

# Quantifying CRISPR off-target effects

Soragia Athina Gkazi<sup>1\*</sup>

1. University College London, Great Ormond Street Institute of Child Health, Molecular and Cellular Immunology, London, UK

\*Corresponding author: Dr Soragia Athina Gkazi, a.s.gkazi@ucl.ac.uk

## Abstract

Recent advances in the era of genetic engineering have significantly improved our ability to make precise changes in the genomes of human cells. Throughout the years, clinical trials based on gene therapies have led to the cure of diseases such as X-linked severe combined immunodeficiency (SCID-X1), adenosine deaminase deficiency (ADA-SCID) and Wiskott-Aldrich syndrome. Despite the success gene therapy has had, there is still the risk of genotoxicity due to the potential oncogenesis introduced by utilising viral vectors. Research has focused on alternative strategies like genome editing without viral vectors as a means to reduce genotoxicity introduced by the viral vectors. Although there is an extensive use of RNA-guided genome editing via the clustered regularly interspaced short palindromic repeats (CRISPR) and associated protein-9 (Cas9) technology for biomedical research, its genome-wide target specificity and its genotoxic side-effects remain controversial. There have been reports of on- and off-target effects created by CRISPR-Cas9 that can include small and large indels and inversions, highlighting the potential risk of insertional mutagenesis. In the last few years, a plethora of *in silico*, *in vitro* and *in vivo* genome-wide assays have been introduced with the sole purpose of profiling these effects. Here we are going to discuss the genotoxic obstacles in gene therapies and give an up-to-date overview of methodologies for quantifying CRISPR-Cas9 effects.

## Introduction

Vector-mediated genotoxicity is a recognised safety concern of viral vector-based gene therapy [1]. Understanding the causative agents in viral vector genotoxicity will aid in improving pre-clinical safety assessment of these vectors. Starting with viral vector-mediated insertional mutagenesis (gain-of-function mutations, gene truncations, genomic translocations etc), the vector design, integration site and insertion profile, become central matters in viral vector-related genotoxicity. Efforts have been made to develop non-viral vector delivery with less insertional mutagenesis risk [2-4]. There has been an increase in investigating new technologies, such as clustered regularly interspaced short palindromic repeats (CRISPR) and associated protein-9 (Cas9), as a more specific targeted genome editor. Several research groups have devised a number of viral and non-viral vectors as appropriate delivery systems, especially since CRISPR-Cas9 allows for broad compatibility due to the small size of its two components.

Since the development of CRISPR-Cas9 system over a decade ago [5, 6] and its first use in genomic engineering research [7-10], this technology has revolutionised the world of gene editing. Nowadays, the CRISPR-Cas9 system has been applied to a number of different projects in order to correct specific genomic sites that are related to disease. The technology utilises short single guide RNA sequences (sgRNA) that bind to a specific target DNA sequence

in the genome and to the Cas9 enzyme which scans the genome for protospacer adjacent motifs (PAMs). Upon recognition of complementary DNA sequence by the sgRNA and its associated PAM, Cas9 proceeds with targeted cleavage of the DNA. The outcome of Cas9-mediated DNA cleavage is a double-stranded break (DSB) within the target DNA (~3-4 nucleotides upstream of PAM). Although the DSB can be repaired through the error-prone non-homologous end joining (NHEJ) pathway resulting in loss of function mutations, researchers have been able to trigger an alternative repair pathway, homologous-directed repair (HDR), which relies on the cell's own repair mechanism to add, delete or change the genetic material upon provision of exogenous customised DNA template. Although the targeting specificity of Cas9 is thought to be tightly controlled by the gRNA and its adjacent PAM, potential off-target cleavage activity could still occur [11-13]. These could be around the cut site or in homologous regions around the genome which Cas9 could recognise. Furthermore, optimising the composition of donor DNA which can be circular or linear double-stranded DNA (dsDNA), or single-stranded oligodeoxynucleotide (ssODN) have shown increased HDR efficiency [14] which can significantly reduce off-target indel generation [15-17]. Validation of all major off-target events is crucial for clinical use. In this review we are going to summarise the main quantification methodologies used to detect CRISPR-Cas9-mediated off-target effects.

### **Tools for quantifying CRISPR-Cas9**

The assays that have been introduced so far for quantifying off-targets in CRISPR-Cas9 edited cells fall into three subcategories; *in silico*, *in vitro* and *in vivo* assays. There are advantages and disadvantages in using any of these and deciding on one heavily relies on the sample material and the research question.

#### *In silico assays*

Computational algorithms that detect potential off-target sites are based massively on the sequence of the gRNA. The first web application described in literature is by Hsu and colleagues [18] which can be used as guidance for the selection and validation of specific target sequences as well as off-target analyses (<http://www.genome-engineering.org/>). They generated a set of sgRNAs targeting multiple sites within two human genomic loci with different trans-activating CRISPR RNA (tracrRNA) 3' truncations. Based on criteria produced by their experimental evidence, the group formulated a computational tool for selecting and validating sgRNAs as well as predicting potential off-target loci. The authors argue in the end that further *in vivo* investigation into the thermodynamics and stability of sgRNA-DNA duplexes as well as exploration of spCas9 variants and orthologs will improve the prediction power and specificity of their tool for off-targets. E-CRISP [19] is the next chronologically available online software which provides gRNA sequence design tools and evaluates off-target effects using alignment methodologies. Many such tools followed including but not limited to Cas-OFFinder [20], COSMID [21], Breaking-Cas [22] (Table 1). Initially, there were limitations with each of these tools in terms of number of mismatches allowed for finding off-targets, but since their creation, they have each updated their versions with more relaxed parameters allowing for greater amounts of off-target identification.

The most exciting, maybe, *in silico* methodology described in literature today is an algorithm that goes beyond predicting the off-target cleavage loci [23]. CRISTA (CRISPR Target Assessment) not only predicts cleavage efficacies but also employs machine learning tools providing a learning process in the patterns that underlie the mechanism of action of the CRISPR-Cas9 system. Abadi S. and colleagues, have described the algorithm as a state-of-the-art approach that considers a number of different features of the CRISPR-Cas9 system, including RNA thermodynamics, sequence similarity and DNA or RNA bulges that would have a pivotal role on CRISPR-Cas9 efficiency and selectivity [23]. Other such tools, utilising deep learning techniques have also been described [24, 25]. By the use of deep convolutional neural networks and deep feedforward neural networks, computational models were trained and tested on any released off-target datasets in CRISPR-Cas9 gene editing. These methodologies have been shown to outperform the current state-of-the-art prediction methods on two different datasets, CRISPOR and GUIDE-seq [25].

A comparative summary table of these *in silico* techniques is shown in Table 1. However, *in silico* predictions can be very broad and often biased, based on the parameters used, therefore, a combinatorial approach employing additional *in vitro* and/or *in vivo* techniques would be needed for evaluation purposes.

#### *In vitro* assays

When Cas9 and similar nucleases cut the genome, they create double-stranded breaks (DSBs). This is the principle used by most of the *in vitro* assays in order to investigate off-target events. They use Cas9 (or other nucleases) to cleave cell-free genomic DNA, sequence the material and then computationally detect DSBs in the sequencing data. By doing so, these assays become quite sensitive in identifying off-targets at a frequency as low as 0.1%. However, the use of cell-free DNA renders them unable to predict off-targets that occur within the cells. A summary of the *in vitro* techniques can be found in Table 2.

The first *in vitro* assay ever described, for quantifying off-target events, is Translocation-Capture sequencing (TC-seq) [26] that has been shown to study chromosomal rearrangements and translocations by infecting cells with a retrovirus expressing specific I-SceI sites with or without activation-induced cytidine deaminase (AIDCA or AID) protein [27, 28]. Genomic DNA from these cells is then isolated and library prepped for sequencing. All AID-dependent chromosomal rearrangements are identified while AID-independent translocations are discarded. Although it is an efficient protocol for studying chromosomal translocations within any given model or environment, PCR amplification errors and PCR biases in GC-rich templates could still occur.

DiGenome-seq was first introduced in 2015 [29] where off-target mutations are identified in cells with Cas9 digested genomes (digenomes). Digenome-seq relies on DNA cleavage rather than binding, and is performed at the genomic level where DNA/RNA bulges are captured and can detect off-targets at of a frequency as low as 0.01% in up to 10 gRNAs at once [30]. Its main advantage is that the DSBs introduced by Cas9 will not be processed by the DNA repair machinery, as opposed to Breaks Labelling, Enrichment on Streptavidin and next generation Sequencing (BLESS) and GUIDE-seq (see *in vivo* assays), increasing the possibility of detecting off-targets. Nonetheless, this technique could also lead to a lot of false positives due to not

being able to map properly sequence reads that are around naturally occurring indel sites. *In vivo* cleavage confirmation would still be required, as would a high skilled bioinformatics analysis [31]. Recently, the same research team that created DiGenome-seq investigated the possibility of CRISPR-Cas9 on- and off- targets being affected by chromatin in eukaryotic cells [32] with an optimised version of DiGenome-seq. Cas9 fusion with chromatin is found to improve its activity by up to several fold [33]. Their work found that chromatin affects genome-wide CRISPR specificity and they hence developed a new tool, DIG-seq, that uses chromatin DNA rather than histone-free DNA *in vitro*.

Another biochemical method to identify off-target harbouring sites is SITE-seq [34, 35]. Using this, extracted and purified genomic DNA is cleaved with Cas9 and Cas9 cleavage sites are biochemically probed and enriched for next generation sequencing. Bioinformatics are then used to identify off-target cleavage sites by selecting for targets with the highest possible activity and specificity. The authors of SITE-seq state that the signature is similar to the one observed with Digenome-seq [34].

The latest addition to the *in vitro* screens for genome-wide CRISPR-Cas9 nuclease off-targets is CIRCLE-seq [36]. In contrast to previously published *in vitro* methods, CIRCLE-seq can be performed using widely accessible next generation sequencing (NGS) technology and requires no reference genome sequence. The protocol involves shearing and circularisation of the purified genomic DNA and degradation of any residual linear DNA. Cas9 nuclease is then used to linearize the circular DNA containing a Cas9 cleavage site, and the cleaved ends are amplified and sequenced to detect off-targets [37]. By enriching for Cas9 nuclease-cleaved genomic DNA before sequencing, CIRCLE-seq becomes more sensitive in locating off-target events. Moreover, it removes the need for larger sample sizes and read depths that introduce background noise rendering identification of low frequency cleavage events even harder as in the case of Digenome-seq [36]. However, careful bioinformatics analysis needs to be carried over due to the amplification bias during Cas9 nuclease-cleaved genomic DNA enrichment.

### *In vivo* assays

In the last 5 years there has been a flurry of development of different *in vivo* assays with the earliest and most cited one being ChIP-seq [13, 38-40]. It uses chromatin immunoprecipitation coupled with high-throughput sequencing to detect Cas9 binding sites and chemical modifications of histone proteins genome-wide. The limiting requirements of ChIP-seq include : (1) high-complexity libraries (80% of 10 million or more reads to be mapped to distinct genomic locations), (2) replicates of two per experiment should be carried out, where either 80% or 40% of the identified targets in replicate one should be among the targets of the replicate two and (3) large numbers of cells (~10 million) [41]. However, there have been advances in the technology since ChIP-seq was first developed. For instance, the Encyclopaedia of DNA Elements (ENCODE) Consortium suggests limiting the genomic regions to be investigated to a few candidate regions, if possible, and then validating them using biological experiments for ChIP-seq to be able to efficiently detect targets in low-complexity libraries [42]. Since ChIP-seq was first used for identifying off-target events, other *in vivo* assays have been established (Table 3).

BLESS is a genome-wide *in vivo* approach to map DSBs at nucleotide resolution by direct *in situ* Breaks Labelling, Enrichment on Streptavidin and next generation Sequencing [43, 44]. The advantages of this methodology are that it can detect DSBs at a nucleotide resolution and does not depend on proteins that bind to DSBs nor single-stranded DNA, minimising bias (Table 3). The disadvantages are that the data have a high background, it only maps unjoined ends and it is susceptible to artefacts associated with cell fixation [31, 45].

The same group that developed the *in vitro* assay CIRCLE-seq [36], had originally introduced an *in vivo* assay, GUIDE-seq, that uses the integration of blunt-ended double-stranded oligodeoxynucleotides (DSOs) into DSBs, genome-wide [45]. The DSO integration sites are then mapped to the exact positions in the genome at nucleotide level using amplification steps and sequencing [45]. Using this method, specific landscapes for the RNA-guided nucleases can be generated globally in living human cells. Moreover, targeted sequencing decreases the costs of sequencing tremendously when compared to other known protocols for the investigation of off-targets. However, as shown in the CIRCLE-seq comparison experiments [36], certain sites are undetectable by GUIDE-seq due to lower read counts. This detection limit could compromise validation of off-targets sites in cells analysed by targeted sequencing, as the lower limit of detection by NGS remains at 0.1% (Table 3). Instead, high-depth targeted amplicon sequencing using genomic DNA from cell-based GUIDE-seq experiments could address this concern [36].

As early as 2011, a research group developed an *in vivo* assay to study genome-wide translocations that occur in haematopoietic malignancies [46]. The assay involved high throughput, genome-wide translocation sequencing (HTGTS) in mammalian cells, particularly identifying translocation events induced by Activation Induced-cytidine Deaminase (AID)-dependent *IgH* class switching [46] and CRISPR-Cas9 [30]. It is proven to have a higher efficiency when compared to whole genome sequencing (WGS) but tends to underestimate the frequency of DSBs [31] and is limited by chromatin accessibility [47]. Moving forward, the research team that first performed HTGTS using Cas9:sgRNA (Cas9: single guide RNA) successfully, introduced a modified version of HTGTS that included a linear amplification mediated PCR step (LAM-PCR) upgrading the technique [47]. The now called LAM-HTGTS, uses “prey” DSBs that bind to “bait” DSBs (Table 3), and their junctions from the isolated genomic DNA are detected in a robust and unbiased manner [31]. It is the only assay, thus far, that can efficiently detect all recurrent DSBs that occur during a period of time in a population of cells [31]. Although this is an exceptionally sensitive method for detecting large genomic rearrangements, it relies on the presence of both a “prey” and a known “bait” DSB, excluding its use on previously isolated genomic DNA without *priori* knowledge of a recurrent DSB that can serve as a “bait”. Moreover, LAM-HTGTS only produces information about the “prey” DSB that binds to a “bait” DSB and misses any data on the “prey” DSB that could persist as a DSB. Finally, it may have a constrained outcome if input material is limited.

The latest development in the *in vivo* predictions for off-target sequences is the use of integrase-defective lentiviral vectors (IDLVs) which was first introduced for the detection of off-target cleavage sites of zinc finger nucleases (ZFNs) [48] but was further optimised for the CRISPR-Cas9 system too [49]. IDLVs have been shown to detect off-targets to a frequency as low as 1%, which is relatively insensitive in comparison to other techniques (DiGenome-seq,

CIRCLE-seq), but is also able to recognise a single nucleotide skipping mutation from sgRNA or its genomic target efficiently [49] (Table 3).

## Summary

- No one tool for identifying off-targets in CRISPR-Cas9 can be used in isolation and hence a combinatorial approach of two or more techniques based on the research hypothesis of each laboratory is advised.
- Technical limitations as a result of small sized amplicons will always arise that NGS could overcome by maintaining high coverage and large read depth to allow for the detection of very rare large indels, an apparent issue in many of the assays. Therefore, choosing the most appropriate sequencing platform and careful analysis of the sequencing data is essential too.
- Special care must be taken when discarding false-positive sequence reads that result from amplification artifacts during PCR. In such cases, including a negative control (with no nuclease expression) at each target site could prove to be beneficial [50].
- Optimising the delivery efficiency could also further reduce off-targets. Studies have shown that a shorter Cas9 expression duration [51], the use of Cas9 nickases [52], Cas9-FokI chimeric proteins [53], proper modifications to residues of the Cas9 protein [54, 55] could greatly increase specificity for genome editing.
- Most importantly, there have been reports on significant, unexpected on-target mutations such as large deletions and complex mutations in mouse embryonic stem cells, mouse hematopoietic progenitors and a human differentiated cell line [56]. This on-target mutagenesis could potentially lead to pathogenicity if used in a clinical setting. Thorough examination of the genome and comprehensive genomic analysis is pivotal for the identification of normal genomes prior to patient administration.

## References

1. David, R.M. and A.T. Doherty, *Viral Vectors: The Road to Reducing Genotoxicity*. *Toxicol Sci*, 2017. **155**(2): p. 315-325.
2. Wang, M., et al., *Efficient delivery of genome-editing proteins using bio-reducible lipid nanoparticles*. *Proc Natl Acad Sci U S A*, 2016. **113**(11): p. 2868-73.
3. Yin, H., et al., *Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo*. *Nat Biotechnol*, 2016. **34**(3): p. 328-33.
4. Zuris, J.A., et al., *Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo*. *Nat Biotechnol*, 2015. **33**(1): p. 73-80.
5. Jore, M.M., et al., *Structural basis for CRISPR RNA-guided DNA recognition by Cascade*. *Nature Structural & Molecular Biology*, 2011. **18**(5): p. 529-U141.
6. Wiedenheft, B., et al., *Structural Basis for DNase Activity of a Conserved Protein Implicated in CRISPR-Mediated Genome Defense*. *Structure*, 2009. **17**(6): p. 904-912.
7. Cho, S.W., et al., *Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease*. *Nature Biotechnology*, 2013. **31**(3): p. 230-232.
8. Cong, L., et al., *Multiplex Genome Engineering Using CRISPR/Cas Systems*. *Science*, 2013. **339**(6121): p. 819-823.
9. Jinek, M., et al., *RNA-programmed genome editing in human cells*. *Elife*, 2013. **2**.

10. Mali, P., et al., *RNA-Guided Human Genome Engineering via Cas9*. Science, 2013. **339**(6121): p. 823-826.
11. Cradick, T.J., et al., *CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity*. Nucleic Acids Research, 2013. **41**(20): p. 9584-9592.
12. Fu, Y.F., et al., *High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells*. Nature Biotechnology, 2013. **31**(9): p. 822-+.
13. Kuscu, C., et al., *Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease*. Nature Biotechnology, 2014. **32**(7): p. 677-+.
14. Song, F. and K. Stieger, *Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks*. Molecular Therapy-Nucleic Acids, 2017. **7**: p. 53-60.
15. Chen, X., et al., *In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting*. Nat Commun, 2017. **8**(1): p. 657.
16. Nakajima, K., et al., *Precise and efficient nucleotide substitution near genomic nick via noncanonical homology-directed repair*. Genome Res, 2018. **28**(2): p. 223-230.
17. Richardson, C.D., et al., *Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA*. Nat Biotechnol, 2016. **34**(3): p. 339-44.
18. Hsu, P.D., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases*. Nat Biotechnol, 2013. **31**(9): p. 827-32.
19. Heigwer, F., G. Kerr, and M. Boutros, *E-CRISP: fast CRISPR target site identification*. Nature Methods, 2014. **11**(2): p. 122-124.
20. Bae, S., J. Park, and J.S. Kim, *Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases*. Bioinformatics, 2014. **30**(10): p. 1473-1475.
21. Cradick, T.J., et al., *COSMID: A Web-based Tool for Identifying and Validating CRISPR/Cas Off-target Sites*. Molecular Therapy-Nucleic Acids, 2014. **3**.
22. Oliveros, J.C., et al., *Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes*. Nucleic Acids Research, 2016. **44**(W1): p. W267-W271.
23. Abadi, S., et al., *A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action*. Plos Computational Biology, 2017. **13**(10).
24. Chuai, G.H., et al., *DeepCRISPR: optimized CRISPR guide RNA design by deep learning*. Genome Biology, 2018. **19**.
25. Lin, J. and K.C. Wong, *Off-target predictions in CRISPR-Cas9 gene editing using deep learning*. Bioinformatics, 2018. **34**(17): p. 656-663.
26. Klein, I.A., et al., *Translocation-Capture Sequencing Reveals the Extent and Nature of Chromosomal Rearrangements in B Lymphocytes*. Cell, 2011. **147**(1): p. 95-106.
27. Jankovic, M., et al., *53BP1 Alters the Landscape of DNA Rearrangements and Suppresses AID-Induced B Cell Lymphoma*. Molecular Cell, 2013. **49**(4): p. 623-631.
28. Pefanis, E., et al., *Noncoding RNA transcription targets AID to divergently transcribed loci in B cells*. Nature, 2014. **514**(7522): p. 389-+.
29. Kim, D., et al., *Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells*. Nature Methods, 2015. **12**(3): p. 237-+.
30. Mei, Y., et al., *Recent Progress in CRISPR/Cas9 Technology*. Journal of Genetics and Genomics, 2016. **43**(2): p. 63-75.

31. Hu, J.Z., et al., *Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing*. Nature Protocols, 2016. **11**(5): p. 853-871.
32. Kim, D. and J.S. Kim, *DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA*. Genome Research, 2018. **28**(12): p. 1894-1900.
33. Ding, X., et al., *Improving CRISPR-Cas9 Genome Editing Efficiency by Fusion with Chromatin-Modulating Peptides*. Crispr Journal, 2019. **2**(1): p. 51-63.
34. Cameron, P., et al., *Mapping the genomic landscape of CRISPR-Cas9 cleavage*. Nature Methods, 2017. **14**(6): p. 600-+.
35. Cameron, P.S., A. H.; Fuller, C. K.; Thompson, M. S.; Cigan, A. M.; Young, J. K.; May, A. P., *SITE-Seq: A Genome-wide Method to Measure Cas9 Cleavage*. Nature Publishing Group, 2017.
36. Tsai, S.Q., et al., *CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets (vol 14, pg 607, 2017)*. Nature Methods, 2018. **15**(5): p. 394-394.
37. Lazzarotto, C.R., et al., *Defining CRISPR-Cas9 genome-wide nuclease activities with CIRCLE-seq*. Nature Protocols, 2018. **13**(11): p. 2615-2642.
38. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome*. Cell, 2007. **129**(4): p. 823-837.
39. Johnson, D.S., et al., *Genome-wide mapping of in vivo protein-DNA interactions*. Science, 2007. **316**(5830): p. 1497-1502.
40. Mikkelsen, T.S., et al., *Genome-wide maps of chromatin state in pluripotent and lineage-committed cells*. Nature, 2007. **448**(7153): p. 553-U2.
41. Furey, T.S., *ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions*. Nature Reviews Genetics, 2012. **13**(12): p. 840-852.
42. Landt, S.G., et al., *ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia*. Genome Research, 2012. **22**(9): p. 1813-1831.
43. Crosetto, N., et al., *Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing*. Nature Methods, 2013. **10**(4): p. 361-+.
44. Ran, F.A., et al., *In vivo genome editing using Staphylococcus aureus Cas9*. Nature, 2015. **520**(7546): p. 186-U98.
45. Tsai, S.Q., et al., *GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases*. Nature Biotechnology, 2015. **33**(2): p. 187-197.
46. Chiarle, R., et al., *Genome-wide Translocation Sequencing Reveals Mechanisms of Chromosome Breaks and Rearrangements in B Cells (vol 147, pg 107, 2011)*. Cell, 2011. **147**(7): p. 1640-1640.
47. Frock, R.L., et al., *Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases*. Nature Biotechnology, 2015. **33**(2): p. 179-186.
48. Gabriel, R., et al., *An unbiased genome-wide analysis of zinc-finger nuclease specificity*. Nat Biotechnol, 2011. **29**(9): p. 816-23.
49. Wang, X.L., et al., *Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors*. Nature Biotechnology, 2015. **33**(2): p. 175-178.
50. Cho, S.W., et al., *Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases*. Genome Research, 2014. **24**(1): p. 132-141.



51. Chen, Y.H., et al., *A Self-restricted CRISPR System to Reduce Off-target Effects (vol 24, pg 1508, 2016)*. *Molecular Therapy*, 2016. **24**(12): p. 2132-2132.
52. Ran, F.A., et al., *Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity (vol 154, pg 1380, 2013)*. *Cell*, 2013. **155**(2): p. 479-480.
53. Tsai, S.Q., et al., *Dimeric CRISPR/Cas-based nucleases for high-precision genome editing in human cells*. *Human Gene Therapy*, 2014. **25**(11): p. A120-A121.
54. Kleinstiver, B.P., et al., *High-Fidelity CRISPR-Cas9 Nucleases with No Detectable Genome-Wide Off-Target Effects*. *Molecular Therapy*, 2016. **24**: p. S288-S288.
55. Slaymaker, I.M., et al., *Rationally engineered Cas9 nucleases with improved specificity*. *Science*, 2016. **351**(6268): p. 84-88.
56. Kosicki, M., K. Tomberg, and A. Bradley, *Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements*. *Nat Biotechnol*, 2018. **36**(8): p. 765-771.

### Table Legends

**Table 1.** A comparative summary for the *in silico* techniques described here.

**Table 2.** A descriptive summary for all *in vitro* techniques discussed here.

**Table3.** *In vivo* off-target detection assays: a summary table.