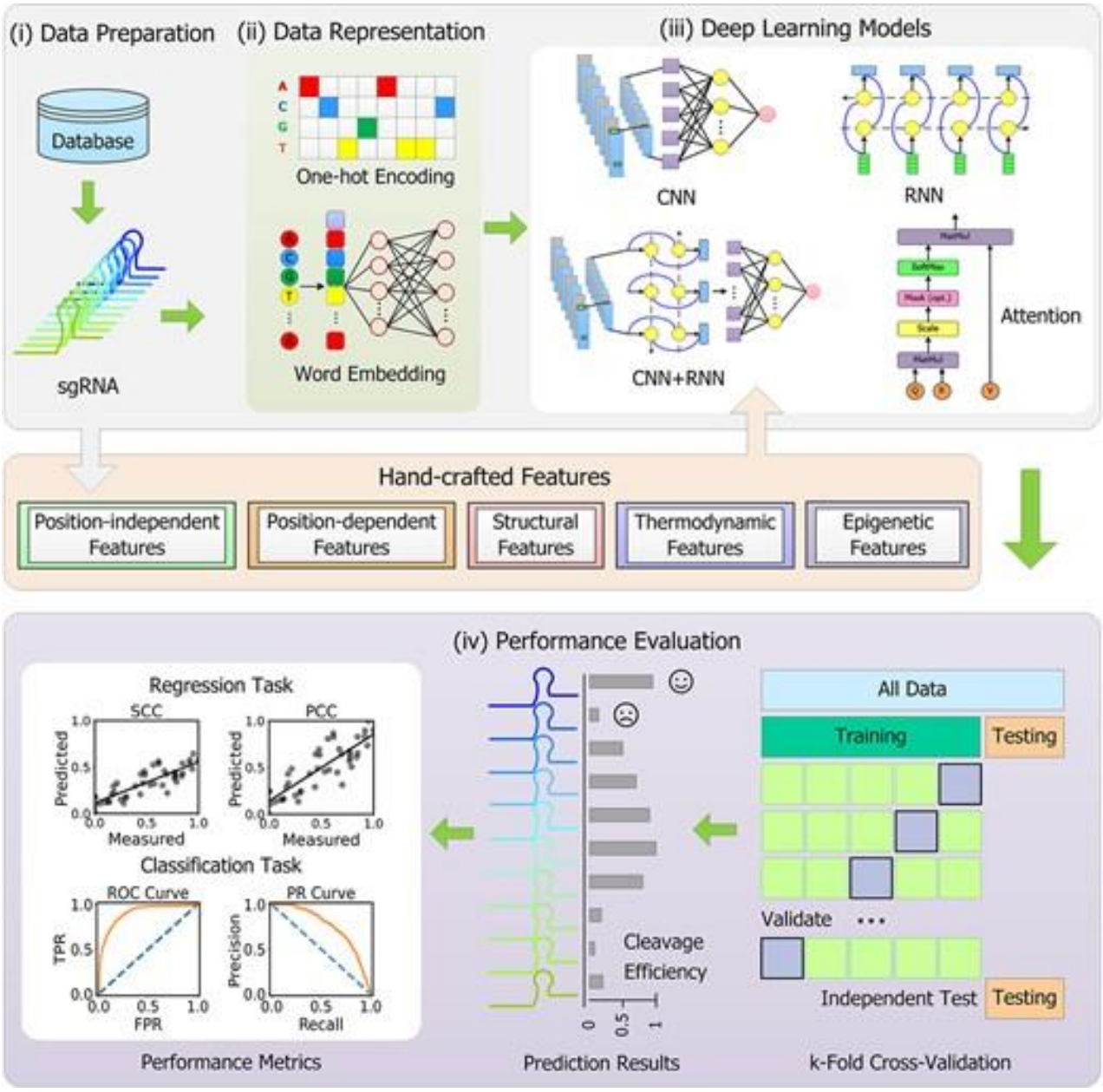

Figure 1 Overview of deep learning for predicting CRISPR/Cas9 sgRNA on-target activity that involved the following steps: (i) data collection and preprocessing; (ii) data representation; (iii) deep learning models. Indirect features extracted from sgRNA sequences could be combined as input for deep learning. (iv) Performance evaluation (Adopted from Toufikuzzaman et al., 2024)

5 CRISPR Off-Target Effect Prediction Model Based on Deep Learning

5.1 DeepCRISPR deep learning framework

DeepCRISPR is one of the earliest models to apply deep learning methods for CRISPR off-target prediction. It uses convolutional neural networks (CNNS) to automatically extract sequence features, no longer relying on traditional manually set metrics. By inputting target sequences and potential off-target sequences, the model can learn the complex relationships among mismatch distribution, PAM environment and sequence context, thereby improving the accuracy and generalization ability of prediction (Chuai et al., 2018).

The advantage of DeepCRISPR lies in its ability to handle large-scale data and achieve automated prediction through an end-to-end approach. Research shows that this model significantly outperforms traditional models such as MIT and CFD in terms of ROC curve and AUC value. However, DeepCRISPR is highly dependent on training data. If the dataset is biased towards certain genomes or species, it may lead to insufficient universality of prediction.

5.2 CRISPR-net model and method

CRISPR-Net is another off-target prediction tool based on deep learning. It adopts a method combining recurrent neural networks (RNN) and convolutional neural networks, which can not only capture local features of sequences but also identify long-term dependencies. CRISPR-Net achieved higher sensitivity and accuracy by training a large amount of real off-target detection data.

Its notable feature is the introduction of sequence position embedding and attention mechanisms, enabling the model to "understand" the degree of influence of different mismatch positions. Compared with DeepCRISPR, CRISPR-Net performs better in cross-species prediction and shows high consistency on human, mouse and plant data. This indicates that deep learning has unique advantages in dealing with complex genomic backgrounds and diverse PAM identification (Zhang et al., 2023).

5.3 Comparison of other deep learning prediction models (such as R-CRISPR, etc.)

In addition to DeepCRISPR and CRISPR-Net, models such as R-CRISPR and ElevatedCRISPR have also been proposed. Most of these models have introduced structural improvements, such as convolutional neural networks (GCN) to capture the spatial features of sequences, or Transformer frameworks to achieve stronger context modeling capabilities (Niu et al., 2021).

Different deep learning models each have their own advantages. For example, R-CRISPR pays more attention to the interaction between mismatch sequences and genomic background and is suitable for studying complex off-target patterns. The Transformer-based model demonstrated higher generalization ability on large-scale datasets. Overall, the introduction of deep learning has significantly improved the accuracy of off-target prediction, but it has also brought about the "black box effect" problem, that is, it is difficult to intuitively explain the biological significance of the prediction results.

6 Model Evaluation Indicators and Experimental Verification Strategies

6.1 Common evaluation indicators of off-target prediction models (ROC curve, etc.)

To measure the performance of different prediction models, researchers usually employ a series of statistical indicators. Receiver operating characteristic (ROC) Curve and area under the curve (AUC) value are the most commonly used evaluation methods to measure the ability of the model to distinguish off-target from non-off-target. In addition, metrics such as accuracy rate, recall rate, and F1 score are also widely applied in various scenarios.

These metrics can reflect the model's performance under different trade-offs. For instance, when emphasizing "comprehensive capture of potential off-targets", the recall rate is even more crucial. In the application scenarios that pursue "reducing false positives", accuracy is even more crucial.

6.2 Experimental verification methods for off-target effect detection

Computational prediction must rely on experimental verification to ensure reliability. The common verification methods currently available include: Guiding seq: Introducing small DNA fragments as tags through double-strand breaks to capture and sequence the real cleavage sites; Digenome-seq: Detection of traces of in vitro cleavage of Cas9 using whole genome sequencing (Charlier et al., 2021); CIRCLE-seq: High sensitivity and low background noise by cyclizing DNA and detecting cleavage products (Tsai et al., 2017).

6.3 Model performance evaluation and comparison

Different prediction methods perform differently in practical applications. The sequence alignment method has obvious advantages in terms of speed and initial screening, but its accuracy is insufficient. Rule-based and machine learning-based methods are relatively balanced, but they rely on feature selection. Deep learning methods perform outstandingly in prediction accuracy, but they have high requirements for data volume and computing resources (Kimata and Satou, 2025).

In the evaluation, researchers usually cross-validate the computational prediction results with the experimental test results to examine the stability and applicability of the model in different genomic contexts. Overall, deep learning

methods represent the future development direction, but how to strike a balance among transparency, interpretability and prediction accuracy remains an urgent problem to be solved in this field.

7 Predictive Analysis of Human *EMX1* Gene Off-target

7.1 Research background of *EMX1* gene editing

The *EMX1* gene (*Empty Spiracles Homeobox 1*) is an important transcription factor related to human development and is often selected as the target model for CRISPR gene editing research. Due to its significant role in the development of the nervous system, researchers pay particular attention to off-target effects when using CRISPR-Cas9 to knockout EMX1. Early studies have shown that even well-designed Sgrnas have potentially dozens of approximate targets in the human genome. Therefore, this gene has become a typical case in off-target prediction and validation studies (Figure 2) (Chaudhari et al., 2020).

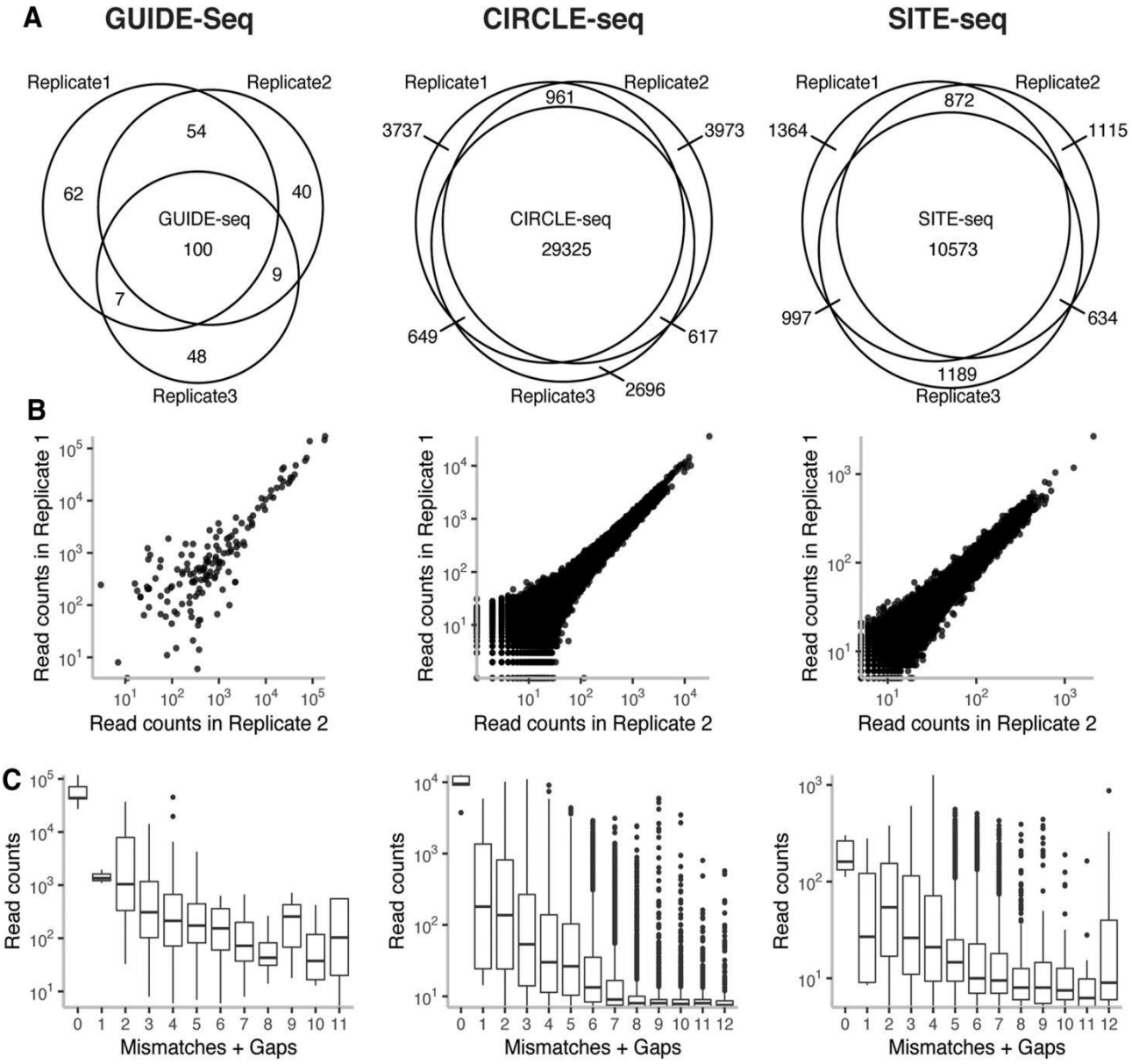

Figure 2 (A) Overlap among three replicates for sites nominated by GUIDE-seq, CIRCLE-seq, and SITE-seq across all eight gRNAs. (B) Correlation of read counts for nominated sites between two replicates for GUIDE-seq (R2 = 0.54–0.67), CIRCLE-seq (R2 = 0.80–0.83), and SITE-seq (R2 = 0.90–0.91) read counts. The plots present data for replicates 1 and 2, and the R2 range is calculated from all three comparisons (1 vs. 2, 2 vs. 3, 1 vs. 3). (C) Distributions of read counts for nominated sites obtained by for GUIDE-seq, CIRCLE-seq, and SITE-seq grouped by the number of mismatches and gaps in the site with respect to gRNA sequence. gRNA, guide RNA (Adopted from Chaudhari et al., 2020)

7.2 Calculation and prediction process of off-target sites

In the off-target prediction of EMX1, researchers usually first use sequence alignment tools (such as Bowtie) to quickly locate possible approximate sequences. Subsequently, the risk score of each candidate site was calculated through the CFD model to screen out high-priority potential off-target sites. Furthermore, the researchers utilized deep learning models such as DeepCRISPR to conduct fine predictions of these candidate sites, thereby narrowing the scope of experimental verification.

The prediction process generally includes the following steps: Input the EMX1 sgRNA sequence and its PAM recognition site; Genome-wide alignment and screening for sequences with high similarity; Calculate the off-target risk score and generate a list of candidate off-target sites; Further optimize the results using a deep learning framework and output a list of highly reliable predictions (Störtz et al., 2023).

7.3 Prediction Results and Experimental Verification Analysis

Through computational prediction, researchers can usually obtain dozens of potential off-target sites. Subsequently, experiments were conducted using GUIDE-seq or CIRCLE-seq for verification. The results showed that there were indeed Cas9 cutting traces at some predicted sites. It is notable that the prediction results of CFD models and DeepCRISPR are often highly consistent with the experimental data, while those relying solely on BLAST or Bowtie produce more false positives than the opposing rules (Lin et al., 2020). This case fully demonstrates that integrating multiple prediction methods and combining them with experimental verification is an effective way to ensure the reliability of off-target assessment (AlJanahi et al., 2020).

8 Future Development Trends

8.1 Multimodal data integration enhances off-target prediction capabilities

An important direction for future off-target prediction is the integration of multimodal data. In addition to the DNA sequence itself, factors such as chromatin open state, DNA methylation, histone modification and three-dimensional genomic structure should also be taken into consideration. By combining these data with deep learning models, the prediction results will be closer to the real intracellular environment (Lin and Wong, 2018).

8.2 Technological and algorithmic improvements for enhancing CRISPR specificity

At the technical level, new Cas variants (such as high-fidelity SpCas9-HF1, eSpCas9) and improved sgRNA design strategies have shown potential in reducing off-target risk (Chen et al., 2017; Wang et al., 2023; Matsumoto et al., 2024). In the future, computational prediction models can be combined with these new technologies to provide customized off-target evaluations for different editing needs. Meanwhile, algorithm development also needs to focus on transparency and interpretability to help researchers understand the biological logic behind the predictions.

8.3 New advances in preclinical safety evaluation of gene editing

As CRISPR technology advances towards clinical application, how to achieve comprehensive safety assessment in the preclinical stage has become a core issue. Computational prediction will be combined with high-throughput sequencing, single-cell analysis and animal models to establish a multi-level off-target assessment system (Tian et al., 2023). This not only helps to reduce the risks in gene therapy, but also provides reliable safety guarantees for agricultural molecular breeding. In the future, establishing an internationally unified prediction and verification standard will be an important guarantee for the healthy development of this field (Cancellieri et al., 2022).

9 Summary and Outlook

This paper systematically reviews the research on the computational prediction of off-target effects in the CRISPR system, unfolding in sequence according to the established outline. Starting from the CRISPR/Cas system and off-target effect mechanisms, this paper introduces sequence alignment methods, rule-based and machine learning-based methods, deep learning prediction frameworks, as well as model evaluation and experimental verification strategies. Through the case study of the human EMX1 gene, the specific process and verification results of computational prediction in practical applications are demonstrated. Overall, the computational method has played a significant role in enhancing the rationality of sgRNA design and reducing experimental risks.

Sequence alignment method, with its simplicity and efficiency, is suitable for preliminary screening, but it lacks consideration of complex biological backgrounds. Rule-based and machine learning-based methods strike a balance between accuracy and efficiency, but are limited by the scale of training data. Deep learning methods, with their powerful feature extraction capabilities, have significantly enhanced prediction accuracy and cross-species generalization ability. However, its reliance on large-scale high-quality data and the "black box effect" issue remain the bottlenecks restricting its wide application.

The future development trends are mainly reflected in three aspects: First, the integration of multimodal data will make the prediction model closer to the real environment within cells, thereby enhancing reliability; Secondly, when combined with new Cas variants and optimized sgRNA strategies, it will promote the minimization of off-target risks. Thirdly, in clinical and agricultural applications, establishing unified safety assessment standards and norms will provide a guarantee for the healthy development of CRISPR technology.

In conclusion, the value of computational prediction in CRISPR off-target research is not only reflected at the theoretical level but also provides a solid support for the safe application of gene editing. With the continuous iteration of algorithms, the continuous accumulation of data, and the increasingly improved experimental verification methods, in the future, we have every reason to believe that off-target prediction will gradually transform from an "auxiliary tool" to a "core guarantee", opening up broader prospects for precise gene editing and molecular breeding.

Acknowledgments

I would like to thank Dr. Xie continuous support throughout the development of this study.

Conflict of Interest Disclosure

The author affirms that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

AlJanahi A., Lazzarotto C., Chen S., Shin T.H., Cordes S., Jabara I., Zhou Y., Young D., Lee B., Yu K., Li Y., Toms B., Tunc I., Hong S.G., Truitt L.L., Klermund J., Kim M.Y., Cathomen T., Gill S., Tsai S.Q., and Dunbar C., 2020, Validation and long-term follow up of CD33 off-targets predicted in vitro and in silico using error-corrected sequencing in rhesus macaques, bioRxiv, 2020: 186858.
https://doi.org/10.1101/2020.07.05.186858

Anuradha B., Pradeep T., and Vikrant N., 2024, Machine learning-driven prediction of CRISPR-Cas9 off-target effects and mechanistic insights, The EuroBiotech Journal, 8(4): 213-229.
https://doi.org/10.2478/ebtj-2024-0020

Cancellieri S., Zeng J., Lin L.Y., Tognon M., Nguyen M.A., Lin J., Bombieri N., Maitland S.A., Ciuculescu M.F., Katta V., Tsai S.Q., Armant M., Wolfe S.A., Giugno R., Bauer D.E., and Pinello L., 2022, Human genetic diversity alters off-target outcomes of therapeutic gene editing, Nature Genetics, 55(1): 34-43.
https://doi.org/10.1038/s41588-022-01257-y

Chao R., and Fei J., 2023, Off-target effects of CRISPR/Cas9 and their solutions, Highlights in Science, Engineering and Technology, 45: 296-301.
https://doi.org/10.54097/hset.v45i.7444

Charlier J., Nadon R., and Makarenkov V., 2021, Accurate deep learning off-target prediction with novel sgRNA-DNA sequence encoding in CRISPR-Cas9 gene editing, Bioinformatics, 37(16): 2299-2307.
https://doi.org/10.1093/bioinformatics/btab112

Chaudhari H., Penterman J., Whitton H.J., Spencer S.J., Flanagan N., Zhang L., Huang E., Khedkar A.S., Toomey J., Shearer C.A., Needham A.W., Ho T.W., Kulman J.D., Cradick T.J., and Kernytsky A., 2020, Evaluation of homology-independent CRISPR-Cas9 off-target assessment methods, The CRISPR Journal, 3(6): 440-453.
https://doi.org/10.1089/crispr.2020.0053

Chen J.S., Dagdas Y., Kleinstiver B.P., Welch M.M., Sousa A.A., Harrington L.B., Sternberg S.H., Joung J.K., Yildiz A., and Doudna J.A., 2017, Enhanced proofreading governs CRISPR-Cas9 targeting accuracy, Nature, 550(7676): 407-410.
https://doi.org/10.1038/nature24268

Choubisa V., 2024, CRISPR-Cas9 off-target predictions using CNN and double CNN: comparative analysis, International Journal of Science and Technology Research Archive, 12(1): 1074-1080.
https://doi.org/10.30574/ijsra.2024.12.1.0927

Chuai G., Ma H., Yan J., Chen M., Hong N., Xue D., Zhou C., Zhu C., Chen K., Duan B., Gu F., Qu S., Huang D., Wei J., and Liu Q., 2018, *DeepCRISPR*: optimized CRISPR guide RNA design by deep learning, Genome Biology, 19(1): 80.
https://doi.org/10.1186/s13059-018-1459-4

Fan Y., and Xu H., 2021, Prediction of off-target effects in CRISPR/Cas9 system by ensemble learning, Current Bioinformatics, 16(9): 1169-1178.
https://doi.org/10.2174/1574893616666210811100938

Guo C., and Zhen D., 2020, Prediction of off-target effects of the CRISPR/Cas9 system for design of sgRNA, In: E3S web of conferences, EDP Sciences, 185: 04018.
https://doi.org/10.1051/e3sconf/202018504018

Guo X., Ma Y., Gao F., and Guo, Y., 2023, Off-target effects in CRISPR/Cas9 gene editing and their potential solutions, Frontiers in Bioengineering and Biotechnology, 11: 1143157.
https://doi.org/10.3389/fbioe.2023.1143157

Kimata K., and Satou K., 2025, Improved CRISPR/Cas9 off-target prediction with DNABERT and epigenetic features, bioRxiv, 2025: 649101.
https://doi.org/10.1101/2025.04.16.649101

Lin J., and Wong K., 2018, Off-target predictions in CRISPR-Cas9 gene editing using deep learning, Bioinformatics, 34(17): i656-i663.
https://doi.org/10.1093/bioinformatics/bty554

Lin J., Zhang Z., Zhang S., Chen J., and Wong K.C., 2020, CRISPR-Net: a recurrent convolutional network quantifies CRISPR off-target activities with mismatches and indels, Advanced Science, 7(13): 1903562.
https://doi.org/10.1002/advs.201903562

Martin F., Sánchez-Hernández S., Gutiérrez-Guerrero A., Pinedo-Gomez J., and Benabdellah K., 2016, Biased and unbiased methods for the detection of off-target cleavage by CRISPR/Cas9: an overview, International Journal of Molecular Sciences, 17(9): 1507.
https://doi.org/10.1038/mtna.2016.90

Matsumoto D., Matsugi E., Kishi K., Inoue Y., Nigorikawa K., and Nomura W., 2024, SpCas9-HF1 enhances accuracy of cell cycle-dependent genome editing, Molecular Therapy-Nucleic Acids, 35(1): 102124.
https://doi.org/10.1016/j.omtn.2024.102124

Naeem M., Majeed S., Hoque M.Z., and Ahmad I., 2020, Latest developed strategies to minimize the off-target effects in CRISPR-Cas-mediated genome editing, Cells, 9(7): 1608.
https://doi.org/10.3390/cells9071608

Niu B., Peng J., Zhang Z., and Shang X., 2021, R-CRISPR: a deep learning network to predict off-target activities with mismatches insertion and deletion in CRISPR-Cas9 system, Genes, 12(12): 1878.
https://doi.org/10.3390/genes12121878

Sherkatghanad Z., Abdar M., Charlier J., and Makarenkov V., 2023, Using traditional machine learning and deep learning methods for on-and off-target prediction in CRISPR/Cas9: a review, Briefings in Bioinformatics, 24(3): bbad131.
https://doi.org/10.1093/bib/bbad131

Störtz F., Mak J., and Minary P., 2023, piCRISPR: physically informed deep learning models for CRISPR/Cas9 off-target cleavage prediction, Artificial Intelligence in the Life Sciences, 3: 100075.
https://doi.org/10.1016/j.ailsci.2023.100075

Sun H., 2023, The challenge facing CRISPR/Cas9 system: off-target effects and their optimization, Highlights in Science, Engineering and Technology, 74: 782-787.
https://doi.org/10.54097/psd28z73

Tian R., Cao C., He D., Dong D., Sun L., Liu J., Chen Y., Wang Y., Huang Z., Li L., Jin Z., Huang Z., Xie H., Zhao T., Zhong C., Hong Y., and Hu Z., 2023, Massively parallel CRISPR off-target detection enables rapid off-target prediction model building, Med, 4(7): 478-492. e6.
https://doi.org/10.1016/j.medj.2023.05.005

Toufikuzzaman M., Abul M., Samee H., and Rahman M.S., 2024, CRISPR-DIPOFF: an interpretable deep learning approach for CRISPR Cas-9 off-target prediction, Briefings in Bioinformatics, 25(2): bbad530.
https://doi.org/10.1093/bib/bbad530

Tsai S.Q., Nguyen N.T., Malagón-López J., Topkar V.V., Aryee M.J., and Joung J.K., 2017, CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets, Nature Methods, 14(6): 607-614.
https://doi.org/10.1038/nmeth.4278

Wang G., Wang C., Chu T., Wu X., Anderson C.M., Huang D., and Li J., 2023, Deleting specific residues from the HNH linkers creates a CRISPR-SpCas9 variant with high fidelity and efficiency, Journal of Biotechnology, 368: 42-52.
https://doi.org/10.1016/j.jbiotec.2023.04.008

Yuan S., 2024, Mitigating the off-target effects in CRISPR/Cas9-mediated genetic editing with bioinformatic technologies, Transactions on Materials, Biotechnology and Life Sciences, 3: 318-326.
https://doi.org/10.62051/dpgwbz03

Zhang G., Luo Y., Dai X., and Dai Z., 2023, Benchmarking deep learning methods for predicting CRISPR/Cas9 sgRNA on- and off-target activities, Briefings in Bioinformatics, 24(6): bbad333.
https://doi.org/10.1093/bib/bbad333

Zhang Y., Long Y., Yin R., and Kwoh C.K., 2020, DL-CRISPR: a deep learning method for off-target activity prediction in CRISPR/Cas9 editing, IEEE Access, 8: 76610-76617.
https://doi.org/10.1109/ACCESS.2020.2989454

Zhang Z., and Jiang Z., 2022, Effective use of sequence information to predict CRISPR-Cas9 off-target, Computational and Structural Biotechnology Journal, 20: 650-661.
https://doi.org/10.1016/j.csbj.2022.01.006