Indian Journal of Biochemistry & Biophysics
Vol. 62, October 2025, pp. 1047-1055
DOI:10.56042/ijbb.v62i10.15918

A Review

Computational tools for CRISPR Off-target detection: An overview

P Supriya[1]*, Tamanna Sharma[1], Swarajya Lakshmi N Bollineni[1], Gowtham Kumar[1], Vyshnavi M[1], P Sumathi[2], M Balakrishnan[1] & Ch Srinivasa Rao[1]

[1]ICAR-National Academy of Agricultural Research Management, Hyderabad-500 030, Telangana, India

[2]Acharya N. G. Ranga Agricultural University, Lam, Guntur-552 034, Andhra Pradesh, India

Received 18 February 2025; *revised 24 March 2025*

The detection of off-targets is crucial for the application of CRISPR technology in both therapeutics and plant genome editing. Off-target effects can lead to unintended genome modifications, potentially disrupting gene function and causing adverse genomic consequences. These off-target regions may feature mismatches, insertions, or deletions and can vary based on their genomic location. Methods for detecting off-targets are generally classified into biased and unbiased categories. Unbiased methods include both *in vitro* and *in vivo* techniques, though biased methods are often preferred due to their time and cost efficiency. Biased methods typically involve *in silico* screening of off-targets, followed by validation through PCR or sequencing, supported by various bioinformatics tools. The advent of artificial intelligence has significantly impacted the CRISPR field. Machine learning and deep learning models, developed using experimental data from unbiased methods, have enhanced the identification of true off-targets. These AI-driven approaches not only improve accuracy but also facilitate the prediction of off-target effects in a more efficient manner. Furthermore, the integration of AI with CRISPR technology holds promise for optimizing guide RNA design, minimizing off-target activity, and tailoring CRISPR systems for specific applications. The combination of CRISPR and AI is expected to advance precision genome editing, paving the way for more reliable therapeutic interventions and agricultural innovations.

Keywords: CRISPR, Deep learning, Genome editing, Guide RNA, Off-target

Introduction

Genome editing is the procedure of making genetic manipulations in the cells of all organisms. It is achieved by making double stranded breaks (DSBs) in the target DNA. Three main nucleases are there for introducing DSBs in target DNA: zinc-finger nucleases (ZFNs), transcription activator-like effect or nucleases (TALENs) and Clustered Regularly-Interspaced Short Palindromic Repeats (CRISPR)[1]. CRISPR-Cas is the most preferred method of genome editing nowadays due to its simplicity and specificity, and the potential to edit large genomes[2]. CRISPR–Cas9 and CRISPR–Cpf1 are the most widely used CRISPR systems for genome editing in plants for different traits such as disease resistance, biotic and abiotic stress tolerance and improved yield potential[3,4].

For performing CRISPR based editing, 20 nucleotide guide RNA sequence (gRNA) and Cas9 protein are introduced into cell where the gRNA binds to its complementary sequence in the genome present adjacent to the protospacer adjacent motif (PAM) sequence, and subsequently cleaved by Cas9 protein. Some mismatches amid the gRNA and DNA are tolerated in CRISPR leading to the non-specific binding of gRNA known as off-targets. Pairing of gRNA with protospacer requires homology in a short region *i.e.* 20 base pairs sequence only which reduces the specificity of getting a single hit in the entire genome. Further, 100% identity is not mandatory for pairing thereby allowing mismatches also. All these factors result in the undesired cleavage or alterations in the genome termed as off-targets[5]. These off-target mutations are the key concern in applying CRISPR technology for genome editing for therapeutics and

*Correspondence:
E-mail:supriya@naarm.org.in; puramsupriya@gmail.com

Abbreviations: BiLSTM, Bidirectional Long Short-Term Memory; BLAST, Basic Local Alignment Search Tool; CHANGE-seq, Circularization for High-Throughput Analysis of Nuclease Genome-wide Effects by Sequencing; CNN, Convolutional Neural Network; CRISPR, Clustered Regularly-Interspaced Short Palindromic Repeats GloVe, Global Vectors for Word Representation; gRNA, Guide RNA; PAM, Protospacer Adjacent Motif; SITE-seq, Selective Enrichment and Identification of Adapter-Tagged DNA Ends by Sequencing; VIVO, Verification of *in vivo* Off-targets

agriculture as they might cause genome instabilities and may disrupt the functionality of the normal genes. There are two kinds of off-targets: one is the off-targets detected in the genomic regions having homology to the target region and other off- targets identified in the genomic region having no sequence similarity with the target[6]. In this review, we summarized about off-targets, methods to detect off-targets and computational tools available for off-target prediction.

Off-targets

Types

The off-target regions are categorized in three classes based on their identified location. The first class includes regions adjacent to PAM sequence 5′-NGG-3′ having mismatches or substitutions. Second class includes the regions adjacent to PAM sequence 5′-NGG-3′ having insertions and/or deletions with respect to target DNA and the third class includes region adjacent to different PAM sequence 5′-NAG-3′ or 5′-NGA-3′. Off-target sites consist of mismatches, RNA bulges or DNA bulges.

Off-targets in plants

Off-target effects are also problematic in plants but they are not studied to much extent as in humans and animals due to their removal in the backcrossing process. However, these off-targets can further add complexity to the mutation analysis especially in polyploidy plants species. Various methods for the detection and reduction of off-targets are basically developed for humans and animals, where the presence of off-targets results in deleterious effects[7]. Figure 1 shows the influence of CRISPR off-targets on genetic mutations, thereby affecting the phenotype of the plant. The off-targets are rarely reported in plants but various studies have reported the presence of off-targets in different crops including rice, cucumber, maize, apple, cotton, soybean, tomato, and wheat. Off-target activity in plants has been predicted by comparing the target site sequence with the whole genome and various bioinformatics tools have been developed to predict off-target sites in plants. Off-targets in plants without a reference genome are difficult to predict with these tools. CRISPR/Cas9 systemCas9 was utilized for performing targeted gene knockout in cucumber, where off-targets were predicted using CRISPR-P program. Five candidate potential off-targets were observed[8]. Also, *in vivo* prediction of off-targets was performed in transgenic rice, where it was concluded that mismatches in the 10-bp long pairing region proximal to PAM can reduce the chances of getting off-targets in plant cells. Further, bioinformatics analysis was done to find the sequences similar to target sequence using BLASTN program[9]. The majority of the off-target prediction in plants has been performed by a biased method using bioinformatics tools followed by PCR or sequencing. However, whole genome re-sequencing (WGRS) is also utilized for the identification of off-targets sites in some crops. Utilization of bioinformatics tools for off-target identification is limited to plants with

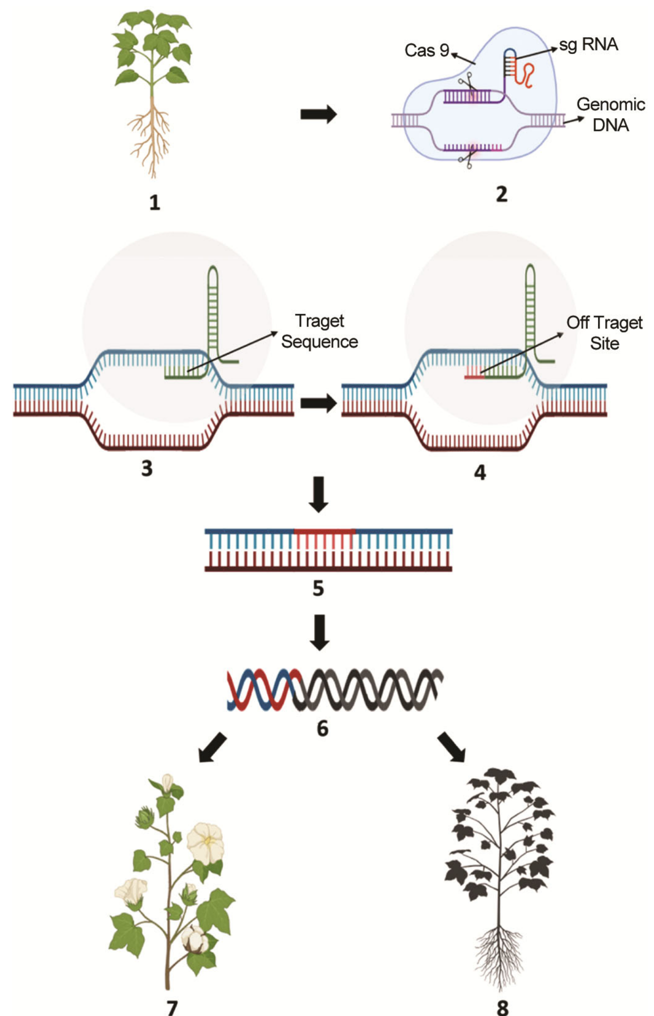

Fig. 1 — The influence of CRISPR off-targets on genetic mutations, thereby affecting the phenotype of the plant. It depicts the following key elements: 1. A plant undergoing CRISPR-mediated genome editing. 2. The CRISPR mechanism at work. 3. The detection of the target sequence. 4. The Cas9 system occasionally binds to sites other than the target loci, resulting in off-target effects. 5. Off-target editing events. 6. Undesired gene activation or deactivation. 7. Normal gene function leading to healthy growth and yield. 8. Abnormal gene function resulting in a malfunctioning plant. This figure was created using Biorender.com

sequenced genomes which results in lower number of predicted off-targets. CRISPR Off-target Predictor (CROP) is the latest developed software to predict off targets that lets the user utilize their own genome of interest. It provides two advantages with respect to other softwares, one is the mapping of different guide variants into the target genome and the flexibility of including any genome as references. Different guide variants are created including substitution, deletion or insertion up to four positions in the gRNA sequence. DNA: RNA hetero-duplex is simulated with mismatch, RNA and DNA bulges and gRNA variants are mapped to the reference genome. These mapped variants are utilized for the estimation of the off-target propensity score of the corresponding gRNA[10]. Due to the lack of unbiased methods for off target prediction, development of learning models is not feasible for plants. A recent study has analysed various studies on the prediction of off-targets in plants. It has been reported that the majority of predictions were performed using *in silico* tools or BLAST and only few studies employed whole genome sequencing. Also, wild-type cas9 or SpCas9 were used in the majority of the studies[11].

Mechanisms of off-target occurrence in plants

Off-target effects refer to unintended modifications at genomic sites that resemble the target sequence but are not the intended editing locations. These unintended edits can lead to undesirable phenotypic changes and affect the stability and safety of the edited plants.

a) Sequence similarity

Off-target effects primarily occur due to the presence of genomic sequences that are similar, but not identical, to the intended target sequence. The guide RNA (gRNA) used in CRISPR-Cas systems can tolerate mismatches, especially if they occur outside the seed region (the 8-12 nucleotides proximal to the protospacer adjacent motif or PAM). This tolerance can result in the Cas9 protein binding to and cleaving at unintended sites[12].

b) PAM sequence flexibility

The PAM sequence is a short DNA sequence following the target DNA sequence necessary for Cas9 binding and cleavage. Variations in PAM sequences can influence off-target activity. For example, the widely used SpCas9 may accept NAG sequences in addition to recognizing the NGG PAM sequence, which increases the possibility of off-target effects[13].

c) Chromatin accessibility

Off-target action may also be impacted by chromatin accessibility. Even if they are not exact matches for the gRNA, areas of the genome that are more accessible and open to the Cas9-gRNA complex have a higher chance of being cleaved. In areas of the genome that are transcriptionally active or have less compact chromatin structures, this may result in off-target consequences[14].

Challenges in plant genome editing

a) Genome complexity

Plant genomes, with higher levels of polyploidy and repetitive sequences, are frequently more complicated than those of other organisms. Because there are more possible off-target sites that could share sequence similarities with the gRNA, this complexity raises the possibility of off-target effects[15].

b) Lack of reference genomes

The lack of high-quality reference genomes for many plant species makes it more difficult to forecast and identify off-target sites. It is difficult to create gRNAs with low off-target potential and to precisely identify off-target effects following genome editing in the absence of complete genomic information[16].

Strategies to minimize off-target effects

Efficient genome editing is based on the specificity of the gRNA that guides the cleavage of the reference genome with the help of cas9. To avoid off-targets, the efficient designing of gRNA is the most important step in CRISPR. The different parameters like length, number of mismatches and GC content of gRNA are important while designing the gRNA. 16-17 nucleotides length, more than 65 % GC content are shown to be efficient gRNA. Also, sequences with more than three mismatches should be avoided to reduce the occurrence of off-targets[17]. Mismatches should be avoided in the seed region for efficient editing. Use of chimeric RNA with partial DNA replacement has also shown reduction in off-target activity in human cells. Modifying the gRNA with the addition of 2′-O-methyl-3′- phosphonoacetate in phosphate backbone has also significantly reduced the off-target activity. Going for other variants of cas protein having more specificity for the target region

of the genome is another strategy to reduce off-targets. SpCas9, derived from *Streptococcus pyogenes*, reported to reduce the activity of off-targets. Also, transient transformation provides less time to CRISPR RNA for off-target activity. Hence, transient transformation with CRISPR RNA /RNP must be performed as compared to stable transformation[18].

Detection of off-targets

Detection of off-targets is the necessary step for performing genome editing using CRISPR. However, screening off-targets from the genome for increasing the specificity of gRNA for its target is a difficult task. Generally, the methodologies used for the detection of off-targets are divided into two categories: biased and unbiased.

Unbiased methods

Unbiased methods represent detection of potential off-targets *in vitro* where the reference genome is incubated with designer nucleases to introduce double strand breaks (DSBs), which are further detected by various assays. The assays are classified into *in vitro* genome-wide assays (*in vitro* detection) and cell-based genome-wide assays (*in vivo* detection).

***a) In vitro* detection**

In this assay, genome editing is performed in the tubes that lacks the cellular environment which may influence the off-target activity of a system. These include GUIDE-seq, CIRCLE-seq, CHANGE-seq, SITE-seq, End-seq, deep sequencing, and Digenome-Seq.

- **i. GUIDE-seq:** Genome-Wide, Unbiased Identification of DSBs Enabled by Sequencing: GUIDE-seq is a popular *in vitro* technique that captures double-strand breaks (DSBs) induced by CRISPR nucleases. It involves the integration of a short, double-stranded oligodeoxynucleotide (dsODN) tag at DSB sites, which can then be sequenced to identify off-target locations. This method is highly sensitive and provides a comprehensive profile of off-target effects across the genome[19].
- **ii. Digenome-seq:** Digested Genome Sequencing: Digenome-seq is an additional potent *in vitro* technique that introduces DSBs with genomic DNA treated with CRISPR-Cas9. Sequencing the digested DNA subsequently reveals both on-target and off-target cleavage sites by mapping the positions of these breaks. This approach allows high-resolution detection of off-target activity without any need for cell-based systems[20].
- **iii. CIRCLE-seq**: (Circularization for *in vitro* Reporting of Cleavage Effects by Sequencing): The latest technique known as CIRCLE-seq, the fragmented genomic DNA is circularized and then treated with CRISPR-Cas9.DSBs are then found by sequencing the circular DNA. Because it does not need cellular transformation and has a high sensitivity and specificity to identify off-target areas, this technique is useful for *in vivo* research[21].

b) *In vivo* detection

Designer nucleases are produced with the reference genome inside the cells in this test *in vivo* off-target detection assays include integrase defective lentiviral vectors (IDLVs), breaks labelling, enrichments on streptavidin and next-generation sequencing (BLESS), chromatin immunoprecipitation high throughput sequencing (Chip–seq), in situ Cas off-targets and their sequencing verification (DISCOVER–seq), *in vivo* off-target verification (VIVO), and genome-wide off-target analysis by two cell embryo injection (GOTI)[22]. The most recent method, known as TEG-seq (tag-enriched GUIDE-seq), involves PCR amplification using certain 5′ phosphorylated primers and differential marking of amplicons that have a double-stranded DNA tag (dsTag) inserted in DSB sites. Phosphorylation of the 5′ end of dsTag-containing amplicons causes them to ligate to a barcoded adaptor (BC-A) for enrichment and amplification[18]. Unbiased methods revealed that off-targets which are present adjacent to alternate PAM sites, may have large number of mismatches and can be cleaved at high frequencies than on-targets[23].

- **i. HTGTS** (High-Throughput, Genome-Wide Translocation Sequencing): When DSBs are mis-repaired and causes non-adjacent DNA ends to connect, chromosomal translocations take place, which is how HTGTS identifies off-target DSBs. This approach offers a comprehensive picture of the possible hazards connected to CRISPR editing by shedding light on off-target activity and the genomic environment in which it takes place[24].
- **ii. BLESS** (Breaks Labelling, Enrichment on Streptavidin and Sequencing): BLESS is a genome-wide, unbiased method that directly labels DSBs in living cells. Biotinylated

oligonucleotides are used to bind to DSBs, which are subsequently sequenced and enriched using streptavidin beads. BLESS offers high-resolution DSB mapping and can be applied to *in vivo* detection of off-target locations[25].

The unbiased techniques are time-consuming, expensive, and provide noisy measurements that often gives false negative results. To overcome these limitations, biased methods based on *in silico* analysis have been developed to identify off-target regions.

Biased methods

Biased methods use the prior knowledge of the target sequence for detection of off-target. These methods are comparatively cheap and faster than unbiased methods.

a) *In silico* predictions

Sequences having similarity with gRNA are predicted through *in silico* tools and are verified by empirical methods. To find regions with high similarity to the gRNA, the prior knowledge is required. This methodology is cost-effective and easier to implement with minimal equipment requirement[26]. Different softwares developed by various research groups for the identification of off-targets are available and are categorized into alignment based, scoring based algorithms and data driven algorithms (Table 1). Following are the different categories of models based on which softwares for off-target prediction are developed.

- **i) Alignment based models:** It is the conventional algorithm based on the sequence homology of gRNA and reference genome on alignment. Data generated by different research groups on CRISPR/cas is used for the creation and refinement of the algorithm. The tools are based on different search engines including BLAST, BWA and Bowtie. CasOT, Target Finder, Cas-OFFinder, FlashFry, E-CRISP, CrisFlash, CasFinder, CRISPR Design Tool and Breaking-cas are the commonly utilized tools for off-targets prediction.
- **ii) Scoring based models:** It is the advanced algorithm based on the selection score of gRNA with mimimal off- target sites. Different softwares are available that utilize scoring strategies MIT specificity score or Cutting frequency determination (CFD). CHOPCHOP and CRISPOR gRNA designing tools are based on MIT score, whereas GuideScan and CRISPRs can are based on CFD score[18].
- **iii) Data driven based models:** These are the learning-based algorithms based on the various aspects of machine learning to effectively predict off-targets. Machine learning as well as deep learning models are applied for off-target prediction. However, deep learning is gaining a lot of attention for the detection of off-targets due to the availability of large data sets from CRISPR studies. Learning model construction has several hurdles such as data collection and integration from diverse cell types and experiments, inadequate labelling of data, data imbalance where true off-target sites are relatively small as compared to other mismatch regions, and insufficient information of leading sequence and epigenetic features affecting gRNA efficacy[27].

The construction of a training dataset using a variety of data inputs is taken into account by a machine learning algorithm, which also includes a set of features that can be used to predict cleavage efficiency. A machine learning-based computational model CRISPcut was developed by Dhanjal and coworkers for the detection of positive off-targets. Gradient boosted regression tree algorithm was utilized for estimating the likelihood of positive off-targets. For determining the cleavage activity at various genomic loci, an experimentally obtained dataset of gRNAs, their targets, and related off-targets as well as accessibility, mismatches, GC content, and position-specific conservation of nucleotides were taken into consideration[28]. CRISPR Target Assessment (CRISTA) is also one of the machine learning based tool trained as a random forest regression model, which was built on cleavage efficiencies described in the experimental studies[23]. CRISPR pred is another tool based on Support Vector Machine (SVM), Linear Regression (LR) and Random Forest (RF) algorithms. Here, sequence as well as secondary structure features are considered for features construction[29]. Effective off-target detection requires a sufficient sampling of a growing dataset of various gRNAs with their target sites and experimentally identified off-target sites in order to extract some general features associated with off-target sites cleaved by the CRISPR/Cas9 system under *in vivo* conditions. Deep learning algorithms have outperformed the existing prediction methods for off-targets. It consists of two steps where the first step is the conversion of

Table 1 — Different softwares/tools available for identification of off-targets

S. No	Tool	Description	Web link/Source
1	CALITAS	Using its Search Reference mode, CALITAS is designed to detect alignments with many gaps in a reference, returning the single best alignment for each off-target site. Genomic variations are also easily included in CALITAS off-target searches.	Command line
2	CasOT	Predicts probable off-target sites based on protospacer adjacent motif and user specified number of mismatches.	http://eendb.zfgenetics.org/casot/
3	COSMID	Detects potential off-targets with user specified number of mismatches and insertions or deletions.	https://crispr.bme.gatech.edu/
4	Crisflash	A potential off-target discovery tool which allows for the inclusion of genetic variant data specific to each individual.	Command line
5	CRISPOR	It identifies guide RNAs in an input sequence and ranks them based on various metrics that assess the likelihood of off-targets in the target genome and forecast on-target activity.	http://crispor.tefor.net/
6	CRISPRitz	A suit of software tools that supports genomic searches by considering genetic variants as well as DNA/RNA bulges and mismatches that the user specifies.	Command line
7	CRISPR-OFF	Predicts potential off-targets with up to 6 mis-matches for the gRNA and also considers PAM sequences for the analysis.	https://rth.dk/resources/crispr/crisproff/
8	CRISTA	Detects potential targets using the Burrows-Wheeler Aligner (BWA) and predicts the cleavage scores for each potential target based on the entire set of characteristics.	https://crista.tau.ac.il/
9	dsNickFury	Screening of off-targets is performed based on potential PAM and user-defined mismatch number.	Command line
10	FlashFry	A quick and adaptable command-line programme that allows users to provide an unconstrained number of mismatches to potential off-targets and off-target scoring criteria.	Command line
11	Off-Spotter	Detects off-target sites based on PAM sequence and number of mismatches.	https://cm.jefferson.edu/Off-Spotter/
PLANTS			
12	Cas-OFFinder	Detects off-targets based on number of mismatches and variations in protospacer-adjacent motif.	http://www.rgenome.net/cas-offinder
13	CCTOP	With appropriate default settings that are simple to adjust, CCTop offers a user-friendly user interface. With respect to their off-target quality, CCTop extracts all putative sgRNA target sites from a specified query sequence, rates them, and provides complete documentation.	https://cctop.cos.uni-heidelberg.de:8043/
14	Chop-Chop	It detects off-target binding of single-guide RNAs (sgRNAs) and TALENs while utilising effective sequence alignment methods to reduce search times.	https://chopchop.cbu.uib.no/
15	CROP	It uses combinations of substitution, deletion, and insertions which are mapped into the reference genome. These mapped variants serve as the basis for scoring and alignment.	https://github.com/vaprilyanto/crop
16	E-CRISP	A computational tool used to design and assess guide RNAs for CRISPR/Cas9. To find target sequences for facilitated genome editing, the web programme employs quick algorithms.	http://www.e-crisp.org/E-CRISP/

DNA and gRNA sequence pairs into vector or matrix representation followed by utilization of different network models to obtain features and make predictions about the sequence pairs. New tools based on deep learning are being developed by different research groups and are utilized for efficient detection of off-targets in the genome. Deep neural networks include recurrent neural networks (RNN), convolutional neural networks (CNN), fully connected neural networks (FNN), and their variants.

On the basis of encoding approach of gRNA sequence, deep learning models are categorised into two categories:

i) **Methods in the spatial domain:** CNN-based methods to detect gRNA on-target activity or off-target effect. gRNA base sequence inputs are compiled into two-dimensional image data, and a convolution layer is utilized to extract potential features in the spatial domain.

ii) **Methods in the temporal domain:** These methods consider the nucleotide in the gRNA sequence as a word, then a trainable matrix (could be either supervised or unsupervised) is used to project the word to the dense real-valued space. This technology is called embedding, which generates the base embedding[30].

DeepCRISPR, AttnToMismatch_CNN, CNN_Std, and CnnCrispr are convolutional neural network (CNN) based methods to predict gRNA off-target activity. Sequence features are automatically recognized and a complex process of feature extraction and off-target prediction is implemented in these methods. AttnToMismatch_CNN and AttnToCrispr_CNN utilize deep learning architectures for sequential analysis: attention-based transformer[31]. DeepCRISPR, is also a deep learning-based tool that utilizes hybrid deep neural networks for model training and prediction. A deep unsupervised representation learning technique has been employed to automatically learn the underlying representation of gRNAs using a complete set of genome-wide unlabelled gRNAs. The learned model was further tuned by a supervised deep neural network using the existing labelled gRNAs. It is the latest tool that predicts both off-target and on-target sites. Elevation is the most efficient tool for the prediction of all cleavage sites in the reference genome[27]. CnnCrispr is a deep learning-based method for prediction of gRNA off-target cleavage propensity. In CnnCrispr, the global and statistical information of input sequences is extracted by the GloVe embedding model with the construction of the co-occurrence matrix of gRNA and its corresponding DNA sequence. Data integration with the deep neural network model leads to the prediction of off-target propensity of a given gRNA at specific DNA fragments[32]. Crispr2vec is based on a metric learning approach in which a neural network is employed to generate an embedding for estimating the distances between sequences[33]. Peng and co-workers reported using an ensemble SVM classifier to predict off-target sites. Ensemble learning is a good strategy to improve the prediction accuracy and stability of the extremely imbalanced collected datasets[34]. CRISPR-IP (CRISPR model based on Identity and Position) is one more deep learning model that identifies the features of base pairs through CNN, utilizes BiLSTM to obtain the position features of base pairs and extracts the sequence pair features through the attention layer. Compared to other deep learning models, the model has shown good performance in predicting the possibility of off-targets for each gRNA-DNA pair[35]. R-CRISPR is another off-target prediction model for gRNA-DNA sequence pair utilizing recurrent convolutional network.

b) Experimental validation

The predicted off-targets need to be verified in the target cell or tissue. PCR based techniques are utilized where the hetero-duplexes formed between wild and mutant stands are cleaved by nucleases T7E1 or surveyor, and the cleaved product is analysed by electrophoresis. Targeted amplicon sequencing is another method of validating the newer methods. In some cases, data generated by biased techniques cannot be captured by traditional methods. Another technique that needs a known primer site for target enrichment is called uni-directional targeted sequencing (UDiTaS), which can find translocations, inversions, and massive deletions that other techniques fail to detect[26].

i) **T7 endonuclease I (T7EI) assay**: The T7EI assay is a widely used method for detecting off-target cleavage. It involves amplifying the genomic regions surrounding predicted off-target sites using PCR, followed by digestion with T7 endonuclease I, which cleaves mismatched heteroduplex DNA. The resulting fragments are analyzed by gel electrophoresis to identify off-target cleavage events[36].

ii) **Surveyor nuclease assay**: Similar to the T7EI assay, the Surveyor nuclease assay detects DNA mismatches by cleaving heteroduplex DNA at mismatch sites. This assay is used to validate predicted off-target sites and provides a simple and rapid means of detecting off-target activity in genomic DNA[37].

iii) **High-throughput sequencing (HTS):** Next-generation sequencing (NGS) techniques can be employed to validate off-target predictions by sequencing the PCR amplicons of potential off-target sites. This method provides precise identification and quantification of off-target mutations, offering a high-resolution approach to off-target validation[38].

iv) **Droplet digital PCR (ddPCR)**: It is a quantitative and sensitive method used for detection of rare off-targets. This involves, making DNA sample into droplets and perform PCR within each droplet. Analyses the droplets to detect and quantify off-target mutations. It is useful for detecting low-frequency off-targets[39].

Biased methods for off-target detection leverage computational predictions and empirical validations to identify and confirm off-target sites. They provide efficient approach with low cost for screening and monitoring of CRISPR application.

Conclusion

CRISPR-Cas, gene editing technique, has revolutionized the plant biotechnology by enabling the

precise modification of gene of interest. This technique can be utilized for different applications like disease resistance or tolerance to particular stress or for enhancing the yield and for improving the quality traits. Despite its advantage of precise gene editing, CRISPR-Cas, the effect of off-target remained a significant concern. This paper provides the overview of off-targets, methods for detection of off-targets in CRISPR gene editing technique. The gRNA sequence/specificity is the most important for successful use of CRISPR technology. The binding of gRNA to untargeted/off-target sites effects the genome editing adversely. The effect of off-target binding is more severe in case of humans and animals compared to plants. For getting the desired results from gene editing the effect of off-targets should be eliminated or on-target specificity should be improved. There are many tools, some are mentioned here, for off-targets detection in plants. Various *in silico* tools, *in vitro* and *in vivo* detection methods are available for detecting the off-targets. Due to insufficient availability of experimental data regarding off-targets, learning based methods that are gaining attention in clinical research, are not yet employed in the off-targets detection of plants. Future work to improve off-target analysis should include increasing the true off-target editing datasets so as to assess new experimental techniques and to train deep learning algorithms to improve the sensitivity of off-target editing quantification.

Acknowledgement

The authors extend their gratitude to ICAR - National Academy of Agricultural Research Management, Hyderabad for providing required infrastructure.

Conflicts of interest

All authors declare no conflict of interest.

References

1 Carroll D, Genome Editing: Past, Present, and Future. *Yale J Biol Med*, 90 (2017) 653.

2 Lino AC, Harper JC, Carney JP & Timlin JA, Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv*, 25 (2018) 1234.

3 Gao C, The future of CRISPR technologies in agriculture. *Nat Rev Mol Cell Biol*, 19 (2018) 275.

4 Jaganathan D, Ramasamy K, Sellamuthu G, Jayabalan S & Venkataraman G, CRISPR for Crop Improvement: An Update Review. *Front Plant Sci*, 9 (2018) 985.

5 Vicente MM, Chaves-Ferreira M, Jorge JMP, Proença JT & Barreto VM, The Off-Targets of Clustered Regularly Interspaced Short Palindromic Repeats Gene Editing. *Front Cell Dev Biol*, 9 (2021).

6 Manghwar H, Li B, Ding X, Hussain A, Lindsey K, Zhang X & Jin S, CRISPR/Cas Systems in Genome Editing: Methodologies and Tools for sgRNA Design, Off-Target Evaluation, and Strategies to Mitigate Off-Target Effects. *Adv Sci*, 7 (2020).

7 Hahn F & Vladimir Nekrasov, CRISPR/Cas precision: do we need to worry about off-targeting in plants?. *Plant Cell Rep*, 38 (2019) 437.

8 Chandrasekaran J, Brumin M, Wolf D, Leibman D, Klap C, Pearlsman M, Sherman A, Arazi T &Gal-On A, Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Mol Plant Pathol*, 17 (2016) 1140.

9 Xie K & Yang Y, RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol Plant*, 6 (2013) 1975.

10 Aprilyanto V, Aditama R & Tanjung ZA, CROP: a CRISPR/Cas9 guide selection program based on mapping guide variants. *Sci Rep*, 11 (2021) 1504.

11 Sturme MHJ, van der Berg JP, Bouwman LMS, De Schrijver A, de Maagd RA, Kleter GA & Battaglia-de Wilde E, Occurrence and Nature of Off-Target Modifications by CRISPR-Cas Genome Editing in Plants. ACS *Agric Sci Technol*, 2 (2022).

12 Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V & Zhang F, DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*, 31 (2013) 827.

13 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA & Charpentier E, A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337 (2012) 816.

14 Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon, DB & Zhang F, Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol*, 32 (2014) 670.

15 Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H & Musunuru K, Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell*, 15 (2014) 27.

16 Zhang F, Wen Y & Guo X, CRISPR/Cas9 for genome editing: progress, implications and challenges. *Hum Mol Genet*, 23 (2018) R40.

17 Hajiahmadi Z, Movahedi A, Wei H, Li D, Orooji Y, Ruan H & Zhuge Q, Strategies to Increase On-Target and Reduce Off-Target Effects of the CRISPR/Cas9 System in Plants. *Int J Mol Sci*, 20 (2019) 3719.

18 Naeem M, Majeed S, Mubasher ZH and Ahmad I, Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing. *Cells*, 9 (2020).

19 Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V & Joung JK, GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol*, 33 (2015)187.

20 Kim D, Kim S, Park J & Kim JS, Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq. *Genome Res*, 25 (2015) 1270.

21 Tsai SQ, Nguyen NT, Malagon-Lopez J, Topkar VV, Aryee MJ & Joung JK, CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods*, 14 (2017) 607.

22 Lin J & Wong KC, Off-target predictions in CRISPR-Cas9 gene editing using deep learning. *Bioinformatics*, 34 (2018) i656.
23 Abadi S, Yan WX, Amar D & Mayrose I, A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. *PLoS Comput Biol*, 13 (2017) e1005807.
24 Hu J, Meyers RM, Dong J, Panchakshari RA, Alt FW & Frock RL, Detecting DNA double-stranded breaks in mammalian genomes by linear amplification–mediated high-throughput genome-wide translocation sequencing. *Nat Protoc*, 11 (2016) 853.
25 Crosetto N, Mitra A, Silva MJ, Bienko M, Dojer N, Wang Q & Bienko M, Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nature Methods*, 10 (2013) 361.
26 Atkins A, Chung CH, Allen AG, Dampier W, Gurrola TE, Sariyer IK, Nonnemacher MR & Wigdahl B, Off-Target Analysis in Gene Editing and Applications for Clinical Translation of CRISPR/Cas9 in HIV-1 Therapy. *Front Genome Ed*, 3 (2021).
27 Chuai&Guohui, Deep CRISPR: optimized CRISPR guide RNA design by deep learning. *Genome Biol*, 19 (2018) 80.
28 Dhanjal JK, Dammalapati S, Pal S & Sundar D, Evaluation of off-targets predicted by sgRNA design tools. *Genomics*, 112 (2020).
29 Rahman MK & Rahman MS, CRISPRpred: A flexible and efficient tool for sgRNAs on-target activity prediction in CRISPR/Cas9 systems. *PLoS One*, 12 (2017).
30 Xiao LM, Wan YQ & Jiang ZR, AttCRISPR: a spacetime interpretable model for prediction of sgRNA on-target activity. *BMC Bioinformatics*, 22 (2021) 589.
31 Vinodkumar PK, Ozcinar C & Anbarjafari G, Prediction of sgRNA Off-Target Activity in CRISPR/Cas9 Gene Editing Using Graph Convolution Network. *Entropy*, 23 (2021) 608.
32 Liu Q, Cheng X & Liu G, Deep learning improves the ability of sgRNA off-target propensity prediction. *BMC Bioinformatics*, 21 (2020).
33 Trivedi TB, Boger R, Kamath GM, Evangelopoulos G, Cate J, Doudna J & Hidary J, Crispr2vec: Machine Learning Model Predicts Off-Target Cuts of CRISPR systems. (2020).
34 Peng H, Zheng Y, Zhao Z, Liu T & Li J, Recognition of CRISPR/Cas9 off-target sites through ensemble learning of uneven mismatch distributions. *Bioinformatics*, 34 (2018) i757.
35 Zhang, Zhong-Rui & Jiang ZR, Effective use of sequence information to predict CRISPR-Cas9 off-target. *Comput Struct Biotechnol J*, 20 (2022) 650.
36 Vouillot L, Thelie A & Pollet N, Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3 (Bethesda)*, 5 (2015) 407.
37 Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J & Gerard GF, Mutation detection using Surveyor™ nuclease. *Biotechniques*, 36 (2004) 702.
38 Zhang XH, Tee LY, Wang XG, Huang QS & Yang SH, Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther Nucleic Acids*, 4 (2015) e264.
39 Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ & Colston BW, High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem*, 83 (2011) 8604.