BioSkryb GENOMICS

WhitePaper

Comprehensive, Single-Cell Multiomic Analysis Is Needed for Safer Cell and Gene Therapies Developed with CRISPR/Cas9 Technologies

Key Takeaways:

- **Gene editing by CRISPR/Cas9 can remove or repair dysfunctional genes.**
- **CRISPR/Cas9 creates on-target edits, which impact the targeted gene, and off-target edits, which occur outside of the targeted region.**
- **Off-target Cas9 activity results in off-target gene editing and mutations that vary between cells.**
- **Precise and accurate single-cell whole genome sequencing, enabled by primary template-directed amplification (PTA) found in ResolveDNA® kits, can accurately detect on- and off-target mutations following CRISPR/Cas9 gene editing.**
- **ResolveOME™, a single-cell whole genome and transcriptome workflow, further enables investigators to draw novel links between genomic edits and gene expression, improving understanding of events that may impact safety and efficacy of cell and gene therapies developed with CRISPR/Cas9.**

To talk to a BioSkryb Genomics scientist email:

science@bioskryb.com

The public health impact of cell and gene therapies is likely to be measured on the same scale as the great medical advances of the previous three centuries: antibiotics, anesthesia, and vaccines. However, significant challenges remain with today's cell and gene therapy technologies, foremost the potential for inducing unintended and variable genomic edits and gene dysregulation, which can reduce the safety and efficacy of these therapies. With more sensitive single-cell and low input DNA analysis methods, cell and gene therapy developers will have the tools to detect off-target genomic edits and transcriptomic expression changes in single cells and population subsets, increasing confidence in therapeutic safety and efficacy.

Gene editing by CRISPR/Cas9 can remove or repair dysfunctional genes

In recent years, CRISPR-Associated Protein 9 (Cas9) has shown incredible promise in genetic engineering, both *in vitro* and *in vivo*. CRISPR/Cas9 systems can be designed to recognize and cleave specific sites in the genome through the use of a guide RNA (gRNA) with approximately 20 nucleotides of sequence complementary to a genomic target site. CRISPR/Cas9 gene modification has already shown its potential in the laboratory: it has been used to inactivate oncogenes and

drug resistance genes in cell culture and has demonstrated the potential to treat conditions ranging from Alzheimer's to broken bones in cell culture and animal models. It has also graduated to clinical trials, with numerous studies enrolling hundreds of patients to verify the feasibility of CRISPR treatments for diseases such as sickle cell, diabetes, and cancer.[1] The first CRISPR-based therapy, Casgevy™ (exagamglogene autotemcel), was recently approved for treatment of certain patients with sickle cell disease or transfusion-dependent ß-thalassemia in several regions including the United States, United Kingdom, and European Union.

There are multiple ways in which CRISPR/Cas9 can be deployed (**Figure 1a**). Two gRNA sequences used in a dual guide system will generate separate double-stranded cuts which

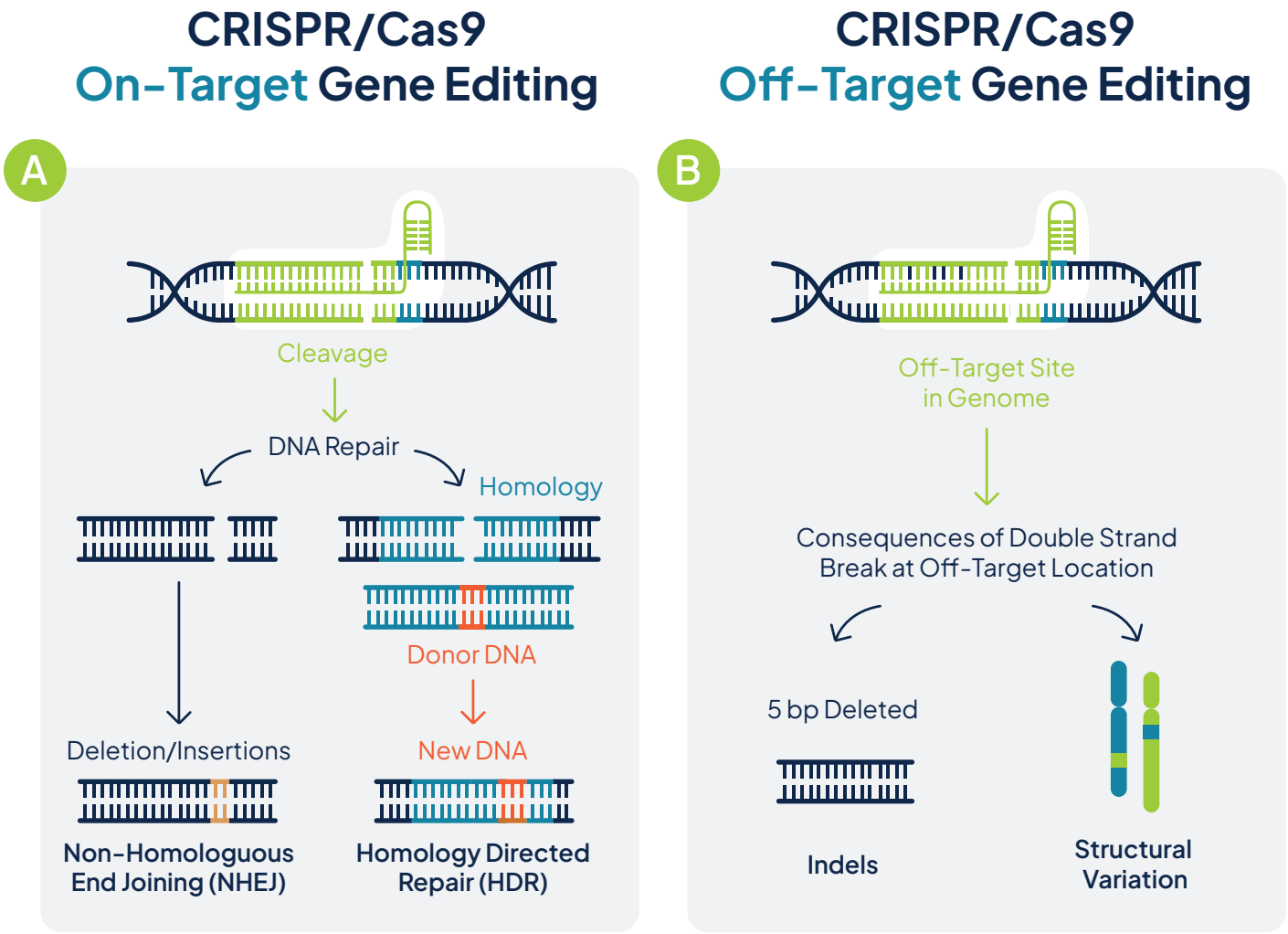

Figure 1. CRISPR/Cas9 mediated gene editing results in on-and off-target events. A) A guide RNA targets the Cas9 nuclease to a genomic locus of interest to generate double-strand breaks (DSBs). DSBs are then repaired by endogenous DNA-repair machinery. Error-prone repair pathways, such as non-homologous end joining (NHEJ), create random indels that can be frame-shifting and cause gene knockouts. Repair by homology-directed repair (HDR) can utilize an exogenous nucleotide donor molecule to create deletions of specified sequences, substitutions, or insertions of sequences. **B)** The Cas9 nuclease can tolerate mismatched bases between the guide RNA and genomic DNA and this can result in DSBs at off-target genomic locations. Following off-target DSBs, error-prone repair pathways can cause indels which can result in frame-shift mutations in a coding sequence, disruption of regulatory elements, and/or structural rearrangements.

can be re-joined by the cell's endogenous repair mechanisms in non-homologous end joining (NHEJ), with the loss of the intervening sequence, thus creating a knockout.[2] The break introduced by a single gRNA will also be repaired by the cell, but errors in the process introduce short, random insertions or deletions (indels), which in a coding sequence can cause a frameshift mutation.[2] CRISPR/Cas9 can also generate knock-ins, by using a single gRNA and an exogenous DNA donor consisting of the desired knock-in sequence flanked by sequences homologous to the insertion site. Homology-directed repair (HDR) will then result in the donor sequence being inserted into the cut site.[2] Additionally, deactivated Cas9 (dCas9), which preserves the site-specific binding capability of Cas9 while removing its endonuclease activity, may be fused to an activator or repressor domain and targeted upstream of a gene to up- or down-regulate its expression.[3]

These approaches all depend on the presence of endogenous 5'-NGG-3' sequences, referred to as protospacer adjacent motif (PAM). PAMs occur frequently throughout the genome, making CRISPR/Cas9 a versatile system for genetic engineering.[2] However, the frequency of these PAM sites, coupled with the imperfect specificity of Cas9, contribute to unintended consequences: off-target editing.

Off-target Cas9 activity results in off-target gene editing and mutations that vary between cells

Cas9 is not perfectly specific and exhibits significant tolerance of base pair mismatches between gRNA and targeted DNA (**Figure 1b**). Genomic sequences with single-nucleotide mismatches to gRNA may be cleaved by Cas9 with efficiencies similar to, or in some cases even higher than, cleavage to on-target sites.[4] Even off-target genomic sequences with up to five mismatched base pairs may be cleaved at a small but significant fraction of on-target efficiency.[4] As this constitutes a significant fraction of the approximately 20-nucleotide homology sequence of a gRNA, the human genome will typically contain thousands of potential off-target cleavage sites for a given gRNA.[4]

Off-target activity often results in small DNA indels.[5] Importantly, these off-target events can vary between individual cells. In single-cell whole genome sequencing following CRISPR/Cas9 genome editing of U2OS cells with a VEGFA gRNA, of the 11 off-target mutations observed, only two were recurrent across single cells.[6] Secondary analysis of data from this study also revealed that precise and comprehensive single-cell whole genome sequencing was able to identify hundreds of additional off-target indels that were not predicted by the off-target prediction algorithm Cas-OFFinder (**Figure 2**).[7] Off-target activity may also occasionally result in large-scale genomic aberrations such as translocations, inversions, deletions, and even chromosome fragmentation.[6, 7, 8] Deciphering these large scale genomic aberrations often requires techniques that examine the entire genome.

Most off-target sites will occur in noncoding regions, due simply to the higher proportion of these regions in the genome. However, mutations and genomic variation in non-coding regions can have dramatic effects on cell physiology and pathobiology. Specifically, indels and single-nucleotide variants (SNVs) in noncoding DNA have been implicated in numerous human diseases and in oncogenesis.[9, 10] Specific functional relationships between variants in noncoding regions and their associated diseases is poorly understood, but it is clear that some noncoding regions, such as lncRNAs, have roles in post-transcriptional gene regulation, and can stabilize, suppress, or promote translation of particular mRNAs.[11] These relationships can only be elucidated through integrated transcriptomic and genomic analysis on a single-cell level.

Characterizing CRISPR/Cas9-based cell and gene therapy products with single-cell genomics and transcriptomics

Gene editing using CRISPR/Cas9 is flexible and affordable; however, the extent of off-target gene editing and the consequences of these events remain poorly understood. Today there is greater scrutiny around characterizing unintended consequences of gene editing. For example, in the United States, the Food and Drug Administration (FDA) is creating guidelines for safety assessments when creating gene products that include human genome editing.[12]

While best practices are being developed by regulatory agencies, what features should investigators require in their assays and analysis plans?

Single-cell resolution. On-and off-target genomic edits vary between individual cells.[6, 7] Only through single-cell analysis can investigators successfully identify the correct on-target edit and catalog the unique off-target events of each individual cell.

Even allelic balance. On-and off-target genomic edits do not always occur in both copies of a gene (alleles). Only through assays demonstrating good allelic balance can investigators be confident in calling zygosity of both on-and off-target genomic edits.

High sensitivity and precision in mutation calling. Indels resulting from gene editing can vary in size, from a single base pair to many. Leveraging an assay with high precision and sensitivity enables researchers to confidently determine whether gene editing has created indels.

Uniform whole genome coverage. Off-target genomic edits occur throughout the genome. An assay with uniform genome coverage enables better detection of off-target mutations.

Bioinformatics analysis support. Single-cell whole genome sequencing has unique bioinformatics considerations. Identifying a collaborator or service provider with expertise in single-cell genomics is a critical aspect in streamlining data analysis.

With these considerations in mind, investigators require a comprehensive solution that enables them to analyze the entire genomes of single cells.

ResolveDNA from BioSkryb Genomics leverages Primary Template-directed Amplification (PTA) to enable precise and accurate single-cell whole genome sequencing. PTA chemistry, available only through BioSkryb Genomics in ResolveDNA and ResolveOME kits, overcomes biases, allelic dropout, low and variable genome coverage, poor reproducibility, and artifacts associated with existing whole genome amplification approaches.[6, 13, 14] This patented chemistry has been used to characterize the variability of on-target edits and catalog the number of off-target indels, copy number variations, and translocations that occur in individual cells following gene editing with CRISPR/Cas9 (**Figure 2**).[6, 7]

For investigators who want to move beyond inference of gene expression and define the mechanisms of gene expression on a single-cell level, BioSkryb Genomics offers ResolveOME.[13] ResolveOME provides single-cell genome and transcriptome amplification in a unified workflow. This approach combines PTA-mediated whole genome amplifications with full-transcript reverse transcription, allowing novel links to be drawn between genomic and transcriptomic changes.

For example, in acute myeloid leukemia cells with acquired resistance that exhibited increased *CEBPA* expression without genomic

copy number increases, ResolveOME enabled the identification of a candidate distal promoter/enhancer SNP approximately 20kb 5' of the *CEBPA* transcriptional start site.[13] Deploying ResolveOME in CRISPR/Cas9 gene edited cells could lead to similar discoveries, notably identifying which off-target mutations drive changes in gene expression and potentially cell type.

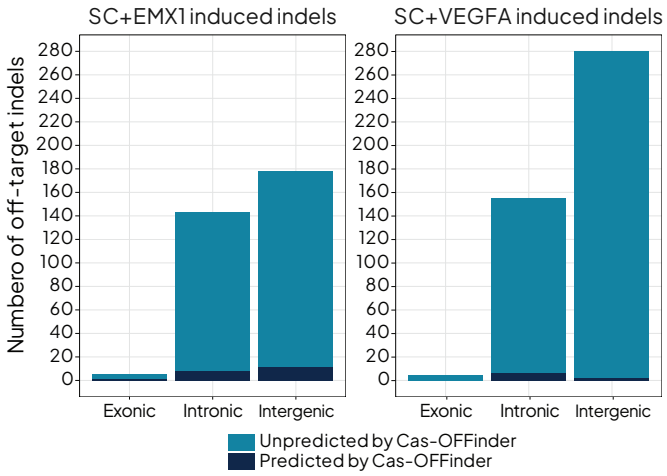

Figure 2: Indels detected by PTA-enabled single cell whole genome sequencing occur much more frequently than predicted by Cas-OFFinder. Secondary analysis of previously published data by Gonzalez-Pena et al.[6] reveals hundreds of indels detected following CRISPR/Cas9 gene editing with one of two guide RNA constructs (EMX1 or VEGFA). N = 15 total single cells per guide treatment comprised of 5 single cells each of CD34+ cord blood cells, U2OS cells, and H9 embryonic stem cells. Total number of indels detected in each group shown. Indels: insertions and deletions. SC: single cell. PTA: primary template-directed amplification.

These complex analyses require bioinformatics support. Supporting the analysis and interpretation of data generated with ResolveDNA and ResolveOME is BaseJumper®. BaseJumper is a bioinformatics platform built for biologists that enables multiomic data analysis. With this suite of products, along with custom service offerings through ResolveServices℠, BioSkryb Genomics provides solutions for more thorough characterization of CRISPR/Cas9 editing-based therapies.

To talk to a BioSkryb Genomics scientist about how ResolveDNA, ResolveOME, ResolveServices, and BaseJumper can empower research programs using CRISPR/Cas9 editing, email a member of our team at:

science@bioskryb.com

References

1. Zheng, R., Zhang, L., Parvin, R., Su, L., Chi, J., Shi, K., Ye, F., Huang, X. (2023). Progress and Perspective of CRISPR-Cas9 Technology in Translational Medicine. Advanced Science. https://doi.org/10.1002 advs.202300195
2. Barrangou, R., Sontheimer, E. J. Marraffini, L. A. (Eds.). (2022). CRISPR Biology and Applications. American Society for Microbiology.
3. Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. Nat Rev Mol Cell Biol. 2016 Jan;17(1):5-15. doi: 10.1038/nrm.2015.2.
4. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol. 2013 Sep;31(9):822-6. doi: 10.1038/nbt.2623.
5. Lin Y, et al. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Research, Volume 42, Issue 11, 17 June 2014, Pages 7473–7485, DOI: https://doi.org/10.1093/nar/gku402
6. Gonzalez-Pena, V., Natarajan, S., Xia, Y., Klein, D., Carter, R., Pang, Y., Shaner, B., Annu, K., Putnam, D., Chen, W., Connelly, J., Pruett-Miller, S., Chen, X., Easton, J., Gawad, C. (2021). Accurate genomic

variant detection in single cells with primary template-directed amplification. Proceedings of the National Academy of Sciences, 118(24). https://doi.org/10.1073/pnas.2024176118
7. Salas-Gonzalez I & Weigman V. (2023, Nov). Quantitative characterization of on-target and off-target variation induced by CRISPR+Cas9 systems at the single-cell resolution. [Poster presentation]. American Society of Human Genetics, USA.
8. Beeke W & Kyle CM. CRISPR nuclease off-target activity and mitigation strategies. Frontiers in Genome Editing, Vol 4. (2022). doi: 10.3389/fgeed.2022.1050507
9. Zhang F, Lupski JR. Non-coding genetic variants in human disease. Hum Mol Genet. 2015 Oct 15;24(R1):R102-10. doi: 10.1093/hmg/ddv259.
10. Deng G, Sui G. Noncoding RNA in Oncogenesis: A New Era of Identifying Key Players. International Journal of Molecular Sciences. 2013; 14(9):18319-18349. https://doi.org/10.3390/ijms140918319
11. Sabatino DE, Bushman FD, Chandler RJ, Crystal RG, Davidson BL, Dolmetsch R, Eggan KC, Gao G, Gil Farina I, Kay MA, McCarty DM, Montini E, Ndu A, and Yuan J. Evaluating the state of the science for adeno associated virus integration: An integrated perspective. Molecular Therapy Vol. 30 No 8 August 2022 DOI:https://doi.org/10.1016/j.ymthe.2022.06.004
12. US Department of Health & Human Services Food and Drug Administration. Human Gene Therapy Products Incorporating Human Genome Editing. March 2022. https://www.fda.gov/media/156894/download
13. Marks, J. R., Zawistowski, J. S., Salas-González, I., Tate, T. A., Morozova, T. v, Blackinton, J. G., Arvapalli, D. M., Velivela, S. D., Harton, G. L., Gawad, C., Weigman, V. J., Hwang, E. S., West, J. A. A. (2023). Unifying comprehensive genomics and transcriptomics in individual cells to illuminate oncogenic and drug resistance mechanisms. BioRxiv, 2022.04.29.489440. https://doi.org/10.1101/2022.04.29.489440
14. Data on file.

To talk to a BioSkryb Genomics scientist email:

science@bioskryb.com

Published by:

BioSkryb GENOMICS

2810 Meridian Pkwy, Suite 110, Durham, NC 27713

bioskryb.com