



# The Off-Targets of Clustered Regularly Interspaced Short Palindromic Repeats Gene Editing

Manuel M. Vicente, Miguel Chaves-Ferreira, João M. P. Jorge, João T. Proença and Vasco M. Barreto\*

DNA Breaks Group, NOVA Medical School (NMS), Centro de Estudos de Doenças Crónicas (CEDOC), NOVA University of Lisbon, Lisbon, Portugal

The repurposing of the CRISPR/Cas bacterial defense system against bacteriophages as simple and flexible molecular tools has revolutionized the field of gene editing. These tools are now widely used in basic research and clinical trials involving human somatic cells. However, a global moratorium on all clinical uses of human germline editing has been proposed because the technology still lacks the required efficacy and safety. Here we focus on the approaches developed since 2013 to decrease the frequency of unwanted mutations (the off-targets) during CRISPR-based gene editing.

## OPEN ACCESS

### Edited by:

Olga Amaral,  
Instituto Nacional de Saúde Doutor  
Ricardo Jorge (INSA), Portugal

### Reviewed by:

Fernando Moreira Simabuco,  
State University of Campinas, Brazil  
Turabe M. H. U. Fazil,  
Nanyang Technological University,  
Singapore

### \*Correspondence:

Vasco M. Barreto  
vasco.barreto@nms.unl.pt

### Specialty section:

This article was submitted to  
Cellular Biochemistry,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

**Received:** 31 May 2021

**Accepted:** 09 August 2021

**Published:** 17 September 2021

### Citation:

Vicente MM, Chaves-Ferreira M,  
Jorge JMP, Proença JT and  
Barreto VM (2021) The Off-Targets of  
Clustered Regularly Interspaced Short  
Palindromic Repeats Gene Editing.  
*Front. Cell Dev. Biol.* 9:718466.  
doi: 10.3389/fcell.2021.718466

**Keywords:** CRISPR/Cas9, off-targets, Cas9 variants, DNA repair, gene editing

## INTRODUCTION

Since 2013, the field of gene editing is undergoing a revolution triggered by landmark publications showing that a ribonucleoprotein involved in the protection of bacteria from bacteriophages can be used both *in vitro* (Gasiunas et al., 2012; Jinek et al., 2012) and in human cell lines (Cong et al., 2013; Mali et al., 2013) to introduce double-strand breaks (DSBs) in DNA at precise locations. This remarkable defense mechanism is made of two main components, namely, a locus with a peculiar structure known as “clustered regularly interspaced short palindromic repeats” (CRISPR), discovered by Mojica et al. (2000) in 20 different microbes in 2000, and a set of CRISPR-associated (Cas) proteins. In a process known as adaptation, the Cas proteins capture the genetic information (the “spacers”) from invading bacteriophages and incorporate them into the CRISPR locus. In the CRISPR system that has been most frequently used for gene editing, the transcript of the CRISPR locus (pre-crRNA) is processed into CRISPR-RNAs (crRNAs) that will associate with a partially complementary trans-activator RNA (tracrRNA) and a Cas endonuclease, thus generating a ribonucleoprotein complex. The crRNA will hybridize to a complementary sequence in the target DNA named protospacer, provided this sequence is flanked by a 2–6 base pair protospacer adjacent motif (PAM) recognized by the nuclease, and the Cas endonuclease will then introduce a DSB at a precise location in the DNA. The targeting of specific sequences via spacers is termed interference.

In eubacteria and notably in archaea, the CRISPR systems are widespread. Three types of systems have been identified, each defined by a signature *Cas* gene: *Cas3* in type I systems, *Cas9* in type II, and *Cas10* in type III. There are excellent articles on the many ways bacteria have diversified their unique adaptive immune system (e.g., Makarova and Koonin, 2015); here, we discuss mainly the components of type II, which, due to its simplicity, has been the basis for most gene editing approaches. For editing purposes, the system was further simplified by the fusion of crRNA with the tracrRNA, generating a molecule known as single guide RNA (sgRNA) (Jinek et al., 2012). Thus, a sgRNA and the Cas9 in the cell are sufficient to introduce a DSB.

Multiple articles have alluded to the path from the early recognition of CRISPR loci as a set of adaptive immunity instructions, by Francisco Mojica, to the 2020 Nobel Prize in Chemistry given to Jennifer Doudna and Michele Charpentier for their repurposing of CRISPR components as a gene-editing tool (e.g., Lander, 2016). The explosion of publications using this flexible, simple, and relatively inexpensive tool and recent clinical applications based on CRISPR gene editing demonstrate the impact of this technological breakthrough, but problems remain. Depending on the context, the technique may lack the necessary efficiency and sequence specificity to mediate flawless editing. These concerns are particularly serious in the case of clinical applications using edited somatic cells (e.g., Lu et al., 2020; Frangoul et al., 2021) and notably in the case of heritable genome editing in humans, i.e., changes introduced in sperm, eggs, or embryos to make genetically modified children, which has already happened at least once despite the legislation in many countries barring all clinical uses of germline editing; early concerns due to the simplicity and potential of the technique followed by a call for a moratorium on germline genome editing signed by prominent figures in the field illustrate the ongoing fear (Baltimore et al., 2015; Lander et al., 2019).

Here we will focus exclusively on one of the main problems of the technique: the unwanted changes in the genetic information that the CRISPR editing produces outside of the target region(s) (Figure 1). These undesired alterations are known in the field as “off-targets.” We will review why off-targets are observed when Cas9-mediated gene editing is used, the several strategies developed to decrease them (Figure 2), and the off-targets of other competing editing technologies, providing quantitative comparisons when the data are available.

## CLUSTERED REGULARLY INTERSPACED SHORT PALINDROMIC REPEATS/Cas9 ON- AND OFF-TARGET ACTIVITY

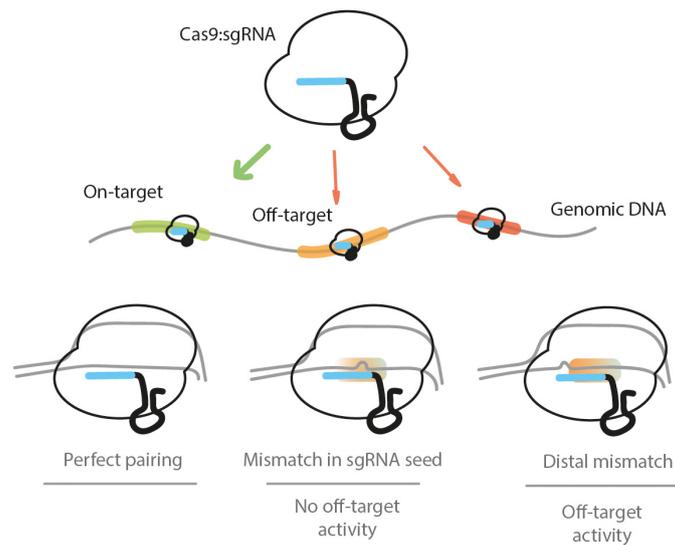
Most CRISPR-based editing involves DNA repair. Targeted DSBs were found to increase the integration frequency of reporters with homology arms by orders of magnitude (Rouet et al., 1994), which has been explored by several gene-editing techniques based on programmable nucleases, of which CRISPR editing is the latest example. In the case of knockins, the DSB will recruit the high-fidelity homology-directed repair (HDR) machinery to integrate a piece of genetic information introduced in the cell as a template (Heyer et al., 2010). In the case of a knockout, no template is introduced, and most strategies are based on the action of the error-prone non-homologous end-joining (NHEJ), which can repair the DSB in a flawed way, introducing deletions or insertions (indels) (Pardo et al., 2009) that lead to frameshifts and, consequently, transcripts that are targeted by the non-sense-mediated mRNA decay (NMD) surveillance pathway.

Unfortunately, the crRNA pairing with the protospacer only involves 20 base pairs and lacks the necessary complexity to always ensure a single hit in a genome with more than 3 billion

base pairs, such as the human genome. Adding to the problem, the pairing of the 20-bp region does not require 100% identity and tolerates mismatches, which vastly increases the number of potential off-targets (Figure 1). Several studies have focused on the promiscuous nature of the binding of the Cas9:sgRNA complex to DNA, providing mechanistic insights on the off-target Cas9 activity. Early work on human cells revealed that a DSB could occur even when the PAM-distal region of the protospacer has up to five mismatches with the sgRNA (Fu et al., 2013; Hsu et al., 2013). Since these initial studies, it has been further found that the specificity of Cas9 is heavily determined by the first 7–12 bp-long seed sequence proximal to the PAM (Jiang and Doudna, 2017). Crucially, the DNA interrogation mechanism by Cas9 was later elucidated, revealing the PAM's absolute requirement for double-stranded DNA unwinding, and the posterior evaluation of sequence similarity between the target sequence and sgRNA. These molecular insights made it clear that off-target activity should only be encountered in PAM-adjacent sequences (Sternberg et al., 2014). In 2014, genome-wide chromatin-immunoprecipitation sequencing (ChIP-seq) was used to map the binding sites of catalytically inactive Cas9 (dCas9). The results validated the PAM requirements for Cas9 binding and showed that targeting was enriched in open chromatin loci, indicating the chromatin structure as one factor mediating Cas9 sequence recognition (Kuscu et al., 2014). The latter finding was independently validated afterward (Yarrington et al., 2018; Janssen et al., 2019).

## OFF-TARGET PREDICTION TOOLS

It is not possible to completely avoid off-targets, but selection of the on-target sgRNAs with the highest efficiency and the smallest number of off-targets or without off-targets in critical genomic regions, such as coding sequences, can reduce off-target frequency/effects. As of the writing of this review, 49 online tools can predict off-target sequences and score the sgRNA efficiency for on-target activity (Torres-Perez et al., 2019). These tools can be alignment- or scoring-based. Alignment-based methods rely on the sequence similarity search of the sgRNA sequence in a given genome and, according to the mechanistic studies on Cas9-mediated DSB sequence recognition requirements, off-target sites are determined. Bowtie, a pre-CRISPR-era NGS-based alignment tool (Langmead et al., 2009), allows very few mismatches in the sgRNA seed sequence, making off-target prediction incomplete (Doench et al., 2016). However, using the Bowtie algorithm, with modifications, the CCTop tool finds off-targets with up to five mismatches in the protospacer sgRNA sequence, and scores each sgRNA according to the likelihood of a stable sgRNA/DNA heteroduplex (Stemmer et al., 2015). CasOFFinder includes no limits in the mismatch base number (Bae et al., 2014). Both tools allow the choice of PAM sequences for different programmable nuclease platforms, but compared to experimentally generated off-targets, the CCTop and CasOFFinder tools only identify a subset of sites, suggesting that sequence similarity alone is not sufficient for complete off-target prediction (Cameron et al., 2017). Scoring-based methods



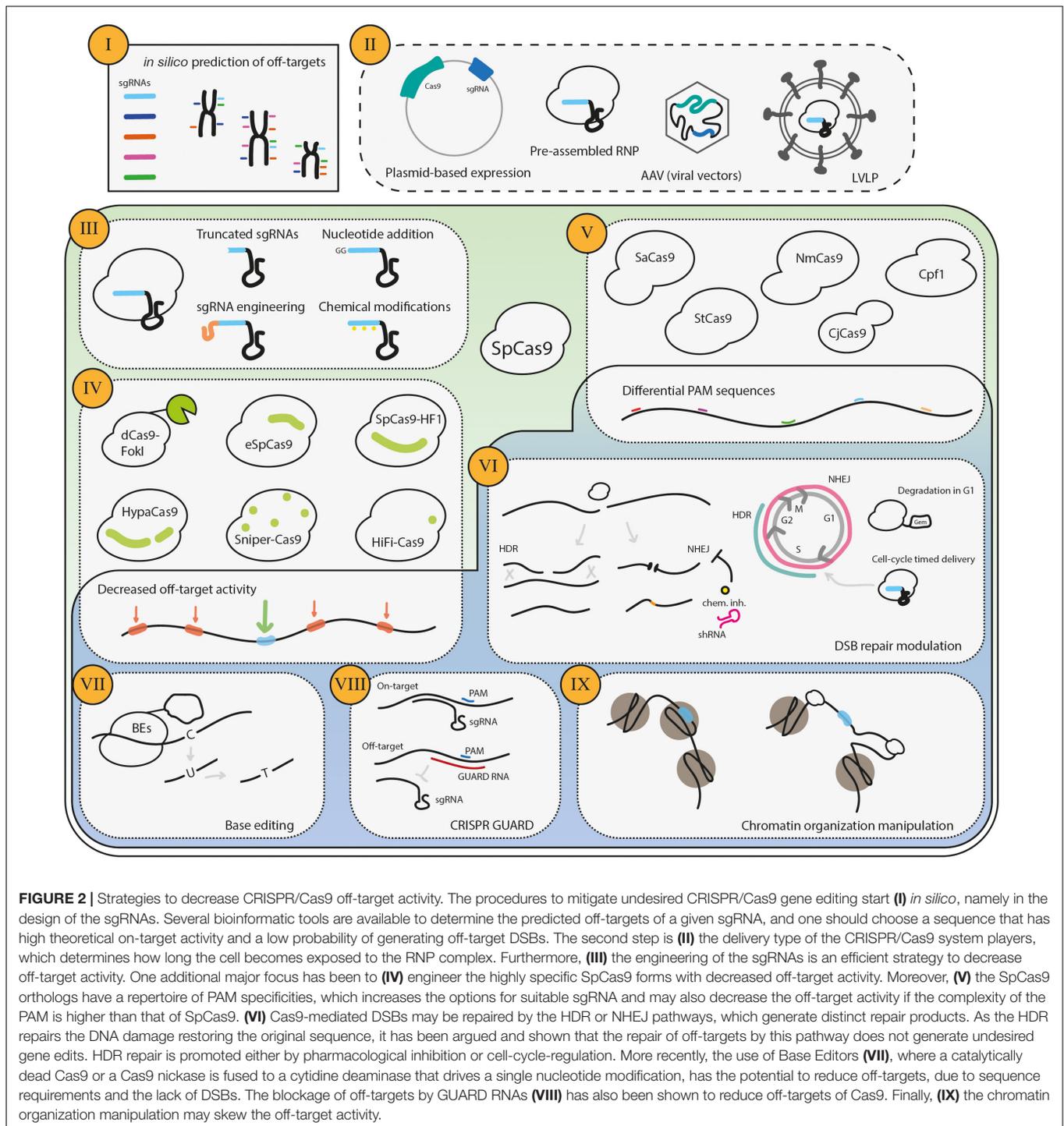
**FIGURE 1** | The Cas9:sgRNA ribonucleoprotein complex targets the genomic DNA when present in the cell nucleus. It will target PAM-adjacent sequences and may encounter several potential recognition sites. If there is a perfect base pairing between the target DNA and the sgRNA, a DSB is induced by Cas9. If there is a mismatch between the sgRNA and the DNA sequences present within the first 7–12 bp (proximal to the PAM) segment, the base pairing will not be sufficient to induce a DSB. If one or more mismatches between the DNA and sgRNA sequences are found in the PAM-distal (5' end of the sgRNA) nucleotides, a DSB can be induced by Cas9, leading to off-target activity. The scenario of sgRNA sequence-independent off-targets is not depicted.

include sgRNA whole-genome alignment and an additional score for each possible sgRNA. According to empirical data, scores are usually attributed to genome context (open vs. closed chromatin, for instance) and specific mismatch type or location within the protospacer. The MIT score, proposed in the early days of the CRISPR era, was derived from the activity evaluation of > 700 sgRNAs in > 100 predicted off-target loci. The data were used to determine a weight matrix for each mismatch position (Hsu et al., 2013). This score was integrated into several tools, such as CHOPCHOP (Labun et al., 2016) and CRISPOR (Haeussler et al., 2016). Another widely used scoring method is the Cutting Frequency Determination (CFD), which was generated from the analysis of the on and off-target activity of a > 27 k sgRNA library. The library included a primary set of all sgRNAs against *Cd33*, regardless of PAM, combined with variant sgRNAs having 1-nucleotide insertions, deletions, and mismatches (Doench et al., 2016). The CFD score is included in the CRISPRScan (Moreno-Mateos et al., 2015), CRISPOR (Haeussler et al., 2016), and GuideScan (Perez et al., 2017) tools. Finally, using empirical datasets, machine-learning-based tools have also been developed. DeepCRISPR provides off-target prediction considering epigenetic sequence factors and was reported to outperform the other available off-target prediction tools (Chuai et al., 2018). Scanning a given single exon from each of 36 different genes, it can be seen that there are major differences in the sgRNA choices of six widely used online tools (Figure 3), which means that these programs also select different constellations of off-targets, particularly in knockout projects, which impose fewer constraints in the selection of the sgRNAs than knockin projects. Systematic comparisons between the tools and methods for sgRNA determination are still needed.

## DETECTION OF OFF-TARGETS: BIASED AND UNBIASED METHODS

Bacteria have evolved for millions of years with their CRISPR systems. In contrast, human cells were exposed to Cas9 during evolution, and their genomes are orders of magnitude larger than those of bacteria. Thus, CRISPR off-target events are unavoidable in large Eukaryotic genomes, and their detection poses several technical challenges. Two types of detection methods have been used. In the biased methods, only sequences predicted by the *in silico* tools mentioned above or resulting from empirical studies are analyzed. This produces a relatively small set of potential off-target sequences that several methods can then analyze, including the detection of mispairings in the amplified region by the T7 endonuclease I, Sanger sequencing, Next Generation Sequencing (NGS), or off-target reporters (Table 1).

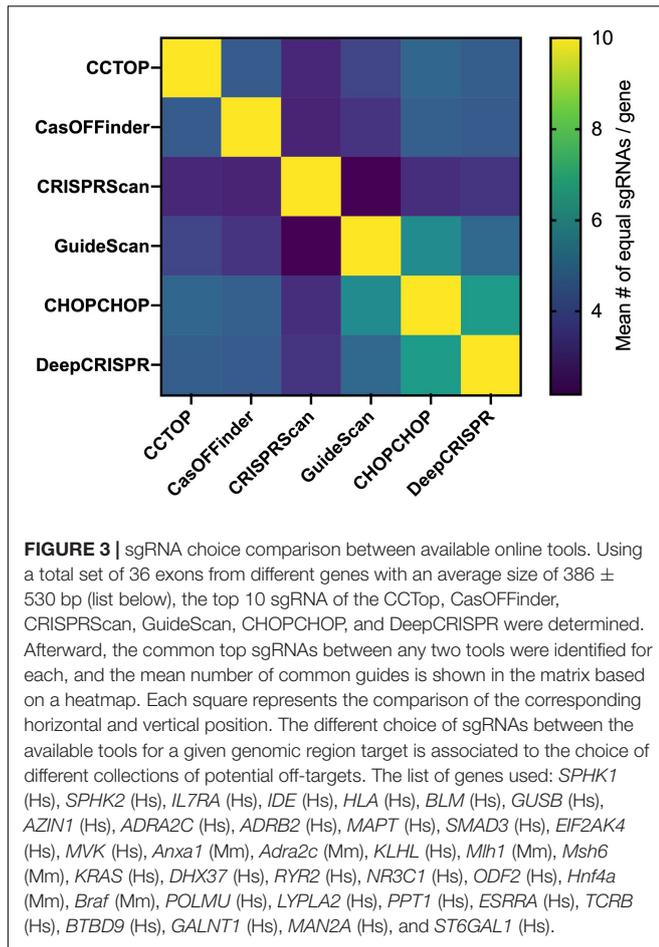
Given the current limitations and disparities in the results produced by the *in silico* approaches to detect the efficiency of the sgRNAs for their on-targets, the prediction of the off-targets is even more problematic, due to the increased number of variables and the lack of data for most species on the epigenetic features that are known to influence the specificity and efficiency of CRISPR editing, explored by programs such as DeepCRISPR (Chuai et al., 2018). Thus, the strategies of the second type, known as the unbiased methods, have gained crucial importance. The unbiased methods teach us about the best criteria to identify off-targets and provide a comprehensive collection of the real off-targets. There is now a plethora of



techniques that we divide here into *in vitro*, ChIP-seq, and DSB-capture methods (Table 1).

One of the *in vitro* methods is Digenome-seq, in which the targeted cells are subjected to whole-genome sequencing. The genomic DNA is digested *in vitro* with the ribonucleoprotein complexes (RNPs) to maximize the detection of off-targets, which are then detected by NGS as fragments with identical 5'

ends, allowing the detection of off-targets with a 0.1% or lower frequency of indels (Kim et al., 2015). The authors concluded that Cas9 could be highly specific, inducing indels at only a few off-targets. Another *in vitro* method providing the same type of data is SITE-Seq, which is based on the cleavage of genomic DNA by an RNP and then the tagging, enrichment, and sequencing of the fragments (Cameron et al., 2017). One



of the limitations of Digenome-seq and SITE-Seq is the high background of random DSBs. Thus, Circularization for *In vitro* Reporting of Cleavage Effects by sequencing (CIRCLE-seq) was eventually developed to solve this problem. Essentially, in this *in vitro* method the DNA is first circularized and then linearized by the RNP being tested. Since only these linearized products will be sequenced, the background of RNP-independent genomic breaks previously present is vastly reduced (Tsai et al., 2017).

ChIP-seq using a catalytically inactive Cas9 revealed that the potential off-targets are enriched in the open chromatin regions, and the number of off-targets varied widely from around 10 to over 1,000 depending on the sgRNA used (Kuscu et al., 2014). DISCOVER-Seq is another ChIP-seq approach, but based on the tracking of the MRE11, a repair factor that is recruited to the Cas9 cut site and allows the identification of the nuclease cutting site with a single-base resolution (Wienert et al., 2019).

Unlike the ChIP-seq methods, the key step in GUIDE-seq is the ligation to the DSBs of double-stranded oligodeoxynucleotides in live cells expressing the RNP followed by NGS (Tsai et al., 2015). It was reported that the majority of off-targets detected had not been detected by the computational methods available at the time. Another method based on the capture of the RNP-dependent DSBs is High-Throughput Gene

Translocation Sequencing (HTGTS), which was originally developed to detect DSBs introduced by natural processes in lymphocytes (Hu et al., 2016). Unlike the previous method, in HTGTS the DSBs to detect are captured *in vivo* by a single DSB generated in the genome by a meganuclease. One additional DSB-capture method is based on the observation that the linear double-stranded genome of an integrase-defective lentiviral vector (IDLV) incorporates preferentially into DSBs by NHEJ. A lentivirus encoding a puromycin-resistant gene was used to tag *in vivo* the DSBs introduced by distinct RNPs and, after selection of the cells with puromycin, the integration sites of the lentivirus were mapped (Wang et al., 2015). Another technique, Breaks Labeling *In Situ* and Sequencing (BLISS), directly tags the DSBs in fixed cells or tissue sections (Yan et al., 2017). It should be noticed that these capture methods do not distinguish between the RNP-dependent breaks and other DSBs, which explains why there was only a 2–3 fold increase in the number of puromycin-resistant colonies relative to the control in the IDLV study (Wang et al., 2015). It is only by mapping the breaks to the genome more likely to be associated with the RNP activity that the off-targets are identified, which makes these capture methods not absolutely unbiased.

Verification of *in vivo* off-targets (VIVO) is a final method worth mentioning because it combines the specificity of the *in vitro* CIRCLE-seq to identify potential off-target sites with the NGS analysis of these regions in cells exposed to RNPs (Akcakaya et al., 2018). Interestingly, the study found that depending on the choice of the sgRNA, the off-targets could be substantial in number or undetectable.

## STRATEGIES TO DECREASE OFF-TARGET ACTIVITY BASED ON sgRNA ENGINEERING

Computational prediction of off-target activity is the first step in minimizing these undesired events. Additional experimental approaches have been proposed to ablate Cas9 off-target activity, based on several aspects of the CRISPR/Cas9 genome editing biology. The first evidence on the sgRNA-target DNA base pairing requirements, mentioned above, inspired initial work on off-target minimization. It was reasoned that truncated sgRNAs would reduce the RNA-DNA interface, increasing the Cas9/sgRNA complex binding sensitivity to mismatches, thus reducing off-target events. The use of truncated sgRNAs reduced the DSB generation in off-target sites by 5,000-fold compared to full-length ones (Fu et al., 2014). Reducing the length of the sgRNA also increases the potential number of off-targets, but 97 potential off-targets were analyzed, including all with a single mismatch, and only one had a detectable number of off-targets. Furthermore, the addition of two guanine nucleotides to the 5' end of the complementary sequence of the sgRNA improved target specificity, reducing off-target events (Cho et al., 2014). More recently, the addition of a hairpin to the 5' end of the sgRNA was shown to elevate the energetic requirements for sgRNA/target DNA base pairing, improving the specificity of the Cas9/sgRNA complex, thus decreasing the off-target activity

**TABLE 1** | Summary table of the methods currently available to detect off-target activity in gene editing applications.

Methods			References
Biased	Advantages	Disadvantages	
T7 endonuclease I	Fast and easy to use	Very limited to predict off-targets	
Sanger sequencing	Relatively fast and easy to use	Limited to amplified regions and low resolution when used in bulk cultures	
Next generation sequencing	High coverage and resolution	Not targeted to off-target discovery (unless it's a whole-genome approach)	
Off-target reporters	Precise measure of off-target activity	Restricted to the reporter sequence	
<b>Unbiased</b>			
<b><i>In vitro</i></b>			
Digenome-seq	Detection of off-targets with a 0.1% or lower frequency of indels	High background of random DSBs, requires reference genome, expensive	Kim et al., 2015
SITE-Seq	Allows the recovery of both high- and low-cleavage off-target sites	High background of random DSBs	Cameron et al., 2017
CIRCLE-seq	Background of RNP-independent genomic breaks is vastly reduced	Limited to cells that can be transfected with the GUIDE-seq dsODN tag	Tsai et al., 2017
<b>Chip-seq</b>			
Chip-seq (with dCas9)	Excellent coverage	Detects only dCas9-associated off-targets	Kuscu et al., 2014
DISCOVER-Seq	Allows the identification of the nuclease cutting site with a single-base resolution	Detects only off-targets associated with adenoviral-delivered Cas9	Wienert et al., 2019
<b>DSB capture</b>			
GUIDE-Seq	Can identify off-targets that are not detected by existing computational methods or CHIP-seq	Efficiency is limited to chromosomal relaxation	Tsai et al., 2015
HTGTS	Able to detect a wide range of broken end structures with nucleotide-level resolution. Sensitive and reproducible, relatively inexpensive and scalable	Only detects DSBs that translocate	Hu et al., 2016
IDLV	Efficient detection of off-targets in primary cells	Lower sensitivity than other methods	Wang et al., 2015
BLISS	Allows the direct labeling of DSBs in fixed cells or tissue sections on a solid surface, highly sensitive, easy scalability, and multiplexing	Detection of off-targets present only in the labeling period	Yan et al., 2017
VIVO	Detects <i>in vivo</i> off-target effects of gene-editing nucleases. Can be used with Cas9 variants and other Cas orthologs	–	Akcakaya et al., 2018

(Kocak et al., 2019). The chemical modification of the ribose-phosphate backbone of a specific site within the sgRNA was also described to reduce off-targets (Ryan et al., 2018).

## ENGINEERED Cas9 VARIANTS TO DECREASE OFF-TARGET ACTIVITY

*Streptococcus pyogenes* Cas9 (SpCas9) is the most broadly used Cas9. Initial studies employed mutated SpCas9 proteins, that work as nickases (nCas9), cleaving only one strand of DNA. To increase the complexity of the sequence that defines the target region, two RNPs with proximal protospacers would theoretically minimize off-target activity. In fact, two sufficiently close DNA strand nicks were able to generate DSBs (Ran et al., 2013) and were shown to decrease off-target activity (Ran et al., 2013; Shen et al., 2014). The use of two nCas9s improved specificity, since the off-targets for each sgRNA are unlikely to be close to

each other and isolated off-target nicks are easily repaired. Two back-to-back papers (Guilinger et al., 2014; Tsai et al., 2014) also explored the concept of DSB generation dependent on two sgRNAs using a dCas9 fused to a *FokI* nuclease to induce DSBs. In both studies, there was a substantial decrease in off-target activity when the dCas9-*FokI* fusions were used. The peptide linkers between dCas9 and *FokI* nuclease were then improved, and by deep sequencing no detectable off-target activity was found. Following these studies, chimeras of RNA-guided *FokI* nucleases have been produced that maintain high editing efficiency, solving the main limitation of this system (Havlicek et al., 2017).

One of the key developments for improved Cas9 proteins was its crystal structure resolution when in complex with the sgRNA and target DNA (Nishimasu et al., 2014). With the protein conformation information, it was possible to determine potential sites for the splitting of Cas9, enabling the controlled reassembly and activation of the protein. The use of this strategy showed no detectable off-target modifications by sequencing of predicted

sites but diminished on-target activity (Zetsche et al., 2015b). The crystal structure also enabled the detection of a positively charged groove responsible for stabilizing the non-targeted DNA strand in the Cas9:sgRNA/DNA complex. Upon charge neutralization of the protein domain, Cas9:sgRNA binding to double-stranded DNA was more dependent on Watson-Crick base pairing between the sgRNA and the target DNA sequence, thus reducing the off-target activity to levels undetectable by whole-genome deep sequencing. This structure-guided engineered Cas9 was coined “enhanced specificity” SpCas9 (eSpCas9) (Slaymaker et al., 2016). In the same year, several amino acids of SpCas9 were identified to have non-specific DNA contacts with the target DNA strand, and once mutated, a decreased tolerance for protospacer mismatches was observed. Moreover, using GUIDE-seq (Tsai et al., 2015), off-target activity was not detected for the mutated SpCas9, named high-fidelity SpCas9 (SpCas9-HF1) (Kleinstiver et al., 2016a). Although highly specific, these two engineered SpCas9 proteins showed reduced on-target activity in some target genes (Kim S. et al., 2017). The mechanism for the increased specificity of these highly specific SpCas9 variants was later found to be due to raised thresholds for conformational activation upon DNA binding of these proteins. With this knowledge, a hyper-accurate Cas9 (HypaCas9) variant was developed that shows improved on-target specificity without compromising on-target activity (Chen et al., 2017). The Sniper-Cas9 variant was then developed using a directed evolution approach, revealing on-target activity level similar to that of SpCas9 and better specificity than the other high-fidelity engineered variants (Lee et al., 2018).

Finally, a HiFi Cas9 variant was generated using an unbiased bacterial screen. It contains a single point mutation in the SpCas9 backbone that increases the on-target activity compared to the other rationally engineered variants, maintaining no detectable off-target activity (Vakulskas et al., 2018).

The race for the generation of Cas9 variants with the desired high on-target and negligible off-target activity appears to still be on a fast pace (Table 2). Further detailed comparisons between these and future engineered Cas9 variants are needed across genomes, delivery methods, and cell types.

## ON- AND OFF-TARGET ACTIVITY OF Cas9 ORTHOLOGS

Since the early days of CRISPR research in bacteria, it was clear that this molecular immune system is ubiquitous (Gasiunas et al., 2020). As Cas9 orthologs began to be discovered and characterized, new tools were developed to induce programmable DSBs, with key differences from the SpCas9-based ones. The SpCas9 ortholog derived from *S. aureus* (SaCas9) has the advantage of being encoded by a small gene (SpCas9 4.1 kb, SaCas9 3.2 kb), which allowed the use of an adeno-associated-virus (AAV) vector for more efficient *in vivo* or *ex vivo* applications (Kumar et al., 2018). Moreover, SaCas9 recognizes a considerably larger PAM sequence than SpCas9, NNGRRT, which theoretically improves specificity. In fact, its application in human cell lines revealed absent off-target activity, evaluated

**TABLE 2 |** Summary table on the Cas9 nucleases currently available for CRISPR/Cas9-mediated gene editing.

Cas9 nucleases			
Engineered Cas9 variants	Characteristics	Off-target activity	References
SpCas9	Most broadly used Cas9		Jinek et al., 2012
nCas9	Mutated SpCas9 D10A protein (deactivation of the RuvC domain). Only cleaves one strand of DNA	1,500 × reduction of off-target effects	Ran et al., 2013
dCas9-FokI	Catalytically inactive Cas9 (D10A and H840A) fused with FokI catalytic domain	No detectable off-targets	Guillinger et al., 2014; Tsai et al., 2014
eSpCas9	Structure-guided engineered Cas9 with the point-mutations K848A, K1003A, and R1060A	Reduced off-target activity to non-detectable levels	Slaymaker et al., 2016
SpCas9-HF1	Mutated version of SpCas9 (N497A, R661A, Q695A, and Q926A)	Off-target activity not detected using GUIDE-seq	Kleinstiver et al., 2016a
HypaCas9	Hyper-accurate Cas9 variant (N692A, M694A, Q695A, and H698A mutations in the REC3 domain)	Better genome-wide specificity relative to both SpCas9-HF1 and eSpCas9	Chen et al., 2017
Sniper-Cas9	Developed using a directed evolution approach	Better specificity than the other high-fidelity engineered variants	Lee et al., 2018
<b>Cas9 orthologs</b>			
SaCas9	Ortholog derived from <i>Staphylococcus aureus</i>	Recognizes a larger PAM sequence, which theoretically improves specificity	Kumar et al., 2018
St1Cas9	Ortholog derived from <i>Streptococcus thermophilus</i>	Evaluation by deep sequencing shows lower off-target activity	Müller et al., 2016
St3Cas9	Ortholog derived from <i>Streptococcus thermophilus</i>	Evaluation by deep sequencing shows lower off-target activity	Müller et al., 2016
NmCas9	Ortholog derived from <i>Neisseria meningitidis</i>	Recognizes a larger PAM sequence and has lower off-target activity compared to SpCas9	Lee et al., 2016
CjCas9	Ortholog derived from <i>Campylobacter jejuni</i>	Lower off-target activity compared to SpCas9	Kim E. et al., 2017

by deep sequencing (Wang et al., 2019). Cas9 proteins from *S. thermophilus*, St1Cas9, and St3Cas9, specific for the PAM sequences NNAGAAW and NGGNG, respectively, also show lower off-target activity evaluated by deep sequencing (Müller et al., 2016). The SpCas9 ortholog from *N. meningitidis* (NmCas9) recognizes a larger PAM sequence, NNNNGHTT, and has lower off-target activity than SpCas9 (Lee et al., 2016). The *C. jejuni* Cas9 (CjCas9) is encoded by an even smaller gene than SaCas9 (2.9 kb) has specificity for two PAM sequences, NNNNACAC and NNNNRY, and lower off-target activity compared to SpCas9 (Kim E. et al., 2017). Despite these lower off-target activities, one should keep in mind that the higher PAM complexity of the less used Cas9 orthologs reduces the number of potential on-targets for a specific application.

Although Cas9 has dominated the gene-editing field, Cas12a (initially described as Cpf1), which was cloned from *Francisella novicida*, has also been used as a tool. Cas12a and Cas9 belong to the class 2 CRISPR-Cas system, but the former has distinctive features, namely a smaller size, the natural use of a single RNA in the RNP complex, a T-rich PAM, and the production of a staggered DSB instead of the blunt break produced by Cas9 (Zetsche et al., 2015a). These breaks were shown to be advantageous for HDR-mediated genome editing (Moreno-Mateos et al., 2017). However, compared to Cas9, the slightly more complex PAM and reports of a lower efficiency (Lee et al., 2019) make it a less appealing tool. More importantly for this review, two careful independent analyses have shown that Cas12a introduces fewer off-targets than Cas9 (Kim et al., 2016; Kleinstiver et al., 2016b; Yan et al., 2017). Biochemical analyses suggest that this difference probably occurs because Cas12a is less tolerant to mismatches along most of the DNA target sequence than Cas9 (Strohkendl et al., 2018). However, *in vitro*, Cas12a and Cas9 have been shown to have strong nicking of DNA sequences containing up to four mismatches, non-specific sequence nicking of dsDNA when Cas12a is bound to a target DNA (Fu et al., 2019; Murugan et al., 2020), and RNA-independent, non-sequence specific DNA cleavage activities (Sundaresan et al., 2017). Whether these effects also occur *in vivo* in Eukaryotic cells is unclear, and nicks are typically less problematic than DSBs, but the sequence non-specific editing would be difficult to avoid using bioinformatic tools. Nevertheless, the discovery of different CRISPR systems is still a highly attractive subject of research and has expanded the tool kit for genome editing.

## CHROMOSOMAL CONTEXTS AND OFF-TARGET ACTIVITY

Chromatin accessibility conditions the efficiency of the CRISPR system in gene editing; notably, the sgRNA binding is facilitated in open chromatin regions (Kuscu et al., 2014; Wu et al., 2014), and the state of the chromatin is often cell cycle-dependent. In addition, HDR is active predominantly during the S and G2 phases of the cell cycle. Thus, it is not surprising that the cell cycle influences both the CRISPR efficiency and the

frequency of off-targets. The CRISPR-based adaptive immune system in bacteria has never had to tackle the challenge of DNA accessibility, as bacteria do not organize their genome in nucleosomes and archaea, despite having histones, maintain a relaxed structure allowing CRISPR to recognize phage DNA even if integrated into its DNA. More importantly, the foreign invading viral or phage DNA is highly accessible before encapsulation, allowing an ample window of opportunity for CRISPR-mediated immunity.

The efficiency of Cas9 editing is reduced in heterochromatic regions (Jensen et al., 2017). Both binding and cleavage are inhibited by nucleosomes (Horlbeck et al., 2016), particularly if the PAM region is located within the nucleosome core (Hinze et al., 2015; Isaac et al., 2016). This has been shown to occur both in human and plant cells (Liu et al., 2019), but not in zebrafish (Moreno-Mateos et al., 2015).

Therefore, quiescent cells or cells otherwise in a state of compacted chromatin hamper the efficiency of targeting in a way that may not be predictable by programs, as nucleosomes sequester putative cleavage sites, particularly on mismatch hits (Yarrington et al., 2018). On the other hand, the off-target effects of Cas9 are greatly reduced in chromatin regions (Hinze et al., 2015; Horlbeck et al., 2016), suggesting that the concerns raised by results of cell-free methods used to compile a list of potential off-target sites are greatly exaggerated, as many of these sites are inaccessible *in vivo* (Kim and Kim, 2018).

Some effort is being put into altering the chromosome accessibility to improve Cas9 activity *in vivo*. One strategy, proxy-CRISPR, consists of using a catalytically dead SpCas9 to bind to locations close to the target region, thus altering the local chromatin structure and exposing previously inaccessible targets to the nuclease activity of FnCas9, CjCas9, NcCas9, and FnCpf1; this strategy was shown to improve editing efficiency on these sites greatly and to lower the off-target activity (Chen et al., 2017). However, this method requires the expression of two CRISPR/Cas systems, making its *in vivo* application difficult. Another method explores Cas9 orthologs fused to chromatin-modulating peptides (CMPs), using a direct approach to remodel the chromatin of the target sites. This method greatly improves editing in refractory sites (Ding et al., 2019). However, for this technique to be used in clinical trials, it has to be determined if the presence of CMPs alters gene expression profiles. A more recent approach uses histone deacetylase (HDAC) inhibitors to promote chromatin accessibility for genome editing. Among the HDAC inhibitors tested, vorinostat (suberoylanilide hydroxamic acid) were the most efficient, raising both NHEJ- and HDR-driven editing, with notable efficiency in the increase of HDR in closed loci (Zhang et al., 2021). The observed increase of NHEJ may hinder tailored gene editing applications, and should be further analyzed.

Besides chromatin accessibility, the repair systems used by Eukaryotic cells change during the cell cycle. HDR is mainly active in mammalian cells during the S phase, when DNA replication leads to the generation of endogenous repair templates, and declines in the late-G2 phase. In contrast, the NHEJ pathway occurs in all phases with special emphasis to G1 (Mao et al., 2008). Therefore, knocking in a DNA sequence tends

to be more efficient during the S and G2 phases. Moreover, one can expect that off-target DSBs generated in the presence of a more proficient HDR pathway are repaired with an endogenous sequence, thus decreasing the frequency of undesired mutations.

A way to increase the use of HDR to deal with the Cas9-induced site-specific DSBs is to synchronize the cells to deliver the Cas9 RNPs at a specific cell cycle phase. Cas9 RNP-mediated editing starts 4 h after nucleofection and lasts for 24 h until the RNPs are degraded (Kim et al., 2014). A careful study determined at which point cells should be synchronized to maximize HDR efficiency while reducing off-target effects. The most effective cell synchronization treatment was Nocodazole, which blocks the cells at the M phase. At this phase, cells have their DNA fully replicated, and their nuclear membrane is broken down (Lin et al., 2014). Researchers have also fused Cas9 to the N-terminal region of Geminin protein, a substrate for the E3 ubiquitin ligase complex APC/Cdh1. This strategy is particularly useful because it does not require cell cycle inhibitors, as it uses the cell cycle regulation machinery, marking Cas9 for degradation via ubiquitination during the G1 phase, thus ensuring the proper expression in time for the maximum HDR potential (Gutschner et al., 2016). More recently, the discovery of bacteriophage proteins that evade prokaryotic adaptive immunity, known as anti-CRISPR proteins (reviewed in Maxwell, 2017), allowed for a similar approach by fusing the anti-CRISPR protein AcrIIA4 to the Chromatin licensing and DNA replication factor 1 (CDT1), which ensures that DNA replication occurs only once per cell division. Gene editing is deactivated by the anti-CRISPR during the G1 and S phases, and during the G2 phase, because the AcrIIA-CDT1 fusion is degraded, Cas9 becomes active. This provided not only a decrease of the off-target effects but, at the same time, a fourfold increase in HDR efficiency (Matsumoto et al., 2020). The efficiency was higher than in previous studies as the Cas9 is only inactivated but not degraded during the G1 phase.

A different strategy is to shift the balance of NHEJ vs. HDR by disrupting key NHEJ players, either using molecular inhibitors, short hairpin RNAs (shRNA), or the co-expression of Cas9 with proteins that mediate the proteasomal degradation of ligase IV (Chu et al., 2015). SCR7 is a ligase 4 inhibitor that has been used in numerous studies with inconsistent results, suggesting that its action is cell-type- or gene-dependent, making it a poor general promoter of HDR for CRISPR (Bischoff et al., 2020). A small molecule, RS-1, which stimulates RAD51 and therefore HDR, improved the efficiency of targeted integrations (Song et al., 2016); other small molecules have been shown to promote HDR directly or indirectly by reducing NHEJ (reviewed in Bischoff et al., 2020). The use of small molecules has also been suggested to help attenuate the off-target problem. For example, by inhibiting NHEJ, the frequency of off-targets is greatly reduced (Zheng et al., 2020), whereas by boosting the HDR efficiency with small molecules targeting can be successfully achieved in diverse cell types with limited toxicity. These techniques open new venues for therapeutic uses in cells that may be more prone to off-target effects than iPSCs. The inhibition of NHEJ could be, however, a source of greater genomic instability in CRISPR applications.

## OFF-TARGET ACTIVITY ACCORDING TO CELL TYPE

The idea that the cell type might influence off-target activity is based on both the evidence that the on-target effects vary with cell type mostly due to chromatin condensation profiles and DSB repair has different efficiencies depending on the cell type. Using ChIP-seq to perform an unbiased genome-wide detection of Cas9 binding sites for different sgRNAs, it was found that open chromatin regions are enriched for off-target sites (Kuscu et al., 2014; Wu et al., 2014). In addition, a pioneer study has sparked the discussion by finding different potential off-target binding in HeKaS3 vs. HEK293T cells. Off-target events might depend on several factors beyond the DNA sequences, including the cell type, DNA-RNA heteroduplex, R-loop expansion, and so on (Duan et al., 2014). More importantly, for translational applications of the technique, the off-target potential of Cas9 has proven to be very limited in human iPSCs, where the DSB repair pathways are working normally, when compared to regular cell lines (Schwank et al., 2013; Smith et al., 2014; Veres et al., 2014).

Another cell-dependent mechanism studied for its effect on Cas9 is CpG site methylation, a cell-type-specific epigenetic mechanism. Despite an early study stating that CpG methylation and chromatin accessibility have little effect on targeting (Hsu et al., 2013), a subsequent analysis determined that CpG sites are negatively correlated with Cas9 binding, particularly when mismatches are present (Wu et al., 2014). We now know that the off-target potential is cell-type based due to the DSB repair system, the state of chromatin condensation, and the methylation pattern of CpG sites.

## OFF-TARGET ACTIVITY ACCORDING TO DELIVERY TYPE

Two back-to-back publications showed that the delivery of RNPs to target cells have reduced off-target effects compared to the use of plasmids encoding Cas9 and sgRNA (Kim et al., 2014; Ramakrishna et al., 2014). RNP delivery was performed using cationic lipid-mediated delivery (lipofectamine) and positively charged nanoparticles, leading to the same results. The RNP binds and cleaves the chromosomal DNA after delivery and is quickly degraded (Banas et al., 2020), leaving little time for the less efficient off-target cleavage, whereas the plasmid expression subsists for days before degradation. The ratio of off-target to on-target mutations has been calculated to be 28-fold lower in RNPs compared to the plasmid-based methods (Lu et al., 2019). In addition, the RNP dose has been calculated to be lower in the RNP complex transfection, resulting in lower cytotoxicity. Lipofectamine delivery was confirmed to work *in vivo* in the inner ear of mice (Zuris et al., 2015), and positively charged nanoparticles (lipidoids) were also confirmed to deliver the editing tools to the brain of mice (Wang et al., 2016).

Plasmid expression requires nuclear transport for RNA transcription and mRNA export for translation. The expressed Cas9 must then complex with the sgRNA and be transported into the nucleus for the editing to occur. To complete these steps, it has

been calculated that approximately 24 h are required (Shen et al., 2019). This time can hinder some experiments dependent on synchronized drug delivery or template donor delivery for HDR. The delivery of Cas9 may also be in the form of RNA for direct translation or protein (ready to work in a ribonucleoprotein complex composed of Cas9 and sgRNA); RNA overcomes the risk of integration and has poor stability compared to plasmid DNA.

Besides the delivery of Cas9 as a protein or as a plasmid, there are several delivery systems available. These can be grouped as viral vectors, physical (i.e., microinjection or electroporation), or chemical methods, (cationic-lipid transfection and lipid nanoparticles), ligand fusion tags, cell-penetrating peptides, and gold nanoparticles, among others. The most common viral vectors are adenoviruses, adeno-associated viruses, and lentiviruses, since these are the ones approved for clinical trials (Chakraborty, 2019).

Viral vectors may lead to unintended Cas9 genome integration, raising important questions on the safety of clinical trials. This topic remains a matter of debate (Chakraborty, 2019). There is also the risk of immunogenicity both from the use of the virus and the expression of the Cas9 protein. Even though the safety of viral approaches has been addressed and improved, the off-targets remain a concern due to the long-term exposure caused by the delivery method; furthermore, the viral vectors have limited packaging capacity.

An approach to avoid the risk of integration, immune response, and the long-term Cas9 expression has been created using a lentivirus-like particle (LVLP) encapsulating up to 100 copies of Cas9 mRNA per LVLP (Lu et al., 2019). Along the same lines, to limit the expression of the nuclease and reach a broad range of target cells, the Cas9/sgRNA RNP was recently packaged into extracellular nanovesicles, leading to an enormous reduction of off-targets compared to the plasmid-based delivery without loss of on-target cleavage (Gee et al., 2020).

## OFF-TARGETS IN OTHER GENE-EDITING TECHNOLOGIES: ZFNs, TALENs, AND BASE EDITORS

Precise gene editing of mammalian cells precedes CRISPR-mediated editing by over 25 years, with the work of Capecchi (Folger et al., 1984) and Smithies (Smithies et al., 1985), who proved that homologous recombination could integrate genetic material. Despite this breakthrough, the low efficiency of integration requires a selection marker or the introduction of a DSB to stimulate DNA repair. Meganucleases of microbial origin capable of recognizing sequences of 14–40 bp can generate DSBs, but the collection of sequences recognized by these enzymes is limited. The ability to introduce a DSB in most regions of the genome emerged with the invention of the chimeric nucleases that use the cleavage domain of the type IIS restriction enzyme *FokI* fused to modules of proteins that bind to specific DNA sequence. The first of these tools to appear were the Zinc Finger Nucleases (ZFN) (Geurts et al., 2009), in which the DNA specificity is given by the association of several finger domains, each contacting primarily with three specific

bp of DNA. In the second type of chimeric nucleases, known as Transcription Activator-Like Effector Nucleases (TALENs) (Miller et al., 2011), the DNA binding specificity is given by the modules of transcription activator-like effectors of bacterial origin, each recognizing a specific bp. ZFNs and TALENs were shown to work with similar efficiencies in iPSCs (Hockemeyer et al., 2009, 2011).

A simple *Pubmed* search shows that the CRISPR technology is now vastly more popular within the research community than the ZFN or TALEN ever were, but in a clinical setting, ultimately, it is the safety of the method that will prevail. There is a recent review comparing these two techniques with CRISPR-mediated editing (Khalil, 2020). Thus, here we will compare the three techniques exclusively concerning the off-targets. In principle, both ZFN and TALEN should have a lower frequency of off-targets than the original CRISPR/Cas9 technique because the *FokI* cleavage domain works as a dimer, which requires the targeting of two neighboring regions in the genome to produce a DSB, thus vastly decreasing the number of off-targets, as explained in a previous section. However, the comparison is not straightforward, due to the different complexity of the sequences targeted by each strategy and the different biochemistry of RNA:DNA and protein:DNA interactions. Nevertheless, both for the chimeric nucleases and CRISPR, increasing the interaction length with the target sequence did not improve the specificity or efficiency (Ran et al., 2013; Yan et al., 2013).

ZFN off-targets with up to three mutations relative to the canonical target sequence were identified using *in vitro* cell-free strategies and confirmed in cell lines and zebrafish (Gupta et al., 2011; Pattanayak et al., 2011). A thorough study using the unbiased IDLV-based methods to tag ZFN-dependent DSBs revealed several off-targets, but overall that ZFNs are highly specific, particularly when obligate heterodimeric *FokI* domains are used. In a TALEN-engineered mouse, no off-targets were detected, although only low-resolution and limited direct methods were used (Gabriel et al., 2011). And a hybrid strategy combining a ZFN and a TALEN seems to improve specificity even further (Yan et al., 2013). Unfortunately, properly controlled studies directly addressing the off-targets produced by CRISPR and the chimeric nucleases are scarce. In a goat genome editing project, CRISPR/Cas9 was reported to have introduced more off-targets than a TALEN approach for the same editing (Zhang et al., 2019); in another set of experiments using IDLV tagging of DSBs, the frequency of off-targets generated by Cas9 and TALENs were directly compared and the results revealed that three different TALENs generated no off-targets, whereas three of six Cas9 RNPs had detectable off-target activity (Wang et al., 2015). Interestingly, besides detecting the known tolerance for a few mismatches, the authors found patterns suggesting that a single-base bulge formed between the sgRNA and the genomic target could also be tolerated. Thus, paradoxically, the data suggest that the wild-type SpCas9 that has been the focus of so much attention introduces more off-targets than the TALENs and, as mentioned above, Cas12a.

The dCas9 (catalytically null version of Cas9) has been fused to many proteins to become a molecular swiss-army knife capable of driving different enzymatic activities or useful structural

properties (e.g., fluorescent proteins) to precise locations in the genome. Some of these strategies have also transformed the original Cas nuclease into base editors by fusing a dCas9 or a Cas nickase with natural cytosine base editors (capable of C-to-T conversions) or engineered adenine base editors (that catalyze A-to-G conversions) to generate single-base editors capable of editing close to a third of the annotated human pathogenic alleles (reviewed in Anzalone et al., 2020). The absence of a potentially dangerous DSB during the editing process is an attractive feature of these tools. However, they use deaminases that have been associated with mutational signatures in several types of cancers. A study in mouse embryos revealed that CRISPR-Cas9 and the adenine base editor had a frequency of mutations close to the spontaneous mutation rate; in contrast, cytosine base editing induced single nucleotide variants at more than 20-fold higher frequencies (Zuo et al., 2019). Thus, base editors have been fine-tuned in recent years through the titration of the expression levels, the use of a “high-fidelity” Cas9 partner (e.g., Rees et al., 2017; Hu et al., 2018; Lee et al., 2018; Doman et al., 2020) or the use of deaminases that preferentially deaminate cytidines in specific motifs (e.g., Gehrke et al., 2018). Nevertheless, besides the Cas9-dependent off-targets, these tools also produce Cas9-independent deamination events that are difficult to measure because of their random nature and rarity. This problem has been addressed independently by two groups, and these efforts include a thorough study of 153 deaminase domains to isolate and further engineer those with the lowest Cas9-independent off-target DNA deamination (Doman et al., 2020; Yu et al., 2020).

## CONCLUSION AND FUTURE PERSPECTIVES

Off-targets are not a problem for basic research using cell lines and animal models. It is sufficient to adopt the experimental design developed several years ago for shRNA studies to ensure that the observed effect is due to the on-target edition. In other words, if two clones obtained with different sgRNAs have the same phenotype, it is very unlikely that the effect is due to a technical artifact because the two guides will have different constellations of off-targets. In addition, rescue experiments are also possible and could be reassuring. It is only in the context of the ongoing clinical trials in humans using somatic cells and in the possible future edition of the germline that the off-targets remain a problem, although not the only one. That is the motivation fueling the many approaches to attenuate the off-target problem.

Two tragic stories haunt the field of gene therapy. Mr. Jesse Gelsinger died in 1999 from an exacerbated innate immune response to an adenoviral vector carrying a gene to correct his partial deficiency of ornithine transcarbamylase (Wilson, 2009). Not long after, out of 20 children suffering from X-Linked Severe Combined Immunodeficiency who were treated by gene therapy in two independent trials, five developed T cell leukemia, which would prove fatal in one case (Cavazzana et al., 2016); all cases resulted from the insertion of the retroviral vector carrying the correct form of the  $\gamma$ c-encoding gene near tumor-promoting

genes. The field eventually recovered from these early failures by developing safer vectors and protocols, and also novel technologies, such as CRISPR, that restore the correct DNA sequence and proper regulation. CRISPR gene editing is now regarded as a much safer gene therapy than the crude pioneering viral-based therapies, which were not gene editing approaches. However, any tragic clinical outcome associated with these novel technologies and whether the prior risk-benefit analyses need to be recalibrated.

CRISPR-mediated gene editing has been described in models of Duchene muscular dystrophy (DMD), a genetic disease caused by mutations in the dystrophin gene (*DMD*), which encodes a cytoskeletal protein essential for muscle cell membrane integrity (Zhang et al., 2017). Cas12a-mediated gene editing was first employed in DMD patient-derived iPSCs and also in a murine model of the disease, with a mutation in Exon 51 of the *DMD* gene that generates a premature stop codon. The Cas12a-mediated gene editing rescued the disease phenotype in both models. The off-target activity was assessed by the T7 Endonuclease assay in 10 potential off-target sites, predicted by Cas-OFFinder, with no detected activity (Zhang et al., 2017). In a follow-up study, the same group applied the same strategy in a canine model of DMD in which the delivery of Cas9:sgRNA was done with adeno-associated viruses, via intramuscular injection (Amoasii et al., 2018). The disease phenotype was rescued with the application of CRISPR-mediated gene editing. Three off-target sites in coding regions were predicted using the MIT CRISPR design tool and deep sequencing of those sequences was performed. There were no genetic alterations in the predicted off-target sites (Amoasii et al., 2018). CRISPR/Cas9-based gene editing has been used in several other animal models of disease, particularly the murine ones. The off-targets have been measured using biased methods and Sanger sequencing of a small number of cloned amplicons from a few potential off-targets, such as in a model of Huntington’s disease (Monteys et al., 2017), biased methods and deep sequencing of amplicons from a modest number of potential off-targets, such as in a murine genetic model of amyotrophic lateral sclerosis (Gaj et al., 2017), and biased and unbiased methods (whole-exome and whole-genome sequencing), such as in a study of an Alzheimer’s disease model (Park et al., 2019). The overall picture is that off-targets and chromosomal rearrangements have not been detected at alarming frequencies in any CRISPR/Cas9-based gene editing to correct genetic deficiencies in animal models of disease. However, the methodology of off-target detection and quantification has not been standardized across these studies.

It is worth keeping in mind that clinical applications of the CRISPR technology in humans started a few years ago, even though the technique is still being perfected. In February 2019, the first report of cancer patients treated with CAR-T cells that had been edited using a combination of a retroviral vector and three CRISPR-knocked-out genes was published (Stadtmauer et al., 2020); although the immunotherapy itself did not provide complete cancer treatment, the engineered cells were well-tolerated. Another CRISPR technology application in the clinic was a phase I clinical trial of edited T cells, where the goal was to disrupt the *PD-1* gene to overcome the high expression of PD-L1

in the malignant cells of refractory non-small-cell lung cancer (Lu et al., 2020). In that study, the off-targets were determined using Cas-OFFinder and indels in those regions were detected by NGS. The off-target mutational frequency was 0.05% in 18 determined off-target sites, which was considered to be safe. Moreover, a CRISPR/Cas9 application in the treatment of the Transfusion-dependent  $\beta$ -thalassemia and sickle cell disease (both severe monogenic diseases) was reported in January of 2021 (Frangoul et al., 2021). CD34+ hematopoietic stem cells were transfected with a Cas9:sgRNA complex targeting the *BCL11A* erythroid-specific enhancer. A total of 223 off-target sites were scanned, and no evidence of off-target activity was detected. Although these reports suggest that the off-targets may not be a problem for clinical applications, three potential technical limitations should be kept in mind. First, any long-term side-effects of the CRISPR editing, such as a slow-growing cancer, will remain undetectable for some time. Second, even when amplicon-based and hybridization capture steps are used to enrich the target genomic regions to be analyzed by NGS, this technique has a resolution limit, particularly when bulk cell populations are analyzed. Third, whole-genome sequencing of large collections of small pools of cells to evaluate in an unbiased way the frequency of off-targets is currently not possible to implement on a routine basis in clinical settings. Notwithstanding these limitations, there are over 60 ongoing clinical trials using CRISPR technology to treat more than 30 different diseases (see CRISPR News Medicine<sup>1</sup>). Standardized measures of off-target analysis in these trials would be welcome.

The many ingenious strategies developed over the last 8 years to decrease the frequency of off-targets that we have described have not exhausted the field. Recently, CRISPR GUARD was reported, a technique based on the targeting of the off-targets by RNPs with short RNAs that do not allow nuclease activity but can outcompete the on-target guide RNA at selected off-targets (Coelho et al., 2020). Although the practicability of this technique remains to be tested, it illustrates how truly novel ideas keep emerging. Such ingenuity will be required to tackle several remaining challenges, including: (1) the generation of improved *in silico* tools that integrate what is known about the tolerance for mismatches but also on single-base bulges and all

<sup>1</sup> <https://crisprmedicineneeds.com/clinical-trials/>

## REFERENCES

- Akcakeya, P., Bobbin, M. L., Guo, J. A., Malagon-Lopez, J., Clement, K., Garcia, S. P., et al. (2018). In vivo CRISPR editing with no detectable genome-wide off-target mutations. *Nature* 561, 416–419. doi: 10.1038/s41586-018-0500-9
- Amoasi, L., Hildyard, J., Li, H., Sanchez-Ortiz, E., Mireault, A., Caballero, D., et al. (2018). Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. *Science* 362, 86–91.
- Anzalone, A. V., Koblan, L. W., and Liu, D. R. (2020). Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38, 824–844. doi: 10.1038/s41587-020-0561-9
- Bae, S., Park, J., and Kim, J.-S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. doi: 10.1093/bioinformatics/btu048

the epigenetic features that influence the chances of a hit in an off-target; (2) more detailed and cell-specific analyses using unbiased methods of the off-targets in cells with therapeutic potential, such as human CD8 T cells and iPSCs of different sources; (3) a better integration of the strategies to reduce off-targets generated at several fronts and the implementation of standards to facilitate comparisons across different studies. This third point seems crucial. Although platforms such as Addgene have facilitated the sharing of molecular tools, it seems that the progress in the field has been delayed by the failure to implement two strategies. First, using the same cell line or animal models, engineered with reporters or not, and common protocol and readouts to compare the results of the different tools would provide quantitative comparisons that would quickly identify the best tools across different platforms. Second, despite the quest for innovation and commercial interests do not create incentives to fully explore the combination of methodologies already described that could lead to cumulative or synergistic decreases in the frequency of off-targets.

## AUTHOR CONTRIBUTIONS

MMV, MC-F, and VMB wrote the manuscript, which was edited by JMPJ and JTP. JMPJ worked in the preparation of the data for **Figure 2**. MMV generated all the figures. All authors contributed to the article and approved the submitted version.

## FUNDING

The work of VMB was funded by iNOVA4Health – UIDB/Multi/04462/2020 and UIDP/Multi/04462/2020, a program financially supported by Fundação para a Ciência e Tecnologia (FCT)/Ministério da Educação e Ciência through national funds, and the FCT grant PTDC/BEX-BCM/5900/2014.

## ACKNOWLEDGMENTS

We thank Thiago Lopes Carvalho for his insights on the several topics addressed and the revision of the text. We also thank the reviewers for the helpful feedback.

- Baltimore, D., Berg, P., Botchan, M., Carroll, D., Charo, R. A., Church, G., et al. (2015). Biotechnology. A prudent path forward for genomic engineering and germline gene modification. *Science* 348, 36–38. doi: 10.1126/science.aab1028
- Banas, K., Rivera-Torres, N., Bialk, P., Yoo, B.-C., and Kmiec, E. B. (2020). Kinetics of nuclear uptake and site-specific DNA cleavage during CRISPR-Directed gene editing in solid tumor cells. *Mol. Cancer Res.* 18, 891–902. doi: 10.1158/1541-7786.MCR-19-1208
- Bischoff, N., Wimberger, S., Maresca, M., and Brakebusch, C. (2020). Improving precise CRISPR genome editing by small molecules: is there a magic potion? *Cells* 9:1318. doi: 10.3390/cells9051318
- Cameron, P., Fuller, C. K., Donohoue, P. D., Jones, B. N., Thompson, M. S., Carter, M. M., et al. (2017). Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods* 14, 600–606. doi: 10.1038/nmeth.4284

- Cavazzana, M., Six, E., Lagresle-Peyrou, C., André-Schmutz, I., and Hachein-Bey-Abina, S. (2016). Gene therapy for X-linked severe combined immunodeficiency: where do we stand? *Hum. Gene Ther.* 27, 108–116. doi: 10.1089/hum.2015.137
- Chakraborty, S. (2019). Sequencing data from Massachusetts General Hospital shows Cas9 integration into the genome, highlighting a serious hazard in gene-editing therapeutics. *F1000Res.* 8:1846. doi: 10.12688/f1000research.20744.1
- Chen, F., Ding, X., Feng, Y., Seebeck, T., Jiang, Y., and Davis, G. D. (2017). Targeted activation of diverse CRISPR-Cas systems for mammalian genome editing via proximal CRISPR targeting. *Nat. Commun.* 8:14958. doi: 10.1038/ncomms14958
- Cho, S. W., Kim, S., Kim, Y., Kweon, J., and Kim, H. S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24, 132–141. doi: 10.1101/gr.162339.113
- Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., et al. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* 33, 543–548. doi: 10.1038/nbt.3198
- Chuai, G., Ma, H., Yan, J., Chen, M., Hong, N., Xue, D., et al. (2018). DeepCRISPR: optimized CRISPR guide RNA design by deep learning. *Genome Biol.* 19:80. doi: 10.1186/s13059-018-1459-4
- Coelho, M. A., De Braekeleer, E., Firth, M., Bista, M., Lukasiak, S., Cuomo, M. E., et al. (2020). CRISPR GUARD protects off-target sites from Cas9 nuclease activity using short guide RNAs. *Nat. Commun.* 11:4132. doi: 10.1038/s41467-020-17952-5
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143
- Ding, X., Seebeck, T., Feng, Y., Jiang, Y., Davis, G. D., and Chen, F. (2019). Improving CRISPR-Cas9 genome editing efficiency by fusion with chromatin-modulating peptides. *CRISPR J.* 2, 51–63. doi: 10.1089/crispr.2018.0036
- Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34, 184–191. doi: 10.1038/nbt.3437
- Doman, J. L., Raguram, A., Newby, G. A., and Liu, D. R. (2020). Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* 38, 620–628. doi: 10.1038/s41587-020-0414-6
- Duan, J., Lu, G., Xie, Z., Lou, M., Luo, J., Guo, L., et al. (2014). Genome-wide identification of CRISPR/Cas9 off-targets in human genome. *Cell Res.* 24, 1009–1012. doi: 10.1038/cr.2014.87
- Folger, K., Thomas, K., and Capecchi, M. R. (1984). Analysis of homologous recombination in cultured mammalian cells. *Cold Spring Harb. Symp. Quant. Biol.* 49, 123–138. doi: 10.1101/sqb.1984.049.01.016
- Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and  $\beta$ -Thalassemia. *N. Engl. J. Med.* 384, 252–260. doi: 10.1056/NEJMoa2031054
- Fu, B. X. H., Smith, J. D., Fuchs, R. T., Mabuchi, M., Curcuru, J., Robb, G. B., et al. (2019). Target-dependent nickase activities of the CRISPR-Cas nucleases Cpf1 and Cas9. *Nat. Microbiol.* 4, 888–897. doi: 10.1038/s41564-019-0382-0
- Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826. doi: 10.1038/nbt.2623
- Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., and Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279–284. doi: 10.1038/nbt.2808
- Gabriel, R., Lombardo, A., Arens, A., Miller, J. C., Genovese, P., Kaepffel, C., et al. (2011). An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* 29, 816–823. doi: 10.1038/nbt.1948
- Gaj, T., Ojala, D. S., Ekman, F. K., Byrne, L. C., Limsirichai, P., and Schaffer, D. V. (2017). In vivo genome editing improves motor function and extends survival in a mouse model of ALS. *Sci. Adv.* 3:eear3952. doi: 10.1126/sciadv.aar3952
- Gasiunas, G., Barrangou, R., Horvath, P., and 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. doi: 10.1073/pnas.1208507109
- Gasiunas, G., Young, J. K., Karvelis, T., Kazlauskas, D., Urbaitis, T., Jasnauskaitė, M., et al. (2020). A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat. Commun.* 11:5512. doi: 10.1038/s41467-020-19344-1
- Gee, P., Lung, M. S. Y., Okuzaki, Y., Sasakawa, N., Iguchi, T., Makita, Y., et al. (2020). Extracellular nanovesicles for packaging of CRISPR-Cas9 protein and sgRNA to induce therapeutic exon skipping. *Nat. Commun.* 11:1334. doi: 10.1038/s41467-020-14957-y
- Gehrke, J. M., Cervantes, O., Clement, M. K., Wu, Y., Zeng, J., Bauer, D. E., et al. (2018). An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* 36, 977–982. doi: 10.1038/nbt.4199
- Geurts, A. M., Cost, G. J., Freyvert, Y., Zeitler, B., Miller, J. C., Choi, V. M., et al. (2009). Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 325:433. doi: 10.1126/science.1172447
- Guillinger, J. P., Thompson, D. B., and Liu, D. R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 32, 577–582. doi: 10.1038/nbt.2909
- Gupta, A., Meng, X., Zhu, L. J., Lawson, N. D., and Wolfe, S. A. (2011). Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic Acids Res.* 39, 381–392. doi: 10.1093/nar/gkq787
- Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F., and Chin, L. (2016). Post-translational regulation of Cas9 during G1 enhances homology-directed repair. *Cell Rep.* 14, 1555–1566. doi: 10.1016/j.celrep.2016.01.019
- Haussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., et al. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* 17:148. doi: 10.1186/s13059-016-1012-2
- Havlicek, S., Shen, Y., Alpugu, Y., Bruntraeger, M. B., Zufir, N. B. M., Phuah, Z. Y., et al. (2017). Re-engineered RNA-Guided FokI-Nucleases for improved genome editing in human cells. *Mol. Ther.* 25, 342–355. doi: 10.1016/j.jymthe.2016.11.007
- Heyer, W.-D., Ehmsen, K. T., and Liu, J. (2010). Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* 44, 113–139. doi: 10.1146/annurev-genet-051710-150955
- Hinz, J. M., Laughery, M. F., and Wyrick, J. J. (2015). Nucleosomes inhibit Cas9 endonuclease activity in vitro. *Biochemistry* 54, 7063–7066. doi: 10.1021/acs.biochem.5b01108
- Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R. C., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* 27, 851–857. doi: 10.1038/nbt.1562
- Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* 29, 731–734. doi: 10.1038/nbt.1927
- Horlbeck, M. A., Witkowsky, L. B., Guglielmi, B., Replogle, J. M., Gilbert, L. A., Villalta, J. E., et al. (2016). Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife* 5:e12677. doi: 10.7554/eLife.12677
- Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi: 10.1038/nbt.2647
- Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63. doi: 10.1038/nature26155
- Hu, J., Meyers, R. M., Dong, J., Panchakshari, R. A., Alt, F. W., and 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. doi: 10.1038/nprot.2016.043
- Isaac, R. S., Jiang, F., Doudna, J. A., Lim, W. A., Narlikar, G. J., and Almeida, R. (2016). Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. *Elife* 5:e13450. doi: 10.7554/eLife.13450
- Janssen, J. M., Chen, X., Liu, J., and Gonçalves, M. A. F. V. (2019). The chromatin structure of CRISPR-Cas9 target DNA controls the balance between mutagenic

- and homology-directed gene-editing events. *Mol. Ther. Nucleic Acids* 16, 141–154. doi: 10.1016/j.omtn.2019.02.009
- Jensen, K. T., Floe, L., Petersen, T. S., Huang, J., Xu, F., Bolund, L., et al. (2017). Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett.* 591, 1892–1901. doi: 10.1002/1873-3468.12707
- Jiang, F., and Doudna, J. A. (2017). CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* 46, 505–529. doi: 10.1146/annurev-biophys-062215-010822
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829
- Khalil, A. M. (2020). The genome editing revolution: review. *J. Genet. Eng. Biotechnol.* 18:68. doi: 10.1186/s43141-020-00078-y
- Kim, D., and Kim, J.-S. (2018). DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA. *Genome Res.* 28, 1894–1900. doi: 10.1101/gr.236620.118
- Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H. R., et al. (2015). Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* 12, 237–243, 1 p following 243. doi: 10.1038/nmeth.3284
- Kim, D., Kim, J., Hur, J. K., Been, K. W., Yoon, S.-H., and Kim, J.-S. (2016). Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* 34, 863–868. doi: 10.1038/nbt.3609
- Kim, E., Koo, T., Park, S. W., Kim, D., Kim, K., Cho, H.-Y., et al. (2017). In vivo genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat. Commun.* 8:14500. doi: 10.1038/ncomms14500
- Kim, S., Bae, T., Hwang, J., and Kim, J.-S. (2017). Rescue of high-specificity Cas9 variants using sgRNAs with matched 5' nucleotides. *Genome Biol.* 18:218. doi: 10.1186/s13059-017-1355-3
- Kim, S., Kim, D., Cho, S. W., Kim, J., and Kim, J.-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012–1019. doi: 10.1101/gr.171322.113
- Kleistiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., et al. (2016a). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495. doi: 10.1038/nature16526
- Kleistiver, B. P., Tsai, S. Q., Prew, M. S., Nguyen, N. T., Welch, M. M., Lopez, J. M., et al. (2016b). Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* 34, 869–874. doi: 10.1038/nbt.3620
- Kocak, D. D., Josephs, E. A., Bhandarkar, V., Adkar, S. S., Kwon, J. B., and Gersbach, C. A. (2019). Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat. Biotechnol.* 37, 657–666. doi: 10.1038/s41587-019-0095-1
- Kumar, N., Stanford, W., de Solis, C., Aradhana, Abraham, N. D., Dao, T.-M. J., et al. (2018). The development of an AAV-Based CRISPR SaCas9 genome editing system that can be delivered to neurons in vivo and regulated via doxycycline and cre-recombinase. *Front. Mol. Neurosci.* 11:413. doi: 10.3389/fnmol.2018.00413
- Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* 32, 677–683. doi: 10.1038/nbt.2916
- Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B., and Valen, E. (2016). CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* 44, W272–W276. doi: 10.1093/nar/gkw398
- Lander, E. S. (2016). The heroes of CRISPR. *Cell* 164, 18–28. doi: 10.1016/j.cell.2015.12.041
- Lander, E. S., Baylis, F., Zhang, F., Charpentier, E., Berg, P., Bourgain, C., et al. (2019). Adopt a moratorium on heritable genome editing. *Nature* 567, 165–168. doi: 10.1038/d41586-019-00726-5
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25
- Lee, C. M., Cradick, T. J., and Bao, G. (2016). The *Neisseria meningitidis* CRISPR-Cas9 system enables specific genome editing in mammalian cells. *Mol. Ther.* 24, 645–654. doi: 10.1038/mt.2016.8
- Lee, J. K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y.-H., et al. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. *Nat. Commun.* 9:3048. doi: 10.1038/s41467-018-05477-x
- Lee, K., Zhang, Y., Kleinstiver, B. P., Guo, J. A., Aryee, M. J., Miller, J., et al. (2019). Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. *Plant Biotechnol. J.* 17, 362–372. doi: 10.1111/pbi.12982
- Lin, S., Staahl, B. T., Alla, R. K., and Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife (Cambridge)* 3:e04766. doi: 10.7554/eLife.04766
- Liu, G., Yin, K., Zhang, Q., Gao, C., and Qiu, J.-L. (2019). Modulating chromatin accessibility by transactivation and targeting proximal dsRNAs enhances Cas9 editing efficiency in vivo. *Genome Biol.* 20:145. doi: 10.1186/s13059-019-1762-8
- Lu, B., Javidi-Parsijani, P., Makani, V., Mehraein-Ghomi, F., Sarhan, W. M., Sun, D., et al. (2019). Delivering SaCas9 mRNA by lentivirus-like bionanoparticles for transient expression and efficient genome editing. *Nucleic Acids Res.* 47:e44. doi: 10.1093/nar/gkz093
- Lu, Y., Xue, J., Deng, T., Zhou, X., Yu, K., Deng, L., et al. (2020). Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer. *Nat. Med.* 26, 732–740. doi: 10.1038/s41591-020-0840-5
- Makarova, K. S., and Koonin, E. V. (2015). Annotation and classification of CRISPR-Cas systems. *Methods Mol. Biol.* 1311, 47–75. doi: 10.1007/978-1-4939-2687-9\_4
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033
- Mao, Z., Bozzella, M., Seluanov, A., and Gorbunova, V. (2008). DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* 7, 2902–2906. doi: 10.4161/cc.7.18.6679
- Matsumoto, D., Tamamura, H., and Nomura, W. (2020). A cell cycle-dependent CRISPR-Cas9 activation system based on an anti-CRISPR protein shows improved genome editing accuracy. *Commun. Biol.* 3, 601–610. doi: 10.1038/s42003-020-01340-2
- Maxwell, K. L. (2017). The Anti-CRISPR story: a battle for survival. *Mol. Cell* 68, 8–14. doi: 10.1016/j.molcel.2017.09.002
- Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., et al. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29, 143–148. doi: 10.1038/nbt.1755
- Mojica, F. J. M., Diez-Villasenor, C., Soria, E., and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* 36, 244–246.
- Monteys, A. M., Ebanks, S. A., Keiser, M. S., and Davidson, B. L. (2017). CRISPR/Cas9 editing of the mutant huntingtin allele in vitro and in vivo. *Mol. Ther.* 25, 12–23. doi: 10.1016/j.ymthe.2016.11.010
- Moreno-Mateos, M. A., Fernandez, J. P., Rouet, R., Vejnar, C. E., Lane, M. A., Mis, E., et al. (2017). CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nat. Commun.* 8:2024. doi: 10.1038/s41467-017-01836-2
- Moreno-Mateos, M. A., Vejnar, C. E., Beaudoin, J.-D., Fernandez, J. P., Mis, E. K., Khokha, M. K., et al. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* 12, 982–988. doi: 10.1038/nmeth.3543
- Müller, M., Lee, C. M., Gasiunas, G., Davis, T. H., Cradick, T. J., Siksnys, V., et al. (2016). *Streptococcus thermophilus* CRISPR-Cas9 systems enable specific editing of the human genome. *Mol. Ther.* 24, 636–644. doi: 10.1038/mt.2015.218
- Murugan, K., Seetharam, A. S., Severin, A. J., and Sashital, D. G. (2020). CRISPR-Cas12a has widespread off-target and dsDNA-nicking effects. *J. Biol. Chem.* 295, 5538–5553. doi: 10.1074/jbc.RA120.012933
- Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., et al. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156, 935–949. doi: 10.1016/j.cell.2014.02.001
- Pardo, B., Gómez-González, B., and Aguilera, A. (2009). DNA repair in mammalian cells: DNA double-strand break repair: how to fix a broken relationship. *Cell. Mol. Life Sci.* 66, 1039–1056. doi: 10.1007/s00018-009-8740-3
- Park, H., Oh, J., Shim, G., Cho, B., Chang, Y., Kim, S., et al. (2019). In vivo neuronal gene editing via CRISPR-Cas9 amphiphilic nanocomplexes alleviates deficits in mouse models of Alzheimer's disease. *Nat. Neurosci.* 22, 524–528. doi: 10.1038/s41593-019-0352-0

- Pattanayak, V., Ramirez, C. L., Joung, J. K., and Liu, D. R. (2011). Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat. Methods* 8, 765–770. doi: 10.1038/nmeth.1670
- Perez, A. R., Pritykin, Y., Vidigal, J. A., Chhangawala, S., Zamparo, L., Leslie, C. S., et al. (2017). GuideScan software for improved single and paired CRISPR guide RNA design. *Nat. Biotechnol.* 35, 347–349. doi: 10.1038/nbt.3804
- Ramakrishna, S., Dad, A.-B. K., Beloor, J., Gopalappa, R., Lee, S.-K., and Kim, H. (2014). Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res.* 24, 1020–1027. doi: 10.1101/gr.171264.113
- Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389. doi: 10.1016/j.cell.2013.08.021
- Rees, H. A., Komor, A. C., Yeh, W.-H., Caetano-Lopes, J., Warman, M., Edge, A. S. B., et al. (2017). Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat. Commun.* 8:15790. doi: 10.1038/ncomms15790
- Rouet, P., Smith, F., and Jasin, M. (1994). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 91, 6064–6068.
- Ryan, D. E., Taussig, D., Steinfeld, I., Phadnis, S. M., Lunstad, B. D., Singh, M., et al. (2018). Improving CRISPR-Cas specificity with chemical modifications in single-guide RNAs. *Nucleic Acids Res.* 46, 792–803. doi: 10.1093/nar/gkx1199
- Schwank, G., Koo, B.-K., Sasselli, V., Dekkers, J. F., Heo, I., Demircan, T., et al. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* 13, 653–658. doi: 10.1016/j.stem.2013.11.002
- Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., et al. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat. Methods* 11, 399–402. doi: 10.1038/nmeth.2857
- Shen, C.-C., Hsu, M.-N., Chang, C.-W., Lin, M.-W., Hwu, J.-R., Tu, Y., et al. (2019). Synthetic switch to minimize CRISPR off-target effects by self-restricting Cas9 transcription and translation. *Nucleic Acids Res.* 47:e13. doi: 10.1093/nar/gky1165
- Slymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88. doi: 10.1126/science.aad5227
- Smith, C., Gore, A., Yan, W., Abalde-Atristain, L., Li, Z., He, C., et al. (2014). Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell* 15, 12–13. doi: 10.1016/j.stem.2014.06.011
- Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A., and Kucherlapati, R. S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* 317, 230–234. doi: 10.1038/317230a0
- Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y. E., and Zhang, J. (2016). RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* 7:10548. doi: 10.1038/ncomms10548
- Stadtmauer, E. A., Fraietta, J. A., Davis, M. M., Cohen, A. D., Weber, K. L., Lancaster, E., et al. (2020). CRISPR-engineered T cells in patients with refractory cancer. *Science* 367:eaba7365. doi: 10.1126/science.aba7365
- Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J., and Mateo, J. L. (2015). CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One* 10:e0124633. doi: 10.1371/journal.pone.0124633
- Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., and Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507, 62–67. doi: 10.1038/nature13011
- Strohkendl, I., Saifuddin, F. A., Rybarski, J. R., Finkelstein, I. J., and Russell, R. (2018). Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a. *Mol. Cell* 71, 816–824.e3. doi: 10.1016/j.molcel.2018.06.043
- Sundaresan, R., Parameshwaran, H. P., Yogesha, S. D., Keilbarth, M. W., and Rajan, R. (2017). RNA-Independent DNA cleavage activities of Cas9 and Cas12a. *Cell Rep.* 21, 3728–3739. doi: 10.1016/j.celrep.2017.11.100
- Torres-Perez, R., Garcia-Martin, J. A., Montoliu, L., Oliveros, J. C., and Pazos, F. (2019). WeReview: CRISPR tools-live repository of computational tools for assisting CRISPR/Cas experiments. *Bioengineering (Basel)* 6:63. doi: 10.3390/bioengineering6030063
- Tsai, S. Q., Nguyen, N. T., Malagon-Lopez, J., Topkar, V. V., Aryee, M. J., and Joung, J. K. (2017). CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* 14, 607–614. doi: 10.1038/nmeth.4278
- Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., et al. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576. doi: 10.1038/nbt.2908
- Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187–197. doi: 10.1038/nbt.3117
- Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., et al. (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* 24, 1216–1224. doi: 10.1038/s41591-018-0137-0
- Veres, A., Gosis, B. S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., et al. (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. doi: 10.1016/j.stem.2014.04.020
- Wang, M., Zuris, J. A., Meng, F., Rees, H., Sun, S., Deng, P., et al. (2016). Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2868–2873. doi: 10.1073/pnas.1520244113
- Wang, X., Wang, Y., Wu, X., Wang, J., Wang, Y., Qiu, Z., et al. (2015). Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat. Biotechnol.* 33, 175–178. doi: 10.1038/nbt.3127
- Wang, Y., Wang, B., Xie, H., Ren, Q., Liu, X., Li, F., et al. (2019). Efficient human genome editing using SaCas9 ribonucleoprotein complexes. *Biotechnol. J.* 14:e1800689. doi: 10.1002/biot.201800689
- Wienert, B., Wyman, S. K., Richardson, C. D., Yeh, C. D., Akcakaya, P., Porritt, M. J., et al. (2019). Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science* 364, 286–289. doi: 10.1126/science.aav9023
- Wilson, J. M. (2009). Lessons learned from the gene therapy trial for ornithine transcarbamylase deficiency. *Mol. Genet. Metab.* 96, 151–157. doi: 10.1016/j.ymgme.2008.12.016
- Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* 32, 670–676. doi: 10.1038/nbt.2889
- Yan, W. X., Mirzazadeh, R., Garnerone, S., Scott, D., Schneider, M. W., Kallas, T., et al. (2017). BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun.* 8:15058. doi: 10.1038/ncomms15058
- Yan, W., Smith, C., and Cheng, L. (2013). Expanded activity of dimer nucleases by combining ZFN and TALEN for genome editing. *Sci. Rep.* 3:2376. doi: 10.1038/srep02376
- Yarrington, R. M., Verma, S., Schwartz, S., Trautman, J. K., and Carroll, D. (2018). Nucleosomes inhibit target cleavage by CRISPR-Cas9 in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 115, 9351–9358. doi: 10.1073/pnas.1810062115
- Yu, Y., Leete, T. C., Born, D. A., Young, L., Barrera, L. A., Lee, S.-J., et al. (2020). Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat. Commun.* 11:2052. doi: 10.1038/s41467-020-15887-5
- Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015a). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038
- Zetsche, B., Volz, S. E., and Zhang, F. (2015b). A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat. Biotechnol.* 33, 139–142. doi: 10.1038/nbt.3149

- Zhang, J., Liu, J., Yang, W., Cui, M., Dai, B., Dong, Y., et al. (2019). Comparison of gene editing efficiencies of CRISPR/Cas9 and TALEN for generation of MSTN knock-out cashmere goats. *Theriogenology* 132, 1–11. doi: 10.1016/j.theriogenology.2019.03.029
- Zhang, J.-P., Yang, Z.-X., Zhang, F., Fu, Y.-W., Dai, X.-Y., Wen, W., et al. (2021). HDAC inhibitors improve CRISPR-mediated HDR editing efficiency in iPSCs. *Sci. China Life Sci.* 64, 1–14. doi: 10.1007/s11427-020-1855-4
- Zhang, Y., Long, C., Li, H., McAnally, J. R., Baskin, K. K., Shelton, J. M., et al. (2017). CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice. *Sci. Adv.* 3:e1602814.
- Zheng, N., Li, L., and Wang, X. (2020). Molecular mechanisms, off-target activities, and clinical potentials of genome editing systems. *Clin. Transl. Med.* 10, 412–426. doi: 10.1002/ctm2.34
- Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., et al. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364, 289–292. doi: 10.1126/science.aav9973
- Zuris, J. A., Thompson, D. B., Shu, Y., Guilinger, J. P., Bessen, J. L., Hu, J. H., et al. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* 33, 73–80. doi: 10.1038/nbt.3081

**Conflict of Interest:** VMB runs a CRISPR-based gene editing service at Nova Medical School, a public university.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Vicente, Chaves-Ferreira, Jorge, Proença and Barreto. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.