BioTechniques
The International Journal of Life Science Methods

CRISPR

CONTENTS

fsg

CRISPR editing validation, immunostaining and DNA sequencing of individual fixed bovine embryos

Bradford W Daigneault‡,[1], Marcela Vilarino[3], Sandeep K Rajput[4], Tristan Frum[2], George W Smith[1] & Pablo J Ross*,‡,[3]

[1]Department of Animal Science, Michigan State University, East Lansing, Michigan, USA; [2]Department of Biochemistry & Molecular Biology, Michigan State University, East Lansing, Michigan, USA; [3]Department of Animal Science, University of California – Davis, Davis, CA, USA; [4]Colorado Center for Reproductive Medicine, Lone Tree, Colorado, USA

BioTechniques 65: 281-283 (November 2018) 10.2144/btn-2018-0051
Keywords: bovine • CRISPR • DNA • embryo • genotype • immunohistochemistry • sequencing

CRISPR technologies used for mammalian embryology have wide implications from basic research to applications in agriculture and biomedicine. Confirmation of successful gene editing following CRISPR/Cas9 delivery is often limited to either protein expression or sequencing analyses of embryos but not both, due to technical challenges. Herein we report an integrative approach for evaluating both protein expression and genotype of single embryos from fixed bovine embryos previously subjected to CRISPR/Cas9 microinjection. The techniques described facilitate investigation of functional genomics in bovine embryos compatible with gene editing in livestock after zygotic CRISPR microinjection. These methods avoid traditional avenues that necessitate the use of gene-edited cell lines followed by nuclear transfer that hinder efficiency, limit physiological relevance and contribute to technical challenges.

Confirmation of genome editing is often required after microinjection of CRISPR/Cas9 components into mammalian zygotes [1–4]. However, functional studies of proteins during embryogenesis associated with gene editing often require embryo fixation [5], thereby hindering the recovery of high-quality DNA suitable for DNA extraction and sequencing for genotype confirmation. We previously established a method for sequencing individual mammalian embryos following CRISPR/Cas9 zygotic microinjection [6]. Expanding upon this technique, our objective was to (i) develop a method to evaluate and recover fixed bovine blastocysts subjected to immunohistochemistry and (ii) extract DNA from fixed and imaged embryos suitable for DNA sequencing. Collectively, our methods represent a significant advancement towards integrating genome-editing strategies into the rapidly expanding field of livestock research [7].

Unless otherwise specified, reagents were purchased from Sigma-Aldrich, MI, USA. All washes and incubations are performed at room temperature on a rotator excluding primary antibody incubation steps. *In vitro*-produced bovine embryos were fixed in 4% paraformaldehyde (Cat. #28906; ThermoFisher Scientific, MA, USA) for 20 min and held in phosphate-buffered saline (PBS; Cat. #AM9625), covered with mineral oil (Cat. #W1503) at 5°C for up to 4 weeks. For immunohistochemistry, embryos were washed 3 × 10 min in washing buffer (WB; PBS + 0.1% Triton X-100) (Cat. #T8787) in 400 µl volumes using AgTech 6-well plates (Cat. #D18) [8]. Embryos were then permeabilized using PBS + 1% Triton X-100 for 30 min followed by a single 10-min wash (WB). Embryos were placed in blocking buffer (BB) WB + 1% BSA (Cat. #A8806) + 10% normal donkey serum compatible with secondary antibody (Cat. #100–151; GemBio, CA, USA) for 2 h followed by a 10-min wash. Overnight incubation with primary antibody (1:300 – independently optimized) was initiated on a rotator at 4°C in an antibody solution (AB) consisting of WB + 1% BSA, in 96-well plates sealed with parafilm. Embryos were then washed (WB) 3 × 10 min and then 3 × 20 min. Secondary antibody incubation in AB (1:500) was performed in 96 well plates for 1 h at room temperature and protected from light. After incubation, embryos were washed (WB) 3 × 10 min and then 3 × 20 min. Finally, DNA was stained with the addition of Hoechst 33258 (5 µg/ml; Cat. #14530) in PBS for 20 min followed by a 20-min wash (WB).

For imaging, the centers of 75 × 25 mm glass slides (Cat. # 2948; Corning, NY, USA) were affixed with a single binder reinforcement label (Cat. #05721; Avery®, CA, USA) and 9 µl of PBS was added to the center of the reinforcement label (Figure 1). A single embryo was placed in the PBS and a 22 × 22 mm coverslip (Cat. # 2845–22, Corning) was gently lowered. Following imaging using an epifluroescent microscope (confocal may increase image quality), embryos were recovered by gently sliding the coverslip until the reinforcement label was partially exposed and then by floating

METHOD SUMMARY

CRISPR/Cas9-microinjected bovine embryos can be fixed and stored for confirmation of induced mutations in single embryos by using both immunohistochemistry for protein analyses and DNA extraction for sequencing.

the coverslip with the addition of up to 60 µl of PBS (Figure 1). Embryos were removed with a pipette and placed in a 400-µl drop of PBS for holding.

DNA extraction of imaged embryos was achieved by placing single embryos in 10 µl of QuickExtract (Cat. #QE0905T; Lucigen, WI, USA) in an Eppendorf tube on ice. Samples were heated at 65°C for 6 min and 95°C for 2 min using a thermocycler and held at 4°C or stored at –20°C for further processing [6].

Two rounds of PCR were applied to templates. Primer concentration and PCR conditions were optimized for individual experiments. The first round of PCR was achieved in 20 µl reactions consisting of 10 µl of GoTaq Hot Start Green Master Mix (2X) (Cat. #M5123; Promega, WI, USA) with the addition of 0.4 µl (10 µM) of forward and reverse primer each and 9.2 µl of template. The second PCR reaction was identical to the first with the exception of 5 µl of PCR product produced from the first reaction, nested primers and 4.2 µl of nuclease-free water. PCR conditions were as follows: 95° for 3 min, followed by 35 cycles of 95° for 30 sec, 56° for 30 sec, 72° for 30 sec and a final extension of 72° for 7 min. Following amplification, 5 µl of PCR product was loaded onto a 0.9% agarose gel containing EtBr (0.5 µg/ml) and 0.5 X TBE running buffer (40 mM Trist-Cl, 45 mM boric acid, 1 mM EDTA) and run at 95V for 1 h with a ladder compatible to expected amplicon size. After confirming amplification of a single band (Figure 1), remaining PCR product was purified using QIAquick PCR Purification Kit (Cat. #28104; Qiagen, Hilden, Germany) following manufacturer specifications. Prior to the final elution step, a single modification was made by adding 10 µl of nuclease-free H_2O to the center of the column. DNA was quantified using a Nanodrop and approximately 100 ng of DNA with a 260/280 nm absorbance ratio of 1.8–1.9 was submitted for Sanger sequencing. Both DNA extraction and Sanger sequencing of fixed and imaged embryos proved highly efficient and repeatable (83 and 95%, respectively, Table 1).

Herein we report a useful technique for evaluating both protein expression and determining the mutation status of single, fixed bovine embryos following CRISPR/Cas9 microinjection. In addition, we adapted imaging techniques to include the use of a cost-effective and readily available binder reinforcement label as an alternative to traditional adhesive spacers. Furthermore, we

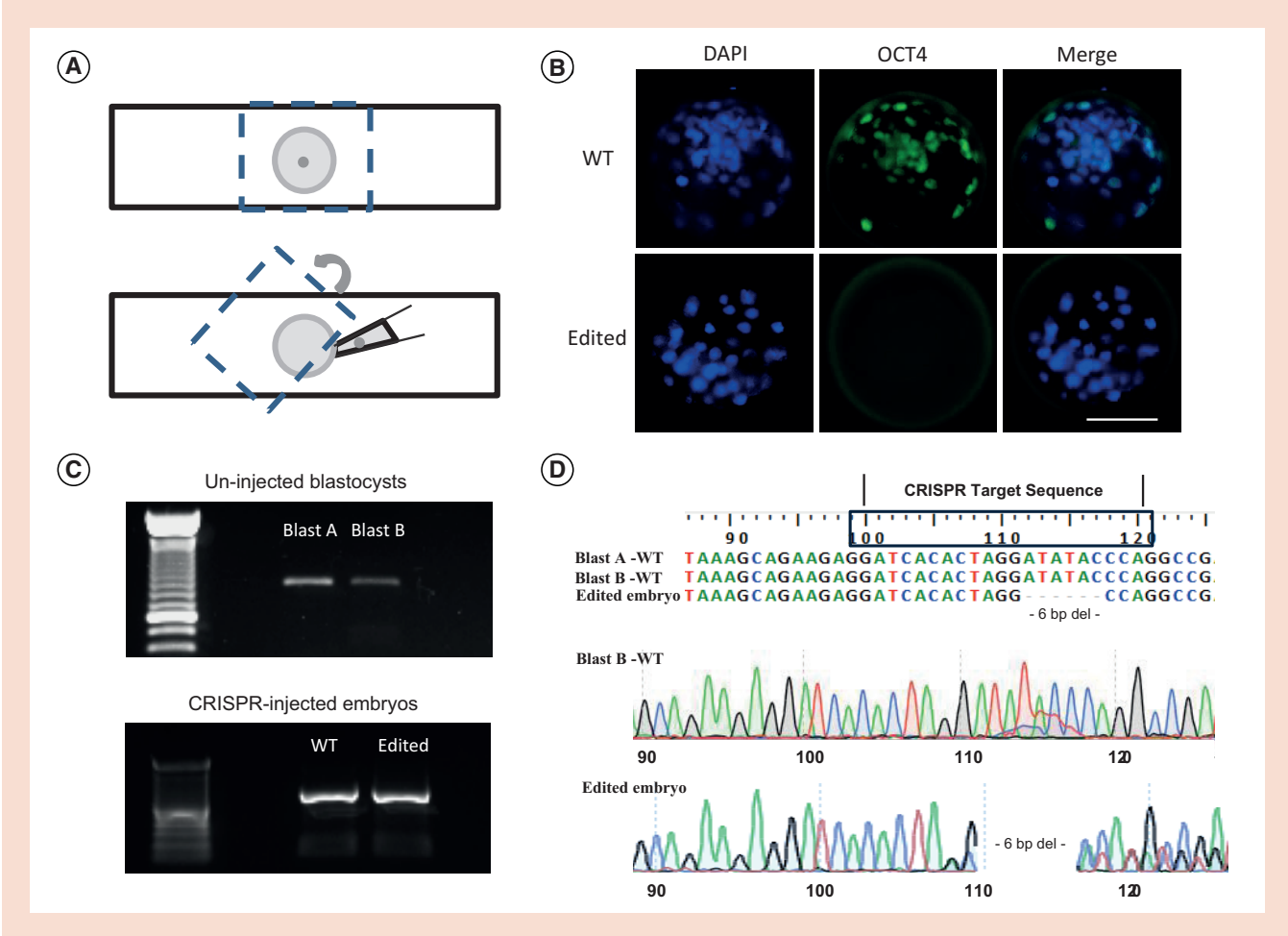


Figure 1. Placement and recovery of fixed bovine embryos for fluorescent imaging followed by DNA recovery for Sanger sequencing. (A) Placement of single embryos in the center of a reinforcement label with phosphate-buffered saline (PBS) on a glass slide with coverslip. Rotation of the coverslip and addition of PBS to recover imaged embryos by aspiration of remaining fluid. **(B)** Epifluorescent immunohistochemistry images of WT and edited, fixed bovine embryos stained with DAPI (blue) and anti-OCT4 (green) – (sc8628) and held in PBS on a glass slide with coverslip. **(C)** Nested PCR product from single bovine blastocysts and CRISPR-injected embryos previously fixed and then subjected to immunohistochemistry. **(D)** Target sequences and respective electopherograms of un-edited bovine blastocysts and edited CRISPR-injected embryos that were fixed, imaged and DNA extracted followed by PCR purification. Scale bar = 100 µm.
WT: Wild-type.

Table 1. DNA amplification and Sanger sequencing efficiency of fixed bovine embryos following immunohistochemistry imaging.

	Imaged	PCR-amplified	Sequenced
No. fixed embryos	23	19	18
Technique	DNA extraction (imaged/amplified)		Sequencing (amplified/sequenced)
Success rate	83%		95%

provide a valuable method to recover sufficient quantity and quality of DNA from single, fixed embryos compatible with sequencing. Collectively, these techniques allow for direct phenotypic and genotypic comparisons and avoid pooled samples with chimeric genotypes potentially confounding data analyses. In addition, the ability to store fixed embryos until further analyses increases experimental flexibility. Finally, these techniques complement the production of gene-edited livestock embryos by allowing for phenotype evaluation and confirmation of mutation status after microinjection, rather than by first transforming cell lines to confirm the desired genotype followed by nuclear transfer, which are both time consuming and technically challenging. In conclusion, we have optimized a method that allows for gene editing confirmation of fixed and stored single embryos by combining both immunohistochemistry and sequencing after CRISRP/Cas9 microinjection of mammalian embryos.

Author contributions

BD designed and performed experiments and prepared the manuscript. MV optimized PCR conditions. TF, SR and GS provided technical guidance and support. PR supervised the study and manuscript preparation.

Open access

Competing & financial interests disclosure

This research was supported by a USDA NIFA AFRI ELI Postdoctoral Fellowship 2016-67012-25254 to BWD and NIFA multistate research project W3171 to PJR. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1. Vilarino M, Rashid ST, Suchy FP *et al.* CRISPR/Cas9 microinjection in oocytes disables pancreas development in sheep. *Sci. Rep.* 7(1), 17472 (2017).
2. Kwon J, Namgoong S, Kim NH. CRISPR/Cas9 as tool for functional study of genes involved in preimplantation embryo development. *PLoS One* 10(3), e0120501 (2015).
3. Fogarty NME, McCarthy A, Snijders KE *et al.* Genome editing reveals a role for OCT4 in human embryogenesis. *Nature* 550(7674), 63–73, doi:10.1038/nature24033 (2017).
4. Bevacqua RJ, Fernandez-Martin R, Savy V *et al.* Efficient edition of the bovine PRNP prion gene in somatic cells and IVF embryos using the CRISPR/Cas9 system. *Theriogenology* 86(8), 1886–1896 e1881 (2016).
5. Muller HA. Immunolabeling of embryos. *Methods Mol. Biol.* 420, 207–218 (2008).
6. Wu J, Vilarino M, Suzuki K *et al.* CRISPR-Cas9 mediated one-step disabling of pancreatogenesis in pigs. *Sci. Rep.* 7(1), 10487 (2017).
7. Ruan J, Xu J, Chen-Tsai RY, Li K. Genome editing in livestock: are we ready for a revolution in animal breeding industry? *Transgenic Res.* 26(6), 715–726 (2017).
8. Chung N, Bogliotti YS, Ding W *et al.* Active H3K27me3 demethylation by KDM6B is required for normal development of bovine preimplantation embryos. *Epigenetics* 1–9 (2018).

First draft submitted: 13 April 2018; Accepted for publication: 9 August 2018

Address correspondence to: Pablo J Ross, Department of Animal Science, University of California – Davis, Davis, CA, USA; pross@ucdavis.edu

‡Authors contributed equally

Reviews

BioTechniques

For reprint orders, please contact: reprints@futuremedicine.com

Guidelines for optimized gene knockout using CRISPR/Cas9

Claude Van Campenhout[1], Pauline Cabochette[2], Anne-Clémence Veillard[1], Miklos Laczik[1], Agnieszka Zelisko-Schmidt[1], Céline Sabatel[1], Maxime Dhainaut[3], Benoit Vanhollebeke[2,4], Cyril Gueydan[‡,5] & Véronique Kruys[*,‡,5]

ABSTRACT

CRISPR/Cas9 technology has evolved as the most powerful approach to generate genetic models both for fundamental and preclinical research. Despite its apparent simplicity, the outcome of a genome-editing experiment can be substantially impacted by technical parameters and biological considerations. Here, we present guidelines and tools to optimize CRISPR/Cas9 genome-targeting efficiency and specificity. The nature of the target locus, the design of the single guide RNA and the choice of the delivery method should all be carefully considered prior to a genome-editing experiment. Different methods can also be used to detect off-target cleavages and decrease the risk of unwanted mutations. Together, these optimized tools and proper controls are essential to the assessment of CRISPR/Cas9 genome-editing experiments.

KEYWORDS:

gene targeting • methodology • quality control

*[1]Diagenode, SA, Liège Science Park, 4102 Seraing, Belgium; [2]Laboratoire de Signalisation Neurovasculaire, Faculté des Sciences, Université libre de Bruxelles (ULB), 12 rue des Profs. Jeener et Brachet, 6041 Gosselies, Belgium; [3]Precision Immunology Institute, Department of Genetics & Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [4]Walloon Excellence in Life Sciences & Biotechnology (WELBIO), Belgium; [5]Laboratoire de Biologie Moléculaire du Gène, Faculté des Sciences, Université libre de Bruxelles (ULB), 12 rue des Profs. Jeener et Brachet, 6041 Gosselies, Belgium; *Author for correspondence: vkruys@ulb.ac.be; [‡]These authors contributed equally*

BioTechniques: 00-00 (June 2019)
10.2144/btn-2018-0187

INTRODUCTION

Engineered nucleases, from zinc-finger nucleases to TALENs and CRISPRs, form a powerful class of genome-editing tools [1]. Among these, the CRISPR/Cas9 system has become the most popular, owing to its ease of use and rapidity. The CRISPR/Cas system was discovered in prokaryotes where it provides adaptive immunity against foreign elements [2]. In 2013, the CRISPR/Cas9 system from *Streptococcus pyogenes* (spCas9, further indicated in the text as Cas9) was successfully adapted for genome editing in eukaryotic cells [3]. Since then, the technique has become extremely popular as it can modify the genome of a large variety of organisms from animals to plants with unprecedented ease.

However, there is still limited predictability of whether the CRISPR system will be able to effectively target a given region of interest. This aspect is of particular importance in the context of CRISPR/cas9-based screens in model organisms and is related to the definition of the target site and the sequence of small guide RNA (sgRNA). Another major hurdle common to all engineered nucleases is the risk of unwanted mutations at sites other than the intended on-target site (off-target effects). The off-target mutations are the consequence of sgRNA binding to DNA sites with less than perfect complementarity [4]. Current strategies to increase targeting specificity notably include: refinements in guide RNA selection, enzyme and guide engineering, and improvements in the delivery method. Here, we describe a series of guidelines to optimize CRISPR/Cas9 efficiency and specificity.

ANALYSIS OF THE TARGET LOCUS

Careful determination of target sites is essential. For many applications, a loss of function may be desirable or even required. Targeting of functional protein domains was recently demonstrated to result in higher proportions of loss-of-function mutations [5]. A common strategy is to select sgRNAs that will target Cas9 nuclease to the N-terminal-coding exons of protein-coding genes. After the action of Cas9 nuclease, the introduction of indels by the error-prone non-homologous end joining repair of double-strand breaks (DSBs) introduces frame-shift mutations and subsequent premature stop codons, leading to mRNA elimination by nonsense-mediated mRNA decay. Genome-editing experiments to generate knockouts should be designed to disrupt exons that are shared by all transcript variants of a given gene. This strategy can also be applied to whole gene families using a sgRNA against exons that are conserved between all family members [6]. The CRISPys algorithm aims to design the optimal sequence to target multiple members of a gene family [7].

The high frequency with which CRISPR-induced mutations can be directed to target genes enables easy isolation of homozygous gene knockouts. Paradoxically, a potential caveat is found in this high efficiency. This holds particularly true in cell lines upon targeting genes essential for cell viability and fitness. In this regard, two distinct genome-wide CRISPR-Cas9-based screens have identified ≈2000 essential genes in the human genome [8]. More recently, Lenoir and colleagues published a database of pooled *in vitro* CRISPR knockout library essentiality screens that can be searched to identify genes that are essential across different human tissues [9].

Genetic screens in zebrafish and mouse have estimated that as many as 30% of genes are embryonic lethal [10,11]. The functional characterization of such essential genes requires the generation of heterozygous knockouts. The generation of hypomorphic alleles with the CRISPR system has been reported by different groups [12,13], but the method is not, at the moment, commonly used.

RNAi or CRISPRi [14] are efficient alternative loss-of-function methods and their effects can be directly evaluated at the transcriptome level. In addition, the development of inducible CRISPR tools provides a solution for genome editing with tight temporal control [15]. They additionally circumvent the mechanisms of genetic compensation that not unfrequently mask the phenotypes of knockout but not knockdown models [16].

Genetic polymorphism in the target region should be carefully assessed as it might have a profound influence on CRISPR/Cas9 efficacy. Although base mismatches (up to five) may be tolerated between the sgRNA and targeted sequences, the PAM and its proximal sequence have a stricter adherence to the consensus [17]. When a sgRNA is selected, the potential presence of a SNP in the PAM and the sgRNA-binding site should be verified as it can abolish Cas9 binding and cleavage. Of note, commonly used laboratory cell lines such as HeLa cells present a variant spectrum that slightly differs from the one found in the human population [18]. In general, sequences found in genomic databases may not exactly correspond to the DNA sequences of the model used for the genome-editing experiment. Sequencing of the target locus prior to sgRNA design will solve this potential pitfall. On the other hand, this PAM constraint can be exploited to target and disrupt heterozygous single-nucleotide mutations in certain dominant autosomal disorders, while leaving the wild-type allele intact [19]. Cell line ploidy is an additional consideration to take into account. Many common laboratory cancer cell lines carry four or more copies of a chromosome. Full knockouts would then require the introduction of mutations in all copies of the target gene. In practice, it is strongly advised to sequence the target loci to verify homozygous knockout when generating mutant clonal cell lines.

Besides the influence of the sequence features, chromatin states also strongly impact Cas9 binding and nuclease activity in vertebrates. Nucleosomes constitute fundamental units of chromatin and their positioning directly impedes Cas9 binding and cleavage *in vitro* and *in vivo*. Highly active sgRNAs for Cas9 are found almost exclusively in regions of low nucleosome occupancy [20]. Higher order chromatin structure (i.e., organization beyond the level of the linear array of nucleosomes) also influences Cas9 binding and enzymatic activity. Several authors showed that Cas9 cleavage efficiency positively correlates with open chromatin based on DNase I hypersensitivity. Along the same line, the activity of Cas9 can be significantly hindered by compact heterochromatin in cells [21]. Interestingly, the engineered Cas9 variants designed to improve specificity, Cas9-HF1 [22] and eSpCas9(1.1) [23], might be even more impacted than Cas9 by the chromatin-related factors [24]. While some gene editing applications have the option to select easy-to-cleave targets, such practice may not be feasible for gene corrections and other potential therapeutic applications. Many CRISPR genome-editing experiments focus on gene targeting and the study of the phenotypic consequences. In these applications, the gene of interest is usually transcriptionally active and the associated chromatin is relatively accessible to Cas9. Nevertheless, chromatin compactness can vary considerably between different genomic sites and from one cell type to another. Gene targeting in model organisms presents an additional challenge as the chromatin landscape is under constant change to ensure coordinated growth and differentiation during early development. Atlases of transcriptional activity (RNA-Seq) and of chromatin accessibility (e.g., ATAC-Seq and ChIP-Seq) are valuable information resources (of note see the ENCODE project: www.encodeproject.org) to predict sgRNA efficiency [25] and have been used to elaborate a predictive algorithm for zebrafish sgRNA selection taking into account chromatin accessibility [26]. Gene editing in mouse and human cells has been greatly facilitated by the publication of the genome-wide Brie and Brunello libraries [27]. These optimized sgRNA libraries respectively target the mouse and human genomes and provide 3–4 sgRNA sequences per gene with predicted high on-target efficiency and low off-target effects.

For more challenging applications such as the editing of heterochromatin-embedded sequences, chromatin manipulation might enhance the CRISPR targeting efficiency. While treatment with chromatin-disrupting drugs does not appear sufficient, transient overexpression of a targeted transcriptional activator might be an effective method to enhance Cas9 editing at closed chromatin regions [21].

DELIVERY METHODS

Introduction of the CRISPR/Cas9 components into cultured cells is often achieved by DNA-based delivery systems such as transfection of plasmids encoding nuclear targeted Cas9 and sgRNA. Transduction with viral particles is also commonly used and is typically more efficient compared with plasmid transfection and is applicable to many cell types including primary cells. Plasmid transfection and viral transduction methods lead to a prolonged or a permanent expression of Cas9, respectively. Extended expression of Cas9 in cells can lead to accumulation of off-targeting events [28]. Indeed, constitutive expression of lentiviral-based Cas9 and sgRNAs leads to an enrichment of predicted off-target sites over time. Reducing the concentration of delivered plasmid during transfection was shown to decrease off-targeting [4]. These data support the idea that controlling the expression of Cas9 and the sgRNAs in order to limit the time of action can reduce genome-wide off-targeting. A doxycycline-inducible promoter allows for transient Cas9 expression and is compatible with lentiviral delivery of the nuclease [29]. Because gene editing results in a permanent change in the genome, CRISPR-mediated editing can be achieved using Cas9 protein/sgRNA ribonucleoprotein (RNP) complexes. Both sgRNAs and nuclear-targeted Cas9 protein can be produced in-house [30] or obtained from different commercial suppliers. RNP complexes can be delivered by a variety of techniques such as lipid-mediated transfection [30], electroporation [28], induced transduction by osmocytosis and propanebetaine (iTOP) [31], microinjection [32] or cell-penetrating peptide-mediated delivery [33]. Uncoupling administration of the sgRNA and Cas9 protein (e.g., in the context of genome-scale screens) can lead to successful gene editing in human primary cells [34] and appears to be more efficient upon delivery of Cas9 protein complexed with a nontargeting gRNA [35]. Finally, biolistic transfer of Cas9/sgRNA RNP complexes or of Cas9- and sgRNA-encoding plasmids appears to be an attractive alternative for cells resistant to other delivery methods, such as plant cells [36].

The sgRNA-Cas9 RNPs were shown to cleave the target chromosomal DNA between 12 and 24 h after delivery and the frequency of gene editing reaches a plateau after 1 day. For plasmid expression of Cas9 and sgRNA, equivalent gene editing levels were only achieved at 3 days after delivery. Furthermore, the Cas9 protein has been shown to be degraded rapidly in cells, within 24–48 h after delivery, compared with several days when continuously expressed from a plasmid [28]. Moreover, the ratio of the indel frequency at the on-target site to off-target sites strongly increases when RNPs are transfected in comparison with plasmid [33]. While off-target effects may be less of a concern in screening applications since any identified 'hits' will be confirmed through follow-up experiments, constitutive expression or high stability of Cas9 nuclease and/or sgRNA may be undesirable for many applications, such as generation of clonal cell lines for a phenotypical study of a specific gene knockout. In addition to the increased potential for off-target effects due to prolonged or constitutive expression of components of the CRISPR-Cas9 system, unwanted incorporation of the plasmid DNA into the cell genome is not uncommon. When the DNA repair pathways are activated after Cas9-mediated double-stand breaks, the risk of foreign DNA integration is increased. The absence of transgene eliminates the risk of unintended DNA integration.

The delivery of RNP complexes also has the major advantage of being easily applicable to a wide range of model organisms and cell types. In *in vivo* contexts, the functionality of the Cas9–sgRNA RNP complexes has been reported as being superior to other delivery methods. In zebrafish, mutagenesis can be performed through microinjection of Cas9-encoding mRNA or of Cas9 protein together with sgRNA into fertilized embryos. Contrary to Cas9-encoding mRNA, RNPs are immediately active upon microinjection and are generally more effective in the fast dividing blastomeres of the zebrafish and Xenopus embryos (Figure 1) [37,38], in which mutagenesis occurring after the first cleavages leads to increased mosaicism. In plants, this DNA-free approach allows generation of marker-free 'cisgenic' variants that could be exempted from current GMO regulations and thereby supersede all the technologies of gene editing in plants [39]. The fact that sgRNAs can be easily synthetized *in vitro* makes it possible to use multiple sgRNAs simultaneously to achieve multigenic targeting [40]. The DNA-free system also suppresses the variability that can arise from the choice of promoter used to drive expression from vector-based CRISPR-Cas9 systems. It is well known that not all promoters are functional in every cell type or cell line, so delivery of Cas9 protein or Cas9 mRNA avoids incompatibilities of certain promoters in specific cells. Codon usage patterns also vary between species and Cas9 derived from DNA or mRNA expression may not yield the expected result as every organism has its own codon bias. Optimization of codon usage is a routine process but can be relatively time-consuming. Codon optimization becomes unnecessary when using Cas9 protein instead of a DNA- or mRNA-based delivery method. Independent of the delivery mode, specific anti-Cas9 antibodies can be used to measure Cas9 expression level and to confirm Cas9 presence in the nucleus (Figure 2).

Lentiviral or plasmid delivery of Cas9 and sgRNA often utilizes a selection gene encoding either a drug-selectable marker (hygromycin, blasticidin or puromycin) or a reporter protein (GFP and NGFR) to isolate cells that are successfully transduced or transfected. When RNPs are transfected or electroporated, alternative strategies can be used, such as surrogate reporters [44]. However, these methods are inefficient for assessing sgRNA efficiency at a large scale because it is both time- and labour-consuming to construct a specific reporter for each individual sgRNA. To avoid specific cloning, the transfection efficiency can also be indirectly evaluated with a dTomato reporter assay [31]. Fluorescent versions of Cas9 such as Cas9–GFP [45] or Cas9–Cy3 [46] can also be used to sort RNP-transfected cells. These latter methods focus on the physical separation of edited cells from unedited cells. An important aspect to consider is that CRISPR experiments lead to a genetic heterogeneity due to the random nature of DNA repair by the NHEJ pathway. As this genetic heterogeneity could yield phenotypic heterogeneity, monoclonal populations should be isolated prior to phenotypic analysis. The first step is to determine the editing efficiency of the entire cell population. This information can indicate how many individual clones should be isolated and checked for editing. If limited dilutions are used to isolate individual cells, it should be realized as soon as possible after termination of the edition process, as nonedited cells could potentially outgrow edited cells.

Gene editing in *in vivo* mouse models was greatly facilitated by the generation of a knock-in (KI) transgenic mouse in which a Cre-inducible Cas9-P2A-GFP cassette was inserted in the Rosa26 locus [47]. Cre-mediated recombination leads to cell- or tissue-specific Cas9 expression, as evidenced by GFP expression. Apart from allowing for gene editing following *in vivo* delivery of sgRNAs, this model can also be used to efficiently edit the genome of primary cells *ex vivo*.

GUIDE RNA EFFICIENCY & SPECIFICITY

The performance of sgRNAs targeting the same gene can vary dramatically. This was recently highlighted in a novel approach to CRISPR genomics where expression of sgRNAs was coupled with specific protein barcodes, allowing for simultaneous multidimensional phenotypic analysis of several dozen knockouts at a single-cell resolution [48]. In a pooled parallel analysis of gene editing efficiency for ten genes (3–4 sgRNAs per gene), the authors demonstrated that the gene knockout at the protein level was highly variable depending on the sgRNAs used. There are many bioinformatic tools available for sgRNA design and some of these tools also apply filters or show 'scores' related to predicted effectiveness. sgRNAs with potential for weak secondary structures are likely to be more efficient than alternatives with strong secondary structures [49]. Nevertheless, no computational tool can guarantee the efficacy of a sgRNA and, when possible, several sgRNAs should be tested. Endonuclease cleavage assays can be used to characterize the *in vitro* efficacy of a particular sgRNA. Experimental validation of sgRNAs before practical application is particularly important to minimize wasted experiments on sgRNAs with poor activity. In these *in vitro* assays, the target DNA site, including its PAM motif, is either inserted into a plasmid or provided in the form of a PCR product. The Cas9 recombinant ▶

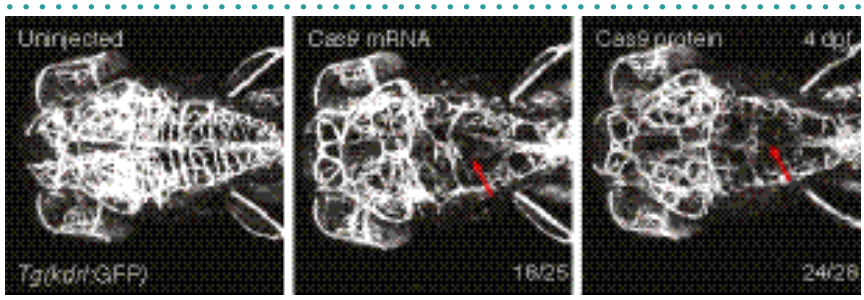


Figure 1. Cas9 Nuclease Protein NLS and Cas9 mRNA injections achieve highly efficient bi-allelic somatic gene disruptions in zebrafish. Tg(kdrl:GFP)^{s843} embryos were injected at the one-cell stage with 300 pg of Cas9 protein (Cas9 Nuclease Protein NLS from Diagenode, cat n° C29010001) or 150 pg of cas9 mRNA *in vitro* transcribed from the XbaI linearized pT3TS-nls-zcas9-nls vector (Addgene #46757) and 30 pg of two sgRNAs targeting reck, an essential regulator of CNS vascularization [41]. Right panel: maximal intensity projections of confocal z-stacks of the cranial vasculature of Tg(kdrl:GFP) larvae at 4 dpf. Red arrows point to the avascular hindbrains. The fraction of injected larvae showing cerebrovascular defects is indicated (18/25 and 24/28).
dpf: Days post-fertilization.

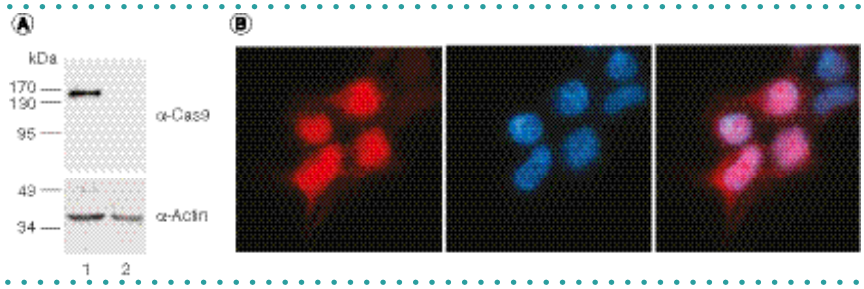


Figure 2. Cas9 expression can be detected by western blot and immunofluorescence. (A) Western blot analysis of HEK293T cells transfected with pSpCas9(BB)-2A-Puro (PX459) V2.0 plasmid (Addgene #62988) [42] encoding Cas9 using Lipofectamine 2000 (Life Technologies). Cell lysates from transfected (lane 1) or untransfected cells (lane 2) were analyzed by immunoblotting with mouse anti-Cas9 (Diagenode Cat No. C15200216). (B) HEK293T cells were transfected with dCas9 plasmid (Addgene #100091) [43]. dCas9 localization was analyzed by immunostaining using anti-Cas9 mouse monoclonal antibody (Diagenode Cat No. C15200216) and Alexa-Fluor 594-conjugated donkey anti-mouse antibody. Nuclei were counterstained with DAPI.

protein and the sgRNA are pre-incubated in a 1:1 molar ratio in the cleavage buffer to reconstitute the Cas9–sgRNA complex prior to the addition of target DNA. Cleavage of plasmid or PCR substrates is monitored by agarose gel electrophoresis with an intercalating dye (Figure 3). The reaction rate can strongly vary in function of DNA source and length (PCR product versus plasmid, circular plasmid versus linear plasmid), optimal enzyme and substrate concentrations, and also reaction time points need to be determined empirically [50]. This *in vitro* test validates sgRNA intrinsic capacity to form cleavage-competent complexes; however, it does not guarantee *in vivo* effectiveness, which also greatly depends on chromatin accessibility, as previously mentioned.

The targeting specificity of Cas9 is believed to be tightly controlled by the 20-nucleotide guide sequence of the sgRNA and the presence of a PAM adjacent to the target sequence in the genome. Nevertheless, potential off-target cleavage activity can still occur on DNA sequence with even 3–5 basepair mismatches in the PAM-distal part of the sgRNA-guiding sequence [4]. Of note, shortening of the sgRNA guide sequence to 17 nucleotides was shown to improve target specificity [51]. Numerous online tools are available to assist in sgRNA design but the correlation between the predictions and the actual measurements varies considerably since sequence homology alone is not fully predictive of off-target sites [52]. These tools also suggest probable off-target sites but the appropriate number of potential sites to experimentally assay remains unclear. Moreover, there are still contradictory conclusions as to the prevalence of off-target effects, from low [53] to high levels of off-targeting [54].

Cleavage at on- and off-target sites can be assessed using various methods, which include mismatch-sensitive enzymes (Surveyor or T7 Endonuclease I assay), restriction fragment length polymorphism (RFLP) analysis, high-resolution melting curve analysis (HRMA) or PCR amplification of the locus of interest followed by sequencing. Surveyor and T7 Endonuclease I specifically cleave heteroduplex DNA mismatch. The T7 Endonuclease I assay outperforms the Surveyor nuclease in terms of sensitivity with deletion substrates, whereas Surveyor is better for detecting single nucleotide changes. The limit of sensitivity for the T7 Endonuclease I assay is around 5% [55]. HRMA utilizes the difference in melting curve of the heteroduplex and mutant homoduplex. A recent report demonstrates that techniques such as targeted NGS, tracking indels by decomposition (TIDE) and indel detection by amplicon analysis (IDAA) outperform nuclease-based methods to detect Cas9-mediated edition in pools of cells [56]. Ultimately, Sanger sequencing of DNA from individual clones is the gold standard for confirming the presence of indels at on-target site but is not easily applicable to off-target detection. Overall, these indels detection methods are relatively straightforward but are low throughput and interrogate one locus at a time.

Unbiased off-target analysis requires the detection of mutations generated in the target cells by the CRISPR/Cas9 system outside their target locus. In theory, WGS of cells before and after editing could be used to study CRISPR/Cas9 specificity. In a clonal population, off-target sites can be determined by the analysis of the new mutations that have been generated outside the intended locus. However, WGS faces its own challenges and might not be easily applicable to the detection of off-target mutations. While sequencing costs continue to drop, a certain degree of bioinformatic expertise is necessary to detect small indels and separate signal from noise. In fact, many spontaneous new mutations may appear during clonal

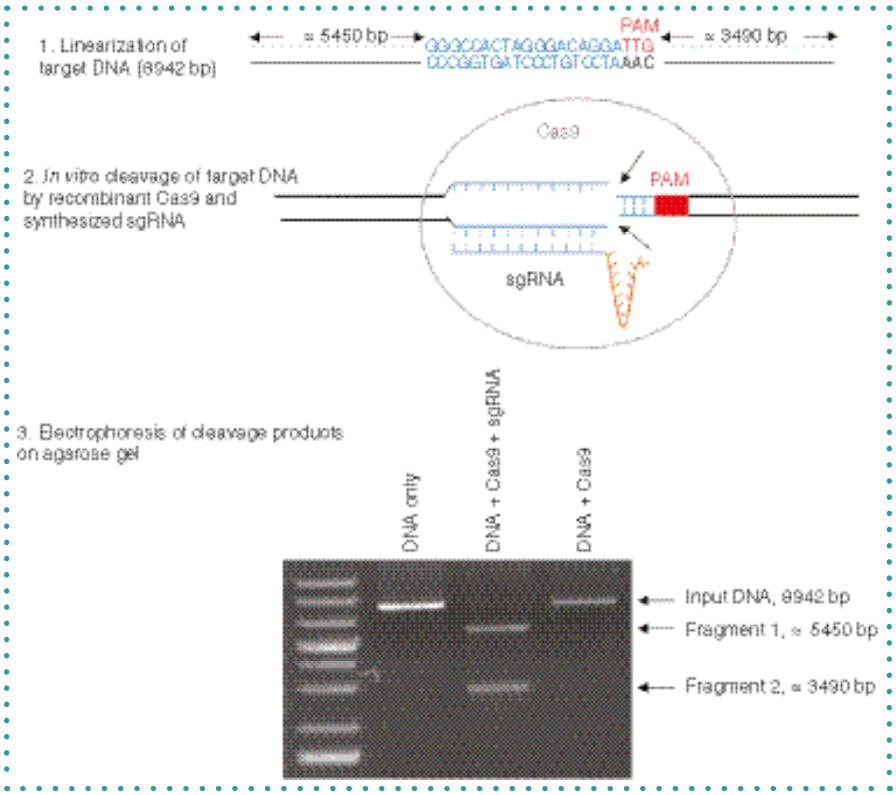


Figure 3. ***In vitro*** **validation of sgRNA by Cas9 cleavage assay. Experimental design of the assay. DNA target is PvuI-linearized CRISPR-SP-Cas9 reporter plasmid (Addgene #62733) [31]. Corresponding sgRNA (GGGCCACUAGGGACAGGAU) was synthesized by** ***in vitro*** **transcription. Target DNA was incubated with (lanes 2 and 3) or without (lane 1) Cas9 recombinant protein (Cas9 Nuclease protein NLS from Diagenode, Cat No. C29010001). Reactions were set-up with a ratio of 20:20:1 (Cas9:sgRNA:DNA target) and incubated for 1 h at 37°C. The products were resolved on 0.8% TAE agarose gel stained with ethidium bromide.**

expansion and it might not be possible to distinguish them from off-target effects. WGS of individual induced pluripotent stem cell clones reveals a large number of indels in the genome that are not the result of Cas9 activity, but rather a consequence of clonal variation or technical artefacts [57]. To circumvent these limitations, several methods have recently been developed to measure Cas9 off-target activity across the genome such as BLESS (labeling of DSBs followed by enrichment and sequencing), HTGTS (high-throughput genome-wide translocation sequencing), GUIDE-Seq (genome-wide unbiased identification of DSBs enabled by sequencing), Digenome-Seq (*in vitro* Cas9-digested WGS) [58], IDLV (detection of off-targets using integrase-deficient lentiviral vectors) [59] and most recently, SITE-Seq (a biochemical method that identifies DNA cut sites) [60] and CIRCLE-Seq (an *in vitro* method for identifying off-target mutations) [61]. Overall, these unbiased methods tend to be less sensitive and have a lower throughput than biased targeted sequencing, in addition to typically requiring higher sequencing coverage and much more complex protocols. These techniques also require manipulation of the genome and might be difficult to apply on some samples (e.g., primary cells, or *in vivo*).

Chromatin immunoprecipitation followed by next-generation sequencing (ChIP-Seq) is a technique of choice for studying protein–DNA interactions. ChIP has been used to pull down the Cas9 nuclease protein together with the DNA fragments to which the nuclease was bound. The immunoprecipitation of Cas9 bound to the genome is technically challenging due to the nuclease activity of Cas9. However, the introduction of two amino-acid changes (D10A and H840A) in Cas9-coding sequence results in a nuclease-inactive DNA-binding protein named 'dead Cas9' (dCas9). Specific enrichment of dCas9 at on-target regions can be evaluated by ChIP-qPCR using ChIP-grade Cas9 antibodies (Figure 4A). Moreover, this approach can be extended to the unbiased analysis of off-target sites by ChIP-Seq (Figure 4B). dCas9-based ChIP-PCR/Seq is thus a powerful approach to score several sgRNAs at once thanks to its rapidity, reduced sequencing cost and high coverage. Moreover, it is of predictive value for sgRNA performance upon association with catalytically active Cas9, although Cas9 DNA-binding and cleavage activities are sometimes uncoupled [62]. As no single method guarantees a complete coverage of off-target sites, multiple approaches should ideally be combined. Therefore, sequence-based *in silico* prediction combined with genome-wide ChIP-Seq dCas9-binding analysis can efficiently identify off-target sites.

Variants of dCas9 have recently been generated that allow repurposing of the system to a variety of applications [64]. Fusing dCas9 to various transcriptional activating or repressing modules proved to be a potent way of regulating gene expression. Moreover, dCas9 can be fused to domains that regulate the epigenetic landscape at endogenous loci. It can also be used to label endogenous loci for live visualization [65] or to edit a single base in the genome [66]. In those applications, the binding specificity of dCas9 fused to various effectors could be tested by dCas9 ChIP-Seq as we describe here.

FUTURE PERSPECTIVE

From the first description of Cas9 derived from *S. pyogenes* for gene editing in 2013, incredible progress has been made to optimize and adapt its use in a wide range of applications. Structural studies of Cas9 led to the generation of several variants such as enhanced specificity Cas9 (eSpcas9), high-fidelity Cas9 (Cas9-HF1) and hyper-accurate Cas9 (HypaCas9), which display increased specificity due to reduced DNA-binding affinity (eSpCas9 and Cas9-HF1) [22,23] or locking of the nuclease domain upon guide/target mismatches (HypaCas9) [67]. In addition, Cas9 nickases (Cas9n) were developed by inactivating the cleavage activity on target or nontarget DNA and have been demonstrated to nick only one DNA strand instead of generating a DSB. DSBs are generated only upon recruitment of a Cas9n pair with two sgRNAs that target opposite strands in close proximity [4,68], ▸

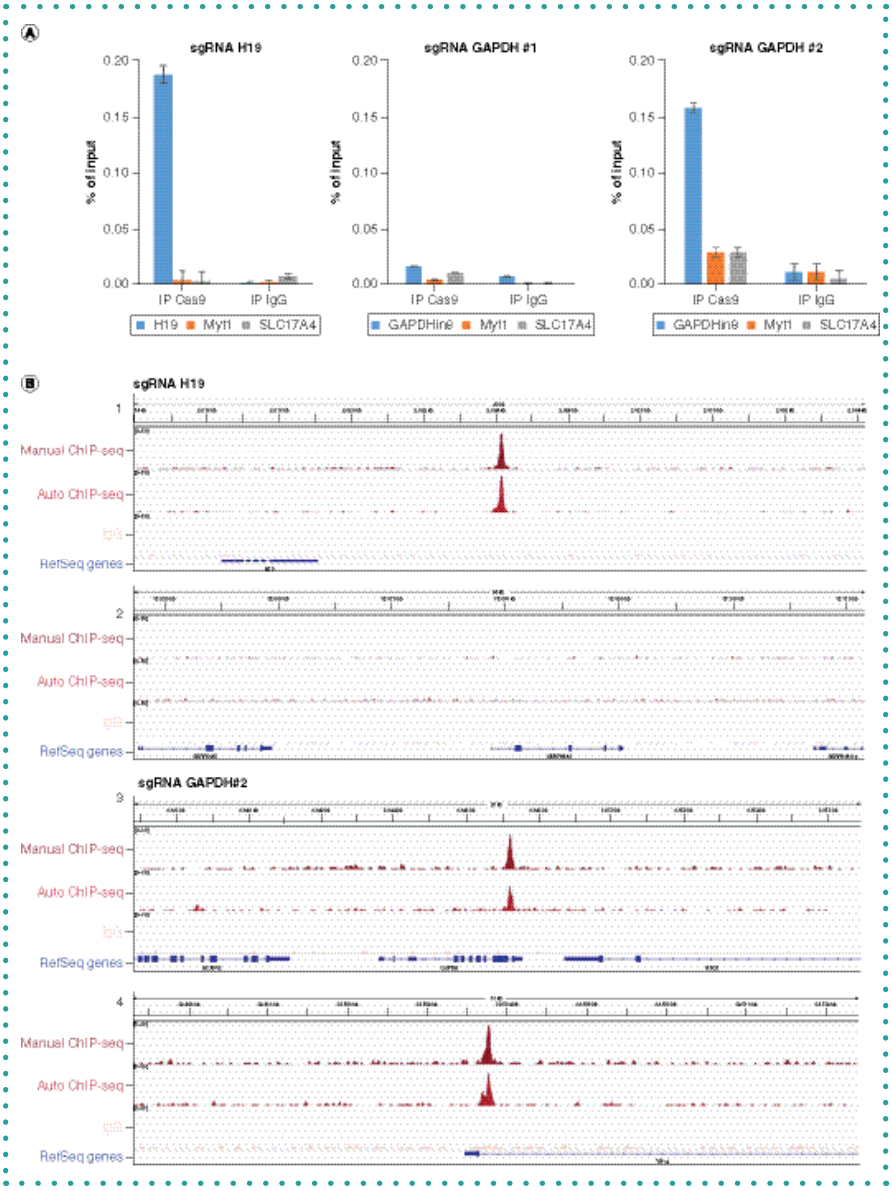


Figure 4. Evaluation of sgRNA specificity by dCas9 ChIP-PCR and ChIP-Seq. Sequences of sgRNA targeting the human H19 (GTCTATCTCTGACAACCCTC) or GAPDH sgRNA#1 GTCTGGCGCCCTCTGGTGGC and sgRNA#2 AAAGACTCGGTCGGTGGTCT) loci were designed using an online selection tool (http://crispr.mit.edu/) and cloned in the pLenti-dCas9-2xAM plasmid (Addgene 92220) [63]. ChIP was performed on sheared chromatin from HEK293T cells using the iDeal ChIP-seq Kit for TF (Diagenode Cat No. C01010170), 5 µl of the polyclonal Cas9 antibody (Diagenode Cat No. C15310258) and 5 µl of control IgG (Diagenode Cat No. C15410206). (A) dCas9 ChIP-PCR analysis of one representative out of three independent experiments is shown (error bars show SD of replicates). Primers specific for the human H19, GAPDH, Myt1 and SLC17A4 were used for the qPCR. The figure shows the recovery, expressed as % of input. (B) dCas9 ChIP-Seq analysis. ChIP was performed manually or with the IP-Star® Compact Automated System (Diagenode Cat. No. B03000002) as described in the manual. Libraries were prepared using the kit MicroPlex (Diagenode Cat. No. C05010012) and sequenced on HiSeq 3000 using SE 50pb reads. The figure shows the read distribution for the manual IP (top), the automated IP (middle) and IgG (bottom) samples. Peaks show distribution of the three datasets in the region surrounding H19 (1) and a representative region of the genome (2) for HEK293T cells expressing a sgRNA for H19. Peaks show distribution of the three datasets in the region surrounding GAPDH (3) and YIPF4 (4) for HEK293T cells expressing a sgRNA targeting GAPDH (sgRNA GAPDH #2).

thereby increasing specificity by double selection. A similar strategy was used to develop a fusion of dCas9 with the catalytic domain of FokI nuclease (fCas9), which induces DSBs only upon dimerization of the FokI domains by sgRNA pairing to complementary strands. With the same aim of reducing Cas9 off-target activity, several Cas9 variants whose editing activity can be irreversibly or reversibly programmed are now also available. Finally, engineered Cas9 variants with novel PAM specificities enlarge the edition spectrum to previously inaccessible sites [69,70].

Several limitations have yet to be addressed to promote Cas9 use in gene therapy. First, the source of Cas9 nucleases, that is, *S. pyogenes* and *Staphylococcus aureus*, are common human pathogens. Recent reports have highlighted pre-existing immunity towards both SpCas9 and SaCas9 in the human population, with a high prevalence of both Cas9-reactive T cells and antibodies [71]. Although it is still unclear whether adeno-associated virus delivery of Cas9 leads to the immune rejection of transduced cells *in vivo*, strategies to control the anti-Cas9 T cell responses, such as transient immunosuppression or engineering Cas9 proteins with mutated T cell epitopes, are being considered [72]. Another limitation of Cas9 for its use in gene therapy resides in its rather large size, which is incompatible for efficient packaging into adeno-associated virus vectors, the most commonly used delivery systems in gene therapy. Although this hurdle can be overcome by the separation of the recognition lobe from the nuclease lobe into two separate vectors, the emergence of Cas9 orthologs of smaller size might provide more efficient alternatives [69]. Beside solving the delivery problem, CRISPR effectors from other bacterial and archeal species offer different substrate specificities or operate according to different mechanisms. This is notably the case of Cas12a (Cpf1), which is structurally different from Cas9, has no requirement for tracer RNA, recognizes a T-rich (TTTN) PAM sequence lying 5′ of the target sequence, and uses a different mechanism for target recognition and cleavage [69]. Cpf1 also possesses the ability to cleave RNA and generate multiple crRNAs from a single pre-crRNA array. This capacity has been harnessed to achieve multiplex gene editing using a single pre-crRNA array, which can both increase knockout efficiency (when using multiple crRNAs targeting the same locus) or easily knockout multiple genes with a single construct [73]. Moreover, gene

editing by Cpf1 results in lower off-target effects than Cas9, as evidenced by genome-wide analysis of edited cells [74]. Finally, the discovery of Cas13a and CasRx as RNA-guided nucleases targeting RNA paves the way to new therapeutic approaches based on RNA editing [69,75].

While several solutions and guidelines to harness CRISPR-Cas9-based gene targeting have been provided, it is expected that therapeutic, industrial and research applications will still place high demand on improving the specificity and efficiency of the CRISPR/Cas9 system. As CRISPR-based gene targeting technology continues to become more sophisticated and diverse, optimized procedures and quality control guidelines should be established.

AUTHOR CONTRIBUTIONS

All the authors contributed to the content of the manuscript; CVC, MD, BV, CG and VK wrote the manuscript. CG and VK share senior authorship.

ACKNOWLEDGMENTS

We thank Romuald Soin and Nadège Delacourt for technical support.

FINANCIAL & COMPETING INTERESTS DISCLOSURE

CVC was funded by a FIRST Entreprise grant from DGO6 from the Walloon Region, Belgium. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

REFERENCES

1. Gaj T, Gersbach CA, Barbas III CF. ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31(7), 397–405 (2013).
2. Barrangou R, Fremaux C, Deveau H *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712 (2007).
3. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346(6213), 1258096 (2014).
4. Hsu PD, Scott DA, Weinstein JA *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31(9), 827–832 (2013).
5. Shi J, Wang E, Milazzo JP *et al.* Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* 33, 661–667 (2015).
6. Endo M, Mikami M, Toki S. Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell. Physiol.* 56(1), 41–47 (2015).
7. Hyams G, Abadi S, Lahav S *et al.* CRISPys: optimal sgRNA design for editing multiple members of a gene family using the CRISPR system. *J. Mol. Biol.* 430(15), 2184–2195 (2018).
8. Wang T, Birsoy K, Hughes NW *et al.* Identification and characterization of essential genes in the human genome. *Science* 350(6264), 1096–1101 (2015).
9. Lenoir WF, Lim TL, Hart T. PICKLES: the database of pooled ***in vitro*** CRISPR knockout library essentiality screens. *Nucleic Acids Res.* 46(D1), D776–D780 (2018).
10. Haffter P, Granato M, Brand M *et al.* The identification of genes with unique and essential functions in the development of the zebrafish, ***Danio rerio***. *Development* 123, 1–36 (1996).
11. Ayadi A, Birling MC, Bottomley J *et al.* Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project. *Mamm. Genome* 23(9–10), 600–610 (2012).
12. Challa AK, Boitet ER, Turner AN *et al.* Novel hypomorphic alleles of the mouse tyrosinase gene induced by CRISPR-Cas9 nucleases cause non-albino pigmentation phenotypes. *PLoS One* 25, 11(5), e0155812 (2016).
13. Goto T, Hara H, Nakauchi H, Hochi S, Hirabayashi M. Hypomorphic phenotype of ***Foxn1*** gene-modified rats by CRISPR/Cas9 system. *Transgenic Res.* 25, 533–544 (2016).
14. Qi LS, Larson MH, Gilbert LA *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152(5), 1173–1183 (2013).
15. Cao J, Wu L, Zhang M *et al.* An easy and efficient inducible CRISPR/Cas9 platform with improved specificity for multiple gene targeting. *Nucleic Acid Res.* 44(19), e149 (2016).
16. El-Brolosy MA, Stainier DYR. Genetic compensation: a phenomenon in search of mechanisms. *PLoS Genet.* 13(7), e1006780 (2017).
17. Zheng T, Hou Y, Zhang P *et al.* Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. *Sci. Rep.* 7, 40638 (2017).
18. Landry JJ, Pyl PT, Rausch T *et al.* The genomic and transcriptomic landscape of a HeLa cell line. *G3 (Bethesda)* 3(8), 1213–1224 (2013).
19. Li Y, Mendiratta S, Ehrhardt K *et al.* Exploiting the CRISPR/Cas9 PAM constraint for single-nucleotide resolution interventions. *PLoS One* 11(1), e0144970 (2016).
20. Horlbeck MA, Witkowsky LB, Guglielmi B *et al.* Nucleosomes impede Cas9 access to DNA ***in vivo*** and ***in vitro***. *Elife* 5, e12677 (2016).
21. Daer RM, Cutts JP, Brafman DA, Haynes KA. the impact of chromatin dynamics on Cas9-mediated genome editing in human cells. *ACS Synth Biol.* 6(3), 428–438 (2017).
22. Kleinstiver BP, Pattanayak V, Prew MS *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529(7587), 490–495 (2016).
23. Slaymaker IM, Gao L, Zetsche B *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* 351(6268), 84–88 (2016).
24. Chen X, Liu J, Janssen JM, Gonçalves MAFV. The chromatin structure differentially impacts high-specificity CRISPR-Cas9 nuclease strategies. *Mol. Ther. Nucleic Acids* 8, 558–563 (2017).
25. Uusi-Mäkelä MIE, Barker HR, Bäuerlein CA *et al.* Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (***Danio rerio***). *PLoS One* 13(4), e0196238 (2018).
26. Chen Y, Zeng S, Hu R *et al.* Using local chromatin structure to improve CRISPR/Cas9 efficiency in zebrafish. *PLoS One* 12(8), e0182528 (2017).
27. Doench JG, Fusi N, Sullender M *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34(2), 184–191 (2016).
28. Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24(6), 1012–1019 (2014).
29. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* 343(6166), 80–84 (2014).
30. Zuris JA, Thompson DB, Shu Y *et al.* Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing ***in vitro*** and ***in vivo***. *Nat. Biotechnol.* 33(1), 73–80 (2015).
31. D'Astolfo DS, Pagliero RJ, Pras A *et al.* Efficient intracellular delivery of native proteins. *Cell* 161(3), 674–690 (2015).
32. Gagnon JA, Valen E, Thyme SB *et al.* Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* 9(5), e98186 (2014).
33. Ramakrishna S, Kwaku Dad AB, Beloor J *et al.* Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res.* 24(6), 1020–1027 (2014).
34. Shifrut E, Carnevale J, Tobin V *et al.* Genome-wide CRISPR screens in primary human T cells reveal key regulators of immune function. *Cell* 175(7), 1958–1971 (2018).
35. Ting PY, Parker AE, Lee JS *et al.* Guide Swap enables genome-scale pooled CRISPR-Cas9 screening in human primary cells. *Nat. Methods.* 15(11), 941–946 (2018).
36. Hamada H, Liu Y, Nagira Y *et al.* Biolistic-delivery-based transient CRISPR/Cas9 expression enables in planta genome editing in wheat. *Sci. Rep.* 8(1), 14422 (2018).
37. Burger A, Lindsay H, Felker A *et al.* Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* 143(11), 2025–37 (2016).
38. Aslan Y, Tadjuidje E, Zorn AM, Cha SW. High-efficiency non-mosaic CRISPR-mediated knock-in and indel mutation in F0 ***Xenopus***. *Development* 144, 2852–2858 (2017).
39. Soda N, Verma L, Giri J. CRISPR-Cas9 based plant genome editing: significance, opportunities and recent advances. *Plant Physiol. Biochem.* 131, 2–11 (2018).
40. Song J, Yang D, Ruan J *et al.* Production of immunodeficient rabbits by multiplex embryo transfer and multiplex gene targeting. *Sci. Rep.* 7(1), 12202 (2017).
41. Eubelen M, Bostaille N, Cabochette P *et al.* A molecular mechanism for Wnt ligand-specific signaling. *Science* 361(6403), eaat1178 (2018).
42. Ran FA, Hsu PD, Wright J *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8(11), 2281–308 (2013).
43. O'Geen H, Ren C, Nicolet CM *et al.* dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Res.* 45(17), 9901–9916 (2017).
44. He Z, Shi X, Liu M *et al.* Comparison of surrogate reporter systems for enrichment of cells with mutations induced by genome editors. *J. Biotechnol.* 221, 49–54 (2016).
45. Mircetic J, Steinebrunner I, Ding L *et al.* Purified Cas9 fusion proteins for advanced genome manipulation. *Small Methods* 1(4),, 1600052 (2017).
46. Kim K, Park SW, Kim JH *et al.* Genome surgery using Cas9 ribonucleoproteins for the treatment of age-related macular degeneration. *Genome Res.* 27(3), 419–426 (2017).
47. Platt RJ, Chen S, Zhou Y *et al.* CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 159(2), 440–455 (2014).
48. Wroblewska A, Dhainaut M, Ben-Zvi B *et al.* Protein barcodes enable high-dimensional single-cell CRISPR screens. *Cell* 175(4), 1141–1155.e16 (2018).
49. Thyme SB, Akhmetova L, Montague TG, Valen E, Schier AF. Internal guide RNA interactions interfere with Cas9-mediated cleavage. *Nat. Commun.* 7, 11750 (2016).
50. Anders C, Jinek M. ***In vitro*** enzymology of Cas9. *Methods Enzymol.* 546, 1–20 (2014).
51. Zhang JP, Li XL, Neises A *et al.* Different effects of sgRNA length on CRISPR-mediated gene knockout efficiency. *Sci. Rep.* 6, 28566 (2016).
52. Haeussler M, Schönig K, Eckert H *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* 17(1), 148 (2016).
53. Kim D, Bae S, Park J *et al.* Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* 12(3), 237–243 (2015).
54. Tsai SQ, Zheng Z, Nguyen NT *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33(2), 187–197 (2015).
55. Vouillot L, Thélie A, Pollet N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3 (Bethesda)* 5(3), 407–415 (2015).

56. Sentmanat MF, Peters ST, Florian CP, Connelly JP, Pruett-Miller SM. A survey of validation strategies for CRISPR-Cas9 editing. *Sci. Rep.* 8(1), 888 (2018).
57. Smith C, Gore A, Yan W *et al.* Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell.* 15(1), 12–13 (2014).
58. Zischewski J, Fischer R, Bortesi L. Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnol. Adv.* 35(1), 95–104 (2017).
59. Wang X, Wang Y, Wu X *et al.* Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat. Biotechnol.* 33(2), 175–178 (2015).
60. Cameron P, Fuller CK, Donohoue PD *et al.* Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods* 14(6), 600–606 (2017).
61. Tsai SQ, Nguyen NT, Malagon-Lopez J *et al.* CIRCLE-seq: a highly sensitive ***in vitro*** screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* 14(6), 607–614 (2017).
62. Wu X, Scott DA, Kriz AJ *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* 32(7), 670–676 (2014).
63. Fujita T, Yuno M, Fujii H. enChIP systems using different CRISPR orthologues and epitope tags. *BMC Res. Notes* 11(1), 154 (2018).
64. Xu X, Qi LS. A CRISPR-dCas toolbox for genetic engineering and synthetic biology. *J. Mol. Biol.* 431(1), 34–47 (2019).
65. Neguembor MV, Sebastian-Perez R, Aulicino F *et al.* (Po)STAC (Polycistronic SunTAg modified CRISPR) enables live-cell and fixed-cell super-resolution imaging of multiple genes. *Nucleic Acids Res.* 46(5), e30 (2018).
66. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533(7603), 420–424 (2016).
67. Chen JS, Dagdas YS, Kleinstiver BP *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550(7676), 407–410 (2017).
68. Cho SW, Kim S, Kim Y *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24(1), 132–141 (2014).
69. Cebrian-Serrano A, Davies B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm. Genome* 28(7–8), 247–261 (2017).
70. Adli M. The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* 9(1), 1911 (2018).
71. Wagner DL, Amini L, Wendering DJ *et al.* High prevalence of ***Streptococcus pyogenes*** Cas9-reactive T cells within the adult human population. *Nat Med.* doi: 10.1038/s41591-018-0204-6 (2018).
72. Crudele JM, Chamberlain JS. Cas9 immunity creates challenges for CRISPR gene editing therapies. *Nat. Commun.* 9(1), 3497 (2018).
73. Zetsche B, Heidenreich M, Mohanraju P *et al.* Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat. Biotechnol.* 35(1), 31–34 (2017).
74. Kim D, Kim J, Hur JK *et al.* Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* 34(8), 863–868 (2016).
75. Konermann S, Lotfy P, Brideau NJ *et al.* Transcriptome engineering with RNA-targeting Type VI-D CRISPR effectors. *Cell* 173(3), 665–676.e14 (2018).

TEG-seq: an ion torrent-adapted NGS workflow for *in cellulo* mapping of CRISPR specificity

Pei-Zhong Tang*,[1], Bo Ding[1], Lansha Peng[1], Vadim Mozhayskiy[1], Jason Potter[1] & Jonathan D Chesnut[1]

[1]*Thermo Fisher Scientific Inc., R&D Synthetic Biology, 5781 Van Allen Way, Carlsbad, CA 92008, USA*

BioTechniques 65: 259-267 (November 2018) 10.2144/btn-2018-0105

Keywords: CRISPR/cas9 • gene editing • GUIDE-seq • off-target • TEG-seq

GUIDE-seq was developed to detect CRISPR/Cas9 off-target. However, as originally reported, it was associated with a high level of nonspecific amplification. In an attempt to improve it, we developed target-enriched GUIDE-seq (TEG-seq). The sensitivity level reached 0.1–10 reads-per-million depending on the NGS platform used, which was equivalent to 0.0002–1% measured by Targeted Amplicon-seq. Application of TEG-seq was demonstrated for the evaluation of various Cas9/gRNA configurations, which suggests delivery of Cas9/gRNA ribonucleoprotein results in significantly fewer off-targets than Cas9/gRNA plasmid. TEG-seq was also applied to 22 gRNAs with relatively high *in silico* ranking score that targeted the biological relevant SNPs. The result indicated the initial selection of gRNAs with high score is important, although it cannot exclude the possibility of off-target.

The evolution of genome editing technologies holds promise for the concept of directly correcting mutations or disrupting abnormal genes in order to treat diseases, particularly inherited genetic disorders. Current databases indicate 135,588 variant–disease associations (VDAs), between 83,002 SNPs and 9169 diseases and phenotypes [1,2] in over 3874 genes [3]. These SNPs/mutations are potentially correctable or disruptable using genome editing tools such as CRISPR-Cas9 [4–7], TALEN [8–11] and ZFN [12–14]. As they are known to induce off-target mutations at sites with homology to the target sites, gene and cell therapeutic applications of these nucleases will require a comprehensive knowledge of their off-target effects to minimize the risk of deleterious outcomes.

Many strategies have been explored to improve the specificity of targeted nucleases. Modifications to the FokI dimerization domain increased the specificity of ZFNs and TALENs by requiring formation of obligate heterodimers to cleave the target DNA in a prescribed orientation and spacing [15,16]. The inactivation of Cas9 nuclease domains to create Cas9 nickases or dCas9–FokI fusions has also increased specificity by requiring two gRNA/Cas9 complexes [17–20]. Likewise, reducing the gRNA length from 20 to 17 nucleotides has been shown to increase specificity [21]. Recently, structure-guided protein engineering [22–24] and phage-assisted continuous evolution [25] have been used to develop novel Cas9 variants that show reduced off-target cleavage. These improvements have reduced initial concerns over the specificity of CRISPR/Cas9 nucleases. However, regardless of the nuclease technology, it is difficult to determine the full spectrum of off-target cleavage events in a complex genome under various experimental conditions. That said, an efficient, unbiased and reliable genome-wide off-target detection method is crucial for the application of genome editing-based gene and cell therapy as well as for benchmarking the fidelity evaluation of different gene editing tools.

Various methods have been developed to identify off-target, nuclease-cleaved sites. These methods rely on the double-strand breaks (DSBs) caused by nucleases that can be marked by a DNA tag either through *in vitro* ligation or *in cellulo* tag integration. The marked DSB site is then amplified and sequenced using next-generation sequencing (NGS). Several *in vitro* methods including BLESS-seq [26], HTGTS [27] and CIRCLE-seq [28] have been developed for off-target detection. These methods, by nature of the fact that the genomic DNA substrate has been removed from a cellular context and stripped of all protein, tend to identify all possible on- and off-target cleavage sites for a particular gRNA. The data analysis can also be challenging due to the potentially high nonspecific sequence reads as a result of the PCR amplification used to enrich the marked DSB. Digenome-seq [29], another *in vitro* method that relies on whole-genome sequencing, may reduce the noise and false positive calls, but likely sacrifices the sensitivity necessary to faithfully identify off-target sites that are cleaved at low frequencies. *In cellulo* approaches, in which off-target cleavage events occur and are tagged in living cells, more closely simulate the cellular nuclease-based gene editing environment. Two methods, IDLV-seq [30] and GUIDE-seq [31] were developed with this goal. IDLV-seq relies

METHOD SUMMARY

TEG-seq is a target-enriched GUIDE-seq using an ion-torrent next-generation sequencing platform for unbiased genome-wide CRISPR off-target detection. It reduces nonspecific amplification, and therefore increases its sensitivity for off-target detection.

on the integration of a linear viral DNA to the DSB site. However, the viral DNA is not chemically protected, which could lead to partial degradation and low efficiency of tag integration into DSBs and low sensitivity of detection. GUIDE-seq uses phosphorothioate-modified double strand DNA oligos (dsOND) as tags to reduce degradation and improve tag integration efficiency at DSBs in living cells. GUIDE-seq also requires an exponential PCR amplification step to enrich the marked DSB sites that may cause higher nonspecific noise and potential false-positive calls.

We report here a modification of the GUIDE-seq method, termed target-enriched GUIDE-seq (TEG-seq), in which 5′ phosphate primers are used for PCR amplification and differentially mark the amplicon containing the DSB site from nonspecifically amplified products. This is followed by preferential magnetic bead enrichment of marked DSB amplicons. These improvements significantly reduce nonspecific amplification and improve sensitivity of DSB detection. We also developed a 96-well format workflow that enables the study of multiple samples in parallel. The sensitivity of this method reached the level of Targeted Amplicon-seq, which has been widely used to validate the known and predicted gRNA off-target sites through direct amplification and NGS [28,31].

Materials & methods

Oligonucleotides, double strand DNA Tag & adaptors

The oligonucleotide sequences for TEG-seq and Targeted Amplicon-seq are listed in Supplementary Table 5. The double strand DNA (dsTag) and adaptors were annealed in the reaction containing 1 x TE buffer, 100 mM NaCl and 50 μM of top and bottom oligos, at 75°C for 5 min and gradually cooled down to room temperature in 20 to 30 min.

gRNA preparation

The gRNAs were synthesized using *in vitro* transcription as described in a previous publication [32]. Briefly, the 80 bp cr/tracrRNA constant region was PCR amplified from the GeneArt® CRISPR Nuclease Vector (1 ng) using the Constant Forward and Universal Reverse oligos (10 μM). The 80 bp cr/tracrRNA PCR products (0.15 μM) were equally mixed with universal forward and reverse oligos (10 μM) as well as target-specific forward and reverse oligos (0.3 μM). The *in vitro* transcription of gRNA was carried out using the TranscriptAid T7 High Yield Transcription Kit (ThermoFisher Scientific, MA, USA) following the manufacturer's protocol.

SpCas9 expression

For the plasmid format, the SpCas9 and gRNA were cloned in the GeneArt CRISPR Nuclease Vector (ThermoFisher Scientific) following the manufacturer's protocol. The recombinant wild type SpCas9 (Platinum™ Cas9) were purchased from ThermoFisher Scientific).

TEG-seq using the ion torrent platform

See the protocol provided in the Supplementary data.

Ion-torrent NGS & data analysis for TEG-seq & Targeted Amplicon-seq

Ion-torrent platform was used for all sequencing. PGM™ with 318-chip or Proton with PI-chip was used for TEG-seq to obtain at least one million reads per sample. Proton with PI-chip or S5 with 540-chip were used for Target Amplicon-seq to obtain at least 10,000 reads per sample. Reads were first aligned to the human genome reference, hg19. Mapped reads were further processed through in-house developed plug-in software, named Motif-Search, which performs the following: sorts barcodes for different samples, stacks and counts reads for all amplicons, and scans the genome and identifies the sequence and genome position for potential off-targets based on the guide and PAM sequence. The candidates for potential CRISPR/Cas9-induced DSB sites were compared with the control sample lacking CRISPR/Cas9 treatment (dsTag only) to determine if the candidates were related to CRISPR/Cas9-induced DSB sites. To compare different samples from various experiments and different NGS runs, reads from all samples were normalized using reads per million (RPM of mapped read).

NGS reads from the Targeted Amplicon-seq experiments were aligned to the corresponding reference PCR sequence. The mapped reads were further processed using the in-house developed plugin, named 'CELFT' (cut efficiency for low frequency target) to visualize the cleavage that contains 'large indel' and/or dsTag integration and calculated percentage of cleavage events. To minimize the false positive for the Targeted Amplicon-seq validation, caused by NGS error, especially the area with homopolymer in cleavage loci, only the large indel for having at least three or more bases was counted as a positive.

Results & discussion

The TEG-seq workflow

GUIDE-seq [31] and other genome-wide off-target detection methods [26–28,30] have been developed to provide relatively unbiased mapping of on- and off-target cleavage events. These technologies depend on the exponential amplification of tagged segments of the genome followed by sequencing and identification of those amplicons. While this technology has proven to be an advancement to the field, it can be associated with high background reads and a technically challenging workflow. Using a similar dsTag as previously reported [31], we tested the GUIDE-seq workflow using ion-torrent NGS (Ion-GUIDE-seq; Supplementary Figure 1 and 'Materials & methods'). The sequencing data showed that only a small portion of amplicons contained the desired dsTag sequence with the majority of the high-read hits not related to CRISPR/Cas9 editing (data not shown). It was hypothesized that the majority of PCR products amplified using the dsTag-specific primers (F or R) paired with the adaptor primer (P1) were actually the P1/P1 product (Supplementary Figure 1). This was confirmed by using the P1 primer alone under various conditions for multiple samples (Supplementary Figure 2). This over-representation of the P1/P1 product likely impaired the sensitivity of Ion-GUIDE-seq for genome-wide off-target detection.

To reduce the undesired P1/P1 product amplified in Ion-GUIDE-seq and enrich for the targeted DSB amplicons for NGS, we developed the TEG-seq sample preparation procedure. As shown in Figure 1, we used 5′-phosphorylated primers paired with the P1 primers for the second-round PCR amplification. This resulted in a 5′-phosphate on the products (5P-F2/P1 and 5P-R2/P1) but not on the P1/P1 product. In the next step, a nonphosphorylated barcoded adaptor (BC-A), containing the ion sequencing primer A was specifically ligated to the phosphorylated product

but not the P1/P1 product, which lacks the 5′phosphate required to ligate the nonphosphorylated barcode adaptor. A third round of PCR was then performed using P1 and 'A-tail' primers to further enrich the P1-A product. The BC-A-tail primer contains an internal spacer that stops the polymerase-mediated nucleotide extension and leaves a single-strand tail at the end of the PCR product. Finally, the dsTag-specific A-P1 product was further enriched from the remaining P1/P1 nonspecific product using a biotinylated oligo complimentary to the tail sequence and captured by streptavidin (SA) magnetic bead selection. The barcoded A-P1 product was then sequenced. By selectively modifying target PCR amplicons using 5′phosophorylated primers followed by bead enrichment for the targeted fragments, TEG-seq reduces nonspecific amplification and therefore likely increases sensitivity. We also used barcoding to multiplex the workflow to improve robustness and throughput.

Validation of TEG-seq accuracy using Targeted Amplicon-seq

To evaluate the sensitivity level of TEG-seq, we compared it with Targeted Amplicon re-sequencing using NGS (Targeted Amplicon-seq), which we believe is currently the most sensitive method to detect nuclease-induced editing at predicted chromosomal locations. The two gRNAs using targeted loci HEK4 and VEG1 (previously identified to have a high and medium number of off-target sites [31] with poor *in silico* ranking score (Supplementary Table 1) were chosen. We detected over 300 and 30 off-target candidates for HEK4 and VEG1 respectively from the TEG-seq experiments and confirmed 253 positive for HEK4 and 28 positive for VEG1 by Targeted Amplico-seq. The rarest off-target event detected by TEG-seq was approximately 1 RPM for HEK4 from 5 million and 4 RPM for VEG1 from 2.4 million total reads (Figure 2A and 2B, Supplementary Table 2). The rarest off-target detected by Targeted Amplicon-seq was 0.0004% for HEK4 and 0.0002% for VEG1. There was a reasonable correlation between TEG-seq and Targeted Amplicon-seq for both HEK4 and VEG1 with the correlation factor R value at 0.9070 and 0.9329 respectively.

Taken together, these data suggest that using NGS platforms such as PGM that produce several million total reads in a single run, TEG-seq is capable of detecting off-target events at levels as low as 1–10 RPM, which corresponds to the cleavage rate as low as 0.001–1% as detected by Targeted Amplicon-seq (Figure 2). The combination of TEG-seq for profiling off-targets followed by validation using Targeted Amplicon-seq could be a good approach to identify appropriately edited cell lines and to validate cells destined for use in cell and gene therapy.

Off-target detection sensitivity under various sequencing depths

The sensitivity of TEG-seq and GUIDE-seq detection is directly proportional to the sequencing depth. The data from our lab and others [28,31] were generated using either Ion-Torrent PGM or Illumina MiSeq, which yield approximately 1–5 million reads per run. The lower limit of detection using the PGM was approximately 1–10 RPM. It is likely that rarer off-target events could be detected by increasing the depth of

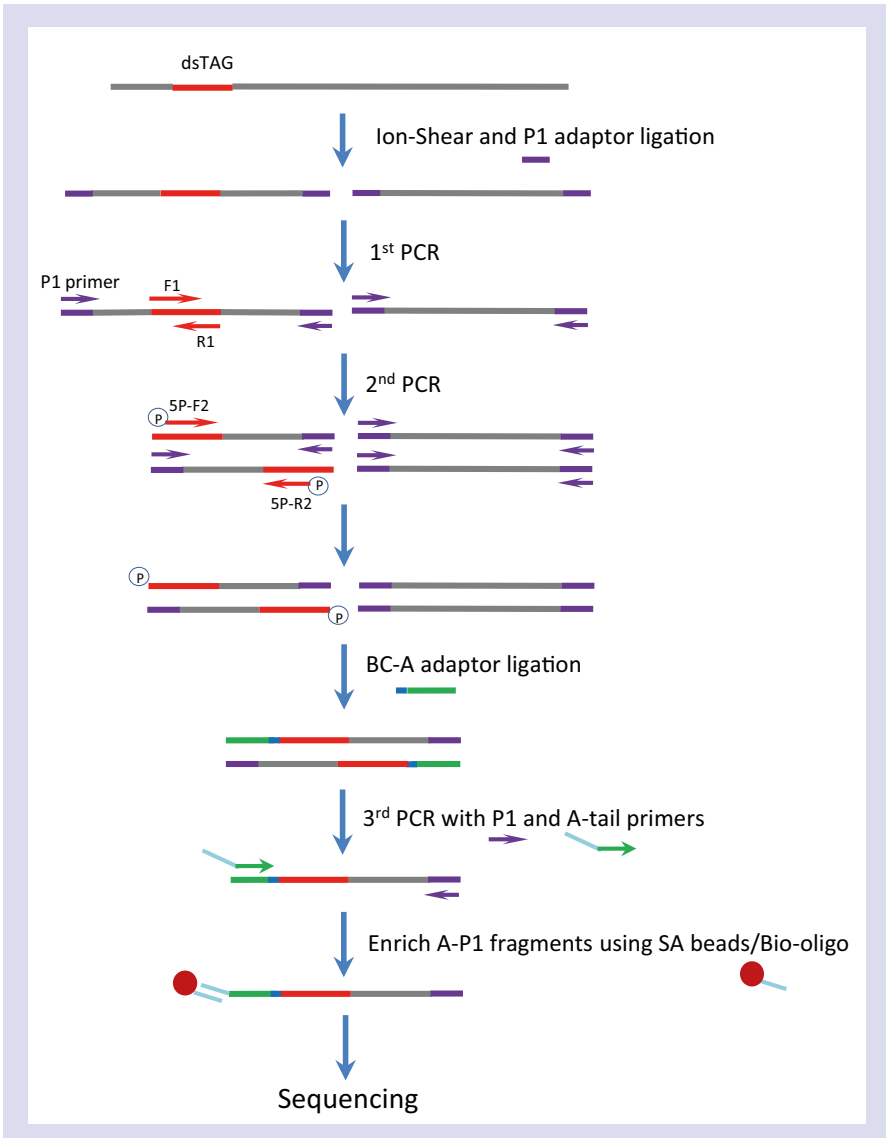


Figure 1. Diagram of ion-TEG-seq workflow. A dsTag was co-transfected with Cas9/gRNA ribonucleoprotein or plasmid that expresses Cas9 and gRNA. The genomic DNA was extracted and fragmented to 200–600 bp using enzyme-based ion-shear. The P1 adaptor was ligated to the fragmented DNA sample. The first PCR was performed in separated tubes using P1/F1 for the forward and P1/R1 for the reverse PCR. In the second (nested) PCR, 5′ phosphorylated primers (5P-F2 or 5P-R2) were used that generated PCR products containing a 5′ phosphate only in the F2/P1 or R2/P1 products, but not P1/P1 product. A nonphosphorylated barcode adaptor (BC-A) was specifically ligated to the F2/P1 or R2/P1 products, but not to the P1/P1 product. A third PCR was performed using P1/A-tail primer followed by a bead enrichment via a biotinylated capture oligo that was complementary to the A-tail primer. The enriched amplicons were then applied to next-generation sequencing.
dsTag: Double strand DNA tag; SA: Streptavidin.

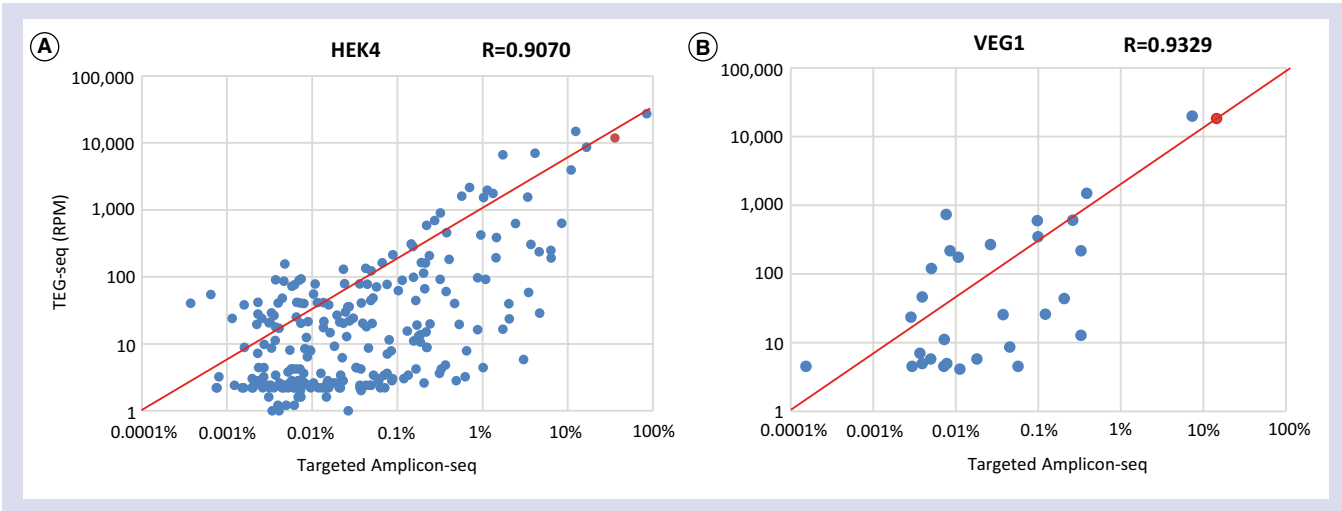


Figure 2. Comparison of off-target detection level between TEG-seq and Targeted Amplicon-seq. (A) HEK4 off-targets detected by TEG-seq is plotted in RPM against the percentage of cleavage detected by Targeted Amplicon-seq. The correlation factor R value is indicated on the upper right side. The on-target activity is indicated by red color. **(B)** VEG1 off-targets detected by TEG-seq in RPM is plotted to the percentage of cleavage detected by Targeted Amplicon-seq. The correlation factor R value is indicated on the upper right side. The on-target activity is indicated by red color.
RPM: Reads-per-million.

sequencing. To verify this, we used the Ion Proton and re-sequenced the HEK4 sample that was delivered with wild-type SpCas9 RNP. The Proton generated 42 million mapped reads, which is approximately 23-fold higher than the number of reads (1.8 million) from PGM. With this increased depth, we detected 32 additional off-target events. The read number per

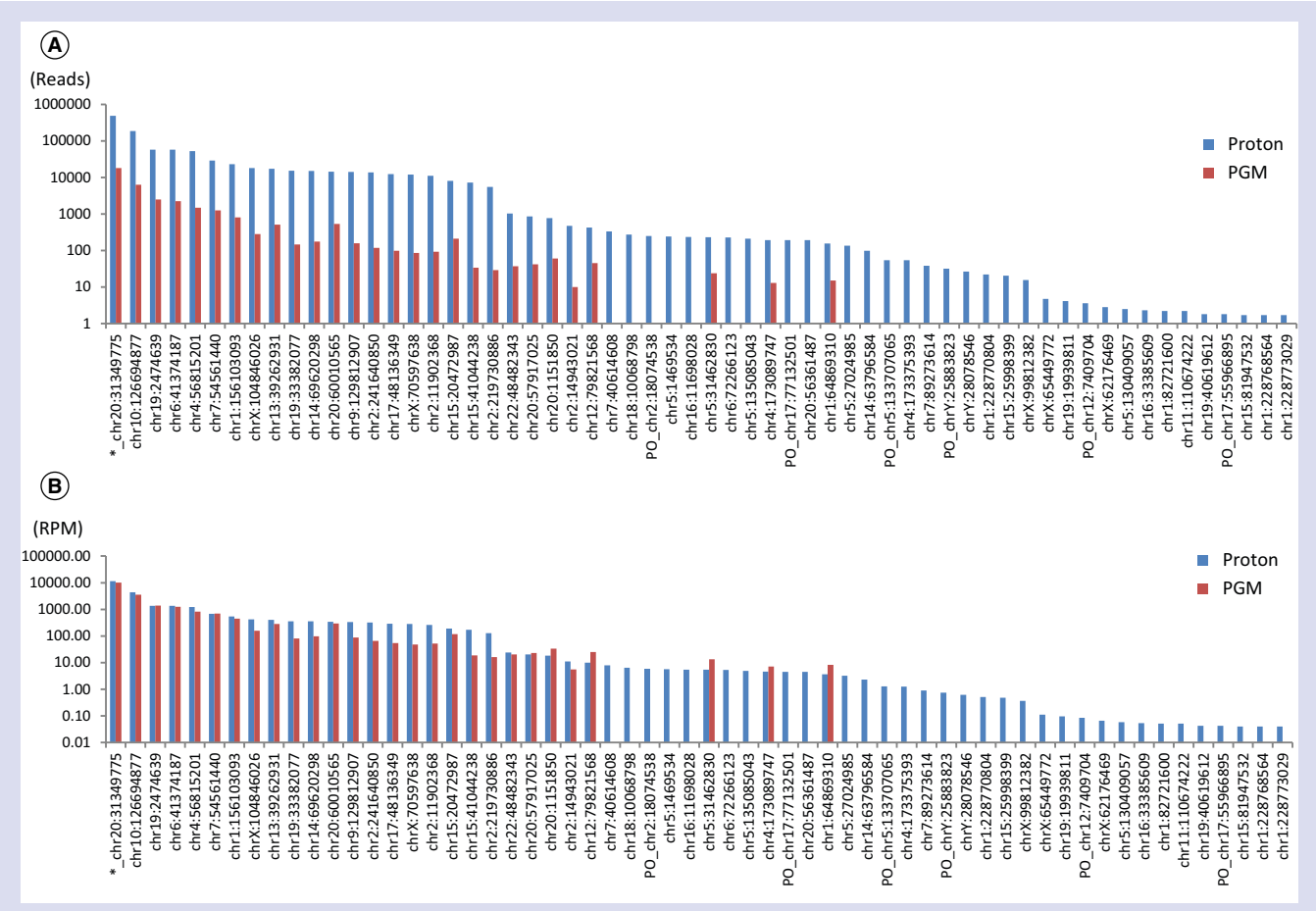


Figure 3. Off-target detection of HEK4 gRNA-directed cleavage at different sequencing depths. (A) Read count of HEK4 on- and off-targets. **(B)** RPM of HEK4 on- and off-targets. Asterisks indicate on-target events.
RPM: Reads-per-million.

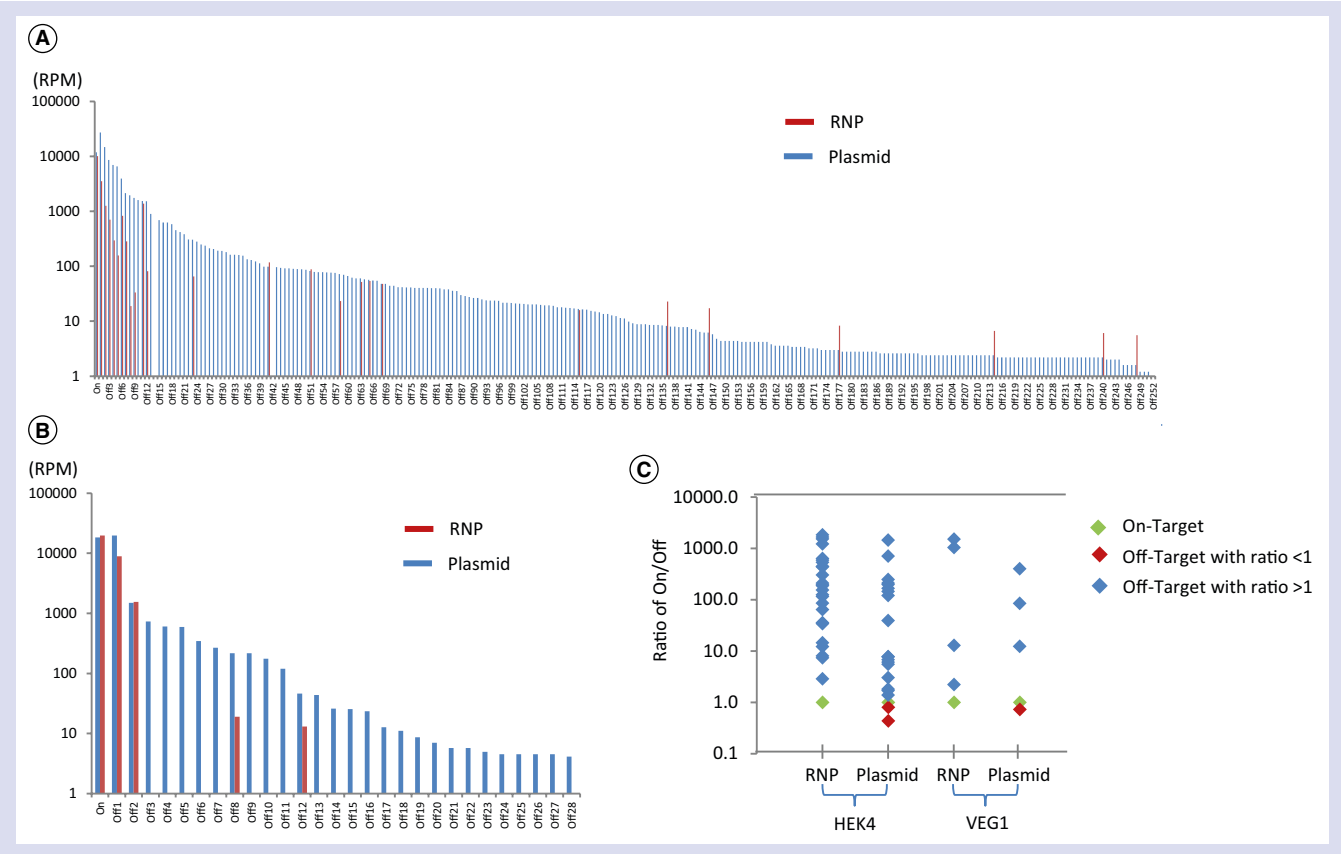


Figure 4. Off-target mapping using Cas9/gRNA plasmid and RNP delivery formats. (A) On- and off-target events detected using plasmid or RNP Cas9 with HEK4 site gRNA. **(B)** On- and off-target events detected using plasmid or RNP Cas9 with VEG1 site gRNAs. **(C)** Pairwise comparison of the on-/off-target ratio for those targets that were positive from both plasmid and RNP delivering formats.
RNP: Ribonucleoprotein; RPM: Reads-per-million.

target from the Proton is approximately a one order of magnitude increase over PGM (Figure 3A). However, their off-target profiles were similar after normalizing the read number to RPM with the additional, more rare 32 off-targets (Figure 3B). All 32 additional off-target events detected by Proton were below 8 RPM and ranged down to 0.04 RPM (Supplementary Table 3). 26 of the 32 additional off-target events detected by the Proton were also detected by PGM when the plasmid format was used. However, six appeared novel and not previously detected in this or other studies published [28,31]. These data indicate that with the throughput of 1–5 million reads generated by the PGM, off-target loci can be detected at the level of 1–10 RPM, which is equivalent to 0.001–1% from Targeted Amplicon-seq. Deeper sequencing platform such as Ion Proton or ion S5 would be recommended to detect lower than 1–10 RPM off-targets.

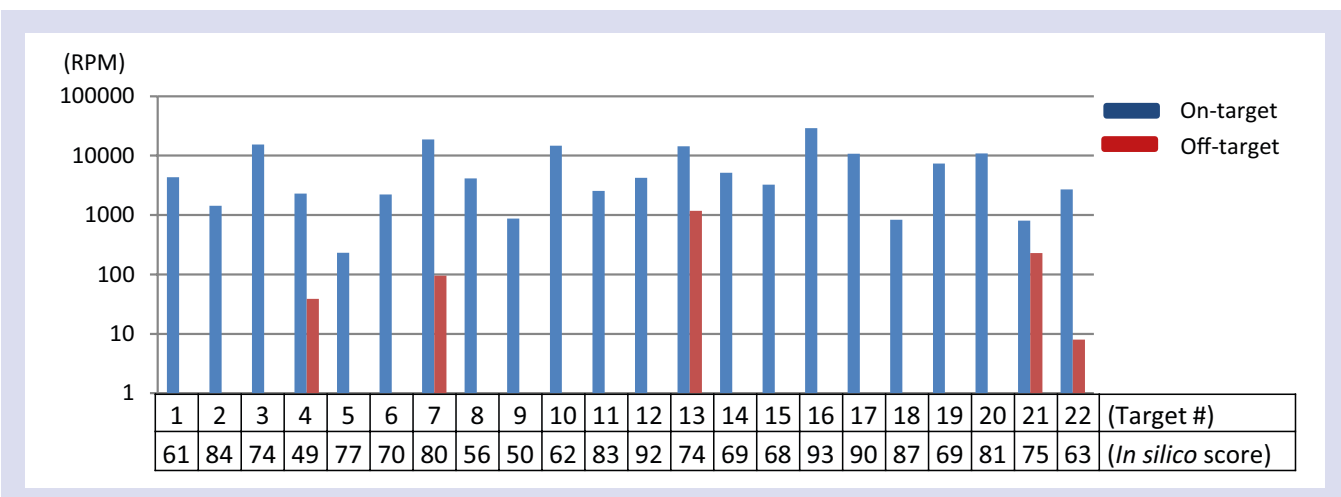


Figure 5. On- and off-targets detected from 22 gRNAs targeting biological relevant SNPs. 22 gRNAs with *in silico* ranking score from MIT CRISPR design tool targeting eight disease-associated SNPs (Supplementary Table 4). Their on- and off-targets were detected by TEG-seq and measured by PRM.

Table 1. The detailed result of five gRNAs detected with a single off-target.

Target no	Gene	gRNA	PAM	MIT *in silico* score	(RPM) TEG-seq	Ratio (On/off)	(%) Amplicon-Seq	MIT *in silico* predicted
4-on	*SMCHD*	GTTTCTCTATCATATAAGAA	TGG	49	2303			
4-off		: : : : : : _ : : C : : : : : : : : : :	TGG		39	59	0.29%	No
7-on	*HBB-2*	GTAACGGCAGACTTCTCCTC	AGG	80	18599			
7-off		:G : : : : C : : : : : : : : _ : : : : :	AGG		95	196	0.20%	No
13-on	*INS*	CTAGGTAGAGAGCTTCCACC	AGG	74	14342			
13-off		GC: : : : : : : : : C : : : : : : : :	AGG		1171	12	0.08%	Yes
21-on	*PAH*	TAGCGAACTGAGAAGGGCCG	AGG	75	801			
21-off		CTTG : : : : : : : : _ : T : : : : :	TGG		228	3.5	30.98%	No
22-on	*PAH*	TTCATCAGGTGCACCCAGAG	AGG	63	2700			
22-off		CG: : : : : : : CA : : : : : : : : :	GGG		8	338	0.01%	Yes

'_' in gRNA sequence is the missing base on genome that forms a 'bulge' mismatch.
RPM: Reads-per-million.

Comparison of genome-wide off-target cleavage using plasmid & RNP CRISPR/Cas9 formats

In comparison with Cas9 delivered to the cell via plasmid DNA, direct transfection of Cas9 protein/gRNA ribonucleoprotein complexes (RNPs) provides a transient burst of activity with no opportunity for direct integration and persistent cleavage activity [32,33]. The Cas9/gRNA RNP is quickly turned over by the cell, which likely lowers the cellular concentration and thus the opportunity for off-target cleavage. CRISPR/Cas9 RNP delivery has grown in popularity but validation of genome-wide off-target effects among wild-type and various 'high fidelity' Cas9 as RNPs has only been reported recently [28]. This type of specificity evaluation benchmarking of various RNPs will be important particularly for therapeutic applications where the Cas9 plasmid format is not optimal. In an effort to more clearly define the differences in specificity between Cas9 plasmid and RNP delivery, we applied TEG-seq to HEK4 and VEG1 gRNAs and found that RNP delivery yielded approximately nine- and six-fold fewer detectable off-target events (25 vs 252 and 4 vs 27) respectively as compared with Cas9 plasmid delivery (Figure 4A & B). In general we found significantly lower off-target RPM with the RNP format than the plasmid format while their RPM of on-targets were similar. More importantly, the ratio of on-/off-target was considerably higher with RNPs than with the plasmid format in a pairwise comparison to those positive off-targets from both formats (Figure 4C). In contrast to plasmid delivery, we did not detect any off-target events from the RNP delivery format that had higher cleavage activity than the on-target (the ratio of on-/off-target <1). Although the results showed marked improvement with RNP delivery, a subset of significantly active off-target sites remain.

These results support the idea that using Cas9 as RNP reduces the number of detectable off-target events and improves overall specificity over plasmid DNA as measured by the ratio of on-target to off-target events. This is important when considering cell line development where a tenfold difference in on-target activity to the next highest off-target event represents an approximate one in ten chance of isolating a cell line with both events occurring in the same genome. Conversely, for therapeutic applications of Cas9 RNPs in nondividing primary cells, complete elimination of off-target events would be desirable as clonal isolation is generally not achievable prior to application of cells to the patient.

Genome wide off-targets on 22 gRNAs with relative high ranking score designed from *in silico* tool

Significant effort has been made to create a rule set to enable better *in silico* prediction of CRISPR specificity [34,35]. However, there remains a relatively large discrepancy between off-targets detected by nonbiased approaches with those predicted using the current *in silico* design tools [28,31]. It's worth noting that the gRNAs (HEK4 and VEG1) evaluated in this study and those in previous work [28,31] had poor *in silico* ranking score (Supplementary Table 1). They were chosen specifically due to their propensity for high off-target cleavage. We felt it pertinent to apply the nonbiased approach to evaluate the gRNAs with relative high *in silico* ranking score to simulate the application scenario of gene correction. We chose 22 gRNAs targeting the SNPs that cause eight genetic disorders (Supplementary Table 4) using the wild-type SpCas9 RNP delivery format. Surprisingly, only five (23%) of 22 resulted in detectable single off-target cleavage events by TEG-seq and confirmed by Targeted Amplicon-seq (Figure 5 and Table 1). The off-target RPMs from TEG-seq were at least one order of magnitude lower compared with their corresponding on-target except the #21. The percentages of off-target cleavage from Amplicon-seq were all <0.29% except the #21. The *in silico* ranking score seem not well-correlated to the off-target result. For example, three targets (#7, #13 and #21) with the detectable off-targets have a relatively high ranking score. Of the five targets with detected off-target, only two were predicted by the design tool; three were not predicted, in which a 'bulge' mismatch was contained. The result implies that the initial selection of gRNA with high *in silico* ranking score is important although it cannot exclude the off-target. The probabilityof having the off-target is low if using Cas9/gRNA RNP delivery format with a high *in silico* score gRNA. To minimize detectable off-target, it is wise to design and test two to three gRNAs for each target.

This study reports an adaptation of the originally reported GUIDE-seq method to the Ion Torrent family of NGS instruments. The new protocol reduces nonspecific sequencing reads by enriching for relevant amplicons

prior to sequencing while also enabling multiplexing of experiments through barcoding. TEG-seq represents an improved tool for validating CRISPR nuclease specificity in a cellular context. The nonbiased detection methods such as TEG-seq could be used for initial specificity screening of gRNAs followed by validation of the final edited cell clones using deep Targeted Amplicon-seq to identify appropriately edited cell lines and to validate cells destined for use in cell and gene therapy.

Author contributions

Pei-Zhong Tang developed the method, designed and performed most of bench work, analyzed and generated data and drafted the manuscript. Lansha Peng performed some bench work. Bo Ding and Vadim Mozhayskiy developed bioinformatic tools. Jason Potter and Jonathan D Chesnut edited the manuscript.

Acknowledgments

The authors would like to thank John Bishop and Steven J Roman for the initial technical advises and discussion.

Open access

Competing & financial interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.future-science.com/doi/full/10.2144/btn-2018-0105

References

1. DisGeNET. www.disgenet.org/web/DisGeNET/menu;jsessionid=4aqu6smfdc4c13gfeo1lsnaxl
2. Clinvar. www.ncbi.nlm.nih.gov/clinvar
3. OMIM Gene Map Statistics. www.omim.org/statistics/geneMap
4. Mali P, Yang L, Esvelt KM *et al.* RNA-guided human genome engineering via Cas9. *Science* 339(6121), 823–826 (2013).
5. Cong L, Ran A, Cox D *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823 (2013).
6. Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24(6), 1012–1019 (2014).
7. Schumann K, Lin S, Boyer E *et al.* Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl Acad. Sci. USA* 112(33), 10437–10442 (2015).
8. Kim Y, Kweon J, Kim A *et al.* A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.* 31(3), 251–258 (2013).
9. Kim YK, Wee G, Park J *et al.* TALEN-based knockout library for human microRNAs. *Nat. Struct. Mol. Biol.* 20(12), 1458–1464 (2013).
10. Miller JC, Tan S, Qiao G *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29(2), 143–148 (2011).
11. Mussolino C, Morbitzer R, Lütge F, Dannemann N, Lahaye T, Cathomen T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 39(21), 9283–9293 (2011).
12. Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. *Science* 300(5620), 764 (2013).
13. Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* 19, 1279–1288 (2009).
14. Kim JS, Lee HJ, Carroll D. Genome editing with modularly assembled zinc-finger nucleases. *Nat. Methods* 7, 91–92 (2010).
15. Miller JC, Holmes MC, Wang J *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* 25(7), 778–785 (2007).
16. Guo J, Gaj T, Barbas CF III. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J. Mol. Biol.* 400(1), 96–107 (2010).
17. Ran FA, Hsu PD, Lin CY *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154(6), 1380–1389 (2013).
18. Mali P, Aach J, Stranges PB *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31(9), 833–838 (2013).
19. Tsai SQ, Wyvekens N, Khayter C *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576 (2014).
20. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 32, 577–582 (2014).
21. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32(6), 279–284 (2014).
22. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88 (2016).
23. Kleinstiver BP, Pattanayak V, Prew MS *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529(7587), 490–495 (2016).
24. Chen JS, Dagdas YS, Kleinstiver BP *et al.* Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature* 550(7676), 407–410 (2017).
25. Hu JH, Miller SM, Geurts MH *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556(7699), 57–63 (2018).
26. Crosetto N, Mitra A, Silva MJ *et al.* Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* 10(4), 361–368 (2013).
27. Frock RL, Hu J, Meyers RM *et al.* Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* 33(2), 179–187 (2015).
28. Tsai SQ, Nguyen NT, Malagon-Lopez J, Topkar VV, Aryee MJ, Joung JK. CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* 14(6), 607–614 (2017).
29. Kim D, Bae S, Park J *et al.* Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods* 12(3), 237–243 (2015).
30. Wang X, Wang Y, Wu X *et al.* Unbiased detection of off-target cleavage by CRISPR-Cas9 and TANENs using integrase-defective lentiviral vectors. *Nat. Biotechnol.* 33(2), 175–179 (2015).
31. Tsai SQ, Zheng Z, Nguyen NT *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33(2), 187–197 (2015).
32. Liang X, Potter J, Kumar S *et al.* Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J. Biotech.* 208, 44–53 (2015).
33. Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24(6), 1012–1019 (2014).
34. Doench JG, Hartenian E, Graham DB *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nat. Biotechnol.* 32(12), 1262–1267 (2014).
35. Doench JG, Fusi N, Sullender M *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34(2), 184–191 (2016).

First draft submitted: 17 July 2018;
Accepted for publication: 7 August 2018

Address correspondence to: Pei-Zhong Tang; Thermo Fisher Scientific Inc., R&D Synthetic Biology, 5781 Van Allen Way, Carlsbad, CA 92008, USA; E-mail: pei-zhong.tang@thermofisher.com

Everything you should know about CRISPR/Cas9

25 JUNE 2018 CRISPR

Eric Kmiec, PhD, Director of the Gene Editing Institute at the Helen F. Graham Cancer Center & Research Institute of Christiana Care Health System, discusses its use of CRISPR to edit DNA outside of the cell for the first time, the hype surrounding CRISPR and its initiative to bring gene editing experiments into colleges and high school laboratories..

UPDATED
3 December, 2018

Eric Kmiec, PhD, Director of the Gene Editing Institute at the Helen F. Graham Cancer Center & Research Institute of Christiana Care Health System

Please can you introduce yourself and tell us about your background?

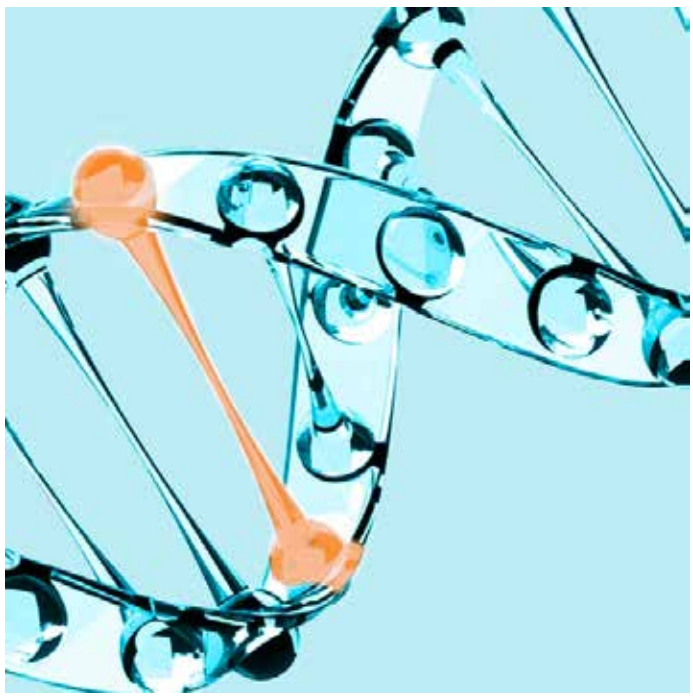
My scientific career began as an undergraduate at Rutgers University (NJ, USA). I then received a Masters from Southern Illinois University (IL, USA) and eventually a PhD at the University of Florida Medical School (FL, USA).

During my PhD, I began working on homologous recombination – the overall process that's used in gene editing – now known as homology directed repair. I studied how this type of recombination event directs evolution and the implications for that. I then began a long career focusing on how the proteins and enzymes actually acted to yield DNA and to pair chromosomes together.

My first faculty appointment was at the University of California, Davis (CA, USA), where I continued in this work and eventually found my way to the University of Delaware (DE, USA), where I spent 12 years as a tenured professor in the Department of Biology.

A few years ago, after meeting with the medical director of the Helen F. Graham Cancer Center & Research Institute at Christiana Care Health System (DE, USA), one of the nation's largest community-based health systems, I was convinced that the only way to truly do translational research in Delaware would be to embed a gene editing program within a community cancer center.

Delaware does not have its own medical school – it's small. So, we utilize Thomas Jefferson University (PA, USA) as the medical school for Delaware. There aren't a lot of impactful clinical opportunities to carry out translational research at the University of Delaware, so I lifted my 13 person lab and moved to the cancer center three years ago. I retained a faculty appointment at the University of Delaware and have also now acquired a faculty appointment at The Wistar Institute



(PA, USA). Right now, there are 15 people in the lab and they range from senior research associates to post docs to graduate students.

What brought you to the field of gene editing?

I was working peripherally in this area in a time I usually call BC – or 'before CRISPR'. The available tools at that time were not as elegant or, frankly, as good as they are now. My transition to the field came about because the perception of what we refer to as gene editing has always been about correcting or repairing genes. The gene editing reaction uses the same enzymes and regulatory factors that pair chromosomes and are involved in DNA replication or recombination activities.

So, while we continue to keep a strong interest in the enzymology of homologous recombination, there's just a natural fit for transitioning to translational medicine, applying gene editing to human cells. Our knowledge of how strands of DNA come together and how the cell deals with DNA pairing reactions is at the center of the entire gene editing reaction. Clinical, agricultural and pharmacological applications are natural extensions of this fundamental work.

The Gene Editing Institute recently published a paper detailing its novel use of CRISPR to edit DNA outside of the cell. Please can you tell us some more about this?

We are dedicated to finding out how CRISPR gene editing works at the enzymatic and molecular level by studying the mechanism of action and the regulation of gene editing inside the cell. Laboratories, like ours, are treating primary cells and animal models with gene editing tools, and in these reaction protocols, the editing tools are introduced into cells, then 24 or 48 hours later, you measure the outcome. In these experiments, you don't really know what's happened inside the cell, you are just happy it worked or disappointed it didn't.

We can guess and, of course, there are sophisticated techniques to do that, but our approach has always been to try and dig out the guts of the reaction and understand its internal workings. We developed a system where we can actually utilize CRISPR and a cell-free extract to carry out exactly the same reactions that are done inside a cell and get essentially the same outcome whether its deletion, insertion or replacement of genes. In my opinion, this work will lead to a fundamental understanding of what makes this whole process work and that will uncover the best approaches to enable gene editing in primary cells to work more efficiently.

From purely a reaction standpoint, the CRISPR/Cas9 complex enters the cell and executes DNA breakage signaling where the edit is to take place, but, in the end, the endogenous protein activity in the cell edits the DNA. Many studies have indicated that breaking the DNA followed by resection of the broken ends is more efficient than the insertion of a piece of corrected DNA that is designated to rescue a genetic mutation. We call this the CRISPR imbalance and it's at the heart of many investigations aimed at improving precision and efficiency. We approach this problem by breaking the enzymatic reactions down into fundamental pieces and putting them back together one by one so that we can understand what causes this genetic imbalance.

Using this experimental strategy, you can manipulate the extract by depleting a particular enzyme thereby defining its role. It then becomes more of a reduction to practice in a biochemical sense. As nucleic acid enzymologists, this is really exciting because we want to optimize the gene editing reaction for repairing genes, not just disabling them.

We recently published this work in the CRISPR Journal and one of the reviewers of our paper said they wished there were more papers like ours. Frankly, so do I. Everyone is racing toward therapy and we're their biggest fans, but going too fast can sometimes lead to overpromising outcomes. I believe that this technology is so extraordinary that it will survive any excess hype that naturally accompanies big discoveries.

Knowing how it works and potentially being able to tweak it a little bit is going to be where major improvements are made. As more clinically oriented scientists get their hands on these regulatory factors and are able to manipulate their activity, it is likely that impactful therapies for genetic disorders will begin to emerge with regularity.

Which techniques do you use in gene editing and what are the pros and cons of these?

We did a lot of early work with ssDNA alone when we didn't have CRISPR, so we employed donor DNA molecules known as oligonucleotides to carry out gene editing. This was successful and there were a number of groups using this approach, but we all found that the frequency with which single bases were repaired in the human chromosome was at a level that would not be clinically relevant.

When programmable nucleases appeared on the scene, that frequency was elevated quite significantly. Recently, many laboratories have moved to utilize adeno-associated viruses as the source of a donor DNA in gene editing reactions aimed at repairing a genetic mutation. Viral DNA overexpression is helpful to this process, but if there's a red flag in that approach, it's that high levels of ssDNA in a cell will likely induce several regulatory proteins that respond to cellular stress.

One of these is p53, a well-known marker for oncogenesis. Eight years ago, we reported that p53 is induced in gene editing reactions where anticancer drugs were used to execute DNA breakage prior to repair. So, I wouldn't be surprised if overloading a cell can lead to some adverse responses that may or may not impact the outcome of gene editing in human cells.

We used to work with TALENs but we don't use them anymore. They're a wonderful reagent, but cumbersome to work with. TALENs are probably more specific than CRISPR, especially when it comes to the current concern of offsite mutagenesis, but this is a tough problem to overcome because you're essentially trying to prove a negative – you're trying to prove that something didn't happen.

I don't see that as a major block to the implementation of CRISPR-directed gene editing in human cells, although skeptics have a significant voice in this field – and we want to hear from them. We've now moved to using Cas12a (Cpf1) instead of Cas9, and other workers are utilizing Cas9 nucleases from other sources. New categories of CRISPR are coming fast, many of which have interesting characteristics.

One emerging concern in the field is that much of the technology developed by major groups is now licensed to, and owned by, a small number of biotechnology companies. While they have, by and large, allowed workers to use them for research purposes without fees, accelerating these discoveries toward clinical application will undoubtedly require license agreements and legalities that will have the unintended consequences of restricting innovation.

About 20 years ago I lived through the promises of gene therapy cures being right around the corner, so now I am a little more skeptical as well as being a bit concerned that overpromising therapeutic outcomes will end up with the same fate. Many of the same laboratories that dominated the early days of viral gene therapy are now heavily engaged in the gene editing technology.

What are the key goals of the gene editing institute?

I see this field as being more than just developing therapeutic strategies for CRISPR. So, I started the Gene Editing Institute with four basic missions.

The first is developing CRISPR as an augmentative therapy for various types of cancer, including lung cancer and melanoma. We are fortunate to be integrated inside a community cancer center. Nationwide, over 85 percent of patients get treated for cancer in community cancer centers, as is the case in Delaware. In addition, these patients are mostly naïve in terms of drug therapy, meaning that they have not been previously treated with a drug in a clinical trial. Our accrual rate to cancer clinical trials is outstanding and I feel strongly that our cancer patients, some in economically challenged situations, deserve to have the option of frontline state-of-the-art therapies.

We have an extremely strong, integrated team including patient advocates, genetic counselors, oncologists and scientists working together to develop effective yet safe protocols for these patients. We are working hard to develop the center as a place where all cancer patients can be treated with the latest therapies. This mission also includes research into the mechanism of action and the regulatory circuitry surrounding human gene editing. To carry out these studies, we've had strong support from the NIH.

The second mission is education. I've been in this field for a long time, so I've lived through the dark ages and now the enlightenment has come. We were able to develop a series of reporter gene systems that helped us understand the mechanism of single agent gene editing. As CRISPR emerged, I thought that one of these systems would be valuable as a teaching tool for a laboratory exercise in gene editing.

So, we launched and developed that idea and won a National Science Foundation award to develop a full curriculum for instruction in gene editing technologies for four-year colleges, community colleges and even high schools. There are plenty of tutorials online, but this is hands on. You can't have a sophisticated experiment in high schools or colleges or community colleges because they don't always have the equipment to carry it out. We therefore developed a strategy that's quite similar to the experiment we published in the CRISPR Journal for these students. They can get an answer from the experiment in 16 hours, which is great, as materials, timing and the students' attention span won't stretch beyond an activity that could potentially take 2-3 weeks.

Our third mission is a simple one. We continue to act as a core facility for research labs around the country, and some overseas, that want to use CRISPR but haven't learned how to do it or how to make it. We provide the tools and a charge-back method for them to use in their research.

The fourth mission is partnership. We have a couple of signed agreements and we're licensing the technology to a couple of biotechnology companies. One of these partnerships has been catalyzed by a grant from the US-Israel Binational Industrial Research and Development Foundation.

We have an excellent relationship with the Israeli biotechnology company, NovellusDx, with whom we are working to improve the precision and efficiency of its well-established cancer diagnostic assay. They are an outstanding company and the scientific strategies surrounding the use of gene editing in cancer diagnostics is just beginning.

NovellusDx remains an extremely innovative leader in this field and has licensed our part of its diagnostic assay directly from Christiana Care. This project is particularly rewarding because it demonstrates the ultimate bench-to-bedside research paradigm. The reaction improvement is essentially the one that we jointly published with scientists from NovellusDx, and we are continuing to utilize it to understand the basic mechanism while NovellusDx is implementing it in a commercial product.

Another key partnership is with The Wistar Institute in Philadelphia (PA, USA). We provide them with the gene editing tools for researchers involved in vaccine development, cancer research and immunotherapy. This relationship allows us to interact on a daily basis with top-line, well-funded research laboratories and the flow of information is bidirectional. They are making us better scientists and I believe we are developing new approaches for their own research.

It's a really exciting time for gene editing, what exciting developments do you think the future holds?

For any genetic molecular translational approach, your first thoughts go immediately to therapy. A number of my colleagues have said that we'll be able to cure all genetic diseases. I'm not sure if we're going to get to that point, but helping people alleviate suffering and pain with a genetic tool is a great goal. I think the future will bring established FDA-approved CRISPR gene editing treatments for both cancer and genetic diseases, such as sickle cell anemia and maybe even cystic fibrosis.

Therapy always comes to the top because it directly affects people's lives. However, there's also remarkable work being done in agriculture and organ transplantation where pigs are being re-engineering to knock out endogenous viruses that can affect transplantation.

A lot of the work we do for academic colleagues around the country centers on knocking out one, two or three genes in a cell line. The ease with which you can create modified cell lines with a trained operator is quite extraordinary. From my perspective, pharmacogenomics has already been impacted quite dramatically and I think drug development and drug discovery will also be part of it.

I believe cancer will be an excellent target for CRISPR as, in most cases, you're not trying to repair a gene. You're trying to destroy oncogenes or genes that are contributing to resistance to chemotherapy or immunotherapy. Cancer is a great target because it's asking CRISPR/Cas9 to do exactly what it does in a bacterial cell and that's to destroy DNA, thereby allowing the tumor cells to be more susceptible to standard treatments.

One of the most extraordinary aspects of CRISPR-directed gene editing is the broad impact it has had on so many disciplines in the public platform, which has enabled important discussions about its use. There have been some new papers published using CRISPR as a diagnostic for HPV as well as articles about a home beer brewing kit. It seems at this point that the only thing CRISPR won't be able to do is open your refrigerator door!

CRISPR upgraded from ‘scissors’ to ‘shredder’

10 APRIL 2019 CRISPR

A novel CRISPR-Cas3 based tool acts more like a ‘shredder’, as opposed to the single-sequence targeting ‘scissors’, with the ability to wipe out long stretches of human DNA.

WRITTEN BY
Abigail Sawyer

UPDATED
13 May, 2019

SOURCE
Dolan AE, Hou Z, Xiao Y et al. Introducing a spectrum of long-range genomic deletions in human embryonic stem cells using Type I CRISPR-Cas. Mol. Cell. doi:10.1016/j.molcel.2019.03.014 (2019)

https://www.uofmhealth.org/news/archive/201904/new-dna-%E2%80%9Cshredder%E2%80%9D-technique-goes-beyond-crispr%E2%80%99s-scissors

https://www.sciencedirect.com/science/article/abs/pii/S1097276519302175?via%3Dihub

An international collaboration of researchers has successfully utilized Type I CRISPR-Cas3 to edit long stretches of DNA in human cells for the first time. This novel tool can target and delete much longer sections than current CRISPR tools, which could prove useful in targeting diseases that are associated with long stretches of DNA.

“Cas3 goes where you want it, travels along the chromosome, and makes a spectrum of deletions tens of kilobases long. This could make it a powerful screening tool to determine what large areas of DNA are most important for a particular disease.”

The tool comprises Type I CRISPR that, along with Type II CRISPR (which includes Cas9), is also found in bacteria. However, Type I has never been utilized in eukaryotic cells, and uses the Cas3 enzyme for shredding DNA. After the successful delivery of the system into human cells, it then deleted portions of targeted DNA that ranged from a few hundred base pairs to 100 kilobases.

“Cas9 is a molecular scissor that goes where you want it and snips once,” commented study leader Yan Zhang (University of Michigan, MI, USA). “But Cas3 goes where you want it, travels along the chromosome, and makes a spectrum of deletions tens of kilobases long. This could make it a powerful screening tool to determine what large areas of DNA are most important for a particular disease.”

The researchers believe this shredding technique could be useful in non-coding regions of DNA, where they can delete long sections to assess what happens. Cas3 also has the ability to travel long distances along chromosomes, which is currently not possible with Cas9. Therefore, it might also be a useful platform in epigenome engineering if the shredding ability is removed.

There were a few challenges faced by the researchers, as they needed to work out how to entice the human stem cells to reveal if any DNA had been deleted – those ‘reporter’ lines developed for CRISPR/Cas9 research were not sensitive enough for the new system if the shredding activity was low. They also needed to develop novel sequencing techniques to reveal which sequences had been shredded.

The team believes the Cas3 tool could be less likely to produce off-target effects than Cas9 systems due to a longer guide RNA sequence and better-defined targeting and degradation steps. However, further research is required to fully assess the new approach.

Solving the off-target effects of CRISPR/Cas9

We spoke to Mark Behlke, the chief scientific officer of Integrated DNA Technologies (IA, USA), about the current limitations of CRISPR/Cas9 techniques and the potential of the novel HiFI Cas9 enzyme to resolve these limitations, as shown in a sickle cell disease-causing mutation.