frontiers | Research Topics

Therapeutic gene correction strategies based on CRISPR systems or other engineered site-specific nucleases

Edited by
Ayal Hendel and Rasmus O. Bak

Published in
Frontiers in Genome Editing
Frontiers in Bioengineering and Biotechnology

ISSN 1664-8714
ISBN 978-2-8325-2345-2
DOI 10.3389/978-2-8325-2345-2

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Therapeutic gene correction strategies based on CRISPR systems or other engineered site-specific nucleases

Topic editors

Ayal Hendel — Bar-Ilan University, Israel

Rasmus O. Bak — Aarhus University, Denmark

Citation

Hendel, A., Bak, R. O., eds. (2023). *Therapeutic gene correction strategies based on CRISPR systems or other engineered site-specific nucleases*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2345-2

Table of contents

frontiers | Frontiers in Genome Editing

TYPE Editorial
PUBLISHED 20 April 2023
DOI 10.3389/fgeed.2023.1203864

OPEN ACCESS

EDITED AND REVIEWED BY
Chen Liang,
McGill University, Canada

*CORRESPONDENCE
Ayal Hendel,
ayal.hendel@biu.ac.il
Rasmus O. Bak,
bak@biomed.au.dk

RECEIVED 11 April 2023
ACCEPTED 11 April 2023
PUBLISHED 20 April 2023

CITATION
Hendel A and Bak RO (2023), Editorial: CRISPR and beyond: Cutting-edge technologies for gene correction in therapeutic applications.
Front. Genome Ed. 5:1203864.
doi: 10.3389/fgeed.2023.1203864

Editorial: CRISPR and beyond: Cutting-edge technologies for gene correction in therapeutic applications

Ayal Hendel[1]* and Rasmus O. Bak[2]*

[1]The Institute for Advanced Materials and Nanotechnology, The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat Gan, Israel, [2]Department of Biomedicine, Aarhus University, Aarhus, Denmark

KEYWORDS

gene editing, gene correction, CRISPR/Cas, HDR, nucleases, prime editing, base editing, perspective

Editorial on the Research Topic
Therapeutic gene correction strategies based on CRISPR systems or other engineered site-specific nucleases

Gene editing promises the ultimate cure for genetic diseases by directly correcting disease-causing variants. However, the first clinical trials have chased the "low hanging fruit" using editing strategies that rely on gene disruption by introducing double-strand DNA breaks that lead to insertions and deletions (indels) by the NHEJ pathway. Since NHEJ is constitutively active throughout the cell cycle and the default DNA repair pathway, this is by far the most efficient type of conventional gene editing as opposed to homology-directed repair (HDR). HDR relies on delivery of an exogenous repair template and this pathway is active only in the S and G2 phases of the cell cycle. These two parameters constitute challenges in clinical use of HDR since exogenous DNA is toxic in most therapeutically relevant cell types and since the inherent competition between NHEJ and HDR can be a bottleneck. However, HDR benefits from enabling precise edits to be made to the genome, thereby representing true gene editing with control over the outcome. Still, in both these modalities the DNA breaks are considered a potential source of genotoxicity due to the possibility of off-target edits and chromosomal aberrations such as translocations and chromothripsis. Next-generation gene editing tools like Base and Prime Editing that rely on DNA single strand nicking reduce the risk of such harmful events but are still limited in the scope of the edits they can generate (Anzalone et al., 2020). The newest types of editors based on CRISPR-associated transposases or CRISPR-directed integrases facilitate larger edits but are still under development and immature for clinical implementation (Yarnall et al., 2022; Tou et al., 2023). This rapidly developing toolbox is expected to broaden the application of CRISPR-based tools and other site-specific engineered nucleases to cure human disease. However, on this venture of realizing precise gene correction there are several unanswered questions and challenges to overcome, some of which we hope to address with this Research Topic on Therapeutic Gene Correction Strategies Based on CRISPR Systems or Other Engineered Site-specific Nucleases. This Research Topic covers a selection of contributions including significant scientific advances in precise genetic engineering as well as expert perspectives on recent advances.

The article by Lu et al. summarizes the recent improvements in nuclease and nickase genome editing approaches for the treatment of genetic diseases (Lu et al.). Furthermore, it highlights the major challenges in the translation of these approaches into clinical applications. The article by Reshetnikov et al. reviews the status of Base Editors for the correction of point mutations in monogenic diseases (Reshetnikov et al.). Cytosine and adenine Base Editors convert C•G to T•A and A•T to G•C, respectively and these tools are described in detail along thorough overviews on studies using these base editors *in vivo* and *in vitro*/*ex vivo* for monogenic retinal, neuromuscular, blood, neurodegenerative, and metabolic disorders.

Transcription activator-like effector domain (TALE) base editors (BEs) are a recent exciting addition to the genome editing toolbox. Boyne et al. in their original paper define the optimal TALE-BEs activity window and demonstrate the feasibility of efficient multiplex gene engineering using a combination of two different molecular tools, a nuclease and a base editor (Boyne et al.). Using such multiplex strategy has numerous important benefits, including better control over the editing outcomes by avoiding the occurrence of translocations that commonly arise when multiple nucleases are employed simultaneously, and it offers the opportunity to achieve more than just multiple knock-outs, as gene knock-ins can also be accomplished at the nuclease target site.

In extension of this, Wolff et al. introduces a new tool for high-efficiency Prime Editing termed piggyPrime (Wolff et al.). This tool makes use of the piggyBac transposon system to facilitate genomic integration of all the Prime Editing genetic components in cells to allow accumulation of prime edits over time leading to up to 100% of targeted alleles in some cell lines. This enables effective generation of transgenic cell lines to model disease-causing genetic variants.

Usher et al. similarly explores the generation of model cell lines carrying disease-causing variants but make use of conventional CRISPR/Cas HDR-based gene editing to install variants into the genome of cell lines (Usher et al.). By compiling data from 95 transfections, they compare HDR parameters such as donor template modifications, concentration, HDR enhancers, and cold shock. They also find that guideRNA efficiency prediction by online algorithms correlate poorly with activity in cells, and they present a workflow for designing and performing gene editing experiments to generate and characterize disease model clonal lines.

The articles by Houghton et al. and Ravendran et al. both describe gene editing approaches to fight monogenic inborn errors of immunity (IEI). Ravendran et al. reviews the DOCK8 immunodeficiency syndrome, which is a type of autosomal recessive hyper IgE syndrome caused by defects in the DOCK8 gene. The authors outline different genome editing strategies that might be applied to cure this devastating immunodeficiency syndrome (Ravendran et al.). Houghton et al. focus on another IEI, X linked lymphoproliferative disease (XLP), caused by mutations or deletions in the SH2D1A gene. The study compares the use of TALENs, CRISPR/Cas9, and CRISPR/Cas12a in combination with AAV6 repair template delivery (Houghton et al.). The components target exon 1 close to the start codon to facilitate integration of a near-full SH2D1A cDNA to be physiologically expressed and regulated by the endogenous promoter. The study shows integration frequencies around 30%-50% in T cells and that this restores SH2D1A gene expression and immune function in patient T cells to levels observed in healthy controls.

While genome editing is a very promising technology, it could, in theory, cause safety issues. Wienert and Cromer discuss the potential for unintended effects of CRISPR nuclease activity in human clinical trials (Wienert and Cromer). In their review they summarize the current sequencing-based solutions that may be able to detect these small and large-scale unintended genome editing effects even at low frequencies of occurrence. They highlight the safety and ethical concerns surrounding *in vivo* delivery of CRISPR tools and the potential for unintended editing in unintended cell types, which could enable germline transmission. Finally, they also outline some advanced potential mitigation strategies that will ensure that the safety of CRISPR keeps pace with its efficacy.

Schmidt et al. and Atkins et al. similarly cover additional important aspects of unintended on- and off-target editing outcomes. Schmidt et al. evaluate on- and off-target editing outcomes in CCR5 CRISPR-Cas9-targeted Mauritian cynomolgus macaque embryos using whole genome sequencing (WGS) analysis (Schmidt et al.). In this first report of WGS analysis of CRISPR-Cas9-targeted non-human primate embryonic cells they identify large deletions at the on-target site and *de novo* mutations at predicted CRISPR/Cas9 off-target sites. These data clearly demonstrate that comprehensive sequencing-based methods are warranted for evaluating editing outcomes in primate embryos and therefore also highlights the risks in human embryo editing. Atkins et al. examine in their review the strengths and limitations of the different classes of off-target cleavage detection techniques (Atkins et al.). Furthermore, they also discuss the clinical relevance of these techniques in the context of assessing the safety of novel CRISPR/Cas9 HIV curative strategies that are currently examined in clinical trials.

In summary, we hope that the original papers and reviews we had the privilege of including in our Research Topic will be a further step in bringing these therapeutic gene correction strategies to patients in need.

Author contributions

AH and RB wrote the editorial paper and approved the submitted version.

Conflict of interest

RB holds equity in Graphite Bio and UNIKUM Tx and is a part-time employee of UNIKUM Therapeutics.

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

Publisher's note

References

Anzalone, A. V., Koblan, L. W., and Liu, D. R. (2020). Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38, 824–844. doi:10.1038/s41587-020-0561-9

Tou, C. J., Orr, B., and Kleinstiver, B. P. (2023). Precise cut-and-paste DNA insertion using engineered type V-K CRISPR-associated transposases. *Nat. Biotechnol.* [Epub ahead of print]. doi:10.1038/s41587-022-01574-x

Yarnall, M. T. N., Ioannidi, E. I., Schmitt-Ulms, C., Krajeski, R. N., Lim, J., Villiger, L., et al. (2022). Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat. Biotechnol.* [Epub ahead of print]. doi:10.1038/s41587-022-01527-4

frontiers in Genome Editing

REVIEW
published: 17 August 2021
doi: 10.3389/fgeed.2021.673022

Off-Target Analysis in Gene Editing and Applications for Clinical Translation of CRISPR/Cas9 in HIV-1 Therapy

Andrew Atkins[1,2], *Cheng-Han Chung*[1,2†], *Alexander G. Allen*[1,2†], *Will Dampier*[1,2], *Theodore E. Gurrola*[1,2], *Ilker K. Sariyer*[3], *Michael R. Nonnemacher*[1,2] *and Brian Wigdahl*[1,2,4*]

[1] *Department of Microbiology and Immunology, Drexel University College of Medicine, Philadelphia, PA, United States,* [2] *Center for Molecular Virology and Translational Neuroscience, Institute for Molecular Medicine and Infectious Disease, Drexel University College of Medicine, Philadelphia, PA, United States,* [3] *Department of Neuroscience and Center for Neurovirology, Temple University Lewis Katz School of Medicine, Philadelphia, PA, United States,* [4] *Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA, United States*

OPEN ACCESS

Edited by:
Ayal Hendel,
Bar-Ilan University, Israel

Reviewed by:
Eric Paul Bennett,
University of Copenhagen, Denmark
Ashley Jacobi,
Integrated DNA Technologies, United States

***Correspondence:**
Brian Wigdahl
bw45@drexel.edu

[†] *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Genome Editing in Infectious Diseases, a section of the journal Frontiers in Genome Editing

Received: *26 February 2021*
Accepted: *21 June 2021*
Published: *17 August 2021*

Citation:
Atkins A, Chung C-H, Allen AG, Dampier W, Gurrola TE, Sariyer IK, Nonnemacher MR and Wigdahl B (2021) Off-Target Analysis in Gene Editing and Applications for Clinical Translation of CRISPR/Cas9 in HIV-1 Therapy. Front. Genome Ed. 3:673022. doi: 10.3389/fgeed.2021.673022

As genome-editing nucleases move toward broader clinical applications, the need to define the limits of their specificity and efficiency increases. A variety of approaches for nuclease cleavage detection have been developed, allowing a full-genome survey of the targeting landscape and the detection of a variety of repair outcomes for nuclease-induced double-strand breaks. Each approach has advantages and disadvantages relating to the means of target-site capture, target enrichment mechanism, cellular environment, false discovery, and validation of bona fide off-target cleavage sites in cells. This review examines the strengths, limitations, and origins of the different classes of off-target cleavage detection systems including anchored primer enrichment (GUIDE-seq), *in situ* detection (BLISS), *in vitro* selection libraries (CIRCLE-seq), chromatin immunoprecipitation (ChIP) (DISCOVER-Seq), translocation sequencing (LAM PCR HTGTS), and *in vitro* genomic DNA digestion (Digenome-seq and SITE-Seq). Emphasis is placed on the specific modifications that give rise to the enhanced performance of contemporary techniques over their predecessors and the comparative performance of techniques for different applications. The clinical relevance of these techniques is discussed in the context of assessing the safety of novel CRISPR/Cas9 HIV-1 curative strategies. With the recent success of HIV-1 and SIV-1 viral suppression in humanized mice and non-human primates, respectively, using CRISPR/Cas9, rigorous exploration of potential off-target effects is of critical importance. Such analyses would benefit from the application of the techniques discussed in this review.

Keywords: HIV-1, off-target, CRISPR, unbiased genome-wide, GUIDE-seq, CIRCLE-seq, DISCOVER-seq, whole genome sequencing

INTRODUCTION

Gene-editing strategies involving engineered nucleases [i.e., zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), meganucleases, and clustered regularly interspaced short palindromic repeat (CRISPR) associated nuclease 9 (Cas9)] have made a substantial impact on biological research and offer great therapeutic potential. While CRISPR/Cas9 is the most versatile of these systems it has also exhibited a propensity for off-target activity (Hockemeyer et al., 2011; Mussolino et al., 2011; Cradick et al., 2013; Fu et al., 2013; Hsu et al., 2013; Mali et al., 2013; Pattanayak et al., 2013; Yang et al., 2013; Cho et al., 2014; Lin et al., 2014; Zhang et al., 2014; Liang et al., 2015; Aryal et al., 2018). Understanding and mitigating the off-target activity in the clinical use of gene therapy is of particular importance because off-target effects may not be limited to transient events but may be pertinent to the lifetime of edited cells. Off-target detection methodologies are necessary because the functionality of gene-editing nucleases in general and the CRISPR system in particular are not fully understood. While some studies have indicated that CRISPR is more susceptible to unintended cleavage events than zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), the versatility of CRISPR targeting has rapidly made it the genome editing tool of choice (Deng et al., 2018; Huang et al., 2018; Panfil et al., 2018; Foss et al., 2019; Gao et al., 2019; Karimian et al., 2019; Li et al., 2019). There is data to suggest that the off-target proclivity of CRISPR guide RNAs (gRNAs) can be overcome with proper design considerations (Cho et al., 2014; Dampier et al., 2014, 2017, 2018; Kim et al., 2015; Aryal et al., 2018; Sullivan et al., 2019). Yet the stringent requirements of targeting fidelity that will be necessary to adapt CRISPR systems to their promising range of clinical applications demand a thorough, sensitive survey of the full genomic impact of each gRNA proposed for such applications. Importantly, sensitivity for off-target detection methods is presented as the minimum frequency of occurrence detectable in a cell population. For example, a method that can detect rare off-target events which occur in one out of 1,000 cells is described as having a sensitivity of 0.1%. Sensitivity is discussed in more detail in section Sensitivity.

The variety of published methods for off-target detection each attempt to improve upon earlier methods in some capacity. Trends in improvement include specificity, sensitivity, and throughput, as well as mechanistic considerations such as how off-target cleavage sites are detected and how those sites are enriched for deep sequencing. In this review we have organized our discussion of techniques based on the underlying mechanistic similarities of the assays. It is important to note however, that consideration of other methodological delineations is critical to a complete understanding of the field. In particular a distinction should be made between nomination and validation. Nomination of off-target sites can be achieved *in silico* based on sequence homology or empirically through experimentation. Nomination is important because broad survey of the full genome is necessary to identify where off-target cleavage may occur. Nomination thereby informs validation methods, which are necessarily site-specific, to confirm that off-target cleavage does in fact occur *in cellulo* and *in vivo*. Off-target sites which are validated are commonly referred to as *bona fide* off-target sites.

It is important to note that while this review focuses on underlying mechanism as the basis for grouping techniques for discussion, there is crossover in terms of detected outcomes and downstream utility for some techniques that are presented in separate sections.

Experimental observation of nuclease-induced off-target cleavage falls broadly into two categories termed: biased and unbiased methods. Biased methods make use of *a priori* knowledge to direct site-specific mutation detection and sequence validation to check for mutations at expected off-target sites, i.e., those with high sequence homology to the gRNA (Hsu et al., 2013; Doench et al., 2016; Tsai and Joung, 2016; Aryal et al., 2018). Conversely, unbiased methods are methods that survey the full genome for cleavage events allowing detection of off-target cleavage events independent of predictions (Koo et al., 2015; Tsai and Joung, 2016). While limited, biased techniques are often easier to implement, have a lower cost, or require minimal equipment. In some cases, the ability to rule out predicted, high-potential off-target sites may be enough for experimental purposes. Well-established biased techniques such as T7E1, Surveyor, and targeted amplicon sequencing are also important benchmarks by which newer methods are validated. In some cases, biased techniques generate data that cannot be captured otherwise. Uni-Directional Targeted Sequencing (UDiTaS) for example, requires a known primer site for target enrichment and is capable of detecting translocations, inversions, and large deletions that are missed by other methods (Giannoukos et al., 2018). With the development of the current range of unbiased techniques capable of surveying the full genome, methods relying on *a priori* knowledge play a smaller role. As genome editing becomes increasingly more precise, moving toward a variety of clinical applications, the need to efficiently survey the whole genome for RNA-guided-nuclease target-affinity precludes the use of biased techniques. Although a wide range of unbiased methods have been developed, there is still no clearly-superior, gold-standard technique (note: All off-target detection methods discussed in this manuscript are presented in **Table 1** with acronym disambiguation).

CRISPR/CAS9 TREATMENT OF HIV-1

Gene-editing strategies have the potential to make a significant impact on human immunodeficiency type 1 (HIV-1) treatment. Recent investigations into the application of the CRISPR/Cas9 system have shown potential in using it as a strategy for curing HIV-1 (Dampier et al., 2014, 2017, 2018; Hu et al., 2014; Kaminski et al., 2016a,b,c; Bella et al., 2018; Dash et al., 2019; Kaushik et al., 2019). HIV curative strategies are challenging because of the rapid establishment of a latent reservoir of infected cells (Siliciano and Greene, 2011). During latency, HIV-1 lies dormant and exhibits minimal expression of viral

TABLE 1 | Methods for detection of off-target CRISPR cleavage.

Method	References	Acronym disambiguation	Description
Surveyor	Guschin et al., 2010	Surveyor assay	Mismatch cleavage
T7E1	Kim et al., 2009	T7E1 assay	Mismatch cleavage
IDAA	Yang et al., 2015	Indel detection by amplicon analysis	DNA capillary electrophoresis
TIDE	Brinkman et al., 2014	Tracking of indels by decomposition	Indel frequency detection
TIDER	Brinkman et al., 2018	Tracking of insertions, deletions and recombination events	Mutation frequency detection
qEva-CRISPR	Dabrowska et al., 2018	Quantitative evaluation of CRISPR/Cas9-mediated editing	Mutation frequency detection
WGS	Smith et al., 2014; Suzuki et al., 2014; Veres et al., 2014; Iyer et al., 2015	Whole-genome sequencing	Whole genome sequencing
Digenome-seq	Kim et al., 2015	*In vitro* nuclease-digested genome sequencing	*In vitro*, genomic DNA cleavage, WGS
Multiplex Digenome-seq	Kim et al., 2016	Multiplex Digenome-seq	*In vitro*, genomic DNA cleavage, WGS
DIG-Seq	Kim and Kim, 2018	DIG-seq	*In vitro*, genomic DNA cleavage, WGS
SITE-Seq	Cameron et al., 2017	Selective enrichment and identification of tagged genomic DNA ends by sequencing	*In vitro*, genomic DNA cleavage
GOTI	Zuo et al., 2019	Genome-wide off-target analysis by two-cell embryo Injection	*In vivo* cleavage, WGS
In vitro selection	Pattanayak et al., 2013	*In vitro* selection with high throughput sequencing	*In vitro*, synthetic library
CIRCLE-seq	Tsai et al., 2017	Circularization for *in vitro* reporting of cleavage effects by sequencing	*In vitro*, genomic library
CHANGE-seq	Lazzarotto et al., 2020	Circularization for high-throughput analysis of nuclease genome-wide effects by sequencing	*In vitro*, genomic library
VIVO	Akcakaya et al., 2018	Verification of *in vivo* off-targets	*In vitro*, genomic library, *in vivo* validation
AMP	Zheng et al., 2014	Anchored multiplex PCR sequencing	Anchored-primer target enrichment
IDLV assay	Wang et al., 2015	Integrase-defective lentiviral vector assay	Anchored-primer target enrichment
GUIDE-seq	Tsai et al., 2015	Genome-wide, unbiased identification of DSBs enabled by sequencing	Anchored-primer target enrichment
iGUIDE	Nobles et al., 2019	Improvement of the GUIDE-seq method	Anchored-primer target enrichment
UDiTaS	Giannoukos et al., 2018	Uni-directional targeted sequencing	Anchored-primer target enrichment
TTISS	Schmid-Burgk et al., 2020	Tagmentation-based tag integration site sequencing	Anchored-primer target enrichment
TC-Seq	Klein et al., 2011	Translocation capture sequencing	Translocation enrichment
HTGTS	Chiarle et al., 2011	High-throughput, genome-wide, translocation sequencing	Translocation enrichment
LAM-PCR HTGTS	Frock et al., 2015; Hu et al., 2016	Linear-amplification-mediated-polymerase chain reaction high-throughput genome-wide translocation sequencing	Translocation enrichment
ChIP-Seq	Iacovoni et al., 2010; Kuscu et al., 2014; Wu et al., 2014; O'Geen et al., 2015	Chromatin immunoprecipitation sequencing	ChIP-seq
DISCOVER-Seq	Wienert et al., 2019, 2020	Discovery of *in situ* Cas off-targets and verification by sequencing	ChIP-seq
BLESS	Crosetto et al., 2013	Direct *in situ* breaks labeling, enrichment on streptavidin and next-generation sequencing	*In situ* end-capture
DSB-Seq	Baranello et al., 2014	Double-strand break sequencing	*In situ* end-capture
END-Seq	Canela et al., 2016	END-Seq	*In situ* end-capture
DSBCapture	Lensing et al., 2016	Double-strand break capture	*In situ* end-capture
BLISS	Yan et al., 2017	Breaks labeling *in situ* and sequencing	*In situ* end-capture
iBLESS	Biernacka et al., 2018	Immobilized-direct *in situ* breaks labeling, enrichment on streptavidin and next-generation sequencing	*In situ* end-capture

proteins which prevents the immune system from clearing the infection. The reservoir is primarily comprised of CD4+ T cells which can be localized to multiple tissues (Murray et al., 2016). Conventional antiretroviral therapy (ART) cannot remove these latently infected cells, which leads to continuous low levels of viral replication (Blankson et al., 2002). Elimination of HIV DNA from infected individuals remains a challenge in medicine.

There are two approaches to HIV-1 treatment using gene-editing nucleases: targeting the provirus in the latent reservoir and targeting host genes necessary for viral entry into cells. Targeting host genes involves targeting genes for CCR5 and CXCR4, either of which can serve as coreceptors allowing entry of the virus into cells (Hou et al., 2015; Xu et al., 2017; Allen et al., 2018). The goal of this approach is ablation of genes by introduction of insertions or deletions (indels) during endogenous repair processes following nuclease cleavage. Targeting the provirus can have two potentially beneficial outcomes, disruption of viral protein production by introduction of indels into proviral sequence during endogenous repair following nuclease cleavage (Liao et al., 2015; Zhu et al., 2015; Ueda et al., 2016; Wang et al., 2016a,b, 2018; Yoder and Bundschuh, 2016; Mefferd et al., 2018; Ophinni et al., 2018; Roychoudhury et al., 2018) or the excision of the provirus or parts of the provirus *via* simultaneous CRISPR/Cas9 cleavage at two target sites (Ebina et al., 2013; Dampier et al., 2014, 2017; Hu et al., 2014; Kaminski et al., 2016a,b; Yin et al., 2016; Bella et al., 2018). In the context of HIV-1 therapy, the long terminal repeat (LTR) has been identified as a promising target (Liao et al., 2015; Panfil et al., 2018). gRNAs designed to target the HIV-1 5′ (LTR), a region which acts as the HIV-1 promoter, can prevent HIV-1 reactivation by causing either transcriptional silencing or proviral excision because identical LTR sequences bookend the HIV-1 provirus (Kaminski et al., 2016a; Bella et al., 2018; Panfil et al., 2018). Additionally, this type of therapy could have the added benefit of targeting both replication competent and incompetent proviruses, which have the potential of generating viral proteins that are toxic to neighboring cells (Pollack et al., 2017; Baxter et al., 2018). Studies using HIV-1 transgenic mice and humanized mice models revealed that CRISPR-based editing resulted in removal of HIV-1 proviral DNA from several major tissues (Kaminski et al., 2016a; Bella et al., 2018). In another set of experiments, editing of HIV-1 proviral DNA by AAV-CRISPR constructs resulted in complete clearance of replication competent virus from ∼40% of animals after the cessation of ART (Dash et al., 2019). In a recent preclinical study, SIV-infected macaques, a well-defined non-human primate model of HIV/AIDS, were treated with AAV9-CRISPR/Cas9 editing constructs targeting LTR and Gag regions of SIV proviral DNA (Mancuso et al., 2020). Remarkably, fragments of integrated SIV proviral DNA were cleaved and removed from viral reservoirs including blood cells and lymphoid tissues leading to a reduction of proviral DNA.

While these observations provide a baseline for the potential use of a CRISPR-based gene editing strategy for the elimination of HIV-1 and a cure of AIDS, evaluation of potential off-target effects becomes highly significant and essential as the field moves closer to clinical translation. The remainder of this review will extensively discuss the landscape of off-target methods that exist today and are commonly used in the field. It will conclude with recommendations for properly assessing the safety of HIV-1 gene therapy.

EARLY TECHNIQUES FOR OFF-TARGET DETECTION ARE BIASED BY A NEED FOR *A PRIORI* KNOWLEDGE

The ability to determine off-target cleavage activity of the CRISPR/Cas9 system is crucial for clinical progression of gene editing. While there has been an influx of off-target sequencing assays developed, many publications rely on amplicon sequencing involving PCR amplification of nominated potential off-target sites followed by sequencing to identify off-target cleavage events in selected regions. This method relies on the use of bioinformatic tools to predict potential off-target sites for gRNAs. Using this knowledge, an investigator can extract genomic DNA from treated cells and amplify the regions that were predicted to have an off-target event. The amplified DNA is then checked for any insertion or deletion (indels) events that may have been caused by the CRISPR/Cas9 system. While effective for off-target site validation, genome-wide empirical nomination is still necessary for comprehensive evaluation of targeting specificity.

There are two methods that have risen in popularity to detect off-target events that still rely on PCR but use a different method of detecting indels. These two methods are called the Surveyor and T7E1 assays (Vouillot et al., 2015). In brief, these assays work by hybridizing two pieces of DNA together: an unaltered sample with one that has been mutated by Cas9 or other gene editing proteins. After hybridization of wild type and mutant DNA strands an enzyme is added that recognizes and cleaves bulges or mismatches in the DNA sequence. These enzymes come from bacterial species and are known as resolvases. Once the cleavage reaction has occurred, the digested DNA is run on an agarose gel and banding patterns and band intensities are used to quantitate the levels of gene editing. These assays do not handle single indels well, meaning that identification of a single nucleotide inserted or deleted by Cas9 can be difficult, and they offer no allelic discrimination with respect to editing events.

In order to detect indels, the method of Indel Detection by Amplicon Analysis (IDAA) is a simple yet effective technique which can detect indels with single base pair resolution (Yang et al., 2015; Carballar-Lejarazu et al., 2020). IDAA involves the amplification of potential nuclease cleavage sites using a three-primer amplification which generates fluorescently labeled amplicons. Detection of indels is achieved using DNA capillary electrophoresis. IDAA is considered a simple and effective method for indel detection and quantification of nuclease editing efficiency. Another way to resolve single indels utilizes bioinformatic tools that compare Sanger sequenced samples. One such tool is called tracking of indels by decomposition (TIDE) (Brinkman et al., 2018) and works by aligning unedited sequences with those that have been edited by Cas9. With the two abi trace files and the gRNA, the program finds where that particular gRNA would cleave the DNA and analyzes the peak heights from the chromatograph to determine if there has been an aberrant nucleotide inserted or deleted, indicating editing at that particular location. This tool has limitations when exploring multiple gRNAs and still requires

hand-tuning. In order to improve on some of the shortcomings of TIDE, a new tool was developed by Synthego called inference of CRISPR edits (ICE) [https://doi.org/10.1101/251082]. By utilizing techniques from the digital signal-processing field, it deconvolutes overlapping signals in the chromatograph allowing it to detect the composition and frequency of multiple editing events. This adaptation expands the utility to allow multiple gRNAs in a single experiment and rapid batch analysis.

Similar improvements to the TIDE methodology include tracking of insertions, deletions and recombination events (TIDER) and quantitative evaluation of CRISPR/Cas9-mediated editing (qEva-CRISPR) which both allow quantitation of mutation frequency, not limited to indels (Brinkman et al., 2018; Dabrowska et al., 2018). While these tools are mainly used to determine on-target events, they can also be used to measure off-target events. This requires a predictive knowledge of where these off-target events might occur and designing primers to those locations. This represents a serious drawback in the applicability of these tools to detect off-target events. The main reason behind this rationale is the need to design primers targeting suspected sites where Cas9 might bind and cleave. Next-generation sequencing (NGS) data from a number of different techniques has shown that using predicted cut sites will not uncover rare off-target events.

GENE-EDITING DESIGN

Precise Genome Editing Using RNA-Guided DNA Nuclease Systems

Following the initial discovery of CRISPR/Cas9 system, major adaptations were made to enable the system to work in human cells. These adaptations included: (1) the codon-optimized sequences of Cas9 that ensured preferable expression by the codon table used in the organism of interest (Cong et al., 2013; Jinek et al., 2013) (2) the attached nuclear localization signals (NLSs) to Cas9 to ensure the nuclear localization of Cas9 in human cells (Cong et al., 2013; Jinek et al., 2013), and (3) a single guide RNA (sgRNA), termed gRNA, constructed to possess both the guiding portion in crRNA and an RNA scaffold derived from tracrRNA (Jinek et al., 2012). These adaptations have enabled the CRISPR/Cas9 system to be programmable to any gene region by changing the protospacer sequence and flexible for use in any organism of interest (Hsu et al., 2014).

Unintentional Cleavage Events Mediated by CRISPR/Cas Nuclease

Evidence of high specificity using nuclease-based genome editing systems is critical for genetic screening in preclinical studies and corresponding transitional research. Since functional DNA is not comprised of random sequence due to evolutionary constraint, identical copies or highly homologous sequences to a designated target could exist in the same genome. Unwanted off-target editing and consequential toxicity has been demonstrated in the use of ZFNs and TALENs (Miller et al., 2007; Szczepek et al., 2007; Guo et al., 2010; Doyon et al., 2011). Soon after the engineered CRISPR/Cas9 was shown to work in human cells, off-target edits induced by CRISPR/Cas9 were addressed using systematic screening approaches (Fu et al., 2013; Hsu et al., 2013; Pattanayak et al., 2013; Qi et al., 2020). Using the Surveyor assay, Cong et al. (2013) showed that some gRNAs bearing up to five mismatches with target sequences induced CRISPR-mediated cleavage. Further experimentation showed that selected gRNAs could induce cleavage events at undesired off-target sites with up to 6 mismatches using the T7E1 assay in three human cell lines. In addition, they did not detect any off-target edits using two selected gRNAs individually at ∼50 tested potential sites each. Another study used synthetic oligomers to generate sequence libraries that contained 10^{12} potential off-target sites derived from the sequence of 4 gRNAs (Pattanayak et al., 2013). The results showed that the cleavage events occurred at synthetic off-target sequences with up to 7 mismatches against treated gRNAs, in agreement with previous studies, showing that incomplete complementarity still induced CRISPR-mediated edits. Together these studies suggested several important concepts: (1) the positions of mismatches affected off-target activity; the mismatches distal to PAM site were better tolerated than those proximal to PAM, (2) off-target edits could occur even with more than six mismatches between gRNA and off-target DNA, and (3) design of gRNAs without detectable off-target events is possible; RNF2 and FACNF gRNA caused no off-target mutations.

Predictive Algorithms for gRNA Selection

The use of computational predictive tools in gRNA design has developed rapidly to accommodate the increasing needs of CRISPR/Cas9 applications. In addition to identifying potential targets, computational tools for gRNA design must rate the exclusivity of those targets in order to avoid the use of gRNAs with off-target proclivity. The search can be as simple as mismatch counts between guide and target. However, recent approaches have adapted sophisticated algorithms into search tools. BLAST serves as the most accessible way to identify off-target sites on the basis of sequence similarity. However, the uniform penalty matrix in BLAST is not sufficient to describe guide-target interaction in the CRISPR/Cas9 system.

Two initial studies utilized similar strategies to characterize off-target activity due to sequence mismatches in the 20-bp complementary target region. One generated a set of gRNA variants that possessed gRNA variants that contained one mismatch against a fixed on-target DNA sequence in the human genome. Hsu et al. quantified the effect of mismatches by high-throughput sequencing of PCR amplicons from the on-target site after treatment of CRISPR/Cas9 (Hsu et al., 2013). Given a 20-bp target site, a set of gRNAs that covered all possible single-mismatch guide sequences were generated systematically such that 3 possible mismatch mutations at each complementary position were synthesized to acquire the contribution of CRISPR/Cas9 activity of each position. The modification frequency at the 20-bp complementary region was used to describe the CRISPR/Cas9 activity at the target site, which was determined by the number of reads that contained either mutations or indels from deep sequencing. The result indicated that the base pairing at the PAM-proximal region tolerated less

mismatches than the PAM-distal region. The authors aggregated sequence modification efficiency over 400 gRNA variants from 15 target sites within EMX1 gene regions, which created the pairwise penalty matrix for each type of mismatch spanning the guide-target binding region. A simplified 20-element matrix for a 20 bp guide-target pair was then used as the basis of the scoring algorithm for a gRNA design tool despite the type of mismatches. This score matrix is referred to as the MIT matrix.

Another experimental test using a larger number of gRNA variants demonstrated the improved prediction of potential off-target loci (Doench et al., 2016). The concept remained the same; given a fixed target DNA sequence, the reduction of CRISPR/Cas9 activity due to guide sequence mutations including 1-nucleotide mismatch, 1-nucleotide deletion, and 1-nucleotide insertion was measured. Over 27,000 unique gRNAs were generated to target the coding sequence of human CD33 regardless of PAM alternatives, along with the perfect match gRNAs for each selected target locus. This set of gRNAs gave a high coverage of every mutation type on each of the 20 guide-target base paring as well as every possible PAM. The goal of this experiment was to understand how CRISPR/Cas9 actively disrupts the expression of an easy-to-detect coding gene with or without guide sequence mutations against target DNA. Therefore, reduction of CD33 expression level on the plasma membrane was used to determine the CRISPR/Cas9 activity instead of deep sequencing. The percent activity was calculated by the mean differences of CD33 detected by phycoerythrin-conjugated anti-CD33 antibody between perfect gRNA and variant gRNA. A table of percent activity for every type of mutation (12 mismatch types × 20 positions and 64 possible PAMs) was used to generate the cutting frequency determination (CFD) score. The CFD score of a guide-target pair with multiple mutations is the multiplication of percent activity for specific mutations.

This data along with subsequent data generated from methods discussed below has led to vast increase in the number of computational techniques for predicting the likelihood of off-target cleavage. The range of computational tools for gRNA design includes E-CRISP (Heigwer et al., 2014), CRISPick (Doench et al., 2014), CHOPCHOP (Montague et al., 2014), CRISPR-ERA (Liu et al., 2015), CRISPOR (Haeussler et al., 2016), GUIDES (Meier et al., 2017), GeneArt (Liang et al., 2017), and uCRISPR (Zhang et al., 2019). More recently published tools tend to use the CFD matrix to evaluate penalty scores (i.e., CRISPOR, GUIDES, CRISPick, and GuideScan) and are therefore more reliable tools than those published before the development of the CDF matrix (i.e., E-CRISP, CHOPCHOP, CRISPR-ERA). More recently, the method uCRISPR has been shown to outperform methods using the MIT and CFD matrices (Zhang et al., 2019).

These tools have been reviewed previously and several publications offer a more in-depth review of this topic (Bolukbasi et al., 2016; Tycko et al., 2016; Manghwar et al., 2020; Wang et al., 2020). While the tools are useful for an initial discounting of egregious target choices, *in silico* predictions should always be confirmed by additional techniques.

UNBIASED TECHNIQUES

Whole Genome Sequencing Is a Feasible but Impractical Method for Off-Target Detection

Whole genome sequencing (WGS) is a straightforward approach to unbiased survey of the full genome for off-target nuclease activity. Endogenous repair mechanisms leave sequence-based evidence of nuclease activity on genomic DNA. Non-homologous end-joining (NHEJ) has been shown to introduce indels during the repair of double-strand breaks induced by nucleases (Cradick et al., 2013; Fu et al., 2013; Hsu et al., 2013; Pattanayak et al., 2013). Other repair outcomes for nuclease-induced DSBs include inversions, translocations, and large deletions (Frock et al., 2015; Hu et al., 2016; Giannoukos et al., 2018). Deep sequencing allows the identification of those repaired sites (**Figure 1A**). WGS ensures a survey of the full genome. There are several advantages to WGS as an off-target detection method. WGS allows an unbiased look at all sites across the genome and has been used to detect unpredicted off-target CRISPR/Cas9 cleavage in clonal cell populations and animal models (Smith et al., 2014; Veres et al., 2014; Dash et al., 2019). WGS detects the behavior of the nucleases in a cellular environment. The signatures of nuclease activity detected by WGS are introduced to the genomic DNA during endogenous repair processes. This is important because cellular features such as chromatin structure have been shown to impact the off-target profile of the CRISPR/Cas9 system (Kuscu et al., 2014; Wu et al., 2014; Chari et al., 2015; Chen et al., 2016, 2017c; Daer et al., 2017; Jensen et al., 2017; Kim and Kim, 2018; Chung et al., 2020). Furthermore, *in vitro* techniques for unbiased off-target detection have demonstrated that CRISPR/Cas9 cleaves more targets *in vitro* compared to targeting within the cellular environment thereby requiring further experimentation to validate the biological relevance of detected targets (Kim et al., 2015; Cameron et al., 2017; Tsai et al., 2017).

WGS has drawbacks though. It is considered inefficient due to a low signal to noise ratio. The vast majority of sequence data collected during WGS represents unedited genomic DNA and the depth-of-coverage for sequence locations of interest is sacrificed to the undisturbed regions. Thus, WGS is limited by throughput, cost, and efficiency compared to whole-genome methods which incorporate target enrichment strategies (e.g., GUIDE-seq) which are discussed in detail later. Nonetheless, the current efficiency of next-generation sequencing does enable this approach. In a study to detect off-target mutations in mice altered with Cas9, a reported 20–25 × depth of coverage was achieved for each sample as a single sequencing library using an Illumina HiSeq 2500 platform (Iyer et al., 2015). Results indicated that a sequencing depth of 10–13X was sufficient to detect 95% of homozygous variants. Other studies report between 33X and 50X coverage as necessary to detect single-nucleotide polymorphisms in human genomes (Bentley et al., 2008; Ajay et al., 2011). Exome sequencing has also been used to assess the targeting specificity of genome editing nucleases (Cho et al., 2014). In a study comparing whole genome sequencing to exome sequencing, the authors

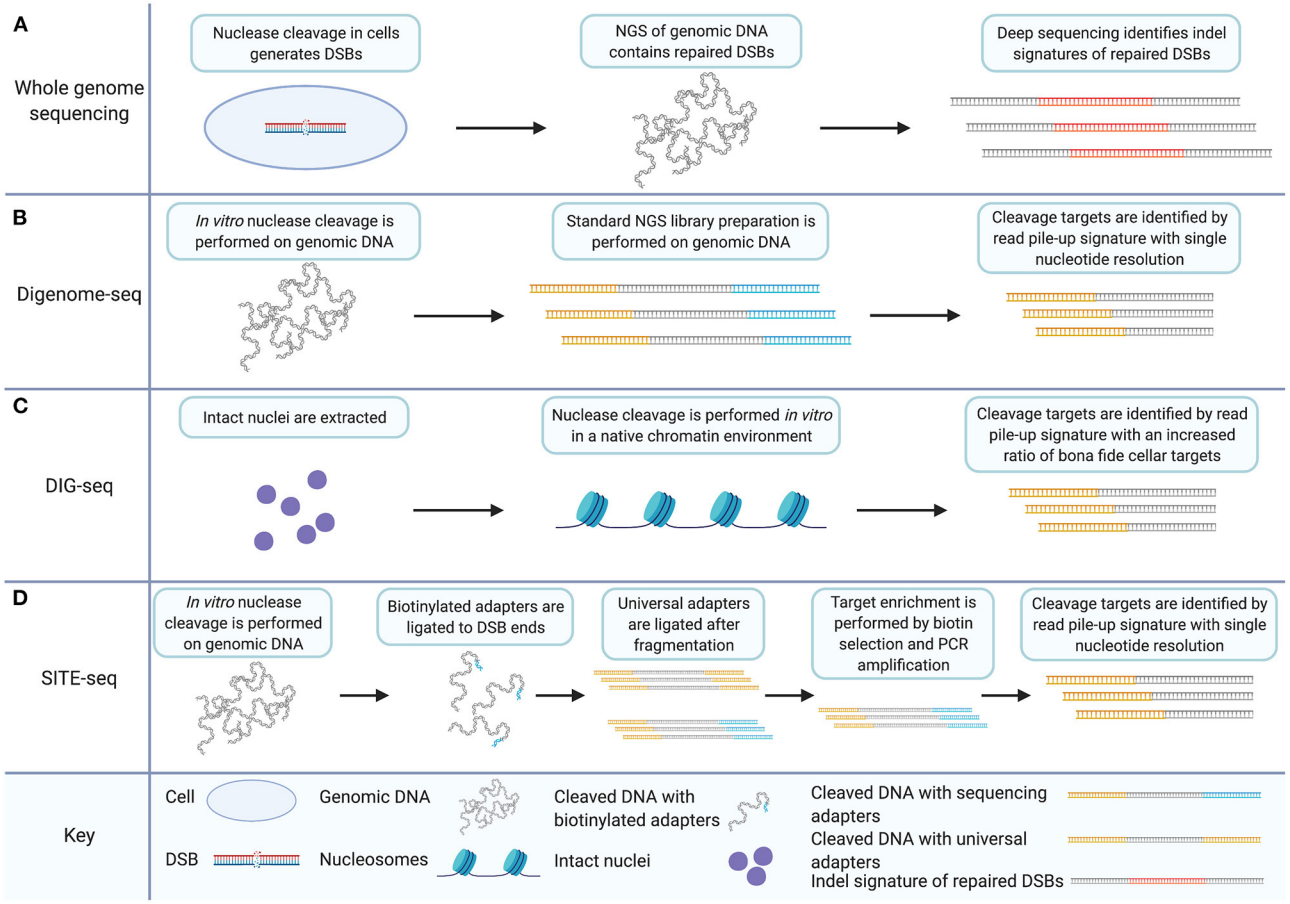


FIGURE 1 | *In vitro* off-target detection methods based on WGS utilize target enrichment strategies that facilitate data analysis and improve signal to noise ratios. **(A)** WGS for CRISPR/Cas9 off-target detection involves nuclease induction of DSBs in cells. Deep sequencing of genomic DNA captures all sequence including indels and genomic rearrangements introduced during DSB repair. **(B)** Digenome-seq involves extraction of genomic DNA and *in vitro* CRISPR/Cas9 treatment which precludes DSB repair. Samples are prepared by standard NGS methods where adapters (in yellow) are added. Cleavage targets are identified by a read pile-up signature with single-nucleotide resolution. **(C)** DIG-seq involves nuclei extraction. Nuclease cleavage is carried out *in vitro* in the native chromatin environment and libraries are prepared for NGS. Cleavage targets are identified by read pile-up, as in standard Digenome-seq. Intact chromatin reduces false discovery and increases the ratio of bona fide cellular targets identified. **(D)** SITE-Seq involves *in vitro* Cas9 digestion, as in Digenome-seq and DIG-seq. Biotinylated adapters (in blue) are then ligated to DSB ends followed by fragmentation and universal adapter ligation. Target enrichment is achieved by biotin selection and PCR amplification. As in Digenome-seq and DIG-seq, cleavage targets are identified by read pile-up during analysis. Created with Biorender.com.

conclude that there is no difference in cost effectiveness between the two approaches with respect to detection of known variants across the exome and that WGS produces better uniformity of read coverage. The results of that study show a mean on-target depth of coverage of 14 × to capture 95% of single-nucleotide variants (SNVs) (Meynert et al., 2014).

Modern methods of off-target detection deliver sensitivity on the order of 0.1% meaning that cleavage events which occur in 1 out of 1,000 cells are detectable (Frock et al., 2015; Kim et al., 2015, 2016; Tsai et al., 2015; Cameron et al., 2017; Yan et al., 2017; Kim and Kim, 2018; Wienert et al., 2019). The studies described above do not pinpoint the depth of coverage in WGS necessary to match genome-wide off-target detection methods which incorporate target enrichment strategies. Furthermore, the metrics reported are not directly comparable to off-target detection sensitivity. What those studies indicate is that WGS sensitivity can be variable depending on experimental conditions and sequencing platform and that exome sequencing does not confer an advantage in this strategy.

A recent study which applied WGS for detection of gene-editing outcomes has implemented a technique termed genome-wide off-target analysis by two-cell embryo injection (GOTI) (Zuo et al., 2019). To implement this method single blastomeres of two-cell mouse embryos were edited with CRISPR/cas9 or a base editor and progeny cells were examined by WGS for SNVs. The results of this study showed that CRISRP/Cas9-induced mutations were not carried through cell division, an important characterization of CRISPR/Cas9 effects. GOTI underscores that WGS still plays an important role for off-target detection in some experimental paradigms. Other off-target detection methods would have been unsuitable for collecting these results.

Despite genome-wide surveillance which made WGS a potential choice for off-target detection, recently published methods for unbiased survey of the whole genome offer greater sensitivity, fewer false-positives, and a better signal-to-noise ratio (Frock et al., 2015; Kim et al., 2015; Tsai et al., 2015, 2017; Cameron et al., 2017; Yan et al., 2017; Kim and Kim, 2018; Wienert et al., 2019). For example, the whole genome sequencing approach has been improved in the form of *in vitro* nuclease-digested genome sequencing (Digenome-seq). Digenome-seq enhances WGS performance as an unbiased off-target detection method (Kim et al., 2015, 2016; Kim and Kim, 2018).

Digenome-seq Enhances WGS Off-Target Detection by Inducing Cleavage *in vitro*

Digenome-seq is an unbiased, *in vitro* off-target cleavage detection technique (Kim et al., 2015). It introduces a change to the WGS approach by implementing nuclease cleavage outside of the cellular environment. Digenome-seq involves the *in vitro* digestion of genomic DNA using CRISPR/Cas9 and the gRNA to be evaluated. The digested genome is then prepared as an ordinary next-generation sequencing library. The alignment of fragment reads from nuclease cleavage sites is distinct from the staggered reads of other fragments because of the absence of sequence repair after nuclease cleavage. This is because endogenous DSBs occur at random locations while the targeted DSBs induced by nuclease cleavage occur at precise sequence locations. Nuclease cleavage sites are distinctly characterized therefore by repeated detection of DSBs at the same sequence location. Digenome-seq achieves target enrichment by introducing a distinct signature to nuclease cleavage targets which improves the resolution of cleavage detection to single-nucleotide precision (**Figure 1B**). This is not achievable using a WGS approach without the *in vitro* digestion of the genome due to the non-specific nature of the indels relied upon for detection (Kim et al., 2015).

There are several published improvements to the Digenome-seq technique. A multiplex version of Digenome-seq has been published, allowing the testing of multiple gRNAs on the same sample simultaneously (Kim et al., 2016). The multiplex method has several modifications. The algorithm used for analysis was modified to allow the identification of cleavage events that leave different end moieties, specifically one or two nucleotide overhangs; the original algorithm only detected blunt ends. This modification reduced false-negatives and identified targets missed by the original Digenome-seq algorithm. False positives were also reduced compared to Digenome-seq by transcribing gRNAs with a plasmid template rather than an oligonucleotide. Plasmid transcripts were reportedly less heterogenous than oligonucleotide transcripts leading to higher fidelity in target recognition. Multiplex analysis of gRNAs was achievable in the Digenome-seq methodology by choosing gRNAs with target sequences differing by at least 11 nucleotides and thus avoiding ambiguity in target detection within the same sample. Multiplex Digenome-seq results were achieved without an increase in depth of coverage. These results demonstrate not only the ability of the technique to detect off-target cleavage from multiple gRNAs simultaneously but also the ability of Cas9 to be directed to multiple targets *in vitro* by multiplexed gRNAs.

Measures of improvement to off-target detection techniques can depend on the specific measurement goals. The unfettered nature of Digenome-seq with respect to chromatin architecture can be viewed as an advantage compared to WGS or techniques such as genome-wide, unbiased identification of DSBs enabled by sequencing (GUIDE-seq) and high-throughput genome-wide translocation sequencing (HTGTS) (Frock et al., 2015; Tsai et al., 2015). GUIDE-seq and HTGTS are mentioned here to make a point of contrast compared to Digenome-seq; both will be discussed in detail in later sections. The distinction allows for the detection of otherwise obscured gRNA off-target affinities. However, DIG-seq, another Digenome-seq modification, can also be considered an improvement of the Digenome-seq method for the opposite reason (Kim and Kim, 2018). DIG-seq is a Digenome-seq based method applied to DNA with chromatin architecture in place. Native chromatin is isolated *via* nuclei extraction and put through the Digenome-seq protocol (**Figure 1C**). DIG-seq is considered an improvement over Digenome-seq under the assumption that the cleavage targets that will be detected under these conditions are of keener interest and greater relevance than the full palate of *in vitro* detected cleavage targets outside of the chromatin architecture. This assumption is upheld by the performance of DIG-seq. DIG-seq performance was compared to two other *in vitro* off-target detection methods: selective enrichment and identification of tagged genomic DNA ends by sequencing (SITE-Seq) and circularization for *in vitro* reporting of cleavage effects by sequencing (CIRCLE-seq), discussed in detail below. Although identifying fewer off-target cleavage sites than CIRCLE-seq or SITE-Seq for the same VEGFA target, DIG-seq had a 62% deep sequencing validation rate compared to 29 and 10% validation rate for the other two techniques, respectively.

SITE-Seq Improves Digenome-seq Methodology With Selective Target Enrichment

SITE-Seq is an unbiased, *in vitro* detection technique for nuclease-induced DSBs (Cameron et al., 2017). SITE-Seq involves the *in vitro* digestion of genomic DNA with CRISPR/Cas9, similar to Digenome-seq. Following 3′ adenylation, DSBs are ligated with biotinylated Illumina-compatible adapters. This leaves a pool of labeled DSBs in genomic DNA, predominately induced by gRNA-guided nuclease cleavage, which allows the selective enrichment of sequence surrounding cleavage sites. Following the initial labeling of the strand break sites, the genomic DNA is fragmented, end-repaired, and 3′ adenylated allowing for another round of Illumina-compatible adapter ligation. Thus, fragments containing sequence from one side of a DSB are exclusively bookended by the P5 and P7 binding-sites necessary for Illumina sequencing. Biotin selection and PCR amplification then lead to a selectively enriched deep-sequencing library comprised predominately of sequences surrounding nuclease cleavage targets.

Similar to Digenome-seq, the technique relies on *in vitro* digestion of genomic DNA and nuclease cleavage targets are distinguished from randomly induced DSBs during sequence analysis by aligned read pileups. SITE-Seq differentiates itself from Digenome-seq particularly by the selective enrichment of nuclease-cleavage targets (**Figure 1D**). This aspect considerably increases the signal-to-noise ratio of the readout compared to Digenome-seq.

SITE-Seq is highly sensitive, around 0.1%. SITE-Seq analysis of the commonly used controls VEGFA and FANCF detected nearly all of the sites identified by Digenome-seq, GUIDE-seq, and (HTGTS). SITE-Seq reportedly detected all previously identified cellular off-targets from preassembled Cas9-gRNA ribonucleoprotein (RNP). Although Digenome-seq sensitivity is equivalent to SITE-Seq, the signal to noise ratio of SITE-Seq is far greater due to the process of enrichment, which allows sequencing of cleavage sites while excluding the remainder of the genomic DNA. However, SITE-Seq shares the problem of a high false-discovery rate with CIRCLE-seq and Digenome-seq. Cellular factors play a role in the off-target activity of nucleases. *In vitro* techniques identify potential off-target sites in the absence of such factors and the sheer quantity of potential sites can inhibit validation of relevant bona fide sites. For example, SITE-Seq identified nine novel off-target sites for VEGFA and two for FANCF in spite of limiting cellular validation to a subset of identified sites. This is touted as a feature in this instance, and it is a good demonstration of the sensitivity of the method. But if the identified potential off-targets for a particular gRNA are too numerous to be efficiently screened for cellular activity, validation of off-target sites becomes a biased technique in spite of the unbiased nature of the assay. A further complication is that the effect of cellular factors and nuclease concentration on off-target cleavage may limit the relevance of validation to the experimental conditions under which it is carried out.

IN VITRO CLEAVAGE LIBRARIES

CIRCLE-seq Is a Highly Sensitive Off-Target Detection Method That Brings Genomic Relevance to *in vitro* Cleavage Libraries

CIRCLE-seq is an unbiased method for detection of off-target CRISPR/Cas9 cleavage (Tsai et al., 2017; Lazzarotto et al., 2018). The method entails fragmentation of genomic DNA *via* sonication, end-repair, and self-ligation of fragments for intramolecular circularization. After circularization, remaining linear DNA is digested using a plasmid-safe ATP-dependent DNase. What remains is a library of circularized fragments of genomic DNA which is then digested using CRISPR/Cas9 and an gRNA to be profiled for off-target affinity. During Cas9 digestion, circles containing on-target and off-target sequence are linearized and are then prepared for next-generation sequencing.

CIRCLE-seq was adapted from earlier published *in vitro* methods for characterizing the off-target profiles of genome-editing nucleases (Pattanayak et al., 2011, 2013). An *in vitro* selection method was introduced to characterize the performance of two zinc-finger nucleases (ZFNs) on a library of 10^{11} sequences. ZFNs targeting human genes for CCR5 and VEGFA were used. VEGFA has become a standard control for evaluation of genome-editing nucleases (Frock et al., 2015; Kim et al., 2015, 2016; Tsai et al., 2015, 2017; Cameron et al., 2017; Yan et al., 2017; Kim and Kim, 2018; Wienert et al., 2019). Both ZFNs were able to cleave target numbers on the order of 10^5 sequences, the majority of which do not arise in the human genome. CCR5-224 also cleaved 37 *in vitro* human sequence targets 10 of which were validated in human K562 cells. The VEGFA-targeting ZFN, VEGFA2468, cleaved 2652 human sequence targets *in vitro*, 32 of which were validated in human K562 cells (Pattanayak et al., 2011).

In a subsequent study, the previous *in vitro* library method for ZFNs was modified to measure CRISPR/Cas9 off-target capacity on an *in vitro* library of 10^{12} sequences (**Figure 2A**) (Pattanayak et al., 2013). Between two gRNAs tested, five off-target human sequences were validated in HEK293T cells. Both ZFNs and the CRISPR/Cas9 system, were shown to exhibit off-target specificity dependent on enzyme concentration with some rare off-target cleavage events occurring only at higher enzyme concentrations (Pattanayak et al., 2011, 2013).

CIRCLE-seq is a further adaptation of the *in vitro* library off-target cleavage detection method (**Figure 2B**). Generating the *in vitro* sequence library from genomic DNA increases the relevance of the library of identified cleavage targets. Additionally, because of the mechanism of cleavage-detection in CIRCLE-seq, each readable fragment contains the sequence from both sides of a given cleavage-site allowing for reference-genome-free off-target sequence identification with single nucleotide resolution. Earlier *in vitro* library methods detected significant background sequence noise, with hundreds of thousands of *in vitro* cleavage targets that are not relevant to the human genome. CIRCLE-seq by contrast, finds only human-genome sequence targets.

At the time of initial publication, CIRCLE-seq was the only unbiased, *in vitro* alternative to Digenome-seq and in some facets of performance CIRCLE-seq exceeds Digenome-seq. In particular, CIRCLE-seq has 180,000-fold higher signal-to-noise ratio than Digenome-seq. CIRCLE-seq owes this increase to the process of enrichment which ensures that only cleavage-target sequences are prepared for deep sequencing. There is however a trade-off between the CIRCLE-seq and Digenome-seq techniques in terms of resource consumption as each CIRCLE-seq sample requires 25 μg of genomic DNA while each Digenome-seq sample requires 1 μg. The high background noise in Digenome-seq can make the identification of rare targets difficult, and it has been suggested that some valid off-target cleavage sites are missed by Digenome-seq because of the filtering thresholds necessary to process excessive background signal (Kim et al., 2015; Tsai et al., 2017). CIRCLE-seq is reportedly more sensitive than Digenome-seq. The error rate of current next-generation sequencing (∼0.1%) is the limiting factor in the detection of rare off-target cleavage events. Both techniques directly detect cleavage events with single nucleotide resolution which is not common to all off-target detection methods (Frock et al., 2015; Tsai et al., 2015).

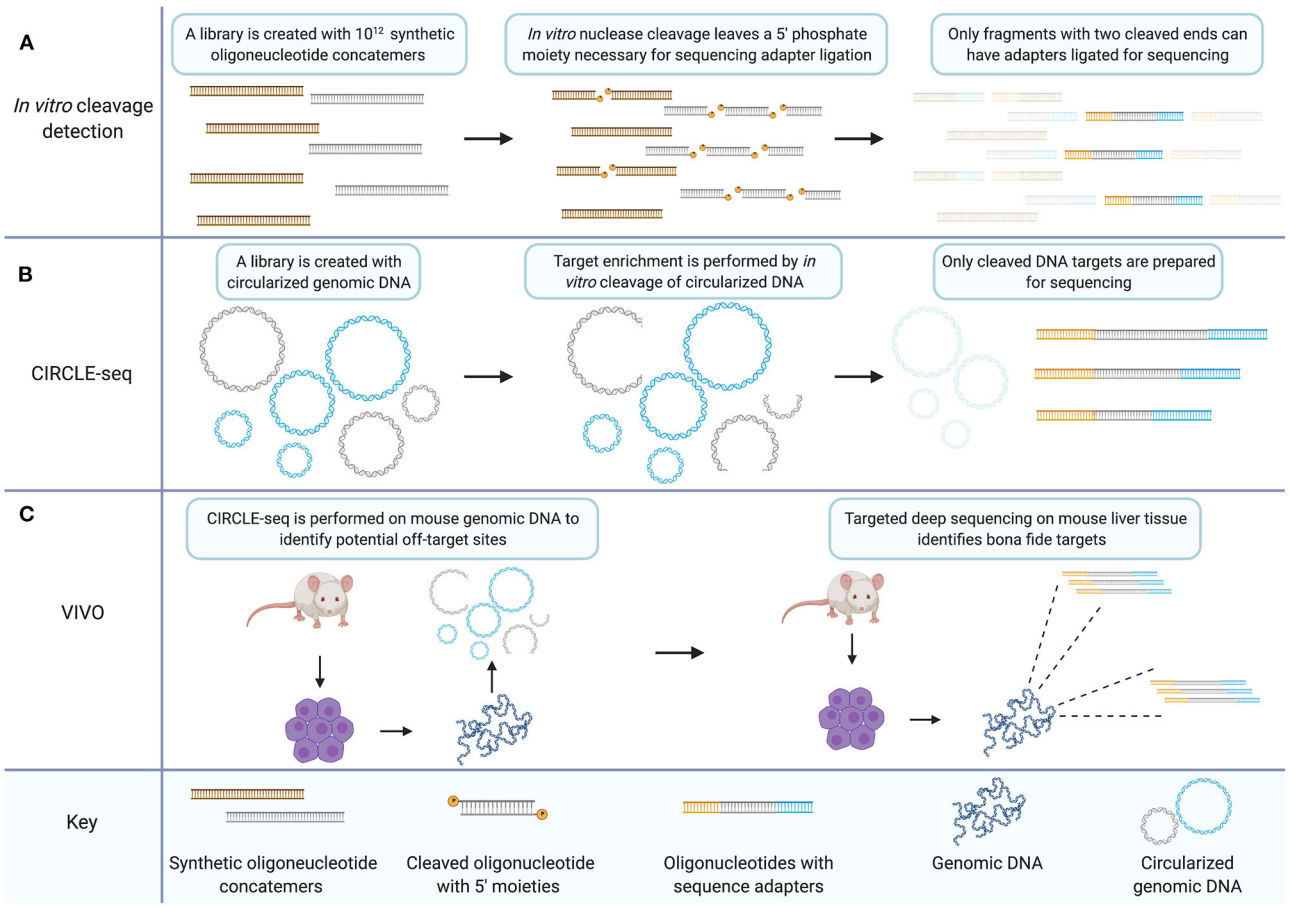


FIGURE 2 | The CIRCLE-seq target enrichment strategy creates an *in vitro* library of CRISPR/Cas9 targets with biological relevance that can inform targeted sequencing for validation of bona fide off-target sites. **(A)** *In vitro* cleavage detection involves generating a library of synthetic oligonucleotide concatemers with potential mismatch targets for a CRISPR gRNA. Target cleavage leaves a 5′ phosphate necessary for ligation of sequencing adapters. Only fragments which have been cleaved at both ends because the target site within the concatemer was recognized by the nuclease two times, are prepared for sequencing. **(B)** CIRCLE-seq involves generation of a circularized *in vitro* library created from genomic DNA. Unrecognized targets remain circularized (in blue) while cleaved target sequences are linearized and can be prepared for sequencing (in gray). **(C)** VIVO involves performing CIRCLE-seq with genomic DNA from mouse primary cells. The CIRCLE-seq results inform targeted deep sequencing with genomic DNA from mouse primary cells derived from CRISPR/Cas9 treated mice thereby validating bona fide off-target sites. Created with Biorender.com.

Recently an updated version of the CIRCLE-seq methodology has been published (Lazzarotto et al., 2020). The modified technique is called circularization for high-throughput analysis of nuclease genome-wide effects by sequencing (CHANGE-seq). CHANGE-seq utilizes a tagmentation reaction in early steps of the protocol which drastically reduces the labor and preparation time for this methodology. Compared to CIRCLE-seq, CHANGE-seq allows more rapid sample processing for higher-throughput experiments and will likely be the preferred method for any experiment using *in vitro* library digestion in the future.

In comparison to cell-based methods for unbiased off-target detection, *in vitro* methods boast some attractive features. *In vitro* methods avoid the need for transfection, which can complicate both inter- and intra-experimental comparisons. Also *in vitro* detection does not rely on endogenous repair pathways like WGS, GUIDE-seq, and HTGTS (Bentley et al., 2008; Ajay et al., 2011; Meynert et al., 2014; Veres et al., 2014; Frock et al., 2015; Iyer et al., 2015; Tsai et al., 2015). GUIDE-seq and HTGTS are mentioned here to make a point of contrast compared to CIRCLE-seq; both will be discussed in detail in later sections. However, *in vitro* techniques also do not give insight into the behavior of gene-editing nucleases in cells. The false positive rate for CIRCLE-seq is reportedly low enough that the sensitivity limits of deep sequencing inhibit its estimation. However, the false discovery rate is high in CIRCLE-seq meaning that CIRCLE-seq frequently identifies off-target sites *in vitro* that are not validated in cellular experiments.

VIVO Utilizes CIRCLE-seq to Identify Deep Sequencing Targets for Validation *in vivo*

The standard for validation of bona fide off-target sites is targeted deep sequencing. A method has been published that is termed verification of *in vivo* off-targets (VIVO) which consists of CIRCLE-seq to identify off-target candidate sites followed by targeted deep sequencing to validate those sites (**Figure 2C**) (Akcakaya et al., 2018). This hybrid technique constitutes a method for validating off-target sites *in vivo* in an animal model. Candidate sites were examined which were identified by CIRCLE-seq in the livers of mice treated with CRISPR/Cas9 in adenoviral vectors using targeted deep sequencing. To do so, they chose a subset of sites from three classes of off-target sequences that they delineate by high, moderate, or low CIRCLE-seq read counts. Results indicate that the probability of validating off-target sites is higher amongst sites that return higher CIRCLE-seq read counts. This agrees with the findings of the originally published CIRCLE-seq method which show that sites with higher CIRCLE-seq read-counts are more likely to be detected by the cell-based method GUIDE-seq (Tsai et al., 2015). Although CIRCLE-seq data sets provide an unbiased genome-wide survey of off-target proclivity for CRISPR/Cas9 gRNAs, the sheer volume of potential off-target sites limited the validation of sites in the VIVO study to a subset of candidates, essentially a biased analysis. Importantly though, off-target sites were validated across all classes in the VIVO study, i.e., high, moderate, and low CIRCLE-seq read counts, underscoring the need for comprehensive analysis of gene-editing nuclease targeting particularly with respect to therapeutic development.

ANCHORED PRIMER ENRICHMENT

GUIDE-seq Combines the Principles of AMP and IDLV With Improved Off-Target Detection Performance

GUIDE-seq is a method for tagging and enriching the sequence surrounding DSBs for deep sequencing (Tsai et al., 2015). Originally published in 2015, the technique remains an important methodology for assessing the targeting fidelity of genome-editing nucleases (Chaudhari et al., 2020). Briefly, cells are transfected with a plasmid coding for Cas9 and a gRNA and co-transfected with a blunt, double-stranded oligodeoxynucleotide (dsODN). The dsODN is then incorporated into DSBs during NHEJ, thus tagging DSB sites with a short, known sequence. Extracted genomic DNA is then fragmented enzymatically or *via* sonication and the resulting fragments undergo end-repair, dA-tailing, and ligation of a universal adapter sequence which is added to both ends of all fragments. Target enrichment is achieved by two rounds of PCR which amplify only fragments containing the dsODN. Thus, the amplified library consists of strands which each contain one half of the sequence surrounding a DSB repaired by NHEJ. GUIDE-seq is conceptually derived from earlier methods. Precursors to GUIDE-seq include anchored multiplex PCR (AMP) and integrase-defective lentiviral vector (IDLV) integration (Gabriel et al., 2011; Zheng et al., 2014; Wang et al., 2015).

AMP is a target enrichment method for deep sequencing applications. Early target enrichment methods include AmpliSeq, TruSeq Amplicon, HaloPlex, and Nested Patch PCR (Varley and Mitra, 2008; Johansson et al., 2011; Do et al., 2013; Yousem et al., 2013). AMP improves on these techniques by enriching targets with only one known primer binding site rather than two (**Figure 3A**). In principle, AMP resembles a much earlier method called rapid amplification of cDNA ends (RACE) which utilizes known DNA sequence to determine the sequence of an adjacent region (Frohman et al., 1988). AMP involves preparation of double-stranded cDNA or sheared genomic DNA using earlier published methods (Zheng et al., 2010, 2011; Neiman et al., 2012). Following end-repair and dA-tailing, sequencing adapters, called universal half-functional adapters, are ligated randomly to the ends of all fragments. Enrichment is accomplished by PCR amplification using anchored primers for known targets. Primers for a second round of PCR are 5′-tagged with sequencing adapters. The resulting libraries have a fully functional pair of adapters for deep sequencing. This results in the selective amplification of targets with only one known primer binding site. Unknown adjacent sequence is then captured, and genomic rearrangements can be identified following deep sequencing.

Detection of IDLV integration has been used to identify on- and off-target cleavage of ZFNs, TALENs, and CRISPR/Cas9 (Gabriel et al., 2011; Wang et al., 2015) IDLV detection takes advantage of the IDLV capability to integrate into DSBs during NHEJ. Integration tags break-sites with known sequence which can be exploited for target enrichment (**Figure 3B**). Targets are amplified for sequencing by linear amplification-mediated (LAM) PCR or non-restrictive LAM (nrLAM) PCR (Schmidt et al., 2007; Gabriel et al., 2009; Paruzynski et al., 2010). IDLV has shortcomings including a low rate of integration and the tendency of IDLVs to sometimes integrate at sites up to 120 bp from the target DSB site (Gabriel et al., 2011; Tsai et al., 2015).

GUIDE-seq technology is a significant advancement over its predecessors. AMP allows the selective amplification of sequence with one side known which was an important step forward from earlier PCR techniques requiring two known primer sites. GUIDE-seq allows selective amplification of a target sequence in which no portion is known by placing the anchor primer on the dsODN (**Figure 3C**). This is essentially the principle behind IDLV detection but the more reliable rate of uptake of the dsODN into DSBs and the precise integration between the two ends of the DSB mark GUIDE-seq as a significant advance over IDLV.

At the time of publication GUIDE-seq set a new benchmark for off-target detection of nuclease-induced DSBs by filling a methodological gap for unbiased survey of the full genome with an effective target enrichment strategy that greatly improved the signal to noise ratio of off-target detection methods utilizing deep sequencing. GUIDE-seq has a detection sensitivity of ∼0.12%, equivalent to that of other current methods (Kim et al., 2015,

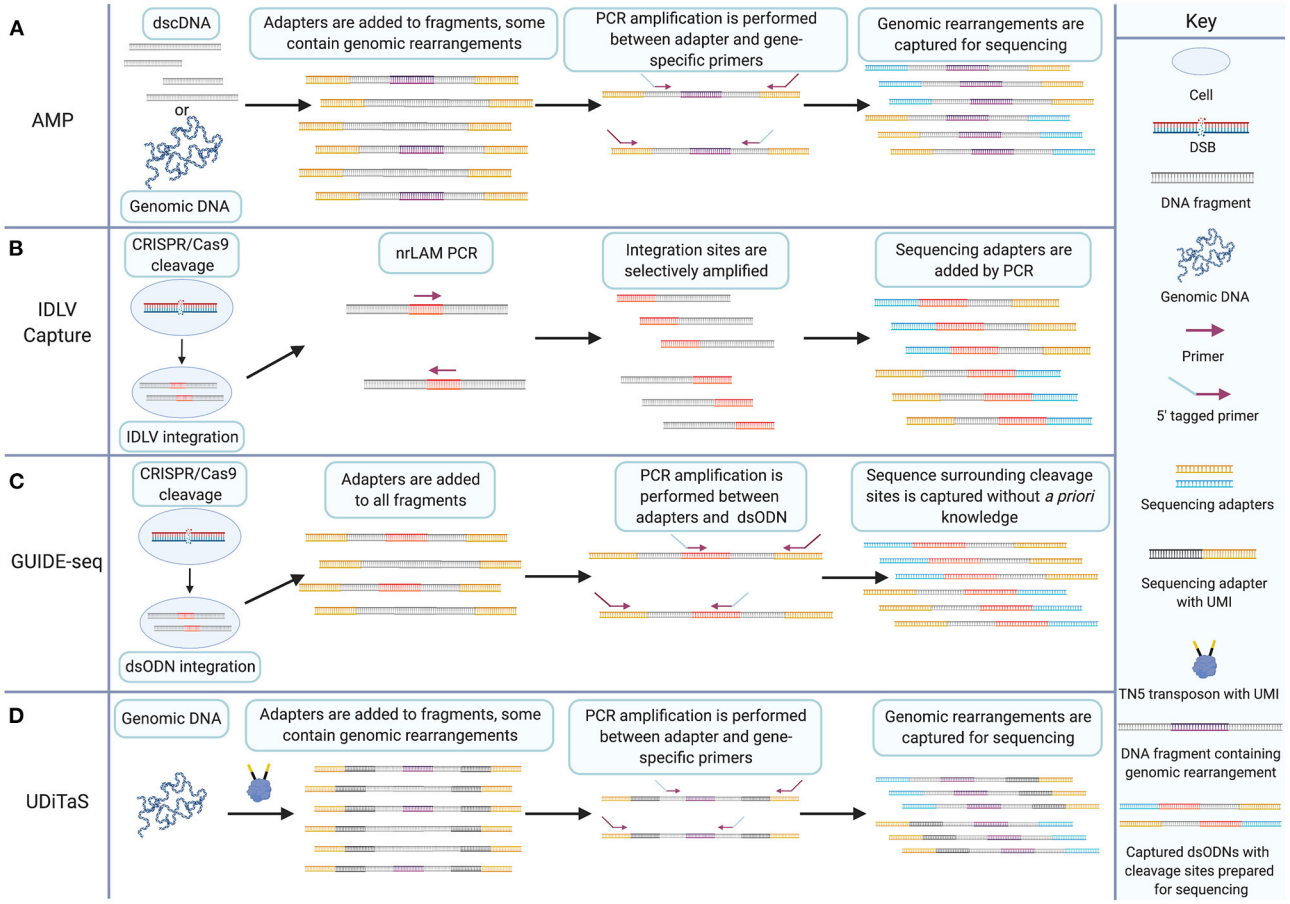


FIGURE 3 | The GUIDE-seq target enrichment strategy combines IDLV capture and AMP for CRISPR/Cas9 cleavage detection without *a priori* knowledge. **(A)** AMP involves adding half-functional adapters (shown in yellow) to both ends of double-stranded cDNA or sheared genomic DNA. Fragments may contain genomic rearrangements that are amplified by PCR between a single anchored gene-specific primer site and a half-functional adapter. 5′ tags on primers enable addition of a second sequencing adapter to amplified target sites. Non-target sites do not get the additional adapter and are excluded from sequencing. **(B)** IDLV capture involves transfection of CRISPR/Cas9 and transduction of IDLV which is integrated into CRISPR/Cas9-induced DSBs during NHEJ, shown in red. nrLAM PCR selectively amplifies cleavage sites from the integrated sequence. Additional rounds of PCR add sequencing adapters (shown in blue and yellow) to the amplicons. **(C)** GUIDE-seq involves transfection of CRISPR/Cas9 and dsODN linkers (shown in red) that are incorporated into cleavage sites during NHEJ. Genomic DNA is fragmented, and half-functional universal adapters (shown in yellow) are added to all fragments. PCR amplification between dsODN and half-functional universal adapters using 5′ tagged primers enables selective amplification of sequence surrounding cleavage sites and the addition of a second adapter necessary for sequencing. **(D)** UDiTaS involves tagmentation of genomic DNA from nuclease-edited cells. Tagmentation fragments DNA and introduces unique molecular indices (UMIs) and adapters. Target enrichment is achieved by selective amplification of fragments between adapters and gene-specific sites. Genomic rearrangements can then be sequence using next-generation sequencing platforms. Created with Biorender.com.

2016; Tsai et al., 2015; Cameron et al., 2017; Kim and Kim, 2018). Furthermore, the biological relevance of GUIDE-seq data tends to be more robust than other methods because DSBs are tagged in the context of a cellular environment, not requiring targeted sequence validation for recognition as a *bona fide* editing site.

However, there are several limitations to the GUIDE-seq method. The dsODN, the key component to the effectiveness of the method, has not been adapted to be administered in an animal model, limiting the range of GUIDE-seq application. In addition, the dsODN has shown cytotoxicity in some primary cells (Wienert et al., 2019). Another limitation of GUIDE-seq is its dependence on the endogenous process of NHEJ to detect and tag cleavage events. DSBs not processed by NHEJ will be missed by the GUIDE-seq method.

iGUIDE Method Reduces Noise in GUIDE-seq Data by Reducing Mispriming Events

A recent update to the GUIDE-seq approach is the iGUIDE method which deals with the problem of mispriming in GUIDE-seq experiments (Nobles et al., 2019). During library preparation, GSP primers can anneal to fragments which lack the dsODN. Amplification can then yield false positive library fragments containing human DNA sequence that were not the sites for nuclease cleavage and dsODN incorporation but functionally resemble true positive library fragments. The iGUIDE method involves the use of a 46 bp dsODN in place of the 34 bp version in the original method. The additional sequence allows filtering of

misprimed library fragments during analysis. Use of the iGUIDE method reportedly reveals features of DSB distribution, such as the stronger tendency for spontaneous DSBs to occur near active genes, which are obfuscated by the noise generated by unfiltered mispriming events (Nobles et al., 2019). To date, the iGUIDE method has gained very little traction and is cited by only a single data paper in the literature. Further discussion in this manuscript will be focused on GUIDE-seq in its originally published form.

TTISS Is a Multiplex GUIDE-seq-Based Method Suitable for Comparison Between Cas9 Variants

Tagmentation-based tag integration site sequencing (TTISS) is a recently published technique which enables a multiplex examination of nucleases and nuclease targets (Schmid-Burgk et al., 2020). The technique is based on GUIDE-seq with some modifications. The protocol is streamlined by utilizing the previously published Tn5 transposase for tagmentation (Picelli et al., 2014). DNA is then purified by spin column and target enrichment is accomplished *via* two nested PCR reactions. TTISS was used to examine the balance between specificity and activity in nine SpCas9 variants including wild-type SpCas9, seven previously published variants, and one novel variant (Kleinstiver et al., 2016; Slaymaker et al., 2016; Chen et al., 2017a; Casini et al., 2018; Hu et al., 2018; Lee et al., 2018; Vakulskas et al., 2018; Schmid-Burgk et al., 2020). The results indicate a trade-off between specificity and activity in general with the precise ratio differing between Cas9 variants. Sequenced targets are attributed to a given Cas9-gRNA pair on the basis of sequence homology. This was effective in the published experiment but could conceivably confound interpretation of some results, limiting the usefulness of TTISS in some contexts. TTISS can reportedly be scaled to accommodate 60 gRNAs per transfection in HEK293T cells. But there is a trade-off in efficiency with 28% fewer off-target sites detected in a multiplexed experiment. The technique is effective for a large-scale screen of Cas9 variants but for a comprehensive look at the full off-target profile of a given Cas9 variant and gRNA-target, the reduced detection efficiency would dictate the use of another technique, e.g., GUIDE-seq or discovery of *in situ* cas off-targets and verification by sequencing (DISCOVER-Seq) (Tsai et al., 2015; Wienert et al., 2019).

UDiTaS Captures Repair Outcomes Missed by Other Methods but Requires *a priori* Knowledge of Target Sites

GUIDE-seq is not the only relevant modification to the AMP methodology. Uni-directional targeted sequencing (UDiTaS) is also a useful DSB detection technique which utilizes universal adapters and anchored primers to characterize the repair outcomes following engineered nuclease cleavage (**Figure 3D**) (Giannoukos et al., 2018). The modifications introduced in UDiTaS increase the robustness and utility of the AMP approach. In particular, UDiTaS introduces enzymatic fragmentation known as tagmentation, for genomic DNA rather than shearing by sonication. This modification addresses the tendency for shearing by sonication to introduce damage to genomic DNA that leads to base miscalling during deep sequencing (Costello et al., 2013; Chen et al., 2017b, 2018). UDiTaS introduces a novel Tn5 transposon which contains an Illumina forward adapter (i5), a barcode, and a UMI. Tagmentation yields a fragmented genomic library with adapters on either end of each fragment. Sequence-specific primers are then used to PCR amplify sites targeted by engineered nucleases. A second round of PCR adds an Illumina reverse adapter (i7), similar to the GUIDE-seq protocol. Not only does tagmentation drastically improve efficiency in hands-on time for library preparation protocols, but it also reportedly showed increased library complexity and increased linearity between expected and measured editing outcomes compared to AMP (Giannoukos et al., 2018).

As an off-target detection technique UDiTaS has limited utility due to its biased nature. Sequence-specific primers target sites of interest which require *a priori* knowledge to design. However, UDiTaS has significant utility in its ability to characterize repair outcomes for nuclease-induced cleavage. This is due to the structure of constructed library segments and the use of site-specific primers. Deep sequencing of UDiTaS will capture the junctions of repaired DSBs and thus structural rearrangements can be identified. These include translocations, inversions and large deletions. GUIDE-seq, by its nature does not detect those repair outcomes. The inserted oligonucleotide, which allows anchored priming without sequence knowledge for cleavage sites in GUIDE-seq, allows the capture of only one half of any repaired DSB junction. Reconstruction of complete cleavage sites is accomplished by mapping during analysis (Tsai et al., 2015). Thus, UDiTaS fills an important gap for data relating to repair outcomes for nuclease induced DSBs. Importantly, one approach to the problem of detecting large deletions is to use long read sequencing technologies (Amarasinghe et al., 2020). However, the accuracy and affordability of short-read sequencing platforms by comparison often make short read next-generation sequencing methods preferable and more accessible. An advantage of UDiTas is that it allows the capture and sequencing of large deletions on short read sequencing platforms. Notably, WGS could also be used to detect translocations, inversions, and large deletions but without targeted enrichment the signal to noise ratio of WGS would be markedly lower. Targeted deep sequencing on the other hand cannot capture translocations and efficient capture of inversions and large deletions would require more *a priori* knowledge for targeted deep sequencing than UDiTaS.

HIGH THROUGHPUT GENOME-WIDE TRANSLOCATION SEQUENCING

HTGTS Is Adapted for Off-Target Detection by Modifications That Enhance Target Enrichment

HTGTS is a method to detect and sequence translocations resulting from DSBs. Originally it was published as a method to study the mechanism of translocation (Chiarle et al., 2011). It has since been adapted as a method to detect off-target cleavage events caused by gene-editing nucleases (Frock et al.,

2015). The original published HTGTS method utilized the I-SceI meganuclease to introduce targeted DSBs to specific c-myc and IgH loci. The sites were selected for their frequent involvement in B cell lymphoma oncogenic translocations (Chiarle et al., 2011). DSBs induced at these known locations were then subsequently fused to other DSBs across the genome by endogenous processes (**Figure 4A**). By exploiting the known sequence of one side of the translocation junction, the sequence of fused sites involved in translocation can then be identified. The original study presented two enrichment chemistries for library preparation to capture the sequence surrounding translocations. Starting with genomic DNA containing translocation fusions with known sequence on one half of the translocation junction, the genomic DNA samples are sheared *via* restriction enzyme digestion. End-repair and adapter ligation are then carried out for all fragments in a sample (**Figure 4B**).

LAM HTGTS Adapts HTGTS for Off-Target Detection

The HTGTS method was repurposed for detection of nuclease off-target activity and protocol modifications were introduced that enhance the adapter-PCR target-enrichment methodology of the original method (**Figure 4C**) (Chiarle et al., 2011; Frock et al., 2015). The modified method is called linear amplification mediated (LAM) high throughput genome-wide translocation sequencing (LAM HTGTS). Applying the HTGTS method, introduced previously, as a nuclease off-target detection method is effectively a function of choosing applicable nucleases to induce desired bait and prey cleavage events. Using the original published method of HTGTS, previously unidentified off-target sites for the I-SceI nuclease were reported. In the updated LAM HTGTS, protocol modifications contribute to the performance of HTGTS as an off-target detection method enabling sensitivity and throughput comparable to other contemporary methods (Frock et al., 2015; Kim et al., 2015, 2016; Tsai et al., 2015; Hu et al., 2016; Cameron et al., 2017; Yan et al., 2017; Kim and Kim, 2018; Wienert et al., 2019).

The two key modifications introduced in the LAM HTGTS protocol are LAM PCR and bridge adapter ligation. LAM PCR is a method of target enrichment for sequences with a single known primer site (Schmidt et al., 2007; Paruzynski et al., 2010). LAM PCR utilizes a 5′ biotinylated primer targeting the known half of each captured junction i.e., one of the two sides of the DSB at the bait site, to linearly amplify across junction sites. Streptavidin selection is then used to magnetically isolate target sequences from genomic DNA. Bridge adapter ligation uses a double-stranded linker with a nucleotide-variable 3′ overhang to facilitate the attachment of adapters to the single-stranded library resulting from linear PCR (**Figure 4C**) (Zhou et al., 2013; Frock et al., 2015; Hu et al., 2016). Implementing these modifications yields 10–50 times more junctions for sequencing compared to the unmodified HTGTS method (Hu et al., 2016).

Performance of LAM HTGTS is comparable to other methods. For gRNAs targeting VEGFA and EMX1, LAM HTGTS identified the same major off-target sites as GUIDE-seq, although the two methods each identified unique subsets of low frequency off-target cleavage sites. This could be due to the cell lines tested but also to differences in the detection methods, which, by nature, may not be able to identify the same low-abundance cleavage sites (Hu et al., 2016). In particular, HTGTS can capture DSBs containing overhang ends, due to the endogenously repaired nature of translocation junctions, while GUIDE-seq only detects blunt-ended cleavage sites, due to the nature of uptake for oligonucleotide linkers (Tsai et al., 2015; Hu et al., 2016).

One drawback to the LAM HTGTS method is the substantial requirement of starting material. Translocations are rare compared to local rejoining events. They occur in 0.1–0.5% of cells in HTGTS libraries. The authors recommend a starting DNA mass between 20 and 100 μg for a single HTGTS library to achieve a 0.5–1.0 $\times$ 10^6 read depth on an Illumina MiSeq (Hu et al., 2016). GUIDE-seq, by contrast requires 800 ng of genomic DNA to achieve comparable detection sensitivity. Although the authors state that the sensitivity of LAM HTGTS could be increased by starting with even more DNA, the input requirements could be prohibitive for this technique on samples of limited abundance.

There is an additional point worth noting, which is made clear by the results presented in the HTGTS publications (Chiarle et al., 2011; Frock et al., 2015; Hu et al., 2016). Even on-target cleavage events can have undesirable consequences. Translocations contribute to genomic instability (Elliott and Jasin, 2002; Ramiro et al., 2006; Kosicki et al., 2018). Also, translocations can result from on-target cleavage events as readily as off-target cleavage events (Chiarle et al., 2011; Frock et al., 2015; Hu et al., 2016; Kosicki et al., 2018). This point highlights the need for detailed characterization of genome-editing systems.

CHROMATIN IMMUNOPRECIPITATION

ChIP-seq

DISCOVER-Seq (described below) is an off-target detection method which selectively amplifies CRISPR/Cas9 cleavage sites by detecting the signature of endogenous DNA repair processes (Wienert et al., 2019). The basis of DISCOVER-Seq is ChIP-seq which entails chromatin immunoprecipitation (ChIP) and subsequent deep sequencing of captured DNA fragments (ChIP-seq). Briefly, ChIP begins with formaldehyde crosslinking of a single-cell suspension (Hoffman et al., 2015). Nuclei are then extracted and fragmented *via* sonication. Fragments of interest can then be isolated—pulled down—using bead-bound antibodies allowing the study of protein-DNA interactions (Kim and Ren, 2006; Wienert et al., 2019). In the ChIP-seq methodology, the pulled-down DNA fragments are then prepared for deep sequencing (Wienert et al., 2019).

ChIP has been extensively employed to capture the sequence surrounding DSBs and characterize the genomic landscape of DSBs. Early studies utilized tiled microarrays with DNA pulled down by ChIP in a method dubbed ChIP-chip (Iacovoni et al., 2010; Szilard et al., 2010; Staszewski et al., 2011). More recent studies have moved to ChIP-seq, utilizing contemporary sequencing methods coupled with ChIP (Kim and Ren, 2006; Frietze and Farnham, 2011; Rodriguez et al., 2012; Barlow

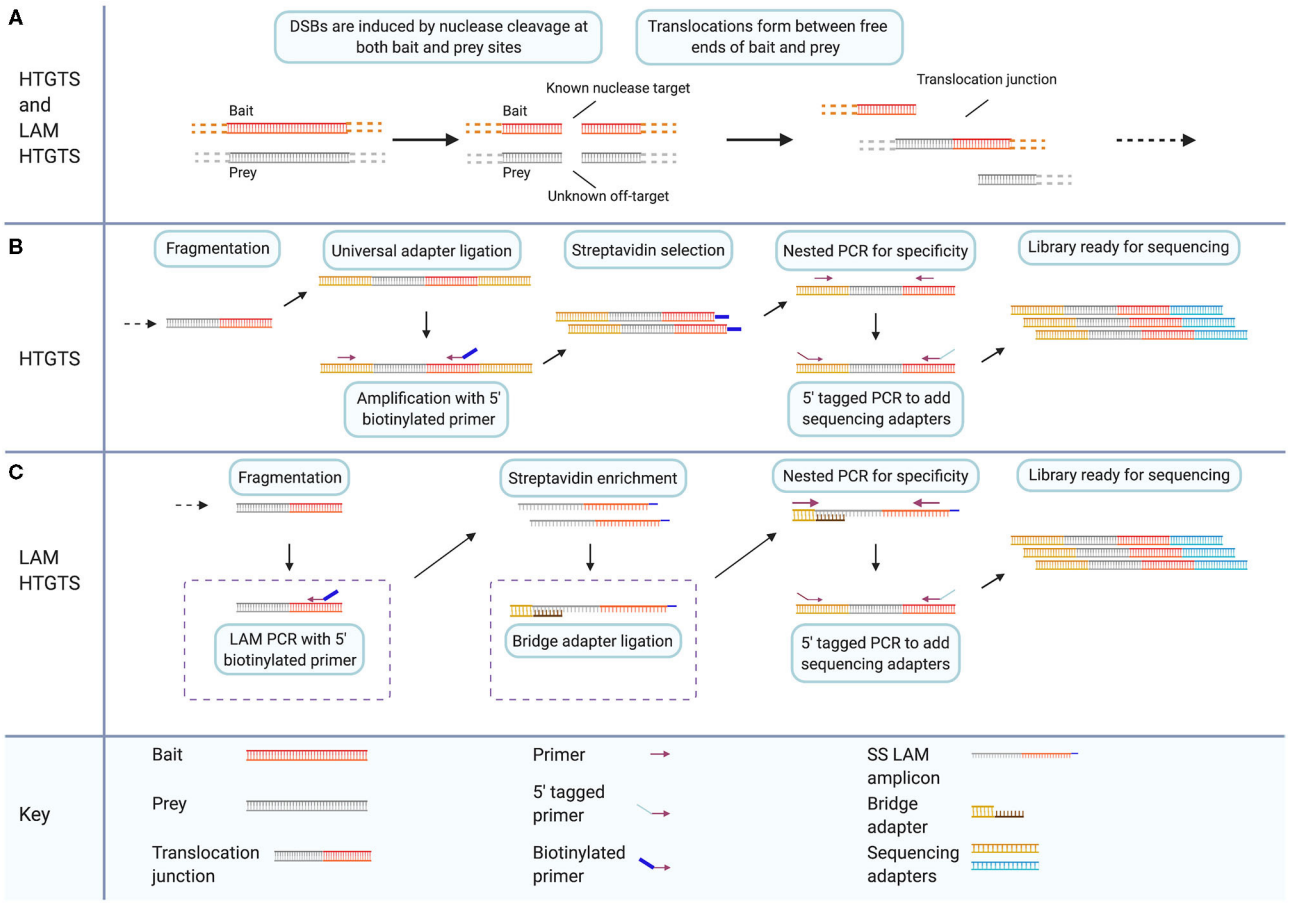


FIGURE 4 | LAM HTGTS has two specific modifications that enhance target enrichment compared to the original method and enable sensitive detection of off-target nuclease cleavage. **(A)** Both HTGTS and LAM HTGTS begin by inducing DSBs through nuclease cleavage in cells for known and unknown sequence targets referred to as bait and prey, respectively, which can form translocation junctions during DSB repair. **(B)** HTGTS involves purification and fragmentation of genomic DNA, ligation of half-functional universal adapters, and PCR amplification of fragments between known bait sequence and universal adapters. Use of a 5′ biotinylated primer during amplification enables Streptavidin enrichment followed by two rounds of PCR for specificity and addition of sequencing adapters. **(C)** LAM HTGTS is similar to the original method with key modifications. One, LAM PCR amplification with 5′ biotinylated primers is followed by Streptavidin enrichment. Two, bridge adapter ligation using and oligo with a 3′ overhang facilitates the further amplification of the single-stranded LAM PCR amplicons which are then prepared for sequencing. Created with Biorender.com.

et al., 2013; Yamane et al., 2013; Zhou et al., 2013; Duan et al., 2014; Kuscu et al., 2014; Wu et al., 2014; Khair et al., 2015; Knight et al., 2015; Madabhushi et al., 2015; O'Geen et al., 2015). γH2AX has been used as a marker for DSBs in ChIP experiments (Iacovoni et al., 2010; Szilard et al., 2010; Rodriguez et al., 2012) DSBs trigger expansive γH2AX binding domains however, and γH2AX can bind kilobases away from the site of a DSB, yielding poor resolution for DSB mapping (Bonner et al., 2008; Iacovoni et al., 2010).

Studies using ChIP-seq to characterize CRISPR/Cas9 off-target proclivity represent early attempts at unbiased survey of Cas9 activity on a genome-wide scale. Multiple studies used ChIP-seq with catalytically inactive Cas9 (dCas9) to pull down Cas9 binding sites (Duan et al., 2014; Kuscu et al., 2014; Wu et al., 2014; Knight et al., 2015; O'Geen et al., 2015). However, ChIP-seq using dCas9 is limited with respect to off-target detection; it has been shown to yield abundant false positives (Kuscu et al., 2014; Wu et al., 2014; Knight et al., 2015; Tsai et al., 2015). For example, only one out of 295 dCas9 binding sites identified by ChIP-seq in mouse embryonic stem cells (mESCs) was identified by targeted sequencing as a bona fide cleavage-target (Wu et al., 2014).

DISCOVER-Seq Adapts ChIP-seq to an Accurate and Sensitive Off-Target Detection Method Comparable to Other Contemporary Methods

DISCOVER-Seq advances the ChIP-seq method by utilizing meiotic recombination 11 homolog 1 (MRE11), a DNA repair protein that is part of the MRE11-RAD50-NBS1 (MRN) complex (**Figure 5**). The MRN complex is involved in DNA damage responses (DDRs) in general, including DSB repair (Connelly

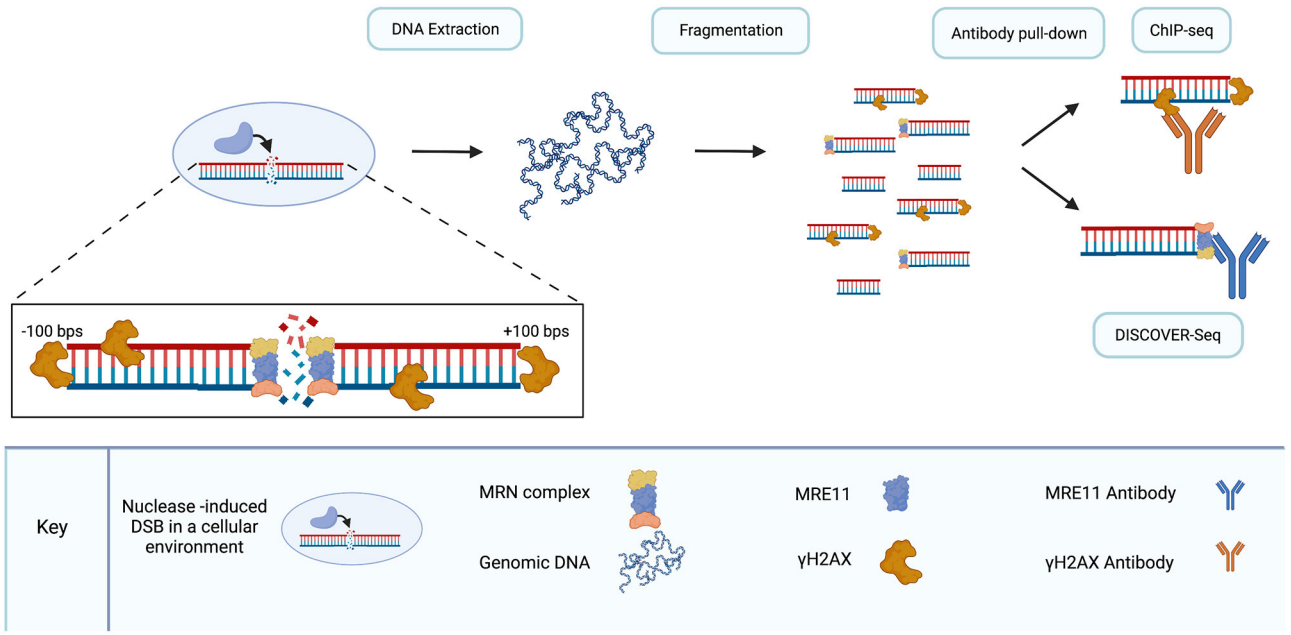


FIGURE 5 | DISCOVER-seq is a specialized version of ChIP-seq. As DSBs are introduced to DNA in living cells, endogenous repair processes recruit proteins to break sites. γH2AX localizes to DSBs within hundreds of base pairs in either direction. The MRN complex, recruited by γH2AX, localizes to the break site. Following genomic DNA extraction and fragmentation, antibody pull-down of fragments enables the sequencing of fragments surrounding DSBs. γH2AX pull-down is used in ChIP-seq and lacks the resolution to precisely locate DSB sites. MRN pull-down, specifically the MRE11 subunit of the MRN complex, enables precise sequencing of DSB sites with single-nucleotide resolution. Use of the MRE11 antibody for pull-down is the distinguishing characteristic of DISCOVER-seq compared to ChIP-seq. Created with Biorender.com.

and Leach, 2002; Moreno-Herrero et al., 2005; Borde, 2007; Oh and Symington, 2018; Syed and Tainer, 2018; Bian et al., 2019). It also has roles in replication stress, handling of dysfunctional telomeres, cellular response to viral infection, and tumorigenesis (Spehalski et al., 2017; Syed and Tainer, 2018; Bian et al., 2019). Notably, the way that the MRE11 subunit in particular handles different DSB end-moieties may dictate whether DSBs are repaired by HR or NHEJ (Shibata et al., 2014; Liao et al., 2016).

MRE11 is optimal for nuclease-cleavage detection because the MRN complex localizes to DSBs, including those created by CRISPR/Cas9, before ends are joined by repair (Syed and Tainer, 2018; Bian et al., 2019; Wienert et al., 2019). MRN is recruited to DSBs by γH2AX. In addition, MRE-11 is ubiquitous and conserved across all taxonomic kingdoms (Connelly and Leach, 2002; van den Bosch et al., 2003; Wienert et al., 2019). Disruption of each individual component of the MRN complex has been shown to be embryonically lethal in mice (Luo et al., 1999; Zhu et al., 2001; Buis et al., 2008) and mutations in the genes of each individual component have been linked to genomic instability in humans (van den Bosch et al., 2003). Expression of MRE11 across a range of tissues in mice has been demonstrated and following induction of DSBs, MRE11-detection peaks in cells before indels are formed (Wienert et al., 2019).

DISCOVER-Seq detects DSBs with single-nucleotide resolution and compares favorably to other off-target detection methods. However, DISCOVER-Seq reportedly has a sensitivity threshold of 0.3%, slightly higher than other contemporary techniques. A VEGFA target was examined in human K562 cells using both DISCOVER-Seq and GUIDE-seq. They identified 49 off-target sites in common between the techniques but also 41 off-targets sites unique to GUIDE-seq and eight off-target sites unique to DISCOVER-Seq (Wienert et al., 2019). This head-to-head comparison suggests that capture of the entirety of the off-target landscape for at least some gRNAs will require multiple methods. Another favorable feature of DISCOVER-Seq compared to GUIDE-seq is that DISCOVER-Seq works in primary induced pluripotent stem cells (iPSCs). DISCOVER-Seq was shown to detect off-target sites in iPSCs and to differentially detect an allelic specificity in primary cells from a Charcot-Marie-Tooth (CMT) patient with a heterozygous mutation. Data was also shown demonstrating that transfection of the dsODN necessary for GUIDE-seq was toxic to iPSCs (Wienert et al., 2019).

Although other techniques may boast greater sensitivity, DISCOVER-Seq is currently one of only two techniques shown to detect off-target events *in vivo* in an animal model (Wienert et al., 2019); VIVO is the other (Akcakaya et al., 2018). DISCOVER-Seq was tested on the same system as VIVO for comparison. A Pcsk9-gP gRNA was delivered *via* adenoviral infection in a murine model. Mice were then sacrificed at 24-, 26-, and 48-h time points. Twenty-seven off-target sites identified by DISCOVER-Seq were validated by amplicon sequencing and had indel rates between 0.9 and 78.1%. An important point of comparison is that

17 of the 27 sites identified by DISCOVER-Seq were identified by the *in vitro* CIRCLE-seq stage of the VIVO method but were not validated due to the high volume of potential sites generated by the CIRCLE-seq method. This is an important point with respect to the efficiency of *in vitro* techniques and the differential utility of currently available off-target detection methods. Unbiased full-genome survey of the off-target landscape is critical for translation of gene-editing to clinical application. And *in vitro* methods are sensitive and thorough means to characterize the activity of targeted nucleases with respect to sequence homology alone. But the need to validate the high volume of targets detected with *in vitro* methods can lead to a biased survey of high-priority or high-probability sites and bona fide off-target loci can be lost among the false positives.

IN SITU END-CAPTURE TECHNIQUES FOR OFF-TARGET DETECTION

In situ end-capture methods are a distinct class of techniques which can detect off-target nuclease cleavage by capturing the free ends of DSBs in fixed cells. A variety of *in situ* methods have been published (Crosetto et al., 2013; Baranello et al., 2014; Dorsett et al., 2014; Canela et al., 2016; Lensing et al., 2016; Yan et al., 2017; Biernacka et al., 2018). These methods can be highly sensitive; END-Seq reportedly has a sensitivity of 0.01% and iBLESS can reportedly detect a single DSB in 100,000 cells in *Saccharomyces cerevisiae* (Canela et al., 2016; Biernacka et al., 2018). However, *in situ* methods are limited to the capture of DSBs at a single timepoint preceding cellular response to the induced damage. These methods also tend to have labor intensive protocols with many technical steps. By nature, this class of techniques are less suitable than other methods discussed in this review for research focused on clinical translation of gene editing technologies and more pertinent to studies of enzyme kinetics or the characterization of end moieties following cleavage events. We therefore have reserved an in-depth treatment of this subject for future consideration.

COMPARISON BETWEEN METHODS

To date there is no off-target detection method optimized for all circumstances. **Table 2** shows the most relevant modern off-target detection methodologies and the important factors that distinguish each technique. Comparisons between methods rely on gene targets that have been used to evaluate engineered nuclease specificity for years and pre-date the development of unbiased genome-wide techniques. These targets are useful as a metric for comparisons between methods but do not generalize to all anticipated applications of each technology. A recent study compared the performance of GUIDE-seq, CIRCLE-seq, and SITE-Seq side-by-side using promiscuous off-target gRNAs (Chaudhari et al., 2020). Results show that each of the three assays performed with similar efficiency at detection of *bona fide* off-target sites. Results also show that GUIDE-seq has the best correlation of assay signal to observed editing but it is the least reproducible across replicates. Overall, this study concludes that GUIDE-seq is good choice for measuring off-target specificity *ex vivo* in a cellular context but CIRCLE-seq is a good choice for experiments which preclude the use of GUIDE-seq (i.e., studies involving *in vivo* nuclease editing).

The common thread between all off-target detection methods is that the read-out is always deep sequencing data. Some methods, such as WGS, require more computational post-processing for analysis than others, such as GUIDE-seq or BLISS, which have published analysis pipelines. One point of distinction between methods which may not be readily apparent, is that there is a difference between single-nucleotide resolution in detection and mapping to single-nucleotide resolution during analysis. Digenome-seq and CIRCLE-seq for example, yield sequence data that has single-nucleotide resolution inherent in the DNA library. GUIDE-seq on the other hand, maps to single-nucleotide resolution during data analysis. Another feature of CIRCLE-seq is that it is a reference-genome free method because each fragment in the library contains both ends of the cleavage site.

In vitro techniques can be useful in experiments where transfections are difficult and characterization of gene-editing performance independent of endogenous repair pathways is desirable. But the end-goal of experimentation can dictate which method is best on a case-by-case basis. CIRCLE-seq and SITE-Seq are sensitive and thorough, capturing high proportions of potential off-target sites for a given gRNA. They are prone to high false-positive rates, often referred to as false discovery. This is an important distinction. With respect to the *in vitro* off-target detection assay, many of the detected sites are true cuts in the DNA. But they are not bona fide off-target sites which occur in living cells. False discovery is a more apt description for such data points. The high rate of false discovery for these methods may be a drawback in some experimental paradigms where the sheer quantity of data from *in vitro* methods precludes comprehensive validation, thereby requiring a biased follow-up analysis. For example, a subset of off-target sites detected by DISCOVER-Seq were captured by VIVO for the same target but were excluded from the validation set (Akcakaya et al., 2018; Wienert et al., 2019). DIG-Seq is a modification for *in vitro* methods which addresses this problem by maintaining chromatin architecture. The fewer sites identified are therefore more likely to have clinical relevance and accordingly a higher validation rate is reported for DIG-Seq compared to CIRCLE-seq and SITE-Seq (Kim and Kim, 2018).

By contrast, some studies are interested in more than identification of cleavage sites. Repair outcomes are also important. HTGTS and UDiTaS can capture translocations and large genomic rearrangements that are missed by other methods. *In situ* techniques offer a distinctly different strategy that can also be construed as an advantage or disadvantage depending on experimental purpose. Based on a study using H2AX and 53BP1 as DSB markers, the majority of DSBs are resolved within an 8 h timeframe (Asaithamby and Chen, 2009). The *in situ* capture of DSBs at a single timepoint may offer a distinct advantage to enzymology studies whereas the sum total of captured events over time may be of greater interest in other studies.

For off-target detection in animal models, DISCOVER-Seq and VIVO are the best options aside from WGS which has a low

TABLE 2 | Important features that influence the utility of off-target detection methods.

Method	References	Category	Sensitivity*	Input DNA	Detection state	Detection mechanism	Edits detected	Genomic context for nuclease	Target enrichment
Digenome-seq	Kim et al., 2015	*In vitro*	0.1%	1 μg	Target-site DSB	Sequence alignment pile-up	DSBs	Cell free DNA	None
DIG-Seq	Kim and Kim, 2018	*In vitro*	0.1%	1 μg	Target-site DSB	Sequence alignment pile-up	DSBs	Chromatin environment	None
SITE-Seq	Cameron et al., 2017	*In vitro*	0.1%	7.5 μg	Target-site DSB	Sequence alignment pile-up	DSBs	Cell free DNA	Streptavidin selection
CIRCLE-seq	Tsai et al., 2017	*In vitro*	0.01%	25 μg	Target-site DSB	*In vitro* cleavage	DSBs	*In vitro* genomic library	Linearization of circularized library
CHANGE-seq	Lazzarotto et al., 2020	*In vitro*	NR**	5 μg	Target-site DSB	*In vitro* cleavage	DSBs	*In vitro* genomic library	Linearization of circularized library
VIVO	Akcakaya et al., 2018	*In vivo*	0.13%	25 μg	Target-site mutation	Targeted sequencing	Repair site mutations (indels)	In organism	Targeted sequencing
GUIDE-seq	Tsai et al., 2015	*Ex vivo*	0.12%	800 ng	Repaired with oligonucleotide incorporation	Oligonucleotide uptake by NHEJ	NHEJ repair sites	Cellular environment	Anchored primer amplification
UDiTaS	Giannoukos et al., 2018	*Ex vivo*	0.1% (all edits), 0.01% (translocations)	50 ng	Repaired with large deletions, inversions, translocations	Targeted sequencing	Repair site mutations (indels), translocations, inversions, large deletions	Cellular environment	Anchored primer amplification
LAM-PCR HTGTS	Frock et al., 2015; Hu et al., 2016	*Ex vivo*	NR**	20–100 μg	Translocation junction	Translocation bait and prey	Translocations	Cellular environment	Anchored primer amplification
DISCOVER-seq	Wienert et al., 2019, 2020	*In vivo*	0.3%	$2 \times 10^6 - 1 \times 10^7$ cells; 40–80 mg homogenized tissue	Unrepaired DSB during DNA damage response (DDR)	MRE11 antibody labeling	Unrepaired DSBs	In organism	ChIP
BLISS	Yan et al., 2017	*In situ*	NR**	100 μL of nuclei suspension from cells on 13 mm coverslips	Unrepaired DSB *in situ*	*In situ* end-capture	Unrepaired DSBs	Cellular environment	Transcription

**Sensitivity is defined as the frequency of occurrence on a per cell basis in a cell population. For example a sensitivity of 0.1% refers to an editing event which occurs in 1 out of 1,000 cells.*
***Not reported (NR) in the cited manuscript.*

signal to noise ratio. While VIVO is more sensitive, DISCOVER-Seq yields a smaller, more clinically relevant data set which may allow an unbiased validation of all identified targets while VIVO may not. However, for off-target detection in a cellular environment, GUIDE-seq is still the most sensitive option which yields the most clinically relevant data. A substantial portion of data (45%) collected by GUIDE-seq was missed by DISCOVER-Seq when looking at the same target. But in some types of primary cells, the dsODN that must be transfected to make GUIDE-seq work, can be cytotoxic (Wienert et al., 2019).

While the different methodologies have distinct mechanisms, there have been several common trends in improvement. Efficiency of each class of technique has been steadily improving. For example, the *in situ* method breaks labeling *in situ* and sequencing (BLISS) is substantially easier and quicker than direct *in situ* breaks labeling, enrichment on streptavidin and next-generation sequencing (BLESS) to carry out without sacrificing sensitivity (Crosetto et al., 2013; Yan et al., 2017). Also, the introduction of Tn5 transposase to replace shearing by sonication has greatly reduced the physical labor involved in library preparation for sequencing. And the sensitivity of all relevant off-target methodologies has been steadily increasing.

SENSITIVITY

Sensitivity is an important measure of comparison for assays measuring the same phenomena. An often-described aspect of techniques in terms of sensitivity is detection of a subset of off-target sites that are unique to a particular method when evaluating the same target, i.e., VEGFA or EMX1. But each technique identifies a subset of off-target sites that others do not, and they cannot all be more sensitive than each other. These technique specific subsets are likely due to genomic context or the specific mechanisms of detection and enrichment. Whether or not a technique detects certain off-target sites that other methods miss differs significantly from the explicit definition of sensitivity as the lower limit of frequency in a cell population that can be detected with statistical confidence. For example, as stated earlier, a sensitivity of 0.1% describes an ability to detect events which occur in 1 out of 1,000 cells. The currently competitive and relevant techniques for off-target detection are primarily limited by the error rate of next-generation sequencing techniques not by the inherent capabilities of the assays. Increasing sensitivity in any of these techniques generally requires more starting material and greater sequencing depth. If sequencing depth is the deciding factor in sensitivity, then methods requiring substantially less starting material than others may be distinctly advantageous.

THROUGHPUT OF OFF-TARGET VALIDATION METHODS

Another area of steady improvement for off-target detection is throughput. This is largely due to improvement in sequencing technology and to target enrichment strategies for off-target cleavage sites. A methodology which is not new but is recently refined and may offer greater throughput for future experiments is rhAmp PCR. rhAmp PCR is used in off-target detection as a validation method that enhances the efficiency and specificity of multiplex PCR by disallowing amplification at sites other than those with exact primer-target homology. Briefly, rhAmp primers require the addition of RNase H2 enzyme to remove a blocking moiety from hybridized primers in order to allow extension. Implementation of rhAmp PCR reduces primer dimers and non-specific amplification (Dobosy et al., 2011). It has been used to facilitate NGS amplicon sequencing allowing higher throughput screening of potential *bona fide* target sites for base editors and CRISPR/Cas9 (Chaudhari et al., 2020; Shapiro et al., 2020; Zeng et al., 2020). Implementation of rhAmp PCR increases the throughput of targeted amplicon screening for bona fide off-target nuclease cleavage.

APPLICATION TO THE HIV GENE EDITING FIELD

Targeting specificity has been considered in the design of gRNAs targeting HIV (Dampier et al., 2014, 2017, 2018; Hu et al., 2014; Kaminski et al., 2016a,b,c; Wang et al., 2016a,b; Bella et al., 2018; Link et al., 2018; Ophinni et al., 2018; Roychoudhury et al., 2018; Darcis et al., 2019; Sullivan et al., 2019; Chung et al., 2020). Some of the studies investigating HIV-1-CRISPR strategies have examined the off-target activities of gRNAs empirically using biased techniques including T7E1 and Surveyor assays, targeted amplicon sequencing and TIDE (Hou et al., 2015; Ji et al., 2016; Saayman et al., 2016; Yoder and Bundschuh, 2016; Lebbink et al., 2017; Kunze et al., 2018; Ophinni et al., 2018; Wang et al., 2018; Campbell et al., 2019). Other studies have used WGS to analyze the specificity of HIV-targeting gRNAs (Hu et al., 2014; Kaminski et al., 2016a,b,c; Xu et al., 2017; Dash et al., 2019). But rigorous examination of targeting specificity using unbiased, genome-wide techniques has not been applied to HIV-targeting gRNAs to date. For studies that predate 2015, this was unavoidable as most of the unbiased, genome-wide approaches have only been developed recently. However, as gene-editing strategies move closer to developing into viable treatment options, the need for high-throughput off-target screening will play an increasingly important role.

Thus far, the limited application of unbiased, genome-wide off-target detection for HIV-targeting gRNAs has been adequate. Most studies have been focused on the considerable need for establishing a proof of concept for the application of this technology and rigorous off-target analysis has not been of paramount importance in establishing the functional aspects of this approach. For example, some studies have established optimal proviral targets for viral deactivation. The LTR is the most common HIV-1 CRISPR/Cas9 target investigated thus far (Ebina et al., 2013; Dampier et al., 2014, 2017; Hu et al., 2014; Zhu et al., 2015; Bialek et al., 2016; Ji et al., 2016; Kaminski et al., 2016a,b,c; Limsirichai et al., 2016; Saayman et al., 2016; Ueda et al., 2016; Wang et al., 2016a,b; Yin et al., 2016; Lebbink et al., 2017; Zhao et al., 2017; Bella et al., 2018; Kunze et al., 2018; Roychoudhury et al., 2018; Campbell et al., 2019; Darcis et al., 2019; Dash et al., 2019; Kaushik et al., 2019; Su et al.,

2020). The ability of CRISPR/Cas9 to deactivate the virus in cell lines, primary cells *ex vivo* and human primary cells in engrafted in mice has also been established (Ebina et al., 2013; Hu et al., 2014; Kaminski et al., 2016a,b; Lebbink et al., 2017; Bella et al., 2018; Ophinni et al., 2018; Campbell et al., 2019; Darcis et al., 2019). Also, the mechanism of that action—mutation, excision, or inversion—has been investigated (Mefferd et al., 2018; Binda et al., 2020). Other studies have characterized viral escape mechanisms and established that a multiplex targeting approach can prevent the emergence of escape mutants (Wang et al., 2016a,b, 2018; Yoder and Bundschuh, 2016; Lebbink et al., 2017; Zhao et al., 2017; Gao et al., 2020). It has also been demonstrated that Tat-driven CRISPR/Cas9 expression can create a negative feedback system that quenches CRISPR/Cas9 production in the absence of viral protein production (Kaminski et al., 2016c). Recently, great strides have been made in demonstrating the utility of the CRISPR/Cas9 system paired with long-acting slow-effective release (LASER) ART in clearing HIV-1 infection from a humanized mouse model (Dash et al., 2019). Additionally, the delivery of CRISPR/Cas9 using AAV vectors has been demonstrated as a viable approach (Kaminski et al., 2016a; Kunze et al., 2018; Dash et al., 2019; Mancuso et al., 2020). Also, an *in vitro* model for magnetically delivering CRISPR/Cas9 across the blood-brain barrier has been developed (Kaushik et al., 2019).

So far, the limited application of off-target analysis has been appropriate to the goals of these proof-of-concept studies. But as CRISPR/Cas9 treatment moves toward clinical application, the gRNAs that are going to be used for clinical treatment will require rigorous off-target analysis. There are published results to uphold this viewpoint. The off-target proclivity of HIV-targeting gRNAs was investigated using targeted amplicon deep-sequencing for the top three off-target candidate sites on each of three gRNAs. No mutations above background level were found at the observed sites. Nonetheless, stable expression of the LTR6 gRNA was found to severely reduce the viability of SupT1 cells (Lebbink et al., 2017). This likely indicates that the off-target screening methodology used was not thorough or sensitive enough to identify all off-target events. These results support the notion that biased targeted examination of potential off-target sites is not sufficient to fully characterize the specificity of gene-editing systems. Results presented in the VIVO and DISCOVER-Seq studies also support this point (Akcakaya et al., 2018; Wienert et al., 2019). DISCOVER-Seq identified bona fide off-target sites for the Pcsk9-gP gRNA that were also identified by VIVO in the *in vitro* CIRCLE-seq phase of the experiment but were not prioritized for further analysis by targeted amplicon sequencing.

As no off-target detection method is ideal in all cases, it is important to consider the factors involved in HIV gene therapy. **Table 3** describes a set of criteria for choosing the ideal method at each stage of the development process. As gRNAs are refined and screened for target specificity with the goal of clinical translation, different off-target detection methodologies are best suited for different phases of evaluation. As described earlier, there are two main phases to this process: nomination and validation. Here a further distinction is made and the evaluation process for gRNAs is described in three phases: discovery, refinement, and validation (**Table 3**). In this paradigm discovery and refinement are two aspects of nomination. In the discovery phase it is important to be able to rapidly and affordably screen potential gRNAs for off-target risks. Computational methods can be employed for this task due to their rapid turn-around time, but moderate false negative rates and high false discovery rates may exclude some good gRNAs. Ideally, SITE-Seq should be used to avoid excluding potentially good candidates. In the refinement stage it is important to have methods that can evaluate the candidates in cells of interest. While DISCOVER-Seq can be used in both *in cellulo* and *in vivo* conditions, it is limited to detecting DSBs that are extant at the time of sampling. With the dynamic nature of these breaks, it is important to understand the accumulated total spectrum of possible targets to produce an appropriate candidate list for validation. GUIDE-seq is the ideal method for this stage. With candidate sites in hand, it is important to validate the entire spectrum of the repair profile in edited cells.

In the validation phase it is important to fully characterize the editing profile of the gRNAs at all on- and off-target sites. The best methods to accomplish this are amplicon sequencing and UDiTaS. At this stage, with the range of targeting sites established using a genome-wide unbiased technique (i.e., GUIDE-seq or DISCOVER-Seq), the use of biased methods requiring *a priori* knowledge is warranted. For this purpose, amplicon sequencing is straight-forward and effective. Whereas GUIDE-seq and DISCOVER-Seq can by their nature only capture sites where editing has occurred, amplicon sequencing reveals the outcome of editing events (e.g., indels) or lack thereof. However, UDiTaS presents several advantages over amplicon sequencing. In addition to capturing both edited and unedited sites, UDiTaS incorporates a UMI thereby allowing quantification of editing efficiency sans PCR bias. Furthermore, as HIV-1 excision therapy will likely require multiple gRNAs delivered simultaneously, it is important to screen for large deletions, a difficult feat for standard amplicon sequencing. UDiTaS solves this problem by utilizing one target specific primer and universal adapters allowing it to capture these alternate repair modalities.

CLOSING REMARKS

The continued development of off-target detection techniques has been a great boon for genome editing. Some studies have found that off-target events are rare in primary cells and animal models, (Smith et al., 2014; Suzuki et al., 2014; Veres et al., 2014; Iyer et al., 2015). And Zuo et al. showed with GOTI that off-targets introduced to a single blastomere in a two-cell mouse embryo are not carried through as cells divide (Zuo et al., 2019). However, these results do not generalize to all gene editing systems or gRNAs. Rather they demonstrate that gene-editing systems have the potential to be highly specific under the proper conditions and provide proof of concept that high-fidelity nuclease targeting can be achieved. But they do not preclude the need for off-target analysis. There is a potent example of gene therapy having serious adverse effects causing lymphocytosis due

TABLE 3 | Criteria for choosing the appropriate off-target detection method for each stage of gRNA development.

	Acceptable	Ideal
Discovery Initial iterations require rapid and low-cost techniques	**Computational Prediction** • Instantaneous • Low cost • High FDR • Moderate FNR	**SITE-Seq** • Rapid • Low cost • High FDR • Low FNR
Refinement Candidate gRNA refinement requires a cellular context and should not rely on *a priori* known candidate sites	**DISCOVER-Seq** • Moderate cost • *In vivo* or *in cellulo* cutting context • Only detects extant DSBs	**GUIDE-seq** • Streamlined for throughput • Moderate cost • Low FDR • Low FNR • Detects cumulative DSBs
Validation Final safety validation of gRNAs for human trials should be evaluated in animal models as well as *ex vivo* tissue samples using methods that capture the entire repair profile	**Amplicon Sequencing** • Low cost • Low FDR • Low FNR • Misses some repair types	**UDiTaS** • Streamlined for throughput • Moderate cost • Low FDR • Low FNR • Detects all repair types

Three phases of gRNA evaluation are presented. For each phase an acceptable method and an ideal method are described with a list of primary attributes for each. For acceptable methods, main drawbacks are shown in red. For ideal methods distinguishing advantages are displayed in green.

to an unforeseen translocation event in one patient (Hacein-Bey-Abina et al., 2003). As new gene-editing systems are developed and more gRNAs are designed, they must be tested empirical and they must also be tested in a variety of conditions. CIRCLE-seq identified 55 sites preferentially cleaved depending on cell type due to the presence of SNVs in the protospacer or PAM underscoring this point (Tsai et al., 2017).

At present it is unclear what the full screening regimen should be to rigorously establish a safety profile for a CRISPR/Cas9 therapeutic. The overlapping portions of data sets for off-target techniques that have been examined on common targets such as VEGFA and EMX1 to facilitate comparison are encouraging with respect to the validity of the methods. But each off-target method has also turned up a subset of bona fide off-target sites which were missed by other methods (Frock et al., 2015; Kim et al., 2015; Tsai et al., 2015, 2017; Cameron et al., 2017; Yan et al., 2017; Kim and Kim, 2018; Wienert et al., 2019). A combination of techniques will be necessary to fully characterize the off-target landscape of any gene-editing system. These strategies will also need to be accompanied by cell viability assays to uphold the results of such screening.

AUTHOR CONTRIBUTIONS

AA, C-HC, AGA, WD, TG, IS, MN, and BW conceptualized the manuscript, contributed to writing, and made critical revisions. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by National Institute of Mental Health (NIMH) R01 MH110360 (Contact PI, BW), NIMH Comprehensive NeuroAIDS Center (CNAC) P30 MH092177 (BW, PI of the Drexel subcontract involving the Clinical and Translational Research Support Core), NIMH T32 MH079785 (Drexel Component PI, BW), and the Ruth L. Kirschstein National Research Service Award T32 MH079785 (BW, Principal Investigator of the Drexel University College of Medicine component and Dr. Olimpia Meucci as Co-Director). AGA was also supported by the Drexel University College of Medicine Dean's Fellowship for Excellence in Collaborative or Themed Research (AGA, fellow; BW, mentor).

REFERENCES

Ajay, S. S., Parker, S. C., Abaan, H. O., Fajardo, K. V., and Margulies, E. H. (2011). Accurate and comprehensive sequencing of personal genomes. *Genome Res.* 21, 1498–1505. doi: 10.1101/gr.123638.111

Akcakaya, P., Bobbin, M. L., Guo, J. A., Malagon-Lopez, J., Clement, K., Garcia, S. P., et al. (2018). *In vivo* CRISPR editing with no detectable genome-wide off-target mutations. *Nature* 561, 416–419. doi: 10.1038/s41586-018-0500-9

Allen, A. G., Chung, C. H., Atkins, A., Dampier, W., Khalili, K., Nonnemacher, M. R., et al. (2018). Gene editing of HIV-1 co-receptors to prevent and/or cure virus infection. *Front. Microbiol.* 9:2940. doi: 10.3389/fmicb.2018.02940

Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., and Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* 21:30. doi: 10.1186/s13059-020-1935-5

Aryal, N. K., Wasylishen, A. R., and Lozano, G. (2018). CRISPR/Cas9 can mediate high-efficiency off-target mutations in mice *in vivo*. *Cell Death Dis.* 9:1099. doi: 10.1038/s41419-018-1146-0

Asaithamby, A., and Chen, D. J. (2009). Cellular responses to DNA double-strand breaks after low-dose gamma-irradiation. *Nucleic Acids Res.* 37, 3912–3923. doi: 10.1093/nar/gkp237

Baranello, L., Kouzine, F., Wojtowicz, D., Cui, K., Przytycka, T. M., Zhao, K., et al. (2014). DNA break mapping reveals topoisomerase II activity genome-wide. *Int. J. Mol. Sci.* 15, 13111–13122. doi: 10.3390/ijms150713111

Barlow, J. H., Faryabi, R. B., Callen, E., Wong, N., Malhowski, A., Chen, H. T., et al. (2013). Identification of early replicating fragile sites that contribute to genome instability. *Cell* 152, 620–632. doi: 10.1016/j.cell.2013.01.006

Baxter, A. E., O'Doherty, U., and Kaufmann, D. E. (2018). Beyond the replication-competent HIV reservoir: transcription and translation-competent reservoirs. *Retrovirology* 15:18. doi: 10.1186/s12977-018-0392-7

Bella, R., Kaminski, R., Mancuso, P., Young, W. B., Chen, C., Sariyer, R., et al. (2018). Removal of HIV DNA by CRISPR from patient blood engrafts in humanized mice. *Mol. Ther. Nucleic Acids* 12, 275–282. doi: 10.1016/j.omtn.2018.05.021

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456, 53–59. doi: 10.1038/nature 07517

Bialek, J. K., Dunay, G. A., Voges, M., Schafer, C., Spohn, M., Stucka, R., et al. (2016). Targeted HIV-1 latency reversal using CRISPR/Cas9-derived transcriptional activator systems. *PLoS ONE* 11:e0158294. doi: 10.1371/journal.pone.0158294

Bian, L., Meng, Y., Zhang, M., and Li, D. (2019). MRE11-RAD50-NBS1 complex alterations and DNA damage response: implications for cancer treatment. *Mol. Cancer* 18:169. doi: 10.1186/s12943-019-1100-5

Biernacka, A., Zhu, Y., Skrzypczak, M., Forey, R., Pardo, B., Grzelak, M., et al. (2018). i-BLESS is an ultra-sensitive method for detection of DNA double-strand breaks. *Commun. Biol* 1:181. doi: 10.1038/s42003-018-0165-9

Binda, C. S., Klaver, B., Berkhout, B., and Das, A. T. (2020). CRISPR-Cas9 dual-gRNA attack causes mutation, excision and inversion of the HIV-1 proviral DNA. *Viruses* 12:330. doi: 10.3390/v12030330

Blankson, J. N., Persaud, D., and Siliciano, R. F. (2002). The challenge of viral reservoirs in HIV-1 infection. *Annu. Rev. Med.* 53, 557–593. doi: 10.1146/annurev.med.53.082901.104024

Bolukbasi, M. F., Gupta, A., and Wolfe, S. A. (2016). Creating and evaluating accurate CRISPR-Cas9 scalpels for genomic surgery. *Nat. Methods* 13, 41–50. doi: 10.1038/nmeth.3684

Bonner, W. M., Redon, C. E., Dickey, J. S., Nakamura, A. J., Sedelnikova, O. A., Solier, S., et al. (2008). GammaH2AX and cancer. *Nat. Rev. Cancer* 8, 957–967. doi: 10.1038/nrc2523

Borde, V. (2007). The multiple roles of the Mre11 complex for meiotic recombination. *Chromosome Res.* 15, 551–563. doi: 10.1007/s10577-007-1147-9

Brinkman, E. K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* 42:e168. doi: 10.1093/nar/gku936

Brinkman, E. K., Kousholt, A. N., Harmsen, T., Leemans, C., Chen, T., Jonkers, J., et al. (2018). Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic Acids Res.* 46:e58. doi: 10.1093/nar/gky164

Buis, J., Wu, Y., Deng, Y., Leddon, J., Westfield, G., Eckersdorff, M., et al. (2008). Mre11 nuclease activity has essential roles in DNA repair and genomic stability distinct from ATM activation. *Cell* 135, 85–96. doi: 10.1016/j.cell.2008.08.015

Cameron, P., Fuller, C. K., Donohoue, P. D., Jones, B. N., Thompson, M. S., Carter, M. M., et al. (2017). Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods* 14, 600–606. doi: 10.1038/nmeth.4284

Campbell, L. A., Coke, L. M., Richie, C. T., Fortuno, L. V., Park, A. Y., and Harvey, B. K. (2019). Gesicle-mediated delivery of CRISPR/Cas9 ribonucleoprotein complex for inactivating the HIV provirus. *Mol. Ther.* 27, 151–163. doi: 10.1016/j.ymthe.2018.10.002

Canela, A., Sridharan, S., Sciascia, N., Tubbs, A., Meltzer, P., Sleckman, B. P., et al. (2016). DNA breaks and end resection measured genome-wide by end sequencing. *Mol. Cell* 63, 898–911. doi: 10.1016/j.molcel.2016.06.034

Carballar-Lejarazu, R., Kelsey, A., Pham, T. B., Bennett, E. P., and James, A. A. (2020). Digital droplet PCR and IDAA for the detection of CRISPR indel edits in the malaria species Anopheles stephensi. *BioTechniques* 68, 172–179. doi: 10.2144/btn-2019-0103

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., et al. (2018). A highly specific SpCas9 variant is identified by *in vivo* screening in yeast. *Nat. Biotechnol.* 36, 265–271. doi: 10.1038/nbt.4066

Chari, R., Mali, P., Moosburner, M., and Church, G. M. (2015). Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nat. Methods* 12, 823–826. doi: 10.1038/nmeth.3473

Chaudhari, H. G., Penterman, J., Whitton, H. J., Spencer, S. J., Flanagan, N., Lei Zhang, M. C., et al. (2020). Evaluation of homology-independent CRISPR-Cas9 off-target assessment methods. *CRISPR J* 3, 440–453. doi: 10.1089/crispr.2020.0053

Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., et al. (2017a). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550, 407–410. doi: 10.1038/nature24268

Chen, L., Liu, P., Evans, T. C. Jr., and Ettwiller, L. M. (2017b). DNA damage is a pervasive cause of sequencing errors, directly confounding variant identification. *Science* 355, 752–756. doi: 10.1126/science.aai8690

Chen, L., Liu, P., Evans, T. C. Jr., and Ettwiller, L. M. (2018). Response to comment on "DNA damage is a pervasive cause of sequencing errors, directly confounding variant identification". *Science* 361. doi: 10.1126/science.aat0958

Chen, X., Liu, J., Janssen, J. M., and Goncalves, M. (2017c). The chromatin structure differentially impacts high-specificity CRISPR-Cas9 nuclease strategies. *Mol. Ther. Nucleic Acids* 8, 558–563. doi: 10.1016/j.omtn.2017.08.005

Chen, X., Rinsma, M., Janssen, J. M., Liu, J., Maggio, I., and Goncalves, M. A. (2016). Probing the impact of chromatin conformation on genome editing tools. *Nucleic Acids Res.* 44, 6482–6492. doi: 10.1093/nar/gkw524

Chiarle, R., Zhang, Y., Frock, R. L., Lewis, S. M., Molinie, B., Ho, Y. J., et al. (2011). Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* 147, 107–119. doi: 10.1016/j.cell.2011.07.049

Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S., et al. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24, 132–141. doi: 10.1101/gr.162339.113

Chung, C. H., Allen, A. G., Sullivan, N. T., Atkins, A., Nonnemacher, M. R., Wigdahl, B., et al. (2020). Computational analysis concerning the impact of DNA accessibility on CRISPR-Cas9 cleavage efficiency. *Mol. Ther.* 28, 19–28. doi: 10.1016/j.ymthe.2019.10.008

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi: 10.1126/science.1231143

Connelly, J. C., and Leach, D. R. (2002). Tethering on the brink: the evolutionarily conserved Mre11-Rad50 complex. *Trends Biochem. Sci.* 27, 410–418. doi: 10.1016/S0968-0004(02)02144-8

Costello, M., Pugh, T. J., Fennell, T. J., Stewart, C., Lichtenstein, L., Meldrim, J. C., et al. (2013). Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res.* 41: e67. doi: 10.1093/nar/gks1443

Cradick, T. J., Fine, E. J., Antico, C. J., and Bao, G. (2013). CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* 41, 9584–9592. doi: 10.1093/nar/gkt714

Crosetto, N., Mitra, A., Silva, M. J., Bienko, M., Dojer, N., Wang, Q., et al. (2013). Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* 10, 361–365. doi: 10.1038/nmeth.2408

Dabrowska, M., Czubak, K., Juzwa, W., Krzyzosiak, W. J., Olejniczak, M., and Kozlowski, P. (2018). qEva-CRISPR: a method for quantitative evaluation of CRISPR/Cas-mediated genome editing in target and off-target sites. *Nucleic Acids Res.* 46: e101. doi: 10.1093/nar/gky505

Daer, R. M., Cutts, J. P., Brafman, D. A., and Haynes, K. A. (2017). The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. *ACS Synth. Biol.* 6, 428–438. doi: 10.1021/acssynbio.5b00299

Dampier, W., Nonnemacher, M. R., Sullivan, N. T., Jacobson, J. M., and Wigdahl, B. (2014). HIV excision utilizing CRISPR/Cas9 technology: attacking the proviral quasispecies in reservoirs to achieve a cure. *MOJ Immunol.* 1. doi: 10.15406/moji.2014.01.00022

Dampier, W., Sullivan, N. T., Chung, C. H., Mell, J. C., Nonnemacher, M. R., and Wigdahl, B. (2017). Designing broad-spectrum anti-HIV-1 gRNAs to target patient-derived variants. *Sci. Rep.* 7:14413. doi: 10.1038/s41598-017-12612-z

Dampier, W., Sullivan, N. T., Mell, J. C., Pirrone, V., Ehrlich, G., C., et al. (2018). Broad spectrum and personalized gRNAs for CRISPR/Cas9 HIV-1 therapeutics. *AIDS Res. Hum. Retroviruses*. 34, 950–960. doi: 10.1089/aid.2017.0274

Darcis, G., Binda, C. S., Klaver, B., Herrera-Carrillo, E., Berkhout, B., and Das, A. T. (2019). The impact of HIV-1 genetic diversity on CRISPR-Cas9 antiviral activity and viral escape. *Viruses* 11. doi: 10.3390/v11030255

Dash, P. K., Kaminski, R., Bella, R., Su, H., Mathews, S., Ahooyi, T. M., et al. (2019). Sequential LASER ART and CRISPR treatments eliminate HIV-1 in a subset of infected humanized mice. *Nat. Commun.* 10:2753. doi: 10.1038/s41467-019-10366-y

Deng, Q., Chen, Z., Shi, L., and Lin, H. (2018). Developmental progress of CRISPR/Cas9 and its therapeutic applications for HIV-1 infection. *Rev. Med. Virol.* 28:e1998. doi: 10.1002/rmv.1998

Do, H., Wong, S. Q., Li, J., and Dobrovic, A. (2013). Reducing sequence artifacts in amplicon-based massively parallel sequencing of formalin-fixed paraffin-embedded DNA by enzymatic depletion of uracil-containing templates. *Clin. Chem.* 59, 1376–1383. doi: 10.1373/clinchem.2012.202390

Dobosy, J. R., Rose, S. D., Beltz, K. R., Rupp, S. M., Powers, K. M., Behlke, M. A., et al. (2011). RNase H-dependent PCR (rhPCR): improved specificity and single nucleotide polymorphism detection using blocked cleavable primers. *BMC Biotechnol.* 11:80. doi: 10.1186/1472-6750-11-80

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34, 184–191. doi: 10.1038/nbt.3437

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., et al. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* 32, 1262–1267. doi: 10.1038/nbt.3026

Dorsett, Y., Zhou, Y., Tubbs, A. T., Chen, B. R., Purman, C., Lee, B. S., et al. (2014). HCoDES reveals chromosomal DNA end structures with single-nucleotide resolution. *Mol. Cell* 56, 808–818. doi: 10.1016/j.molcel.2014.10.024

Doyon, Y., Vo, T. D., Mendel, M. C., Greenberg, S. G., Wang, J., Xia, D. F., et al. (2011). Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods* 8, 74–79. doi: 10.1038/nmeth.1539

Duan, J., Lu, G., Xie, Z., Lou, M., Luo, J., Guo, L., et al. (2014). Genome-wide identification of CRISPR/Cas9 off-targets in human genome. *Cell Res.* 24, 1009–1012. doi: 10.1038/cr.2014.87

Ebina, H., Misawa, N., Kanemura, Y., and Koyanagi, Y. (2013). Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci. Rep.* 3:2510. doi: 10.1038/srep02510

Elliott, B., and Jasin, M. (2002). Double-strand breaks and translocations in cancer. *Cell. Mol. Life Sci.* 59, 373–385. doi: 10.1007/s00018-002-8429-3

Foss, D. V., Hochstrasser, M. L., and Wilson, R. C. (2019). Clinical applications of CRISPR-based genome editing and diagnostics. *Transfusion* 59, 1389–1399. doi: 10.1111/trf.15126

Frietze, S., and Farnham, P. J. (2011). Transcription factor effector domains. *Subcell. Biochem.* 52, 261–277. doi: 10.1007/978-90-481-9069-0_12

Frock, R. L., Hu, J., Meyers, R. M., Ho, Y. J., Kii, E., and Alt, F. W. (2015). Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* 33, 179–186. doi: 10.1038/nbt.3101

Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. U.S.A.* 85, 8998–9002. doi: 10.1073/pnas.85.23.8998

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826. doi: 10.1038/nbt.2623

Gabriel, R., Eckenberg, R., Paruzynski, A., Bartholomae, C. C., Nowrouzi, A., Arens, A., et al. (2009). Comprehensive genomic access to vector integration in clinical gene therapy. *Nat. Med.* 15, 1431–1436. doi: 10.1038/nm.2057

Gabriel, R., Lombardo, A., Arens, A., Miller, J. C., Genovese, P., Kaeppel, C., et al. (2011). An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* 29, 816–823. doi: 10.1038/nbt.1948

Gao, Q., Dong, X., Xu, Q., Zhu, L., Wang, F., Hou, Y., et al. (2019). Therapeutic potential of CRISPR/Cas9 gene editing in engineered T-cell therapy. *Cancer Med.* 8, 4254–4264. doi: 10.1002/cam4.2257

Gao, Z., Fan, M., Das, A. T., Herrera-Carrillo, E., and Berkhout, B. (2020). Extinction of all infectious HIV in cell culture by the CRISPR-Cas12a system with only a single crRNA. *Nucleic Acids Res.* 48, 5527–5539. doi: 10.1093/nar/gkaa226

Giannoukos, G., Ciulla, D. M., Marco, E., Abdulkerim, H. S., Barrera, L. A., Bothmer, A., et al. (2018). UDiTaS, a genome editing detection method for indels and genome rearrangements. *BMC Genomics* 19:212. doi: 10.1186/s12864-018-4561-9

Guo, J., Gaj, T., and Barbas, C. F. 3rd (2010). Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J. Mol. Biol.* 400, 96–107. doi: 10.1016/j.jmb.2010.04.060

Guschin, D. Y., Waite, A. J., Katibah, G. E., Miller, J. C., Holmes, M. C., and Rebar, E. J. (2010). A rapid and general assay for monitoring endogenous gene modification. *Methods Mol. Biol.* 649, 247–256. doi: 10.1007/978-1-60761-753-2_15

Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., et al. (2003). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* 348, 255–256. doi: 10.1056/NEJM200301163480314

Haeussler, M., Schonig, K., Eckert, H., Eschstruth, A., Mianne, J., Renaud, J. B., et al. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* 17:148. doi: 10.1186/s13059-016-1012-2

Heigwer, F., Kerr, G., and Boutros, M. (2014). E-CRISP: fast CRISPR target site identification. *Nat. Methods* 11, 122–123. doi: 10.1038/nmeth.2812

Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* 29, 731–734. doi: 10.1038/nbt.1927

Hoffman, E. A., Frey, B. L., Smith, L. M., and Auble, D. T. (2015). Formaldehyde crosslinking: a tool for the study of chromatin complexes. *J. Biol. Chem.* 290, 26404–26411. doi: 10.1074/jbc.R115.651679

Hou, P., Chen, S., Wang, S., Yu, X., Chen, Y., Jiang, M., et al. (2015). Genome editing of CXCR4 by CRISPR/cas9 confers cells resistant to HIV-1 infection. *Sci. Rep.* 5:15577. doi: 10.1038/srep15577

Hsu, P. D., Lander, E. S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278. doi: 10.1016/j.cell.2014.05.010

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi: 10.1038/nbt.2647

Hu, J., Meyers, R. M., Dong, J., Panchakshari, R. A., Alt, F. W., and Frock, R. L. (2016). Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc.* 11, 853–871. doi: 10.1038/nprot.2016.043

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63. doi: 10.1038/nature26155

Hu, W., Kaminski, R., Yang, F., Zhang, Y., Cosentino, L., Li, F., et al. (2014). RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc. Natl. Acad. Sci. U.S.A.* 111, 11461–11466. doi: 10.1073/pnas.1405186111

Huang, J., Wang, Y., and Zhao, J. (2018). CRISPR editing in biological and biomedical investigation. *J. Cell. Physiol.* 233, 3875–3891. doi: 10.1002/jcp.26141

Iacovoni, J. S., Caron, P., Lassadi, I., Nicolas, E., Massip, L., Trouche, D., et al. (2010). High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. *EMBO J.* 29, 1446–1457. doi: 10.1038/emboj.2010.38

Iyer, V., Shen, B., Zhang, W., Hodgkins, A., Keane, T., Huang, X., et al. (2015). Off-target mutations are rare in Cas9-modified mice. *Nat. Methods* 12, 479–479. doi: 10.1038/nmeth.3408

Jensen, K. T., Floe, L., Petersen, T. S., Huang, J., Xu, F., Bolund, L., et al. (2017). Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett.* 591, 1892–1901. doi: 10.1002/1873-3468.12707

Ji, H., Jiang, Z., Lu, P., Ma, L., Li, C., Pan, H., et al. (2016). Specific reactivation of latent HIV-1 by dCas9-SunTag-VP64-mediated guide RNA Targeting the HIV-1 promoter. *Mol. Ther.* 24, 508–521. doi: 10.1038/mt.2016.7

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10.1126/science.1225829

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. *Elife* 2:e00471. doi: 10.7554/eLife.00471.009

Johansson, H., Isaksson, M., Sorqvist, E. F., Roos, F., Stenberg, J., Sjoblom, T., et al. (2011). Targeted resequencing of candidate genes using selector probes. *Nucleic Acids Res.* 39:e8. doi: 10.1093/nar/gkq1005

Kaminski, R., Bella, R., Yin, C., Otte, J., Ferrante, P., Gendelman, H. E., et al. (2016a). Excision of HIV-1 DNA by gene editing: a proof-of-concept *in vivo* study. *Gene Ther.* 23, 690–695. doi: 10.1038/gt.2016.41

Kaminski, R., Chen, Y., Fischer, T., Tedaldi, E., Napoli, A., Zhang, Y., et al. (2016b). Elimination of HIV-1 genomes from human T-lymphoid cells by CRISPR/Cas9 gene editing. *Sci. Rep.* 6:22555. doi: 10.1038/srep28213

Kaminski, R., Chen, Y., Salkind, J., Bella, R., Young, W. B., Ferrante, P., et al. (2016c). Negative feedback regulation of HIV-1 by gene editing strategy. *Sci. Rep.* 6:31527. doi: 10.1038/srep31527

Karimian, A., Azizian, K., Parsian, H., Rafieian, S., Shafiei-Irannejad, V., Kheyrollah, M., et al. (2019). CRISPR/Cas9 technology as a potent molecular tool for gene therapy. *J. Cell. Physiol.* 234, 12267–12277. doi: 10.1002/jcp.27972

Kaushik, A., Yndart, A., Atluri, V., Tiwari, S., Tomitaka, A., Gupta, P., et al. (2019). Magnetically guided non-invasive CRISPR-Cas9/gRNA delivery across blood-brain barrier to eradicate latent HIV-1 infection. *Sci. Rep.* 9:3928. doi: 10.1038/s41598-019-40222-4

Khair, L., Baker, R. E., Linehan, E. K., Schrader, C. E., and Stavnezer, J. (2015). Nbs1 ChIP-Seq identifies off-target DNA double-strand breaks induced by AID in activated splenic B cells. *PLoS Genet.* 11:e1005438. doi: 10.1371/journal.pgen.1005438

Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H. R., et al. (2015). Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* 12, 237–243. doi: 10.1038/nmeth.3284

Kim, D., and Kim, J. S. (2018). DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA. *Genome Res.* 28, 1894–1900. doi: 10.1101/gr.236620.118

Kim, D., Kim, S., Kim, S., Park, J., and Kim, J. S. (2016). Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq. *Genome Res.* 26, 406–415. doi: 10.1101/gr.199588.115

Kim, H. J., Lee, H. J., Kim, H., Cho, S. W., and Kim, J. S. (2009). Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* 19, 1279–1288. doi: 10.1101/gr.089417.108

Kim, T. H., and Ren, B. (2006). Genome-wide analysis of protein-DNA interactions. *Annu. Rev. Genomics Hum. Genet.* 7, 81–102. doi: 10.1146/annurev.genom.7.080505.115634

Klein, I. A., Resch, W., Jankovic, M., Oliveira, T., Yamane, A., Nakahashi, H., et al. (2011). Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* 147, 95–106. doi: 10.1016/j.cell.2011.07.048

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., et al. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495. doi: 10.1038/nature16526

Knight, S. C., Xie, L., Deng, W., Guglielmi, B., Witkowsky, L. B., Bosanac, L., et al. (2015). Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science* 350, 823–826. doi: 10.1126/science.aac6572

Koo, T., Lee, J., and Kim, J. S. (2015). Measuring and reducing off-target activities of programmable nucleases including CRISPR-Cas9. *Mol. Cells* 38, 475–481. doi: 10.14348/molcells.2015.0103

Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 765–771. doi: 10.1038/nbt.4192

Kunze, C., Borner, K., Kienle, E., Orschmann, T., Rusha, E., Schneider, M., et al. (2018). Synthetic AAV/CRISPR vectors for blocking HIV-1 expression in persistently infected astrocytes. *Glia* 66, 413–427. doi: 10.1002/glia.23254

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* 32, 677–683. doi: 10.1038/nbt.2916

Lazzarotto, C. R., Malinin, N. L., Li, Y., Zhang, R., Yang, Y., Lee, G., et al. (2020). CHANGE-seq reveals genetic and epigenetic effects on CRISPR-Cas9 genome-wide activity. *Nat. Biotechnol.* 38, 1317–1327. doi: 10.1038/s41587-020-0555-7

Lazzarotto, C. R., Nguyen, N. T., Tang, X., Malagon-Lopez, J., Guo, J. A., Aryee, M. J., et al. (2018). Defining CRISPR-Cas9 genome-wide nuclease activities with CIRCLE-seq. *Nat. Protoc.* 13, 2615–2642. doi: 10.1038/s41596-018-0055-0

Lebbink, R. J., de Jong, D. C., Wolters, F., Kruse, E. M., van Ham, P. M., Wiertz, E. J., et al. (2017). A combinational CRISPR/Cas9 gene-editing approach can halt HIV replication and prevent viral escape. *Sci. Rep.* 7:41968. doi: 10.1038/srep41968

Lee, J. K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y. H., et al. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. *Nat. Commun.* 9:3048. doi: 10.1038/s41467-018-05477-x

Lensing, S. V., Marsico, G., Hansel-Hertsch, R., Lam, E. Y., Tannahill, D., and Balasubramanian, S. (2016). DSBCapture: *in situ* capture and sequencing of DNA breaks. *Nat. Methods* 13, 855–857. doi: 10.1038/nmeth.3960

Li, T., Zhu, L., Xiao, B., Gong, Z., Liao, Q., and Guo, J. (2019). CRISPR-Cpf1-mediated genome editing and gene regulation in human cells. *Biotechnol. Adv.* 37, 21–27. doi: 10.1016/j.biotechadv.2018.10.013

Liang, X., Potter, J., Kumar, S., Ravinder, N., and Chesnut, J. D. (2017). Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *J. Biotechnol.* 241, 136–146. doi: 10.1016/j.jbiotec.2016.11.011

Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., et al. (2015). Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J. Biotechnol.* 208, 44–53. doi: 10.1016/j.jbiotec.2015.04.024

Liao, H. K., Gu, Y., Diaz, A., Marlett, J., Takahashi, Y., Li, M., et al. (2015). Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat. Commun.* 6:6413. doi: 10.1038/ncomms7413

Liao, S., Tammaro, M., and Yan, H. (2016). The structure of ends determines the pathway choice and Mre11 nuclease dependency of DNA double-strand break repair. *Nucleic Acids Res.* 44, 5689–5701. doi: 10.1093/nar/gkw274

Limsirichai, P., Gaj, T., and Schaffer, D. V. (2016). CRISPR-mediated activation of latent HIV-1 expression. *Mol. Ther.* 24, 499–507. doi: 10.1038/mt.2015.213

Lin, Y., Cradick, T. J., Brown, M. T., Deshmukh, H., Ranjan, P., Sarode, N., et al. (2014). CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res.* 42, 7473–7485. doi: 10.1093/nar/gku402

Link, R. W., Nonnemacher, M. R., Wigdahl, B., and Dampier, W. (2018). Prediction of human immunodeficiency virus type 1 subtype-specific off-target effects arising from CRISPR-Cas9 gene editing therapy. *CRISPR J.* 1, 294–302. doi: 10.1089/crispr.2018.0020

Liu, H., Wei, Z., Dominguez, A., Li, Y., Wang, X., and Qi, L. S. (2015). CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. *Bioinformatics* 31, 3676–3678. doi: 10.1093/bioinformatics/btv423

Luo, G., Yao, M. S., Bender, C. F., Mills, M., Bladl, A. R., Bradley, A., et al. (1999). Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. *Proc. Natl. Acad. Sci. U.S.A.* 96, 7376–7381. doi: 10.1073/pnas.96.13.7376

Madabhushi, R., Gao, F., Pfenning, A. R., Pan, L., Yamakawa, S., Seo, J., et al. (2015). Activity-induced DNA breaks govern the expression of neuronal early-response genes. *Cell* 161, 1592–1605. doi: 10.1016/j.cell.2015.05.032

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., et al. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833–838. doi: 10.1038/nbt.2675

Mancuso, P., Chen, C., Kaminski, R., Gordon, J., Liao, S., Robinson, J. A., et al. (2020). CRISPR based editing of SIV proviral DNA in ART treated non-human primates. *Nat. Commun.* 11:6065. doi: 10.1038/s41467-020-19821-7

Manghwar, H., Li, B., Ding, X., Hussain, A., Lindsey, K., Zhang, X., et al. (2020). CRISPR/Cas systems in genome editing: methodologies and tools for sgRNA design, off-target evaluation, and strategies to mitigate off-target effects. *Adv. Sci.* 7:1902312. doi: 10.1002/advs.201902312

Mefferd, A. L., Bogerd, H. P., Irwan, I. D., and Cullen, B. R. (2018). Insights into the mechanisms underlying the inactivation of HIV-1 proviruses by CRISPR/Cas. *Virology* 520, 116–126. doi: 10.1016/j.virol.2018.05.016

Meier, J. A., Zhang, F., and Sanjana, N. E. (2017). GUIDES: sgRNA design for loss-of-function screens. *Nat. Methods* 14, 831–832. doi: 10.1038/nmeth.4423

Meynert, A. M., Ansari, M., FitzPatrick, D. R., and Taylor, M. S. (2014). Variant detection sensitivity and biases in whole genome and exome sequencing. *BMC Bioinformatics* 15:247. doi: 10.1186/1471-2105-15-247

Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y. L., Rupniewski, I., et al. (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* 25, 778–785. doi: 10.1038/nbt1319

Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M., and Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* 42, W401–407. doi: 10.1093/nar/gku410

Moreno-Herrero, F., de Jager, M., Dekker, N. H., Kanaar, R., Wyman, C., and Dekker, C. (2005). Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. *Nature* 437, 440–443. doi: 10.1038/nature03927

Murray, A. J., Kwon, K. J., Farber, D. L., and Siliciano, R. F. (2016). The latent reservoir for HIV-1: how immunologic memory and clonal expansion contribute to HIV-1 persistence. *J. Immunol.* 197, 407–417. doi: 10.4049/jimmunol.1600343

Mussolino, C., Morbitzer, R., Lutge, F., Dannemann, N., Lahaye, T., and Cathomen, T. (2011). A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 39, 9283–9293. doi: 10.1093/nar/gkr597

Neiman, M., Sundling, S., Gronberg, H., Hall, P., Czene, K., Lindberg, J., et al. (2012). Library preparation and multiplex capture for massive parallel sequencing applications made efficient and easy. *PLoS ONE* 7: e48616. doi: 10.1371/journal.pone.0048616

Nobles, C. L., Reddy, S., Salas-McKee, J., Liu, X., June, C. H., Melenhorst, J. J., et al. (2019). iGUIDE: an improved pipeline for analyzing CRISPR cleavage specificity. *Genome Biol.* 20:14. doi: 10.1186/s13059-019-1625-3

O'Geen, H., Henry, I. M., Bhakta, M. S., Meckler, J. F., and Segal, D. J. (2015). A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic Acids Res.* 43, 3389–3404. doi: 10.1093/nar/gkv137

Oh, J., and Symington, L. S. (2018). Role of the Mre11 complex in preserving genome integrity. *Genes* 9. doi: 10.3390/genes9120589

Ophinni, Y., Inoue, M., Kotaki, T., and Kameoka, M. (2018). CRISPR/Cas9 system targeting regulatory genes of HIV-1 inhibits viral replication in infected T-cell cultures. *Sci. Rep.* 8:7784. doi: 10.1038/s41598-018-26190-1

Panfil, A. R., London, J. A., Green, P. L., and Yoder, K. E. (2018). CRISPR/Cas9 genome editing to disable the latent HIV-1 provirus. *Front. Microbiol.* 9:3107. doi: 10.3389/fmicb.2018.03107

Paruzynski, A., Arens, A., Gabriel, R., Bartholomae, C. C., Scholz, S., Wang, W., et al. (2010). Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nat. Protoc.* 5, 1379–1395. doi: 10.1038/nprot.2010.87

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., and Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* 31, 839–843. doi: 10.1038/nbt.2673

Pattanayak, V., Ramirez, C. L., Joung, J. K., and Liu, D. R. (2011). Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. *Nat. Methods* 8, 765–770. doi: 10.1038/nmeth.1670

Picelli, S., Bjorklund, A. K., Reinius, B., Sagasser, S., Winberg, G., and Sandberg, R. (2014). Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* 24, 2033–2040. doi: 10.1101/gr.177881.114

Pollack, R. A., Jones, R. B., Pertea, M., Bruner, K. M., Martin, A. R., Thomas, A. S., et al. (2017). Defective HIV-1 proviruses are expressed and can be recognized by cytotoxic T lymphocytes, which shape the proviral landscape. *Cell Host Microbe* 21, 494–506 e494. doi: 10.1016/j.chom.2017.03.008

Qi, J., Ding, C., Jiang, X., and Gao, Y. (2020). Advances in developing CAR T-cell therapy for HIV cure. *Front. Immunol.* 11:361. doi: 10.3389/fimmu.2020.00361

Ramiro, A. R., Jankovic, M., Callen, E., Difilippantonio, S., Chen, H. T., McBride, K. M., et al. (2006). Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. *Nature* 440, 105–109. doi: 10.1038/nature04495

Rodriguez, R., Miller, K. M., Forment, J. V., Bradshaw, C. R., Nikan, M., Britton, S., et al. (2012). Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* 8, 301–310. doi: 10.1038/nchembio.780

Roychoudhury, P., De Silva Feelixge, H., Reeves, D., Mayer, B. T., Stone, D., Schiffer, J. T., et al. (2018). Viral diversity is an obligate consideration in CRISPR/Cas9 designs for targeting the HIV reservoir. *BMC Biol.* 16:75. doi: 10.1186/s12915-018-0544-1

Saayman, S. M., Lazar, D. C., Scott, T. A., Hart, J. R., Takahashi, M., Burnett, J. C., et al. (2016). Potent and targeted activation of latent HIV-1 using the CRISPR/dCas9 activator complex. *Mol. Ther.* 24, 488–498. doi: 10.1038/mt.2015.202

Schmid-Burgk, J. L., Gao, L., Li, D., Gardner, Z., Strecker, J., Lash, B., et al. (2020). Highly parallel profiling of Cas9 variant specificity. *Mol. Cell.* 78, 794–800. doi: 10.1016/j.molcel.2020.02.023

Schmidt, M., Schwarzwaelder, K., Bartholomae, C., Zaoui, K., Ball, C., Pilz, I., et al. (2007). High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nat. Methods* 4, 1051–1057. doi: 10.1038/nmeth1103

Shapiro, J., Iancu, O., Jacobi, A. M., McNeill, M. S., Turk, R., Rettig, G. R., et al. (2020). Increasing CRISPR efficiency and measuring its specificity in HSPCs using a clinically relevant system. *Mol. Ther. Methods Clin. Dev.* 17, 1097–1107. doi: 10.1016/j.omtm.2020.04.027

Shibata, A., Moiani, D., Arvai, A. S., Perry, J., Harding, S. M., Genois, M. M., et al. (2014). DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. *Mol. Cell* 53, 7–18. doi: 10.1016/j.molcel.2013.11.003

Siliciano, R. F., and Greene, W. C. (2011). HIV latency. *Cold Spring Harb. Perspect. Med.* 1:a007096. doi: 10.1101/cshperspect.a007096

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88. doi: 10.1126/science.aad5227

Smith, C., Gore, A., Yan, W., Abalde-Atristain, L., Li, Z., He, C., et al. (2014). Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell* 15, 12–13. doi: 10.1016/j.stem.2014.06.011

Spehalski, E., Capper, K. M., Smith, C. J., Morgan, M. J., Dinkelmann, M., Buis, J., et al. (2017). MRE11 promotes tumorigenesis by facilitating resistance to oncogene-induced replication stress. *Cancer Res.* 77, 5327–5338. doi: 10.1158/0008-5472.CAN-17-1355

Staszewski, O., Baker, R. E., Ucher, A. J., Martier, R., Stavnezer, J., and Guikema, J. E. (2011). Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig Loci in activated B cells. *Mol. Cell* 41, 232–242. doi: 10.1016/j.molcel.2011.01.007

Su, H., Sravanam, S., Gorantla, S., Kaminski, R., Khalili, K., Poluektova, L., et al. (2020). Amplification of replication competent HIV-1 by adoptive transfer of human cells from infected humanized mice. *Front. Cell. Infect. Microbiol.* 10:38. doi: 10.3389/fcimb.2020.00038

Sullivan, N. T., Dampier, W., Chung, C. H., Allen, A. G., Atkins, A., Pirrone, V., et al. (2019). Novel gRNA design pipeline to develop broad-spectrum CRISPR/Cas9 gRNAs for safe targeting of the HIV-1 quasispecies in patients. *Sci. Rep.* 9:17088. doi: 10.1038/s41598-019-52353-9

Suzuki, K., Yu, C., Qu, J., Li, M., Yao, X., Yuan, T., et al. (2014). Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell Stem Cell* 15, 31–36. doi: 10.1016/j.stem.2014.06.016

Syed, A., and Tainer, J. A. (2018). The MRE11-RAD50-NBS1 complex conducts the orchestration of damage signaling and outcomes to stress in DNA replication and repair. *Annu. Rev. Biochem.* 87, 263–294. doi: 10.1146/annurev-biochem-062917-012415

Szczepek, M., Brondani, V., Buchel, J., Serrano, L., Segal, D. J., and Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.* 25, 786–793. doi: 10.1038/nbt1317

Szilard, R. K., Jacques, P. E., Laramee, L., Cheng, B., Galicia, S., Bataille, A. R., et al. (2010). Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX. *Nat. Struct. Mol. Biol.* 17, 299–305. doi: 10.1038/nsmb.1754

Tsai, S. Q., and Joung, J. K. (2016). Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat. Rev. Genet.* 17, 300–312. doi: 10.1038/nrg.2016.28

Tsai, S. Q., Nguyen, N. T., Malagon-Lopez, J., Topkar, V. V., Aryee, M. J., and Joung, J. K. (2017). CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* 14, 607–614. doi: 10.1038/nmeth.4278

Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187–197. doi: 10.1038/nbt.3117

Tycko, J., Myer, V. E., and Hsu, P. D. (2016). Methods for optimizing CRISPR-Cas9 genome editing specificity. *Mol. Cell* 63, 355–370. doi: 10.1016/j.molcel.2016.07.004

Ueda, S., Ebina, H., Kanemura, Y., Misawa, N., and Koyanagi, Y. (2016). Anti-HIV-1 potency of the CRISPR/Cas9 system insufficient to fully inhibit viral replication. *Microbiol. Immunol.* 60, 483–496. doi: 10.1111/1348-0421.12395

Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., et al. (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* 24, 1216–1224. doi: 10.1038/s41591-018-0137-0

van den Bosch, M., Bree, R. T., and Lowndes, N. F. (2003). The MRN complex: coordinating and mediating the response to broken chromosomes. *EMBO Rep.* 4, 844–849. doi: 10.1038/sj.embor.embor925

Varley, K. E., and Mitra, R. D. (2008). Nested Patch PCR enables highly multiplexed mutation discovery in candidate genes. *Genome Res.* 18, 1844–1850. doi: 10.1101/gr.078204.108

Veres, A., Gosis, B. S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., et al. (2014). Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell* 15, 27–30. doi: 10.1016/j.stem.2014.04.020

Vouillot, L., Thelie, A., and Pollet, N. (2015). Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3* 5, 407–415. doi: 10.1534/g3.114.015834

Wang, G., Zhao, N., Berkhout, B., and Das, A. T. (2016a). A combinatorial CRISPR-Cas9 attack on HIV-1 DNA extinguishes all infectious provirus in infected T cell cultures. *Cell. Rep.* 17, 2819–2826. doi: 10.1016/j.celrep.2016.11.057

Wang, G., Zhao, N., Berkhout, B., and Das, A. T. (2016b). CRISPR-Cas9 can inhibit HIV-1 replication but NHEJ repair facilitates virus escape. *Mol. Ther.* 24, 522–526. doi: 10.1038/mt.2016.24

Wang, J., Zhang, X., Cheng, L., and Luo, Y. (2020). An overview and metanalysis of machine and deep learning-based CRISPR gRNA design tools. *RNA Biol.* 17, 13–22. doi: 10.1080/15476286.2019.1669406

Wang, X., Wang, Y., Wu, X., Wang, J., Wang, Y., Qiu, Z., et al. (2015). Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat. Biotechnol.* 33, 175–178. doi: 10.1038/nbt.3127

Wang, Z., Wang, W., Cui, Y. C., Pan, Q., Zhu, W., Gendron, P., et al. (2018). HIV-1 employs multiple mechanisms to resist Cas9/single guide RNA targeting the viral primer binding site. *J. Virol.* 92. doi: 10.1128/JVI.01135-18

Wienert, B., Wyman, S. K., Richardson, C. D., Yeh, C. D., Akcakaya, P., Porritt, M. J., et al. (2019). Unbiased detection of CRISPR off-targets *in vivo* using DISCOVER-Seq. *Science* 364, 286–289. doi: 10.1101/469635

Wienert, B., Wyman, S. K., Yeh, C. D., Conklin, B. R., and Corn, J. E. (2020). CRISPR off-target detection with DISCOVER-seq. *Nat. Protoc.* 15, 1775–1799. doi: 10.1038/s41596-020-0309-5

Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* 32, 670–676. doi: 10.1038/nbt.2889

Xu, L., Yang, H., Gao, Y., Chen, Z., Xie, L., Liu, Y., et al. (2017). CRISPR/Cas9-mediated CCR5 ablation in human hematopoietic stem/progenitor cells confers HIV-1 resistance *in vivo*. *Mol. Ther.* 25, 1782–1789. doi: 10.1016/j.ymthe.2017.04.027

Yamane, A., Robbiani, D. F., Resch, W., Bothmer, A., Nakahashi, H., Oliveira, T., et al. (2013). RPA accumulation during class switch recombination represents 5'-3′ DNA-end resection during the S-G2/M phase of the cell cycle. *Cell. Rep.* 3, 138–147. doi: 10.1016/j.celrep.2012.12.006

Yan, W. X., Mirzazadeh, R., Garnerone, S., Scott, D., Schneider, M. W., Kallas, T., et al. (2017). BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun.* 8:15058. doi: 10.1038/ncomms15058

Yang, H., Wang, H., Shivalila, C. S., Cheng, A. W., Shi, L., and Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370–1379. doi: 10.1016/j.cell.2013.08.022

Yang, Z., Steentoft, C., Hauge, C., Hansen, L., Thomsen, A. L., Niola, F., et al. (2015). Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res.* 43:e59. doi: 10.1093/nar/gkv126

Yin, C., Zhang, T., Li, F., Yang, F., Putatunda, R., Young, W. B., et al. (2016). Functional screening of guide RNAs targeting the regulatory and structural HIV-1 viral genome for a cure of AIDS. *AIDS* 30, 1163–1174. doi: 10.1097/QAD.0000000000001079

Yoder, K. E., and Bundschuh, R. (2016). Host double strand break repair generates HIV-1 strains resistant to CRISPR/Cas9. *Sci. Rep.* 6:29530. doi: 10.1038/srep29530

Yousem, S. A., Dacic, S., Nikiforov, Y. E., and Nikiforova, M. (2013). Pulmonary Langerhans cell histiocytosis: profiling of multifocal tumors using next-generation sequencing identifies concordant occurrence of BRAF V600E mutations. *Chest* 143, 1679–1684. doi: 10.1378/chest.12-1917

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A. H., Shea, D., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* 26, 535–541. doi: 10.1038/s41591-020-0790-y

Zhang, D., Hurst, T., Duan, D., and Chen, S. J. (2019). Unified energetics analysis unravels SpCas9 cleavage activity for optimal gRNA design. *Proc. Natl. Acad. Sci. U.S.A.* 116, 8693–8698. doi: 10.1073/pnas.1820523116

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., et al. (2014). The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol. J.* 12, 797–807. doi: 10.1111/pbi.12200

Zhao, N., Wang, G., Das, A. T., and Berkhout, B. (2017). Combinatorial CRISPR-Cas9 and RNA interference attack on HIV-1 DNA and RNA can lead to cross-resistance. *Antimicrob. Agents Chemother.* 61. doi: 10.1128/AAC.01486-17

Zheng, Z., Advani, A., Melefors, O., Glavas, S., Nordstrom, H., Ye, W., et al. (2010). Titration-free massively parallel pyrosequencing using trace amounts of starting material. *Nucleic Acids Res.* 38:e137. doi: 10.1093/nar/gkq332

Zheng, Z., Advani, A., Melefors, O., Glavas, S., Nordstrom, H., Ye, W., et al. (2011). Titration-free 454 sequencing using Y adapters. *Nat. Protoc.* 6, 1367–1376. doi: 10.1038/nprot.2011.369

Zheng, Z., Liebers, M., Zhelyazkova, B., Cao, Y., Panditi, D., Lynch, K. D., et al. (2014). Anchored multiplex PCR for targeted next-generation sequencing. *Nat. Med.* 20, 1479–1484. doi: 10.1038/nm.3729

Zhou, Z. X., Zhang, M. J., Peng, X., Takayama, Y., Xu, X. Y., Huang, L. Z., et al. (2013). Mapping genomic hotspots of DNA damage by a single-strand-DNA-compatible and strand-specific ChIP-seq method. *Genome Res.* 23, 705–715. doi: 10.1101/gr.146357.112

Zhu, J., Petersen, S., Tessarollo, L., and Nussenzweig, A. (2001). Targeted disruption of the Nijmegen breakage syndrome gene NBS1 leads to early embryonic lethality in mice. *Curr. Biol.* 11, 105–109. doi: 10.1016/S0960-9822(01)00019-7

Zhu, W., Lei, R., Le Duff, Y., Li, J., Guo, F., Wainberg, M. A., et al. (2015). The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology* 12:22. doi: 10.1186/s12977-015-0150-z

Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., et al. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364, 289–292. doi: 10.1126/science.aav9973

Author Disclaimer: The contents of the paper were solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

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

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

frontiers in Genome Editing

BRIEF RESEARCH REPORT
published: 12 November 2021
doi: 10.3389/fgeed.2021.786893

piggyPrime: High-Efficacy Prime Editing in Human Cells Using piggyBac-Based DNA Transposition

*Jonas Holst Wolff, Jakob Haldrup, Emil Aagaard Thomsen, Sofie Andersen and Jacob Giehm Mikkelsen**

Department of Biomedicine, Aarhus University, Aarhus, Denmark

Prime editing is a novel genome editing technology that allows a wide range of tailored genomic alterations. Prime editing does not involve homologous recombination, but suffers from low efficacy. Here, we demonstrate piggyPrime, a transfected single-vector system based on piggyBac DNA transposition for genomic integration of all prime editing components in human cells allowing easy and effective transgenesis with prime editing efficacies up to 100% in cell lines.

Keywords: prime editing, CRISPR, prime editing guide RNA, pegRNA, piggyBac, DNA transposition

OPEN ACCESS

Edited by:
Ayal Hendel,
Bar-Ilan University, Israel

Reviewed by:
Mustapha Aouida,
Hamad bin Khalifa University, Qatar
Adi Barzel,
Tel Aviv University, Israel
Mario Amendola,
Institut National de la Santé et de la Recherche Médicale (INSERM), France

****Correspondence:***
Jacob Giehm Mikkelsen
giehm@biomed.au.dk

Specialty section:
This article was submitted to Genome Editing Tools and Mechanisms, a section of the journal Frontiers in Genome Editing

Received: *30 September 2021*
Accepted: *26 October 2021*
Published: *12 November 2021*

Citation:
Wolff JH, Haldrup J, Thomsen EA, Andersen S and Mikkelsen JG (2021) piggyPrime: High-Efficacy Prime Editing in Human Cells Using piggyBac-Based DNA Transposition. Front. Appl. Math. Stat. 3:786893. doi: 10.3389/fgeed.2021.786893

INTRODUCTION

Prime editing, based on the CRISPR-Cas9 system (Jinek et al., 2012; Jinek et al., 2013; Ran et al., 2013), enables precise editing of the genome by supporting targeted insertions, deletions, or any of the 12 possible single-base substitutions. Gene editing by prime editing involves neither donor templates nor double-stranded breaks (Anzalone et al., 2019). These unique properties of prime editing are based on the delivery of a prime editor (PE), consisting of a Cas9-reverse transcriptase fusion protein (hereafter referred to as PE2), along with the prime editing guide RNA (pegRNA) that specifies both the genomic target as well as the desired edit to be written directly into the genome. Prime editing has tremendous potential for treatment of disease-causing mutations, as well as generation of disease models, both *in vitro* and *in vivo* (Schene et al., 2020; Jang et al., 2021; Kim et al., 2021; Liu et al., 2021; Park et al., 2021; Petri et al., 2021; Qian et al., 2021). However, the use of prime editing is currently challenged by low efficacy, leading to time-consuming optimization and/or screening approaches in order to achieve satisfactory editing activities (Liu et al., 2020; Schene et al., 2020; Chemello et al., 2021; Kim et al., 2021; Petri et al., 2021).

Stable integration of gene cassettes encoding conventional CRISPR effectors, like Cas9 and single guide RNAs, into the genome of mammalian cells is widely used across life science research, including for generation of model cell lines and in CRISPR screens (Shalem et al., 2014; Holmgaard et al., 2017; Thomsen et al., 2020). For prime editing, effective integration of the PE2-expressing cassette into the genome of mammalian cells is challenged by the large size of the PE2 coding sequence (6351 bp). This makes the use of viral vectors difficult due to restricted packaging capabilities (Kumar et al., 2001), and so far, the PE2 system has only been integrated into the genome of mammalian cells by delivering intein-split PE2 cassettes using two separate lentiviral vectors (Anzalone et al., 2019). Here, we present piggyPrime, a non-viral, single-vector system for easy and efficient integration of all prime editing components in human cells, utilizing the large integration capacity of the piggyBac transposon system. Importantly, prolonged expression of PE2 and pegRNA facilitated by DNA transposition supports increased levels of prime editing, providing thus a novel approach for effective transgenesis.

MATERIALS AND METHODS

Plasmid Constructions

DNA amplification was performed using Phusion High-Fidelity PCR Master Mix (ThermoFisher Scientific) unless otherwise stated. All oligoes for pegRNAs and nicking sgRNAs were from Integrated DNA Technologies. Cloning oligoes for previously published pegRNAs were derived using pegIT (Anderson et al., 2021), which was also used to design the ngRNA used for the HBB(E7V) target. pCMV-PE2 (Addgene plasmid no. 132775) and pU6-pegRNA-HEK3-CTTins (Addgene plasmid no. 132778) were gifts from David Liu (Anzalone et al., 2019). pCMV-hyPBase is described elsewhere (Yusa et al., 2011). For generation of pPBT-PE2-PGK-Blast (Addgene plasmid no. 173219), CMV-PE2 was first amplified from pCMV-PE2 using Platinum SuperFi II High-Fidelity DNA Polymerase (ThermoFisher Scientific) and assembled into a HindII-/NdeI-digested pPBT-EFS-Cas9-P2A-mCherry (unpublished) using NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs) to generate pPBT-PE2. Fragments containing PGK-Blast and a bGH poly A signal were then amplified from pCW-Cas9-Blast (Addgene plasmid no. 83481) and pCMV-PE2, respectively, using PCR and assembled into XbaI-digested pPBT-PE2. For generation of pPBT-pegRNA-Puro, a modified pegRNA Golden Gate cloning cassette was amplified from pU6-pegRNA-GG-acceptor (Addgene plasmid no. 132777), with primers designed to convert the BsaI restriction sites to BsmBI sites, which allows pegRNAs to be cloned as described elsewhere (Anzalone et al., 2019), but with the use of BsmBI (New England Biolabs). The EF-1α promoter was then amplified from lentiGuide-Puro (Addgene plasmid no. 52963) and assembled with the pegRNA cloning cassette into a HindIII-/SmaI-digested pPBT-EFS-Cas9-P2A-mCherry backbone using NEBuilder. The resulting plasmid, pPBT-peRNA_GG-Puro (Addgene plasmid no. 173220), was then subjected to pegRNA Golden Gate cloning of the HEK3-CTTins pegRNA using the protocol described elsewhere (Anzalone et al., 2019), but with the use of BsmBI-v2 (New England Biolabs) (**Supplementary Note S2**). For generation of piggyPrime vectors, pPBT-PE2-PuroTK-pegRNA-GG (Addgene plasmid no. 173222) was first generated, into which pegRNAs can easily be cloned analogous to cloning of pegRNAs into pPBT-pegRNA_GG-Puro. For this, the M-MLV RT was amplified from pCMV-PE2, P2A-PuroTK-pA was amplified from pPBT-EFS-Cas9-P2A-PuroTK (unpublished) and the modified pegRNA Golden Gate cassette was amplified from pPBT-pegRNA_GG-Puro. pPBT-EFS-Cas9-P2A-PuroTK contains a mutation within the PuroTK gene that removes a BsmBI restriction site. The three fragments were then assembled into a BamHI-/SmaI-digested pPBT-PE2-PGK-Blast using NEBuilder. All piggyPrime vectors were subsequently generated by Golden Gate assembly of pegRNAs using BsmBI-v2 (**Supplementary Note S2**). For generation of multiplexed piggyPrime vectors, the HBB(E7V)-piggyPrime vector was linearized using XbaI (ThermoFisher Scientific). The HEK3-CTTins pegRNA expression cassette was then amplified from the HEK3-CTTins piggyPrime vector and inserted into the HBB(E7V)-piggyPrime using NEBuilder. Same procedure was used to insert the HBB(E7V) nicking sgRNA instead, which was amplified from a pU6-HBB(E7V)-ngRNA plasmid. The primers used for cloning of all plasmids are listed in **Supplementary Table S3**.

Cell Culture Conditions and Transfection

HEK293T and HeLa cells (ATCC) were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5% Fetal Calf Serum (FCS) and 1% penicillin/streptomycin (P/S). K562 cells (ATCC) were maintained in RPMI-1640 medium (Sigma-Aldrich) supplemented with 10% FCS and 1% P/S. All cells were incubated, maintained, and cultured at 37°C with 5% CO_2. For transfection of HEK293T and HeLa, 5×10^4 cells were seeded in 24-well plates 18–24 h prior to transfection. Transfection was performed using 1,000 ng of plasmid DNA and 2.5 µL TurboFect™ Transfection Reagent (ThermoFisher Scientific) following the manufacturer's instructions. For transfection of K562 cells, 1×10^5 cells were seeded in 24-well plates and transfected with 1,000 ng plasmid DNA using 3 µL Lipofectamine™ 2000 (ThermoFisher Scientific) according to manufacturer's instructions. For co-transfections using the PE3 system (**Figure 1C**), a plasmid ratio of 3:1 (pegRNA:ngRNA) was used. For co-transfections of piggyBac vectors and pCMV-hyPBase, a plasmid ratio of 9:1 was used (vector:hyPBase). In all transfection experiments, medium was changed 16 h after transfection, and unless otherwise specified, cells were harvested 72 h after transfection. Selection medium (5 µg/ml blasticidin and/or 1 µg/ml puromycin (ThermoFisher Scientific)) was applied to indicated experiments at day 3 after transfection and maintained for the full duration of all experiments. Cells were passaged as required.

Genomic DNA Extraction and Analysis of Prime Editing Events

Genomic DNA was extracted by addition of 100–300 µL lysis buffer (10 mM Tris-HCl, pH 7.5; 0.05% SDS; 25 µg/ml proteinase K (ThermoFisher Scientific)) depending on confluency of cells at time of harvest. Lysis mixture was incubated 2 h at 37°C followed by enzyme inactivation at 80°C for 30 min. PCR was then performed on 150 ng of extracted genomic DNA using Phusion High-Fidelity PCR Master Mix (ThermoFisher Scientific) with primers listed in **Supplementary Table S2**. PCR amplicons were purified by 1% agarose gel electrophoresis using E.Z.N.A® Gel Extraction Kits (Omega Bio-Tek). Sanger sequencing of PCR amplicons was then performed by Eurofins Genomics, and prime editing events were analyzed using DECODR (Bloh et al., 2021).

Copy-Number Determination Using ddPCR

For copy-number (CN) determination of integrated piggyPrime cassettes, genomic DNA was harvested by ethanol precipitation

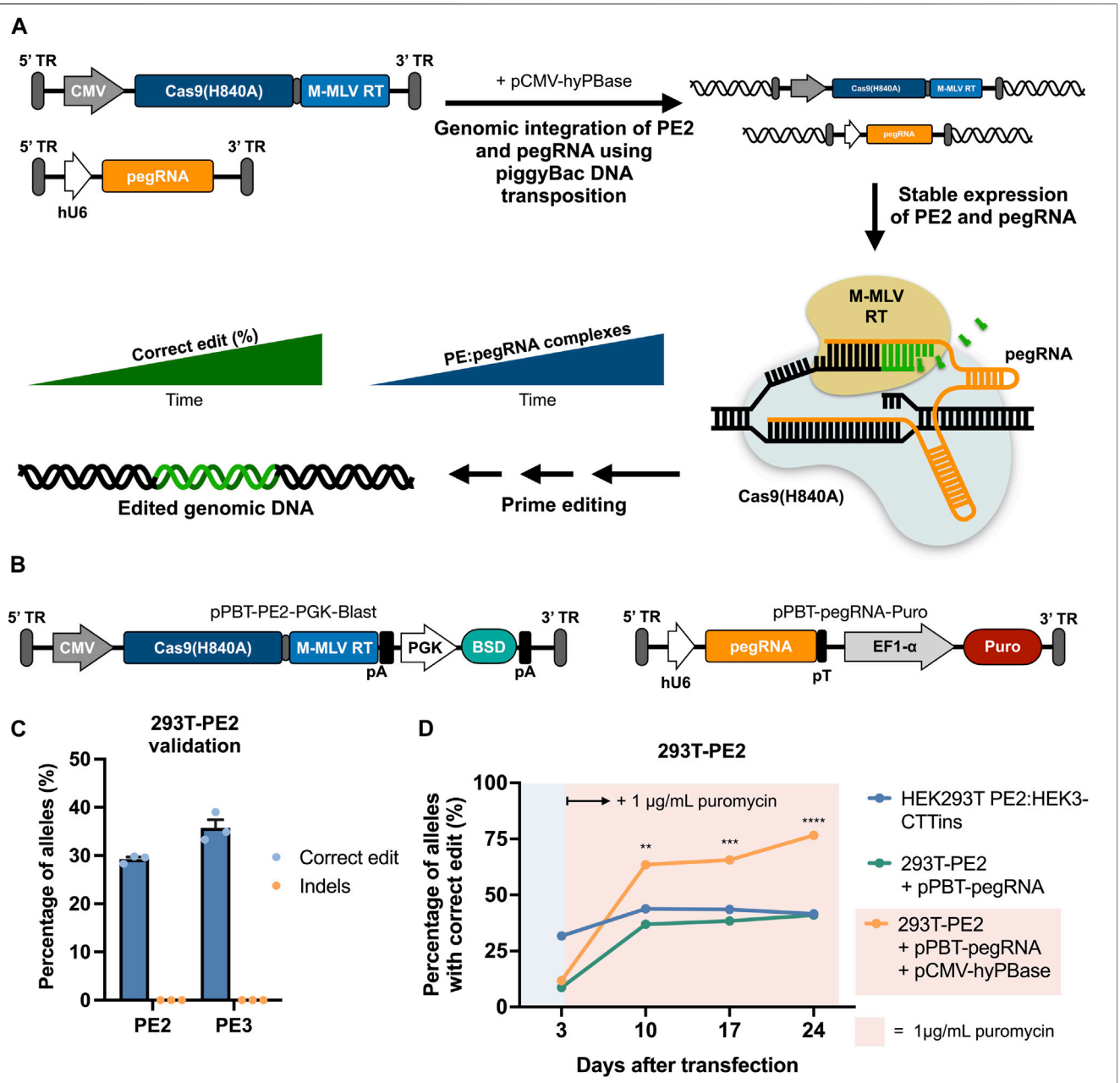


FIGURE 1 | | Prime editing by piggyBac-mediated transposition of PE2 and pegRNA cassettes using separate piggyBac vectors. **(A)** Graphical overview of prime editing based on piggyBac DNA transposition, demonstrating the concept of potent editing as a result of stable expression of both PE2 and pegRNA. **(B)** Schematic representation of the piggyBac vectors carrying either PE2 (left) or a pegRNA (right). **(C)** Integration of the Cas9-RT expression cassette alone into HEK293T cells facilitates correct editing at the HEK3 target when transfected with a HEK3-CTTins pegRNA (PE2) as well as when transfected with an additional nicking sgRNA (PE3). Editing rates were determined 3 days after transfection. **(D)** Integration of the HEK3-CTTins pegRNA expression cassette into 293T-PE2 cells (293T-PE2 + pPBT-pegRNA + pCMV-hyPBase) results in increased prime editing over time compared to cells with transiently expressed PE2 and pegRNA (HEK293T PE2:HEK3-CTTins) or transiently expressed pPBT-pegRNA-Puro (293T-PE2 + pPBT-pegRNA). Puromycin was applied at day 3 after transfection only to cells with stably integrated pegRNA cassette. Data and error bars show mean (n = 3) ± sd. Statistical significance was calculated using multiple unpaired t-tests with correction for multiple testing (**p < 0.002, ***p < 0.0002, ****p < 0.0001).

followed by HindIII (ThermoFisher Scientific) digestion for 1 h at 37°C. Quantitative Droplet Digital PCR (ddPCR) was then performed on a QX200TM Droplet DigitalTM PCR System with ddPCR Supermix for Probes (No dUTP) (BioRad) using 25 ng of digested genomic DNA as input. Primers and probes used are listed in **Supplementary Table S4** and targets the Puromycin and albumin (ALB) genes. Data were analyzed using QuantaSoft™ Analysis Pro.

pegRNA Expression-Levels ddPCR

Total RNA was extracted from cells using Roche High Puro miRNA Isolation Kit (Roche Applied Science) and subjected to DNase I treatment (ThermoFisher Scientific). 250 ng RNA was then used for cDNA synthesis using Maxima H Minus cDNA Synthesis Master Mix (ThermoFisher Scientific). Quantitative Droplet Digital PCR (ddPCR) was then performed on a QX200TM Droplet DigitalTM PCR System with ddPCR Supermix for Probes (No dUTP) (BioRad) with 1/8 diluted DNase I treated RNA as input. Primers and probes used are listed in **Supplementary Table S4**.

Statistical Analysis

For all graphs, mean (n = 3) and standard deviation (sd) were calculated and plotted using GraphPad Prism 9. Statistical analysis was performed using multiple unpaired t-tests with correction for multiple testing.

RESULTS

Stable Genomic Integration of Prime Editing Effectors Using piggyBac DNA Transposition Enables Increased Editing in HEK293T Cells

We first sought to establish evidence of functional integration of both PE2- and pegRNA-expressing cassettes with the use of piggyBac DNA transposition and validate that expression of the prime editing components from integrated cassettes was sufficient to confer targeted prime editing (**Figure 1A**). To do this, we constructed a piggyBac vector (pPBT-PE2-PGK-Blast) carrying a PE2 cassette driven by a CMV promoter as well as a blasticidin resistance gene driven by a PGK promoter (**Figure 1B**). We then transfected HEK293T cells with this vector along with plasmid DNA encoding a hyperactive piggyBac transposase (Yusa et al., 2011) (hyPBase), selected for blasticidin-resistant cells, and validated functional prime editing in these cells (referred to as 293T-PE2) after subsequent transfection with a plasmid encoding a pegRNA designed to incorporate a three-nucleotide CTT insertion at the *HEK3* genomic site (Anzalone et al., 2019) (HEK3-CTTins pegRNA) (**Figure 1C**). This resulted in editing rates that were comparable with previous reported editing efficacies using the same pegRNA (Anzalone et al., 2019). We then constructed a piggyBac vector containing the HEK3-CTTins pegRNA expression cassette (pPBT-pegRNA-Puro) (**Figure 1B**), integrated this into 293T-PE2 cells using hyPBase and measured the edit rates at fixed timepoints following transfection and in the presence of puromycin (**Figure 1D**). At day 10 after transfection, the cells with both PE2 and pegRNA cassettes integrated into the genome showed a markedly higher editing rate compared to wild-type HEK29T cells co-transfected with pCMV-PE2 and pU6-pegRNA (63.5 ± 1.4% vs 43.8 ± 2.3%) (**Figure 1D**). From day 10 until day 24 after transfection, the cells with integrated PE2 and pegRNA cassettes showed an increase in correct editing, whereas transiently transfected cells did not show an increase in editing, indicating that long-term expression allowed for targeted edits to accumulate. Furthermore, we did not detect any indel formation, even 24 days after transfection (**Supplementary Figure S1**).

Integration of all Prime Editing Components Using All-In-One Single Vectors Allows for Potent Editing at Target Sites

Next, we sought to develop a single-vector system that could deliver all the components of the prime editing system to cells in an easy and adaptable way. For this, we constructed a piggyBac vector containing PE2, a puromycin resistance gene, and a pegRNA Golden Gate cloning cassette for easy pegRNA cloning (**Figure 2A**; **Supplementary Note S2**). We then constructed piggyBac-PE2-pegRNA vectors (hereafter referred to as piggyPrime vectors) using five different pegRNAs (**Supplementary Note S1**), integrated them into the genome of both HEK293T and HeLa cells using hyPBase and established puromycin-resistant cell lines. At day 3 after transfection, all cell lines showed minimal prime editing activity with no detectable editing in the majority of piggyPrime-transfected cells (**Figure 2B**). However, at day 10 after transfection, editing could be observed in all cell lines, with editing rates ranging from 46 ± 1.5% to 98 ± 4% in HEK293T cells (**Figure 1B**) and from 32 ± 4% to 69 ± 0.6% in HeLa cells (**Figure 2C**). In accordance with our initial findings, editing rates kept increasing at all targets at day 17 and 24 after transfection, with some targets reaching 100% editing (**Figure 2B**). At day 24, the average editing rates across all 5 pegRNAs were 84.7 ± 15.5% and 63.6 ± 17.1% in HEK293T and HeLa cells, respectively. Furthermore, we successfully integrated the HEK3-CTTins and HBB(E7V) piggyPrime vectors into K562 cells by transfection and observed up to 46 ± 0.7% correct editing at day 24, despite the fact that editing could not be detected at day 3 after transfection (**Figure 2D**). For all targets across all cell lines, we did not detect any indel formation at day 24 after transfection (**Supplementary Figure S2**).

Additionally, we measured the level of pegRNA present in selected cell lines to confirm that the PE2-pegRNA cassette was stably integrated and expressed. We transfected K562 cells with piggyPrime vectors either with or without co-transfection of hyPBase-encoding plasmid DNA and found that only cells co-transfected with hyPBase had pegRNA levels that were detectable by ddPCR at day 14 (**Figure 2E**; **Supplementary Figure S3A**). Furthermore, only the cells that were co-transfected with hyPBase showed an increase in prime editing activity from day 3 to 14 (**Figure 2F**; **Supplementary Figures S3B,C**). We also determined the copy-number of the integrated HEK3-CTTins piggyPrime vectors and found that, on average, 6 and 5.4 copies were present in HEK293T and HeLa cells, respectively (**Figure 2G**). Hence, this confirmed that PE2 and pegRNAs were indeed stably expressed from integrated transposons, and that prolonged expression of PE2 and pegRNA was necessary to achieve increased levels of prime editing.

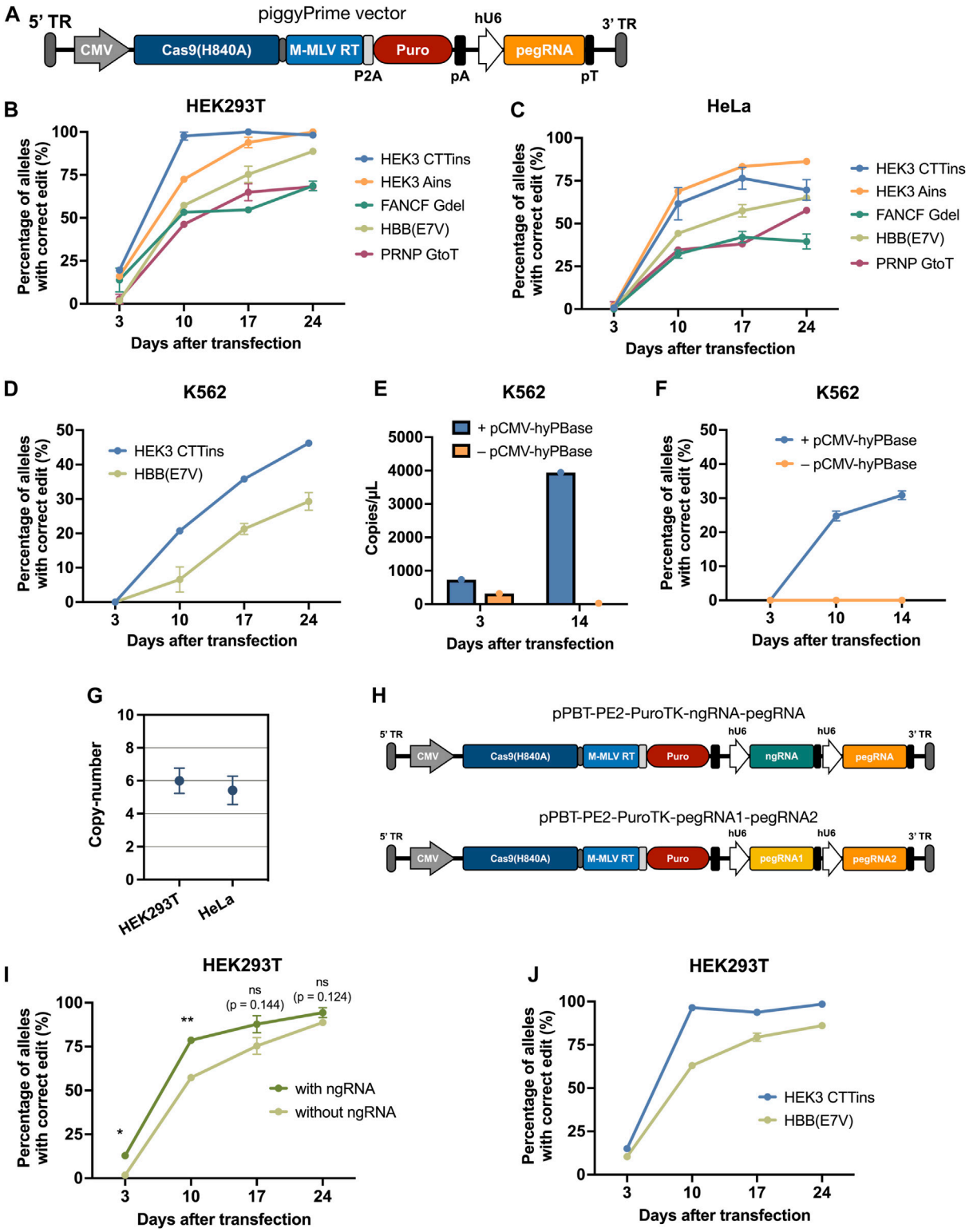


FIGURE 2 | | Effective prime editing by piggyBac-mediated integration of all prime editing components using a single-vector system. **(A)** Schematic overview of the piggyPrime vector, consisting of (from the right) a 5' terminal repeat (TR), CMV promotor, Cas9(H840A)-linker-M-MLV-RT (PE2), P2A, PuroR, bGH pA, hU6 promotor, pegRNA, and 3' TR. **(B,C)** Integration of piggyPrime vectors containing the HEK3-CTTins, HEK3-Ains, HBB(E7V), FANCF-6Gdel, and PRNP-GtoT pegRNAs into HEK293T **(B)** and HeLa **(C)** cells resulting in increasing correct editing at target sites over time. Puromycin was applied at day 3 after transfection. **(D)** piggyPrime vectors can also successfully be integrated into the genome of K562 cells by transfection, resulting in increased correct editing over time. **(E)** pegRNA levels were

(*Continued*)

FIGURE 2 | determined at day 3 and 14 in K562 cells transfected with the HEK3-CTTins piggyPrime vector either with or without hyPBase. Only cells co-transfected with the piggyPrime vector and hyPBase-encoding plasmid DNA showed detectable pegRNA levels at day 14. **(F)** Editing rates were determined at multiple time points in K562 cells transfected with the HEK3-CTTins piggyPrime vector either with or without hyPBase. Only cells co-transfected with the piggyPrime vector and hyPBase-encoding plasmid DNA showed detectable correct editing. **(G)** The average copy-number of HEK3-CTTins piggyPrime vectors was determined in HEK293T and HeLa cells using ddPCR. **(H)** Schematic overview of piggyPrime vectors carrying both a nicking sgRNA (ngRNA) and a pegRNA (top) or dual pegRNAs (bottom). **(I)** Integration of HBB(E7V) piggyPrime vector carrying both a ngRNA and a pegRNA cassette resulted in markedly increased editing compared to HBB(E7V) piggyPrime vectors without a ngRNA. **(J)** Integration of piggyPrime vectors carrying both the HEK3-CTTins and HBB(E7V) pegRNA resulted in correct editing at both target sites, without compromising editing efficacy. Data and error bars show mean (n = 3) ± sd. Statistical significance was calculated using multiple unpaired t-tests with correction for multiple testing (*p < 0.03, **p < 0.002).

An increase in prime editing efficacy can be achieved by delivering a conventional sgRNA (termed a "nicking sgRNA" or "ngRNA") in addition to the pegRNA, a method termed PE3 or PE3b (Anzalone et al., 2019). Therefore, we additionally constructed an HBB(E7V) piggyPrime vector capable of integrating a ngRNA cassette in addition to the PE2 and the HBB(E7V) pegRNA (**Figure 2H**). Notably, including a ngRNA within the HBB(E7V) piggyPrime vector led to markedly higher editing rates at the HBB target in HEK293T cells at day 3 and 10 after transfection (**Figure 2I**). We also constructed a piggyPrime vector encoding both HBB(E7V) and HEK3-CTTins pegRNAs (**Figure 2H**), which led to successful multiplexed editing of both the HBB and HEK3 target, without compromising editing efficacy at any of the targets (**Figure 2J**).

DISCUSSION

This study demonstrates, to the best of our knowledge, the first single-vector system capable of genomic integration of all the components of both the PE2 and PE3/PE3b system resulting in potent editing of up to 100% of targeted alleles. Editing efficacies were generally found to be lower in HeLa and K562 cells compared to HEK293T cells, which may partially reflect differences in plasmid transfection rates. Such differences between cell lines are in accordance with previous studies also reporting lower efficacy of prime editing in both HeLa and K562 cells (Anzalone et al., 2019; Nelson et al., 2021). Some concern could be raised in regard to generating cell lines with constitutively expressed genome editing effectors; however, the prime editing system has been shown to confer far less off-target editing than conventional CRISPR-Cas9, thus minimizing concerns of such undesired editing (Anzalone et al., 2019). Alternatively, potential unwanted off-target editing or by-product formation could be addressed by using controllable expression systems in the piggyPrime vector context. Additionally, in all experiments we were unable to detect any indel formation at day 24 after transfection by Sanger sequencing, even when including a ngRNA, which has been shown to increase indel formation of the PE system (Anzalone et al., 2019).

PiggyPrime represents a groundbreaking new approach for effective generation of transgenic model cell lines harboring disease-causing genetic variants or genes with desired genetic alterations. piggyPrime does not depend on homologous recombination and only requires a single pegRNA to be designed and cloned into a piggyPrime vector in a one-step Golden Gate cloning, thus aiding universal use. If desired, the piggyPrime system can furthermore accommodate a nicking sgRNA in addition to the pegRNA, which is expected to increase efficacy at targets of interest. Alternatively, the system can be multiplexed allowing two and potentially more pegRNAs to be integrated to generate complex disease models harboring multiple mutations. Our findings demonstrate potent prime editing leading to effective transgenesis in cells with prolonged expression of key prime editing components, suggesting that a longer time frame for prime editing is crucial for improved efficacy and common use. piggyPrime is easy adaptable to most proliferating cell types and is likely to become the standard technology for generation of cells with tailored genetic edits throughout the scientific community.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JW and JM conceived the study. JW designed, performed, and analyzed the experiments. ET assisted in construction and design of plasmids. SA assisted in ddPCR experiments. JH performed ddPCR. JW and JM wrote the manuscript. All authors have read and approved the manuscript.

FUNDING

This work was made possible through support from The Lundbeck Foundation (Grant R324-2019-1832) and the Innovation Fund Denmark (PASCAL-MID, Grant 8056-00010A).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2021.786893/full#supplementary-material

REFERENCES

Anderson, M. V., Haldrup, J., Thomsen, E. A., Wolff, J. H., and Mikkelsen, J. G. (2021). pegIT - a Web-Based Design Tool for Prime Editing. *Nucleic Acids Res.* 49, W505–W509. doi:10.1093/nar/gkab427

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-Replace Genome Editing without Double-Strand Breaks or Donor DNA. *Nature* 576 (7785), 149–157. doi:10.1038/s41586-019-1711-4

Bloh, K., Kanchana, R., Bialk, P., Banas, K., Zhang, Z., Yoo, B.-C., et al. (2021). Deconvolution of Complex DNA Repair (DECODR): Establishing a Novel Deconvolution Algorithm for Comprehensive Analysis of CRISPR-Edited Sanger Sequencing Data. *CRISPR J.* 4 (1), 120–131. doi:10.1089/crispr.2020.0022

Chemello, F., Chai, A., Li, H., Rodriguez-Caycedo, C., Sanchez-Ortiz, E., Atmanli, A., et al. (2021). Precise Correction of Duchenne Muscular Dystrophy Exon Deletion Mutations by Base and Prime Editing. *Sci. Adv.* 7 (18), eabg4910. doi:10.1126/sciadv.abg4910

Holmgaard, A., Askou, A. L., Benckendorff, J. N. E., Thomsen, E. A., Cai, Y., Bek, T., et al. (2017). *In Vivo* knockout of the Vegfa Gene by Lentiviral Delivery of CRISPR/Cas9 in Mouse Retinal Pigment Epithelium Cells. *Mol. Therapy-Nucleic Acids* 9, 89–99. doi:10.1016/j.omtn.2017.08.016

Jang, H., Jo, D. H., Cho, C. S., Shin, J. H., Seo, J. H., Yu, G., et al. (2021). Application of Prime Editing to the Correction of Mutations and Phenotypes in Adult Mice with Liver and Eye Diseases. *Nat. Biomed. Eng.* doi:10.1038/s41551-021-00788-9

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337 (6096), 816. doi:10.1126/science.1225829

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-Programmed Genome Editing in Human Cells. *eLife* 2, e00471. doi:10.7554/eLife.00471

Kim, Y., Hong, S.-A., Yu, J., Eom, J., Jang, K., Yoon, S., et al. (2021). Adenine Base Editing and Prime Editing of Chemically Derived Hepatic Progenitors Rescue Genetic Liver Disease. *Cell Stem Cell* 28, 1614–1624. doi:10.1016/j.stem.2021.04.010

Kumar, M., Keller, B., Makalou, N., and Sutton, R. E. (2001). Systematic Determination of the Packaging Limit of Lentiviral Vectors. *Hum. Gene Ther.* 12 (15), 1893–1905. doi:10.1089/104303401753153947

Liu, P., Liang, S.-Q., Zheng, C., Mintzer, E., Zhao, Y. G., Ponnienselvan, K., et al. (2021). Improved Prime Editors Enable Pathogenic Allele Correction and Cancer Modelling in Adult Mice. *Nat. Commun.* 12 (1), 2121. doi:10.1038/s41467-021-22295-w

Liu, Y., Li, X., He, S., Huang, S., Li, C., Chen, Y., et al. (2020). Efficient Generation of Mouse Models with the Prime Editing System. *Cel Discov.* 6 (1), 27. doi:10.1038/s41421-020-0165-z

Nelson, J. W., Randolph, P. B., Shen, S. P., Everette, K. A., Chen, P. J., Anzalone, A. V., et al. (2021). Engineered pegRNAs Improve Prime Editing Efficiency. *Nat. Biotechnol.* 184 (22), 5635.5652.e29. doi:10.1038/s41587-021-01039-7

Park, S.-J., Jeong, T. Y., Shin, S. K., Yoon, D. E., Lim, S.-Y., Kim, S. P., et al. (2021). Targeted Mutagenesis in Mouse Cells and Embryos Using an Enhanced Prime Editor. *Genome Biol.* 22 (1), 170. doi:10.1186/s13059-021-02389-w

Petri, K., Zhang, W., Ma, J., Schmidts, A., Lee, H., Horng, J. E., et al. (2021). CRISPR Prime Editing with Ribonucleoprotein Complexes in Zebrafish and Primary Human Cells. *Nat. Biotechnol.*, 1–5. doi:10.1038/s41587-021-00901-y

Qian, Y., Zhao, D., Sui, T., Chen, M., Liu, Z., Liu, H., et al. (2021). Efficient and Precise Generation of Tay–Sachs Disease Model in Rabbit by Prime Editing System. *Cel Discov.* 7 (1), 50. doi:10.1038/s41421-021-00276-z

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome Engineering Using the CRISPR-Cas9 System. *Nat. Protoc.* 8 (11), 2281–2308. doi:10.1038/nprot.2013.143

Schene, I. F., Joore, I. P., Oka, R., Mokry, M., van Vugt, A. H., van Boxtel, R., et al. (2020). Prime Editing for Functional Repair in Patient-Derived Disease Models. *Nat. Commun.* 11 (1), 1–8. doi:10.1038/s41467-020-19136-7

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., et al. (2014). Genome-scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* 343 (6166), 84–87. doi:10.1126/science.1247005

Thomsen, E. A., Rovsing, A. B., Anderson, M. V., Due, H., Huang, J., Luo, Y., et al. (2020). Identification of BLNK and BTK as Mediators of Rituximab-Induced Programmed Cell Death by CRISPR Screens in GCB-Subtype Diffuse Large B-Cell Lymphoma. *Mol. Oncol.* 14 (9), 1978–1997. doi:10.1002/1878-0261.12753

Yusa, K., Zhou, L., Li, M. A., Bradley, A., and Craig, N. L. (2011). A Hyperactive piggyBac Transposase for Mammalian Applications. *Proc. Natl. Acad. Sci.* 108 (4), 1531–1536. doi:10.1073/pnas.1008322108

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

frontiers in Genome Editing

HYPOTHESIS AND THEORY
published: 17 March 2022
doi: 10.3389/fgeed.2022.793010

CRISPR/Cas-Based Gene Editing Strategies for DOCK8 Immunodeficiency Syndrome

*Sujan Ravendran, Sabina Sánchez Hernández, Saskia König and Rasmus O. Bak**

Department of Biomedicine, Aarhus University, Aarhus, Denmark

Defects in the DOCK8 gene causes combined immunodeficiency termed DOCK8 immunodeficiency syndrome (DIDS). DIDS previously belonged to the disease category of autosomal recessive hyper IgE syndrome (AR-HIES) but is now classified as a combined immunodeficiency (CID). This genetic disorder induces early onset of susceptibility to severe recurrent viral and bacterial infections, atopic diseases and malignancy resulting in high morbidity and mortality. This pathological state arises from impairment of actin polymerization and cytoskeletal rearrangement, which induces improper immune cell migration-, survival-, and effector functions. Owing to the severity of the disease, early allogenic hematopoietic stem cell transplantation is recommended even though it is associated with risk of unintended adverse effects, the need for compatible donors, and high expenses. So far, no alternative therapies have been developed, but the monogenic recessive nature of the disease suggests that gene therapy may be applied. The advent of the CRISPR/Cas gene editing system heralds a new era of possibilities in precision gene therapy, and positive results from clinical trials have already suggested that the tool may provide definitive cures for several genetic disorders. Here, we discuss the potential application of different CRISPR/Cas-mediated genetic therapies to correct the DOCK8 gene. Our findings encourage the pursuit of CRISPR/Cas-based gene editing approaches, which may constitute more precise, affordable, and low-risk definitive treatment options for DOCK8 deficiency.

Keywords: gene editing (CRISPR-Cas9), CRISPR/Cas 9, hematopoietic stem cell, DOCK8 immunodeficiency syndrome, DOCK8 deficiency, DOCK8, gene therapy, Primary immunodeficiency

OPEN ACCESS

Edited by:
David Hagin, Tel Aviv Sourasky Medical Center, Israel

Reviewed by:
Alexandra Freeman, National Institutes of Health (NIH), United States
Caroline Y. Kuo, David Geffen School of Medicine at UCLA, University of California, Los Angeles, United States

****Correspondence:***
Rasmus O. Bak
bak@biomed.au.dk

Specialty section:
This article was submitted to Genome Editing in Blood Disorders, a section of the journal Frontiers in Genome Editing

Received: *11 October 2021*
Accepted: *14 February 2022*
Published: *17 March 2022*

Citation:
Ravendran S, Hernández SS, König S and Bak RO (2022) CRISPR/Cas-Based Gene Editing Strategies for DOCK8 Immunodeficiency Syndrome. Front. Genome Ed. 4:793010. doi: 10.3389/fgeed.2022.793010

INTRODUCTION

Primary immunodeficiencies (PIDs) include more than 400 rare congenital monogenic disorders characterized by impairment of immunity, susceptibility to infectious diseases, autoimmunity, autoinflammatory diseases, allergy and/or malignancy (Tangye et al., 2020). In recent years, there has been an increase in the recognition and diagnosis of previously undefined genetically caused abnormalities in the immune system (Tangye et al., 2020). This has been made possible through the completion of the Human Genome project in the early 2000s, improved definition of clinical phenotypes, and advancement of cost-effective and time-efficient sequencing through implementation of next generation DNA sequencing technologies (Meyts et al., 2016; Bousfiha et al., 2020; Tangye et al., 2020; Gates et al., 2021).

Among these disorders is DOCK8 immunodeficiency syndrome (DIDS) also known as DOCK8 deficiency. Until recently DOCK8 deficiency was termed DOCK8-related Hyper Immunoglobulin E

(IgE) Syndrome (HIES), as it is characterized by elevated IgE levels, eosinophilia, and recurrent infections. The majority of HIES is caused by either autosomal dominant inheritance (AD-HIES) of mutations in the signal transducer and activator of transcriptase 3 (STAT3) gene (Grimbacher et al., 1999), or autosomal recessive inheritance (AR-HIES) primarily of mutations in the guanine-nucleotide exchange factor dedicator of cytokinesis 8 (DOCK8) gene (Su et al., 2011). However, an increased insight into the functionality of DOCK8 illuminates its impact on both the T- and B-cell compartment of the immune system, which has promoted the reclassification as a DOCK8-related combined immunodeficiency (CID). DOCK8 deficiency is a severe disorder with early onset of morbidity and high mortality rates exceeding those associated with STAT3 HIES (Aydin et al., 2015; Tsilifis et al., 2021). Since the majority of the clinical manifestations of DOCK8 deficiency pertain to the immune system, hematopoietic stem cell transplantation (HSCT) in early childhood is encouraged. However, this is challenged by the need for a HLA-matched donor and associated with adverse events such as immune rejection and graft-versus-host disease (Copelan, 2006; Aydin et al., 2019).

The ideal treatment of DOCK8 deficiency would be correcting the disease-causing mutation in the patient's own cells, thereby restoring the DOCK8 functionality. This would circumvent the obstacle of identifying HLA compatible donors for allogeneic transplantation and eliminate the associated risks. During the past two decades, genetic therapies have shown promising results for an expanding numbers of genetic disorders (Booth et al., 2019; Porteus, 2019). Meanwhile, precise genome editing tools were developed and applied in a range of pre-clinical and even a few clinical gene therapy studies. In particular, the discovery of the CRISPR/Cas system as a highly versatile genome editing platform accelerated the development of genome editing methods (Bak et al., 2018a). This ultimately led to the 2020 Nobel Prize in Chemistry awarded to Emmanuelle Charpentier and Jennifer Doudna for 'the development of a method for genome editing'. CRISPR/Cas offers unprecedented simplicity in facilitating genome editing, and has proven highly precise and efficient (Jensen et al., 2019; Porteus, 2019). Since the first injections of CRISPR gene edited cells into patients in 2016 (Cyranoski, 2016), there have been published reports on only a few clinical trials. These trials have marked important milestones by providing evidence on safety, whereas recent clinical trials on sickle cell disease, β-thalassemia, and transthyretin amyloidosis have been the first to demonstrate therapeutic and potentially curative potential (Frangoul et al., 2021; Gillmore et al., 2021).

In preclinical studies, CRISPR/Cas gene editing has shown tremendous potential in a wide range of diseases, but has so far not been applied to DOCK8. Here, we will elaborate on why gene editing is within the realm of possibility for treating DOCK8 deficiency. First, we briefly present our current understanding of the genetic, molecular, and cellular mechanisms involved in DOCK8 deficiency. Second, we portray the common disease manifestations and discuss current diagnostic and treatment approaches. Third, after describing recent advancements in the field of genome editing and discussing advantages and disadvantages of the different precise gene editing platforms, we define suitable CRISPR/Cas strategies for treating, which may constitute a definitive cure for DOCK8 deficiency. Finally, we give a concise summary of hurdles and challenges for using gene editing in the clinical setting.

THE GENETICS OF DOCK8 DEFICIENCY

The large DOCK8 gene is located on the short arm of chromosome 9, includes 48 exons, spans over 250 kilobases, and encodes a protein of approximately 190 kDa. Bi-allelic loss-of-function mutations in the DOCK8 gene is associated with DOCK8 deficiency (Zhang et al., 2010; Database resources of the, 2018). DOCK8 deficiency is estimated to affect less than one person per million, but the exact prevalence is unknown (Biggs et al., 2017). The disease was not recognized until 2009, and only about 200 cases have been described world-wide so far, which have been identified predominantly in populations with consanguineous marriage (Zhang et al., 2009; Biggs et al., 2017).

To get a collected overview on the different patient mutations, we performed a comprehensive data collection of **60** disease-causing DOCK8 variants described in the literature and registered in the ClinVar database. These variants are represented in **Figure 1** and listed in **Supplementary Table S1**. Even though no specific mutation hotspot regions were identified, the majority of disease-causing mutations in DOCK8 were deletions which cover 61.5% of the variants and range from a few base pairs to deletions spanning several hundred base pairs. The high propensity for deletions has been hypothesized to be partly caused by the occurrence of repetitive genomic sequences leading to abnormal recombinations in this region (Engelhardt et al., 2009). The pathogenic variants in DOCK8 are predominantly loss of function, thus abolishing the expression of DOCK8, but occurrences of DOCK8 duplication has been shown to associate with neurodevelopmental conditions (Jing et al., 2014).

As with a few other primary immunodeficiencies, there have been reported cases of somatic reversion leading to partial re-expression of DOCK8 protein in some cell linages. These occurrences of "natural gene-therapy" may reflect the location of DOCK8 within a recombination hotspot, promoting either a somatic repair of a point-mutation, recombination-mediated gene conversion, or recombination-mediated intragenic single crossover (Pillay et al., 2021). Jing et al. observed some clinical improvement in seventeen patients with somatic reversion, with significant improvement in overall survival and age-stratified morbidity. However, these improvements were insufficient for disease elimination presumed to be due to inadequate DOCK8 re-establishment, particularly within the T cells (Pillay et al., 2021). In contrast, Pillay et al. identified three patients with biallelic compound heterozygous DOCK8 germline variants, who displayed significant DOCK8 expression in their lymphocyte

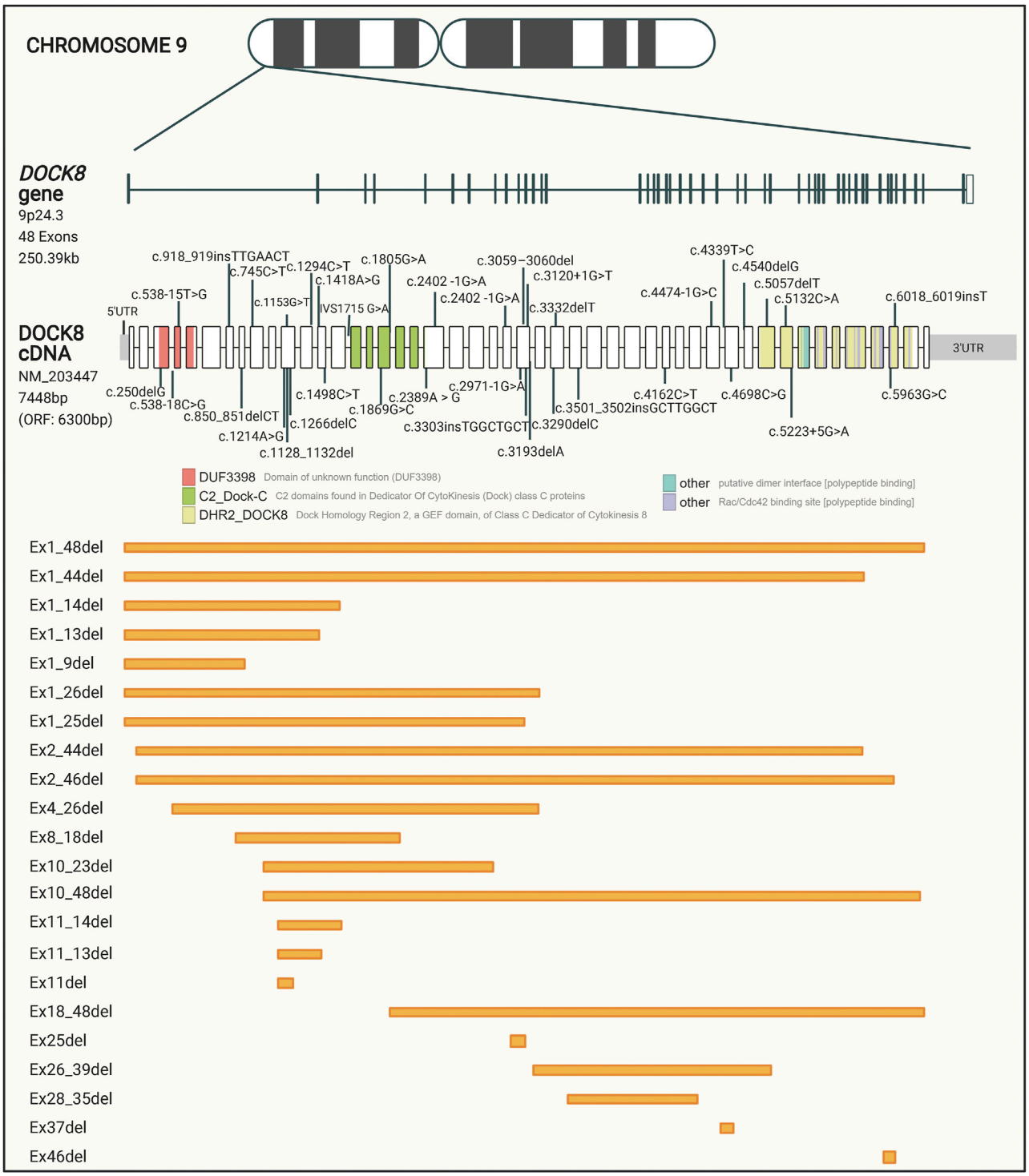


FIGURE 1 | Schematic representation of reported patient mutations in the DOCK8 gene. The DOCK8 gene, composed of 48 exons, is located on the short arm of chromosome nine and spans over 250 kb. The distribution of mutations associated with DOCK8 deficiency collected from the ClinVar database is represented along the DOCK8 cDNA. Boxes represent the 48 exons of the gene and different colors indicate major domain-encoding regions. Out of a total of 1,139 DOCK8 variants reported to date, 60 have been found in patients where DOCK8 deficiency has been diagnosed. Orange boxes represent deletions of one or more exons.

subsets ranging from 10% of all B cells to 75% of all $CD8^+$ T cells, while myeloid cells did not express DOCK8. DNA sequencing analyses revealed that one pathogenic allele had been genetically repaired, which was hypothesized to have occurred in either a single common lymphoid progenitor cell or a single hematopoietic stem cell. In all three patients, the somatic reversion improved

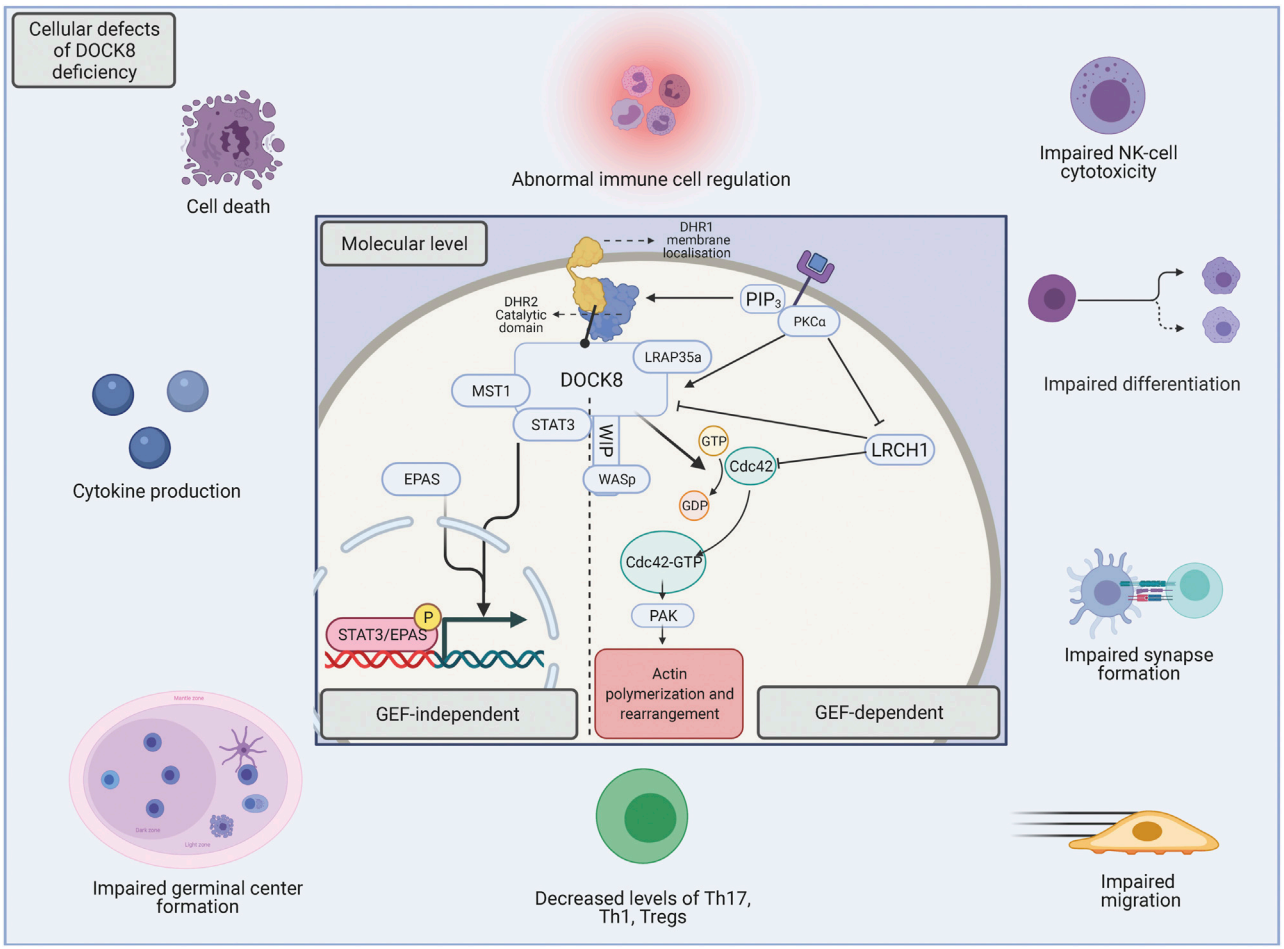


FIGURE 2 | The underlying molecular foundation for DOCK8 deficiency. The perturbation of DOCK8 expression disturbs a broad spectrum of immune cell functions such as differentiation, survival, migration, activation, immunotolerance and -function (McGhee and Chatila, 2010). The basis of the various functions of DOCK8 can be divided into either GEF-dependent actin regulation or functions within GEF-independent pathways. When chemokines bind to extracellular receptors, phosphoinositide 3 kinase (PI3K) is activated and initiates the production of phosphatidylinositol (3,4,5)-triphosphate (PIP3), which recruits DOCK8 via the DHR-1 subunit, which consequently leads to membrane-adjacent GEF activity (Xu et al., 2017; Sakurai et al., 2021). In addition, when chemokines bind extracellular receptors, PKCα is activated, which phosphorylate DOCK8 for dissociation from Leuchine Rich Repeats And Calponin Homology Domain Containing (LRCH1) (2), thus diminishing its inhibitory impact (Xu et al., 2017). The catalytic DHR2 domain of DOCK8 interacts with the nucleotide-free form of the Rho GTPase Cdc42, and mediates activation through GDP-GTP exchange (Harada et al., 2012) and leads to down-stream regulation of several biological activities such as cell morphology, -survival, -signaling and -cytoskeletal dynamics, all mediated through p21-activated kinases (PAK) (Bokoch, 2003). In addition to the aforementioned functions, DOCK8 loss in the GEF-independent pathways leads to nuclear translocation of EPAS1 promoting IL-31 production (Yamamura et al., 2017). Furthermore, DOCK8 associates with the transcription factor STAT3 and facilitates activation-induced STAT3 translocation to the nucleus. Here, the guanine nucleotide exchange function of DOCK8 is also necessary for optimal STAT3 phosphorylation and Th17 differentiation (Su et al., 2019).

survival, differentiation, and function of lymphocytes and provided great clinical improvement to the patients (Namekata et al., 2014). Such single progenitor/stem cell reversions signify that only modest DOCK8 correction frequencies by gene therapy in autologous hematopoietic stem cells or lymphoid progenitor cells could provide significant clinical benefit to the patients.

THE ROLE OF DOCK8

Until recently, the molecular mechanism of DOCK8 and its influence in cell homeostasis was unknown and unexplored. However, recent discoveries have shed light on these, and we present here these recent discoveries with a focus on the immunoregulatory influence of DOCK8.

Molecular Homeostasis of DOCK8

DOCK8 belongs to the subfamily of DOCK proteins, which are atypical guanine nucleotide exchange factors (GEFs), which to date consists of 11 proteins numerically named from DOCK1 to 11 (Côté and Vuori, 2007). DOCK proteins activate small G proteins (guanine nucleotide-binding proteins), which are GTPases involved in signal transduction. G proteins bind GTP in their on-state but hydrolyse GTP to GDP and then transition

into an off-state. Re-activation requires dissociation of GDP and binding of GTP—an exchange that is facilitated by DOCK proteins. DOCK proteins consist of two conserved protein domains known as Dock Homology Region 1 and 2, (DHR1 and -2) (Côté et al., 2005). The DHR1 domain is located upstream of the DHR2 domain and mediates the binding to phosphoinositide (PI), which leads to localized GEF activity near the cell membrane (Sakurai et al., 2021; Coˆte´ and Vuori, 2002). The DHR2 domain interacts with the nucleotide-free form of Rho GTPases such as RhoA, Rac, or Cdc42, depending on the DOCK protein. This interaction induces the catalytic activation of the GTPases mediated by GDP-GTP exchange (Harada et al., 2012).

The DHR2 domain of DOCK8 acts as a Cdc42-specific GEF (Kunimura et al., 2020) and therefore regulates Cdc42-specific activities such as cytoskeleton remodelling and actin polymerization, which in turn influence diverse signalling pathways and controls cellular morphology, migration, and protein trafficking (Begum et al., 2004; Li and Gundersen, 2008; Kumari et al., 2014) (**Figure 2**). This has particularly been reported in T and B cells where Cdc42 have been shown to be implicated in cytoskeletal remodelling necessary for functional T cell activation and cytokine secretion as well as defects in B cell receptor signalling and differentiation into plasma cells (Chemin et al., 2012; Burbage et al., 2015; Su and Orange, 2020) Mutations in Cdc42 are associated with an unusual broad spectrum of diverse abnormal phenotypical characteristics which alters morphological appearance and somatic/non-somatic functions. In some patients, cases of immunodeficiency have been reported although the phenotypic spectrum associated with Cdc42 mutations seems wider than that of DOCK8 deficiency (Al-Herz et al., 2016). This would be explained by the ubiquitous expression of Cdc42 whereas DOCK8 expression is largely confined to cells of the immune system, leading to the immune-specific phenotypical characteristics of DOCK8 deficiency[1].

DOCK8 specifically associates with the transcription factor STAT3, which is mutated in AD-HIES. This interaction facilitates activation-induced STAT3 translocation to the nucleus, and the guanine nucleotide exchange function of DOCK8 is also necessary for optimal STAT3 phosphorylation and Th17 differentiation (Su et al., 2019). This functional relationship between DOCK8 and STAT3 explains the phenotypic overlap between DOCK8 deficiency and AD-HIES.

Immunological Impairment

There is a large diversity in immunophenotypical appearance of DOCK8 deficiency patients, which reflects the prominent role of DOCK8 in several key immunological processes (McGhee and Chatila, 2010) either in a cytoskeleton-dependent or -independent immune response in both innate and adaptive immunity (**Figure 2**). DOCK8 therefore serves critical roles in several immune cell types to preserve a broad immune response against bacterial, viral, and fungal agents as well as to sustain self-tolerance.

DOCK8 regulates actin cytoskeletal rearrangement (Dustin, 2002), which has been deemed crucial for facilitating adhesion and formation and functionality of the immunological synapses. This interaction between an immune cell and an antigen presenting cell is mediated by surface components such as the lymphocyte function associated-1 (LFA-1) and the counter receptor Intercellular Adhesion Molecule (ICAM-1), which plays an essential role in the complex cascade of molecular events inducing optimal function and homeostasis of immune cells (Janssen et al., 2016). In the absence of DOCK8, a significant impairment of LFA-1/ICAM-1 binding capacity is observed in $CD8^+$ T cells, Regulatory T cells (Tregs), B cells, T follicular helper cells (Tfh), and T helper (Th) cells explaining some of the broad implications of DOCK8 deficiciency (Randall et al., 2011; Randall et al., 2012; Zhang et al., 2016; Janssen et al., 2017; Janssen et al., 2020).

DOCK8 deficiency causes reduced humoral immunity and self-tolerance. The germinal centers, located in secondary lymphoid organs, facilitate the selection and maturation of antigen-activated B-cell clones and provide an optimal immunological response to infections or immunization (Meshaal et al., 2018). However, in absence of DOCK8, the migration of Tfh cells into the germinal center is impaired (Zhang et al., 2016). This may play a significant part in the impaired maturation of B cells into memory cells (Randall et al., 2009; Caracciolo et al., 2014) and reduced persistence of the germinal centers (Biram et al., 2019). Furthermore, the compromised immunological synapse formation (Zhang et al., 2016) and deficient LFA-1 polarization consequently results in reduced production of high affinity IgG antibodies (Jabara et al., 2012; Zhang et al., 2016; Tang et al., 2019), reduced receptor repertoire, and antibody avidity (Janssen et al., 2014).

A heightened immune response, caused by abnormal regulation of T helper and Tregs and increased IgE production, consequently leading to atopic diseases, is common in DOCK8 deficient patients (Aydin et al., 2015). This may partly be due to the significant numerical reduction of Tregs (Caracciolo et al., 2014; Du et al., 2021). Tregs, a subset of $CD4^+$ T cells, provide an essential negative immunomodulatory function in immune homeostasis and maintaining immune tolerance towards self-antigens (Singh et al., 2017). In the absence of DOCK8, the capacity of Tregs to suppress the proliferation of T cells is absent causing autoimmunity (Shi et al., 2018; Du et al., 2021). This may be attributed to an impaired function of the role of DOCK8 in IL-2 signalling (Randall et al., 2011; Shi et al., 2018), impaired Treg migration (Randall et al., 2021), and defective thymocyte differentiation to Treg (Janssen et al., 2021). Susceptibility towards atopic diseases may in addition be caused by the bias towards Th2 (Engelhardt et al., 2015), lack of peripheral B cell tolerance, and increases in autoantibody production (Du et al., 2021).

One of the key features of DOCK8 deficiency patients is their predisposition for cutaneous infections (Zhang et al., 2014; Aydin et al., 2015). This may stem from abnormal trafficking of immune cells to the skin as DOCK8-mutated T-cells and NK-cells show

[1]Biogps.org

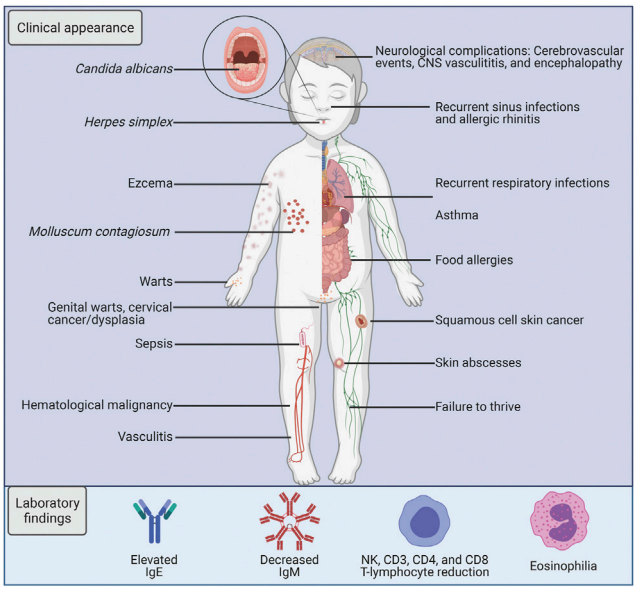


FIGURE 3 | Characteristics of DOCK8 deficiency. A schematic illustration listing the key clinical and laboratory findings in DOCK8 deficiency patients.

impaired morphological integrity leading to cytothripsis during prolonged migration through confined spaces (Lambe et al., 2011). The increased susceptibility to non-skin centred viruses may reflect the progressive lymphopenia, particular of the T cell population, and atypical functionality of T cells due to impaired persistence and recall of antigen-stimulated CD8 T cells, irregular synapse formation with the antigen presenting cells, altered differentiation and impaired proliferation of T cells and DCs (Zhang et al., 2010; Keles et al., 2014; Janssen et al., 2020). In addition, the decreased circulating plasmacytoid dendritic cells and impaired migration of dendritic inhibits normal trafficking to lymph nodes leading to insufficient dendritic cell accumulation in parenchyma for optimal conditions for T-cell priming (Mizesko et al., 2013; Krishnaswamy et al., 2015; Krishnaswamy et al., 2017; Kunimura et al., 2020).

The Natural Killer (NK) cell population exerts an essential antiviral effect by enforcing cellular death of virus-infected cells and is essential for tumour surveillance. The cytotoxic effector function of NK cells in DOCK8 deficient patients is also defective due to impaired lytic synapse formation, abnormal actin accumulation and granule polymerisation (Crawford et al., 2013). Furthermore, DOCK8 is involved in the development of Natural Killer T (NKT) cells and their cytokine production meaning that DOCK8 deficient patients have immature NKT cells and/or NKT cells that display compromised survival (Tangye et al., 2017).

The quantity and function of Th17 T cells is also diminished in DOCK8 deficiency patients (Milner et al., 2008; Caracciolo et al., 2014; Su et al., 2019). Decreased differentiation of Th17 T cells has been accentuated as one of the primary characteristics of HIES (Sandquist and Kolls, 2018). Reduced Th17 differentiation leads to suboptimal activity of anti-fungal and anti-bacterial immunity mainly due to impaired recruitment of neutrophils (Keles et al., 2016). This is partially due to memory $CD4^+$ T cells favouring the production of Th2 cytokines at the expense of Th1 and Th17 promoting cytokines (Milner et al., 2008). In addition, intrinsic factors inhibiting Th17 differentiation due to impaired STAT3 phosphorylation, translocation, and transcriptional activity is also implicated (Al-Herz et al., 2016).

CLINICAL MANIFESTATIONS OF DOCK8 DEFICIENCY

The initial diagnosis of DOCK8 deficiency is based on the clinical characteristics in combination with the laboratory immunological findings, with final verification through genetic analysis. DOCK8 deficiency was described clinically for the first time in 2009 and is characterized by early-onset and severe morbidity. Cohort studies have reported around 50% probability to survive beyond 20 years of age with a mean age at death of 9–12 years (Aydin et al., 2015; Zhang et al., 2014). The disease primarily presents with atopic disease, upper and/or lower respiratory infection, frequent viral cutaneous infections, and malignancy (**Figure 3**) (Zhang et al., 2009; Zhang et al., 2014; Aydin et al., 2015; Haskologlu et al., 2020). Mortality occurs primarily due to infectious agents particularly affecting the skin and respiratory tracts, followed by malignancy, and less commonly CNS vasculitis (Aydin et al., 2015). Almost obligatory findings in these patients are eczema and markedly elevated IgE levels, and there is a high frequency of atopic diseases like food allergies and asthma (Chu et al., 2012; Broides et al., 2017).

Flow cytometric analysis is used to analyse intracellular expression levels of DOCK8, but it is also used to identify B cell maturation arrest and altered frequencies of $CD4^+$ and $CD8^+$ T cells (Caracciolo et al., 2014). The circulating peripheral blood of DOCK8-deficient patients is characterised by a decreased $CD4^+$ T cell count and a shift in the $CD8^+$ T cell compartment towards a more exhausted phenotypic subset (Janssen et al., 2020). The B cell compartment displays an increase in naive B cells and a reduction in memory B cells (Tang et al., 2019). However, differential counts of lymphoid cells can show a significant heterogeneity in abnormal findings, thus warranting additional clinical parameters for diagnosis (Alsum et al., 2013).

DOCK8-deficiency leads to a predisposition of cancer, particularly subtypes of haematological or epithelial origin which are often virally-induced either by Epstein Barr Virus (EBV)-driven leiomyosarcomas and lymphomas, and/or Human Papilloma Virus (HPV)-associated squamous cell carcinomas (Papan et al., 2014; Zhang et al., 2014; Aydin et al., 2015; Haskologlu et al., 2020) In the aforementioned cohort studies, 8–17% of patients had developed malignancies during the follow-up period, which included cases of lymphoma (Burkitt and non-Hodgkin lymphoma), squamous cell carcinoma, and sarcoma (Zhang et al., 2014; Aydin et al., 2015).

To summarize, physicians are encouraged to be vigilant about the clinical manifestations of DOCK8-related primary immunodeficiency. It is mainly characterized as HIES and enhances the susceptibility of recurrent viral and bacterial infection, atopic disease, and higher probability of malignancy.

CURRENT TREATMENT STRATEGIES

Present management of DOCK8 deficiency includes frequent screening for disease progression and treatment of complications through administration of immunoprophylaxis, antiviral and antibacterial treatments prior to a definitive cure through HSCT, if a compatible donor can be identified.

The evidence supporting allogeneic HSCT for treatment of DOCK8 has been described through multiple reports (Chu et al., 2012; Aydin et al., 2019). HSCT is performed after an initial myeloablative or reduced intensity conditioning regimen consisting primarily of chemotherapy and occasionally with additional radiotherapy (Aydin et al., 2019). The purpose of conditioning is to induce adequate immunosuppression and ablation of the recipient's hematopoietic stem cells. Administration of interferon alpha has shown efficacious as a rescue therapy for viral infections such as Herpex Simplex virus (HSV) and Human Papilloma virus (HPV) (Al-Zahrani et al., 2014; Gernez et al., 2018). In addition, patients awaiting stem cell transplantation benefit from immunoglobulin replacement therapy (IVIG) and prophylactic treatment (Bazinet and Popradi, 2019). Antibodies in DOCK8 patients have been shown to display reduced avidity, which is why IVIG is recommended despite normo- or hyperphysiological antibody levels (Janssen et al., 2014). Furthermore, prophylactic treatment is given post-transplantation to prevent infection and non-infectious complications in the period until immune reconstitution (Devetten and Armitage, 2007).

The detection and aggressive treatment of infectious disease is paramount to avert fatal progression leading to death. However, a large retrospective report consisting of 136 patients with DOCK8 immunodeficiency accentuates the severe disease progression. Unfortunately, even with early intervention with aggressive therapies or prophylaxis such as anti-bacterial, fungal, viral, immunomodulatory, and immunoglobulin replacement treatment, 63% of the patients succumb by their fourth decade of life (Aydin et al., 2015). Therefore, early allogeneic HSCT is clearly indicated, which is also the sole possibility for a curative treatment. Advancements in HSCT have vastly increased the post-HSCT survival for DOCK8 deficiency patients. Hence, patients undergoing HSCT between 1995 and 2010 had a 2-years overall survival of 57 *versus* 92% for patients transplanted between 2011 and 2015 (Aydin et al., 2019). However, HSCT still entails several risks of severe adverse events, particularly from haploidentical relatives or unrelated donors which pose risks of graft-versus host disease and graft failure (Fox et al., 2018). Reports show that among DOCK8 deficient patients undergoing HSCT, 33% develop acute graft versus host disease (Aydin et al., 2019). Furthermore, HSCT is not always available due to disease progression because of late or misguided diagnosis, and the lack of HLA-compatible donors (Broder et al., 2017; Slatter and Gennery, 2018; Gavrilova, 2019).

Finally, HSCT is associated with high expenses and its use is gradually increasing thus indicating the need for novel therapies that reduce the overall medical cost associated with HSCT and its side effects (Morgan et al., 2017; Mayerhoff et al., 2019). Autologous HSCT with genetically modified stem cells may constitute a promising therapy with fewer adverse outcomes and higher availability to patients with DOCK8 deficiency (Kim et al., 1996; Talib and Shepard, 2020).

THE THERAPEUTIC PROMISES OF GENOME EDITING

Our ability to precisely rewrite and manipulate the instructions encoded in the genome has greatly expanded over the last few decades. By using programmable nucleases, such as Zinc Finger Nucleases (ZFNs) (Miller et al., 2007; Szczepek et al., 2007; Christian et al., 2010), TALE nucleases (TALENs) (Miller et al., 2011; Mussolino et al., 2011; Jinek et al., 2012) and RNA-guided Cas nucleases (Cong et al., 2013; Tebas et al., 2014), researchers can direct the creation of double-strand breaks (DSBs) to specific sites in the DNA and hereby exploit the cellular DNA repair machinery to introduce desired genetic modifications. Induced DSBs are repaired through either Non-Homologous End-joining (NHEJ), an error-prone repair mechanism that induces insertion or deletions (INDELs) at the DNA breakpoint, or Homology-Directed Repair (HDR), a precise repair pathway that uses homologous repair templates to copy from during repair of the DSB, allowing the inclusion of foreign DNA sequences into a specific locus (**Figure 4A**).

Using programmable nucleases, gene disruption or remodeling of regulatory sequences can be easily achieved by NHEJ. These are strategies that have already been used in clinical

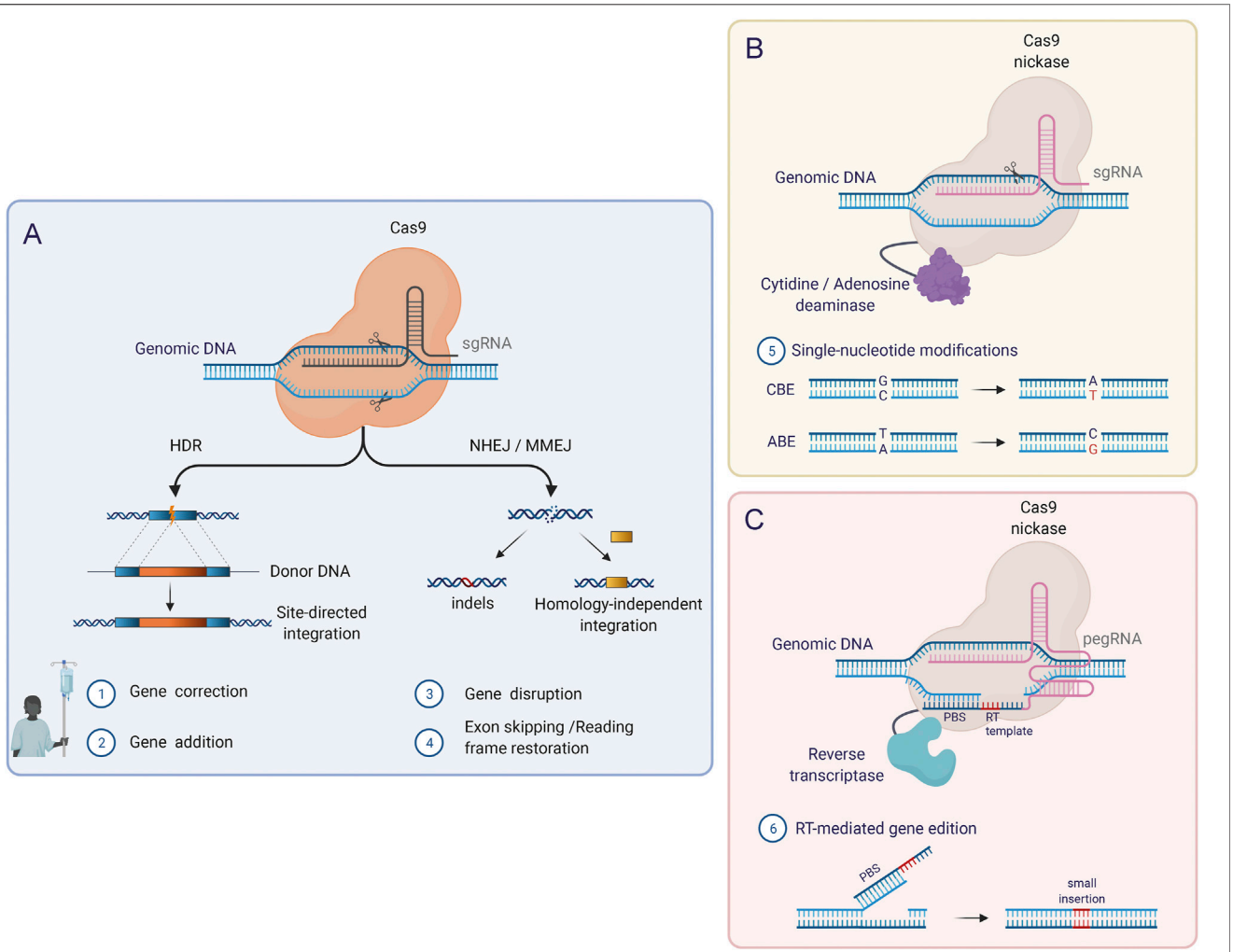


FIGURE 4 | Genome editing tools based on CRISPR/Cas. **(A)** The original CRISPR/Cas RNA-guided nuclease system induces DSBs in the genome in a targeted manner and requires binding of the single guide RNA molecule (sgRNA) to the target DNA as well as the recognition of a specific PAM sequence adjacent to the target sequence. Using this system, the two major DNA repair pathways (HDR and NHEJ) can be exploited to introduce modifications at the target locus. **(B)** Base Editors (BEs) combine Cas nickases lacking one nuclease domain with DNA deaminases. BEs mediate single-nucleotide conversion, which enables the correction of point mutations. Two types of BEs can be distinguished: Cytosine base editors (CBEs), which mediate C-G to T-A conversions, and adenine base editors (ABEs) which induce A-T to G-C conversions. **(C)** Prime Editors (PEs) can induce point mutations, small insertions, and small deletions and consist of a Cas nickase fused to a reverse transcriptase (RT) domain. In this system, a reformulated prime editing gRNA (pegRNA) confers the specificity (like the sgRNAs) but additionally contains the template for the desired DNA modifications in the 3′ end. After the induction of a single-strand break (SSB) by the Cas9 nickase, the 3′ end of the pegRNA (primer binding site; PBS) hybridizes with the free 3′ DNA end and acts as a reverse transcription template.

trials with promising results (Bushman, 2007; Xu et al., 2019; Stadtmauer et al., 2020; Frangoul et al., 2021; Gillmore et al., 2021). However, for most autosomal recessive PIDs, the addition, replacement, or correction of the affected gene is required thereby necessitating the use of the HDR pathway. This can be achieved by direct correction of the genetic mutation or through the integration of the complete or partial open reading frame cDNA sequence either directly after the endogenous promoter or into a safe harbor site in the genome with a heterologous promoter.

The ability of hematopoietic stem cells (HSCs) to perpetually self-renew and differentiate into all hematopoietic lineages makes them an ideal therapeutic target for gene therapy for treating blood and immune system diseases, including PIDs. Several gene therapy clinical trials have been performed in HSCs since the first one in 1990, and they have mainly been performed using retroviral vectors that integrate the transgene into the chromosomes of HSCs in a semi-random fashion (Booth et al., 2016). Despite clinical trials for several PIDs showing curative potential of these gene therapies, there are also multiple reports of patients developing leukemia due to insertional mutagenesis caused by the retroviral vector (Dever et al., 2016) These events were caused by vector integration close to cancer-related genes such as LMO2 and transactivation of these genes by the strong viral promoter/enhancer elements present in the retroviral vectors. Although lentiviral vectors with self-inactivating mechanisms have proven

a safer alternative, they have not entirely eliminated the risks associated with random vector integration, thereby rationalizing the pursuit for safer alternatives such as precise genome editing. Undoubtedly, the advent of the CRISPR/Cas gene editing system supports the onward march towards precision gene therapy.

Initially the CRISPR/Cas system was comprised of the Cas9 endonuclease and two small RNAs: CRISPR RNA (crRNA) and *trans*-activating RNA (tracrRNA). While crRNA confers specificity to a complementary region in the genome and thereby serves to guide Cas9 to its target, the tracrRNA acts as a Cas9 binding handle to enable the formation of a ribonucleoprotein (RNP) complex. In seminal work from the 2020 Nobel prize winners Jennifer Doudna and Emmanuelle Charpentier they merged the crRNA and tracrRNA into a single guide molecule (sgRNA) thereby reducing the system from three to two components (Cong et al., 2013).

The therapeutic potential of targeted gene editing in long-term repopulating HSCs (LT-HSCs) was first demonstrated in a preclinical study for X-linked Severe Combined Immunodeficiency (SCID-X1) using ZFNs delivered by mRNA electroporation and repair template delivery by an integration-defective lentiviral vector (IDLVs). In this study, Genovese et al. achieved targeted integration of a partial IL2RG cDNA comprising a super-exon of exons 5–8 of IL2RG into exon 5 of the endogenous IL2RG gene. Thus, transcription occurs from the endogenous IL2RG promoter and exon 4 splices with the newly inserted super-exon to generate a correct and full-lenght IL2RG reading frame, thereby providing a platform with the potential to correct all SCID-X1 IL2RG mutations downstream of exon 4. Long-term engraftment of the targeted HSCs in transplanted NSG mice was confirmed and they were also able to correct the defective IL2RG gene in HSCs from a patient with SCID-X1. The CRISPR/Cas system was similarly applied in HSCs for the first time to correct the mutation in the β-globin gene responsible for Sickle Cell Disease (SCD) (Lattanzi et al., 2021). Here, precise correction of the disease-causing mutation was performed with similar evidence of long-term engraftment in mice and reconstitution of functional β-globin (Mohrin et al., 2010). However, both studies also revealed what is now considered the largest challenge for applying precise HDR-mediated gene editing in HSCs, which is the low efficiencies of HDR-mediated editing in the long-term (LT)-HSC compartment compared to the progenitor cell population of the total $CD34^+$ cells. This is mainly believed to be caused by HDR only being active in the late S and G2 phases of the cell cycle whereas NHEJ is the prevailing repair mechanism in quiescent cells (Wilson et al., 2008). This poses a paradoxical challenge in HSC-based gene editing since otherwise quiescent HSCs must be forced into cycling, but cycling is known to be associated with loss of stem cell properties (Song et al., 2016). Hence, NHEJ-focused HSC therapies have shown higher efficiencies in HSCs, confirmed in a recent clinical trial (Frangoul et al., 2021), but is generally more difficult to apply for recessive disorders where expression of a functional gene must be restored.

To enhance HDR frequencies, researchers have been working on different strategies that include the use of repair pathway-modulating small molecules that promote HDR or inhibit NHEJ (Lin et al., 2014; Chu et al., 2015; Robert et al., 2015; Nambiar et al., 2019; De Ravin et al., 2021; Fu et al., 2021), cell cycle synchronization to ensure S/G2 status upon editing (Charpentier et al., 2018; Shin et al., 2020), and the development of novel engineered Cas9 variants for example fusing HDR-promoting or NHEJ-inhibiting proteins to Cas9 (Jayavaradhan et al., 2019; Ferrari et al., 2020). New protocols also transiently inhibit the p53 pathway to achieve high percentages of HDR editing in LT-HSCs (Vavassori et al., 2021). The most advanced example of this is a recent study correcting the CD40 ligand gene (CD40LG), which has deactivating mutations in X-linked hyper-IgM syndrome type I (HIGM1) (Schiroli et al., 2019). Here, the authors introduce mRNA encoding a dominant negative p53 variant (GSE53) along the CRISPR/Cas9 gene editing components. Prior studies have shown that during gene editing in HSCs, DSBs and the presence of adeno-associated virus (AAV) vector genomes, used to deliver the HDR repair template, activate p53, which constrains HSC yield, proliferation, and engraftment of gene-edited HSCs (Cromer et al., 2021). With the addition of GSE53, the authors showed up to 30% CD40LG correction frequencies (cDNA integration) in LT-HSCs.

The HDR pathway also provides the possibility of replacing entire gene sequences or large genomic regions. This was recently showcased in LT-HSCs where a DNA repair template was designed in a way that the copy-paste mechanism replaced the one pf the α-globin genes (HBA1) with that of β-globin. This approach could prove therapeutic in patients with β-thalassemia to normalize the balance between α and β chains, thus restoring adult hemoglobin functionality (Maresca et al., 2013). Moreover, this study showed that whole gene replacement is possible in LT-HSCs, thereby providing an additional genome editing strategy for genetic diseases.

As alternatives to HDR, novel gene correction approaches based on HDR-independent targeted gene integration (Sakuma et al., 2016; Yao et al., 2017; Porto et al., 2020), base editing (Anzalone et al., 2019), and more recently prime editing (Newby et al., 2021) try to overcome the inherent limitations of HDR-mediated genome editing. In base and prime editing, the functional properties of Cas9 can be extended by fusing new effector domains to catalytically inactive Cas9 protein or Cas9 nickase. In this way, without requiring DSBs or donor DNA templates, Base Editors (BEs) mediate single-nucleotide conversions in a targeted manner, while Prime Editors (PEs) write new genetic information into a specific nicked locus directed by a small template present on the sgRNA (**Figures 4B,C**) (Newby et al., 2021). These are promising alternatives for genome editing, and BEs have already shown as high as 68% base editing of the β-globin gene in human LT-HSCs evaluated 16 weeks after transplantation into mice (Vaidyanathan et al., 2018) Similar evidence for PEs must be provided to reinforce their applicability in HSCs.

CRISPR/Cas9 Delivery Strategies in HSCs

Delivery has for a long time been the main hurdle for advancing gene therapy. Since HSCs were the first stem cells to be discovered, purified, and used for therapy (bone marrow transplants), HSCs were also obvious first choice for gene

therapy since *ex vivo* gene therapy is much simpler than *in vivo* gene therapy.

In general, three different approaches exist to introduce the two components of the CRISPR/Cas-system into cells. In the first approach, DNA such as plasmid DNA is used, encoding the Cas9 and sgRNA. Plasmid delivery is associated with rather slow onset of editing, when compared to the other modalities (Vaidyanathan et al., 2018). The second "all-RNA" approach delivers mRNA encoding Cas9 along in vitro-transcribed or chemically synthesized sgRNAs (Laustsen et al., 2019). Lastly, a recombinant Cas9 protein precomplexed to the sgRNA as an RNP complex can be delivered into cells (Genovese et al., 2014; Lino et al., 2018; Vakulskas et al., 2018). However, in most primary cells, the cost-effective plasmid delivery approach leads to high undesirable cytotoxicity (Cromer et al., 2018; Lino et al., 2018). All-RNA delivery using Cas9 mRNA and sgRNAs is better tolerated by primary cells, even though they still induce a higher innate immune response than RNP delivery (Hendel et al., 2015). Hence, the preferred delivery format for HSC gene editing is using RNP complexes with sgRNAs that are chemically synthetized with modified nucleotides at both ends to protects them from degradation by exonucleases (Laustsen et al., 2019). The delivery mode of choice in HSCs is using electroporation, which relies on short pulses of electrical current to induce small pores in the cell membrane that allows diffusion of macromolecules (Gundry et al., 2017). Combined with Cas9 RNP, this mode has shown exceptionally high on-target efficiency. At the same time, the short half-life of the RNP complex provides a hit-and-run modality that reduces the risk of off-target activity at sites that resemble the intended target in the genome (Genovese et al., 2014; De Ravin et al., 2017; Lino et al., 2018; Vaidyanathan et al., 2018).

In addition to the Cas9 and sgRNA, HDR requires the introduction of a repair template. Here, non-viral repair templates like chemically synthesized single-stranded oligodeoxynucleotide (ssODN) of up to 200 nt have shown effective in HSC gene editing (DeWitt et al., 2016; Romero et al., 2019). However, ssODNs have been reported to be less efficient and induce higher toxicity compared to repair template delivery approaches that rely on viral vectors. At the same time, ssODNs suffer from size constraints associated with chemical DNA synthesis (Roth et al., 2018). However, recent work establishes evidence for surpassing the size constrains of ssODNs enabling delivery of >1 kb long dsDNA co-electroporated with Cas9 RNP complexes in primary human T cells with a tolerable toxicity profile (Wang et al., 2015). The applicability of this platform in HSC is intriguing, but needs further investigation. Lentiviral vectors, which have been employed in numerous clinical trials for *ex vivo* HSC gene therapy, have also been employed as repair template for HDR. Here, the natural integrating mechanism of lentiviral vectors is removed by introducing an inactivating mutation in the viral Integrase enzyme to generate integration-defective lentiviral vectors (IDLVs) suitable for donor DNA delivery. IDLVs do have higher carrying capacity than AAV vectors, but have been shown to be inferior to AAVs when used for repair template delivery, where specifically AAV serotype 6 has proven effective in HSCs with tolerable cytotoxicity (Grieger and Samulski, 2005). Despite the relatively low packaging capacity of AAV vectors at around 4.5 kb (Bak and Porteus, 2017), this is sufficient for most genome editing purposes, and an approach splitting a large transgene between two AAV donors that undergo sequential HDR at the target locus has been devised to overcome this limit (Bak et al., 2018b). Hence, the combination of RNP complexes with simultaneous delivery of DNA donor template using AAV6 currently represents the most promising technology for versatile gene editing in HSCs (Hacein-Bey-Abina et al., 2002; Bak et al., 2017; Wilkinson et al., 2021).

TOWARDS A CURATIVE CRISPR/CAS9-BASED GENE EDITING APPROACH FOR DOCK8-RELATED PRIMARY IMMUNODEFICIENCY

Retroviral gene therapy in autologous HSCs has provided clinical benefit in several PIDs including SCID-X1 (Aiuti et al., 2002; Gaspar et al., 2004; Hacein-Bey-Abina et al., 2010; Gaspar et al., 2011), Adenosine deaminase deficiency (ADA-SCID) (Aiuti et al., 2009; Boztug et al., 2010; Cicalese et al., 2016; Scott et al., 2017), Wiskott Aldrich Syndrome (WAS) (Ott et al., 2006; Aiuti et al., 2013; Hacein-Bey Abina et al., 2015), and X-linked chronic granulomatous disease (X-CGD) (Cavazzana-Calvo et al., 2010; Kang et al., 2010). However, integration of the transgene in these approaches occurs semi-randomly into the patient's genome, and as mentioned earlier this can lead to oncogenic transformation due to insertional mutagenesis. Even vectors with lower genotoxic potential, such as lentiviral vectors (LVs), can give rise to insertional mutagenesis, which is a risk that needs to be considered in clinical applications (Brown et al., 1998; Modlich et al., 2009). Furthermore, transgene expression levels often differ from the physiological levels of the affected gene due to the use of a constitutive promoter that does not allow tissue-specific or temporal regulation of expression. Unregulated gene expression can for some diseases be detrimental exemplified by the CD40LG gene, which must be expressed at controlled levels as evidenced by preclinical mouse studies where *ex vivo* gene therapy in a mouse model of X-linked hyper-IgM syndrome with CD40LG-encoding murine gamma-retroviral vectors in HSCs led to lymphoproliferative disorder assumingly as a consequence of unregulated expression of the CD40L transgene (Glessner et al., 2017). DOCK8 deficiency has not been approached with retro- or lentiviral gene delivery, but copy number variation analyses have identified DOCK8 duplications to be significantly associated with a spectrum of neuropsychiatric disorders (Jing et al., 2014). Even though a direct link from DOCK8 CNV to immunological defects has not been established, this might suggest that elevated levels of DOCK8 gene expression can impede normal cellular function. Tight DOCK8 expression in different mature immune cell subsets and regulated DOCK8 expression during hematopoiesis would be impossible to establish with LVs carrying constitutive heterologous promoters and would require full delineation of the regulatory

mechanisms that govern DOCK8 expression and reconstruction of a DOCK8 promoter suitable for LV use. Hence, LV-mediated gene delivery may not be a therapeutic option in DOCK8 deficiency, whereas precise gene editing approaches may be optimally suited for such diseases. In the following section, we summarize essential considerations for gene editing and describe different gene editing strategies and their potential use for correcting DOCK8-mediated immunodeficiency.

Potential Gene Correction Strategies for DOCK8 Deficiency

Utilizing the Non-Homologous End-Joining Pathway

An NHEJ-based treatment strategy for DOCK8 deficiency would be highly desirable since the NHEJ pathway is much more active in HSCs compared to the HDR pathway. However, due to the autosomal recessive nature of DOCK8 deficiency, an NHEJ-based strategy that introduces INDELs in the genome is challenging to apply to DOCK8. For disease-causing variants where the reading frame is disrupted, one option for the NHEJ pathway is to use the "reframing" approach which relies on introduced INDELs to restore the correct reading frame (**Figure 4A**). However, CRISPR/Cas-generated INDELs occur in a semi-stochastic fashion, which means that a population of edited cells will contain a mix of different INDELs, which are specific to the sgRNA used. Hence, this approach depends on the availability of a sgRNA in the vicinity of the mutation, which creates reframing INDELs. Depending on the type and location of these reframing INDELs there may be a loss or addition of amino acids to the reading frame, which may in some rare instances establish a dominant gain of function variant. However, the frequency of reframing INDELs may only need to be low, as evidenced by the aforementioned cases of somatic reversion establishing the possibility that correction of a single lymphoid progenitor or stem cell may be sufficient to provide therapeutic benefit. Depending on the type and location of these reframing INDELs there may be a loss or addition of amino acids to the reading frame, which may perturb protein function. However, for some patient mutations, this approach may be applied. Recent preclinical studies have in fact used reframing to correct mutations in HSCs from patients with Fanconi Anemia where there is a great survival adantage (Román-Rodríguez et al., 2019). The downside of this approach is that it cannot be generalized but would need a different sgRNA for each patient mutation. Also, for longer gene deletions, this approach would not be feasible and 61.5% of patients harbour deletions. This individualized approach is costly to develop and therefore difficult to bring to clinical trials.

Base and Prime Editing

BEs are promising new tools for gene editing, but they can only address a subset of mutations (**Figure 4B**). In theory, Cytosine BEs enable correction of 26% of all known pathogenic SNP variants, while the Adenine BEs could potentially correct 28%. However, in DOCK8 deficient patients, only a subset of approximately 26% of patients carry pathogenic SNPs in DOCK8. Therefore, even the pursuit to develop individualized base editors for specific patient mutations would only be possible for a subset of the patients. Similar challenges exist for the recently developed PE (Newby et al., 2021) (**Figure 4C**). The most distinctive attribute of this technology and advantage over BEs is its ability to make any small sequence changes. Like the BEs, this occurs without inducing a DSB - thus mitigating the error-prone NHEJ pathway and the low rates of HDR in post-miotic cells (Scholefield and Harrison, 2021). However, PE is limited to make insertions of less than 80 bp and deletions smaller than 50 bp (Kim et al., 2014). This only enables correction of around 15% of the current DOCK8 patient mutations. Furthermore, there have not been any reports of efficient PE in HSCs.

Homology-Independent Targeted Insertion

Homology-independent targeted insertion (HITI) is a gene editing approach that enables integration of DNA at a specific target site without relying on homologous sequencies in the DNA donor (**Figure 5F**). Instead, it uses the NHEJ pathway to integrate the linear DNA donor at the break site through an end-joining mechanism, but its efficiency has been reported to be less than 5% in most cases (Suzuki and Izpisua Belmonte, 2018). The applicability of HITI in $CD34^+$ HSPC might be high as these quiescent cells are often in the G0/G1 phases of the cell cycle where NHEJ is the primary DNA repair pathway (Suzuki and Izpisua Belmonte, 2018). Recently, Hanan Bloomer et. al. utilized the HITI system reaching an average of 21% integration in long-term repopulating HSC in mouse xenotransplantation studies (Bloomer et al., 2021). However, the integration mechanism allows transgene cassette insertion in both orientation and the NHEJ mechanism can also lead to end-trimming of the DNA donor template and/or the genomic target site. Hence, this pathway does not lead to exact genome editing outcomes, which might impede its application for DOCK8 deficiency. At present, the technology would need to mature and further studies would need to be conducted to proof its applicability in HSCs.

Homology-Directed Repair

The most versatile method in terms of possibilities of gene editing outcome is utilizing the HDR pathway, which requires co-delivery of a homologous DNA repair template. In this template, the new DNA sequence is flanked by DNA sequences that are homologous to the target gene sequence around the cut site (**Figure 4A**). HDR enables various kinds of genetic alterations ranging from single base pair changes to whole cDNA insertion strategies. In contrast to gene delivery by retroviral vectors, HDR preserves endogenous DOCK8 gene regulatory elements with the possibility to re-establish physiological gene expression levels.

Single base pair correction has previously been shown in $CD34^+$ HSCs and is particularly advanced for Sickle Cell Disease (Magis et al., 2019; Lattanzi et al., 2021) This disease is one of the most prevalent genetic disorders and is caused by a single nucleotide substitution that changes a glutamic acid into valine. A direct base pair correction approach is highly desirable since no major perturbations are made to the gene and all regulatory elements of the promoter, untranslated regions, splice elements, and introns are maintained (**Figure 5A**). This approach could also be applied to DOCK8 mutations but suffers from the same challenge as base and prime editors as mutation-specific CRISPR/Cas reagents must be developed, constituting a major financial burden.

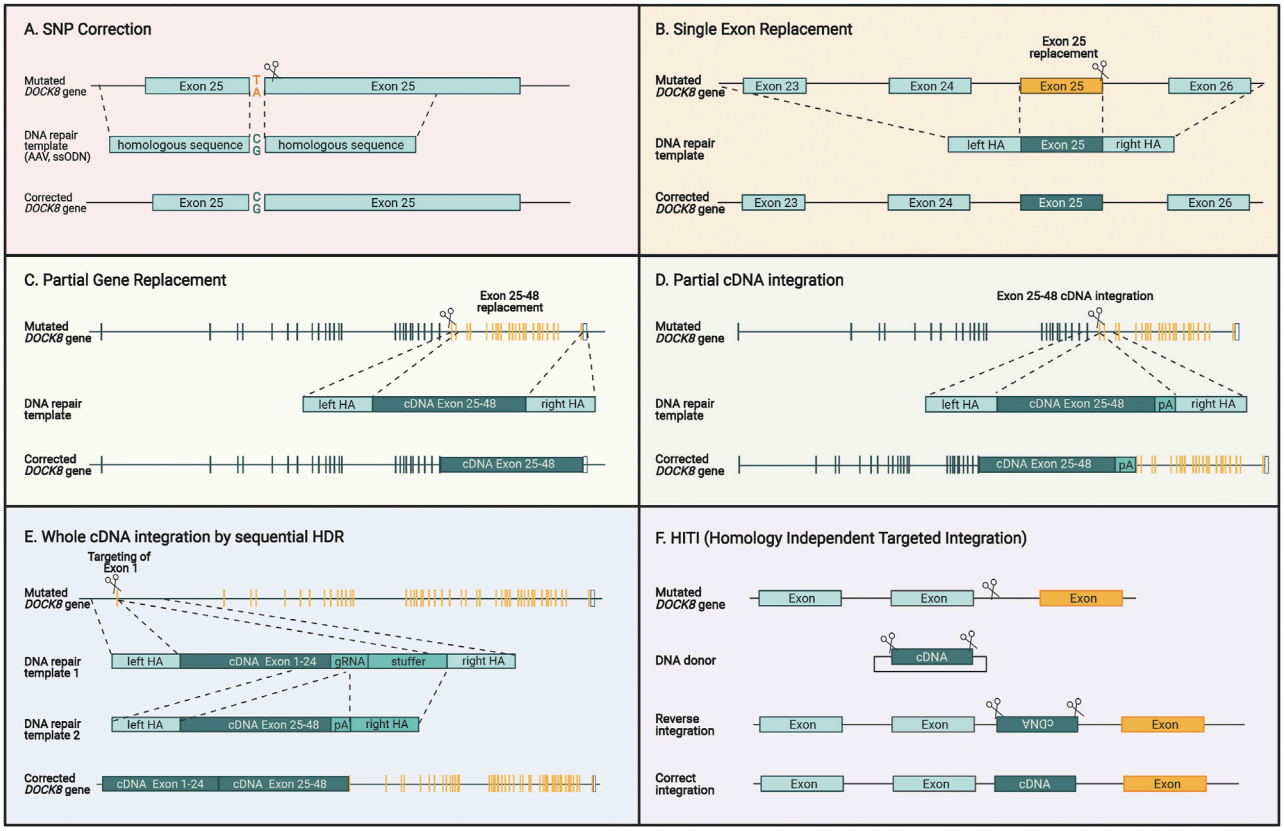


FIGURE 5 | Different genome editing strategies for restoring DOCK8 gene expression. **(A)** Specific correction of SNPs by providing DNA repair templates with homologous sequences surrounding the mutation. **(B)** Single exons can be replaced by introducing a double strand break close to the end of the exon in addition to supplying a DNA repair template encoding homologous sequences surrounding the exon. **(C)** Multiple exons can be replaced by cutting close to the region which is intended to be replaced and providing a DNA repair template with homology arms that are homologous to the adjacent regions. **(D)** A cDNA sequence covering several exons can be introduced upstream to the mutated sequence and the homology arms contain the sequences surrounding the cut site. This way the cDNA will be fused directly to the previous exons, and the downstream exon will be inactivated. **(E)** To replace the entire DOCK reading frame, a two-step HDR approach can be applied due. For utilizing AAV vectors for repair template delivery, this two-step approach is necessary due to the large size of the gene and the limited capacity of AAV vectors. **(F)** HITI is different to the previous HDR-based strategies since it uses the NHEJ pathway to insert a DNA sequence without homology arms directly at the cut site. Since this insertion is not exact, intronic regions are targeted to insert a desired sequence that maintains the normal splicing mechanisms and keeps endogenous expression intact. For all figures, the yellow sequences designate exons where patient mutations will be corrected by the specific approach.

A more universal approach would be to insert part of the DOCK8 reading frame into the endogenous DOCK8 locus thereby spanning larger segments of the gene and potentially covering several patient mutations. This could be used to replace single exons, multiple exons, and potentially all exons (**Figures 5B–D**). Such strategies have been used preclinically before for multiple hematopoietic diseases including β-thalassemia (Lattanzi et al., 2021), X-SCID (Cromer et al., 2018; Pavel-Dinu et al., 2019), hyper-IgM syndrome (Hubbard et al., 2016; Kuo et al., 2018; Schiroli et al., 2019), and X-CGD (De Ravin et al., 2016; DeWitt et al., 2016). Since AAV is the preferred vector for delivery of the repair template, DOCK8 represents a particular challenge due to its large ORF size of 6.3 kb. The maximum size of AAV packaging is around 4.7kb, which must include the homology arms, which are normally 2 x 400bp leaving around 3.9 kb for the cDNA to be inserted. For DOCK8, this would allow inserting approximately 60% of the DOCK8 ORF. To include as many patient mutations as possible, the optimal region to target would be exons 25–48 encompassing 39% of the known patient mutations (**Figure 5D**).

We have previously devised an HDR strategy for integrating large gene segments exceeding the capacity of AAV6 vectors. Here, the large transgene is delivered using two separate AAV repair template vectors (Bak et al., 2018b). This makes use of two consecutive HDR steps that first integrate one half of the gene and then the next half, hence overcoming the capacity limit for a single AAV, which potentially would enable a universal HDR approach to target all known DOCK8 mutations present in the reading frame (**Figure 5E**) (Balakrishnan and Jayandharan, 2014). This approach has recently shown promise as a curative correction strategy for cystic fibrosis, which involves the large 4.4 kb CFTR ORF (Vaidyanathan et al., 2021).

One key aspect is tailoring the donor design and targeting strategies for optimal expression of the target gene. Studies show that for some genes, mere integration of the full reading frame cDNA into the start codon of the endogenous locus leads to

suboptimal gene expression levels. Physiological transgene expression can be reached by improving several steps including biallelic integration rates, cDNA codon usage, and inclusion of transcriptional and posttranscriptional regulatory elements. Some studies have used the cDNA of the reading frame only, thereby excluding the 3′UTR from the constructs. This exclusion has advantages during HDR since the 3′UTR sequence in the DNA donor template cannot be diverged from the endogenous 3′UTR like is possible for the reading frame using synonymous codons. This "internal homology" creates two sites of homology between the DNA donor template and the genome: (1) the site of the DSB to which the homology arms have homology and (2) the 3′UTR which is distant from the DSB when designed for the region around the endogenous start codon. This double homology creates the possibility of unpredictable HDR events, which is the reason why the exclusion of the 3′ UTR is preferred. However, 3′ UTRs are known to include several regulatory elements like miRNA binding sites and AU-rich elements that can be stabilizing or destabilising to the mRNA. An example of this is gene editing with CD40LG replacement by HDR where the inclusion of the 3′ UTR is essential to ensure high gene expression levels (Hubbard et al., 2016; Kuo et al., 2018). Studies have also shown that retaining intron one or the terminal intron in the cDNA can contribute positively to efficient expression levels (Gray et al., 2021; Sweeney et al., 2021).

The Societal Challenges of Bringing CRISPR/Cas9 Gene Therapies for DOCK8 Deficiency to Patients

The European Commission defines rare diseases as those with a prevalence below five of every 2,000 people (<0.25%) (European Commission (2021, 2021). Although no official definition of "ultra-rare" disease has been established, the European Union defines orphan medicinal products as those that address individuals affected by severe or life-threatening diseases which affect no more than 5 persons in 10,000 (<0.05%) withing the European union. Therefore, based on the current literature entailing DOCK8 deficiency diagnoses, this disease may be defined as an ultra-rare disease.

Several financial, logistical, and ethical questions arise when considering orphan drugs for rare diseases. There are few to no financial incentives for biopharmaceutical companies to venture into ultra-rare disorders due to high development costs and the prospect of low revenues, which is only circumvented with soaring treatment prices. This was demonstrated, with the recently approved one time curative gene replacement therapy for Spinal Muscular Atrophy, which cost more than two million US dollars (Dean et al., 2021). Furthermore, non-economic aspects such as the need for accelerated approval of orphan drugs and how to encourage cooperation between countries and stakeholders in the pursuit of bringing orphan drugs to the marked is worth considering (Kacetl et al., 2020). There has been rapid growth in orphan drug policy establishment, and it is important that governments establish incentives that promote research and development for these indications (Chan et al., 2020). For ultra-rare disease gene therapy development, it is pivotal that translational research and clinical trials are performed in international collaborations to promote access to patient samples and ultimately to centralize clinical trials and coordinate logistical challenges. While this might be possible in high-income countries, several stakeholder are now recognizing that the single largest challenge will be providing access to novel and expensive treatment modalities for patients from low- and middle-income countries and patients from disadvantaged communities and ethnic groups (Gene, 2021). Because gene therapies involve complex procedures during the GMP-compliant manufacturing of a living cell product (**Figure 6**), there is a need for advanced infrastructure and highly educated personnel. This demand constitutes a major bottleneck for many institutions with the desire to treat patients with novel gene therapies, and they often fall short in meeting the demands from patients who are in critical need to gain access to potential life-saving therapies. In the future, semi-automated closed cell manufacturing systems might make it easier to implement gene therapies locally (Adair et al., 2016).

As the possible applications of gene therapy exceed beyond the field of research, ethical concerns arise and are discussed to establish a framework for appropriate application of genetic therapies. One important ethical consideration in gene therapy is the distinction between somatic and germline gene editing. The 2018 reports of gene edited Chinese twin babies sparked multiple calls for a global moratorium on clinical uses of human germline editing (Lander et al., 2019). Important discussions arose from this event concerning ethical issues such as patient safety, missing consent from the unborn child, uncertainty of the monitoring period of adverse events, how to justify the defying of natural order, and the potential future implementation of gene therapy as a preventive treatment (Araki and Ishii, 2014). While these discussions on germline gene therapy are ongoing, it is important to remember that somatic gene therapy is bringing about an increasing number of success stories, thereby challenging the need for germ line therapies.

Current Technological Challenges of Gene Editing Reaching the Clinical Setting

Human cells have acquired several mechanisms to detect and correct genomic lesions as each cell experiences several thousand DNA lesions daily (Carusillo and Mussolino, 2020). Most gene editing technologies piggyback on these mechanisms, but also suffer from the adverse events of activating DNA damage responses. This is not only caused by DNA double strand breaks but are also invoked by exposure to AAV vectors. This can lead to cumulative p53 pathway activation which can negatively impact engraftment of edited HSCs (Cromer et al., 2021).

Off-target INDEL induction, translocations, chromothripsis, large on-target INDELs, and off-target integration of DNA donor template are other non-intended consequences that can potentially lead to adverse events and must therefore be evaluated (Urnov, 2021). Off-target activity is based on following three elements: the uniqueness of the target site, the chromatin state of the genome, and nuclease exposure duration and efficiency. Cas9 specificity is dependent on target homology with the 20-nt spacer region of the gRNA, but this sequence can tolerate a mismatch of

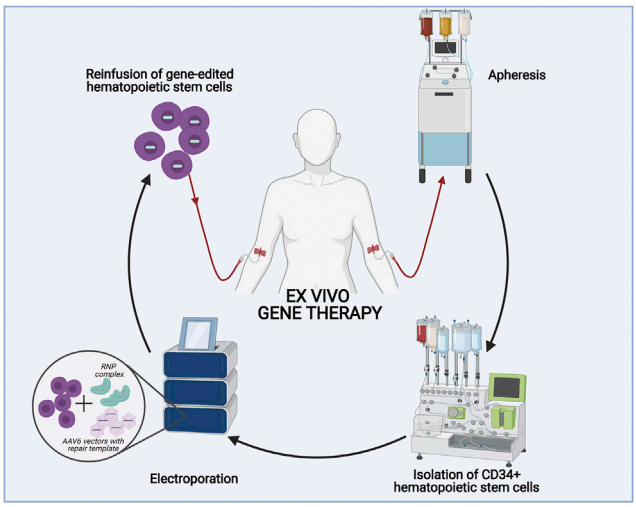


FIGURE 6 | Principles of *ex vivo* gene editing. For a patient specific treatment, $CD34^+$ hematopoietic stem cells are isolated from a patient's blood through apheresis followed by an automated cell processing system that enrich $CD34^+$ cells. Afterwards, the genome editing tools (Cas9 and sgRNA) are delivered to the cells by electroporation and the DNA repair templates are delivered by addition of AAV6 vectors to the cell culture. The genetically modified cells can be cryopreserved before infusion into the patient, which has typically undergone myeloablative conditioning prior to infusion.

several bases, consequently enabling possible binding to secondary unintended regions (Mali et al., 2013). In addition, off-target activity is more common to occur at open chromatin regions rather than at closed chromatin region (Singh et al., 2015; Kim and Kim, 2018). Minimizing the off-target activity can theoretically be reached through either increasing the nuclease dissociation from Watson-Crick base paired genomic regions or reducing its cleavage rate (Bisaria et al., 2017). Several Cas9 variants have been engineered with such overall reduction in DNA affinity, thereby maintaining on-target affinity within a window of maximal cleavage while reducing off-target affinity with concomitant reduction in cleavage (Genovese et al., 2014). Several unbiased detection methods have been developed to identify off-target site, like GUIDE-seq and DISCOVER-Seq (Tsai et al., 2015; Wienert et al., 2019) while other methods like CAST-seq evaluates translocations and other gross rearrangements (Turchiano et al., 2021). Overall, with careful sgRNA design and implementation of novel technologies like high-fidelity Cas9 proteins, these adverse genomic events can be reduced and are often benchmarked against existing lentiviral vectors that integrate their cargo semi-randomly in the genome. Off-target activity and insertional mutagenesis remain important concerns in gene therapy and more clinical trials and long-term follow-up is the only means to truly gauge the proportions and clinical relevance of these events.

CONCLUSION AND FUTURE DIRECTION

In the near future, we expect several genetic diseases such as DOCK8 deficiency, which lack efficient and safe treatment regimens, to be treatable with appropriate gene therapy strategies. The great advance compared to conventional pharmaceutical approaches is that gene therapy directly corrects the underlying genomic abnormality of the disease. This provides the possibility to cure these diseases rather than treat symptoms. Given the vast diversity in gene-editing tools, we emphasize the requirement of selecting proper therapeutic strategies that match the underlying genetics of the disease. These strategies have distinct advantages, limitations, and potential adverse effects and for many indications the choices might not be straight-forward. Therefore, when pursuing a curative treatment for DOCK8 deficiency it is imperative that several modalities are explored at the developmental stage to maximize the therapeutic effects while considering disease mechanism, mutation locations, and strategies for delivery and correction. We suggest that efficient and patient-universal correction may be achieved by exploiting the HDR pathway. Additionally, *ex vivo* introduction of repair templates using AAV donor templates currently seems to be the most proper delivery mechanism and may be designed to replace single or multiple exons with hot-spot mutations or perhaps even include a full cDNA through a two-step HDR strategy. However, further experimental work is warranted to evaluate the efficiency of the different strategies in pursuit of a definitive cure for DOCK8 deficiency. Meanwhile, intensive research in genome editing leads to a continuous emergence of technologies that are bound to enable future breakthroughs in clinical gene therapy. Already, CRISPR/Cas9-based clinical trials have proven successful for sickle cell disease, beta-thalassemia, and transthyretin amyloidosis (Frangoul et al., 2021; Gillmore et al., 2021), and many efforts are focused on optimizing the conditioning regimen, such as the development of anti-CD117 antibodies that deplete HSCs in a targeted and safe manner (Kwon et al., 2019). With these advances and accumulation of experience from clinical trials, the future looks bright for bringing more CRISPR/Cas9-based gene therapies to patients that are safer and more efficient.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SR took lead in writing the manuscript with input from all co-authors. SK was lead in generating figures with input from all authors. ROB supervised the writing process and edited the final manscript for submission.

ACKNOWLEDGMENTS

ROB gratefully acknowledges the support from a Lundbeck Foundation Fellowship (R238-2016-534 3349), the Independent Research Fund Denmark (0134-00113B, 0242-00009B, and 9144-00001B), an AIAS-COFUND (Marie Curie) fellowship from Aarhus Institute of Advanced Studies (AIAS) co-funded by Aarhus University's Research Foundation and the European Union's seventh Framework Program under grant agreement no 609033, the Novo Nordisk Foundation (NNF19OC0058238 and NNF17OC0028894), Innovation Fund Denmark (8056-00010B), the Carlsberg Foundation (CF20-0424 and CF17-0129), Slagtermester Max Wørzner og Hustru Inger Wørzners Mindelegat, the AP Møller Foundation, the Riisfort Foundation, and a Genome Engineer Innovation Grant from Synthego. All figures were created with BioRender.com.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2022.793010/full#supplementary-material

REFERENCES

Adair, J. E., Waters, T., Haworth, K. G., Kubek, S. P., Trobridge, G. D., Hocum, J. D., et al. (2016). Semi-Automated Closed System Manufacturing of Lentivirus Gene-Modified Haematopoietic Stem Cells for Gene Therapy. *Nat. Commun.* 7, 1–10. doi:10.1038/ncomms13173

Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M. P., Baricordi, C., et al. (2013). Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science* 341 (6148), 1233151. doi:10.1126/science.1233151

Aiuti, A., Cattaneo, F., Galimberti, S., Benninghoff, U., Cassani, B., Callegaro, L., et al. (2009). Gene Therapy for Immunodeficiency Due to Adenosine Deaminase Deficiency. *N. Engl. J. Med.* 360 (5), 447–458. doi:10.1056/NEJMoa0805817

Aiuti, A., Slavin, S., Aker, M., Ficara, F., Deola, S., Mortellaro, A., et al. (2002). Correction of ADA-SCID by Stem Cell Gene Therapy Combined with Nonmyeloablative Conditioning. *Science* 296 (5577), 2410–2413. doi:10.1126/science.1070104

Al-Herz, W., Chu, J. I., van der Spek, J., Raghupathy, R., Massaad, M. J., Keles, S., et al. (2016). Hematopoietic Stem Cell Transplantation Outcomes for 11 Patients with Dedicator of Cytokinesis 8 Deficiency. *J. Allergy Clin. Immunol.* 138 (3), 852–859. doi:10.1016/j.jaci.2016.02.022

Al-Zahrani, D., Raddadi, A., Massaad, M., Keles, S., Jabara, H. H., Chatila, T. A., et al. (2014). Successful Interferon-Alpha 2b Therapy for Unremitting Warts in a Patient with DOCK8 Deficiency. *Clin. Immunol.* 153 (1), 104–108. doi:10.1016/j.clim.2014.04.005

Alsum, Z., Hawwari, A., Alsmadi, O., Al-Hissi, S., Borrero, E., Abu-staiteh, A., et al. (2013). Clinical, Immunological and Molecular Characterization of DOCK8 and DOCK8-like Deficient Patients: Single center Experience of Twenty Five Patients. *J. Clin. Immunol.* 33 (1), 55–67. doi:10.1007/s10875-012-9769-x

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace Genome Editing without Double-Strand Breaks or Donor DNA. *Nature* 576 (7785), 149–157. doi:10.1038/s41586-019-1711-4

Araki, M., and Ishii, T. (2014). International Regulatory Landscape and Integration of Corrective Genome Editing into *In Vitro* Fertilization. *Reprod. Biol. Endocrinol.* 12 (1), 108–112. doi:10.1186/1477-7827-12-108

Aydin, S. E., Freeman, A. F., Al-Herz, W., Al-Mousa, H. A., Arnaout, R. K., Aydin, R. C., et al. (2019). Hematopoietic Stem Cell Transplantation as Treatment for Patients with DOCK8 Deficiency. *J. Allergy Clin. Immunol. Pract.* 7 (3), 848–855. doi:10.1016/j.jaip.2018.10.035

Aydin, S. E., Kilic, S. S., Kilic, S. S., Aytekin, C., Kumar, A., Porras, O., et al. (2015). DOCK8 Deficiency: Clinical and Immunological Phenotype and Treatment Options - a Review of 136 Patients. *J. Clin. Immunol.* 35 (2), 189–198. doi:10.1007/s10875-014-0126-0

Bak, R. O., Dever, D. P., and Porteus, M. H. (2018). CRISPR/Cas9 Genome Editing in Human Hematopoietic Stem Cells. *Nat. Protoc.* 13 (2), 358–376. doi:10.1038/nprot.2017.143

Bak, R. O., Dever, D. P., Reinisch, A., Cruz Hernandez, D., Majeti, R., and Porteus, M. H. (2017). Multiplexed Genetic Engineering of Human Hematopoietic Stem and Progenitor Cells Using CRISPR/Cas9 and AAV6. *Elife* 6, 1–19. doi:10.7554/eLife.27873

Bak, R. O., Gomez-Ospina, N., and Porteus, M. H. (2018). Gene Editing on Center Stage. *Trends Genet.* 34 (8), 600–611. doi:10.1016/j.tig.2018.05.004

Bak, R. O., and Porteus, M. H. (2017). CRISPR-mediated Integration of Large Gene Cassettes Using AAV Donor Vectors. *Cel Rep.* 20 (3), 750–756. doi:10.1016/j.celrep.2017.06.064

Balakrishnan, B., and Jayandharan, G. (2014). Basic Biology of Adeno-Associated Virus (AAV) Vectors Used in Gene Therapy. *Cgt* 14 (2), 86–100. doi:10.2174/1566523214666140302193709

Bazinet, A., and Popradi, G. (2019). A General Practitioner's Guide to Hematopoietic Stem-Cell Transplantation. *Curr. Oncol.* 26 (3), 187–191. doi:10.3747/co.26.5033

Begum, R., Nur-E-Kamal, M. S. A., and Zaman, M. (2004). The Role of Rho GTPases in the Regulation of the Rearrangement of Actin Cytoskeleton and Cell Movement. *Exp. Mol. Med.* 36 (4), 358–366. doi:10.1038/emm.2004.47

Biggs, C. M., Keles, S., and Chatila, T. A. (2017). DOCK8 Deficiency: Insights into Pathophysiology, Clinical Features and Management. *Clin. Immunol.* 181 (181), 75–82. doi:10.1016/j.clim.2017.06.003

Biram, A., Davidzohn, N., and Shulman, Z. (2019). T Cell Interactions with B Cells during Germinal center Formation, a Three-step Model. *Immunol. Rev.* 288 (1), 37–48. doi:10.1111/imr.12737

Bisaria, N., Jarmoskaite, I., and Herschlag, D. (2017). Lessons from Enzyme Kinetics Reveal Specificity Principles for RNA-Guided Nucleases in RNA Interference and CRISPR-Based Genome Editing. *Cel Syst.* 4 (1), 21–29. doi:10.1016/j.cels.2016.12.010

Bloomer, H., Smith, R. H., Hakami, W., and Larochelle, A. (2021). Genome Editing in Human Hematopoietic Stem and Progenitor Cells via CRISPR-Cas9-Mediated Homology-independent Targeted Integration. *Mol. Ther.* 29 (4), 1611–1624. doi:10.1016/j.ymthe.2020.12.010

Bokoch, G. M. (2003). Biology of the P21-Activated Kinases. *Annu. Rev. Biochem.* 72, 743–781. doi:10.1146/annurev.biochem.72.121801.161742

Booth, C., Gaspar, H. B., and Thrasher, A. J. (2016). Treating Immunodeficiency through HSC Gene Therapy. *Trends Mol. Med.* 22 (4), 317–327. doi:10.1016/j.molmed.2016.02.002

Booth, C., Romano, R., Roncarolo, M. G., and Thrasher, A. J. (2019). Gene Therapy for Primary Immunodeficiency. *Hum. Mol. Genet.* 28 (R1), R15–R23. doi:10.1093/hmg/ddz170

Bousfiha, A., Jeddane, L., Picard, C., Al-Herz, W., Ailal, F., Chatila, T., et al. (2020). Human Inborn Errors of Immunity: 2019 Update of the IUIS Phenotypical Classification. *J. Clin. Immunol.* 40 (1), 66–81. doi:10.1007/s10875-020-00758-x

Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P. P., Díez, I. A., Dewey, R. A., et al. (2010). Stem-Cell Gene Therapy for the Wiskott-Aldrich Syndrome. *N. Engl. J. Med.* 363 (20), 1918–1927. doi:10.1056/NEJMoa1003548

Broder, M. S., Quock, T. P., Chang, E., Reddy, S. R., Agarwal-Hashmi, R., Arai, S., et al. (2017). The Cost of Hematopoietic Stem-Cell Transplantation in the United States. *Am. Health Drug Benefits* 10 (7), 366–374.

Broides, A., Mandola, A. B., Levy, J., Yerushalmi, B., Pinsk, V., Eldan, M., et al. (2017). The Clinical and Laboratory Spectrum of Dedicator of Cytokinesis 8 Immunodeficiency Syndrome in Patients with a Unique Mutation. *Immunol. Res.* 65 (3), 651–657. doi:10.1007/s12026-016-8883-x

Brown, M. P., Topham, D. J., Sangster, M. Y., Zhao, J., Flynn, K. J., Surman, S. L., et al. (1998). Thymic Lymphoproliferative Disease after Successful Correction of CD40 Ligand Deficiency by Gene Transfer in Mice. *Nat. Med.* 4 (11), 1253–1260. doi:10.1038/3233

Burbage, M., Keppler, S. J., Gasparrini, F., Martínez-Martín, N., Gaya, M., Feest, C., et al. (2015). Cdc42 Is a Key Regulator of B Cell Differentiation and Is Required for Antiviral Humoral Immunity. *J. Exp. Med.* 212 (1), 53–72. doi:10.1084/jem.20141143

Bushman, F. D. (2007). Retroviral Integration and Human Gene Therapy. *J. Clin. Invest.* 117 (8), 2083–2086. doi:10.1172/JCI32949

Caracciolo, S., Moratto, D., Giacomelli, M., Negri, S., Lougaris, V., Porta, F., et al. (2014). Expansion of CCR4+ Activated T Cells Is Associated with Memory B Cell Reduction in DOCK8-Deficient Patients. *Clin. Immunol.* 152 (1-2), 164–170. doi:10.1016/j.clim.2014.03.008

Carusillo, A., and Mussolino, C. (2020). DNA Damage: From Threat to Treatment. *Cells* 9 (7), 1665. doi:10.3390/cells9071665

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., et al. (2010). Transfusion independence and HMGA2 Activation after Gene Therapy of Human β-thalassaemia. *Nature* 467 (7313), 318–322. doi:10.1038/nature09328

Chan, A. Y. L., Chan, V. K. Y., Olsson, S., Fan, M., Jit, M., Gong, M., et al. (2020). Access and Unmet Needs of Orphan Drugs in 194 Countries and 6 Areas: A Global Policy Review with Content Analysis. *Value in Health* 23 (12), 1580–1591. doi:10.1016/j.jval.2020.06.020

Charpentier, M., Khedher, A. H. Y., Menoret, S., Brion, A., Lamribet, K., Dardillac, E., et al. (2018). CtIP Fusion to Cas9 Enhances Transgene Integration by Homology-dependent Repair. *Nat. Commun.* 9 (1), 1133. doi:10.1038/s41467-018-03475-7

Chemin, K., Bohineust, A., Dogniaux, S., Tourret, M., Guégan, S., Miro, F., et al. (2012). Cytokine Secretion by CD4+ T Cells at the Immunological Synapse Requires Cdc42-dependent Local Actin Remodeling but Not Microtubule Organizing center Polarity. *J.I.* 189 (5), 2159–2168. doi:10.4049/jimmunol.1200156

Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., et al. (2010). Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics* 186 (2), 757–761. doi:10.1534/genetics.110.120717

Chu, E. Y., Freeman, A. F., and Jing, H. (2012). Cutaneous Manifestations of DOCK8 Deficiency Syndrome. *Arch. Dermatol.* 148 (1), 79–84. doi:10.1001/archdermatol.2011.262

Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., et al. (2015). Increasing the Efficiency of Homology-Directed Repair for CRISPR-Cas9-Induced Precise Gene Editing in Mammalian Cells. *Nat. Biotechnol.* 33 (5), 543–548. doi:10.1038/nbt.3198

Cicalese, M. P., Ferrua, F., Castagnaro, L., Pajno, R., Barzaghi, F., Giannelli, S., et al. (2016). Update on the Safety and Efficacy of Retroviral Gene Therapy for Immunodeficiency Due to Adenosine Deaminase Deficiency. *Blood* 128 (1), 45–54. doi:10.1182/blood-2016-01-688226

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339 (6121), 819–823. doi:10.1126/science.1231143

Copelan, E. A. (2006). Hematopoietic Stem-Cell Transplantation. *N. Engl. J. Med.* 354, 1813–1826. doi:10.1056/nejmra052638

Côté, J.-F., Motoyama, A. B., Bush, J. A., and Vuori, K. (2005). A Novel and Evolutionarily Conserved PtdIns(3,4,5)P3-Binding Domain Is Necessary for DOCK180 Signalling. *Nat. Cel Biol.* 7 (8), 797–807. doi:10.1038/ncb1280

Côté, J.-F., and Vuori, K. (2007). GEF what? Dock180 and Related Proteins Help Rac to Polarize Cells in New Ways. *Trends Cel Biol.* 17 (8), 383–393. doi:10.1016/j.tcb.2007.05.001

Coˆte´, J.-F., and Vuori, K. (2002). Identification of an Evolutionarily Conserved Superfamily of DOCK180-Related Proteins with Guanine Nucleotide Exchange Activity. *J. Cel Sci.* 115 (24), 4901–4913. doi:10.1242/jcs.00219

Crawford, G., Enders, A., Gileadi, U., Stankovic, S., Zhang, Q., Lambe, T., et al. (2013). DOCK8 Is Critical for the Survival and Function of NKT Cells. *Blood* 122 (12), 2052–2061. doi:10.1182/blood-2013-02-482331

Cromer, M. K., Camarena, J., Martin, R. M., Lesch, B. J., Vakulskas, C. A., Bode, N. M., et al. (2021). Gene Replacement of α-globin with β-globin Restores Hemoglobin Balance in β-thalassemia-derived Hematopoietic Stem and Progenitor Cells. *Nat. Med.* 27, 677–687. doi:10.1038/s41591-021-01284-y

Cromer, M. K., Vaidyanathan, S., Ryan, D. E., Curry, B., Lucas, A. B., Camarena, J., et al. (2018). Global Transcriptional Response to CRISPR/Cas9-AAV6-Based Genome Editing in CD34+ Hematopoietic Stem and Progenitor Cells. *Mol. Ther.* 26 (10), 2431–2442. doi:10.1016/j.ymthe.2018.06.002

Cyranoski, D. (2016). CRISPR Gene-Editing Tested in a Person for the First Time. *Nature* 539, 479. doi:10.1038/nature.2016.20988

Database resources of the National Center for Biotechnology Information (2018). Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 46 (D1), D8–D13. doi:10.1093/nar/gkx1095

De Ravin, S. S., Brault, J., Meis, R. J., Liu, S., Li, L., Pavel-Dinu, M., et al. (2021). Enhanced Homology-Directed Repair for Highly Efficient Gene Editing in Hematopoietic Stem/Progenitor Cells. *Blood* 137, 2598–2608. doi:10.1182/blood.2020008503

De Ravin, S. S., Li, L., Wu, X., Choi, U., Allen, C., Koontz, S., et al. (2017). CRISPR-Cas9 Gene Repair of Hematopoietic Stem Cells from Patients with X-Linked Chronic Granulomatous Disease. *Sci. Transl. Med.* 9, 9. doi:10.1126/scitranslmed.aah3480

De Ravin, S. S., Reik, A., Liu, P.-Q., Li, L., Wu, X., Su, L., et al. (2016). Targeted Gene Addition in Human CD34+ Hematopoietic Cells for Correction of X-Linked Chronic Granulomatous Disease. *Nat. Biotechnol.* 34 (4), 424–429. doi:10.1038/nbt.3513

Dean, R., Jensen, I., Cyr, P., Miller, B., Maru, B., Sproule, D. M., et al. (2021). An Updated Cost-Utility Model for Onasemnogene Abeparvovec (Zolgensma) in Spinal Muscular Atrophy Type 1 Patients and Comparison with Evaluation by the Institute for Clinical and Effectiveness Review (ICER). *J. Market Access Health Pol.* 9 (1), 1889841. doi:10.1080/20016689.2021.1889841

Dever, D. P., Bak, R. O., Reinisch, A., Camarena, J., Washington, G., Nicolas, C. E., et al. (2016). CRISPR/Cas9 β-globin Gene Targeting in Human Haematopoietic Stem Cells. *Nature* 539 (7629), 384–389. doi:10.1038/nature20134

Devetten, M., and Armitage, J. O. (2007). Hematopoietic Cell Transplantation: Progress and Obstacles. *Ann. Oncol.* 18 (9), 1450–1456. doi:10.1093/annonc/mdm064

DeWitt, M. A., Magis, W., Bray, N. L., Wang, T., Berman, J. R., Urbinati, F., et al. (2016). Selection-free Genome Editing of the Sickle Mutation in Human Adult Hematopoietic Stem/progenitor Cells. *Sci. Transl. Med.* 8 (360), 360ra134. doi:10.1126/scitranslmed.aaf9336

Du, Y., Fang, Q., and Zheng, S.-G. (2021). "Regulatory T Cells: Concept, Classification, Phenotype, and Biological Characteristics," in *T Regulatory Cells in Human Health and Diseases*. Editor S-G. Zheng (Singapore: Springer), 1–31. doi:10.1007/978-981-15-6407-9_1

Dustin, M. L. (2002). Membrane Domains and the Immunological Synapse: Keeping T Cells Resting and Ready. *J. Clin. Invest.* 109 (2), 155–160. doi:10.1172/JCI0214842

Engelhardt, K. R., Gertz, M. E., Keles, S., Schäffer, A. A., Sigmund, E. C., Glocker, C., et al. (2015). The Extended Clinical Phenotype of 64 Patients with Dedicator of Cytokinesis 8 Deficiency. *J. Allergy Clin. Immunol.* 136 (2), 402–412. doi:10.1016/j.jaci.2014.12.1945

Engelhardt, K. R., McGhee, S., Winkler, S., Sassi, A., Woellner, C., Lopez-Herrera, G., et al. (2009). Large Deletions and point Mutations Involving the Dedicator of Cytokinesis 8 (DOCK8) in the Autosomal-Recessive Form of Hyper-IgE Syndrome. *J. Allergy Clin. Immunol.* 124, 1289–1302. doi:10.1016/j.jaci.2009.10.038

European Commission (2021). Rare Diseases. Availableat: https://ec.europa.eu/info/research-and-innovation/research-area/health-research-and-innovation/rare-diseases_en (Accessed February 11, 2021).

Ferrari, S., Jacob, A., Beretta, S., Unali, G., Albano, L., Vavassori, V., et al. (2020). Efficient Gene Editing of Human Long-Term Hematopoietic Stem Cells Validated by Clonal Tracking. *Nat. Biotechnol.* 38 (11), 1298–1308. doi:10.1038/s41587-020-0551-y

Fox, T. A., Chakraverty, R., Burns, S., Carpenter, B., Thomson, K., Lowe, D., et al. (2018). Successful Outcome Following Allogeneic Hematopoietic Stem Cell Transplantation in Adults with Primary Immunodeficiency. *Blood* 131 (8), 917–931. doi:10.1182/blood-2017-09-807487

Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., et al. (2021). CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β-Thalassemia. *N. Engl. J. Med.* 384 (3), 252–260. doi:10.1056/NEJMoa2031054

Fu, Y.-W., Dai, X.-Y., Wang, W.-T., Yang, Z.-X., Zhao, J.-J., Zhang, J.-P., et al. (2021). Dynamics and Competition of CRISPR-Cas9 Ribonucleoproteins and AAV Donor-Mediated NHEJ, MMEJ and HDR Editing. *Nucleic Acids Res.* 49 (2), 969–985. doi:10.1093/nar/gkaa1251

Gaspar, H. B., Cooray, S., Gilmour, K. C., Parsley, K. L., Adams, S., Howe, S. J., et al. (2011). Long-term Persistence of a Polyclonal T Cell Repertoire after Gene Therapy for X-Linked Severe Combined Immunodeficiency. *Sci. Transl. Med.* 3 (97), 97ra79. doi:10.1126/scitranslmed.3002715

Gaspar, H. B., Parsley, K. L., Howe, S., King, D., Gilmour, K. C., Sinclair, J., et al. (2004). Gene Therapy of X-Linked Severe Combined Immunodeficiency by Use of a Pseudotyped Gammaretroviral Vector. *The Lancet* 364 (9452), 2181–2187. doi:10.1016/S0140-6736(04)17590-9

Gates, A. J., Gysi, D. M., Kellis, M., and Barabási, A.-L. (2021). A Wealth of Discovery Built on the Human Genome Project - by the Numbers. *Nature* 590 (7845), 212–215. doi:10.1038/d41586-021-00314-6

Gavrilova, T. (2019). Considerations for Hematopoietic Stem Cell Transplantation in Primary Immunodeficiency Disorders. *Wjt* 9 (3), 48–57. doi:10.5500/wjt.v9.i3.48

Gene (2021). Gene Therapies Should Be for All. *Nat. Med.* 27 (8), 1311. doi:10.1038/s41591-021-01481-9

Genovese, P., Schiroli, G., Escobar, G., Di Tomaso, T., Firrito, C., Calabria, A., et al. (2014). Targeted Genome Editing in Human Repopulating Haematopoietic Stem Cells. *Nature* 510 (7504), 235–240. doi:10.1038/nature13420

Gernez, Y., Baker, M. G., and Maglione, P. J. (2018). Humoral Immunodeficiencies: Conferred Risk of Infections and Benefits of Immunoglobulin Replacement Therapy. *Transfusion* 58 (December), 3056–3064. doi:10.1111/trf.15020

Gillmore, J. D., Gane, E., Taubel, J., Kao, J., Fontana, M., Maitland, M. L., et al. (2021). CRISPR-Cas9 *In Vivo* Gene Editing for Transthyretin Amyloidosis. *N. Engl. J. Med.* 385, 493–502. doi:10.1056/nejmoa2107454

Glessner, J. T., Li, J., Li, J., Wang, D., March, M., Lima, L., et al. (2017). Copy Number Variation Meta-Analysis Reveals a Novel Duplication at 9p24 Associated with Multiple Neurodevelopmental Disorders. *Genome Med.* 9 (1), 1–11. doi:10.1186/s13073-017-0494-1

Gray, D. H., Villegas, I., Long, J., Santos, J., Keir, A., Abele, A., et al. (2021). Optimizing Integration and Expression of Transgenic Bruton's Tyrosine Kinase for CRISPR-Cas9-Mediated Gene Editing of X-Linked Agammaglobulinemia. *CRISPR J.* 4 (2), 191–206. doi:10.1089/crispr.2020.0080

Grieger, J. C., and Samulski, R. J. (2005). Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *J. Virol.* 79 (15), 9933–9944. doi:10.1128/jvi.79.15.9933-9944.2005

Grimbacher, B., Holland, S. M., Gallin, J. I., Greenberg, F., Hill, S. C., Malech, H. L., et al. (1999). Hyper-IgE Syndrome with Recurrent Infections - an Autosomal Dominant Multisystem Disorder. *N. Engl. J. Med.* 340 (9), 692–702. doi:10.1056/NEJM199903043400904

Gundry, M. C., Dever, D. P., Yudovich, D., Bauer, D. E., Haas, S., Wilkinson, A. C., et al. (2017). Technical Considerations for the Use of CRISPR/Cas9 in Hematology Research. *Exp. Hematol.* 54, 4–11. doi:10.1016/j.exphem.2017.07.006

Hacein-Bey Abina, S., Gaspar, H. B., Blondeau, J., Caccavelli, L., Charrier, S., Buckland, K., et al. (2015). Outcomes Following Gene Therapy in Patients with Severe Wiskott-Aldrich Syndrome. *JAMA* 313 (15), 1550–1563. doi:10.1001/jama.2015.3253

Hacein-Bey-Abina, S., Hauer, J., Lim, A., Picard, C., Wang, G. P., Berry, C. C., et al. (2010). Efficacy of Gene Therapy for X-Linked Severe Combined Immunodeficiency. *N. Engl. J. Med.* 363 (4), 355–364. doi:10.1056/NEJMoa1000164

Hacein-Bey-Abina, S., Le Deist, F., Carlier, F., Bouneaud, C., Hue, C., De Villartay, J.-P., et al. (2002). Sustained Correction of X-Linked Severe Combined Immunodeficiency by *Ex Vivo* Gene Therapy. *N. Engl. J. Med.* 346 (16), 1185–1193. doi:10.1056/NEJMoa012616

Harada, Y., Tanaka, Y., Terasawa, M., Pieczyk, M., Habiro, K., Katakai, T., et al. (2012). DOCK8 Is a Cdc42 Activator Critical for Interstitial Dendritic Cell Migration during Immune Responses. *Blood* 119 (19), 4451–4461. doi:10.1182/blood-2012-01-407098

Haskologlu, S., Kostel Bal, S., Islamoglu, C., Aytekin, C., Guner, S., Sevinc, S., et al. (2020). Clinical, Immunological Features and Follow up of 20 Patients with Dedicator of Cytokinesis 8 (DOCK8) Deficiency. *Pediatr. Allergy Immunol.* 31 (5), 515–527. doi:10.1111/pai.13236

Hendel, A., Bak, R. O., Clark, J. T., Kennedy, A. B., Ryan, D. E., Roy, S., et al. (2015). Chemically Modified Guide RNAs Enhance CRISPR-Cas Genome Editing in Human Primary Cells. *Nat. Biotechnol.* 33 (9), 985–989. doi:10.1038/nbt.3290

Hubbard, N., Hagin, D., Sommer, K., Song, Y., Khan, I., Clough, C., et al. (2016). Targeted Gene Editing Restores Regulated CD40L Function in X-Linked Hyper-IgM Syndrome. *Blood* 127 (21), 2513–2522. doi:10.1182/blood-2015-11-683235

Jabara, H. H., McDonald, D. R., Janssen, E., Massaad, M. J., Ramesh, N., Borzutzky, A., et al. (2012). DOCK8 Functions as an Adaptor that Links TLR-MyD88 Signaling to B Cell Activation. *Nat. Immunol.* 13 (6), 612–620. doi:10.1038/ni.2305

Janssen, E., Kumari, S., Tohme, M., Ullas, S., Barrera, V., Tas, J. M. J., et al. (2017). DOCK8 Enforces Immunological Tolerance by Promoting IL-2 Signaling and Immune Synapse Formation in Tregs. *JCI Insight* 2 (19), 1–18. doi:10.1172/jci.insight.94298

Janssen, E., Morbach, H., Ullas, S., Bannock, J. M., Massad, C., Menard, L., et al. (2014). Dedicator of Cytokinesis 8-deficient Patients Have a Breakdown in Peripheral B-Cell Tolerance and Defective Regulatory T Cells. *J. Allergy Clin. Immunol.* 134 (6), 1365–1374. doi:10.1016/j.jaci.2014.07.042

Janssen, E., Tohme, M., Butts, J., Giguere, S., Sage, P. T., Velázquez, F. E., et al. (2020). DOCK8 Is Essential for LFA-1-dependent Positioning of T Follicular Helper Cells in Germinal Centers. *JCI Insight* 5, 5. doi:10.1172/jci.insight.134508

Janssen, E., Tohme, M., Hedayat, M., Leick, M., Kumari, S., Ramesh, N., et al. (2016). A DOCK8-WIP-WASp Complex Links T Cell Receptors to the Actin Cytoskeleton. *J. Clin. Invest.* 126 (10), 3837–3851. doi:10.1172/JCI85774

Janssen, E., Wilkie, H., and Geha, R. S. (2021). Macabre TH2 Skewing in DOCK8 Deficiency. *J. Allergy Clin. Immunol.* 148 (1), 73–75. doi:10.1016/j.jaci.2021.02.025

Jayavaradhan, R., Pillis, D. M., Goodman, M., Zhang, F., Zhang, Y., Andreassen, P. R., et al. (2019). CRISPR-Cas9 Fusion to Dominant-Negative 53BP1 Enhances HDR and Inhibits NHEJ Specifically at Cas9 Target Sites. *Nat. Commun.* 10 (1), 2866. doi:10.1038/s41467-019-10735-7

Jensen, T. I., Axelgaard, E., and Bak, R. O. (2019). Therapeutic Gene Editing in Haematological Disorders withCRISPR/Cas9. *Br. J. Haematol.* 185 (5), 821–835. doi:10.1111/bjh.15851

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337 (6096), 816–821. doi:10.1126/science.1225829

Jing, H., Zhang, Q., Zhang, Y., Hill, B. J., Dove, C. G., Gelfand, E. W., et al. (2014). Somatic Reversion in Dedicator of Cytokinesis 8 Immunodeficiency Modulates Disease Phenotype. *J. Allergy Clin. Immunol.* 133 (6), 1667–1675. doi:10.1016/j.jaci.2014.03.025

Kacetl, J., Marešová, P., Maskuriy, R., and Selamat, A. (2020). Ethical Questions Linked to Rare Diseases and Orphan Drugs - A Systematic Review. *Rmhp* 13, 2125–2148. doi:10.2147/RMHP.S260641

Kang, E. M., Choi, U., Theobald, N., Linton, G., Long Priel, D. A., Kuhns, D., et al. (2010). Retrovirus Gene Therapy for X-Linked Chronic Granulomatous Disease Can Achieve Stable Long-Term Correction of Oxidase Activity in Peripheral Blood Neutrophils. *Blood* 115 (4), 783–791. doi:10.1182/blood-2009-05-222760

Keles, S., Charbonnier, L. M., Kabaleeswaran, V., Reisli, I., Genel, F., Gulez, N., et al. (2016). Dedicator of Cytokinesis 8 Regulates Signal Transducer and Activator

of Transcription 3 Activation and Promotes TH17 Cell Differentiation. *J. Allergy Clin. Immunol.* 138 (5), 1384–1394. doi:10.1016/j.jaci.2016.04.023. DOCK8

Keles, S., Jabara, H. H., Reisli, I., McDonald, D. R., Barlan, I., Hanna-Wakim, R., et al. (2014). Plasmacytoid Dendritic Cell Depletion in DOCK8 Deficiency: Rescue of Severe Herpetic Infections with IFN-α 2b Therapy. *J. Allergy Clin. Immunol.* 133 (6), 1753–1755. doi:10.1016/j.jaci.2014.03.032

Kim, D., and Kim, J.-S. (2018). DIG-seq: A Genome-wide CRISPR Off-Target Profiling Method Using Chromatin DNA. *Genome Res.* 28 (12), 1894–1900. doi:10.1101/gr.236620.118

Kim, S., Kim, D., Cho, S. W., Kim, J., and Kim, J.-S. (2014). Highly Efficient RNA-Guided Genome Editing in Human Cells via Delivery of Purified Cas9 Ribonucleoproteins. *Genome Res.* 24 (6), 1012–1019. doi:10.1101/gr.171322.113

Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996). Hybrid Restriction Enzymes: Zinc finger Fusions to Fok I Cleavage Domain. *Proc. Natl. Acad. Sci.* 93 (3), 1156–1160. doi:10.1073/pnas.93.3.1156

Krishnaswamy, J. K., Gowthaman, U., Zhang, B., Mattsson, J., Szeponik, L., Liu, D., et al. (2017). Migratory CD11b + Conventional Dendritic Cells Induce T Follicular Helper Cell-dependent Antibody Responses. *Sci. Immunol.* 2 (18), 1–14. doi:10.1126/sciimmunol.aam9169

Krishnaswamy, J. K., Singh, A., Gowthaman, U., Wu, R., Gorrepati, P., Sales Nascimento, M., et al. (2015). Coincidental Loss of DOCK8 Function in NLRP10-Deficient and C3H/HeJ Mice Results in Defective Dendritic Cell Migration. *Proc. Natl. Acad. Sci. USA* 112 (10), 3056–3061. doi:10.1073/pnas.1501554112

Kumari, S., Curado, S., Mayya, V., and Dustin, M. L. (2014). T Cell Antigen Receptor Activation and Actin Cytoskeleton Remodeling. *Biochim. Biophys. Acta (Bba) - Biomembranes* 1838 (2), 546–556. doi:10.1016/j.bbamem.2013.05.004

Kunimura, K., Uruno, T., and Fukui, Y. (2020). DOCK Family Proteins: Key Players in Immune Surveillance Mechanisms. *Int. Immunol.* 32 (1), 5–15. doi:10.1093/intimm/dxz067

Kuo, C. Y., Long, J. D., Campo-Fernandez, B., de Oliveira, S., Cooper, A. R., Romero, Z., et al. (2018). Site-Specific Gene Editing of Human Hematopoietic Stem Cells for X-Linked Hyper-IgM Syndrome. *Cel Rep.* 23 (9), 2606–2616. doi:10.1016/j.celrep.2018.04.103

Kwon, H.-S., Logan, A. C., Chhabra, A., Pang, W. W., Czechowicz, A., Tate, K., et al. (2019). Anti-human CD117 Antibody-Mediated Bone Marrow Niche Clearance in Nonhuman Primates and Humanized NSG Mice. *Blood* 133 (19), 2104–2108. doi:10.1182/blood-2018-06-853879

Lambe, T., Crawford, G., Johnson, A. L., Crockford, T. L., Bouriez-Jones, T., Smyth, A. M., et al. (2011). DOCK8 Is Essential for T-Cell Survival and the Maintenance of CD8+ T-Cell Memory. *Eur. J. Immunol.* 41 (12), 3423–3435. doi:10.1002/eji.201141759

Lander, E. S., Baylis, F., Zhang, F., Charpentier, E., Berg, P., Bourgain, C., et al. (2019). Adopt a Moratorium on Heritable Genome Editing. *Nature* 567 (7747), 165–168. doi:10.1038/d41586-019-00726-5

Lattanzi, A., Camarena, J., Lahiri, P., Segal, H., Srifa, W., Vakulskas, C. A., et al. (2021). Development of β-globin Gene Correction in Human Hematopoietic Stem Cells as a Potential Durable Treatment for Sickle Cell Disease. *Sci. Transl. Med.* 13, 13. doi:10.1126/scitranslmed.abf2444

Laustsen, A., and Bak, R. O. (2019). "Electroporation-Based CRISPR/Cas9 Gene Editing Using Cas9 Protein and Chemically Modified sgRNAs: Methods and Protocols," in *Methods in Molecular Biology (Clifton, N.J.)*. Editor Y. Luo (New York, NY: United States: Springer New York), 1961, 127–134. doi:10.1007/978-1-4939-9170-9_9

Li, R., and Gundersen, G. G. (2008). Beyond Polymer Polarity: How the Cytoskeleton Builds a Polarized Cell. *Nat. Rev. Mol. Cel Biol.* 9 (11), 860–873. doi:10.1038/nrm2522

Lin, S., Staahl, B. T., Alla, R. K., and Doudna, J. A. (2014). "Enhanced Homology-Directed Human Genome Engineering by Controlled Timing of CRISPR/Cas9 Delivery,"*Elife*. Editor D. Weigel, 3, e04766. doi:10.7554/eLife.04766

Lino, C. A., Harper, J. C., Carney, J. P., and Timlin, J. A. (2018). Delivering Crispr: A Review of the Challenges and Approaches. *Drug Deliv.* 25 (1), 1234–1257. doi:10.1080/10717544.2018.1474964

Magis, W., DeWitt, M. A., Wyman, S. K., Vu, J. T., Heo, S.-J., Shao, S. J., et al. (2019). High-level Correction of the Sickle Mutation Amplified *In Vivo* during Erythroid Differentiation. *bioRxiv*, 432716. doi:10.1101/432716

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., et al. (2013). CAS9 Transcriptional Activators for Target Specificity Screening and Paired Nickases for Cooperative Genome Engineering. *Nat. Biotechnol.* 31 (9), 833–838. doi:10.1038/nbt.2675

Maresca, M., Lin, V. G., Guo, N., and Yang, Y. (2013). Obligate Ligation-Gated Recombination (ObLiGaRe): Custom-Designed Nuclease-Mediated Targeted Integration through Nonhomologous End Joining. *Genome Res.* 23 (3), 539–546. doi:10.1101/gr.145441.112

Mayerhoff, L., Lehne, M., Hickstein, L., Salimullah, T., Prieur, S., Thomas, S. K., et al. (2019). Cost Associated with Hematopoietic Stem Cell Transplantation: A Retrospective Claims Data Analysis in Germany. *J. Comp. Effectiveness Res.* 8 (2), 121–131. doi:10.2217/cer-2018-0100

McGhee, S. A., and Chatila, T. A. (2010). DOCK8 Immune Deficiency as a Model for Primary Cytoskeletal Dysfunction. *Dis. Markers* 29 (3-4), 151–156. doi:10.3233/DMA-2010-0740

Meshaal, S. S., El Hawary, R. E., Eldash, A., Grimbacher, B., Camacho-Ordonez, N., Abd Elaziz, D. S., et al. (2018). Diagnosis of DOCK8 Deficiency Using Flow Cytometry Biomarkers: an Egyptian Center Experience. *Clin. Immunol.* 195 (195), 36–44. doi:10.1016/j.clim.2018.07.011

Meyts, I., Bosch, B., Bolze, A., Boisson, B., Itan, Y., Belkadi, A., et al. (2016). Exome and Genome Sequencing for Inborn Errors of Immunity. *J. Allergy Clin. Immunol.* 138 (4), 957–969. doi:10.1016/j.jaci.2016.08.003

Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y.-L., Rupniewski, I., et al. (2007). An Improved Zinc-finger Nuclease Architecture for Highly Specific Genome Editing. *Nat. Biotechnol.* 25 (7), 778–785. doi:10.1038/nbt1319

Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., et al. (2011). A TALE Nuclease Architecture for Efficient Genome Editing. *Nat. Biotechnol.* 29 (2), 143–148. doi:10.1038/nbt.1755

Milner, J. D., Brenchley, J. M., Laurence, A., Freeman, A. F., Hill, B. J., Elias, K. M., et al. (2008). Impaired TH17 Cell Differentiation in Subjects with Autosomal Dominant Hyper-IgE Syndrome. *Nature* 452 (7188), 773–776. doi:10.1038/nature06764

Mizesko, M. C., Banerjee, P. P., Monaco-Shawver, L., Mace, E. M., Bernal, W. E., Sawalle-Belohradsky, J., et al. (2013). Defective Actin Accumulation Impairs Human Natural Killer Cell Function in Patients with Dedicator of Cytokinesis 8 Deficiency. *J. Allergy Clin. Immunol.* 131 (3), 840–848. doi:10.1016/j.jaci.2012.12.1568

Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M. H., et al. (2009). Insertional Transformation of Hematopoietic Cells by Self-Inactivating Lentiviral and Gammaretroviral Vectors. *Mol. Ther.* 17 (11), 1919–1928. doi:10.1038/mt.2009.179

Mohrin, M., Bourke, E., Alexander, D., Warr, M. R., Barry-Holson, K., Le Beau, M. M., et al. (2010). Hematopoietic Stem Cell Quiescence Promotes Error-Prone DNA Repair and Mutagenesis. *Cell Stem Cell.* 7 (2), 174–185. doi:10.1016/j.stem.2010.06.014

Morgan, R. A., Gray, D., Lomova, A., and Kohn, D. B. (2017). Hematopoietic Stem Cell Gene Therapy: Progress and Lessons Learned. *Cell Stem Cell.* 21 (5), 574–590. doi:10.1016/j.stem.2017.10.010

Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T., and Cathomen, T. (2011). A Novel TALE Nuclease Scaffold Enables High Genome Editing Activity in Combination with Low Toxicity. *Nucleic Acids Res.* 39 (21), 9283–9293. doi:10.1093/nar/gkr597

Nambiar, T. S., Billon, P., Diedenhofen, G., Hayward, S. B., Taglialatela, A., Cai, K., et al. (2019). Stimulation of CRISPR-Mediated Homology-Directed Repair by an Engineered RAD18 Variant. *Nat. Commun.* 10 (1), 3395. doi:10.1038/s41467-019-11105-z

Namekata, K., Kimura, A., Kawamura, K., Harada, C., and Harada, T. (2014). Dock GEFs and Their Therapeutic Potential: Neuroprotection and Axon Regeneration. *Prog. Retin. Eye Res.* 43, 1–16. doi:10.1016/j.preteyeres.2014.06.005

Newby, G. A., Yen, J. S., Woodard, K. J., Mayuranathan, T., Lazzarotto, C. R., Li, Y., et al. (2021). Base Editing of Haematopoietic Stem Cells Rescues Sickle Cell Disease in Mice. *Nature* 595 (7866), 295–302. doi:10.1038/s41586-021-03609-w

Ott, M. G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., et al. (2006). Correction of X-Linked Chronic Granulomatous Disease by Gene Therapy, Augmented by Insertional Activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat. Med.* 12 (4), 401–409. doi:10.1038/nm1393

Papan, C., Hagl, B., Heinz, V., Albert, M. H., Ehrt, O., Sawalle-Belohradsky, J., et al. (2014). Beneficial IFN-α Treatment of Tumorous Herpes Simplex Blepharoconjunctivitis in Dedicator of Cytokinesis 8 Deficiency. *J. Allergy Clin. Immunol.* 133 (5), 1456–1458. doi:10.1016/j.jaci.2014.02.008

Pavel-Dinu, M., Wiebking, V., Dejene, B. T., Srifa, W., Mantri, S., Nicolas, C. E., et al. (2019). Gene Correction for SCID-X1 in Long-Term Hematopoietic Stem Cells. *Nat. Commun.* 10 (1), 1634. doi:10.1038/s41467-019-09614-y

Pillay, B. A., Fusaro, M., Gray, P. E., Statham, A. L., Burnett, L., Bezrodnik, L., et al. (2021). Somatic Reversion of Pathogenic DOCK8 Variants Alters Lymphocyte Differentiation and Function to Effectively Cure DOCK8 Deficiency. *J. Clin. Invest.* 131, 131. doi:10.1172/JCI142434

Porteus, M. H. (2019). A New Class of Medicines through DNA Editing. *N. Engl. J. Med.* 380 (10), 947–959. doi:10.1056/nejmra1800729

Porto, E. M., Komor, A. C., Slaymaker, I. M., and Yeo, G. W. (2020). Base Editing: Advances and Therapeutic Opportunities. *Nat. Rev. Drug Discov.* 19 (12), 839–859. doi:10.1038/s41573-020-0084-6

Randall, K. L., Lambe, T., Johnson, A. L., Treanor, B., Kucharska, E., Domaschenz, H., et al. (2012). Europe PMC Funders Group DOCK8 Mutations Cripple B Cell Immune Synapse, Germinal Centers and Long-Lived Antibody Production *Nat. Immunol.* 10 (12), 1283–1291. doi:10.1038/ni.1820.DOCK8

Randall, K. L., Chan, S. S.-Y., Ma, C. S., Fung, I., Mei, Y., Yabas, M., et al. (2011). DOCK8 Deficiency Impairs CD8 T Cell Survival and Function in Humans and Mice. *J. Exp. Med.* 208 (11), 2305–2320. doi:10.1084/jem.20110345

Randall, K. L., Lambe, T., Johnson, A. L., Treanor, B., Kucharska, E., Domaschenz, H., et al. (2009). Dock8 Mutations Cripple B Cell Immunological Synapses, Germinal Centers and Long-Lived Antibody Production. *Nat. Immunol.* 10 (12), 1283–1291. doi:10.1038/ni.1820

Randall, K. L., Law, H. D., Ziolkowski, A. F., Wirasinha, R. C., Goodnow, C. C., and Daley, S. R. (2021). DOCK8 Deficiency Diminishes Thymic T-regulatory Cell Development but Not Thymic Deletion. *Clin. Transl Immunol.* 10 (1), 1–8. doi:10.1002/cti2.1236

Robert, F., Barbeau, M., Éthier, S., Dostie, J., and Pelletier, J. (2015). Pharmacological Inhibition of DNA-PK Stimulates Cas9-Mediated Genome Editing. *Genome Med.* 7 (1), 93. doi:10.1186/s13073-015-0215-6

Román-Rodríguez, F. J., Ugalde, L., Álvarez, L., Díez, B., Ramírez, M. J., Risueño, C., et al. (2019). NHEJ-mediated Repair of CRISPR-Cas9-Induced DNA Breaks Efficiently Corrects Mutations in HSPCs from Patients with Fanconi Anemia. *Cell Stem Cell.* 25 (5), 607–621. doi:10.1016/j.stem.2019.08.016

Romero, Z., Lomova, A., Said, S., Miggelbrink, A., Kuo, C. Y., Campo-Fernandez, B., et al. (2019). Editing the Sickle Cell Disease Mutation in Human Hematopoietic Stem Cells: Comparison of Endonucleases and Homologous Donor Templates. *Mol. Ther.* 27 (8), 1389–1406. doi:10.1016/j.ymthe.2019.05.014

Roth, T. L., Puig-Saus, C., Yu, R., Shifrut, E., Carnevale, J., Li, P. J., et al. (2018). Reprogramming Human T Cell Function and Specificity with Non-viral Genome Targeting. *Nature* 559 (7714), 405–409. doi:10.1038/s41586-018-0326-5

Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.-I. T., and Yamamoto, T. (2016). MMEJ-assisted Gene Knock-In Using TALENs and CRISPR-Cas9 with the PITCh Systems. *Nat. Protoc.* 11 (1), 118–133. doi:10.1038/nprot.2015.140

Sakurai, T., Kukimoto-Niino, M., Kunimura, K., Yamane, N., Sakata, D., Aihara, R., et al. (2021). A Conserved PI(4,5)P2-binding Domain Is Critical for Immune Regulatory Function of DOCK8. *Life Sci. Alliance* 4 (4), e202000873. doi:10.26508/lsa.202000873

Sandquist, I., and Kolls, J. (2018). Update on Regulation and Effector Functions of Th17 Cells. *F1000Res* 7 (0), 205–208. doi:10.12688/f1000research.13020.1

Schiroli, G., Conti, A., Ferrari, S., della Volpe, L., Jacob, A., Albano, L., et al. (2019). Precise Gene Editing Preserves Hematopoietic Stem Cell Function Following Transient P53-Mediated DNA Damage Response. *Cell Stem Cell.* 24 (4), 551–565. doi:10.1016/j.stem.2019.02.019

Scholefield, J., and Harrison, P. T. (2021). Prime Editing - an Update on the Field. *Gene Ther.* 28, 396–401. doi:10.1038/s41434-021-00263-9

Scott, O., Kim, V. H.-D., Reid, B., Pham-Huy, A., Atkinson, A. R., Aiuti, A., et al. (2017). Long-Term Outcome of Adenosine Deaminase-Deficient Patients-A Single-Center Experience. *J. Clin. Immunol.* 37 (6), 582–591. doi:10.1007/s10875-017-0421-7

Shi, H., Liu, C., Tan, H., Li, Y., Nguyen, T.-L. M., Dhungana, Y., et al. (2018). Hippo Kinases Mst1 and Mst2 Sense and Amplify IL-2R-STAT5 Signaling in Regulatory T Cells to Establish Stable Regulatory Activity. *Immunity* 49 (5), 899–914. doi:10.1016/j.immuni.2018.10.010

Shin, J. J., Schröder, M. S., Caiado, F., Wyman, S. K., Bray, N. L., Bordi, M., et al. (2020). Controlled Cycling and Quiescence Enables Efficient HDR in Engraftment-Enriched Adult Hematopoietic Stem and Progenitor Cells. *Cel Rep.* 32 (9), 108093. doi:10.1016/j.celrep.2020.108093

Singh, A. K., Eken, A., Hagin, D., Komal, K., Bhise, G., Shaji, A., et al. (2017). DOCK8 Regulates Fitness and Function of Regulatory T Cells through Modulation of IL-2 Signaling. *JCI Insight* 2 (19), 1–15. doi:10.1172/jci.insight.94275

Singh, R., Kuscu, C., Quinlan, A., Qi, Y., and Adli, M. (2015). Cas9-chromatin Binding Information Enables More Accurate CRISPR Off-Target Prediction. *Nucleic Acids Res.* 43 (18), e118. doi:10.1093/nar/gkv575

Slatter, M. A., and Gennery, A. R. (2018). Hematopoietic Cell Transplantation in Primary Immunodeficiency - Conventional and Emerging Indications. *Expert Rev. Clin. Immunol.* 14 (2), 103–114. doi:10.1080/1744666X.2018.1424627

Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y. E., and Zhang, J. (2016). RS-1 Enhances CRISPR/Cas9- and TALEN-Mediated Knock-In Efficiency. *Nat. Commun.* 7 (1), 10548. doi:10.1038/ncomms10548

Stadtmauer, E. A., Fraietta, J. A., Davis, M. M., Cohen, A. D., Weber, K. L., Lancaster, E., et al. (2020). CRISPR-engineered T Cells in Patients with Refractory Cancer. *Science* 367, 367. doi:10.1126/science.aba7365

Su, H. C., Jing, H., Angelus, P., and Freeman, A. F. (2019). Insights into Immunity from Clinical and Basic Science Studies of DOCK8 Immunodeficiency Syndrome. *Immunol. Rev.* 287 (1), 9–19. doi:10.1111/imr.12723

Su, H. C., Jing, H., and Zhang, Q. (2011). DOCK8 Deficiency. *Ann. N. Y Acad. Sci.* 1246 (1), 26–33. doi:10.1111/j.1749-6632.2011.06295.x

Su, H. C., and Orange, J. S. (2020). The Growing Spectrum of Human Diseases Caused by InheritedCDC42 Mutations. *J. Clin. Immunol.* 40 (4), 551–553. doi:10.1007/s10875-020-00785-8

Suzuki, K., and Izpisua Belmonte, J. C. (2018). *In Vivo* genome Editing via the HITI Method as a Tool for Gene Therapy. *J. Hum. Genet.* 63 (2), 157–164. doi:10.1038/s10038-017-0352-4

Sweeney, C. L., Pavel-Dinu, M., Choi, U., Brault, J., Liu, T., Koontz, S., et al. (2021). Correction of X-CGD Patient HSPCs by Targeted CYBB cDNA Insertion Using CRISPR/Cas9 with 53BP1 Inhibition for Enhanced Homology-Directed Repair. *Gene Ther.* 28, 373–390. doi:10.1038/s41434-021-00251-z

Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D. J., and Cathomen, T. (2007). Structure-based Redesign of the Dimerization Interface Reduces the Toxicity of Zinc-finger Nucleases. *Nat. Biotechnol.* 25 (7), 786–793. doi:10.1038/nbt1317

Talib, S., and Shepard, K. A. (2020). Unleashing the Cure: Overcoming Persistent Obstacles in the Translation and Expanded Use of Hematopoietic Stem Cell-Based Therapies. *Stem Cell Transl Med* 9 (4), 420–426. doi:10.1002/sctm.19-0375

Tang, W., Dou, Y., Qin, T., Ding, Y., Tang, X., Zhao, X., et al. (2019). Skewed B Cell Receptor Repertoire and Reduced Antibody Avidity in Patients with DOCK8 Deficiency. *Scand. J. Immunol.* 89 (6), e12759–9. doi:10.1111/sji.12759

Tangye, S. G., Al-Herz, W., Bousfiha, A., Chatila, T., Cunningham-Rundles, C., Etzioni, A., et al. (2020). Human Inborn Errors of Immunity: 2019 Update on the Classification from the International Union of Immunological Societies Expert Committee. *J. Clin. Immunol.* 40 (1), 24–64. doi:10.1007/s10875-019-00737-x

Tangye, S. G., Pillay, B., Randall, K. L., Avery, D. T., Phan, T. G., Gray, P., et al. (2017). Dedicator of Cytokinesis 8-deficient CD4 + T Cells Are Biased to a T H 2 Effector Fate at the Expense of T H 1 and T H 17 Cells. *J. Allergy Clin. Immunol.* 139 (3), 933–949. doi:10.1016/j.jaci.2016.07.016

Tebas, P., Stein, D., Tang, W. W., Frank, I., Wang, S. Q., Lee, G., et al. (2014). Gene Editing ofCCR5in Autologous CD4 T Cells of Persons Infected with HIV. *N. Engl. J. Med.* 370 (10), 901–910. doi:10.1056/NEJMoa1300662

Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., et al. (2015). GUIDE-seq Enables Genome-wide Profiling of Off-Target Cleavage by CRISPR-Cas Nucleases. *Nat. Biotechnol.* 33 (2), 187–197. doi:10.1038/nbt.3117

Tsilifis, C., Freeman, A. F., and Gennery, A. R. (2021). STAT3 Hyper-IgE Syndrome-An Update and Unanswered Questions. *J. Clin. Immunol.* 41 (5), 864–880. doi:10.1007/s10875-021-01051-1

Turchiano, G., Andrieux, G., Klermund, J., Blattner, G., Pennucci, V., el Gaz, M., et al. (2021). Quantitative Evaluation of Chromosomal Rearrangements in Gene-Edited Human Stem Cells by CAST-Seq. *Cell Stem Cell.* 28 (6), 1136–1147. doi:10.1016/j.stem.2021.02.002

Urnov, F. D. (2021). CRISPR-Cas9 Can Cause Chromothripsis. *Nat. Genet.* 53 (6), 768–769. doi:10.1038/s41588-021-00881-4

Vaidyanathan, S., Azizian, K. T., Haque, A. K. M. A., Henderson, J. M., Hendel, A., Shore, S., et al. (2018). Uridine Depletion and Chemical Modification Increase Cas9 mRNA Activity and Reduce Immunogenicity without HPLC Purification. *Mol. Ther. - Nucleic Acids* 12, 530–542. doi:10.1016/j.omtn.2018.06.010

Vaidyanathan, S., Baik, R., Chen, L., Bravo, D. T., Suarez, C. J., Abazari, S. M., et al. (2021). Targeted Replacement of Full-Length CFTR in Human Airway Stem Cells by CRISPR/Cas9 for Pan-Mutation Correction in the Endogenous Locus. *bioRxiv*. doi:10.1101/2021.02.26.432961

Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., et al. (2018). A High-Fidelity Cas9 Mutant Delivered as a Ribonucleoprotein Complex Enables Efficient Gene Editing in Human Hematopoietic Stem and Progenitor Cells. *Nat. Med.* 24 (8), 1216–1224. doi:10.1038/s41591-018-0137-0

Vavassori, V., Mercuri, E., Marcovecchio, G. E., Castiello, M. C., Schiroli, G., Albano, L., et al. (2021). Modeling, Optimization, and Comparable Efficacy of T Cell and Hematopoietic Stem Cell Gene Editing for Treating hyper-IgM Syndrome. *EMBO Mol. Med.* 13 (3), 1–25. doi:10.15252/emmm.202013545

Wang, J., Exline, C. M., Declercq, J. J., Llewellyn, G. N., Hayward, S. B., Li, P. W.-L., et al. (2015). Homology-driven Genome Editing in Hematopoietic Stem and Progenitor Cells Using ZFN mRNA and AAV6 Donors. *Nat. Biotechnol.* 33 (12), 1256–1263. doi:10.1038/nbt.3408

Wienert, B., Wyman, S. K., Richardson, C. D., Yeh, C. D., Akcakaya, P., Porritt, M. J., et al. (2019). Unbiased Detection of CRISPR Off-Targets *In Vivo* Using DISCOVER-Seq. *Science* 364 (6437), 286–289. doi:10.1126/science.aav9023

Wilkinson, A. C., Dever, D. P., Baik, R., Camarena, J., Hsu, I., Charlesworth, C. T., et al. (2021). Cas9-AAV6 Gene Correction of Beta-Globin in Autologous HSCs Improves Sickle Cell Disease Erythropoiesis in Mice. *Nat. Commun.* 12 (1), 1–9. doi:10.1038/s41467-021-20909-x

Wilson, A., Laurenti, E., Oser, G., van der Wath, R. C., Blanco-Bose, W., Jaworski, M., et al. (2008). Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. *Cell.* 135 (6), 1118–1129. doi:10.1016/j.cell.2008.10.048

Xu, L., Wang, J., Liu, Y., Xie, L., Su, B., Mou, D., et al. (2019). CRISPR-edited Stem Cells in a Patient with HIV and Acute Lymphocytic Leukemia. *N. Engl. J. Med.* 381 (13), 1240–1247. doi:10.1056/NEJMoa1817426

Xu, X., Han, L., Zhao, G., Xue, S., Gao, Y., Xiao, J., et al. (2017). LRCH1 Interferes with DOCK8-Cdc42-Induced T Cell Migration and Ameliorates Experimental Autoimmune Encephalomyelitis. *J. Exp. Med.* 214 (1), 209–226. doi:10.1084/jem.20160068

Yamamura, K., Uruno, T., Shiraishi, A., Tanaka, Y., Ushijima, M., Nakahara, T., et al. (2017). The Transcription Factor EPAS1 Links DOCK8 Deficiency to Atopic Skin Inflammation via IL-31 Induction. *Nat. Commun.* 8, 8. doi:10.1038/ncomms13946

Yao, X., Wang, X., Hu, X., Liu, Z., Liu, J., Zhou, H., et al. (2017). Homology-mediated End Joining-Based Targeted Integration Using CRISPR/Cas9. *Cell Res.* 27 (6), 801–814. doi:10.1038/cr.2017.76

Zhang, Q., Davis, J. C., Dove, C. G., and Su, H. C. (2010). Genetic, Clinical, and Laboratory Markers for DOCK8 Immunodeficiency Syndrome. *Dis. Markers* 29 (3-4), 131–139. doi:10.3233/DMA-2010-0737

Zhang, Q., Davis, J. C., Lamborn, I. T., Freeman, A. F., Jing, H., Favreau, A. J., et al. (2009). Combined Immunodeficiency Associated withDOCK8Mutations. *N. Engl. J. Med.* 361 (21), 2046–2055. doi:10.1056/nejmoa0905506

Zhang, Q., Dove, C. G., Hor, J. L., Murdock, H. M., Strauss-Albee, D. M., Garcia, J. A., et al. (2014). DOCK8 Regulates Lymphocyte Shape Integrity for Skin Antiviral Immunity. *J. Exp. Med.* 211 (13), 2549–2566. doi:10.1084/jem.20141307

Zhang, Y., Garcia-Ibanez, L., and Toellner, K.-M. (2016). Regulation of Germinal center B-Cell Differentiation. *Immunol. Rev.* 270 (1), 8–19. doi:10.1111/imr.12396

Conflict of Interest: ROB holds equity in Graphite Bio and UNIKUM Therapeutics and is a part-time employee of UNIKUM Therapeutics.

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

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

frontiers | Frontiers in Genome Editing

ORIGINAL RESEARCH
published: 23 May 2022
doi: 10.3389/fgeed.2022.828489

Genome Editing With TALEN, CRISPR-Cas9 and CRISPR-Cas12a in Combination With AAV6 Homology Donor Restores T Cell Function for XLP

Benjamin C. Houghton[1], *Neelam Panchal*[1], *Simone A. Haas*[2,3], *Kay O. Chmielewski*[2,3,4], *Markus Hildenbeutel*[2,3], *Thomas Whittaker*[1], *Claudio Mussolino*[2,3], *Toni Cathomen*[2,3], *Adrian J Thrasher*[1] *and Claire Booth*[1]*

[1]*Molecular and Cellular Immunology, UCL Great Ormond Street Institute of Child Health, London, United Kingdom,* [2]*Institute for Transfusion Medicine and Gene Therapy, Medical Center – University of Freiburg, Freiburg, Germany,* [3]*Center for Chronic Immunodeficiency, Faculty of Medicine, University of Freiburg, Freiburg, Germany,* [4]*Faculty of Biology, University of Freiburg, Freiburg, Germany*

OPEN ACCESS

Edited by:
Rasmus O. Bak,
Aarhus University, Denmark

Reviewed by:
Amanda Mary Dudek,
Stanford University, United States
Dimitrios Laurin Wagner,
Charité - Universitätsmedizin Berlin, Germany
Oscar Quintana Bustamante,
CIEMAT/CIBERER/IIS-FJD, Spain

***Correspondence:**
Claire Booth
C.booth@ucl.ac.uk

Specialty section:
This article was submitted to Genome Editing in Blood Disorders, a section of the journal Frontiers in Genome Editing

Received: *03 December 2021*
Accepted: *06 April 2022*
Published: *23 May 2022*

Citation:
Houghton BC, Panchal N, Haas SA, Chmielewski KO, Hildenbeutel M, Whittaker T, Mussolino C, Cathomen T, Thrasher AJ and Booth C (2022) Genome Editing With TALEN, CRISPR-Cas9 and CRISPR-Cas12a in Combination With AAV6 Homology Donor Restores T Cell Function for XLP. Front. Genome Ed. 4:828489. doi: 10.3389/fgeed.2022.828489

X-linked lymphoproliferative disease is a rare inherited immune disorder, caused by mutations or deletions in the *SH2D1A* gene that encodes an intracellular adapter protein SAP (Slam-associated protein). SAP is essential for mediating several key immune processes and the immune system - T cells in particular - are dysregulated in its absence. Patients present with a spectrum of clinical manifestations, including haemophagocytic lymphohistiocytosis (HLH), dysgammaglobulinemia, lymphoma and autoimmunity. Treatment options are limited, and patients rarely survive to adulthood without an allogeneic haematopoietic stem cell transplant (HSCT). However, this procedure can have poor outcomes in the mismatched donor setting or in the presence of active HLH, leaving an unmet clinical need. Autologous haematopoeitic stem cell or T cell therapy may offer alternative treatment options, removing the need to find a suitable donor for HSCT and any risk of alloreactivity. SAP has a tightly controlled expression profile that a conventional lentiviral gene delivery platform may not be able to fully replicate. A gene editing approach could preserve more of the endogenous regulatory elements that govern SAP expression, potentially providing a more optimum therapy. Here, we assessed the ability of TALEN, CRISPR-Cas9 and CRISPR-Cas12a nucleases to drive targeted insertion of *SAP* cDNA at the first exon of the *SH2D1A* locus using an adeno-associated virus serotype 6 (AAV6)-based vector containing the donor template. All nuclease platforms were capable of high efficiency gene editing, which was optimised using a serum-free AAV6 transduction protocol. We show that T cells from XLP patients corrected by gene editing tools have restored physiological levels of *SAP* gene expression and restore SAP-dependent immune functions, indicating a new therapeutic opportunity for XLP patients.

Keywords: TALEN, CRISPR (clustered regularly interspaced short palindromic repeat)/Cas9 (CRISPR associated protein 9)-mediated genome editing, Cas12a, X-linked lymphoproliferative disease (XLP), primary immunodefciencies, AAV6, T cell therapy, homology-directed repair

INTRODUCTION

X linked lymphoproliferative disease (XLP) is a rare inherited immune system disorder, affecting 1–2:1,000,000 births (Booth et al., 2011). It arises due to mutations and deletions in the *SH2D1A* gene, which encodes an intracellular adaptor protein SLAM-associated protein (SAP) that is critical for relaying signals received at the cell surface by SLAM family receptors (Booth et al., 2011; Cannons et al., 2011). SAP is a small 128 amino acid cytoplasmic protein consisting of a single Src homology-2 (SH2) domain and a short C-terminal tail (Sayos et al., 1998). By binding to specific tyrosine-based motifs in the cytoplasmic tail of SLAM family receptors - such as SLAM, 2B4, NTB-A, Ly9 CD84 and CRACC - *via* an arginine residue in the SH2 domain (Cannons et al., 2011), SAP can recruit additional proteins that can activate downstream signaling cascades.

In the absence of SAP, several immune functions are affected, including reduced T cell and NK cell cytotoxicity, a lack of NKT cell development, defective CD4 T follicular helper (T_{FH}) cell help to B cells leading to abnormal humoral function, and a reduced sensitivity to restimulation-induced cell death (RICD) that contributes to unconstrained immune responses to viral infection (Ma et al., 2006). These deficits give rise to a range of clinical manifestations, including haemophagocytic lymphohistiocytosis (HLH), dysgammaglobulinemia, lymphoma and autoimmunity (Panchal et al., 2018a).

Treatment options for XLP patients are limited and the only curative therapy is a haematopoietic stem cell transplant (HSCT), however, outcomes can be poor in the mismatched donor setting, as patients are at risk of graft-vs. host disease (GvHD). An autologous gene corrective approach could fulfil an unmet clinical need for patients lacking a well-matched donor. We have previously shown that lentiviral gene addition can restore SAP protein expression and immune function when delivered to haematopoietic stem cells (HSCs) and T cells, in several *in vitro* and *in vivo* models (Rivat et al., 2013; Panchal et al., 2018b). However, SAP has a tightly controlled expression profile, limited to T_{convs} (not T_{reg}), NK and NKT cells. Within T cell subsets, SAP expression levels are upregulated after TCR engagement (Mehrle et al., 2005) and alter with memory or effector phenotypes (Hale et al., 2013), indicating an importance of finely tuned control and giving rise to concern that uncontrolled expression of this important signaling molecule in a conventional gene therapy procedure could cause further dysregulation. Although there is no direct evidence that irregular SAP expression is pathogenic, elevated (Geng et al., 2021) or decreased (Liu et al., 2021; Yang et al., 2021) levels are seen in several diseases, and SLAM signaling pathway alterations are implicated in autoimmunity (Comte et al., 2018; Dragovich and Mor, 2018; Malaer et al., 2019; Gartshteyn et al., 2021).

We hypothesised that a gene editing approach, using site-specific nucleases and a homology-directed repair (HDR) template to place a corrective SAP cDNA under the control of the full native promoter, could harness more of the endogenous regulatory elements that govern SAP expression, to potentially provide an optimal therapy. Genome editing is now a clinical reality due to the advent of highly site-specific and efficient DNA nucleases, including zinc-finger nucleases (ZFN), transcription activator–like effector-nucleases (TALENs) and CRISPR-Cas systems. All these platforms have entered the clinic, and the data from these trials is building an understanding of their safety profile and clinical efficacy (Tebas et al., 2014; Qasim et al., 2017; Frangoul et al., 2020; Stadtmauer et al., 2020; Gillmore et al., 2021).

TALENs are comprised of two customisable protein DNA binding domains, each expressed as a fusion protein to a *Fok*I endonuclease domain that upon dimerization creates a staggered DNA double stand break when the DNA binding domains are targeted to proximal loci on the DNA. Cas9 and Cas12a nucleases are guided by short RNAs that bind to their specific DNA loci *via* Watson-Crick base pairing. DNA cleavage is dependent on the presence of a protospacer adjacent motif (PAM) that takes the sequence NGG for *Streptococcus pyogenes* Cas9 (SpCas9) and TTTV for *Acidaminococcus* sp. Cas12a (AsCas12a). Cas9 is now the most widely used platform, creating a blunt end cut 3 nucleotides upstream of the PAM. However, Cas12a is an attractive novel editing platform, due to its creation of a 5bp overhang (downstream of PAM, from 18 nucleotides on the non-target strand, to 23 on the target strand) that may more efficiently stimulate HDR. The double strand break created by the nucleases can be harnessed to seamlessly insert therapeutic sequences by supplying a HDR template. For haematopoietic cells, delivery of this repair template *via* a genome of an adeno-associated virus with a serotype 6 (AAV6) capsid has shown the greatest efficiency, and capable of correcting immune dysfunctions in several primary immune deficiencies including SCID-X1, CD40L X-CGD and Wiskott Aldrich syndrome (Genovese et al., 2014; Kuo et al., 2018; Pavel-Dinu et al., 2019; Rai et al., 2020; Sweeney et al., 2021).

Most manifestations of XLP result from defects in T cell immunity. T cells offer an attractive therapeutic candidate, due to their accessibility and amenability to genome editing, and have an established safety record in many clinical trials for HIV and cancer immunotherapy (Tebas et al., 2014; Qasim et al., 2017; Panchal, et al., 2021). We have previously demonstrated that adoptive transfer of gene corrected T cells rescues a murine model of SAP deficiency alongside cellular and humoral abnormalities in XLP patient T cells (Panchal et al., 2018b). Here we set out to show proof of concept for gene editing platforms to correct T cell function using technologies which could then be transferred to HSC correction.

In order to allow treatment of the greatest number of patients, we designed a targeted integration of a *SAP* cDNA close to the initiation codon of the first exon. We identified TALEN, CRISPR-Cas9 and CRISPR-Cas12a nucleases, which were able to mediate high efficiency DNA cleavage in primary human T cells, with minimal off-target nuclease activity. When nuclease activity was coupled with an AAV6 HDR template, over 45% of XLP patient T cells showed targeted insertion of *SAP* cDNA. A serum-free transduction protocol optimised the editing procedure by reducing the AAV6 dose required. Finally, we show that gene edited XLP patient T cells have a restored, physiological level of SAP protein expression and restored SAP-dependent immune functions equal to that of healthy controls in immunoassays of T:B signaling, restimulation-induced cell death (RICD) and cytotoxicity. These data indicate that an

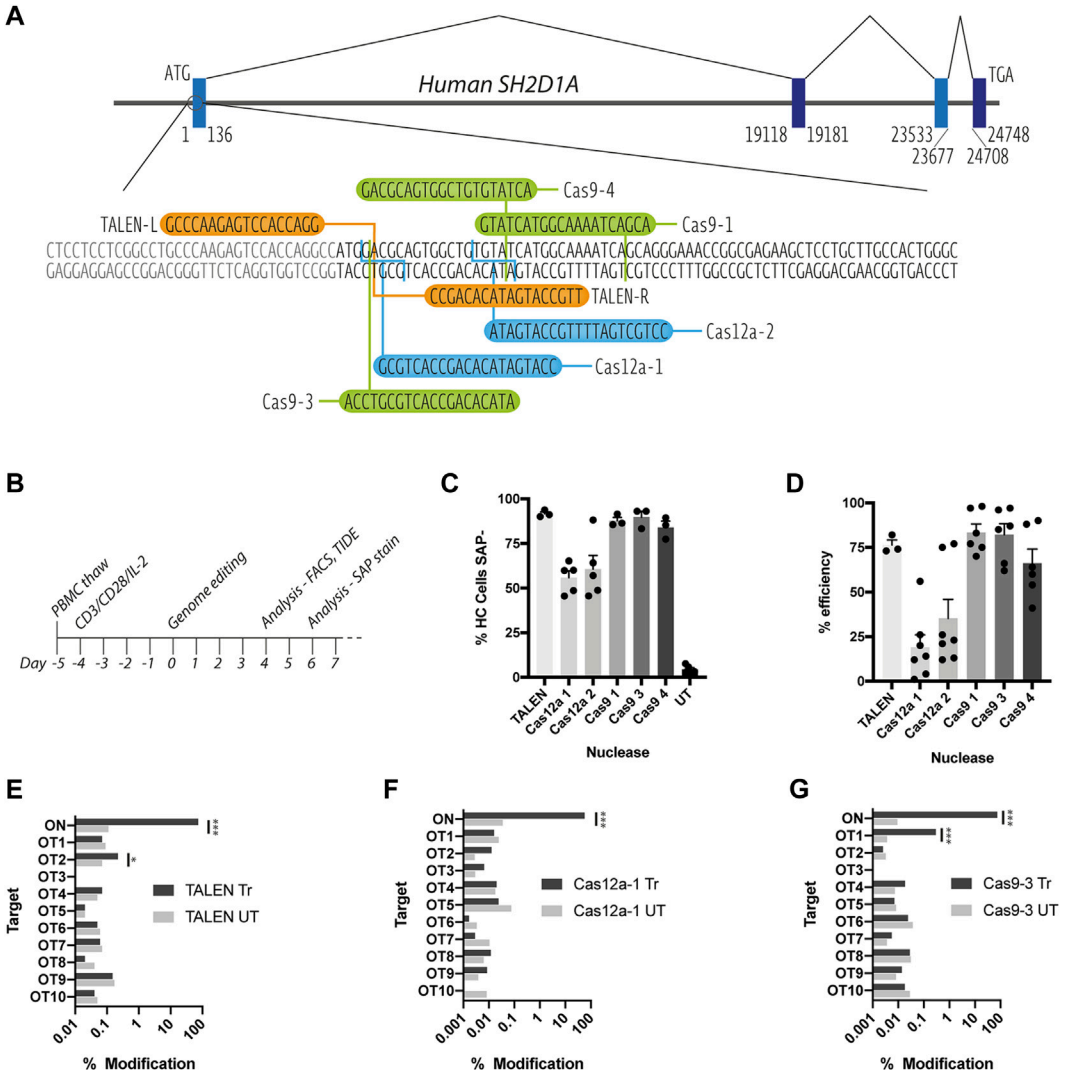


FIGURE 1 | Assessment of on and off-target TALEN-, Cas9-and Cas12a-mediated activity. **(A)** Schematic diagram of the *SH2D1A* locus showing the start of the protein coding region of exon 1, TALEN-L and -R binding sites (orange), Cas9 (green) and Cas12a (blue) gRNA targets **(B)** T cell genome editing experimental timeline **(C)** Healthy control (HC) T cells nucleofected with TALEN mRNA, Cas9 or Cas12a RNPs, assessed for SAP expression by intracellular SAP staining and flow cytometry at d6 post nucleofection (n = 3–6, mean, SEM) **(D)** % INDEL frequency by TIDE analysis of sanger sequencing data of PCR amplicons amplified from nuclease treated T cell genomic DNA (n = 3–7, mean, SEM) **(E–G)** NGS-generated data of modifications at on target (ON) and *in silico* predicted off-target loci (OT1-10) for [E] TALEN [F] Cas12a-1 and [G] Cas9-3 nucleases (n = 1, treated (Tr) vs. untreated (UT), OT sites marked $^{*}p < 0.0001$).

autologous gene editing strategy could present a highly effective therapy for XLP patients.

RESULTS

Targeting the *SH2D1A* Locus at High Efficiency in T Cells by TALEN, CRISPR-Cas9 and CRISPR-Cas12a

To determine the feasibility of a gene correction strategy for XLP, we firstly aimed to determine the optimal nuclease platform for creating the site-specific DNA double strand break. A TALEN pair, three S. p.Cas9 guide RNAs (gRNA) and two A.s.Cas12a crRNAs were identified targeting loci early in exon1 of the *SH2D1A* gene (**Figure 1A**). Stimulated PBMCs were nucleofected with either *in vitro*-transcribed TALEN mRNA, or ribonucleoprotein (RNP) complexes of Cas9 (Alt-R® S. p.HiFi Cas9 V3, IDT) or Cas12a (Alt-R® AsCas12a Ultra, IDT) protein and their respective guides, according to the timeline (**Figure 1B**).

All three nuclease platforms mediated efficient gene editing with TALEN and CRISPR-Cas9 nucleofection resulting in >90%, and CRISPR-Cas12a >50% knock down of SAP protein, when assessed using intracellular staining and flow cytometry (**Figure 1C** and **Supplementary Figure S1**). This data was supported by TIDE analysis detection of insertions and deletions (INDELs) created by non-homologous end-joining DNA repair of DNA double strand breaks (**Figure 1D**). TALEN, Cas12a-1 and Cas9-3 were selected for

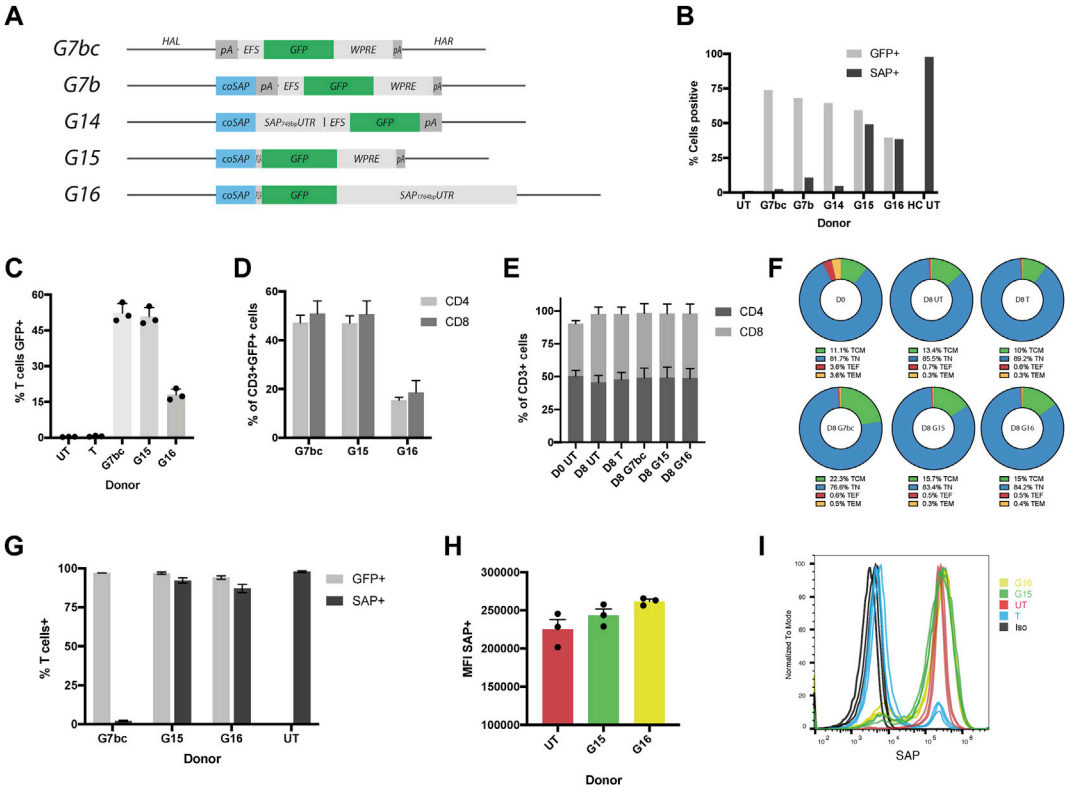


FIGURE 2 | HDR donors mediate efficient integration and drive physiological levels of SAP expression. **(A)** Schematic diagrams of HDR donors designed for targeted insertion at the *SH2D1A* locus: G7bc (Homology arm left (HAL)-bovine growth hormone poly adenylation signal (bGHpA), EF1aShort (EFS), enhanced green fluorescent protein (GFP), woodchuck post-transcriptional regulatory element (WPRE), SV40pA, homology arm right (HAR)); G7b (HAL, codon optimised SAP cDNA (coSAP), bGHpA, EFS, GFP, WPRE, SV40pA, HAR); G14 (HAL, coSAP, 745bp SAP proximal 3′untranslated region (SAP_{745bp}UTR), EFS, GFP, WPRE, SV40pA, HAR); G15 (HAL, coSAP, P2A, GFP, WPRE, SV40pA, HAR); G16 (HAL, coSAP, P2A, GFP, 1764bp proximal 3′untranslated region (SAP_{1764bp}UTR), HAR). All constructs are flanked at the 5 and 3′ by inverted terminal repeats (ITRs) of an AAV2 viral genome (not shown). **(B)** Stimulated XLP patient T cells nucleofected with TALEN mRNA and transduced with HDR donors G7bc, G7b, G14, G15 and G16 alongside untreated patient (UT) and healthy control cells (HC), graph shows %GFP+ (unstained cells) and %SAP+ (intracellularly stained cells), analysed by flow cytometry (n = 1) **(C–I)** Healthy control T cells edited with TALEN and HDR donors, FACS-sorted on GFP + for G/H, n = 3, mean, SEM throughout) **(C)** % T cells positive for GFP (mean: G7bc, 52%; G15,50%; G16,18%) compared to untreated (UT) and TALEN-only (T) controls. **(D)** % of CD3+GFP + T cells expressing either CD4 or CD8 on cell surface **(E)** % $CD4^+$ and $CD8^+$ in $CD3^+$ T cell gate at day 0 (D0) and day 8 (D8 = 4 days post edit). **(F)** CD62L/CD45RA T cell memory phenotype staining in $CD4^+$ cell subset: T central memory (TCM) = CD62L + CD45RA-, Naïve (TN) = CD62L + CD45RA+, T effector (TF) = CD62L-CD45RA-,T effector memory (TEM) = CD62L-CD45RA- **(G)** Post-FACS-sorting %GFP+ and %SAP+ in healthy control T cells edited with G7bc, G15 or G16 HDR donors analysed by flow cytometry on unstained cells and intracellularly SAP stained cells respectively **(H)** Mean fluorescent intensity (MFI) of SAP + cells compared to untreated healthy control (UT) **(I)** Representative flow histogram shows comparable MFI between gene edited (G15, G16) and untreated healthy control cells. Cells treated with TALEN only (T) have both SAP+ and SAP- population, the latter overlaying the isotype control (Iso).

further testing due their proximity to the start codon of the *SAP* coding sequence.

Minimal Detection of off Target Nuclease Activity Across Nuclease Platforms

To assess on and off-target cutting at more depth, we performed targeted next-generation sequencing (NGS) at the on-target (ON) locus and the top 10 *in silico* predicted off-target sites (OT1-10) for TALENs, Cas12a-1 and Cas9-3 nucleases (Determined by PROGNOS and Benchling online software, respectively, see **Supplementary Table S1**). NGS confirmed high efficiency modification at the on-target locus across all platforms giving a modification rate of 74, 57, and 75% for TALEN, Cas12a-1 and Cas9-3 respectively (**Figures 1E–G**). We detected off-target activity at two intronic loci (TALEN OT2 and Cas9-3 OT1) at low frequency (0.22 and 0.29% respectively) (**Figures 1E–G**). TALEN OT2 is in the third intron of the *TET1* (Ten-eleven translocation methylcytosine dioxygenase 1) gene while Cas9-3 OT1 occurs in the 24th intron of *RPTOR* (Regulatory-associated protein of mTOR). In both cases, the distance from the corresponding exon boundaries (up/downstream) are 21900/43664bp and 451/14567bp away, respectively (**Supplementary Figure S2**).

SH2D1A-Targeted DNA Breaks can be Harnessed for HDR-Driven Insertion of *SAP* cDNA

To determine if we could harness the HDR pathway to insert a corrective *SAP* cDNA under its native promoter, we designed a

series of donor templates for delivery *via* AAV6 vector (**Figure 2A**). Donors were constructed into an AAV vector genome using Gibson assembly, to maintain the transition from promoter regions into a codon optimised *SAP* cDNA without restriction enzyme sequences. To allow for simplified identification of positively recombined clones, a *GFP* reporter gene was included, either as a separate expression cassette (G7b, G14), or co-expressed with SAP and cleaved during translation *via* a P2A self-cleaving peptide (G15, G16). In addition, we wanted to compare the endogenous *SH2D1A* 3′untranslated region (UTR) to a woodchuck post-transcriptional element (WPRE). We therefore cloned either the full annotated UTR (1764bp, G16) or a shorter 745bp fragment that contains all the validated microRNA binding sites for the SAP3′UTR, and fills the AAV packaging capacity of the two-cassette format (G14) (**Figure 2A**). A *GFP*-only donor was constructed as a negative control, alongside giving the option to create a functional knockout of *SAP* in healthy control cells, which could provide an additional *in vitro* model (G7bc).

Donor constructs were tested by transducing stimulated XLP patient-derived T cells 15 min post-electroporation of TALEN mRNA (1×10^5 GC/cell). HDR efficiency was assessed by measuring the levels of GFP and SAP protein expression at day 5 post procedure. All donors were able to mediate HDR efficiently, as observed by GFP expression in approximately 50% of the cells, analysed by flow cytometry (**Figure 2B**). Intracelluar SAP protein staining also analysed by flow cytometry showed that G15 and G16 were able to restore SAP protein expression at the same efficiency indicated by GFP fluorescence, as anticipated from the co-translational design. However, HDR donors containing the separate GFP expression cassette (G7b/G14) were not able to restore SAP protein expression, potentially due to suppression of the *SH2D1A* promoter arising from the downstream EFS promoter (Villemure et al., 2001) (**Figure 2B**, **Supplementary Figure S3**).

To confirm this result, we performed further experiments in T cells from several healthy controls using the G7bc and G15 and G16 donors. Flow cytometry analysis confirmed high efficiency editing rates of 52, 50 and 18% respectively in $CD3^+$ T cells (**Figure 2C**, **Supplementary Figure S4A**). T cell phenotype was further interrogated using flow cytometry markers CD4, CD8, CD45RA and CD62L. Importantly for an XLP therapeutic, the GFP + population contained both $CD4^+$ and $CD8^+$ T cells, and these populations were not skewed by the culture period or editing procedure (**Figures 2D,E**). Memory phenotype was similarly unaffected between edited and untouched controls (**Figure 2F**, **Supplementary Figure S5**).

To assess the degree to which HDR donors can replicate endogenous levels of SAP expression, we sorted GFP + cells from the bulk population and performed SAP staining. We found there was no significant difference in the level of SAP expression between untouched controls and cells that had been edited with either G15 or G16 HDR donor (**Figure 2H**). As expected, cells that had been edited with the G7bc donor and FACS-sorted had a complete knockout of SAP protein (**Figure 2G**, **Supplementary Figure S6**). From these data we concluded that the higher rates of editing achieved with the G15 donor, coupled with a highly favourable expression profile merited its use for all further experiments. To confirm that the SAP-containing donors were integrating at the correct loci we performed an in/out PCR, using a forward primer that binds upstream of the 5' limit of the left homology arm and a reverse primer binding within the codon optimized SAP cDNA. As expected, a 953bp band was evident in cells treated with TALEN and G15 or G16 donors (**Supplementary Figure S4B**). We also performed ddPCR on the same amplicon, obtaining copy numbers (relative to an albumin control amplicon of similar length) consistent with the purity of cells selected using FACS-sorting (**Supplementary Table S2**).

Optimisation of Transduction Conditions Allows Reduction of AAV6 HDR Template Dose

It has been reported in several studies that AAV6 is bound and neutralised by galectin 3 binding protein (G3BP) found in human and bovine serum, and that bovine serum-free culture conditions can enhance transduction of HSC and T cells (Denard et al., 2012; Song et al., 2013; Denard et al., 2018; Rogers et al., 2021). To test this in our hands, we transduced Jurkat cells at a range of doses ($0–2 \times 10^5$ GC/cell) in either culture media with or without foetal bovine serum (FBS). Those in FBS-free media were supplemented to full serum after 4 h. We observed significantly more GFP expression at reduced vector doses in the FBS-free conditions (**Supplementary Figure S7**). To investigate this in the context of a gene editing procedure we nucleofected T cells with TALEN, Cas9-3 or Cas12a-1 nucleases prior to transduction with AAV at a range of vector doses within 15 min, in T cell culture media with or without 5% human serum (HS). As before, those in low HS media were supplemented to full serum culture at 4 h. Across all nuclease platforms we found that cells transduced in the absence of HS had improved rates of HDR over a range of AAV6 vector doses (**Figures 3A–C**). Furthermore, we observed that the TALENs were able to mediate the highest rate of HDR up to 50%, while Cas9 and Cas12a had a similar efficiency with approximately 30% GFP positive cells.

When we tested HS-free transduction conditions 2 h prior to nucleofection with TALEN pairs, the effect was more pronounced, with no significant loss of GFP at 33-fold less AAV dose in the HS-free transduction conditions (**Figure 3D**, **Supplementary Figure S8**). As the TALEN pair offered preferential rates of HDR, particularly when using an HS-free transduction protocol, we used this platform for subsequent experiments in XLP patient T cells.

SAP Protein Expression is Restored to the Level of Healthy Controls in Gene Edited XLP Patient T Cells

To ascertain if gene editing could restore SAP protein expression and SAP-dependent immune function to XLP patient T cells, PBMCs from 5 XLP patients were edited with TALEN and G15 HDR donor as before. HDR efficiency was measured by GFP expression, alongside T cell phenotyping markers on day 3 post

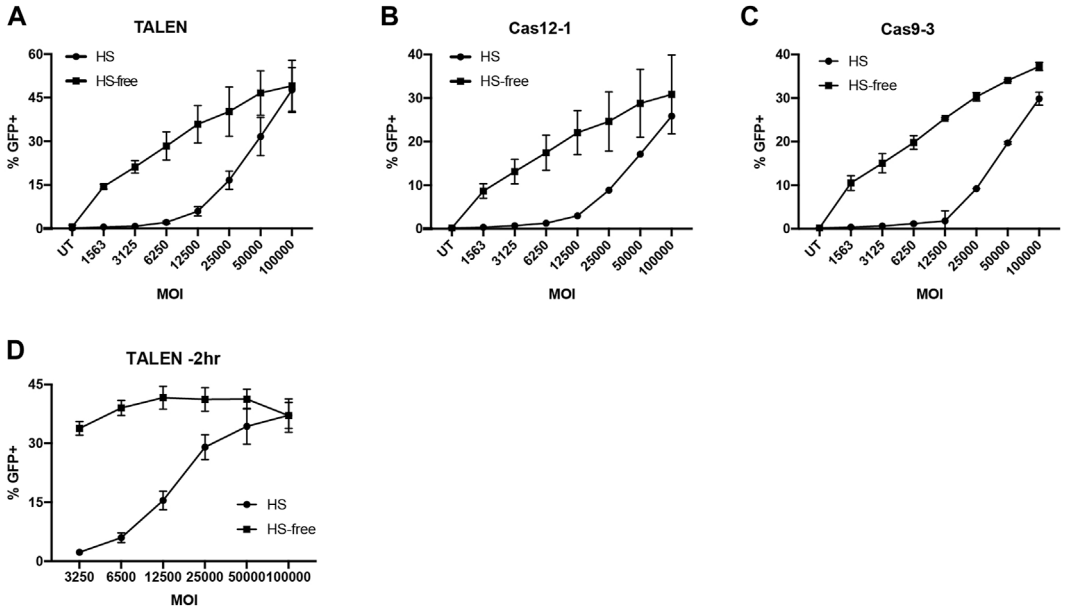


FIGURE 3 | Serum free AAV6 transduction maintains HDR efficiency at reduced MOI. **(A–C)** %GFP + T cells transduced with G7bc AAV in HS or HS-free conditions 15 min post-nucleofection with **(A)** TALEN, **(B)** Cas12a-1 RNP or **(C)** Cas9-3 RNP (N = 2, mean, SD). **(D)** %GFP + T cells transduced with G7bc AAV at a range of MOIs 2h prior to TALEN nucleofection, in human serum (HS) or HS-free transduction conditions (n = 3, mean, SEM).

edit. We achieved an average of 45 and 46% gene corrected lymphocytes, as indicated by GFP fluorescence and intracellular SAP staining respectively (**Figure 4A**, **Supplementary Figure S9**). Crucially, levels of SAP protein expression were highly similar to that of healthy controls. Furthermore, cells that had been transduced with HDR AAV but not electroporated had similar SAP protein to unedited patient cells indicating low levels of expression originating from non-integrated HDR vector (**Figure 4B**). Similarly to healthy T cell editing, we observed an equal distribution between CD4 and CD8 positive T cells (**Figure 4C**). Memory phenotyping markers within CD4 and CD8 T cell subsets were more varied in these samples, however, they remained broadly consistent between patient and healthy control groups throughout the experiment, and edited samples expanded to a similar extent as healthy and unedited patient cells (**Figures 4D,E**, **Supplementary Figure S10**). With this promising data we moved to test SAP-dependent immune function in three *in vitro* models.

Gene Correction Restores SAP-dependent Immune Functions to XLP Patient T Cells

It has been shown that XLP T cells have reduced sensitivity to RICD, a process that in heathy individuals constrains immune responses by triggering apoptosis in cells restimulated through the TCR (Ruffo et al., 2016; Panchal et al., 2018b). To determine if gene edited cells can regain sensitivity to RICD, we cultured the gene edited T cells for 10 days, then restimulated with a range of concentrations of OKT3 antibody for 24 h before quantification of cell death using PI viability dye in flow cytometry following an established protocol (Katz and Snow, 2013). RICD in the bulk population was restored to approximately 50% that of healthy controls, matching the rate of HDR in the bulk population. When the analysis was performed only on corrected cells (GFP+), the amount of cell death equaled that of healthy controls, indicating sensitivity to RICD has been fully restored in the gene edited cells (**Figure 4F**, **Supplementary Figure S11**).

An absence of SAP in T_{FH} cells leads to humoral deficiencies, due to insufficient levels of T:B cell signaling in germinal centres. It has been shown that this interaction can be modelled *in vitro* by culturing naïve T cells with naïve B cells in the presence of SEB for 10 days before assessing levels of cytokine and immunoglobulin in the culture supernatant by ELISA (Ma et al., 2015; Panchal et al., 2018b). Naïve XLP patient T cells that had undergone a gene editing procedure showed restored T:B cell signaling, with B cell secretion of IgG and IgM corrected to the level of healthy controls, and significantly improved T_{FH} IL-21 secretion, compared to untreated patient cells (**Figures 4G–I**).

XLP patient T cells exhibit defective cytotoxicity against Epstein-Barr virus (EBV)-infected B cells (Sharifi et al., 2004; Dupré et al., 2005), which can contribute to the development of HLH. To test if gene editing could restore this function, we generated EBV-specific cytotoxic T lymphocytes (CTLs), by stimulating PBMC with irradiated LCL (Lymphoblastoid cell line, generated in house using EBV-B95.8 and healthy donor PBMC). Despite restimulated T cells showing significant loss of viability during the editing process, we were able to show in one patient sample that, even at low effector:target cell ratio, gene edited XLP patient T cells had similar killing activity to healthy control cells when challenged in a chromium^{51}-release cytotoxicity assay (**Figure 4J**). To support this data, CTLs generated from healthy controls and edited with either G15 or

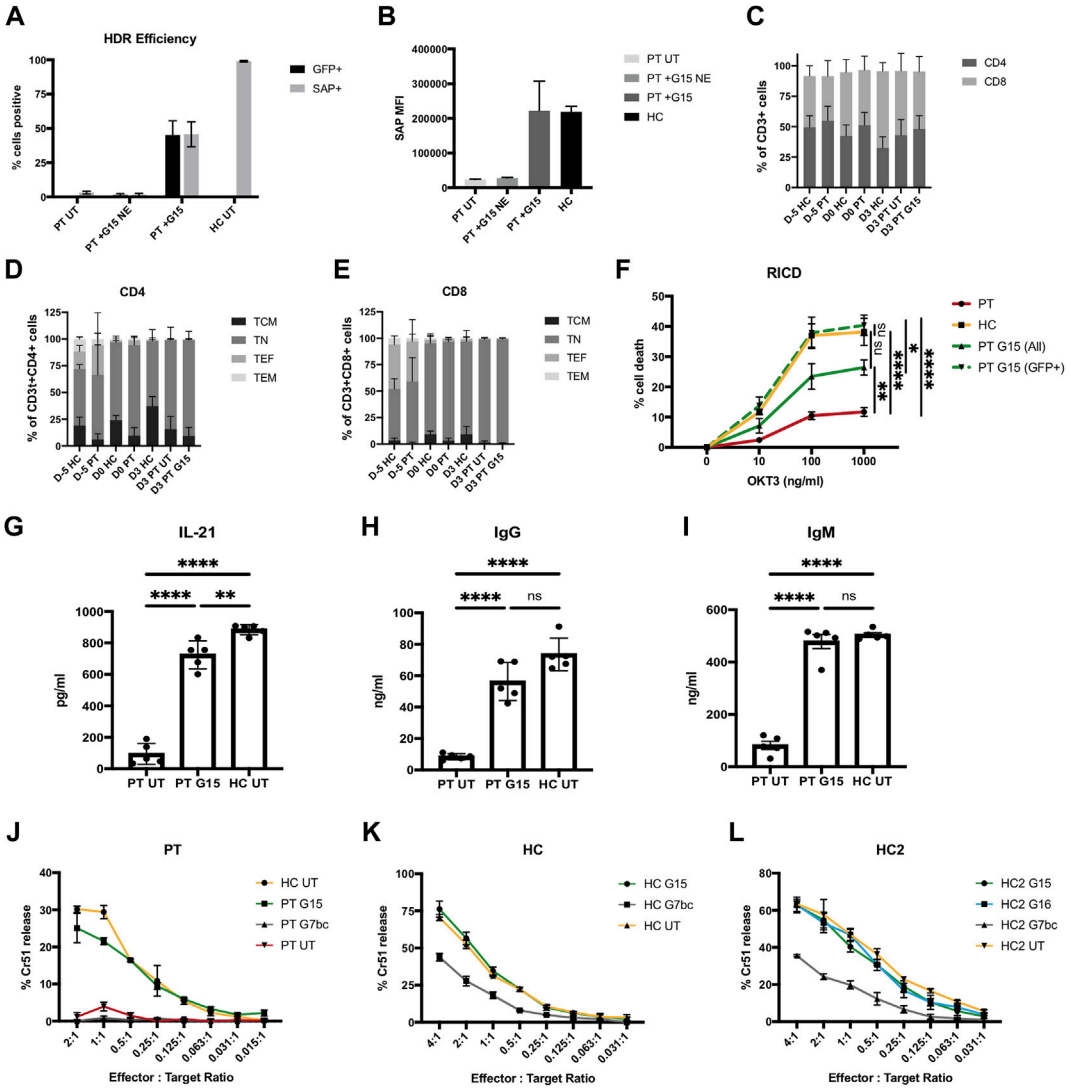


FIGURE 4 | Genome editing restores SAP gene expression and function to XLP patient T cells. **(A–I)** PBMC harvested from 5 unrelated XLP patients stimulated and edited with TALEN and G15 HDR donor supplied 2 h prior to nucleofection in HS-free conditions, MOI2.5 × 10^4. **(A)** %GFP+ (unstained cells, d3 post edit) and SAP+ (intracellularly stained, d6 post edit) lymphocytes analysed by flow cytometry; untouched patient (PT UT) and patient transduced without electroporation with nuclease (PT G15 NE), compared to cells receiving G15 and TALEN (PT + G15) and healthy control cells (HC). **(B)** MFI of SAP expressing lymphocytes. **(C)** % of $CD3^+$ T cells expressing CD4 and CD8 on cell surface at point of thawing (D-5), editing (D0), and FACS analysis (D3) for healthy controls (HC), untreated patient (PT UT) or gene corrected with TALEN and G15 HDR donor (PT G15). **(D)** CD45RA CD62L T cell memory phenotype within the $CD4^+$ T cell population. [E] CD45RA CD62L T cell memory phenotype within the $CD8^+$ T cell population. [F] % cell death of lymphocytes restimulated with either 0, 10, 100, or 1,000 ng/ml OKT3 analysed by flow cytometry analysis by gating on the live bulk population (ALL) or live GFP + cells (GFP+), data normalised to non-restimulated condition with equation given in methods (mean, SEM). Stars denote Tukeys post-hoc test on one-way ANOVA performed on 1000 ng/ml OKT condition only: PT vs. HC, PT vs. PT G15 (GFP+), ****$p < 0.0001$; PT vs. PT G15 (All), **$p = 0.0058$; HC vs. PT G15 (All), ns $p = 0.0660$; HC vs. PT G15 (GFP+), ns $p = 0.9524$; PT G15 (All) vs. PT G15 (GFP+),*$p = 0.0160$. **(G–H)** Quantification of IL-21, IgG and IgM present in supernatant of 10 days co-culture of untreated patient cell (PT), the bulk population (unsorted) of gene edited patient cells (PT G15) or untreated healthy control (HC) naïve T cells with allogenic B cells in the presence of SEB, by means of ELISA (n = 5, mean, SEM). Stars denote Tukeys post-hoc test on one-way ANOVA **[G]** PT UT vs. PT G15/PT UT vs. HC UT, ****$p < 0.0001$; PT G15 vs. HC UT **$p = 0.0081$ [H] PT UT vs. PT G15/PT UT vs. HC UT, ****$p <$ 0.0001; PT G15 vs. HC UT, ns $p = 0.0727$ **(I)** PT UT vs. PT G15/PT UT vs. HC UT, ****$p < 0.0001$; PT G15 vs. HC UT, ns $p = 0.7534$. **(J–L)** FACS-sorted (GFP+) CTL cytotoxic lymphocytes challenged against $Chromium^{51}$-labelled EBV-B cell lymphoblastoid cell line targets at a series of effector:target cell ratios, graphs show % Chromium release compared to complete cell lysis, samples analysed in triplicate (n = 1, mean, SD). **(J)** Patient cells edited with G15 or G7bc HDR donor compared to untreated patient and healthy control. **(K)** Healthy control CTLs edited with either G15 or G7bc HDR donors and FACS-sorted on GFP + cells prior to assay, compared to untreated HC cells. **(L)** An additional healthy control CTL line (HC2) edited with G7bc, G15 and G16 HDR donors.

G7bc HDR donors were GFP-sorted and similarly challenged. In two separate experiments, CTLs edited with SAP containing donor G15 showed equal killing to unedited healthy control CTLs, while those edited to create a SAP knockout using donor G7bc (**Figure 2G**) showed reduced killing activity (**Figures 4J,L**).

DISCUSSION

HSC therapy is widely used to treat primary immunodeficiencies, including XLP. However, GvHD remains a significant risk in the mismatched donor setting, leaving an unmet need that could be fulfilled by an autologous gene correction approach. We have previously shown that lentiviral vectors can be used to deliver a corrective copy of SAP cDNA into HSCs and T cells, to restore immune function *in vitro* and *in vivo* models of XLP. However, SAP has a tightly restricted profile that is challenging to replicate using this technology, which may be of particular importance when developing an HSC gene therapy approach. In this study, we aimed to use gene editing technology to create a site-specific insertion of SAP cDNA, hypothesising that this would harness more of the endogenous DNA regulatory mechanisms that govern SAP expression, to provide more physiological expression pattern and therefore a more optimal therapy in both the HSC and T cell setting.

Genome editing is centered on the creation of a site-specific DNA break, and there has been significant improvement in the specificity and efficiency of nucleases in recent years that has brought this technology into the clinic (Qasim et al., 2017; Frangoul et al., 2020). We opted to test three nuclease technologies, TALEN, CRISPR-S. p.HiFi Cas9 V3 and the more recently identified high activity mutant CRISPR-A.s.Cas12a Ultra (Zhang et al., 2021) to determine an optimal editor for the *SH2D1A* locus. We found all platforms capable of creating double strand DNA breaks at the *SH2D1A* locus at high efficiency. Cas9 creates a blunt cut, while TALENs and Cas12a create staggered cut that may more efficiently stimulate HDR (Zetsche et al., 2015). Despite generating INDELs at roughly half that of the Cas9 RNPs, our Cas12a RNPs were able to mediate similar levels of HDR as Cas9. In addition, TALENs created similar INDELs to Cas9 but were able to mediate more HDR, supporting this hypothesis.

We used targeted NGS to investigate on and off-target nuclease activity at sites predicted by *in silico* prediction software. On target amplicons confirmed the highly efficiency genome modification mediated by all three nuclease platforms. Modifications at off-target loci occurred at low frequency, and were not detected in exons, indeed the nearest exon boundary was over 450bp away for Cas9-3 OT1 (*RPTOR*) and over 20 kb for TALEN OT2 (*TET1*), suggesting the risk of mutagenesis in coding regions remains low. However, it may be possible to mitigate the TALEN off target activity by using a recently described *FokI* mutant (Miller et al., 2019). Miller and others identified several single residue substitutions in *Fok*I that exhibited a significant reduction in off-target activity, up to 1000-fold in some case. The substitutions lead to a reduction in catalytic rate, allowing dissociation from mismatched targets before cleavage occurs. Although several variants may need to be tested for each locus, this technology could be readily transferred to the TALEN platform (Miller et al., 2019).

It has previously been shown that using a RNP delivery format and a high-fidelity Cas9 variant can reduce off target cutting activity of Cas9, yet, despite adopting these technologies, we found evidence of low level off-target cutting at the Cas9-3 OT1 site. Mitigation may be possible through introducing internal 2′-O-methyl-3′-phosphonoacetate modifications to gRNA, however, in the absence of established design principles, optimal modifications must be determined empirically for each guide (Ryan et al., 2018; Vakulskas and Behlke, 2019) Further Cas9 protein engineering is likely to identify additional variants with improved target specificity. Indeed, a recent study used kinetics-guided cryo-electron microscopy to identify a kinked duplex structure, formed between DNA and gRNA with mismatches only at the PAM-distal positions 18–20, that supports an active Cas9 configuration (Bravo et al., 2022). A targeted mutation (creating SuperFi-Cas9) was able ablate this structure and reduce off target activity of sgRNA with PAM-distal mismatcheds. Although at an early stage, these studies could be particularly useful to our application, as Cas9-3 OT1 mismatches occur at gRNA positions 19 and 20.

As both *TET1* and *RPTOR* are important for haematopoiesis and T-cell differentiation and function, there is a possibility of mutation of intronic regulatory signals, further studies of mutations at these OT loci may be required, alongside unbiased genome-wide screening for additional loci not predicted in silico, prior to translation to the clinic (Ko et al., 2015; Wang et al., 2016), Little is known about mechanisms involved in SAP gene expression. Consensus binding site for Ets-1 and Ets-2 have been identified as being important for core promoter activity (Okamoto et al., 2004), and we have previously attempted to generate lentiviral vectors incorporating this motif by cloning proximal sections of the promoter. The expression output of these vectors was significantly below physiological levels, indicating that other motifs must be important (Panchal et al., 2018b). Our exon1 gene editing approach allows for the all the upstream promoter regulatory mechanisms to be preserved. At the 3′end of SH2D1A encoding mRNA, RNA binding protein sites have been identified that could have a role in regulating SH2D1A RNA stability in the cells (Okamoto et al., 2004). In addition, several studies have validated microRNA binding sites in the SAP3′UTR that have a role in regulation of SAP (Ding et al., 2012), most prominently miRNA-31 (Heide et al., 2016; Ripamonti et al., 2017), which is known to play a role in regulating IL-2 secretion (Fan et al., 2012) and IFN signalling in CD8 T cells (Moffett et al., 2017). To take advantage of these regulatory motifs, we generated an HDR donor that incorporated the full length 3′UTR, as this strategy has been shown to improve the expression profile in other T cell editing studies (Hubbard et al., 2016). However, we did not see any notable improvement in our studies but were hampered by reduced HDR efficiency. AAV vector titre was also reduced, likely due to the full-length UTR slightly exceeding the packaging capacity of AAV particles, leading to packaging of incomplete genomes. Further studies using shorter sections of UTR, looking in different T cell populations at different time points post-activation could provide more insight into the importance of these motifs. Our construct using the woodchuck post transcriptional element (WPRE) gave rise to expression equal to that of healthy controls, and furthermore, is widely used in gene therapy vectors for primary immune deficiency (Gaspar

et al., 2006; Marangoni et al., 2009; Kohn et al., 2020), presenting a clearer path to clinical translation.

We were motivated to optimise the transduction protocols by the high cost associated with performing gene editing at clinical scale. It has been reported that AAV6 is bound and precipitated by G3BP in FBS and HS (Denard et al., 2012) leading to a reduction in transduction efficiency. Here, we show that transducing T cells in media without HS leads to significantly improved gene editing at reduced viral doses, particularly when transduction is performed prior to nucleofection. Similar findings have recently been reported where T cells are cultured using FBS, supporting this approach (Wang et al., 2015; Rogers et al., 2021). In our study, the serum-free culture period had no effect on T cell phenotype markers. This approach could offer significant savings in viral production costs needed for each clinical product. However, further optimisations may be possible, either at the level of AAV transduction—such as human serum albumin (HSA) or polyvinyl alcohol (PVA) that have increased HSC transduction (Wang et al., 2017; Yang et al., 2020), or at the level of the proteins involved in DNA repair pathway by using small molecules (Li et al., 2017; Ma et al., 2018; Bischoff et al., 2020).

In the absence of SAP, XLP patient T cells exhibit deficits in several immune functions, including reduced sensitivity to RICD, insufficient T:B cell signalling and a lack a cytotoxicity against EBV-infected B cell targets. Here, we have shown for the first time that gene editing XLP patient T cells can restore SAP-dependent immune functions to the level of healthy control cells in assays that model these deficits. Although ultimately the number of corrected T cells needed for clinical benefit can only be fully determined in clinical trial, we are guided by patients that show somatic reversion of mutations. Such patients may have a SAP replete population as low as 5–8%, which leads to significantly prolonged survival without transplant (up to 40 years reported) suggesting a significant reduction in the risk of developing HLH and possibly malignancy, even with low levels of corrected cells (Palendira et al., 2012; Hoshino et al., 2018). We have also seen in our previous work using a lentiviral strategy for T cell correction in an XLP murine modal that 20–40% correction can lead to complete functional recovery (Panchal et al., 2018b). Drawing on these results, the rates of HDR reached in this work (30–50%) would be at a level that could restore a clinically relevant level of immune function in patients.

However, while a T cell product could treat several of the most severe clinical manifestations of XLP and offer a life-saving bridge to HSCT, studies have shown that NKT cells (which are absent in XLP patients) have an important role in tumour surveillance (Das et al., 2013; Weng et al., 2014), and SAP-replete NK cells can aid with clearing EBV-infected cells (Parolini et al., 2000). Here we have demonstrated across different gene editing platforms that it is possible to target the SAP gene locus with high levels of efficiency and specificity and promote corrective HDR in the context of primary T cells, leading to functional rescue. We are now transferring these technologies to HSCs with the aim of developing a gene edited HSC therapy for patients with XLP lacking a suitable donor for HSCT.

MATERIALS AND METHODS

Human Samples

Informed consent was obtained from all human subjects, including 6 unrelated XLP patients that have confirmed mutations in SH2D1A. Pt1—exon 2 deletion; Pt2—exon 2 deletion; Pt3—c.163c > t; Pt4—c.163C > T, p.Arg55X; Pt5 –exon 2-4 deletion; Pt6—exon1 deletion, which was not amenable to gene editing. Pt1 was used for testing HDR donor constructs and EBV-B cell cytotoxicity, Pts1-5 we were used in T:B co-culture assay and analysis of MFI in corrected cells, Pt1,2,3,5 corrected cells were used for RICD assay, Pt1-6 as untreated (UT) controls.

Cell Culture

Jurkat T cells and lymphoblastoid cell lines were cultured in RPMI containing 10% Foetal bovine serum (FBS) and 1% Penicillin-streptomycin (pen-strep) and passaged twice weekly using 1:10 dilution. HEK293Ts were maintained in DMEM supplemented with 10% FBS and 1% pen-strep and passaged twice weekly. Cells were washed with PBS and released from the culture flask with Trypsin-EDTA (all reagents ThermoFisher Scientific) before collection and neutralisation in complete DMEM and seeding back into culture flask at a 1:10 dilution.

T Cell Stimulation and Culture

Human peripheral blood mononuclear cells (PBMC) were harvested from whole blood using Ficoll-Paque density centrifugation (GE Healthcare). PBMC were cultured in TexMACS™ Medium (Miltneyi) supplemented with 5% human serum (Sigma) and 1% pen-strep. PBMC were stimulated with Human T-Activator CD3/CD28 Dynabeads (Gibco) at 1:1 ratio, in the presence of 100 U/ml IL-2 (Proleukin) in G-Rex® 24 plates (Wilson Wolf).

TALEN mRNA Synthesis

TALEN pairs were identified and constructed as previously described (Mussolino et al., 2011). Plasmid constructs were linearised at the 3′ of the expression cassette using restriction enzyme digest, then purified (Qiagen). mRNA was produced using the T7 mScript™ Standard mRNA Production System according to manufacturer's protocol. Briefly, linearised DNA template was transcribed, treated with DNAse1 and purified (RNeasy Mini-kit, Qiagen), before further reactions for addition of a polyA tail using the supplied enzymes, and Cap 1 structure capping. After a further purification, mRNA integrity was assessed using TapeStation (Agilent) and quantified on NanoDrop Microvolume Spectrophotometer. Left and right TALEN arms were combined at 1:1 ratio, aliquoted (7.5 µg each arm) and stored at −20°C.

CRISPR-Cas9 and CRISPR-Cas12a Design and RNP Assembly

Potential Cas9 and Cas12a target sites were identified using Benchling online software (www.benchling.com). Cas9 (Alt-R®

S. p. HiFi Cas9 Nuclease V3, IDT) and Cas12a (Alt-R® A.s. Cas12a (Cpf1) Ultra, IDT) proteins were complexed to their respective RNA guides (Cas9 synthetic gRNA, Merck; Cas12a crRNA, IDT) at a protein:guide molar ratio of 1:2.5 (Vakulskas et al., 2018) and 1:2 respectively, at room temperature for 10 min immediately prior to nucleofection.

T Cell Editing

T cells were cultured as described above. Dynabeads were removed using a DynaMag™-15 Magnet (Invitrogen), and cells washed in PBS and counted. For serum-free transduction prior to nucleofection, cells were washed again and resuspended in TexMACs media with pen-strep and IL-2 but without human serum. Nucleases were delivered into cells *via* electroporation using the Lonza 4D nucleofector, buffer P3, program EO-115 - typically 1.5-3 million cells in the 1 ml cuvette, or 0.5-1 million the 20 µl cuvette.

Determining Nuclease Efficiency Using TIDE/ICE Analysis

Genomic DNA was harvested from edited cells at day 5 post nucleofection (Qiagen). PCR amplicons were generated using the following primer pair for all nucleases (Fwd: TGGCCTCTGAGT AAACCGCA, Rev: AGCGAGGGATTGAGGCGAAA, product length: 718bp, Tm: 69°C) using Q5 polymerase (NEB). After PCR purification (Qiagen), amplicons were sanger sequenced (Eurofins genomics) using the forward primer. The resulting ab.1 file was the input to the online TIDE tool (Brinkman et al., 2014) (TALENs), or ICE software (Cas9/Cas12a, Synthego) which generated the % modification score.

Assessment of Cutting Activity at Predicted off-Target Loci

The top 10 most highly predicted loci for off-target nuclease activity of Cas9-3 and Cas12a-1 were identified by Benchling online software, while TALEN sites were predicted using PROGNOS (http://bao.rice.edu/cgi-bin/prognos/prognos.cgi) (Fine et al., 2014). PCR amplicons were designed to generate 150-200bp with the expected cut site in the centre. DNA from male healthy donor T cells edited with each nuclease platform (alongside untreated controls) was extracted at day 3 post nucleofection (Qiagen) and used as a template for on-target and off-target PCR reactions. The amplicons were then prepared for Illumina next generation sequencing by performing end repair, adapter ligation and bar coding using the NEBNext® UltraTM II DNA Library Prep Kit (NEB) according to the manufacturer's instructions. Libraries were quantified using the ddPCR™ Library Quantification Kit for Illumina TruSeq (Biorad), before sequencing using MiSeq Reagent Kit v2, 500cycles on an Illumina MiSeq platform (Illumina). The generated paired-end reads were processed using the command line version of the CRISPResso2 pipeline (Clement et al., 2019), obtained editing frequencies were compared to the untreated control samples using a one-sided Fisher's exact test. *, **, and *** indicate $p < 0.05$, $p < 0.01$, $p < 0.001$, respectively.

AAV6 Production and Purification

An AAV2 genome plasmid was kindly provided by Professor Amit Nathwani (UCL), and the pDGM6 (RRID: Addgene_110660) AAV6 packaging plasmid by the Russel lab (University of Washington). Homology arms were amplified from healthy control genomic DNA, both right and left 850bp long in all donor constructs (Right homology arm ChrX: 124345793-124346642, left homology arm ChrX:124346644-124347493). Constructs were cloned using HiFi assembly (NEB), allowing the left homology arm sequence to run directly into a codon optimised SAP cDNA without restriction enzyme sequences. SH2D1A 3′UTR sequences were also amplified from genomic DNA, all other elements were amplified from previously described lentiviral plasmids (Rivat et al., 2013; Panchal et al., 2018b).

AAV6 particles were produced in HEK293T cells *via* co-transfection of pDGM6 and the HDR genome plasmid using polyethylenimine (PEI, Sigma-Aldrich) 24 h after plating in complete DMEM media (DMEM (Gibco) 10%FBS, 1% penstrep). Media was replaced after 4 h and replaced again after 24 h with 2% DMEM. After a further 48 h, the cells were released into the media by scraping, before centrifuging to separate cell pellet and culture supernatant for processing separately. AAV6 particles were precipitated from cell media using ammonium sulphate on ice for 30min before collection by centrifugation and resuspended in TD buffer (1xPBS, 1mM $MgCl_2$, 2.5 mM KCl). The cell pellet was resuspended in TD and freeze/thaw ×4 in the presence of 0.5% deoxycholic acid (VWR), before centrifugation and harvesting the supernatant. Both fractions were incubated with benzonase 50 U/ml (Novagen) before combining prior to iodixanol-gradient centrifugation. AAV6 particles were harvested from the 40–60% gradient interface and stored at 4°C. Titration was performed using the QuickTiter™ AAV Quantitation Kit (Cell Biolabs).

In/out PCR and ddPCR

Primers for detection of G15 and G16 donors at the SH2D1A locus were designed using the NIH Primer-BLAST tool (Ye et al., 2012) (Fwd: TGGACAAAATGCTGAAAGGTGG; Rev: GTCTCTCTG CTGATCTTGCCG, amplicon length 953bp, Tm: 64°C). For ddPCR, the same primers were used with the addition of a probe CCAGGGCTCCGGAGTCAGGC (5′6-FAM, Internal ZEN and 3′ Iowa Black FQ, IDT). The genomic reference amplicon primers targeted albumin (Fwd GCTGTCATCTCT TGTGGGCTG, Rev CACAAATTTGGAAACAGAACAGGC ATT, amplicon length 1035bp) and probe CCTGTCATGCCC ACACAAATCTCTCC (5′HEX, Internal ZEN and 3′ Iowa Black FQ, IDT). Droplets were generated and analysed according to the manufacturer's instructions (QX200 system, Bio-Rad). The cycling conditions were (95°C 10 min initiation, 50x (94°C 1min, 60°C 30 s, 72°C 6 min) 98°C 10 min, store 12 C).

Detection of SAP Protein Expression by Intracellular Staining and Flow Cytometry

Intracellular staining of SAP was performed using the IntraPrep Permeabilizaton Reagent (Beckman). The primary antibody was

either mouse anti-human SH2D1A antibody (Clone 1C9 - Abnova Cat# H00004068-M01, RRID:AB_425532), or isotype control (Novus Cat# NB110-7082, RRID:AB_790752). The secondary was Goat anti-mouse polyclonal immunoglobulins RPE Goat F (ab')2 (Dako).

T Cell Phenotyping

The following T cell phenotyping panel was used to determine Tcm, Tem, Tscm and Tnaive populations - CD3 PE-Cy7 (Clone UCHT1), CD4 APC-Cy7 (Clone RPA-T4), CD8 APC (Clone RPA-T8), CD62L BV510 (Clone DREG-56), CD45RA BV650 (Clone HI100), CD95 BV711 (Clone DX2) (BD Biosciences).

T:B Cell Co-Culture Assay

PBMC were stimulated and edited as described and cultured for 7 days, before naïve T cell magnetic bead selection according to the manufacturers protocol (Miltneyi). T cells were plated at 5.0×10^4 cells/well in a round-bottom 96 well plate. Naïve B cells were isolated from tonsillar mononuclear cells (Milteny) and added to the T cells at a 1:1 ratio, in addition to 150 ng/ml staphylococcal enterotoxin B (SEB). Cells were co-cultured for 10 days, before the harvesting the supernatant for ELISA analysis of IL-21, IgG and IgM (eBioscience).

RICD Assay

RICD was assessed using an established protocol (Katz and Snow, 2013). Briefly, edited T cells were cultured for 10 days, before washing with PBS, counting and plating at 5.0×10^4 cells/well in triplicate in a 96 well plate in 100 µl media. Dilutions of OKT3 antibody (Tonbo Biosciences) were prepared at 2000 ng/ml, 200 ng/ml, and 20 ng/ml, and 100 µl added to the cells to make final concentrations of 1000 ng/ml, 100 ng/ml and 10 ng/ml in the wells.

After 24 h, PI was added (final concentration 1 µg/ml) before running a fixed volume of cell suspension from each well in flow cytometry (Cytoflex, Beckman). The data was analysed by determining the number of live cells (PI-) and comparing unstimulated controls to stimulated conditions by using the formula: % cell loss = [1-(# PI-restimulated cells/# PI- untreated cells)]x100.

Generation of EBV-B Cell Specific Cytotoxic T Lymphocytes (CTLs) and Chromium^{51}-Release Cytotoxicity Assay

Lymphoblastoid cell lines (LCL) were previously created from EBV-seropositive healthy donors using EBV B95.8 supernatant (Ricciardelli et al., 2014; Panchal et al., 2018b). EBV-CTLs were generated by stimulating PBMC with 40Gy-irradiated LCL, initially at 40:1 PBMC:LCL ratio in the absence of IL-2 for 10 days, then subsequently at weekly intervals at 4:1 ratio in the presence of 100 U/ml IL-2. CTLs were edited after the third stimulation and sorted by flow cytometry for corrected cells (GFP+) before a fourth restimulation. To test cytotoxic capability, LCLs were labelled with Chromium^{51} ($Na_2{}^{51}CrO_4$, Perkin Elmer) for 1 h at 37°C. An initial effector:target ratio was determined based on available effector cell numbers, then a twofold serial dilution was used to determine cytotoxic range. Labelled target cells were washed and added to the plate (5000/well) for 4 h at 37°C. 50 µl culture supernatant was taken and added to 150 µl scintillation fluid and incubated for 12 h at room temperature. Cr^{51} release was determined using a beta counter (PerkinElmer).

Statistical Analysis

Statistical analysis was performed with Graphpad Prism 9 software.

Data is presented as the mean ± SEM or SD as denoted in the figure legend. Statistical analysis was performed using Prism 9 software (Graphpad Software Inc.), details of statistical tests used, including all p values are indicated in the relevant figure legend.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/sra/PRJNA799193.

AUTHOR CONTRIBUTIONS

BH—experiment design and execution; manuscript preparation; NP—assistance with experiments; KC/SH—NGS run and analysis; TW—ddPCR run and analysis; CM/TC—TALEN design and genome editing expertise; AT—part funder and project supervision, CB—project PI.

FUNDING

This work has been made possible due to funding granted by the Wellcome Trust (102769/Z/13/Z and 217112/Z/19/Z) and Action Medical Research (GN2814). In addition, this work is supported by the NIHR GOSH BRC. The views expressed are those of the authors and not necessarily those of the Wellcome Trust, AMR, the NHS, the NIHR or the Department of Health.

ACKNOWLEDGMENTS

We would like to thank all colleagues and collaborators - particularly Christos Georgiadis, Ulrike Mock, Ida Ricciardelli, Ayad Eddaoudi, Stuart Tangye and Yenan Bryceson - that have given time, ideas or materials to this project. We would also like to thank all XLP patients and their families, and our funders - the Wellcome Trust, Action Medical Research and the NIHR GOSH BRC. The views expressed are those of the authors and not necessarily those of the Wellcome Trust, AMR, the NHS, the NIHR or the Department of Health.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2022.828489/full#supplementary-material

REFERENCES

Bischoff, N., Wimberger, S., Maresca, M., and Brakebusch, C. (2020). Improving Precise CRISPR Genome Editing by Small Molecules: Is There a Magic Potion? *Cells* 9, 1318. doi:10.3390/cells9051318

Booth, C., Gilmour, K. C., Veys, P., Gennery, A. R., Slatter, M. A., Chapel, H., et al. (2011). X-Linked Lymphoproliferative Disease Due to SAP/SH2D1A Deficiency: A Multicenter Study on the Manifestations, Management and Outcome of the Disease. *Blood* 117, 53–62. doi:10.1182/blood-2010-06-284935

Bravo, J. P. K., Liu, M.-S., Hibshman, G. N., Dangerfield, T. L., Jung, K., McCool, R. S., et al. (2022). Structural Basis for Mismatch Surveillance by CRISPR-Cas9. *Nature* 603, 343–347. doi:10.1038/s41586-022-04470-1

Brinkman, E. K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy Quantitative Assessment of Genome Editing by Sequence Trace Decomposition. *Nucleic Acids Res.* 42, e168. doi:10.1093/nar/gku936

Cannons, J. L., Tangye, S. G., and Schwartzberg, P. L. (2011). SLAM Family Receptors and SAP Adaptors in Immunity. *Annu. Rev. Immunol.* 29, 665–705. doi:10.1146/annurev-immunol-030409-101302

Clement, K., Rees, H., Canver, M. C., Gehrke, J. M., Farouni, R., Hsu, J. Y., et al. (2019). CRISPResso2 Provides Accurate and Rapid Genome Editing Sequence Analysis. *Nat. Biotechnol.* 37, 224–226. doi:10.1038/s41587-019-0032-3

Comte, D., Karampetsou, M. P., Humbel, M., and Tsokos, G. C. (2018). Signaling Lymphocyte Activation Molecule Family in Systemic Lupus Erythematosus. *Clin. Immunol.* 204, 57–63. doi:10.1016/j.clim.2018.11.001

Das, R., Bassiri, H., Guan, P., Wiener, S., Banerjee, P. P., Zhong, M. C., et al. (2013). The Adaptor Molecule SAP Plays Essential Roles During Invariant NKT Cell Cytotoxicity and Lytic Synapse Formation. *Blood* 121 (17), 3386–95. doi:10.1182/blood-2012-11-468868

Denard, J., Beley, C., Kotin, R., Lai-Kuen, R., Blot, S., Leh, H., et al. (2012). Human Galectin 3 Binding Protein Interacts with Recombinant Adeno-Associated Virus Type 6. *J. Virol.* 86, 6620–6631. doi:10.1128/jvi.00297-12

Denard, J., Rouillon, J., Leger, T., Garcia, C., Lambert, M. P., Griffith, G., et al. (2018). AAV-8 and AAV-9 Vectors Cooperate with Serum Proteins Differently Than AAV-1 and AAV-6. *Mol. Ther. - Methods Clin. Dev.* 10, 291–302. doi:10.1016/j.omtm.2018.08.001

Ding, S., Liang, Y., Zhao, M., Liang, G., Long, H., Zhao, S., et al. (2012). Decreased microRNA-142-3p/5p Expression Causes CD4+ T Cell Activation and B Cell Hyperstimulation in Systemic Lupus Erythematosus. *Arthritis Rheum.* 64, 2953–2963. doi:10.1002/art.34505

Dragovich, M. A., and Mor, A. (2018). The SLAM Family Receptors: Potential Therapeutic Targets for Inflammatory and Autoimmune Diseases. *Autoimmun. Rev.* 17, 674–682. doi:10.1016/j.autrev.2018.01.018

Duprè, L., Andolfi, G., Tangye, S. G., Clementi, R., Locatelli, F., Aricò, M., et al. (2005). SAP Controls the Cytolytic Activity of CD8+ T Cells against EBV-Infected Cells. *Blood* 105, 4383–4389. doi:10.1182/blood-2004-08-3269

Fan, W., Liang, D., Tang, Y., Qu, B., Cui, H., Luo, X., et al. (2012). Identification of microRNA-31 as a Novel Regulator Contributing to Impaired Interleukin-2 Production in T Cells from Patients with Systemic Lupus Erythematosus. *Arthritis Rheum.* 64, 3715–3725. doi:10.1002/art.34596

Fine, E. J., Cradick, T. J., Zhao, C. L., Lin, Y., and Bao, G. (2014). An Online Bioinformatics Tool Predicts Zinc finger and TALE Nuclease Off-Target Cleavage. *Nucleic Acids Res.* 42, e42. doi:10.1093/nar/gkt1326

Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., et al. (2020). CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β-Thalassemia. *N. Engl. J. Med.* 384, 252–260. doi:10.1056/nejmoa2031054

Gartshteyn, Y., Askanase, A. D., and Mor, A. (2021). SLAM Associated Protein Signaling in T Cells: Tilting the Balance toward Autoimmunity. *Front. Immunol.* 12, 654839. doi:10.3389/fimmu.2021.654839

Gaspar, H. B., Bjorkegren, E., Parsley, K., Gilmour, K. C., King, D., Sinclair, J., et al. (2006). Successful Reconstitution of Immunity in ADA-SCID by Stem Cell Gene Therapy Following Cessation of PEG-ADA and Use of Mild Preconditioning. *Mol. Ther.* 14, 505–513. doi:10.1016/j.ymthe.2006.06.007

Geng, L., Yang, J., Tang, X., Peng, H., Tian, J., Hu, Z., et al. (2021). SLAM/SAP Decreased Follicular Regulatory T Cells in Patients with Graves' Disease. *J. Immunol. Res.* 2021, 1–11. doi:10.1155/2021/5548463

Genovese, P., Schiroli, G., Escobar, G., Di Tomaso, T., Firrito, C., Calabria, A., et al. (2014). Targeted Genome Editing in Human Repopulating Haematopoietic Stem Cells. *Nature* 510, 235–240. doi:10.1038/nature13420

Gillmore, J. D., Gane, E., Taubel, J., Kao, J., Fontana, M., Maitland, M. L., et al. (2021). CRISPR-Cas9 *In Vivo* Gene Editing for Transthyretin Amyloidosis. *N. Engl. J. Med.* 385, 493–502. doi:10.1056/nejmoa2107454

Hale, J. S., Youngblood, B., Latner, D. R., Mohammed, A. U. R., Ye, L., Akondy, R. S., et al. (2013). Distinct Memory CD4+ T Cells with Commitment to T Follicular Helper- and T Helper 1-Cell Lineages Are Generated after Acute Viral Infection. *Immunity* 38, 805–817. doi:10.1016/j.immuni.2013.02.020

Heide, V. V. D., Möhnle, P., Rink, J., Briegel, J., and Kreth, S. (2016). Down-regulation of MicroRNA-31 in CD4+ T Cells Contributes to Immunosuppression in Human Sepsis by Promoting TH2 Skewing. *Anesthesiology* 124, 908–922. doi:10.1097/aln.0000000000001031

Hoshino, A., Yang, X., Tanita, K., Yoshida, K., Ono, T., Nishida, N., et al. (2018). Modification of Cellular and Humoral Immunity by Somatically Reverted T Cells in X-Linked Lymphoproliferative Syndrome Type 1. *J. Allergy Clin. Immun.* 143, 421–424. doi:10.1016/j.jaci.2018.07.044

Hubbard, N., Hagin, D., Sommer, K., Song, Y., Khan, I., Clough, C., et al. (2016). Targeted Gene Editing Restores Regulated CD40L Function in X-Linked Hyper-IgM Syndrome. *Blood* 127, 2513–2522. doi:10.1182/blood-2015-11-683235

Katz, G., and Snow, A. L. (2013). Immune Homeostasis, Methods and Protocols. *Methods Mol. Biol.* 979, 15–25. doi:10.1007/978-1-62703-290-2

Ko, M., An, J., and Rao, A. (2015). DNA Methylation and Hydroxymethylation in Hematologic Differentiation and Transformation. *Curr. Opin. Cel Biol.* 37, 91–101. doi:10.1016/j.ceb.2015.10.009

Kohn, D. B., Booth, C., Booth, C., Kang, E. M., Pai, S.-Y., Shaw, K. L., et al. (2020). Lentiviral Gene Therapy for X-Linked Chronic Granulomatous Disease. *Nat. Med.* 26, 200–206. doi:10.1038/s41591-019-0735-5

Kuo, C. Y., Long, J. D., Campo-Fernandez, B., de Oliveira, S., Cooper, A. R., Romero, Z., et al. (2018). Site-Specific Gene Editing of Human Hematopoietic Stem Cells for X-Linked Hyper-IgM Syndrome. *Cel Rep.* 23, 2606–2616. doi:10.1016/j.celrep.2018.04.103

Li, G., Zhang, X., Zhong, C., Mo, J., Quan, R., Yang, J., et al. (2017). Small Molecules Enhance CRISPR/Cas9-mediated Homology-Directed Genome Editing in Primary Cells. *Sci. Rep.* 7, 8943. doi:10.1038/s41598-017-09306-x

Liu, B., Zeng, L., Shao, Y., and Fu, R. (2021). Expression and Function of SLAMF6 in CD8+ T Lymphocytes of Patients with Severe Aplastic Anemia. *Cell Immunol.* 364, 104343. doi:10.1016/j.cellimm.2021.104343

Ma, C. S., Pittaluga, S., Avery, D. T., Hare, N. J., Maric, I., Klion, A. D., et al. (2006). Selective Generation of Functional Somatically Mutated IgM+CD27+, but Not Ig Isotype-Switched, Memory B Cells in X-Linked Lymphoproliferative Disease. *J. Clin. Invest.* 116, 322–333. doi:10.1172/jci25720

Ma, C. S., Wong, N., Rao, G., Avery, D. T., Torpy, J., Hambridge, T., et al. (2015). Monogenic Mutations Differentially Affect the Quantity and Quality of T Follicular Helper Cells in Patients With Human Primary Immunodeficiencies. *J. Allergy Clin. Immunol.* 136 (4), 993–1006.e1. doi:10.1016/j.jaci.2015.05.036

Ma, X., Chen, X., Jin, Y., Ge, W., Wang, W., Kong, L., et al. (2018). Small Molecules Promote CRISPR-Cpf1-Mediated Genome Editing in Human Pluripotent Stem Cells. *Nat. Commun.* 9, 1303. doi:10.1038/s41467-018-03760-5

Malaer, J. D., Marrufo, A. M., and Mathew, P. A. (2019). 2B4 (CD244, SLAMF4) and CS1 (CD319, SLAMF7) in Systemic Lupus Erythematosus and Cancer. *Clin. Immunol.* 204, 50–56. doi:10.1016/j.clim.2018.10.009

Marangoni, F., Bosticardo, M., Charrier, S., Draghici, E., Locci, M., Scaramuzza, S., et al. (2009). Evidence for Long-Term Efficacy and Safety of Gene Therapy for Wiskott-Aldrich Syndrome in Preclinical Models. *Mol. Ther.* 17, 1073–1082. doi:10.1038/mt.2009.31

Mehrle, S., Frank, S., Schmidt, J., Schmidt-Wolf, I. G., and Märten, A. (2005). SAP and SLAM Expression in Anti-CD3 Activated Lymphocytes Correlates with Cytotoxic Activity. *Immunol. Cel Biol* 83, 33–39. doi:10.1111/j.1440-1711.2005.01302.x10.1111/j.1440-1711.2004.01302.x

Miller, J. C., Patil, D. P., Xia, D. F., Paine, C. B., Fauser, F., Richards, H. W., et al. (2019). Enhancing Gene Editing Specificity by Attenuating DNA Cleavage Kinetics. *Nat. Biotechnol.* 37, 945–952. doi:10.1038/s41587-019-0186-z

Moffett, H. F., Cartwright, A. N. R., Kim, H.-J., Godec, J., Pyrdol, J., Äijö, T., et al. (2017). The microRNA miR-31 Inhibits CD8+ T Cell Function in Chronic Viral Infection. *Nat. Immunol.* 18, 791–799. doi:10.1038/ni.3755

Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T., and Cathomen, T. (2011). A Novel TALE Nuclease Scaffold Enables High

Genome Editing Activity in Combination with Low Toxicity. *Nucleic Acids Res.* 39, 9283–9293. doi:10.1093/nar/gkr597

Okamoto, S., Ji, H., Howie, D., Clarke, K., Gullo, C., Manning, S., et al. (2004). Expression of theSH2D1A Gene Is Regulated by a Combination of Transcriptional and Post-Transcriptional Mechanisms. *Eur. J. Immunol.* 34, 3176–3186. doi:10.1002/eji.200324755

Palendira, U., Low, C., Bell, A. I., Ma, C. S., Abbott, R. J. M., Phan, T. G., et al. (2012). Expansion of Somatically Reverted Memory CD8+ T Cells in Patients with X-Linked Lymphoproliferative Disease Caused by Selective Pressure from Epstein-Barr Virus. *J. Exp. Med.* 209, 913–924. doi:10.1084/jem.20112391

Panchal, N., Booth, C., Cannons, J. L., and Schwartzberg, P. L. (2018a). X-Linked Lymphoproliferative Disease Type 1: A Clinical and Molecular Perspective. *Front. Immunol.* 9, 666. doi:10.3389/fimmu.2018.00666

Panchal, N., Ghosh, S., and Booth, C. (2021). T Cell Gene Therapy to Treat Immunodeficiency. *Br. J. Haematol.* 192, 433–443. doi:10.1111/bjh.17070

Panchal, N., Houghton, B., Diez, B., Ghosh, S., Ricciardelli, I., Thrasher, A. J., et al. (2018b). Transfer of Gene Corrected T Cells Corrects Humoral and Cytotoxic Defects in X-Linked Lymphoproliferative Disease (XLP1). *J. Allergy Clin. Immun.* 142, 235–245. doi:10.1016/j.jaci.2018.02.053

Parolini, S., Bottino, C., Falco, M., Augugliaro, R., Giliani, S., Franceschini, R., et al. (2000). X-Linked Lymphoproliferative Disease. *J. Exp. Med.* 192, 337–346. doi:10.1084/jem.192.3.337

Pavel-Dinu, M., Wiebking, V., Dejene, B. T., Srifa, W., Mantri, S., Nicolas, C. E., et al. (2019). Gene Correction for SCID-X1 in Long-Term Hematopoietic Stem Cells. *Nat. Commun.* 10, 1634. doi:10.1038/s41467-019-09614-y

Qasim, W., Zhan, H., Samarasinghe, S., Adams, S., Amrolia, P., Stafford, S., et al. (2017). Molecular Remission of Infant B-ALL after Infusion of Universal TALEN Gene-Edited CAR T Cells. *Sci. Transl. Med.* 9, eaaj2013. doi:10.1126/scitranslmed.aaj2013

Rai, R., Romito, M., Rivers, E., Turchiano, G., Blattner, G., Vetharoy, W., et al. (2020). Targeted Gene Correction of Human Hematopoietic Stem Cells for the Treatment of Wiskott - Aldrich Syndrome. *Nat. Commun.* 11, 4034. doi:10.1038/s41467-020-17626-2

Ricciardelli, I., Blundell, M. P., Brewin, J., Thrasher, A., Pule, M., and Amrolia, P. J. (2014). Towards Gene Therapy for EBV-Associated Posttransplant Lymphoma with Genetically Modified EBV-Specific Cytotoxic T Cells. *Blood* 124, 2514–2522. doi:10.1182/blood-2014-01-553362

Ripamonti, A., Provasi, E., Lorenzo, M., De Simone, M., Ranzani, V., Vangelisti, S., et al. (2017). Repression of miR-31 by BCL6 Stabilizes the Helper Function of Human Follicular Helper T Cells. *Proc. Natl. Acad. Sci. U.S.A.* 114, 12797–12802. doi:10.1073/pnas.1705364114

Rivat, C., Booth, C., Alonso-Ferrero, M., Blundell, M., Sebire, N. J., Thrasher, A. J., et al. (2013). SAP Gene Transfer Restores Cellular and Humoral Immune Function in a Murine Model of X-Linked Lymphoproliferative Disease. *Blood* 121, 1073–1076. doi:10.1182/blood-2012-07-445858

Rogers, G. L., Huang, C., Clark, R. D. E., Seclén, E., Chen, H.-Y., and Cannon, P. M. (2021). Optimization of AAV6 Transduction Enhances Site-Specific Genome Editing of Primary Human Lymphocytes. *Mol. Ther. Methods Clin. Dev.* 23, 198–209. doi:10.1016/j.omtm.2021.09.003

Ruffo, E., Malacarne, V., Larsen, S. E., Das, R., Patrussi, L., Wülfing, C., et al. (2016). Inhibition of Diacylglycerol Kinase α Restores Restimulation-Induced Cell Death and Reduces Immunopathology in XLP-1. *Sci. Transl. Med.* 8, 321ra7. doi:10.1126/scitranslmed.aad1565

Ryan, D. E., Taussig, D., Steinfeld, I., Phadnis, S. M., Lunstad, B. D., Singh, M., et al. (2018). Improving CRISPR-Cas Specificity with Chemical Modifications in Single-Guide RNAs. *Nucleic Acids Res.* 46, 792–803. doi:10.1093/nar/gkx1199

Sayos, J., Wu, C., Morra, M., Wang, N., Zhang, X., Allen, D., et al. (1998). The X-Linked Lymphoproliferative-Disease Gene Product SAP Regulates Signals Induced through the Co-Receptor SLAM. *Nature* 395, 462–469. doi:10.1038/26683

Sharifi, R., Sinclair, J. C., Gilmour, K. C., Arkwright, P. D., Kinnon, C., Thrasher, A. J., et al. (2004). SAP Mediates Specific Cytotoxic T-Cell Functions in X-Linked Lymphoproliferative Disease. *Blood* 103, 3821–3827. doi:10.1182/blood-2003-09-3359

Song, L., Kauss, M. A., Kopin, E., Chandra, M., Ul-Hasan, T., Miller, E., et al. (2013). Optimizing the Transduction Efficiency of Capsid-Modified AAV6 Serotype Vectors in Primary Human Hematopoietic Stem Cells *In Vitro* and in a Xenograft Mouse Model *In Vivo*. *Cytotherapy* 15, 986–998. doi:10.1016/j.jcyt.2013.04.003

Stadtmauer, E. A., Fraietta, J. A., Davis, M. M., Cohen, A. D., Weber, K. L., Lancaster, E., et al. (2020). CRISPR-Engineered T Cells in Patients with Refractory Cancer. *Science* 367, eaba7365. doi:10.1126/science.aba7365

Sweeney, C. L., Pavel-Dinu, M., Choi, U., Brault, J., Liu, T., Koontz, S., et al. (2021). Correction of X-CGD Patient HSPCs by Targeted CYBB cDNA Insertion Using CRISPR/Cas9 with 53BP1 Inhibition for Enhanced Homology-Directed Repair. *Gene Ther.* 28, 373–390. doi:10.1038/s41434-021-00251-z

Tebas, P., Stein, D., Tang, W. W., Frank, I., Wang, S. Q., Lee, G., et al. (2014). Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV. *N. Engl. J. Med.* 370, 901–910. doi:10.1056/nejmoa1300662

Vakulskas, C. A., and Behlke, M. A. (2019). Evaluation and Reduction of CRISPR Off-Target Cleavage Events. *Nucleic Acid Ther.* 29, 167–174. doi:10.1089/nat.2019.0790

Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., et al. (2018). A High-Fidelity Cas9 Mutant Delivered as a Ribonucleoprotein Complex Enables Efficient Gene Editing in Human Hematopoietic Stem and Progenitor Cells. *Nat. Med.* 24, 1216–1224. doi:10.1038/s41591-018-0137-0

Villemure, J.-F., Savard, N., and Belmaaza, A. (2001). Promoter Suppression in Cultured Mammalian Cells Can Be Blocked by the Chicken β-Globin Chromatin Insulator 5′HS4 and Matrix/Scaffold Attachment Regions. *J. Mol. Biol.* 312, 963–974. doi:10.1006/jmbi.2001.5015

Wang, J., DeClercq, J. J., Hayward, S. B., Li, P. W.-L., Shivak, D. A., Gregory, P. D., et al. (2015). Highly Efficient Homology-Driven Genome Editing in Human T Cells by Combining Zinc-Finger Nuclease mRNA and AAV6 Donor Delivery. *Nucleic Acids Res.* 44, e30. doi:10.1093/nar/gkv1121

Wang, M., Sun, J., Crosby, A., Woodard, K., Hirsch, M. L., Samulski, R. J., and Li, C. (2017). Direct Interaction of Human Serum Proteins With AAV Virions to Enhance AAV Transduction: Immediate Impact on Clinical Applications. *Gene Ther.* 24 (1), 49–59. doi:10.1038/gt.2016.75

Wang, X., Chu, Y., Wang, W., and Yuan, W. (2016). mTORC Signaling in Hematopoiesis. *Int. J. Hematol.* 103, 510–518. doi:10.1007/s12185-016-1944-z

Weng, X., Liao, C. M., Bagchi, S., Cardell, S. L., Stein, P. L., and Wang, C. R. (2014). The Adaptor Protein SAP Regulates Type II NKT-Cell Development, Cytokine Production, and Cytotoxicity Against Lymphoma. *Eur. J. Immunol.* 44 (12), 3646–57. doi:10.1002/eji.201444848

Yang, H., Qing, K., Keeler, G. D., Yin, L., Mietzsch, M., Ling, C., et al. (2020). Enhanced Transduction of Human Hematopoietic Stem Cells by AAV6 Vectors: Implications in Gene Therapy and Genome Editing. *Mol. Ther. Nucleic Acids* 20, 451–458. doi:10.1016/j.omtn.2020.03.009

Yang, J., Geng, L., Ma, Y., Tang, X., Peng, H., Tian, J., et al. (2021). SLAMs Negatively Regulate IL-21 Production in Tfh-Like Cells from Allergic Rhinitis Patients. *J. Asthma Allergy* 14, 361–369. doi:10.2147/jaa.s291879

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T. L. (2012). Primer-BLAST: A Tool to Design Target-Specific Primers for Polymerase Chain Reaction. *Bmc Bioinformatics* 13, 134. doi:10.1186/1471-2105-13-134

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* 163, 759–771. doi:10.1016/j.cell.2015.09.038

Zhang, L., Zuris, J. A., Viswanathan, R., Edelstein, J. N., Turk, R., Thommandru, B., et al. (2021). AsCas12a Ultra Nuclease Facilitates the Rapid Generation of Therapeutic Cell Medicines. *Nat. Commun.* 12, 3908. doi:10.1038/s41467-021-24017-8

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

frontiers | Frontiers in Genome Editing

BRIEF RESEARCH REPORT
published: 05 July 2022
doi: 10.3389/fgeed.2022.932434

Optimizing CRISPR/Cas9 Editing of Repetitive Single Nucleotide Variants

*Inga Usher[1], Lorena Ligammari[1], Sara Ahrabi[2], Emily Hepburn[3], Calum Connolly[3], Gareth L. Bond[4], Adrienne M. Flanagan[1,5] and Lucia Cottone[1]**

[1]Department of Pathology (Research), UCL Cancer Institute, University College London, London, United Kingdom, [2]Department of Haematology, UCL Cancer Institute, University College London, London, United Kingdom, [3]UCL Medical School, University College London, London, United Kingdom, [4]Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, United Kingdom, [5]Department of Histopathology, Royal National Orthopaedic Hospital, London, United Kingdom

CRISPR/Cas9, base editors and prime editors comprise the contemporary genome editing toolbox. Many studies have optimized the use of CRISPR/Cas9, as the original CRISPR genome editing system, in substituting single nucleotides by homology directed repair (HDR), although this remains challenging. Studies describing modifications that improve editing efficiency fall short of isolating clonal cell lines or have not been validated for challenging loci or cell models. We present data from 95 transfections using a colony forming and an immortalized cell line comparing the effect on editing efficiency of donor template modifications, concentration of components, HDR enhancing agents and cold shock. We found that *in silico* predictions of guide RNA efficiency correlated poorly withactivity in cells. Using NGS and ddPCR we detected editing efficiencies of 5–12% in the transfected populations which fell to 1% on clonal cell line isolation. Our data demonstrate the variability of CRISPR efficiency by cell model, target locus and other factors. Successful genome editing requires a comparison of systems and modifications to develop the optimal protocol for the cell model and locus. We describe the steps in this process in a flowchart for those embarking on genome editing using any system and incorporate validated HDR-boosting modifications for those using CRISPR/Cas9.

Keywords: CRISPR, CRISPR/Cas9, genome editing, prime editing, homology directed repair (HDR), cell line, stem cells

OPEN ACCESS

Edited by:
Hua Xiang,
Institute of Microbiology CAS, China

Reviewed by:
Aftab Ahmad,
University of Agriculture, Pakistan
Yawen Fu,
Loma Linda University, United States

****Correspondence:***
Lucia Cottone
l.cottone@ucl.ac.uk

Specialty section:
This article was submitted to Genome Editing Tools and Mechanisms, a section of the journal Frontiers in Genome Editing

Received: *29 April 2022*
Accepted: *31 May 2022*
Published: *05 July 2022*

Citation:
Usher I, Ligammari L, Ahrabi S, Hepburn E, Connolly C, Bond GL, Flanagan AM and Cottone L (2022) Optimizing CRISPR/Cas9 Editing of Repetitive Single Nucleotide Variants. Front. Genome Ed. 4:932434. doi: 10.3389/fgeed.2022.932434

1 INTRODUCTION

CRISPR-associated protein 9 (Cas9) genome editing systems have revolutionized biological research. The ability to target and modify genetic loci offers the potential to investigate the functional consequences of variants discovered through genome wide association studies (Gallagher and Chen-Plotkin, 2018; Chen et al., 2022) and to correct the >75,000 known human pathogenic variants listed in Clinvar (Landrum and Kattman, 2018). The original CRISPR/Cas9 system has been available for more than a decade and extensive research has characterized how CRISPR/Cas9 works and how this can be optimized for maximal efficiency, with over 18,000 publications related to the term "CRISPR" listed on PubMed. CRISPR/Cas9 components are widely commercially available as recombinant components and plasmids. It is possible to achieve efficiencies of up to 80% in human cells when utilizing CRISPR/Cas9 to disrupt the DNA sequence and generate gene knock-outs (Guo et al., 2018). However, when using CRISPR/Cas9 to substitute nucleotides (knock-in), the efficiency is far lower, often in the order of 1%, although efficiencies of up to ~50% are reported following protocol modifications (Paquet et al., 2016; Kwart et al., 2017; Guo et al., 2018; Okamoto et al., 2019;

Maurissen and Woltjen, 2020) (reviewed in Liu et al., 2019a). More recently, it was demonstrated in yeast that retron-based CRISPR/Cas9 could boost multiplexing knock-in experiments, opening new possibilities for human cells (Zhao et al., 2022). Nevertheless nucleotide substitution by CRISPR/Cas9 relies on the non-dominant DNA repair pathway, homology directed repair (HDR). CRISPR/Cas9's reliance on DNA damage repair limits the efficiency of substituting nucleotides and can introduce off-target mutations.

Prime editing (PE) emerged in 2019 to address the low knock-in efficiency and propensity for off-target effects of CRISPR/Cas9 and is available in its latest iteration with PE4 and 5 (Anzalone et al., 2019; Chen et al., 2021). PE utilizes the guide RNA-directed "search" ability of CRISPR/Cas9 but builds on the accuracy of the "replace" ability by avoiding a double stranded break and reliance on cellular DNA repair. Base editors represent another alternative to CRISPR/Cas9 for single nucleotide substitution but are limited by their editing of bystander nucleotides making them unsuitable for repetitive nucleotides (Rees and Liu, 2018). Together, CRISPR/Cas9, PE and base editors comprise the main genome editing toolbox available to the contemporary researcher.

The efficiency of genome editing varies with the target locus, distance to the protospacer adjacent motif (PAM) and mismatch repair proficiency of the cell model (Chen et al., 2021; Ferreira da Silva et al., 2022). For example, the editing efficiency of CRISPR/Cas9 falls as distance from the PAM increases and stem cells show relative resistance to transfection compared to immortalized cell lines (Madsen and Semple, 2019). PE shows improved efficiency and reduced off-target changes compared to CRISPR/Cas9, however PE necessitates delivery as a plasmid, the design of more components (pegRNA spacer, extension and ngRNA for PE3b) and more extensive optimization compared to CRISPR/Cas9. PE efficiency is highly variable (up to 50-fold difference) depending on the genetic background of the cell model and component design (Anzalone et al., 2019). These factors make it likely that CRISPR/Cas9-mediated HDR will still be employed for repetitive target loci or cell models that are not amenable to PE, where a plasmid delivery is not desirable or where the rate of CRISPR/Cas9-mediated HDR is sufficiently high to avoid the need for extensive design and optimization processes. Selecting the right system for an experiment will depend on the target locus, intended substitution and cell model.

In this study we edited three single nucleotide loci implicated in the development of sarcomas using the CRISPR/Cas9 system. These loci are challenging to edit because they are repetitive, precluding the use of base editors, and distant from a PAM (>15 nucleotides), making editing by CRISPR/Cas9 or other Cas enzymes challenging. We applied several previously reported optimizing modifications across 95 transfections in induced pluripotent stem cells (iPSC), representing a colony forming model, and an adherent cell model to validate their utility in editing these challenging loci. Editing outcomes were characterized at each stage using contemporary technologies. While PE was not available at the time of our experiments, we have addressed how we would incorporate its use. Finally, we synthesize a flowchart that can be adapted to any subsequent CRISPR genome editing system.

2 MATERIALS AND METHODS

2.1 Induced Pluripotent Stem Cells

2.1.1 Cell Culture

The human episomal line of induced pluripotent stem cells (A18945, Gibco; Thermo Fisher Scientific, Inc, Waltham, MA, United States) was maintained in feeder-free culture on Geltrex (A1413202, Invitrogen, Thermo Fisher Scientific) and Essential 8 Flex (E8 Flex) medium (A28585, Gibco, Thermo Fisher Scientific) supplemented with 0.5% of Penicillin (10,000 U/ml). Cells were cultured at 37 °C in a humidified atmosphere with 5% CO2. Cells were passaged by incubation for 5 min at 37 °C with Dulbecco's phosphate buffered saline (DPBS)-EDTA 0.5 mM pH 8.00 (Invitrogen, Thermo Fisher Scientific, 14190250 and 15575020).

2.1.2 CRISPR/Cas9 Editing of iPSC

All CRISPR/Cas9 components were purchased through IDT. Single-stranded oligodeoxynucleotides (ssODN) were ordered as Alt-R™ HDR Donor Oligos.

1. The gRNA was prepared by duplexing Alt-R® CRISPR-Cas9 crRNA and Alt-R® CRISPR-Cas9 tracrRNA (IDT, 1072532).
2. The ribonucleoprotein (RNP) was formed using 0.78 µl Alt-R® CRISPR-Cas9 gRNA, 1.02 µl Alt-R® S. p. Cas9 Nuclease V3 (IDT, 1081058) and 1.2 µl PBS (total 3 µl).
3. Cells were detached, counted and transfected using the Lonza™ P3 Primary Cell 4D-Nucleofector™ X (Lonza, V4XP-3024) and electroporation program CA137.
4. To prepare the electroporation mixture, 0.5×10^6 cells were resuspended in 20 µl Lonza electroporation buffer, 1 µl of RNP complex and 0.5 µl Alt-R™ HDR Donor Oligo.
5. Following transfection, cells were recovered: (i) at 37°C, (ii) at 32 °C for 24 h then at 37 °C, (iii) with Alt-R® CRISPR-Cas9 HDR enhancer (IDT, 1081072) in E8 Flex without antibiotics for 24 h, (iv) with DMSO at 1% in E8 Flex without antibiotics for 24 h.

2.1.3 Colony Picking

For all stressful steps, such as single cell dissociation and colony picking, E8 Flex medium with RevitaCell™ was used for 2 h before and until colonies formed. In all other steps E8 Flex was used without RevitaCell™.

Cells were detached with Accutase® (Innovative Cell Technologies, Inc, San Diego, CA AT104), counted and 700–1,000 cells per dish were plated. Medium-sized colonies were picked using a P200 pipette. The aspirated colony was transferred to a 96 well plate and triturated 10 times. When cells were 50–70% confluent they were split 1:2, half were seeded for genomic DNA extraction and half for subculture or freezing.

2.1.4 Mirror Plates for Freezing and for DNA Extraction

Colonies expanded in 96 well plates were washed with 100 µl DPBS and detached with 30 µl of Accutase®. A mirror plate for

freezing was prepared containing 50 µl of 2X freezing medium (E8 Flex plus 20% DMSO). Cells were collected with 70 µl of E8 Flex and 50 µl were transferred to the mirror plate and stored at −80°C. The remaining 50 µl of suspension was kept for DNA extraction: the plate was spun at 1950 RCF for 30 min at 4 °C, the medium was removed, and the plate stored at −80°C. Upon thawing, 30–50 µl of Lucigen QuickExtract™ DNA Extraction Solution (LGC, Middlesex, United Kingdom, QE09050) was added, then triturated and heated at 65°C for 6 min then 98°C for 2 min and used directly for genotyping.

2.2 U-CH1 Chordoma Cell Line

2.2.1 Cell Culture

The human U-CH1 chordoma cell line (ATCC® CRL-3217™, www.chordomafoundation.org) was grown as previously described (Scheipl et al., 2016). Cell authentication was regularly performed by Short Tandem Repeat fingerprinting (Culture Collections, Public Health England, United Kingdom) (Supplementary Table 2).

2.2.2 CRISPR/Cas9 Editing

1. The gRNA was prepared by duplexing Alt-R® CRISPR-Cas9 crRNA and Alt-R® CRISPR-Cas9 tracrRNA (IDT, 1072532).
2. 3.9 µl Alt-R® CRISPR-Cas9 gRNA, 5.1 µl of 10 mg/ml (or 10 µg/µl = 50 µg = 300 pmol) Alt-R® S. p. Cas9 Nuclease V3 (IDT, 1081058) and 5.9 µl sterile DPBS (total 15 µl) were combined for the RNP.
3. Cells were detached, counted, and transfected using the Lonza Amaxa® Cell Line Nucleofector® Kit V (Lonza, Basel, Switzerland VCA-1003) using electroporation program A30. To prepare the electroporation mixture, 2×10^6 cells were resuspended in 60 µl Lonza electroporation buffer, with 10 µl of RNP complex (final Cas9 concentration ~100 pmol) and 3 µl of modified Alt-R™ HDR Donor Oligo.
4. Following transfection, the cells were recovered in medium without antibiotics.

2.2.3 Analysis of editing outcomes using digital droplet polymerase chain reaction (ddPCR)

A common primer set and probes for each allele were designed: a hexachlorofluorescein (HEX) probe for the parental allele (A/T) and fluorescein amidite (FAM) probe for the edited allele (G/C). ddPCR assays were designed using primer3plus (Untergasser et al., 2007) and the BioRad Droplet Digital™ PCR Applications Guide. ddPCR experiments were carried out using the BioRad QX200 ddPCR supermix for probes (no dUTP) workflow, Automated Droplet Generator, BioRad Automated Droplet Generation Oil for Probes (BioRad, Hercules, California, United States; #1864110), Eppendorf vapo. protect thermocycler and QX200 Automated Droplet Reader. Results were analyzed using the BioRad QuantaSoft™ Analysis Pro Software using rare event detection.

2.2.4 Flow Cytometry Activated Cell Sorting (FACS)

Transfected U-CH1 cells were recovered for 36 h before single cell sorting into collagen-coated 96 well plates using a BD FACS Aria Fusion Cell Sorter™ (Becton Dickinson, Franklin Lakes, New Jersey, United States) running FACSDiva Software version 6. Cells were incubated with TOPRO3+ (Garvey et al., 2016) before sorting to allow the exclusion of dead cells and the top 10% of ATTO-550 positive cells were selected.

2.2.5 Mirror Plate for DNA Extraction

Cells were washed with 100 µl DPBS and detached with 50 µl of Accutase® (Innovative Cell Technologies, AT104) at 37 °C for 10–15 min 50µl of medium was added, 50 µl were taken for genomic DNA extraction using Lucigen QuickExtract™ (LGC, QE09050) while the other 50 µl were replated for subculture.

2.3 Techniques Common to Both Cell Models

Regular testing was performed to exclude *mycoplasma* contamination using the EZ-PCR *Mycoplasma* Test Kit (K1-0210, Geneflow, Lichfield, Staffordshire, United Kingdom).

2.3.1 Genotyping Using Illumina MiSeq™ Next Generation Sequencing (NGS)

DNA was extracted using Zymo Column Extraction (Zymo Research, Irvine, California, United States , D3024) (for bulk transfections) or Lucigen QuickExtract™ DNA Extraction Solution (for picked colonies). PCR was performed with Kapa Hifi HotStart polymerase (for <500 base pair product) (Kapa Biosystems, Roche Molecular Systems, Inc, Pleasanton, California, United States , KR0370): 12.5 µl 2X KAPA HiFi HotStart ReadyMix, 0.75 µl 10 µM Forward Primer (with MiSeq™ adapter, **Supplementary Table S3**), 0.75 µl 10 µM Reverse Primer (with MiSeq™ adapter), 2 µl DNA and PCR-grade water up to 25 µl. 3 µl of DNA extracted in Lucigen was used for PCR. The PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN Ltd, Manchester, England). MiSeq™ was performed in-house.

2.3.2 Analysis of Editing Outcomes Using MiSeq™ Data

FASTQ files were analyzed using Cas Analyser (Park et al., 2017) with the following parameters: Nuclease type = single nuclease, comparison range (R) = 40, Minimum frequency (n) = 1 and no optional wild type marker. Rates of unedited and edited outcomes (NHEJ ± substitution and HDR) were calculated by number of reads containing outcome/total number of reads. Indels were changes in sequence length compared to the reference sequence.

We defined the outcomes as follows:

- Wild type/unedited (90–100% of reads match the reference sequence)
- (Homozygous) knock-out (90–100% of reads show indels)
- Heterozygous knock-in (40–60% of the reads match the reference and 40–60% show the knock-in)
- Homozygous knock-in (90–100% of reads show the knock-in)
- Combined/mixed repair (40–60% of the reads show indels and 40–60% match the reference sequence or show the knock-in).

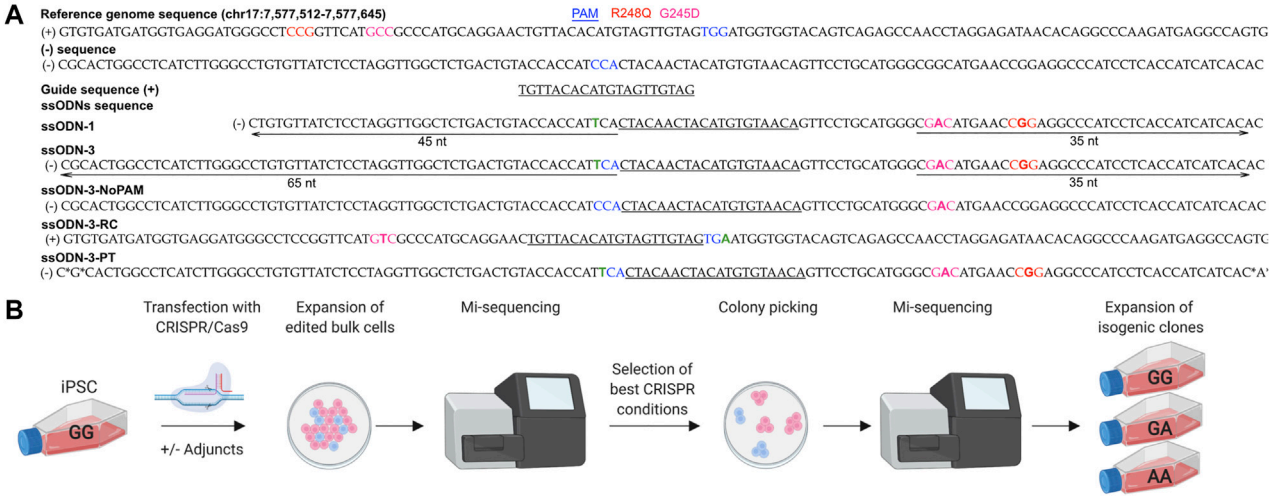


FIGURE 1 | Workflow for early characterization of bulk populations by MiSeq followed by clonal line isolation by colony picking. **(A)** Design of CRISPR/Cas9 components. Five ssODN designs were compared. *: phosphorothioate bonds. Green "T" indicates PAM modification to prevent cleavage of the donor. **(B)** iPSCs were transfected under different experimental conditions and screened by MiSeq NGS. Populations with the highest rates of HDR were selected for colony picking. Individual colonies showing accurate repair by MiSeq were expanded as isogenic lines.

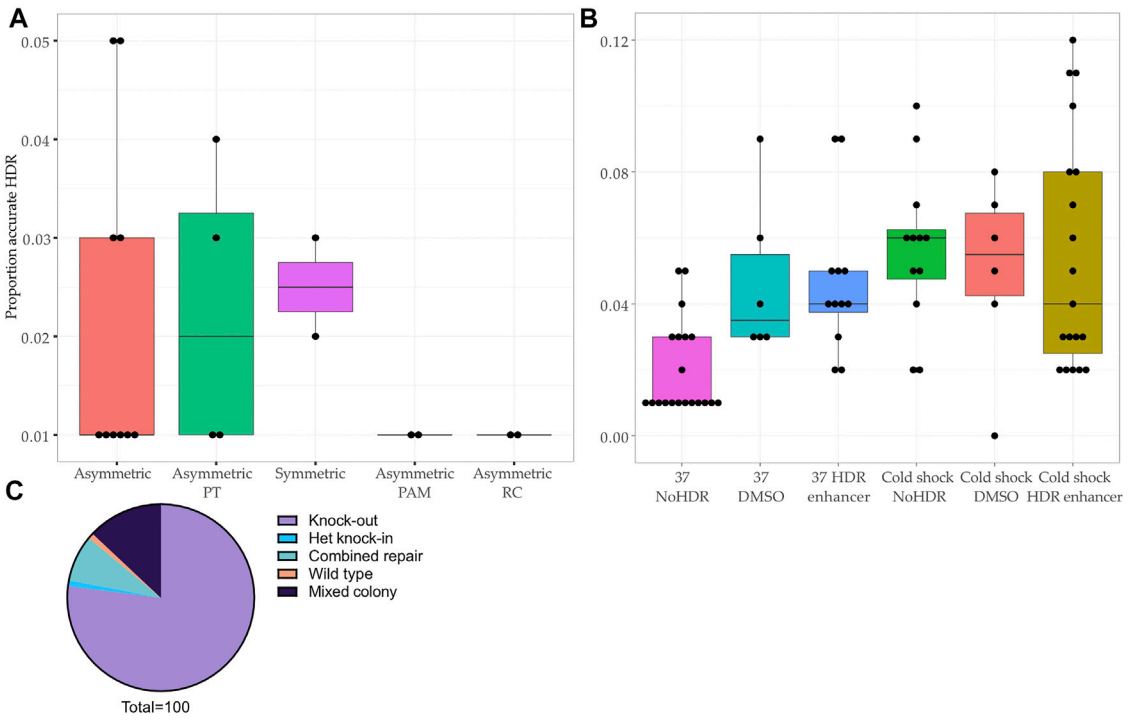


FIGURE 2 | Comparison of HDR efficiency associated with protocol modifications at the bulk population level and editing outcomes in clonal lines. **(A–B)** Proportion of reads showing accurate repair by HDR in bulk population DNA transfected with **(A)** different ssODNs (p-value <0.45, Kruskal–Wallis test) and **(B)** different experimental conditions (p-value < 0.01, Kruskal–Wallis test). Asymmetric PAM: asymmetric donor without blocking mutation in PAM. Asymmetric PT: addition of phosphorothioate nucleotides to asymmetric ssODN. Asymmetric RC: reverse complement of asymmetric ssODN. HDR: addition of Alt-R™ HDR Enhancer after transfection. DMSO: addition of DMSO after transfection. NoHDR: no HDR enhancer or DMSO after transfection. **(C)** Editing outcomes in 100 colonies picked from two transfections showing population HDR rates of 11 and 12%.

2.3.3 Genotyping Using Sanger Sequencing

PCR was performed: 12.5 µl AmpliTaq Gold™ 360 Master Mix (Thermo Fisher Scientific, 4398881), 0.5 µl 10 µM forward-reverse primer mix (**Supplementary Table S3**), 10 µl water plus 2 µl Lucigen QuickExtract™ DNA. PCR products were cleaned using the ExoSAP-IT™ Express PCR Product Cleanup Reagent (Applied Biosystems, Thermo Fisher Scientific, 15563677) and sent for Sanger sequencing (Source BioScience, Nottingham, United Kingdom).

2.3.4 Genotyping Using TaqMan™ qPCR

TaqMan™ genotyping was performed for rs2305089 (Applied Biosystems, 4351379): 5 µl TaqMan™ genotyping mastermix, 0.5 µl Taqman™ primer/probe mix (C__11223433_10), 3.5 µl water, 1 µl DNA. Results were analysed using the Genotyping application on the Thermo Fisher Connect™ cloud.

2.4 Data Analysis

All analysis and statistics were performed using R version 4.0.5 (2021–03–31) and GraphPad Prism version 8.0.0 for Windows (GraphPad Software, San Diego, California United States , www.graphpad.com). Cartoons were created with Biorender.com.

3 RESULTS

3.1 Editing of *TP53* SNVs in iPSC or Other Colony Forming Cell Model

We tested the ability of CRISPR/Cas9 modifications to introduce the germline pathogenic G245D and R248Q variants in *TP53* which lie in proximity, distant from a PAM and in repetitive sequences, into iPSC (**Figure 1A**). iPSCs tolerate transfection and single cell sorting poorly and are expensive to maintain in culture making it important to characterize editing outcomes early in the workflow (**Figure 1B**). *TP53* is expressed in iPSCs making it likely that the chromatin will be open allowing access of the RNP (Liu et al., 2019b).

3.2 Preliminary Checks of Cell Model

iPSCs accumulate genetic alterations during cell culture including SNVs at G245D, R248Q and other loci in *TP53* (Amps et al., 2011; Merkle et al., 2017; Liu et al., 2019b). We therefore first ensured that a genetically pure population, free of *TP53* SNVs, was utilized for transfection by picking 20 clonal sublines grown from the parental iPSC line; and checked five by Sanger sequencing which were found to be free of SNVs in the region around the G245D and R248Q loci. One of these sublines was selected for transfection (**Supplementary Figure S1**).

3.3 Design of Components Targeting *TP53* in iPSCs

We designed 6 guide RNAs (gRNAs) for CRISPR/Cas9 using online tools E-CRISP Heigwer et al., 2014, CHOPCHOP Labun et al., 2019 and the IDT Alt-R™ CRISPR HDR Design Tool (https://eu.idtdna.com/pages/tools/alt-r-crispr-hdr-design-tool) (**Supplementary Figure S2B**). All gRNAs were comparable in their likelihood of on- and off-target effects when assessed *in silico*. When tested in cells, only one gRNA generated a double strand break (DSB) on Sanger sequencing (**Supplementary Figure S2C** and **Figure 1A**) and did so at >15 nucleotides from the SNVs limiting our choice of gRNA and editing efficiency (Paquet et al., 2016; Kwart et al., 2017) (**Figure 1A** and **Supplementary Figure S2B**).

Designing PE components is more complex. PegRNA design is a major factor determining editing efficiency and several primer binding sites (PBS) and reverse transcriptase (RT) template combinations are possible when designing the pegRNA but only a fraction of these will achieve optimal efficiency (Anzalone et al., 2019). Various tools have been developed to aid design including PEGfinder (Chow et al., 2020), PE-designer (Hwang et al., 2021) and PrimeDesign (Hsu et al., 2021). For the G245D locus in *TP53*, 19 pegRNA designs were generated using PrimeDesign which could be combined with >10 PE gRNAs (Hsu et al., 2021) (data not shown). However, the low tolerance of iPSC to multiple transfections discouraged the use of the PE system for this study.

3.4 Optimizing the HDR Efficiency of CRISPR/Cas9 in iPSCs

HDR-enhancing modifications to the CRISPR/Cas9 protocol have been extensively investigated (**Supplementary Table S1**). As we could not modify the cut-to-mutation distance, we tested if modifying ssODN design or post-transfection experimental conditions would improve the rate of HDR. We performed 75 individual transfections of iPSCs (**Figures 1A,B**) (**Supplementary Table S1**) which tolerated electroporation with 40–50% viability after transfection, and compared editing outcomes in the bulk populations using MiSeq (reviewed in Sledzinski et al., 2020).

The editing efficiency was variable, ranging from 0 to 12% for the knock-in without indels (accurate HDR) (**Figures 2A,B**). We corroborated previous reports (Richardson et al., 2016; Liang et al., 2017; O'Brien et al., 2019) that asymmetry of the homology arms improves accurate HDR (range 1–5%, mean 2.0%, 10 samples) as did addition of phosphorothioated nucleotides (range 1–4%, mean 2.3%, 4 samples) (Papaioannou et al., 2009; Gutierrez-Triana et al., 2018) and arms of equal length (range 2–3%, mean 2.5%, 2 samples) (**Figure 2A**). The introduction of a silent mutation of the PAM (Paquet et al., 2016; Okamoto et al., 2019), or the reverse complement (RC) of the asymmetric design (Richardson et al., 2016) (**Figure 1A**, **Supplementary Table S1** and **Supplementary Table S3**) did not improve the HDR efficiency (**Figure 2A**), however only two experiments were performed for these conditions.

Next, using the asymmetric ssODN, we proceeded to modify post-transfection conditions by adding 1% dimethyl sulfoxide (DMSO) or the IDT Alt-R™ HDR Enhancer (Stratigopoulos et al., 2018) and culturing cells at 32°C (cold shock) (Guo et al., 2018) (**Supplementary Table S1**). The most effective protocol included cold shock and Alt-R™ HDR Enhancer, a finding consistent with previous studies (Skarnes et al., 2019; Di Stazio et al., 2021) (**Figure 2B**). We confirmed that DMSO increases HDR, making it a cost-effective alternative to

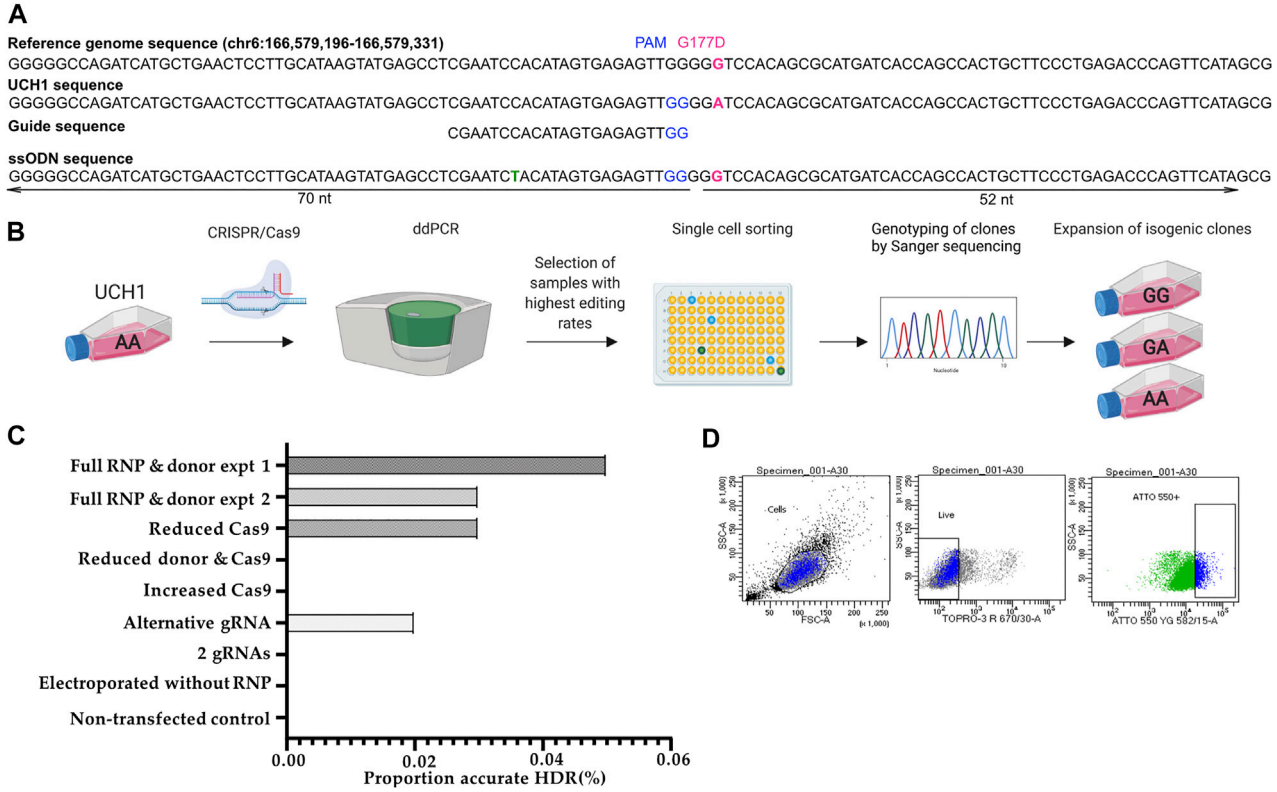


FIGURE 3 | Workflow for early single cell sorting and high throughput genotyping. **(A)** Design of CRISPR/Cas9 components targeting the G177D SNV in *TBXT*. **(B)** After transfection, U-CH1 bulk populations were screened by ddPCR. Populations showing HDR (2–10%) were single cell sorted using FACS. DNA was extracted from expanded clonal cell populations by establishing a "mirror plate" and used directly for genotyping by TaqMan™ qPCR or ddPCR **(C)** Bar plot of proportion of reads showing accurate HDR when concentrations of CRISPR/Cas9 components are varied. **(D)** Dot plots showing the gating strategy for sorting U-CH1 cells based on TOPRO3 and ATTO-550 fluorescence.

commercial HDR enhancers (Stratigopoulos et al., 2018) (**Figure 2B**).

3.5 Isolating Cell Lines Reduces the Editing Efficiency Observed in the Bulk Population by NGS

We chose two transfected populations which showed a promising rate of accurate HDR (11 and 12%) for single cell line isolation (**Figure 1B** and **Figure 2C**). After colony picking and expansion, most colonies were repaired by NHEJ (77/100, 77%) or showed a mixed repair (13/100, 13%). The final editing efficiency was 1% despite the initially promising rate of HDR at the level of the bulk population.

Finally, quality assurance of the clonal lines was undertaken to assess the homogeneity/purity of the population. Using MiSeq we established that the clonal lines were pure populations free of off-target alterations in the ~200 base pairs surrounding the variants. No off-target sites were predicted by the gRNA design tools.

Our data show that *in silico* predictions of gRNA efficiency correlated poorly with activity in our iPSC. Rates of HDR vary between transfections and can be boosted with asymmetric donors, with or without PT modifications, and HDR-enhancing modifications. The isolation of clonal cell lines from bulk populations resulted in a significant attrition of HDR efficiency.

3.6 Editing of *TBXT* in the U-CH1 Immortalized Cancer Cell Line

Many cell models grow as adherent cultures that tolerate single cell sorting and are relatively cheap to culture. The emphasis for these models is high throughput generation and genotyping of cell lines. U-CH1 is a cell model of the rare bone cancer, chordoma, and is associated with the G177D SNV in *TBXT*. We employed the U-CH1 chordoma cell line, which expresses *TBXT* at high levels and likely to be in euchromatin, to investigate the functional impact of the G177D SNV (Kelley et al., 2010; Pillay et al., 2012).

3.7 Design of CRISPR/Cas9 Components for the U-CH1 Chordoma Cell Line

We ensured a pure population free of SNVs in the region surrounding the G177D variant using MiSeq to avoid the time taken for single cell sorting (**Supplementary Figure S4A**).

Four candidate gRNAs were designed and assessed using the IDT CRISPR-Cas9 guide RNA design checker. Of the four

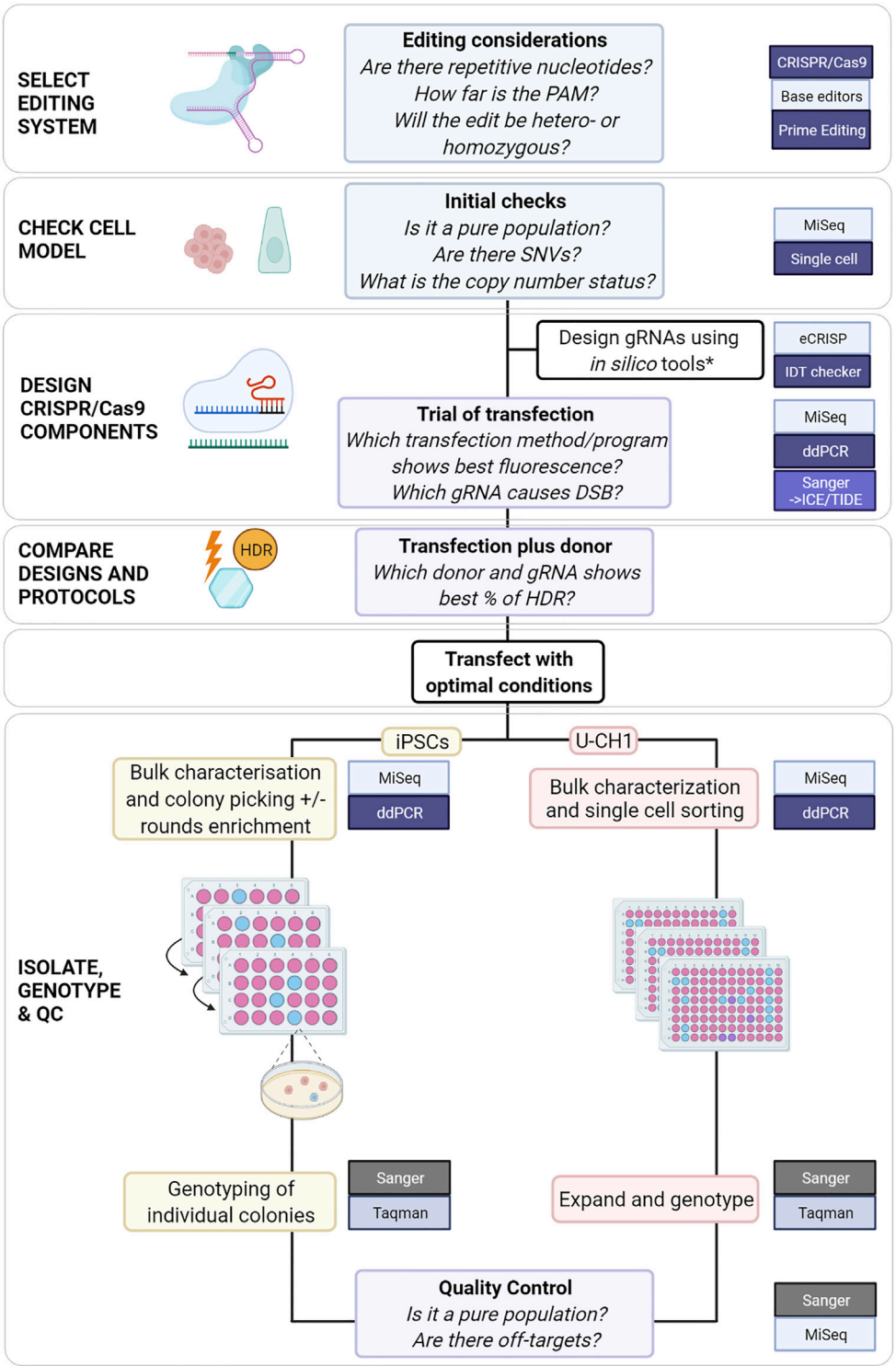


FIGURE 4 | Proposed comprehensive flowchart for editing stem cells or immortalized cell lines using CRISPR/Cas9. *(Heigwer et al., 2014; Labun et al., 2019; Sledzinski, Nowaczyk and Olejniczak, 2020) and (Schubert et al., 2021) for HDR-specific ssODN design.

gRNAs predicted to be effective *in silico*, two caused a DSB when tested in cells (**Figures 3A,B** and **Supplementary Figure S4B–C**).

We tested whether varying Cas9 (2.05 µl (17 pmol final concentration) versus 10.2 µl (404 pmol final concentration)), donor template concentration or combining gRNAs would

improve the editing efficiency as measured by MiSeq (Lin et al., 2014). We established that the RNP complexed with a single gRNA, at recommended concentrations, showed the best performance (**Figure 3C**).

3.8 Screening of Transfected Bulk Populations Using Digital Droplet PCR

After transfection we screened the bulk population for HDR using digital droplet PCR (ddPCR) as a faster alternative to MiSeq before proceeding to single cell sorting, allowing us to maintain the cells in culture during screening as U-CH1 cells grow slowly (**Figure 3B**). The partitioning technology of ddPCR enables the detection of rare knock-in events at a frequency as low as 0.1% (Miyaoka et al., 2014, Miyaoka et al., 2018). Screening of 10 transfected populations by ddPCR showed an editing efficiency of 2–10%. The transfected cells tolerated single cell sorting by FACS 24–36 h after transfection (**Figure 3B**) then required 3–5 weeks to expand sufficiently for subculture into a "mirror" plate for genotyping. The Lucigen one-step DNA extraction protocol was utilized, allowing direct input of DNA into TaqMan™ qPCR or ddPCR for high throughput genotyping.

3.9 High Throughput Screening of Hundreds of U-CH1 Clones

We screened ~500 transfected clonal lines and isolated seven cell lines which were free of indels: four wild type, two heterozygous knock-ins and one homozygous knock-in, giving an overall HDR efficiency of 0.6% (**Figure 3D**).

The relationship between HDR efficiency and cut-to-mutation distance (Paquet et al., 2016; Kwart et al., 2017) was again observed: the silent blocking mutation introduced into the ssODN was successfully edited more frequently than the G177D SNV and in a homozygous fashion, in contrast to the G177D SNV which was only edited in one allele.

Finally, we checked the isolated clones by sequencing 1,000 bp around the site of the edit by Sanger sequencing: we confirmed the presence of the introduced edit and ensured no off-target alterations were present (**Supplementary Figure S6**). No other off target sites were predicted by the design tools.

In summary, we show the results of screening hundreds of potential edited clones using high throughput technologies but highlight a significant attrition in efficiency between bulk populations and single cell line isolation. Varying the concentration of CRISPR/Cas9 components did not improve editing efficiency in U-CH1.

4 DISCUSSION

When planning a genome editing experiment, the methods available to the scientist have expanded substantially over the past decade. Each system, CRISPR/Cas9, base editors and PE, offers solutions for introducing different alterations: CRISPR/Cas9 for knock-outs and knock-ins, base editors for substitutions in non-repetitive nucleotides and PE for substitutions, transversions and indels. As genome editing systems require increasingly complex components and modified systems emerge, preliminary experiments become increasingly important for informing final experimental design (Anzalone et al., 2019). We propose a flowchart that incorporates critical steps and tools utilized at each stage to edit colony-forming cells or adherent cell lines (**Figure 4**).

Bioinformatic platforms are indispensable for designing CRISPR/Cas9 components but the accuracy of predictions can be affected by the chromatin organization of the target locus and varies by cell model. This was confirmed by our results that show that *in silico* predictions of gRNAs efficiency did not show the same performance in cells. Even when experiments are repeated there was several-fold variation in the HDR efficiency, potentially related to factors such as inherited genetic variability and cell cycle phase, highlighting the importance of characterizing bulk populations before single cell line isolation.

Cut-to-mutation distance is a major factor in determining CRISPR/Cas9 HDR efficiency and zygosity of the resulting edit (Paquet et al., 2016; Kwart et al., 2017). This was confirmed in our experiments with the frequent homozygous knock-in of the silent PAM mutation, close to the cut site, compared to the heterozygous knock-in of the target locus, distant from the cut site. This may have accounted for the successful homozygous knock-in in U-CH1 where the mutation is closer to the cut site than in iPSCs. If a homozygous knock-in is required, gRNAs that cut close to the target locus are required. If these gRNAs are not effective in preliminary testing, it is worthwhile considering Cas9 nucleases with different PAM requirements or using PE. We were able to validate the beneficial effect of an asymmetric donor, HDR enhancer and cold shock, the combination of which increased the low baseline HDR efficiency several fold. Asymmetric donors are thought to influence annealing and release of strands being repaired (O'Brien et al., 2019) while HDR enhancers block NHEJ (Pinder, Maurissen, reviewed in Bischoff) to favor HDR and cold shock is thought to affect G2/M transition or persistence of the RNP (Guo, Maurissen). Tools that aid optimal design of ssODNs may be employed (O'Brien et al., 2019; Schubert et al., 2021).

Despite these optimizations boosting the HDR efficiency to a promising mean rate of 11.5% in the bulk population, a single cell line was isolated from transfected iPSCs with a similar picture in U-CH1. This may be related to the low overall editing efficiency, consistent over repeated transfections, compared to previous optimization studies. Although some studies showed impressive rates of HDR, it is important to note that they instead utilized a bulk population-based read out without isolating cell lines (Zhang et al., 2017; Skarnes et al., 2019; Di Stazio et al., 2021) or they edited the more robust HEK293 cell model (Richardson et al., 2016; Liang et al., 2017; Di Stazio et al., 2021). The significant attrition of editing efficiency is an important consideration when modest editing efficiencies are achieved.

Characterization of the factors affecting the efficiency of PE is less advanced but initial studies suggest optimizing the

melting temperature of the PBS, using a dual pegRNA strategy (Lin et al., 2021) and disrupting the action of the mismatch repair pathway (Ferreira da Silva et al., 2022). Analysis of PE outcomes using MiSeq data is a newer concept and one tool is available, PE-Analyzer, with more likely to follow (Hwang et al., 2021). Sanger sequencing represents a cheap alternative for analyzing the composition of editing outcomes in bulk populations. Tools include Synthego's ICE (Inference of CRISPR Edits) (Hsiau et al., 2019) and Tracking Indels by Decomposition (TIDE) (Brinkman et al., 2014). Sanger Sequencing tools designed for CRISPR/Cas9 have been applied to PE experiments highlighting their versatility (Ferreira da Silva et al., 2022).

Given the challenges of editing the loci described in our work, a more complex system such as PE might be worth exploring in future studies. PE could be compared to CRISPR/Cas9 using the methods described to determine whether the trade-off between the relative simplicity of CRISPR/Cas9 is balanced by increased efficiency. Compared to the extensive characterization of CRISPR/Cas9, the optimization of PE is a growing field. A significant advance is the manipulation of mismatch repair pathways in PE (Chen et al., 2021; Ferreira da Silva et al., 2022). At present pegRNA design is the major determinant of PE efficiency (Anzalone et al., 2019). Finding the most efficiency combination of PBS and RT designs is an important preliminary step.

For those who will attempt HDR by CRISPR/Cas9 based on its relative simplicity and the robustness of the CRISPR/Cas9 recombinant components compared to PE, if an acceptable efficiency is attained after clonal isolation, the modifications we have validated may be employed to boost efficiency.

5 CONCLUSION

For SNVs for which a cut site can be generated at < 15 nucleotides, it may be simpler and faster to use CRISPR/Cas9 with protocol modifications to achieve HDR. For more challenging loci, such as those presented in this study, PE could be considered but would require more extensive optimization. For all genome editing systems the efficiency varies with cell model and target locus amongst other factors. Preliminary testing will inform the choice of system and protocol modifications. We propose a flowchart which could be used to guide the planning of CRISPR/Cas9 experiments to edit SNVs (**Figure 4**).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found online at https://www.ebi.ac.uk/ena with accession number PRJEB52736.

AUTHOR CONTRIBUTIONS

Conceptualization, IU, SA, GB, AF, and LC; methodology, IU, SA, LL, and LC; formal analysis, IU, EH, and CC; investigation, IU, LL, and LC; writing—original draft preparation, IU and LC; writing—review and editing, LL, SA, EH, CC, AF, and LC; visualization, IU and LC; supervision, LC and AF; funding acquisition, IU, GB, AF, and LC. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was funded by the Tom Prince Trust (AF), The Pathological Society of Great Britain and Ireland (grant number TSGS041903) (IU and AF), Cancer Research UK (University College London (CRUK UCL) Centre Award C416/A25145) (AF, LC, and GB), the Royal National Orthopaedic Hospital NHS Trust R&D Department, Bone Cancer Research Trust, Chordoma Foundation. AF was supported by the National Institute for Health Research, the University College London Hospitals Biomedical Research Centre, and the Cancer Research United Kingdom University College London Experimental Cancer Medicine Centre. IU was supported by Chordoma United Kingdom. SA was supported by the Hardy Keinan fellowship.

ACKNOWLEDGMENTS

We acknowledge Sunniyat Rahman, Samuel Weeks, Daniel Wetterskog, Yang Li, Iben Lyskjaer and Florian Merkle for their invaluable advice in setting up our protocols.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2022.932434/full#supplementary-material

REFERENCES

Amps, K., Andrews, P. W., Anyfantis, G., Armstrong, L., Avery, S., Baharvand, H., et al. (2011). Screening Ethnically Diverse Human Embryonic Stem Cells Identifies a Chromosome 20 minimal Amplicon Conferring Growth Advantage. *Nat. Biotechnol.* 29 (12), 1132–1144. doi:10.1038/NBT.2051

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace Genome Editing without Double-Strand Breaks or Donor DNA. *Nature* 576 (7785), 149–157. doi:10.1038/S41586-019-1711-4

Bio-Rad Laboratories, Inc. (2022). Bulletin 6407 Ver B, [Online]. Available at: http://www.biorad.com/webroot/web/pdf/lsr/literature/Bulletin_6407.pdf (Accessed February 01, 2022).

Brinkman, E. K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy Quantitative Assessment of Genome Editing by Sequence Trace Decomposition. *Nucleic Acids Res.* 42 (22), e168. doi:10.1093/nar/gku936

Chen, P. J., Hussmann, J. A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.-F., et al. (2021). Enhanced Prime Editing Systems by Manipulating Cellular Determinants of Editing Outcomes. *Cell.* 184 (22), 5635–5652.e29. doi:10.1016/J.CELL.2021.09.018

Chen, R., Yang, Z., Liu, J., Cai, X., Huo, Y., Zhang, Z., et al. (2022). Functional Genomic Analysis Delineates Regulatory Mechanisms of GWAS-Identified Bipolar Disorder Risk Variants. *Genome Med.* 14 (1), 53. doi:10.1186/S13073-022-01057-3

Chow, R. D., Chen, J. S., Shen, J., and Chen, S. (2020) 'pegFinder: A pegRNA Designer for CRISPR Prime Editing', bioRxiv, p. 2020.05.06.081612. doi:10.1101/2020.05.06.081612

Di Stazio, M., Foschi, N., Athanasakis, E., Gasparini, P., and d'Adamo, A. P. (2021). Systematic Analysis of Factors that Improve Homologous Direct Repair (HDR) Efficiency in CRISPR/Cas9 Technique. *PLoS ONE* 16 (3), e0247603. March 2021. doi:10.1371/JOURNAL.PONE.0247603

Ferreira da Silva, J., Oliveira, G. P., Arasa-Verge, E. A., Kagiou, C., Moretton, A., Timelthaler, G., et al. (2022). Prime Editing Efficiency and Fidelity Are Enhanced in the Absence of Mismatch Repair. *Nat. Commun.* 13 (1). doi:10.1038/S41467-022-28442-1

Gallagher, M. D., and Chen-Plotkin, A. S. (2018). The Post-GWAS Era: From Association to Function. *Am. J. Hum. Genet.* 102 (5), 717–730. doi:10.1016/J.AJHG.2018.04.002

Garvey, C. M., Spiller, E., Lindsay, D., Chiang, C.-T., Choi, N. C., Agus, D. B., et al. (2016). A High-Content Image-Based Method for Quantitatively Studying Context-dependent Cell Population Dynamics. *Sci. Rep.* 6, 29752. doi:10.1038/SREP29752

Guo, Q., Mintier, G., Ma-Edmonds, M., Storton, D., Wang, X., Xiao, X., et al. (2018). 'Cold Shock' Increases the Frequency of Homology Directed Repair Gene Editing in Induced Pluripotent Stem Cells. *Sci. Rep.* 8 (1), 2080. doi:10.1038/S41598-018-20358-5

Gutierrez-Triana, J. A., Tavhelidse, T., Thumberger, T., Thomas, I., Wittbrodt, B., Kellner, T., et al. (2018). Efficient Single-Copy HDR by 5' Modified Long dsDNA Donors. *eLife* 7, 39468. doi:10.7554/ELIFE.39468

Heigwer, F., Kerr, G., and Boutros, M. (2014). E-CRISP: Fast CRISPR Target Site Identification. *Nat. Methods* 11 (2), 122–123. doi:10.1038/NMETH.2812

Hsiau, T., Conant, D., Rossi, N., Maures, T., Waite, K., Yang, J., et al. (2019, Inference of CRISPR Edits from Sanger Trace Data), bioRxiv, p. 251082. doi:10.1101/251082

Hsu, J. Y., Grünewald, J., Szalay, R., Shih, J., Anzalone, A. V., Lam, K. C., et al. (2021). PrimeDesign Software for Rapid and Simplified Design of Prime Editing Guide RNAs. *Nat. Commun.* 12 (11), 1–6. doi:10.1038/s41467-021-21337-7

Hwang, G.-H., Jeong, Y. K., Habib, O., Hong, S.-A., Lim, K., Kim, J.-S., et al. (2021). PE-designer and PE-Analyzer: Web-Based Design and Analysis Tools for CRISPR Prime Editing. *Nucleic Acids Res.* 49 (W1), W499–W504. doi:10.1093/NAR/GKAB319

Kelley, M. J., Brüderlein, S., Sommer, J. B., Meltzer, P. S., Li, S., Osada, T., et al. (2010). Molecular Characterization of Putative Chordoma Cell Lines. *Sarcoma* 2010, 14. No Title. doi:10.1155/2010/630129

Kwart, D., Paquet, D., Teo, S., and Tessier-Lavigne, M. (2017). Precise and Efficient Scarless Genome Editing in Stem Cells Using CORRECT. *Nat. Protoc.* 12 (2), 329–354. doi:10.1038/NPROT.2016.171

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., and Valen, E. (2019). CHOPCHOP V3: Expanding the CRISPR Web Toolbox beyond Genome Editing. *Nucleic Acids Res.* 47 (W1), W171–W174. doi:10.1093/NAR/GKZ365

Landrum, M. J., and Kattman, B. L. (2018). ClinVar at Five Years: Delivering on the Promise. *Hum. Mutat.* 39 (11), 1623–1630. doi:10.1002/HUMU.23641

Liang, X., Potter, J., Kumar, S., Ravinder, N., and Chesnut, J. D. (2017). Enhanced CRISPR/Cas9-mediated Precise Genome Editing by Improved Design and Delivery of gRNA, Cas9 Nuclease, and Donor DNA. *J. Biotechnol.* 241, 136–146. doi:10.1016/J.JBIOTEC.2016.11.011

Lin, Q., Jin, S., Zong, Y., Yu, H., Zhu, Z., Liu, G., et al. (20212021). High-efficiency Prime Editing with Optimized, Paired pegRNAs in Plants. *Nat. Biotechnol.* 3939 (88), 923–927. doi:10.1038/s41587-021-00868-w

Lin, S., Staahl, B. T., Alla, R. K., and Doudna, J. A. (2014). Enhanced Homology-Directed Human Genome Engineering by Controlled Timing of CRISPR/Cas9 Delivery. *eLife* 3, e04766. doi:10.7554/ELIFE.04766

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., et al. (2019a). Methodologies for Improving HDR Efficiency. *Front. Genet.* 10 (JAN), 691. doi:10.3389/FGENE.2018.00691

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., et al. (2019b). Methodologies for Improving HDR Efficiency. *Front. Genet.* 9 (JAN). doi:10.3389/FGENE.2018.00691

Madsen, R. R., and Semple, R. K. (2019). Luminescent Peptide Tagging Enables Efficient Screening for CRISPR-Mediated Knock-In in Human Induced Pluripotent Stem Cells. *Wellcome Open Res.* 4, 37. doi:10.12688/WELLCOMEOPENRES.15119.3

Maurissen, T. L., and Woltjen, K. (2020). Synergistic Gene Editing in Human iPS Cells via Cell Cycle and DNA Repair Modulation. *Nat. Commun.* 11 (1), 16643. doi:10.1038/S41467-020-16643-5

Merkle, F. T., Ghosh, S., Kamitaki, N., Mitchell, J., Avior, Y., Mello, C., et al. (2017). Human Pluripotent Stem Cells Recurrently Acquire and Expand Dominant Negative P53 Mutations. *Nature* 545 (7653), 229–233. doi:10.1038/NATURE22312

Miyaoka, Y., Chan, A. H., Judge, L. M., Yoo, J., Huang, M., Nguyen, T. D., et al. (2014). Isolation of Single-Base Genome-Edited Human iPS Cells without Antibiotic Selection. *Nat. Methods* 11 (3), 291–293. doi:10.1038/NMETH.2840

Miyaoka, Y., Mayerl, S. J., Chan, A. H., and Conklin, B. R. (2018). Detection and Quantification of HDR and NHEJ Induced by Genome Editing at Endogenous Gene Loci Using Droplet Digital PCR. *Methods Mol. Biol.* 1768, 349–362. doi:10.1007/978-1-4939-7778-9_20

O'Brien, A. R., Wilson, L. O. W., Burgio, G., and Bauer, D. C. (2019). Unlocking HDR-Mediated Nucleotide Editing by Identifying High-Efficiency Target Sites Using Machine Learning. *Sci. Rep.* 9 (1), 39142. doi:10.1038/S41598-019-39142-0

Okamoto, S., Amaishi, Y., Maki, I., Enoki, T., and Mineno, J. (2019). Highly Efficient Genome Editing for Single-Base Substitutions Using Optimized ssODNs with Cas9-RNPs. *Sci. Rep.* 9 (1), 41121. doi:10.1038/S41598-019-41121-4

Papaioannou, I., Disterer, P., and Owen, J. S. (2009). Use of Internally Nuclease-Protected Single-Strand DNA Oligonucleotides and Silencing of the Mismatch Repair Protein, MSH2, Enhances the Replication of Corrected Cells Following Gene Editing. *J. Gene Med.* 11 (3), 267–274. doi:10.1002/JGM.1296

Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., et al. (2016). Efficient Introduction of Specific Homozygous and Heterozygous Mutations Using CRISPR/Cas9. *Nature* 533 (7601), 125–129. doi:10.1038/NATURE17664

Park, J., Lim, K., Kim, J.-S., and Bae, S. (2017). Cas-analyzer: An Online Tool for Assessing Genome Editing Results Using NGS Data. *Bioinformatics* 33 (2), 286–288. doi:10.1093/BIOINFORMATICS/BTW561

Pillay, N., Plagnol, V., Tarpey, P. S., Lobo, S. B., Presneau, N., Szuhai, K., et al. (2012). A Common Single-Nucleotide Variant in T Is Strongly Associated with Chordoma. *Nat. Genet.* 44 (11), 1185–1187. doi:10.1038/NG.2419

Rees, H. A., and Liu, D. R. (2018). Base Editing: Precision Chemistry on the Genome and Transcriptome of Living Cells. *Nat. Rev. Genet.* 19 (12), 770–788. doi:10.1038/S41576-018-0059-1

Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., and Corn, J. E. (2016). Enhancing Homology-Directed Genome Editing by Catalytically Active and Inactive CRISPR-Cas9 Using Asymmetric Donor DNA. *Nat. Biotechnol.* 34 (3), 339–344. doi:10.1038/NBT.3481

Scheipl, S., Barnard, M., Cottone, L., Jorgensen, M., Drewry, D.H., Zuercher, W.J., et al. (2016). EGFR Inhibitors Identified as a Potential Treatment for Chordoma in a Focused Compound Screen. *J. Pathol.* 239 (3), 320–334. doi:10.1002/PATH.4729

Schubert, M. S., Thommandru, B., Woodley, J., Turk, R., Yan, S., Kurgan, G., et al. (20212021). Optimized Design Parameters for CRISPR Cas9 and Cas12a Homology-Directed Repair. *Sci. Rep.* 11 (11), 1–15. doi:10.1038/s41598-021-98965-y

Skarnes, W. C., Pellegrino, E., and McDonough, J. A. (2019). Improving Homology-Directed Repair Efficiency in Human Stem Cells. *Methods* 164-165 (165), 18–28. doi:10.1016/J.YMETH.2019.06.016

Sledzinski, P., Nowaczyk, M., and Olejniczak, M. (2020). Computational Tools and Resources Supporting CRISPR-Cas Experiments. *Cells* 9 (5), 1288. doi:10.3390/CELLS9051288

Stratigopoulos, G., De Rosa, M. C., LeDuc, C. A., Leibel, R. L., and Doege, C. A. (2018). DMSO Increases Efficiency of Genome Editing at Two Non-coding Loci. *PLoS ONE* 13 (6), e0198637. doi:10.1371/JOURNAL.PONE.0198637

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., and Leunissen, J.A.M. (2007). Primer3Plus, an Enhanced Web Interface to Primer3. *Nucleic Acids Res.* 35 (suppl. 2), W71–W74. doi:10.1093/nar/gkm306

Zhang, J.-P., Li, X.-L., Li, G.-H., Chen, W., Arakaki, C., Botimer, G. D., et al. (2017). Efficient Precise Knockin with a Double Cut HDR Donor after CRISPR/Cas9-mediated Double-Stranded DNA Cleavage. *Genome Biol.* 18 (1). doi:10.1186/S13059-017-1164-8

Zhao, B., Chen, S.-A. A., Lee, J., and Fraser, H. B. (2022). Bacterial Retrons Enable Precise Gene Editing in Human Cells. *CRISPR J.* 5 (1), 31–39. doi:10.1089/CRISPR.2021.0065

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

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

frontiers | Frontiers in Genome Editing

TYPE Review
PUBLISHED 26 July 2022
DOI 10.3389/fgeed.2022.892769

OPEN ACCESS

EDITED BY
Ayal Hendel,
Bar-Ilan University, Israel

REVIEWED BY
Beeke Wienert,
Graphite Bio, Inc., United States
Karim Benabdellah,
Andalusian Autonomous Government of Genomics and Oncological Research (GENYO), Spain
Mollie Schubert,
Integrated DNA Technologies, United States

*CORRESPONDENCE
Jacques-P. Tremblay,
jacques-p.tremblay@crchudequebec.ulaval.ca

SPECIALTY SECTION
This article was submitted to Genome Editing Tools and Mechanisms, a section of the journal Frontiers in Genome Editing

RECEIVED 09 March 2022
ACCEPTED 08 July 2022
PUBLISHED 26 July 2022

CITATION
Lu Y, Happi Mbakam C, Song B, Bendavid E and Tremblay J-P (2022), Improvements of nuclease and nickase gene modification techniques for the treatment of genetic diseases. *Front. Genome Ed.* 4:892769. doi: 10.3389/fgeed.2022.892769

Improvements of nuclease and nickase gene modification techniques for the treatment of genetic diseases

Yaoyao Lu[1,2], Cedric Happi Mbakam[1,2], Bo Song[1,2], Eli Bendavid[1,2] and Jacques-P. Tremblay[1,2]*

[1]CHU de Québec Research Center, Laval University, Quebec City, QC, Canada, [2]Department of Molecular Medicine, Laval University, Quebec City, QC, Canada

Advancements in genome editing make possible to exploit the functions of enzymes for efficient DNA modifications with tremendous potential to treat human genetic diseases. Several nuclease genome editing strategies including Meganucleases (MNs), Zinc Finger Nucleases (ZFNs), Transcription Activator-like Effector Nucleases (TALENs) and Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR associated proteins (CRISPR-Cas) have been developed for the correction of genetic mutations. CRISPR-Cas has further been engineered to create nickase genome editing tools including Base editors and Prime editors with much precision and efficacy. In this review, we summarized recent improvements in nuclease and nickase genome editing approaches for the treatment of genetic diseases. We also highlighted some limitations for the translation of these approaches into clinical applications.

KEYWORDS

gene editing, ZFN, TALEN, CRISPR-cas, Cytidine Base Editor (CBE), Adenosine Base Editor (ABE), prime editing

1 Introduction

Mutations of a single or several nucleotides in human genome are responsible for major hereditary health problems (Benusiglio et al., 2021; Samuelson et al., 2021; Xiao et al., 2021; Chen et al., 2022; Cobo et al., 2022). To date, about 7,000 hereditary diseases are estimated to be caused by monogenic mutations (Claussnitzer et al., 2020). Homologous recombination (HR) has long been proposed as an avenue to treat human genetic diseases and its efficiency can be increased by inducing the DNA double-strand breaks (DSBs) (Mao et al., 2008). A significant step forward in gene therapy has been the discovery of enzymes known as nucleases. These enzymes enable gene editing technologies to modify a specific DNA sequence within the natural cell environment for the correction of hereditary diseases (Durai, 2005; Cermak et al., 2011; Hwang et al., 2013).

Over the past decades, significant technology development have empowered bio-engineers with tools such as nuclease-mediated Meganucleases (MNs) (Epinat et al.,

TABLE 1 The comparison of nuclease-mediated technologies.

	Mega-nuclease	ZFN	TALEN	CRISPR/Cas9
Enzyme	endonuclease	Fok1-nuclease	Fok1-nuclease	Cas9 nuclease
Target site	LAGLIDADG proteins	Zinc-finger binding sites	RVD tandem repeat region of TALE protein	PAM/spacer sequence
Recognition sequence size	12–45 bp	9–18 bp	14–20 bp	3–8 bp/20 bp
Targeting limitations	MN cleaving site	Difficult to target non-G-rich sites	5′ targeted base must be a T for each TALEN monomer	Targeted site must precede a PAM sequence
Advantage	1) High specificity	1) Small protein size	1) High specificity	1) Easy to engineer
	2) Relatively easy to deliver *in vivo*	2) Relatively easy *in vivo* delivery	2) Relatively easy to engineer	2) Easy to multiplex
Disadvantage	1) Complex to engineer	1) Expensive	1) Difficult to multiplex	1) Lower specificity
	2) Difficult to multiplex	2)Time-consuming	2) Not applicable for methylcytosine DNA	2) Limited *in vivo* delivery
	3) The target loci need to be engineered into genome	3) Difficult to select the target sequence	3) Limited *in vivo* delivery	
		4) All the ZF domains should be active	4) All the TALEs should be active	

2003), Zinc Finger Nucleases (ZFNs) (Porteus and Baltimore, 2003; Miller et al., 2007), Transcription Activator-Like Effector Nucleases (TALENs) (Method of the Year 2011, 2012) and Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR associated protein 9 (CRISPR-Cas9) (Jinek et al., 2012a; Ran et al., 2013b; Shalem et al., 2014b). These technologies grant us access to the genome for an accurate base-to-base modification without DSBs by deploying techniques such as Base editing and Prime editing (Gaudelli, 2017; Komor et al., 2017; Anzalone et al., 2019). Theoretically, these gene editing technology enable the replacement of single or multiple bases in any gene of interest at any given location. However, the gene editing effectiveness are influenced by at least three factors: (i) The type of gene editing (e.g., DNA base pair conversion, deletion, insertion, or a combination of the three above changes), (ii) the availability of a gene editing technology to achieve its desired outcome, (iii) the efficiency of the gene editing process. More importantly, the main bottleneck, the acquisition of tissue-specific edits and unwanted genome modification events, is still remains.

To explore how to edit the genome, in this review, we particularly focused on the mechanisms, the limitations and optimizations of six high-profile gene editing technologies as well as the recent progress of various types of genome editing tools used in clinical or preclinical research.

2 Nuclease-based genome engineering technologies

In the past decade, various nuclease gene editing technologies have been developed and widely used. These technologies empowered scientists to modify specific sequences in the genome of diverse organisms (Tzfira et al., 2003; Urnov et al., 2010). The most common nucleases-mediated gene editing technologies (Table 1) are MNs (Figure 1A), ZFNs (Figure 1B), TALENs (Figure 1C), and CRISPR-Cas9 (Figure 1D). These technologies combined specific DNA target recognition sequences and programmable endonucleases to induce the desired genomic DNA sequence alterations by introduction of DNA DSB resulting in insertions, deletions, gene replacements and nucleotide substitutions (Lo et al., 2013; Miyaoka et al., 2016). In eukaryotic cells, double-strand DNA cleavage by nucleases triggers two major DNA repair mechanisms including: (i) The non-homologous end joining (NHEJ) and microhomology-mediated end joining through re-ligation of the ends pathways, and (ii) the homology-directed repair (HDR) generated by repairing through a separate donor DNA template (Ranjha et al., 2018). Genome editing takes advantage of these DNA repair processes to produce desired genomic alteration in cell cultures and organisms. (Valerie and Povirk, 2003; Iliakis et al., 2004). However, the DNA DSBs can cause undesired outcomes such as insertions and deletions (Indels) as well as p53 activation (Naeem et al., 2020; Ihry et al., 2018).

2.1 Meganucleases

MNs also referred to as homing endonucleases recognize a large DNA stretches (12–40 base pairs) to facilitate cleavage in the most genomes (Rouet et al., 1994b). MNs are generally encoded by introns or inteins to promote homing of their respective genetic elements into intron or intein-free

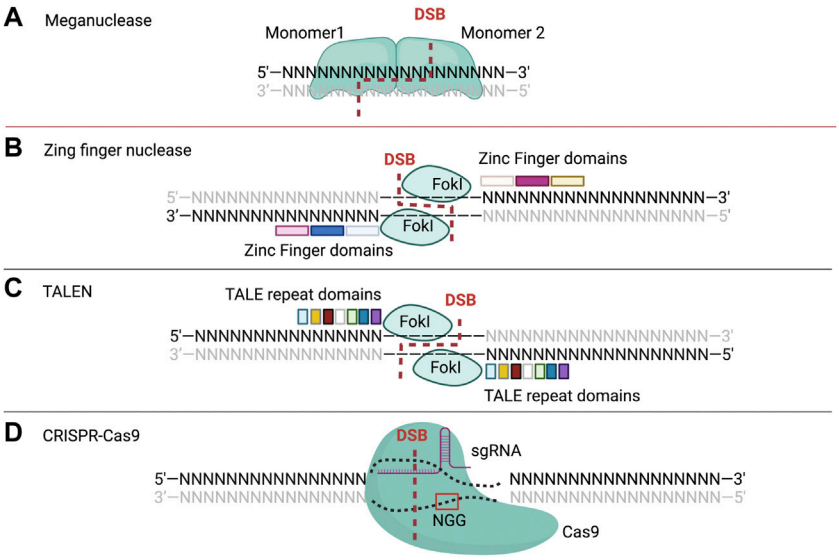


FIGURE 1
The **(A)** is depicting the two monomer domains (monomer 1 and monomer 2) of mega nuclease which cleave at the binding sites, resulting in the DNA double strand break. The **(B)** shows the two zinc finger domains of ZFNs made of three ZF motifs distinct binding sites (sequence of 3 nucleotides) and cleaving domains (FokI). The **(C)** represents TALE repeat domains (shown in colored squares) pair with distinct binding DNA nucleotides and cleaving domains (FokI). The **(D)** shows the single guide RNA complexed with Cas9 to open and cleave the DNA sequence, trough the recognition of the PAM sequence (by Cas9) which is NGG for SpCas9, and the target sequence (by sgRNA, single guide RNA). N represents any nucleotide amongst A, T, C, and G nucleotide.

homologous allelic sites (Epinat et al., 2003). One of the distinctive features of these MNs is their high specificity, due to the tight coupling of their binding site. This tight coupling recognizes a single locus within the yeast nuclear or mitochondrial genome (Paques and Duchateau, 2007). The LAGLIDADG proteins are the most well-studied homing endonuclease. They interact with their targets by nonspecific interactions between the ß strands and the backbone of the target DNA through the recognition of a sequence of 2–4 bp region. Consequently, several new engineered endonuclease variations derived from the following homing endonuclease: (i) I-CreI which was discovered in the chloroplast genome of *Chlamydomonas reinhardtii*, and (ii) I-SceI which is present in the mitochondria of *Saccharomyces cerevisiae* (Prieto et al., 2007; Gao et al., 2010; Zekonyte et al., 2021; Lee et al., 2022). These engineered endonucleases enable *in vivo* and *in vitro* genetic modifications. Due to the tiny molecular weight of modified MNs makes the *in vivo* delivery possible. However, the editing efficiency of this strategy is low compared to later developed nuclease-mediated technologies (Aubert et al., 2016; Zhu et al., 2017). Furthermore, reengineering MNs to expand the spectrum of DNA target sequences is complex and laborious and therefore vastly limit its application (Gouble et al., 2006). Recently, Zekonyte et al. (2021) used the engineered I-CreI meganuclease administered *via* intravenous (IV) injection of AAV9 into mice for the correction of m.5024C>T mutation in the mt-tRNA^{Ala} gene, as a curative method for disorders caused by heteroplasmic mitochondrial DNA mutation. This resulted to the elimination of mutant mitochondrial DNA followed by the restoration of mt-tRNA^{Ala} level. Moreover, this approach has also been used in pig model of autosomal dominant Retinitis Pigmentosa (adRP) for the correction of P23H mutation in the rhodopsin (RHO) gene (Jalligampala et al., 2021) and in non-human primates for the modification of the proprotein convertase subtilisin/kexin type 9 (PCSK9) gene responsible for hypercholesterolemia (Wang et al., 2018a; Wang et al., 2021).

2.1.1 Meganuclease related limitations and perspectives

Since the introduction of MNs application, unexpected drawbacks are constantly being discovered. Some of these challenges are: (i) The targeted locus must contain a specific MN cleavage site for each endonuclease whereas the microbial self-splicing intervening sequence could specifically duplicate into recipient alleles of their host gene lacking such sequence (Rouet et al., 1994a; Stoddard, 2011), (ii) low efficacy (Chapdelaine et al., 2010), and (iii) potential genotoxicity (Suzuki et al., 2020). The natural repertory of homing endonucleases is limited to a finite number of proteins, most of them still being hypothetical or uncharacterized. Thus, for other protein, the cleavage site for each meganuclease has to be

inserted into the target genome. Because of this flaw, the application of this technology is significantly limited. However, some engineered enzymes originated from meganuclease I-CreI (Arnould et al., 2011), I-SceI (Siegl et al., 2010) and I-DmoI (a monomeric meganuclease from the hyperthermophilic archaeon *Desulfurococcus mobilis*) (Molina et al., 2016) capable of cleaving DNA in specific genomic sites have been generated. I-CreI plays a critical role in the localization and occupancy of the catalytic metal ions, which is crucial for the DNA cleavage (Prieto et al., 2018). Wang et al. (2022) developed a transgenic Xenopus tropicalis line which is used for evaluating the potential effects of I-SceI mediated transgenesis and further understanding its mechanisms. The fusion of transcription activator-like effector (TALE) DNA-binding domains to MNs dramatically increases the efficiency by 35-fold compared to standalone MNs to modify T-cells receptor alpha (Boissel et al., 2014a). Furthermore, due to specific sites targeted by high cleavage specificity and the long length of the sequences (Petersen and Niemann, 2015b; Izmiryan et al., 2016), low off-target effects were detected in MNs because of the structure of meganucleases and the delivery methods. Some strategies could reduce the off-target by combining meganuclease with TALarray or I-TevI (a GIY-YIG enzyme) (Boissel et al., 2014b; Wang et al., 2018b).

2.2 Zinc finger nucleases

Zinc Finger Proteins (ZFPs) are artificially synthetic engineered hybrid heterodimeric proteins for site specific genome editing. ZFPs include a sequence of 3–6 peptides (called Zinc Finger (ZF) domains), each binding to a specific sequence of 3–6 base pairs for a specific attachment to a gene sequence (Durai et al., 2005). Two ZFPs are required to fuse with *Flavobacterium okeanokoites* endonuclease I (FokI) to induce a DSB at a specific genomic site. ZFNs are frequently used for gene silencing and knockout (Santiago et al., 2008; Gutschner et al., 2011; Gaj et al., 2012; Sun et al., 2018). Therefore, ZFNs have emerged as a versatile tool for gene targeting in various mammalian cells and organisms for the treatment of hereditary diseases (Almeida and Matos, 2019) and creation of animal models for diseases (Petersen and Niemann, 2015a).

ZFNs represent the first gene editing method applied for clinical treatment of diseases. ZFNs have been used to modify autologous $CD4^+$ T-cells to inhibit the function of the human chemokine (C-C motif) receptor 5 gene (CCR5) receptor and reduce the infection of these cells by HIV (Tebas et al., 2014). The results showed that infusion of genetically modified $CD4^+$ T-cells was well tolerated, and the HIV viral load has been decreased in blood level of most patients. Another experiment also showed that the HIV-specific $CD8^+$ T-cell responses are substantially restored (Tebas et al., 2021). The ZFN approach was also successfully used for the treatment of ß-hemoglobinopathies (Hoban et al., 2015; Chang et al., 2017; Psatha et al., 2018; Smith et al., 2019). Indeed, the phase I/II clinical trial (NCT03432364) sponsored by Sangamo Therapeutics Inc., aimed to assess the safety, tolerability and efficacy of ST-400 for the treatment of transfusion-dependent beta-thalassemia. The ST-400 are patient's hematopoietic stem cells genetically modified by ZFNs to disrupt a specific and precise sequence of the enhancer of the BCL11A gene in order to boost the expression of fetal hemoglobin (HbF) (Bauer and Orkin, 2015; Masuda et al., 2016). A cohort of six participants will be completed by November 2022. In other studies involving two patients manifesting different genotypic profiles showed a prompt hematopoietic reconstitution with long term increased HbF levels; however, serious adverse events (e.g., Hypersensitivity) have been recorded with one patient as a result of reengineered ST-400 (Smith et al., 2019). Scientists at Bioverative Inc., a Sanofi company, are conducting a phase I/II clinical trials (NCT03653247) in a cohort of eight patients to evaluate the safety, tolerability and efficacy of autologous hematopoietic stem cell transplantation using BIVV003 for the treatment of severe Sickle Cell Disease (SCD) in adults. This trial was supported by encouraging preclinical results that showed a robust long-term engraftment of *ex-vivo* modified hematopoietic stem and progenitor cells (HSPC) from patients. In other phase I/II clinical trials conducted by Sangamo Therapeutics Inc., the UCSF Benioff Children's Hospital delivered SB-318 (NCT02702115) and SB-913 (NCT03041324) into participants to insert the corrected copy of α-L-iduronidase (IDUA) and iduronate-2-sulfatase (IDS) transgenes respectively into the albumin locus to provide permanent liver specific expression of iduronidase. Preliminary results showed the evidence of albumin-IDS mRNA transcripts in liver and the hepatocytes are able to generate active IDS enzyme (Muenzer et al., 2019). The summary of ongoing clinical trials is presented in Table 2.

2.2.1 ZFN related limitations and perspectives

Although ZFNs exhibited their ability to modify a specific gene in mammalian cells, the strategy faces three major limitations. 1) Cannot cut arbitrary gene sequences (Durai, 2005). 2) A ZFN coding gene must be engineered for each specific target site (Porter et al., 2019). 3) The likelihood of off-target gene editing is another drawback of the ZFN technology (Pattanayak et al., 2011). At the beginning, the targeted sequence should contain 5′-GNN, 5′-ANN, 5′-CNN or 5′TNN (Dreier et al., 2001, 2005). Therefore, this technology is costly, laborious, time consuming and requires highly trained researchers for protein engineering (Porter et al., 2019). Recently, the liaison between the ZFP and FokI cleavage domain have been substituted to increase the number of distinct zinc-finger arrays enabling cleavage at a target genomic site (Paschon et al., 2019). In comparison to classical ZFNs, this technique reduced the off-target effect, boosted modification activities, and is more precise. More significantly, it can target and cleave at any intended base (Paschon et al., 2019). In addition, different strategies to engineer the ZFNs and reduce the off-target

TABLE 2 Summary of ZFNs ongoing clinical trials.

Disease	Trial number	Sponsor	Status	Drug	Phase	Completion date
Mucopolysaccharidosis II	NCT03041324	Sangamo Therapeutics	Terminated	SB-913	I/II	May-2021
Mucopolysaccharidosis I	NCT02702115	Sangamo Therapeutics	Terminated	SB-318	I/II	Nov-2021
Hemophilia B	NCT02695160	Sangamo Therapeutics	Terminated	SB-FIX	I	Apr-2021
Beta-thalassemia	NCT03432364	Sangamo Therapeutics	Active	ST-400	I/II	Nov-2022
Mucopolysaccharidosis, Hemophilia	NCT04628871	Sangamo Therapeutics	Enrolling	SB-913, SB-318, SB-FIX	NA	Jan-2030

mutations have also been developed (Ji et al., 2018; Miller et al., 2019a). The assembly of a high specific ZFNs system is highly complex. Researchers isolated naturally occurring ZF modules with different sequence specificities to engineer ZF modules with altered DNA binding specificities (Xiong et al., 2013). Furthermore, several methods are available to increase the specificity and reduced the cellular toxicity of this system by improving the ZFN architecture to develop FokI nuclease domain variants, which could result in a 3,000-fold reduction in off-target indels. (Miller et al., 2007; Miller et al., 2019b).

2.3 Transcription activator-like effector nucleases

A class of naturally occurring DNA binding proteins called the Transcription Activator-Like Effector (TALE) has been identified in the plant pathogen Xanthomonas. These TALEs regulate the transcription of several host target genes (Sanjana et al., 2012). TALENs are artificial engineered proteins combining the DNA-binding properties of a TALE protein and the DNA cleavage of the FokI endonuclease (Moscou and Bogdanove, 2009). The central region of TALEs is composed of 34 amino acid repeats amongst which 32 are constant and 2 are variable and recognized as repeat variable diresidues (RVDs) (Scholze and Boch, 2011). RVDs are involved in the DNA target recognition (Boch et al., 2009). Two TALENs target binding sequences are required to form a FokI dimer that induces a DSB (Feng et al., 2014).

TALENs have successfully been used for the modification of T-cell receptors for the treatment of leukemia (Qasim et al., 2017; Benjamin et al., 2020). UCART19, a CAR-T-cell product engineered with TALENs, was tested in children and adults in phase I clinical trials (NCT02808442 and NCT02746952) to cure advanced lymphoid malignancies and refractory B-cell acute lymphoblastic leukemia (B-ALL). These trials demonstrated the potential of UCART19 in patients with aggressive leukemia, but significant adverse events such as cytokine release syndrome, acute graft-versus-host of the skin and infectious complications have also been observed (Benjamin et al., 2020). There are many phase I trials sponsored by Cellectis Inc., using programmed allogenic engineered T-cells expressing different CARs such as UCART123 (NCT04106076, NCT03203369), UCARTCS1A (NCT04142619) and UCART22 (NCT04150497) respectively to treat acute myeloid leukemia, blastic plasmacytoid dendritic cell neoplasm (BPDCN), multiple myeloma and $CD22^{+}$ B cell acute lymphoblastic leukemia. Ongoing clinical trials using this approach are summarized in Table 3.

2.3.1 TALENs related limitations and perspectives

Compared to MNs and ZFNs, TALENs exhibits high efficiency, low off-target effects and are proven to target the mitochondrial DNA (Mussolino et al., 2014). Nonetheless, there are some constraints that prevent a more widespread deployment. 1) The repetitive sequences of TALEs make them difficult to construct using polymerase chain reaction (PCR) (Cermak et al., 2011). 2) TALENs are unable to target a methylated DNA, because the methylation of cytosine can potentially abrogate TALE binding and alter recognition by its normal RVD (Deng et al., 2019). Different approaches have been proposed to bypass the challenges associated with the TALEs repetitive sequences are the following: (i) design the ligation-independent cloning techniques (Reyon et al., 2012), (ii) the high-throughput solid-phase assembly (Schmid-Burgk et al., 2013), (iii) the Golden gate cloning (Cermak et al., 2015), and (iv) alternative one-day TALE assembly (Zhang et al., 2020). To improve the efficiency of gene editing, a new bicistronic TALEN termed T2A using classical TALEN coding sequences linked to different reporter molecules by 2A "self-cleaving peptide" has been developed. This improvement could help each TALEN monomer to transcribe from the same reading frame in order to increase the gene editing efficacy. (Mariano et al., 2014; Martín-Fernández et al., 2020). Additionally, Zhang et al. (2017) used deciphered TALEs for 5-hydroxymethylcytosine and 5-methylcytosine to achieve methylation-dependent genome editing and gene activation *in vivo*.

2.4 Clustered regularly interspaced short palindromic repeats—CRISPR associated protein 9

CRISPR-Cas9 system is a sophisticated gene editing tool that revolutionized the genome engineering field and generated excitement for the potential of novel therapeutic approaches

TABLE 3 Summary of TALENs ongoing clinical trials.

Disease	Trial number	Sponsor	Statute	Drug	Phase	Completion date
Myeloma	NCT03190278	Cellectis S.A.	recruiting	UCART123v1.2	I	Oct-2022
Leukemia	NCT04150497	Cellectis S.A.	recruiting	UCART22	I	Oct-2022
Myeloma	NCT04142619	Cellectis S.A.	recruiting	UCARTCS1A	I	Nov-2022

TABLE 4 Summary of CRISPR ongoing clinical trials.

Disease	Trial number	Sponsor	Drug	Phase	Completion date
Leukemia and lymphoma	NCT03398967	Chinese PLA General Hospital	CD19, CD20, CD22 CAR-T-cells	I/II	May-2022
Leukemia	NCT04557436	Great Ormond Street Hospital for Children NHS Foundation Trust	PBLTT52CAR19	I	June-2022
Gastrointestinal cancer	NCT04426669	Intima Bioscience, Inc.	NA	I/II	Oct-2022
Lymphoma	NCT04767308	Huazhong University of Science and Technology	CT125A	I	Dec-2023
β-thalassemia	NCT04925206	EdiGene (GuangZhou) Inc.	ET-01	I	June 2024
Leber congenital amaurosis	NCT03872479	Editas Medicine, Inc.	EDIT-101	I/II	Mar-2024
β-thalassemia	NCT03655678	Vertex Pharmaceuticals Incorporated	CTX001	II/III	Aug-2024
Sickle cell disease	NCT03745287	Vertex Pharmaceuticals Incorporated	CTX001	II/III	Oct-2024
Transthyretin amyloidosis	NCT04601051	Intellia Therapeutics	NTLA-2001	I	Nov-2024
Leukemia	NCT04037566	Xijing Hospital	XYF19 CAR-T	I	Aug-2025
Myeloid leukemia	NCT05066165	Intellia Therapeutics	NTLA-5001	I/II	Sep-2025
N-H lynphoma	NCT04637763	Caribou Biosciences, Inc.	CB-010	I	Sep-2025
Hereditary angioedema	NCT05120830	Intellia Therapeutics	NTLA-2002	I/II	Dec-2025
Sickle cell disease	NCT04819841	Graphite Bio, Inc.	GPH101	I/II	May-2026
Sickle cell disease	NCT05329649	Vertex Pharmaceuticals Incorporated	CTX001	III	May-2026
Leukemia	NCT04035434	CRISPR Therapeutics AG	CTX110	I	Aug-2026
Sickle cell disease	NCT04774536	Mark Walters, MD	CRISPR_SCD001	I/II	Dec-2026
Myeloma	NCT04244656	CRISPR Therapeutics AG	CTX120	I	Janv-2027
Carcinoma	NCT04438083	CRISPR Therapeutics AG	CTX130	I	Avr-2027
Lymphoma	NCT04502446	CRISPR Therapeutics AG	CTX130	I	May-2027
Sickle cell disease	NCT04208529	Vertex Pharmaceuticals Incorporated	CTX001		Sep-2039

to treat human diseases. The CRISPR-Cas system is divided into class I and class II (Makarova et al., 2015). The class I uses multi-protein complexes for nucleic acid cleavage and is subdivided into CRISPR-Cas types I, III, and IV. The class II uses a single protein effector domain for the cleavage and is subdivided in CRISPR-Cas type II, V, and VI. The CRISPR-Cas9 belongs to type II system, which is simple to use and thus become the most widely utilized tool for biological research and translational applications (Tang and Fu, 2018).

Over the past decade, CRISPR-Cas9 system has been modified and adapted to become a versatile tool for genome editing in eukaryotes (Ran et al., 2013a; Perez-Pinera et al., 2013; Shalem et al., 2014a; Gao et al., 2019; Casas-Mollano et al., 2020). The system requires a Cas9 nuclease and a single guide RNA (sgRNA) adapted from CRISPR RNA (crRNA), which specifies the target site (the spacer sequence), fused with a trans-activating RNA (tracrRNA), which forms a complex with the crRNA (the scaffold sequence) (Jinek et al., 2012a). The sgRNA forms a stable ribonucleoprotein complex with Cas9 nuclease which initially attaches to a protospacer adjacent motif (PAM) to initiate the first conformational changes of the protein. The targeting activity is driven by 20 nucleotides of RNA-DNA base-pairing between the target DNA strand protospacer and the complementary RNA strand and through interactions between the non-target DNA

strand PAM. Subsequently, mediates the second conformational change of the protein which then becomes active. Once activated, the HNH domain of Cas9 cleaves the DNA strand to which the sgRNA is attached and the RuvC domain cleaves at the PAM strand (Jinek et al., 2012b). One of the advantages of CRISPR/Cas9 is that it only requires a sgRNA to specify the DNA sequence where a DSB needs to be generated.

There are many clinical trials based on CRISPR-Cas9 and half of these trials (Phase I and Phase II) have already been successfully completed, the ongoing clinical trials are summarized in Table 4. In a phase I trial (NCT03399448), Pennsylvania University used a multiplex CRISPR-Cas9 to knockout the TCRα, TCRβ, PD-1 genes to treat various malignancies (Stadtmauer et al., 2020). The results showed that the modified T cells engrafted in patients at stable levels for at least 9 months and were barely immunogenic, indicating the feasibility of CRISPR-Cas9 gene editing for cancer immunotherapy (Stadtmauer et al., 2020). In addition to that, some clinical trials moved the technology from ZFNs or TALENs to CRISPR-Cas9 (Yi and Li, 2016; Lu et al., 2020). This was done for example to knockout PD-1 and CD52 for the different types of cancer by electroporating Cas9 and a sgRNA to edit the cells *ex vivo*. The results indicated that the clinical applications of CRISPR-Cas9 gene-edited T-cells are generally safe and feasible (Lu et al., 2020; Frangoul et al., 2021; Modarai et al., 2021). The disruption of erythroid enhancer of the BCL11A gene by CRISPR-Cas9 for the treatment of ß-thalassemia has been observed with serious adverse events. The reported serious adverse events are sepsis and pneumonia in the presence of neutropenia, vaso-occlusive liver disease, abdominal pain and cholelithiasis after CTX001 (autologous CRISPR-Cas9–edited $CD34^+$ HSPCs) intravenous (IV) administration, although the substantially raised of hemoglobin levels in fetal blood cells (Frangoul et al., 2021). Two clinical trials are currently conducted to evaluate the safety and efficacy of CRISPR-Cas9 strategies either to restore ß-globin expression in ß-thalassemic patients harboring the IVS-2-654C>T mutation (NCT04205435) or to de-repress γ-globin in transfusion dependent thalassemia (TDT) patients (NCT04211480), but no results are yet available. Recently, another study prepared NTLA-2001, an *in vivo* gene-editing therapeutic agent, made by lipid nanoparticles encapsulating messenger RNA of Cas9 protein and single guide RNA targeting misfolding transthyretin (TTR) responsible of transthyretin amyloidosis. The result showed that the TTR is durably knockout after a single dose of NTLA-2001 (Gillmore et al., 2021; Philippidis, 2021).

2.4.1 Clustered regularly interspaced short palindromic repeats—CRISPR associated protein 9 limitations and perspectives

The CRISPR-Cas9 technology is associated with some limitation: (i) There are some off-target effects, (ii) the PAM requirement initially limits the editing scope to only sequence near an NGG for canonical SpCas9, and (iii) a more efficient method of delivering the CRISPR/Cas9 system for *in vivo* applications, which includes a protein of about 160 kDa, is required (Woo et al., 2015). Fortunately, the impact of these shortcomings is reduced with the continuous improvement of this technology. Amongst these improvements, Cas9 variants (Kleinstiver et al., 2016a; Slaymaker et al., 2016; Chen et al., 2017; Liu et al., 2020) and engineered sgRNA (Doench et al., 2014; Kocak et al., 2019; Nelson et al., 2022a) have been created to improve specificity and efficiency. Subsequently, the Cas9 variants requiring different PAM sequences which contributes to solve the PAM restriction challenge has also been developed. These variants are described as follows: (i) VQR (D1135V/R1335Q/T1337R) variant which recognizes NGA sequence and (ii) EQR (D1135E/R1335Q/T1337R) variant which recognizes NGA PAM sequence instead of an NGG sequence (Kleinstiver et al., 2015), (iii) SpCas9-NG variant which enables the recognition of NG PAM sequence (Nishimasu et al., 2018), (iv) xCas9 binding to NG, GAT as well as GAA (Hu et al., 2018) and the newest variant (v) SpRY which recognizes the NYN PAM and nearly eliminates the PAM restriction (Walton et al., 2020). These PAM flexibilities significantly increased the genome accessibility. For the delivery part, some CRISPR/Cas9 tools were successfully used to edit cells *in vitro*. This *ex vivo* approach may have certain safety benefits especially regarding off-target gene editing. However, some versions of CRISPR-Cas9 with a sgRNA cannot be efficiently delivered *in vivo* due to their size. Split viruses have been developed to resolve this problem but with reduced expression of the fusion protein. The limitations and optimizations of CRISPR/Cas9 and other limitations such as immunotoxicity, DNA-damage toxicity, etc. have also been detailed. (Tang and Gu, 2020; Uddin et al., 2020; Liu et al., 2021a; Wang et al., 2021b; Khatibi et al., 2021; Maximiano et al., 2021).

3 Nickase-based genome engineering technologies

DSBs at targeted genomic loci could be associated with serious undesirable effects, including p53 activation (Haapaniemi et al., 2018), translocations (Ghosh et al., 2018), off-target mutation (Vakulskas and Behlke, 2019) and complex undesired products (Paquet et al., 2016). Furthermore, half of all known disease associated gene variants are point mutations (Landrum et al., 2016). Therefore, the Cas9 nickases emerged as useful tools with a targetable property (Cong et al., 2013). The Cas9n D10A and Cas9n H840A are Cas9 variants mediating the cleavage of a single strand of the DNA respectively in the gRNA complementary or non-complementary DNA strand. These Cas9 nickases have been fused with various enzymes, to develop new gene editing technologies, including Base editors [Cytidine base editor (CBE)], Adenosine base editor (ABE) and Prime editor (PE) (Figure 2).

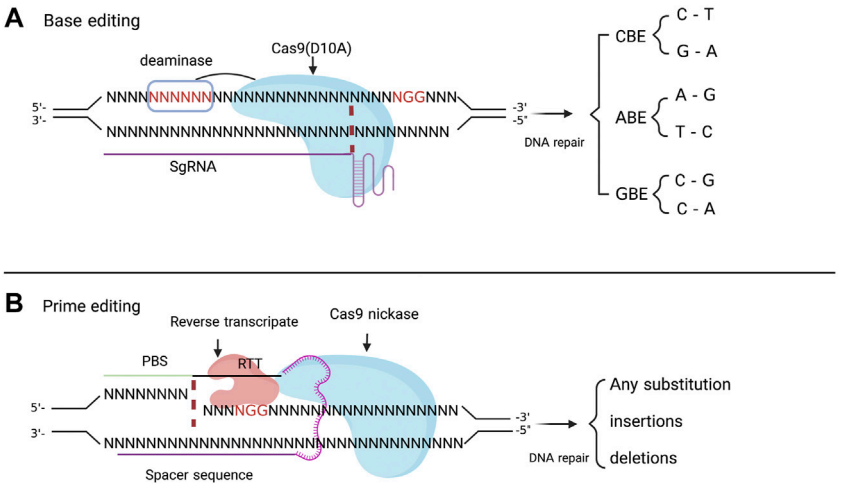


FIGURE 2
(A) represents a basic principle of base editor which is made of sgRNA that target a specific DNA sequence, the Cas9 nickase (D10A) which in interaction with sgRNA binds his recognition domain (PAM sequence) and cleaves the non-PAM DNA strand. The Cas9 is linked to a deaminase which modifies the targeted nucleotide in a window of 5 nucleotides (shown in red color) in the spacer sequence through CBE, ABE, and GBE. The **(B)** depicts the prime editing principle makes of pegRNA which include a spacer sequence, a primer binding site (PBS) and reverse transcript template (RTT), and Cas9 nickase fused with a reverse transcriptase. The pegRNA recognizes the target sequence and provide the desired sequence for modification. Once the Cas9n cleaves the DNA sequence, the reverse transcriptase uses the RTT as template for the synthesis of a new sequence containing the desired edit (gene substitution, insertion and deletion).

3.1 Base editor

Base editors are mainly used to target point mutations that may result in an altered DNA sequence with novel or enhanced functions and gene inactivation (Gaudelli, 2017; Gapinske et al., 2018; Zhang et al., 2019). Briefly DNA base editing requires two main components: a Cas9n fused with a deaminase and a sgRNA which binds to a specific DNA sequence (Komor et al., 2016). Three (3) types of base editors have been created including the Cytidine Base Editor (CBE) (Komor et al., 2016), Adenosine Base Editor (ABE) (Gaudelli, 2017) and Glycosylase Base Editors (GBE) (Zhao et al., 2021a) (Figure 2A).

Theoretically, base editors can function in both dividing and non-dividing cells to address several of single-nucleotide polymorphisms (SNPs) associated with human diseases.

Base editing is being used to study and treat hereditary diseases in a variety of cell types (Zeng et al., 2020; Tekel et al., 2021) and organisms (Kim et al., 2021c), including animal models of human hereditary diseases (Liu et al., 2018; Caso and Davies, 2022). Alternatively, some preclinical experiments demonstrated the prospects of using the base editing technologies to treat diseases. Due to the base editor, hA3A-BE3 creates C-to-T conversion at NGN PAM sites efficiently. Wang et al. (2020) showed that the binding site of BCL11A (TGACCA: −114 to −119) represents an ideal target site for binding with hA3A-BE3 to induce a mutation and raise the level of fetal γ-globin expression to ameliorate the ß-hemoglobinopathies. Furthermore, a transformer BE (tBE) system to eliminate unintended mutations was developed and delivered by AAV into mice creating a premature stop codon in proprotein convertase subtilisin/kexin type 9 (PCSK9) gene which significantly reduced the serum PCSK9 and cholesterol (Wang et al., 2021a). Newby et al. (2021) used mRNA encoding the Base editor to treat hematopoietic stem and progenitor cells (HSPCs) from patient with SCD generating high percentage conversion of the SCD allele (HBB^S) into Makassar ß-globin (HBB^G), which is a non-pathogenic variant. It has been shown that base edited CAR T-cells for combinational therapy against T-cell malignancies permitted to enhance molecular remission prior to allo-HSCT for T-cell malignancies (Georgiadis et al., 2021).

3.1.1 BE system limitations and optimizations

The BE system is mainly affected by three aspects: (i) the product purity, (ii) the off-target mutation and 3) the editing window of adjacent sites. Product purity refers to the percentage of edited sequencing reads [reads in which the targeted C has been converted into T but also C to R (G or A)] (Zhang et al., 2018). Furthermore, the product purity is also associated with the frequencies of Indels which are formed during the gene modifying process. Other studies have shown that over-expression of Uracil DNA glycosylase inhibitor (UGI) increased the purity of BE products in human cells (Komor et al., 2017; Jiang et al., 2018; Jang et al., 2021) In

addition to UGI, Gehrke et al. (2018) used an engineered human enzyme known as APOBEC3A (eA3A) to develop the eA3A-BE3 base editor which improved target accuracy and reduced bystander mutations. To reduce the Indel formations, the BE4-Gam and CBEmax base editors have been developed (Komor et al., 2017; Huang et al., 2019). The genome targeting function of the base editor caused substantial off-target editing in genomic DNA and RNA (Gaudelli, 2017). Different bioinformatic tools have been developed to overcome the off-target problems including the CRISPR-Cas9/Cpf1 (Kim et al., 2020; Lee et al., 2020), BE-Designer, BE-Analyzer (Hwang and Bae. 2021) and Digenome-seq (Kim et al., 2021b). Furthermore, the engineered deaminases (Rees et al., 2017; Zhou et al., 2019) and gRNA sequences (Kleinstiver et al., 2016b; Hu et al., 2021) have also been built to significantly reduce the off target in the BE system. The range of the editing window for base editing varies according to different application (Kim et al., 2017; Wang et al., 2018b, Wang et al., 2020a; Huang et al., 2019; Tan et al., 2019; Dang et al., 2021; Fu et al., 2021). There are two solutions to set the best range for editing window. 1) When only one specific base pair is required to be changed accurately, the editing window should be minimized to increase the target base accuracy. 2) When CBE system is used to introduce premature stop codons, to produce large-scale saturation mutations, to screen gene function, to locate key amino acid positions in protein domains, etc., a large editing activity window is more advantageous.

3.2 Prime editing

Genome editing with base editors effectively induced C→T, G→A, A→G, T→C, C→G, and C→A base substitutions without inducing DSB (Gaudelli, 2017; Zhao et al., 2021b; Kurt et al., 2021). However, they are unable to correct variants beyond these six transition mutations, or other modifications like insertions and deletions of DNA fragments which are successfully achieved by Prime editing (PE). PE uses an engineered Cas9 nickase fused to a reverse transcriptase (RT) enzyme and a modified sgRNA known as prime editing guide RNA (pegRNA) (Figure 2B) (Anzalone et al., 2019). Recent efforts stepwise improved the efficiency of PE system to PEmax system (Chen et al., 2021) and engineer pegRNA known as epegRNA (Nelson et al., 2022b). PE makes possible the accurate insertion up to 1 kb (Wang et al., 2022a) and the deletion of up to 10 kb (Choi et al., 2022) DNA fragment.

As base editing, prime editing has not yet entered clinical trials due to its immature development. However, the potential of PE has so far been demonstrated during the past 2 years *in vitro* (Surun et al., 2020; Habib et al., 2022; Happi Mbakam et al., 2022; Petri et al., 2022; Tremblay et al., 2022) and in animal models (Liu et al., 2021b; Jang et al., 2022; Zheng et al., 2022; Zhi et al., 2022). Unfortunately, the gene modification efficiency is very low in some models. This may be due to the use of the split viruses to overcome the size of PE, which cannot be packed into a single viral delivery vector. Prime editing has the potential to increase the safety and expand the scope of genome-editing in T-cells showing that method is adaptable to enhance the efficiency of CAR T-cell therapy by concurrently introducing additional complex gene edits into T-cells (Petri et al., 2022).

3.2.1 Prime editing limitations and optimizations

The defects of Prime editing are somehow similar to classical CRISPR/Cas9 and base editing as described above, such as the PAM restriction, off-targets and large molecular weight delivery hindering. However, some new progresses have been obtained. Kweon et al. (2021) developed PE2 variants by using various SpCas9 variants named PE2-VQR, PE2-VRQR, PE2-NG, PE2-SpG, and PE2-SpRY. The PE2-SpRY enables targeting 94.4% of pathogenic variants. Furthermore, some studies demonstrated that the PE is not always efficient due to some unknown factors (Li et al., 2020; Lin et al., 2020; Xu et al., 2020). Recently, further research showed that the design of the pegRNA strongly affects the efficiency of prime editors. Researchers demonstrated that optimization of pegRNA sequence widely improved PE efficiency in different cell lines and the efficiency of installing or correcting disease-associated mutations (Jiang et al., 2020, 2022; Kim et al., 2021b; Lin et al., 2021; Nelson et al., 2021).

4 Conclusion

During the past decade, gene editing technologies got tremendous improvement in optimizations and applications. With the continuous optimizations of these technologies ZFNs, TALENs and CRISPR-Cas9 have already entered human clinical trials. To date, the majority of clinical applications of these technologies are focused on *ex vivo* gene editing therapeutics. *Ex vivo* editing is highly effective for many medical conditions, such as sickle cell disease, but genome editing should ideally be used for diseases that require *in vivo* cell modification. Having said that, the *in vivo* applications of CRISPR technologies is challenged by issues such as off-target editing, inefficiency, and the stimulation of counterproductive immune responses. Current research addressing these issues may launch new avenues for clinical applications of those nuclease-mediated technologies. Moreover, the novel innovations such as Base editing and Prime editing are still at the pre-clinical stage. All these approaches require DNA strand break that evokes DNA damage responses. Therefore, more efforts are needed to address these limitations including wide off-target events, genome stability, transcription-activation systems and cell

proliferation to accelerate the treatment of genetic and infectious diseases.

Author contributions

YL and CHM conceived the manuscript, YL, CHM, EB, BS, and J-PT wrote the manuscript, J-PT supervised the work.

Funding

The work is supported by the grants from the Canadian Institute of Health Research and from the Foundation for Cell and Gene Therapy. YL is supported by the China Scholarship Council (CSC) Grant.

Conflict of interest

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

Publisher's note

References

Almeida, M. J., and Matos, A. (2019). Designer nucleases: Gene-editing therapies using CCR5 as an emerging target in HIV. *Curr. HIV Res.* 17, 306–323. doi:10.2174/1570162X17666191025112918

Anzalone, A. v., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157. doi:10.1038/s41586-019-1711-4

Arnould, S., Delenda, C., Grizot, S., Desseaux, C., Paques, F., Silva, G. H., et al. (2011). The I-CreI meganuclease and its engineered derivatives: Applications from cell modification to gene therapy. *Protein Eng. Des. Sel.* 24, 27–31. doi:10.1093/protein/gzq083

Aubert, M., Madden, E. A., Loprieno, M., Feelixge, H. S. D., Stensland, L., Huang, M.-L., et al. (2016). *In vivo* disruption of latent HSV by designer endonuclease therapy. *JCI Insight* 1, e88468. doi:10.1172/jci.insight.88468

Bauer, D. E., and Orkin, S. H. (2015). Hemoglobin switching's surprise: The versatile transcription factor BCL11A is a master repressor of fetal hemoglobin. *Curr. Opin. Genet. Dev.* 33, 62–70. doi:10.1016/j.gde.2015.08.001

Benjamin, R., Graham, C., Yallop, D., Jozwik, A., Mirci-Danicar, O. C., Lucchini, G., et al. (2020). Genome-edited, donor-derived allogeneic anti-CD19 chimeric antigen receptor T cells in paediatric and adult B-cell acute lymphoblastic leukaemia: Results of two phase 1 studies. *Lancet* 396, 1885–1894. doi:10.1016/S0140-6736(20)32334-5

Benusiglio, P. R., Fallet, V., Sanchis-Borja, M., Coulet, F., and Cadranel, J. (2021). Lung cancer is also a hereditary disease. *Eur. Respir. Rev.* 30, 210045. doi:10.1183/16000617.0045-2021

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., et al. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509–1512. doi:10.1126/science.1178811

Boissel, S., Jarjour, J., Astrakhan, A., Adey, A., Gouble, A., Duchateau, P., et al. (2014a). megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Res.* 42, 2591–2601. doi:10.1093/nar/gkt1224

Boissel, S., Jarjour, J., Astrakhan, A., Adey, A., Gouble, A., Duchateau, P., et al. (2014b). megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Res.* 42, 2591–2601. doi:10.1093/nar/gkt1224

Casas-Mollano, J. A., Zinselmeier, M. H., Erickson, S. E., and Smanski, M. J. (2020). CRISPR-cas activators for engineering gene expression in higher eukaryotes. *CRISPR J.* 3, 350–364. doi:10.1089/crispr.2020.0064

Caso, F., and Davies, B. (2022). Base editing and prime editing in laboratory animals. *Lab. Anim.* 56, 35–49. doi:10.1177/0023677221993895

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., et al. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39, e82. doi:10.1093/nar/gkr218

Cermak, T., Starker, C. G., and Voytas, D. F. (2015). Efficient design and assembly of custom TALENs using the Golden Gate platform. *Methods Mol. Biol.* 1239, 133–159. doi:10.1007/978-1-4939-1862-1_7

Chang, K.-H., Smith, S. E., Sullivan, T., Chen, K., Zhou, Q., West, J. A., et al. (2017). Long-term engraftment and fetal globin induction upon BCL11A gene editing in bone-marrow-derived CD34 + hematopoietic stem and progenitor cells. *Mol. Ther. Methods Clin. Dev.* 4, 137–148. doi:10.1016/j.omtm.2016.12.009

Chapdelaine, P., Pichavant, C., Rousseau, J., Pâques, F., and Tremblay, J. P. (2010). Meganucleases can restore the reading frame of a mutated dystrophin. *Gene Ther.* 17, 846–858. doi:10.1038/gt.2010.26

Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., et al. (2017). Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature* 550, 407–410. doi:10.1038/nature24268

Chen, P. J., Hussmann, J. A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.-F., et al. (2021). Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell.* 184, 5635–5652.e29. e29. doi:10.1016/j.cell.2021.09.018

Chen, W., Dai, G., Qian, Y., Wen, L., He, X., Liu, H., et al. (2022). PIK3CA mutation affects the proliferation of colorectal cancer cells through the PI3K-MEK/PDK1-GPT2 pathway. *Oncol. Rep.* 47, 11. doi:10.3892/or.2021.8222

Choi, J., Chen, W., Suiter, C. C., Lee, C., Chardon, F. M., Yang, W., et al. (2022). Precise genomic deletions using paired prime editing. *Nat. Biotechnol.* 40, 218–226. doi:10.1038/s41587-021-01025-z

Claussnitzer, M., Cho, J. H., Collins, R., Cox, N. J., Dermitzakis, E. T., Hurles, M. E., et al. (2020). A brief history of human disease genetics. *Nature* 577, 179–189. doi:10.1038/s41586-019-1879-7

Cobo, I., Tanaka, T., Glass, C. K., and Yeang, C. (2022). Clonal hematopoiesis driven by DNMT3A and TET2 mutations: Role in monocyte and macrophage biology and atherosclerotic cardiovascular disease. *Curr. Opin. Hematol.* 29, 1–7. doi:10.1097/MOH.0000000000000688

Cong, L., Ran, F. A., Cox, D., Lin, S. L., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi:10.1126/science.1231143

Dang, L., Zhou, X., Zhong, X., Yu, W., Huang, S., Liu, H., et al. (2021). Correction of the pathogenic mutation in TGM1 gene by adenine base editing in mutant embryos. *Mol. Ther.* 30, 175–183. doi:10.1016/j.ymthe.2021.05.007

Deng, P., Carter, S., and Fink, K. (2019). Design, construction, and application of transcription activation-like effectors. *Methods Mol. Biol.*, 47–58. doi:10.1007/978-1-4939-9065-8_3

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., et al. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* 32, 1262–1267. doi:10.1038/nbt.3026

Dreier, B., Beerli, R. R., Segal, D. J., Flippin, J. D., and Barbas, C. F. (2001). Development of zinc finger domains for recognition of the 5′-ANN-3′ family of DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* 276, 29466–29478. doi:10.1074/jbc.M102604200

Dreier, B., Fuller, R. P., Segal, D. J., Lund, C. v., Blancafort, P., Huber, A., et al. (2005). Development of zinc finger domains for recognition of the 5′-CNN-3′

family DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* 280, 35588–35597. doi:10.1074/jbc.M506654200

Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M. H., Chandrasegaran, S., et al. (2005). Zinc finger nucleases: Custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res.* 33, 5978–5990. doi:10.1093/nar/gki912

Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M. H., and Chandrasegaran, S. (2005). Zinc finger nucleases: Custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res.* 33, 5978–5990. doi:10.1093/nar/gki912

Epinat, J. C., Arnould, S., Chames, P., Rochaix, P., Desfontaines, D., Puzin, C., et al. (2003). A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res.* 31, 2952–2962. doi:10.1093/nar/gkg375

Feng, Y., Zhang, S., and Huang, X. (2014). A robust TALENs system for highly efficient mammalian genome editing. *Sci. Rep.* 4, 3632. doi:10.1038/srep03632

Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and β-thalassemia. *N. Engl. J. Med.* 384, 252–260. doi:10.1056/NEJMoa2031054

Fu, J., Li, Q., Liu, X., Tu, T., Lv, X., Yin, X., et al. (2021). Human cell based directed evolution of adenine base editors with improved efficiency. *Nat. Commun.* 12, 5897. doi:10.1038/s41467-021-26211-0

Gaj, T., Guo, J., Kato, Y., Sirk, S. J., and Barbas, C. F. (2012). Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat. Methods* 9, 805–807. doi:10.1038/nmeth.2030

Gao, H., Smith, J., Yang, M., Jones, S., Djukanovic, V., Nicholson, M. G., et al. (2010). Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J.* 61, 176–187. doi:10.1111/j.1365-313X.2009.04041.x

Gao, Z., Herrera-Carrillo, E., and Berkhout, B. (2019). A single H1 promoter can drive both guide RNA and endonuclease expression in the CRISPR-cas9 system. *Mol. Ther. Nucleic Acids* 14, 32–40. doi:10.1016/j.omtn.2018.10.016

Gapinske, M., Luu, A., Winter, J., Woods, W. S., Kostan, K. A., Shiva, N., et al. (2018). CRISPR-SKIP: Programmable gene splicing with single base editors. *Genome Biol.* 19, 107. doi:10.1186/s13059-018-1482-5

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi:10.1038/nature24644

Gehrke, J. M., Cervantes, O., Clement, M. K., Wu, Y., Zeng, J., Bauer, D. E., et al. (2018). An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* 36, 977–982. doi:10.1038/nbt.4199

Georgiadis, C., Rasaiyaah, J., Gkazi, S. A., Preece, R., Etuk, A., Christi, A., et al. (2021). Base-edited CAR T cells for combinational therapy against T cell malignancies. *Leukemia* 35, 3466–3481. doi:10.1038/s41375-021-01282-6

Ghosh, R., Das, D., and Franco, S. (2018). "The role for the DSB response pathway in regulating chromosome translocations65–87. doi:10.1007/978-981-13-0593-1_6

Gillmore, J. D., Gane, E., Taubel, J., Kao, J., Fontana, M., Maitland, M. L., et al. (2021). CRISPR-Cas9 *in vivo* gene editing for transthyretin amyloidosis. *N. Engl. J. Med.* 385, 493–502. doi:10.1056/NEJMoa2107454

Gouble, A., Smith, J., Bruneau, S., Perez, C., Guyot, V., Cabaniols, J. P., et al. (2006). Efficient *in toto* targeted recombination in mouse liver by meganuclease-induced double-strand break. *J. Gene Med.* 8, 616–622. doi:10.1002/jgm.879

Gutschner, T., Baas, M., and Diederichs, S. (2011). Noncoding RNA gene silencing through genomic integration of RNA destabilizing elements using zinc finger nucleases. *Genome Res.* 21, 1944–1954. doi:10.1101/gr.122358.111

Haapaniemi, E., Botla, S., Persson, J., Schmierer, B., and Taipale, J. (2018). CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* 24, 927–930. doi:10.1038/s41591-018-0049-z

Habib, O., Habib, G., Hwang, G. H., and Bae, S. (2022). Comprehensive analysis of prime editing outcomes in human embryonic stem cells. *Nucleic Acids Res.* 50, 1187–1197. doi:10.1093/nar/gkab1295

Happi Mbakam, C., Rousseau, J., Tremblay, G., Yameogo, P., and Tremblay, J. P. (2022). Prime editing permits the introduction of specific mutations in the gene responsible for duchenne muscular dystrophy. *Int. J. Mol. Sci.* 23, 6160. doi:10.3390/ijms23116160

Hoban, M. D., Cost, G. J., Mendel, M. C., Romero, Z., Kaufman, M. L., Joglekar, A. v., et al. (2015). Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* 125, 2597–2604. doi:10.1182/blood-2014-12-615948

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63. doi:10.1038/nature26155

Hu, Z., Wang, Y., Liu, Q., Qiu, Y., Zhong, Z., Li, K., et al. (2021). Improving the precision of base editing by bubble hairpin single guide RNA. *MBio* 12, e00342-21. doi:10.1128/mBio.00342-21

Huang, T. P., Zhao, K. T., Miller, S. M., Gaudelli, N. M., Oakes, B. L., Fellmann, C., et al. (2019). Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat. Biotechnol.* 37, 626–631. doi:10.1038/s41587-019-0134-y

Hwang, G. H., and Bae, S. (2021). Web-based base editing toolkits: BE-Designer and BE-analyzer. *Methods Mol. Biol.* 2189, 81–88. doi:10.1007/978-1-0716-0822-7_7

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., et al. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 227–229. doi:10.1038/nbt.2501

Ihry, R. J., Worringer, K. A., Salick, M. R., Frias, E., Ho, D., Theriault, K., et al. (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 24, 939–946. doi:10.1038/s41591-018-0050-6

Iliakis, G., Wang, H., Perrault, A. R., Boecker, W., Rosidi, B., Windhofer, F., et al. (2004). Mechanisms of DNA double strand break repair and chromosome aberration formation. *Cytogenet. Genome Res.* 104, 14–20. doi:10.1159/000077461

Izmiryan, A., Danos, O., and Hovnanian, A. (2016). Meganuclease-mediated COL7A1 gene correction for recessive dystrophic epidermolysis bullosa. *J. Investig. Dermatol.* 136, 872–875. doi:10.1016/j.jid.2015.11.028

Jalligampala, archana, Noel, Jennifer, JamesFransen, W., Wang, Wei, Jabbar, Maha, Hasan, Nazarul, et al. (2021). *Rho 1-2 meganuclease, an allele-specific gene-editing therapy, rejuvenates rod photoreceptor structure and function in a pig model of autosomal dominant Retinitis Pigmentosa (adRP)*. Investigative Ophthalmology & Visual Science, 1478.

Jang, H., Jo, D. H., Cho, C. S., Shin, J. H., Seo, J. H., Yu, G., et al. (2022). Application of prime editing to the correction of mutations and phenotypes in adult mice with liver and eye diseases. *Nat. Biomed. Eng.* 6, 181–194. doi:10.1038/s41551-021-00788-9

Jang, H. K., Jo, D. H., Lee, S. N., Cho, C. S., Jeong, Y. K., Jung, Y., et al. (2021). High-purity production and precise editing of DNA base editing ribonucleoproteins. *Sci. Adv.* 7, eabg2661. doi:10.1126/sciadv.abg2661

Ji, H., Lu, P., Liu, B., Qu, X., Wang, Y., Jiang, Z., et al. (2018). Zinc-finger nucleases induced by HIV-1 tat excise HIV-1 from the host genome in infected and latently infected cells. *Mol. Ther. Nucleic Acids* 12, 67–74. doi:10.1016/j.omtn.2018.04.014

Jiang, T., Zhang, X. O., Weng, Z., and Xue, W. (2022). Deletion and replacement of long genomic sequences using prime editing. *Nat. Biotechnol.* 40, 227–234. doi:10.1038/s41587-021-01026-y

Jiang, W., Feng, S., Huang, S., Yu, W., Li, G., Yang, G., et al. (2018). BE-PLUS: A new base editing tool with broadened editing window and enhanced fidelity. *Cell. Res.* 28, 855–861. doi:10.1038/s41422-018-0052-4

Jiang, Y. Y., Chai, Y. P., Lu, M. H., Han, X. L., Lin, Q., Zhang, Y., et al. (2020). Prime editing efficiently generates W542L and S621I double mutations in two ALS genes in maize. *Genome Biol.* 21, 257. doi:10.1186/s13059-020-02170-5

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., Charpentier, E., et al. (2012a1979). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi:10.1126/science.1225829

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., Charpentier, E., et al. (2012b1979). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi:10.1126/science.1225829

Khatibi, S., Sahebkar, A., and Aghaee-Bakhtiari, S. H. (2021). CRISPR genome editing technology and its application in genetic diseases: A review. *Curr. Pharm. Biotechnol.* 22, 468–479. doi:10.2174/1389201021666200621161610

Kim, D., Kang, B. C., and Kim, J. S. (2021a). Identifying genome-wide off-target sites of CRISPR RNA-guided nucleases and deaminases with Digenome-seq. *Nat. Protoc.* 16, 1170–1192. doi:10.1038/s41596-020-00453-6

Kim, H. J., Oh, S. Y., and Lee, S. J. (2020). Single-base genome editing in corynebacterium glutamicum with the help of negative selection by target-mismatched CRISPR/Cpf1. *J. Microbiol. Biotechnol.* 30, 1583–1591. doi:10.4014/jmb.2006.06036

Kim, H. K., Yu, G., Park, J., Min, S., Lee, S., Yoon, S., et al. (2021b). Predicting the efficiency of prime editing guide RNAs in human cells. *Nat. Biotechnol.* 39, 198–206. doi:10.1038/s41587-020-0677-y

Kim, Y. B., Komor, A. C., Levy, J. M., Packer, M. S., Zhao, K. T., Liu, D. R., et al. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* 35, 371–376. doi:10.1038/nbt.3803

Kim, Y., Hong, S.-A., Yu, J., Eom, J., Jang, K., Yoon, S., et al. (2021c). Adenine base editing and prime editing of chemically derived hepatic progenitors rescue genetic liver disease. *Cell. Stem Cell.* 28, 1614–1624.e5. e5. doi:10.1016/j.stem.2021.04.010

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., et al. (2016b). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495. doi:10.1038/nature16526

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., et al. (2016a). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495. doi:10.1038/nature16526

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. v., Nguyen, N. T., Zheng, Z., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485. doi:10.1038/nature14592

Kocak, D. D., Josephs, E. A., Bhandarkar, V., Adkar, S. S., Kwon, J. B., Gersbach, C. A., et al. (2019). Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat. Biotechnol.* 37, 657–666. doi:10.1038/s41587-019-0095-1

Komor, A. C. E., Zhao, K. T., Packer, M. S., Gaudelli, N. M., Waterbury, A. L., Koblan, L. W., et al. (2017). Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci. Adv.* 3, eaao4774. doi:10.1126/sciadv.aao4774

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi:10.1038/nature17946

Kurt, I. C., Zhou, R., Iyer, S., Garcia, S. P., Miller, B. R., Langner, L. M., et al. (2021). CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* 39, 41–46. doi:10.1038/s41587-020-0609-x

Kweon, J., Yoon, J. K., Jang, A. H., Shin, H. R., See, J. E., Jang, G., et al. (2021). Engineered prime editors with PAM flexibility. *Mol. Ther.* 29, 2001–2007. doi:10.1016/j.ymthe.2021.02.022

Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., et al. (2016). ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 44, D862–D868. doi:10.1093/nar/gkv1222

Lee, H. J., Kim, H. J., and Lee, S. J. (2020). CRISPR-Cas9-mediated pinpoint microbial genome editing aided by target-mismatched sgRNAs. *Genome Res.* 30, 768–775. doi:10.1101/gr.257493.119

Lee, K. Z., Mechikoff, M. A., Parasa, M. K., Rankin, T. J., Pandolfi, P., Fitzgerald, K. S., et al. (2022). Repurposing the homing endonuclease I-SceI for positive selection and development of gene-editing technologies. *ACS Synth. Biol.* 11, 53–60. doi:10.1021/acssynbio.1c00340

Li, H., Li, J., Chen, J., Yan, L., and Xia, L. (2020). Precise modifications of both exogenous and endogenous genes in rice by prime editing. *Mol. Plant* 13, 671–674. doi:10.1016/j.molp.2020.03.011

Lin, Q., Jin, S., Zong, Y., Yu, H., Zhu, Z., Liu, G., et al. (2021). High-efficiency prime editing with optimized, paired pegRNAs in plants. *Nat. Biotechnol.* 39, 923–927. doi:10.1038/s41587-021-00868-w

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., et al. (2020). Prime genome editing in rice and wheat. *Nat. Biotechnol.* 38, 582–585. doi:10.1038/s41587-020-0455-x

Liu, G., Lin, Q., Jin, S., and Gao, C. (2021a). The CRISPR-Cas toolbox and gene editing technologies. *Mol. Cell.* 82, 333–347. doi:10.1016/j.molcel.2021.12.002

Liu, M.-S., Gong, S., Yu, H.-H., Jung, K., Johnson, K. A., Taylor, D. W., et al. (2020). Engineered CRISPR/Cas9 enzymes improve discrimination by slowing DNA cleavage to allow release of off-target DNA. *Nat. Commun.* 11, 3576. doi:10.1038/s41467-020-17411-1

Liu, P., Liang, S. Q., Zheng, C., Mintzer, E., Zhao, Y. G., Ponnienselvan, K., et al. (2021b). Improved prime editors enable pathogenic allele correction and cancer modelling in adult mice. *Nat. Commun.* 12, 2121. doi:10.1038/s41467-021-22295-w

Liu, Z., Chen, M., Chen, S., Deng, J., Song, Y., Lai, L., et al. (2018). Highly efficient RNA-guided base editing in rabbit. *Nat. Commun.* 9, 2717. doi:10.1038/s41467-018-05232-2

Lo, T. W., Pickle, C. S., Lin, S., Ralston, E. J., Gurling, M., Schartner, C. M., et al. (2013). Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. *Genetics* 195, 331–348. -+. doi:10.1534/genetics.113.155382

Lu, Y., Xue, J., Deng, T., Zhou, X., Yu, K., Deng, L., et al. (2020). Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer. *Nat. Med.* 26, 732–740. doi:10.1038/s41591-020-0840-5

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736. doi:10.1038/nrmicro3569

Mao, Z., Bozzella, M., Seluanov, A., and Gorbunova, V. (2008). DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell. Cycle* 7, 2902–2906. doi:10.4161/cc.7.18.6679

Mariano, A., Xu, L., and Han, R. (2014). Highly efficient genome editing via 2A-coupled co-expression of two TALEN monomers. *BMC Res. Notes* 7, 628. doi:10.1186/1756-0500-7-628

Martín-Fernández, J. M., Fleischer, A., Vallejo-Diez, S., Palomino, E., Sánchez-Gilabert, A., Ruiz, R., et al. (2020). New bicistronic TALENs greatly improve genome editing. *Curr. Protoc. Stem Cell. Biol.* 52, e104. doi:10.1002/cpsc.104

Masuda, T., Wang, X., Maeda, M., Canver, M. C., Sher, F., Funnell, A. P. W., et al. (2016). Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science* 351, 285–289. doi:10.1126/science.aad3312

Maximiano, M. R., Tavora, F., Prado, G. S., Dias, S. C., Mehta, A., Franco, O. L., et al. (2021). CRISPR genome editing technology: A powerful tool applied to developing agribusiness. *J. Agric. Food Chem.* 69, 6379–6395. doi:10.1021/acs.jafc.1c01062

Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y. L., Rupniewski, I., et al. (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* 25, 778–785. doi:10.1038/nbt1319

Miller, J. C., Patil, D. P., Xia, D. F., Paine, C. B., Fauser, F., Richards, H. W., et al. (2019a). Enhancing gene editing specificity by attenuating DNA cleavage kinetics. *Nat. Biotechnol.* 37, 945–952. doi:10.1038/s41587-019-0186-z

Miller, J. C., Patil, D. P., Xia, D. F., Paine, C. B., Fauser, F., Richards, H. W., et al. (2019b). Enhancing gene editing specificity by attenuating DNA cleavage kinetics. *Nat. Biotechnol.* 37, 945–952. doi:10.1038/s41587-019-0186-z

Miyaoka, Y., Chan, A. H., and Conklin, B. R. (2016). Detecting single-nucleotide substitutions induced by genome editing. *Cold Spring Harb. Protoc.*, pdb.top090845. doi:10.1101/pdb.top090845

Modarai, S. R., Kanda, S., Bloh, K., Opdenaker, L. M., and Kmiec, E. B. (2021). Precise and error-prone CRISPR-directed gene editing activity in human CD34+ cells varies widely among patient samples. *Gene Ther.* 28, 105–113. doi:10.1038/s41434-020-00192-z

Molina, R., Besker, N., Marcaida, M. J., Montoya, G., Prieto, J., D'Abramo, M., et al. (2016). Key players in I-DmoI endonuclease catalysis revealed from structure and dynamics. *ACS Chem. Biol.* 11, 1401–1407. doi:10.1021/acschembio.5b00730

Moscou, M. J., and Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501. doi:10.1126/science.1178817

Muenzer, J., Prada, C. E., Burton, B., Lau, H. A., Ficicioglu, C., Foo, C. W. P., et al. (2019). Champions: A phase 1/2 clinical trial with dose escalation of SB-913 ZFN-mediated *in vivo* human genome editing for treatment of mps II (hunter syndrome). *Mol. Genet. Metabolism* 126, S104. doi:10.1016/j.ymgme.2018.12.263

Mussolino, C., Alzubi, J., Fine, E. J., Morbitzer, R., Cradick, T. J., Lahaye, T., et al. (2014). TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res.* 42, 6762–6773. doi:10.1093/nar/gku305

Nelson, J. W., Randolph, P. B., Shen, S. P., Everette, K. A., Chen, P. J., Anzalone, A. v., et al. (2021). Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* 40, 402–410. doi:10.1038/s41587-021-01039-7

Nelson, J. W., Randolph, P. B., Shen, S. P., Everette, K. A., Chen, P. J., Anzalone, A. v., et al. (2022a). Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* 40, 402–410. doi:10.1038/s41587-021-01039-7

Nelson, J. W., Randolph, P. B., Shen, S. P., Everette, K. A., Chen, P. J., Anzalone, A. v., et al. (2022b). Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* 40, 402–410. doi:10.1038/s41587-021-01039-7

Newby, G. A., Yen, J. S., Woodard, K. J., Mayuranathan, T., Lazzarotto, C. R., Li, Y., et al. (2021). Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature* 595, 295–302. doi:10.1038/s41586-021-03609-w

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262. doi:10.1126/science.aas9129

Paques, F., and Duchateau, P. (2007). Meganucleases and DNA double-strand break-induced recombination: Perspectives for gene therapy. *Curr. Gene Ther.* 7, 49–66. doi:10.2174/156652307779940216

Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., et al. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125-129doi:10.1038/nature17664

Paschon, D. E., Lussier, S., Wangzor, T., Xia, D. F., Li, P. W., Hinkley, S. J., et al. (2019). Diversifying the structure of zinc finger nucleases for high-precision genome editing. *Nat. Commun.* 10, 1133. doi:10.1038/s41467-019-08867-x

Pattanayak, V., Ramirez, C. L., Joung, J. K., and Liu, D. R. (2011). Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. *Nat. Methods* 8, 765–770. doi:10.1038/nmeth.1670

Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., et al. (2013). RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nat. Methods* 10, 973–976. doi:10.1038/nmeth.2600

Petersen, B., and Niemann, H. (2015a). Advances in genetic modification of farm animals using zinc-finger nucleases (ZFN). *Chromosome Res.* 23, 7–15. doi:10.1007/s10577-014-9451-7

Petersen, B., and Niemann, H. (2015b). Molecular scissors and their application in genetically modified farm animals. *Transgenic Res.* 24, 381–396. doi:10.1007/s11248-015-9862-z

Petri, K., Zhang, W., Ma, J., Schmidts, A., Lee, H., Horng, J. E., et al. (2022). CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat. Biotechnol.* 40, 189–193. doi:10.1038/s41587-021-00901-y

Philippidis, A. (2021). First positive clinical data for *in vivo* genome editing in humans opens "new era of medicine. *Hum. Gene Ther.* 32, 783–786. doi:10.1089/hum.2021.29170.bfs

Porter, S. N., Levine, R. M., and Pruett-Miller, S. M. (2019). A practical guide to genome editing using targeted nuclease technologies. *Compr. Physiol.* 9, 665–714. doi:10.1002/cphy.c180022

Porteus, M. H., and Baltimore, D. (2003). Chimeric nucleases stimulate gene targeting in human cells. *Science* 300, 763. doi:10.1126/science.1078395

Prieto, J., Redondo, P., Lopez-Mendez, B., D'Abramo, M., Merino, N., Blanco, F. J., et al. (2018). Understanding the indirect DNA read-out specificity of I-CreI Meganuclease. *Sci. Rep.* 8, 10286. doi:10.1038/s41598-018-28599-0

Prieto, J., Redondo, P., Padró, D., Arnould, S., Epinat, J.-C., Pâques, F., et al. (2007). The C-terminal loop of the homing endonuclease I-CreI is essential for site recognition, DNA binding and cleavage. *Nucleic Acids Res.* 35, 3262–3271. doi:10.1093/nar/gkm183

Psatha, N., Reik, A., Phelps, S., Zhou, Y., Dalas, D., Yannaki, E., et al. (2018). Disruption of the BCL11A erythroid enhancer reactivates fetal hemoglobin in erythroid cells of patients with β-thalassemia major. *Mol. Ther. Methods Clin. Dev.* 10, 313–326. doi:10.1016/j.omtm.2018.08.003

Qasim, W., Zhan, H., Samarasinghe, S., Adams, S., Amrolia, P., Stafford, S., et al. (2017). Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* 9, eaaj2013. doi:10.1126/scitranslmed.aaj2013

Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., et al. (2013a). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell.* 154, 1380–1389. doi:10.1016/j.cell.2013.08.021

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., Zhang, F., et al. (2013b). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308. doi:10.1038/nprot.2013.143

Ranjha, L., Howard, S. M., and Cejka, P. (2018). Main steps in DNA double-strand break repair: An introduction to homologous recombination and related processes. *Chromosoma* 127, 187–214. doi:10.1007/s00412-017-0658-1

Rees, H. A., Komor, A. C., Yeh, W. H., Caetano-Lopes, J., Warman, M., Edge, A. S. B., et al. (2017). Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat. Commun.* 8, 15790. doi:10.1038/ncomms15790

Reyon, D., Khayter, C., Regan, M. R., Joung, J. K., and Sander, J. D. (2012). Engineering designer transcription activator--like effector nucleases (TALENs) by REAL or REAL-fast assembly. *Curr. Protoc. Mol. Biol.* 100, Unit 12.15. doi:10.1002/0471142727.mb1215s100

Rouet, P., Smih, F., and Jasin, M. (1994a). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 91, 6064–6068. doi:10.1073/pnas.91.13.6064

Rouet, P., Smih, F., and Jasin, M. (1994b). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.* 14, 8096–8106. doi:10.1128/Mcb.14.12.8096

Samuelson, C., Radtke, S., Zhu, H., Llewellyn, M., Fields, E., Cook, S., et al. (2021). Multiplex CRISPR/Cas9 genome editing in hematopoietic stem cells for fetal hemoglobin reinduction generates chromosomal translocations. *Mol. Ther. Methods Clin. Dev.* 23, 507–523. doi:10.1016/j.omtm.2021.10.008

Sanjana, N. E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G., Zhang, F., et al. (2012). A transcription activator-like effector toolbox for genome engineering. *Nat. Protoc.* 7, 171–192. doi:10.1038/nprot.2011.431

Santiago, Y., Chan, E., Liu, P.-Q., Orlando, S., Zhang, L., Urnov, F. D., et al. (2008). Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* 105, 5809–5814. doi:10.1073/pnas.0800940105

Scholze, H., and Boch, J. (2011). TAL effectors are remote controls for gene activation. *Curr. Opin. Microbiol.* 14, 47–53. doi:10.1016/j.mib.2010.12.001

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., et al. (2014a1979). Genome-scale CRISPR-cas9 knockout screening in human cells. *Science* 343, 84–87. doi:10.1126/science.1247005

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelson, T., et al. (2014b1979). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84–87. doi:10.1126/science.1247005

Siegl, T., Petzke, L., Welle, E., and Luzhetskyy, A. (2010). I-SceI endonuclease: A new tool for DNA repair studies and genetic manipulations in streptomyces. *Appl. Microbiol. Biotechnol.* 87, 1525–1532. doi:10.1007/s00253-010-2643-y

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., Zhang, F., et al. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88. doi:10.1126/science.aad5227

Smith, A. R., Schiller, G. J., Vercellotti, G. M., Kwiatkowski, J. L., Krishnamurti, L., Esrick, E. B., et al. (2019). Preliminary results of a phase 1/2 clinical study of zinc finger nuclease-mediated editing of BCL11A in autologous hematopoietic stem cells for transfusion-dependent beta thalassemia. *Blood* 134, 3544. doi:10.1182/blood-2019-125743

Stadtmauer, E. A., Fraietta, J. A., Davis, M. M., Cohen, A. D., Weber, K. L., Lancaster, E., et al. (2020). CRISPR-engineered T cells in patients with refractory cancer. *Science*, eaba7365. doi:10.1126/science.aba7365

Stoddard, B. L. (2011). Homing endonucleases: From microbial genetic invaders to reagents for targeted DNA modification. *Structure* 19, 7–15. doi:10.1016/j.str.2010.12.003

Sun, Z., Wang, M., Han, S., Ma, S., Zou, Z., Ding, F., et al. (2018). Production of hypoallergenic milk from DNA-free beta-lactoglobulin (BLG) gene knockout cow using zinc-finger nucleases mRNA. *Sci. Rep.* 8, 15430. doi:10.1038/s41598-018-32024-x

Surun, D., Schneider, A., Mircetic, J., Neumann, K., Lansing, F., Paszkowski-Rogacz, M., et al. (2020). *Efficient Generation and Correction of Mutations in Human iPS Cells Utilizing mRNAs of CRISPR Base Editors and Prime*, 11. doi:10.3390/genes11050511*Genes. (Basel)*

Suzuki, S., Ohta, K., Nakajima, Y., Shigeto, H., Abe, H., Kawai, A., et al. (2020). Meganuclease-based artificial transcription factors. *ACS Synth. Biol.* 9, 2679–2691. doi:10.1021/acssynbio.0c00083

Tan, J., Zhang, F., Karcher, D., and Bock, R. (2019). Engineering of high-precision base editors for site-specific single nucleotide replacement. *Nat. Commun.* 10, 439. doi:10.1038/s41467-018-08034-8

Tang, L. C., and Gu, F. (2020). Next-generation CRISPR-cas for genome editing: Focusing on the Cas protein and PAM. *Yi Chuan* 42, 236–249. doi:10.16288/j.yczz.19-297

Tang, Y., and Fu, Y. (2018). Class 2 CRISPR/Cas: An expanding biotechnology toolbox for and beyond genome editing. *Cell. Biosci.* 8, 59. doi:10.1186/s13578-018-0255-x

Tebas, P., Jadlowsky, J. K., Shaw, P. A., Tian, L., Esparza, E., Brennan, A. L., et al. (2021). CCR5-edited CD4+ T cells augment HIV-specific immunity to enable post-rebound control of HIV replication. *J. Clin. Investig.* 131, 144486. doi:10.1172/JCI144486

Tebas, P., Stein, D., Tang, W. W., Frank, I., Wang, S. Q., Lee, G., et al. (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* 370, 901–910. doi:10.1056/NEJMoa1300662

Tekel, S. J., Brookhouser, N., Standage-Beier, K., Wang, X., and Brafman, D. A. (2021). Cytosine and adenosine base editing in human pluripotent stem cells using transient reporters for editing enrichment. *Nat. Protoc.* 16, 3596–3624. doi:10.1038/s41596-021-00552-y

Tremblay, G., Rousseau, J., Mbakam, C. H., and Tremblay, J. P. (2022). Insertion of the Icelandic mutation (A673T) by prime editing: A potential preventive treatment for familial and sporadic alzheimer's disease. *CRISPR J.* 5, 109–122. doi:10.1089/crispr.2021.0085

Tzfira, T., Frankman, L. R., Vaidya, M., and Citovsky, V. (2003). Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiol.* 133, 1011–1023. doi:10.1104/pp.103.032128

Uddin, F., Rudin, C. M., and Sen, T. (2020). CRISPR gene therapy: Applications, limitations, and implications for the future. *Front. Oncol.* 10, 1387. doi:10.3389/fonc.2020.01387

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., and Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11, 636–646. doi:10.1038/nrg2842

Vakulskas, C. A., and Behlke, M. A. (2019). Evaluation and reduction of CRISPR off-target cleavage events. *Nucleic Acid. Ther.* 29, 167–174. doi:10.1089/nat.2019.0790

Valerie, K., and Povirk, L. F. (2003). Regulation and mechanisms of mammalian double-strand break repair. *Oncogene* 22, 5792–5812. doi:10.1038/sj.onc.1206679

Walton, R. T., Christie, K. A., Whittaker, M. N., and Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296. doi:10.1126/science.aba8853

Wang, F., Zeng, Y., Wang, Y., and Niu, Y. (2020a). The development and application of a base editor in biomedicine. *Biomed. Res. Int.* 2020, 2907623. doi:10.1155/2020/2907623

Wang, J., He, Z., Wang, G., Zhang, R., Duan, J., Gao, P., et al. (2022a). Efficient targeted insertion of large DNA fragments without DNA donors. *Nat. Methods* 19, 331–340. doi:10.1038/s41592-022-01399-1

Wang, J., Lu, C., and Wei, S. (2022b). *Whole-genome sequencing identifies I-SceI-mediated transgene integration sites in Xenopus tropicalis snai2: eGFP line*. *G3 (Bethesda)*. doi:10.1093/g3journal/jkac037

Wang, L., Breton, C., Warzecha, C. C., Bell, P., Yan, H., He, Z., et al. (2021a). Long-term stable reduction of low-density lipoprotein in nonhuman primates following *in vivo* genome editing of PCSK9. *Mol. Ther.* 29, 2019–2029. doi:10.1016/j.ymthe.2021.02.020

Wang, L. R., Li, L. X., Ma, Y. L., Hu, H. D., Li, Q., Yang, Y., et al. (2020b). Reactivation of gamma-globin expression through Cas9 or base editor to treat beta-hemoglobinopathies. *Cell. Res.* 30, 276–278. doi:10.1038/s41422-019-0267-z

Wang, L., Smith, J., Breton, C., Clark, P., Zhang, J., Ying, L., et al. (2018a). Meganuclease targeting of PCSK9 in macaque liver leads to stable reduction in serum cholesterol. *Nat. Biotechnol.* 36, 717–725. doi:10.1038/nbt.4182

Wang, X., Li, J., Wang, Y., Yang, B., Wei, J., Wu, J., et al. (2018b). Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. *Nat. Biotechnol.* 36, 946–949. doi:10.1038/nbt.4198

Wang, Y., Wang, H., Jian, Y., Luo, Z., Shao, H., and Zhang, W. (2021b). Strategies for optimization of the CRISPR-based genome editing system for enhanced editing specificity. *Hum. Gene Ther.* doi:10.1089/hum.2021.283

Woo, J. W., Kim, J., Kwon, S. il, Corvalán, C., Cho, S. W., Kim, H., et al. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat. Biotechnol.* 33, 1162–1164. doi:10.1038/nbt.3389

Xiao, T., Wu, B., Cao, Y., Liu, R., Cheng, G., Wang, L., et al. (2021). Genetic identification of pathogenic variations of the DMD gene: A retrospective study from 10, 481 neonatal patients based on next-generation sequencing data. *Ann. Transl. Med.* 9, 766. doi:10.21037/atm-20-7102

Xiong, K., Li, S., Zhang, H., Cui, Y., Yu, D., Li, Y., et al. (2013). Targeted editing of goat genome with modular-assembly zinc finger nucleases based on activity prediction by computational molecular modeling. *Mol. Biol. Rep.* 40, 4251–4256. doi:10.1007/s11033-013-2507-5

Xu, W., Zhang, C., Yang, Y., Zhao, S., Kang, G., He, X., et al. (2020). Versatile nucleotides substitution in plant using an improved prime editing system. *Mol. Plant* 13, 675–678. doi:10.1016/j.molp.2020.03.012

Yi, L., and Li, J. (2016). CRISPR-Cas9 therapeutics in cancer: Promising strategies and present challenges. *Biochim. Biophys. Acta* 1866, 197–207. doi:10.1016/j.bbcan.2016.09.002

Zekonyte, U., Bacman, S. R., Smith, J., Shoop, W., Pereira, C. v., Tomberlin, G., et al. (2021). Mitochondrial targeted meganuclease as a platform to eliminate mutant mtDNA *in vivo*. *Nat. Commun.* 12, 3210. doi:10.1038/s41467-021-23561-7

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A. H., Shea, D., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* 26, 535–541. doi:10.1038/s41591-020-0790-y

Zhang, H., Pan, H., Zhou, C., Wei, Y., Ying, W., Li, S., et al. (2018). Simultaneous zygotic inactivation of multiple genes in mouse through CRISPR/Cas9-mediated base editing. *Development* 145, dev168906. doi:10.1242/dev.168906

Zhang, M., Zhou, C., Wei, Y., Xu, C., Pan, H., Ying, W., et al. (2019). Human cleaving embryos enable robust homozygotic nucleotide substitutions by base editors. *Genome Biol.* 20, 101. doi:10.1186/s13059-019-1703-6

Zhang, S., Wang, J., and Wang, J. (2020). One-day TALEN assembly protocol and a dual-tagging system for genome editing. *ACS Omega* 5, 19702–19714. doi:10.1021/acsomega.0c02396

Zhang, Y., Liu, L., Guo, S., Song, J., Zhu, C., Yue, Z., et al. (2017). Deciphering TAL effectors for 5-methylcytosine and 5-hydroxymethylcytosine recognition. *Nat. Commun.* 8, 901. doi:10.1038/s41467-017-00860-6

Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M. A., et al. (2021a). Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* 39, 35–40. doi:10.1038/s41587-020-0592-2

Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M. A., et al. (2021b). Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* 39, 35–40. doi:10.1038/s41587-020-0592-2

Zheng, C. W., Liang, S. Q., Liu, B., Liu, P. P., Kwan, S. Y., Wolfe, S. A., et al. (2022). A flexible split prime editor using truncated reverse transcriptase improves dual-AAV delivery in mouse liver. *Mol. Ther.* 30, 1343–1351. doi:10.1016/j.ymthe.2022.01.005

Zhi, S., Chen, Y., Wu, G., Wen, J., Wu, J., Liu, Q., et al. (2022). Dual-AAV delivering split prime editor system for *in vivo* genome editing. *Mol. Ther.* 30, 283–294. doi:10.1016/j.ymthe.2021.07.011

Zhou, C., Sun, Y., Yan, R., Liu, Y., Zuo, E., Gu, C., et al. (2019). Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* 571, 275–278. doi:10.1038/s41586-019-1314-0

Zhu, C., Bortesi, L., Baysal, C., Twyman, R. M., Fischer, R., Capell, T., et al. (2017). Characteristics of genome editing mutations in cereal crops. *Trends Plant Sci.* 22, 38–52. doi:10.1016/j.tplants.2016.08.009

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 10 August 2022
DOI 10.3389/fbioe.2022.942440

OPEN ACCESS

EDITED BY
Ayal Hendel,
Bar-Ilan University, Israel

REVIEWED BY
Sjaak Philipsen,
Erasmus Medical Center, Netherlands
Eric Kmiec,
Christiana Care Health System, United States
Donald B. Kohn,
University of California, Los Angeles, United States

*CORRESPONDENCE
Elena I. Leonova,
e.leonova@spbu.ru

SPECIALTY SECTION
This article was submitted to Preclinical Cell and Gene Therapy, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 12 May 2022
ACCEPTED 14 July 2022
PUBLISHED 10 August 2022

CITATION
Reshetnikov VV, Chirinskaite AV, Sopova JV, Ivanov RA and Leonova EI (2022), Translational potential of base-editing tools for gene therapy of monogenic diseases.
Front. Bioeng. Biotechnol. 10:942440.
doi: 10.3389/fbioe.2022.942440

Translational potential of base-editing tools for gene therapy of monogenic diseases

Vasiliy V. Reshetnikov[1,2], Angelina V. Chirinskaite[3], Julia V. Sopova[3,4], Roman A. Ivanov[1] and Elena I. Leonova[3,5]*

[1]Department of Biotechnology, Sirius University of Science and Technology, Sochi, Russia, [2]Department of Molecular Genetics, Institute of Cytology and Genetics, Novosibirsk, Russia, [3]Center of Transgenesis and Genome Editing, St. Petersburg State University, St. Petersburg, Russia, [4]Laboratory of Amyloid Biology, St. Petersburg State University, St. Petersburg, Russia, [5]Scientific Center for Genetics and Life Sciences, Sirius University of Science and Technology, Sochi, Russia

Millions of people worldwide have rare genetic diseases that are caused by various mutations in DNA sequence. Classic treatments of rare genetic diseases are often ineffective, and therefore great hopes are placed on gene-editing methods. A DNA base–editing system based on nCas9 (Cas9 with a nickase activity) or dCas9 (a catalytically inactive DNA-targeting Cas9 enzyme) enables editing without double-strand breaks. These tools are constantly being improved, which increases their potential usefulness for therapies. In this review, we describe the main types of base-editing systems and their application to the treatment of monogenic diseases in experiments *in vitro* and *in vivo*. Additionally, to understand the therapeutic potential of these systems, the advantages and disadvantages of base-editing systems are examined.

KEYWORDS

base editing, dCas9, nCas9, prime editor, monogenic disease, gene therapy

Therapy for genetic diseases

To date, ~6,500 genetic diseases with annotated phenotypes have been identified and affect more than 100 million people worldwide (Nguengang Wakap et al., 2020). Treatments of such diseases have been limited mostly to symptomatic and supportive care. The main interventions into metabolic disorders are aimed at substrate restriction, replacement of deficient products, and inhibition of formation and removal of toxic metabolites (Yue et al., 2019). Enzyme replacement therapy, pharmacotherapy, and hematopoietic stem cell transplantation are used for these purposes (Chen and Altman, 2017; Li, 2018; Taylor et al., 2019). Genetic therapeutic strategies include gene replacement therapy, which requires targeted transfer of exogenous genetic material into human cells; mRNA correction (an antisense oligonucleotide, small interfering RNA, microRNA, or RNA editing); *cis*-regulation therapy; and gene-editing technology (Chen W. et al., 2020; Matharu and Ahituv, 2020; Reshetnikov et al., 2022). Recent advances in gene therapy are based on the use of nucleases such as ZFN, TALENS, and Cas9, which can precisely introduce double-strand breaks, that are repaired by the cell's repair systems

(Gaj et al., 2013; Guo et al., Forthcoming 2021). Nevertheless, the numerous off-target effects associated with DSBs, delivery challenges, and immunogenicity preclude the use of these tools in clinical practice (Cui et al., 2021; Guo et al., Forthcoming 2021), despite it has been used in clinical trials (Ou et al., 2020; Frangoul et al., 2021).

Recent advances in gene-editing technology made it possible to edit DNA without a DSB. This approach became feasible after a catalytically inactive DNA-targeting Cas9 enzyme (dCas9) was obtained, which together with single guide RNAs allows to localize effector domains to specific DNA sequences to either repress (CRISPRi) or activate (CRISPRa) transcription of a target gene(s) (Gilbert et al., 2014). CRISPRa has been successfully employed to treat diseases and eliminate haploinsufficiency in mice (Matharu et al., 2019; Colasante et al., 2020). Aside from the inactive Cas9, nCas9 has been obtained, which has a nickase activity and can create only a single-strand break at target sites (Cong et al., 2013). Fusion of nCas9 and APOBEC1 cytidine deaminase or TadA adenine deaminase has helped to devise cytosine and adenine base-editing systems, respectively (Rees and Liu, 2018). These tools can edit approximately 60% of known pathogenic mutations (Rees and Liu, 2018). Until recently, this state of affairs has been a shortcoming of the editors in question, but the development of prime editing tools, which can correct various types of mutations (transversion, insertion, or deletion), has removed these limitations (Anzalone et al., 2019). Here we describe the results of *in vitro* and *in vivo* research on animal models of rare genetic diseases, the main prospects and shortcomings of these tools, and current progress in their clinical application.

Evolution of DNA-editing systems

The CRISPR-Cas9 system is an adaptive-immune-system component in bacteria and archaea and targets viral or plasmid dsDNA molecules (Wiedenheft et al., 2012). In laboratory practice, the most widely used Cas9 nuclease is Cas9 from the bacterium *Streptococcus pyogenes* (*Sp*Cas9) (Sander and Joung, 2014); however, some other analogs, for example, *Staphylococcus aureus* Cas9 (*Sa*Cas9) can also be used (Cebrian-Serrano and Davies, 2017; Matharu et al., 2019). The Cas9 nuclease is directed by guide RNAs (either a complex of tracrRNA with crRNA or a fusion single guide RNA) to a target dsDNA sequence containing a short stretch of nucleotides (downstream of the target sequence) termed the protospacer adjacent motif or PAM (for SpCas9, the PAM is 5′-NGG-3′, where N stands for any nucleotide). Upon recognition of a PAM and binding to the target sequence, DNA opens and the "R-loop" is formed (Jore et al., 2011). Cas9 activates and using RuvC-like and HNH domains makes two nicks in two complementary strands at the target locus, resulting in a DSB (Jinek et al., 2012). The CRISPR/Cas9 system has opened up numerous opportunities for genome editing in different organisms, and now there are many reports on its various applications [for review see (Mengstie and Wondimu, 2021)]; in particular, this system is used to create animal models of human diseases (Leonova and Gainetdinov, 2020). It has found many applications in biotechnology, including cracking the challenge of antibiotic resistance (Matharu et al., 2019; Novick, 2021, 202; Zohra et al., 2021). For instance, the use of a CRISPR-Cas9 system targeted against resistance genes has helped to reduce the resistance to β-lactames in *E. coli* (Kim et al., 2016) and *K. pneumonia* (Hao et al., 2020) and to lower the number of antibiotic-resistant *E. faecalis* strains (Rodrigues et al., 2019).

The CRISPR-Cas12 system is another editing system of bacterial origin for targeted DSB introduction. Cas12 nucleases are guide RNA–targeted DNA-specific endonucleases recognizing a PAM (for Cas12 proteins, it is usually T-rich, for example, for Cas12a, the PAM is 5′-TTTV-3', where V is for G/C/A) (Chen P. et al., 2020). Unlike Cas9, Cas12 needs only one short crRNA for targeting (Zetsche et al., 2015). By now, at least 11 types of Cas12 proteins have been discovered: Cas12a (formerly known as Cpf1), Cas12b, Cas12c, Cas12d, Cas12e, Cas12f (also known as Cas14), Cas12g, Cas12h, Cas12i, Cas12j, and Cas12k (Tong et al., 2021). It has been shown that upon recognizing a target and making the first cut, Cas12 proteins stay and exert a nonspecific endonuclease activity toward surrounding DNA molecules, which is called collateral activity (Chen et al., 2018). Cas12 proteins have been widely used for gene editing and transcriptional regulation [for review see (Tong et al., 2021)]. In addition, with Cas13, the Cas12 nuclease is employed in (mostly viral) nucleic-acid detection systems like DETECTR or SHERLOCK (Gootenberg et al., 2018); in particular, these systems have been proposed for COVID-19 detection (Safari et al., 2021).

It should be noted that Cas9 introduces DSB with the formation of blunt ends, while Cas12 introduces sticky ends. Anyway, both systems can activate similar repair systems: Ku-dependent non-homologous DNA end joining (NHEJ), a Polymerase θ-mediated end joining (TMEJ or microhomology-mediated end joining (MMEJ)), and homology-directed repair (HDR). The molecular mechanisms of preference in cellular repair post-CRISPR/Cas9 cleavage are still unclear (Wyatt et al., 2016). The choice what kind of repair pathway will be activated depends on many factors, such as the phase of the cell cycle, chromatin structure and the CRISPR/Cas construction (Leonova and Gainetdinov, 2020; Vítor et al., 2020). For example, MMEJ seems to be most active during the M and early S phases in dividing cells (Yanik et al., 2018). The most important factor in determining which double-strand break repair pathway will be used is whether or not the 5' termini of broken ends are resected. Ends with little (about 10 n.t.) or no single stranded overhang are typically rejoined by Ku-dependent NHEJ. In contrast, TMEJ assumes prominence as the extent of 5′ > 3′ resection exceeds 45 nt (Yousefzadeh et al., 2014; Yanik

et al., 2018). Repair of DSB by different mechanisms leads to many random indels thereby making the DNA-editing process inaccurate. To address the challenge of making single-nucleotide DNA editing precise and efficient, Cas12 and Cas9 proteins have been modified (nuclease domains activities have been either fully eliminated or turned to nickases) and then fused with adenosine or cytidine deamination enzymes.

First *Sp*Cas9 nickases (endonuclease variants where Cas9 cuts either the paired (targeted) or unpaired (non-targeted) DNA target strand but not both) have been obtained by introducing amino acid (a.a.) substitutions into *Sp*Cas9 nuclease domains: residue D10 in the RuvC-like domain or residue H840 in the HNH domain has been replaced with alanine (Sapranauskas et al., 2011). Incubation of these variants of the *Sp*Cas9 nuclease (hereafter referred to as nCas9) in complex with guide RNA and plasmid DNA results in nicked open circular plasmids, whereas wild-type *Sp*Cas9 produces a linear DNA product. Furthermore, it has been found that the RuvC-like domain cleaves an unpaired DNA strand, while HNH cleaves the paired strand (Jinek et al., 2012). When both mutations (D10A and H840A) are introduced into *Sp*Cas9, the nuclease activity is eliminated, but the targeting activity remains. This catalytically inactive nuclease is called dead Cas9 or dCas9.

Cytidine base editors

The cytidine deaminase reaction in DNA leads to the cytidine-to-uridine transition giving rise to a functional G-to-A substitution. AID/APOBEC cytidine deaminases are well known and are normally found in jawed vertebrates. These enzymes can bind and deaminate RNA and single-strand DNA (ssDNA). In humans, this family includes several cytidine deaminases: AID, APOBEC1, APOBEC3 (a subfamily with seven members: A, B, C, D, F, G, and H), APOBEC2, and APOBEC4 (Salter et al., 2016). Apart from APOBEC, in genetic engineering, researchers use activation-induced deaminase (AID) (from vertebrates) and lamprey CDA1 and CDA1-like proteins (Muramatsu et al., 1999; Pancer et al., 2004). Despite the low sequence identity between human AID/APOBEC and CDA1 (and CDA1-like) proteins and because all these enzymes are functional cytidine deaminases involved in adaptive immunity, it is believed that CDA1 and CDA1-like proteins may be affiliated with the AID/APOBEC family of proteins (Holland et al., 2018).

The first Cas9-targeted DNA-specific cytidine base editor (CBE) was created by Komor and others in 2016 (Komor et al., 2016). Rat APOBEC1 (rAPOBEC1) was fused to the N terminus of dCas9 through the XTEN linker resulting in rAPOBEC1-XTEN-dCas9 chimeric protein (Figure 1). This editor manifested more than 50% effectiveness of DNA deamination *in vitro*; however, *in vivo*, its effectiveness is drastically lower (0.8–7.7%) due to the cellular response to U-G heteroduplex DNA: activation of uracil DNA glycosylase (UDG), which catalyzes the removal of U from DNA in cells and initiates base excision repair (Kunz et al., 2009), thus leading to C recovery at the target site.

To address this problem, a UDG protein inhibitor (UGI) from bacteriophage PBS1 (Mol et al., 1995) has been fused to the C terminus of BE1 and utilized as a second-generation base editor (BE2) (Komor et al., 2016). Deamination efficiency of BE2 (rAPOBEC1-dCas9-UGI) has been assessed *in vivo* on six genomic loci. The results were promising: a threefold rise (in comparison to BE1) led to ≤20% C-to-U conversion effectiveness in HEK293 cells.

Mismatch repair (MMR) machinery uses nicked heteroduplex DNA as a good substrate for PCNA loading and subsequent endonuclease activation on the incised strand (Pluciennik et al., 2010); therefore, introducing a nick into the nonedited DNA strand near the nucleotide mismatch (base-editing result) may increase the repair of the wild-type strand and elevate the amounts of edited DNA. Thus, to further increase the base editor conversion degree *in vivo*, the Cas9 nuclease should create a nick in the nonedited strand of target DNA. Base-editing efficiency of BE3 in human cells has turned out to be even higher than that of BE2 and in some cases reached 75% (Komor et al., 2016); however, a slightly increased indel rate was observed after BE3 treatment in comparison with BE1 or BE2. Off-target activity of this system was reported to be low and mostly due to Cas9 off-target effects.

In 2016, the Target-AID cytidine base editor was constructed by Nishida and others (Nishida et al., 2016). The first Target-AID system was based on the targeting activity of dCas9 and the cytidine deamination activity of an AID/APOBEC family protein: an AID lamprey ortholog called CDA1. Two proteins were fused through a long (100 a.a.) peptide linker. This system has shown only a 2% mutation rate in yeast cells (Nishida et al., 2016). To raise editing rates, Target-AID has gone through evolution similar to BE systems: firstly, dCas9 was replaced with nCas9(D10A) increasing effectiveness up to 35%. Next, this protein was fused to UGI, which raised the mutagenesis rate up to 74%. The latest Target-AID system acts in a similar fashion but not identically to BE3: in Target-AID, modifications were preferably introduced 15–19 bases upstream of the PAM (overlapping with the BE3 effective editing window); in contrast to rAPOBEC1, CDA1 seems to lack sequence preferences, and therefore the Target-AID system has good potential for therapeutic use owing to a wider range of target sequences.

PAM recognition by Cas9 is a factor lowering the practical potential of Cas-based systems by narrowing the spectrum of targets. To solve this problem, Kleinstiver and coworkers (Kleinstiver et al., 2015) have mutagenized Cas9 in the PAM recognition domain. The resulting mutants were named *Sp*CasVQR (containing D1135V/R1335Q/T1337R mutations) and *Sp*Cas9EQR (containing D1135E/R1335Q/T1337R

mutations) and recognized respectively NGAN (also NGNG but with generally lower efficiency) and NGAG PAMs. Additionally, a quadruple mutant of *Sp*Cas9VRER (D1135V/G1218R/R1335E/T1337R) was obtained. It manifested the highest activity toward an NGCG PAM and minimal activity toward an NGG PAM (Kleinstiver et al., 2015). In 2018, Nishimasu with colleagues modified *Sp*Cas9 for nonclassic PAM recognition. Introduction of several mutations (R1335V/L1111R/D1135V/G1218R/E1219F/A1322R/T1337R) into *Sp*Cas9 (the obtained Cas9 variant was designated as *Sp*Cas9-NG) has led to relaxed 5′-NG-3′ PAM recognition (Nishimasu et al., 2018).

Kim with colleagues (Kim et al., 2017) have developed some BE3-modifications regarding PAM recognition. The *Sa*Cas9 nuclease is smaller than SpCas9 and recognizes another PAM: NNGRRT (Ran et al., 2015), thereby potentially expanding the number of available target sites for cytidine base editing. A nickase version of *Sa*Cas9 (*Sa*Cas9n) was fused with rAPOBEC1 and UGI, and this protein was named *Sa*BE3. The efficiency of this system on target sites in general exceeded that of BE3 (Kim et al., 2017). The *Sp*Cas9 protein of BE3 was replaced with above-mentioned mutated Cas9 proteins (VQR, EQR, or VRER Cas9 variants) to set up VQR-BE3, EQR-BE3, and VRER-BE3 systems, which should target NGAN, NGAG, or NGCG PAMs, respectively. The efficiency of editing by these proteins in HEK293 cells is up to 50% while having a low off-target activity (Kim et al., 2017). Mutating an APOBEC1 active-center residue (W80Y/F) narrowed the editing window to three nucleotides. Likewise, mutations in the substrate-binding domain of APOBEC1 (R126E or R132E) narrowed the editing window. Proteins with double mutations (W80Y R126E, W80Y R132E, or R126E R132E) have the editing window ~2 nt wide, thereby showing more predictable and precise editing, whereas triple mutants have almost threefold lower maximal editing yields, with the editing window narrowed to almost two to one nucleotide (depending on the target locus). When combined, the two innovations (window-modulating mutations in APOBEC1 and VQR-BE3) allowed editing with a narrowed activity window and greater positional selectivity of target sites containing an NGA PAM (Kim et al., 2017).

In 2017, Komor and others (Komor et al., 2017) developed BE3s involving different AID/APOBEC family members (AID, CDA1, or APOBEC3G) to address the problem of sequence context preferences of rAPOBEC1. It was reported that AID-BE3 and CDA-BE3 are efficient when the nucleotide one bp upstream of the target C is G; however, overall (non-GC) editing rates were lower in comparison with BE3. Furthermore, deamination by AID-BE3 and CDA-BE3 was more accurate (the product was purer) in comparison with BE (Komor et al., 2017). Not only deaminases but also the mutual position of BE parts could influence efficiency, accuracy, and robustness of cytidine deamination. For instance, extending the linker length to 32 a.a. Between proteins nCas9 and rAPOBEC1 gave a 1.2-fold increase in reaction efficiency. Extending the linker length between nCas9 and UGI to 9 a.a. Led to a 1.3-fold decrease in non-T product formation, with no apparent changes in C-to-T editing. Insertion of another copy of UGI into the C terminus of BE3 induced a more than twofold increase in product purity relative to BE3. Combining these three improvements has led to the development of the fourth generation of base editors: BE4 (Komor et al., 2017). Compared to BE3, BE4 offers a 2.3-fold decrease in byproduct amounts as well as 2.3-fold lower indel formation.

Next, Rees with colleagues (Rees et al., 2017) modified BE3 to reduce off-target effects and created HF-BE3, a base editor containing high-fidelity Cas9 variant HF-Cas9 (containing four point mutations [N497A, R661A, Q695A, and Q926A] for elimination of nonspecific Cas9–DNA interactions). In comparison with BE3, HF-BE3 shows 37-fold less off-target editing with only a slight reduction in on-target editing efficiency. Successful delivery of the HF-BE3 system using ribonucleoproteins into the mouse ear and zebrafish embryo and generation of C-to-T substitutions *in vivo* has been reported (Rees et al., 2017).

In 2021, Liu and others (Liu et al., 2021) designed a cytidine editing system based on the *Neisseria meningitides* Cas9 (Nme2-Cas9) specific to cytidine dinucleotide PAM (N4CC), thus enlarging the target sequence pool and offering compact size (1,082 a.a.) and natural high fidelity. A cytidine base editor with *Nme*2-Cas9 was created by replacing n*Sp*Cas9 from the BEmax editor with a nickase version of *Nme*2-Cas9 (D16A). The obtained editor was designated as n*Nme*2CBE. Compared to n*Sp*-CBE, the newly developed editor showed comparable editing efficiency and a smaller amount of off-target products (Liu et al., 2021).

The problem of the size of base editors has been addressed differently: some researches propose to use orthologs of SpCas9 (Ran et al., 2015; Liu et al., 2021), but in ref. (Levy et al., 2020), another approach is utilized. Adeno-associated virus (AAV) delivery is size-limited; accordingly, for ABE or CBE systems to be split, it was decided to use a *trans*-splicing intein enabling CBE and ABE division into halves thereby enabling dual AAV packaging of base editors. The assembly of this split-intein CBE was conducted in several steps: fusion of each split DnaE intein half from *Nostoc punctiforme* (*Npu*) to each half of the original BE3, followed by dividing it within the SpCas9 sequence immediately before Cys574. This split base editor construct is called *Npu*-BE3 and has a good on-target base editing rate of approximately 34% in HEK293T cells. A BE4max-based *Npu*-BE4max construct has also been developed. Codon usage optimization and a nuclear localization signal (NLS) resulted in higher base-editing efficiency (44%) than that of *Npu*-BE4 involving IDT (Integrated DNA Technologies) codon optimization (22%). A rational version of the base-editing system for AAV delivery consisted of a spliced NLS- and codon-optimized APOBEC fused to the Cas9 nickase and UGI and is referred to as CBE3.9max (Levy et al., 2020). This base

editor has gone through a number of modifications until optimized v5 AAV split-CBE3.9max manifested 56% base-editing efficiency in HEK293 cells. *In vivo* (in a mouse), this construct had organ-dependent moderate efficiency varying from 4% in skeletal muscles to 21% in the liver (Levy et al., 2020).

Besides Cas9, another Cas nuclease family member has been repurposed for targeted base editing, which is Cas12. Li and others have fused catalytically dead Cas12a from *L. bacterium* (dLbCas12) with rAPOBEC1 and a uracil DNA glycosylase inhibitor; thus, a dCas12 targeted base editor was obtained (Li et al., 2018). It showed high editing activity (up to 70% efficiency toward some cytosines). In mammalian cells, efficiency dropped down to 20% on average. The main editing window of this base editor ranges from position 8 to 13 (assuming that the base next to the PAM is position 1). Introducing mutations W90Y and R126E in APOBEC has narrowed the editing window to 10–12 positions of the spacer (Li et al., 2018). dCas12BE has undergone numerous modifications (Wang X. et al., 2020) starting from fusion to various AID/APOBEC family proteins: rAPOBEC1, hAPOBEC3A, hAPOBEC3B, or hAID (referred to rA1, hA3A, hA3B, and hAID, respectively). hA3A-dCas12a-BE has the highest editing efficiency among the aforementioned nucleases. Later, mutations W98Y, W104A, and P134Y have been introduced into hA3A-dCas12a-BE, and relative efficiency has been assessed: $hA3A^{W104A}$-dCas12a-BE, $hA3A^{W98Y/W104A}$-dCas12a-BE, and $hA3A^{W104A/P134Y}$-dCas12a-BE perform active editing. Next, similarly to ref. (Koblan et al., 2018), codons in dCas12-BEs have been optimized for mammalian expression. Editing windows of the obtained mutated hA3A-dCas12a-BE-ops were shown to be ~15 bp long. Introducing the Y132D or Y130F mutation into the hA3A region of dCas12BE leads to editing-window narrowing, increased accuracy, and a lower frequency of indel formation. Consequently, $hA3A_{W104A/Y132D}$-dCas12a-BE-op and $hA3A^{W98Y/W104A/Y130F}$-dCas12a-BE-op were called BEACON1 and BEACON2, respectively, and have shown editing productivity similar to that of AncBE4max (while creating much fewer indels) in the cell. Furthermore, the BEACONs have been tested *in vivo*: C-to-T editing efficiency in mouse organs ranges from 51% to 71% (Wang X. et al., 2020).

In general, CBE architectures have gradually evolved to improve editing efficiency and product purity, to lower the indel rate, and to broaden PAM recognition specificity in a native environment of a target sequence.

Adenine base editors

Another class of base editors is adenine base editors or ABEs. There are no natural adenine deaminases acting on DNA, and to make DNA adenine deamination possible, RNA-specific deaminases should be modified. Gaudelli with colleagues have devised an adenine base editor converting adenine to inosine in DNA, resulting in a T-to-G substitution (Gaudelli et al., 2017). They used directed evolution to create a DNA-specific form of RNA-specific adenine deaminase TadA. TadA is a tRNA adenine deaminase converting adenine to inosine (I) in the ssRNA of the anticodon loop of $tRNA_{Arg}$ (Kim et al., 2006). Some APOBECs share homology with TadA, and it is reported that APOBECs possess RNA- and DNA-binding properties. Therefore, it has been hypothesized that some mutant TadA (TadA*) enzymes are able to bind and edit DNA. As a consequence, mutations A106V and D108N have been incorporated into the TadA deaminase, and the obtained protein has been subsequently fused through the XTEN linker to nCas9(D10A) and a C-terminal NLS (Figure 1). The resulting protein serves as the ABE1.2 DNA base editor. Editing efficiency in cells is only 3.2%, and editing is performed mostly at the fifth protospacer position (generally ~fourth to ninth position, assuming that the PAM is positions 21–23) (Gaudelli et al., 2017).

This inefficient but working DNA-specific adenine editor has given rise to the evolution of ABE systems. Incorporation of mutations D147Y and E155V into TadA* (giving the ABE2.1 system via replacement of the precise version of deaminase) has led to a twofold to sevenfold increase in editing efficiency as compared with ABE1.2 at six genomic loci tested. An ABE2.6 variant with a prolonged XTEN linker (to 32 a.a.) has slightly higher (relative to ABE2.1) editing efficiency: 14%. Because normally, TadA operates as a homodimer (Losey et al., 2006), TadA* (version 2.1) has been fused to the N terminus of ABE2.1, and the efficiency of the obtained ABE2.9 system is 7.5% higher, resulting in an editing efficiency of 20%. Three new TadA mutations (L84F, H123Y, and I157F) have been applied to ABE2.9 to generate the ABE3.1 system showing 1.6-fold better performance than ABE2.9 does; however, a distinct sequence preference was observed. To solve this problem, various mutations were introduced into the TadA protein. Four mutations (H36L, R51L, S146C, and K157N) in ABE3.1 led to ABE5.1, which shows decreased editing efficiency in HEK293T cells. This intermediate system was modified by fusing wild-type TadA to the N terminus of ABE 5.1, thus giving rise to ABE 5.3 (with average editing efficiency of 39%) and broadening sequence compatibility. Introducing P48S into TadA* (5.3) resulted in the ABE6.3 system with elevated average DNA-editing efficiency, by 1.3-fold. Mutations W23R, P48A, and R152P in ABE6.3 resulted in the ABE7.10 system showing improved editing efficiency, up to 58%, at six loci in HEK293 cell lines; this performance is 29-fold better than that of the ABE1.2 system. Subsequent ABE7.10 analysis has revealed that the indel percentage is extremely low (<0.1%) and off-target activity is almost absent, suggesting that systems eliminating inosine from DNA are less active than these toward uracil. These seven sequential evolution rounds of ABEs are giving the scientists a lot of information about the editing principles of ABE systems (Gaudelli et al., 2017).

Another approach to addressing the editing efficiency problem is used in ref. (Koblan et al., 2018). These researchers found that the stability and magnitude of base editor expression are factors influencing base modification. Codon usage optimization and NLS presence were hypothesized as factors impairing base modification. It was demonstrated that bipartite NLS (bpNLS) presence at both the N and C termini of BE4 induces a 1.3-fold improvement in its editing efficiency. Bis-bpNLS BE4 with GenScript codon usage was named BE4max and had 1.8-fold higher editing efficiency as compared to bis-bpNLS BE4 with IDT codons, and manifested approximately 89% editing efficiency in HEK293 cells. An analogous approach has been chosen for adenine base editor ABE7.10: SV40 NLS substitution by bis-bpNLS improved editing efficiency 1.5–2.0-fold, and GenScript codon optimization yielded 1.3- to 7.9-fold higher editing efficiency of this base editor called ABEmax in comparison to IDT. ABEmax has remained an extremely accurate and robust editor, but its indel rate is 1.7% compared to <0.1% of ABE7.10 (Koblan et al., 2018).

Reports of the development of SpCas9-NG along with ABEmax have inspired Huang and others (Huang S. et al., 2019) to develop a fusion ABEmax-SpCas9-NG system called NG-ABEmax. High editing efficiency and NG-PAM recognition offer great potential for splice site modification in order to modulate RNA splicing in the cell. It was demonstrated that ABEmax-NG effectively recognizes all types of NG(N) PAMs and efficiently performs DNA editing *in vitro* and *in vivo*, whereas ABEmax recognizes the classic NGG PAM with high efficiency and the NGA PAM with modest efficiency (Huang S. et al., 2019). Meanwhile, T. Huang with colleagues have created analogous systems based on other Cas9 nucleases with altered PAM recognition: VRQR-SpCas9 (PAM: NGA) and VRER-SpCas9 (PAM: NGCG) (Kleinstiver et al., 2015, 2016), yielding VRQR-ABEmax and VRER-ABEmax, respectively (Huang T. P. et al., 2019). These editors were tested in HEK293 cells at six genomic loci. VRQR-ABEmax manifested 35% editing efficiency, being 3.2-fold better than ABEmax. VRER-ABEmax showed conversion efficiency averaging 40%: a 7.0-fold improvement over ABEmax (Huang T. P. et al., 2019). ABE7.10 evolution continues, and in 2020, two papers got published describing another generation of ABEs. Using phage-assisted noncontinuous and continuous evolution (Richter et al., 2020), investigators have obtained a next-generation ABE: ABE8e, which contains eight additional mutations leading to a dramatic activity boost as compared with ABE7.10 without increasing off-target activity.

Gaudelli and others (Gaudelli et al., 2020) have evolved ABE7.10 into 40 new ABE8 variants. Compared to ABE7.10, ABE8 performs ~1.5-fold more efficient editing at canonical positions (A5–A7) in the protospacer and ~3.2-fold more efficient editing at noncanonical positions (A3–A4 and A8–A10). Additionally, ABE8 recognizes classic PAM (NGG), and its editing efficiency is 4.2-fold higher at non-NGG PAM variants as compared to ABE7.10. ABE8s have base-editing capacity even at sites previously difficult to target. ABE8s can achieve 98–99% target modification in primary T cells, meaning that these editors are a promising tool for cell therapy applications (Gaudelli et al., 2020).

Some ABEs perform off-target RNA editing. It has been hypothesized that the reason lies in the wtTadA domain of the editor. In ref. (Grünewald et al., 2019), researchers deleted the wtTadA domain of ABEmax thereby obtaining the miniABEmax construct. The undesirable off-target RNA editing declined but not dramatically: 1.5-fold. Introducing mutation K20A/R21A or V82G into TadA* led to lowering of nontarget adenine modification rates with on-target efficiency rates being slightly higher for miniABEmax (V82G). There was also an interesting observation that miniABEmax (V82G) possesses an imprecise C-to-G base-editing activity within the editing windows of some DNA on-target sites (Grünewald et al., 2019).

Similarly to ref. (Grünewald et al., 2019), Xu and others (Xu et al., 2021) have developed a TadA-less adenine base editor using the SpCas9-NG nuclease. The heterodimeric adenine deaminase domain (ecTadA-ecTadA*) in ABE-NG either with the originally evolved ecTadA monomer or its high-fidelity version (ecTadA-V82G) for minimization the of the off-target RNA editing activity gave the miniABE-NG (iABE-NG) editor system. The on-target DNA editing activity of miniABE-NG is higher than that of ABE7.10-NG; however, miniABE (V82G)-NG has remarkably lower on-target editing activity when compared to ABE7.10-NG. An attempt to improve the on-target DNA editing efficiency of high-fidelity miniABE (V82G)-NG without increasing its low off-target RNA editing activity was made. The A56G mutation resulted in the miniABE (GG)-NG editor featuring completely restored on-target DNA editing activity with remaining low off-target activity (Xu et al., 2021).

Dual base editors

Having both adenine and cytidine deamination activities in one system seems to be a nice and desirable prospect. One of the ways to do so is to combine an existing CBE and ABE. Such a dual deaminase has been devised by Grünewald and others (Grünewald et al., 2020). This system called SPACE (synchronous programmable adenine and cytosine editor) consists of the miniABEmax (V82G) editor fused with Target-AID CBE (Figure 1). SPACE can carry out A-to-G editing at 25 out of 26 genomic sites edited by miniABEmax-V82G alone, but cytidine editing is performed at all target loci as compared to Target-AID. The efficiency of adenine editing by SPACE is somewhat lower relative to miniABEmax-V82G (13% versus 18.1%, respectively), whereas C-to-T editing efficiency rates of SPACE and Target-AID are quite similar (22% versus 24%, respectively). The frequency of unwanted

indels induced by SPACE is quite low (on average 1.44%) (Grünewald et al., 2020).

A similar approach was used in ref. (Zhang et al., 2020), i.e., a dual base editor system. The researchers fused ABE7.10 with BE3. Two deaminases were fused with nCas9 and UGI. For enhancing efficiency, a number of modifications were made: codon optimization was applied to hAID and TadA domains; two bipartite NLSes were added to the editor; a rigid 15-mer (EAAAKEAAAKEAAAK) linker was chosen for fusion; and finally, two copies of uracil DNA glycosylase inhibitor (UGI) were added thus resulting in the A&C-BEmax system. The A-to-G editing window of A&C-BEmax is not changed as compared to ABEmax, while the C-to-T editing window widened to 16 nucleotides in comparison with that of AID-B4 (10 nucleotides). A&C-BEmax is an efficient DNA editor showing simultaneously different A/C mutation rates on the same allele, varying from 2% to 30%, with the percentage of alleles bearing only C-to-T or A-to-G mutations varying from 5.3% to 82.6% and from 0.2% to 10%, respectively. Conversion rates are higher when adenines are at position 6 or 7. In HeLa cells at all examined targets, various base-editing efficiency (reaching 20–60% depending on target) is observed (Zhang et al., 2020).

Prime base editors (PEs)

These are conceptually new base editors allowing to directly introduce any possible substitution (both transitions and transversions) into a desired site. The first PE (PE1) was created through a fusion of nSpCas9(H840A) with Moloney murine leukemia virus reverse transcriptase (RT) (Anzalone et al., 2019). nSpCas9(H840A) fused with RT is guided to target DNA using special prime editing guide RNA (pegRNA) (Figure 1). The latter has several functions: it guides the base editor to a target DNA, interacts with DNA, bears desirable base modifications, and primes the reverse-transcription reaction (by means of a primer-binding sequence or PBS). The principle underlying the base editing by PEs is the following: nSpCas9(H840A) guided to a target locus makes a nick, a 3′ ssDNA flap is bound by pegRNA PBS, and this strand serves as a primer for reverse transcription, which extends the 3′ ssDNA flap and incorporates pegRNA-coded base substitutions into the DNA strand. The 5′ flap is excised, and a 3′ flap-favored base is incorporated. Yeast application of this system shows only modest editing efficiency.

In the PE2 system, mutations D200N, L603W, T330P, T306K, and W313F are introduced into RT, resulting in thermostability and processivity improvement leading to a 1.6- to 5.1-fold higher mutagenesis rate relative to PE1 (Anzalone et al., 2019). To improve favorability of the repair of the nonedited strand, a strategy entailing nick introduction into the nontargeted strand has been proposed. The additional guide RNA has been suggested for directing the Cas9 H840A nickase (a part of the PE system) to incise the genomic DNA at a nearby site still not causing a DSB. Application of this method has manifested elevated editing efficiency, up to 55%, with nicks positioned at approximately 40–90 bp on the 3′ side of the pegRNA-induced nick (Anzalone et al., 2019).

Here we focus only on some modifications of the editing systems that have been utilized *in vitro* and *in vivo* to correct point mutations. A number of other BE, ABE, and PE modifications are listed in another review (Yang et al., 2021).

All in all, here we outlined the main events in the evolution of base editors from BE cytidine base editors (able to make one type of substitution in a strictly controlled sequence with off-target editing) to PEs able to edit multiple nucleotides at various loci with high efficiency and a low off-target rate.

Site-directed DNA base editing for therapy of monogenic diseases

Correction of genetic point mutations via DNA-editing approaches has become widespread in animal models (*in vivo* research). Although these genome-editing tools have not yet been tested clinically, the data from animal research show their possible usefulness for the treatment of various rare monogenic diseases. The examples of *in vivo* and *in vitro*/*ex vivo* DNA editing for therapy of monogenic diseases are listed in Table 1, Table 2 respectively.

In vivo

Eye diseases

DNA-editing techniques have been successfully applied *in vivo* to correct the *Rpe65* gene mutation that is the cause of Leber congenital amaurosis (Jo et al., 2021; Suh et al., 2021). *Rpe65* codes an enzyme that is essential for the conversion of vitamin A from all-*trans*-retinol to 11-*cis*-retinal: the chromophore of the visual pigments. Consequently, a loss of the functional RPE65 enzyme leads to severe visual impairment from birth or in the first several months of life but does not affect other tissues and organs (Chao et al., 1993). In two refs. (Jo et al., 2021; Suh et al., 2021), in murine strain *rd12* (carrying a nonsense mutation in exon 3: c.130C > T; p.R44X), which manifests the first signs of retinal degeneration at ~3 weeks of age (Pang et al., 2005), investigators were able to rescue retinal and visual function. Suh and others (Suh et al., 2021) have tested the ABEmax system, which they have delivered into retinal cells by means of a lentivirus and achieved the following: RPE65 expression is restored in 32% of retinal cells, the total amount of 11-*cis*-retinal is 30% of the level in wild-type mice, and there is a 34% reduction in the concentration of all-*trans*-retinyl esters. In another study (Jo et al., 2021), investigators used the

NG-ABEmax system as well as a dual AAV with trans-splicing intein as a vector for delivery to retinal cells; however, the efficiency of DNA editing was lower (13%) than that in ref. (Suh et al., 2021).

Hearing loss–related diseases

Neonatal injection of dual AID-CBEmax AAVs into the inner ears of deaf Baringo mice carrying point mutation A545G in the *Tmc1* gene, coding for transmembrane channel-like one protein, has helped to restore inner hair cell sensory transduction and hair cell morphology and transiently rescued low-frequency hearing 4 weeks after the injection (Yeh et al., 2020).

Neuromuscular disorders

Adenine base editing has been successfully used to correct the *DMD* gene (dystrophin) mutations that are associated with Duchenne muscular dystrophy (Ryu et al., 2018; Xu et al., 2021). Loss of dystrophin leads to progressive muscle weakness and wasting, which eventually leads to respiratory disturbances, cardiomyopathy, and death before the age of 30 (Mercuri and Muntoni, 2013). The first DNA editing by means of the ABE7.10 system was performed in 2018 (Ryu et al., 2018): researchers performed intramuscular administration of two *trans*-splicing AAV vectors into the *tibialis anterior* muscle in *Dmd* knockout mice (carrying a nonsense mutation in exon 20) and evaluated therapy efficacy at 8 weeks postinjection. Postmortem histological analysis of the *tibialis anterior* showed that dystrophin expression was restored in 17% of myofibers, and deep sequencing analysis revealed that the efficiency of editing was ~3.3%. These results are encouraging because ~4% of normal dystrophin expression is sufficient to improve muscle function (Putten et al., 2013). In 2021 single systemic administration of the iABE-NGA system delivered by means of two AAVs (via tail vein injection) was performed to restore dystrophin in *mdx*^{4cv} mice, which carry a premature stop codon in exon 53 (Xu et al., 2021). A distinctive feature of the study is that it covered two time points (~5 weeks and ~9 months postadministration), which helped to evaluate the long-term impact of systemic ABE editing therapy. The results indicate that the effects of single administration of the systemic ABE editing therapy even strengthen with time in some tissues: there was 45.9% restoration of dystrophin levels in the heart at 5 weeks compared to wild-type mice and 95.9% at 9 months, ~10% restoration in the gastrocnemius at 5 weeks and ~5% at 9 months, as well as ~4% restoration in the diaphragm at 5 weeks and ~8% at 9 months. Similar results of systemic DNA editing in various organs have been obtained after a single retro-orbital injection of the CBE3.9max system, which is designed to edit a silent mutation in the murine *Dnmt1* locus (editing efficiency up to 59% in the brain, up to 38% in the liver and retina, up to 9% in skeletal muscle, and up to 20% in the heart) (Levy et al., 2020). Finally, PE editors have also been used successfully for the treatment of muscular dystrophy (Chemello et al., 2021).

Furthermore, base editors have been employed to treat spinal muscular atrophy (Lin et al., 2020). The latter is a progressive motor neuron disease (caused by a mutation in the *SMN1* gene) with onset during infancy and causes motor impairments and death in the first years of life. Neonatal injection of the miniABEmax system into lateral ventricles of SMNΔ7 SMA mice yielded an editing efficiency of 3–5% on postnatal day 7.

Blood disorders

Base-editing tools have been successfully applied to treat β-thalassemia *in vitro*, *ex vivo*, and *in vivo* (Gaudelli et al., 2017, 2020; Liang et al., 2017; Gehrke et al., 2018; Koblan et al., 2018; Wang L. et al., 2020; Miller et al., 2020; Richter et al., 2020; Zeng et al., 2020; Zhang et al., 2020; Antoniou et al., 2021; Li et al., 2021; Newby et al., 2021). β-Thalassemia develops due to deficient production of β-globin and is characterized by microcytic hypochromic anemia and abnormal results on a peripheral-blood smear. Reactivation of γ-globin expression is associated with lowered morbidity and mortality and significantly relieved disease symptoms. Therefore, therapeutic strategies against β-thalassemia are based both on the correction of a mutation in the *HBB* gene and on the introduction of mutations that disrupt binding sites of repressor proteins or create gain-of-function binding sites for activators, thereby derepressing γ-globin expression. For example, in an *in vitro* experiment on HUDEP-2 (ΔGγ) cells (Zhang et al., 2020) using the A&C-BEmax system, investigators disrupted the BCL11A binding site (strong transcription repression element) in the promoter of γ-globin genes (*HBG1* and *HBG2*) and generated a GATA1-binding site (active as enhancer) *de novo* in the promoter. Zhang and others have been able to achieve over 40% editing efficiency and nearly sixfold enhancement of γ-globin mRNA expression as compared to its expression in HUDEP-2 cells. *Ex vivo* ribonucleoprotein electroporation of the A3A$^{(N57Q)}$-BE3 system into human-peripheral-blood-mobilized CD34^{+} hematopoietic stem and progenitor cells also successfully disrupts a GATA1-binding motif and reduces BCL11A expression. In addition, those authors edited the HBB −28A>G promoter mutation. Due to this multiplex approach, substantial efficiency of DNA editing and upregulation of β- and γ-globins were achieved (Zeng et al., 2020). An *in vivo* experiment (Li et al., 2021) has been performed on β-YAC$^{+/-}$/CD46$^{+/+}$ mice, which were obtained by crossing mice carrying a yeast artificial chromosome (β-YAC) bearing the wild-type 248-kbp human β-globin locus with homozygous

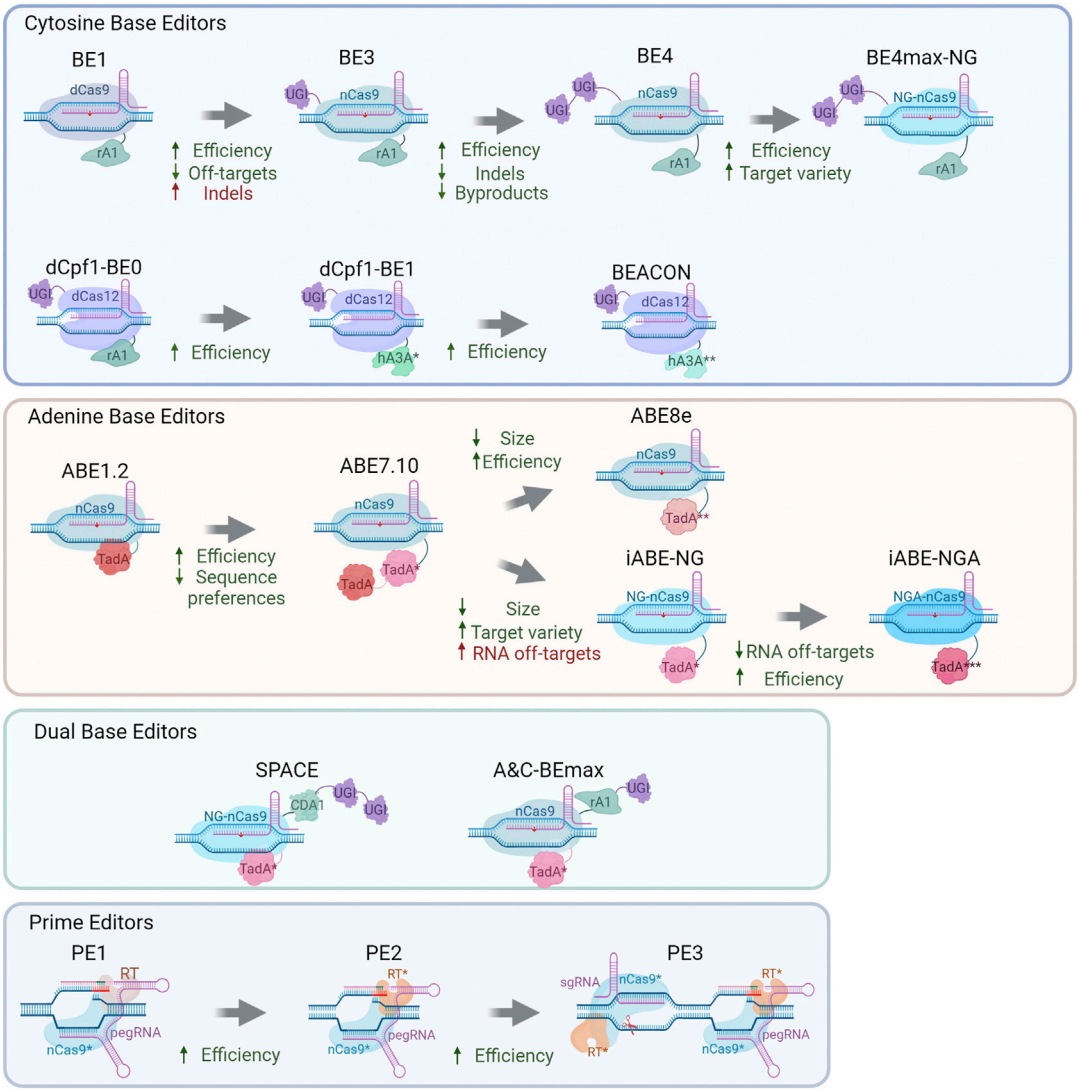


FIGURE 1
| A brief overview of base-editing systems.

transgenic mice expressing human CD46. This approach is based on transduction of peripheral $CD34^+$ hematopoietic stem/progenitor cells (for intravenous injection) with an adenovirus containing ABE vectors. $CD34^+$ hematopoietic stem/progenitor cells are mobilized beforehand in β-$YAC^{+/-}$/$CD46^{+/+}$ mice by subcutaneous injections. The transduced cells return to bone marrow, where they persist long-term. Base editing in hematopoietic stem/progenitor cells by means of an ABE vector led to efficient γ-globin induction, which persisted for 16 weeks after the introduction of the genetic construct. Another successful example of directed base-editing is the *in vivo* correction of mutated GTG (Val) codon encoding amino acid 6 in β-globin gene (HBB^S), which leads to sickle cell disease (SCD) (Newby et al., 2021). Authors combine the engineered Cas9-NRCH nickase, that recognizes a CACC protospacer-adjacent motif, with deoxyadenosine deaminase TadA-8e to generate ABE8e-NRCH. This base editor was transfected into human $CD34^+$ HSPCs from SCD donors or mice $HBB^{S/S}$ HSPCs via electroporation of ABE8e-NRCH and sgRNA in RNA or RNP forms. The editing resulted in formation of Makassar β-globin (HBB^G), a non-pathogenic variant with alanine in sixth position. Edited human cells were transplanted in immunodeficient NBSGW mice that support multilineage engraftment of human hematopoietic cells. This resulted in a decrease of β^S from 96 ± 0.28% of total β-like globin protein in unedited erythroblasts to 40 ± 2.3% in edited erythroblasts. The amount of β^G in edited cells reached 58 ± 2.8%. The electroporation of ABE8e-NRCH RNP into mice $HBB^{S/S}$

TABLE 1 *In vivo* DNA editing for therapy of monogenic diseases

Strain	Model	Delivery system	Editing systems	Target gene	Tissue	References
rd12 mice	Leber congenital amaurosis	AAV	NG-ABEmax	*Rpe65*	retina	Jo et al. (2021)
rd12 mice	Leber congenital amaurosis	Lentivirus	ABEmax	*Rpe65*	retina	Suh et al. (2021)
Baringo mice	deafness	AAV	AID-CBEmax	*Tmc1*	inner ears	Yeh et al. (2020)
Dmd knockout mice	Duchenne muscular dystrophy	AAV	ABE7.10	*Dmd*	skeletal muscles	Ryu et al. (2018)
mdx4cv mice	Duchenne muscular dystrophy	AAV	AAV-iNG	*Dmd*	heart, gastrocnemius, diaphragm and muscles	Xu et al. (2021)
ΔE51 mice	Duchenne muscular dystrophy	AAV	ABEmax PE	*Dmd*	muscles	Chemello et al. (2021)
SC-SMA$^{\Delta 7}$ mice	spinal muscular atrophy	Plasmid	miniABEmax	*Smn2*	lateral ventricles	Lin et al. (2020)
β-YAC/CD46 mice	β-hemoglobinopathies	Adenovirus	ABE	*HBG1 and HBG2 promoter*	bone marrow cells	Li et al. (2021)
Npc1tm (^{I1061T}) mice	Niemann–Pick disease type C	AAV	CBE3.9max ABEmax	*Npc1*	cortex cerebellum	Levy et al. (2020)
G93A-SOD1 mice	Amyotrophic lateral sclerosis	AAV	BE4	*Sod1*	spinal cord	Lim et al. (2020)
Fah−/− mice	Hereditary tyrosinaemia type I	LPN	ABE6.3 RA6.3	*Fah*	liver	Song et al. (2020)
Fah−/− mice	Hereditary tyrosinemia type 1	AAV	BE3	*Hpd*	liver	Rossidis et al. (2018)
NSG-PiZ mice	alpha-1 antitrypsin deficiency	LPN	BE4	*Serpina1*	liver	Packer et al. (2022)
(Pah)enu2 mice	phenylketonuria	AAV	BE3	*Pah*	liver	Villiger et al. (2018)
B6.BTBR-Pahenu2	phenylketonuria	AAV	BE-PLUS	*Pah*	liver	Zhou et al. (2022)
HGPS mice	Hutchinson–Gilford progeria syndrome	AAV	ABE7.10max-VRQR	*LMNA*	aorta bone muscle liver	Koblan et al. (2021)
Idua-^{W392X} mice	Hurler syndrome	AAV	ABEmax	*Idua*	liver	Bose et al. (2021)

HSPCs followed by transplantation into irradiated adult recipient mice led to expression of βG that made up 75–82% of total β-like globin protein. Moreover, transplantation of base-edited *HBB*$^{S/S}$ HSPCs restored all tested blood parameters to levels similar to those of healthy control mice (Newby et al., 2021). Of note, such *in vivo* rodent experiments on human cell lines allow investigators to adapt editing tools for future clinical trials.

Neurodegenerative disorders

In this field, the first step toward the treatment of neurodegenerative diseases was recently made in *Npc1*^{I1061T} mice, which are a model of Niemann–Pick disease, also known as neurodegenerative ataxia (Levy et al., 2020). A nonsense mutation in the intracellular cholesterol transporter (*Npc1*) gene leads to impaired cholesterol trafficking and accumulation of cholesterol inside cells. Niemann–Pick disease features ataxia, motor impairment, progressive intellectual disability, and dementia (Praggastis et al., 2015). A single retro-orbital injection of the CBE3.9max system into *Npc1*^{I1061T} mice prolonged the survival of Purkinje neurons and caused a 10% increase in the lifespan of the mice as compared with untreated mice.

Base-editing tools have also been tested in the treatment of amyotrophic lateral sclerosis in *SOD1*$^{(G93A)}$ mice, which are characterized by an especially aggressive course of the neurodegenerative disease and have an average lifespan of ~120 days (Lim et al., 2020). Amyotrophic lateral sclerosis is an autosomal dominant disease, and some cases are associated with a defective protein, superoxide dismutase 1 (SOD1), whose accumulation leads to a loss of motor neurons in the spinal cord and brain (Rosen, 1993). Intrathecal injection of dual AAV particles encoding a split-intein CBE system improved motor functions, reduced mutant-SOD1 reactive inclusions in the spinal cord, and increased the animals' lifespan by ~11%.

TABLE 2 *In vitro/ex vivo* DNA editing for therapy of monogenic diseases.

Cell line/Primary cells	Model	Delivery system	Editing systems	Target gene	References
LCL^{HFEC282Y}	haemochromatosis	Plasmid	ABE7.10	*HFE*	Gaudelli et al. (2017)
HEK293T$^{HBG1/HBG2\ (-113mut,-175mut\ and\ -198mut)}$	β-hemoglobinopathies	Plasmid	ABEmax	*HBG1/2*	Koblan et al. (2018)
HEK293T$^{HBG1/HBG2\ (-175mut\ and\ -198mut)}$	β-hemoglobinopathies	Plasmid	ABE8e	*HBG1/2*	Richter et al. (2020)
CD34$^+$ cells from donors with Sickle-cell disease	β-hemoglobinopathies	Plasmid	ABE8 variants	*HBG1/2*	Gaudelli et al. (2020)
Fibroblast cells from β-thalassemia patients	β-hemoglobinopathies	Plasmid	BE3	*HBB*	Liang et al. (2017)
β-thalassemia patient-derived erythroid precursor cells	β-hemoglobinopathies	Plasmid	eA3A-BE3	*HBB*	Gehrke et al. (2018)
CD34$^+$ cells from a β-thalassemia patient	β-hemoglobinopathies	RNP	hA3A-BE3	*HBB*	Liren Wang et al. (2020)
HUDEP-2 cells	β-hemoglobinopathies	Plasmid	A&C-BEmax	*HBG1/2* promoter	Zhang et al. (2020)
CD34$^+$ hematopoietic stem and progenitor cells derived from β-thalassemia patient	β-hemoglobinopathies	RNP	A3A (N57Q)-BE3	*HBG1/2* and *HBB* promoter	Zeng et al. (2020)
CD34$^+$ cells from donors with Sickle-cell disease	β-hemoglobinopathies	RNP	ABE8e-NRCH	*HBB*	Newby et al. (2021)
HEK293T $^{HBB\ (G6V)}$	β-hemoglobinopathies	Plasmid	ABE8e-NRCH	*HBB*	Miller et al. (2020)
HEK293T$^{HBG1/HBG2\ (-198T/C)}$	β-hemoglobinopathies	Plasmid	ABE7.10	*HBG1/2*	Gaudelli et al. (2017)
CuFi-3 (CFTR R553X) primary cells derived from Cystic fibrosis-affected individuals	Cystic fibrosis	RNP	ABE7.10-NG	*CFTR*	Krishnamurthy et al. (2021)
FBN1^{T7498C} cells	Marfan Syndrome	Plasmid	BE3	*FBN1*	Zeng et al. (2018)
Human embryos FBN1^{T7498C} (2d)		Microinjected mRNA of BE3 and sgRNA into zygotes			
chemically derived hepatic progenitors (CdHs)	Hereditary tyrosinemia type 1	Plasmid	ABEmax PE3	*Fuh*	Kim et al. (2021)
Mouse astrocytes (APOE4)	Alzheimer's disease	Plasmid	BE3	*APOE*	Komor et al. (2016)
HEK293T and SH-SY5Y cells	Alzheimer's disease	Plasmid	Target-AID	*APP*	Guyon et al. (2021)
HEK293T$^{HBB(E6V)}$	Sickle cell disease	Plasmid	PE3	*HBB*	Anzalone et al. (2019)
HEK293T$^{HEXA\ (1278+TATC)}$	Tay-Sachs syndrome	Plasmid	PE3	*HEXA*	Anzalone et al. (2019)
HEK293T$^{PRNP(G127V)}$	Prion disease	Plasmid	PE3	*PRNP*	Anzalone et al. (2019)
Patient-derived fibroblasts harboring the MPDU1^{L119P}	congenital disorder of glycosylation type 1f	Plasmid	BE4max	*MPDU1*	Koblan et al. (2018)
N2a neuroblastoma cells	Chronic pain	Plasmid	BE4max	*SCN9a*	Koblan et al. (2018)
Derived from children with progeria	Hutchinson–Gilford progeria syndrome	lentivirus	ABE7.10max-VRQR	*LMNA*	Koblan et al. (2021)

Nevertheless, overall editing efficiency was ~1.2%, and only ~6.5% of spinal-cord cells were successfully transduced by both AAV vectors, suggesting that there is some room for improvement of the therapy efficacy (Lim et al., 2020).

Metabolic disorders

Various DNA-editing approaches have been implemented for the treatment of type I hereditary tyrosinemia, which is attributed to loss of function of fumarylacetoacetate hydrolase (FAH) (Rossidis et al., 2018; Song et al., 2020; Kim et al., 2021). FAH deficiency impairs tyrosine catabolism, induces accumulation of toxic metabolic intermediates in the liver, and has a cytotoxic effect on hepatocytes (Grompe, 2001). *Fah$^{mut/mut}$* mice have a mutation in exon eight of the *Fah* gene, whereas treatment of adult mice with either the ABE6.3 or RA6.3 system restores the expression of functional FAH in ~1% and ~4% of hepatocytes, respectively (Song et al., 2020), and as a consequence, *Fah$^{mut/mut}$* mice do not experience the characteristic weight loss after discontinuation of administration of a tyrosine-catabolic pathway inhibitor: 2-(2-nitro-4-trifluoromethylbenzoyl)-1,3-cyclohexanedione. Those authors applied two nonviral strategies for systemic delivery

of the editing systems via tail vein injection: hydrodynamics-based transfection of plasmid DNA and lipid nanoparticle–mediated delivery of mRNAs, and the efficiency of the former was almost an order of magnitude higher. In ref. (Rossidis et al., 2018), scientists attempted to edit DNA during the embryonic period. Unlike previous studies aimed at restoring the function of a protein, here the focus was on introducing a nonsense mutation into 4-hydroxyphenylpyruvate dioxygenase (*Hpd*). Inactivation of HPD in $Fah^{mut/mut}$ mice prevents the accumulation of toxic metabolites of tyrosine in the liver. An adenoviral vector was used to deliver the BE3 editing system, which was injected into the vitelline vein on fetal day 16. The efficiency of base editing in the liver was found to gradually increase: 14% on postnatal day 1, 37% on postnatal day 30, and 40% on postnatal day 90. In this way, in these mice, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and serum bilirubin levels and the number of hepatocytes were restored. Therefore, embryonic DNA editing holds promise as a therapeutic modality for complex genetic disorders identified during prenatal screening.

In vivo models based on NSG-PiZ mice have been utilized to successfully correct alpha-1 antitrypsin deficiency, which is characterized by a lung disease and/or liver disease (Packer et al., 2022). Mutations in the SERPINA1 gene induces misfolding of the protein product and accumulation of toxic aggregates within hepatocytes, along with insufficient inhibition of neutrophil elastase in lungs (Fregonese and Stolk, 2008). Tail vein injection of the BE4 system as lipid nanoparticle–based formulation of RNA into adolescent NSG-PiZ mice has exerted pronounced effects already at 1 week after treatment (histological changes in the liver and biochemical alterations in blood serum), and these characteristics only improved at 12 and 32 weeks after this therapy (Packer et al., 2022).

Another example of the use of base editing for the correction of mutations that lead to metabolic disorders is a treatment of phenylketonuria in $(Pah)^{enu2}$ mice (Villiger et al., 2018; Zhou et al., 2022). Phenylketonuria is characterized by phenylalanine hydroxylase deficiency and impaired metabolism of L-Phe, resulting in systemic hyperphenylalaninemia. Without an appropriate dietary therapy, this condition causes damage to the central nervous system and induces severe intellectual disability (Blau et al., 2010). A three-stage study included application of the BE3 system to cultured cells, to liver organoids, and to $(Pah)^{enu2}$ mice. In the *in vivo* experiment, L-Phe blood levels and mRNA correction rates were time- and dose-dependent, peaking at 26 weeks after injection (conversion up to 63%). It must be pointed out that only a high concentration of the AAV (5×10^{11} *vg* vs. 1×10^{11} *vg*) was able to return the blood level of L-Phe to the physiological range (below 120 μmol/L). In another study (Zhou et al., 2022), researchers performed intravenous injection of AAV vectors carrying the BE3-PLUS editing system on the second postnatal day and achieved a sustained dose-dependent reduction in blood L-Phe levels up to 24 weeks of age.

Of note, base editors are also employed to inactivate genes that have unwanted functions. A vivid example is the *PCSK9* gene: a loss-of-function mutation in this gene results in reductions of the enzyme and low-density lipoprotein cholesterol levels; the latter change has a protective effect, i.e., reduces the risk of atherosclerotic cardiovascular disease (Rao et al., 2018). Base editors that are delivered *in vivo* using different delivery systems (a lipid nanoparticle–based adenoviral vector or AAV) can efficiently knock down *PCSK9* in the liver after a single infusion, with concomitant stable reductions in blood levels of PCSK9 and low-density lipoprotein cholesterol in mice and monkeys (Carreras et al., 2019; Musunuru et al., 2021; Rothgangl et al., 2021).

Other genetic disorders

Comprehensive research on DNA-editing tools has been conducted to treat Hutchinson–Gilford progeria syndrome (HGPS) in mice (Koblan et al., 2021). Hutchinson–Gilford progeria is an autosomal dominant disease featuring rapid aging, cardiovascular disease, and early death. The illness is caused by a point mutation in the lamin A (*LMNA*) gene; this mutation leads to a mis-splicing event and to the formation of a truncated nonfunctional progerin protein (De Sandre-Giovannoli et al., 2003). Investigators conducted both an *in vitro* experiment on fibroblast lines derived from patients with HGPS and an experiment on a mouse model transgenic for human *LMNA* (HGPS mice); these mice exhibit symptoms of cardiovascular complications and have a life expectancy of ~215 days. The *in vitro* experiment involving lentiviral delivery of the ABE7.10max-VRQR system yielded up to 90% genomic correction of the *LMNA* mutation and a ≤90% reduction in progerin levels both at 10 and 20 days after administration. The experiment on mice was carried out via systemic retro-orbital injections of the ABE7.10max-VRQR system. In the *in vivo* experiment, in contrast to the *in vitro* assay, the delivery was implemented using two AAV9 vectors with trans-splicing inteins, which have broad tissue tropism. The researchers tested several time points of treatment administration—a single injection on the third or 14th postnatal day—as well as long-term effects (analysis of the results at ages 6 weeks and 6 months). The results indicated that regardless of age (6 weeks or 6 months), the efficiency of DNA editing in the target organs (the heart, liver, aorta, and bone) persists and reaches 10–60% depending on the organ, whereas the amount of progerin decreases by ≤ 90%. Histological analysis showed that the treatment with the ABE system significantly alleviates morphological manifestations of the disease: a modestly reversed loss of the hypodermal fat layer

and an increase in the number of vascular smooth muscle cells in the aorta by more than threefold. Besides, the treated HGPS mice had a lifespan 1.8–2.4 times that of untreated HGPS mice. Moreover, the results of the ABE therapy on the 14th postnatal day were significantly better in various parameters as compared with the therapy on postnatal day 3. On the other hand, as the authors themselves stated, such effects may be explained by the ~10-fold higher dose of AAVs (the dose was calculated based on body weight). Nevertheless, questions about the optimal age for the therapy of various genetic diseases and about the optimal dose of the therapeutic vector remain open.

A nonsense mutation in the *IDUA* gene leads to the absence of the corresponding enzyme (α-l-iduronidase) and a buildup of large sugar polymers (glycosaminoglycans) in lysosomes, thus inducing one form of type 1 mucopolysaccharidosis (Hurler syndrome). In ref. (Bose et al., 2021), *in utero* and postnatal base editing by ABE improved cardiac function and survival of Idua-W392X mice. The adult mice demonstrated normalization of biochemical, histological, and neurobehavioral parameters, with a more pronounced recovery in the mice treated embryonically.

In vitro

DNA-editing technologies for the treatment of facioscapulohumeral muscular dystrophy, cystic fibrosis, prion diseases, sickle cell disease, Alzheimer's disease, and Tay-Sachs and Marfan syndromes have so far been tested only on *in vitro* models (cell culture) (Anzalone et al., 2019; Guyon et al., 2021; Krishnamurthy et al., 2021; Šikrová et al., 2021). The first study on a cytosine base editor (BE3) was published in 2016 (Komor et al., 2016). The editing was targeted to point mutations in the *APOE* gene, whose sequence alterations significantly increase the risk of Alzheimer's disease. Nucleofection of the BE3 system into immortalized mouse astrocytes—in which the endogenous *Apoe* gene was replaced by human *APOE4*—resulted in 58–75% efficiency of DNA editing. Editing of another point mutation that is also associated with the risk of Alzheimer's disease—a substitution in the amyloid precursor protein (*APP*) gene—by means of the BE3 system reduced the amounts of $A\beta_{40}$ and $A\beta_{42}$ peptides *in vitro* by more than 20% (Guyon et al., 2021).

Another research project (Šikrová et al., 2021) is based on immortalized myoblasts derived from individuals susceptible to facioscapulohumeral muscular dystrophy (two subtypes: FSHD1 and FSHD2). To suppress unwanted expression of DUX4, an approach was used involving the introduction of a mutation into the functional polyadenylation signal (ATTAAA) in an exon with the help of the ABEmax system. The findings revealed a significant decrease in *DUX4* mRNA levels (10–1,000-fold downregulation).

A mutation in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene causes an inherited disorder that involves severe damage to the lungs, digestive system, and other organs. One study (Krishnamurthy et al., 2021) — performed on both the CuFi-3$^{(CFTR\ R553X)}$ cell line and primary human airway epithelial cells with specific CFTR mutations—has revealed successful DNA editing (up to 80% efficiency) and restoration of CFTR anion channel function by means of ABE7.10-NG systems.

An example of successful DNA editing during the embryonic period involves the heterozygous mutation T7498C in the fibrillin gene (*FBN1*) in human embryos (Zeng et al., 2018). Pathogenic *FBN1* mutations cause Marfan syndrome, which is an autosomal dominant disease that affects the skeletal, ocular, and cardiovascular systems. Researchers microinjected mRNA of the BE3 system into zygotes and after 2 days evaluated the effectiveness of the therapy. Additionally, an experiment was conducted on modified cell line HEK293T (FBN1T^{7498C}). In both cases, high efficiency of targeted editing was achieved (40–90%).

Finally, the use of the PE3 system in cell lines with various mutations—HEK293T$^{HBB(E6V)}$ (a model of sickle cell disease), HEK293T$^{HEXA(1278+TATC)}$ (a model of Tay-Sachs syndrome), and HEK293T$^{PRNP(G127V)}$ (a model of a prion disease)—has yielded high editing efficiency (31–53%) and low numbers of indels (<5%) (Anzalone et al., 2019). Collectively, these data imply that the PE3 system can either introduce or correct transversion, insertion, or deletion mutations.

Translational potential of site-directed DNA-editing systems for gene therapy of monogenic diseases

There are few successfully implemented clinical trials of the CRISPR-Cas9 system, and site-directed editing systems based on dCas9 or nCas9 for the treatment of rare monogenic disease have so far been tested *in vivo* only on rodents. The absence of clinical trials of these systems is probably due to the fact that they were discovered relatively recently. The first successful case of CRISPR-Cas9 application *in vivo* involves the treatment of patients with hereditary transthyretin amyloidosis, which is characterized by accumulation of amyloid fibrils in tissues. Intravenous administration of CRISPR-Cas9 prevented the synthesis of the defective protein through frameshift mutations (trial registration # NCT04601051) (Gillmore et al., 2021). Furthermore, the CRISPR-Cas9 system is currently being tested in a clinical trial (NCT03872479) on 18 patients with type 10 Leber congenital amaurosis and is aimed at removing a point mutation in the *CEP290* gene. CRISPR-Cas9 has found broader applications in *ex vivo* clinical studies (Arnold, 2021; He et al., 2021). In particular, CRISPR-Cas9–edited HSPCs with inactivated BCL11A (a transcription factor responsible for the repression of fetal hemoglobin expression) (clinical trials NCT03745287 and NCT00842634) had significantly ameliorated the manifestations of sickle cell disease and

transfusion-dependent β-thalassemia, and the effects were stable for more than a year (Frangoul et al., 2021). These examples indicate that a Cas9-based tool, similar to dCas9-based and nCas9-based systems, can be utilized to restore a normal gene sequence, to create a landing site for a transcription factor, and to inactivate a gene. Of note, clinical-trial data so far are still preliminary and derive from an analysis of a very small number of patients with limited follow-up. The issue of long-term consequences of possible off-target effects and indels remains unresolved too; these are natural outcomes of DSB repair (Lin et al., 2014; Meyenberg et al., 2021). It has been shown, that base editing of human HSPCs avoided p53 activation (Newby et al., 2021). The activation of p53 leads to different cellular outcomes such as cell cycle arrest and apoptosis; the former facilitates DNA repair and promotes cell survival (Zhang et al., 2011). Overall, replacing CRISPR-Cas9 editors with nCas9-based or dCas9-based DNA-editing systems, which do not produce a DSB, looks promising. An analysis of clinical trials in diseases against which CRISPR-Cas9 has been used to date suggests that such a replacement is possible from a functional point of view, but the limitations of dCas9-based and nCas9-based systems should be addressed.

The main requirement for the use of base editors is targetability, whose disruption gives rise to off-target effects (Komor et al., 2016; Gaudelli et al., 2017). The number of off-target effects depends on PAM specificity, single guide RNA design, deaminase DNA- or RNA-binding capacity, and Cas variants (Zuo et al., 2019; Huang et al., 2021). It is worth mentioning that the deaminases that form the basis of DNA base editors have an activity toward RNA bases, and the APOBEC family of deaminases and DNA editors based on them are especially nonspecific (Levy et al., 2020). TadA deaminase, which is a part of ABE systems, is more specific, and therefore ABE systems lead to significantly fewer *de novo* single-nucleotide variants (Levy et al., 2020). On the other hand, ABE systems also tend to convert cytosine to guanine or thymine, and these substitutions occur independently of adenosine conversions (Kim et al., 2019). This ABE-mediated cytosine conversion is single-guide-RNA–dependent and may be minimized via improvement of the guide RNA by chemical modifications (Kim et al., 2019).

In addition, the number of off-target effects strongly depends on the method of delivery of the genetically engineered construct (Lin et al., 2022). The main means of delivery of CRISPR-Cas9–based editors are DNA constructs (plasmids or genetic cassettes of a viral vector) or ribonucleoprotein complexes, which differ in editing efficiency and lifetime. The advantage of base-editing systems involving CRISPR-Cas9 over gene-adding strategies and RNA editors is a permanent effect. Accordingly, treatment with base editors in the form of short-lived ribonucleoproteins can produce a stable therapeutic outcome that can last for life.

Conclusion

Findings from *in vivo* experiments suggest that even single systemic administration of a base editor can have long-term numerous effects on many tissues and organs, thereby significantly increasing life expectancy, thus making this strategy feasible even in the treatment of the most complicated genetic diseases. Rapid developments in base-editing systems are intended to reduce off-target effects and raise editing efficiency. One of the most promising approaches in this context is the improvement of bioinformatic approaches [for instance, based on a deep learning algorithm that is capable of predicting base-editing outcomes (Marquart et al., 2021)] that would help to select optimal editing tools on the basis of individual genetic characteristics of a patient. We believe that the optimization of base-editing tools and the design of new bioinformatic approaches will enable the testing of these tools in clinical trials in the next 5 years.

Author contributions

Conceptualization: EIL, and JVS; formal analysis, EIL, VVR, and JVS; writing: VVR, AVC, JVS; writing—review and editing: RAI, EIL, VVR, and JVS. visualization, AVC; supervision JVS; EIL, RAI; project administration, EIL. All authors have read and agreed to the published version of the manuscript.

Funding

The work is supported by the Sirius University of Science and Technology, project: GTH-RND-2112.for VVR, RAI and by a St. Petersburg State University grant for the development of scientific research ID 92561695 to JVS, AVC, and EIL and ID 93025998 to JVS.

Acknowledgments

The English language was corrected by shevchuk-editing.com. Pictures were created with Biorender.com.

Conflict of interest

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

Publisher's note

References

Antoniou, P., Miccio, A., and Brusson, M. (2021). Base and prime editing technologies for blood disorders. *Front. Genome Ed.* 3, 618406. doi:10.3389/fgeed.2021.618406

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157. doi:10.1038/s41586-019-1711-4

Arnold, C. (2021). Record number of gene-therapy trials, despite setbacks. *Nat. Med.* 27, 1312–1315. doi:10.1038/s41591-021-01467-7

Blau, N., van Spronsen, F. J., and Levy, H. L. (2010). Phenylketonuria. *Lancet* 376, 1417–1427. doi:10.1016/S0140-6736(10)60961-0

Bose, S. K., White, B. M., Kashyap, M. V., Dave, A., De Bie, F. R., Li, H., et al. (2021). *In utero* adenine base editing corrects multi-organ pathology in a lethal lysosomal storage disease. *Nat. Commun.* 12, 4291. doi:10.1038/s41467-021-24443-8

Carreras, A., Pane, L. S., Nitsch, R., Madeyski-Bengtson, K., Porritt, M., Akcakaya, P., et al. (2019). *In vivo* genome and base editing of a human PCSK9 knock-in hypercholesterolemic mouse model. *BMC Biol.* 17, 4. doi:10.1186/s12915-018-0624-2

Cebrian-Serrano, A., and Davies, B. (2017). CRISPR-cas orthologues and variants: Optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm. Genome* 28, 247–261. doi:10.1007/s00335-017-9697-4

Chao, D. L., Burr, A., and Pennesi, M. (1993). "RPE65-Related leber congenital amaurosis/early-onset severe retinal dystrophy," in *GeneReviews®*. Editors M. P. Adam, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, and K. W. Gripp (Seattle (WA): University of Washington, Seattle). Available at: http://www.ncbi.nlm.nih.gov/books/NBK549574/(Accessed July 5, 2022).

Chemello, F., Chai, A. C., Li, H., Rodriguez-Caycedo, C., Sanchez-Ortiz, E., Atmanli, A., et al. (2021). Precise correction of Duchenne muscular dystrophy exon deletion mutations by base and prime editing. *Sci. Adv.* 7, eabg4910. doi:10.1126/sciadv.abg4910

Chen, B., and Altman, R. B. (2017). Opportunities for developing therapies for rare genetic diseases: Focus on gain-of-function and allostery. *Orphanet J. Rare Dis.* 12, 61. doi:10.1186/s13023-017-0614-4

Chen, J. S., Ma, E., Harrington, L. B., Da Costa, M., Tian, X., Palefsky, J. M., et al. (2018). CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* 360, 436–439. doi:10.1126/science.aar6245

Chen, P., Zhou, J., Wan, Y., Liu, H., Li, Y., Liu, Z., et al. (2020). A Cas12a ortholog with stringent PAM recognition followed by low off-target editing rates for genome editing. *Genome Biol.* 21, 78. doi:10.1186/s13059-020-01989-2

Chen, W., Hu, Y., and Ju, D. (2020). Gene therapy for neurodegenerative disorders: Advances, insights and prospects. *Acta Pharm. Sin. B* 10, 1347–1359. doi:10.1016/j.apsb.2020.01.015

Colasante, G., Qiu, Y., Massimino, L., Di Berardino, C., Cornford, J. H., Snowball, A., et al. (2020). *In vivo* CRISPRa decreases seizures and rescues cognitive deficits in a rodent model of epilepsy. *Brain* 143, 891–905. doi:10.1093/brain/awaa045

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823. doi:10.1126/science.1231143

Cui, Z., Liu, H., Zhang, H., Huang, Z., Tian, R., Li, L., et al. (2021). The comparison of ZFNs, TALENs, and SpCas9 by GUIDE-seq in HPV-targeted gene therapy. *Mol. Ther. - Nucleic Acids* 26, 1466–1478. doi:10.1016/j.omtn.2021.08.008

De Sandre-Giovannoli, A., Bernard, R., Cau, P., Navarro, C., Amiel, J., Boccaccio, I., et al. (2003). Lamin A truncation in hutchinson-gilford progeria. *Science* 300, 2055. doi:10.1126/science.1084125

Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and β-thalassemia. *N. Engl. J. Med.* 384, 252–260. doi:10.1056/NEJMoa2031054

Fregonese, L., and Stolk, J. (2008). Hereditary alpha-1-antitrypsin deficiency and its clinical consequences. *Orphanet J. Rare Dis.* 3, 16. doi:10.1186/1750-1172-3-16

Gaj, T., Gersbach, C. A., and Barbas, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31, 397–405. doi:10.1016/j.tibtech.2013.04.004

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A●T to G●C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi:10.1038/nature24644

Gaudelli, N. M., Lam, D. K., Rees, H. A., Solá-Esteves, N. M., Barrera, L. A., Born, D. A., et al. (2020). Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat. Biotechnol.* 38, 892–900. doi:10.1038/s41587-020-0491-6

Gehrke, J. M., Cervantes, O., Clement, M. K., Wu, Y., Zeng, J., Bauer, D. E., et al. (2018). An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* 36, 977–982. doi:10.1038/nbt.4199

Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., et al. (2021). Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* 159, 647–661. doi:10.1016/j.cell.2014.09.029

Gillmore, J. D., Gane, E., Taubel, J., Kao, J., Fontana, M., Maitland, M. L., et al. (2021). CRISPR-Cas9 *in vivo* gene editing for transthyretin amyloidosis. *N. Engl. J. Med.* 385, 493–502. doi:10.1056/NEJMoa2107454

Gootenberg, J. S., Abudayyeh, O. O., Kellner, M. J., Joung, J., Collins, J. J., Zhang, F., et al. (2018). Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* 360, 439–444. doi:10.1126/science.aaq0179

Grompe, M. (2001). The pathophysiology and treatment of hereditary tyrosinemia type 1. *Semin. Liver Dis.* 21, 563–572. doi:10.1055/s-2001-19035

Grünewald, J., Zhou, R., Iyer, S., Lareau, C. A., Garcia, S. P., Aryee, M. J., et al. (2019). CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.* 37, 1041–1048. doi:10.1038/s41587-019-0236-6

Grünewald, J., Zhou, R., Lareau, C. A., Garcia, S. P., Iyer, S., Miller, B. R., et al. (2020). A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. *Nat. Biotechnol.* 38, 861–864. doi:10.1038/s41587-020-0535-y

Guo, N., Liu, J.-B., Li, W., Ma, Y.-S., and Fu, D. (Forthcoming 2021). The power and the promise of CRISPR/Cas9 genome editing for clinical application with gene therapy. *J. Adv. Res.* doi:10.1016/j.jare.2021.11.018

Guyon, A., Rousseau, J., Bégin, F.-G., Bertin, T., Lamothe, G., Tremblay, J. P., et al. (2021). Base editing strategy for insertion of the A673T mutation in the APP gene to prevent the development of AD *in vitro*. *Mol. Ther. - Nucleic Acids* 24, 253–263. doi:10.1016/j.omtn.2021.02.032

Hao, M., He, Y., Zhang, H., Liao, X.-P., Liu, Y.-H., Sun, J., et al. (2020). CRISPR-Cas9-Mediated carbapenemase gene and plasmid curing in carbapenem-resistant *enterobacteriaceae*. *Antimicrob. Agents Chemother.* 64, e00843–20. doi:10.1128/AAC.00843-20

He, X., Urip, B. A., Zhang, Z., Ngan, C. C., and Feng, B. (2021). Evolving AAV-delivered therapeutics towards ultimate cures. *J. Mol. Med.* 99, 593–617. doi:10.1007/s00109-020-02034-2

Holland, S. J., Berghuis, L. M., King, J. J., Iyer, L. M., Sikora, K., Fifield, H., et al. (2018). Expansions, diversification, and interindividual copy number variations of AID/APOBEC family cytidine deaminase genes in lampreys. *Proc. Natl. Acad. Sci. U. S. A.* 115, E3211–E3220. doi:10.1073/pnas.1720871115

Huang, S., Liao, Z., Li, X., Liu, Z., Li, G., Li, J., et al. (2019). Developing ABEmax-NG with precise targeting and expanded editing scope to model pathogenic splice site mutations *in vivo*. *iScience* 15, 640–648. doi:10.1016/j.isci.2019.05.008

Huang, T. P., Newby, G. A., and Liu, D. R. (2021). Precision genome editing using cytosine and adenine base editors in mammalian cells. *Nat. Protoc.* 16, 1089–1128. doi:10.1038/s41596-020-00450-9

Huang, T. P., Zhao, K. T., Miller, S. M., Gaudelli, N. M., Oakes, B. L., Fellmann, C., et al. (2019). Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat. Biotechnol.* 37, 626–631. doi:10.1038/s41587-019-0134-y

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., Charpentier, E., et al. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi:10.1126/science.1225829

Jo, D. H., Jang, H.-K., Cho, C. S., Han, J. H., Ryu, G., Jung, Y., et al. (2021). Therapeutic adenine base editing corrects nonsense mutation and improves visual function in a mouse model of Leber congenital amaurosis. *bioRxiv-Genetics* [Preprint]. Available at: https://www.biorxiv.org/content/10.1101/2021.01.07.425822v1. doi:10.1101/2021.01.07.425822

Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* 18, 529–536. doi:10.1038/nsmb.2019

Kim, H. S., Jeong, Y. K., Hur, J. K., Kim, J.-S., and Bae, S. (2019). Adenine base editors catalyze cytosine conversions in human cells. *Nat. Biotechnol.* 37, 1145–1148. doi:10.1038/s41587-019-0254-4

Kim, J.-S., Cho, D.-H., Park, M., Chung, W.-J., Shin, D., Ko, K. S., et al. (2016). CRISPR/Cas9-mediated re-sensitization of antibiotic-resistant *Escherichia coli* harboring extended-spectrum β-lactamases. *J. Microbiol. Biotechnol.* 26, 394–401. doi:10.4014/jmb.1508.08080

Kim, J., Malashkevich, V., Roday, S., Lisbin, M., Schramm, V. L., Almo, S. C., et al. (2006). Structural and kinetic characterization of *Escherichia coli* TadA, the wobble-specific tRNA deaminase. *Biochemistry* 45, 6407–6416. doi:10.1021/bi0522394

Kim, Y. B., Komor, A. C., Levy, J. M., Packer, M. S., Zhao, K. T., Liu, D. R., et al. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* 35, 371–376. doi:10.1038/nbt.3803

Kim, Y., Hong, S.-A., Yu, J., Eom, J., Jang, K., Yoon, S., et al. (2021). Adenine base editing and prime editing of chemically derived hepatic progenitors rescue genetic liver disease. *Cell. Stem Cell.* 28, 1614–1624. e5. doi:10.1016/j.stem.2021.04.010

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., et al. (2016). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495. doi:10.1038/nature16526

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485. doi:10.1038/nature14592

Koblan, L. W., Doman, J. L., Wilson, C., Levy, J. M., Tay, T., Newby, G. A., et al. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* 36, 843–846. doi:10.1038/nbt.4172

Koblan, L. W., Erdos, M. R., Wilson, C., Cabral, W. A., Levy, J. M., Xiong, Z.-M., et al. (2021). *In vivo* base editing rescues Hutchinson–Gilford progeria syndrome in mice. *Nature* 589, 608–614. doi:10.1038/s41586-020-03086-7

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi:10.1038/nature17946

Komor, A. C., Zhao, K. T., Packer, M. S., Gaudelli, N. M., Waterbury, A. L., Koblan, L. W., et al. (2017). Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci. Adv.* 3, eaao4774. doi:10.1126/sciadv.aao4774

Krishnamurthy, S., Traore, S., Cooney, A. L., Brommel, C. M., Kulhankova, K., Sinn, P. L., et al. (2021). Functional correction of CFTR mutations in human airway epithelial cells using adenine base editors. *Nucleic Acids Res.* 49, 10558–10572. doi:10.1093/nar/gkab788

Kunz, C., Saito, Y., and Schär, P. (2009). DNA repair in mammalian cells: Mismatched repair: Variations on a theme. *Cell. Mol. Life Sci.* 66, 1021–1038. doi:10.1007/s00018-009-8739-9

Leonova, E. I., and Gainetdinov, R. R. (2020). CRISPR/Cas9 technology in translational biomedicine. . *Physiol. biochem.* 54, 354–370. doi:10.33594/000000224

Levy, J. M., Yeh, W.-H., Pendse, N., Davis, J. R., Hennessey, E., Butcher, R., et al. (2020). Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. *Nat. Biomed. Eng.* 4, 97–110. doi:10.1038/s41551-019-0501-5

Li, C., Georgakopoulou, A., Mishra, A., Gil, S., Hawkins, R. D., Yannaki, E., et al. (2021). *In vivo* HSPC gene therapy with base editors allows for efficient reactivation of fetal γ-globin in β-YAC mice. *Blood Adv.* 5, 1122–1135. doi:10.1182/bloodadvances.2020003702

Li, M. (2018). Enzyme replacement therapy: A review and its role in treating lysosomal storage diseases. *Pediatr. Ann.* 47, e191–e197. doi:10.3928/19382359-20180424-01

Li, X., Wang, Y., Liu, Y., Yang, B., Wang, X., Wei, J., et al. (2018). Base editing with a Cpf1–cytidine deaminase fusion. *Nat. Biotechnol.* 36, 324–327. doi:10.1038/nbt.4102

Liang, P., Ding, C., Sun, H., Xie, X., Xu, Y., Zhang, X., et al. (2017). Correction of β-thalassemia mutant by base editor in human embryos. *Protein Cell.* 8, 811–822. doi:10.1007/s13238-017-0475-6

Lim, C. K. W., Gapinske, M., Brooks, A. K., Woods, W. S., Powell, J. E., Zeballos, C., et al. (2020). Treatment of a mouse model of ALS by *in vivo* base editing. *Mol. Ther.* 28, 1177–1189. doi:10.1016/j.ymthe.2020.01.005

Lin, X., Chen, H., Lu, Y.-Q., Hong, S., Hu, X., Gao, Y., et al. (2020). Base editing-mediated splicing correction therapy for spinal muscular atrophy. *Cell. Res.* 30, 548–550. doi:10.1038/s41422-020-0304-y

Lin, Y., Cradick, T. J., Brown, M. T., Deshmukh, H., Ranjan, P., Sarode, N., et al. (2014). CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res.* 42, 7473–7485. doi:10.1093/nar/gku402

Lin, Y., Wagner, E., and Lächelt, U. (2022). Non-viral delivery of the CRISPR/Cas system: DNA versus RNA versus RNP. *Biomater. Sci.* 10, 1166–1192. doi:10.1039/D1BM01658J

Liu, Z., Chen, S., Jia, Y., Shan, H., Chen, M., Song, Y., et al. (2021). Efficient and high-fidelity base editor with expanded PAM compatibility for cytidine dinucleotide. *Sci. China Life Sci.* 64, 1355–1367. doi:10.1007/s11427-020-1775-2

Losey, H. C., Ruthenburg, A. J., and Verdine, G. L. (2006). Crystal structure of *Staphylococcus aureus* tRNA adenosine deaminase TadA in complex with RNA. *Nat. Struct. Mol. Biol.* 13, 153–159. doi:10.1038/nsmb1047

Marquart, K. F., Allam, A., Janjuha, S., Sintsova, A., Villiger, L., Frey, N., et al. (2021). Predicting base editing outcomes with an attention-based deep learning algorithm trained on high-throughput target library screens. *Nat. Commun.* 12, 5114. doi:10.1038/s41467-021-25375-z

Matharu, N., and Ahituv, N. (2020). Modulating gene regulation to treat genetic disorders. *Nat. Rev. Drug Discov.* 19, 757–775. doi:10.1038/s41573-020-0083-7

Matharu, N., Rattanasopha, S., Tamura, S., Maliskova, L., Wang, Y., Bernard, A., et al. (2019). CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* 363, eaau0629. doi:10.1126/science.aau0629

Mengstie, M. A., and Wondimu, B. Z. (2021). Mechanism and applications of CRISPR/Cas-9-Mediated genome editing. *Biol. Targets Ther.* 15, 353–361. doi:10.2147/BTT.S326422

Mercuri, E., and Muntoni, F. (2013). Muscular dystrophies. *Lancet* 381, 845–860. doi:10.1016/S0140-6736(12)61897-2

Meyenberg, M., Ferreira da Silva, J., and Loizou, J. I. (2021). Tissue specific DNA repair outcomes shape the landscape of genome editing. *Front. Genet.* 12, 728520. doi:10.3389/fgene.2021.728520

Miller, S. M., Wang, T., Randolph, P. B., Arbab, M., Shen, M. W., Huang, T. P., et al. (2020). Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nat. Biotechnol.* 38, 471–481. doi:10.1038/s41587-020-0412-8

Mol, C. D., Arvai, A. S., Sanderson, R. J., Slupphaug, G., Kavli, B., Krokan, H. E., et al. (1995). Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: Protein mimicry of DNA. *Cell.* 82, 701–708. doi:10.1016/0092-8674(95)90467-0

Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O., et al. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* 274, 18470–18476. doi:10.1074/jbc.274.26.18470

Musunuru, K., Chadwick, A. C., Mizoguchi, T., Garcia, S. P., DeNizio, J. E., Reiss, C. W., et al. (2021). *In vivo* CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* 593, 429–434. doi:10.1038/s41586-021-03534-y

Newby, G. A., Yen, J. S., Woodard, K. J., Mayuranathan, T., Lazzarotto, C. R., Li, Y., et al. (2021). Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature* 595, 295–302. doi:10.1038/s41586-021-03609-w

Nguengang Wakap, S., Lambert, D. M., Olry, A., Rodwell, C., Gueydan, C., Lanneau, V., et al. (2020). Estimating cumulative point prevalence of rare diseases: Analysis of the orphanet database. *Eur. J. Hum. Genet.* 28, 165–173. doi:10.1038/s41431-019-0508-0

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., et al. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353, aaf8729. doi:10.1126/science.aaf8729

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262. doi:10.1126/science.aas9129

Novick, R. P. (2021). Antibacterial particles and predatory bacteria as alternatives to antibacterial chemicals in the era of antibiotic resistance. *Curr. Opin. Microbiol.* 64, 109–116. doi:10.1016/j.mib.2021.09.016

Ou, L., Przybilla, M. J., Ahlat, O., Kim, S., Overn, P., Jarnes, J., et al. (2020). A highly efficacious PS gene editing system corrects metabolic and neurological complications of mucopolysaccharidosis type I. *Mol. Ther.* 28, 1442–1454. doi:10.1016/j.ymthe.2020.03.018

Packer, M. S., Chowdhary, V., Lung, G., Cheng, L.-I., Aratyn-Schaus, Y., Leboeuf, D., et al. (2022). Evaluation of cytosine base editing and adenine base editing as a potential treatment for alpha-1 antitrypsin deficiency. *Mol. Ther.* 30, 1396–1406. doi:10.1016/j.ymthe.2022.01.040

Pancer, Z., Amemiya, C. T., Ehrhardt, G. R. A., Ceitlin, J., Larry Gartland, G., Cooper, M. D., et al. (2004). Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature* 430, 174–180. doi:10.1038/nature02740

Pang, J.-J., Chang, B., Hawes, N. L., Hurd, R. E., Davisson, M. T., Li, J., et al. (2005). Retinal degeneration 12 (rd12): A new, spontaneously arising mouse model for human leber congenital amaurosis (LCA). *Mol. Vis.* 11, 152–162.

Pluciennik, A., Dzantiev, L., Iyer, R. R., Constantin, N., Kadyrov, F. A., and Modrich, P. (2010). PCNA function in the activation and strand direction of MutLα endonuclease in mismatch repair. *Proc. Natl. Acad. Sci.* 107, 16066–16071. doi:10.1073/pnas.1010662107

Praggastis, M., Tortelli, B., Zhang, J., Fujiwara, H., Sidhu, R., Chacko, A., et al. (2015). A murine Niemann-Pick C1 I1061T knock-in model recapitulates the pathological features of the most prevalent human disease allele. *J. Neurosci.* 35, 8091–8106. doi:10.1523/JNEUROSCI.4173-14.2015

Putten, M., Hulsker, M., Young, C., Nadarajah, V. D., Heemskerk, H., Weerd, L., et al. (2013). Low dystrophin levels increase survival and improve muscle pathology and function in dystrophin/utrophin double-knockout mice. *FASEB J.* 27, 2484–2495. doi:10.1096/fj.12-224170

Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., et al. (2015). *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186–191. doi:10.1038/nature14299

Rao, A. S., Lindholm, D., Rivas, M. A., Knowles, J. W., Montgomery, S. B., Ingelsson, E., et al. (2018). Large-scale phenome-wide association study of PCSK9 variants demonstrates protection against ischemic stroke. *Circ. Genom. Precis. Med.* 11, e002162. doi:10.1161/CIRCGEN.118.002162

Rees, H. A., Komor, A. C., Yeh, W.-H., Caetano-Lopes, J., Warman, M., Edge, A. S. B., et al. (2017). Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat. Commun.* 8, 15790. doi:10.1038/ncomms15790

Rees, H. A., and Liu, D. R. (2018). Base editing: Precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 19, 770–788. doi:10.1038/s41576-018-0059-1

Reshetnikov, V. V., Chirinskaite, A. V., Sopova, J. V., Ivanov, R. A., and Leonova, E. I. (2022). Cas-based systems for RNA editing in gene therapy of monogenic diseases: *In Vitro* and *in Vivo* application and translational potential. *Front. Cell Dev. Biol.* 10, 903812. doi:10.3389/fcell.2022.903812

Richter, M. F., Zhao, K. T., Eton, E., Lapinaite, A., Newby, G. A., Thuronyi, B. W., et al. (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* 38, 883–891. doi:10.1038/s41587-020-0453-z

Rodrigues, M., McBride, S. W., Hullahalli, K., Palmer, K. L., and Duerkop, B. A. (2019). Conjugative delivery of CRISPR-cas9 for the selective depletion of antibiotic-resistant enterococci. *Antimicrob. Agents Chemother.* 63, e01454-19. doi:10.1128/AAC.01454-19

Rosen, D. (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 364, 362. doi:10.1038/364362c0

Rossidis, A. C., Stratigis, J. D., Chadwick, A. C., Hartman, H. A., Ahn, N. J., Li, H., et al. (2018). *In utero* CRISPR-mediated therapeutic editing of metabolic genes. *Nat. Med.* 24, 1513–1518. doi:10.1038/s41591-018-0184-6

Rothgangl, T., Dennis, M. K., Lin, P. J. C., Oka, R., Witzigmann, D., Villiger, L., et al. (2021). *In vivo* adenine base editing of PCSK9 in macaques reduces LDL cholesterol levels. *Nat. Biotechnol.* 39, 949–957. doi:10.1038/s41587-021-00933-4

Ryu, S.-M., Koo, T., Kim, K., Lim, K., Baek, G., Kim, S.-T., et al. (2018). Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. *Nat. Biotechnol.* 36, 536–539. doi:10.1038/nbt.4148

Safari, F., Afarid, M., Rastegari, B., Borhani-Haghighi, A., Barekati-Mowahed, M., Behzad-Behbahani, A., et al. (2021). CRISPR systems: Novel approaches for detection and combating COVID-19. *Virus Res.* 294, 198282. doi:10.1016/j.virusres.2020.198282

Salter, J. D., Bennett, R. P., and Smith, H. C. (2016). The APOBEC protein family: United by structure, divergent in function. *Trends biochem. Sci.* 41, 578–594. doi:10.1016/j.tibs.2016.05.001

Sander, J. D., and Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347–355. doi:10.1038/nbt.2842

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., Siksnys, V., et al. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* 39, 9275–9282. doi:10.1093/nar/gkr606

Šikrová, D., Cadar, V. A., Ariyurek, Y., Laros, J. F. J., Balog, J., van der Maarel, S. M., et al. (2021). Adenine base editing of the DUX4 polyadenylation signal for targeted genetic therapy in facioscapulohumeral muscular dystrophy. *Mol. Ther. - Nucleic Acids* 25, 342–354. doi:10.1016/j.omtn.2021.05.020

Song, C.-Q., Jiang, T., Richter, M., Rhym, L. H., Koblan, L. W., Zafra, M. P., et al. (2020). Adenine base editing in an adult mouse model of tyrosinaemia. *Nat. Biomed. Eng.* 4, 125–130. doi:10.1038/s41551-019-0357-8

Suh, S., Choi, E. H., Leinonen, H., Foik, A. T., Newby, G. A., Yeh, W.-H., et al. (2021). Restoration of visual function in adult mice with an inherited retinal disease via adenine base editing. *Nat. Biomed. Eng.* 5, 169–178. doi:10.1038/s41551-020-00632-6

Taylor, M., Khan, S., Stapleton, M., Wang, J., Chen, J., Wynn, R., et al. (2019). Hematopoietic stem cell transplantation for mucopolysaccharidoses: Past, present, and future. *Biol. Blood Marrow Transpl.* 25, e226–e246. doi:10.1016/j.bbmt.2019.02.012

Tong, B., Dong, H., Cui, Y., Jiang, P., Jin, Z., Zhang, D., et al. (2021). The versatile type V CRISPR effectors and their application prospects. *Front. Cell. Dev. Biol.* 8, 622103. doi:10.3389/fcell.2020.622103

Villiger, L., Grisch-Chan, H. M., Lindsay, H., Ringnalda, F., Pogliano, C. B., Allegri, G., et al. (2018). Treatment of a metabolic liver disease by *in vivo* genome base editing in adult mice. *Nat. Med.* 24, 1519–1525. doi:10.1038/s41591-018-0209-1

Vítor, A. C., Huertas, P., Legube, G., and de Almeida, S. F. (2020). Studying DNA double-strand break repair: An ever-growing toolbox. *Front. Mol. Biosci.* 7, 24. doi:10.3389/fmolb.2020.00024

Wang, L., Li, L., Ma, Y., Hu, H., Li, Q., Yang, Y., et al. (2020). Reactivation of γ-globin expression through Cas9 or base editor to treat β-hemoglobinopathies. *Cell. Res.* 30, 276–278. doi:10.1038/s41422-019-0267-z

Wang, X., Ding, C., Yu, W., Wang, Y., He, S., Yang, B., et al. (2020). Cas12a base editors induce efficient and specific editing with low DNA damage response. *Cell. Rep.* 31, 107723. doi:10.1016/j.celrep.2020.107723

Wiedenheft, B., Sternberg, S. H., and Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338. doi:10.1038/nature10886

Wyatt, D. W., Feng, W., Conlin, M. P., Yousefzadeh, M. J., Roberts, S. A., Mieczkowski, P., et al. (2016). Essential roles for polymerase θ-mediated end joining in the repair of chromosome breaks. *Mol. Cell.* 63, 662–673. doi:10.1016/j.molcel.2016.06.020

Xu, L., Zhang, C., Li, H., Wang, P., Gao, Y., Mokadam, N. A., et al. (2021). Efficient precise *in vivo* base editing in adult dystrophic mice. *Nat. Commun.* 12, 3719. doi:10.1038/s41467-021-23996-y

Yang, L., Tang, J., Ma, X., Lin, Y., Ma, G., Shan, M., et al. (2021). Progression and application of CRISPR-Cas genomic editors. *Methods* 194, 65–74. doi:10.1016/j.ymeth.2021.03.013

Yanik, M., Ponnam, S. P. G., Wimmer, T., Trimborn, L., Müller, C., Gambert, I., et al. (2018). Development of a reporter system to explore MMEJ in the context of replacing large genomic fragments. *Mol. Ther. - Nucleic Acids* 11, 407–415. doi:10.1016/j.omtn.2018.03.010

Yeh, W.-H., Shubina-Oleinik, O., Levy, J. M., Pan, B., Newby, G. A., Wornow, M., et al. (2020). *In vivo* base editing restores sensory transduction and transiently improves auditory function in a mouse model of recessive deafness. *Sci. Transl. Med.* 12, eaay9101. doi:10.1126/scitranslmed.aay9101

Yousefzadeh, M. J., Wyatt, D. W., Takata, K., Mu, Y., Hensley, S. C., Tomida, J., et al. (2014). Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. *PLoS Genet.* 10, e1004654. doi:10.1371/journal.pgen.1004654

Yue, W. W., Mackinnon, S., and Bezerra, G. A. (2019). Substrate reduction therapy for inborn errors of metabolism. *Emerg. Top. Life Sci.* 3, 63–73. doi:10.1042/ETLS20180058

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A. H., Shea, D., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* 26, 535–541. doi:10.1038/s41591-020-0790-y

Zeng, Y., Li, J., Li, G., Huang, S., Yu, W., Zhang, Y., et al. (2018). Correction of the Marfan syndrome pathogenic FBN1 mutation by base editing in human

cells and heterozygous embryos. *Mol. Ther.* 26, 2631–2637. doi:10.1016/j.ymthe.2018.08.007

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell.* 163, 759–771. doi:10.1016/j.cell.2015.09.038

Zhang, X.-P., Liu, F., and Wang, W. (2011). Two-phase dynamics of p53 in the DNA damage response. *Proc. Natl. Acad. Sci. U. S. A.* 108, 8990–8995. doi:10.1073/pnas.1100600108

Zhang, X., Zhu, B., Chen, L., Xie, L., Yu, W., Wang, Y., et al. (2020). Dual base editor catalyzes both cytosine and adenine base conversions in human cells. *Nat. Biotechnol.* 38, 856–860. doi:10.1038/s41587-020-0527-y

Zhou, L., Su, J., Long, J., Tao, R., Tang, W., Qin, F., et al. (2022). A universal strategy for AAV delivery of base editors to correct genetic point mutations in neonatal PKU mice. *Mol. Ther. - Methods & Clin. Dev.* 24, 230–240. doi:10.1016/j.omtm.2022.01.001

Zohra, T., Numan, M., Ikram, A., Salman, M., Khan, T., Din, M., et al. (2021). Cracking the challenge of antimicrobial drug resistance with CRISPR/Cas9, nanotechnology and other strategies in ESKAPE pathogens. *Microorganisms* 9, 954. doi:10.3390/microorganisms9050954

Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., et al. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364 (6437), 289–292. Epub 2019/03/02. doi:10.1126/science.aav9973

frontiers | Frontiers in Genome Editing

TYPE Review
PUBLISHED 10 November 2022
DOI 10.3389/fgeed.2022.1050507

OPEN ACCESS

EDITED BY
Ayal Hendel,
Bar-Ilan University, Israel

REVIEWED BY
Toni Cathomen,
Medical Center—University of Freiburg, Germany
Eli J. Fine,
National Resilience, United States
Dimitrios Laurin Wagner,
Berlin Center for Advanced Therapies (BeCAT), Charité—Universitätsmedizin Berlin, Germany

*CORRESPONDENCE
M. Kyle Cromer,
kyle.cromer@ucsf.edu

SPECIALTY SECTION
This article was submitted to Genome Editing in Blood Disorders, a section of the journal Frontiers in Genome Editing

RECEIVED 21 September 2022
ACCEPTED 26 October 2022
PUBLISHED 10 November 2022

CITATION
Wienert B and Cromer MK (2022), CRISPR nuclease off-target activity and mitigation strategies.
Front. Genome Ed. 4:1050507.
doi: 10.3389/fgeed.2022.1050507

CRISPR nuclease off-target activity and mitigation strategies

Beeke Wienert[1] and M. Kyle Cromer[2,3,4]*

[1]Graphite Bio, Inc., South San Francisco, CA, United States, [2]Department of Surgery, University of California, San Francisco, San Francisco, CA, United States, [3]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA, United States, [4]Eli and Edythe Broad Center for Regeneration Medicine, University of California, San Francisco, San Francisco, CA, United States

The discovery of CRISPR has allowed site-specific genomic modification to become a reality and this technology is now being applied in a number of human clinical trials. While this technology has demonstrated impressive efficacy in the clinic to date, there remains the potential for unintended on- and off-target effects of CRISPR nuclease activity. A variety of *in silico*-based prediction tools and empirically derived experimental methods have been developed to identify the most common unintended effect—small insertions and deletions at genomic sites with homology to the guide RNA. However, large-scale aberrations have recently been reported such as translocations, inversions, deletions, and even chromothripsis. These are more difficult to detect using current workflows indicating a major unmet need in the field. In this review we summarize potential sequencing-based solutions that may be able to detect these large-scale effects even at low frequencies of occurrence. In addition, many of the current clinical trials using CRISPR involve *ex vivo* isolation of a patient's own stem cells, modification, and re-transplantation. However, there is growing interest in direct, *in vivo* delivery of genome editing tools. While this strategy has the potential to address disease in cell types that are not amenable to *ex vivo* manipulation, *in vivo* editing has only one desired outcome—on-target editing in the cell type of interest. CRISPR activity in unintended cell types (both on- and off-target) is therefore a major safety as well as ethical concern in tissues that could enable germline transmission. In this review, we have summarized the strengths and weaknesses of current editing and delivery tools and potential improvements to off-target and off-tissue CRISPR activity detection. We have also outlined potential mitigation strategies that will ensure that the safety of CRISPR keeps pace with efficacy, a necessary requirement if this technology is to realize its full translational potential.

KEYWORDS

Gene therapy, off-target activity, *in vivo* delivery, genome editing, CRISPR/Cas9, next-generating sequencing

Introduction

Gene therapy and off-target genome editing

Gene therapy to correct, add, or modify genes holds great promise for many genetic disorders, including hemoglobinopathies, immunodeficiencies, and lysosomal storage disorders. Historically, gene therapy referred to viral-mediated gene addition, however this has the potential to disrupt essential genes or activate oncogenes due to semi-random genomic integration (Hacein-Bey-Abina et al., 2008). Gene editing tools such as CRISPR-Cas, TALENs, mega nucleases, or zinc finger nucleases have thus emerged as exciting alternatives due to the ability to target them to specific sites in the genome. Among these, the more straightforward and modular design of CRISPR guide RNAs (gRNA), which direct the Cas protein to a complementary site in the genome, has made them the preferred tool for both research and clinical applications. Ongoing clinical trials using CRISPR-modified cells have published results without any adverse events for both genome editing in T cells (Lu et al., 2020; Stadtmauer et al., 2020) and hematopoietic stem and progenitor cells (HSPCs) (Frangoul et al., 2021). In addition, the first clinical trial using CRISPR-Cas9 to treat transthyretin amyloidosis by editing hepatocytes *in vivo* has reported disease phenotype improvements in a small group of patients (Gillmore et al., 2021). These early clinical trials highlight the immense potential of CRISPR-Cas9 to treat disease, albeit lacking long-term follow-up data to support safety in humans.

One concern of clinical genome editing is the potential to cause unintended DNA alterations that may have a detrimental effect on cellular function (Figure 1). These undesired consequences can stem from on-target or off-target edits causing unwanted insertions and deletions (indels) or larger rearrangements (structural variants (SVs)) such as translocations, inversions, and duplications. The field has made great progress in developing methods to detect undesired editing events *in silico*, in cell-free DNA *in vitro*, and in live cells *ex vivo* (Blattner et al., 2020), but often it is challenging to link genomic alterations to their impact on cellular health and function. For example, off-target indels occurring in a gene desert may have no phenotypic effect, while some indels at the on-target site may lead to aberrant mRNA and protein products (Tuladhar et al., 2019) that significantly impact cell function (Lindeboom et al., 2019). As CRISPR-Cas9 genome editing moves towards *in vivo* therapeutic applications, making this link becomes even more critical as rare events could be detrimental if occurring in an oncogenic context. In addition, *in vivo* applications carry the risk of both on- and off-target genome edits in an unintended cell type such as the germline or other tissues.

This field is rapidly evolving and new technologies to quantify unintended modifications are continuously being developed and evaluated. Furthermore, there is no broad consensus on the most appropriate measures needed to be taken to comprehensively assess the frequency and risks of CRISPR off-target activity, both for publication in high-impact journals or pre-clinically for the FDA. This review will therefore summarize the state of the field in terms of the current methods to evaluate on- and off-target gene edits, recent advances in method development for both *ex vivo* and *in vivo* editing workflows, and strategies for mitigation and reduction of off-target and off-tissue edits altogether.

Methods to find off-targets sites in genome editing applications

One main advantage of CRISPR-Cas editing over viral genome addition is that it is specifically targeted to a gene locus rather than dependent on random integration. However, as the genome encompasses billions of base pairs, it is possible that the CRISPR-targeted sequence has a near-match elsewhere. Indeed, it has been shown that Cas9 and other

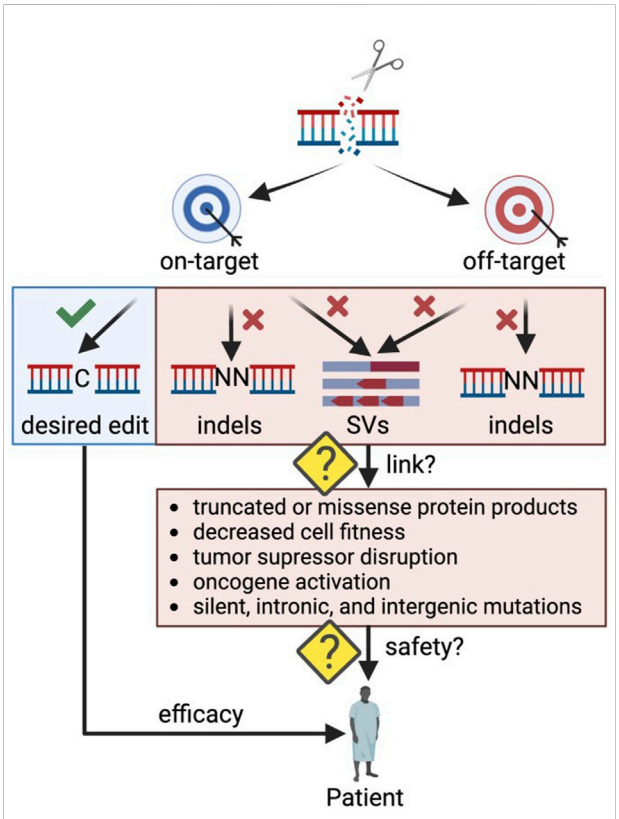


FIGURE 1
Outstanding questions to ensure safety in therapeutic genome editing applications. Indels: insertions and deletions; SVs: structural variants. Created with BioRender.com.

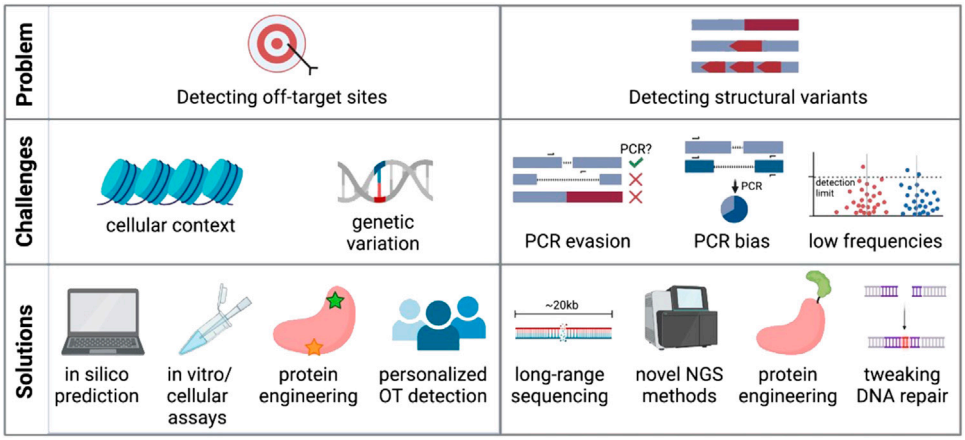


FIGURE 2
Challenges and potential solutions of current off-target detection methods. Created with BioRender.com.

nucleases will often cut highly homologous sequences depending on the location of the mismatch and the genomic context (Mali et al., 2013a; Fu et al., 2013; Hsu et al., 2013; Pattanayak et al., 2013). Finding these off-target sites is critical so that the risk of unintended genomic events can be assessed and minimized (Figure 2).

Many computational tools are available to identify highly homologous genomic sequences and thus predict potential off-target sites for CRISPR-Cas9 activity (Bae et al., 2014; Cradick et al., 2014; Heigwer et al., 2014; Montague et al., 2014; Xiao et al., 2014; Zhu et al., 2014; Singh et al., 2015; Stemmer et al., 2015; Concordet and Haeussler, 2018; Aprilyanto et al., 2021). The application of machine learning to large experimental datasets has further improved the predictive power of these bioinformatic tools (Allen et al., 2018; Shen et al., 2018; Xiang et al., 2021), although these algorithms are biased towards the input gRNAs and reference genomes used to build their predictions. After identifying possible regions of off-target activity, screening must be done in genome-edited cells to confirm whether these sites show signs of CRISPR activity. This is most often done using targeted sequencing of candidate sites with standard sequencing panels achieving detection of variants at or below 5% frequency (Starks et al., 2021), while the detection limit for indels might be lower (0.2-1%) depending on sequencing depth (Chaudhari et al., 2020).

Alternatively, a number of experimental methods have been developed to find off-target sites that may not have been bioinformatically predicted based on homology to the gRNA. These methods vary widely in their approach and even starting material, using cell-free genomic DNA *in vitro* (Kim et al., 2015; Cameron et al., 2017; Tsai et al., 2017; Kim and Kim, 2018; Lazzarotto et al., 2020), in intact live cells *ex vivo* (Crosetto et al., 2013; Tsai et al., 2015; Yan et al., 2017; Wienert et al., 2019; 2020; Zhu et al., 2019; Dobbs et al., 2022), and *in vivo* animal models (Akcakaya et al., 2018; Wienert et al., 2019; Liang et al., 2022). Methods that use cell- and nucleosome-free DNA generally report the highest number of off-targets, many of which cannot be verified in a cellular context (Cromer et al., 2022a). Furthermore, methods such as GUIDE-Seq have been shown to identify more off-target sites in immortalized cell lines than when assaying primary cells (Shapiro et al., 2020). This highlights the importance of chromatin context and DNA repair factors in determining therapeutically relevant off-target activity. And even when using intact cells as input into these assays, conclusions drawn from immortalized cell lines with accumulated variants, distorted karyotypes, and dysfunctional DNA repair pathways (Mittelman and Wilson, 2013; Passerini et al., 2016) may confound the clinical relevance of identified off-target edits.

In spite of these myriad strategies for detection of off-target indels, recent work has shown that *ex vivo* editing in HSPCs elicits very few bona fide off-target sites (<1 true off-target site per gRNA) when using clinically relevant workflows (Cromer et al., 2022b). Of bona fide sites, all were highly homologous to the target sequence and predicted by the majority of *in silico* methods included in the study. However, several sources have shown that genetic variation amongst people can impact off-target activity (Lessard et al., 2017; Lazzarotto et al., 2020; Cromer et al., 2022a). Therefore, implementing a personalized patient-specific workflow in gene therapy products may be needed to circumvent the issue. Common SNPs can be taken into account for *in silico* prediction, *in vitro* methods could be personalized by using genomic DNA extracted from a patient sample (Tsai et al., 2017; Lazzarotto et al., 2020), and some cellular methods are amenable to use on primary cells from patients (Wienert et al., 2019). However, this type of personalized

off-target analysis is limited by cost, logistical feasibility, and the availability of patient material.

Taken together, researchers have a broad range of tools available that allow them to identify potential off-target sites *in silico*, in cell-free DNA *in vitro*, in live cells *ex vivo*, and in animal models. By applying these tools sensibly in the experimental design of therapeutic genome editing strategies, off-target gene editing can be identified, and measures can be taken to minimize these unintended events.

Methods to detect structural variation at double-strand breaks

While the off-target detection methods described above are most useful for identifying localized effects of DNA double-strand breaks (DSBs), larger scale off-target effects have been observed. These include gross chromosomal rearrangements such as translocations (Bothmer et al., 2020; Stadtmauer et al., 2020; Samuelson et al., 2021), chromothripsis (Leibowitz et al., 2021), and aneuploidy (Amendola et al., 2022; Nahmad et al., 2022). Translocation events most often occur as a consequence of: 1) on-target cleavage and recombination with a homologous region of the genome (Turchiano et al., 2021); 2) simultaneous cleavage at an on-target and off-target sequence (Lattanzi et al., 2021); or 3) following multiple on-target cleavage events in multiplexed editing workflows (Qasim et al., 2017; Bothmer et al., 2020; Stadtmauer et al., 2020; Samuelson et al., 2021; Diorio et al., 2022). In addition, large-scale deletions either surrounding the cut site or of the distal end of the chromosome can occur (Cullot et al., 2019), as well as copy-neutral loss-of-heterozygosity (Boutin et al., 2021).

Although targeted amplicon sequencing is commonly used to report on small indels at the cut site, most standard sequencing methods only allow sequencing of relatively short amplicons (hundreds of base pairs). Detecting and quantifying large-scale, multi-kilobase events with PCR-based sequencing methods thus remains challenging (Figure 2B). This is due to several reasons: 1) any deletion that eliminates primer binding sites would not be efficiently amplified with standard sequencing methods and would be missed; 2) if the primer binding sites are preserved larger deletions could skew the PCR reaction towards shorter amplicons and overestimate deletion events; and 3) other undesired on-target events may include inversions, gene duplications, and large insertions (Skryabin et al., 2020) that may also evade detection by PCR-based methods. New No-Amp long-range sequencing protocols avoid PCR and instead use CRISPR-Cas9 to enrich for the sequence of interest up to tens of kilobases. This PCR-free strategy circumvents size bias and can identify large deletions and other structural variants at the target site. However, limited read depth can make it difficult to detect and quantify low-frequency events and low base-calling accuracy of some methods may not achieve single base pair resolution (Lang et al., 2020). However, if these sequencing methods are able to improve and become cheaper, they may become the standard for evaluating structural variants after genome editing.

In addition to large deletions, other genomic abnormalities remain technically challenging to capture, especially when occurring at low frequency. To identify translocations of the on-target site with other genomic regions, several assays have been developed (Zheng et al., 2014; Hu et al., 2016; Qasim et al., 2017; Giannoukos et al., 2018; Yin et al., 2019; Turchiano et al., 2021) which use a sequence at the on-target site as "bait" and next-generation sequencing and bioinformatics to identify the "prey." This allows identification of genomic sequences that have been fused to the on-target site. Another bioinformatic approach analyzes multiplexed-PCR data for on- and off-target sites using a pipeline specific for translocation detection (Amit et al., 2021) which could allow these events to be quantified from pre-existing data.

The delivery modality of the genome editing tools can also introduce unintended effects. For instance, if the nuclease or DNA repair template is delivered by adeno-associated virus (AAV), there is the possibility that non-homologous integration of inverted terminal repeats (ITRs) could occur (Hanlon et al., 2019; Nelson et al., 2019), however this may have minimal effect on adjacent genes if the template is promoter-less. The introduction of DSBs by a nuclease increases the amount of non-homologous integration of AAV vector sequences (Miller et al., 2003; Miller et al., 2004), however the overall frequency seems to be determined by the genomic context and can range between 0.06% and 12.5% of total events (Hanlon et al., 2019). These rare events can be captured with long-range sequencing methods or by a recently developed next-generation sequencing method, named ITR-Seq (Breton et al., 2020), which can identify and quantify ITR integrations on a genome-wide basis independent of the on-target site.

In summary, the field has made great progress in developing methods that can identify structural variants including deletions, inversions, duplications, insertions, and translocations. However, absolute quantification of these events remains challenging due to their low frequency of occurrence. While promising, novel long-range sequencing strategies are still lacking read depth and quality compared to traditional sequencing methods. As structural variants are diverse it is currently not possible for a single assay to capture all possible events, but future advances in sequencing technology could allow for this to become a reality.

Methods to identify unintended editing events *in vivo*

Most CRISPR-based therapies currently in the clinic rely on isolation of patient-derived stem cells, *ex vivo* modification, and re-transplantation. This approach thereby addresses the limited availability of matched donors and risk of immune rejection or

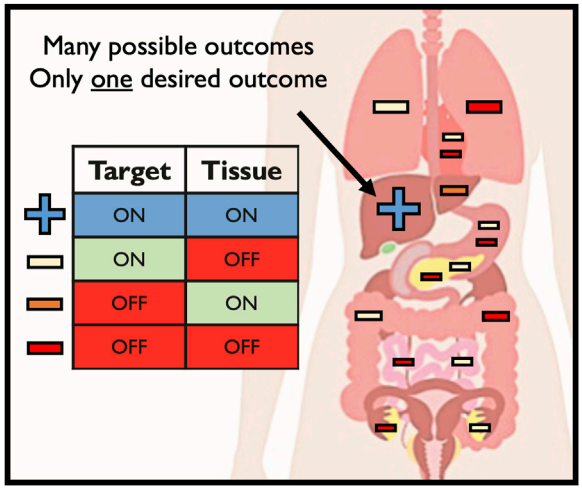


FIGURE 3
Possible outcomes of in vivo genome editing.

graft-versus-host-disease associated with allogeneic transplantation. However, these strategies are only compatible with cell types that may be safely isolated, modified *ex vivo*, and transplanted back into the patient, such as HSPCs. Therefore, the next frontier will be to deliver genome-editing components to modify cells directly where they reside in the body.

Toward this end, many delivery modalities have been developed and optimized to transduce clinically relevant cell types *in vivo* (Long et al., 2016; Goldstein et al., 2019; Mangeot et al., 2019; Gillmore et al., 2021). These platforms are now being used to package and deliver CRISPR-based editing tools *in vivo*, which has shown initial success in the first human clinical trials. One of the most prominent of these trials was conducted by Intellia where a liver-tropic lipid nanoparticle (LNP) was used to deliver Cas9 mRNA along with a gRNA specific to the *TTR* gene in order to treat transthyretin amyloidosis (Gillmore et al., 2021). This strategy effectively lowered serum TTR levels up to 87% from baseline in human patients, serving as a landmark study for efficacy of Cas9 to achieve a clinical endpoint. While *in vivo* Cas9 delivery was found to be quite effective in this instance, there was limited data collected to confirm safety aside from the absence of severe adverse events in these patients.

When delivering editing tools *in vivo*, there is only a single desired outcome—on-target editing at the intended CRISPR cleavage site in the intended target tissue (Figure 3). However, a number of unintended consequences can occur following delivery of editing tools to patients *in vivo*, such as: 1) off-target genomic activity in the intended target tissue (off-target, on-tissue); 2) on-target genomic activity in unintended tissue types (on-target, off-tissue); and 3) off-target genomic activity in unintended tissue types (off-target, off-tissue). Off-tissue events in the gonads are of particular clinical and ethical concern since these could result in changes to the germline which may be transmitted to a patient's offspring (Turocy et al., 2021). Despite these fears and the use of methods to detect off-target Cas9 activity *in vivo* in animal models (Akcakaya et al., 2018; Wienert et al., 2020; Liang et al., 2022), no study to date has investigated the frequency of unintended events following delivery of editing tools to human patients in a clinical context *in vivo*. In the seminal TTR Cas9-LNP paper, the only investigation into off-target activity was done by performing GUIDE-Seq *ex vivo* in hepatocytes (Gillmore et al., 2021). While this is helpful in locating sites of potential off-target activity in the patient's genome, these results were not validated *in vivo*. In the simplest form, clinically routine liver biopsies could have been performed pre- and post-delivery (perhaps at sites both near and far from the hepatic artery where the LNP would have entered the liver) to quantify the frequency of on- and off-target activity at the on-tissue site. However, this approach would yield little insight into CRISPR activity outside the liver, even though this LNP was reported to edit the spleen, adrenal glands, and ovaries at detectable frequencies. While the liver may be easily biopsied, this is not a routine procedure for many other tissues, particularly the ovaries. This therefore presents a major barrier to ensuring the safety of *in vivo* CRISPR delivery.

A potential source of genetic material that could be obtained in a minimally invasive fashion to determine the frequency of intended and unintended editing events following *in vivo* CRISPR delivery is cell-free genomic DNA (cfDNA). CfDNA is primarily derived from dying cells that release their genomic material into the bloodstream. Because of this, assaying cfDNA may be a powerful approach for detecting potentially pathogenic effects of CRISPR delivery, either in terms of genotoxic or oncogenic editing events. While cfDNA has primarily been used in the diagnostic space to detect occurrence/relapse of cancer (Bronkhorst et al., 2019), this technology is sensitive enough to sample maternal blood to discover *de novo* mutations in the fetus during pregnancy (Kitzman et al., 2012). In fact, a proof-of-concept study used cfDNA to map insertion sites following *in vivo* delivery of lentiviral vectors (Cesana et al., 2021). Therefore, a comparable strategy could be employed to quantify the frequency of on- and off-target cleavage activity following *in vivo* CRISPR delivery. However, unlike the workflow developed for mapping lentiviral insertions which relies on sequencing outward from the vector backbone, mapping sites of CRISPR activity may be aided by defining high-likelihood regions of activity. This could be done using *in silico* or empirical off-target detection methods defined above, and candidate regions could then be probed for indels by targeted deep sequencing of cfDNA. An alternative approach that would not require defining CRISPR off-target sites *a priori* could be to use translocations as a surrogate for on- and off-target activity by adapting technologies such as LAM-HTGTS (Hu et al., 2016), CAST-Seq (Turchiano et al., 2021), or PEM-Seq (Yin et al., 2019) to use patient cells or cfDNA as input. Furthermore, if using

cfDNA isolated from peripheral blood is successful, a similar approach could be used to detect occurrence of CRISPR activity in cerebrospinal fluid (CSF) as well to quantify the ability of genome editing tools to cross the blood-brain barrier and edit cells residing in the brain, which may not be safely biopsied.

While cfDNA presents an opportunity to quantify on-target and off-target editing, it may give little insight into the tissue of origin. To shed light on this without invasive biopsies, the use of cell-free RNA (cfRNA) is a possibility. Analogous to the use of cfDNA, workflows to isolate cfRNA from the peripheral blood have been developed that allow insight into the tissue of origin due to the predominant expression patterns of cells releasing DNA and RNA into the bloodstream (Cheung et al., 2019). While the degree to which tissue-of-origin could be gleaned from this approach has yet to be fully explored, the investigation of on- and off-target CRISPR activity at expected cleavage sites in cfRNA could determine whether intended or unintended genome editing results in changes to the transcriptome. Since sites of CRISPR off-target activity typically reside in intergenic regions of the genome with no known functional significance (Cromer et al., 2022b), it is possible that CRISPR activity will be detected in cfDNA, but not in transcribed cfRNA. This could be an important measure to assay efficacy and safety of *in vivo* editing immediately following therapeutic delivery as well as over time. The combined cfDNA/RNA approach could also be an effective way to detect pathogenic clonal expansion of edited cells following treatment. In this specific use case, it may not be necessary to have identified the initiating driver genomic event, but even oligoclonality of passenger events—such as a particular indel at the on- or common off-target sites—could allow us to infer that clonal expansion is occurring. Importantly, the strategies proposed are most likely to capture and monitor the frequency of small site-specific indels, and more sophisticated methods (like those described earlier) may be needed to identify large genomic aberrations in cfDNA/RNA.

Linking genomic outcomes to off-target significance

Even when we can successfully identify off-target CRISPR effects, determining if an unintended editing event is of concern to the patient's health remains challenging (Figure 1). Broadly speaking, off-target genomic events are most likely to either elicit no effect or result in a loss or gain of fitness. While loss of fitness will likely result in drop out of the cell carrying the undesired event, gains of fitness are of greater concern due to the possibility of oncogenicity. Although site-specific off-target effects are infrequent (Cromer et al., 2022b), in the event that they do occur, the likelihood of directly disrupting another gene is small (only 1% of the human genome is coding DNA and only 7.2% of predicted off-target sites for exon-targeting gRNAs fall in exonic regions). And while modifications to non-coding DNA sequences may alter gene expression patterns or modify elements with as-yet-unknown important functions (ENCODE Project Consortium, 2012) interpreting non-coding genomic disruptions is difficult. As our understanding of the function of non-coding regions of DNA improves, we may better predict the impact of off-targets modifications in the future. Until then, we must rely on methods that can measure oncogenicity and toxicity from off-target modifications events *in vitro* or *in vivo*.

The most conventional approach to assess tumorigenicity of cell products is implanting cells at an ectopic site in immunodeficient mice followed by monitoring for tumor growth and other adverse events (Human Gene Therapy Products Incorporating Human Genome Editing | FDA). One major caveat of this method is that it has limited sensitivity, depends on the animal model used, and may miss low frequency events (Sato et al., 2019). Alternative *in vivo* approaches have developed technology to track clonality of cell-based products following *ex vivo* HSPC editing and transplantation through barcodes included in the HDR template (Ferrari et al., 2021; Sharma et al., 2021) or by tracking indel diversity (Magis et al., 2022). These technologies can identify clones that have expanded abnormally and hint towards genomic events that led to the oncogenic transformation. Currently these approaches have been limited to research applications but could potentially also be incorporated in therapeutic workflows in the future. However, *in vivo* studies are time- and cost-intensive and can slow down the drug development process immensely. Thus, *in vitro* studies that measure oncogenicity or genomic instability would be preferred, though these may not properly recapitulate *in vivo* activity. While performing whole genome sequencing on every cell product for every patient would ensure an unbiased approach of variant discovery across the whole genome, the currently limited coverage per base pair would miss low-frequency events. Using an intermediate approach of exome sequencing the most commonly mutated oncogenes and tumor suppressors increases read depth significantly and could provide a feasible alternative to assess the safety of *ex vivo* gene therapy drug products (Cromer et al., 2022a).

While the above work focuses on unintended off-target effects, even unintended *on*-target effects can lead to adverse effects. For instance, when a therapeutic editing approach targets a coding sequence—like knocking out a pathogenic gene to correct a disease phenotype (Gillmore et al., 2021)—an array of indels will form at the cut site. A recent report has shown that Cas9-induced indels can result in the formation of disrupted, non-natural mRNAs, which can be translated into aberrant protein products (Tuladhar et al., 2019). This study found that indels can induce internal ribosome entry sites to produce alternative mRNAs or induce exon skipping by disrupting exon-splicing enhancers. The same study also provided a bioinformatic tool to help design gRNAs to avoid such events (Tuladhar et al., 2019). Since truncated protein products could exert a dominant negative function (Savas

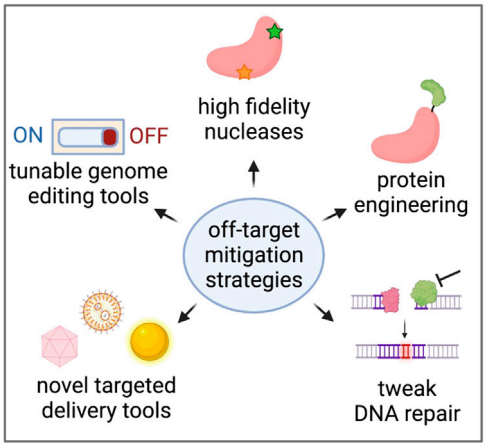


FIGURE 4
Strategies to reduce unintended genome editing events. Created with BioRender.com.

et al., 2006), potential undesired translated proteins should be studied carefully. Properly characterizing the genome-edited cell population by combining on-target amplicon sequencing with mRNA sequencing and proteomics may allow us to identify and develop strategies to reduce the occurrence of such events.

Taken together, linking genomic events to oncogenicity is difficult and currently available *in vitro* and *in vivo* assays often lack sensitivity. Progress has been made to develop barcoding technologies that can track transformed cells and next-generation sequencing methods such as exome and RNA sequencing can also help identify oncogenic events.

Approaches to reduce off-target and off-tissue editing

As we learn more about the types of editing events that can occur at on- and off-target sites, many researchers are developing methods to reduce off-target effects altogether. These efforts range from protein engineering to make nucleases more specific to the discovery of novel, more specific delivery vehicles of genome editing reagents *in vivo* (Figure 4).

Careful nuclease selection and gRNA selection is often the first step when designing a *de novo* genome editing strategy (Lee et al., 2016). While a number of CRISPR nucleases have been discovered (Swarts and Jinek, 2018; Li and Peng, 2019), the majority of clinical efforts to date have used one of the original enzymes characterized, *Streptococcus pyogenes* Cas9 (SpCas9) (Jinek et al., 2012). This nuclease is one of the most common due to its relatively unrestrictive protospacer adjacent motif (PAM) and its high frequency of cleavage activity at a wide variety of loci in a range of cell types across a number of organisms, from humans to Arabidopsis (Mali et al., 2013b; Miki et al., 2018). While this nuclease typically has few genomic sites of off-target activity, some of these can be cut at high frequencies (>30% of alleles harboring indels), depending on the specificity of the particular gRNA (Cromer et al., 2022a). To address this, more specific versions of Cas9 have been engineered which reduce off-target activity by > 20-fold (Chen et al., 2017; Vakulskas et al., 2018; Bravo et al., 2022; Kulcsar et al., 2022). In doing so, incorporation of these higher-fidelity Cas9 variants has been shown to reduce the risk of large-scale genomic rearrangements (Turchiano et al., 2021). In addition to engineering more specific nucleases, a study that fused Cas9 to the exonuclease TREX2 in order to prevent perfect DNA repair reported significantly fewer large deletions and nearly eliminated chromosomal translocations during multiplex editing in T cells (Yin et al., 2022).

The format in which Cas9 is delivered—most often DNA, mRNA, or ribonucleoprotein (RNP)—will alter Cas9 expression and duration of exposure. This in turn has been shown to impact off-target activity, with short half-life RNP and mRNA resulting in lower off-target activity than longer-lived DNA formulations (Cameron et al., 2017; Lu et al., 2019; Zhang et al., 2021). In addition, tunable/inducible control strategies have been incorporated to regulate CRISPR expression using bioavailable small molecules, (Truong et al., 2015; Zhao et al., 2018), light (Nihongaki et al., 2015), and even magnets (Hsu and Hu, 2019). Similarly, other groups have developed self-inactivating Cas9 and AAV delivery vectors that may prevent prolonged exposure to genome editing tools and therefore reduce the likelihood of unintended activity or genomic events (Li et al., 2019; Ibraheim et al., 2021). However, depending on the tools these strategies are applied to, there remains the potential for off-target activity or large-scale genomic rearrangements following creation of DSBs. Furthermore, in their current forms, most approaches are only compatible with *ex vivo* editing workflows where high efficiency delivery of large payloads is possible.

All nuclease-based genome editing applications rely on DSB resolution, therefore modifying the cell's natural DNA damage repair pathways has emerged as a strategy for increasing the ratio of intended to unintended genomic events (Yeh et al., 2019; Xue and Greene, 2021). For example, transiently inhibiting non-homologous end joining (NHEJ)-mediators such as 53BP1 or DNA-dependent protein kinase catalytic subunit can decrease indels and increase precise genome editing outcomes through homology-directed repair (Robert et al., 2015; Canny et al., 2018; Riesenberg and Maricic, 2018; Riesenberg et al., 2019). While some work has been done to determine which DNA repair enzymes are responsible for the formation of small indels (Hussmann et al., 2021), less is known about the factors that promote large deletions. Recently, a study that used a clonal library of embryonic stem cells deficient for DNA repair genes

found that inhibition of NHEJ-mediating enzymes increased frequencies of large deletions, while inhibition of microhomology-mediated end joining-mediating proteins decreased them (Kosicki et al., 2022). Of course, it is crucial to ensure that temporarily inhibiting DNA repair enzymes does not affect other regions of the cell's genome. Another study has shown that the presence of an HDR template such as a single-stranded oligodeoxynucleotide or AAV donor can reduce the frequency of large deletions by 50%–80% (Wen et al., 2021), emphasizing the importance of testing unintended editing outcomes in the context of both the nuclease and the DNA donor.

While Cas9 nuclease technology continues to improve, recent editing tools replace this nuclease with a nickase to introduce single base pair changes or small site-specific modifications, most commonly in the form of single or dual nickase editing (Ran et al., 2013), base editing (Komor et al., 2016), or prime editing (Anzalone et al., 2019). Although these tools avoid formation of DSBs and likely reduce the frequency of large-scale genomic rearrangements, there is still the possibility of unintended off-target activity. In the case of base and prime editors, this arises from the tethering of Cas9 with deaminases and reverse transcriptases, respectively. In fact, some studies have reported that base editors can initiate off-target activity at sites with little homology to the gRNA (Jin et al., 2019; Zuo et al., 2019; Lei et al., 2021) and in a significant proportion of cellular mRNA (Grünewald et al., 2019a). In addition, base editor-induced modifications are often single nucleotide variants which are more difficult to detect by next-generation sequencing than localized indels introduced by traditional CRISPR nucleases. Ongoing efforts continue to engineer improved versions of these base editors to reduce off-target activity (Rees et al., 2017; Grünewald et al., 2019b; Li et al., 2022).

The above advances primarily concern the editing tools themselves, which is most likely to boost on-target effects and reduce unintended off-target consequences. However, these improvements will likely have limited impact on the ratio of on-tissue to off-tissue activity following *in vivo* delivery of editing tools. Toward this end, many groups are working to improve specificity of the delivery modalities themselves. This includes screening for vectors or nanoparticles that have specific tissue tropisms, such as those optimized to cross the blood-brain-barrier, to transduce vascular tissue, and more (White et al., 2004; Choudhury et al., 2016; Sago et al., 2018; Boehnke et al., 2022). There also are efforts to conjugate cell type-specific antibodies to delivery vectors to improve targeted *in vivo* delivery (Tombácz et al., 2021). While preliminary, this approach may be an effective means to improve on-tissue editing when injecting delivery vectors systemically.

In the early stages of development, but with great translational potential, are strategies to encode logic into cells (i.e., to introduce DNA code capable of responding to a given cellular state). As with CRISPR, many of these efforts use RNA-based homology to facilitate downstream expression of transgenes in the presence of a user-defined RNA sequence (Green et al., 2014, 2017; Kaseniit et al., 2022). Several proof-of-concept studies demonstrated that this technology could be used to encode complex logic into cells, such as multi-input OR, AND, and NOT gates. While much of this work was done in *E. coli* or human cell lines, if an analogous system was ported to clinically relevant primary cells it could allow genome editing tools to only be expressed in cells with a particular gene expression profile—effectively reducing or eliminating off-tissue activity.

Concluding remarks

All the above efforts have been aimed at reducing unintended off-target and off-tissue activity. However, because millions or billions of cells are transplanted with *ex vivo* therapies, and billions or trillions of cells may be transduced with *in vivo* delivery vectors, any degree of unintended activity has the potential to be deleterious. Jennifer Doudna stated that one day she hopes to see CRISPR become a "standard of care" (Jennifer Doudna and William Kearney). If this is ever to become a reality, how do we make these therapies safe enough to be delivered routinely to patients?

While any manipulation to the genome opens the possibility for unwanted genetic events, we believe advances in off-target/off-tissue detection methods and improvements in genome editing tools and delivery modalities will ultimately allow personalized medicine to become a reality. As the development of advanced tools allows us to introduce increasingly complex features (and even logic) into cells, we will likely have to establish increasingly complex safety mechanisms as well. These may include automatic or inducible safety switches that provide a necessary safeguard in the instance of adverse clinical events (Di Stasi et al., 2011; Liang et al., 2018; Martin et al., 2020). And while initial data from CRISPR-based therapies in the clinic (both *ex vivo* and *in vivo*) has shown incredible promise, as greater numbers of patients receive genome editing treatments, we must ensure that editing *safety* keeps pace with *efficacy*. If CRISPR is ever to become the standard-of-care, then all of us—basic biologists, synthetic biologists, bioinformaticists, and clinicians—will have to combine efforts to ensure that genome editing therapies are as safe as possible.

Author contributions

BW and MKC generated figures and wrote the manuscript.

Funding

MKC is funded by an American Society of Gene and Cell Therapy Career Development Award as well as National Heart, Lung, and Blood Institute R01 (HL135607).

Conflict of interest

BW is current employee and equity holder in Graphite Bio, Inc., a for-profit company.

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

Publisher's note

References

Akcakaya, P., Bobbin, M. L., Guo, J. A., Malagon-Lopez, J., Clement, K., Garcia, S. P., et al. (2018). *In vivo* CRISPR editing with no detectable genome-wide off-target mutations. *Nature* 561, 416–419. doi:10.1038/s41586-018-0500-9

Allen, F., Crepaldi, L., Alsinet, C., Strong, A. J., Kleshchevnikov, V., De Angeli, P., et al. (2018). Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat. Biotechnol.* 37, 64–72. doi:10.1038/nbt.4317

Amendola, M., Brusson, M., and Miccio, A. (2022). CRISPRthripsis: The risk of CRISPR/Cas9-induced chromothripsis in gene therapy. *Stem Cells Transl. Med.* 11, 1003–1009. doi:10.1093/stcltm/szac064

Amit, I., Iancu, O., Levy-Jurgenson, A., Kurgan, G., McNeill, M. S., Rettig, G. R., et al. (2021). CRISPECTOR provides accurate estimation of genome editing translocation and off-target activity from comparative NGS data. *Nat. Commun.* 12, 3042. doi:10.1038/s41467-021-22417-4

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157. doi:10.1038/s41586-019-1711-4

Aprilyanto, V., Aditama, R., Tanjung, Z. A., Utomo, C., and Liwang, T. (2021). Crop: A CRISPR/Cas9 guide selection program based on mapping guide variants. *Sci. Rep.* 11, 1504. doi:10.1038/s41598-021-81297-2

Bae, S., Park, J., and Kim, J.-S. (2014). Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. doi:10.1093/bioinformatics/btu048

Blattner, G., Cavazza, A., Thrasher, A. J., and Turchiano, G. (2020). Gene editing and genotoxicity: Targeting the off-targets. *Front. Genome Ed.* 2, 613252. doi:10.3389/fgeed.2020.613252

Boehnke, N., Straehla, J. P., Safford, H. C., Kocak, M., Rees, M. G., Ronan, M., et al. (2022). Massively parallel pooled screening reveals genomic determinants of nanoparticle delivery. *Science* 377, eabm5551. doi:10.1126/science.abm5551

Bothmer, A., Gareau, K. W., Abdulkerim, H. S., Buquicchio, F., Cohen, L., Viswanathan, R., et al. (2020). Detection and modulation of DNA translocations during multi-gene genome editing in T cells. *CRISPR J.* 3, 177–187. doi:10.1089/crispr.2019.0074

Boutin, J., Rosier, J., Cappellen, D., Prat, F., Toutain, J., Pennamen, P., et al. (2021). CRISPR-Cas9 globin editing can induce megabase-scale copy-neutral losses of heterozygosity in hematopoietic cells. *Nat. Commun.* 12, 4922. doi:10.1038/s41467-021-25190-6

Bravo, J. P. K., Liu, M.-S., Hibshman, G. N., Dangerfield, T. L., Jung, K., McCool, R. S., et al. (2022). Structural basis for mismatch surveillance by CRISPR-Cas9. *Nature* 603, 343–347. doi:10.1038/s41586-022-04470-1

Breton, C., Clark, P. M., Wang, L., Greig, J. A., and Wilson, J. M. (2020). ITR-Seq, a next-generation sequencing assay, identifies genome-wide DNA editing sites *in vivo* following adeno-associated viral vector-mediated genome editing. *BMC Genomics* 21, 239. doi:10.1186/s12864-020-6655-4

Bronkhorst, A. J., Ungerer, V., and Holdenrieder, S. (2019). The emerging role of cell-free DNA as a molecular marker for cancer management. *Biomol. Detect. Quantif.* 17, 100087. doi:10.1016/j.bdq.2019.100087

Cameron, P., Fuller, C. K., Donohoue, P. D., Jones, B. N., Thompson, M. S., Carter, M. M., et al. (2017). Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods* 14, 600–606. doi:10.1038/nmeth.4284

Canny, M. D., Moatti, N., Wan, L. C. K., Fradet-Turcotte, A., Krasner, D., Mateos-Gomez, P. A., et al. (2018). Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR-Cas9 genome-editing efficiency. *Nat. Biotechnol.* 36, 95–102. doi:10.1038/nbt.4021

Cesana, D., Calabria, A., Rudilosso, L., Gallina, P., Benedicenti, F., Spinozzi, G., et al. (2021). Retrieval of vector integration sites from cell-free DNA. *Nat. Med.* 27, 1458–1470. doi:10.1038/s41591-021-01389-4

Chaudhari, H. G., Penterman, J., Whitton, H. J., Spencer, S. J., Flanagan, N., Lei Zhang, M. C., et al. (2020). Evaluation of homology-independent CRISPR-Cas9 off-target assessment methods. *The CRISPR Journal* 3, 440–453. doi:10.1089/crispr.2020.0053

Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., et al. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550, 407–410. doi:10.1038/nature24268

Cheung, K. W. E., Choi, S.-Y. R., Lee, L. T. C., Lee, N. L. E., Tsang, H. F., Cheng, Y. T., et al. (2019). The potential of circulating cell free RNA as a biomarker in cancer. *Expert Rev. Mol. diagn.* 19, 579–590. doi:10.1080/14737159.2019.1633307

Choudhury, S. R., Fitzpatrick, Z., Harris, A. F., Maitland, S. A., Ferreira, J. S., Zhang, Y., et al. (2016). *In vivo* selection yields AAV-B1 capsid for central nervous system and muscle gene therapy. *Mol. Ther.* 24, 1247–1257. doi:10.1038/mt.2016.84

Concordet, J.-P., and Haeussler, M. (2018). Crispor: Intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res.* 46, W242–W245. doi:10.1093/nar/gky354

Cradick, T. J., Qiu, P., Lee, C. M., Fine, E. J., and Bao, G. (2014). Cosmid: A web-based tool for identifying and validating CRISPR/cas off-target sites. *Mol. Ther. Nucleic Acids* 3, e214. doi:10.1038/mtna.2014.64

Cromer, M. K., Barsan, V. V., Jaeger, E., Wang, M., Hampton, J. P., Chen, F., et al. (2022a). Ultra-deep sequencing validates safety of CRISPR/Cas9 genome editing in human hematopoietic stem and progenitor cells. *Nat. Commun.* 13, 4724. doi:10.1038/s41467-022-32233-z

Cromer, M. K., Majeti, K. R., Rettig, G. R., Murugan, K., Kurgan, G. L., Hampton, J. P., et al. (2022b). Comparative analysis of CRISPR off-target activity discovery tools following *ex vivo* editing of $CD34^+$ hematopoietic stem and progenitor cells. *BioRxiv*. doi:10.1101/2022.09.09.507306

Crosetto, N., Mitra, A., Silva, M. J., Bienko, M., Dojer, N., Wang, Q., et al. (2013). Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* 10, 361–365. doi:10.1038/nmeth.2408

Cullot, G., Boutin, J., Toutain, J., Prat, F., Pennamen, P., Rooryck, C., et al. (2019). CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun.* 10, 1136. doi:10.1038/s41467-019-09006-2

Di Stasi, A., Tey, S.-K., Dotti, G., Fujita, Y., Kennedy-Nasser, A., Martinez, C., et al. (2011). Inducible apoptosis as a safety switch for adoptive cell therapy. *N. Engl. J. Med.* 365, 1673–1683. doi:10.1056/NEJMoa1106152

Diorio, C., Murray, R., Naniong, M., Barrera, L., Camblin, A., Chukinas, J., et al. (2022). Cytosine base editing enables quadruple-edited allogeneic CART cells for T-ALL. *Blood* 140, 619–629. doi:10.1182/blood.2022015825

Dobbs, F. M., van Eijk, P., Fellows, M. D., Loiacono, L., Nitsch, R., and Reed, S. H. (2022). Precision digital mapping of endogenous and induced genomic DNA breaks by INDUCE-seq. *Nat. Commun.* 13, 3989. doi:10.1038/s41467-022-31702-9

Doudna, Jennifer, and Kearney, William.2020 An interview with nobel prize-winner jennifer Doudna. Issues Sci. Technol.XXXVI Available at: https://issues.org/jennifer-doudna-interview/ (Accessed August 17, 2022).

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi:10.1038/nature11247

FDA Human gene therapy products incorporating human genome editing | FDA.2022 Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/human-gene-therapy-products-incorporating-human-genome-editing (Accessed August 31, 2022).

Ferrari, S., Beretta, S., Jacob, A., Cittaro, D., Albano, L., Merelli, I., et al. (2021). BAR-Seq clonal tracking of gene-edited cells. *Nat. Protoc.* 16, 2991–3025. doi:10.1038/s41596-021-00529-x

Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and β-thalassemia. *N. Engl. J. Med.* 384, 252–260. doi:10.1056/NEJMoa2031054

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826. doi:10.1038/nbt.2623

Giannoukos, G., Ciulla, D. M., Marco, E., Abdulkerim, H. S., Barrera, L. A., Bothmer, A., et al. (2018). UDiTaS^{TM}, a genome editing detection method for indels and genome rearrangements. *BMC Genomics* 19, 212. doi:10.1186/s12864-018-4561-9

Gillmore, J. D., Gane, E., Taubel, J., Kao, J., Fontana, M., Maitland, M. L., et al. (2021). CRISPR-Cas9 *in vivo* gene editing for transthyretin amyloidosis. *N. Engl. J. Med.* 385, 493–502. doi:10.1056/NEJMoa2107454

Goldstein, J. M., Tabebordbar, M., Zhu, K., Wang, L. D., Messemer, K. A., Peacker, B., et al. (2019). *In situ* modification of tissue stem and progenitor cell genomes. *Cell Rep.* 27, 1254–1264. e7. doi:10.1016/j.celrep.2019.03.105

Green, A. A., Kim, J., Ma, D., Silver, P. A., Collins, J. J., and Yin, P. (2017). Complex cellular logic computation using ribocomputing devices. *Nature* 548, 117–121. doi:10.1038/nature23271

Green, A. A., Silver, P. A., Collins, J. J., and Yin, P. (2014). Toehold switches: De-novo-designed regulators of gene expression. *Cell* 159, 925–939. doi:10.1016/j.cell.2014.10.002

Grünewald, J., Zhou, R., Garcia, S. P., Iyer, S., Lareau, C. A., Aryee, M. J., et al. (2019a). Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* 569, 433–437. doi:10.1038/s41586-019-1161-z

Grünewald, J., Zhou, R., Iyer, S., Lareau, C. A., Garcia, S. P., Aryee, M. J., et al. (2019b). CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.* 37, 1041–1048. doi:10.1038/s41587-019-0236-6

Hacein-Bey-Abina, S., Garrigue, A., Wang, G. P., Soulier, J., Lim, A., Morillon, E., et al. (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* 118, 3132–3142. doi:10.1172/JCI35700

Hanlon, K. S., Kleinstiver, B. P., Garcia, S. P., Zaborowski, M. P., Volak, A., Spirig, S. E., et al. (2019). High levels of AAV vector integration into CRISPR-induced DNA breaks. *Nat. Commun.* 10, 4439. doi:10.1038/s41467-019-12449-2

Heigwer, F., Kerr, G., and Boutros, M. (2014). E-CRISP: Fast CRISPR target site identification. *Nat. Methods* 11, 122–123. doi:10.1038/nmeth.2812

Hsu, M.-N., and Hu, Y.-C. (2019). Local magnetic activation of CRISPR. *Nat. Biomed. Eng.* 3, 83–84. doi:10.1038/s41551-019-0354-y

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832. doi:10.1038/nbt.2647

Hu, J., Meyers, R. M., Dong, J., Panchakshari, R. A., Alt, F. W., and Frock, R. L. (2016). Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc.* 11, 853–871. doi:10.1038/nprot.2016.043

Hussmann, J. A., Ling, J., Ravisankar, P., Yan, J., Cirincione, A., Xu, A., et al. (2021). Mapping the genetic landscape of DNA double-strand break repair. *Cell* 184, 5653–5669.e25. e25. doi:10.1016/j.cell.2021.10.002

Ibraheim, R., Tai, P. W. L., Mir, A., Javeed, N., Wang, J., Rodríguez, T. C., et al. (2021). Self-inactivating, all-in-one AAV vectors for precision Cas9 genome editing via homology-directed repair *in vivo*. *Nat. Commun.* 12, 6267. doi:10.1038/s41467-021-26518-y

Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., et al. (2019). Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science* 364, 292–295. doi:10.1126/science.aaw7166

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi:10.1126/science.1225829

Kaseniit, K. E., Katz, N., Kolber, N. S., Call, C. C., Wengier, D. L., Cody, W. B., et al. (2022). Modular and programmable RNA sensing using ADAR editing in living cells. *BioRxiv*. doi:10.1101/2022.01.28.478207

Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H. R., et al. (2015). Digenome-seq: Genome-wide profiling of CRISPR-cas9 off-target effects in human cells. *Nat. Methods* 12 (237–43), 237–243. 1 p following 243. doi:10.1038/nmeth.3284

Kim, D., and Kim, J.-S. (2018). DIG-Seq: A genome-wide CRISPR off-target profiling method using chromatin DNA. *Genome Res.* 28, 1894–1900. doi:10.1101/gr.236620.118

Kitzman, J. O., Snyder, M. W., Ventura, M., Lewis, A. P., Qiu, R., Simmons, L. E., et al. (2012). Noninvasive whole-genome sequencing of a human fetus. *Sci. Transl. Med.* 4, 137ra76. doi:10.1126/scitranslmed.3004323

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi:10.1038/nature17946

Kosicki, M., Allen, F., Steward, F., Tomberg, K., Pan, Y., and Bradley, A. (2022). Cas9-induced large deletions and small indels are controlled in a convergent fashion. *Nat. Commun.* 13, 3422. doi:10.1038/s41467-022-30480-8

Kulcsar, P. I., Talas, A., Welker, E., and Krausz, S. L. (2022). SuperFi-Cas9 exhibits extremely high fidelity but reduced activity in mammalian cells. *BioRxiv*. doi:10.1101/2022.05.27.493683

Lang, D., Zhang, S., Ren, P., Liang, F., Sun, Z., Meng, G., et al. (2020). Comparison of the two up-to-date sequencing technologies for genome assembly: HiFi reads of pacific biosciences sequel II system and ultralong reads of oxford nanopore. *Gigascience* 9, giaa123–7. doi:10.1093/gigascience/giaa123

Lattanzi, A., Camarena, J., Lahiri, P., Segal, H., Srifa, W., Vakulskas, C. A., et al. (2021). Development of β-globin gene correction in human hematopoietic stem cells as a potential durable treatment for sickle cell disease. *Sci. Transl. Med.* 13. doi:10.1126/scitranslmed.abf2444

Lazzarotto, C. R., Malinin, N. L., Li, Y., Zhang, R., Yang, Y., Lee, G., et al. (2020). CHANGE-seq reveals genetic and epigenetic effects on CRISPR-Cas9 genome-wide activity. *Nat. Biotechnol.* 38, 1317–1327. doi:10.1038/s41587-020-0555-7

Lee, C. M., Cradick, T. J., Fine, E. J., and Bao, G. (2016). Nuclease target site selection for maximizing on-target activity and minimizing off-target effects in genome editing. *Mol. Ther.* 24, 475–487. doi:10.1038/mt.2016.1

Lei, Z., Meng, H., Lv, Z., Liu, M., Zhao, H., Wu, H., et al. (2021). Detect-seq reveals out-of-protospacer editing and target-strand editing by cytosine base editors. *Nat. Methods* 18, 643–651. doi:10.1038/s41592-021-01172-w

Leibowitz, M. L., Papathanasiou, S., Doerfler, P. A., Blaine, L. J., Sun, L., Yao, Y., et al. (2021). Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat. Genet.* 53, 895–905. doi:10.1038/s41588-021-00838-7

Lessard, S., Francioli, L., Alfoldi, J., Tardif, J.-C., Ellinor, P. T., MacArthur, D. G., et al. (2017). Human genetic variation alters CRISPR-Cas9 on- and off-targeting specificity at therapeutically implicated loci. *Proc. Natl. Acad. Sci. U. S. A.* 114, E11257–E11266. doi:10.1073/pnas.1714640114

Li, A., Lee, C. M., Hurley, A. E., Jarrett, K. E., De Giorgi, M., Lu, W., et al. (2019). A self-deleting AAV-CRISPR system for *in vivo* genome editing. *Mol. Ther. Methods Clin. Dev.* 12, 111–122. doi:10.1016/j.omtm.2018.11.009

Li, A., Mitsunobu, H., Yoshioka, S., Suzuki, T., Kondo, A., and Nishida, K. (2022). Cytosine base editing systems with minimized off-target effect and molecular size. *Nat. Commun.* 13, 4531. doi:10.1038/s41467-022-32157-8

Li, Y., and Peng, N. (2019). Endogenous CRISPR-cas system-based genome editing and antimicrobials: Review and prospects. *Front. Microbiol.* 10, 2471. doi:10.3389/fmicb.2019.02471

Liang, Q., Monetti, C., Shutova, M. V., Neely, E. J., Hacibekiroglu, S., Yang, H., et al. (2018). Linking a cell-division gene and a suicide gene to define and improve cell therapy safety. *Nature* 563, 701–704. doi:10.1038/s41586-018-0733-7

Liang, S.-Q., Liu, P., Smith, J. L., Mintzer, E., Maitland, S., Dong, X., et al. (2022). Genome-wide detection of CRISPR editing *in vivo* using GUIDE-tag. *Nat. Commun.* 13, 437. doi:10.1038/s41467-022-28135-9

Lindeboom, R. G. H., Vermeulen, M., Lehner, B., and Supek, F. (2019). The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat. Genet.* 51, 1645–1651. doi:10.1038/s41588-019-0517-5

Long, C., Amoasii, L., Mireault, A. A., McAnally, J. R., Li, H., Sanchez-Ortiz, E., et al. (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351, 400–403. doi:10.1126/science.aad5725

Lu, B., Javidi-Parsijani, P., Makani, V., Mehraein-Ghomi, F., Sarhan, W. M., Sun, D., et al. (2019). Delivering SaCas9 mRNA by lentivirus-like bionanoparticles for transient expression and efficient genome editing. *Nucleic Acids Res.* 47, e44. doi:10.1093/nar/gkz093

Lu, Y., Xue, J., Deng, T., Zhou, X., Yu, K., Deng, L., et al. (2020). Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer. *Nat. Med.* 26, 732–740. doi:10.1038/s41591-020-0840-5

Magis, W., DeWitt, M. A., Wyman, S. K., Vu, J. T., Heo, S.-J., Shao, S. J., et al. (2022). High-level correction of the sickle mutation is amplified *in vivo* during erythroid differentiation. *iScience* 25, 104374. doi:10.1016/j.isci.2022.104374

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., et al. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833–838. doi:10.1038/nbt.2675

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013b). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi:10.1126/science.1232033

Mangeot, P. E., Risson, V., Fusil, F., Marnef, A., Laurent, E., Blin, J., et al. (2019). Genome editing in primary cells and *in vivo* using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nat. Commun.* 10, 45. doi:10.1038/s41467-018-07845-z

Martin, R. M., Fowler, J. L., Cromer, M. K., Lesch, B. J., Ponce, E., Uchida, N., et al. (2020). Improving the safety of human pluripotent stem cell therapies using genome-edited orthogonal safeguards. *Nat. Commun.* 11, 2713. doi:10.1038/s41467-020-16455-7

Miki, D., Zhang, W., Zeng, W., Feng, Z., and Zhu, J.-K. (2018). CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nat. Commun.* 9, 1967. doi:10.1038/s41467-018-04416-0

Miller, D. G., Petek, L. M., and Russell, D. W. (2004). Adeno-associated virus vectors integrate at chromosome breakage sites. *Nat. Genet.* 36, 767–773. doi:10.1038/ng1380

Mittelman, D., and Wilson, J. H. (2013). The fractured genome of HeLa cells. *Genome Biol.* 14, 111. doi:10.1186/gb-2013-14-4-111

Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M., and Valen, E. (2014). CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* 42, W401–W407. doi:10.1093/nar/gku410

Nahmad, A. D., Reuveni, E., Goldschmidt, E., Tenne, T., Liberman, M., Horovitz-Fried, M., et al. (2022). Frequent aneuploidy in primary human T cells after CRISPR-Cas9 cleavage. *Nat. Biotechnol.* doi:10.1038/s41587-022-01377-0

Nelson, C. E., Wu, Y., Gemberling, M. P., Oliver, M. L., Waller, M. A., Bohning, J. D., et al. (2019). Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. *Nat. Med.* 25, 427–432. doi:10.1038/s41591-019-0344-3

Nihongaki, Y., Kawano, F., Nakajima, T., and Sato, M. (2015). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat. Biotechnol.* 33, 755–760. doi:10.1038/nbt.3245

Passerini, V., Ozeri-Galai, E., de Pagter, M. S., Donnelly, N., Schmalbrock, S., Kloosterman, W. P., et al. (2016). The presence of extra chromosomes leads to genomic instability. *Nat. Commun.* 7, 10754. doi:10.1038/ncomms10754

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., and Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* 31, 839–843. doi:10.1038/nbt.2673

Qasim, W., Zhan, H., Samarasinghe, S., Adams, S., Amrolia, P., Stafford, S., et al. (2017). Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* 9, eaaj2013. doi:10.1126/scitranslmed.aaj2013

Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389. doi:10.1016/j.cell.2013.08.021

Rees, H. A., Komor, A. C., Yeh, W.-H., Caetano-Lopes, J., Warman, M., Edge, A. S. B., et al. (2017). Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat. Commun.* 8, 15790. doi:10.1038/ncomms15790

Riesenberg, S., Chintalapati, M., Macak, D., Kanis, P., Maricic, T., and Pääbo, S. (2019). Simultaneous precise editing of multiple genes in human cells. *Nucleic Acids Res.* 47, e116. doi:10.1093/nar/gkz669

Riesenberg, S., and Maricic, T. (2018). Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells. *Nat. Commun.* 9, 2164. doi:10.1038/s41467-018-04609-7

Robert, F., Barbeau, M., Éthier, S., Dostie, J., and Pelletier, J. (2015). Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med.* 7, 93. doi:10.1186/s13073-015-0215-6

Sago, C. D., Lokugamage, M. P., Paunovska, K., Vanover, D. A., Monaco, C. M., Shah, N. N., et al. (2018). High-throughput *in vivo* screen of functional mRNA delivery identifies nanoparticles for endothelial cell gene editing. *Proc. Natl. Acad. Sci. U. S. A.* 115, E9944–E9952. doi:10.1073/pnas.1811276115

Samuelson, C., Radtke, S., Zhu, H., Llewellyn, M., Fields, E., Cook, S., et al. (2021). Multiplex CRISPR/Cas9 genome editing in hematopoietic stem cells for fetal hemoglobin reinduction generates chromosomal translocations. *Mol. Ther. Methods Clin. Dev.* 23, 507–523. doi:10.1016/j.omtm.2021.10.008

Sato, Y., Bando, H., Di Piazza, M., Gowing, G., Herberts, C., Jackman, S., et al. (2019). Tumorigenicity assessment of cell therapy products: The need for global consensus and points to consider. *Cytotherapy* 21, 1095–1111. doi:10.1016/j.jcyt.2019.10.001

Savas, S., Tuzmen, S., and Ozcelik, H. (2006). Human SNPs resulting in premature stop codons and protein truncation. *Hum. Genomics* 2, 274–286. doi:10.1186/1479-7364-2-5-274

Shapiro, J., Iancu, O., Jacobi, A. M., McNeill, M. S., Turk, R., Rettig, G. R., et al. (2020). Increasing CRISPR efficiency and measuring its specificity in hspcs using a clinically relevant system. *Mol. Ther. Methods Clin. Dev.* 17, 1097–1107. doi:10.1016/j.omtm.2020.04.027

Sharma, R., Dever, D. P., Lee, C. M., Azizi, A., Pan, Y., Camarena, J., et al. (2021). The TRACE-Seq method tracks recombination alleles and identifies clonal reconstitution dynamics of gene targeted human hematopoietic stem cells. *Nat. Commun.* 12, 472. doi:10.1038/s41467-020-20792-y

Shen, M. W., Arbab, M., Hsu, J. Y., Worstell, D., Culbertson, S. J., Krabbe, O., et al. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* 563, 646–651. doi:10.1038/s41586-018-0686-x

Singh, R., Kuscu, C., Quinlan, A., Qi, Y., and Adli, M. (2015). Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. *Nucleic Acids Res.* 43, e118. doi:10.1093/nar/gkv575

Skryabin, B. V., Kummerfeld, D.-M., Gubar, L., Seeger, B., Kaiser, H., Stegemann, A., et al. (2020). Pervasive head-to-tail insertions of DNA templates mask desired CRISPR-Cas9-mediated genome editing events. *Sci. Adv.* 6, eaax2941. doi:10.1126/sciadv.aax2941

Stadtmauer, E. A., Fraietta, J. A., Davis, M. M., Cohen, A. D., Weber, K. L., Lancaster, E., et al. (2020). CRISPR-engineered T cells in patients with refractory cancer. *Science* 367, eaba7365. doi:10.1126/science.aba7365

Starks, E. R., Swanson, L., Docking, T. R., Bosdet, I., Munro, S., Moore, R. A., and Karsan, A. (2021). Assessing limit of detection in clinical sequencing. *J. Mol. Diagn.* 23, 455–466. doi:10.1016/j.jmoldx.2020.12.010

Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J., and Mateo, J. L. (2015). Cctop: An intuitive, flexible and reliable crispr/cas9 target prediction tool. *PLoS ONE* 10, e0124633. doi:10.1371/journal.pone.0124633

Swarts, D. C., and Jinek, M. (2018). Cas9 versus Cas12a/Cpf1: Structure-function comparisons and implications for genome editing. *Wiley Interdiscip. Rev. RNA* 9, e1481. doi:10.1002/wrna.1481

Tombácz, I., Laczkó, D., Shahnawaz, H., Muramatsu, H., Natesan, A., Yadegari, A., et al. (2021). Highly efficient CD4+ T cell targeting and genetic recombination using engineered CD4+ cell-homing mRNA-LNPs. *Mol. Ther.* 29, 3293–3304. doi:10.1016/j.ymthe.2021.06.004

Truong, D.-J. J., Kühner, K., Kühn, R., Werfel, S., Engelhardt, S., Wurst, W., et al. (2015). Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res.* 43, 6450–6458. doi:10.1093/nar/gkv601

Tsai, S. Q., Nguyen, N. T., Malagon-Lopez, J., Topkar, V. V., Aryee, M. J., and Joung, J. K. (2017). CIRCLE-Seq: A highly sensitive *in vitro* screen for genome-wide CRISPR-cas9 nuclease off-targets. *Nat. Methods* 14, 607–614. doi:10.1038/nmeth.4278

Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187–197. doi:10.1038/nbt.3117

Tuladhar, R., Yeu, Y., Tyler Piazza, J., Tan, Z., Rene Clemenceau, J., Wu, X., et al. (2019). CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nat. Commun.* 10, 4056. doi:10.1038/s41467-019-12028-5

Turchiano, G., Andrieux, G., Klermund, J., Blattner, G., Pennucci, V., El Gaz, M., et al. (2021). Quantitative evaluation of chromosomal rearrangements in gene-edited human stem cells by CAST-Seq. *Cell Stem Cell* 28, 1136–1147.e5. e5. doi:10.1016/j.stem.2021.02.002

Turocy, J., Adashi, E. Y., and Egli, D. (2021). Heritable human genome editing: Research progress, ethical considerations, and hurdles to clinical practice. *Cell* 184, 1561–1574. doi:10.1016/j.cell.2021.02.036

Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., et al. (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* 24, 1216–1224. doi:10.1038/s41591-018-0137-0

Wen, W., Quan, Z.-J., Li, S.-A., Yang, Z.-X., Fu, Y.-W., Zhang, F., et al. (2021). Effective control of large deletions after double-strand breaks by homology-directed repair and dsODN insertion. *Genome Biol.* 22, 236. doi:10.1186/s13059-021-02462-4

White, S. J., Nicklin, S. A., Büning, H., Brosnan, M. J., Leike, K., Papadakis, E. D., et al. (2004). Targeted gene delivery to vascular tissue *in vivo* by tropism-modified adeno-associated virus vectors. *Circulation* 109, 513–519. doi:10.1161/01.CIR.0000109697.68832.5D

Wienert, B., Wyman, S. K., Richardson, C. D., Yeh, C. D., Akcakaya, P., Porritt, M. J., et al. (2019). Unbiased detection of CRISPR off-targets *in vivo* using DISCOVER-Seq. *Science* 364, 286–289. doi:10.1126/science.aav9023

Wienert, B., Wyman, S. K., Yeh, C. D., Conklin, B. R., and Corn, J. E. (2020). CRISPR off-target detection with DISCOVER-seq. *Nat. Protoc.* 15, 1775–1799. doi:10.1038/s41596-020-0309-5

Xiang, X., Corsi, G. I., Anthon, C., Qu, K., Pan, X., Liang, X., et al. (2021). Enhancing CRISPR-Cas9 gRNA efficiency prediction by data integration and deep learning. *Nat. Commun.* 12, 3238. doi:10.1038/s41467-021-23576-0

Xiao, A., Cheng, Z., Kong, L., Zhu, Z., Lin, S., Gao, G., et al. (2014). CasOT: A genome-wide cas9/gRNA off-target searching tool. *Bioinformatics* 30, 1180–1182. doi:10.1093/bioinformatics/btt764

Xue, C., and Greene, E. C. (2021). DNA repair pathway choices in CRISPR-cas9-mediated genome editing. *Trends Genet.* 37, 639–656. doi:10.1016/j.tig.2021.02.008

Yan, W. X., Mirzazadeh, R., Garnerone, S., Scott, D., Schneider, M. W., Kallas, T., et al. (2017). BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun.* 8, 15058. doi:10.1038/ncomms15058

Yeh, C. D., Richardson, C. D., and Corn, J. E. (2019). Advances in genome editing through control of DNA repair pathways. *Nat. Cell Biol.* 21, 1468–1478. doi:10.1038/s41556-019-0425-z

Yin, J., Liu, M., Liu, Y., Wu, J., Gan, T., Zhang, W., et al. (2019). Optimizing genome editing strategy by primer-extension-mediated sequencing. *Cell Discov.* 5, 18. doi:10.1038/s41421-019-0088-8

Yin, J., Lu, R., Xin, C., Wang, Y., Ling, X., Li, D., et al. (2022). Cas9 exo-endonuclease eliminates chromosomal translocations during genome editing. *Nat. Commun.* 13, 1204. doi:10.1038/s41467-022-28900-w

Zhang, S., Shen, J., Li, D., and Cheng, Y. (2021). Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics* 11, 614–648. doi:10.7150/thno.47007

Zhao, C., Zhao, Y., Zhang, J., Lu, J., Chen, L., Zhang, Y., et al. (2018). HIT-Cas9: A CRISPR/cas9 genome-editing device under tight and effective drug control. *Mol. Ther. Nucleic Acids* 13, 208–219. doi:10.1016/j.omtn.2018.08.022

Zheng, Z., Liebers, M., Zhelyazkova, B., Cao, Y., Panditi, D., Lynch, K. D., et al. (2014). Anchored multiplex PCR for targeted next-generation sequencing. *Nat. Med.* 20, 1479–1484. doi:10.1038/nm.3729

Zhu, L. J., Holmes, B. R., Aronin, N., and Brodsky, M. H. (2014). CRISPRseek: A bioconductor package to identify target-specific guide RNAs for CRISPR-cas9 genome-editing systems. *PLoS ONE* 9, e108424. doi:10.1371/journal.pone.0108424

Zhu, Y., Biernacka, A., Pardo, B., Dojer, N., Forey, R., Skrzypczak, M., et al. (2019). qDSB-Seq is a general method for genome-wide quantification of DNA double-strand breaks using sequencing. *Nat. Commun.* 10, 2313. doi:10.1038/s41467-019-10332-8

Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., et al. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364, 289–292. doi:10.1126/science.aav9973

frontiers | Frontiers in Genome Editing

TYPE Correction
PUBLISHED 20 January 2023
DOI 10.3389/fgeed.2022.1112956

OPEN ACCESS

EDITED AND REVIEWED BY
Ayal Hendel,
Bar-Ilan University, Israel

*CORRESPONDENCE
M. Kyle Cromer,
kyle.cromer@ucsf.edu

SPECIALTY SECTION
This article was submitted to Genome Editing in Blood Disorders, a section of the journal Frontiers in Genome Editing

RECEIVED 30 November 2022
ACCEPTED 02 December 2022
PUBLISHED 20 January 2023

CITATION
Wienert B and Cromer MK (2023), Corrigendum: CRISPR nuclease off-target activity and mitigation strategies. *Front. Genome Ed.* 4:1112956. doi: 10.3389/fgeed.2022.1112956

Corrigendum: CRISPR nuclease off-target activity and mitigation strategies

Beeke Wienert[1] and M. Kyle Cromer[2,3,4]*

[1]Graphite Bio, Inc., South San Francisco, CA, United States, [2]Department of Surgery, University of California, San Francisco, San Francisco, CA, United States, [3]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA, United States, [4]Eli and Edythe Broad Center for Regeneration Medicine, University of California, San Francisco, San Francisco, CA, United States

KEYWORDS

gene therapy, off-target activity, *in vivo* delivery, genome editing, CRISPR/Cas9, next-generating sequencing

A Corrigendum on
CRISPR nuclease off-target activity and mitigation strategies

by Wienert B and Cromer MK (2022). Front. Genome Ed. 4:1050507. doi: 10.3389/fgeed.2022.1050507

In the published article, "Dobbs et al., 2022" was cited incorrectly in the article. The citation has now been inserted at the correct location in the section, "**Methods to find off-targets sites in genome editing applications**", paragraph three, and should read: "These methods vary widely in their approach and even starting material, using cell-free genomic DNA *in vitro* (**Kim et al., 2015**; **Cameron et al., 2017**; **Tsai et al., 2017**; **Kim and Kim, 2018**; **Lazzarotto et al., 2020**), in intact live cells *ex vivo* (Crosetto et al., 2013; Tsai et al., 2015; Yan et al., 2017; Wienert et al., 2019; 2020; Zhu et al., 2019; Dobbs et al., 2022), and *in vivo* animal models (Akcakaya et al., 2018; Wienert et al., 2019; Liang et al., 2022)."

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

Publisher's note

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

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 10 November 2022
DOI 10.3389/fbioe.2022.1033669

OPEN ACCESS

EDITED BY
Ayal Hendel,
Bar-Ilan University, Israel

REVIEWED BY
Elena I. Leonova,
St. Petersburg State University, Russia
Adi Barzel,
Tel Aviv University, Israel

*CORRESPONDENCE
Alexandre Juillerat,
alexandre.juillerat@cellectis.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Preclinical Cell and Gene Therapy, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 31 August 2022
ACCEPTED 30 September 2022
PUBLISHED 10 November 2022

CITATION
Boyne A, Yang M, Pulicani S, Feola M, Tkach D, Hong R, Duclert A, Duchateau P and Juillerat A (2022), Efficient multitool/multiplex gene engineering with TALE-BE. *Front. Bioeng. Biotechnol.* 10:1033669. doi: 10.3389/fbioe.2022.1033669

Efficient multitool/multiplex gene engineering with TALE-BE

Alex Boyne[1†], Ming Yang[1†], Sylvain Pulicani[2†], Maria Feola[1], Diane Tkach[1], Robert Hong[1], Aymeric Duclert[2], Philippe Duchateau[2] and Alexandre Juillerat[1*]

[1]Cellectis Inc., New York, NY, United States, [2]Cellectis, Paris, France

TALE base editors are a recent addition to the genome editing toolbox. These molecular tools are fusions of a transcription activator-like effector domain (TALE), split-DddA deaminase halves, and an uracil glycosylase inhibitor (UGI) that have the distinct ability to directly edit double strand DNA, converting a cytosine (C) to a thymine (T). To dissect the editing rules of TALE-BE, we combined the screening of dozens of TALE-BE targeting nuclear genomic loci with a medium/high throughput strategy based on precise knock-in of TALE-BE target site collections into the cell genome. This latter approach allowed us to gain in depth insight of the editing rules *in cellulo*, while excluding confounding factors such as epigenetic and microenvironmental differences among different genomic loci. Using the knowledge gained, we designed TALE-BE targeting CD52 and achieved very high frequency of gene knock-out (up to 80% of phenotypic CD52 knock out). We further demonstrated that TALE-BE generate only insignificant levels of Indels and byproducts. Finally, we combined two molecular tools, a TALE-BE and a TALEN, for multiplex genome engineering, generating high levels of double gene knock-out (~75%) without creation of translocations between the two targeted sites.

KEYWORDS

gene editing, base editors, TALE, t-cells, cell engineering

Introduction

Base editing, one of the most recent advances in the field of genome editing is a technology that allows the introduction of point mutations (transitions C>T or A>G and cytosine transversion C>G) in defined loci of a targeted DNA sequence (Koblan et al., 2018; Anzalone et al., 2020; Kurt et al., 2021; Zhao et al., 2021). Base editors create mutations by deamination at the targeted bases (C or A), which are subsequently converted into T or G during the DNA repair process. Such a process does not create DNA double strand breaks as does CRISPR/Cas9, or other engineered nucleases, and is a promising therapeutic strategy for genetic diseases.

In contrast to the Cas9 and Cpf1 base editing platforms that operate predominantly on ssDNA, a newly discovered bacterial deaminase (DddA) catalyzes the deamination of cytidine within double strand DNA molecules and

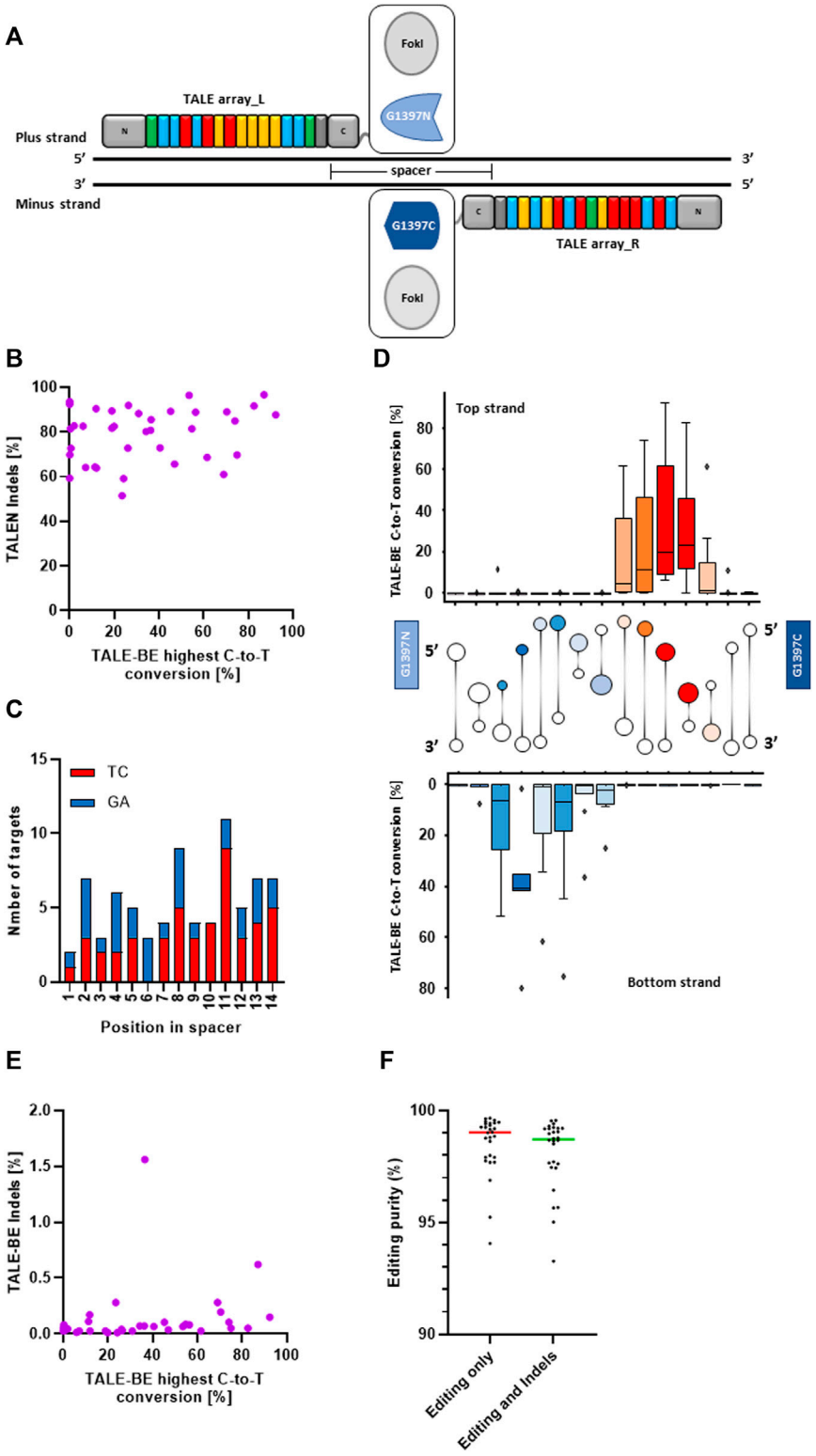


FIGURE 1
(A). Schematic representation of a TALEN and a TALE-BE. **(B).** Average of the highest C-to-T conversion frequencies of 37 base editors (among all edited bases within the target spacer region) versus the frequencies of indels created by the matching TALEN (*N* = 2, independent T-cells donors). **(C).** Repartition of TC/GA in the spacer of the 37 TALE-BE. **(D).** C-to-T conversion frequencies of the target C (top) or G (bottom) at different positions within the 15 bp TALE-BE spacer (box represent 5%–95% percentile, whiskers represent min and max) with cartoon of DNA double helix drawn

(*Continued*)

FIGURE 1 (Continued)
between top and bottom strands to aid visualization of edited positions within the spacer. Colors of cartoon are selected to match those of the box plot, with darker colors indicating more editing. **(E).** Average of the highest C-to-T conversion frequencies of 37 base editors versus indels frequencies generated within the spacer. **(F).** Editing purity (median) within the cell population. Left: C-to-T conversion within the C-to-A/G/T population. Right: C-to-T conversion within the C-to-A/G/T + Indels population. For all panels: *N* = 2, independent T-cells donors.

allows for the development of designer base editors with alternative DNA targeting platforms (Mok et al., 2020). These TALE-BE, and very recently Zinc finger BE (Lim et al., 2022), were used for several applications including the creation of mutations in nuclear DNA (Mok et al., 2020), and, unlike Cas9 and Cpf1 base editors, also mitochondrial DNA (Lee et al., 2021; Qi et al., 2021; Sabharwal et al., 2021; Guo et al., 2022; Liu et al., 2022; Wei et al., 2022), or chloroplast (Kang et al., 2021; Nakazato et al., 2021), generating inheritable modifications and rendering the TALE-BE the first functional base editing tool available for these latter cellular compartments. Indeed, these new TALE-BE expand the base editing toolbox, providing additional ways to target specific sites for correction. For instance TALE-BE, like other TALE-derived editors, should bring with them enhanced ability to access hard to edit loci (Jain et al., 2021). And by utilizing different rules for targeting, and interacting with, the genome, these editors will open up additional sites outside of the scope of previously described Cas9 and Cpf1 base editing platforms. However, despite these successful applications, more detailed and comprehensive study are necessary to fully promote the development of these new molecular tools. In particular, while it also requires high modification frequencies to ensure compatibility with product development, multiplex gene editing using base editors represents a promising strategy to avoid unintended translocations between edited loci.

Here, we studied the determinants to achieve high editing frequencies using TALE-BE. We combined the screening of several TALE-BE targeting various endogenous loci with the development of a medium/high throughput cell-based assay that would avoid biases due to confounding effects such as epigenomic factors or modifications (Valton et al., 2012; Bose et al., 2021). The accumulated knowledge enabled the definition of TALE-BE design guidelines that were further applied to nuclear base editing, allowing for very efficient knock-out of CD52 in primary T-cells (up to 87% phenotypically and 86% editing at the genomic level). We further demonstrated the possibility of combining a TALEN (TALE nuclease) and a TALE-BE (TALE Base editor) to perform a double gene KO of *TRAC* and *CD52* (75% double negative cell population), a combination of target genes used for allogeneic CAR T-cell adoptive therapies. Such combinations of molecular tools open the way to simultaneous multiplex gene engineering with more controllable outcomes.

Results

Comparison of TALEN and TALE-BE efficiencies on nuclear genomic loci

To consolidate our understanding of the determinants for efficient editing by TALE-BE, we first identified a set of 37 TALEN (TALE nucleases, Figure 1A), originating from a previously described backbone (Valton et al., 2012; Juillerat et al., 2014; Gautron et al., 2017), that showed high activity in primary T-cells (median indels = 82% and s.d. = 12) (Figure 1A, Supplementary Figure S1A). These 37 target sequences (Supplementary Table S1) were also carefully selected to target regions with different chromatin states in T cells as found by ChromHMM (Ernst and Kellis 2017, Supplementary Table S1). In addition, the spacer sequence, the DNA sequence between the two TALE binding sites, was also kept constant to 15 bp as it was previously shown to accommodate both TALEN and TALE-BE (Juillerat et al., 2014; Mok et al., 2020). Finally, the sequence of the spacers contained various numbers of, but homogeneously distributed, Cs, Gs, TCs or GAs as previous studies demonstrated a strong editing preference for 5′-TC contexts (Figure 1C; Supplemntary Figures S1B,C) (Mok et al., 2020). We then produced the corresponding 37 TALE-BE by replacing the FokI catalytic domain with the DddAtox split and an uracil glycosylase inhibitor (UGI). We focused on the so-called G1397 split, in which the deaminase is split at the C-terminus of G1397 residue, as this fusion previously showed improved editing activity (Mok et al., 2020) (Figure 1A). We then compared, in primary T-cells, the maximum editing within the spacer for a given TALE-BE to the Indel frequencies created by the corresponding TALEN counterpart (Figure 1B). The lack of correlation (Spearman correlation = 0.16, *p*-value = 0.33) between the two data sets (TALEN vs. TALE-BE editing frequencies) confirms that the key determinant for efficient editing is related to the positioning of the target cytosine within the spacer. Indeed, analysis of editing efficiency in function of the position within the spacer highlighted a defined 4–5 bp editing window on both, top and bottom strands, occupying approximately the same three dimensional space when visualized along the DNA double helix (Mok et al., 2022) (Figure 1D).

Interestingly, low frequencies of Indels (small insertion and deletions, <0.5%) were observed for 35 out of 37 base editors (Indel frequencies: median = 0.06% and s.d. = 0.17, Figure 1E). The Indels observed at the target site moderately correlated with

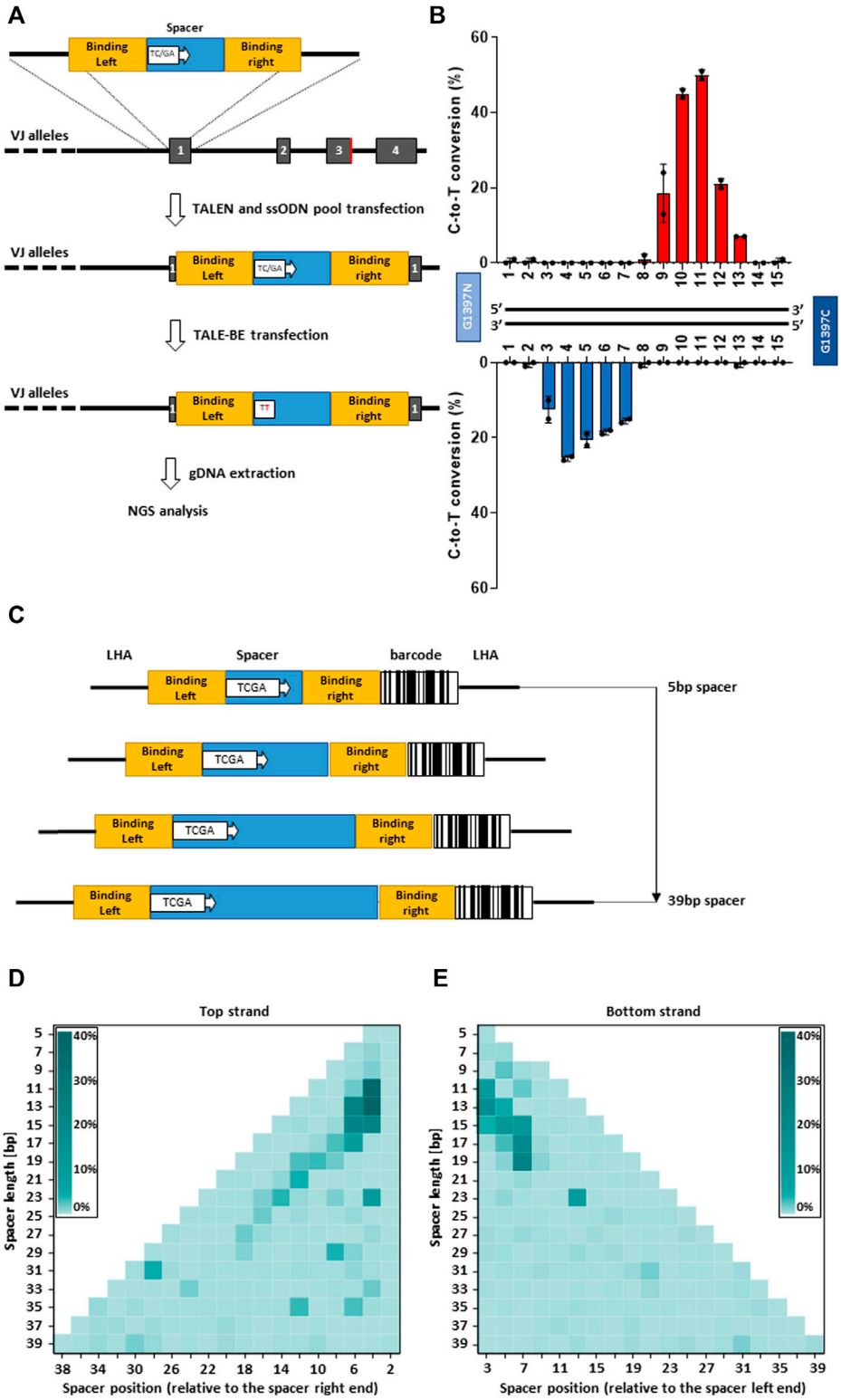


FIGURE 2
(A). Scheme of the strategy to generate artificial base editor target sites. In a first step a pool of ssODN encoding various base editor spacer sequences is inserted into TRAC locus. In a second step the TALE-BE is transfected. Two days post transfection the genomic DNA is collected, and the inserted sequence is analyzed by NGS. **(B)**. Mean C-to-T conversion frequencies of the target Cs (top) or Gs (bottom) at different positions within the 15 bp TALE-BE spacer. **(C)**. Schematic representation of the ssODN pool collection with spacer length ranging from 5 to 39 bp. **(D)**. heatmap of C-to-T conversion when the TC was present on the top strand in function of the spacer length. **(E)**. heatmap of C-to-T conversion when the TC was present on the bottom strand in function of the spacer length. For all panels: *N* = 2, independent T-cells donors.

editing frequency within the spacers (Spearman correlation = 0.44, p-value = 0.007)) (Figure 1E). In addition, we measured low byproduct (C-to-A/G) editing within the editing window, overall indicative of a very high final purity of the edited cell populations (C-to-T within C-to-A/G/T: median = 99.6% and s.d. = 0.9; C-to-T within C-to-A/G/T + Indels: median = 99.5% and s.d. = 1.1, Figure 1F; Supplementary Figure S1D).

Definition of the optimal base editing window

To comprehensively investigate DddA-derived cytosine base editors (TALE-BE), we designed an experimental setup allowing us to screen, in a defined genomic context, for base editing efficiency in a medium to high throughput format. We, and others, have previously demonstrated the possibility to precisely insert *via* homologous recombination, a short (50–200 bp) sequence, into the genome of primary T cells using a single strand oligonucleotide (ssODN) as a template (Roth et al., 2018; Yang et al., 2021). We decided to generate a pool of cells, containing predefined TALE-BE target sequences precisely inserted into a chosen genomic locus in the *TRAC* gene (Yang et al., 2021). Each of the TALE-BE targets contained a unique TC or GA (target for the DddA deaminase) within the spacer sequence flanked by two fixed TALE binding sequences (RVD-L and RVD-R, Figure 2A). This setup allows the uniform TALE-BE binding to the artificial target sites, excluding editing variability caused by 1) different DNA binding affinities from different TALE array protein and 2) the impact of epigenomic factors, such as chromosome relaxation around the artificial BE target sites. We first designed a collection of ssODN that contain two fixed TALE array protein binding sites from one of the most active TALE-BE identified previously on endogenous genomic loci (T-25 TALE-BE, Supplementary Table S1). The two TALE binding sequences were separated by various 15 bp spacer sequences (similar length to the previous collection of TALE-BE targeting endogenous loci). To generate the pool of cells harboring the collection of BE targets, the 30 ssODN (15 TC and 15 GA, Supplementary Table S2) oligonucleotides were mixed in equal amounts and transfected in primary T-cells by electroporation simultaneously with a TALEN targeting *TRAC* (Valton et al., 2015). In the second step, 2 days post transfection of the ssODN pool, the mRNAs encoding the T-25 TALE-BE were vectorized by electroporation. The genomic DNA of transfected cells was then harvested at day 2 post TALE-BE transfection for editing analysis (Figure 2A). The NGS analysis showed that the ssODNs were efficiently and homogenously integrated at the *TRAC* locus (read number: median = 1667.5, mean = 1686.2, s.d. = 351.7, Supplementary Figure S2A). The control sample treated without TALE-BE showed low frequencies of background mutations, whereas the samples treated with TALE-BE showed detectable and reproducible levels of C-to-T conversion (Supplementary Figures S2B,C).

The analysis further highlighted editing windows comparable to those observed with the 37 TALE-BE targeting endogenous sequences (Figure 2B, Figure 1D). Indeed, when comparing the relative activity of the two data sets, we found strong overlap for the primary site of editing at the TC of the top strand of the editing window (Pearson corr. coeff. = 0.82, Figure S2D). The overlap was less robust, but still apparent, for the secondary site of editing at the TC of the bottom strand (Pearson corr. coeff. = 0.69). Altogether, the similarity of the two datasets further validate this artificial pooled spacer approach and that the positional rules for editing are not locus dependent.

As the artificial pooled spacer approach demonstrated the possibility to dissect TALE-BE editing profiles, we expanded our ssODN collection to spacers with odd number lengths, spanning from 5 to 39 bp (i.e. 5, 7, 9, 11 … 37, 39 bp). To reduce the number of ssODN needed for the collection, a TCGA quadruplex target sequence was incorporated in the spacer at every other position (Figure 2C). This design, containing 191 unique ssODNs (Supplementary Table S3), allowed us to simultaneously interrogate editing efficiencies on both strands with a single ssODN. Additionally, to facilitate the sequence analysis, a unique barcode was added to each construct (Figure 2C). Upon filtering the NGS data to remove the reads in which the barcode conflicted with the spacer sequence, we obtained a high and homogenous representation of each ssODN (read number: median = 545, mean = 3522.6, s.d. = 7122.5, Supplementary Figure S3A). As with the previous collection (15 bp spacer), low frequencies of mutations were observed without the TALE-BE, while C-to-T conversion was robustly measured with the TALE-BE, either on the plus or minus strand (Supplementary Figure S3B). Analysis of the data indicated that, for our TALE-BE scaffold, a spacer length ranging from 11 to 17 bp to achieve optimal editing, with a 4–7 bp editing windows on the different spacers (Figure 2D, Figure 2E).

Efficient generation of CD52 gene knock out with TALE-BE and multiplexed cell engineering

In the context of allogeneic CAR-T therapies, *CD52* can be knocked out, in combination with *TRAC*, *via* TALEN-gene editing to create resistance to alemtuzumab, a CD52-targeting monoclonal antibody used in lymphodepleting regimens (Poirot et al., 2015; Qasim et al., 2017). In order to improve the outcome of simultaneous multiplex gene knock-out, and prevent possible translocations arising with the contemporaneous use of multiple nucleases (Poirot et al., 2015), we sought to develop two base editing approaches, targeting either a splice site or the signal sequence, to promote efficient CD52 knock-outs.

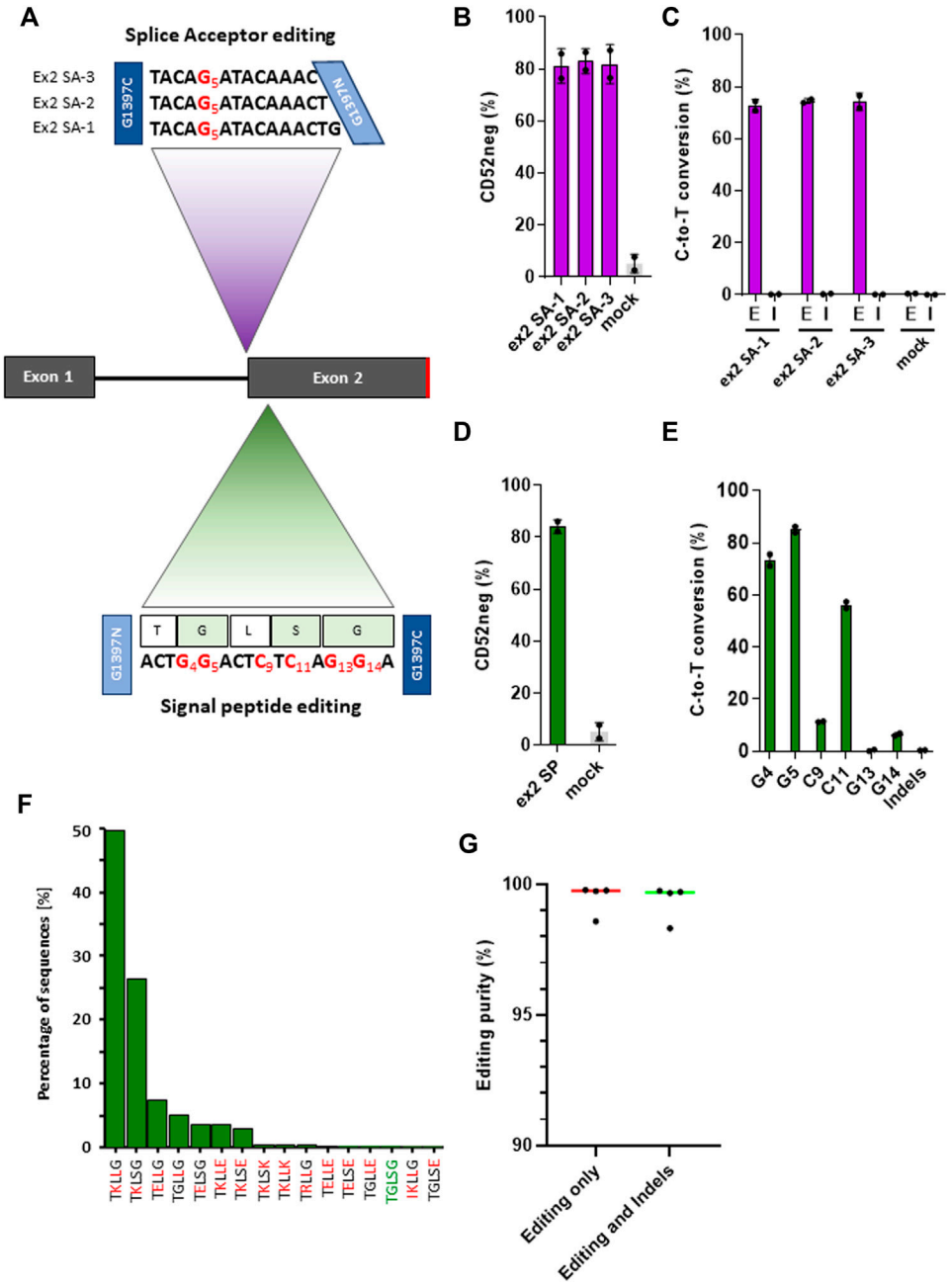


FIGURE 3
(A). Schematic representation of target spacer sequence for TALE-BE targeting CD52. Top: TALE-BE (ex2 SA-1, two and 3) designed to edit CD52 exon2 splice acceptor site. The conserved G of the splice site, targeted by the TALE-BE) is depicted in red. Bottom: TALE-BE targeting CD52 signal peptide sequence in exon 2 (SP). TALE-BE spacer sequence (with targetable Cs or Gs numbered) and peptide sequence are depicted. **(B)**. CD52 negative cell frequency in the TALE-BE treated (targeting CD52 exon2 splice acceptor site) or mock electroporated PBMC populations, 6 days post electroporation, measured by flow cytometry. **(C)**. Editing **(E)** frequencies (C-to-T conversion) of the conserved G of the exon two acceptor splice site (Editing) and the indel **(I)** frequencies within the target locus, measured by NGS 6 days post transfection. **(D)**. CD52 negative cell frequency in the TALE-BE treated (targeting CD52 signal sequence in exon2) or mock electroporated PBMC populations, 6 days post electroporation, measured by flow cytometry. **(E)**. Editing frequencies (C-to-T conversion) at different position within the TALE-BE target spacer (CD52 signal sequence in exon2) and indel frequencies within the target locus, measured by NGS at Day 6 post transfection. **(F)**. Frequencies of peptide species created by the TALE-BE targeting the CD52 signal sequence in exon 2. The first 16 most abundant species are presented. Mutation relative to the native signal peptide are in red. (*N* = 2, independent T-cells donors). **(G).** Editing purity within the cell population. Aggregate of the four TALE-BE targeting CD52. Left: C-to-T conversion within the C-to-A/G/T population. Right: C-to-T conversion within the C-to-A/G/T + Indels population. For all panels: *N* = 2, independent T-cells donors.

Because the CD52 gene only has two exons, and the exon two contains the sequence coding for the mature peptide, splice site mutations at the intron 1/exon 2 junction conserved motif would presumably disrupt RNA splicing, causing the retention of intron 1 and ultimately leading to the loss of CD52 (Figure 3A). Taking into consideration the sequences surrounding the target splice site and some constraints in the TALE-BE design: T at the extremities of the target, 16 bp TALE binding sequences, 11–17 bp spacer length, we were able to identify 34 potential base editors targeting the conserved G within the splice site. We then narrowed down the BE candidate list, according to the TC positioning data we gathered from the artificial pooled spacer experiments, to three lead BEs (Figure 3A). Primary T cells were transfected with mRNA encoding these three pairs of TALE-BE. Seven days post transfection, phenotypic CD52 knock-out was monitored by flow cytometry and splice site editing was measured by NGS. We observed high level of phenotypic knock-out for the three TALE-BE (Figure 3B; Supplementary Figure S4, ex2 SA-1 mean 81.1% ± 4.7%, ex2 SA-2 mean 83% ± 3.4% and ex2 SA-3 mean 81.9%, ± 5.3%), correlating with editing levels (ex2 SA-1 mean 72.6%, ± 1.7%, ex2 SA-2 mean 74.5%, ± 0.6%. and ex2 SA-3 mean 74.2%, ± 2.3%, Figure 3C). As expected from our previous datasets, NGS data analysis results showed very low levels of Indels at these sites (ex2 SA-1 mean 0.16%, ± 0.05%; ex2 SA-2 mean 0.28%, ± 0.06%; ex2 SA-3 mean 0.12%, ± 0.02%, Mock transfected mean 0.01%, ± 0.005%; Figure 3C).

Mutations in signal peptides has been shown to disrupt the processing and the translocation of nascent peptides and thus impair the surface expression of certain genes (Wiren et al., 1989). We thus evaluated a second CD52 KO approach by designing a TALE-BE targeting CD52 signal sequence. This approach could potentially lead to 1) a silent mutation at Leu23 residue and 2) several amino acid changes (Gly22Lys, Ser24Leu and Gly25Lys) (Figure 3A). We anticipated that such changes in the residues, mutating a hydrophobic glycine to a highly charged lysine, and a polar serine to a hydrophobic leucine in the signal peptide, would significantly impact the ability for the signal peptide to correctly direct translocation. Indeed, 6 days post TALE-BE mRNA transfection (ex2 SP), we observed by flow cytometry an average of 84.2% (± 1.8%) CD52 negative cells (Figure 3D; Supplementary Figure S4). The NGS sequencing analysis revealed that all six positions were mutated, albeit at different levels (mean editing frequencies: G [4]: 73.65 ± 1%, G [5]: 85.65 ± 0.7%, C [9]: 11.4 ± 0.1% C [11]: 56.5 ± 0.9%, G [13]: 0.6 ± 0.1, G [14]:6.5 ± 0.5%) (Figure 3E). Altogether, we identified editing leading to potentially 34 different species (at the protein level), including the WT, present in different proportions (Figure 3F; Supplementary Table S4).

To evaluate possible off-target editing of the four CD52 TALE-BEs (targeting the splice site and the signal sequence), we generated *in silico* a list of potential off-site targets for these BEs, covering altogether 307 unique sites (207 experimentally investigated, Supplementary Table S5). Targeted amplicon sequencing using a multiplexed amplicon sequencing assay (Chaudhari et al., 2020) was performed and the analysis did not demonstrate evidence of editing (95 sites for ex SA-1, 72 sites for ex2 SA-2, 89 sites for ex2 SA-3 and 35 sites for ex2 SP, Supplementary Table S6, 1-2 independent T-cells donors).

Having achieved very high phenotypic knock-out (median CD52 negative population: 82.1%) and editing purity (median = 99.7 and s.d. = 0.6, Figure 3G), we further evaluated the possibility to perform multiplex gene editing using two different molecular tools, a nuclease and a base editor. As a proof of concept, we combined a TALEN targeting *TRAC* and either a TALEN or a base editor (TALE-BE ex2 SP) targeting CD52, a combination of edits reported for the generation of allogeneic CAR T-cells (Poirot et al., 2015). We detected high and similar levels of phenotypic double gene knock-out by flow cytometry in both TALEN/TALEN and TALEN/TALE-BE treated samples (79% and 75% respectively, Supplementary Figure S5). However, translocations between the two targeted loci, detected by multiplexed amplicon sequencing, (Amit et al., 2021), were only observed in the TALEN/TALEN treated sample (479 reads out of 224,406 for the TALEN/TALEN sample; 0 reads out of 144,323 for the TALEN/BE sample, N = 1, 1 single T-cell donor). Overall, we believe that such a multi-tool approach for multiplex gene editing has the potential to streamline development of products by easing QCs (absence of translocation) and more globally improve safety of multiplex cell engineering.

Discussion

Base editing represents one of the newest gene editing technologies. Recently, the TALE scaffold was demonstrated to be compatible with the creation of a new class of DddA-derived cytosine base editors. In this study, we combined the screening of several base editors targeting various endogenous loci with the development of a simple and robust medium-throughput approach to interrogate the determinants of editing by TALE-BE. This medium-throughput screening strategy is taking advantage of the highly efficient and precise TALEN mediated ssODN knock-in in primary T cells and allowed to assess the TALE-BE editing efficiency on hundreds of different targets *in cellulo*. Because all BE artificial target sequences are inserted into the same predefined locus in the genome, this method allowed us to focus on how target/spacer sequence variations could affect TALE-BE while excluding factors such as DNA binding affinities or epigenetic variations. The experimental results pointed out an optimal 13–17 bp spacer length window for editing, with the G1397C-bearing arm of the TALE-BE being placed 4–7 bp down the 3' direction of the target TC for the best editing activity.

Although CRISPR/Cas cytosine base editors do not introduce the intended mutations through double strand break (DSB) repair by the non-homologous end joining (NHEJ) pathway, significant levels of unwanted indels creation has often been reported (Thuronyi et al., 2019; Doman et al., 2020). Since the first report of such designer base editors, improvements of editing efficiencies were often obtained at the cost of editing purity (Indels, byproduct mutation or bystander mutations) or *vice versa* (Komor et al., 2017; Thuronyi et al., 2019; Doman et al., 2020; Tran et al., 2022). Often this is achieved by the restoring of partial function to the dead Cas9 employed to target the base editing domain to the site of interest. By restoring one of two catalytic sites (adding back the catalytic His at position 840 while retaining the inactivating Asp10Ala), researchers were able to increase editing by nicking the non-edited strand (Komor et al., 2016). While extremely precise introduction of the intended mutation (high purity of the final product) is a prerequisite for application such as gene correction, bystander and byproduct edits might be of a lesser concern for gene disruption applications. However, generation of DSBs by base editors may raise greater concerns as CRISPR/Cas base nucleases have been recently associated with major on-target genome instability or chromosomal abnormalities (Weisheit et al., 2020; Alanis-Lobato et al., 2021; Boutin et al., 2021; Papathanasiou et al., 2021; Boutin et al., 2022; Geng et al., 2022; Nahmad et al., 2022; Sánchez-Rivera et al., 2022). In this study we only found marginal byproduct mutation (C-to-A/G), and more importantly low Indel creation, by TALE-BE looking at dozens of these molecular tools, even at high editing frequencies (>80% in bulk population), occurring at levels similar to, or lower than, those reported by other groups (Mok et al., 2020; Lim et al., 2022). One possibility for these low levels of Indels might be the result of a lack of reliance on opposite strand nicking for improved editing by TALE-BE. By avoiding the need for a nick on one strand opposite an edited abasic site on the other, these editors may simply do less potentially DSB-causing damage at the target site. Additional studies will need to be carried out to further elucidate whether these are the resulting byproducts of an imperfect DNA repair pathway, or perhaps the results of rare double strand breaks occurring during the replication process. Furthermore, a careful design of the BE positioning, would allow to prevent or minimize bystander mutations.

Base editors have been used to edit or mutate conserved genetic elements such as enhancers (Zeng et al., 2020), start codons (Wang et al., 2020), splice sites (Kluesner et al., 2021), branch points (Yuan et al., 2018), and conserved active sites (Hanna et al., 2021). We estimated that ~46,000 (46,608) splice sites in the genome could potentially be targeted by TALE-BE, impacting 15,279 different transcripts, representing 76.57% of all the transcripts in human genome (Frankish et al., 2021) and, overall, indicating that splice site editing could be a viable approach for gene knock-out by TALE-BE. To demonstrate the feasibility of such an approach, we designed highly efficient TALE-BE targeting the conserved G of the intron 1/exon 2 junction splice site of the CD52 gene. We further demonstrated that as an alternative to splice site editing, targeting the signal peptide can also lead to efficient surface protein knock-out.

Base editors represent promising molecular tools for multiplex gene engineering, though currently limited to knock-out or gene corrections. Here we demonstrated the feasibility of efficient multiplex gene engineering using a combination of two different molecular tools, a nuclease, and a base editor. Such a multiplex/multitool strategy presents several advantages. First, it prevents creation of translocations often observed with the simultaneous use of several (>1) nucleases (Poirot et al., 2015; Webber et al., 2019; Samuelson et al., 2021). And second, it allows the possibility to go beyond multiple knock-outs while still allowing gene knock-in at the nuclease target site, altogether extending the scope of possible application, while better controlling the engineered cell population outcome (e.g., absence of translocations). The precise positional rules we have determined for TALE-BE, low frequency of unwanted indels generated, and the increased accessibility to additional cell compartments allowing access beyond the traditional nuclear targets, expands the potential scope of such a TALE-based multiplex/multitool strategy beyond the capabilities of most other non-TALE editing tools.

Materials and methods

T cell culture

Cryopreserved human PBMCs were acquired from ALLCELLS. PBMCs were cultured in X-vivo-15 media (Lonza Group), containing 20 ng/ml human IL-2 (Miltenyi Biotec), and 5% human serum AB (Seralab). Human T cell activator TransAct (Miltenyi Biotec) was used to activate T cells at 25 µL TransAct per million $CD3^+$ cells the day after thawing the PBMCs. TransAct was kept in the culture media for 72 h.

Small scale mRNA production

Plasmid of the 37 TALE-BE and 37 matching TALEN, containing a T7 promoter and a polyA sequence, were produced as non-clonal after assembly (transformant was directly inoculated for culture and plasmid preparation). The plasmids were then linearized with SapI (NEB) and mRNA was produced by *in vitro* transcription (NEB HiScribe ARCA, NEB).

Small scall TALEN and TALE-BE testing (37 endogenous targets and TRAC/ CD52 multiplex engineering)

T cells activated with TransAct (Miltenyi Biotec) for 3 days were passaged into fresh complete media containing 20 ng/ml human IL-2 (Miltenyi Biotec), and 5% human serum AB (Seralab) 10–12 h s before transfection.

The harvested cells were washed once with warm PBS. 1E6 PBS washed cells were pelleted and resuspended in 20 µL Lonza P3 primary cell buffer (Lonza). 1 µg/arm/million cells of mRNA for TALEN or TALE-BE was mixed with the cells and then the cell mixture was electroporated using the Lonza 4D-Nucleofector under the EO115 program for stimulated human T cells. After electroporation, 80 µL warm complete media was added to the cuvette to dilute the electroporation buffer, the mixture was then carefully transferred to 400 ml pre-warmed complete media in 48-well plates. TALEN transfected cells were incubated at 30°C for an overnight culture and then transferred back to 37°C incubator. TALE-BE transfected cells were incubated at 37°C throughout the process. Cells were harvested at Day 6 post transfection for gDNA extraction and NGS analysis.

Large scale TALEN and TALE-BE mRNA production (CD52 targeting BEs)

Plasmids encoding the TRAC TALEN contained a T7 promoter and a polyA sequence. The TALEN mRNA from the TRAC TALEN plasmid was produced by Trilink. Sequence targeted by the TRAC TALEN (17-bp recognition sites, upper case letters, separated by a 15-bp spacer): (TTC CTCCTACTCACCATcagcctcctggttatGGTACAGGTAAGA GCAA).

The TALEN mRNA from the CD52 TALEN plasmid was produced by Trilink. Sequence targeted by the CD52 TALEN (17-bp recognition sites, upper case letters, separated by a 15-bp spacer): (TTCCTCCTACTCACCATcagcctcctggttatGGTACAG GTAAGAGCAACGCCTGGCA).

Plasmids encoding TALE-BE T-25 and CD52 TALE-BE contained a T7 promoter and a polyA sequence (TALE-BE Sequence and target sequence in Supplementary info). Sequence verified plasmids were linearized with SapI (NEB) befor *in vitro* mRNA synthesis. mRNA was produced with NEB HiScribe™ T7 Quick High Yield RNA Synthesis Kit (NEB). The 5′capping reaction was performed with ScriptCap™ m7G Capping System (Cellscript). Antarctic Phosphatase (NEB) was used to treat the capped mRNA and the final cleanups was performed with Mag-Bind TotalPure NGS beads (Omega bio-tek) and Invitrogen DynaMag-2 Magnet (ThermoFisher).

ssODN repair template transfection

The ssODN pool targeting the TRAC locus (Supplementary Table S2, S3) were ordered from Integrated DNA Technologies (IDT) and resuspended in ddH2O at 50 pmol/µl.

T cells activated with TransACT for 3 days were passaged into fresh complete media containing 20 ng/ml human IL-2 (Miltenyi Biotec), and 5% human serum AB (Seralab) 10–12 h s before transfection.

The harvested cells were washed once with warm PBS. 1E6 PBS washed cells were pelleted and resuspended in 20 µl Lonza P3 primary cell buffer (Lonza). 200 pmol ssODN pool and 1µg/arm of TRAC TALEN were mixed with the cell and then the cell mixture was electroporated using the Lonza 4D-Nucleofector under the EO115 program for stimulated human T cells. After electroporation, 80 µl warm complete media was added to the cuvette to dilute the electroporation buffer, the mixture was then carefully transferred to 400 ml pre-warmed complete media in 48-well plates. Cells transfected with ssODN and TALEN were then incubated at 30°C until 24 h s post TALEN transfection before transfer back to 37°C.

Cells with ssODN KI were cultured for 2 days before harvesting for TALE-BE treatment. The harvested cells were washed once with warm PBS. 1E6 PBS washed cells were pelleted and resuspended in 20 µL Lonza P3 primary cell buffer (Lonza). 1µg/arm of TALE-BE T-25 were mixed with the cell and then the cell mixture was electroporated using the Lonza 4D-Nucleofector under the EO115 program for stimulated human T cells. After electroporation, 80 µL warm complete media was added to the cuvette to dilute the electroporation buffer, the mixture was then carefully transferred to 400 ml pre-warmed complete media in 48-well plates. Cells transfected with TALE-BE incubated at 37°C for two more days before harvesting for gDNA extraction and NGS analysis.

Large scale CD52 TALE-BE testing

T cells activated with TransACT for 3 days were passaged into fresh complete media containing 20 ng/ml human IL-2 (Miltenyi Biotec), and 5% human serum AB (Seralab) 10–12 h s before transfection.

The harvested cells were washed twice with Cytoporation Media T (BTXpress, 47-0002). 5E6 washed cells were pelleted and resuspended in 180 µL Cytoporation Media T. 2µg/arm/ million cells of TALE-BE mRNA was mixed with the cells to a final volume of 200 µL and then the cell/mRNA mixture was electroporated using the BTX Pulse Agile in 0.4 cm gap cuvettes. After electroporation, 180 µL warm complete media was added to the cuvette to dilute the electroporation buffer, and the mixture was then carefully transferred to 2 ml pre-warmed complete media in 12-well plates. TALE-BE transfected cells were

incubated at 37°C throughout the process. Cells were harvested at Day 6 post transfection for gDNA extraction and NGS analysis.

Genomic DNA extraction

Cells were harvested and washed once with PBS. Genomic DNA extraction was performed using Mag-Bind Blood & Tissue DNA HDQ kits (Omega Bio-Tek) following the manufacturer's instructions.

Targeted PCR and NGS

100 µg genomic DNA was used per reaction in a 50 µL reaction with Phusion High-Fidelity PCR Master Mix (NEB). The PCR condition was set to 1 cycle of 30 s at 98°C; 30 cycles of 10 s at 98°C, 30s at 60°C, 30 s at 72°C; 1 cycle of 5 min at 72°C; hold at 4°C. The PCR product was then purified with Omega NGS beads (1:1.2 ratio) and eluted into 30 µL of 10 mM Tris buffer pH7.4. The second PCR which incorporates NGS indices was then performed on the purified product from the first PCR. 15 ul of the first PCR product were set in a 50 µL reaction with Phusion High-Fidelity PCR Master Mix (NEB). The PCR condition was set to 1 cycle of 30 s at 98°C; 8 cycles of 10 s at 98°C, 30 s at 62°C, 30s at 72°C; 1 cycle of 5 min at 72°C; hold at 4°C. Purified PCR products were sequenced on MiSeq (Illumina) on a 2 × 250 nano V2 cartridge.

Flow cytometry

TRAC KO was monitored using an anti-TCRa/b antibody (Biolegend, #306732, clone IP26, BV605). CD52 KO was monitored using an anti-52 antibody (BD Biosciences, #563609, Clone 4C8, AlexaFlour488). Flow cytometry was performed on BD FACSCanto (BD Biosciences) and data analysis processed with FlowJo. Cell population was first gated for lymphocytes (SSC-A vs. FSC-A) and singlets (FSC-H vs. FSC-A). The lymphocyte gate was further analyzed for expression of CD52 and -TCRa/b expression from this gated population.

In Silico off-site prediction

To evaluate possible off-target editing of the CD52 TALE-BEs, we generated *in silico* a list of potential off site targets of these BEs. That list was generated as follow. The TALE-BE have two binding sequences of 17 bp separated by a spacer. These binding sequences begin necessarily by a T. Hence, we first selected as potential targets all genomic sequences starting with a T, ending with an A, and having a size comprised between 27 bp and 67 bp (both included), allowing for spacers ranging from 10 to 40 bp. Then, the number of mismatches between the binding sequences of the potential target versus the actual TALE-BE target was counted. If that total number was greater than 8, the potential target was removed. Finally, all potential targets lacking a G in the left half of the spacer, or a C in the right half of the spacer (editing windows) were discarded.

Off-site and translocation multiplexed amplicon sequencing

rhAmp primers were designed on the on-target and/or off-target sites established by an *in silico* off-site prediction. Locus-specific forward and reverse primers were obtained from Integrated DNA Technologies (IDT) either in ready to use pools or individually plated, and use accordingly to IDT protocol for RNase H2-dependent multiplex assay amplification (1 cycle of s at 95°C 10 min; 14 cycles of 15s at 95°C followed by 8 min at 65°C; 1 cycle of 15 min at 99.5°C; hold at 4°C) followed by a universal PCR to add indexes (i5 or i7) for NGS (1 cycle of s at 95°C 3 min; 24 cycles of 15 s at 95°C followed by 30 s at 60°C and 30s at 72°C; 1 cycle of 1 min at 72°C; hold at 4°C). Purified PCR amplicons were sequenced on a NextSeq (Illumina) on a NextSeq 500/550 Mid Output Kit (150 cycles) cartridge.

Off-site and translocation multiplexed amplicon analysis

rhAmp sequencing reads were retrieved, trimmed for quality and aligned against the human genome (assembly GRCh38). Then, sequences that were aligned at positions corresponding to the generated potential off-sites were retrieved, if more than 200 reads were aligned. From these sequences, C>T and G>A and unedited C and G were counted. A χ^2 test of independent variables between the control and treated samples was performed. Sites with a p-value less than 5% were selected.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

AB, MY, AB, PD, and AJ conceived of the study and designed the experiments. AB, MY, MF, DT, and RH performed the experiments. AB, MY, SP, AD, and AJ analyzed the experiments. AB, MY, SP, AD, PD, and AJ

wrote the manuscript with support from all authors. All authors contributed to the article and approved the submitted version.

Funding

The authors declare that this study was funded by Cellectis.

Conflict of interest

AB, MY, SP, MF, DT, RH, AD, PD, and AJ are currently employed by the company Cellectis or former employees of the company Cellectis.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.1033669/full#supplementary-material

References

Alanis-Lobato, G., Zohren, J., McCarthy, A., Fogarty, N. M. E., Kubikova, N., Hardman, E., et al. (2021). Frequent loss of heterozygosity in CRISPR-Cas9-edited early human embryos. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2004832117. doi:10.1073/pnas.2004832117

Amit, I., Iancu, O., Levy-Jurgenson, A., Kurgan, G., McNeill, M. S., Rettig, G. R., et al. (2021). CRISPECTOR provides accurate estimation of genome editing translocation and off-target activity from comparative NGS data. *Nat. Commun.* 12, 3042. doi:10.1038/s41467-021-22417-4

Anzalone, A. V., Koblan, L. W., and Liu, D. R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38, 824–844. doi:10.1038/s41587-020-0561-9

Bose, S. K., White, B. M., Kashyap, M. V., Dave, A., De Bie, F. R., Li, H., et al. (2021). *In utero* adenine base editing corrects multi-organ pathology in a lethal lysosomal storage disease. *Nat. Commun.* 12, 4291–4316. doi:10.1038/s41467-021-24443-8

Boutin, J., Cappellen, D., Rosier, J., Amintas, S., Dabernat, S., Bedel, A., et al. (2022). ON-target adverse events of CRISPR-cas9 nuclease: More chaotic than expected. *Cris. J.* 5, 19–30. doi:10.1089/crispr.2021.0120

Boutin, J., Rosier, J., Cappellen, D., Prat, F., Toutain, J., Pennamen, P., et al. (2021). CRISPR-Cas9 globin editing can induce megabase-scale copy-neutral losses of heterozygosity in hematopoietic cells. *Nat. Commun.* 12, 4922. doi:10.1038/s41467-021-25190-6

Chaudhari, H. G., Penterman, J., Whitton, H. J., Spencer, S. J., Flanagan, N., Lei Zhang, M. C., et al. (2020). Evaluation of homology-independent CRISPR-cas9 off-target assessment methods. *Cris. J.* 3, 440–453. doi:10.1089/crispr.2020.0053

Doman, J. L., Raguram, A., Newby, G. A., and Liu, D. R. (2020). Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* 38, 620–628. doi:10.1038/s41587-020-0414-6

Ernst, J., and Kellis, M. (2017). Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* 12, 2478–2492. doi:10.1038/nprot.2017.124

Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J. E., Mudge, J. M., et al. (2021). Gencode 2021. *Nucleic Acids Res.* 49, D916–D923. doi:10.1093/nar/gkaa1087

Gautron, A. S., Juillerat, A., Guyot, V., Filhol, J. M., Dessez, E., Duclert, A., et al. (2017). Fine and predictable tuning of TALEN gene editing targeting for improved T cell adoptive immunotherapy. *Mol. Ther. - Nucleic Acids* 9, 312–321. doi:10.1016/j.omtn.2017.10.005

Geng, K., Merino, L. G., Wedemann, L., Martens, A., Sobota, M., Sanchez, Y. P., et al. (2022). CRISPR/Cas9 deletions induce adverse on-target genomic effects leading to functional DNA in human cells. *bioRxiv* 2021, 07.01.450727. doi:10.1101/2021.07.01.450727

Guo, J., Chen, X., Liu, Z., Sun, H., Zhou, Y., Dai, Y., et al. (2022). DdCBE mediates efficient and inheritable modifications in mouse mitochondrial genome. *Mol. Ther. - Nucleic Acids* 27, 73–80. doi:10.1016/j.omtn.2021.11.016

Hanna, R. E., Hegde, M., Fagre, C. R., DeWeirdt, P. C., Sangree, A. K., Szegletes, Z., et al. (2021). Massively parallel assessment of human variants with base editor screens. *Cell* 184, 1064–1080. e20. doi:10.1016/j.cell.2021.01.012

Jain, S., Shukla, S., Yang, C., Zhang, M., Fatma, Z., Lingamaneni, M., et al. (2021). TALEN outperforms Cas9 in editing heterochromatin target sites. *Nat. Commun.* 12, 606–610. doi:10.1038/s41467-020-20672-5

Juillerat, A., Dubois, G., Valton, J., Thomas, S., Stella, S., Maréchal, A., et al. (2014). Comprehensive analysis of the specificity of transcription activator-like effector nucleases. *Nucleic Acids Res.* 42, 5390–5402. doi:10.1093/nar/gku155

Kang, B. C., Bae, S. J., Lee, S., Lee, J. S., Kim, A., Lee, H., et al. (2021). Chloroplast and mitochondrial DNA editing in plants. *Nat. Plants* 7, 899–905. doi:10.1038/s41477-021-00943-9

Kluesner, M. G., Lahr, W. S., Lonetree, C. lin, Smeester, B. A., Qiu, X., Slipek, N. J., et al. (2021). CRISPR-Cas9 cytidine and adenosine base editing of splice-sites mediates highly-efficient disruption of proteins in primary and immortalized cells. *Nat. Commun.* 12, 2437–2512. doi:10.1038/s41467-021-22009-2

Koblan, L. W., Doman, J. L., Wilson, C., Levy, J. M., Tay, T., Newby, G. A., et al. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* 36, 843–846. doi:10.1038/nbt.4172

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi:10.1038/nature17946

Komor, A. C., Zhao, K. T., Packer, M. S., Gaudelli, N. M., Waterbury, A. L., Koblan, L. W., et al. (2017). Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci. Adv.* 3, eaao4774–10. doi:10.1126/sciadv.aao4774

Kurt, I. C., Zhou, R., Iyer, S., Garcia, S. P., Miller, B. R., Langner, L. M., et al. (2021). CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* 39, 41–46. doi:10.1038/s41587-020-0609-x

Lee, H., Lee, S., Baek, G., Kim, A., Kang, B. C., Seo, H., et al. (2021). Mitochondrial DNA editing in mice with DddA-TALE fusion deaminases. *Nat. Commun.* 12, 1190–1215. doi:10.1038/s41467-021-21464-1

Lim, K., Cho, S. I., and Kim, J. S. (2022). Nuclear and mitochondrial DNA editing in human cells with zinc finger deaminases. *Nat. Commun.* 13, 366–410. doi:10.1038/s41467-022-27962-0

Liu, Y., Zhou, J., Lan, T., Zhou, X., Yang, Y., Li, C., et al. (2022). Elimination of Cas9-dependent off-targeting of adenine base editor by using TALE to separately guide deaminase to target sites. *Cell Discov.* 8, 28–12. doi:10.1038/s41421-022-00384-4

Mok, B. Y., de Moraes, M. H., Zeng, J., Bosch, D. E., Kotrys, A. V., Raguram, A., et al. (2020). A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature* 583, 631–637. doi:10.1038/s41586-020-2477-4

Mok, B. Y., Kotrys, A. V., Raguram, A., Huang, T. P., Mootha, V. K., and Liu, D. R. (2022). CRISPR-free base editors with enhanced activity and expanded targeting scope in mitochondrial and nuclear DNA. *Nat. Biotechnol.* 40, 1378–1387. doi:10.1038/s41587-022-01256-8

Nahmad, A. D., Reuveni, E., Goldschmidt, E., Tenne, T., Liberman, M., Horovitz-Fried, M., et al. (2022). Frequent aneuploidy in primary human T cells after CRISPR–Cas9 cleavage. *Nat. Biotechnol.* [Epub ahead of print]. doi:10.1038/s41587-022-01377-0

Nakazato, I., Okuno, M., Yamamoto, H., Tamura, Y., Itoh, T., Shikanai, T., et al. (2021). Targeted base editing in the plastid genome of *Arabidopsis thaliana*. *Nat. Plants* 7, 906–913. doi:10.1038/s41477-021-00954-6

Papathanasiou, S., Markoulaki, S., Blaine, L. J., Leibowitz, M. L., Zhang, C. Z., Jaenisch, R., et al. (2021). Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nat. Commun.* 12, 5855. doi:10.1038/s41467-021-26097-y

Poirot, L., Philip, B., Schiffer-Mannioui, C., Le Clerre, D., Chion-Sotinel, I., Derniame, S., et al. (2015). Multiplex genome-edited T-cell manufacturing platform for "off-the-shelf" adoptive T-cell immunotherapies. *Cancer Res.* 75, 3853–3864. doi:10.1158/0008-5472.CAN-14-3321

Qasim, W., Zhan, H., Samarasinghe, S., Adams, S., Amrolia, P., Stafford, S., et al. (2017). Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* 9, eaaj2013. doi:10.1126/scitranslmed.aaj2013

Qi, X., Chen, X., Guo, J., Zhang, X., Sun, H., Wang, J., et al. (2021). Precision modeling of mitochondrial disease in rats via DdCBE-mediated mtDNA editing. *Cell Discov.* 7, 95. doi:10.1038/s41421-021-00325-7

Roth, T. L., Puig-Saus, C., Yu, R., Shifrut, E., Carnevale, J., Li, P. J., et al. (2018). Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* 559, 405–409. doi:10.1038/s41586-018-0326-5

Sabharwal, A., Kar, B., Restrepo-Castillo, S., Holmberg, S. R., Mathew, N. D., Kendall, B. L., et al. (2021). The FusX TALE base editor (FusXTBE) for rapid mitochondrial DNA programming of human cells *in vitro* and zebrafish disease models *in vivo*. *Cris. J.* 4, 799–821. doi:10.1089/crispr.2021.0061

Samuelson, C., Radtke, S., Zhu, H., Llewellyn, M., Fields, E., Cook, S., et al. (2021). Multiplex CRISPR/Cas9 genome editing in hematopoietic stem cells for fetal hemoglobin reinduction generates chromosomal translocations. *Mol. Ther. - Methods & Clin. Dev.* 23, 507–523. doi:10.1016/j.omtm.2021.10.008

Sánchez-Rivera, F. J., Diaz, B. J., Kastenhuber, E. R., Schmidt, H., Katti, A., Kennedy, M., et al. (2022). Base editing sensor libraries for high-throughput engineering and functional analysis of cancer-associated single nucleotide variants. *Nat. Biotechnol.* 40, 862–873. doi:10.1038/s41587-021-01172-3

Thuronyi, B. W., Koblan, L. W., Levy, J. M., Yeh, W. H., Zheng, C., Newby, G. A., et al. (2019). Continuous evolution of base editors with expanded target compatibility and improved activity. *Nat. Biotechnol.* 37, 1070–1079. doi:10.1038/s41587-019-0193-0

Tran, M. T. N., Kc, R., and Hewitt, A. W. (2022). A taxonomic and phylogenetic classification of diverse base. *Cris. J.* 5, 311–328. doi:10.1089/crispr.2021.0095

Valton, J., Dupuy, A., Daboussi, F., Thomas, S., Maréchal, A., Macmaster, R., et al. (2012). Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J. Biol. Chem.* 287, 38427–38432. doi:10.1074/jbc.C112.408864

Valton, J., Guyot, V., Marechal, A., Filhol, J. M., Juillerat, A., Duclert, A., et al. (2015). A multidrug-resistant engineered CAR T cell for allogeneic combination immunotherapy. *Mol. Ther.* 23, 1507–1518. doi:10.1038/mt.2015.104

Wang, X., Liu, Z., Li, G. L., Dang, L., Huang, S., He, L., et al. (2020). Efficient gene silencing by adenine base editor-mediated start codon mutation. *Mol. Ther.* 28, 431–440. doi:10.1016/j.ymthe.2019.11.022

Webber, B. R., Lonetree, C. lin, Kluesner, M. G., Johnson, M. J., Pomeroy, E. J., Diers, M. D., et al. (2019). Highly efficient multiplex human T cell engineering without double-strand breaks using Cas9 base editors. *Nat. Commun.* 10, 5222. doi:10.1038/s41467-019-13007-6

Wei, Y., Li, Z., Xu, K., Feng, H., Xie, L., Li, D., et al. (2022). Mitochondrial base editor DdCBE causes substantial DNA off-target editing in nuclear genome of embryos. *Cell Discov.* 8, 27–394. doi:10.1038/s41421-022-00391-5

Weisheit, I., Kroeger, J. A., Malik, R., Klimmt, J., Crusius, D., Dannert, A., et al. (2020). Detection of deleterious on-target effects after HDR-mediated CRISPR editing. *Cell Rep.* 31, 107689. doi:10.1016/j.celrep.2020.107689

Wiren, K. M., Ivashkiv, L., Ma, P., Freeman, M. W., Potts, J. T., and Kronenberg, H. M. (1989). Mutations in signal sequence cleavage domain of preproparathyroid hormone alter protein translocation, signal sequence cleavage, and membrane binding properties. *Mol. Endocrinol.* 3, 240–250. doi:10.1210/mend-3-2-240

Yang, M., Tkach, D., Boyne, A., Kazancioglu, S., Duclert, A., Poirot, L., et al. (2021). Optimized two-step electroporation process to achieve efficient nonviral-mediated gene insertion into primary T cells. *FEBS Open Bio* 12, 38–50. doi:10.1002/2211-5463.13292

Yuan, J., Ma, Y., Huang, T., Chen, Y., Peng, Y., Li, B., et al. (2018). Genetic modulation of RNA splicing with a CRISPR-guided cytidine deaminase. *Mol. Cell* 72, 380–394.e7. e7. doi:10.1016/j.molcel.2018.09.002

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A. H., Shea, D., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* 26, 535–541. doi:10.1038/s41591-020-0790-y

Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M. A., et al. (2021). Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* 39, 35–40. doi:10.1038/s41587-020-0592-2

frontiers | Frontiers in Genome Editing

TYPE Original Research
PUBLISHED 12 January 2023
DOI 10.3389/fgeed.2022.1031275

OPEN ACCESS

EDITED BY
Ayal Hendel,
Bar-Ilan University, Israel

REVIEWED BY
Adam Ameur,
Uppsala University, Sweden
Gavin Kurgan,
Integrated DNA Technologies, United States
Roberto Nitsch,
AstraZeneca, Sweden

*CORRESPONDENCE
Jenna Kropp Schmidt,
jkropp@wisc.edu

†These authors share senior authorship

SPECIALTY SECTION
This article was submitted to Genome Editing in Infectious Diseases, a section of the journal Frontiers in Genome Editing

RECEIVED 29 August 2022
ACCEPTED 15 December 2022
PUBLISHED 12 January 2023

CITATION
Schmidt JK, Kim YH, Strelchenko N, Gierczic SR, Pavelec D, Golos TG and Slukvin II (2023), Whole genome sequencing of *CCR5* CRISPR-Cas9-edited Mauritian cynomolgus macaque blastomeres reveals large-scale deletions and off-target edits. *Front. Genome Ed.* 4:1031275. doi: 10.3389/fgeed.2022.1031275

Whole genome sequencing of *CCR5* CRISPR-Cas9-edited Mauritian cynomolgus macaque blastomeres reveals large-scale deletions and off-target edits

Jenna Kropp Schmidt[1]*, Yun Hee Kim[1], Nick Strelchenko[1], Sarah R. Gierczic[1], Derek Pavelec[2], Thaddeus G. Golos[1,3,4†] and Igor I. Slukvin[1,5,6†]

[1]Wisconsin National Primate Research Center, University of Wisconsin-Madison, Madison, WI, United States, [2]University of Wisconsin Biotechnology Center, University of Wisconsin-Madison, Madison, WI, United States, [3]Department of Comparative Biosciences, University of Wisconsin-Madison, Madison, WI, United States, [4]Department of Obstetrics and Gynecology, University of Wisconsin-Madison, Madison, WI, United States, [5]Department of Pathology and Laboratory Medicine, University of Wisconsin-Madison, Madison, WI, United States, [6]Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, WI, United States

Introduction: Genome editing by CRISPR-Cas9 approaches offers promise for introducing or correcting disease-associated mutations for research and clinical applications. Nonhuman primates are physiologically closer to humans than other laboratory animal models, providing ideal candidates for introducing human disease-associated mutations to develop models of human disease. The incidence of large chromosomal anomalies in CRISPR-Cas9-edited human embryos and cells warrants comprehensive genotypic investigation of editing outcomes in primate embryos. Our objective was to evaluate on- and off-target editing outcomes in *CCR5* CRISPR-Cas9-targeted Mauritian cynomolgus macaque embryos.

Methods: DNA isolated from individual blastomeres of two embryos, along with paternal and maternal DNA, was subjected to whole genome sequencing (WGS) analysis.

Results: Large deletions were identified in macaque blastomeres at the on-target site that were not previously detected using PCR-based methods. *De novo* mutations were also identified at predicted CRISPR-Cas9 off-target sites.

Discussion: This is the first report of WGS analysis of CRISPR-Cas9-targeted nonhuman primate embryonic cells, in which a high editing efficiency was coupled with the incidence of editing errors in cells from two embryos. These data demonstrate that comprehensive sequencing-based methods are

Abbreviations: WGS, Whole genome sequencing; NHP, Non-human primate; HIV, Human immunodeficiency virus; HSC, Hematopoietic stem cell; MCM, Mauritian cynomolgus macaque; RNP, Ribonucleoprotein; CNV, copy number variation; WT, Wild-type; HET, Heterozygous; HOM, Homozygous; DEL, Deletion; InDel, Insertions and/or deletions; SNVs, Single nucleotide variants.

warranted for evaluating editing outcomes in primate embryos, as well as any resultant offspring to ensure that the observed phenotype is due to the targeted edit and not due to unidentified off-target mutations.

KEYWORDS

CRISPR- Cas9, whole genome sequencing (WGS), embryo, macaque, CCR5

1 Introduction

Advances in genome editing, particularly using CRISPR-Cas9 technology, have facilitated the introduction and correction of disease-associated mutations in animal and cell culture models. Non-human primates (NHPs) are superior for modeling human diseases as they share similar aspects of immune, neuro-, and reproductive physiology and are ideal for transplant and neurodevelopmental disorder research. The interest in creating NHP models of human disease has been augmented by the need to better define the etiology of a disease and for the development of treatments and therapeutics (Dray et al., 2018; Abbott et al., 2019; Moshiri et al., 2019; Tapmeier et al., 2021; Ozirmak et al., 2022). For example, resistance to human immunodeficiency virus (HIV) has been observed in human patients with a 32 base pair deletion in the *CCR5* gene (*CCR5*-Δ32). CCR5 serves as an HIV co-receptor (Dean et al., 1996; Liu et al., 1996; Samson et al., 1996). Transplantation of hematopoietic stem cells (HSCs) containing the *CCR5*-Δ32 mutation to human HIV patients led to the cure of HIV infection in some but not all cases (Hütter et al., 2009; Allers et al., 2011; Henrich et al., 2014; Cummins et al., 2017). Generating NHPs with *CCR5*-deletions would aid in determining the mechanisms of HIV elimination following transplantation of allogeneic HSCs with *CCR5* mutations and the development of clinical protocols for reproducible HIV cure (Schmidt et al., 2022b).

Genome editing approaches to create gene disruption in NHP embryos have been successful, yet evidence of CRISPR-Cas9-induced chromosomal anomalies in mammalian cells and embryos warrants further investigation of embryonic editing outcomes in primate embryos. CRISPR-Cas9 editing has been shown to result in large scale deletions (up to 6 kb) and whole chromosome loss leading to genomic instability in mouse embryonic stem cells (Kosicki et al., 2018) and embryos (Adikusuma et al., 2018; Papathanasiou et al., 2021). CRISPR-Cas9 editing in human embryos has resulted in the loss of the targeted allele (Zuccaro et al., 2020) and also segmental chromosome losses (Alanis-Lobato et al., 2021). Moreover, large scale deletions in human embryos at an off-target site were also observed (Zuccaro et al., 2020). Loss of heterozygosity surrounding the on-target site is another consequence of CRISPR-Cas9 targeting observed in human embryos (Zuccaro et al., 2020; Alanis-Lobato et al., 2021). Collectively, these studies have revealed undesired on- and off-target mutations that arise when using wild-type Cas9 for gene correction in human and mouse embryos.

We previously demonstrated CRISPR-Cas9 editing of *CCR5* in Mauritian cynomolgus macaque (MCM, *Macaca fascicularis*) embryos using PCR-based methods to confirm successful targeting of the locus (Schmidt et al., 2020). The objective of the present study was to comprehensively evaluate on- and off-target editing in CRISPR-Cas9-edited MCM embryos using whole genome sequencing (WGS) methods to survey individual blastomeres. Molecular analysis revealed large-scale deletions contributing to greater mosaicism within individual embryos than was previously identified using PCR-based methods. Given that large-scale on- and off-target mutations might hinder establishment of a viable pregnancy, further optimization of macaque embryo editing to avoid targeting errors would be essential to facilitate generation of novel NHP models for human diseases.

2 Materials and methods

2.1 Animals

Methods for deriving the MCM embryos analyzed in the present study were previously reported (Schmidt et al., 2020). Parental DNA was obtained from a female (12 years) and male (6 years) MCM used in that study. All procedures were performed in accordance with the NIH Guide for the Care and Use of Laboratory Animals and under the approval of the University of Wisconsin-Madison College of Letters and Sciences and Vice Chancellor's Office for Research and Graduate Education Institutional Animal Care and Use Committee.

2.2 Isolation of CRISPR-Cas9 injected embryo DNA

In vitro fertilized Mauritian cynomolgus macaque embryos were produced as previously described (Schmidt et al., 2020). Briefly, one-cell stage embryos were microinjected with Cas9 complexed with two sgRNAs targeting exon 2 of the *CCR5* gene to form the ribonucleoprotein (RNP). Embryos were cultured individually in a microwell of a CultureCoin MIRI-TL dish (Esco Medical, Denmark) containing 25 μL of Global medium overlaid with mineral oil and the culture dish was placed in a MIRI TL Time-Lapse incubator (Esco Medical) to monitor embryo

development. Individual blastomeres from two embryos arrested at the 6-cell and 9-cell stage were isolated. The zona pellucida was removed by treatment with 1 mg/ml of activated pronase E (Sigma-Aldrich, cat no: P2730) and the embryo was washed in calcium- and magnesium-free PBS with .2% EDTA and 1 mg/ml human albumin (MP Biomedicals, cat no: 823051). Individual blastomeres were then dissociated by gently pipetting the embryo. DNA was amplified from individual blastomeres using a REPLI-G single cell kit (Qiagen, cat no: 150343).

2.3 PCR analysis of CCR5 on-target editing in individual blastomeres

PCR amplifications were performed as previously described to assess the targeted *CCR5* region (Schmidt et al., 2020) using primers that either amplified short (613 bp) or long (2,925 bp) amplicons surrounding the predicted cut sites. PCR products were run on 1.2%–1.5% agarose gels at 120 V. On-target *CCR5* editing was determined by visualizing either the expected wild-type (unmodified) *CCR5* PCR amplicon size of 613 bp or a biallelic mutation producing a product of 415 bp. To evaluate large-scale deletions near the on-target site, a long-range *CCR5* PCR was performed. PCR primer sequences and expected amplicon sizes are listed in Supplementary Table S1. PCR reactions were performed using the Q5 Hot start High-Fidelity DNA polymerase kit following manufacturer recommendations. Gel electrophoresis was performed using standard methods to visualize the amplicons.

2.4 Isolation of parental DNA

Blood draws from both the oocyte and semen donors of the in vitro-produced embryos were performed to obtain parental DNA from peripheral blood mononuclear cells. Genomic DNA was isolated from blood cells using a Quick-DNA Miniprep kit (Zymo Research, cat no: D3024).

2.5 DNA quality assessment

DNA quality was assessed at the University of Wisconsin Biotechnology Center's by the NexGen DNA Sequencing Core using an Agilent Femto Pulse system (Agilent, Santa Clara, CA) to confirm a uniform yield of DNA product with the average product length of greater than 9.4 kb.

2.6 Whole genome sequencing and analysis

Whole genome sequencing was performed by the University of Wisconsin-Madison Biotechnology Center using the Illumina short-read platform and a NovaSeq 6000 instrument. Reads were trimmed to remove sequencing adapters and low quality base calls using the software skewer (Jiang et al., 2014) and then mapped to the *Macaca fascicularis* reference genome, M_fascicularis_5.0, using an Illumina Dynamic Read Analysis for GENomics (DRAGEN) Bio-IT platform version 3.7. Small variant and calling was performed using DRAGEN. Variants from control (parental) samples were used to filter and identify *de novo* mutations. Variant annotation was performed using SNEPeff tool that will predict synonymous or non-synonymous amino acid changes, gains or losses of start/stop codons, and frame shifts due to insertions or deletions. Structural analysis was performed using Parliament2 (Zarate et al., 2020) and only those called by at least two callers were included. Variants with lower quality that were filtered out as well as non-filtered variants are included as potential candidate mutations. *De novo* structural variants were those identified in blastomeres that were not present in the parental sequences. Short read sequencing is not ideally suited for calling structural variants, hence low quality-filtered out variants as well non-filtered variants are both included as potential candidates. Integrated Genomics Viewer software (https://software.broadinstitute.org/software/igv/download) was used to view the WGS data.

2.7 Off-target analysis

Potential CRISPR-Cas9 off-target regions were identified using the Cas-OFFinder tool (Bae et al., 2014) (http://www.rgenome.net/cas-offinder/) and allowing for three mismatches. Regions of interest were then evaluated in the WGS dataset to see if *de novo* mutations were present in individual blastomeres compared to the parental DNA. The presence of *de novo* mutations in three predicted off-target genes were assessed by Sanger sequencing of PCR amplicons containing the region of interest for individual blastomeres. DNA obtained from a wild-type cynomolgus macaque iPSC line was sequenced in parallel. PCR reactions were performed using the Q5 Hot start High-Fidelity DNA polymerase kit following manufacturer recommendations and the reactions were cleaned up using a Gel extraction and PCR clean up kit (IBI, cat no: IB47010). PCR primer sequences are listed in Supplementary Table S1. Sanger sequencing reactions were carried out by Functional Biosciences Inc., Madison, Wisconsin and the sequencing data was analyzed using the 4Peaks (https://nucleobytes.com/4peaks/index.html) application.

3 Results

3.1 WGS of individual blastomeres produces variable sequence coverage

To functionally delete *CCR5* in macaque embryos, we designed gRNAs that would encompass a 24-bp deletion that

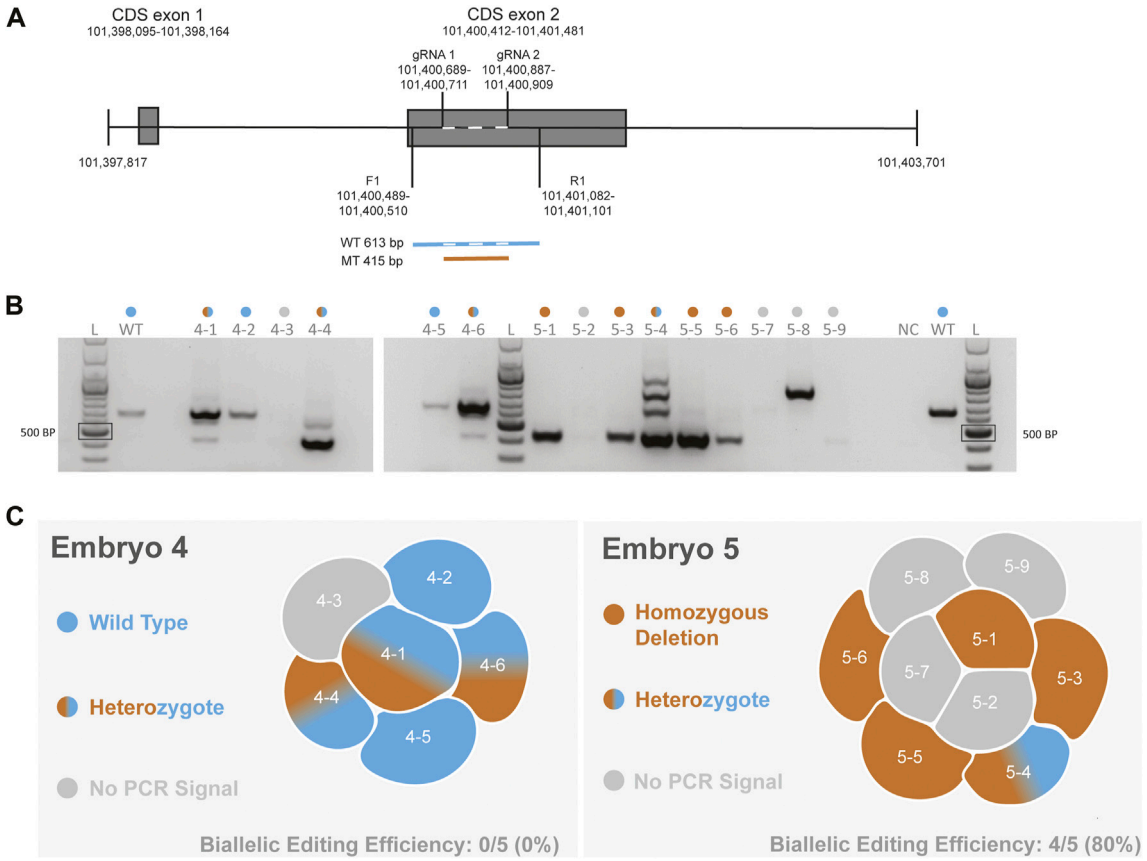
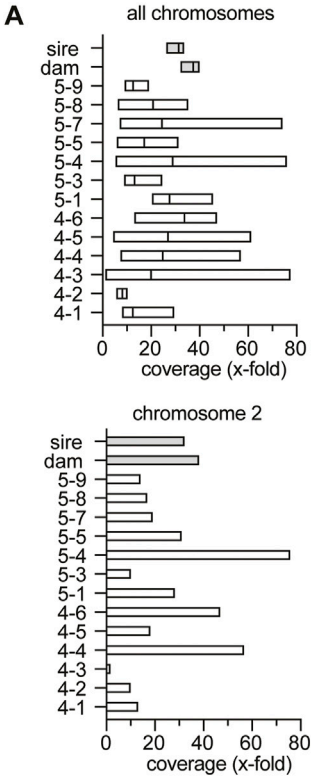


FIGURE 1
Evaluation of *CCR5* editing in MCM embryos using PCR-based methods **(A)**. Schematic of the *CCR5* gene including gRNA targeting and the forward (F1) and reverse (R1) primer. Wild-type (WT) product is 613 bp in length, whereas a biallelic deletion mutation (MT) produces a 415 bp product. Dashed line with the wild-type sequence indicates the targeting region **(B)**. Gel electrophoresis images of PCR products from blastomeres of embryos 4 and 5. A positive control reaction with DNA from an unmanipulated control embryo and a no template negative control (NC) were included. The colored dots above each lane indicate the editing outcome as indicated in 1C. The PCR and gel electrophoresis results were provided in our initial report describing *CCR5* editing in MCM embryos (Schmidt et al., 2020) **(C)**. Diagram summarizing the PCR-based editing outcome and biallelic editing efficiency within each embryo.

has previously been shown to be essential for expressing *CCR5* in non-human primates (Chen et al., 1998). Our previous cell-based editing experiments confirmed successful on-target editing with functional deletion of the *CCR5* gene in both human (Kang et al., 2015) and macaque (D'Souza et al., 2022) iPSCs. A schematic diagram of the targeting region is shown in Figure 1A. In our initial report describing targeting of this region in MCM embryos, PCR-based methods were used to evaluate CRISPR-Cas9 targeting of the *CCR5* locus (Schmidt et al., 2020). Two embryos were dissociated into individual blastomeres and DNA was isolated for PCR evaluation and single-cell WGS. PCR and gel electrophoresis revealed editing mosaicism in each embryo (Figures 1B,C), although PCR signal was undetected in one and three blastomeres from embryo 4 and 5, respectively (Figure 1B).

Single-cell DNA amplification and WGS was performed on DNA from all six blastomeres of embryo 4 and 8 of 9 blastomeres from embryo 5. In addition, DNA isolated from peripheral blood mononuclear cells from the sire and dam of the embryos was sequenced in parallel. Chromosomal coverage varied across individual blastomeres and chromosomes ranging from .81–77.77-fold coverage, whereas the parental sequence coverage was at a depth of ~30x, as expected for somatic cells. A sequencing coverage of 30x is interpreted as the genome was sequenced ~30 times. Figure 2A shows the mean coverage and range of sequence depth across chromosomes for each sample and illustrates the variability in coverage in blastomeres compared to parental DNA isolated from peripheral blood cells. The *CCR5* gene resides on chromosome 2, a chromosome that greatly varied in sequence depth coverage across blastomeres (Figure 2A). Sequence coverages for each sample by chromosome are listed in Supplementary Table S2. Blastomere 5-2 had an atypical distribution of GC content and was excluded from all analyses. For the remaining blastomeres, the sequence coverage at on- and off-target regions was taken

A



B Comparison of *CCR5* editing outcomes by PCR versus WGS analysis

Blastomere	PCR	WGS
4-1	HET DEL/multiple bands	confirms PCR
4-2	wild-type	confirms PCR
4-3	no signal	low coverage, inconclusive
4-4	HET DEL	confirms PCR
4-5	wild-type	confirms PCR
4-6	HOM DEL/multiple bands	confirms PCR
5-1	HOM DEL	confirms PCR
5-2	no signal	poor sequencing quality
5-3	HOM DEL	low coverage, inconclusive
5-4	HET DEL/multiple bands	HOM DEL
5-5	HOM DEL	confirms PCR
5-7	no signal	low coverage, inconclusive
5-8	no signal/unexpected band	HOM DEL
5-9	no signal	low coverage, inconclusive

C Sequence coverage of blastomeres containing a deletion not detected by PCR

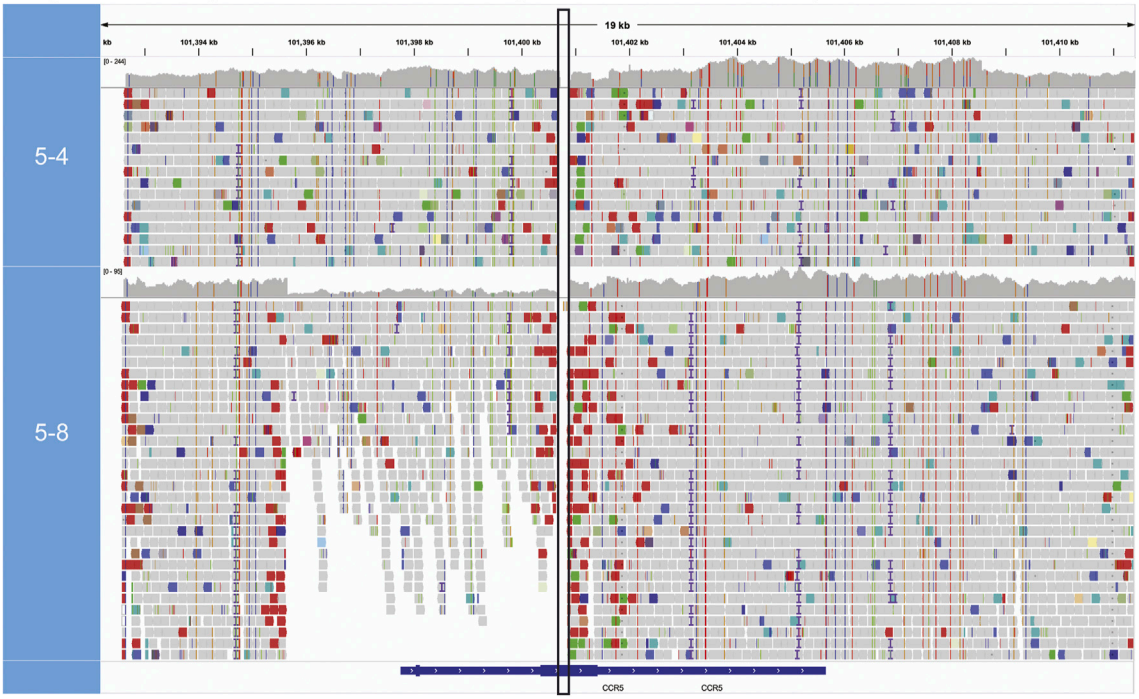


FIGURE 2
WGS identification of genotypes not identified by PCR-based methods **(A)**. The mean and range in sequencing coverage depth (x-fold coverage) are plotted for each blastomere and parental DNA. The mean and range for all chromosomes is shown in the upper graph and the lower graph shows chromosome 2; *CCR5* resides on chromosome 2 **(B)**. Comparison of *CCR5* editing outcomes by PCR versus WGS analysis. Blue text indicates cells where WGS identified deletions not previously identified by PCR. HET: heterozygous, HOM: homozygous **(C)**. Sequence coverage at the *CCR5* targeting region in blastomeres 5-4 and 5-8 where homozygous/biallelic edits were observed by WGS. The WGS viewer software indicates potential deletions with red bars and in the alignment tracks of 5-8, these are present around the target region and were minimal to absent in the parental coverage map. The vertical black box indicates the expected deleted region between the gRNA target sites.

into consideration and when limited sequences were observed at the region of interest the WGS result was deemed as not conclusive.

3.2 WGS identified additional on-target deletions

WGS confirmed the genotypes identified using PCR-based methods for most of the blastomeres that had detectable PCR signal (7 out of 9) and determined the genotype for one blastomere in which the *CCR5* region could not be amplified by PCR (Figure 2B). Sequences that spanned the target region are indicative of wild-type sequences, whereas a deletion was inferred if there was a break in the sequence coverage. Representative examples of wild-type (WT), heterozygous (HET) and homozygous deletion (HOM DEL) genotypes as determined by WGS are illustrated in Supplementary Figure S1. WGS sequence coverage at the *CCR5* targeting region also identified deletions not previously observed using PCR-based methods (Figure 2B). If there was poor sequence coverage at the target site, the WGS genotype could not be determined and was deemed inconclusive. Blastomere 5-4 was identified to be HET by PCR-based methods, however, when looking at the sequencing alignment, no sequences spanned the region between the gRNA sites indicating that the blastomere contained a biallelic deletion (Figure 2C). Regardless of the unexpected gel band pattern of 5-4, the HOM DEL was confirmed *via* Sanger sequencing of the amplicons isolated from the three lowest bands in the agarose gel; each amplicon contained the expected ~200 bp deletion and no WT sequences were detected (Supplementary Material S2). PCR using the standard primer pair previously failed to identify the genotype of blastomere 5-8, whereas WGS coverage indicated deletions spanning the gRNA sites and revealed large-scale deletions that encompassed the PCR primer sequences (Figure 2C). Moreover, the depth of coverage was reduced by approximately half at the 5' end, confirming that one allele contained a large-scale deletion upstream of the first gRNA site.

3.3 Identification and validation of large-scale deletions at the CCR5 targeting region

De novo structural variants that were not present in the parental DNA but were within individual blastomeres of both embryos 4 and 5 were identified by WGS (Table 1). As short read sequencing platforms are not ideally suited for identifying structural variants, those that were called in at least two variant callers are listed. The on-target deletion between the gRNA sites was identified as a variant in 4-4. Large-scale deletions that span the target region were identified in blastomeres 4-6, 5-4, and a similar deletion of ~5.2 kb was seen in both 5-5 and 5-8. Several inversions were also identified and within each embryo an inversion was unique to a pair of blastomeres.

The PCR primer sequences used for initially genotyping blastomeres were within the deleted regions, hence these deletions could not be identified using PCR-based methods. The positions of the large-scale deletions identified by WGS in blastomeres 4-6, 5-5, and 5-8 are illustrated in Figure 3A. PCR-based methods using primers flanking each deleted region followed by gel electrophoresis confirmed the presence of the large-scale deletions in 4-6, 5-5, and 5-8 (Figure 3B). The large-scale deletions were confirmed by Sanger sequencing of PCR amplicons for the deletions detected in 4-6, 5-5, and 5-8 with one exception (Supplementary Material S1). Poor sequencing data did not allow for verification of the ~5.2 kb deletion in the 5.8 blastomere whereas the deleted sequence was confirmed in blastomere 5.5 (Supplementary Material S1).

3.4 WGS detection of on-target INDELs in blastomere 5-4

The presence of insertions and/or deletions (INDELs) within the on-target region was evaluated. Single nucleotide variants (SNVs) were identified and considered to be *de novo* mutations if they were present in the blastomere but not detected in either parent. The number of *de novo* SNVs detected in each blastomere is provided in Supplementary Table S3. Homozygous insertions of 13 bp, 5 bp, 11 bp, and 33 bp were identified near the gRNA one cut site in blastomere 5-4 that were not identified in the cynomolgus macaque reference genome nor the parental DNA sequences (Figure 4). This was the only blastomere with INDEL formation near a cut site.

3.5 Off-target edits detected at loci with restricted sequence homology to the gRNAs

To assess the feasibility of detecting off-target mutations by WGS, we utilized the *in silico* Cas-OFFinder tool to identify off-target sites based on sequence homology to the gRNA sequences and the total number of off-target sites when allowing for up to 9 mismatches between the gRNA and off-target sequence is provided in Supplementary Table S4. While off-target editing can occur at sites with up to 4 mismatches (Pattanayak et al., 2013; Cromer et al., 2022), we focused on assessing the WGS data at 93 predicted off-target regions that contained 2 or 3 mismatches which occurs more commonly. *De novo* mutations within individual blastomeres were identified by WGS in 16 predicted off-targeted sites of which 7 were located within genes and nine were located in

TABLE 1 ***De novo*** **structural variants at the** ***CCR5*** **locus identified in individual blastomeres.**

	Blastomere	Variant type	Position	Size (kb)
Embryo 4	4-4	inversion	25,740,108–123,812,023	98,071.915
		inversion	91,767,007–127,108,482	35,341.475
		deletion	101,400,706–101,400,904	.198
	4-5	inversion	25,740,424–123,812,038	98,071.915
	4-6	duplication	82,604,203–135,851,822	53,247.619
		deletion	101,400,160–101,400,916	.756
Embryo 5	5-3	inversion	51,210,725–141,809,530	90,598.805
	5-4	deletion	14,388,873–142,671,745	128,282.872
		inversion	91,766,840–127,108,372	35,341.532
	5-5	inversion	91,766,814–127,108,375	35,341.561
		deletion	101,395,673–101,400,914	5.214
	5-8	deletion	101,395,673–101,400,914	5.214
		deletion	101,400,695–101,401,620	.925

intergenic regions (Table 2). The specific mutations identified by WGS at each off-target site are provided in Supplementary Material S2.

Blastomeres with mutations at off-target regions identified by WGS were subjected to PCR and Sanger sequencing, with the exception that all blastomeres of embryo 5 were sequenced for the *SFMTB2* region. Sanger sequences are shown in Table 3. The T to G point mutation at the predicted region in the *NFASC* gene was identified in blastomeres of both embryos 4 and 5. This mutation is likely a spontaneously occurring mutation as the parental genotype was T/T, embryo 1 blastomeres were G/G and embryo 5 blastomeres carried either a T/T, G/G or T/G genotype. *De novo* deletions near the predicted cut site of the off-target regions in the *SFMBT2* and *LIPC* genes suggests off-target editing by the CRISPR-Cas9 RNP (Table 3). Furthermore, blastomeres of both embryos show deletions in *SFMBT2* and pairs of blastomeres in embryo 5 have the same 7 or 8 bp deletion suggestive that the editing occurred in a previous cell division similar to mutations introduced by on-target CRISPR-Cas9 targeting.

3.6 Structural variants detected at off-target sites

The presence of structural variants at predicted off-target sites was investigated for the off-target regions in the *NFASC*, *SFMBT2,* and *LIPC* genes. Large-scale deletions, inversions and duplications were identified at these sites by WGS (Table 4). An inversion and deletion at the *NFASC* off-target region was observed in each blastomere of embryos 4 and 5, and an inversion identified at the *LIPC* site was shared between blastomeres 4-5 and 4-6. A smaller deletion of 210 bp was detected in blastomere 4-1. Fewer off-target structural variants were shared among blastomeres within an embryo compared to those that were identified at the on-target site. A summary of on- and off-target mutations identified by WGS in each blastomere is provided in Supplementary Table S5.

4 Discussion

In the present study, comprehensive assessment of CRISPR-Cas9 targeting in MCM blastomeres by WGS confirmed editing mosaicism, revealed undesired on- and off-target editing events in NHP embryos, including large scale deletions, and resolved genotypes at the on-target sites that were previously undetected using PCR-based methods. INDELs were observed at on- and predicted off-target sites, where sequence disruption was confirmed by Sanger sequencing for two off-target regions. WGS analysis also provided insight into the timing of CRISPR-Cas9 targeting as identical structural variants and *de novo* mutations were shared in pairs of blastomeres but were not identified in the majority of blastomeres suggesting that editing was delayed and did not occur at the one-cell stage. While CRISPR-Cas9 can introduce mutations at disease-associated loci in NHP embryos, the occurrence of unexpected editing events requires rigorous assessment of not only embryos, but also the offspring to confirm that any resulting phenotype is not due to off-target effects.

Undesired editing events at the on-target site, including large-scale deletions, have been observed in human and

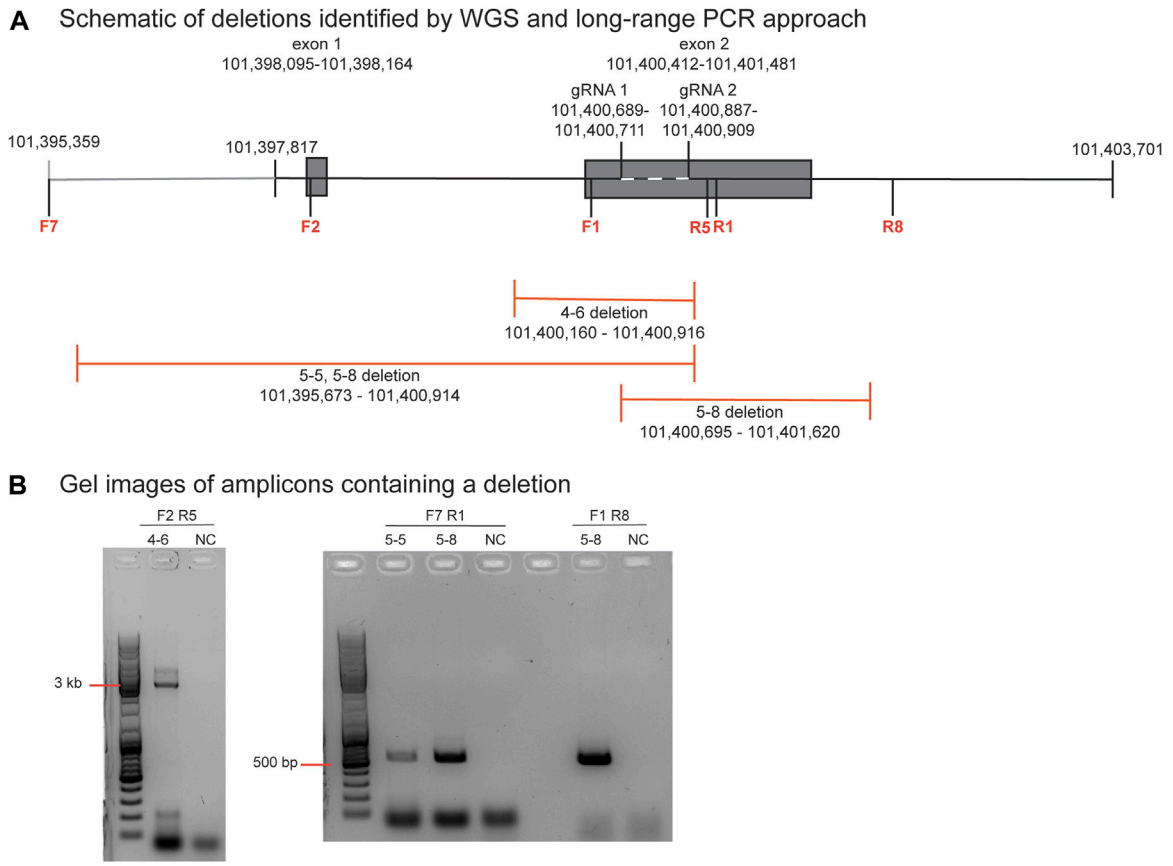


FIGURE 3
Identification and verification of a large-scale deletion identified by WGS in blastomere 4-6 **(A)**. Schematic diagram of deletions identified by WGS in blastomeres 4-6, 5-5, and 5-8, and the positions of the primers to evaluate the ~756 bp deletion in blastomere 4-6. Primers are denoted as F or R and their sequences are listed in Supplementary Table S1 **(B)**. Gel electrophoresis image of the PCR amplicons generated by primers that flank each deletion. A no template negative control (NC) reaction was also included.

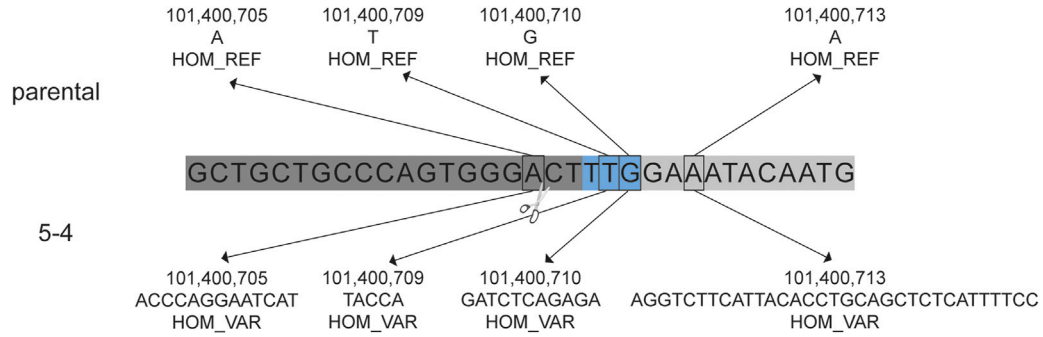


FIGURE 4
Insertions in blastomere 5-4 at the CCR5 target site. *De novo* insertions identified in blastomere 5-4 that were not identified in the parental DNA. The gRNA 1 sequence is indicated by the dark gray bar and the blue indicates the PAM sequence (5′-TTG). Homozygous (HOM) variants (VAR), specifically insertions, were detected near the predicted DNA cut site (indicated by scissors) that were not detected in the cynomolgus macaque reference genome (REF) nor the parental DNA sequence.

TABLE 2 ***De novo*** **mutations identified by WGS in predicted off-target regions.**

	Chr	Position	Gene	Sequence (5′-3′)	Blastomeres with *de novo* variants
gRNA 1	2	101,400,689–101,400,711	*CCR5**	GCTGCTGCCCAGTGGGACTT	
	1	63,497,146–63,497,168	*NFASC*	TCAGCTGCCCAGTGTGACTT	4-4, 4-5, 5-3, 5-4, 5-7, 5-8, 5-9
	1-scaffold	378,063–378,085	intergenic_ke145507.1	GCTGCTGCCCAGTAGTGCTT	4-4, 4-6
	10	31,267,650–31,267,672	intergenic_CM001928.1	TCTGCTGCCCAGTGGGCCTC	5-1
	19	40,019,282–40,019,304	*DLL3*	GCCCCTGCACAGTGGGACTT	4-4, 4-6, 5-3
gRNA 2	2	101,400,887–101,400,909	*CCR5**	GCTGTGTTTGCCTCTCTCCC	
	7	35,664,631–35,664,653	*LIPC*	GCTGTGCTTGCCTCTTTCCC	4-5, 4-6, 5-1
	9	7,497,178–7,497,200	*SFMBT2*	GCTGGCTTTGCCTCTCTCCC	4-3, 4-6, 5-5, 5-7, 5-8
	11	129,547,583–129,547,605	intergenic_CM001929.1	GCTGTGCTTCCCTCTCTCCC	4-5, 4-6, 5-1, 5-5
	14	14,725,950–14,725,972	intergenic_CM001932.1	GCTGTGTTCTCCTCTCTCCC	4-5, 4-6, 5-4
	1	199,634,457–199,634,479	intergenic_CM001919.1	GCTGTGGTTTCCTCTCTCCT	5-3
	3	53,360,905–53,360,927	intergenic_CM001921.1	CCTGTGATTACCTCTCTCCC	4-3, 4-5, 5-1, 5-5
	4	2,457,033–2,457,055	intergenic_CM001922.1	GCTGTGTCTCCCTCTCCCCC	5-1
	5	188,649,889–188,649,911	intergenic_CM001923.1	AATGTGTTTGCCTCTTTCCC	4-3, 5-1, 5-4, 5-7, 5-8, 5-9
	7	28,263,588–28,263,610	*CYP19A1*	GGAGTGTTGGCCTCTCTCCC	5-7
	12	62,901,723–62,901,745	intergenic_CM001930.1	GCTGTGTCCGCCTCTCTCGC	4-1, 4-3, 5-3, 5-5, 5-8
	17	62,039,482–62,039,504	*NDFIP2*	GCTGTCTTTGCCACTCTCAC	4-3
	19	49,110,480–49,110,502	*CCDC114*	GCTGTGCATCCCTCTCTCCC	5-1, 5-3, 5-4, 5-8

A total of 93 regions were assessed that contained 2-3 mismatches. *Denotes on-target gene sequence. Chr: chromosome.

mouse embryos targeted with wild-type Cas9 nuclease (Adikusuma et al., 2018; Zuccaro et al., 2020; Alanis-Lobato et al., 2021; Papathanasiou et al., 2021). In the present study, the expected 198 bp deletion as well as 756 bp, 925 bp and ~5.2 kb deletions were detected at the *CCR5* on-target site through WGS structural variant analysis of individual NHP embryonic cells. Deletions have been previously described in reports of CRISPR-Cas9 targeting in NHP embryos that were transferred to surrogates and produced edited offspring, including an ~11.5 kb deletion in *SHANK3* in one cynomolgus monkey (Zhao et al., 2017), a ~7.2 kb deletion in *PINK1* in two rhesus monkeys (Yang et al., 2019), and 920 bp at the *OCT4* knock-in site in a cynomolgus monkey (Cui et al., 2018). The *SHANK3* mutant died *in utero* at 125 days of gestation (term is 165 days) and the two *PINK1* mutants were triplets that died days after birth. These studies demonstrated that implantation and pregnancies can be achieved despite the presence of large-scale editing errors. Furthermore, on-target deletions have now been observed across studies in both NHP embryos and tissues where different genes were targeted by wild-type CRISPR-Cas9, which necessitates refinement of genome editing tools for creating precise disease-associated mutations.

In the current study, we identified pairs of blastomeres of the same embryo which contained similar structural variants, however these variants were not identical suggesting that they arose from separate editing events (e.g., blastomeres 5-5 and 5-8 shared a ~5.2 kb deletion and 5-4 and 5-5 shared a ~35,341 kb inversion near the *CCR5* targeting site). Mosaicism in blastomere genotypes suggests that editing was delayed and did not occur at the one-cell stage as the deletion was not detected in all cells of the embryo. Editing mosaicism has been observed in tissues of edited NHPs produced from embryo transfer of CRISPR-Cas9 targeted embryos (Niu et al., 2014; Chen et al., 2015; Tsukiyama et al., 2019; Zhou et al., 2019).

A goal of the present study was to assess the feasibility of using WGS to assess off-target effects in individual blastomeres. An *in silico* based approach guided the nomination of potential off-target regions based on sequence homology to the gRNAs allowing up to three mismatches. Following *CCR5* targeting, mutations were observed in the genes *SFMBT2* and *LIPC* by WGS and the introduction of INDELS was confirmed by Sanger sequencing. Blastomeres of both embryos displayed sequence disruption *via* INDEL formation with the presence of 4, 7 or 9 bp deletions in some of the cells at the predicted off-target site within the *SFMBT2* gene. These results confirmed that off-target editing could be assessed by

TABLE 3 Sanger-sequencing of amplicons containing predicted off-target mutations identified by WGS.

Gene	Sample	Sequence (5′-3′)
NFASC	WT	TCAGCTGCCCAGTGTGACTTGGG
	4-4	TCAGCTGCCCAGGGTGACTTGGG
	4-5	TCAGCTGCCCAGGGTGACTTGGG
	5-3	TCAGCTGCCCAGTGTGACTTGGG
	5-4	TCAGCTGCCCAGGGTGACTTGGG
	5-7	TCAGCTGCCCAGGGTGACTTGGG
	5-8	TCAGCTGCCCAGGGTGACTTGGG
	5-9	TCAGCTGCCCAGGGTGACTTGGG
SFMBT2	WT	GCTGTGTTTGCCTCTCTCCCAGG
	4-3	GCTGGCTTTGCCTCTC----GGG
	4-6	GNTGGNTTTGCCTCTCTNCCAGG
	5-1	NNNGNNTTNNCNNNNNCNCCAGG
	5-3	GCTGGCTTTG--------CCAGG
	5-4	NNGGCNTTNNCATTTTTCCCAGG
	5-5	GCTGGCTTTG--------CCAGG
	5-6	NNNGGCTTTNCCTTTTTCCCAGG
	5-7	GCTGGCTTTGCC-------CGGG
	5-8	GNNGGATTTGCC-------CGGG
LIPC	WT	GCTGTGCTTGCCTCTTTCCCTGG
	4-5	GCTGTG---G-----NTNNCTGG
	4-6	GNNGNGCNNGCNTCTTTCCCTGG
	5-1	GCTGTGCTTGCCTCTTTCCCTGG

WT, wild-type sequence; underline text: predicted cut site; bolded font: PAM, sequence; red font: denotes mutation; dash(-): deletion; N, not conclusive base determination.

WGS in individual blastomeres, although we used a biased *in silico* method that relied on assessing targets with sequence homology and did not evaluate potential targets with greater than three mismatches. Additional *in silico* nominated targets should be evaluated to fully assess the impact of off-target editing as CRISPR-Cas9 cleavage can occur at off-target sites with up to four mismatches (Pattanayak et al., 2013; Cromer et al., 2022). Moreover, unbiased methods that survey the whole genome without prior knowledge or prediction of sequence homology would be more informative, yet there is not a current superior method or technique for this analysis (Chaudhari et al., 2020; Atkins et al., 2021). *In vitro*-based off-target analysis methods in the future could be adapted for single-cells, but with current use of whole genome amplification (WGA) and an incomplete reference genome, a large number of false positives might be called due to errors incurred during WGA or due to differences in the reference assembly.

Relatively few off-target mutations have been identified in studies that have generated edited NHP offspring by transfer of CRISPR-Cas9 targeted embryos to surrogate embryo recipients. A 2 bp deletion in one off-target region was reported in an edited cynomolgus monkey (Cui et al., 2018) and one intronic and two intergenic INDELS were identified in two edited rhesus monkeys (Wang et al., 2018). Luo et al. (2019) identified *de novo* mutations that the authors thought were not introduced by CRISPR-Cas9 and rather could be attributed to natural spontaneous generational mutations or that were due to technical noise. In these previous NHP studies, WGS analysis was performed on DNA obtained from cells or tissues of live offspring or miscarried fetuses whereas here we reported WGS on individual blastomeres of NHP embryos. The higher incidence of off-target editing in this study could be explained by the editing efficiency of the RNP or it is possible that such significant off-target errors could have been embryonic lethal and therefore not present in offspring that survived to or near term. Editing has shown to be more rapid and efficient when targeting with an RNP versus Cas9 mRNA in NHP embryos (Midic et al., 2017). To mitigate potential off-target editing events, the following strategies could be implemented: 1) microinjection of the RNP at the time of fertilization (Lamas-Toranzo et al., 2019), 2) use of a Cas9 nuclease modified to improve specificity (Huang et al., 2022), or 3) use of a base or prime editing approach that does not result in a double-stranded DNA break (Zeballos and Gaj, 2020).

Off-target INDELs and segmental chromosome errors introduced by CRISPR-Cas9 have been observed in human embryos (Zuccaro et al., 2020). The formation of INDELs has been reported at a predicted off-target site that had two mismatches to the gRNA sequence (Zuccaro et al., 2020). Moreover, segmental chromosome errors were detected near predicted off-target sites and were often restricted to one cell, hence the authors concluded that the events likely occurred during the second or third cell cycle. While the present study did not focus on whole or segmental chromosomal errors, structural variants were detected at off-target sites that could be investigated in the future.

Large-scale mutations introduced by CRISPR-Cas9 editing in human and mouse embryos have shown to contribute to genomic instability through segmental and whole chromosomal loss (Adikusuma et al., 2018; Zuccaro et al., 2020; Alanis-Lobato et al., 2021; Papathanasiou et al., 2021). Unrepaired double-stranded DNA breaks at the CRISPR-Cas9 cleavage site have shown to result in fragmented chromosomes leading to chromosome mis-segregation and micronuclei formation in human cell lines and mouse embryos (Leibowitz et al., 2021; Papathanasiou et al., 2021). In human cleavage stage embryos, failure to replicate the genome before entry into mitotic divisions contributes to poorer embryo quality due to aneuploidy associated with chromosomal fragmentation and the formation of a micronucleus (Palmerola et al., 2022). Human and NHP embryos naturally have a higher incidence of aneuploidy where

TABLE 4 ***De novo*** **structural variants at predicted off-target sites identified in individual blastomeres by WGS.**

Gene	Blastomere	Variant type	Position	Size (kb)
NFASC	4-3	deletion	60,461,014–144,877,317	84,416.3
	4-4	inversion	29,288,909–135,582,062	103,293.2
		inversion	40,415,565–127,456,502	87,040.9
	5-3	inversion	44,728,954–174,674,237	129,945.3
	5-4	inversion	40,415,565–127,456,502	87,040.9
	5-7	deletion	63,036,915–113,225,838	50,188.9
	5-8	duplication	32,797,257–199,946,823	167,149.6
		deletion	60,460,921–144,877,320	84,416.4
SFMBT2	4-3	duplication	4,656,592–131,420,663	126,764.1
		duplication	631,150–44,472,044	43,840.9
		deletion	6,341,129–44,439,744	38,098.6
LIPC	4-1	deletion	35,664,439–35,664,648	0.2
	4-3	inversion	18,877,162–149,973,064	131,095.9
	4-4	inversion	22,447,357–70,039,192	47,591.8
	4-5	inversion	29,303,656–142,197,757	112,894.1
	4-6	duplication	1,727,262–94,864,685	93,137.4
		inversion	29,303,656–142,197,757	112,894.1
	5-1	deletion	27,706,507–56,330,973	28,624.5
	5-5	duplication	9,799,825–73,825,441	64,025.6
	5-7	inversion	18,272,609–49,503,847	31,231.2
	5-9	inversion	1,842,595–98,494,590	96,652.0

partial or whole chromosomes encapsulated by micronuclei may be present in fragmented cells of the embryos (Daughtry et al., 2019; Palmerola et al., 2022). Cellular fragmentation was previously observed in >65% of *in vitro* fertilized cleavage stage rhesus macaque embryos, and when analyzing individual blastomeres from 50 embryos, 73.5% showed whole and/or partial chromosomal losses or gains (Daughtry et al., 2019). The incidence of chromothripsis was not assessed in the present study given the limitations in interpreting WGA artifacts in light of an incomplete reference genome, in addition to the difficulty in discerning whether chromosome loss was due to CRISPR-Cas9 targeting or naturally-occurring cellular fragmentation events. As CRISPR-Cas9 editing errors can result in chromosomal disruption and/or elimination (Leibowitz et al., 2021; Papathanasiou et al., 2021) it is plausible that the embryo development may be negatively impacted and could explain the poor embryo transfer rate observed in our previous study (Schmidt et al., 2020) and in general the low live birth rates of CRISPR-Cas9 targeted NHPs (Schmidt et al., 2022a; Schmidt et al., 2022b).

4.1 Limitations of the study

Single-cell WGA can introduce amplification bias potentially limiting the interpretation of the variants identified in this study. Several studies comparing commercially available sc-WGA kits revealed differences in reproducibility, error rates, target coverage, read depth distribution and allele drop out, however, the REPLI-G sc-WGA kit used in the current study was shown to have a high mapping rage (>90%), be reproducible and have a lower error rate (Borgström et al., 2017; Biezuner et al., 2021) compared to other kits. In the present study, sequence coverage was variable across chromosomes within and across individual blastomeres compared to the parental DNA that did not undergo WGA. A limitation to the present study is the lack of analysis of unmanipulated control blastomeres to assess the rate of errors or artifacts incurred due to WGA. Translocation events were not called for this reason and only structural variants identified by two callers were reported in this study. While it is uncertain whether the structural variants identified

by WGS are due to CRISPR-Cas9 targeting, the variants were detected in pairs of blastomeres that were processed through independent WGA reactions suggesting that the mutation occurred during an early cleavage division or that similar sites of the genome are reproducibly prone to amplification errors during the WGA process. Recent advances in WGA technology such as primary template-directed amplification (Gonzalez-Pena et al., 2021) or linear amplification through transposon insertion (Chen et al., 2017) have shown to amplify single-cell genomes with more uniformity and accuracy. While these technologies are not incorporated into a commercial kit, they could be implemented in future single blastomere WGS studies for greater accuracy and reproducibility.

5 Conclusion

Overall, utilizing a WGS approach to determine CRISPR-Cas9 editing outcomes allows for the identification of edits not identified by PCR. In this study, WGS revealed the incidence of on-target large-scale deletions and INDEL formation at off-target sites. Imprecise editing could hinder the development of an NHP disease model that both genocopies and phenocopies the disease. The consequences of undesired editing events on gene expression of off-target and/or neighboring genes was not evaluated in this study, but should be considered in future studies. Based on evidence from human and mouse embryonic targeting by CRISPR-Cas9, it is likely that chromosomal damage incurred early in embryo development could negatively impact embryo viability. It remains unclear whether a reduced concentration or volume of the RNP delivered to the one-cell embryo would have a dose-dependent impact on on-target errors. Additional studies are needed to optimize embryonic editing by wild-type Cas9 and/or to use alternative next-generation Cas9 nucleases that do not create a double-stranded DNA break (Komor et al., 2016; Zeballos and Gaj, 2020). Regardless, WGS analysis should be implemented to thoroughly characterize editing genotypes in NHP models generated through this technology.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA880597.

Ethics statement

The animal study was reviewed and approved by the University of Wisconsin-Madison College of Letters and Sciences and Vice Chancellor's Office for Research and Graduate Education Institutional Animal Care and Use Committee.

Author contributions

JS, NS, TG, and IS contributed to the conception and design of the study. NS, YK, SG, and JS collected and analyzed the data. DP performed the bioinformatic analysis. JS, TG, and IS drafted the manuscript. The publication's contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Funding

This research was funded by NIH grants R24OD021322 awarded to IS and TG, P51OD011106 awarded to the Wisconsin National Primate Research Center, University of Wisconsin-Madison, and K99 HD099154-01 awarded to JS. This research was conducted at a facility constructed with support from Research Facilities Improvement Program grant numbers RR15459-01 and RR020141-01.

Acknowledgments

The authors would like to thank Dr. Jeffrey Rogers for his guidance in the analysis and for his insightful review of the manuscript. We would also like to acknowledge Logan Keding for his contribution to the graphic design of Figure 1.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2022.1031275/full#supplementary-material

References

Abbott, D., Rogers, J., Dumesic, D., and Levine, J. (2019). Naturally occurring and experimentally induced rhesus macaque models for polycystic ovary syndrome: Translational gateways to clinical application. *Med. Sci. (Basel, Switz.* 7, 107. doi:10.3390/medsci7120107

Adikusuma, F., Adikusuma, F., Piltz, S., and Corbett, M. A. (2018). Large deletions induced by Cas9 cleavage. *Nature* 560, E8–E9. doi:10.1038/s41586-018-0380-z

Alanis-Lobato, G., Zohren, J., Mccarthy, A., Fogarty, N., Kubikova, N., Hardman, E., et al. (2021). Frequent loss of heterozygosity in CRISPR-Cas9-edited early human embryos. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2004832117. doi:10.1073/pnas.2004832117

Allers, K., Hütter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E., et al. (2011). Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. *Blood* 117, 2791–2799. doi:10.1182/blood-2010-09-309591

Atkins, A., Chung, C., Allen, A., Dampier, W., Gurrola, T., Sariyer, I., et al. (2021). Off-target analysis in gene editing and applications for clinical translation of CRISPR/Cas9 in HIV-1 therapy. *Front. genome Ed.* 3, 673022. doi:10.3389/fgeed.2021.673022

Bae, S., Park, J., and Kim, J. (2014). Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinforma. Oxf. Engl.* 30, 1473–1475. doi:10.1093/bioinformatics/btu048

Biezuner, T., Raz, O., Amir, S., Milo, L., Ada, R., Fried, Y., et al. (2021). Comparison of seven single cell whole genome amplification commercial kits using targeted sequencing. *Sci. Rep.* 11, 17171. doi:10.1038/s41598-021-96045-9

Borgström, E., Paterlini, M., Mold, J., Frisen, J., and Lundeberg, J. (2017). Comparison of whole genome amplification techniques for human single cell exome sequencing. *PloS one* 12, e0171566. doi:10.1371/journal.pone.0171566

Chaudhari, H., Penterman, J., Whitton, H., Spencer, S., Flanagan, N., Lei Zhang, M., et al. (2020). Evaluation of homology-independent CRISPR-cas9 off-target assessment methods. *CRISPR J.* 3, 440–453. doi:10.1089/crispr.2020.0053

Chen, C., Xing, D., Tan, L., Li, H., Zhou, G., Huang, L., et al. (2017). Single-cell whole-genome analyses by linear amplification via transposon insertion (LIANTI). *Sci. (New York, N.Y.)* 356, 189–194. doi:10.1126/science.aak9787

Chen, Y., Zheng, Y., Kang, Y., Yang, W., Niu, Y., Guo, X., et al. (2015). Functional disruption of the dystrophin gene in rhesus monkey using CRISPR/Cas9. *Hum. Mol. Genet.* 24, 3764–3774. doi:10.1093/hmg/ddv120

Chen, Z., Kwon, D., Jin, Z., Monard, S., Telfer, P., Jones, M., et al. (1998). Natural infection of a homozygous delta24 CCR5 red-capped mangabey with an R2b-tropic simian immunodeficiency virus. *J. Exp. Med.* 188, 2057–2065. doi:10.1084/jem.188.11.2057

Cromer, M., Majeti, K., Rettig, G., Murugan, K., Kurgan, G., Hampton, J., et al. (2022). *Comparative analysis of CRISPR off-target activity discovery tools following* ex vivo *editing of CD34⁺ hematopoietic stem and progenitor cells*. bioRxiv.

Cui, Y., Niu, Y., Zhou, J., Chen, Y., Cheng, Y., Li, S., et al. (2018). Generation of a precise Oct4-hrGFP knockin cynomolgus monkey model via CRISPR/Cas9-assisted homologous recombination. *Cell Res.* 28, 383–386. doi:10.1038/cr.2018.10

Cummins, N., Rizza, S., Litzow, M., Hua, S., Lee, G., Einkauf, K., et al. (2017). Extensive virologic and immunologic characterization in an HIV-infected individual following allogeneic stem cell transplant and analytic cessation of antiretroviral therapy: A case study. *PLoS Med.* 14, e1002461. doi:10.1371/journal.pmed.1002461

D'Souza, S., Kumar, A., Weinfurter, J., Park, M., Maufort, J., Tao, L., et al. (2022). Generation of SIV-resistant T cells and macrophages from nonhuman primate induced pluripotent stem cells with edited CCR5 locus. *Stem Cell Rep.* 17, 953–963. doi:10.1016/j.stemcr.2022.03.003

Daughtry, B., Rosenkrantz, J., Lazar, N., Fei, S., Redmayne, N., Torkenczy, K., et al. (2019). Single-cell sequencing of primate preimplantation embryos reveals chromosome elimination via cellular fragmentation and blastomere exclusion. *Genome Res.* 29, 367–382. doi:10.1101/gr.239830.118

Dean, M., Carrington, M., Winkler, C., Huttley, G., Smith, M., Allikmets, R., et al. (1996). Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia growth and development study, multicenter AIDS cohort study, multicenter hemophilia cohort study, san francisco city cohort, ALIVE study. *Sci. (New York, N.Y.)* 273, 1856–1862. doi:10.1126/science.273.5283.1856

Dray, B., Raveendran, M., Harris, R., Benavides, F., Gray, S., Perez, C., et al. (2018). Mismatch repair gene mutations lead to lynch syndrome colorectal cancer in rhesus macaques. *Genes & cancer* 9, 142–152. doi:10.18632/genesandcancer.170

Gonzalez-Pena, V., Natarajan, S., Xia, Y., Klein, D., Carter, R., Pang, Y., et al. (2021). Accurate genomic variant detection in single cells with primary template-directed amplification. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2024176118. doi:10.1073/pnas.2024176118

Henrich, T., Hanhauser, E., Marty, F., Sirignano, M., Keating, S., Lee, T., et al. (2014). Antiretroviral-free HIV-1 remission and viral rebound after allogeneic stem cell transplantation: Report of 2 cases. *Ann. Intern. Med.* 161, 319–327. doi:10.7326/M14-1027

Huang, X., Yang, D., Zhang, J., Xu, J., and Chen, Y. (2022). Recent advances in improving gene-editing specificity through CRISPR-cas9 nuclease engineering. *Cells* 11, 2186. doi:10.3390/cells11142186

Hütter, G., Nowak, D., Mossner, M., Ganepola, S., Müssig, A., Allers, K., et al. (2009). Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N. Engl. J. Med.* 360, 692–698. doi:10.1056/NEJMoa0802905

Jiang, H., Lei, R., Ding, S. W., and Zhu, S. (2014). Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinforma.* 15, 182. doi:10.1186/1471-2105-15-182

Kang, H., Minder, P., Park, M., Mequitta, W., Torbett, B., and Slukvin, I. I. (2015). CCR5 disruption in induced pluripotent stem cells using CRISPR/Cas9 provides selective resistance of immune cells to CCR5-tropic HIV-1 virus. *Mol. Ther. Nucleic Acids* 4, e268. doi:10.1038/mtna.2015.42

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi:10.1038/nature17946

Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 765–771. doi:10.1038/nbt.4192

Lamas-Toranzo, I., Galiano-Cogolludo, B., Cornudella-Ardiaca, F., Cobos-Figueroa, J., Ousinde, O., and Bermejo-Álvarez, P. (2019). Strategies to reduce genetic mosaicism following CRISPR-mediated genome edition in bovine embryos. *Sci. Rep.* 9, 14900. doi:10.1038/s41598-019-51366-8

Leibowitz, M. L., Papathanasiou, S., Doerfler, P., Blaine, L., Sun, L., Yao, Y., et al. (2021). Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat. Genet.* 53, 895–905. doi:10.1038/s41588-021-00838-7

Liu, R., Paxton, W., Choe, S., Ceradini, D., Martin, S., Horuk, R., et al. (1996). Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 86, 367–377. doi:10.1016/s0092-8674(00)80110-5

Luo, X., He, Y., Zhang, C., He, X., Yan, L., Li, M., et al. (2019). Trio deep-sequencing does not reveal unexpected off-target and on-target mutations in Cas9-edited rhesus monkeys. *Nat. Commun.* 10, 5525. doi:10.1038/s41467-019-13481-y

Midic, U., Hung, P. H., Vincent, K. A., Goheen, B., Schupp, P. G., Chen, D. D., et al. (2017). Quantitative assessment of timing, efficiency, specificity and genetic mosaicism of CRISPR/Cas9-mediated gene editing of hemoglobin beta gene in rhesus monkey embryos. *Hum. Mol. Genet.* 26, 2678–2689. doi:10.1093/hmg/ddx154

Moshiri, A., Chen, R., Kim, S., Harris, R. A., Li, Y. M., Raveendran, M., et al. (2019). A nonhuman primate model of inherited retinal disease. *J. Clin. Investigation* 129, 863–874. doi:10.1172/JCI123980

Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., et al. (2014). Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell* 156, 836–843. doi:10.1016/j.cell.2014.01.027

Ozirmak, L. N., Gray, S., Bowen, C., Reyes-Uribe, L., Dray, B., Deng, N., et al. (2022). Comparative molecular genomic analyses of a spontaneous rhesus macaque model of mismatch repair-deficient colorectal cancer. *PLoS Genet.* 18, e1010163. doi:10.1371/journal.pgen.1010163

Palmerola, K., Amrane, S., De Los Angeles, A., Xu, S., Wang, N., De Pinho, J., et al. (2022). Replication stress impairs chromosome segregation and preimplantation development in human embryos. *Cell* 185, 2988–3007.e20. doi:10.1016/j.cell.2022.06.028

Papathanasiou, S., Markoulaki, S., Blaine, L., Leibowitz, M., Zhang, C., Jaenisch, R., et al. (2021). Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nat. Commun.* 12, 5855. doi:10.1038/s41467-021-26097-y

Pattanayak, V., Lin, S., Guilinger, J., Ma, E., Doudna, J., and Liu, D. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* 31, 839–843. doi:10.1038/nbt.2673

Samson, M., Libert, F., Doranz, B., Rucker, J., Liesnard, C., Farber, C., et al. (1996). Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 382, 722–725. doi:10.1038/382722a0

Schmidt, J., Jones, K., Van Vleck, T., and Emborg, M. (2022a). Modeling genetic diseases in nonhuman primates through embryonic and germline modification: Considerations and challenges. *Sci. Transl. Med.* 14, eabf4879. doi:10.1126/scitranslmed.abf4879

Schmidt, J. K., Strelchenko, N., Park, M. A., Kim, Y. H., Mean, K. D., Schotzko, M. L., et al. (2020). Genome editing of CCR5 by CRISPR-Cas9 in Mauritian cynomolgus macaque embryos. *Sci. Rep.* 10, 18457. doi:10.1038/s41598-020-75295-z

Schmidt, J., Reynolds, M., Golos, T., and Slukvin, I. (2022b). CRISPR/Cas9 genome editing to create nonhuman primate models for studying stem cell therapies for HIV infection. *Retrovirology* 19, 17. doi:10.1186/s12977-022-00604-5

Tapmeier, T., Rahmioglu, N., Lin, J., De Leo, B., Obendorf, M., Raveendran, M., et al. (2021). Neuropeptide S receptor 1 is a nonhormonal treatment target in endometriosis. *Sci. Transl. Med.* 13, eabd6469. doi:10.1126/scitranslmed.abd6469

Tsukiyama, T., Kobayashi, K., Nakaya, M., Iwatani, C., Seita, Y., Tsuchiya, H., et al. (2019). Monkeys mutant for PKD1 recapitulate human autosomal dominant polycystic kidney disease. *Nat. Commun.* 10, 5517. doi:10.1038/s41467-019-13398-6

Wang, S., Ren, S., Bai, R., Xiao, P., Zhou, Q., Zhou, Y., et al. (2018). No off-target mutations in functional genome regions of a CRISPR/Cas9-generated monkey model of muscular dystrophy. *J. Biol. Chem.* 293, 11654–11658. doi:10.1074/jbc.AC118.004404

Yang, W., Liu, Y., Tu, Z., Xiao, C., Yan, S., Ma, X., et al. (2019). CRISPR/Cas9-mediated PINK1 deletion leads to neurodegeneration in rhesus monkeys. *Cell Res.* 29, 334–336. doi:10.1038/s41422-019-0142-y

Zarate, S., Carrol, A., Mahmoud, M., Krasheninina, O., Jun, G., Salerno, W. J., et al. (2020). Parliament2: Accurate structural variant calling at scale. *GigaScience* 9, giaa145. giaa145. doi:10.1093/gigascience/giaa145

Zeballos, M. A., and Gaj, T. (2020). Next-generation CRISPR technologies and their applications in gene and cell therapy. *Trends Biotechnol.* 39, 692–705.

Zhao, H., Tu, Z., Xu, H., Yan, S., Yan, H., Zheng, Y., et al. (2017). Altered neurogenesis and disrupted expression of synaptic proteins in prefrontal cortex of SHANK3-deficient non-human primate. *Cell Res.* 27, 1293–1297. doi:10.1038/cr.2017.95

Zhou, Y., Sharma, J., Ke, Q., Landman, R., Yuan, J., Chen, H., et al. (2019). Atypical behavior and connectivity in *SHANK3*-mutant macaques. *Nature* 570, 326–331. doi:10.1038/s41586-019-1278-0

Zuccaro, M. V., Xu, J., Mitchell, C., Marin, D., Zimmerman, R., Rana, B., et al. (2020). Allele-specific chromosome removal after Cas9 cleavage in human embryos. *Cell* 183, 1650–1664. e15. doi:10.1016/j.cell.2020.10.025

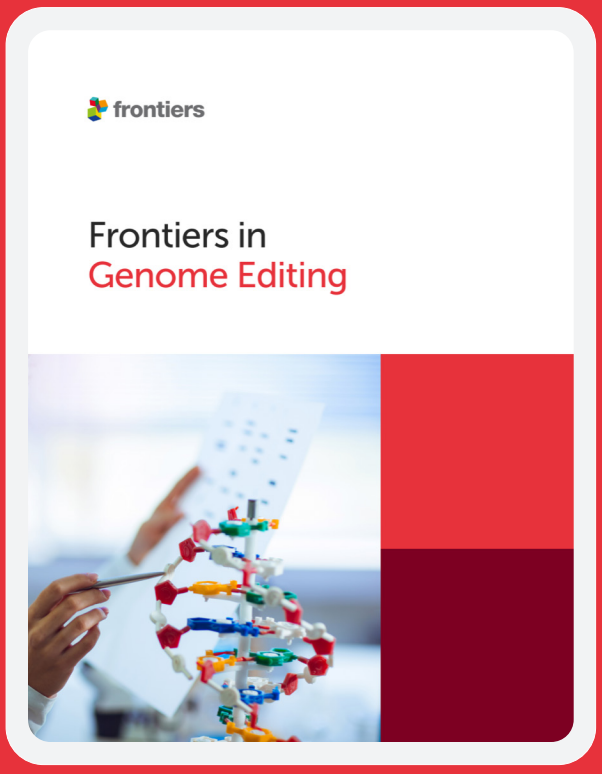
Frontiers in
Genome Editing

Explores the applications of genome and epigenome editing technologies

A multidisciplinary journal that explores cutting-edge research on existing and emerging genome editing technologies and their implications across wide-ranging areas of research - from human health to agriculture.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact