Journal of Genetics and Genomics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

Journal homepage: www.journals.elsevier.com/journal-of-genetics-and-genomics/

Review

# Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing

Jinjing Li [a, b, 1], Shunyan Hong [a, b, 1], Wanjin Chen [a, b, **], Erwei Zuo [c, ***], Hui Yang [d, *]

[a] Department of Neurology and Institute of Neurology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China
[b] Fujian Key Laboratory of Molecular Neurology, Fujian Medical University, Fuzhou, 350005, China
[c] Lingnan Guangdong Laboratory of Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, 518124, China
[d] Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Research Center for Brain Science and Brain-Inspired Intelligence, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, 200031, China

ARTICLE INFO

*Article history:*
Received 7 October 2019
Received in revised form
5 November 2019
Accepted 14 November 2019
Available online xxx

*Keywords:*
Gene editing
CRISPR-Cas9
Base editors
Off-target

ABSTRACT

CRISPR-mediated genome editing is a revolutionary technology for genome manipulation that uses the CRISPR-Cas systems and base editors. Currently, poor efficiency and off-target problems have impeded the application of CRISPR systems. The on-target efficiency has been improved in several advanced versions of CRISPR systems, whereas the off-target detection still remains a key challenge. Here, we outline the different versions of CRISPR systems and off-target detection strategies, discuss the merits and limitations of off-target detection methods, and provide potential implications for further gene editing research.

Copyright © 2019, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. All rights reserved.

## Introduction

Efficient genome editing in cells and organisms holds great potential for both basic and applied science. Early genome editing tools, including zinc-finger nucleases (ZFNs) and transcriptional activator-like effector nucleases (TALENs), rely mainly on protein-DNA interactions (Joung and Sander, 2013; Urnov et al., 2010). Although these early gene editing tools enabled researchers to manipulate the genome, their manipulation was complex, making it difficult to apply this technology to a wide variety of organisms. The clustered regularly interspaced short palindromic repeats (CRISPR)-associated nuclease (Cas), an adaptive immune response in archaea, has been developed into a revolutionary tool used in gene editing (Doudna and Charpentier, 2014; Jinek et al., 2012). Cas9, CRISPR-associated protein 9 endonuclease derived from *Streptococcus pyogenes*, produces site-specific double-strand breaks (DSBs) in the genome under the direction of single-guide RNAs (sgRNAs). The DSBs of the genomic DNA activate endogenous repair mechanisms via homology-directed repair (HDR), nonhomologous end-joining (NHEJ), or microhomology-mediated end joining (MMEJ) (Sander and Joung, 2014).

Many researchers have recently combined the CRISPR-Cas9 systems' components with deaminase enzymes to achieve specific base editing. Base editing converts one base pair to another base pair at the target DNA or RNA locus without the need for DSBs or a donor template. DNA base editors (BEs) comprise a catalytically manipulated Cas9 variants (dCas9, Cas9n) fused to a deaminase enzyme, including cytidine deaminase (CBEs) (Komor et al., 2016) and adenosine deaminase (ABEs) (Gaudelli et al., 2017). BEs work efficiently and precisely, producing fewer undesired indels to genomic DNA than standard genome editing methods.

The CRISPR-Cas systems and BEs can both generate off-target activities that can be quite different from on-target sites and these may cause the functional disruption of normal genes or lead to the activation of susceptible genes (Fu et al., 2013; Kuscu et al., 2014; Sander and Joung, 2014; Zuo et al., 2019). However, the

* Corresponding author.
** Corresponding author. Department of Neurology and Institute of Neurology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China.
*** Corresponding author.
*E-mail addresses:* wanjinchen75@fjmu.edu.cn (W. Chen), erweizuo@163.com (E. Zuo), huiyang@ion.ac.cn (H. Yang).
[1] These authors contributed equally.

https://doi.org/10.1016/j.jgg.2019.11.002
1673-8527/Copyright © 2019, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. All rights reserved.

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002

detection of off-target activity associated with the CRISPR-Cas systems and BEs have been very challenging. In this review, we summarize different versions of the CRISPR-Cas systems and BEs, generalize the different strategies generated to identify off-target effects, and discuss the merits and the limitations of these off-target analysis methods.

## Development of the CRISPR-Cas systems

The CRISPR system is an adaptive immune system of archaea and various bacteria, and it has played a revolutionary role in the field of gene editing. Unlike previous gene editing systems (ZFNs and TALENs), CRISPR-Cas9, a kind of RNA-guided endonuclease (RGEN), uses the combination of DNA and RNA to guide endonucleases (Cas) to target sites. The RNA used to bind DNA is known as sgRNA, and sgRNA generally consists of ~20 nucleotides at the 5' end of the protospacer adjacent motif (PAM). These nucleotides guide the Cas protein to target genomic sites and induce precise DSBs (Doudna and Charpentier, 2014; Friedland et al., 2013). Subsequently, DSBs can then be repaired by three DNA repair pathways: NHEJ, HDR, and MMEJ. HDR is a precise but infrequent method for correcting harmful mutations in the presence of homologous recombination templates. In contrast, NHEJ and MMEJ are error-prone repair methods that introduce insertions or deletions at targeted site, but these methods are dominant in most mammalian cells (Mao et al., 2008). Thus, NHEJ is commonly used to construct gene knockdown and gene knockout models, or to produce exon skipping events using frameshift mutations. In most cases, HDR is less efficient than the NHEJ and MMEJ pathways in most cell lines (Mali et al., 2013b; Song and Stieger, 2017). Cas nucleases derived from different bacteria have been discovered and applied to gene editing, among which *Streptococcus pyogenes* (SpCas9) and *Staphylococcus aerous* (SaCas9) are the most commonly used. To produce a specific DSB, each CRISPR system differs in terms of sgRNA design and PAM requirements: NGG for SpCas9 and NNGRRT for SaCas9. The use of different PAM choices greatly broadens the targeting range of the CRISPR system in gene editing.

Classical SpCas9 uses two nuclease domains, RuvC1 and HNH (Gasiunas et al., 2012; Jinek et al., 2012), and mutation of D10A or H840A can inactivate these two domains, respectively, producing nickases (known as Cas9n) that can cleave complementary or noncomplementary DNA chains by producing single strand breaks (Mali et al., 2013a; Shen et al., 2014). Some researchers have found that a paired-nicking strategy with Cas9n can introduce HDR-mediated DNA repair similar to classical SpCas9 at various sites, but this method introduces fewer base insertions or deletions mediated by NHEJ (Ran et al., 2013; Xu et al., 2014). Catalytically dead Cas9 (dCas9), which is inactivated in both nuclease domains (HNH and RuvC), retains the DNA binding ability while losing the DNA cleavage ability. dCas9 fused with transcriptional activation domains form an RNA-guided transcriptional activator that can target the promoter region of specific genes (Konermann et al., 2015; Maeder et al., 2013; Zhou et al., 2018). Conversely, dCas9 fused to transcriptional suppressors results in decreased expression of targeted genes (Gilbert et al., 2013; Qi et al., 2013).

Early versions of CRISPR-Cas9 systems were far from ideal. Many attempts were made to evolve the performance of these systems (Table 1). To facilitate the packaging of Cas9 in viral vectors, smaller endonucleases were constantly being discovered and used. SaCas9 (Ran et al., 2015), using the NNGRRT PAM, is about 1 kb smaller than SpCas9 (4.2 kb), which makes it more suitable for packaging into viral vectors for therapeutic applications. KKH SaCas9, a variant of SaCas9, contains the E782K/N968K/R1015H substitutions with NNNRRT PAM, and these modifications increased the SaCas9 targeting range by 2–4 folds (Kleinstiver et al., 2015). *Streptococcus thermophilus* Cas9 (St1Cas9) (Esvelt et al., 2013) and *Neisseria meningitidis* Cas9 (NmCas9) (Hou et al., 2013) belong to the Class 2 Type II immune system. They are about 1 kb smaller (3.4 and 3.2 kb) than SpCas9, with more complex PAM sequences of NNAGAAW and NNNNGMTT. The longer PAM of St1Cas9 and NmCas9 decreases undesired editing at potential off-target sites, although at the cost of reduced targeting range. Another group found *Acidaminococcus*sp. *BV3L6* (AsCpf1, 3.9 kb) and *Lachnospiraceae bacterium ND2006* (LbCpf1, 3.7 kb), two single CRISPR-derived RNA (crRNA)-guided endonucleases that belong to the Cpf1 family, contained a TTTN PAM and were active both *in vitro* and in mammalian cells (Zetsche et al., 2015). Moreover, Cpf1 endonucleases are thought to increase the efficiency of HDR because they are able to produce staggered ends, and this may facilitate gene editing (Zetsche et al., 2015).

SpCas9 has been the most commonly used endonuclease in gene editing, and many variants were scrutinized for the purpose of broadening the targeting range, improving Cas9 cutting efficiency, and reducing the possibility of off-target activity. The engineered VQR (D1135V/R1335Q/T1337R), EQR (D1135E/R1335Q/T1337R), and VRER (D1135V/G1218R/R1335E/T1337R) SpCas9 variants recognize NGA, NGAG, and NGCG, respectively, and they greatly expand the targeting range of CRISPR-mediated gene editing (Hirano et al., 2016; Kleinstiver et al., 2015). SpCas9-HF1, a high-fidelity variant harboring alterations (N497A/R661A/Q695A/Q926A) designed to reduce nonspecific DNA contacts, retains the same high on-target efficiency as wild-type SpCas9 and rarely produces off-target events (Kleinstiver et al., 2016). Furthermore, using structure-guided protein engineering, Feng Zhang's laboratory developed a new SpCas9 variant (K848A/K1003A/R1060A) termed enhanced specificity SpCas9 (eSpCas9) that improved the specificity of SpCas9 and reduced off-target effects by weakening nontarget strand binding and encouraging rehybridization to the target DNA strand (Slaymaker et al., 2016). To further solve the off-target problem without compromising relative cleavage activity, several variants with high fidelity were engineered, including HeFspCas9, EvoCas9, HypaCas9, and sniper-Cas9. (Casini et al., 2018; Chen et al., 2017; Kulcsar et al., 2017; Lee et al., 2018). Nishimasu et al. reported an engineered SpCas9 variant (SpCas9-NG), with amino acid alterations of R1335A/L1111R/D1135V/G1218R/E1219F/A1322R/T1337R, that enable relaxed NG PAM recognition and increase activity at the NGH (H = A, T, or C) PAM (Nishimasu et al., 2018). Subsequently, a new variant named xCas9, created by phage-assisted continuous evolution, enables SpCas9 to reach more relaxed PAM recognition, including NG, GAA, and GAT (Hu et al., 2018). Notably, xCas9 greatly broadened the targeting range of SpCas9 and showed higher specificity than SpCas9.

In addition to explore Cas9 variants and engineering Cas9, researchers also investigated the impact of length and other features of guide RNAs (gRNAs) on the off-target frequency. It has been reported that the truncated gRNAs (17 or 18 nucleotides) can increase the specificity dramatically without sacrificing on-target gene editing efficiency (Fu et al., 2014). Ryan et al. have developed a novel way to enhance specificity by inducing chemical modifications (2′-O-methyl-3′-phosphonoacetate, or MP) in gRNAs at specific sites in their DNA recognition sequence (Ryan et al., 2018). This method reveals a unique way for improving specificity while maintaining high on-target efficiency. Another group focusing on gRNA modifications showed that incorporation of next-generation bridged nucleic acids (2′,4′-BNANC[N–Me]) and locked nucleic acids (LNA) at specific sites within crRNAs could greatly enhance the target specificity *in vitro* and in cells (Cromwell et al., 2018). Yin et al. revealed that off-target activity could be decreased by replacing RNA nucleotides with DNA nucleotides in crRNA (Yin

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002

**Table 1**
The main features, purposes of improvements and PAMs of various CRISPR systems.

| Nuclease | Feature of improvements | PAM | Major purpose of improvements |
|---|---|---|---|
| SpCas9 | NLS-SpCas9-NLS structure, *Streptococcus pyogenes* derived Cas9 protein | NGG | Classical Cas9 |
| SaCas9 | *Staphylococcus aerous* derived Cas9 protein replacement | NNGRRT | Reduce size |
| SaKKHCas9 | E782K/N968K/R1015H in SaCas9 protein | NNNRRT | Reduce size and extend PAM |
| St1Cas9 | *Streptococcus thermophiles* derived Cas9 protein replacement | NNAGAAW | Improve specificity |
| NmCas9 | *Neisseria meningitides* derived Cas9 protein replacement | NNNNGMTT | Improve specificity |
| AsCpf1 | *Acidaminococcus* derived nuclease | TTTN | Increase the efficiency of HDR |
| LbCpf1 | *Lachnospiraceae* derived nuclease | TTTN | Increase the efficiency of HDR |
| VQR-SpCas9 | D1135V/R1335Q/T1337R in SpCas9 | NGA | Extend PAM |
| EQR-SpCas9 | D1135E/R1335Q/T1337R in SpCas9 | NGAG | Extend PAM |
| VRER-SpCas9 | D1135V/G1218R/R1335E/T1337R in SpCas9 | NGCG | Extend PAM |
| SpCas9-HF1 | N497A/R661A/Q695A/Q926A in SpCas9 | NGG | Improve specificity |
| eSpCas9 | K848A/K1003A/R1060A in SpCas9 | NGG | Improve specificity |
| HeFSpCas9 | N497A/R661A/Q695A/K848A/Q926A/K1003A/R1060A in SpCas9 | NGG | Improve specificity |
| EvoSpCas9 | M495V/Y515N/K526E/R661Q in SpCas9 | NGG | Improve specificity |
| HypaCas9 | N692A/M694A/Q695A/H698A in SpCas9 | NGG | Improve specificity |
| Sniper-Cas9 | F539S/M763I/K890N in SpCas9 | NGG | Improve specificity |
| xCas9(3.7) | A262T/R324L/S409I/E480K/E543D/M694I/E1219V in SpCas9 | NG/GAA/GAT | Extend PAM |
| SpCas9-NG | R1335A/L1111R/D1135V/G1218R/E1219F/A1322R/T1337R in SpCas9 | NG | Extend PAM |

et al., 2018). This strategy of partial DNA replacement maintains high on-target efficiency when applied to both crRNA and sgRNA, as well as with crRNA of Cpf1. It was also reported that delivery of RGN-ribonucleoproteins (RNPs) showed positive effects on improving editing efficiency and specificity (Lone et al., 2018; Rees et al., 2017).

## Off-target activities of CRISPR-Cas9

Programmable endonucleases can induce DSBs at target sites in genomic DNA, but can also generate unwanted cleavages outside of on-target sites and produce off-target mutations. Cleavage at off-target sites can lead to chromosomal rearrangements, including deletions, insertions, and translocations, which may result in the disruption of normal gene expression and the activation of oncogenes (Cho et al., 2014; Choi and Meyerson, 2014; Veres et al., 2014). In general, off-target events can be divided into two categories: sgRNA-dependent off-target and sgRNA-independent off-target events. The sgRNA-dependent off-target events are because of the similarity of sequences between the sgRNA and the off-target sites (Cong et al., 2013; Fu et al., 2013; Hsu et al., 2013b). Generally, prime-template pairing can tolerate mismatches up to 8 bases, and such mispairing can result in abnormal cleavages outside the target site of CRISPR/Cas9. On the contrary, sgRNA-independent off-target events are base alterations in the sites that differ from sgRNA sequences, which are more difficult to define. Detecting and preventing the off-target events remains a key challenge in basic research and therapeutic applications. Various methods have been applied for detecting off-target sites. These methods can be classified as biased and unbiased; the biased method is off-target detection of predicted sites, and the unbiased method is off-target analysis over the entire genome.

### *Biased detection of off-target activities by CRISPR-Cas9*

The earliest off-target detection method predicted potential off-target sites based on the sequence of sgRNA. Many online tools, such as the Feng Zhang lab's Target Finder (http://crispr.mit.edu/) (Hsu et al., 2013a), the CasFinder (http://arep.med.harvard.edu/CasFinder/) (Couvin et al., 2018), and Cas-OFFinder (http://www.rgenome.net/cas-offinder/) (Doench et al., 2016), can be used for prediction of off-target sites. Potential off-target sites predicted by these online tools were amplified by polymerase chain reaction (PCR) and then analyzed by a T7EI (T7 endonuclease I) assay, Sanger sequencing, or high-throughput sequencing. These biased detection methods are simple, rapid, and cost-effective for the detection of off-target mutations, and thereby the most widely used methods to detect mutations induced by CRISPR systems. The T7EI assay recognizes and cleaves heteroduplexes at mismatches DNA and was initially used to detect off-target activities. However, this assay suffers poor sensitivity; it cannot detect off-target mutations that occur at frequencies less than 1%, and it is not practical cost-effective for high-throughput screening (Mashal et al., 1995). Sanger sequencing is used to verify the deletions, insertions or alterations of base pairs in potential off-target sites, but this method is time consuming and inefficient, and because of the low-sensitivity, it is unable to determine the extent of editing of the off-target sites. Although the T7EI assay and Sanger sequencing detect mutations, they cannot always reveal details of the mutation structure. Another method is to use high-throughput sequencing to detect the editing of potential off-target sites after PCR amplification. This method is more sensitive than Sanger sequencing and could enable accurate and high-throughput measurement of hundreds of off-target sites at once.

### *Unbiased detection of off-target cleavage*

Unbiased off-target analyses can detect unwanted cleavages on the whole-genome level; these methods are based on deep sequencing and include whole genome sequencing (WGS), cross-linking chromatin immunoprecipitation of endonuclease-mutant Cas9 (ChIP-dCas9), breaks labeling, enrichments on streptavidin and next-generation sequencing (BLESS), linear amplification-mediated high-throughput genome-wide translocation sequencing (LAM-HTGTS), genome-wide, unbiased identification of DSBs enabled by sequencing (GUIDE-seq), digested genome sequencing (Digenome-seq), circularization for *in vitro* reporting of cleavage effects by sequencing (CIRCLE-seq), selective enrichment and identification of tagged genomic DNA ends by sequencing (SITE-seq), and discovery of *in situ* Cas off-targets and verification by sequencing (DISCOVER-seq) (Fig. 1).

#### *Whole genome sequencing (WGS)*

Whole genome sequencing, the earliest unbiased way to detect off-target mutations, is viable to detect not only small indels and single nucleotide polymorphisms (SNPs) but also structural

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002

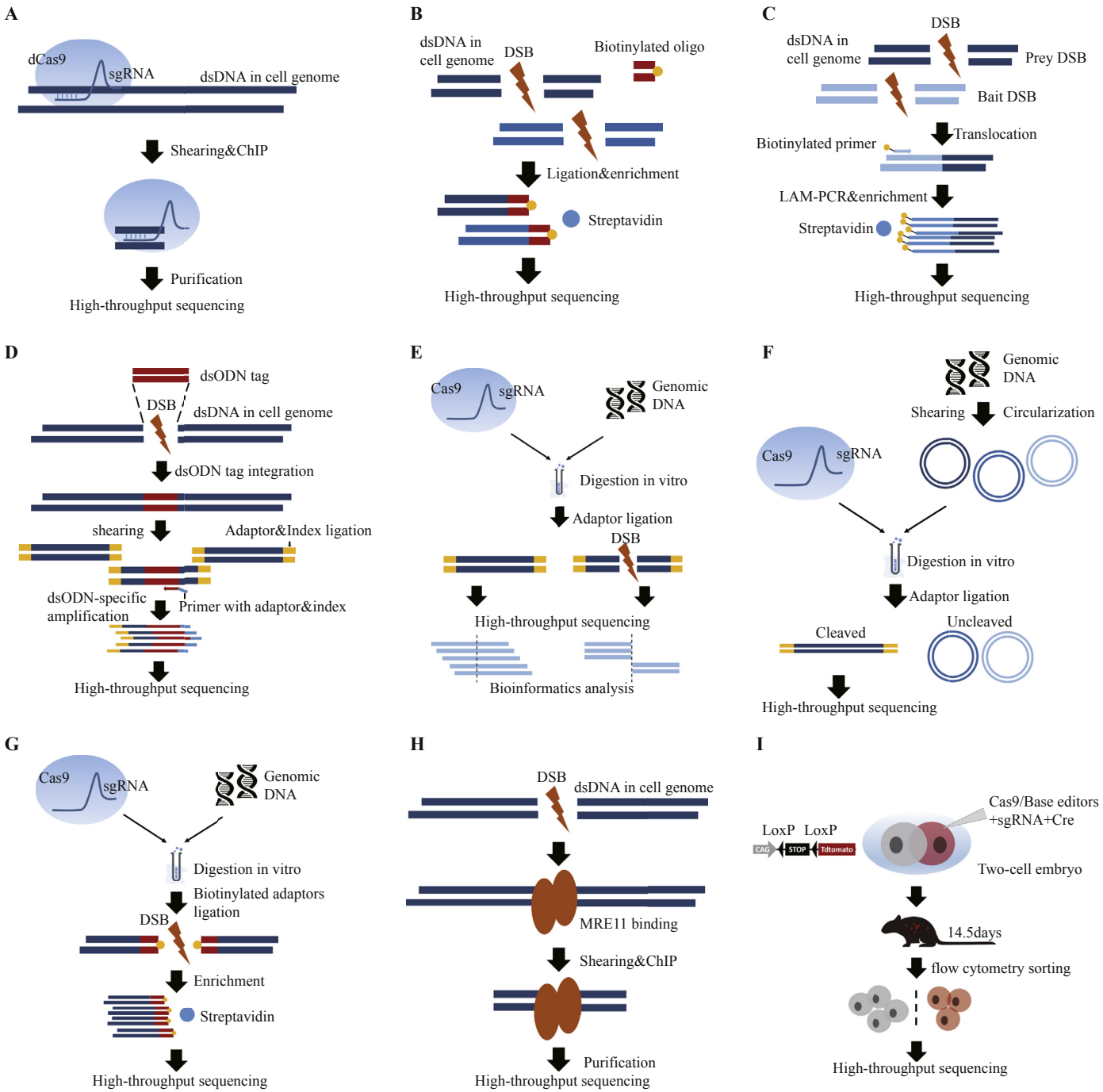



**Fig. 1.** Overview of unbiased off-target detection workflow. **A**: Cross-linking chromatin immunoprecipitation of endonuclease-mutant Cas9 (ChIP-dCas9). **B**: Breaks labeling, enrichments on streptavidin and next-generation sequencing (BLESS). **C**: Linear amplification-mediated high-throughput genome-wide translocation sequencing (LAM-HTGTS). **D**: Genome-wide, unbiased identification of DSBs enabled by sequencing (GUIDE-seq). **E**: Digested genome sequencing (Digenome-seq). **F**: Circularization for *in vitro* reporting of cleavage effects by sequencing (CIRCLE-seq). **G**: Selective enrichment and identification of tagged genomic DNA ends by sequencing (SITE-seq). **H**: Discovery of *in situ* Cas off-targets and verification by sequencing (DISCOVER-seq). **I**: Genome-wide off-target analysis by two-cell embryo injection (GOTI).

variants such as inversions and rearrangements (Koo et al., 2015; Schaefer et al., 2017; Wu et al., 2014a). The off-target mutations can be detected by sequencing the whole genome before and after gene editing and determined by simple analysis of new mutations generated outside the on-target locus compared with the original population. This method can be used to analyze a single clone, genome edited animals, and several species of plants. However, whole-genome sequencing lacks the sensitivity required to detect off-target mutations in large populations and may miss many low-frequency off-target sites.

*Cross-linking chromatin immunoprecipitation of endonuclease-mutant Cas9 (ChIP-dCas9)*

Cross-linking chromatin immunoprecipitation (ChIP) is a standard method to localize nuclear proteins in the genome (Park, 2009). The combination of ChIP with next-generation sequencing technology (NGS) (ChIP-seq) enables the detection of the protein-DNA interactions at the genome level (Furey, 2012). The interaction between the dCas9-sgRNA complex and DNA ensured stable binding of the complex to the DNA locus without DSBs which enables the application of ChIP-seq for off-target detection. ChIP-

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002

dCas9 provides a readout of Cas9 binding sites in the genome, which include both on-target and off-target sites. Although ChIP-seq provides important information on potential CRISPR/Cas9 system binding sites, it does not always correlate with true off-target sites. Some groups observed little to none of cleavages in off-target sites predicted by ChIP-seq (Cencic et al., 2014; Kuscu et al., 2014; Wu et al., 2014b). The application of dCas9 rather than wild-type Cas9 and the cross-linking that stabilized the weak and transient interactions between dCas9-sgRNA complex and DNA may alter specificity in off-target detection (Teytelman et al., 2013; Wu et al., 2014a).

*Breaks labeling, enrichments on streptavidin and next-generation sequencing (BLESS)*

BLESS is a genome-wide *in vivo* method for detecting endonuclease-induced DSBs in fixed cells by direct *in situ* break labeling using biotinylated oligonucleotides. The labeling DSBs are then enriched and sequenced (Crosetto et al., 2013). The advantages of BLESS are that it can detect DSBs instantaneously and do not rely on endogenous cellular DNA damage repair machinery nor single-stranded DNA. One obvious limitation of BLESS is that it can only map unjoined ends present at a single moment, and it cannot detect existing DSBs before the cells are permeabilized, which may miss many off-target cleavage sites.

*Linear amplification-mediated high-throughput genome-wide translocation sequencing (LAM-HTGTS)*

LAM-HTGTS is an unbiased method to identify the chromosomal translocations caused by end joining between DSBs induced by endonucleases in cultured mammalian cells (Hu et al., 2016). HTGTS relies on fusion of "bait" DSBs and other DSBs ("prey" DSBs), which may cause chromosomal translocations. Bait-prey junctions are cloned from genomic DNA by LAM-PCR and ligated to bridge adapters, followed by PCR amplification of the single-stranded DNA junction library and subsequent NGS analysis. LAM-HTGTS can detect chromosomal translocations induced by DSBs that tend to be missed by other methods with high sensitivity, but the original method cannot detect the small insertions, deletions, or SNPs. As the rearrangement of chromosomes is relatively rare as a result of DNA DSBs, this method may miss a large number of off-target sites and may be better used when combined with other methods for performing off-target identification.

*Genome-wide, unbiased identification of DSBs enabled by sequencing (GUIDE-seq)*

GUIDE-seq is based on the incorporation of captured end-protected double-stranded oligodeoxynucleotides (dsODN) into the DSBs, after which the integrated dsODNs are amplified and sequenced to detect off-target sites. GUIDE-seq is able to capture off-target activities beyond bioinformatics predictions with indel frequencies as low as 0.03% (Tsai et al., 2015). Similar to BLESS, GUIDE-seq can only detect DSBs generated at special moments which account for 30%–50% of total DSBs. In addition, GUIDE-seq can detect nuclease-independent DNA breakpoint "hotspots." Thus, GUIDE-seq is used to study genome-wide DNA damages and whole-genome on- and off-target effects of the endonucleases.

*Digested genome sequencing (digenome-seq)*

Digenome-seq is a method for detecting endonuclease off-target activities *in vitro*. This method cleaves the genomic DNA *in vitro* using high concentrations of Cas9-sgRNA ribonucleoprotein (RNP) complexes which are combined by the endonuclease and the corresponding sgRNA; the resulting geomic fragments are sequenced by NGS (Kim et al., 2015). Thus, Digenome-seq can detect off-target sites with frequencies lower than 0.1% (Kim et al., 2015). The main advantage of the Digenome-seq is that the genomic DNA is digested *in vitro* with a high concentration of RNP, which can maximize the cleavage and increase the probability of off-target detection. In addition, this method does not rely on the intracellular environment; the operation is relatively simple and does not require PCR amplification. However, the cell-free environment may be a disadvantage of this method. The off-target found by the method may be different from other in vivo methods, which may lead to the occurrence of false positives (Fu et al., 2016).

*Circularization for in vitro reporting of cleavage effects by sequencing (CIRCLE-seq)*

CIRCLE-seq is an unbiased and sensitive method to detect off-target mutations induced by endonucleases on the genome-wide level. Similar to Digenome-seq, this off-target detection strategy was also performed in vitro (Tsai et al., 2017) and differs from in vivo detecting methods that rely on efficient transfection or an intracellular environment. CIRCLE-seq facilitates the in vitro nuclease-mediated cleavage of genomic DNA with high enrichment, enabling the detection of off-target mutations with low frequency (Lazzarotto et al., 2018). In addition, this method can be performed using widely accessible NGS technology without the need of reference genomic sequences. Briefly, genomic DNA is circularized and purified and then treated with Cas9-sgRNA RNP complexes. The resulting DNA fragments are ligated to adapters for NGS analysis. Results of off-target activities for the gRNA of the *HBB* gene showed that CIRCLE-seq could find 26 of the 29 off-target sites found by Digenome-seq, and also found 156 new sites, suggesting the high sensitivity of the method. However, each CIRCLE-seq sample needs a large amount of genomic DNA for circularization (~25 μg), which may be a limiting factor, depending on the cell source of genomic DNA.

*Selective enrichment and identification of tagged genomic DNA ends by sequencing (SITE-seq)*

SITE-seq is also an unbiased biochemical method using Cas9-sgRNA RNP complexes to detect off-target mutations within genomic DNA *in vitro* (Cameron et al., 2017). The cleaved genomic DNA is biochemically probed and enriched for NGS. Bioinformatics is used to analyze the most likely off-target sites. SITE-seq showed a higher sensitivity for off-target mutations when identified at lower concentrations in cells. Like Digenome-seq and CIRCLE-seq, SITE-seq is also an *in vitro* off-target detection method, independent of the cell's internal environment. Compared with Digenome-seq, SITE-seq requires the ligation of the cleaved DNA with biotinylated adapters and two rounds of enzymatic fragmentation are needed for fragment enrichment.

*Discovery of in situ cas off-targets and verification by sequencing (DISCOVER-seq)*

DISCOVER-seq is a recently discovered unbiased off-target detection method that is capable of *in vivo* detection. This strategy relies on endogenous DNA repair mechanisms that recruit DNA repair factor MRE11 to the sites of DSBs in the genome at the appropriate time (Wienert et al., 2019). The ChIP-seq technique is then used to analyze the MRE11 locations on the genome and identify the off-target sites. This method showed high accuracy in both the human K562 cell line and the mouse B16–F10 cell line, but the sensitivity was slightly lower than that of GUIDE-seq. Compared with GUIDE-seq, DISCOVER-seq avoids the cytotoxicity caused by transfection of the tag dsODNs, enabling sensitive detection of off-target sites in human induced pluripotent stem cells (hiPSCs) or other patient-derived cells.

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002

## Development of the base editors

Point mutations are a causative factor in many genetic syndromes, but traditional gene editing techniques can only correct point mutations via HDR. Generally, HDR repair is inefficient and usually introduces random insertions and deletions at the dsDNA breakpoint. Base editors that combine cytosine or adenosine deaminase with CRISPR systems have been developed recently to induce targeted base alterations in the DNA and RNA of many species (Table 2).

The first generation of base editors (BE1) combined cytosine deaminase (APOBEC1) (Harris and Liddament, 2004; Yang et al., 2017) and dCas9 (inactivated by mutations of D10A and H840A) under the guidance of sgRNA to facilitate the direct conversion of a base from C·G to T·A (Komor et al., 2016). BE1 was able to convert cytidine to thymidine within the window of −16 to −12 bases from the PAM of the sgRNA. The second-generation base editor (BE2) system added a uracil-DNA glycosylase inhibitor (UGI) to the BE1 system, which enhanced the editing efficiency of C to T replacement by three-fold, through inhibiting the base excision repair (BER) mediated error-free repair (Komor et al., 2016). To further improve the editing efficiency of BE2, Cas9n (D10A), which can nick the target strand, was applied to replace dCas9 and was combined with cytidine deaminase and a UGI, forming a third-generation base editor (BE3) with a six-fold increase in efficiency over BE2 (Komor et al., 2016). Compared with BE2, BE3 introduced more indels owing to the cleavage of the target strand (1.1% *vs* 0.1%). Notably, the base editing efficiency of the BE3 system has been improved, and BE3 has been broadly used for genome editing in bacteria, plants, mammalian cells, mice, and human embryos. To better inhibit endogenous base excision repair, Komor et al. fused another UGI copy into BE3 to develop BE4, which showed higher editing efficiency than BE3 (Komor et al., 2017).

Recently, Koblan et al. and Zafra et al. improved the expression of BEs by optimizing base editor codons (BE4max) (Koblan et al., 2018), and thereby improved the editing efficiency of basic editing in mammalian cells (~1.7- to 9-fold). In addition, AncBE4-max was developed using ancestral sequence reconstruction, which showed higher editing efficiency than BE4max at various sites (Zafra et al., 2018). DNA modifications, such as DNA methylation, greatly inhibit base editing efficiency. Wang et al. developed a human hA3A-derived BE system (hA3A-BE3) which could induce the highest editing efficiency in highly methylated regions to date (Wang et al., 2018). In parallel, many groups replaced SpCas9 in BE3 with various engineered endonucleases that recognize different PAM sequences, such as SaCas9, KKH SaCas9, VQR, EQR, VRER, xCas9, and SpCas9-NG (Esvelt et al., 2011; Hu et al., 2018; Kim et al., 2017; Nishimasu et al., 2018). These new BEs show different PAM preferences that greatly broaden the targeting range of the BEs. Moreover, to overcome the limitation of the CRISPR–Cas9 BEs for its G/C-rich PAM sequence, Li et al. fused the rat APOBEC1 in a noncatalytically active version of the *Lachnospiraceae* bacterial Cpf1 (Li et al., 2018). This editor (CRISPR-Cpf1-based BE) recognizes T-rich PAM sequences and catalyzes C to T conversion in human cells while inducing low levels of indels and off-target editing. Another BE named enAsCas12a-HF1 optimized by Kleinstiver et al. improves the efficiency of C to T conversion and reduces off-target effects (Kleinstiver et al., 2019). Meanwhile, a series of BE3s with narrower editing windows (YE1-BE3, YE2-BE3, EE-BE3, and YEE-BE3) were developed by introducing different engineered APOBEC1s (Kim et al., 2017). To improve the editing specificity, HF-BE3 was engineered by introducing mutations into nCas9 to decrease its DNA affinity (Rees et al., 2017). Recently, unexpected RNA off-target events have driven the investigation of novel variants of BEs, including BE3(APOBEC1$^{R33A}$), BE3(APOBEC1$^{R33A/K34A}$) (Grunewald et al., 2019), and BE3(APOBEC1$^{W90A+R126E}$) (Zhou et al., 2019), which can effectively decrease or eliminate the unwanted RNA editing activity. Zhou et al. also replaced APOBEC1 with human APOBEC3A (hA3A), which reportedly possessed only DNA but no RNA affinity (Wang et al., 2018). In addition, the point mutation R128A or Y130F was induced in h3A3 and consequently decreased the RNA off-target events to baseline levels (Zhou et al., 2019).

**Table 2**
Outline of structures, PAMs, and major purposes of improvement of base editors.

| CBE | Structure | ABE | Structure | PAM | Major purpose of improvements |
|---|---|---|---|---|---|
| BE3 | rA1-SpCas9n-UGI | ABE7.10 | TadA-TadA*- SpCas9n | NGG | |
| BE4 | rA1-SpCas9n-2XUGI | | | NGG | Increase efficiency |
| BE4max | rA1-codon optimized SpCas9n-2XUGI | ABEmax | TadA-TadA*- codon optimized SpCas9n | NGG | Increase efficiency |
| AncBE4max | rA1-ancestral SpCas9n-2XUGI | | | NGG | Increase efficiency |
| hA3A-BE3 | hA3A-SpCas9n-UGI | | | NGG | Increase efficiency |
| Sa-BE3 | rA1-SaCas9n-UGI | Sa-ABE | TadA-TadA*- SaCas9n | NNGRRT | Reduce size |
| SaKKH-BE3 | rA1-SaKKHCas9n-UGI | SaKKH-ABE | TadA-TadA*- SaKKHCas9n | NNNRRT | Reduce size and extend PAM |
| VQR-BE3 | rA1-VQR SpCas9n-UGI | VQR-ABE | TadA-TadA*- VQRSpCas9n-UGI | NGA | Extend PAM |
| VRER-BE3 | rA1-VRER SpCas9n-UGI | VRER-ABE | TadA-TadA*-VRERSpCas9n-UGI | NGCG | |
| xBE3 | rA1-xSpCas9n-UGI | xABE | TadA-TadA*-xSpCas9n | NG | Extend PAM |
| SpCas9-NG-BE3 | rA1-SpCas9n NG -UGI | | | NG | Extend PAM |
| dCpf1-BE | rA1-dCpf1-UGI | | | TTTV | Extend PAM |
| AsCas12a-HF1 | rA1-enAsCas12a variant -UGI | | | TTTV | Improve specificity |
| YE1-BE3 | rA1$^{W90Y/R126E}$ -variants-SpCas9n-UGI | | | NGG | Improve specificity |
| YE2-BE3 | rA1$^{W90Y/R132E}$ -variants-SpCas9n-UGI | | | | |
| EE-BE3 | rA1$^{R126E/R132E}$ -variants-SpCas9n-UGI | | | | |
| YEE-BE3 | rA1$^{W90Y/R126E/R132E}$ -variants-SpCas9n-UGI | | | | |
| HF-BE3 | rA1-SpCas9n HF-UGI | | | NGG | Improve specificity |
| BE3 (APOBEC1$^{R33A}$) BE3 (APOBEC1$^{R33A/K34A}$) BE3 (APOBEC1$^{W90A/R126E}$) | rA1 variants-SpCas9n-UGI | ABEmaxAW | TadA$^{E59A\ or\ E59Q}$-TadA*$^{V106W}$-codon optimized | NGG | Reduce RNA off-target |
| hA3A$^{R128A\ or\ Y130F}$ -BE3 | hA3A variants-SpCas9n-UGI | ABE7.10$^{F148A}$ | SpCas9nTadA variant-TadA* varient-SpCas9n | NGG | Reduce RNA off-target |

CBEs and ABEs use similar improvement strategies and are listed in the same row. rA1, rat APOBEC1; hA3A, human APOBEC3A.

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002

In addition to the cytosine base editors (CBEs), the ABEs have recently been discovered, which can induce AT to GC base conversion. Liu's research team developed a series of ABEs based on *E. coli* TadA (ecTadA) for converting AT to GC in gDNA (Gaudelli et al., 2017). ABEs 7.1 to 7.10 were produced, and they significantly improved editing efficiency. Several versions of ABEs have been developed since the milestone report from Liu group (Gaudelli et al., 2017). By replacing nCas9 with nxCas9 in ABE7.10, xABE was created, which extends the targeting of ABEs to genomic regions containing NG, GAA, or GAT PAM sequences. Other Cas9 variants (e.g., SaCas9, KKH SaCas9, VQR, and VRER) with altered PAM specificity were also used to develop new ABEs that recognize the PAM sequences NGA, NNGRRT, NNNRRT, NGA, and NGCG (Hu et al., 2018; Hua et al., 2018, Hua et al., 2019; Yang et al., 2018). Similar to BE4max, a high efficiency version called ABEmax was engineered to facilitate the editing process by optimizing nuclear localization signal (NLS) and codon usage of nCas9 (Koblan et al., 2018). ABEs can also lead to unwanted SNPs in RNA transcripts induced by deaminase TadA. Thus, an improved version called ABEmaxAW was developed to reduce RNA off-target effects of ABEmax by introducing E59A or E59Q in TadA and V106W in TadA* (Rees et al., 2019). However, this variant still decreased the on-target efficiency of ABEmax, which reinforces the idea that the mutation strategy needs to balance the conflict of on-target efficiency and off-target activity. Recently, a new variant ABE7.10 ($TadA^{F148A}$+$TadA^{*F148A}$) was reported, which can decrease SNV numbers in RNA to almost baseline levels and maintain relatively high on-target efficiency (Zhou et al., 2019).

## Off-target analysis of the base editors

### *High-throughput sequencing (HTS)*

Similar to the application in CRISPR-Cas9 off-target detection, WGS provides global observation of genome sequences, which can cover SNPs induced by base editors. However, the limited sensitivity is still a barrier for WGS in the application of low-frequency mutation detection. Specifically, for base editors, the RNA binding activity of deaminases leads to the concern about off-target events at the RNA level (Conticello, 2008; Sowden et al., 1996). Recent studies based on RNA-seq revealed that DNA base editors can induced numerous RNA mutations (Grunewald et al., 2019; Rees et al., 2019). Interestingly, RNA off-target events were found to be essentially random and sgRNA-independent, which can be ascribed to the overexpression of deaminases (Zhou et al., 2019). RNA-seq can be used to evaluate the unwanted RNA editing of new base editors, especially when applying BEs to adeno-associated virus systems for therapeutic purposes.

### *Genome-wide off-target analysis by two-cell embryo injection (GOTI)*

GOTI is a newly reported method used to detect genome wide gene editing of off-target mutations (Zuo et al., 2019). Compared with other off-target detection methods, GOTI can detect single-nucleotide variants (SNVs) *in vivo*. This method can be applied to evaluate the off-target mutations induced by CRISPR-Cas9, CBEs, and ABEs. Intriguingly, Yang's group found that numerous *de novo* SNVs were induced by BE3, which was not reported in previous studies. Unlike BE3, ABE7.10 induced no increase in SNVs, probably owing to the lack of DNA-binding ability of TadA. This result is consistent with a similar study in rice—a species without the interference of SNVs present in different individuals (Jin et al., 2019).

## Concluding remarks

Efficiency and specificity are two problems that cannot be evaded with gene editing tools. The expectation of applying the CRISPR system to cure diseases in humans raises a higher demand for the safety of editing tools. Thus, precise detection of off-target activities is an essential step in evaluating the practicability of a new CRISPR system. Biased methods provide a convenient but low-sensitivity way to detect off-target effects that rely on bioinformatics algorithms, and unbiased methods detect off-target effects via different editing-related events, such as DSBs, repair factor recruitment, and SNPs. Considering different merits and limitations of these strategies, two or more complementary strategies should be applied simultaneously when performing off-target analysis.

## Declaration of interests

The authors declare no competing interests.

## Acknowledgments

This work was supported by the grant 81771230 (W.C.), 31922048 (E.Z.) and 31522037 (H.Y.) from the National Natural Science Foundation of China. We sincerely appreciate the guidance and help from Dr. Miao Zhao, Dr. Ying-qian Lu and Dr. Xiang Lin.

## References

Cameron, P., Fuller, C.K., Donohoue, P.D., Jones, B.N., Thompson, M.S., Carter, M.M., Gradia, S., Vidal, B., Garner, E., Slorach, E.M., Lau, E., Banh, L.M., Lied, A.M., Edwards, L.S., Settle, A.H., Capurso, D., Llaca, V., Deschamps, S., Cigan, M., Young, J.K., May, A.P., 2017. Mapping the genomic landscape of CRISPR-Cas9 cleavage. Nat. Methods 14, 600–606.

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., Lorenzin, F., Prandi, D., Romanel, A., Demichelis, F., Inga, A., Cereseto, A., 2018. A highly specific SpCas9 variant is identified by *in vivo* screening in yeast. Nat. Biotechnol. 36, 265.

Cencic, R., Miura, H., Malina, A., Robert, F., Ethier, S., Schmeing, T.M., Dostie, J., Pelletier, J., 2014. Protospacer adjacent motif (PAM)-distal sequences engage CRISPR Cas9 DNA target cleavage. PLoS One 9, e109213.

Chen, J.S., Dagdas, Y.S., Kleinstiver, B.P., Welch, M.M., Sousa, A.A., Harrington, L.B., Sternberg, S.H., Joung, J.K., Yildiz, A., Doudna, J.A., 2017. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. Nature 550, 407.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., Kim, J.S., 2014. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. 24, 132–141.

Choi, P.S., Meyerson, M., 2014. Targeted genomic rearrangements using CRISPR/Cas technology. Nat. Commun. 5, 3728.

Cong, L., Ran, F.A., Cox, D., Lin, S.L., Barretto, R., Habib, N., Hsu, P.D., Wu, X.B., Jiang, W.Y., Marraffini, L.A., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823.

Conticello, S.G., 2008. The AID/APOBEC family of nucleic acid mutators. Genome Biol. 9, 229.

Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Neron, B., Rocha, E.P.C., Vergnaud, G., Gautheret, D., Pourcel, C., 2018. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res. 46, W246–W251.

Cromwell, C.R., Sung, K., Park, J., Krysler, A.R., Jovel, J., Kim, S.K., Hubbard, B.P., 2018. Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease specificity. Nat. Commun. 9.

Crosetto, N., Mitra, A., Silva, M.J., Bienko, M., Dojer, N., Wang, Q., Karaca, E., Chiarle, R., Skrzypczak, M., Ginalski, K., Pasero, P., Rowicka, M., Dikic, I., 2013. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. Nat. Methods 10, 361–365.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H.W., Listgarten, J., Root, D.E., 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. 34, 184–191.

Doudna, J.A., Charpentier, E., 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346, 1258096.

Esvelt, K.M., Carlson, J.C., Liu, D.R., 2011. A system for the continuous directed evolution of biomolecules. Nature 472, 499–503.

Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., Church, G.M., 2013. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat. Methods 10, 1116–1121.

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002

Fu, B.X., St Onge, R.P., Fire, A.Z., Smith, J.D., 2016. Distinct patterns of Cas9 mismatch tolerance in vitro and in vivo. Nucleic Acids Res. 44, 5365–5377.
Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M., Calarco, J.A., 2013. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. Nat. Methods 10, 741–743.
Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., Sander, J.D., 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. 31, 822–826.
Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., Joung, J.K., 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. 32, 279–284.
Furey, T.S., 2012. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. Nat. Rev. Genet. 13, 840–852.
Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U. S. A. 109, E2579–E2586.
Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., Liu, D.R., 2017. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551, 464–471.
Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z.R., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442–451.
Grunewald, J., Zhou, R.H., Garcia, S.P., Iyer, S., Lareau, C.A., Aryee, M.J., Joung, J.K., 2019. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. Nature 569, 433.
Harris, R.S., Liddament, M.T., 2004. Retroviral restriction by APOBEC proteins. Nat. Rev. Immunol. 4, 868–877.
Hirano, S., Nishimasu, H., Ishitani, R., Nureki, O., 2016. Structural basis for the altered PAM specificities of engineered CRISPR-cas9. Mol. Cell 61, 886–894.
Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.F., Sontheimer, E.J., Thomson, J.A., 2013. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. Proc. Natl. Acad. Sci. U. S. A. 110, 15644–15649.
Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffini, L.A., Bao, G., Zhang, F., 2013a. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 31, 827–832.
Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y.Q., Fine, E.J., Wu, X.B., Shalem, O., Cradick, T.J., Marraffini, L.A., Bao, G., Zhang, F., 2013b. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 31, 827.
Hu, J., Meyers, R.M., Dong, J., Panchakshari, R.A., Alt, F.W., Frock, R.L., 2016. Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. Nat. Protoc. 11, 853–871.
Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., Liu, D.R., 2018. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 556, 57–63.
Hua, K., Tao, X., Yuan, F., Wang, D., Zhu, J.K., 2018. Precise A.T to G.C base editing in the rice genome. Mol. Plant 11, 627–630.
Hua, K., Tao, X., Zhu, J.K., 2019. Expanding the base editing scope in rice by using Cas9 variants. Plant Biotechnol. J. 17, 499–504.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.
Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J.L., Zhang, F., Gao, C., 2019. Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. Science 364, 292–295.
Joung, J.K., Sander, J.D., 2013. TALENs: a widely applicable technology for targeted genome editing. Nat. Rev. Mol. Cell Biol. 14, 49–55.
Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H.R., Hwang, J., Kim, J.I., Kim, J.S., 2015. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat. Methods 12, 237–243, 1 p following 243.
Kim, Y.B., Komor, A.C., Levy, J.M., Packer, M.S., Zhao, K.T., Liu, D.R., 2017. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat. Biotechnol. 35, 371–376.
Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., Joung, J.K., 2016. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529, 490–495.
Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P., Li, Z., Peterson, R.T., Yeh, J.R., Aryee, M.J., Joung, J.K., 2015. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature 523, 481–485.
Kleinstiver, B.P., Sousa, A.A., Walton, R.T., Tak, Y.E., Hsu, J.Y., Clement, K., Welch, M.M., Horng, J.E., Malagon-Lopez, J., Scarfo, I., Maus, M.V., Pinello, L., Aryee, M.J., Joung, J.K., 2019. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. Nat. Biotechnol. 37, 276–282.
Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., Liu, D.R., 2018. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. 36, 843–846.
Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., Liu, D.R., 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420–424.
Komor, A.C., Zhao, K.T., Packer, M.S., Gaudelli, N.M., Waterbury, A.L., Koblan, L.W., Kim, Y.B., Badran, A.H., Liu, D.R., 2017. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. Sci. Adv. 3, eaao4774.
Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., Nureki, O., Zhang, F., 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517, 583–588.
Koo, T., Lee, J., Kim, J.S., 2015. Measuring and reducing off-target activities of programmable nucleases including CRISPR-cas9. Mol. Cells 38, 475–481.
Kulcsar, P.I., Talas, A., Huszar, K., Ligeti, Z., Toth, E., Weinhardt, N., Fodor, E., Welker, E., 2017. Crossing enhanced and high fidelity SpCas9 nucleases to optimize specificity and cleavage. Genome Biol. 18, 190.
Kuscu, C., Arslan, S., Singh, R., Thorpe, J., Adli, M., 2014. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat. Biotechnol. 32, 677–683.
Lazzarotto, C.R., Nguyen, N.T., Tang, X., Malagon-Lopez, J., Guo, J.A., Aryee, M.J., Joung, J.K., Tsai, S.Q., 2018. Defining CRISPR-Cas9 genome-wide nuclease activities with CIRCLE-seq. Nat. Protoc. 13, 2615–2642.
Lee, J.K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y.H., Lee, K., Jung, I., Kim, D., Kim, S., Kim, J.S., 2018. Directed evolution of CRISPR-Cas9 to increase its specificity. Nat. Commun. 9.
Li, X., Wang, Y., Liu, Y., Yang, B., Wang, X., Wei, J., Lu, Z., Zhang, Y., Wu, J., Huang, X., Yang, L., Chen, J., 2018. Base editing with a Cpf1-cytidine deaminase fusion. Nat. Biotechnol. 36, 324–327.
Lone, B.A., Karna, S.K.L., Ahmad, F., Shahi, N., Pokharel, Y.R., 2018. CRISPR/Cas9 system: a bacterial tailor for genomic engineering. Genet. Res. Int. 2018, 3797214.
Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., Joung, J.K., 2013. CRISPR RNA-guided activation of endogenous human genes. Nat. Methods 10, 977–979.
Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L.H., Church, G.M., 2013a. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. 31, 833.
Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., Church, G.M., 2013b. RNA-guided human genome engineering via Cas9. Science 339, 823–826.
Mao, Z., Bozzella, M., Seluanov, A., Gorbunova, V., 2008. Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair (Amst) 7, 1765–1771.
Mashal, R.D., Koontz, J., Sklar, J., 1995. Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. Nat Genet 9, 177–183.
Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H., Oura, S., Holmes, B., Tanaka, M., Seki, M., Hirano, H., Aburatani, H., Ishitani, R., Ikawa, M., Yachie, N., Zhang, F., Nureki, O., 2018. Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science 361, 1259–1262.
Park, P.J., 2009. ChIP-seq: advantages and challenges of a maturing technology. Nat. Rev. Genet. 10, 669–680.
Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173–1183.
Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., Koonin, E.V., Sharp, P.A., Zhang, F., 2015. *In vivo* genome editing using *Staphylococcus aureus* Cas9. Nature 520, 186–191.
Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., Zhang, F., 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154, 1380–1389.
Rees, H.A., Komor, A.C., Yeh, W.H., Caetano-Lopes, J., Warman, M., Edge, A.S.B., Liu, D.R., 2017. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. Nat. Commun. 8, 15790.
Rees, H.A., Wilson, C., Doman, J.L., Liu, D.R., 2019. Analysis and minimization of cellular RNA editing by DNA adenine base. Sci. Adv. 5, eaax5717.
Ryan, D.E., Taussig, D., Steinfeld, I., Phadnis, S.M., Lunstad, B.D., Singh, M., Vuong, X., Okochi, K.D., McCaffrey, R., Olesiak, M., Roy, S., Yung, C.W., Curry, B., Sampson, J.R., Bruhn, L., Dellinger, D.J., 2018. Improving CRISPR-Cas specificity with chemical modifications in single-guide RNAs. Nucleic Acids Res. 46, 792–803.
Sander, J.D., Joung, J.K., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32, 347–355.
Schaefer, K.A., Wu, W.H., Colgan, D.F., Tsang, S.H., Bassuk, A.G., Mahajan, V.B., 2017. Unexpected mutations after CRISPR-Cas9 editing *in vivo*. Nat. Methods 14, 547–548.
Shen, B., Zhang, W.S., Zhang, J., Zhou, J.K., Wang, J.Y., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X.X., Skarnes, W.C., 2014. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat. Methods 11, 399.
Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., Zhang, F., 2016. Rationally engineered Cas9 nucleases with improved specificity. Science 351, 84–88.
Song, F., Stieger, K., 2017. Optimizing the DNA donor template for homology-directed repair of double-strand breaks. Mol. Ther. Nucleic Acids 7, 53–60.
Sowden, M., Hamm, J.K., Smith, H.C., 1996. Overexpression of APOBEC-1 results in mooring sequence-dependent promiscuous RNA editing. J. Biol. Chem. 271, 3011–3017.

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002

ARTICLE IN PRESS



Teytelman, L., Thurtle, D.M., Rine, J., van Oudenaarden, A., 2013. Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. Proc. Natl. Acad. Sci. U. S. A. 110, 18602–18607.

Tsai, S.Q., Nguyen, N.T., Malagon-Lopez, J., Topkar, V.V., Aryee, M.J., Joung, J.K., 2017. CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets. Nat. Methods 14, 607–614.

Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., Aryee, M.J., Joung, J.K., 2015. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat. Biotechnol. 33, 187–197.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., Gregory, P.D., 2010. Genome editing with engineered zinc finger nucleases. Nat. Rev. Genet. 11, 636–646.

Veres, A., Gosis, B.S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., Erdin, S., Cowan, C.A., Talkowski, M.E., Musunuru, K., 2014. 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, 27–30.

Wang, X., Li, J., Wang, Y., Yang, B., Wei, J., Wu, J., Wang, R., Huang, X., Chen, J., Yang, L., 2018. Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. Nat. Biotechnol. 36, 946–949.

Wienert, B., Wyman, S.K., Richardson, C.D., Yeh, C.D., Akcakaya, P., Porritt, M.J., Morlock, M., Vu, J.T., Kazane, K.R., Watry, H.L., Judge, L.M., Conklin, B.R., Maresca, M., Corn, J.E., 2019. Unbiased detection of CRISPR off-targets *in vivo* using DISCOVER-Seq. Science 364, 286–289.

Wu, X., Kriz, A.J., Sharp, P.A., 2014a. Target specificity of the CRISPR-Cas9 system. Quant Biol 2, 59–70.

Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S., Jaenisch, R., Zhang, F., Sharp, P.A., 2014b. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat. Biotechnol. 32, 670–676.

Xu, T., Li, Y., Van Nostrand, J.D., He, Z., Zhou, J., 2014. Cas9-based tools for targeted genome editing and transcriptional control. Appl. Environ. Microbiol. 80, 1544–1552.

Yang, B., Li, X., Lei, L., Chen, J., 2017. APOBEC: from mutator to editor. J. Genet. Genomics 44, 423–437.

Yang, L., Zhang, X., Wang, L., Yin, S., Zhu, B., Xie, L., Duan, Q., Hu, H., Zheng, R., Wei, Y., Peng, L., Han, H., Zhang, J., Qiu, W., Geng, H., Siwko, S., Zhang, X., Liu, M., Li, D., 2018. Increasing targeting scope of adenosine base editors in mouse and rat embryos through fusion of TadA deaminase with Cas9 variants. Protein Cell 9, 814–819.

Yin, H., Song, C.Q., Suresh, S., Kwan, S.Y., Wu, Q., Walsh, S., Ding, J., Bogorad, R.L., Zhu, L.J., Wolfe, S.A., Koteliansky, V., Xue, W., Langer, R., Anderson, D.G., 2018. Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. Nat. Chem. Biol. 14, 311–316.

Zafra, M.P., Schatoff, E.M., Katti, A., Foronda, M., Breinig, M., Schweitzer, A.Y., Simon, A., Han, T., Goswami, S., Montgomery, E., Thibado, J., Kastenhuber, E.R., Sanchez-Rivera, F.J., Shi, J., Vakoc, C.R., Lowe, S.W., Tschaharganeh, D.F., Dow, L.E., 2018. Optimized base editors enable efficient editing in cells, organoids and mice. Nat. Biotechnol. 36, 888–893.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V., Zhang, F., 2015. Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-cas system. Cell 163, 759–771.

Zhou, H., Liu, J., Zhou, C., Gao, N., Rao, Z., Li, H., Hu, X., Li, C., Yao, X., Shen, X., Sun, Y., Wei, Y., Liu, F., Ying, W., Zhang, J., Tang, C., Zhang, X., Xu, H., Shi, L., Cheng, L., Huang, P., Yang, H., 2018. *In vivo* simultaneous transcriptional activation of multiple genes in the brain using CRISPR-dCas9-activator transgenic mice. Nat. Neurosci. 21, 440–446.

Zhou, C.Y., Sun, Y.D., Yan, R., Liu, Y.J., Zuo, E.W., Gu, C., Han, L.X., Wei, Y., Hu, X.D., Zeng, R., Li, Y.X., Zhou, H.B., Guo, F., Yang, H., 2019. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. Nature 571, 275–278.

Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L., Steinmetz, L.M., Li, Y., Yang, H., 2019. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. Science 364, 289–292.

Please cite this article as: Li, J et al., Advances in detecting and reducing off-target effects generated by CRISPR-mediated genome editing, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2019.11.002