Universiteit Leiden
The Netherlands

Shepherding precision gene editing with CRISPR-Cas9 variants and adenoviral vectors

Tasca, F.

Citation

Tasca, F. (2023, June 15). *Shepherding precision gene editing with CRISPR-Cas9 variants and adenoviral vectors*. Retrieved from https://hdl.handle.net/1887/3620378

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3620378

Note: To cite this publication please use the final published version (if applicable).

Shepherding precision gene editing with CRISPR-Cas9 variants and adenoviral vectors

Francesca Tasca

Shepherding precision gene editing with CRISPR-Cas9 variants and adenoviral vectors
Francesca Tasca

Cover: Modified immunofluorescence picture of patient-derived myotube showing the rescue of dystrophin expression upon treatment, taken by Francesca Tasca. Technical details on the immunostaining can be found in Chapter 3.

Layout:Francesca Tasca

Printing: GildePrint

The research presented in this thesis was performed at the Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands. This research was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 765269 (IMGENE - Improving Genome Editing Efficiency), the Prinses Beatrix Spierfonds (W.OR16-13 and W.OR21-01), the China Scholarship Council-Leiden University Joint Scholarship Programme, and the DutchDuchenne Parent Project (17.012).

Copyright © Francesca Tasca, 2023, Leiden, the Netherlands. All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means without written permission of the copyright owner.

Shepherding precision gene editing with CRISPR-Cas9 variants and adenoviral vectors

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 15 juni 2023
klokke 13.45 uur

door

Francesca Tasca

geboren te Marostica, Italie
in 1993

Promotor: Prof. Dr. R.C. Hoeben
Co-promotor: Dr. M.A.F.V. Gonçalves

Leden Promotiecommissie:
Prof. Dr. A.M. Aartsma-Rus
Prof. Dr. M.J.T.H. Goumans
Prof. Dr. N. Geijsen
Prof. Dr. F. Larcher (Universidad Carlos III de Madrid, Spain)
Prof. Dr. J.G. Mikkelsen (Aarhus University Department of Biomedicine, Denmark)
Prof. Dr. E. Mastrobattista (University of Utrecht, Netherlands)

Contents

Chapter 1 General Introduction

Chapter 2 Adenoviral vectors meet gene editing: a rising partnership for the genomic engineering of human stem cells and their progeny

Chapter 3 High-capacity adenovector delivery of forced CRISPR-Cas9 heterodimers fosters precise chromosomal deletions in human cells

Chapter 4 Large-scale genome editing based on high-capacity adenovector particles and CRISPR-Cas9 nucleases rescues full-length dystrophin synthesis in DMD muscle cells

Chapter 5 Expanding the editable genome and CRISPR-Cas9 versatility using DNA cutting-free gene targeting based on *in trans* paired nicking

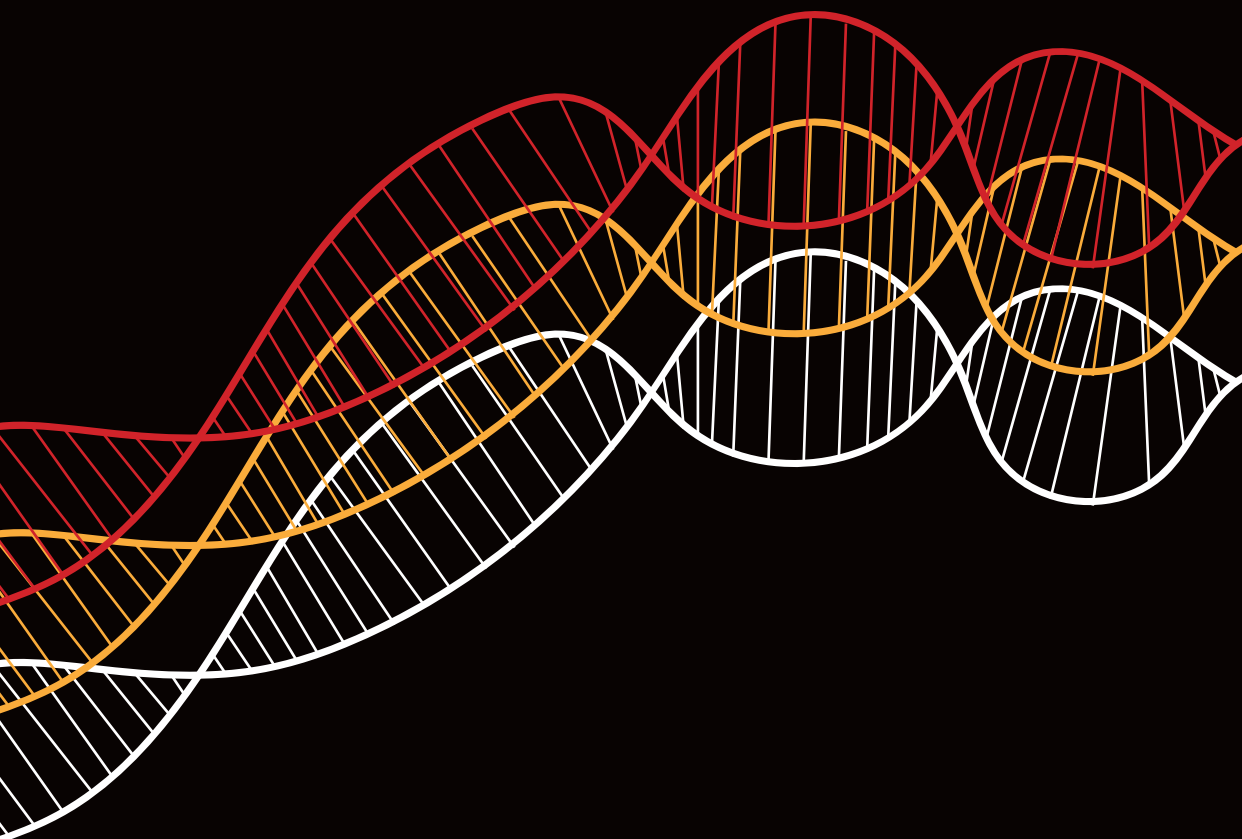
Chapter 6 Conclusions and Final remarks

Chapter 7 Nederlandse Samenvatting

Curriculum Vitae

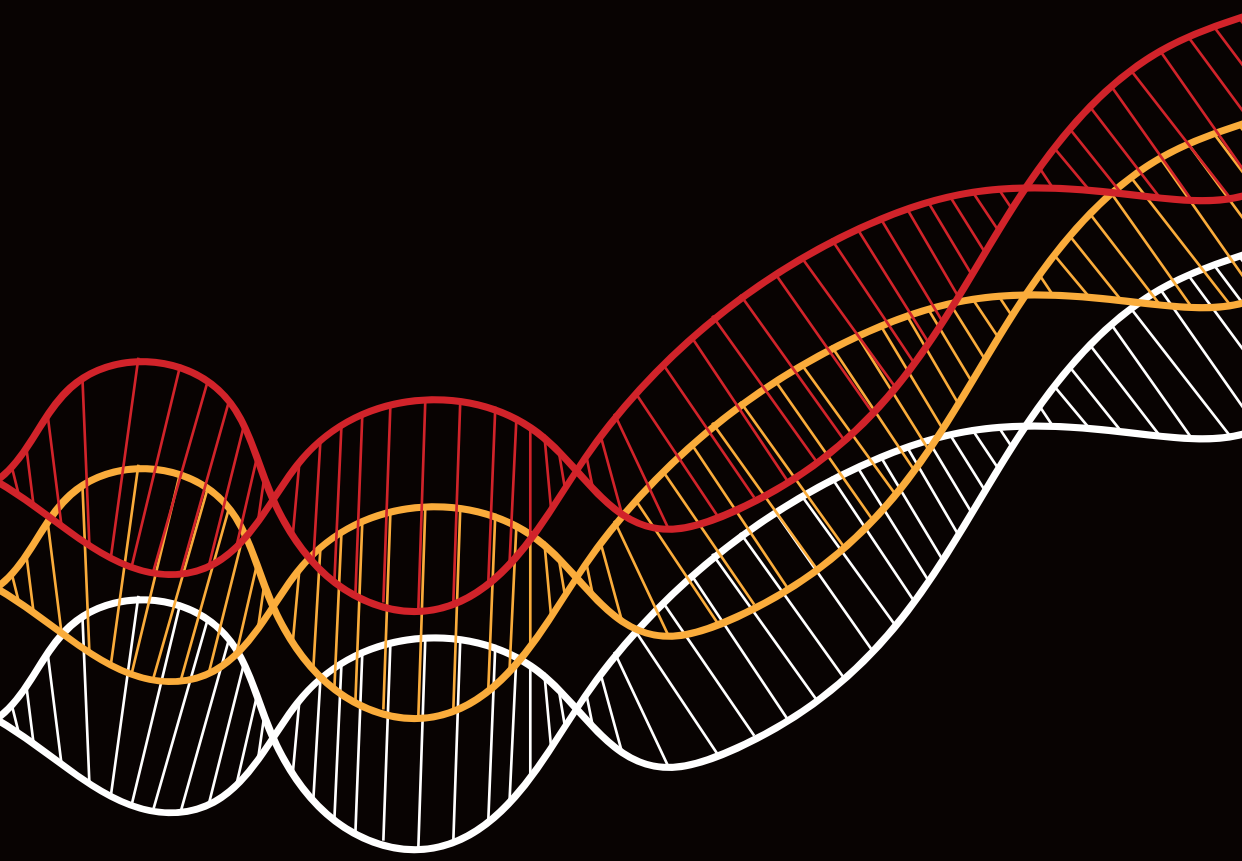
List of Publications

Acknowledgements



Chapter 1

General Introduction



General Introduction

The past two decades have seen the growing development and consequent vast application of next-generation genome editing (GE) tools in fundamental and applied research. Nowadays GE based on RNA-guided nucleases (RGNs) (e.g., engineered CRISPR-Cas9 nucleases) are the most common tools for targeted genetic modification of eukaryotic cells. Normally, RGN-induced double-stranded DNA breaks (DSBs) are employed with or without donor DNA constructs to obtain specific GE goals through targeted DSB repair involving homology-directed repair (HDR) or non-homologous end joining (NHEJ) pathways, respectively. Nevertheless, GE technologies are in need of increased efficiency and accuracy, especially looking forward to translation into diverse clinical applications. The work presented in this thesis focuses on investigating new principles for improving the efficiency and accuracy of GE, particularly in cells with high therapeutic potential, such as induced pluripotent stem cells (iPSCs). In particular, these principles entailed (i) modifying the structure of donor HDR substrates and RGNs, and (ii) integrating third-generation adenoviral vector and RGN technologies. **Chapter 2** provides an updated review on the initial efforts and recent progress in the field of GE and adenoviral vector systems. Moreover, **Chapter 2** outlines the use of the latter gene delivery tools for GE of human stem cells and their progeny towards the fulfilment of their therapeutic potential.

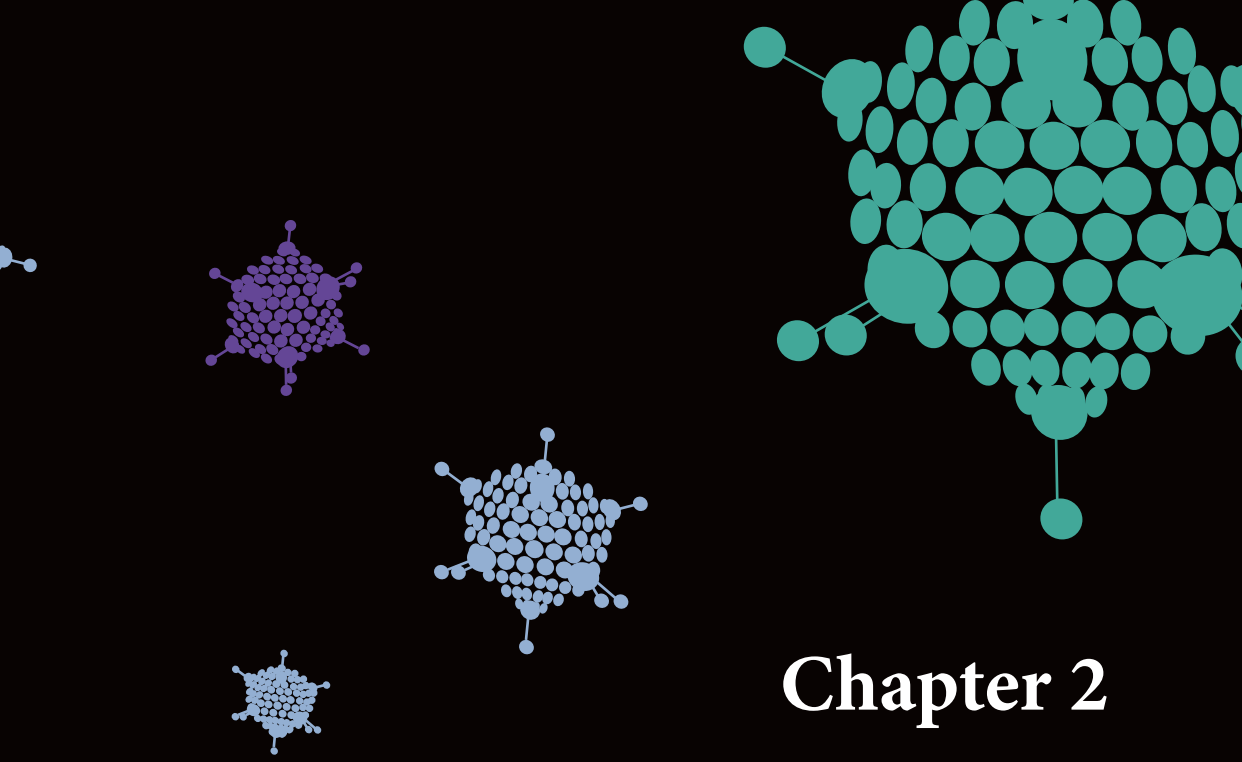
Despite noticeable advances in the GE field, described in **Chapter 2**, several hurdles still hinder its full potential. The work described in **Chapter 3** and **Chapter 4** is based on employing fully viral gene-deleted adenoviral vectors (*aliases* third-generation or high-capacity adenoviral vectors) as delivery agents of modified donor HDR substrates and/or RGNs whose combined effects result in improved efficiency and specificity of NHEJ- and HDR-mediated GE strategies (**Chapter 3** and **Chapter 4**, respectively). In **Chapter 3**, it is established that Cas9 heterodimer fusion proteins achieve precise deletion of specific chromosomal DNA stretches in a more controlled manner than conventional, independently acting, Cas9 monomers. With the strategy developed in **Chapter 4** it was instead possible to obtain large-scale (up to 14.6 kb) precise gene knock-ins in hard-to-transfect stem and progenitor cells when using a third-generation adenoviral vector system for the combined delivery of donor DNA templates and high-specificity CRISPR-Cas9 nucleases. Together, **Chapter 3** and **Chapter 4**, highlight the versatility of converting third-generation adenoviral vectors into delivery agents of GE tools and applying these in diverse therapeutically relevant cell types, i.e., iPSCs and myoblasts. The insights derived from these studies were obtained in the context of research directed at treating Duchenne muscular dystrophy (DMD).

DMD is a progressive muscle-wasting disease caused by mutations in the vast (i.e., circa 2.4-Mb) dystrophin-encoding *DMD* gene. In most cases, DMD-directed gene therapies aim at the production of partially functional dystrophins, either through mutation-specific DMD repair strategies resulting in the expression of Becker-like dystrophins, or the overexpression of microdystrophins. In **Chapter 3**, the investigated NHEJ-mediated multiplexing GE

strategy, based on Cas9 heterodimer fusion proteins and coordinated action of the resulting RGN pairs, is applied to repair defective DMD alleles resulting in the expression of Becker-like dystrophins. In **Chapter 4**, stable expression of fully functional dystrophin molecules upon targeted insertion of full-length dystrophin expression units into a "safe harbor" chromosomal locus is instead investigated to achieve complementation of DMD-causing mutations regardless of their type or location.

Finally, **Chapter 5** stresses the need to further enhance the specificity and fidelity of GE procedures and explores the benefits of progressing towards DSB-free GE strategies. The work detailed in this chapter focuses on the use of a single-strand DNA break (SSB)-based GE strategy (i.e., *in trans* paired nicking) to edit particularly sensitive genomic regions and cells. This SSB-based GE strategy proved to be successful in making seamless edits and potentially expands the editable genome to tracts previously not possible to modify effectively due to their repetitive nature and/or essentiality for proper cell function or viability.

All together the work presented in this thesis broadens the horizon of possible GE applications, including those directed at gene therapies, by investigating the feasibility of using adenoviral vectors to test novel GE approaches and by exploring the utility of an emerging DSB-free GE strategy with a seamless and scarless character.



Chapter 2

Adenoviral vectors meet gene editing: a rising partnership for the genomic engineering of human stem cells and their progeny

Francesca Tasca*[1], Qian Wang*[1], Manuel A.F.V. Gonçalves[1]

Cells, 2020, **9**(4): 953

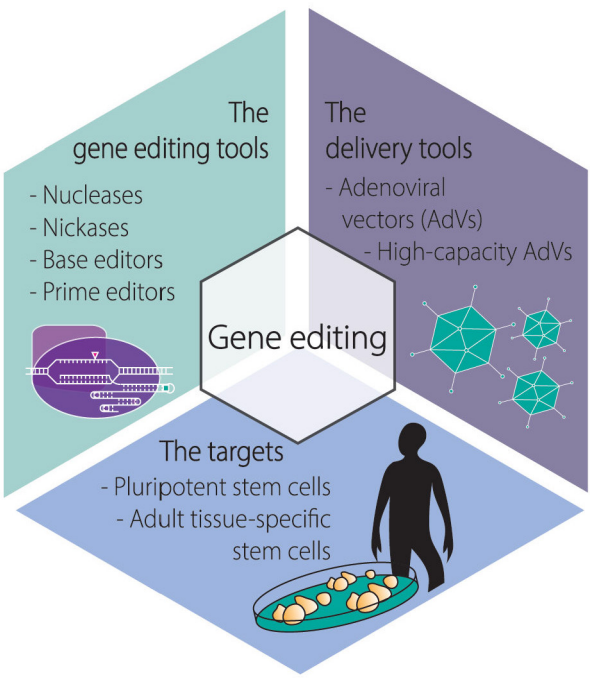
* These authors contributed equally to this work

[1]Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the

Abstract

Gene editing permits changing specific DNA sequences within the vast genomes of human cells. Stem cells are particularly attractive targets for gene editing interventions as their self-renewal and differentiation capabilities consent studying cellular differentiation processes, screening small-molecule drugs, modeling human disorders, and testing regenerative medicines. To integrate gene editing and stem cell technologies, there is a critical need for achieving efficient delivery of the necessary molecular tools in the form of programmable DNA-targeting enzymes and/or exogenous nucleic acid templates. Moreover, the impact that the delivery agents themselves have on the performance and precision of gene editing procedures is yet another critical parameter to consider. Viral vectors consisting of recombinant replication-defective viruses are under intense investigation for bringing about efficient gene-editing tool delivery and precise gene-editing in human cells. In this review, we focus on the growing role that adenoviral vectors are playing in the targeted genetic manipulation of human stem cells, progenitor cells, and their differentiated progenies in the context of in vitro and ex vivo protocols. As preamble, we provide an overview on the main gene editing principles and adenoviral vector platforms and end by discussing the possibilities ahead resulting from leveraging adenoviral vector, gene editing, and stem cell technologies.

Graphical abstract



1. Introduction

1.1. The Main Gene Editing Principles Based on Programmable Nucleases and Their Key Pros and Cons

Commonly, gene editing is triggered after programmable nucleolytic enzymes bind to predefined chromosomal sequences and locally generate double-stranded or single-stranded DNA breaks (DSBs or SSBs, respectively). The ensuing mending of these chromosomal breaks by cellular DNA repair mechanisms leads to the installation of targeted genomic changes whose extent can span from single to thousands of base pairs (bps).

Gene editing endeavors can disable a coding sequence (knockout) or remove specific genomic tracts. Moreover, they can equally restore a coding sequence or insert into specific genomic locations new genetic information (knock-in) present in exogenous (donor) DNA molecules. Typically, DNA editing strategies that knock-out or restore endogenous coding sequences involve the transfer of programmable nucleases that generate frameshifting insertions and deletions (indels) after the repair of targeted DSBs by non-homologous end joining (NHEJ) pathways. These include, classic NHEJ (cNHEJ) and alternative NHEJ (alt-NHEJ) pathways such as microhomology-mediated end-joining (MMEJ) and single-strand annealing (SSA) [1]. The cNHEJ is the most active and fast-acting of the DNA repair pathways in mammalian cells often resulting in no or limited end-processing by exonucleases prior to ligation of chromosomal ends [1]. Importantly, chromosomal ligation products containing indels can be generated [1], especially in the presence of a programmable nuclease that re-cleaves precisely ligated products until an indel disrupts its target site and becomes "fixed" in the cell population. It is also noteworthy mentioning that; (i) the target site sequences, (ii) the class of programmable nuclease employed, and (iii) the type of repair mechanism engaged in DSB repair, all contribute to different indel profiles which vary considerably in length and nucleotide composition [1, 2]. Yet, depending to some extent on microhomologies, the targeting of specific sequences by a programmable nuclease can yield specific indels in a high frequency of modified alleles [3–7].

Indels resulting from NHEJ-mediated repair of targeted DSBs can be exploited for disrupting non-coding elements (e.g., splicing motifs to induce exon-skipping) or reframing coding sequences that rescue endogenous gene expression via bypassing preexisting nonsense mutations (i.e., premature stop codons) [8, 9]. Alternatively, indels can be exploited for disrupting coding sequences that knockout endogenous gene expression via installing stop codons that induce nonsense-mediated mRNA decay (NMD) [9–11]. However, it is important to mention recent research demonstrating the existence of an evolutionary conserved NMD-dependent mechanism in which the presence of a nonsense mutation in a gene can activate transcription of related genes whose products functionally complement the mutant gene [12, 13]. Another cautionary note concerns other recent findings in which DSB-derived indels in coding sequences can generate transcripts yielding various types of aberrant gene products

[14]. Therefore, these recently characterized processes, involving either genetic compensation responses triggered by indel-derived nonsense mutations or indels as such, have the potential of hindering the creation of robust gene knockout phenotypes and predictable gene editing outcomes. For a more thorough and predictable removal of pre-existing genetic information, so-called multiplexing gene editing approaches can be deployed instead. In this case, two different programmable nucleases work in concert to generate a pair of intrachromosomal DSBs that lead to the excision of the intervening DNA sequence after end-to-end NHEJ ligation of the chromosomal termini [9, 15–17]. Alternatively, two programmable nucleases designed for generating a pair of inter-chromosomal DSBs can direct the assembly of specific translocations to, for instance, confirm or study the involvement of these translocations in cellular transformation events and, ultimately, cancer emergence [9, 18].

Normally, knocking-in gene editing strategies encompass the delivery of programmable nucleases together with exogenous donor DNA that is inserted at the site-specific DSB via either homology-independent pathways (e.g., NHEJ) [19] or homology-directed DNA repair (HDR) [9–11]. Generally, HDR-mediated knock-ins are more precise than those resulting from homology-independent processes in that they lack extraneous footprints at the border between endogenous and exogenous DNA. Indeed, instead of direct exogenous-to-endogenous DNA ligations via NHEJ or MMEJ, whose junction products often contain differently sized indels or specific footprints, DSB repair through HDR is a higher fidelity process [1, 20]. This process involves genetic exchange between donor and target sequences and includes extensive exonucleolytic processing of chromosomal breaks, single-strand invasions, and DNA synthesis over DSB-repairing donor templates [20]. Ultimately, these molecular interactions result in accurate "copy-pasting" of the foreign genetic information into a specific locus [9–11]. Yet, HDR-mediated gene editing is normally less frequent than gene editing based on DNA repair mechanisms that are independent of large tracts of homology between target and donor DNA templates (e.g., cNHEJ and MMEJ). In fact, as aforementioned, cNHEJ is the main DSB repair mechanism in mammalian cells [1, 20]. Further contributing to the differences in knocking-in frequencies obtained through gene editing involving cNHEJ versus HDR is the fact that the former pathway is active throughout the cell cycle; whereas the latter is only operative during the S and late G2 phases, when normally sister chromatids are available as sources of endogenous DNA-repairing templates [1, 20]. For this reason, gene editing involving the recruitment of the HDR pathway is unsuitable in non-cycling cells, such as, quiescent human hematopoietic stem cells (hHSCs) and terminally differentiated cells. Another consideration concerns the steep decline in HDR-mediated gene editing frequencies as the length of the exogenous DNA increases and the extent of continuous homology between target and donor DNA decreases [21]. Therefore, the choice of the DSB repair pathway to exploit, and hence the designing of the DSB-repairing substrates to use, is contingent upon the specific application(s). For instance, knocking-in large genetic payloads into introns of safe harbor loci (e.g., *AAVS1* and *CCR5*) for achieving stable and homogeneous transgene expression in

cell populations may be best pursued via selecting HDR-independent gene editing strategies; whereas knocking-in donor DNA into coding sequences for modeling or repairing genetic defects in stem or progenitor cells is best accomplished through precise HDR-dependent gene editing.

1.2. The Main Programmable Nuclease Platforms and Their Key Pros and Cons

Under regular conditions, HDR-mediated gene knock-ins are very rare events in human cells, with typical frequencies varying between 10^{-6} and 10^{-7} [22–24]. The finding that site-specific DSBs made by homing endonucleases at chromosomally embedded recombinant sequences could stimulate HDR by several orders of magnitude, was a powerful stimulus for the development of programmable nucleases [25–27].

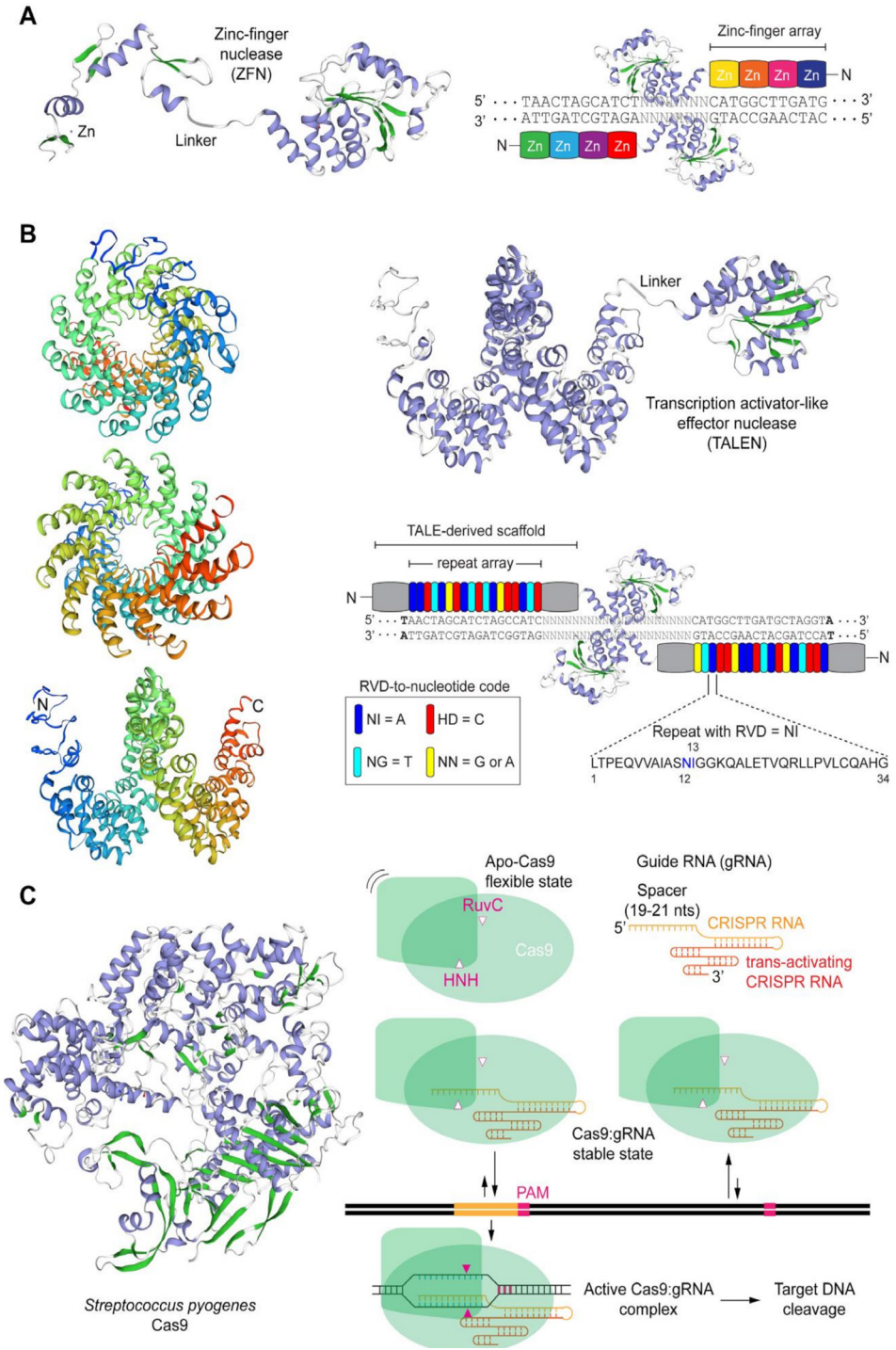
The crucial feature of programmable nucleases is their capability of binding to and cleaving at predefined DNA sequences, including those located within large genomes [9–11, 28]. Nowadays the main classes of programmable nucleases are, in chronological order of appearance, zinc-finger nucleases (ZFNs) [29], transcription activator-like effector (TALE) nucleases (TALENs) [30–34], and RNA-guided nucleases (RGNs) [35–38]. Naturally, the development of programmable nuclease technologies was invariably grounded on fundamental insights obtained from a broad range of biological systems, spanning from vertebrate cells and phytopathogenic bacteria, in the case of ZFNs [39] and TALENs [40, 41], respectively, to bacteria and archaea, in the case of RGNs [42, 43].

ZFNs and TALENs are modular proteins that present an overall similar architecture (**Figure 1A** and **1B**). In particular, they consist of a customizable DNA-binding domain fused through a flexible linker to a non-specific nuclease domain, typically that of the type IIS FokI restriction enzyme whose catalytic activity is dependent on dimerization [44]. Resulting from their comparable generic architectures, ZFNs and TALENs act in a similar fashion in that members of ZFN and TALEN pairs bind in close proximity to each other on opposite DNA strands of a bipartite target sequence leading to site-specific DSBs at the spacer region after local dimerization of the FokI nuclease domains (**Figure 1A** and **1B**). The DNA-binding domains of ZFNs and TALENs consist of arrays of engineered zinc-finger motifs and TALE repeats, respectively, with each zinc-finger motif usually binding to nucleotide triplets and each TALE repeat binding to single nucleotides within their respective double-stranded target sites (**Figure 1A** and **1B**). Cys_2-His_2 zinc-fingers are found in metazoans where they serve as motifs in RNA and DNA binding proteins whose wide roles include transcriptional and epigenetic regulation of target genes [45, 46]. Native TALE proteins are found in certain phytopathogenic bacteria (e.g., *Xanthomonas* sp.) where they serve as virulence factors once injected into host plant cells via type III secretory apparatuses [47]. The binding of zinc-finger motifs to specific triplets can be substantially affected by flanking nucleotides [48]. This sequence context dependency contributes to making highly specific ZFNs a laborious task requiring complex protein engineering methodologies that may include several rounds

of optimization and/or screening and selection of ZFN candidates from large zinc-finger libraries [48]. In contrast, the binding of TALE repeats to their cognate nucleotides does not seem to be substantially influenced by neighboring sequences [49]. This limited sequence context dependency aids the assembly of functional and highly specific TALENs whose designing flexibility and genomic space coverage is superior to that of ZFNs [49]. DNA binding of TALEs are, however, significantly hindered by cytosine methylation [50, 51] and Krüppel-associated box-induced heterochromatin [52]. Importantly, the former epigenetic modification can be elegantly surpassed by incorporating non-canonical TALE repeats within TALE arrays [51]. Native RGNs are found in many bacteria and archaea where they form adaptive immune systems against invading agents, e.g., bacteriophages and foreign plasmids [53]. Engineered RGNs, such as those based on the prototypic clustered regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated 9 (Cas9) system, from *Streptococcus pyogenes* [35–38], operate differently from ZFNs and TALENs in that target DNA cleavage does not depend exclusively on protein-DNA binding but also on RNA-DNA hybridization. In particular, RGNs, consisting of a sequence-specific single guide RNA (gRNA) coupled to an invariant nuclease, first recognize so-called protospacer adjacent motifs (PAMs) on the DNA via PAM-interacting domains in the nuclease component [10, 54]. In the case of the *S. pyogenes* Cas9 the PAM reads NGG. Typically, in instances in which the 19-21 deoxyribonucleotides "upstream" from the PAM are complementary to the 5' end the gRNA, DSB formation ensues through the concerted action of the HNH and RuvC-like nuclease domains of Cas9 (**Figure 1C**). The events leading to DSB formation upon initial Cas9-PAM interrogation include, PAM-proximal DNA unwinding, R-loop formation and expansion via increasing gRNA:DNA annealing which subsequently triggers HNH translocation and pairing with the RuvC-like domain. Ultimately, HNH-RuvC pairing catalyzes phosphodiester bond cleavage of both DNA chains, predominantly three base-pairs upstream from the PAM (**Figure 1C**) [10, 43, 55].

Crucially, RGNs can cut DNA at unintended genomic sequences (off-target sites) especially if mismatches between gRNA and DNA sequences locate at PAM-distal positions [56–60]. Furthermore, albeit to a lesser extent than NGG, *S. pyogenes* Cas9 can also effectively engage non-canonical PAMs (e.g., NAG), which further contributes to off-target activities [57, 60, 61]. Therefore, similarly to their programmable nuclease predecessors, the application of RGNs warrants careful assessment of potential off-target sites, especially if directed toward clinical testing. Indeed, judiciously chosen gRNAs can, *per se*, greatly reduce off-target activities in vitro and in vivo [62, 63]. As TALENs, targeted DNA cleavage by RGNs is also hindered to some extent by epigenetic mechanisms underpinning specific heterochromatic states [52, 64–66]. However, in contrast to TALENs, RGNs do not seem affected by DNA methylation [57].

The fact that readdressing RGNs to new target sites simply comprises modifying the 5' end of the gRNA component, and hence does not require de novo protein engineering as


A
Zinc-finger nuclease (ZFN)
Zn
Linker
Zinc-finger array
Zn Zn Zn Zn
N
5' · · · TAACTAGCATCTNNNNNNNCATGGCTTGATG · · · 3'
3' · · · ATTGATCGTAGANNNNNNNGTACCGAACTAC · · · 5'
N Zn Zn Zn Zn
B
Linker
Transcription activator-like effector nuclease (TALEN)
TALE-derived scaffold
repeat array
N
5' · · · TAACTAGCATCTAGCCATCNNNNNNNNNNNNNNNNNNNCATGGCTTGATGCTAGGTA · · · 3'
3' · · · ATTGATCGTAGATCGGTAGNNNNNNNNNNNNNNNNNNNGTACCGAACTACGATCCAT · · · 5'
N
RVD-to-nucleotide code
NI = A
HD = C
NG = T
NN = G or A
Repeat with RVD = NI
13
LTPEQVVAIASNIGGKQALETVQRLLPVLCQAHG
1
12
34
N
C
C
Apo-Cas9 flexible state
RuvC
Cas9
HNH
Guide RNA (gRNA)
Spacer (19-21 nts)
5'
CRISPR RNA
3'
trans-activating CRISPR RNA
Cas9:gRNA stable state
PAM
Active Cas9:gRNA complex
Target DNA cleavage
Streptococcus pyogenes Cas9

◂ **Figure 1. Schematics of the main programmable nuclease platforms.** (**A**) Zinc-finger nucleases (ZFNs). ZFNs are chimeric modular DNA-binding proteins consisting of the FokI nuclease domain fused through a flexible linker to an array of 3–6 artificial Cys_2-His_2 zinc-finger motifs. Each zinc-finger motif acquires its structure through tetrahedral coordination of 2 cysteines in β-sheets and 2 histidines in α-helixes by zinc ions. ZFN monomers of a working ZFN pair bind on opposite DNA strands in a tail-to-tail configuration leading to local FokI nuclease domain dimerization and ensuing site-specific double-stranded DNA breaks (DSB) formation within the spacer sequence. (**B**) Transcription activator-like effector (TALE) nucleases (TALENs). TALENs are chimeric modular DNA-binding proteins comprising the FokI nuclease domain fused through a flexible linker to a series of typically 17.5 repeats derived from TALE proteins. TALE proteins contain a translocation and transcriptional activation domain separated by a central array of typically 33-35 isomorphic repeats. The repeats harbor at amino acid positions 12 and 13 highly polymorphic residues named repeat variable di-residues (RVDs) that bind to specific nucleotides. The structure of 17.5 TALE repeats from an engineered TALEN monomer are depicted in frontal and lateral views. TALEN monomers of a working TALEN pair bind on opposite DNA strands in a tail-to-tail configuration resulting in local FokI nuclease domain dimerization and ensuing site-specific DSB formation within the spacer sequence. (**C**) RNA-guided CRISPR-Cas9 nucleases. Engineered CRISPR-Cas9 nucleases are sequence-specific ribonucleoprotein complexes consisting of a Cas9 protein with two nucleases domains (i.e., HNH and RuvC-like) bound to a single guide RNA (gRNA) formed by a sequence customizable CRISPR RNA (crRNA) fused to a constant trans-activating CRISPR RNA (tracrRNA) scaffold moiety to which the *S. pyogenes* Cas9 enzyme binds to. Target sequences of Cas9:gRNA complexes consist of the protospacer-adjacent motif (PAM) NGG placed next to an usually 20 nucleotide-long sequence complementary to the 5'-terminal end of the crRNA (spacer). The tertiary protein structures shown, each of which derived from the primary amino acid sequences of specific ZFN, TALE and Cas9 reagents, were homology-modeled through the SWISS-MODEL server. β-sheets and α-helixes are colored in green and violet, respectively.

ZFNs and TALENs do, confers these CRISPR-based nucleases with unsurpassed versatility and ease-of-use. Such features have fueled the primacy of RGNs amongst current programmable nuclease platforms. In fact, since the initial adaptation of natural CRISPR-Cas9 systems into genome engineering tools [35–38], RGN technologies are diversifying, being combined and adapted, at increasing rates [67]. For instance, structure-guided rational design and directed evolution approaches are producing new Cas9 variants whose features include; recognition of alternative PAMs that broaden the range of targetable genomic sites and improved target site specificities [67]. In parallel with these developments, phylogenetic analyses and mining of metagenomic datasets are unearthing components that make-up the highly diverse universe of CRISPR systems which, in addition to DNA, also target and degrade invading RNA [53]. Many of these components end up being successfully converted into reagents for (epi)genome and transcriptome modification or modulation in mammalian cells [67–69].

1.3. A Brief Overview on the Biology of Adenoviruses and Their Recombinant Types

Adenoviruses are a diverse group of viruses from the *Adenoviridae* family that have been evolving in a wide range of vertebrates, including humans, where they cause mild ailments, e.g., in the respiratory and gastrointestinal tracts [70–72]. Human adenoviruses belong to the *Mastadenovirus* genus with over 55 different serotypes identified so far. The various serotypes are grouped in species A through G based on phylogenetic, genome structure and hemagglutination criteria. Structurally, adenoviral particles (virions) consist of a non-enveloped icosahedral protein capsid displaying protruding fibers [70–72] (**Figure 2**). A linear double-stranded DNA genome with terminal proteins bound to their 5' ends is packaged inside each virion capsid consisting of 240 trimers of the hexon protein, 12 pentamers of

the penton base polypeptide and 12 trimeric fiber proteins that protrude from each of the 12 capsid vertices (**Figure 2**). Each homo-trimeric fiber consists of a basal tail domain that docks within the penton base axis, a slender shaft region and an apical globular knob domain responsible for the initial attachment of the virion to host cell receptors (**Figure 2**). In addition to the major capsomers hexon, penton base, and fiber, the adenoviral capsid also contains other so-called minor proteins some of which are thought to be important for cementing the virion structure [72, 73]. Adenovirus serotypes present broad cellular tropisms owing to their usage of a wide range of cell surface receptors. Identified primary attachment receptors include, the coxsackie and adenovirus receptor (CAR) used by the prototypic serotypes 2 and 5 from species C [74, 75] and CD46 and desmoglein-2 engaged by species B serotypes [76, 77]. Certain serotypes engage instead glycans and polysialic acids as primary attachment moieties [78, 79]. The natural diversity of adenoviruses and their corresponding wide range of host-cell receptors is permitting; (i) constructing new vectors based on rare serotypes that can escape pre-exiting immunity to adenoviruses prevalent in the human population, for anti-cancer and vaccination purposes [80]; and (ii) changing the tropism of established vectors based on species C adenovirus serotype 5 into those of other serotypes so that cells with therapeutic relevance lacking CAR can be efficiently transduced [81]. For instance, genetic retargeting of vector particles through the exchange of the apical regions of the adenovirus serotype 5 fiber (**Figure 2**) for those of species B adenovirus serotype 35 or 50 permits efficient transduction of $CAR^{low}/CD46^{high}$ hHSCs [82,83], human mesenchymal stromal cells (hMSCs) [84, 85] and human muscle progenitor cells [86].

The processes through which adenoviruses introduce their genomes into host-cell nuclei have been most extensively studied in the case of serotype 5 [87]. Briefly, after the initial attachment to the host cell, endocytosis via clathrin-coated vesicles is triggered by interactions between RGD motifs in penton bases and cellular integrins (e.g., $\alpha_v\beta_5$). Subsequently, incoming fiberless virions escape lysosomal degradation via the lowering of the pH in endosomes that permits remodeled capsid components to lyse the vesicle membranes. Once in the cytosol, the remodeled nucleocapsids bind to motor proteins dynein/dynactin that transports them along the microtubule network until they dock at the nuclear pore complex and release the packaged DNA into the nucleoplasm [87].

The most thoroughly used adenoviral vectors (AdVs) are deleted in the transcriptional units *E1A* and *E1B* that make-up the early region 1 (*E1*) (**Figure 3**). The production of these first-generation, *E1*-deleted, AdVs takes place in packaging cell lines (e.g., HEK293 and PER.C6) that express, and hence complement, *in trans* the *E1* gene products [88, 89]. The deletion of *E1*, firstly, blunts the activation of the regular adenoviral gene expression program preventing the replication of vector particles in transduced cells and, secondly, creates room for the packaging of approximately 5.0 kb of exogenous DNA in adenoviral capsids. Since the *E3* region is dispensable for replication in cell culture systems, some vector designs combine deletions in *E1* with deletions in *E3* that permit the packaging of up to 8 kb of exogenous

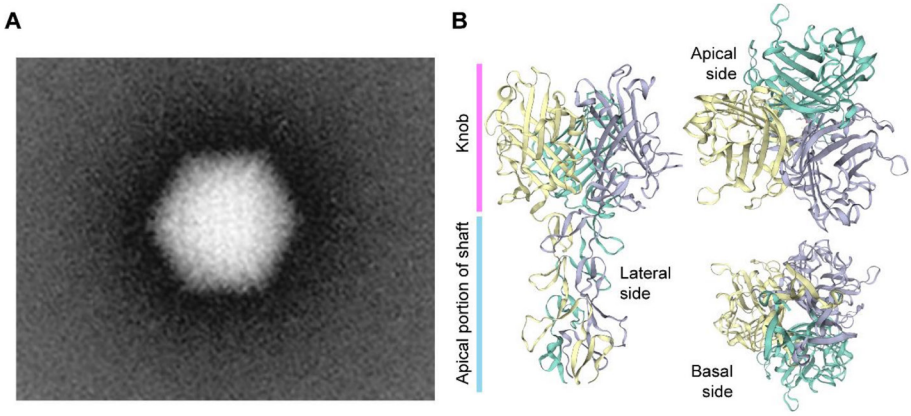


Figure 2. Adenovirus particle and the structure of its cell receptor-interacting fibers. (**A**) Transmission electron microscopy image of an adenovirus particle (virion). The icosahedral shape of the non-enveloped virion capsid can be discerned (~90 nm). A few of the twelve slender protruding fibers with their apical globular knob domains responsible for the initial interaction with the host-cell coxsackie and adenovirus receptor (CAR), can equally be discerned. (**B**) Three-dimensional model of the apical regions of the adenovirus serotype 5 fiber. The fiber is a homotrimer of the polypeptide encoded by the *L5* open reading frame and consists of the tail (not shown), the rod-like shaft and the globular knob domains. The tail anchors the fiber to the adenovirus capsid via non-covalent binding to the penton base proteins; the shaft projects the knob away from the capsid facilitating its interaction with CAR on the surface of host cells. The quaternary protein structure was homology-modeled using the SWISS-MODEL server and is depicted in different angles.

DNA [71]. As it came to be known, the *E1* deletion does not fully prevent residual expression from some of the transcriptional units that remain in vector genomes [71]. The resulting leaky synthesis of viral gene products leads to vector dose-dependent cytotoxicity in vitro and short-lived transgene expression in vivo (2-3 weeks) due to the clearance of transduced cells by the immune system [90]. For this reason, *E1*-deleted AdVs, in particular those based on serotypes with low seroprevalence in the human population, are being applied in clinical trials not for gene therapies requiring prolonged transgene expression but as vaccination agents instead, e.g., against hemorrhagic fever and AIDS caused by Ebola and HIV-1 infections, respectively [91, 92].

Second-generation AdVs combine deletions in *E1* or *E1* and *E3* with deletions in other early regions, i.e., *E4* or *E2* (**Figure 3**). Therefore, these vectors are generated in specialized packaging cell lines that complement *in trans* the respective missing gene products [71]. Although second-generation AdVs are more crippled than first-generation AdVs, at high vector doses, leaky synthesis of viral gene products can still be detected which also correlates with short-term transgene expression in vivo [71, 93].

To abrogate altogether leaky viral gene expression in transduced cells and, at the same time, maximize the size of foreign DNA that can be incorporated in adenoviral capsids, high-capacity adenoviral vectors (HC-AdVs) were developed [71] (**Figure 3**). These third-generation AdVs (a.k.a. “gutless” or helper-dependent AdVs) lack all viral coding sequences

retaining from the parental virus genome exclusively the short *cis*-acting inverted terminal repeats (ITRs) (103-bp each) and packaging elements needed for, respectively, vector DNA replication and encapsidation in producer cells (**Figure 3**). The need for complementing *in trans* the full set of adenoviral gene products, makes the production of HC-AdVs more complex than that of their earlier generation counterparts. In particular, HC-AdV particles are assembled in *E1*-complementing cell lines that express a site-specific recombinase (e.g., Cre or FLP) [71, 94, 95]. These producer cell lines are transduced with an *E1*-deleted helper AdV that expresses *in trans* the viral gene products necessary for the replication and packaging of HC-AdV genomes into adenoviral capsids. Crucially, the packaging signals of the helper genomes are flanked by recognition sequences for the site-specific recombinase so that the vast majority of assembled AdV capsids contain HC-AdV DNA in detriment of helper DNA owing to the selective recombinase-mediated removal of the packaging elements from the latter templates. Normally, besides the adenoviral *cis*-acting elements and the foreign DNA of interest, HC-AdV genomes also contain a so-called "stuffer" DNA segment to increase the HC-AdV DNA length to at least ~28 kb and, in doing so, guarantee vector genome stability during replication in producer cells [94].

2. Adenoviral Vector-Based Gene Editing in Human Adult Stem Cells and Their Progeny

2.1. Targeted Gene Disruption

Various viral vector systems initially developed for transgene expression and gene therapy undertakings, have also started to be investigated and coopted as gene editing agents (for a review on their features and main pros and cons, see, ref. 9). In fact, all three classes of replication-defective AdV systems (**Figure 3**) are included in these gene-editing research efforts, that are covered next.

E1-deleted AdVs based on serotype 5 displaying apical fiber motifs from CD46-interacting serotype 35 (AdV5/35) have been tested for conferring resistance to HIV-1 infection. In particular, AdV5/35 vectors encoding *CCR5*-specific ZFNs were applied for NHEJ-mediated generation of human $CD4^+$ T cells with reduced amounts of the transmembrane HIV-1 co-receptor protein CCR5 [96]. The ex vivo cell transduction protocol resulted in 40%-60% disruption of *CCR5* alleles in these cells. Importantly, transplantation experiments in immunodeficient NOD/Shi-scid/$gc^{-/-}$ (NOG) mice led to a 3-fold enrichment of $CD4^+$ T cells with *CCR5* knockout alleles in animals infected with HIV-1, suggesting selection for gene-modified cells. Notably, next generation sequencing (NGS) analysis of transduced $CD4^+$ T cells revealed a substantial ZFN-induced off-target activity (i.e., 5.39% indels) at the neighboring and highly sequence identical *CCR2* locus [96]. Building on this principle but aiming at a longer protective effect against HIV-1 infection, another study focused on targeting adult hematopoietic stem/progenitor cells (HSPCs). In this work, AdV5/35-mediated delivery of

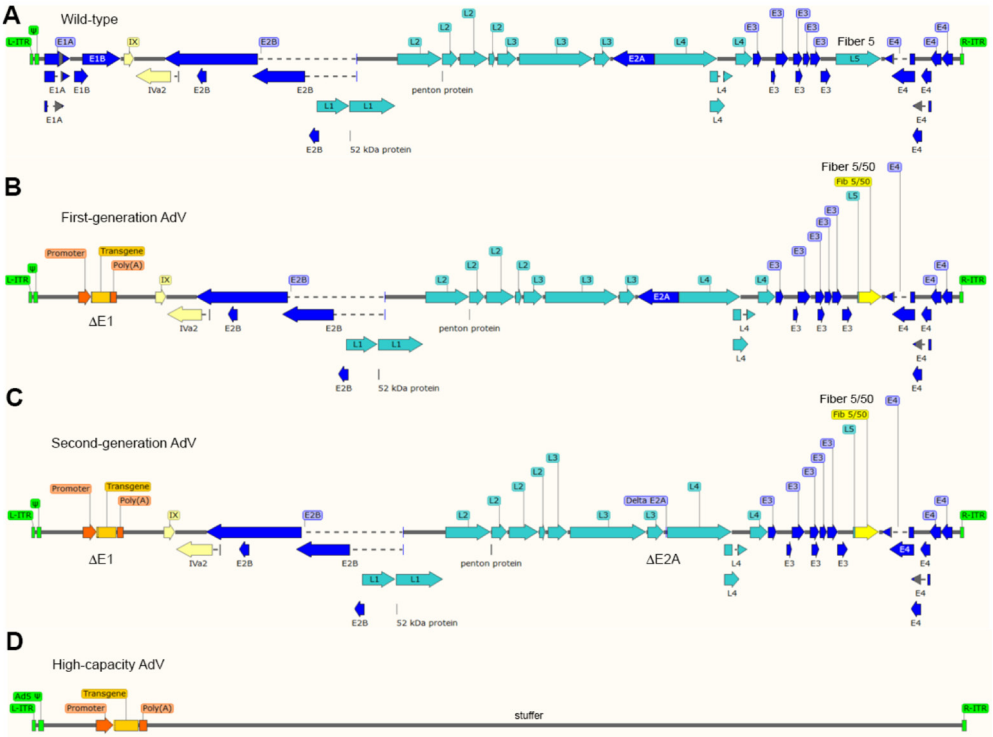


Figure 3. Schematics of wild-type and recombinant adenoviruses. (**A**) Genome structure of the prototypic human adenovirus serotype 5 drawn in relation to the genome structures of (**B**) first-generation (*E1*-deleted), (**C**) second-generation (*E1*- and *E2A*-deleted), and (**D**) third-generation or high-capacity (fully viral gene deleted) adenoviral vectors. The vectors contain a typical expression unit (transgene) consisting of a coding sequence of interest under the transcriptional control of a heterologous promoter and a polyadenylation signal. The first- and second-generation vector genomes encode chimeric fibers consisting of the basal shaft sequence of the human adenovirus serotype 5 linked to the apical shaft and knob domains from the CD46-interacting human adenovirus serotype 50 (yellow arrows). The non-coding *cis*-acting elements involved in vector genome replication and encapsidation are the inverted terminal repeats (ITRs) and packaging signal (Ψ), respectively. The latter signal and the "left" and "right" ITRs (L-ITR and R-ITR, respectively) are depicted in green. Regulatory functions necessary for activating the viral gene expression program are encoded by the early (E) regions *E1A*, *E1B*, *E2A*, *E3* and *E4* (dark blue arrows). The structural proteins required for assembling mature virions are encoded by the late (L) regions *L1* through *L5* (light blue arrows). The *L5* open reading frame (ORF) yields the cell surface receptor-interacting fibers. The full activation of the late viral gene expression program takes place after the onset of viral DNA replication. The ORFs coding for the intermediate proteins IX and IVa2 are also shown (light yellow arrows). Other adenoviral ORFs, e.g., small non-coding RNAs VAI and VAII are not depicted. The SnapGene software (version 5.0.7) was used for generating the different diagrams on the basis of the human adenovirus serotype 5 source sequence retrieved from GenBank accession number: AY601635.1.

CCR5-specific ZFNs into HSPC-enriched CD34^{+} cells led to target allele knockout frequencies above 25%. However, these knockout levels were only obtained in the presence of protein kinase C (PKC) activators, an expedient used to presumably improve vector transduction and/or ZFN expression [97]. Moreover, low, yet detectable, off-target activity at *CCR2* and at three other non-coding sequences located elsewhere in the genome were observed via NGS analysis. Subsequent cell transplantation assays in immunodeficient NOD/SCID/gc$^{-/-}$ (NSG) mice showed a vector dose-dependent reduction in the levels of human cell engraftment as

measured by CD45$^+$ cell counts in animals infused with HSPCs treated with PKC activators and *CCR5*-targeting ZFNs [97]. To avoid the toxicity caused by PKC activators, Maier and co-workers tested instead anti-CD3/CD28 stimulation as an adjuvant for improving transduction of T lymphocytes by an AdV5/35 vector encoding *CCR5*-specific ZFNs [98]. When compared to the experimental group exposed to PKC activation, this method enhanced the frequencies of target gene knockout by almost 3-fold (up to 32%). Importantly, ZFN-associated toxicity was not detectable with levels of off-target *CCR2* disruption in transduced T lymphocytes remaining below 4%, as estimated through genotyping assays based on mismatch-sensing nucleases and DNA fluorescence densitometry [98].

The generation of AdVs encoding ZFNs is challenging due to cytotoxicity caused by transgene overexpression in producer cells. To overcome this limitation, Saydaminova and colleagues exploited miRNA-dependent downregulation of transgene expression in 293-Cre packaging cells. This strategy permitted generating tropism-modified HC-AdVs encoding *CCR5*-specific ZFNs at high yields and without vector genome rearrangements. Importantly, miRNA profiling guaranteed that the endogenous miRNA suppressing ZFN synthesis in producer cells was not expressed in hHSC-enriched CD34$^+$ target cells [99]. Transduction of the erythroleukemia cell line MO7e and primary CD34$^+$ cells with the resulting HC-AdV coding for the *CCR5*-specific ZFNs led to 43.6% and 13% indel formation, respectively, at *CCR5* as determined by mismatch-sensing nuclease assays. Cell transplantation experiments in immunodeficient NOG mice revealed, however, that human CD34$^+$ cells transduced with the ZFN-encoding HC-AdV engrafted in the bone marrow at 3-fold lower levels than their non-transduced counterparts (i.e., 2.12% versus ~6%, respectively) [99].

A *CCR5*-specific ZFN pair delivered ex vivo into autologous CD4$^+$ T cells of AIDS patients by an *E1*-deleted AdV5/35 vector formed the basis for the first clinical testing of a programmable nuclease [100]. The infusion of 10 billion cells, of which 11%-28% were *CCR5*-disrupted, was shown to be safe. Moreover, edited cells persisted after transplantation with a mean half-life of 48 weeks and, tantalizingly, upon an interruption of anti-retroviral therapy, the rates with which *CCR5*-disrupted cells declined were significantly slower than those of unmodified cells [100]. Outstanding questions following from this landmark study are the feasibility in achieving sufficient numbers of cells with bi-allelic *CCR5* knockout without inducing cytotoxicity and with minimal ZFN-induced off-target effects. Finally, the combination of genetically retargeted AdV5/35 vectors and ZFN technologies has also been used for knocking out endogenous T-cell receptor genes and the primary HIV-1 receptor gene *CXCR4* in T cells [101, 102].

In addition to ZFNs, the AdV platform is equally suitable for the delivery of TALENs into human somatic cells, e.g., muscle progenitor cells and hMSCs. In fact, Holkers and co-workers demonstrated that, in contrast to HIV-1-based lentiviral vectors, transgenes encoding TALENs can be transferred intact into human cells by AdVs [103]. Indeed, lentiviral vectors encoding TALENs suffer substantial genetic rearrangements in the form of deletions of

various sizes that occur within the direct repeats corresponding to the TALE DNA-binding domains (**Figure 1B**). These deletions are likely caused by frequent reverse transcriptase template switching events taking place within the TALE repetitive tracts. Thus, the transfer of transgenes coding for TALE-based proteins through standard and integration-defective lentiviral vectors (IDLVs) requires substantial coding sequence optimization for minimizing sequence identity among repeats [104, 105].

It is also noteworthy to mention that, although IDLVs permit transient expression of ZFNs and sequence optimized TALENs in human cells, the yields necessary for robust targeted DSB formation might not be reached due to epigenetic silencing mechanisms directed at IDLV genomes involving histone deacetylases [106, 107]. In contrast, functional assays revealed that AdVs expressing TALENs allow for robust targeted DSB formation in several human cell types, e.g., muscle progenitor cells and hMSCs [103]. Second-generation AdVs deleted in *E1* and *E2A* and displaying apical motifs from CD46-interacting serotype 50 (AdV5/50) were used in these proof-of-concept experiments validating the AdV platform for the delivery of functional TALENs into human cells [103]. Follow-up experiments using first-generation and second-generation fiber-modified AdVs encoding TALENs and *S. pyogenes* Cas9 addressed to sequences flanking the major *DMD* mutational hotspot triggered large deletions comprising multiple exons (>500 kb) in patient-derived muscle progenitor cells [17]. These maneuvers designed for repairing *DMD* alleles causing Duchenne muscular dystrophy (DMD), led to the synthesis of in-frame mRNA transcripts encoding a truncated yet potentially functional Becker-like dystrophin protein [17].

Currently, the integration of AdV and programmable nuclease technologies for gene editing in somatic cells is dominated by the delivery and testing of RGNs. The first viral vector-mediated delivery of RGN components into mammalian cells consisted of using fiber-modified *E1*- and *E2A*- deleted AdVs expressing Cas9 or gRNAs directed to either a chromosomally integrated *EGFP* reporter or to the *AAVS1* safe harbor locus located in the human chromosome 19 at position 19q13.3-qter. In co-transduction experiments, robust targeted DSB formation was achieved at *AAVS1* in several cell types including human muscle progenitor cells and hMSCs [108]. In another study, co-transduction of human lung microvascular endothelial cells with an *E1*-deleted AdV and a lentiviral vector encoding Cas9 and a *TIE2*-specific gRNA, respectively, induced up to 90% of target gene disruption. Direct phenotypic analysis of *TIE2*-edited cell populations showed a persistent increase in endothelial cell permeability when compared to control cells [109].

In addition to NHEJ-mediated target gene disruption for basic biology studies, AdV-mediated RGN delivery is also being explored for modifying genes underlying human disorders. In this regard, to facilitate the delivery of RGN components, Maggio and colleagues co-packaged Cas9 and gRNA expression units within single particles of fiber-modified *E1*- and *E2A*-deleted AdVs [17]. In these experiments, testing "all-in-one" AdV-mediated transfer of RGN components, *DMD* exons 51 and 53 were separately targeted for resetting the *DMD*

reading frame in muscle progenitor cells derived from DMD patients [17]. In a follow-up study, fiber-modified *E1*- and *E2A*-deleted AdVs encoding Cas9 and gRNA pairs targeting *DMD* introns 52 and 53 or introns 43 and 54 were assembled for triggering single or multiple exon deletions, respectively [110]. The latter dual RGN-encoding vector permitted removal of the aforementioned major *DMD* mutational hotspot in up to 18% of target alleles in patient-derived muscle progenitor cells [110]. More recently, fiber-modified HC-AdVs were applied for the delivery of optimized high-specificity dual RGNs equally targeting *DMD* introns 43 and 54. The transduction of muscle progenitor cells isolated from DMD patients with these CD46-targeting HC-AdV particles resulted in the removal of the major *DMD* mutational hotspot in up to 42% of target alleles resulting in the direct detection of Becker-like dystrophin synthesis in differentiated muscle cell populations [111].

A study by Li and coworkers documented over 30% indel formation at *CCR5* in $CD4^+$ T cells that had been pretreated with a PKC activator and subsequently selected for RGN expression after exposure to *E1*-deleted AdV5/35 particles encoding EGFP-tagged RGNs. Significantly, the authors obtained evidence for the acquisition of resistance of *CCR5*-edited $CD4^+$ T cells to two different HIV-1 strains in vitro [112].

Disruption of binding motifs for the *HBG* repressor protein BCL11A is a promising strategy to reactivate *HBG* expression and fetal γ-globin synthesis to complement the absence of functional adult β-globin in β-thalassemic and sickle cell disease (SCD) patients. In this regard, transduction of mobilized peripheral blood $CD34^+$ cells from healthy donors with fiber-modified HC-AdVs encoding *HBG*-specific RGNs led to around 20% of target motif disruption in these cells [113]. Moreover, no indels were observed in the top 10 candidate off-target sites, as assessed by mismatch-sensing nuclease assays and, importantly, the erythroid differentiation capability of the gene-edited hematopoietic progenitors was maintained [113]. Cell transplantation assays in lethally irradiated immunodeficient mice revealed indel frequencies ranging from 19% to 25% at *HBG* alleles in human $CD45^+$ cells isolated from bone marrow at 10 weeks post-transplantation. Upon in vitro differentiation of these bone marrow-derived $CD45^+$ cells, the frequencies of γ-$globin^+$ cells were ~50% and ~27% in the transduced and non-transduced groups, respectively, as determined by flow cytometry [113]. In addition, β-YAC/CD46 mice were also used in this study to overcome the known block on human erythrocytic lineage differentiation in NSG mice. β-YAC/CD46 mice contain a human DNA fragment encompassing the entire 82-kb human β-globin locus and express the human CD46 receptor which permits transducing mouse cells with HC-AdV particles displaying adenovirus serotype 35 fibers. Hence, this mouse model allows in vivo evaluation of *HBG* reactivation in mature circulating erythrocytes. Bone-marrow Lin^- cells isolated from β-YAC/CD46 mice were transduced with the fiber-modified HC-AdVs encoding *HBG*-specific RGNs and were subsequently transplanted into lethally irradiated C57BL/6 recipient mice. At 10 weeks post-transplantation, there was a ~5-fold reduction of *HBB* mRNA and a ~30-fold increase in *HBG* mRNA levels in red blood cells when compared to controls. These

results indicate that a switch in the balance of adult to fetal globin expression was achieved [113]. In another study, Li and co-workers using fiber-modified HC-AdVs encoding RGNs targeting *BCL11A* gene enhancer or BCL11A protein binding sequences obtained over 20% indel formation at these motifs in CD34$^+$ cells [114]. Interestingly, however, in vitro colony-forming unit (CFU) assays based on semi-solid methyl-cellulose medium showed a reduction in the number of multi-lineage progenitors derived from vector-transduced cells [114]. In addition, cell transplantation assays in irradiation-conditioned NSG mice demonstrated that engraftment rates of CD45$^+$ cells in mice receiving grafts transduced with RGN-encoding HC-AdVs were 5- to 10-fold lower than those transplanted with non-transduced cells or cells transduced with a control vector encoding exclusively Cas9 [114]. The low numbers of CFUs in vitro and engraftment rates in vivo indicated RGN-induced cytotoxic effects. In line with this data, Schiroli and colleagues found through single-cell transcriptomics analysis that DSBs induced by ZFNs and RGNs can activate a P53-dependent DNA damage response in HSPCs [115]. To shorten the duration of RGN activity, bacteriophage anti-CRISPR (Acr) peptides AcrIIA2 and A4, were exploited to inhibit long-term Cas9 activity [114]. Sequential transfer of *BCL11A* enhancer-specific RGNs and Acr peptides via tropism-modified HC-AdV transductions with an interval of 48 hours led to 37.9% indel formation in the human umbilical cord blood-derived erythroid progenitor cell line HUDEP-2 [114]. Flow cytometry and qRT-PCR analyses showed a switch of *HBB* to *HBG* expression in the edited HUDEP-2 populations. After applying a similar sequential HC-AdV transduction protocol to CD34$^+$ cells followed by transplantation of vector-treated cells into irradiation-conditioned NSG mice, Li and coworkers observed comparable levels of CD45$^+$ cell engraftment in mice receiving non-transduced and vector-transduced cells. Indel frequencies at the *BCL11A* gene enhancer and BCL11A protein binding site ranged from 8.5% to 27% and from 10.5% to 21%, respectively, in CD45$^+$ cells isolated from bone marrow, as measured by mismatch-sensing nuclease assays. Finally, in vitro differentiation of isolated CD45$^+$ cells into erythroid cells, revealed a ~1.4-fold increase in the percentage of g-globin$^+$ cells in the edited over the control groups [114].

2.2. Targeted Gene Integration

As aforesaid, HDR leads to precise genomic DNA editing in the presence of exogenous donor templates that can be designed for gene knock-ins, gene knockouts or gene correction. Therefore, AdVs are also being utilized for transferring programmable nucleases together with donor templates into human cells. In this context, Coluccio and colleagues combined AdV-mediated ZFN delivery with the transfer of donor HDR substrates in AdVs or IDLVs for testing homology-directed gene insertion in human keratinocytes [116]. In this study, *AAVS1*-specific ZFNs were delivered by an *E1*-deleted AdV5/35 vector, whereas the donor, containing a reporter gene flanked by *AAVS1*-targeting homologous sequences, was transferred via either vesicular stomatitis virus glycoprotein G-pseudotyped IDLV or *E1*-deleted AdV5/50 particles. Transduction of HaCaT cells, a human keratinocyte cell line, with ZFN-encoding AdV particles together with IDLV or AdV donors led to chromosomal

transgene integration frequencies of 20% and 1%, respectively [116]. However, combining AdV5/35 and IDLV vectors for introducing into human primary keratinocytes *AAVS*1-specific ZFNs and donor templates, respectively, resulted in substantially lower frequencies of stable transgene insertion (i.e., 0.3%), presumably in part due to the observed inefficient transduction of these target cells by IDLV particles [116]. In another study, investigating homology-directed gene targeting, Holkers and coworkers combined the transfer of HDR substrates in AdV or IDLV particles with AdV-mediated delivery of TALENs instead [117]. In particular, *AAVS1*-specific TALENs were delivered by an *E1*-deleted AdV5/50 vector, whereas the donor, containing a reporter gene flanked by *AAVS1*-targeting sequences, was transferred via either IDLV or *E1*- and *E2A*-deleted AdV5/50 particles. Transduction of human muscle progenitor cells with TALEN-encoding AdVs together with IDLV or AdV donors led to chromosomal transgene integration frequencies of 9.1% and 1.24%, respectively. These data together with that of Collucio and coworkers indicate that IDLV donors lead to higher frequencies of DSB-dependent gene knock-ins than those achieved by AdV donors. However, isolation of genetically modified muscle progenitor cells (n = 214 clones) followed by clonal analysis using junction PCR assays demonstrated that a large proportion of IDLV-modified cells contained random insertions (13.4%) or inaccurate *AAVS1* insertions (44.3%), of whom a substantial fraction corresponded to head-to-tail donor DNA concatemers (38.5%). In contrast, neither random insertions nor inaccurate *AAVS1* insertions were detected in the randomly isolated AdV-modified cells [117]. Thus, although free-ended IDLV genomes lead to higher frequencies of genetically modified cells than protein-capped AdV genomes, the latter genomes result in more specific and accurate HDR-mediated donor DNA insertion [28, 117]. The relevance of the donor DNA structure to the specificity and accuracy of gene targeting was demonstrated by experiments in which the excision of HDR substrates from the context of protein-capped AdV genomes resulted in an increase in random donor DNA insertions, as determined by clonal analysis using junction PCR assays [117]. Presumably, albeit more efficacious for generating populations of genetically modified cells, linear free-ended DNA is prone to homology-independent capture at chromosomal DSBs (targeted or otherwise) through illegitimate recombination processes comprising end-to-end DNA ligations.

Li and colleagues applied HC-AdV5/35 vectors for delivering into human $CD34^+$ cells *AAVS1*-specific RGNs and donor DNA templates encoding EGFP and the positive selectable marker $mgmt^{P140K}$ [118]. The latter gene product confers resistance to O6BG/bis-chloroethylnitrosourea (BCNU). In this study, *AAVS1* gRNA target sites flanked the donor template for enhancing the frequencies of genetically modified cells via RGN-induced donor DNA excision. Co-transduction of human $CD34^+$ cells with both AdVs resulted in 0.9% of $EGFP^+$ hematopoietic cell clones as determined by CFU assays. Further characterization of these colonies (n=14) showed accurate insertion of the donor DNA at the *AAVS1* locus. The delivery of *AAVS1*-specific RGNs and *AAVS1*-targeting donor templates into murine Lin^- cells,

isolated from the bone marrow of human *AAVS1*/CD46 transgenic mice, was done through their ex vivo co-transduction with HC-AdV5/35 particles. As controls, parallel samples of Lin^- cells were exposed exclusively to one of the two vectors. Subsequently, vector-transduced Lin^- cells were transplanted into lethally irradiated C57BL/6 mice. Notably, in these experiments, no significant differences in engraftment rates were observed in mice receiving cells treated with the different HC-AdV5/35 regimens. At 4 weeks post-transplantation, an average of 1.1% and <0.2% of $EGFP^+$ peripheral blood mononuclear cells (PBMCs) were measured in the experimental and control groups, respectively. After three rounds of BCNU selection an enrichment in $EGFP^+$ cell marking was observed that varied from ~20 to ~100%, depending on the recipient mouse analyzed. Importantly, multilineage $EGFP^+$ cell marking was stably maintained for 16 weeks in secondary recipients demonstrating genetic modification of *bona fide* murine HSCs. Building on these data and experimental settings, Li and colleagues went on to test HDR-mediated knock-in of a γ-globin-coding transgene at the human *AAVS1* locus in murine Lin^- cells isolated from AAVS1/CD46 transgenic mice. The transgene was placed under the regulation of a mini-β-globin locus control region for preferential expression in erythroid cells. Lin^- cells transduced with HC-AdV5/35 particles were transplanted into lethally irradiated C57BL/6 mice and were subsequently subjected to three rounds of BCNU selection. At 16 weeks post-transplantation, the level of γ-globin was on average 20.52% and 22.33% of that of adult mouse β-globin as measured by high-performance liquid chromatography and qRT-PCR analyses, respectively [118].

The cumulative data from these investigations on the use of AdV systems for gene editing of adult stem cells and their progeny bodes well for their application in basic research and biotechnologies, including for the development of genetic therapies targeting acquired and inherited disorders.

3. Human Embryonic Stem Cells (hESCs) and Human Induced Pluripotent Stem Cells (hiPSCs) Genome Editing

Human pluripotent stem cells (hPSCs) renown rose ever since the first isolation of human embryonic stem cells (hESCs) from pre-implantation embryos in 1998 [119]. Under well-defined culture conditions, hESCs are able to self-renew and can replicate for long periods in vitro while maintaining their full potential to differentiate into any somatic cell type derived from the three embryonic germ layers; endoderm, ectoderm, and mesoderm. These unique features of self-renewal and pluripotency facilitate studying cell differentiation processes and creating in vitro models of human disorders ("disease-in-a-dish"). In addition, hESCs hold the promise of revolutionizing regenerative medicine through the establishment of innovative stem cell therapies and represent invaluable tools for drug screening and development. Nevertheless, the therapeutic application of hESCs is limited not only by technical challenges but also ethical concerns stemming from their human-embryo origins [120]. For this reason, the generation of human induced pluripotent stem cells (hiPSCs) represented a fundamental turning point in this field of biomedical research [121]. This revolutionizing discovery took place in 2006, when Takahashi, Yamanaka and colleagues discovered that a cocktail of four transcription factors (i.e., KLF4, c-MYC, OCT4, and SOX2) was capable of reprogramming somatic, terminally differentiated cells, "back" to an hESC-like state [122, 123]. Indeed, for the most part, hiPSCs maintain the characteristics of hESCs, including their defining features of self-renewal and pluripotency. Crucially, cellular reprogramming overcomes the ethical concerns associated with hESCs and offers the possibility for generating and differentiating hiPSCs from virtually any individual into tissue-specific cell types. These capabilities permit in vitro disease modeling and drug screenings [124, 125]. Moreover, hiPSCs open the perspective for autologous cell transplantation therapies for repairing tissues and organs affected by injuries or, when combined with gene-editing technologies, inherited disorders [124, 125] (**Figure 4**). Indeed, the advances made in gene editing technologies are greatly impacting hPSC-based research [126]. Firstly, gene editing of hiPSCs is an important steppingstone towards their clinical translation, in that targeted correction of patient-derived hiPSCs might pave the way for the development of personalized regenerative medicines of otherwise untreatable genetic diseases [126] (**Figure 4**). Secondly, gene editing contributes to the establishment of clear genotype-phenotype associations by permitting the generation of isogenic pairs of hiPSC lines that share the same genetic background and differ exclusively in specific well-defined DNA sequences. These isogenic hiPSC pairs can be obtained either via correcting a genetic defect in a patient-derived hiPSC line or introducing mutations causing a genetic defect in a wild-type hiPSC line (**Figure 4**).

Several studies employing engineered ZFNs, TALENs, and RGNs, have shown the utility of these molecular tools for gene editing in hPSCs [127]. The off-target effects and unpredictable genomic changes resulting from the repair of DSBs made by programmable nucleases are, however, major concerns in the gene editing field, especially in its application

to stem cells [56–60]. In this regard, recent developments on genome engineering strategies based on sequence- and strand-specific nucleases (nickases) as such [61, 128–130] or on the fusion of these nickases to cytidine or adenine deaminases (i.e., base editors) [131] or reverse transcriptases (i.e., prime editors) [132] is gaining momentum. In part, this momentum derives from the fact that these tools open up the perspective for efficient, DSB-free, genetic modification of stem cells whose sensibility to DSBs is particularly acute [115, 133, 134]. Next to gene editing strategies based on nucleases and nickases, there are also gene editing approaches that rely on the exclusive delivery of exogenous HDR substrates into hPSCs. In this case, stringent positive and negative selection schemes are often necessary for the isolation of properly targeted cells as HDR events are very rare in the absence of DSBs at target DNA [25–27] or SSBs at target and donor DNA [61, 128–130, 135]. Moreover, to ameliorate the inefficiency of HDR in the absence of targeted DNA lesions, whenever possible, donor templates are endowed with long sequences homologous to target genomic regions. Indeed, extensive homologous sequences, normally spanning several thousands of bps flanking the desired exogenous DNA are exploited for obtaining site-specific gene insertion through spontaneous HDR. However, regardless of their dependency on or independency from nucleases or nickases, and derivatives thereof, a main challenge for operational gene editing in adult stem cells and hPSCs remains the need for delivering the necessary molecular tools in an efficient and, ideally, non-cytotoxic manner. To this end, various viral and non-viral delivery systems are being explored [9, 136]. We will next highlight the contributions of HC-AdV technology for gene editing in hiPSCs and hESCs

3.1. High-Capacity Adenoviral Vector (HC-AdV)-Based Gene Editing in hESCs and hiPSCs

HC-AdV-based gene editing of PSCs involving exclusively donor DNA delivery was initially applied in murine ESCs for achieving HDR-mediated correction of *Hprt* alleles [137]. Soon thereafter, Suzuki and coworkers tested HC-AdVs for gene editing in hESCs [138]. These authors started by comparing HC-AdVs displaying serotype 5 or serotype 35 fibers for transducing hESCs by measuring through flow cytometry the frequencies of cells transiently expressing the Venus fluorescent protein reporter. Both viral vectors showed a clear multiplicity of infection (MOI)-dependent increase in transduction efficiencies that reached over 90% of target cells. The highest gene transfer levels were obtained with the tropism-modified vector. Notably, at a low to moderate MOI range, i.e., 10–300 transducing units per cell (TU/cell), cytotoxic effects were not significantly different from mock-transduced cells. Subsequently, HC-AdVs displaying conventional serotype 5 fibers were employed at a MOI of 300 TU/cell to deliver an *HRPT1*-targeting construct with long regions of homology (i.e., 14.3 kb and 9.2 kb) designed to insert a *neomycin phosphotransferase* (neo^R) cassette. Cells stably expressing the neo^R gene product acquire resistance to the aminoglycoside antibiotic G418 (also known as geneticin) (**Figure 5**). In addition to the positive selection marker gene neo^R, in order to minimize the expansion of cells with ectopic vector DNA integration, the vector genome

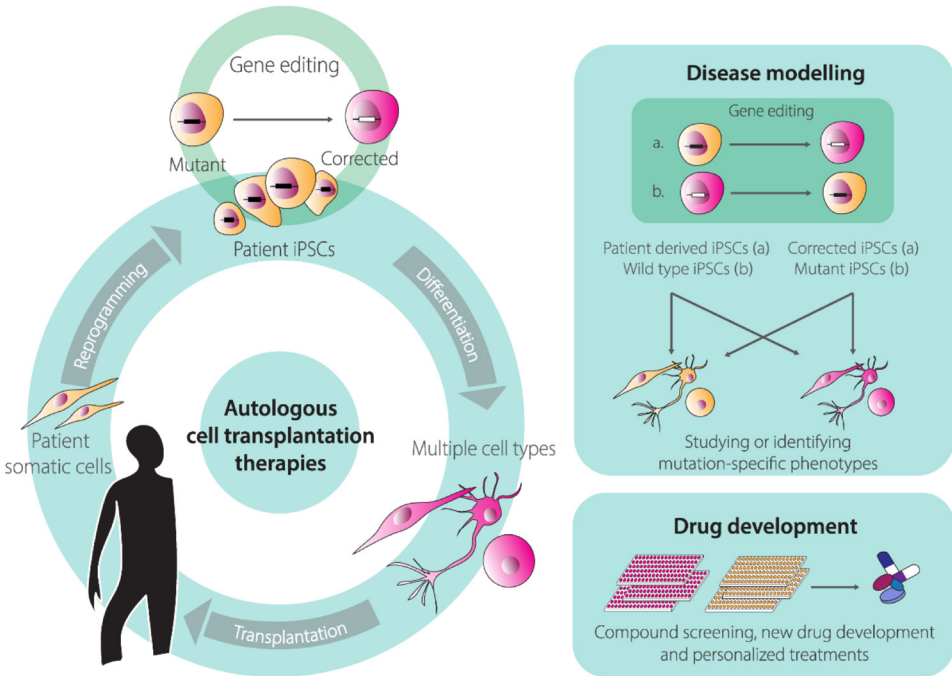


Figure 4. Illustration of human induced pluripotent stem cell (hiPSC)-based research and development activities enablwed by genome editing technologies. *Ex vivo* reprogramming of patient-derived somatic cells into hiPSCs followed by their genetic correction, expansion, and directed differentiation into specialized cells types opens the perspective for the development of innovative autologous cell therapies. Generation of hiPSC lines sharing the same genetic background and differing from each other at predefined genetic loci can be accomplished via either (a) targeted correction of specific mutations in patient-derived hiPSCs or (b) targeted installation of specific mutations in wild-type, healthy donor-derived, hiPSCs. The resulting pairs of isogenic hiPSC lines form tractable experimental systems for the controlled and robust establishment of genotype-phenotype associations during disease modeling and for high-throughput screens aiming at assessing drug toxicities and/or identifying new drug candidates.

also contained a negative selection cassette external to the homology regions expressing the Herpes Simplex Virus type 1 thymidine kinase (HSV1-tk) (**Figure 5**). Therefore, in case of HDR-independent or random chromosomal integration of HC-AdV DNA, stable HSV1-tk synthesis converts the pro-drug ganciclovir (GCV) into a phosphorylated cytotoxic product that leads to cell death (**Figure 5**). Among 5.1×10^6 transduced hESCs, 136 colonies were G418-resistant and, of these, 31 were G418/GCV double-resistant. PCR and Southern blot analyses further demonstrated that of the 31 double-resistant colonies, 14 were correctly targeted at *HPRT1* [138]. Importantly, HC-AdV transductions led to significantly higher gene transfer efficiencies than those obtained by "naked" DNA transfections based on electroporation and FuGENE HD. Moreover, when compared to the electroporation of the same *HPRT1*-targeting construct, HC-AdV donor delivery proved to be ~300 fold more efficient in terms of the frequencies of precisely edited cells obtained [138].

Building on these promising findings, a follow-up study investigated a similar HC-AdV-based gene editing approach in both hESCs and hiPSCs [139]. In this study, the authors explored different gene editing settings, i.e., (i) knock-in of a donor *neo*^{R} cassette at the housekeeping *HPRT1* locus, (ii) knock-in of a donor *neo*^{R} cassette designed for conditional knock-out of target genes located at different genomic positions, and (iii) knock-in of a donor *EGFP* cassette at a transcriptionally inactive *HB9* locus. Firstly, *HRPT1*-targeting experiments for knocking-in the donor *neo*^{R} cassette in two distinct hiPSC lines led to 20% and 7% of correctly targeted clones after positive-negative G418/GCV selection [139]. Significantly, control experiments involving the electroporation of the linearized *HPRT1*-targeting HC-AdV plasmid led to 0% of correctly targeted clones. Secondly, *neo*^{R} cassette knock-in experiments at *KU80, LIG1,* and *LIG3* led to 81%, 34%, and 42% gene targeting frequencies, respectively. Subsequently, the *lox*P-flanked *neo*^{R} cassette was excised in ~25% of the targeted cells through transient Cre delivery and target gene knockouts were confirmed through clonal analyses using Southern blotting, RT-qPCR, and western blotting. Finally, HC-AdV-mediated *EGFP* knock-in at the transcriptionally inactive *HB9* locus led to 23% and 57% of accurate gene targeting in hiPSC and hESC lines, respectively. Other studies confirmed that silent loci are accessible to HC-AdV-based gene editing. For example, to trace gene expression during cell differentiation, HC-AdVs were employed to knock-in live-cell reporter genes into *ALB* and *OC* alleles to monitor the differentiation of hESCs and/or hiPSCs along the hepatic and osteogenic lineages, respectively [140, 141].

3.2. HC-AdV-Based Gene Editing for Targeted Gene Correction in Human Pluripotent Stem Cells (hPSCs)

HC-AdVs are also being investigated for targeted correction of disease-causing mutations in hPSCs (**Figure 5**). Initial experiments targeted mutations underlying Hutchinson–Gilford progeria syndrome (HGPS) and atypical Werner syndrome (AWS) in hiPSCs [142]. HGPS and AWS are laminopathies whose mutations in the exon 11 of the *LMNA* gene include C1824T and A1733T, respectively. These mutations affect the nuclear structure resulting in premature aging. By exploiting the large cloning capacity of HC-AdV particles, HGPS and AWS can potentially be tackled by a single large *LMNA*-targeting construct covering different mutations. Similar to previous work [138], upon HC-AdV donor DNA transduction of hiPSCs and positive-negative G418/GCV selection, integration of the *neo*^{R} cassette at the *LMNA* target site between exons 10 and 11 ranged from 78% to 100%, as assessed through PCR and Southern blot analyses [142]. Correction of the 1-bp substitutions C1824T and A1733T located in exon 11 of *LMNA* in HGPS-hiPSCs and AWS-hiPSCs, respectively, was verified through DNA sequencing of targeted clones. This analysis revealed that 12 out of 25 HGPS-hiPSC clones and 35 out of 65 AWS-hiPSCs clones were accurately repaired. Subsequently, the *neo*^{R} cassette, flanked by *FRT* sites, was excised by transient expression of FLPe recombinase leading to wild-type *LMNA* expression and subsequent rescue of the HGPS phenotype, as determined by the restoration of normal nuclear architecture and cell senescence programs

[142]. Next, in addition to confirming the pluripotency of gene-edited hiPSCs, the authors meticulously investigated the genetic and epigenetic integrity of the corrected cells. In particular, correctly targeted hiPSCs showed a normal karyotype, expressed pluripotency markers and exhibited demethylation of the promoter of the pluripotency gene *OCT4* [142]. Moreover, genome-wide single nucleotide polymorphism (SNP), DNA microarray, and genome-wide DNA methylation analyses indicated a generic maintenance of the genetic background, global gene expression patterns, and global epigenetic states, respectively, in gene-edited cells using parental hiPSC lines as references [142]. In another study, HC-AdV-based gene editing was applied to correct the A→T transversion at nucleotide 20 in exon 1 of the β-globin-encoding *HBB* gene in hiPSCs obtained from SCD patients [143]. In these experiments, the positive-negative G418/GCV selection resulted in an average of 85% of colonies with *neo*^{R} targeted insertions with an average of 81% of these colonies presenting the desired *HBB* gene correction [143].

The previously described gene editing experiments targeting *LMNA* [142] and *HBB* [143], demonstrated that vector DNA-derived SNPs could be found in the correctly targeted clones at positions 4.4-kb and 3.6-kb away from the *neo*^{R} insertion site within *LMNA* and *HBB* alleles, respectively. On the basis of these results, the authors postulated that the HC-AdV platform might be valuable for repairing mutations found in a relatively broad target region, increasing its potential as a versatile gene correction tool. As an example, a single *LMNA*-targeting HC-AdV could potentially repair over 200 *LMNA* mutations associated with laminopathies [142].

Two subsequent studies sought to formally investigate; (i) the extent of homology between endogenous target and exogenous HC-AdV donor templates required for efficient gene editing [144]; and (ii) the relationship between the distance from the knock-in target site and the incorporation of polymorphic markers located along the region of homology [145]. In both studies, HC-AdV targeting constructs were directed to the *CFTR* locus in a hiPSC line harboring the heterozygous mutations ΔF508 and ΔI507 in exon 10 of the target gene. To investigate the effect of the extent of homology on the efficiency of HC-AdV-based gene editing, a set of five different HC-AdVs containing differently sized wild-type *CTFR* sequences were tested [144]. The homology regions spanned total lengths of 23.8 kb, 21.4 kb, 14.8 kb, 9.6 kb, and 5.6 kb. Transduction of hiPSCs with the various HC-AdV donors followed by G418 and GCV double selection led to the emergence of colonies that were subsequently subjected to Southern blot analysis for determining the frequencies of targeted events. The HC-AdV donor construct carrying 23.8 kb of sequence homology to genomic DNA led to 97.4%-100% of gene-targeted clones; whilst the HC-AdV donor construct bearing 5.6 kb of sequence homology to genomic DNA yielded 50% of gene-targeted clones [144]. Together, these data lend additional support to a direct correlation between the length of homology between target and donor DNA and the frequency of HDR-mediated gene targeting [21].

In order to investigate the extent of exchange of homologous sequences between

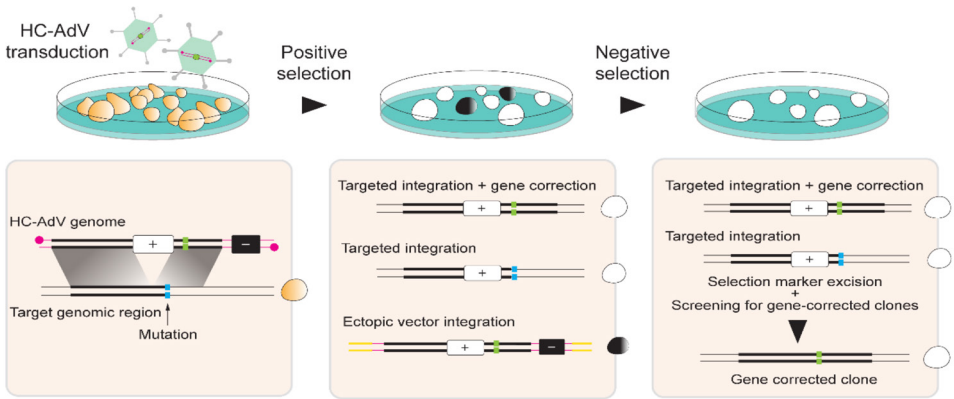


Figure 5. DSB-independent gene editing based on HC-AdV donor DNA transduction and positive-negative cell selection protocols. HC-AdV genomes contain a positive selection cassette, e.g., *neo*^{R} (white box) flanked by extensive human DNA sequences that are homologous to a target genomic region except for specific nucleotide(s) (left panel). In this example, donor and acceptor templates bear wild-type and mutant allelic sequences of a target gene (green and cyan boxes, respectively) so that, after recombination, involving outward homologous regions, gene correction ensues (middle panel, upper diagram). Next to these wanted outcomes there are also unwanted ones in the form of homologous and non-homologous recombination events resulting in no gene correction and random chromosomal donor DNA integration (middle panel, central, and bottom diagrams, respectively). Cells containing these different types of genetic modifications survive and multiply in the presence of a cell-killing drug that is broken-down by the positive-selection gene product. Selective elimination of cells with random HC-AdV donor DNA insertions is accomplished owing to the presence of a suicide negative selection cassette located outside the homology regions, e.g., *HSV-tk* (black box), that convers a prodrug substrate into a cell-killing product. The positive selection marker can subsequently be removed by site-specific recombinases, e.g., Cre and FLP that leave *lox*P and *FRT* site footprints, respectively, in the genome. Alternatively, transposon/transposase systems, e.g., footprint-free PiggyBac variants can be used that ultimately achieve scarless genomic modifications. Finally, genotyping screens permit identifying cells containing correctly targeted alleles (right panel).

target and donor DNA templates, twelve 2-bp insertions were introduced along the 23.8 kb homology region in a *CTFR*-targeting HC-AdV construct [145]. Upon HC-AdV-mediated gene targeting, each of these 2-bp insertions convert an endogenous restriction enzyme recognition site into that of another allowing for straightforward assessment of the extent of recombination between target and donor DNA sequences. As assessed through Southern blot analysis, 89.5% of drug-resistant hiPSC clones were correctly targeted at *CFTR* alleles [145]. Furthermore, PCR and restriction enzyme fragment length analyses of the drug-selected hiPSC clones showed that the closest marker to the insertion site (i.e., 208 bp) was incorporated in 100% of the analyzed clones. Conversely, the most distant marker to the insertion site (i.e., 11.2 kb) was incorporated in only 21.7% of the analyzed clones, suggesting that the vicinity of polymorphic markers to the insertion site is proportional to their genomic incorporation rate. Interestingly, 4.8% of the clones presented all the twelve restriction enzyme markers. This data suggests that HC-AdV-based gene editing can be used to introduce genetic information distributed over a wide range of homologous DNA in hiPSCs (i.e., at least up to 22.2 kb) [145].

As aforementioned, HC-AdV-based gene editing is equally applicable for establishing tractable in vitro disease models comprising pairs of isogenic hPSC lines whose genomes differ at well-defined locations (Figure 4). Indeed, HC-AdV-based gene editing has been explored for modeling various human disorders, including; Parkinson's disease [146], Fanconi

anemia [147], retinitis pigmentosa [148], and Werner syndrome [149].

Combining HC-AdV and programmable nuclease technologies offers the prospect for improving gene editing frequencies. In this regard, Suzuki and colleagues used HC-AdVs to deliver donor templates alone or together with TALEN expression units [150]. The TALEN and donor HDR substrates were tailored for targeting *HBB* alleles underlying SCD in hiPSC lines. Transduction of SCD patient-derived hiPSCs with the "all-in-one" HC-AdV resulted in an increase in gene-targeting frequencies when compared to those achieved by HC-AdV delivery of donor DNA templates alone [150]. Specifically, among 2×10^5 cells transduced with the "all-in-one" HC-AdV, 28 G418-resistant clones were analyzed and of these 86% were correctly targeted. Conversely, among 9×10^6 cells transduced with an HC-AdV delivering exclusively donor DNA, 134 G418-resistant clones were analyzed with only 22% of these being correctly targeted [150].

The cumulative data on HC-AdV-based gene editing in hPSCs bodes well for its application in basic research, drug screening, disease modeling, and eventually, development of autologous cell therapies for inherited disorders (**Figure 4**).

4. Conclusions and Outlook

Rapid advancements in the gene editing and stem cell fields are contributing to broaden the range of options for addressing scientific questions and developing candidate gene and cell therapies. To support the integration of these fields, and hence further widen their reach, it is crucial to develop delivery systems that permit introducing programmable DNA-targeting enzymes and donor nucleic acid templates into target cells in an efficient and versatile manner. Moreover, additional parameters that need to be taken into consideration concern the effects that the delivery systems themselves have on the ultimate performance and accuracy of gene editing procedures. In the case of gene-editing tool delivery through viral vector systems, it is important that vector genomes transporting donor templates or encoding programmable DNA-targeting enzymes are refractory to (i) structural rearrangements [103], (ii) epigenetic silencing mechanisms [106, 107], and (iii) capture at chromosomal DSBs via illegitimate recombination processes [117, 151, 152].

Recent developments on genomic engineering comprise the progression from chromosomal cutting to chromosomal non-cutting approaches based on nicking Cas9 variants and on these variants fused to heterologous DNA-modifying moieties. These new gene editing principles include; (i) HDR-mediated chromosomal insertion of exogenous DNA spanning from single bps to whole transgenes through SSB formation at target and donor DNA [61, 128–130], and (ii) donor DNA-free in situ installation of genetic changes through base editing [131] or prime editing [132]. Base editors, comprising a Cas9 nickase covalently linked to a cytidine or adenine deaminase, induce C→T or A→G transitions, respectively [153, 154]. These conversions occur within so-called "editing windows" located in target sequences defined by a standard gRNA [131, 153, 154]. Prime editors, consisting

of a Cas9 nickase covalently linked to an engineered oncoretroviral reverse transcriptase (RT), in addition to transitions, also generate defined indels and transversions, e.g., A→C, G→T, T→A, and C→G [132]. The exact genetic modification depends on the designing of an extended gRNA dubbed prime editor gRNA (pegRNA). The pegRNA is formed by the standard gRNA sequences crRNA and tracrRNA (**Figure 1C**) covalently linked to a RT primer binding site (PBS) and a RT template sequence bearing the intended edit. After nicking, the PBS locally anneals to the 3'-ended DNA flap that primes RT synthesis over the RT template. The resulting DNA copy of the edit ultimately becomes incorporated at the genomic target site upon a series of cellular processing steps responsible for removing DNA flaps that do not hybridize to target sequences [132].

The SSB-mediated gene editing approaches are opening the perspective for modifying complex genomes with unprecedented precision while minimizing unwanted events characteristic of DSB-mediated gene editing procedures. In addition to off-target mutagenesis [56–61], unwanted genome-modifying events include translocations [60, 61] and unpredictable genomic "scars" at target sequences in the form of indels and larger structural rearrangements resulting from site-specific DSB repair via prevalent NHEJ pathways [60, 155]. Not surprisingly, however, new gene editing approaches and technologies bring to the fore their own sets of shortcomings that need to be carefully assessed and resolved. For instance, base editors can yield off-target editing at the genome and transcriptome levels [156]; whereas primer editing can install target-site mutations derived from RT synthesis into the pegRNA scaffold [132]. Although the optimization of gene editing tools and strategies should ideally take place in the target cell types of interest, each of which bearing its specific epigenome, these investigations are rendered difficult due to the fact that latest-generation gene editing tools are becoming even larger than the original Cas9:gRNA complexes. Indeed, prime editors and base editors consist of a bulky Cas9 nickase fused to one and two, respectively, heterologous proteins that must work together as large macromolecular machines [67, 131, 132, 153, 154]. Therefore, there is a pressing need for developing and testing delivery vehicles that can introduce such large machines into primary human cells so that their performance and interaction with human (epi)genomes can be thoroughly investigated. In this context, the research reviewed herein on the testing and use of AdV systems for the targeted genetic modification of stem cells, progenitor cells, and their progeny, supports the view that these agents will become increasingly applied for achieving flexible gene-editing tool delivery and precise gene-editing outcomes in human cells. Defining features underpinning the suitability of AdVs for investigating new gene-editing modalities include their efficient transduction of cycling and quiescent cells, amenability to tropism modifications, high genetic stability and strict episomal nature. Moreover, in the case of HC-AdVs, the absence of viral genes and vast packaging capacity (i.e., up to 36 kb) makes this platform particularly suited for ferrying into cells large genetic payloads for testing precision gene-editing principles based on the recruitment of the HDR pathway or the delivery of DNA-editing fusion constructs, e.g., base

and prime editors.

Acknowledgements

The authors thank Prof. Rob Hoeben (Department of Cell and Chemical Biology, LUMC, The Netherlands) for providing the transmission electron micrograph of the adenovirus serotype 5 virion and for critically reading the manuscript. The authors also thank Dr. Marcella Dias Brescia and Hidde Zittersteijn (both from the Department of Cell and Chemical Biology, LUMC, The Netherlands) for critically reading the manuscript.

Additional information

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 765269 (IMGENE – Improving Genome Editing Efficiency). The research in our laboratory is further supported by the Dutch Prinses Beatrix Spierfonds (W.OR11-18) and the Duchenne Parent Project NL (17.012). Q.W. is the recipient of a Ph.D. fellowship from the China Scholarship Council-Leiden University Joint Scholarship Program.

Conflict of interest statement: None declared.

References

1. Chang, H.H.Y.; Pannunzio, N.R.; Adachi, N.; Lieber, M.R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell. Biol.* **2017**, *18*, 495–506.
2. Kim, Y.; Kweon, J.; Kim, J.S. TALENs and ZFNs are associated with different mutation signatures. *Nat. Methods* **2013**, *10*, 185.
3. Allen, F.; Crepaldi, L.; Alsinet, C.; Strong, A.J.; Kleshchevnikov, V.; De Angeli, P.; Páleníková, P.; Khodak, A.; Kiselev, V.; Kosicki, M.; et al. Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat. Biotechnol.* **2019**, *37*, 64–72.
4. Ata, H.; Ekstrom, T.L.; Martínez-Gálvez, G.; Mann, C.M.; Dvornikov, A.V.; Schaefbauer, K.J.; Ma, A.C.; Dobbs, D.; Clark, K.J.; Ekker, S.C. Robust activation of microhomology-mediated end joining for precision gene editing applications. *PLoS Genet*. **2018**, *14*, e1007652.
5. Taheri-Ghahfarokhi, A.; Taylor, B.J.M.; Nitsch, R.; Lundin, A.; Cavallo, A.L.; Madeyski-Bengtson, K.; Karlsson, F.; Clausen, M.; Hicks, R.; Mayr, L.M.; et al. Decoding non-random mutational signatures at Cas9 targeted sites. *Nucleic Acids Res*. **2018**, *46*, 8417–8434.
6. Shen, M.W.; Arbab, M.; Hsu, J.Y.; Worstell, D.; Culbertson, S.J.; Krabbe, O.; Cassa, C.A.; Liu, D.R.; Gifford, D.K.; Sherwood, R.I. Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **2018**, *563*, 646–651.
7. van Overbeek, M.; Capurso, D.; Carter, M.M.; Thompson, M.S.; Frias, E.; Russ, C.; Reece-Hoyes, J.S.; Nye, C.; Gradia, S.; Vidal, B.; et al. DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks. *Mol. Cell* **2016**, *63*, 633–646.
8. Ousterout, D.G.; Perez-Pinera, P.; Thakore, P.I.; Kabadi, A.M.; Brown, M.T.; Qin, X., Fedrigo, O.; Mouly, V.; Tremblay, J.P.; Gersbach, C.A. Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol. Ther*. **2013**, *21*, 1718–1726.
9. Chen, X.; Gonçalves, M.A. Engineered Viruses as Genome Editing Devices. *Mol. Ther*. **2016**, *24*, 447–457.
10. Doudna, J.A.; Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **2014**, *346*, 1258096.
11. Chandrasegaran, S.; Carroll, D. Origins of Programmable Nucleases for Genome Engineering. *J. Mol. Biol.* **2016**, *428*, 963–989.
12. El-Brolosy, M.A.; Kontarakis, Z.; Rossi, A.; Kuenne, C.; Günther, S.; Fukuda, N.; Kikhi, K.; Boezio, G.L.M.; Takacs, C.M.; et al. Genetic compensation triggered by mutant mRNA degradation. *Nature* **2019**, *568*, 193–197.
13. Ma, Z.; Zhu, P.; Shi, H.; Guo, L.; Zhang, Q.; Chen, Y.; Chen, S.; Zhang, Z.; Peng, J.; Chen, J. PTC-bearing mRNA

elicits a genetic compensation response via Upf3a and COMPASS components. *Nature* **2019**, *568*, 259–263.
14. Tuladhar, R.; Yeu, Y.; Piazza, J.T.; Tan, Z.; Clemenceau, J.R.; Wu, X.; Barrett, Q.; Herbert, J.; Mathews, D.H.; Kim, J.; et al. CRISPR–Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nat. Commun*. **2019**; *10*, 4056.
15. Söllü, C.; Pars, K.; Cornu, T.I.; Thibodeau-Beganny, S.; Maeder, M.L.; Joung, J.K.; Heilbronn, R.; Cathomen, T. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. *Nucleic Acids Res*. **2010**, *38*, 8269–8276.
16. Ousterout, D.G.; Kabadi, A.M.; Thakore, P.I.; Majoros, W.H.; Reddy, T.E.; Gersbach, C.A. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat. Commun*. **2015**, *6*, 6244.
17. Maggio, I.; Stefanucci, L.; Janssen, J.M.; Liu, J., Chen, X.; Mouly, V.; Gonçalves, M.A. Selection-free gene repair after adenoviral vector transduction of designer nucleases: Rescue of dystrophin synthesis in DMD muscle cell populations. *Nucleic Acids Res*. **2016**, *44*, 1449–1470.
18. Choi, P.S.; Meyerson, M. Targeted genomic rearrangements using CRISPR/Cas technology. *Nat. Commun*. **2014**, *24*, 5:3728.
19. Yamamoto, Y.; Gerbi, S.A. Making ends meet: Targeted integration of DNA fragments by genome editing. *Chromosoma* **2018**, *127*, 405–420.
20. Scully, R.; Panday, A.; Elango, R.; Willis, N.A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell. Biol.* **2019**, *20*, 698–714.
21. Deng, C.; Capecchi, M.R. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell. Biol.* **1992**, *12*, 3365–3371.
22. Porter, A.C.; Itzhaki, J.E. Gene targeting in human somatic cells. Complete inactivation of an interferon-inducible gene. *Eur. J. Biochem*. **1993**, *218*, 273–281.
23. Ganguly, A.; Smelt, S.; Mewar, R.; Fertala, A.; Sieron, A.L.; Overhauser, J.; Prockop, D.J. Targeted insertions of two exogenous collagen genes into both alleles of their endogenous loci in cultures human cells: The insertions are directed by relatively short fragments containing the promoters and the 5' ends of the genes. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7365–7369.
24. Thyagarajan, B.; Johnson, B.L.; Campbell, C. The effect of target site transcription on gene targeting in human cells in vitro. *Nucleic Acids Res*. **1995**, *23*, 2784–2790.
25. Lukacsovich, T.; Yang, D.; Waldman, A.S. Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. *Nucleic Acids Res*. **1994**, *22*, 5649–5657.
26. Rouet, P.; Smih, F.; Jasin, M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6064–6068.
27. Rouet, P.; Smih, F.; Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol*. **1994**, *14*, 8096–8106.
28. Maggio, I.; Gonçalves, M.A. Genome editing at the crossroads of delivery, specificity, and fidelity. *Trends Biotechnol*. **2015**, *33*, 280–291.
29. Kim, Y.G.; Cha, J.; Chandrasegaran, S. Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1156–1160.
30. Christian, M.; Cermak, T.; Doyle, E.L.; Schmidt, C.; Zhang, F.; Hummel, A.; Bogdanove, A.J.; Voytas, D.F. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **2010**, *186*, 757–761.
31. Li, T.; Huang, S.; Zhao, X.; Wright, D.A.; Carpenter, S.; Spalding, M.H.; Weeks, D.P.; Yang, B. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res*. **2011**, *39*, 6315–6325.
32. Mahfouz, M.M.; Li, L.; Shamimuzzaman, M.; Wibowo, A.; Fang, X.; Zhu, J.K. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2623–2628.
33. Miller, J.C.; Tan, S.; Qiao, G.; Barlow, K.A.; Wang, J.; Xia, D.F.; Meng, X.; Paschon, D.E.; Leung, E.; Hinkley, S.J.; et al. A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol*. **2011**, *29*, 143–148.
34. Mussolino, C.; Morbitzer, R.; Lütge, F.; Dannemann, N.; Lahaye, T.; Cathomen, T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res*. **2011**, *39*, 9283–9293.
35. Cho, S.W.; Kim, S.; Kim, J.M.; Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol*. **2013**, *31*, 230–232.
36. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **2013**, *339*, 819–823.
37. Jinek, M.; East, A.; Cheng, A.; Lin, S.; Ma, E.; Doudna, J. RNA-programmed genome editing in human cells. *Elife* **2013**, *2*, e00471.
38. Mali, P.; Yang, L.; Esvelt, K.M.; Aach, J.; Guell, M.; DiCarlo, J.E.; Norville, J.E.; Church, G.M. RNA-Guided human genome engineering via Cas9. *Science* **2013**, *339*, 823–826.

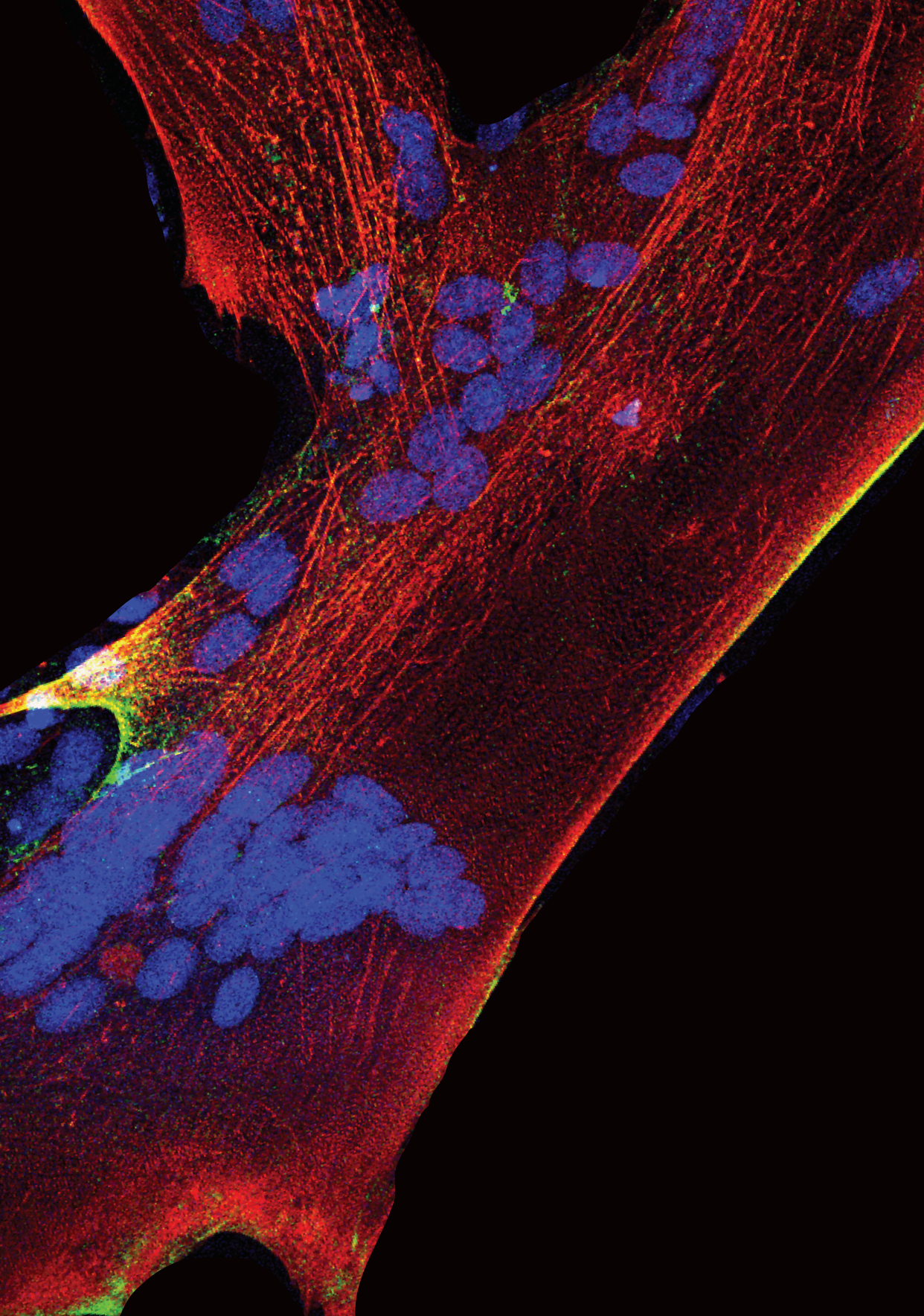
39. Miller, J.; McLachlan, A.D.; Klug, A. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J*. **1985**, *4*, 1609–1614.
40. Boch, J.; Scholze, H.; Schornack, S.; Landgraf, A.; Hahn, S.; Kay, S.; Lahaye, T.; Nickstadt, A.; Bonas, U. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **2009**, *326*, 1509–1512.
41. Moscou, M.J.; Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. *Science* **2009**, *326*, 1501.
42. Gasiunas, G.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2579–E2586.
43. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821.
44. Bitinaite, J.; Wah, D.A.; Aggarwal, A.K.; Schildkraut, I. FokI dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10570–10575.
45. Klug, A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu. Rev. Biochem*. **2010**, *79*, 213–231.
46. Ecco, G.; Imbeault, M.; Trono, D. KRAB zinc finger proteins. *Development* **2017**, *144*, 2719–2729.
47. Boch, J.; Bonas, U. Xanthomonas AvrBs3 family-type III effectors: Discovery and function. *Annu. Rev. Phytopathol*. **2010**, *48*, 419–436.
48. Cathomen, T.; Joung, K.J. Zinc-finger Nucleases: The next generation emerges. *Mol. Ther*. **2008**, *16*, 1200–1207.
49. Mussolino, C.; Cathomen, T. TALE nucleases: Tailored genome engineering made easy. *Curr. Opin. Biotechnol*. **2012**, *23*, 644–650.
50. Bultmann, S.; Morbitzer, R.; Schmidt, C.S.; Thanisch, K.; Spada, F.; Elsaesser, J.; Lahaye, T.; Leonhardt, H. Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res*. **2012**, *40*, 5368–5377.
51. Valton, J.; Dupuy, A.; Daboussi, F.; Thomas, S.; Maréchal, A.; Macmaster, R.; Melliand, K.; Juillerat, A.; Duchateau, P. Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J. Biol. Chem*. **2012**, *287*, 38427–38432.
52. Chen, X.; Rinsma, M.; Janssen, J.M.; Liu, J.; Maggio, I.; Gonçalves MA. Probing the impact of chromatin conformation on genome editing tools. *Nucleic Acids Res*. **2016**, *44*, 6482–6492.
53. Makarova, K.S.; Wolf, Y.I.; Iranzo, J.; Shmakov, S.A.; Alkhnbashi, O.S.; Brouns, S.J.J.; Charpentier, E.; Cheng, D.; Haft, D.H.; Horvath, P.; et al. Evolutionary classification of CRISPR-Cas systems: A burst of class 2 and derived variants. *Nat. Rev. Microbiol*. **2020**, *18*, 67–83.
54. Anders, C.; Niewoehner, O.; Duerst, A.; Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **2014**, *513*, 569–573.
55. Shibata, M.; Nishimasu, H.; Kodera, N.; Hirano, S.; Ando, T.; Uchihashi, T.; Nureki, O. Real-space and real-time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy. *Nat. Commun*. **2017**, *8*, 1430.
56. Cradick, T.J.; Fine, E.J.; Antico, C.J.; Bao, G. CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res*. **2013**, *41*, 9584–9592.
57. Hsu, P.D.; Scott, D.A.; Weinstein, J.A.; Ran, F.A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E.J.; Wu, X.; Shalem, O.; et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol*. **2013**, *31*, 827–832.
58. Fu, Y.; Foden, J.A.; Khayter, C.; Maeder, M.L.; Reyon, D.; Joung, J.K.; Sander, J.D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol*. **2013**, *31*, 822–826.
59. Lin, Y.; Cradick, T.J.; Brown, M.T.; Deshmukh, H.; Ranjan, P.; Sarode, N.; Wile, B.M.; Vertino, P.M.; Stewart, F.J.; Bao, G. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res*. **2014**, *42*, 7473–7485.
60. Frock, R.L.; Hu, J.; Meyers, R.M.; Ho, Y.J.; Kii, E.; Alt, F.W. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol*. **2015**, *33*, 179–186.
61. Chen, X.; Tasca, F.; Wang, Q.; Liu, J.; Janssen, J.M.; Brescia, M.D.; Bellin, M.; Szuhai, K.; Kenrick, J.; Frock, R.L.; et al. Expanding the editable genome and CRISPR-Cas9 versatility using DNA cutting-free gene targeting based on in trans paired nicking. *Nucleic Acids Res*., **2020**, *48*, 974–995.
62. Cho, S.W.; Kim, S.; Kim, Y.; Kweon, J.; Kim, H.S.; Bae, S.; Kim, J.S. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. **2014**, *24*, 132–141.
63. Akcakaya, P.; Bobbin, M.L.; Guo, J.A.; Malagon-Lopez, J.; Clement, K.; Garcia, S.P.; Fellows, M.D.; Porritt, M.J.; Firth, M.A.; Carreras, A.; et al. In vivo CRISPR editing with no detectable genome-wide off-target mutations. *Nature*, **2018**, *561*, 416–419.
64. Chen, X., Liu, J., Janssen, J.M., Gonçalves, M.A. The chromatin structure differentially impacts high-specificity CRISPR-Cas9 nuclease strategies. *Mol. Ther. Nucleic Acids*, **2017**, *8*, 558–563.
65. Daer, R.M., Cutts, J.P., Brafman, D.A., Haynes, K.A. The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. *ACS Synth. Biol*., **2017**, *6*, 428–438.
66. Jensen, K.T.; Fløe, L.; Petersen, T.S.; Huang, J.; Xu, F.; Bolund, L.; Luo, Y.; Lin, L. Chromatin accessibility

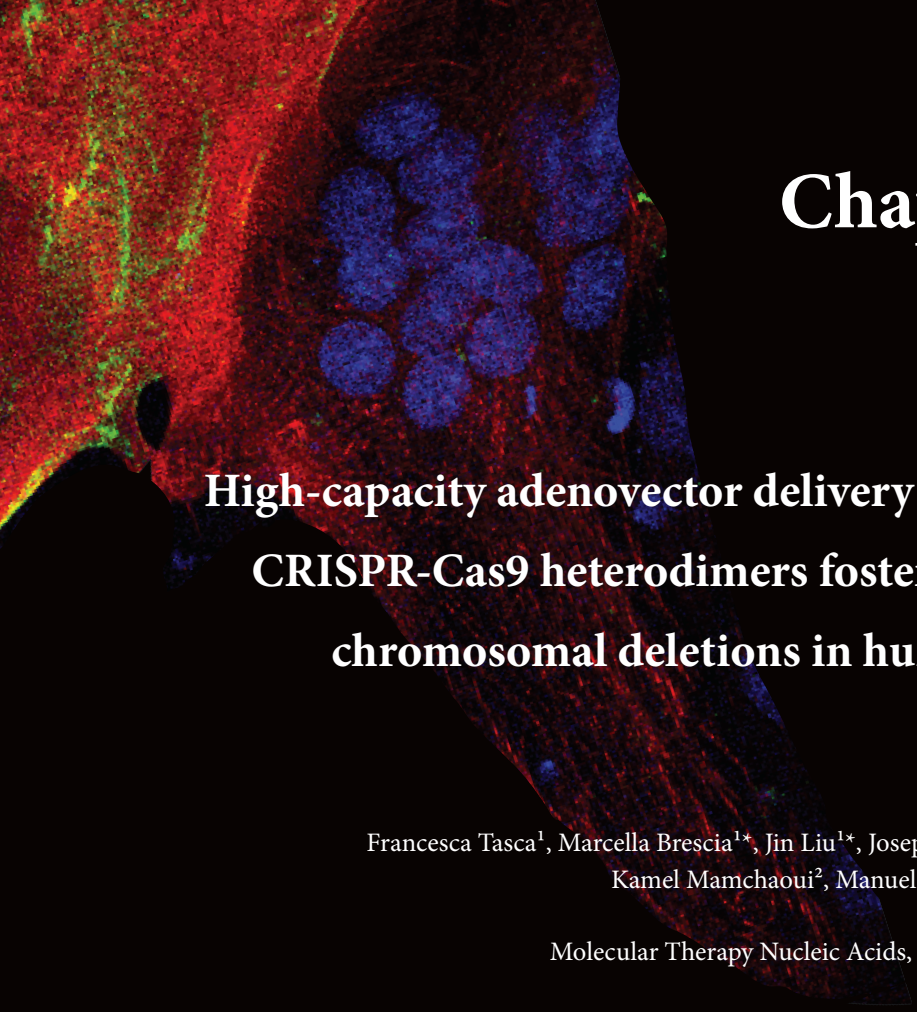
and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett.*, **2017**, *591*, 1892–1901.
67. Chen, X.; Gonçalves, M.A. DNA, RNA, and Protein Tools for Editing the Genetic Information in Human Cells. *iScience* **2018**, *6*, 247–263.
68. Brooks, A.K.; Gaj, T. Innovations in CRISPR technology. *Curr. Opin. Biotechnol.* **2018**, *52*, 95–101.
69. Hajizadeh Dastjerdi, A.; Newman, A.; Burgio, G. The Expanding Class 2 CRISPR Toolbox: Diversity, Applicability, and Targeting Drawbacks. *BioDrugs* **2019**, *33*, 503–513.
70. Wold, W.S.M.; Horwitz, M.S. Adenoviruses. *Fields Virology* (5th ed.), **2007**, Lippincott Williams & Wilkins, Philadelphia, PA.
71. Gonçalves, M.A.; de Vries, A.A. Adenovirus: From foe to friend. *Rev. Med. Virol.* **2006**, *16*, 167–186.
72. Russell, W.C. Adenoviruses: Update on structure and function. *J. Gen. Virol.* **2009**, *90*, 1–20.
73. Vellinga, J.; van der Heijdt, S.; Hoeben, R.C. The adenovirus capsid: Major progress in minor proteins. *J. Gen. Virol.* **2005**, *86*, 1581–1588.
74. Bergelson, J.M.; Cunningham, J.A.; Droguett, G.; Kurt-Jones, E.A.; Krithivas, A.; Hong, J.S.; Horwitz, M.S.; Crowell, R.L.; Finberg, R.W. Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science* **1997**, *275*, 1320–1323.
75. Tomko, R.P.; Xu, R.; Philipson, L. HCAR and MCAR: The human and mouse cellular receptors for subgroup C adenoviruses and group B coxsackieviruses. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 3352–3356.
76. Gaggar, A.; Shayakhmetov, D.M.; Lieber, A. CD46 is a cellular receptor for group B adenoviruses. *Nat. Med.* **2003**, *9*, 1408–1412.
77. Wang, H.; Li, Z.Y.; Liu, Y.; Persson, J.; Beyer, I.; Möller, T.; Koyuncu, D.; Drescher, M.R.; Strauss, R.; Zhang, X.B.; et al. Desmoglein 2 is a receptor for adenovirus serotypes 3, 7, 11 and 14. *Nat. Med.* **2011**, *17*, 96–104.
78. Nilsson, E.C.; Storm, R.J.; Bauer, J.; Johansson, S.M.; Lookene, A.; Ångström, J.; Hedenström, M.; Eriksson, T.L.; Frängsmyr, L.; Rinaldi, S.; et al. The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. *Nat. Med.* **2011**, *17*, 105–109.
79. Arnberg, N., Edlund, K., Kidd, A.H., Wadell, G. Adenovirus type 37 uses sialic acid as a cellular receptor. *J. Virol.* **2000**, *74*, 42–48.
80. Alonso-Padilla, J.; Papp, T.; Kaján, G.L.; Benkő, M.; Havenga, M.; Lemckert, A.; Harrach, B.; Baker, A.H. Development of Novel Adenoviral Vectors to Overcome Challenges Observed With HAdV-5-based Constructs. *Mol. Ther.* **2016**, *24*, 6–16.
81. Barry, M.A.; Rubin, J.D.; Lu, S.C. Retargeting adenoviruses for therapeutic applications and vaccines. *FEBS Lett.* **2020**. doi: 10.1002/1873-3468.13731.
82. Shayakhmetov, D.M.; Papayannopoulou, T.; Stamatoyannopoulos, G.; Lieber, A. Efficient gene transfer into human CD34+ cells by a retargeted adenovirus vector. *J. Virol.*, **2000**, *74*, 2567–2583.
83. Knaän-Shanzer, S.; Van Der Velde, I.; Havenga, M.J.E.; Lemckert, A.A.C.; De Vries, A.A.F.; Valerio, D. Highly efficient targeted transduction of undifferentiated human hematopoietic cells by adenoviral vectors displaying fiber knobs of subgroup B. *Hum. Gene Ther.*, **2001**, *12*, 1989–2005.
84. Gonçalves, M.A.; de Vries, A.A.; Holkers, M.; van de Watering, M.J.M.; van der Velde, I.; van Nierop, G.P.; Valerio, D.; Knaän-Shanzer, S. Human mesenchymal stem cells ectopically expressing full-length dystrophin can complement Duchenne muscular dystrophy myotubes by cell fusion. *Hum. Mol. Genet.*, **2005**, *15*, 213–221.
85. Knaän-Shanzer, S.; van de Watering, M.J.M.; van der Velde, I.; Gonçalves, M.A.F.V.; Valerio, D.; de Vries, A.A.F. Endowing human adenovirus serotype 5 vectors with fiber domains of species B greatly enhances gene transfer into human mesenchymal stem cells. *Stem Cells* **2005**, *23*, 1598–1607.
86. Gonçalves, M.A.; Holkers, M.; Cudré-Mauroux, C.; Van Nierop, G.P.; Knaän-Shanzer, S.; Van Der Velde, I.; Valerio, D.; De Vries, A.A. Transduction of myogenic cells by retargeted dual high-capacity hybrid viral vectors: Robust dystrophin synthesis in duchenne muscular dystrophy muscle cells. *Mol. Ther.*, **2006**, *13*, 976–986.
87. Meier, O.; Greber, U.F. Adenovirus endocytosis. *J. Gene Med.* **2004**, *6*, S152–S163.
88. Graham, F.L.; Smiley, J.; Russell,W.C.; Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **1977**, *36*, 59–74.
89. Fallaux, F.J.; Bout, A.; van der Velde, I.; van den Wollenberg, D.J.; Hehir, K.M.; Keegan, J.; Auger, C.; Cramer, S.J.; van Ormondt, H.; van der Eb, A.J.; et al. New helper cells and matched early region 1-deleted adenovirus vectors prevent generation of replication-competent adenoviruses. *Hum. Gene Ther.* **1998**, *9*, 1909–1917.
90. Shirley, J.L.; de Jong, Y.P.; Terhorst, C.; Herzog, R.W. Immune Responses to Viral Gene Therapy Vectors. *Mol. Ther.* **2020**, *28*, 709–722.
91. Shukarev, G.; Callendret, B.; Luhn, K.; Douoguih, M. A two-dose heterologous prime-boost vaccine regimen eliciting sustained immune responses to Ebola Zaire could support a preventive strategy for future outbreaks. *Hum. Vaccin. Immunother.* **2017**, *13*, 266–270.
92. Barouch, D.H.; Tomaka, F.L.; Wegmann, F.; Stieh, D.J.; Alter, G.., Robb, M.L.; Michael, N.L.; Peter, L.; Nkolola, J.P.; Borducchi, E.N.; et al. 2018. Evaluation of a mosaic HIV-1 vaccine in a multicentre, randomised, double-blind, placebo-controlled, phase 1/2a clinical trial (APPROACH) and in rhesus monkeys (NHP 13-19).

Lancet **2018**, *392*, 232–243.
93. Havenga, M.J.; Holterman, L.; Melis, I.; Smits, S.; Kaspers, J.; Heemskerk, E.; van der Vlugt, R.; Koldijk, M.; Schouten, G.J.; Hateboer, G.; et al. Serum-free transient protein production system based on adenoviral vector and PER.C6 technology: High yield and preserved bioactivity. *Biotechnol. Bioeng*. **2008**, *100*, 273–283.
94. Parks, R.J.; Graham, F.L. A helper-dependent system for adenovirus vector production helps define a lower limit for efficient DNA packaging. *J. Virol*. **1997**, *71*, 3293–3298.
95. Jager, L.; Hausl, M.A.; Rauschhuber, C.; Wolf, N.M.; Kay, M.A.; Ehrhardt, A. A rapid protocol for construction and production of high-capacity adenoviral vectors. *Nat. Protoc*. **2009**, *4*, 547–564.
96. Perez, E.E.; Wang, J.; Miller, J.C.; Jouvenot, Y.; Kim, K.A.; Liu, O.; Wang, N.; Lee, G.; Bartsevich, V.V.; Lee, Y.L.; et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **2008**, *26*, 808–816.
97. Li, L.; Krymskaya, L.; Wang, J.; Henley, J.; Rao, A.; Cao, L.-F.; Tran, C.-A.; Torres-Coronado, M.; Gardner, A.; Gonzalez, N.; et al. Genomic Editing of the HIV-1 Coreceptor CCR5 in Adult Hematopoietic Stem and Progenitor Cells Using Zinc Finger Nucleases. *Mol. Ther.* **2013**, *21*, 1259–1269.
98. Maier, D.A.; Brennan, A.L.; Jiang, S.; Binder-Scholl, G.K.; Lee, G.; Plesa, G.; Zheng, Z.; Cotte, J.; Carpenito, C.; Wood, T.; et al. Efficient clinical scale gene modification via zinc finger nuclease-targeted disruption of the HIV co-receptor CCR5. *Hum. Gene Ther.* **2013**, *24*, 245–258.
99. Saydaminova, K.; Ye, X.; Wang, H.; Richter, M.; Ho, M.; Chen, H.; Xu, N.; Kim, J.S.; Papapetrou, E.; Holmes, M.C.; et al. Efficient genome editing in hematopoietic stem cells with helper-dependent Ad5/35 vectors expressing site-specific endonucleases under microRNA regulation. *Mol. Ther. Methods Clin. Dev.* **2015**, *1*, 14057.
100. Tebas, P.; Stein, D.; Tang, W.W.; Frank, I.; Wang, S.Q.; Lee, G.; Spratt, S.K.; Surosky, R.T.; Giedlin, M.A.; Nichol, G.; et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* **2014**, *370*, 901–910.
101. Provasi, E.; Genovese, P.; Lombardo, A.; Magnani, Z.; Liu, P.Q.; Reik, A.; Chu, V.; Paschon, D.E.; Zhang, L.; Kuball, J.; et al. Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer. *Nat. Med.* **2012**, *18*, 807–815.
102. Yuan, J.; Wang, J.; Crain, K.; Fearns, C.; Kim, K.A.; Hua, K.L.; Gregory, P.D.; Holmes, M.C.; Torbett, B.E. Zinc-finger Nuclease Editing of Human cxcr4 Promotes HIV-1 CD4+ T Cell Resistance and Enrichment. *Mol. Ther.* **2012**, *20*, 849–859.
103. Holkers, M.; Maggio, I.; Liu, J.; Janssen, J.M.; Miselli, F.; Mussolino, C.; Recchia, A.; Cathomen, T.; Gonçalves, M.A. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res.* **2013**, *41*, e63.
104. Yang, L.; Guell, M.; Byrne, S.; Yang, J.L.; De Los Angeles, A.; Mali, P.; Aach, J.; Kim-Kiselak, C.; Briggs, A.W.; Rios, X.; et al. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* **2013**, *41*, 9049–9061.
105. Fang, Y.; Stroukov, W.; Cathomen, T.; Mussolino, C. Chimerization Enables Gene Synthesis and Lentiviral Delivery of Customizable TALE-Based Effectors. *Int. J. Mol. Sci.* **2020**, *21*, 795.
106. Pelascini, L.P.; Maggio, I.; Liu, J.; Holkers, M.; Cathomen, T.; Gonçalves, M.A. Histone deacetylase inhibition rescues gene knockout levels achieved with integrase-defective lentiviral vectors encoding zinc-finger nucleases. *Hum. Gene Ther. Methods* **2013**, *24*, 399–411.
107. Pelascini, L.P.; Janssen, J.M.; Gonçalves, M.A. Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. *Hum. Gene Ther.* **2013**, *24*, 78–96.
108. Maggio, I.; Holkers, M.; Liu, J.; Janssen, J.M.; Chen, X.; Gonçalves, M.A. Adenoviral vector delivery of RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells. *Sci. Rep.* **2014**, *4*, 5105.
109. Gong, H.; Liu, M.; Klomp, J.; Merrill, B.J.; Rehman, J.; Malik, A.B. Method for Dual Viral Vector Mediated CRISPR-Cas9 Gene Disruption in Primary Human Endothelial Cells. *Sci. Rep.* **2017**, *7*, 42127.
110. Maggio, I.; Liu, J.; Janssen, J.M.; Chen, X.; Gonçalves, M.A. Adenoviral vectors encoding CRISPR/Cas9 multiplexes rescue dystrophin synthesis in unselected populations of DMD muscle cells. *Sci. Rep.* **2016**, *6*, 37051.
111. Brescia, M.; Janssen, J.M.; Liu, J.; Gonçalves, M.A. High-capacity adenoviral vectors permit robust and versatile testing of DMD gene repair tools and strategies in human cells. *Cells* **2020**, *9*, 869.
112. Li, C.; Guan, X.; Du, T.; Jin, W.; Wu, B.; Liu, Y.; Wang, P.; Hu, B.; Griffin, G.E.; Shattock, R.J.; et al. Inhibition of HIV-1 infection of primary CD4+ T-cells by gene editing of CCR5 using adenovirus-delivered CRISPR/Cas9. *J. Gen. Virol.* **2015**, *96*, 2381–2393.
113. Li, C.; Psatha, N.; Sova, P.; Gil, S.; Wang, H.; Kim, J.; Kulkarni, C.; Valensisi, C.; Hawkins, R.D.; Stamatoyannopoulos, G.; et al. Reactivation of gamma-globin in adult beta-YAC mice after ex vivo and in vivo hematopoietic stem cell genome editing. *Blood* **2018**, *131*, 2915–2928.
114. Li, C.; Psatha, N.; Gil, S.; Wang, H.; Papayannopoulou, T.; Lieber, A. HDAd5/35(++) Adenovirus Vector Expressing Anti-CRISPR Peptides Decreases CRISPR/Cas9 Toxicity in Human Hematopoietic Stem Cells. *Mol.*

Ther. Methods Clin. Dev. **2018**, *9*, 390–401.
115. Schiroli, G.; Conti, A.; Ferrari, S.; Della Volpe, L.; Jacob, A.; Albano, L.; Beretta, S.; Calabria, A.; Vavassori, V.; Gasparini, P.; et al. Precise Gene Editing Preserves Hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. *Cell Stem Cell* **2019**, *24*, 551–565.
116. Coluccio, A.; Miselli, F.; Lombardo, A.; Marconi, A.; Malagoli Tagliazucchi, G.; Gonçalves, M.A.; Pincelli, C.; Maruggi, G.; Del Rio, M.; Naldini, L.; et al. Targeted gene addition in human epithelial stem cells by zinc-finger nuclease-mediated homologous recombination. *Mol. Ther.* **2013**, *21*, 1695–1704.
117. Holkers, M.; Maggio, I.; Henriques, S.F.; Janssen, J.M.; Cathomen, T.; Gonçalves, M.A. Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat. Methods* **2014**, *11*, 1051–1057.
118. Li, C.; Mishra, A.S.; Gil, S.; Wang, M.; Georgakopoulou, A.; Papayannopoulou, T.; Hawkins, R.D.; Lieber, A. Targeted Integration and High-Level Transgene Expression in AAVS1 Transgenic Mice after In Vivo HSC Transduction with HDAd5/35++ Vectors. *Mol. Ther.* **2019**, *27*, 2195–2212.
119. Thomson, J.A.; Itskovitz-Eldor, J.; Shapiro, S.S.; Waknitz, M.A.; Swiergiel, J.J.; Marshall, V.S.; Jones, J.M. Embryonic stem cell lines derived from human blastocysts. *Science* **1998**, *282*, 1145–1147.
120. de Miguel-Beriain, I. The ethics of stem cells revisited. *Adv. Drug Deliv. Rev.* **2015**, *82-83*, 176-180.
121. Hu, J.; Wang, J. From embryonic stem cells to induced pluripotent stem cells-Ready for clinical therapy? *Clin. Transplant* **2019**, *33*, e13573.
122. Takahashi, K.; Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**, *126*, 663–676.
123. Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **2007**, *131*, 861–872.
124. Doss MX, Sachinidis, A. Current Challenges of iPSC-Based Disease Modeling and Therapeutic Implications. *Cells* **2019**, *8*, 403.
125. Kimbrel, E.A.; Lanza, R. Pluripotent stem cells: The last 10 years. *Regen. Med.* **2016**, *11*, 831–847.
126. Hockemeyer, D.; Jaenisch, R. Induced Pluripotent Stem Cells Meet Genome Editing. *Cell Stem Cell* **2016**, *18*, 573–586.
127. Li, M.; Suzuki, K.; Kim, N.Y.; Liu, G.H.; Izpisua Belmonte, J.C. A cut above the rest: Targeted genome editing technologies in human pluripotent stem cells. *J. Biol. Chem.* **2014**, *289*, 4594–4599.
128. Chen, X.; Janssen, J.M.; Liu, J.; Maggio, I.; t Jong, A.E.J.; Mikkers, H.M.M.; Gonçalves, M. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat. Commun.* **2017**, *8*, 657.
129. Nakajima, K.; Zhou, Y.; Tomita, A.; Hirade, Y.; Gurumurthy, C.B.; Nakada, S. Precise and efficient nucleotide substitution near genomic nick via noncanonical homology-directed repair. *Genome Res.* **2018**, *28*, 223–230.
130. Hyodo, T.; Rahman, M.L.; Karnan, S.; Ito, T.; Toyoda, A.; Ota, A.; Wahiduzzaman, M.; Tsuzuki, S.; Okada, Y.; Hosokawa, Y.; et al. Tandem Paired Nicking Promotes Precise Genome Editing with Scarce Interference by p53. *Cell Rep.* **2020**, *30*, 1195–1207.
131. Rees, H.A.; Liu, D.R. Base editing: Precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **2018**, *19*, 770–788.
132. Anzalone, A.V.; Randolph, P.B.; Davis, J.R.; Sousa, A.A.; Koblan, L.W.; Levy, J.M.; Chen, P.J.; Wilson, C.; Newby, G.A.; Raguram, A.; et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **2019**, *576*, 149–157.
133. Haapaniemi, E.; Botla, S.; Persson, J.; Schmierer, B.; Taipale, J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **2018**, *24*, 927–930.
134. Ihry, R.J.; Worringer, K.A.; Salick, M.R.; Frias, E.; Ho, D.; Theriault, K.; Kommineni, S.; Chen, J.; Sondey, M.; Ye, C.; et al. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **2018**, *24*, 939–946.
135. Gonçalves, M.A.; van Nierop, G.P.; Holkers, M.; de Vries AA. Concerted nicking of donor and chromosomal acceptor DNA promotes homology-directed gene targeting in human cells. *Nucleic Acids Res.* **2012**, *40*, 3443–3455.
136. Rui, Y.; Wilson, D.R.; Green, J.J. Non-Viral Delivery To Enable Genome Editing. *Trends Biotechnol.* **2019**, *37*, 281–293.
137. Ohbayashi, F.; Balamotis, M.A.; Kishimoto, A.; Aizawa, E.; Diaz, A.; Hasty, P.; Graham, F.L.; Caskey, C.T.; Mitani, K. Correction of chromosomal mutation and random integration in embryonic stem cells with helper-dependent adenoviral vectors. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13628–13633.
138. Suzuki, K.; Mitsui, K.; Aizawa, E.; Hasegawa, K.; Kawase, E.; Yamagishi, T.; Shimizu, Y.; Suemori, H.; Nakatsuji, N.; Mitani, K. Highly efficient transient gene expression and gene targeting in primate embryonic stem cells with helper-dependent adenoviral vectors. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13781–13786.
139. Aizawa, E.; Hirabayashi, Y.; Iwanaga, Y.; Suzuki, K.; Sakurai, K.; Shimoji, M.; Aiba, K.; Wada, T.; Tooi, N.; Kawase, E.; et al. Efficient and accurate homologous recombination in hESCs and hiPSCs using helper-dependent adenoviral vectors. *Mol. Ther.* **2012**, *20*, 424–431.
140. Umeda, K.; Suzuki, K.; Yamazoe, T.; Shiraki, N.; Higuchi, Y.; Tokieda, K.; Kume, K.; Mitani, K.; Kume, S. Albu-

min gene targeting in human embryonic stem cells and induced pluripotent stem cells with helper-dependent adenoviral vector to monitor hepatic differentiation. *Stem Cell Res.* **2013**, *10*, 179–194.
141. Sone, T.; Shin, M.; Ouchi, T.; Sasanuma, H.; Miyamoto, A.; Ohte, S.; Tsukamoto, S.; Nakanishi, M.; Okano, H.; Katagiri, T.; et al. Dual usage of a stage-specific fluorescent reporter system based on a helper-dependent adenoviral vector to visualize osteogenic differentiation. *Sci. Rep.* **2019**, *9*, 9705.
142. Liu, G.H.; Suzuki, K.; Qu, J.; Sancho-Martinez, I.; Yi, F.; Li, M.; Kumar, S.; Nivet, E.; Kim, J.; Soligalla, R.D.; et al. Targeted gene correction of laminopathy-associated LMNA mutations in patient-specific iPSCs. *Cell Stem Cell* **2011**, *8*, 688–694.
143. Li, M.; Suzuki, K.; Qu, J.; Saini, P.; Dubova, I.; Yi, F.; Lee, J.; Sancho-Martinez, I.; Liu, G.H.; Izpisua Belmonte, J.C. Efficient correction of hemoglobinopathy-causing mutations by homologous recombination in integration-free patient iPSCs. *Cell Res.* **2011**, *21*, 1740–1744.
144. Palmer, D.J.; Grove, N.C.; Ing, J.; Crane, A.M.; Venken, K.; Davis, B.R.; Ng, P. Homology Requirements for Efficient, Footprintless Gene Editing at the CFTR Locus in Human iPSCs with Helper-dependent Adenoviral Vectors. *Mol. Ther. Nucleic Acids* **2016**, *5*, e372.
145. Palmer, D.J.; Grove, N.C.; Turner, D.L.; Ng, P. Gene Editing with Helper-Dependent Adenovirus Can Efficiently Introduce Multiple Changes Simultaneously over a Large Genomic Region. *Mol. Ther. Nucleic Acids* **2017**, *8*, 101–110.
146. Liu, G.H.; Qu, J.; Suzuki, K.; Nivet, E.; Li, M.; Montserrat, N.; Yi, F.; Xu, X.; Ruiz, S.; Zhang, W.; et al. Progressive degeneration of human neural stem cells caused by pathogenic LRRK2. *Nature* **2012**, *491*, 603–607.
147. Liu, G.H.; Suzuki, K.; Li, M.; Qu, J.; Montserrat, N.; Tarantino, C.; Gu, Y.; Yi, F.; Xu, X.; Zhang, W.; et al. Modelling Fanconi anemia pathogenesis and therapeutics using integration-free patient-derived iPSCs. *Nat. Commun.* **2014**, *5*, 4330.
148. Yoshida, T.; Ozawa,,Y.; Suzuki, K.; Yuki, K.; Ohyama, M.; Akamatsu, W.; Matsuzaki, Y.; Shimmura, S.; Mitani, K.; Tsubota, K.; et al. The use of induced pluripotent stem cells to reveal pathogenic gene mutations and explore treatments for retinitis pigmentosa. *Mol. Brain* **2014**, *7*, 45.
149. Zhang, W.; Li, J.; Suzuki, K.; Qu, J.; Wang, P.; Zhou, J.; Liu, X.; Ren, R.; Xu, X.; Ocampo, A.; et al. Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science* **2015**, *348*, 1160–1163.
150. Suzuki, K.; Yu, C.; Qu, J.; Li, M.; Yao, X.; Yuan, T.; Goebl, A.; Tang, S.; Ren, R.; Aizawa, E.; et al. Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell Stem Cell* **2014**, *15*, 31–36.
151. Miller, D.G.; Petek, L.M.; Russell, D.W. Adeno-associated virus vectors integrate at chromosome breakage sites. *Nat. Genet.* **2004**, *36*, 767.
152. Hanlon, K.S.; Kleinstiver, B.P.; Garcia, S.P.; Zaborowski, M.P.; Volak, A.; Spirig, S.E.; Muller, A.; Sousa, A.A.; Tsai, S.Q.; Bengtsson, N.E. High levels of AAV vector integration into CRISPR-induced DNA breaks. *Nat. Commun.* **2019**, *10*, 1–11.
153. Komor, A.C.; Kim, Y.B.; Packer, M.S.; Zuris, J.A.; Liu, D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **2016**, *533*, 420–424.
154. Gaudelli, N.M.; Komor, A.C., Rees, H.A.; Packer, M.S., Badran; A.H., Bryson, D.I.; Liu, D.R. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **2017**, *551*, 464–471.
155. Kosicki, M.; Tomberg, K.; Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **2018**, *36*, 765–771.
156. Park, S.; Beal, P.A. Off-Target Editing by CRISPR-Guided DNA Base Editors. *Biochemistry* **2019**, *58*, 3727–3734.





Chapter 3

High-capacity adenovector delivery of forced CRISPR-Cas9 heterodimers fosters precise chromosomal deletions in human cells

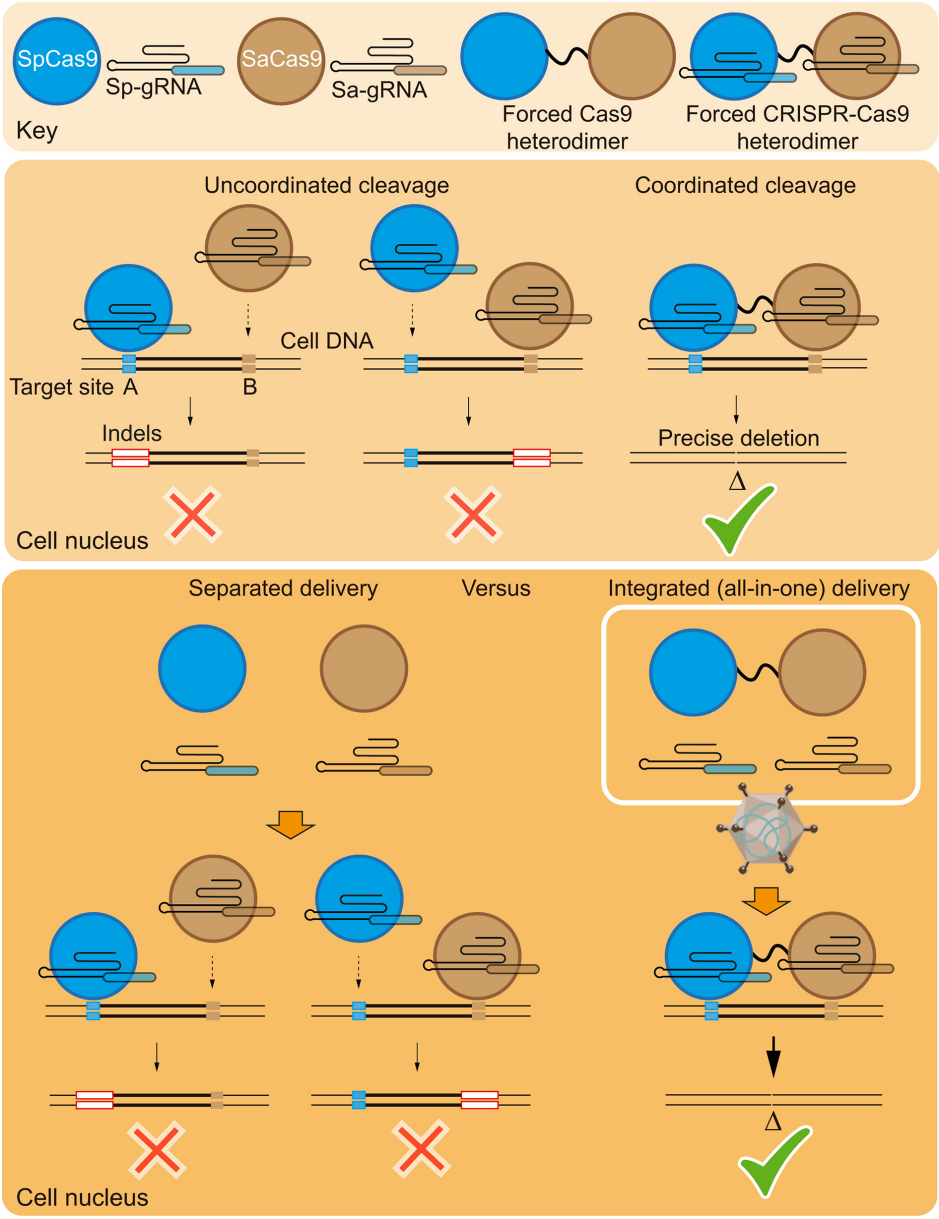
Francesca Tasca[1], Marcella Brescia[1*], Jin Liu[1*], Josephine M. Janssen[1], Kamel Mamchaoui[2], Manuel A.F.V. Gonçalves[1]

Molecular Therapy Nucleic Acids, 2023, 31: 746-762

* These authors contributed equally to this work
[1]Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands
[2]Sorbonne Université, Inserm, Institut de Myologie, Centre de Recherche en Myologie, Paris, France

Illustration: Modified from immunofluorescence picture of patient-derived myotubes after treatment. (See Figure 5D for technical details)

Graphical Abstract



Abstract

Genome editing based on dual CRISPR-Cas9 complexes (multiplexes) permits removing specific genomic sequences in living cells leveraging research on functional genomics and genetic therapies. Delivering the required large and multicomponent reagents in a synchronous and stoichiometric manner remains, however, challenging. Moreover, uncoordinated activity of independently acting CRISPR-Cas9 multiplexes increases the complexity of genome editing outcomes. Here, we investigate the potential of fostering precise multiplexing genome editing using high-capacity adenovector particles (AdVPs) for the delivery of Cas9 ortholog fusion constructs alone (forced Cas9 heterodimers) or together with their cognate guide RNAs (forced CRISPR-Cas9 heterodimers). We demonstrate that the efficiency and accuracy of targeted chromosomal DNA deletions achieved by single AdVPs encoding forced CRISPR-Cas9 heterodimers is superior to that obtained when the various components are delivered separately. Finally, all-in-one AdVP delivery of forced CRISPR-Cas9 heterodimers triggers robust DMD exon 51 splice site excision resulting in reading frame restoration and selection-free detection of dystrophin in muscle cells derived from Duchenne muscular dystrophy patients. In conclusion, AdVPs promote precise multiplexing genome editing through the integrated delivery of forced CRISPR-Cas9 heterodimer components which, in comparison to split conventional CRISPR-Cas9 multiplexes, engage target sequences in a more coordinated fashion.

Introduction

Diverse types of genetic modifications spanning from single base-pairs to mega-bases have been identified as disease-causing genotypes. During the past decade, gene therapies based on prokaryotic type II clustered regularly interspaced short palindromic repeat (CRISPR)-associated Cas9 (CRISPR-Cas9) systems have started to be investigated and tested for the correction or complementation of such genotypes [1-2]. Engineered CRISPR-Cas9 nucleases are ribonucleoprotein complexes consisting of a Cas9 endonuclease and a single guide RNA (gRNA) [3]. The Cas9 protein cleaves target DNA upon recognition of a protospacer adjacent motif (PAM) located next to ~20 nucleotide stretches complementary to the 5' end of the gRNA (protospacer). PAM nucleotide sequences differ from one CRISPR-Cas9 system to another. For instance, the PAM of the prototypic and most used *Streptococcus pyogenes* Cas9 (SpCas9) nuclease is NGG [4]. Its ortholog Cas9 nuclease from *Staphylococcus aureus* (SaCas9) recognizes instead the longer PAM consensus sequence NNGRRT (R=A or G) [5]. The generation of a site-specific double-strand DNA break (DSB) by Cas9:gRNA complexes elicits endogenous DNA repair pathways that can be exploited for targeted genetic modifications [3]. In mammalian cells, a prevalent DNA repair pathway that arises in response to a DSB is the classical non-homologous end joining (NHEJ) [6]. This pathway results in end-to-end ligation of DSBs that, when inaccurate, yields small insertions and deletions (indels) that can lead to targeted gene knockouts [3]. In contrast to homology-directed DSB repair, the prevalence of NHEJ throughout the cell cycle makes its exploitation for genome editing purposes possible in both dividing and post-mitotic cells or tissues [6].

Delivering dual CRISPR-Cas9 complexes formed by a Cas9 nuclease and two gRNAs (multiplexes) addressed to neighbouring target sites, induces NHEJ-mediated intrachromosomal deletions encompassing the sequences located between the site-specific DSBs [7-9]. Of notice, this multiplexing genome editing principle has been explored in the first *in vivo* CRISPR-based gene editing clinical trial (BRILLIANCE Phase 1/2). In this trial, patients suffering from Leber congenital amaurosis 10, a severe retinal dystrophy caused by a cryptic exon in *CEP290*, received sub-retinal injections of an advanced medicinal therapy product consisting of a pair of adeno-associated viral (AAV) vectors (EDIT-101). Together, these vectors express *Staphylococcus aureus* CRISPR-Cas9 multiplexes designed for *CEP290* reading frame restoration through NHEJ-mediated excision of the disease-causing cryptic exon [10]. However, a general consideration regarding multiplexing genome editing concerns the fact that, next to the intended chromosomal deletions, complex genomic modifications normally emerge with much higher frequencies [8, 9, 11]. The uncoordinated action of the individual CRISPR-Cas9 components contributes to these unintended genome editing endpoints. Unintended bystander products mostly comprise indels at either or both target sites and imprecise deletions in which indels locate at the junction of end-to-end chromosomal termini ligations [8, 9, 11]. Co-transfection of plasmid constructs encoding dual gRNAs and covalently linked Cas9 nucleases has been shown to heighten the accuracy of targeted DNA

deletions following tandem DSB formation [12]. Additionally, owing to orthogonal gRNA-Cas9 interactions, covalently linked Cas9 orthologs (orthogonal Cas9-Cas9 chimeras) [12], here dubbed forced Cas9 heterodimers, ensure the formation of functional CRISPR-Cas9 multiplexes which further maximize the accumulation of precise deletions over unintended genomic modifications (**Figure S1**).

In this study, we hypothesized that advanced multiplexing genome editing approaches based on forced Cas9 heterodimers should profit from synchronous and stoichiometric assembly of the attendant reagents in target cells. However, delivering the required large and multicomponent reagents in such a fashion is challenging, especially in hard-to-transfect cell types such as those with potential or established therapeutic potential. Although viral vectors have a proven track-record in achieving efficient and non-cytotoxic delivery of genome editing tools into hard-to-transfect cells, *in vitro* and *in vivo*, commonly used AAV vectors cannot deliver large genetic cargoes due to their limited DNA packaging capacity (<4.7 kb) [13, 14]. Therefore, we sought to investigate the potential of high-capacity adenoviral vectors, henceforth dubbed adenovector particles (AdVPs), for genome editing involving conventional and advanced multiplexing strategies based on split and forced CRISPR-Cas9 heterodimers, respectively. Indeed, AdVPs congregate a valuable set of features for this purpose, namely, (i) lack of viral genes; (ii) vast packaging capacity (up to 36 kb); (iii) high genetic stability; (iv) amenability to straightforward cell-tropism modifications; and (v) efficient transduction of dividing and post-mitotic cells [15-17]. AdVPs achieved efficient transfer of forced Cas9 heterodimers alone or together with their cognate gRNAs (forced CRISPR-Cas9 heterodimers) into muscle progenitor cells (myoblasts) from healthy (wild-type) and Duchenne muscular dystrophy (DMD) individuals. Importantly, as *S. pyogenes* Cas9:gRNA complexes can present high off-target activities, we assembled next-generation forced Cas9 heterodimers in which the *S. pyogenes* protein component is eCas9^{4NLS}, a variant of the high-specificity SpCas9 nuclease eSpCas9(1.1) [18] whose improved performance results from having 2 extra nuclear localization signals [19]. Moreover, a dual gRNA pair in which the *S. pyogene*s gRNA component has an optimized scaffold [20] was used to direct forced CRISPR-Cas9 heterodimers to the repair of defective alleles underlying DMD.

DMD (MIM #310200) is a lethal muscle-wasting X-linked disorder caused by loss-of-function mutations in the vast (~2.4 Mb) dystrophin-encoding *DMD* gene (prevalence: ~1 in 4700 boys [21]. Although *DMD* segmental duplications and point mutations give rise to this pathology, most DMD causing mutations consist of intragenic deletions comprising one or more exons that disrupt the mRNA reading frame [22]. Crucially, it is known that in-frame deletions within the *DMD* gene can yield internally truncated dystrophin proteins whose partial functionality causes a less severe muscular dystrophy, named Becker muscular dystrophy (BMD) (MIM #300376). Hence, targeted removal of reading frame-disrupting mutations that result in in-frame mRNA transcripts encoding shorter, yet partially functional, Becker-like dystrophins have therapeutic potential [22, 23]. The assembly of Becker-like

dystrophins has been achieved via multiplexing genome editing strategies in DMD patient-derived muscle progenitor cells (myoblasts) [11, 24-27], induced pluripotent stem cells [28] and dystrophic Dmd^{mdx} mice [29-32]. These experiments involved the use of different agents to deliver dual programmable nucleases based on zinc-fingers, transcription activator-like effectors and CRISPR systems (reviewed in ref. 33).

In this study, we build on the AdVP platform to demonstrate that transferring forced Cas9 heterodimers rather than each Cas9 component separately, increases the frequency of precise targeted DNA deletions while decreasing the extent of unintended genomic modifications. Significantly, single AdVPs assembled for all-in-one transfer of forced CRISPR-Cas9 heterodimers (i.e., forced Cas9 heterodimers and their respective gRNAs), further improves the performance of multiplexing genome editing. Finally, transduction experiments using AdVPs encoding forced CRISPR-Cas9 heterodimers combined with DNA- and protein-level assays established robust *DMD* exon 51 splice site motif excision resulting in reading frame restoration and dystrophin synthesis in unselected DMD patient-derived muscle cell populations.

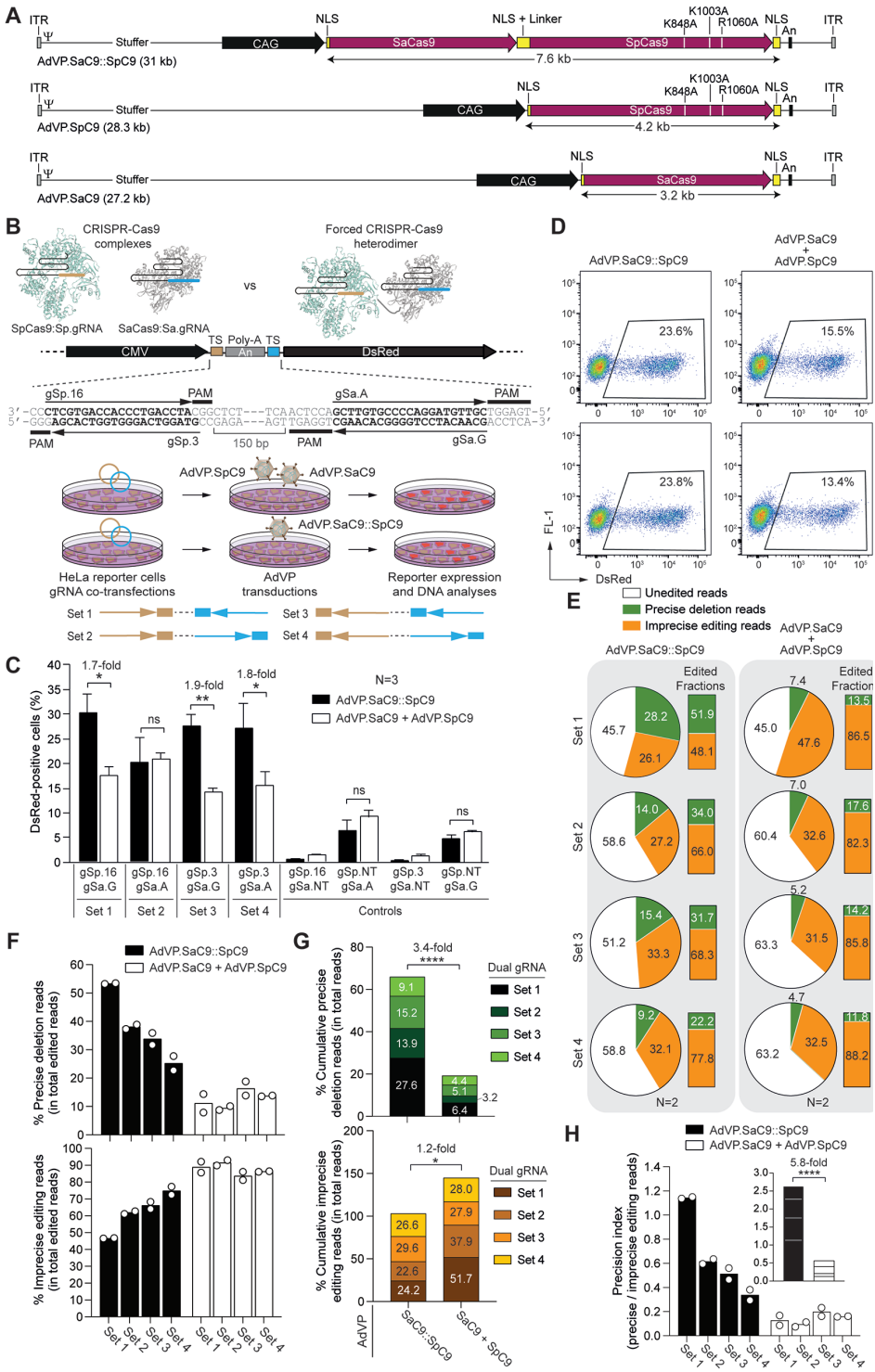
Results

AdVP delivery of forced Cas9 heterodimers facilitates targeted DNA deletions

AdVP capsids have potential for packaging and delivering full-length forced Cas9 heterodimer constructs alone or together with their cognate gRNA units. Moreover, through interactions with ubiquitously expressed CD46 receptors, AdVPs with fiber motifs from species B adenoviruses, such as those from type-50, transduce otherwise refractory coxsackievirus and adenovirus receptor-negative cells with established and potential therapeutic relevance, including stem cells and progenitor cells from the hematopoietic and skeletal muscle systems, respectively [34-36]. Finally, compared to earlier generation viral gene-containing adenoviral vectors, AdVPs have dampened cytotoxicity *in vitro* and immunogenicity *in vivo*, which bodes well for their potential clinical translation involving endogenous gene repair paradigms [16]. Hence, we started by assembling AdVPs displaying type-50 fibers and encoding a forced Cas9 heterodimer (SaC9::SpC9) consisting of the *S. aureus* SaCas9 nuclease (SaC9) [5] fused through a flexible linker to eCas9^{4NLS} (SpC9). SpC9 is a variant of the *S. pyogenes* high-specificity eSpCas9(1.1) nuclease [18] whose enhanced activity results from having 2 additional nuclear localization signals [19]. The resulting vector AdVP.SaC9::SpC9 was produced together with control AdVP.SaC9 and AdVP.SpC9 vectors encoding, respectively, SaC9 and SpC9 separately (**Figure 1A**). To test AdVP-assisted multiplexing genome editing strategies aimed at targeted chromosomal DNA deletions, we first generated HeLa.DsRed$^{TS.An.TS}$ reporter cells containing a conditional *DsRed* expression unit at the *AAVS1* safe harbor locus. In these cells, DsRed protein synthesis ensues upon the deletion of a polyadenylation signal (An) situated between the reporter and a constitutively

active CMV promoter (**Figure 1B**). Four different sets of gRNA pairs, each tailored for the targeted excision of the intervening An sequence, were tested. The dual gRNAs were designed to engage their bipartite target sequences in different orientations and to have a constant spacing between their target sites (i.e., ~150 bp) (**Figure 1B**). Multiplexing genome editing experiments in HeLa.DsRed$^{TS.An.TS}$ cells were initiated by co-transfecting expression plasmids for specific gRNA pairs. Subsequently, the transfected cells were exposed to equivalent functional units of AdVP.SaC9::SpC9 or AdVP.SaC9 and AdVP.SpC9. As negative controls, HeLa.DsRed$^{TS.An.TS}$ cells subjected to the same AdVP transduction conditions, were initially co-transfected with plasmids expressing dual gRNAs in which one member consisted of an irrelevant non-targeting gRNA. In cells exposed to 3 out of the 4 dual gRNA sets tested, higher levels of DsRed expression were detected by flow cytometry in cells transduced with AdVP.SaC9::SpC9 than in cells co-transduced with AdVP.SaC9 and AdVP.SpC9 (**Figure 1C** and **1D**). One gRNA pair (i.e., set 2) yielded similar frequencies of DsRed-positive cells upon transduction with AdVP.SaC9::SpC9 and co-transduction with AdVP.SaC9 plus AdVP.SpC9 (**Figure 1C**). These data demonstrate that the forced Cas9 heterodimer construct SaC9::SpC9 is functional and capable of inducing robust RNA-programmable targeted DNA deletions regardless of the relative orientations of the bipartite target sequences of dual gRNAs.

Next, we sought to investigate the types of DNA editing events registered in AdVP-treated reporter cells through amplicon deep sequencing analysis. This genotyping analysis revealed that, independently of the dual gRNA set used, transfer of forced Cas9 heterodimers led to higher levels of precise DNA deletions than those obtained through the separate delivery of Cas9 heterodimer moieties (**Figure 1E** and **1F**). Conversely, imprecise DNA modifications consisting in large part of indels at each target site and at chromosomal junctions were detected at lower frequencies in reporter cells exposed to SaC9::SpC9 than in cells subjected to SaC9 and SpC9 (**Figure 1E** and **1F**). Indeed, cumulative analysis of precise and imprecise DNA modifications corresponding to the dual gRNA set aggregate showed a robust increase in the former and a modest yet statistically significant reduction in the latter events in cells transduced with AdVP.SaC9::SpC9 (**Figure 1G**). As a result, the precision index represented by the ratios between accurate and imprecise genome editing reads, was most favorable by a 5.8-fold factor in HeLa.DsRed$^{TS.An.TS}$ cells transduced with AdVP.SaC9::SpC9 than in cells independently transduced with AdVP.SaC9 and AdVP.SpC9 (**Figure 1H**).


A
ITR
Stuffer
CAG
NLS
SaCas9
NLS + Linker
SpCas9
K1003A
K848A
R1060A
NLS
An
ITR
AdVP.SaC9::SpC9 (31 kb)
7.6 kb
AdVP.SpC9 (28.3 kb)
4.2 kb
AdVP.SaC9 (27.2 kb)
3.2 kb
B
CRISPR-Cas9 complexes
Forced CRISPR-Cas9 heterodimer
vs
SpCas9:Sp.gRNA
SaCas9:Sa.gRNA
CMV
TS
Poly-A
TS
DsRed
gSp.16
PAM
gSa.A
PAM
gSp.3
150 bp
PAM
gSa.G
AdVP.SpC9
AdVP.SaC9
AdVP.SaC9::SpC9
HeLa reporter cells gRNA co-transfections
AdVP transductions
Reporter expression and DNA analyses
Set 1
Set 2
Set 3
Set 4
C
DsRed-positive cells (%)
1.7-fold
1.9-fold
1.8-fold
N=3
AdVP.SaC9::SpC9
AdVP.SaC9 + AdVP.SpC9
ns
gSp.16 gSa.G
gSp.16 gSa.A
gSp.3 gSa.G
gSp.3 gSa.A
gSp.16 gSa.NT
gSp.NT gSa.A
gSp.3 gSa.NT
gSp.NT gSa.G
Set 1
Set 2
Set 3
Set 4
Controls
D
AdVP.SaC9::SpC9
AdVP.SaC9 + AdVP.SpC9
23.6%
15.5%
23.8%
13.4%
FL-1
DsRed
E
Unedited reads
Precise deletion reads
Imprecise editing reads
AdVP.SaC9::SpC9
AdVP.SaC9 + AdVP.SpC9
Edited Fractions
Set 1
Set 2
Set 3
Set 4
N=2
F
AdVP.SaC9::SpC9
AdVP.SaC9 + AdVP.SpC9
% Precise deletion reads (in total edited reads)
% Imprecise editing reads (in total edited reads)
G
3.4-fold
****
Dual gRNA
Set 1
Set 2
Set 3
Set 4
% Cumulative precise deletion reads (in total reads)
1.2-fold
*
% Cumulative imprecise editing reads (in total reads)
AdVP
SaC9::SpC9
SaC9 + SpC9
H
AdVP.SaC9::SpC9
AdVP.SaC9 + AdVP.SpC9
5.8-fold
****
Precision index (precise / imprecise editing reads)

◂ **Figure 1. AdVP delivery of forced Cas9 heterodimers incudes robust targeted DNA deletions. (A)** Schematics of AdVP genomes. AdVP.SaCas9 and AdVP.SpC9 encode, respectively, *S. aureus* SaCas9 (SaC9) and eCas9^{4NLS} (SpC9), a variant of the high-specificity *S. pyogenes* eSpCas9(1.1) nuclease. The point mutations conferring enhanced specificity to eSpCas9(1.1) are specified. AdVP.SaC9::SpC9 encodes a forced Cas9 heterodimer consisting of SaCas9 fused to eCas9^{4NLS} (SaC9::SpC9). The synthesis of all nucleases is driven from the hybrid CAG regulatory sequences. An, polyadenylation signal; NLS, nuclear localization signal (NLS); ITR and Ψ, adenoviral inverted terminal repeats and packaging signal *cis*-acting elements required for vector DNA replication and encapsidation, respectively. **(B)** Workflow of the functional readout for assessing multiplexing gene-editing. Reporter HeLa.DsRed$^{TS.An.TS}$ cells were used for tracking multiplexing gene-editing outcomes with conventional (split) and forced (linked) Cas9 heterodimers. Upper panel, HeLa.DsRed$^{TS.An.TS}$ cells encode a DsRed reporter whose expression is dependent on the elimination of a polyadenylation signal (An) located between the CMV promoter and the reporter ORF. The protospacer and protospacer adjacent motif (PAM) sequences corresponding to the target sites (TS) for SpCas9 gRNAs (i.e., gSp.16 and gSp.3) and SaCas9 gRNAs (i.e., gSa.A and gSa.G), are indicated. Lower panel, generic experimental design. HeLa.DsRed$^{TS.An.TS}$ are exposed to different combinations of gRNA pairs (dual gRNAs) via plasmid co-transfections and subsequently are co-transduced with AdVP.SaC9 and AdVP.SpC9 or transduced with AdVP.SaC9::SpC9. Multiplexing gene-editing outcomes at the cellular and molecular levels are evaluated through flow cytometry and amplicon deep sequencing analyses, respectively. **(C)** Quantification of targeted DNA deletions. HeLa.DsRed$^{TS.An.TS}$ cells were transfected with four different combinations of dual gRNAs (i.e., Set 1 through 4) and then transduced with AdVP.SaC9::SpC9 or co-transduced with AdVP.SaC9 and AdVP.SpC9. The target sites of each CRISPR-Cas9 complex are represented in the upper schematics, with boxes and arrows indicating PAM and gRNA protospacer orientations, respectively. Target DNA deletion frequencies were measured by DsRed-directed flow cytometry at 3 days post-transduction. Data are shown as mean ± SEM of three independent biological replicates. Significant differences between the indicated datasets were determined by two-tailed Student's *t* tests; **P<0.01, *P<0.05; P>0.05 was considered non-significant (ns). **(D)** Representative flow cytometry dot plots of HeLa.DsRed$^{TS.An.TS}$ cells transfected and transduced with the specified reagents. **(E-G)** Characterization of gene editing outcomes through amplicon deep sequencing. The precise deletion and imprecise editing read frequencies within the total edited read counts and the cumulative precise deletion and imprecise editing read frequencies within the total read counts in reporter cells exposed to the indicated gene editing reagents are shown in panels **F** and **G**, respectively. The next-generation sequencing analysis was performed on genomic DNA from two independent biological replicates (~50,000 paired-end reads per sample). Significant differences between the indicated datasets were determined by two-way ANOVA; *P<0.05; ****P<0.0001. **(H)** Multiplexing DNA editing precision in reporter cells. gRNA set-specific and cumulative (inset) precision index plot corresponding to the ratios between precise and imprecise editing read frequencies in HeLa.DsRed$^{TS.An.TS}$ cells transfected with gRNA sets 1 through 4 and transduced or co-transduced with the indicated AdVPs.

AdVP delivery of forced Cas9 heterodimers achieves robust endogenous *DMD* gene repair

To investigate the performance of conventional (untethered) vis-à-vis forced (tethered) Cas9 heterodimers at an endogenous human locus, we selected a dual gRNA composed of *S. pyogenes* gRNA gSp^{IN50} and *S. aureus* gRNA gSa^{EX51} ($dgRNA^{D51}$) designed for NHEJ-mediated excision of the *DMD* exon 51 splice acceptor (SA) coding motif upon site-specific cleavage at intron 50 and exon 51, respectively (**Figure 2A**, top panel). Targeted SA elimination is expected to induce exon skipping during pre-mRNA processing resulting in reading frame restoration and ensuing Becker-like dystrophin expression in muscle cells from 13-14% of DMD patients [22]. Among these are patients with deletions of *DMD* exons 48 through 50 (D48-50). To this end, we started by transducing myoblasts from a donor with the *DMD* D48-50 genotype (DMD.1 myoblasts) stably expressing $dgRNA^{D51}$ (DMD.1.$dgRNA^{D51}$ myoblasts) with AdVP.SaC9::SpC9 or with AdVP.SaC9 and AdVP.SpC9, at total multiplicities of infection (MOI) of 25, 50 and 100 genome copies per cell (GC $cell^{-1}$) (**Figure 2A**, bottom panel). Parallel cultures of DMD.1.$dgRNA^{D51}$ myoblasts individually transduced with AdVP.SaC9 or AdVP.SpC9 at the same total MOI provided for negative controls. Amplicons diagnostic for the 171-bp genomic deletion encompassing the *DMD* exon 51 SA coding motif were readily detected in cells receiving conventional and forced Cas9 heterodimers (**Figure S2**). In addition to these deletion-specific PCR products, genotyping assays based on incubating PCR amplicons with the mismatch sensing T7 endonuclease I (T7EI) also detected a prevalence of imprecise indels resulting from the activity of SaC9:gSa^{EX51} and SpC9:gSp^{IN50} complexes at their target sequences (**Figure S2**).

To characterize genome editing outcomes in the form of precise deletions versus unintended genomic modifications comprising indels at target sites or chromosomal deletion junctions, amplicon deep sequencing analysis was performed on AdVP-treated DMD.1.$dgRNA^{D51}$ myoblasts. Consistent with the transduction experiments in HeLa.$DsRed^{TS.An.TS}$ cells, this sensitive genotyping analysis showed that forced Cas9 heterodimers yielded precise deletions and imprecise genomic modifications at higher and lower rates, respectively, than those induced by untethered Cas9 heterodimers (**Figure 2B-2E**) which, in turn, resulted in an ~1.8-fold increase in the precision index for AdVP.SaC9::SpC9 (**Figure 2F**). Equally consistent with the next-generation sequencing and T7EI genotyping assays in reporter HeLa.$DsRed^{TS.An.TS}$ cells and engineered DMD.1.$dgRNA^{D51}$ myoblasts (**Figure 1E-1H** and **Figure S2**, respectively), the majority of genomic modifications consisted of imprecise DNA editing products independently of the MOI used as indicated by precision indexes below 1 (**Figure 2F**).

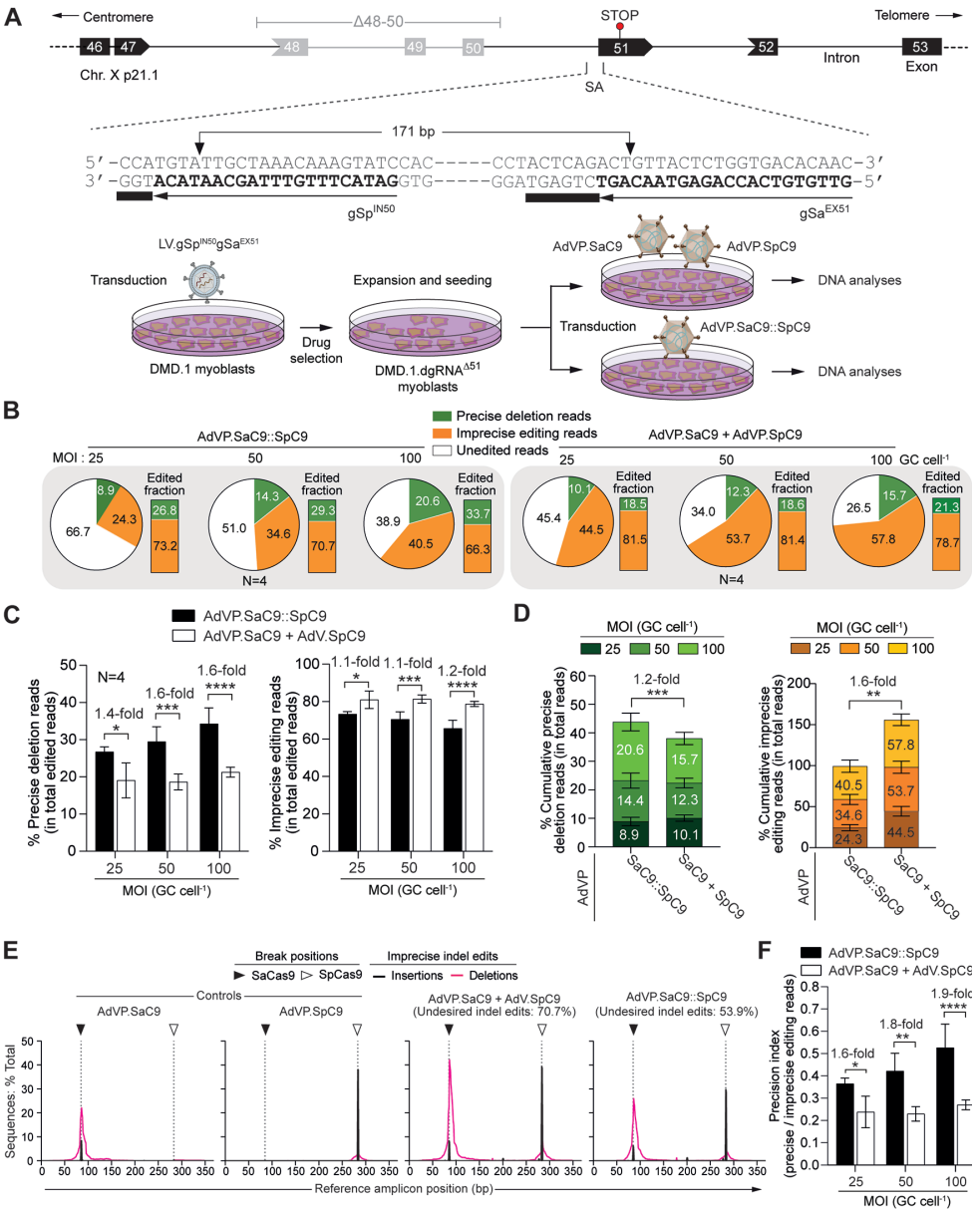


Figure 2. AdVP delivery of forced Cas9 heterodimers achieves efficient *DMD* gene repair. (A) Schematics of the *DMD* gene repair strategy. *DMD* gene repair approaches based on the excision of the exon 51 splice acceptor (SA) motif after AdVP delivery of conventional or forced Cas9 heterodimers. Upper panel, the concomitant generation of DSBs at *DMD* intron 50 and exon 51 sequences by SpC9:gSp^{IN50} and SaC9:gSa^{EX51} complexes, respectively, leads to the removal of the intervening sequence containing the exon 51 SA motif. The ligation of the resulting chromosomal termini by NHEJ yields in-frame *DMD* transcripts coding for Becker-like dystrophins in muscle cells with out-of-frame *DMD* deletions. Lower panel, generic experimental design. Multiplexing *DMD* gene editing experiments were carried out in DMD.1 myoblasts (Δ48-50) stably expressing gSp^{IN50} and gSa^{EX51} upon transduction with a lentivector encoding both gRNAs and a drug-resistance gene (i.e., DMD.1.dgRNA$^{\Delta 51}$ myoblasts). Multiplexing gene-editing outcomes in DMD.1.dgRNA$^{\Delta 51}$ myoblasts transduced with AdVP.SaC9::SpC9 or co-transduced with AdVP.SaC9 and AdVP.SpC9

◂ were evaluated through amplicon deep sequencing. **(B-D)** Quantification of *DMD* editing outcomes in DMD.1.dgRNA$^{\Delta 51}$ myoblasts. The proportions between sequencing reads derived from unedited endogenous *DMD* alleles, precise deletions, and unintended edits in DMD.1.dgRNA$^{\Delta 51}$ myoblasts exposed to AdVP.SaC9::SpC9 or to AdVP.SaC9 and AdVP.SpC9 at the indicated total MOIs, are displayed in panel **B**. The precise deletion and imprecise editing read frequencies within the total edited read counts are plotted in panel **C** (left and right graph, respectively). The cumulative precise deletion and imprecise editing read frequencies within the total read counts in DMD.1.dgRNA$^{\Delta 51}$ exposed to three MOI of the indicated AdVPs are shown in panel **D** (left and right graph, respectively). The next-generation sequencing data was derived from four biological replicates with 50,000 paired-end reads analyzed per sample. **(E)** Representative indel profiles in DMD.1.dgRNA$^{\Delta 51}$ myoblasts transduced with the indicated AdVPs at a total MOI of 100 GC cell-1. The frequencies, types and distributions of unintended indel "footprints" detected within the *DMD* target region are plotted. MOI, multiplicity of infection; GC cell^{-1}, genome copies per cell. Significant differences between the indicated datasets were determined by two-way ANOVA; ****P<0.0001, ***P<0.001, **P<0.01, *P<0.05. **(F)** Multiplexing DNA editing precision in DMD.1.dgRNA$^{\Delta 51}$ myoblasts. Precision index plot corresponding to the ratios between precise-deletion and unintended-edit read frequencies in AdVP-treated DMD.1.dgRNA$^{\Delta 51}$ myoblasts.

All-in-one AdVP delivery of forced CRISPR-Cas9 heterodimers enhances *DMD* gene repair precision

To streamline and further investigate AdVP-assisted multiplexing *DMD* gene repair, we next generated vector AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ for all-in-one delivery of forced CRISPR-Cas9 heterodimers comprising SaC9::SpC9 and dgRNA^{D51} (**Figure 3A**). AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ was produced to an high titer (i.e., 9.8×10^{10} GC ml^{-1}) and, importantly, restriction fragment length analysis of DNA isolated from purified particles revealed that packaged vector genomes retained their structural integrity (**Figure 3B**). Moreover, synthesis of forced full-length Cas9 heterodimers was established by western blot analysis of wild-type myoblasts transduced with AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ (**Figure 3C**). Interestingly, confocal immunofluorescence microscopy disclosed that, in contrast to SpC9 and SaC9, the SpC9::SaC9 fusion product was, in most cells, prevalently found in the cytoplasm despite having more NLS motifs than its constituent subunits (**Figure S3**). This sub-cellular distribution suggests that the large size of SpC9::SaC9 is a contributing factor to its relatively lower nuclear translocation capability. Regardless, T7EI genotyping assays on wild-type myoblasts transduced with AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ showed a clear dose-dependent increase in the frequency of targeted DNA deletions (**Figure S4**), and lower amounts of imprecise genome editing events than those observed in engineered DMD.1.gRNA^{D51} myoblasts transduced with AdVP.SaC9::SpC9 (compare **Figure S2** with **Figure S4**). Amplicon deep sequencing analysis of wild-type myoblasts transduced with AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ confirmed that targeted DNA deletions occurred in an AdVP dose-dependent manner (**Figure 3D** and **3E**). Significantly, regardless of the MOI used, the distribution between precise and imprecise genome editing events (**Figure 3D** and **3F**) was more balanced than that previously observed in DMD.1.gRNA^{D51} myoblasts exposed to either single AdVP.SaC9::SpC9 or to dual AdVP.SaC9 and AdVP.SpC9 vectors (**Figure 2B**). Indeed, all-in-one AdVP delivery of forced CRISPR-Cas9 heterodimer components resulted in precision indexes at or above 1 (**Figure 3G**).

Based on these previous data, we next sought to compare side-by-side the performance of *DMD* gene repair resulting from delivering multiplexing CRISPR genome editing

components as independent or integrated units. To this end, we transduced myoblasts from a second donor with the *DMD* D48-50 genotype (i.e., DMD.2 myoblasts) with AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ and, in parallel, transduced these myoblasts stably expressing dual gRNA^{D51} (i.e., DMD.2.dgRNA^{D51}) with either AdVP.SaC9::SpC9 or with AdVP.SaC9 plus AdVP.SpC9 at equivalent total MOI (**Figure 4A**). Amplicon deep sequencing analysis showed a progressive increase in precise DNA deletions and concomitant decrease in imprecise DNA modifications in cell populations transduced with AdVP.SaC9 and AdVP.SpC9, AdVP.SaC9::SpC9 and AdVP.SaC9::SaC9.dgRNA$^{\Delta E51}$ (**Figure 4B**). This trend was observed at all MOI tested and reached statistical significance in most transduction group comparisons (**Figure 4B**). Moreover, cumulative genome editing outcome analysis corresponding to the aggregate of the various MOI confirmed that transductions with AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ yielded increased precise and decreased imprecise genomic edits when compared to those triggered by co-transductions with AdVP.SaC9 and AdVP.SpC9 (i.e., 1.5- and 2.7-fold, respectively) (**Figure 4C**). As a result, the precision indexes representing the relation between precise and imprecise editing read frequencies of AdVP.SaC9::SpC9.dgRNA^{D51} transductions were higher than those of AdVP.SaC9 and AdVP.SpC9 co-transductions by a factor of ~3 to 4 fold (**Figure 4D**). Moreover, AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ also led to higher precise DNA deletions frequencies than AdVP.SaC9::SpC9 (i.e., 1.7 fold) (**Figure 4C**). In contrast, AdVP.SaC9::SaC9 transductions while capable of reducing imprecise edits by 2.4 fold, were not capable of increasing the precise DNA editing fractions when compared to AdVP.SaC9 and AdVP.SpC9 co-transductions (**Figure 4C**). Collectively, these data indicate that the performance of multiplexing genome editing procedures can profit from integrated all-in-one delivery of forced CRISPR-Cas9 heterodimer units. Importantly, amplicon deep sequencing analysis of unmodified DMD.2 myoblasts and engineered DMD.2.dgRNA^{D51} myoblasts transduced at the highest total MOI of 200 GC cell^{-1} with AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ and AdVP.SaC9 together with AdVP.SpC9, respectively, showed background indel frequencies at top-ranked candidate off-target sequences for gSp^{IN50} and Sa^{EX51} (**Figure 4E**).

Encouraged by the collective data on the activity and specificity of forced CRISPR-Cas9 heterodimers, we proceeded to assess *de novo* assembly and expression of Becker-like dystrophin molecules upon multiplexing *DMD* gene repair with AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ (**Figure 5A**). To this end, DMD.2 myoblasts were transduced with AdVP.SaC9::SaC9.dgRNA$^{\Delta 51}$ and subsequently induced to differentiate into syncytial myotubes. Fluorescence microscopy and western blot analyses readily led to the detection of Becker-like dystrophin synthesis specifically in the cultures containing myotubes differentiated from AdVP.SaC9::SaC9.gRNA$^{\Delta E51}$-transduced myoblasts (**Figure 5B** and **5C**, respectively). Finally, co-detection of the late skeletal muscle marker sarcomeric α-actinin and Becker-like dystrophin in myotubes by confocal microscopy confirmed the differentiation capacity of the DMD muscle progenitor cells edited through AdVP delivery of forced CRISPR-Cas9 heterodimers addressed to the *DMD* intron 50-exon 51 junction (**Figure 5D**).

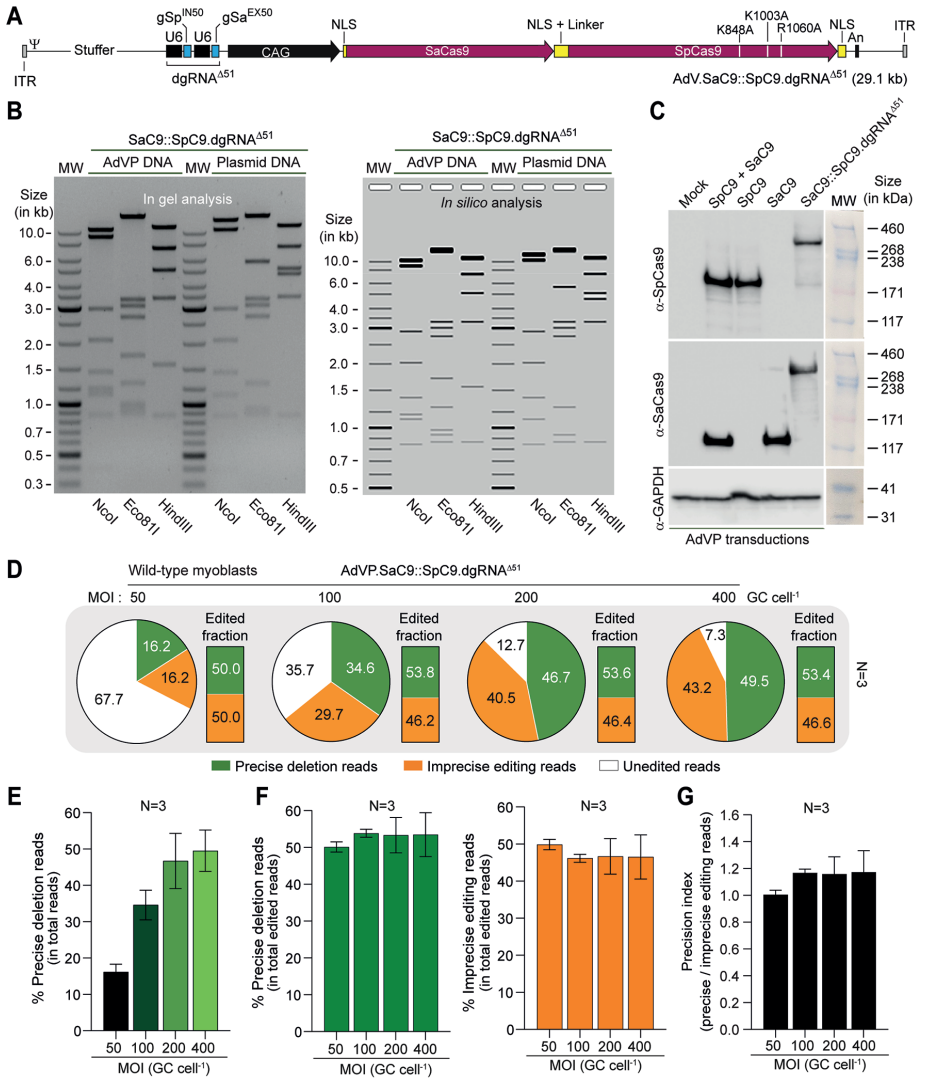


Figure 3. All-in-one AdVP delivery of forced CRISPR-Cas9 heterodimers promotes precise *DMD* gene repair. (A) Schematics of AdVP for all-in-one delivery of forced CRISPR-Cas9 heterodimers. AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ encodes the forced Cas9 heterodimer SaC9::SpC9 composed of the SaCas9 nuclease covalently linked through flexible linkers to the high-specificity SpCas9 nuclease variant eCas9^{4NLS}. The synthesis of SaC9::SpC9 and of the dual gRNA pair gSp^{IN50} and gSa^{EX51} (dgRNA$^{\Delta 51}$) is controlled via the hybrid CAG regulatory sequences and the human *U6* promoter, respectively. An, polyadenylation signal; NLS, nuclear localization signal (NLS); ITR and Ψ, adenoviral inverted terminal repeats and packaging signal *cis*-acting elements for vector DNA replication and encapsidation, respectively. **(B)** Assessing AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ DNA integrity. Restriction fragment length analysis (RFLA) of vector DNA isolated from purified AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ particles. *In silico* and in gel RFLA analyses are presented. Marker, GeneRuler DNA Ladder molecular weight mix. Parental circular plasmid served as additional molecular weight references. **(C)** Assessing full-length SaC9::SpC9 synthesis upon AdVP transduction. Western blot analysis was performed on wild-type myoblasts exposed to AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ or to AdVP.SaC9 and AdVP.SpC9 at a total MOI of 400 GC cell^{-1} at three days post-transduction. Detection of GAPDH provided for protein loading controls. MW, HiMark Pre-Stained Protein Standard molecular weight marker. **(D-G)** Deep sequencing analysis of *DMD* editing outcomes in wild-type myoblasts. The proportions

◂ between sequencing reads corresponding to unmodified *DMD* alleles, precise deletions, and unintended edits in wild-type myoblasts transduced with AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ at the indicated MOIs, are presented in panel **D**. The precise deletion read frequencies within the total read counts and within the total edited read counts are depicted in panel **E** and in the left graph of panel **F**, respectively; whilst the unintended editing read frequencies within the total edited read counts are presented in the right graph of panel **F**. **(G)** Multiplexing DNA editing precision in wild-type myoblasts. Precision index plot corresponding to the relation between deletion to imprecise editing reads in wild-type myoblasts transduced with AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ at the indicated MOIs. MOI, multiplicity of infection; GC cell^{-1}, genome copies per cell.

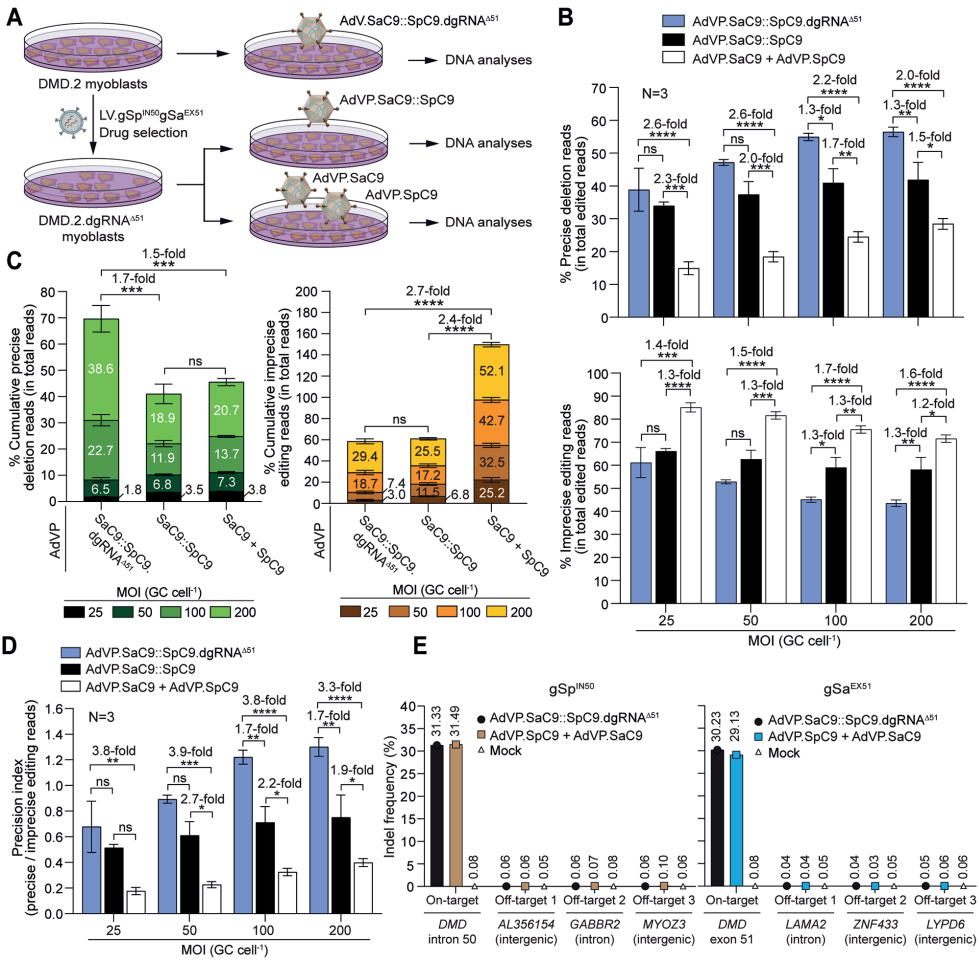


Figure 4. All-in-one AdVP delivery of forced CRISPR-Cas9 heterodimers enhances *DMD* gene repair precision. (A) Schematics of the experimental design. Multiplexing *DMD* gene editing experiments were performed in parental DMD.2 myoblasts (Δ48-50) and their derivatives stably expressing gSp^{IN50} and gSa^{EX51} upon transduction with a lentivector encoding both gRNAs and a drug-resistance gene (i.e., DMD.2.dgRNA$^{\Delta 51}$ myoblasts). Multiplexing gene-editing outcomes in parental DMD.2 myoblasts transduced with AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ were compared with those registered in engineered DMD.2.dgRNA$^{\Delta 51}$ myoblasts transduced with AdVP.SaC9::SpC9 or co-transduced with AdVP.SaC9 and AdVP.SpC9 at different total MOIs. **(B** and **C)** Amplicon deep sequencing analysis of *DMD* editing outcomes upon split versus all-in-one delivery of CRISPR-Cas9 multiplex components. The precise deletion and unintended editing read frequencies within the total edited read counts are shown in panel **B** (top and bottom graph, respectively); whilst the cumulative precise deletion and unintended editing read frequencies within the total read counts obtained through the three different AdVP transduction conditions are presented in panel **C**. Bars and error bars correspond to mean ± SEM from three biological replicates (50,000 paired-end reads per sample). **(D)** Multiplexing DNA editing precision upon split versus

◄ all-in-one delivery of CRISPR-Cas9 multiplex components. Precision index plot corresponding to the ratios between precise deletion and unintended editing read frequencies in myoblasts subjected to the AdVP transduction conditions depicted in panel A. Significant differences between the indicated datasets were determined by two-way ANOVA; ****P<0.0001, ***P<0.001, **P<0.01, *P<0.05. P>0.05 was considered non-significant (ns). MOI, multiplicity of infection; GC cell-1, genome copies per cell. **(E)** Assessing off-target DNA cleavage upon all-in-one AdVP transduction. Parental DMD.2 myoblasts and engineered DMD.2.dgRNA$^{\Delta 51}$ myoblasts were exposed to AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ and to AdVP.SpC9 and AdVP.SaC9, respectively, at a total MOI of 200 GC cell^{-1}. DNA cleaving activities at the dual gRNA target sites in *DMD* intron 50 and *DMD* exon 51 and at three top-ranked candidate off-target sites for gSp^{EX51} (i.e., *AL356154, GABBR2* and *MYOZ3*) and gSa^{IN50} (i.e., *LAMA2, ZNF433* and *LYPD6*) were quantified by amplicon deep sequencing using 50,000 paired-end reads per sample.

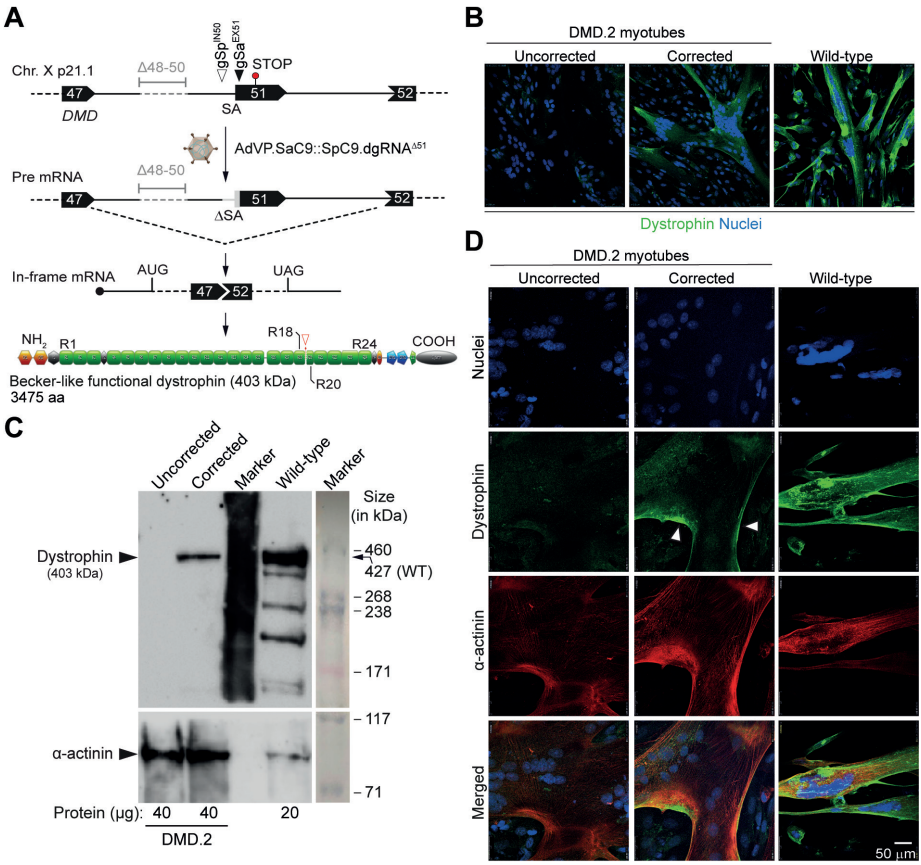


Figure 5. All-in-one AdVP delivery of forced CRISPR-Cas9 heterodimers rescues dystrophin synthesis in DMD muscle cells. (A) Schematics of the *DMD* gene repair strategy and outcome. AdVP delivery of forced CRISPR-Cas9 heterodimers addressed to *DMD* intron 50 and exon 51 results in NHEJ-mediated excision of the exon 51 splice acceptor (SA) site leading to in-frame *DMD* transcript assembly and Becker-like dystrophin expression in muscle cells with out-of-frame *DMD* deletions **(B)** Detection of dystrophin by fluorescence microscopy. Confocal immunofluorescence microscopy analysis on muscle cells differentiated from DMD.2 patient-derived myoblasts transduced with AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ at an MOI of 200 GC cell^{-1} (corrected). Mock-transduced DMD.2 myoblasts (uncorrected) and healthy donor-derived myoblasts (wild-type) subjected to the same myogenic differentiated conditions served as negative and positive controls, respectively. Immunostaining was done at 10 days post-transduction and nuclei were labelled with the DNA dye DAPI. **(C)** Detection of dystrophin by western blot analysis. Western blotting was performed on muscle cells differentiated from DMD.2 patient-derived myoblasts transduced with AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ at an MOI of 200 GC cell^{-1} (corrected). Mock-transduced DMD.2 myoblasts (uncorrected) and healthy donor-derived myoblasts (wild-type) treated with the same myogenic differentiated conditions served as negative and positive controls, respectively. Differentiated wild-type muscle cells

◄ all-in-one delivery of CRISPR-Cas9 multiplex components. Precision index plot corresponding to the ratios between precise deletion and unintended editing read frequencies in myoblasts subjected to the AdVP transduction conditions depicted in panel A. Significant differences between the indicated datasets were determined by two-way ANOVA; ****P<0.0001, ***P<0.001, **P<0.01, *P<0.05. P>0.05 was considered non-significant (ns). MOI, multiplicity of infection; GC cell-1, genome copies per cell. **(E)** Assessing off-target DNA cleavage upon all-in-one AdVP transduction. Parental DMD.2 myoblasts and engineered DMD.2.dgRNA$^{\Delta 51}$ myoblasts were exposed to AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ and to AdVP.SpC9 and AdVP.SaC9, respectively, at a total MOI of 200 GC cell^{-1}. DNA cleaving activities at the dual gRNA target sites in *DMD* intron 50 and *DMD* exon 51 and at three top-ranked candidate off-target sites for gSp^{EX51} (i.e., *AL356154, GABBR2* and *MYOZ3*) and gSa^{IN50} (i.e., *LAMA2, ZNF433* and *LYPD6*) were quantified by amplicon deep sequencing using 50,000 paired-end reads per sample. provided controls for the expression of endogenous full-length dystrophin and the skeletal muscle differentiation marker sarcomeric α-actinin. Marker, HiMark Pre-Stained Protein Standard molecular weight marker. **(D)** Assessing the differentiation capacity of AdVP-edited muscle cells. Confocal microscopy co-detection of dystrophin and sarcomeric α-actinin was performed on muscle cells differentiated from DMD.2 myoblasts initially transduced with AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ at an MOI of 200 GC cell^{-1} (corrected). Mock-transduced DMD.2 myoblasts (uncorrected) and healthy donor-derived myoblasts (wild-type) exposed to the same myogenic differentiated conditions served as negative and positive controls, respectively. The Co-immunostaining was done at 10 days post-transduction and nuclei were labelled with the DNA dye DAPI.

Discussion

Rapid progression in the genome editing field is contributing to widen the options for tackling complex scientific questions and developing candidate gene and cell therapies. The integration of gene delivery and gene editing systems is expected to advance this field by enabling efficient and predictable genetic modification of relevant target cell types *in vitro* and *in vivo*. Yet, an emerging trend concerns the fact that precision genome editing is increasingly underpinned by large and multicomponent reagents that render the application of commonly used delivery agents, such as AAV, cumbersome or ineffective [33]. As a case in point, forced Cas9 heterodimer proteins and cognate orthogonal dual gRNAs constitute robust and versatile genome editing tools owing to the cooperative action of the resulting CRISPR-Cas9 multiplexes. Indeed, plasmid co-transfection experiments in HEK293T cells demonstrated that forced Cas9 heterodimers besides promoting the generation of precise DNA deletions, permit cleaving otherwise refractory target sites with non-canonical PAMs, thus enlarging the genome editing targeting range [12]. Presumably these attributes result from the increased effective concentration of CRISPR-Cas9 multiplexes at target sequences in a synchronous fashion (**Figure S1**). However, the specificity conferred by forced Cas9 heterodimers based on wild-type nucleases is not superior to that of untethered Cas9 proteins [12]. In addition, productive and reproducible delivery of Cas9 fusion constructs in hard-to-transfect cells is severely limited due to their large size.

In this study, we have hence introduced a next-generation forced Cas9 heterodimer architecture (SaC9::SpC9) consisting of SaCas9 fused through a flexible linker to the enhanced high-specificity eCas9^{4NLS} [19] and, using this tool, demonstrate the capacity of AdVPs to package functional Cas9 fusion constructs together with dual gRNA units. In doing so, we overcome the bottlenecks associated with the production and delivery of forced CRISPR-Cas9 heterodimer components for multiplexing gene editing purposes in hard-to-transfect dividing and non-dividing cells. AdVP transduction experiments comparing the performance of separated versus integrated delivery of multiplexing genome editing components supports

the importance of introducing such components in an integrated and stoichiometric fashion (**Figure S5**).

Hence, it is possible that the efficiency as well as the accuracy of other advanced genome editing principles based on large and multicomponent reagents will equally profit from integrated all-in-one delivery approaches. Prime editing (PE) [37], for instance, depends on the coordinated action between a Cas9 nickase fused to an engineered reverse transcriptase (prime editor) and, often, dual gRNAs in which one or both members are extended as PE gRNAs (pegRNAs). In fact, recently, dual PE complexes consisting of a prime editor protein and two pegRNAs specifying bipartite target sequences and edits of interest, have been successfully tailored for inserting or deleting specific DNA tracts upon plasmid co-transfections in diverse cell lines [38-43]. Previously, by exploiting the cell-cycle independent AdVP platform to transduce dividing and post-mitotic cells alike, work in our laboratory had demonstrated that the activity of single PE complexes is significantly promoted in cycling cells [44]. Hence, it should be informative to investigate dual PE- and NHEJ-based multiplexing genome editing endpoints in *ex vivo* and *in vivo* settings using AdVP delivery as these settings are normally associated with cycling and post-mitotic cells, respectively.

Candidate *in vivo* and *ex vivo* DMD genetic therapies under investigation present their own sets of pros and cons [33]. Current experimental cell therapies for muscular dystrophies based on the transplantation of *ex vivo* corrected myogenic cells present important bottlenecks, including limited cell survival, migration, and tissue engraftment [45, 46]. On the other hand, *ex vivo* approaches offer a controlled gene-correction environment, bypass vector-neutralizing antibodies, and minimize direct contact between the patient and immunogenic components [33]. In this study, AdVP delivery of forced CRISPR-Cas9 heterodimers was directed to the targeted excision of the *DMD* exon 51 splice acceptor motif in DMD muscle progenitor cells to evaluate *DMD* reading frame restoration and ensuing detection of dystrophin molecules. This NHEJ-based *DMD* gene editing strategy, amenable to 13-14% of the DMD patient population [22], readily led to the synthesis of Becker-like dystrophins in syncytial muscle cell populations differentiated from unselected DMD myoblasts.

In conclusion, in this study, we have introduced a next-generation forced Cas9 heterodimer construct for RNA-programmable installation of chromosomal deletions in an efficient, specific and accurate manner. In addition, we demonstrate the feasibility in using AdVPs to package forced CRISPR-Cas9 constructs for stoichiometric delivery of synchronously acting multiplexing gene editing components into human cells. Finally, by capitalizing on AdVPs to similarly transfer genome editing tools independently of their size, we found that integrated rather than separated delivery of forced CRISPR-Cas9 heterodimer units can foster the precision of targeted DNA deletions at the expense of unintended genomic modifications. Collectively, our data provides insights that can guide the future development and application of genome editing principles that depend on the balanced delivery and coordinated action of individual parts.

Methods

Cells

HeLa cells (American Type Culture Collection), EGFP expressing HeLa cell-derived H27 cells [47], and HeLa.DsRed$^{TS.An.TS}$ cells, conditionally expressing a DsRed reporter, were cultured in Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific; Cat. No.: 41966-029) supplemented with 5% (v/v) fetal bovine serum ultra-low endotoxin (FBS; Biowest; Cat. No.: S1860-500). The culture conditions for the human wild-type and *DMD*-defective myoblasts #6594 and #AB1098, herein referred to as DMD.1 and DMD.2 myoblasts, respectively, have been previously described [48, 59]. In brief, these muscle progenitor cells were grown in Skeletal Muscle Cell Growth Medium (Ready-to-use; PromoCell; Cat. No.: C-23060) supplemented with 20% FBS, 1× Glutamax (Thermo Fisher Scientific; Cat. No.: 35050) and 100 U ml^{-1} penicillin/streptomycin or in F10 medium (ThermoFisher Scientific; Cat. No.:41550021) supplemented with 10% FBS (Gibco™; Cat. No.: 10500064), 10 ng µl^{-1} recombinant human basic fibroblast growth factor (FGF; Peprotech; Cat. No.: 100-18B), 1 µM dexamethasone (Sigma-Aldrich; Cat. No.: D2915) and 100 U ml^{-1} penicillin/streptomycin. The PEC3.30 AdVP packaging cells [27] were maintained in high-glucose DMEM supplemented with 10% FBS, 10 mM $MgCl_2$ and 0.4 µg ml^{-1} puromycin (Thermo Fisher Scientific; Cat. No.: A11138-03). HEK293T cells were maintained in DMEM supplemented with 5% FBS and 100 U ml^{-1} penicillin/streptomycin. The cells used in this study were mycoplasma-free and were kept at 37°C in humidified-air atmospheres with 10% CO_2 (i.e., HeLa, HeLa.DsRed$^{TS.An.TS}$, and HEK293T cells), 5% CO_2 (i.e., human myoblasts), or at 39°C in a humidified-air atmosphere with 10% CO_2 (i.e., PEC3.30 cells).

Recombinant DNA

The construct AV44_pCAG.Cas9^{D10A}.gRNA^{S1} encodes the *S. pyogenes* nicking enzyme SpCas9^{D10A} together with *AAVS1*-targeting gRNA^{S1} [50]. This construct together with AQ02_pDonor.AAVS1.CMV.TS.An.TS.DsRed was employed to knock-in at the *AAVS1* safe harbour locus the conditional *DsRed* cassette via in trans paired nicking [50], resulting in the fluorescence-based reporter cell line HeLa.DsRed$^{TS.An.TS}$. Plasmid AW42_pLV.gSp^{IN50}gSa^{EX51} was used for the assembly of lentivector particles LV.gSp^{IN50}gSa^{EX51} employed for the generation of myoblasts constitutively expressing a dual gRNA consisting of gSp^{IN50} and gRNA gSa^{EX51} (dgRNA^{D51}) targeting *DMD* intron 50 and *DMD* exon 51, respectively. The former and latter gRNAs are compatible with the orthogonal Cas9 proteins SpCas9 and SaCas9, respectively. The gRNA expressing plasmids AZ43_gSp.16, BA21_gSp.3, AM51_gSp.ISceI, BB10_gSa.G and AV73_gSa.A were assembled by inserting the annealed oligonucleotides listed in **Table S1** into the BveI-digested *S. pyogenes* gRNA acceptor construct AY56_Sp.gRNA-acceptor [19], or into the Esp3I-digested *S. aureus* gRNA acceptor construct BPK2660 (Addgene plasmid #70709) [51], herein named, AZ46_Sa.gRNA-acceptor. The *S. pyogenes* gRNA expression

units used in this study have mutations in the scaffold coding sequence that disrupt a cryptic RNA polymerase III terminator and extend a stabilizing gRNA duplex which, together, can contribute to improved DNA editing activities [20], AM51_gSp.ISceI encodes a non-targeting gRNA, i.e., $gRNA^{NT}$. This gRNA is irrelevant in human cells as it addresses *S. pyogenes* Cas9 proteins to the recognition sequence of the *S. cerevisiae* I-SceI homing endonuclease. AM51_gSp.ISceI and AZ46_Sa.gRNA-acceptor served as negative controls in transfection experiments. The annotated maps and nucleotide sequences of all the constructs generated for this study are available in **Figure S6-S11.**

Cell transfections

HeLa cells were seeded at a density of 5×10^4 cells per well of 24-well plates (Greiner Bio-One). Next, transfections were initiated by adding 1 mg ml^{-1} 25 kDa linear polyethyleneimine (PEI, Polysciences) solution (pH 7.4) to the different plasmid mixtures diluted in 150 mM NaCl (Merck) to a final volume of 50 µl. The amounts of PEI and DNA (in ml and ng, respectively) as well as the compositions of each of the DNA mixtures corresponding to the different transfection reactions are specified in **Tables S2, S3** and **S4**. After the addition of the PEI polycation, the transfection reactions were immediately and vigorously vortexed for 10 sec, after which, DNA-PEI complexes were allowed to form for 15 min at room temperature (RT). The resulting DNA-PEI complexes were subsequently added directly into the culture media of the target cells and, after 6 hours, the transfection media were substituted by regular culture media.

Generation of a fluorescence-based reporter cell line

The fluorescence-based reporter system $HeLa.DsRed^{TS.An.TS}$ consists of HeLa cells genetically modified with a conditional DsRed reporter cassette and a puromycin-resistance gene. Expression of the DsRed fluorescent protein is dependent on the elimination of a polyadenylation signal (polyA) located between a CMV promoter and the DsRed ORF. NHEJ-mediated polyA sequence deletion is made possible in the presence of CRISPR-Cas9 nucleases designed to cleave at target sites flanking the transcription termination sequence. The single cell-derived clone $HeLa.DsRed^{TS.An.TS}$ was obtained by targeted integration of the reporter cassette into the *AAVS1* safe harbour locus [52, 53]. The integration of the exogenous cassette was achieved through a seamless gene knock-in strategy named in trans paired nicking [50-54] based on simultaneous single-strand break (SSB) formation at donor and acceptor DNA by CRISPR-Cas9 nickases, in this case, $Cas9^{D10A}$:$gRNA^{S1}$ complexes encoded by AV44_pCAG.$Cas9^{D10A}$.$gRNA^{S1}$ [50]. The annotated map and nucleotide sequence of the donor construct AQ02_Donor.S1.CMV.TS.An.TS.DsRed employed to generate the $HeLa.DsRed^{TS.An.TS}$ reporter cells are available in the **Supplementary Information**. The generation of this fluorescence-based reporter cell line was initiated by transfecting HeLa cells following the PEI-based protocol above-described and the transfection scheme depicted in **Table S2**. After three days, the cells were transferred to a new plate containing regular growth medium

and one day later, the growth medium was supplemented with 1 μg ml^{-1} of puromycin. Parental mock-transfected cells served as negative controls during the drug selection procedure. A puromycin-resistant single cell-derived clone was expanded and employed in all the genome editing experiments involving sequential plasmid transfections and AdVP transductions.

Production and characterization of AdVPs

The AdVP molecular clones AO76_pHC-Ad.CAG.SaCas9, AW71_pHC-Ad.CAG.eCas9^{4NLS}, AW78_pHC-Ad.CAG.SaCas9.link.eCas9^{4NLS} and X65_pHC-Ad.SaCas9.link.eCas9^{4NLS}.dgRNA^{D51} were assembled through standard recombinant DNA techniques and then used for the production of the fiber-modified AdVPs AdVP.SaC9, AdVP.SpC9, AdVP.SaC9::SpC9 and AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$, respectively. The annotated maps and relevant nucleotide sequences of the AdVP genomes are available in the **Supplementary Information**. The protocols used in the generation, purification and characterization of the resulting fiber-modified AdVP stocks have been described in detail before [27, 44]. In brief, to initiate AdVP production, PEC3.30 producer cells expressing bacteriophage P1 Cre recombinase and adenovirus type 5 E1- and E2A-encoded proteins, were seeded at a density of 1.6×10^6 cells per well of 6-well plates (Greiner Bio-One). The following day, 6.25 μg of MssI-linearized AdVP plasmid clones were diluted in a total volume of 200 μl of 150 mM NaCl to which 20.6 μl of a 1 mg ml^{-1} solution of 25-kDa linear polyethyleneimine (PEI; Polysciences) were added. The transfection solutions were then immediately and thoroughly mixed in a vortex and subsequently incubated for 25 min at RT to let DNA-PEI complexes form before being added in a dropwise manner to the medium of the producer cells. Six hours post-transfection the medium was replaced with fresh medium containing E1-deleted helper AdV vector AdV.SRα.LacZ.1.50[55] at a multiplicity of infection (MOI) of 40 infections units (IU) per cell. The helper vector enables the expression of the proteins necessary for the replication and assembly of the AdV particles. Additionally, by transferring the cells to the permissive temperature of 34°C, expression of a thermosensitive version of the adenovirus DNA-binding protein (DBP) ensues in the PEC3.30 cells further contributing to vector complementation. Producer cells were harvested upon helper-triggered full cytopathic effect (CPE) and then were subjected to three cycles of freezing and thawing in liquid N_2 and 37°C water baths, respectively. After centrifugation for 10 min at 2,000 ×g, the supernatants containing the vector particles were recovered and employed in three subsequent amplifications rounds in producer cells co-transduced with helper AdV.SRα.LacZ.1.50. The AdVPs retrieved from the last propagation round, involving 20 T175-cm^2 culture flasks, were purified by sequential block and continuous CsCl buoyant density ultracentrifugation steps. Finally, the purified AdVPs were de-salted by ultrafiltration through Amicon Ultra-15 100K MWCO filters (MerckMillipore; Cat. No: UFC910024).

The transducing unit titers of purified AdVP stocks were determined through qPCR assays using the iQ™ SYBR® Green Supermix (Bio-Rad; Cat. No.: L010171C) and the primers

targeting the AdVPs packaging signal (ψ) listed in **Table S5**. Three 3-fold serial dilutions of the vector genomes extracted from the purified AdVP stocks using the DNeasy Blood & Tissue kit (QIAGEN; Cat. No.: 69506) were diluted 1:100 and employed as qPCR templates. Eight 10-fold serial dilutions of a linearized parental plasmid stock containing 1×10^7 GC μl^{-1} were used as qPCR templates to generate a standard curve. The qPCR primers, cycling conditions and reaction components are specified in **Tables S5** and **S6**. Data analysis was performed by using the Bio-Rad CFX Manager 3.1 software and the titers were calculated based on the Ct values of standard curves and extracted AdVP genome dilutions. The AdVP titers obtained are listed in **Table S7**. Additionally, the functional titers of AdVP.SaC9, AdVP.SpC9 and AdVP.SaC9::SpC9 were assessed by using an assay based on flow cytometric quantification of *EGFP* knockout frequencies in H27 indicator cells following gRNA and Cas9 nuclease delivery. To express the appropriate *EGFP*-targeting *S. pyogenes* and *S. aureus* gRNAs, H27 cells were first transfected according to the protocol described above and the transfection scheme depicted in **Table S3**. Next, these cells were transduced with a range of AdVP stock dilutions. Three days post-transduction the percentages of reporter-negative cells were determined through flow cytometry and used to calculate AdVP stock titers in gene knockout units per ml (**Table S8**).

The assessment of the structural integrity of packaged vector genomes in purified AdVP stocks of AdVP.SaC9::SpC9.dgRNA$^{\Delta 51}$ was essentially carried out as described previously [56]. In brief, 80 µl of purified AdVPs were treated with 8 µl of 10 mg ml^{-1} DNase I (Sigma-Aldrich; Cat. No.: 10104159001) at 37 °C for 30 min. Next, 2.4 µl of 0.5 M EDTA (pH 8.0), 6 µl of 10% (w/v) sodium dodecyl sulphate (SDS) and 1.5 µl of 20 mg ml^{-1} proteinase K (Thermo Fisher Scientific; Cat. No.: EO0491) were added to inactivate the DNase I activity. Following an incubation at 55 °C for 1 h, vector DNA isolation was completed by using the QIAEX II Gel Extraction Kit (QIAGEN; Cat. No.: 20021) according to the manufacturer's instructions. The isolated vector genomes were then subjected to restriction enzyme fragment analysis (RFLA) by using the Gel-Doc XR+ system (Bio-Rad) and the Image Lab 6.0.1 software (Bio-Rad). Parental AdVP plasmid clones digested with the same restriction enzymes applied to vector genomes served as molecular weight references. The *in-silico* restriction patterns corresponding to intact plasmid and vector DNA were made with the aid of the SnapGene (version 5.3.1) software.

Production of lentivector particles

The lentivector LV.gSp^{IN50}gSa^{EX51} was assembled according to previously detailed protocols [57,58]. In brief, HEK293T cells were seeded in 175-cm^2 culture flasks (Greiner Bio-One) and, the next day, wete transfected with a 30-µg DNA mixture composed of lentivector shuttle plasmid AW42_pLV.gSp^{IN50}gSa^{EX51}, packaging plasmid psPAX2 (Addgene plasmid #12260; a gift from Didier Trono), and VSV-G-pseudotyping plasmid pLP/VSVG (Invitrogen) at 2:1:1 (size-normalized for molecule copy number) diluted in 150 mM NaCl to a final 1-ml volume. Next, after receiving 90 µl of a 1 mg ml^{-1} PEI solution (25 kDa PEI; Polysciences),

the transfection mixture was immediately and vigorously vortex for approximately 10 sec. After 15-20 min at RT the DNA-PEI complexes were diluted in 19 ml of culture medium and directly added to the HEK293T producer cells. After 24 hours, the transfection medium was replaced by fresh DMEM supplemented with 5% FBS and, at 3 days post-transfection, the producer-cell conditioned medium was harvested and the cellular debris were removed by centrifugation and filtration through 0.45-µm pore-sized HT Tuffryn membrane filters (Pall Life Sciences; Cat. No. PN4184). The lentiviral vector particle titer in the clarified supernatant was assessed by employing the RETROTEK HIV-1 p24 antigen ELISA kit (ZeptoMetrix; Cat. No.: 0801111). On the basis of the resulting physical particle concentration of 416 ng $p24^{gag}$ ml^{-1} the functional lentivector dose applied for generating dual gRNA expressing DMD myoblasts was estimated by converting 1 ng of p24gag antigen to 2500 lentiviral vector transducing units [59].

Generation of myoblasts expressing dual gRNAs

The generation of DMD.1 and DMD.2 myoblasts constitutively expressing gSa^{EX51} and gSp^{IN50} was done via transduction with $LV.gSp^{IN50}gSa^{EX51}$. In brief, cells were seeded in regular growth medium at a density of 5×10^4 cells per well of 24-well plates. The following day the cells were exposed to medium containing the lentiviral vector at an MOI of 5 TU $cell^{-1}$. After 2-3 days, the cells were transferred to a new plate containing regular growth medium and, one day later, the medium of DMD.1 and DMD.2 myobalsts cell culture was supplemented with 20 and 50 µg ml^{-1} of hygromycin B (Invitrogen; Cat. Nr.: 10687010), respectively. Parental mock-transduced cells served as negative controls during the drug selection procedure.

Transduction experiments

Transduction experiments in $HeLa.DsRed^{TS.An.TS}$ cells were initiated by seeding the cells in wells of 24-well plates at a density of 5×10^4 cells per well. The next day, the cells were exposed to the appropriate gRNA constructs by using the PEI-based transfection protocol described above, and the transfection mixtures indicated in **Table S3.** After 6 hours the transfection medium was replaced by 500 µl of regular culture medium containing equivalent functional units of AdVPs. At 3 days post-transduction the cells were analyzed through DsRed-directed flow cytometry and were collected for the quantification and characterization of targeted genome-modifying events.

Transduction experiments in human myoblasts were initiated by seeding the myoblasts in wells of 24-well plates at a density of 5×10^4 cells per well. The next day, the medium was replaced by 500 µl of medium containing AdVPs at different MOIs. At 3 days post-transduction, the myoblasts were transferred into wells of 6-well plates and, after reaching confluency, the myoblasts were collected for genomic DNA extraction to quantify and characterize genome-modifying events via next-generation deep sequencing analysis.

Cell differentiation assays

Human myoblasts were transferred in regular culture medium into wells of 24- or 6-well

plates pre-coated with a 0.1% (w/v) gelatin solution (Sigma-Aldrich; Cat. Nr.: G13393). After reaching full confluency, the muscle progenitor cells were exposed to myogenic differentiation medium consisting of phenol red-free DMEM (ThermoFisher Scientific; Cat. No.: 11880-028) supplemented with 100 μg ml^{-1} human holo-transferrin (Sigma-Aldrich; Cat. No.: T0665), 10 μg ml^{-1} human insulin (Sigma-Aldrich; Cat. No.: I9278) and 100 U ml^{-1} penicillin/streptomycin. The differentiation of post-mitotic myotubes was assessed 4-5 days later by confocal immunofluorescence microscopy and western blot analyses using the antibodies specified in **Table S9** and **S10**, respectively.

Flow cytometry

The frequencies of edited HeLa.DsRed$^{TS.An.TS}$ cells were determined by reporter-directed flow cytometry using a BD LSR II flow cytometer (BD Biosciences). In brief, after large-volume PBS washes and trypsin treatments, the reporter cells were collected by centrifugation at 300 ×g for 5 min and cell pellets were resuspended in PBS containing 0.5% bovine serum albumin (BSA) and 2 mM EDTA (pH 8.0) (FACS buffer). Mock-transduced cells served as control to establish background fluorescence levels. A minimum of 10,000 single live-cells were acquired per sample and the resulting data were analyzed with the aid of the FlowJo 10.6.0 software (TreeStar).

Target DNA cleaving assays

Targeted DSB formation in transduced cells was assessed by using genotyping assays based on the mismatch sensing T7EI enzyme. To this end, genomic DNA from mock-transduced and AdVP-transduced cells was isolated by using the DNeasy Blood & Tissue kit (QIAGEN; Cat. No.: 69506) following the manufacturer's recommendations. The *DMD*-specific PCR amplifications were performed with the Phusion High-Fidelity DNA Polymerase system (ThermoFisher; Cat. No.: F-530). The primers, PCR mixtures and cycling parameters used are specified in **Tables S11** and **S12**. The resulting PCR amplicons were denaturated and reannealed by applying the thermocycling program indicated in **Table S13**. T7EI-based DNA cleaving assays were done as follows. First, 8 μl of each PCR mixture was incubated in 15-μl reactions consisting of 1× NEBuffer 2 (New England Biolabs; Cat.No.: B7002S) and 5 U of T7EI (New England Biolabs; Cat. No.: M0302). Next, after 15-min incubations at 37°C, the DNA samples were subjected to electrophoresis through 2% (w/v) agarose gels in 1× Tris–acetate–EDTA (TAE) buffer. The resulting ethidium bromide-stained DNA species were then detected by using a Molecular Imager Gel-DocTM XR+ system (Bio-Rad) and analyzed via the Image Lab 6.0.1 software (Bio-Rad).

Confocal immunofluorescence microscopy

Undifferentiated myoblasts and differentiated myotubes were analyzed through immunofluorescence microscopy analysis. Cells cultured on glass coverslips were fixed with 4% paraformaldehyde (PFA) for 10 min. Next, after three washes with PBS, the cells were permeabilized in 0.5% (v/v) Triton X-100 in TBS (50 mM Tris-HCl pH 7.5, 100mM NaCl) at

RT for 5 min, after which, they were washed thrice for 10 min with 0.1% Triton X-100 in TBS. Subsequently, the cells were incubated overnight at 4°C with a blocking Antibody Dilution Solution (Abdil) consisting of 0.1% Triton X-100, 2% BSA and 0.1% sodium azide in TBS. The specimens were then exposed for 2 hours at RT to the appropriate primary antibodies diluted in the blocking solution (**Table S9**). After three 5-min washes with 0.1% Triton X-100 in TBS, the specimens were incubated with fluorochrome-conjugated secondary antibodies diluted in blocking solution for 1 h in the dark at RT (**Table S9**). Finally, after three 10-min washes with 0.1% Triton X-100 in TBS, ProLong™ Diamond Antifade Mountant reagent containing DAPI (ThermoFisher Scientific; Cat. No.: P36971) was used for mounting the specimens. Immunofluorescence microscopy images were acquired by using an upright Leica SP8 confocal microscope equipped with Leica hybrid detectors HyD. All images were analyzed through the LAS X (Leica Microsystems) and ImageJ (NIH, US National Institutes of Health) software packages.

Western blotting

Cultures of differentiated myotubes were processed for western blot analysis as follows. After 4 to 5 days in differentiation medium, the myotube-containing cultures were lysed on ice for 30 min by incubation in 50 µl of RIPA buffer (ThermoFisher Scientific; Cat. No.: 89900) supplemented with a protease inhibitor cocktail (cOmplete Mini, Sigma-Aldrich; Cat. No.: 11836153001). The resulting cell lysates were then passed through a 30-gauge syringe several times. Protein quantification was carried out by using the Pierce BCA Protein Assay Kit (ThermoFisher Scientific, Cat. No: 23225), following the manufacturer's instructions. Next, the indicated amounts of total protein diluted in 4× sample buffer (Bio-Rad; Cat. No.: 161-0791) and 20× reducing agent (Bio-Rad; Cat. No.: 161-0792), were incubated at 95°C for 5 min. Protein samples and 15 µl of HiMark Prestained Protein Standard (ThermoFisher Scientifi; Cat. No.: LC5699) were loaded in a 3–8% Criterion XT Tris-Acetate precast gel (Bio-Rad; Cat. No.: 3450130). The polyacrylamide gel was then placed in a Criterion Cell containing XT Tricine running buffer (Bio-Rad; Cat. No.: 1610790) and run for 30 min at 75 V (0.07 A) and for 1.5 h at 150 V (0.12 A). Subsequently, the resolved proteins were transferred to polyvinylidene difluoride (PVDF) membranes with the aid of a Trans-Blot Turbo Midi PVDF pack (Bio-Rad; Cat. No.: 1704157) and a Trans-Blot Turbo system (Bio-Rad) according to the manufacturer's recommendations for high-molecular-weight proteins (2.5 A, 25 V, 10 min). The PVDF membranes were then blocked for 2 h at RT in 5% non-fat dry milk (Campina Elk; Cat. No.: 112349) dissolved in TBS with 0.1% (v/v) Tween 20 (TBST). Next, the membranes were incubated overnight at 4°C with primary antibodies (**Table S10**) diluted in 5% non-fat dry milk. After three 10-min washes in TBST, the membranes were incubated for 2 hours at RT with the appropriate secondary antibodies (**Table S10**) conjugated to horseradish peroxidase (IgG-HRP) diluted in 5% non-fat dry milk. Proteins were detected by using horseradish peroxidase substrate Clarity Western ECL (ThermoFisher Scientific; Cat. No.: 1705061) following the manufacturer's specifications. The protein lysate samples

employed for Cas9 detection were instead retrieved at 3 days post-transduction. The samples were lysed in Laemmli buffer consisting of 8.0% glycerol, 3% SDS and 200 mM Tris-HCl (pH 6.8) and were subsequently incubated for 5 minutes at 100°C. Next, the samples underwent the same procedures as described above.

Next-generation sequencing for on-target and off-target analyses

HeLa cells, human myoblasts and human iPSCs that underwent various AdVP-based genome editing approaches were analyzed by amplicon deep sequencing to quantify and characterize the resulting genome-editing events at on- and off-target sites. Genomic DNA isolated with the aid of the DNeasy Blood & Tissue kit reagents and protocol was used as input to a previously described amplicon deep sequencing analyses pipeline [44, 60]. In brief, the *DMD* exon 51 target region was first amplified with primers containing adapter tag overhangs using Phusion High-Fidelity Polymerase (Thermo Fisher Scientific; Cat. No.: #F-530L). The primers, PCR mixtures and cycling parameters used are specified in **Tables S14** and **S15.** The resulting amplicons were subsequently purified with AMPure XP beads (Beckman Coulter; Cat. No.: A63881) and subjected to PCR barcoding using Illumina tag-specific primer pairs with unique sequence combinations for demultiplexing and sample identification. The cycling parameters, primers and PCR mixtures used for the preparation of barcoded amplicons are indicated in **Tables S15**, **S16** and **S17**, respectively. Next, the samples were again purified with AMPure XP beads and the concentrations of barcoded amplicons were determined by using the Qubit dsDNA HS assay kit (Thermo Fisher Scientific; Cat. No.: Q32854) and a Qubit2.0 fluorometer. Finally, purified amplicons were pooled in equal molar ratios and were then subjected to Illumina MiSeq deep sequencing for retrieving 50,000 paired-end reads. Finally, after demultiplexing and adapter trimming of the paired-end MiSeq raw reads (R1 and R2 fastq files) with Cutadapt version 2.10 [61] alignment of amplicon sequences to reference sequences was carried out by using the CRISPResso2 software [62]. The scripts applied in each CRISPResso2 analyses round are available in **Figure S12-S15.**

Statistical analysis

Statistical analyses were performed with the aid of the GraphPad Prism software (version 8.0.1) on datasets derived from independent biological replicates. Statistical significances were calculated with the tests indicated in the various figure legends. *P*-values lower than 0.05 were considered statistically significant.

Additional Information

Supplementary Data: Supplementary Data can be found online at doi: 10.1016/j.omtn.2023.02.025.

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement

no. 765269. This work was also supported by the Prinses Beatrix Spierfonds (W.OR16-13) and the Dutch Duchenne Parent Project (17.012). Authors of this paper are members of the European Reference Network – Neuromuscular diseases (ERN EURO-NMD).

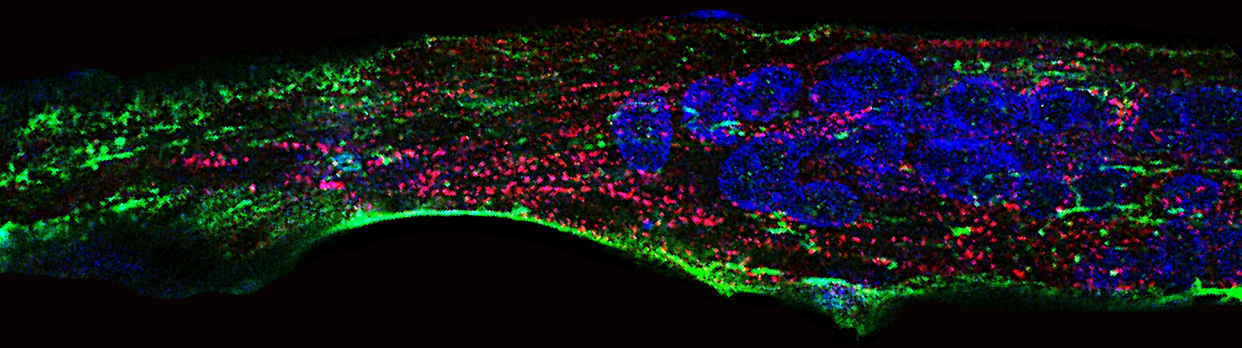
Conflict of interest statement: None decleared.

References

1. Ernst, M.P.T.; Broeders, M.; Herrero-Hernandez, P.; Oussoren, E.; van der Ploeg, A.T.; Pijnappel, W.W.M.P. Ready for Repair? Gene Editing Enters the Clinic for the Treatment of Human Disease. *Mol. Ther. Methods Clin. Dev.* **2020**, *18*, 532-557.
2. Kan, M.J.; Doudna, J.A. Treatment of Genetic Diseases With CRISPR Genome Editing. *JAMA.* **2020**, 328, 980-981.
3. Doudna, J.A.; Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **2014**, *346*, 1258096.
4. Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pirzada, Z.A.; Eckert M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **2011**, *471*, 602-607.
5. Ran, F.A.; Cong, L.; Yan, W.X.; Scott, D.A.; Gootenberg, J.S.; Kriz, A.J.; Zetsche, B.; Shalem, O.; Wu, X.; Makarova, K.S.; et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature.* **2015**, 520, 186-191.
6. Pannunzio, N.R.; Watanabe, G.; Lieber, M.R. Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *J. Biol. Chem.* **2018**, *293*, 10512-10523.
7. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **2013**, *339*, 819-823.
8. Mandal, P.K; Ferreira, L.M.; Collins, R.; Meissner, T.B.; Boutwell, C.L.; Friesen, M.; Vrbanac, V.; Garrison, B.S.; Stortchevoi, A.; Bryder, D.; et al. Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* **2014**, *15*, 643-652.
9. Canver, M.C.; Bauer, D.E.; Dass, A.; Yien, Y.Y.; Chung, J.; Masuda, T.; Maeda, T.; Paw, B.H.; Orkin, S.H. Characterization of genomic deletion efficiency mediated by clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J. Biol. Chem.* **2017**, *292*, 2556.
10. Maeder, M.L.; Stefanidakis, M.; Wilson, C.J.; Baral, R.; Barrera, L.A.; Bounoutas, G.S.; Bumcrot, D.; Chao, H.; Ciulla, D.M.; DaSilva, J.A.; et al. Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. *Nat. Med.* **2019**, *25*, 229-233.
11. Maggio, I.; Stefanucci, L.; Janssen, J.M.; Liu, J.; Chen, X.; Mouly, V.; Gonçalves, M.A.F.V. Selection-free gene repair after adenoviral vector transduction of designer nuclease: rescue of dystrophin synthesis in DMD muscle cell populations. *Nucleic Acids Res.* **2016**, *44*, 1449-1470.
12. Bolukbasi, M.F.; Liu, P.; Luk, K.; Kwok, S.F.; Gupta, A.; Amrani, N.; Sontheimer, E.J.; Zhu, L.J.; Wolfe, S.A. Orthogonal Cas9-Cas9 chimeras provide a versatile platform for genome editing. *Nat Commun.* **2018**, *9*, 4856.
13. Chen, X.; Gonçalves, M.A.F.V. Engineered Viruses as Genome Editing Devices. *Mol. Ther.* **2016**, *24*, 447-457.
14. Gonçalves, M.A.F.V. Adeno-associated virus: from defective virus to effective vector. *Virol. J.* **2005**, *2*, 43.
15. Gao, J.; Mese, K.; Bunz, O.; Ehrhardt, A. State-of-the-art human adenovirus vectorology for therapeutic approaches. *FEBS Lett.* **2019**, *593*, 3609-3622.
16. Tasca, F.; Wang, Q.; Gonçalves, M.A.F.V. Adenoviral Vectors Meet Gene Editing: A Rising Partnership for the Genomic Engineering of Human Stem Cells and Their Progeny. *Cells* **2020**, *9*, 953.
17. Ricobaraza, A.; Gonzalez-Aparicio, M.; Mora-Jimenez, L.; Lumbreras, S.; Hernandez-Alcoceba, R. High-capacity adenoviral vectors: Expanding the scope of gene therapy. *Int. J. Mol. Sci.* **2020**, *21*, 3643.
18. Slaymaker, I.M.; Gao, L.; Zetsche, B.; Scott, D.A.; Yan, W.X.; Zhang, F. Rationally engineered Cas9 nucleases with improved specificity. *Science* **2016**, *351*, 4-88.
19. Maggio, I.; Zittersteijn, H.A.; Wang, Q.; Liu, J.; Janssen, J.M.; Ojeda, I.T.; van der Maarel, S.M.; Lankester, A.C.; Hoeben, R.C.; Gonçalves, M.A.F.V. Integrating gene delivery and gene-editing technologies by adenoviral vector transfer of optimized CRISPR-Cas9 components. *Gene Ther.* **2020**, *27*, 209-225.
20. Dang, Y.; Jia, G.; Choi, J.; Ma, H.; Anaya, E.; Ye, C.; Shankar, P.; Wu, H. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol.* **2015**, *16*, 280.
21. Duan, D.; Goemans, N.; Takeda, S.; Mercuri, E.; Aartsma-Rus, A. Duchenne muscular dystrophy. *Nat. Rev. Dis. Primers.* **2021**, *7*, 13.
22. Bladen, C.L.; Salgado, D.; Monges, S.; Foncuberta, M.E.; Kekou, K.; Kosma, K.; Dawkins, H.; Lamont, L.,; Roy, A.J.; Chamova, T.; et al. The TREAT-NMD DMD Global Database: analysis of more than 7,000 Duchenne

muscular dystrophy mutations. *Hum. Mutat.* **2015**, *36*, 395-402.
23. Chemello, F.; Bassel-Duby, R.; Olson, E.N. Correction of muscular dystrophies by CRISPR gene editing. *J. Clin. Invest.* **2020**, *130*, 2766-2776.
24. Ousterout, D.G.; Kabadi, A.M.; Thakore, P.I.; Perez-Pinera, P.; Brown, M.T.; Majoros, W.H.; Reddy, T.E.; Gersbach, C.A. Correction of dystrophin expression in cells from Duchenne muscular dystrophy patients through genomic excision of exon 51 by zinc finger nucleases. *Mol Ther.* **2015**, *23*, 523-532.
25. Ousterout, D.G.; Kabadi, A.M.; Thakore, P.I.; Majoros, W.H.; Reddy, T.E.; Gersbach, C.A. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause duchenne muscular dystrophy. *Nat. Commun.* **2015**, *6*, 6244.
26. Maggio, I.; Liu, J.; Janssen, J.M.; Chen, X.; Gonçalves, M.A.F.V. Adenoviral vectors encoding CRISPR/Cas9 multiplexes rescue dystrophin synthesis in unselected populations of DMD muscle cells. Sci. Rep. **2016**, *6*, 37051.
27. Brescia, M.; Janssen, J.M.; Liu, J.; Gonçalves, M.A.F.V. High-Capacity Adenoviral Vectors Permit Robust and Versatile Testing of *DMD* Gene Repair Tools and Strategies in Human Cells. *Cells* **2020**, *9*, 869.
28. Young, C.S.; Hicks, M.R.; Ermolova, N.V.; Nakano, H.; Jan, M.; Younesi, S.; Karumbayaram, S.; Kumagai-Cresse, C.; Wang, D.; Zack, J.A.; et al. A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle Cells. *Cell Stem Cell.* **2016**, *18*, 533-540.
29. Long, C.; Amoasii, L.; Mireault, A.A.; McAnally, J.R.; Li, H.; Sanchez-Ortiz, E.; Bhattacharyya, S.; Shelton, J.M.; Bassel-Duby, R.; Olson, E.N. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **2016**, *351*, 400-403.
30. Nelson, C.E.; Hakim, C.H.; Ousterout, D.G.; Thakore, P.I.; Moreb, E.A.; Castellanos, R.M.; Madhavan, S.; Pan, X.; Ran, F.A.; Yan, W.X.; et al. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **2016**, *351*, 403-407.
31. Tabebordbar, M.; Zhu, K.; Cheng, J.K.W.; Chew, W.L.; Widrick, J.J.; Yan, W.X.; Maesner, C.; Wu, E.Y.; Xiao, R.; Ran, F.A.; et al. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **2016**, *351*, 407-411.
32. Xu, L.; Park, K.H.; Zhao, L.; Xu, J.; El Refaey, M.; Gao, Y.; Zhu, H.; Ma, J.; Han, R. CRISPR-mediated Genome Editing Restores Dystrophin Expression and Function in mdx Mice. *Mol. Ther.* **2016**, *4*, 564-569.
33. Maggio, I.; Chen, X.; Gonçalves, M.A.F.V. The emerging role of viral vectors as vehicles for *DMD* gene editing. *Genome Med.* **2016**, *8*, 59.
34. Knaän-Shanzer, S.; Van Der Velde, I.; Havenga, M.J.; Lemckert, A.A.; De Vries, A.A.; Valerio, D. Highly efficient targeted transduction of undifferentiated human hematopoietic cells by adenoviral vectors displaying fiber knobs of subgroup B. *Hum. Gene Ther.* **2001**, *12*, 1989-2005.
35. Gonçalves, M.A.F.V.; de Vries, A.A.; Holkers, M.; van de Watering, M.J.; van der Velde, I.; van Nierop, G.P.; Valerio, D.; Knaän-Shanzer, S. Human mesenchymal stem cells ectopically expressing full-length dystrophin can complement Duchenne muscular dystrophy myotubes by cell fusion. Hum. Mol. Genet. **2006**, *15*, 213-221.
36. Gonçalves, M.A.F.V.; Holkers, M.; Cudré-Mauroux, C.; van Nierop, G.P.; Knaän-Shanzer, S.; van der Velde, I.; Valerio, D.; de Vries, A.A. Transduction of myogenic cells by retargeted dual high-capacity hybrid viral vectors: robust dystrophin synthesis in Duchenne muscular dystrophy muscle cells. *Mol. Ther.* **2006**, *13*, 976-986.
37. Anzalone, A.V.; Randolph, P.B.; Davis, J.R.; Sousa, A.A.; Koblan, L.W.; Levy, J.M.; Chen, P.J.; Wilson, C.,; Newby, G.A.; Raguram, A.; et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **2019**, *576*, 149-157.
38. Anzalone, A.V.; Gao, X.D.; Podracky, C.J.; Nelson, A.T.; Koblan, L.W.; Raguram, A.; Levy, J.M.; Mercer, J.A.M.; Liu, D.R. Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat. Biotechnol.* **2022**, *40*, 731-740.
39. Choi, J.; Chen, W.; Suiter, C.C.; Lee, C.; Chardon, F.M.; Yang, W.; Leith, A.; Daza, R.M.; Martin, B.; Shendure, J. Precise genomic deletions using paired prime editing. *Nat. Biotechnol.* **2022**, *40*, 218-226.
40. Liang, T.; Zhang, X.O.; Weng, Z.; Xue, W Deletion and replacement of long genomic sequences using prime editing. *Nat. Biotechnol.* **2022**, *40*, 227-234.
41. Tao, R.; Wang, Y.; Hu, Y.; Jiao, Y.; Zhou, L.; Jiang, L.; Li, L.; He, X.; Li, M.; Yu, Y.; et al. WT-PE: Prime editing with nuclease wild-type Cas9 enables versatile large-scale genome editing. *Signal Transduct Target Ther.* **2022**, *7*, 108.
42. Wang, J.; He, Z.; Wang, G.; Zhang, R.; Duan, J.; Gao, P.; Lei, X.; Qiu, H.; Zhang, C.; Zhang, Y.; et al. Efficient targeted insertion of large DNA fragments without DNA donors. *Nat. Methods.* **2022**, *19*, 331-340.
43. Zhuang, Y.; Liu, J.; Wu, H.; Zhu, Q.; Yan, Y.; Meng, H.; Chen, P.R.; Yi, C. Increasing the efficiency and precision of prime editing with guide RNA pairs. *Nat. Chem. Biol.* **2022**, *18*, 29-37.
44. Wang, Q.; Liu, J.; Janssen, J.M.; Tasca, F.; Mei, H.; Gonçalves, M.A.F.V. Broadening the reach and investigating the potential of prime editors through fully viral gene-deleted adenoviral vector delivery. *Nucleic Acids Res.*

2021, *49*, 11986-12001.
45. Biressi, S.; Filareto, A.; Rando, T.A. Stem cell therapy for muscular dystrophies. J. Clin. Invest. **2020**, *130*, 5652-5664.
46. Boyer, O.; Butler-Browne, G.; Chinoy, H.; Cossu, G.; Galli, F.; Lilleker, J.B.; Magli, A.; Mouly, V.; Perlingeiro, R.C.R.; Previtali, S.C.; et al. Myogenic Cell Transplantation in Genetic and Acquired Diseases of Skeletal Muscle. *Front. Genet.* **2021**, *12*, 702547.
47. Gonçalves, M.A.F.V.; van der Velde, I.; Knaän-Shanzer, S.; Valerio, D.; de Vries, A.A. Stable transduction of large DNA by high-capacity adeno-associated virus/adenovirus hybrid vectors. *Virology* **2004**, *321*, 287-296.
48. Mamchaoui, K.; Trollet, C.; Bigot, A.; Negroni, E.; Chaouch, S.; Wolff, A.; Kandalla, P.K.; Marie, S.; Di Santo, J.; St Guily, J.L.; et al. Immortalized pathological human myoblasts: towards a universal tool for the study of nuromuscular disorders. *Skelet. Muscle* ***2011****, 1*, 34.
49. Thorley, M.; Duguez, S.; Mazza, E.M.C.; Valsoni, S.; Bigot, A.; Mamchaoui, K.; Harmon, B.; Voit, T.; Mouly, V.; Duddy, W. Skeletal muscle characteristics are preserved in hTERT/cdk4 human myogenic cell lines. *Skelet. Muscle* **2016**, *6*, 43.
50. Chen, X.; Janssen, J.M.; Liu, J.; Maggio, I.; 't Jong, A.E.J.; Mikkers, H.M.M.; Gonçalves, M.A.F.V. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat. Commun.* **2017**, *8*, 657.
51. Kleinstiver, B.P.; Prew, M.S.; Tsai, S.Q.; Nguyen, N.T.; Topkar, V.V.; Zheng, Z., Joung, J.K. Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* ***2015****, 33*, 1293-1298.
52. Papapetrou, E.P.; Schambach, A. Gene insertion into genomic safe harbors for human gene therapy. *Mol. Ther.* **2016**, *24*, 678-84.
53. Pavani, G.; Amendola, M. Targeted Gene Delivery: Where to Land. *Front. Genome Ed.* **2021**, *2*, 609650.
54. Chen, X.; Tasca, F.; Wang, Q.; Liu, J.; Janssen, J.M.; Brescia, M.D.; Bellin, M.; Szuhai, K.; Kenrick, J.; Frock, R.L.; et al. Expanding the editable genome and CRISPR-Cas9 versatility using DNA cutting-free gene targeting based on in trans paired nicking. *Nucleic Acids Res.* **2020**, *48*, 974-995.
55. Janssen, J.M.; Liu, J.; Skokan, J.; Gonçalves, M.A.F.V.; de Vries AA. Development of an AdEasy-based system to produce first- and second-generation adenoviral vectors with tropism for CAR- or CD46-positive cells. *J. Gene Med.* **2013**, *15*, 1-11.
56. Holkers, M.; Maggio, I.; Henriques, S.F.D.; Janssen, J.M.; Cathomen, T.; Gonçalves, M.A.F.V. Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat. Methods.* **2014**, *11*, 1051–7.
57. Pelascini, L.P.L.; Gonçalves, M.A.F.V. Lentiviral vectors encoding zinc-finger nucleases specific for the model target locus HPRT1. *Methods Mol. Bio.* **2014**, *1114*, 181-99.
58. Pelascini, L.P.L.; Janssen, J.M.; Gonçalves, M.A.F.V. Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. *Hum. Gene Ther.* **2013**, *24*, 78-96.
59. Zufferey, R.; Dull, T.; Mandel, R.J.; Bukovsky, A.; Quiroz, D.; Naldini, L.; Trono, D. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J. Virol.* **1998**, *72*, 9873-9880.
60. Wang, Q.; Liu, J.; Janssen, J.M.; Le Bouteiller, M.; Frock, R.L.; Gonçalves, M.A.F.V. Precise and broad scope genome editing based on high-specificity Cas9 nickases. *Nucleic Acids Res.* **2021**, *49*, 1173-1198.
61. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2011**, *17*, 10–12.
62. Clement, K.; Rees, H.; Canver, M.C.; Gehrke, J.M.; Farouni, R.; Hsu, J.Y.; Cole, M.A.; Liu, D.R.; Joung, J.K.; Bauer, D.E.; et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **2019**, *37*, 224-226.

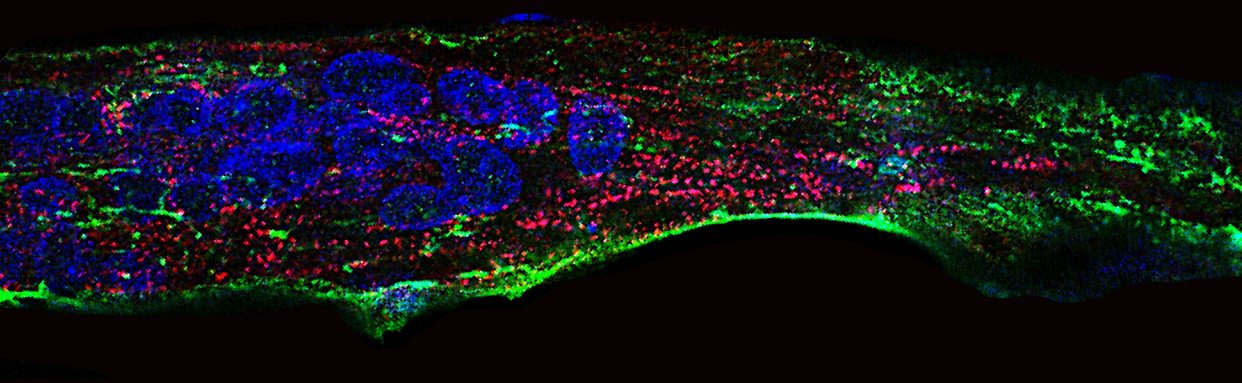


Chapter 4

Large-scale genome editing based on high-capacity adenovector particles and CRISPR-Cas9 nucleases rescues full-length dystrophin synthesis in DMD muscle cells

Francesca Tasca[1], Marcella Brescia[1], Qian Wang[1], Jin Liu[1], Josephine M. Janssen[1], Karoly Szuhai[1], Manuel A.F.V. Gonçalves[1]

Nucleic Acids Research, 2022, **50**(13):7761-7782



[1]Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands

Illustration: Modified from immunofluorescence picture of patient-derived myotubes after treatment. (See Figure 4C for techincal details)

Abstract

Targeted chromosomal insertion of large genetic payloads in human cells leverages and broadens synthetic biology and genetic therapy efforts. Yet, obtaining large-scale gene knock-ins remains particularly challenging especially in hard-to-transfect stem and progenitor cells. Here, fully viral gene-deleted adenovector particles (AdVPs) are investigated as sources of optimized high-specificity CRISPR-Cas9 nucleases and donor DNA constructs tailored for targeted insertion of full-length dystrophin expression units (up to 14.8-kb) through homologous recombination (HR) and homology-mediated end joining (HMEJ). In muscle progenitor cells, donors prone to HMEJ yielded higher CRISPR-Cas9-dependent genome editing frequencies than HR donors, with values ranging between 6%-34%. In contrast, AdVP transduction of HR and HMEJ substrates in induced pluripotent stem cells (iPSCs) resulted in similar CRISPR-Cas9-dependent genome editing levels. Notably, when compared to regular iPSCs, in p53 knockdown iPSCs, CRISPR-Cas9-dependent genome editing frequencies increased up to 6.7-fold specifically when transducing HMEJ donor constructs. Finally, single DNA molecule analysis by molecular combing confirmed that AdVP-based genome editing achieves long-term complementation of DMD-causing mutations through the site-specific insertion of full-length dystrophin expression units. In conclusion, AdVPs are a robust and flexible platform for installing large genomic edits in human cells and p53 inhibition fosters HMEJ-based genome editing in iPSCs.

Introduction

Genome editing is a fast-evolving field with increasing impact in basic science, biotechnology, and medicine [1, 2]. Particularly versatile genome editing strategies permit incorporating exogenous donor sequences into endogenous loci subjected to double-strand DNA breaks (DSBs) made by engineered CRISPR-Cas9 nucleases [3-9]. This versatility stems from the amenability of these gene knock-in strategies to genomic modifications spanning from single base-pairs to large transgene(s); and from the straightforward designing of CRISPR-Cas9 nucleases with high activities and specificities [3-9]. Indeed, in contrast to earlier programmable nucleases, CRISPR-Cas9 nucleases are protein engineering-free, in that they consist of sequence-customizable guide RNA (gRNA) and immutable RNA-programmable Cas9 proteins that cleave target sequences upon gRNA-DNA hybridization [9, 10]. Hence, targeted DNA knock-ins can be accomplished by delivering CRISPR-Cas9 nucleases together with donor DNA constructs whose designs favour site-specific DSB repair through either; non-homologous end joining (NHEJ) or homology-directed repair (HDR) pathways [11, 12], i.e., homologous recombination (HR), microhomology-mediated end joining (MMEJ) and, more recently, homology-mediated end joining (HMEJ) [13, 14]. In contrast to the lack of homology to target sequences in NHEJ-prone donors [15, 16], MMEJ, HMEJ and HR donors present increasingly larger homology tracts flanking the foreign DNA of interest, with each homology arm typically spanning 20-30 bp, ~900 bp and 0.5-2.0 kb, respectively. Moreover, diversely from HR donors, donors tailored for ectopic NHEJ, MMEJ and HMEJ, have their targeting modules flanked by CRISPR-Cas9 cleaving sites. This "double-cut" arrangement ensures exogenous DNA release from construct backbones in cell nuclei, fostering gene knock-ins via the processing and alignment of donor and target DNA termini [11, 12].

When compared to NHEJ and MMEJ donors, more exact and properly oriented chromosomal integration of exogenous DNA is achieved through HMEJ and HR donor designs [15]. In addition, it is well-established that the efficiency and precision of ectopic HR profits from extending homology tracts especially when aiming at chromosomal insertion of larger genetic payloads [13, 17, 18]. Normally, HMEJ donors yield higher gene knock-in frequencies than HR, NHEJ or MMEJ donors [13, 14], however, the performance of HMEJ donors containing homology lengths considerably longer than the typical ~900-bp, has not been assessed. Equally of notice, HMEJ-based genome editing, similarly to other strategies based on "double-cut" donors, can take place in HR refractory non-dividing cells, turning it into a high-potential approach for *in vivo* applications [14, 19]. Notwithstanding, unwanted chromosomal insertion of prokaryotic backbone sequences is associated with donor plasmid delivery, especially when applying the "double-cut" genome editing strategies [20]. Critically, plasmids harbouring large transgenes and/or homology tracts transfect poorly even in easy-to-transfect cells which, often, demands complex and time-consuming cell selection procedures.

Viral vectors instead achieve efficient delivery of genome editing tools into hard-to-transfect cell types [21]. However, commonly used adeno-associated viral vectors cannot

deliver large transgenes nor large homology tracts due to their limited packaging capacity (<4.7 kb) [21, 22]. There is, therefore, a pressing need for alternative DNA delivery systems allowing the efficacious investigation and application of novel genome editing principles independently of the size of the attendant tools. In this regard, high-capacity adenoviral vectors (also named third-generation adenoviral vectors), henceforth dubbed adenovector particles (AdVPs), congregate a valuable set of features, namely, (i) lack of viral genes; (ii) vast packaging capacity (up to 36 kb), (iii) high genetic stability; (iv) amenability to straightforward cell-tropism modifications; and (v) efficient transduction of dividing and non-dividing cells [23-25].

Here, we demonstrate that AdVPs are suitable for engineering large-scale genomic edits in human stem and progenitor cells upon the delivery of optimized high-specificity CRISPR-Cas9 nucleases and donor constructs tailored for HR or HMEJ. In parallel, these tools were applied for testing the rescue of the genetic defect underlying Duchenne muscular dystrophy (DMD) in human myogenic cells, i.e., muscle progenitor cells (myoblasts) and induced pluripotent stem cells (iPSCs). DMD (MIM #310200) is a lethal and frequent muscle-wasting X-linked disorder (prevalence: ~1 in 4700 boys) caused by a multitude of diverse types of mutations scattered along the enormous *DMD* gene (~2.4 Mb). These mutations disrupt striated muscle-specific dystrophin isoforms (427-kDa) encoded in 14-kb mRNA transcripts with 79 exons [26]. The absence of cytoskeleton-to-dystrophin glycoprotein complex (DGC) linkages in muscle cells results in sarcolemma fragility and impaired cell signalling. Eventually, this leads to the replacement of damaged muscle with fibrotic and adipose tissues [26]. Currently, the vast majority of DMD-directed genetic therapies are mutation-specific and/or yield only partially functional micro-dystrophins or shortened Becker-like dystrophins [26, 27]. Complementation of DMD-causing mutations regardless of their type or location via stable expression of full-length dystrophin offers the perspective for more effective and broadly applicable approaches, including those involving *ex vivo* correction and autologous transplantation of stem/progenitor cells with myogenic capacity [28, 29].

Notably, iPSCs represent a particularly valuable cell source for the development of DMD-targeting genetic therapies [28, 29]. Indeed, iPSCs derived from reprogrammed human somatic cells are capable of unlimited self-renewal *in vitro* and, under proper stimuli, differentiate into specific cell types [30], including skeletal and cardiac muscle cells. These unique features support *in vitro* disease modelling and, in combination with genome editing technologies, the development of candidate autologous cell therapies. Yet, in this cell type, genome editing mostly involves small-scale edits delivered in oligonucleotide or plasmid DNA substrates. The former substrates can yield small but high-frequency genome editing; the latter, in contrast, normally require linkage to laborious positive-selection genes whose removal depends on laborious supplementation of site-specific recombinases or transposases. Critically, homology-directed installation of larger edits renders the iPSC genome editing process even more challenging due to the difficulty in transferring the correspondingly sizable

genetic payloads into these cells in an efficient and non-cytotoxic manner [9, 10].

In this study, we show that AdVP delivery of donors prone to HMEJ and HR, together with matched CRISPR-Cas9 complexes, achieves targeted integration of transgenes encoding full-length, hence fully functional, dystrophin in HeLa cells, myoblasts and iPSCs. In myoblasts, HMEJ donors led to significantly higher frequencies of site-specific transgene integration than HR donors. Via additional AdVP transduction experiments, we further found that HMEJ-based genome editing is compromised in iPSCs yet, it can be enhanced via p53 inhibition. Importantly, edited myoblasts kept stable recombinant full-length dystrophin protein synthesis and differentiation capability. Finally, we confirmed that CRISPR-Cas9-dependent stable full-length dystrophin expression is, in most cases, the result of the precise chromosomal insertion of HMEJ and HR donor sequences at a commonly used safe harbour locus, i.e., the adeno-associated virus integration site 1 (*AAVS1*) at 19q13.4-qter.

Results

AdVPs achieve all-in-one delivery of optimized CRISPR-Cas9 complexes inducing robust and specific DNA cleavage

By lacking only a few viral ORFs, first- and second-generation adenoviral vectors do not permit exploiting the full DNA packaging capacity of adenoviral capsids, i.e., 36-kb [21, 53]. In addition, at high multiplicities of infection (MOI), "leaky" expression from vector resident viral ORFs contributes to cytotoxic effects *in vitro* and immune responses *in vivo* [53]. Therefore, in this study, we selected fully viral gene-deleted AdVPs [21, 23-25] for investigating large-scale cell engineering strategies based on the recruitment of homology-directed gene targeting processes. And, to broaden the target cell range, AdVPs were endowed with adenovirus type-50 fibers that, by engaging the ubiquitously expressed CD46 receptor [24, 54], permit efficient transduction of otherwise refractory human myoblasts [55] and other coxsackievirus and adenovirus receptor (CAR)-negative cells with high therapeutic relevance. Hence, we first sought to generate a CD46-binding AdVP, namely AdVP.eCas9^{4NLS}gRNA^{S1}, for all-in-one transfer of optimized CRISPR-Cas9 components targeting *AAVS1* loci at 19q13.4-qter. We decided to target *AAVS1* owing to its common use as a safe harbor for transgene insertion and stable expression in a wide range of human cell types [56-58]. AdVP.eCas9^{4NLS}gRNA^{S1} encodes a variant of the high-specificity eSpCas9(1.1) nuclease [59], called eCas9^{4NLS}, whose improved performance derives from having 2 extra nuclear localization signals (NLS) [60]; and an *AAVS1*-specific gRNA, named gRNA^{S1}, that harbors an optimized Cas9-binding scaffold [61] (**Figure 1A**). Importantly, AdVP.eCas9^{4NLS}gRNA^{S1} particles were produced at high titers (**Table S3**) and contained structurally intact DNA with no evidence for rearranged or truncated species, as shown by restriction fragment length analysis (RFLA) (**Figures 1B**). To assess the functionality of AdVP.eCas9^{4NLS}gRNA^{S1} in delivering active eCas9^{4NLS}:gRNA^{S1} complexes into human cells, we transduced cervical carcinoma HeLa cells, wild-type human myoblasts and myoblasts derived from two separate

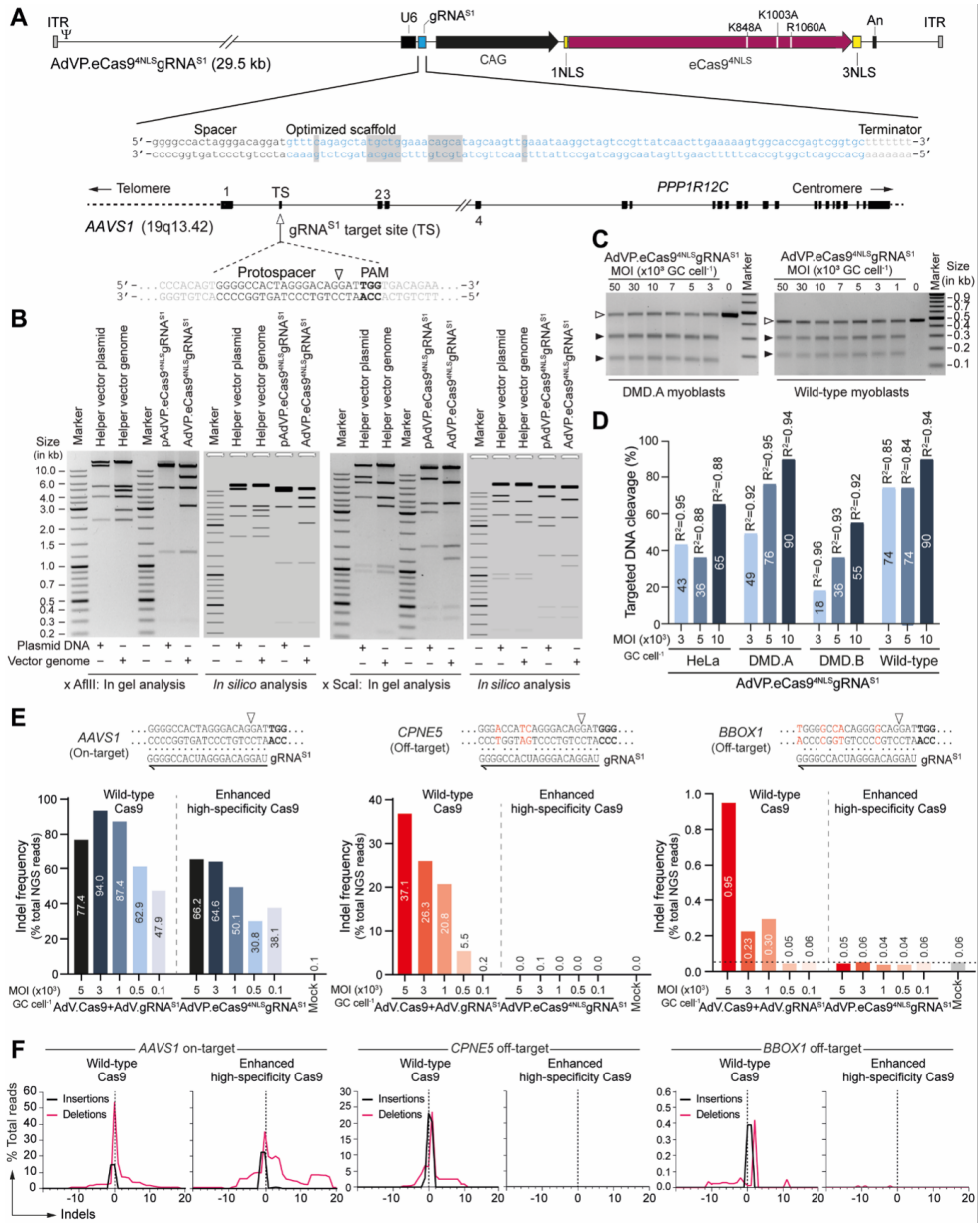


Figure 1. Characterization and testing of AdVP for all-in-one delivery of optimized *AAVS1*-targeting RGN complexes. (**A**) Schematics of AdVP genome encoding eCas9^{4NLS}:gRNA^{S1} complexes and *AAVS1* target site. Enhanced high-specificity eCas9^{4NLS} nuclease and optimized gRNA^{S1} synthesis are driven by hybrid CAG regulatory sequences and the human U6 gene promoter, respectively. The point mutations conferring enhanced specificity to the optimized eCas^{4NLS} nuclease are specified. The point mutations and insertions in the gRNA scaffold coding sequence that maximize the expression of full-length RNA molecules with a stabilizing extended stem-loop are shaded. NLS, nuclear localization signal (NLS); ITR and Ψ, adenoviral inverted terminal repeats and packaging signal cis-acting elements for vector DNA replication and encapsidation, respectively. The *AAVS1* target site locates in the first intron of *PPP1R12C* (19q13.42). The gRNA^{S1} protospacer and protospacer adjacent motif (PAM) sequences are highlighted. Open arrowheads indicate the DNA cleavage position. (**B**) Assessing AdVP.eCas9^{4NLS}gRNA^{S1} DNA integrity. Restriction fragment

◄ length analysis (RFLA) of vector DNA isolated from purified AdVP.eCas9^{4NLS}gRNA^{S1} particles. In-silico and in-gel RFLA analyses are presented. Parental circular plasmid pAdVP.eCas9^{4NLS}gRNA^{S1} and helper vector DNA used to assemble AdVP.eCas9^{4NLS}gRNA^{S1} particles served as molecular weight references. (**C**) Probing AdVP.eCas9^{4NLS}gRNA^{S1} functionality. The indicated target cells were exposed to different amounts of AdVP.eCas9^{4NLS}gRNA^{S1} and site-specific DNA cleavage was assessed at 3 days post-transduction through the detection of DSB-derived indels using the mismatch-sensing T7EI enzyme. Solid and open arrowheads point to DNA species derived from T7EI-digested and undigested amplicons, respectively. MOI, multiplicity of infection in genome copies per cell (GC cell^{-1}); Marker, GeneRuler DNA Ladder Mix. (**D**) Quantification of target DNA cleavage. The eCas9^{4NLS}:gRNA^{S1} activities were measured by deconvolution of Sanger sequence traces corresponding to *AAVS1*-specific amplicons derived from the indicated target cells exposed to three doses of AdVP.eCas9^{4NLS}gRNA^{S1}. (**E**) Assessing the specificity of Cas9:gRNA^{S1} versus eCas9^{4NLS}:gRNA^{S1} complexes. Wild-type myoblasts were exposed to AdVP.eCas9^{4NLS}gRNA^{S1} or to AdVP.Cas9 and AdVP.gRNA^{S1} at the indicated total MOI. DNA cleaving activities at the *AAVS1* target site and at two validated off-target sites (i.e., *CPNE5* and *BBOX1*) were quantified by amplicon deep sequencing of DSB-derived indels at three days post-transduction (~50,000 paired-end reads per sample). Nucleotide mismatch positions between gRNA^{S1} spacer and off-target *CPNE5* and *BBOX1* sequences is highlighted in red. (**F**) Characterization of nuclease-induced indel footprints by amplicon deep sequencing. Wild-type myoblasts were exposed to AdVP.eCas9^{4NLS}gRNA^{S1} or to AdVP.Cas9 and AdVP.gRNA^{S1} at an MOI of 5×10^3 GC cell^{-1} each. The types and distributions of indels detected within *AAVS1*, *CPNE5* and *BBOX1* are plotted.

DMD patients (hereinafter named DMD.A and DMD.B myoblasts) at MOI ranging from 1×10^3 to 50×10^3 genome copies per cell (GC cell^{-1}). As detected through a T7 endonuclease I (T7EI)-based genotyping assay at three days post-transduction, eCas9^{4NLS}:gRNA^{S1} readily led to DSB formation at *AAVS1* in all cell types tested (**Figure 1C** and **Figure S2**). After applying AdVP.eCas9^{4NLS}gRNA^{S1} at 3×10^3, 5×10^3 and 10×10^3 GC cell^{-1}, targeted DNA cleaving activities ranging from 43% to 65%, 49% to 90%, 18% to 55% and 74% to 90% were measured by Sanger sequence deconvolution [45] in HeLa cells, DMD.A myoblasts, DMD.B myoblasts and wild-type myoblasts, respectively (**Figure 1D**). To assess the specificities of regular Cas9:gRNA and optimized eCas9^{4NLS}:gRNA^{S1} complexes, wild-type myoblasts were co-transduced with adenovectors AdV.Cas9 and AdV.gRNA^{S1} [46] or transduced with AdVP.eCas9^{4NLS}gRNA^{S1}, respectively. Consistent with earlier findings derived from orthogonal HTGTS assays in HEK293T cells [34], off-target DNA cleavage at *CPNE5* and *BBOX1* was readily detected by amplicon deep sequencing in human myoblasts exposed to Cas9:gRNA^{S1} complexes (**Figures 1E** and **1F**). Crucially, independently of the AdVP dosages used, off-target DNA cleavage at the same off-target sites was virtually undetected in human myoblasts that had been instead exposed to optimized eCas9^{4NLS}:gRNA^{S1} complexes (**Figures 1E** and **1F**). Interestingly, supporting the view that Cas9 variants can lead to altered indel footprint profiles, at the highest vector doses applied, deletion size distributions induced by eCas9^{4NLS}:gRNA^{S1} were wider than those triggered by Cas9:gRNA^{S1} (**Figure S3**).

Altogether, these data indicate that AdVP.eCas9^{4NLS}gRNA^{S1} is a reliable tool to efficiently and specifically induce targeted DSBs at the commonly used *AAVS1* safe harbor locus in human cells.

HMEJ donors yield higher CRISPR-Cas9-dependent genome editing frequencies than HR donors after AdVP delivery

Next, we sought to couple site-specific genomic DNA cleavage to AdVP delivery of donor DNA matched to *AAVS1* targeting through HR and HMEJ (**Figure 2A**). To this end, AdVP.EGFP::DYS and AdVP.EGFP::DYS^{TS}, encoding full-length dystrophin fused to

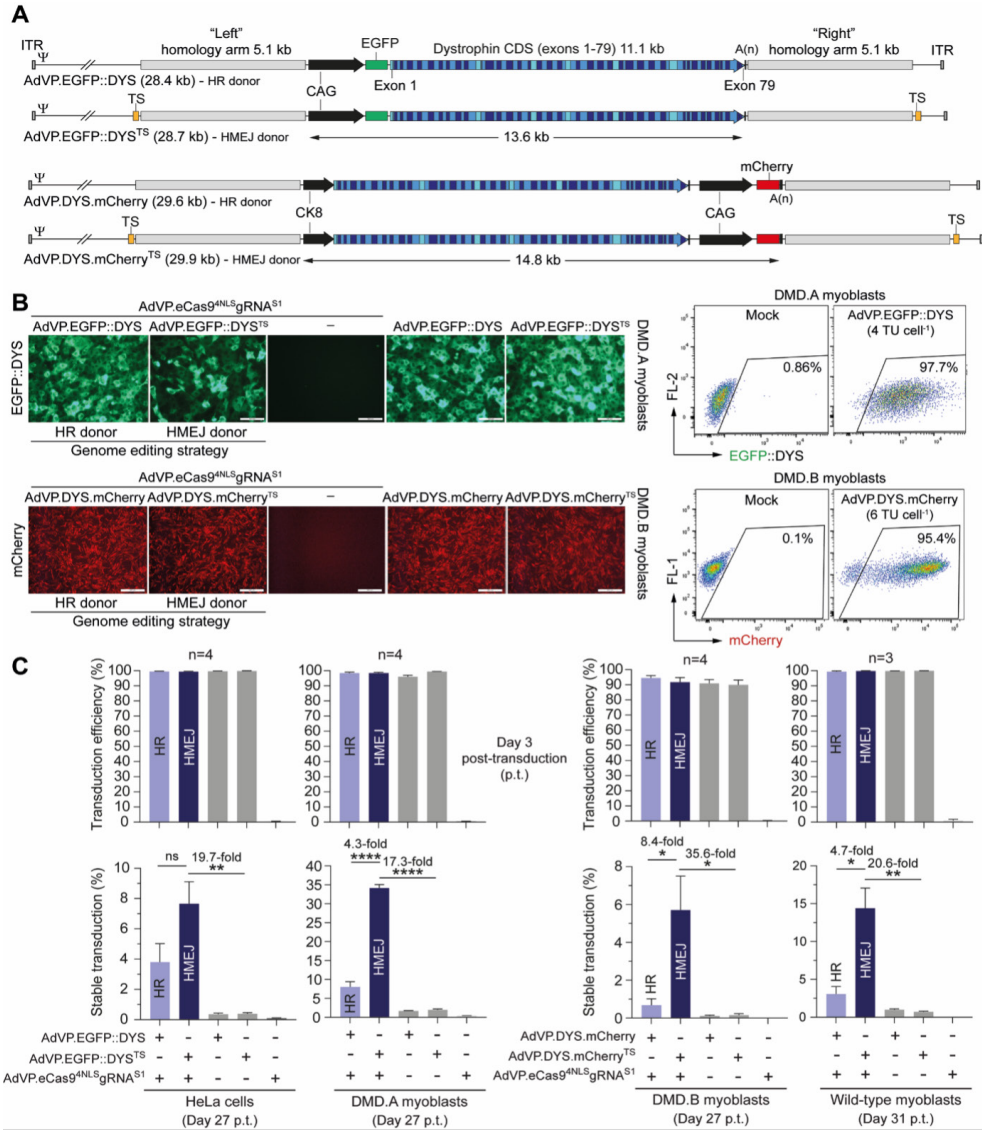


Figure 2. Assembly and testing of AdVP donors designed for homology-directed genome editing. (**A**) Schematics of AdVP donor structures. In AdVP.EGFP::DYS and AdVP.EGFP::DYS^{TS}, the CAG promoter drives the synthesis of a fusion product between EGFP and the human full-length dystrophin (EGFP::DYS). In AdVP.DYS.mCherry and AdVP.DYS.mCherry^{TS}, striated muscle-specific CK8 and constitutive CAG promoters drive the synthesis of full-length dystrophin (DYS) and the mCherry live-cell reporter, respectively. The recombinant DNA in the various vectors is flanked by sequences homologous to the human *AAVS1* safe harbor locus, for testing homology-directed gene targeting upon site-specific DSB formation. AdVP.EGFP::DYS^{TS} and AdVP.DYS.mCherry^{TS} differ from AdVP.EGFP::DYS and AdVP.DYS.mCherry in that they have their targeting modules flanked by the gRNA^{S1} target site (TS). This arrangement guarantees targeted DSB formation at endogenous and exogenous DNA sequences for generating donor substrates amenable to HMEJ. ITR and Ψ, inverted terminal repeats and packaging elements, respectively. (**B**) Testing AdVP donors in human myoblasts cells. Transduction efficiencies were determined by reporter-directed fluorescence microscopy and flow cytometry at three days post-transduction (left and right panels, respectively). Representative micrographs and dot plots of DMD myoblasts from two

◂ different sources (**A** and **B**) transduced with the indicated AdV donors are shown. Controls consisted of cells transduced solely with each AdVP donor or only with AdVP.eCas9^{4NLS}gRNA^{S1}. (**C**) Assessing stable transduction frequencies upon AdPV-mediated delivery of HR and HMEJ donors. Transduction efficiencies (top graphs) and stable transduction levels (bottom graphs) reached in the indicated cell types were determined by reporter-directed flow cytometry at three days and over three weeks post-transduction, respectively. The AdVP transduction conditions used in these experiments are listed in **Table S20**. Data are presented as mean ± SEM of either four or three independent biological replicates. Significant differences between the indicated datasets were determined by Student's t-tests; ****P<0.0001, **P<0.01; *P<0.05. P>0.05 was considered non-significant (ns).

the enhanced green fluorescent protein, were generated (**Figure 2A**). To test the versatility of AdVP-based genome editing strategies, AdV.DYS.mCherry and AdV.DYS.mCherry^{TS} were also assembled (**Figure 2A**). The former and latter vector have essentially the same structure as AdV.EGFP::DYS and AdV.EGFP::DYS^{TS}, respectively, except that, instead of a single expression unit, two independent expression units drive the synthesis of full-length dystrophin and reporter mCherry proteins (**Figure 2A**). Of notice, recombinant dystrophin and mCherry synthesis are under the control of late striated muscle-specific regulatory elements and a constitutively active promoter (i.e., CK8 and CAG, respectively) (**Figure 2A**). This bicistronic design guarantees therapeutic gene expression specifically in differentiated muscle cells and offers the possibility for isolating stem/progenitor cells with myogenic capacity by constitutive expression of a clinically applicable selection marker, e.g., truncated nerve growth factor receptor [62]. AdVP.EGFP::DYS and AdVP.DYS.mCherry have their expression cassettes surrounded by 5.1 kb of DNA sequences identical to the genomic DNA flanking the gRNA^{S1} target site (homology arms) to favor DSB-induced transgene insertion through HR. In order to engage not only HR but also HMEJ processes, AdVP.EGFP::DYS^{TS} and AdVP.DYS.mCherry^{TS} contain the gRNA^{S1} target site at the outward extremities of the 5.1 kb homology arms. This arrangement assures eCas9^{4NLS}:gRNA^{S1}-mediated donor DNA release from AdVP genomes, to facilitate its targeted insertion through HMEJ.

Like for AdVP.eCas9^{4NLS}gRNA^{S1}, the integrity of vector genomes in all AdVP donors was confirmed by RFLA of DNA isolated from purified vector particles (**Figure S4**). Physical particle titers were determined by quantification of packaged vector genome copies (**Table S3**), while functional particle titers of purified preparations of AdVP.EGFP::DYS, AdVP.EGFP::DYS^{TS}, AdVP.DYS.mCherry and AdVP.DYS.mCherry^{TS} were in turn determined by end-point titration assays on HeLa cells followed by quantification of transduced cells by reporter-directed flow cytometry. The functional particle titers of AdVP.EGFP::DYS, AdVP.EGFP::DYS^{TS}, AdVP.DYS.mCherry and AdVP.DYS.mCherry^{TS} were, respectively, 5.23×10^{8} Hela-transducing units per ml (TU ml^{-1}), 3.92×10^{8} TU ml^{-1}, 7.38×10^{10} TU ml^{-1} and 7.42×10^{10} TU ml^{-1} (**Table S3**). Importantly, transduction of human myoblasts with these vectors resulted in transgene expression in most of the target cells independent of their individual origins (**Figure 2B**).

Next, to assess stable transgene expression levels, human cells were initially transduced with AdVP.eCas9^{4NLS}gRNA^{S1} together with AdVP donors whose genomes are insensitive (i.e., AdVP.EGFP::DYS and AdVP.DYS.mCherry) or susceptible (i.e., AdVP.EGFP::DYS^{TS} and AdVP.DYS.mCherry^{TS}) to site-specific DNA cleavage. Controls consisted of cells transduced

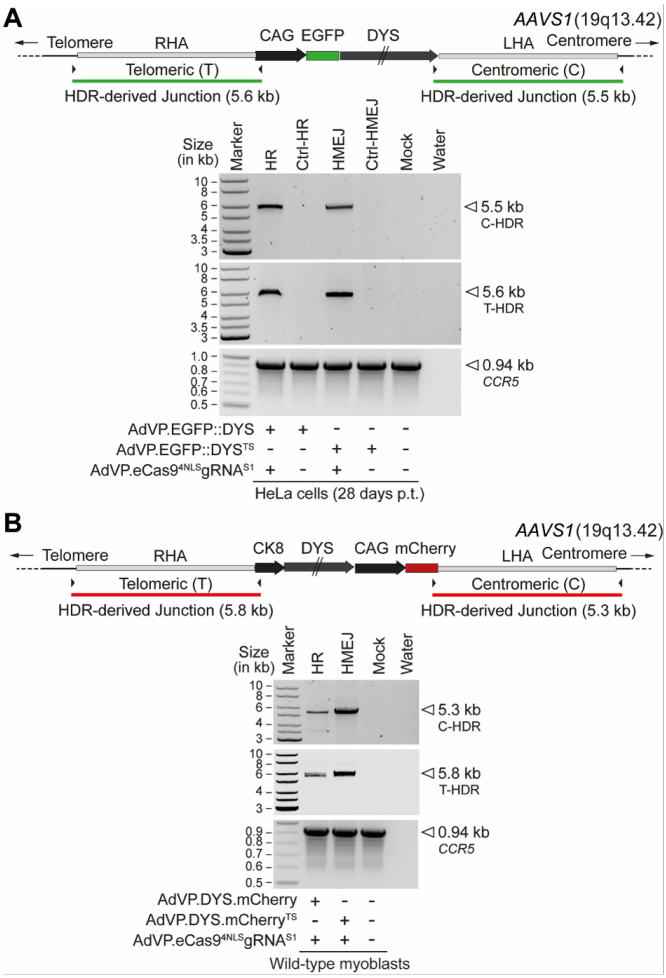


Figure 3. Targeted chromosomal DNA integration in human cells upon AdVP transduction. (**A**) AdVP-based genome editing in HeLa cells. Long-range junction PCR detection of *AAVS1*-targeted insertions in unselected HeLa cell populations exposed to the indicated donor and nuclease constructs at 8 TU cell^{-1} and 3×10^3 GC cell^{-1}, respectively. (**B**) AdVP-based genome editing in human myoblasts. Long-range junction PCR detection of *AAVS1*-targeted insertions in human myoblasts genetically modified through co-transduction with the indicated donor and nuclease constructs at 102 TU cell^{-1} and 5×10^3 GC cell^{-1}, respectively. Amplicons diagnostic for telomeric-sided transgenic-*AAVS1* junctions (T-HDR) and centromeric-sided transgenic-*AAVS1* junctions (C-HDR) junctions are depicted. *CCR5* served as an internal control template. Marker, GeneRuler DNA Ladder Mix. RHA and LHA, "right" and "left" homology arms, respectively.

with each AdVP individually. Dual color fluorescence microscopy analyses confirmed the capacity of AdVPs delivering separately nuclease and donor DNA constructs to co-transduce target cells (**Figure S5**), and western blot analysis established transient nuclease expression in dividing cell populations, as shown by the rapid decline in nuclease amounts after a peak at 2 days post-transduction (**Figure S6**).

At three days post-transduction, over 90% of target cells exposed to AdVP donors

expressed reporter proteins, as determined through flow cytometry (**Figure 2C** , top graphs). The transduced cell populations were then sub-cultured for more than 3 weeks to remove episomal vector DNA (**Figure S7**) and, during this period, were monitored through reporter-directed flow cytometry (**Figure S8**). This analysis revealed a clear CRISPR-Cas9-dependent increase in stable transduction levels (**Figure 2C** bottom graphs and **Figure S8**). In addition, although the frequencies of stably transduced cells varied in a target cell-dependent manner, AdVP donors delivering templates susceptible to HMEJ (i.e., AdVP.EGFP::DYS^{TS} and AdVP.DYS.mCherry^{TS}) invariably led to higher stable transduction levels than those transferring templates strictly susceptible to HR (i.e., AdVP.EGFP::DYS and AdVP.DYS.mCherry) (**Figure 2C**, bottom graphs and **Figure S8**). These flow cytometry datasets were consistent with those obtained through qPCR tracing of HR and HMEJ donor DNA upon AdVP transductions (**Figure S7**). Finally, junction PCR assays revealed that the clear CRISPR-Cas9-dependent increase in stable transduction levels (**Figure 2C** bottom graphs, and **Figures S7** and **S8**), was accompanied by the detection of targeted chromosomal insertion of HR and HMEJ donor DNA in cells exposed to eCas9^{4NLS}:gRNA^{S1} complexes (**Figure 3**).

Long-term constitutive expression of full-length dystrophin was confirmed through confocal microscopy and western blot analyses of genetically corrected DMD muscle cell populations prior to, and after, myogenic differentiation (**Figures 4A, 4B** and **S9**). Co-detection in myotubes of the late skeletal muscle-specific marker sarcomeric α-actinin and full-length dystrophin by confocal microscopy ascertained the differentiation capacity of DMD muscle progenitor cells endowed with constitutive and *CK8*-regulated expression units (**Figure 4C** and **S9**, respectively). Taken together, these data establish that HMEJ donors delivered in the context of AdVP genomes yield higher CRISPR-Cas9-mediated genetic correction than their HR counterparts in muscle progenitor cells and that the AdVP platform can be tailored for the permanent complementation of DMD-causing mutations in patient-derived muscle cells.

Large-scale engineering of iPSCs through AdVP-based HR and HMEJ genome editing

Combining genome editing and iPSC technologies is appealing for establishing "disease-in-a-dish" models, building robust multi-component synthetic gene circuits, and investigating autologous cell therapies [63-65]. In this context, the use of AdVPs as sources of donor DNA has been hitherto mostly explored for mutation-specific gene correction through spontaneous HR [24]. Despite the association of large homology tracts to short exogenous sequences, these programmable nuclease-independent AdVP genome editing approaches yield low frequencies of engineered iPSCs, as generally determined by the ratios between drug-resistant to transduced-cell numbers (i.e., 10^{-6}-10^{-5}). Moreover, the performance of HMEJ donors in iPSCs and its comparison with that of conventional HR donors requires investigation.

Therefore, we next set out to test AdVP-based HR and HMEJ strategies for large-scale engineering of iPSCs and, to this end, DMD patient-derived iPSCs (DMD iPSCs) were transduced with AdVP.eCas9^{4NLS}gRNA^{S1} in combination with AdV.DYS.mCherry

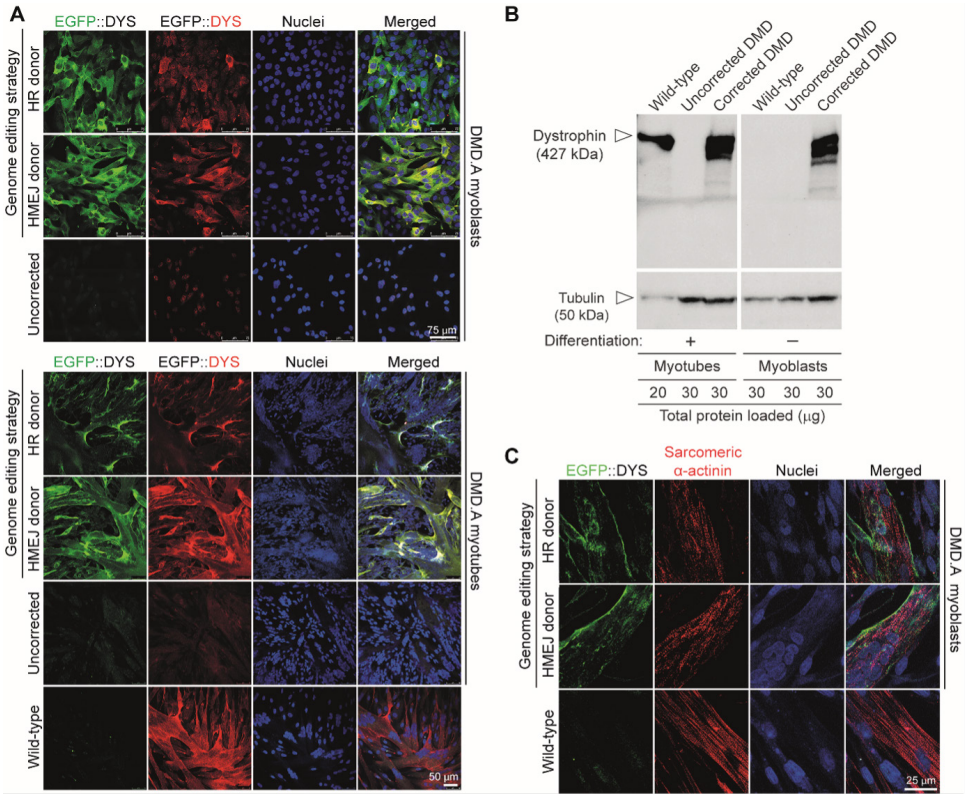


Figure 4. Rescue of full-length dystrophin synthesis in AdVP-corrected DMD muscle cells. (**A**) Detection of dystrophin by fluorescence microscopy. Fluorescence microscopy analyses on DMD patient-derived muscle cells stably expressing EGFP::DYS after AdVP delivery of eCas9^{4NLS}:gRNA^{S1} together with HR or HMEJ donor templates. These analyses were done before and after muscle cell differentiation (top and bottom panels, respectively). Uncorrected DMD muscle cells and healthy donor-derived muscle cells served as negative and positive controls, respectively. (**B**) Detection of dystrophin by western blot analysis. Western blotting was performed on uncorrected and corrected DMD muscle cells after (+) and before (-) myogenic differentiation. Differentiated wild-type muscle cells served as a reference for endogenous full-length dystrophin expression. Detection of tubulin provided for protein loading controls. (**C**) Detection of sarcomeric α-actinin by immunofluorescence microscopy. Assessing the differentiation capacity of AdVP-edited DMD myoblasts by co-detection of fluorescence signals specific for EGFP::DYS and the late muscle cell marker sarcomeric α-actinin. Nuclei were labelled with the DNA dye DAPI.

or AdV.DYS.mCherry^{TS}. As controls, DMD iPSCs were exposed to each vector separately. Transduction levels of ~80% were measured through mCherry-directed flow cytometry at four days post-transduction (**Figure 5A** top graph). Subsequent sub-culturing followed by mCherry-directed flow cytometry at 19 days post-transduction revealed a CRISPR-Cas9-dependent increase in the frequencies of stably expressing DMD iPSCs (**Figure 5A** bottom graph). These frequencies (~0.1%) were substantially lower than those detected in human myoblasts exposed to the same vector combinations (**Figure 2C,** bottom graphs), which in turn correlated with the detection of lower amounts of targeted DSB-derived indels in DMD iPSCs (**Figure S10**). Moreover, differently from the transduction experiments in HeLa cells and

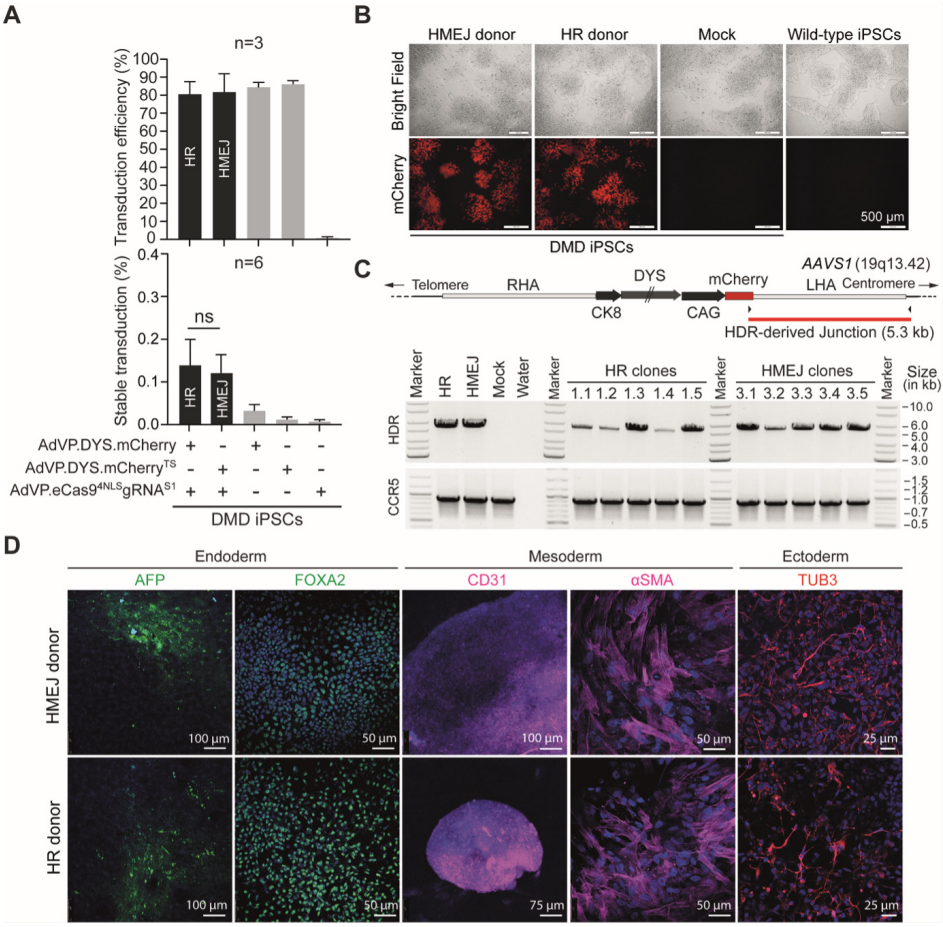


Figure 5. Testing AdVP-based HR and HMEJ genome editing strategies in iPSCs. (**A**) Testing AdVP donors encoding full-length dystrophin in DMD patient-derived iPSCs. Transduction efficiencies (top graph) and stable transduction levels (bottom graph) in iPSCs with a deletion of DMD exons 45 through 50 (DMD iPSCs), were determined by mCherry-directed flow cytometry at 4 and 19 days, respectively, after transduction with the indicated AdVPs. Bars and error bars correspond to mean ± SEM, respectively. Student's t-test showed non-significant (ns) differences between the indicated datasets (n=6 biological replicates; P>0.05). (**B**) Stably transduced DMD iPSCs. Representative micrographs of mCherry-sorted DMD iPSCs initially transduced with AdVP.eCas9^{4NLS}gRNA^{S1} at 5×10^3 GC cell^{-1} together with AdVP.DYS.mCherry (HR donor) or with AdVP.DYS.mCherry^{TS} (HMEJ donor) at 102 TU cell^{-1} each. Controls consisted of wild-type iPSCs and parental mock-transduced DMD iPSCs. (**C**) Establishing targeted chromosomal integrations. Junction PCR analysis on DMD iPSC populations genetically modified through AdVP-based HR or HMEJ genome editing strategies and randomly isolated iPSC clone derivatives. Amplicons diagnostic for centromere-sided transgenic-*AAVS1* junctions are depicted. *CCR5* served as an internal control template. Marker, GeneRuler DNA Ladder Mix. (**D**) Characterization of AdVP-edited DMD iPSCs. The pluripotency of genome-edited DMD iPSCs was ascertained by spontaneous differentiation and immunofluorescence detection of markers covering the three embryonic germ layers. Nuclei were identified by DAPI staining.

human myoblasts (**Figure 2C**), in DMD iPSCs, HR and HMEJ donors performed comparably (**Figure 5A** bottom graph). Notwithstanding, DMD iPSCs genetically modified through the AdVP-based HR and HMEJ strategies (**Figure 5B**), were confirmed to have undergone homology-directed chromosomal insertion of the exogenous DNA, as demonstrated by

junction PCR analysis on bulk populations and randomly isolated cell clones (**Figure 5C** and **S11**). Moreover, the resulting genome-edited iPSCs remained pluripotent as shown by their ability to differentiate along the three embryonic germ layers (**Figure 5D**). Finally, further supporting their self-renewal and multilineage potential, genome-edited iPSCs could differentiate into cardiomyocyte-like cells, as established by the detection of the striated and cardiac muscle markers sarcomeric α-actinin and cardiac troponin I, respectively, as well as by the acquisition of a spontaneously beating phenotype (**Figure S12** and **Supplementary Data File 1**).

TP53 downregulation facilitates HMEJ-based genome editing in iPSCs

It is well-known that stem cells are particularly sensitive to DSBs and that inhibiting the p53-dependent DNA damage response can increase the frequencies of stem cells edited through CRISPR-Cas9-induced HR [66, 67]. However, the impact of p53 on HMEJ-mediated genome editing strategies, in which cells are exposed to chromosomal and episomal DSBs, requires investigation. To this end, we started by generating independent iPSC lines (i.e., DMD iPSCs and wild-type iPSCs) containing low and normal amounts of p53. This p53 modulation was accomplished via lentiviral vector expression of short-hairpin RNAs targeting human *TP53* (shp53) and bacterial *LacZ* (shLacZ) transcripts, respectively. A robust shp53-dependent *TP53* knockdown, concomitant with significant downregulation of three canonical p53-responsive genes (i.e., *p21*, *FAS* and *PUMA*), was established through RT-qPCR analyses (**Figure 6A**). This strict shp53-mediated gene silencing was corroborated at the protein level by western blot analysis (**Figure S13**).

Next, the various iPSC lines were co-transduced with AdVP.eCas9^{4NLS}gRNA^{S1} and AdV.DYS.mCherry or with AdVP.eCas9^{4NLS}gRNA^{S1} and AdV.DYS.mCherry^{TS} and, once again, controls consisted of cells transduced with each AdVP individually. Transductions with AdVP donors resulted in transgene expression in the vast majority of DMD iPSCs and wild-type iPSCs as monitored by direct fluorescence microscopy (**Figure 6B**) and quantified by flow cytometry at 3 days post-transduction (**Figure 6C** top graphs). Moreover, iPSCs with regular and reduced amounts of p53 were transduced similarly well (**Figure 6C** top graphs). After over 2.5 weeks of sub-culturing, mCherry-directed flow cytometry revealed that the CRISPR-Cas9-dependent increase in stable transduction levels was, on a per experimental setting basis, roughly one order of magnitude higher in wild-type than in DMD iPSCs (**Figure 6C** bottom graphs). In itself, this result underscores the notion that the susceptibility of different iPSC lines to genome editing interventions can vary greatly. Interestingly, p53 knockdown led to significant and non-significant enhancing effects on the performance of the HMEJ- and HR-based genome editing strategies, respectively. Together, these data indicate that the observed blunted performance of HMEJ-based genome editing in iPSCs can be rescued to a significant degree via interfering with p53 function (**Figure 6C** bottom graphs).

Finally, we performed COBRA-FISH molecular karyotyping to probe the genetic stability of iPSCs with regular or reduced p53 levels and subjected to HR- or HMEJ-based genome

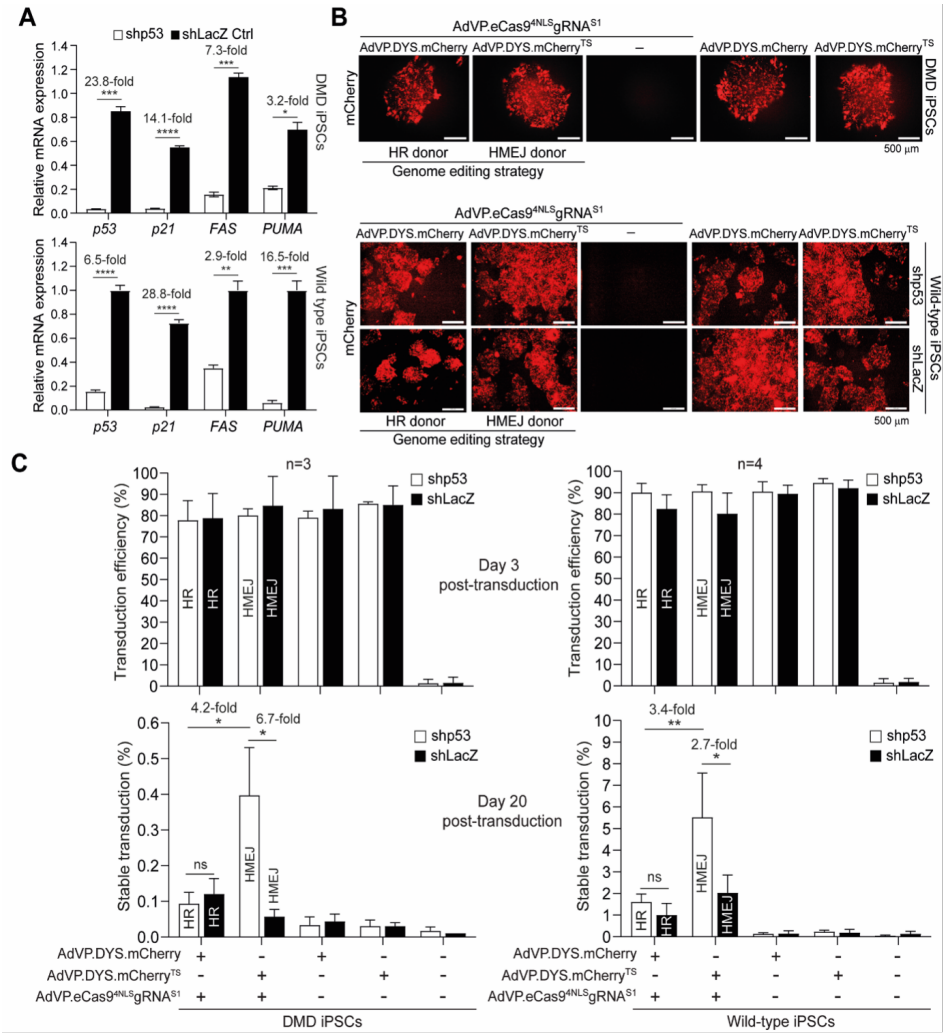


Figure 6. Effect of p53 knockdown on AdVP-based HR and HMEJ genome editing strategies in iPSCs. (**A**) Functional p53 knockdown in human iPSCs. Quantitative RT-qPCR analysis of the indicated p53-responsive genes in dystrophin-defective and wild-type iPSCs stably expressing shRNAs targeting p53 (shp53) or control LacZ transcripts (shLacZ). Data are plotted as mean ± SEM of three technical replicates. Significance between the indicated datasets was calculated with Student's t-tests; ****P<0.0001, ***P<0.001, **P<0.01, *P<0.05 (**B**) Probing AdVP transduction efficiencies in normal and p53 knockdown iPSCs. Direct fluorescent microscopy analysis on the indicated iPSCs at 2 days after transduction with AdVP.eCas9^{4NLS}gRNA^{S1} at 5×10^3 GC cell^{-1} together with AdVP.DYS.mCherry (HR donor) or with AdVP.DYS.mCherry^{TS} (HMEJ donor) at 100 TU cell^{-1} each. Controls consisted of cells individually transduced with AdVP donors or AdVP.eCas9^{4NLS}gRNA^{S1}. (**C**) Testing AdVP-based HR and HMEJ genome editing strategies in normal and p53 knockdown iPSCs. Transduction efficiencies (top graph) and stable transduction levels (bottom graph) in wild-type and DMD iPSCs expressing shp53 or control shLacZ were assessed by mCherry-directed flow cytometry at 3 and 20 days, respectively. AdVP.eCas9^{4NLS}gRNA^{S1} and AdVP donors were applied at 5×10^3 GC cell^{-1} and 10^2 TU cell^{-1}, respectively. Data are presented as mean ± SEM of at least three biological replicates. Significance between the indicated datasets was calculated with Student's t-tests. **P<0.01, *P<0.05. P>0.05 was considered non-significant (ns).

editing upon AdVP transduction. Regardless of the iPSC population tested, at the level of COBRA-FISH karyotyping resolution, neither numerical alterations (monosomy or trisomy) nor structural alterations (i.e., translocations, insertions, or deletions) were detected and a prevalence of cells with 2N was observed (**Figure S14A**). DNA content analysis of actively cycling iPSC populations using propidium iodide-directed flow cytometry confirmed the prevalence of 2N followed by 4N cell fractions (**Figure S14B**). Besides DNA replication, the acquisition of a complete 4N ploidy number, or higher, can occur through endoreduplication whereby an extra round or more of DNA synthesis is not followed by cytokinesis yielding endopolyploidy cells. Nonetheless, clearly, future p53 inhibiting agents selected for fostering HMEJ-based genome editing should act in a strictly transient fashion due to the notorious role of *TP53* as a tumor-suppressor gene supporting genomic stability.

Characterization of large-scale genomic edits in human myoblasts enabled with AdVPs

Junction PCR analysis established homology-directed chromosomal insertion of exogenous DNA through AdVP-based HR and HMEJ genome editing in HeLa cells (**Figure 3A**), human myoblasts (**Figure 3B**) and iPSCs (**Figure 5C** and **Figure S11**). Additionally, next to targeted chromosomal insertions, interphase FISH revealed the presence of off-target insertions in wild-type and DMD.B myoblasts (**Figure S15**). To complement these data, we next applied molecular combing to characterize genome editing events underlying stable expression of full-length, hence fully functional, dystrophin in human muscle cells. AdVP delivery of CRISPR-Cas9 nucleases and donor DNA constructs designed to favor HR and HMEJ resulted in 8.0±1.4% and 34.1±0.9% of EGFP::DYS-positive DMD.A myoblasts, respectively (**Figure 2C** bottom graphs). These unselected cell populations served as source material for the molecular combing analysis with mock-transduced cell populations acting as controls. Molecular combing analysis consists of fluorescence-based detection of probe hybridization to recombinant and human sequences in stretched single chromosome fibers (**Figure 7A** and **Figure S1**). The probe coverage measurements in chromosome fibers from mock-transduced cultures (**Figure S1B**) and in transgene-negative chromosome fibers from AdVP-transduced cultures (**Figure 7B** left panels), generally coincided with the DNA lengths expected for unmodified *AAVS1* alleles. Importantly, the same measurements in transgene-positive fibers mostly corresponded to precisely targeted *AAVS1* alleles (**Figure 7B** right panels and **Figure S16**). Indeed, this analysis revealed that in the cell fractions genetically modified with HR and HMEJ donor sequences, CRISPR-Cas9-induced *AAVS1* integration occurred at frequencies of 71.5% and 93.5%, respectively (**Figure 7C**). Amongst the targeted HR and HMEJ donor DNA insertions, 62.3% and 73.3% were precise, corresponding to 87.1% and 78.4% of the total targeted events, respectively (**Figure 7C**). In addition to these precise single-copy DNA insertions, genome-edited cell populations also contained lower fractions of multiple-copy targeted DNA insertions (**Figure 7C** and **Figure S16**). Moreover, albeit at a low frequency, site-specific chromosomal insertions consistent with direct end-to-end ligation

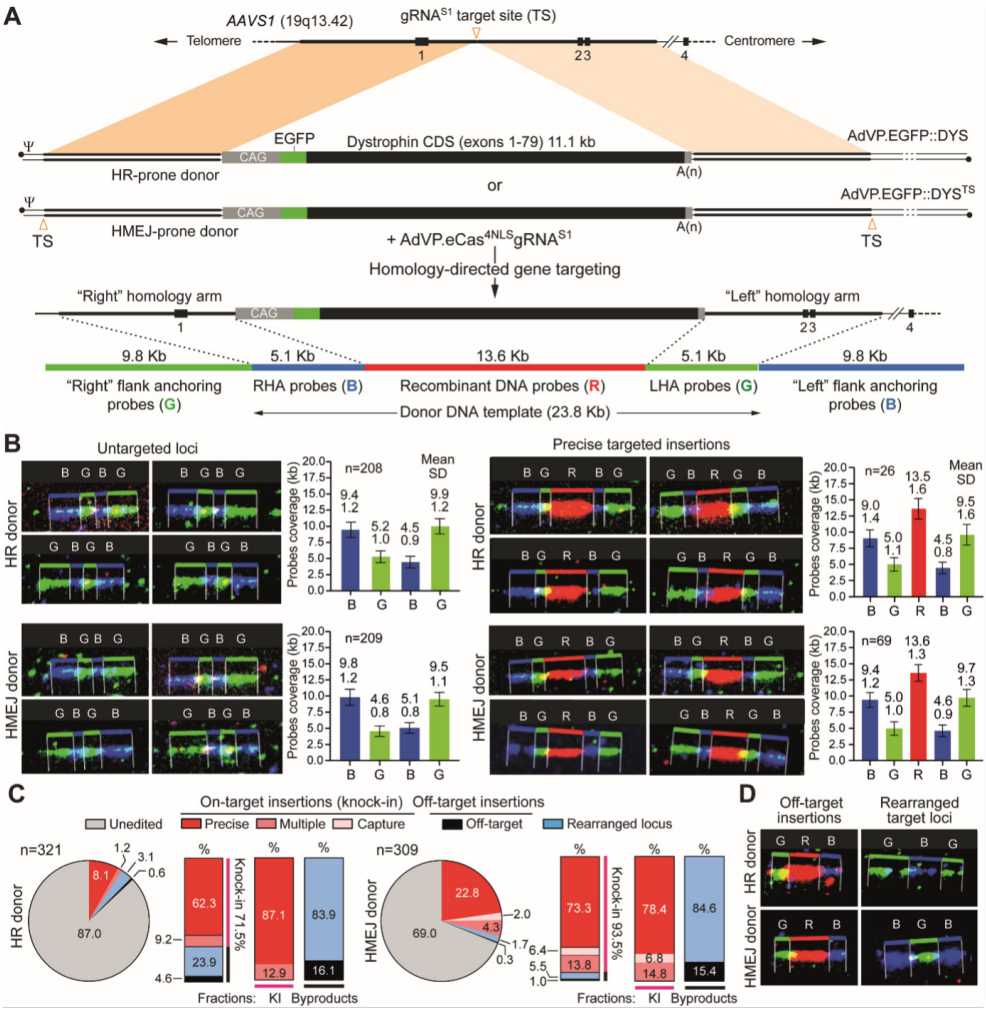


Figure 7. Establishing and characterizing AdVP-mediated targeted chromosomal insertion of large DNA at *AAVS1*. (**A**) Molecular combing set-up for measuring and mapping genome editing events. Schematic representation of homology-directed gene targeting substrates (i.e. donor DNA templates in AdVP.EGFP::DYS and AdVP.EGFP::DYS^{TS}) and precise DNA knock-in product triggered by gRNA^{S1}-directed DSB formation. The sizes and coverage of the probes used to stain recombinant DNA and genomic regions adjacent to the *AAVS1* target site are shown. G (green), "right" flank anchoring probes or "left" homology arm (LHA) probes; B (blue), "Left" flank anchoring probes or "right" homology arm (RHA) probes; R (red), recombinant DNA probes. (**B**) Validation of precisely targeted DNA insertions. Concordance between single-molecule measurements and theoretical values expected for precise genome editing events and unmodified target alleles in populations of DMD.A myoblasts genetically modified with HR (8% EGFP::DYS+ cells) and HMEJ (34% EGFP::DYS+ cells) donor templates (top and bottom panels, respectively). The former and latter unselected myoblast populations were generated through co-transduction with AdVP.eCas9^{4NLS}gRNA^{S1} (10^4 GC cell^{-1}) and AdVP.EGFP::DYS (4 TU cell^{-1}) or AdVP.eCas9^{4NLS}gRNA^{S1} (10^4 GC cell^{-1}) and AdVP.EGFP::DYS^{TS} (4 TU cell^{-1}), respectively. The mean ± S.D. values (upper and lower numerals, respectively) correspond to the sizes (in kb) of the four *AAVS1* sections and the single recombinant DNA section hybridizing to the respective probes. The graphs were assembled by measuring the hybridization signals derived from the indicated numbers of individual fluorescently labelled loci. (**C**) Cumulative quantification of genome editing events. Absolute and relative molecular combing signal distributions (pie charts and parts of whole bars, respectively) corresponding to homology-dependent insertions (precise), multiple copy insertions and homology-independent insertions (capture) detected in unselected DMD.A myoblast populations stably transduced after delivery of AdVP.eCas9^{4NLS}gRNA^{S1} together with AdVP.

◂ EGFP::DYS (HR donor) or AdVP.EGFP::DYS^{TS} (HMEJ donor). (**D**) Detection of genome editing byproducts. Single-molecule detection of off-target DNA insertion events and rearranged target alleles in DMD.A myoblasts genetically modified with HR or HMEJ donor DNA delivered by AdVPs.

of excised HMEJ donor DNA to *AAVS1*, were detected in human myoblasts initially co-transduced with AdVP.eCas9^{4NLS}gRNA^{S1} and AdVP.EGFP::DYS^{TS} (i.e. homology-independent recombinant DNA insertions) (**Figure 7C** right panel and **Figure S16**). As expected, these homology-independent genome editing events, presumably resulting from NHEJ-mediated "capture" of donor DNA at site-specific *AAVS1* breaks, were not detected in human myoblasts initially co-transduced with AdVP.eCas9^{4NLS}gRNA^{S1} and AdVP.EGFP::DYS (**Figure 7C** left panel). Importantly, regardless of the targeted genome-editing precision, all probe coverage measurements were consistent with the integration of complete dystrophin expression units (**Figure 7C**). Finally, off-target donor DNA insertions and *AAVS1* DNA rearrangements, possibly caused by unbalanced translocation elicited by CRISPR-Cas9-induced DSBs, were also observed in myoblast populations subjected to both genome editing strategies (**Figure 7D** and **Figure S16**). In conclusion, molecular combing analysis confirmed that genetic modification through AdVP-based HR and HMEJ genome editing strategies mostly involves exogenous DNA integration at *AAVS1* with most of these events corresponding to precise chromosomal DNA insertions.

Discussion

In this study, we report that AdVP delivery of CRISPR-Cas9 nucleases together with homology-directed repair templates tailored for HR or HMEJ achieves efficient and targeted insertion of large DNA cargoes (up to 14.8 kb) into *AAVS1*, a prototypic safe-harbour locus [56-58]. Therapeutic gene knock-in into safe harbour loci is a flexible genome editing concept in that it offers the possibility for correcting recessive disease phenotypes independently of the underlying mutation(s); and the associated tools might, in principle, be directed to other conditions once validated in a specific setting. The added versatility of this approach stems from engineering AdVPs endowed with CAR-independent fibers, whose engagement with an alternative receptor(s) ensures efficient transduction of diverse human cell types [23, 24]. In this instance, targeted chromosomal integration of full-length dystrophin expression units in iPSCs and CAR-negative myoblasts was achieved using CD46-targeting AdVPs. This data raises the prospect of predictable and mutation-independent genetic complementation of patient-derived stem or progenitor cells with myogenic capability [28, 29].

In the presence of CRISPR-Cas9 nucleases, HMEJ templates in AdVP genomes yielded higher frequencies of genome-edited cells than their HR counterparts in human myoblasts. Indeed, single-molecule visualization of exogenous and endogenous sequences by molecular combing revealed that chromosomal DNA insertions occurred mostly in a targeted and precise fashion. Undesirable genome-editing events were nonetheless detected in the form of multiple or imprecise targeted insertions, off-target insertions, and on-target rearrangements. Hence, future research should be directed at identifying the parameters that minimize unwanted

outcomes while further fostering precise gene targeting. A possibility concerns extending the homology arms in donor constructs beyond the current 10.2 kb of total target-donor DNA homology. Related to this, AdVP transduction experiments entailing the sole delivery of HR donors and positive/negative cell-selection regimens in iPSCs showed an increased *CFTR* mutation correction frequency via extending total target-donor homologies from 9.6 kb to 21.4 kb [18]. In particular, homologies spanning 9.6 kb and 21.4 kb yielded 0.7×10^{-5} and 2.4×10^{-5} G418-resistant iPSCs (i.e., 3.4-fold increase), of which 75% and 100%, respectively, were shown to be correctly targeted after ganciclovir selection against random insertions [18]. Of notice, recombinant DNA spanning 21-23 kb of target-site homologous sequences plus a full-length dystrophin expression unit "fit" within AdVP capsids. The dual AdVP system investigated here should further allow identifying optimal dosages of each genome editing component by achieving efficient and segregated delivery of CRISPR-Cas9 nucleases and donor templates into different human cell types.

Adult stem or progenitor cells and reprogrammed iPSCs are particularly appealing targets for AdVP-assisted genome editing owing to the relevance of these cell types in the development of disease and synthetic biology models or candidate autologous cell therapies [63-65], such as those directed at striated muscle disorders [28, 29, 64]. In this regard, earlier data showed that AdVP transduction of HR donors alone and together with *HBB*-specific TALENs yielded genome editing frequencies of 1.5×10^{-5} and 1.4×10^{-4}, respectively, as determined by counting G418-resistant iPSC clones from a sickle cell disease patient (i.e., 11-fold increment) [68]. More recently, the enhancing effect of site-specific DSBs on AdVP-assisted iPSC genome editing was extended to the use of CRISPR-Cas9 nucleases. In particular, AdVP delivery of HR donors alone and together with *CFTR*-specific CRISPR-Cas9 complexes led to genome editing frequencies of 1.5×10^{-5} and 1.8×10^{-3}, respectively, as assessed by counting puromycin-resistant iPSC clones from a cystic fibrosis patient (i.e., 117-fold increment) [69].

In the present work, AdVP delivery of *AAVS1*-specific CRISPR-Cas9 nucleases and tailored HR templates yielded ~0.1% and ~1.0% of genome-edited DMD and wild-type iPSCs, respectively, as determined by reporter-directed flow cytometry and junction PCR assays, with the latter assay confirming HDR-derived gene knock-ins in randomly isolated iPSC clones. Crucially, these high frequencies of accurate genome-editing events were obtained using neither gene trapping nor positive/negative marker genes often necessary for selecting correctly targeted cell fractions [24].

DSBs, including those made by CRISPR nucleases, readily trigger p53-dependent cell cycle arrest and apoptosis in stem cells, which greatly hinders the recovery of cells precisely edited through HR [66, 67]. This outcome, shown to be aggravated in the presence of AAV donor constructs [67], can be alleviated through p53 inhibition [66, 67]. We found that after AdVP delivery, the normally higher CRISPR-Cas9-dependent genome editing levels achieved by engaging HMEJ over HR donors in muscle progenitor cells and HeLa cells are instead cancelled in iPSCs. Nevertheless, we report that the performance of iPSC genome editing

involving HMEJ donor templates is rescued via p53 knockdown. These data suggest that gene knock-in strategies using such "double-cut" donor designs might profit from transient p53 inhibition, especially so in highly DSB-sensitive cell types. Indeed, notwithstanding the risks associated with interfering with a key tumor suppressor protein, transient inhibition of p53 function at the post-transcriptional or post-translational levels is starting to be explored in stem and progenitor cells for enhancing HR-based gene knock-ins [67, 70].

In conclusion, using DMD as a target disease model, we report that AdVPs are a robust platform for delivering and installing large genetic payloads in human cells and for investigating the performance of different targeted gene insertion approaches irrespectively of genome-editing tools and transgene sizes.

Material and methods

Cells

Human cervix carcinoma HeLa cells (American Type Culture Collection) were cultured in Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific; Cat. No.: 41966-029) supplemented with 5% (v/v) fetal bovine serum ultra-low endotoxin (FBS; Biowest; Cat. No.: S1860-500). The origins and culture conditions for the human wild-type myoblasts [31] as well as for the *DMD*-defective myoblasts, herein referred to as myoblasts DMD.A [32] and DMD.B [31] have been previously described [33]. In brief, the cells were grown in Skeletal Muscle Cell Growth Medium (Ready-to-use; PromoCell; Cat. No.: C-23060) supplemented with 20% FBS, 1× Glutamax (Thermo Fisher Scientific; Cat. No.: 35050-061) and 100 U ml^{-1} penicillin/streptomycin (Thermo Fisher Scientific; Cat. No.: 15140-122). The *DMD*-defective human induced pluripotent stem cells (iPSCs) used in this work, CENSOi001-B (herein referred to as DMD iPSCs), were purchased from the European Bank for induced pluripotent Stem Cells (EBiSC). The generation and characterization of the wild-type iPSCs LUMC0020iCTRL [34, 35] and LUMC0072iCTRL01 [36], were detailed elsewhere. The iPSCs were cultured in mTeSR medium (STEMCELL Technologies; Cat. No.: 85850) or in feeder-free Essential 8 Medium (E8; ThermoFisher Scientific; Cat. No.: A1517001) both supplemented with 25 U ml^{-1} penicillin and 25 μg ml^{-1} of streptomycin. The DMD iPSCs were cultured in plates coated with Matrigel (Corning Matrigel hESC-Qualified Matrix; Corning; Cat. No: 354277) when cultured with mTeSR medium or in plates coated with Vitronectin Recombinant Human Protein (VTN-N; ThermoFisher Scientific; Cat. No.: A14700) when cultured with E8 medium according to the manufacturer guidelines. When ready for sub-culturing, to let cell-cell dissociation occur, the iPSCs were washed with phosphate-buffered saline (PBS; pH 7.4) solution and then incubated with 0.5 mM ethylenediaminetetraacetic acid (EDTA; Invitrogen Cat. No.: 15575020) in PBS at 37°C for 5 min. After the removal of the EDTA solution, the cells were seeded in new wells with the proper medium supplemented with a 1:200 dilution of the ROCK inhibitor-containing solution RevitaCell (ThermoFisher Scientific; Cat. No.: A2644501). The PEC3.30 AdVP packaging cells [33] were kept in high-

glucose DMEM supplemented with 10% FBS, 10 mM $MgCl_2$ and 0.4 μg ml^{-1} puromycin (Thermo Fisher Scientific; Cat. No.: A11138-03). Finally, HEK293T cells were maintained in DMEM supplemented with 10% FBS and 100 U ml^{-1} penicillin/streptomycin. The cells used in this study were mycoplasma-free and were kept at 37°C in humidified-air atmospheres with 10% CO_2 (HeLa cells, HEK293T) or 5% CO_2 (Myoblasts and iPSCs) and at 39°C in humidified-air atmospheres with 10% CO_2 (PEC3.30 cells).

Production and characterization of adenoviral vectors

The constructs AO75_pHC.Ad.EGFP::DYS, AQ77_pHC.Ad.EGFP::DYS^{TS}, BE14_pHC.Ad.DYS.mCherry, BE15_pHC.Ad.DYS.$mCherry^{TS}$ and AX70_pHC.Ad.$eCas9^{4NLS}gRNA^{S1}$ were the molecular clones used for the production of the fiber-modified AdVPs AdVP.EGFP::DYS, AdVP.EGFP::DYS^{TS}, AdVP.DYS.mCherry, AdVP.DYS.$mCherry^{TS}$ and AdVP.$eCas9^{4NLS}gRNA^{S1}$ respectively. The annotated maps and nucleotide sequences of the AdVP genomes are available in the **Supplementary Information**. The construct encoding the human full-length dystrophin fused to EGFP (EGFP::DYS) and the construct containing the synthetic 436-bp striated muscle-specific *CK8* promoter were detailed elsewhere [37, 38]. The protocols used in the generation and purification of the resulting fiber-modified AdVP stocks have been described in detail before [33]. In brief, the bacteriophage P1 Cre recombinase- and adenovirus type 5 E1-expressing PEC3.30 producer cells, were seeded at a density of 1.4×10^6 cells per well of 6-well plates (Greiner Bio-One) one day prior to transfection in medium deprived of puromycin. Subsequently, 6.25 μg of MssI-linearized AO75_pHC.Ad.EGFP::DYS or AQ77_pHC.Ad.EGFP::DYS^{TS} were diluted in a total volume of 200 μl of 150 mM NaCl to which were added 20.6 μl of a 1 mg ml^{-1} solution of 25-kDa linear polyethyleneimine (PEI; Polysciences). After vigorous mixing in a vortex for about 10 sec, the solutions were incubated for 25 min at room temperature (RT) to let DNA-PEI complexes form. Finally, the DNA-PEI complexes were directly added to the medium of the producer cells. Six hours later, the transfection medium was substituted by fresh medium and containing E1-deleted helper AdV vector AdV.SRα.LacZ.1.50 [39] at a multiplicity of infection (MOI) of 7.5 infections units (IU) per cell. The helper vector drives the expression of the proteins necessary for the replication and assembly of the AdV particles. In addition, the PEC3.30 cells express a thermosensitive version of the adenovirus DNA-binding protein (DBP) that further contributes to vector complementation once producer cells are placed at the permissive temperature of 34°C. Upon helper-triggered emergence of about 80-100% of cytopathic effect (CPE), producer cells were harvested and subjected to three cycles of freezing and thawing in liquid N_2 and 37°C water baths, respectively. Next, cellular debris was removed by centrifugation for 10 min at 2,000×g and supernatants containing the vector particles were collected. Assembled vector particles present in the clarified supernatants were subsequently amplified through four rounds of propagation in producer cells transduced with helper AdV.SRα.LacZ.1.50. The last round of propagation involved seventeen T175-cm^2 culture flasks. The resulting AdV particles were then purified by sequential block and continuous CsCl buoyant density

ultracentrifugation steps and were de-salted by ultrafiltration through Amicon Ultra-15 100K MWCO filters (MerckMillipore; Cat. No: UFC910024). The production of AdVP.DYS.mCherry, AdVP.DYS.mCherry^{TS} and AdVP.eCas9^{4NLS}gRNA^{S1} involved essentially the same procedure applied for the production of AdVP.EGFP::DYS, AdVP.EGFP::DYS^{TS}, except that 1.6×10^6 PEC3.30 cells per well were seeded for the initial transfection step and, six hours later, the transfection media were substituted by fresh media containing helper AdV.SRα.LacZ.1.50 at an MOI of 30 IU cell^{-1}. The functional titers of the purified AdVP stocks AdVP.EGFP::DYS, AdVP.EGFP::DYS^{TS}, AdVP.DYS.mCherry, and AdVP.DYS.mCherry^{TS} were calculated by transducing HeLa cells with a range of vector stock dilutions. Three days post-transduction the percentages of reporter-positive cells were evaluated through flow cytometry. The resulting titers were 5.23×10^8 HeLa-transducing units (TU) ml^{-1}, 3.92×10^8 TU ml^{-1}, 7.38×10^{10} TU ml^{-1} and 7.42×10^{10} TU ml^{-1} for AdVP.EGFP::DYS, AdVP.EGFP::DYS^{TS}, AdVP.DYS.mCherry, and AdVP.DYS.mCherry^{TS}, respectively. The titer of AdVP.eCas9^{4NLS}gRNA^{S1} was assessed by using the Quant-iT™ PicoGreen™ dsDNA Assay Kit reagents (Thermo Fisher Scientific; Cat. No.: P11496A) as detailed elsewhere [39]. The resulting titer of AdV.Cas9.gRNA^{S1} was 2.82×10^{12} genome copies (GC) ml^{-1}. In addition, the AdVP physical titers were also determined through qPCR assays. Firstly, the vector DNA was isolated from purified AdVP stocks by using the DNeasy Blood & Tissue kit (QIAGEN; Cat. No.: 69506) and diluted 1:100. Next, three 3-fold serial dilutions of the extracted vector genomes were employed for qPCR using the iQ™ SYBR® Green Supermix (Bio-Rad; Cat. No.: L010171C) and the primers targeting the AdVPs packaging signal (ψ) listed in **Table S1**. As standard curve 8 serial 10-fold dilutions of a linearized parental plasmid stock containing 1×10^7 GC µl^{-1} was used as qPCR template. The qPCR primers, cycling conditions and components are specified in **Tables S1** and **S2**. Data analysis was performed by using the Bio-Rad CFX Manager 3.1 software and the titers were calculated based on the Ct values of standard curve and viral vector dilutions. The physical and functional AdVP titers are summarized in **Table S3**. The structural integrity of vector genomes packaged in purified adenoviral capsids was carried out essentially as described previously [33]. In brief, 80-100 µl of purified AdVPs were treated with 8 µl of 10 mg ml^{-1} DNaseI (Sigma-Aldrich; Cat. No.: 10104159001) at 37 °C for 30 min. Next, the DNase I was inactivated by adding 2.4 µl of 0.5 M EDTA solution (pH 8.0), 6 µl of 10% (w/v) sodium dodecyl sulphate (SDS) and 1.5 µl of 20 mg ml^{-1} proteinase K (Thermo Fisher Scientific; Cat. No.: EO0491). The resulting mixtures were then incubated at 55 °C for 1 h and vector DNA isolation was completed by using the QIAEX II Gel Extraction Kit (QIAGEN; Cat. No.: 20021) following the manufacturer's instructions. Finally, the isolated vector genomes were subjected to restriction enzyme fragment analysis (RFLA) by using the Gel-Doc XR+ system (Bio-Rad) and the Image Lab 6.0.1 software (Bio-Rad). Parental plasmids, digested with the same restriction enzymes applied to vector genomes, served as molecular weight references. The *in-silico* restriction patterns corresponding to intact plasmid and vector DNA were made with the aid of SnapGene (version 5.2.4) software.

Production of lentiviral vectors

The oligonucleotide pairs used for assembling lentiviral vector transfer plasmids pLV.Neo.shp53 and pLV.Neo.shLacZ encoding shRNAs shp53 and shLacZ, respectively, are listed in **Table S4**. After annealing, the oligonucleotide pairs were inserted through cohesive end ligation into the BbsI- and XhoI-treated expression cassette from a previously published doxycycline-regulated lentiviral vector system harbouring a *TetO*-modified human *H1* promoter [40]. The corresponding lentiviral vectors LV.Neo.shp53 and LV.Neo.shLacZ were generated according to previously detailed protocols [41, 42]. In brief, one day prior to transfection HEK293T cells were seeded in 175-cm^2 culture flasks (Greiner Bio-One). Next, 30 µg of DNA composed of lentiviral vector shuttle, packaging, and VSV-G-pseudotyping plasmids at a ratio of 2:1:1 (size-normalized for molecule copy number) and 90 µl of 1 mg ml^{-1} PEI solution (25 kDa PEI; Polysciences) were diluted in 1 ml of a 150 mM NaCl. Upon vigorous mixing in a vortex for about 10 sec and incubation for 25 min at room temperature (RT), the DNA-PEI complexes were directly added to the medium of the producer cells. The packaging and pseudotyping constructs used were psPAX2 (Addgene #12260) and pLP/VSVG (Invitrogen), respectively. After 24 hours the transfection medium was replaced by fresh DMEM supplemented with 5% FBS. At 3 days post-transfection, the producer-cell conditioned medium was harvested, and the cellular debris was removed by centrifugation and filtration through 0.45-µm pore-sized HT Tuffryn membrane filters (Pall Life Sciences; Cat. No.: PN4184). The lentiviral vector particle titers in the clarified supernatants were assessed by employing the RETROTEK HIV-1 p24 antigen ELISA kit (ZeptoMetrix; Cat. No.: 0801111). The resulting titers for LV.Neo.shp53 and LV.Neo.shLacZ were 590 ng $p24^{gag}$ ml^{-1} and 660 ng $p24^{gag}$ ml^{-1}, respectively.

Generation of p53 knockdown iPSCs

The generation of DMD iPSCs and wild-type iPSCs conditionally expressing a short hairpin RNA (shRNA) controlling *p53* downregulation, shp53, was initiated by lentiviral vector transduction with LV.Neo.shp53. As control, parallel cell cultures were transduced with LV.Neo.shLacZ expressing a shRNA targeting *LacZ*, shLacZ. These lentiviral vectors encode a neomycin resistant (*Neo*) gene and the shRNAs under the control of a doxycycline-inducible promoter. In brief, cells were seeded in regular growth medium at a density of 6×10^4 cells per well of 24-well plates. The following day DMD iPSCs and wild-type iPSCs (LUMC0020iCTRL) were exposed to medium containing each lentiviral vector at an MOI of 1 and 0.1 TU ml^{-1}, respectively. After three days, the cells were transferred to a new plate containing regular growth medium and one day later, the medium was supplemented with 50 µg ml^{-1} of G418 sulfate (Millipore; Cat. No.: 345810). Parental mock-transduced cells served as negative controls during the drug selection procedure. The resulting stably transduced iPSCs were cultured in medium supplemented with 10 ng ml^{-1} Doxycycline (Sigma-Aldrich; Cat. No.: D9891) to activate *p53* downregulation as confirmed through western blot analysis.

To activate shRNA expression, doxycycline was added three days prior to the start of the transduction experiments and was kept throughout the duration of these experiments.

Quantitative polymerase chain reaction (qPCR)

Episomal donor DNA amounts were traced through *EGFP*-directed qPCR analysis. In brief, HeLa cells were transduced with AdVP.EGFP::DYS and AdVP.EGFP::DYS^{TS} alone, at an MOI of 8 TU ml^{-1}, or together with AdVP.eCas9^{4NLS}gRNA^{S1} at an MOI of 3000 GC cell^{-1}. Next, DNA was extracted at 3-, 12-, and 28-days post-transduction with the DNeasy Blood & Tissue kit (QIAGEN; Cat. No.: 69506). Extracted DNA from mock- and AdVP-transduced cells was subsequently subjected to qPCR using the iQ™ SYBR® Green Supermix (Bio-Rad; Cat. No.: L010171C) together with the primers targeting *EGFP* and *GAPDH* listed in **Table S1**. The cycling conditions and components of qPCR mixtures are specified in **Tables S1** and **S2**. The signal outputs were detected by using the CFX Connect Real-Time PCR Detection System (Bio-Rad) and the resulting data was analysed via the Bio-Rad CFX Manager 3.1 software.

Reverse transcriptase-qPCR (RT-qPCR)

The knockdown of *TP53* expression, resulting in downregulation of the p53 target genes *p21*, *PUMA* and *FAS*, was assessed by RT-qPCR analysis of cells expressing shp53 incubated for at least two days in the presence of 10 ng ml^{-1} Doxycycline (Sigma-Aldrich; Cat. No.: D9891). Cells expressing shLacZ instead of shp53 served as control. Total RNA from cell cultures was extracted by using the NucleoSpin RNA kit essentially according to the manufacturer's instructions (Macherey Nagel; Cat. No.: 740955). Next, reverse transcription was conducted with the SuperScript III Reverse Transcriptase Kit (Invitrogen; Cat. No.: 18080-044). In brief, 1000 ng of RNA was incubated with 1 µl of 100 ng µl^{-1} random primers and 1 µl of 10mM dNTPs Mix in a 13-µl reaction volume. The reaction mixtures were incubated for 5 min at 65°C followed by 2 min at 4°C. Next, to each tube was added 4 µl 5× First-Strand Buffer, 2 µl of 20 U ml^{-1} RiboLock RNase Inhibitor (Thermo Fisher Scientific; Cat. No.: EO0381), 1 µl 0.1 M DTT and 1 µl of SuperScript™ III RT (200 units µl^{-1}). Finally, the samples were incubated at 25°C for 5 min followed by incubation at 50°C for 1 h. Eventually, the reactions were terminated by heating at 70°C for 15 min. The resulting cDNA templates were then diluted 5-fold in nuclease-free water and 1-µl samples of diluted cDNA were subjected to qPCR by using the iQ™ SYBR® Green Supermix (Bio-Rad; Cat. No.: L010171C) and the primers listed in **Table S1**. *GAPDH* transcripts served as RT-qPCR targets for gene expression normalization. The signal outputs were detected by using the CFX Connect Real-Time PCR Detection System (Bio-Rad) and data analysis was performed using Bio-Rad CFX Manager 3.1 software. The qPCR mixture components and cycling conditions are specified in **Tables S1** and **S2**, respectively.

Transduction experiments

Transduction experiments in HeLa cells and human myoblasts were initiated by seeding the cells in wells of 24-well plates at a density of 4×10^4 and 5×10^4 cells per well, respectively.

The next day, the cells were incubated in 500 µl of medium containing AdVPs at different MOI. Transduction efficiencies were determined at 3 days post-transductions by reporter-directed flow cytometry and direct fluorescence microscopy analyses. Transduction experiments in iPSCs were performed as follows. RevitaCell was added at a dilution of 1:200 to the cells for 30 minutes prior to the transfer procedure to maximize cell survival. The cells were dissociated to obtain cell suspensions having 1×10^5 cells in 100 µl of medium supplemented with RevitaCell (1:200). The cell suspensions were subsequently added to a V-bottom 96-well plate (Greiner Bio-One) containing AdVPs in 100 µl of medium. After 1 hour of incubation at 37°C in a humidified-air atmosphere with 5% CO_2, the cells were transferred to a 12-well plate containing medium supplemented with RevitaCell (1:200). Two hours later, fresh medium supplemented with RevitaCell (1:200) was added. Transduction efficiencies were assessed at 3- or 4-days post-transduction by reporter-directed flow cytometry and direct fluorescence microscopy analyses.

Cell differentiation assays

The human myoblasts were transferred in a regular culture medium into wells pre-coated with a 0.1% (w/v) gelatin solution (Sigma-Aldrich; Cat. No.: G13393). After reaching full confluency, the muscle progenitor cells were exposed to myogenic differentiation medium consisting of phenol red-free DMEM (ThermoFisher Scientific; Cat. No.: 11880-028) supplemented with 100 µg ml^{-1} human holo-transferrin (Sigma-Aldrich; Cat. No.: T0665), 10 µg ml^{-1} human insulin (Sigma-Aldrich; Cat. No.: I9278) and 100 U ml^{-1} penicillin/streptomycin. Post-mitotic myotubes were detected 4-6 days later by immunofluorescence staining and western blot analyses with antibodies specific for the late muscle-specific markers sarcomeric α-actinin or myosin heavy-chain. The differentiation of iPSCs along the three embryonic germ layers was induced by using a spontaneous differentiation protocol as follows. The iPSCs were first treated with 0.5 mM EDTA in PBS for 1 min at 37°C and were subsequently gently dissociated into large cell clumps by scrapping. The resulting cell clumps were then cultured in suspension for 24 h on low-attachment plates at 37°C. Next, the iPSCs were seeded on glass coverslips coated with VTN-N in E8 medium containing RevitaCell (1:200) and, the following day, this medium was substituted by DMEM/F12 growth medium (Thermo Fisher Scientific; Cat. No.: 31331-028) containing 20% FBS. The DMEM/F12 medium was replenished every 2–3 days. After 3 weeks under differentiation conditions, the iPSCs were processed for confocal immunofluorescence microscopy for the detection of markers characteristic of the endoderm, mesoderm and ectoderm lineages. The markers tested were α-fetoprotein (AFP), forkhead box protein A2 (FOXA2), α-smooth muscle actin (α-SMA), endothelial cell adhesion molecule-1 (CD31), and tubulin β3 class III (TUBB3). The cardiomyogenic differentiation of iPSCs was carried out following the protocol for cardiac lineage differentiation based on stepwise supplementation of iPSC medium with specific small molecules as described elsewhere [43]. In brief, iPSCs kept in E8 medium supplemented with RevitaCell (1:200) were seeded in 12-well plates coated with Matrigel

(Corning; Cat. No.: 734-1440) or Geltrex (ThermoFisher; Cat. No.: A1413301) at different cell densities, ranging from 5×10^4 to 9×10^5 cells. At 24 hours after seeding the medium was replaced with a modified LI-BPEL (mBEL) medium [43] supplemented with 5 μM CHIR 99021 (Axon Medchem; Cat. No.: Axon1386). Two days later, the medium was replenished with mBEL medium supplemented with 5μM XAV 939 (Tocris; Cat. No.: 3748/10) and 0.25 μM IWP-L6 (AbMole; Cat. No.: M2781). Finally, after an additional 2-day period the medium was replenished with mBEL medium supplemented with Insulin-Transferrin-Selenium-Ethanolamine (ITS -X) (1:1000) (ThermoFisher; Cat. No.: 51500056). Next, cells were kept in culture under differentiation conditions for up to 30-35 days in the presence of mBEL medium supplemented with ITS-X (1:1000) that was replenished every 2-3 days. Areas of beating cardiomyocytes were detected from day 10 onwards. Finally, at the selected differentiation endpoints, the detection of the cell-specific markers cardiac troponin I and sarcomeric α-actinin was performed by immunofluorescence microscopy analyses.

Flow cytometry

Cell transduction frequencies were determined by reporter-directed flow cytometry using a BD LSR II flow cytometer (BD Biosciences). In brief, HeLa cells and human myoblasts were washed with PBS and, after trypsin treatment and centrifugation at 300×g for 5 min, they were collected in PBS containing 0.5% BSA and 2 mM EDTA (pH 8.0) (FACS buffer). iPSCs were washed with PBS and treated with 0.5 mM EDTA in PBS for 5 min at 37°C for gentle dissociation. To obtain single-cell suspensions, the cells were collected by centrifugation at 200×g for 2 min and were then resuspended in 200 μL of Accutase (STEMCELL Technologies; Cat. No.: 07920) and incubated for 10 minutes at 37°C. After centrifugation at 300×g for 5 min, iPSCs were also resuspended in FACS buffer. Mock-transduced cells served as a control to establish background fluorescence levels. Data were analysed with the aid of FlowJo 10.6.0 software (TreeStar).

Chromosomal DNA content analysis

The DNA content in iPSCs with regular or knocked-down p53 levels genetically modified via HR- or HMEJ-based genome editing strategies, was assessed as follows. After the addition of 200 μl of Accutase (STEMCELL Technologies; Cat. No.: 07920), the cells were incubated for 10 minutes at 37°C to obtain single-cell suspensions. Next, after centrifugation at 300×g for 5 min, the cells were resuspended in 70% ethanol and incubated at 4°C overnight. After one wash with PBS, the cells were resuspended in PBS containing 50 μg ml^{-1} RNAse A (ThermoFisher; Cat. No.: EN0531) and 20 μg ml^{-1} propidium iodide (Abcam; Cat. No.: ab14083). Finally, after an overnight incubation period at 4°C the four iPSC populations were washed twice with PBS, resuspended in FACS buffer, and analysed for ploidy number by propidium iodide-directed flow cytometry.

COBRA-FISH karyotyping of iPSCs

Combined binary ratio labelling multicolour FISH-based molecular karyotyping

(COBRA-FISH) was performed to determine the karyotype of iPSCs with regular or knocked-down p53 levels genetically modified via HR- or HMEJ-based genome editing strategies. The detailed COBRA-FISH protocol applied to these cells has been published [44]. In brief, metaphase suspensions were dropped on microscopy glass slides and airdried overnight. Slides with metaphase chromosomes were pre-treated at 37 °C for 10 min with 100 μg ml^{-1} RNase A (Roche; Cat. No.: 10154105103) in 2× saline-sodium citrate (SSC; Sigma-Aldrich; Cat. No.: S0902) and were then incubated with 0.005% pepsin (Sigma-Aldrich; Cat. No.: P0525000) in 0.1 M HCl for 5 min at 37 °C. After a 10-min fixation at room temperature with 1% (v/v) formaldehyde (Merck; Cat. No.: 1.03999.1000) in PBS (pH 7.4), the specimens were dehydrated by three 3-min incubations in ethanol at increasing concentrations, i.e., 70%, 90% and 100%. Next, the coverslips were air-dried and exposed to whole-chromosome painting probes fluorescently labelled with the dyes 7-diethylaminocoumarin-3-carboxylic acid (DEAC), Cy3, Cy5 and rhodamine green by using the Universal Linkage System (ULS) kit (Kreatech Biotechnology). After DNA denaturation at 80 °C for 75 sec, hybridizations were let to proceed for 3 days at 37 °C in a humidified chamber. Unbound probes were eliminated by first washing with 0.1% Tween-20 (Promega; Cat. No.: PRH5152) in 2× SSC, and then with 50% formamide (Merck; Cat. No.: 1.09684.1000) in 2× SSC pH 7.0 at 44 °C followed by incubation at 60 °C in 0.1× SSC. Each washing step was done twice for 5 min. After dehydration by exposure to the above-mentioned increasing concentrations of ethanol, the specimens were air-dried and embedded in Citifluor AF1/DAPI (400 ng ml^{-1}) solution (Aurion; Cat. No.: E17970). Stained chromosomes were visualised with a Leica DMRA fluorescence microscope (Leica, Wetzlar, Germany) and images were captured with the aid of a CoolSnap HQ2 camera (Photometrics, Tucson, USA). A minimum of 35 metaphase cells were analysed from each sample and reported in the International System for Human Cytogenomic Nomenclature (ISCN) format.

Cell sorting and clonal expansion

Sorting of genetically modified cells was done using a BD FACS Aria III flow cytometer (BD Biosciences), following the removal of episomal DNA through sub-culturing of transduced myoblasts and iPSCs. The retrieved reporter-positive cells were collected in a 1:1 mixture of regular medium and FBS. Single cell-derived iPSC colonies were isolated by seeding at low-density single-cell suspensions of mCherry-positive cells in wells of 6-well plates containing culture medium supplemented with Revitacell (1:200), 50 μM α-thioglycerol and 20 nM of cathocuprione disulphonate. The iPSC single-cell suspensions were obtained by incubation in Gentle Dissociation Buffer (StemCell; Cat.No.: 07174) for 10 min at 37°C and filtering through a Sterile Cell Strainer 70 μm nylon mesh filter. The growing single cell-derived colonies were subsequently monitored and selected by using direct fluorescence microscopy.

Characterization of genome-modifying events by long-range junction PCR

mCherry-positive iPSCs, wild-type myoblasts and DMD.B myoblasts generated after HR or HMEJ-based genome editing strategies were sorted at ~3 weeks post-transduction as whole populations or single cell-derived clones following the cell sorting and clonal expansion procedures described above. Genomic DNA from each sample was obtained with DNeasy Blood & Tissue kit (QIAGEN; Cat. No.: 69506) following the manufacturer's recommendations. Next, conventional and long-range junction PCR analyses were performed with GoTaq G2 DNA Polymerase (Promega; Cat. No.: M7801) and Platinum SuperFi II DNA Polymerase (ThermoFisher; Cat. No.:12361010), respectively. The PCR screening of the mCherry-positive cell populations and single cell-derived clones was performed using the PCR mixtures and cycling parameters indicated in **Table S5** and **S6**, respectively. Genomic DNA of HeLa cells exposed to AdVP.EGFP::DYS and AdVP.EGFP::DYS^{TS} alone, at an MOI of 8 TU ml^{-1}, or together with AdVP.eCas9^{4NLS}gRNA^{S1} at an MOI of 3×10^3 GC cell^{-1} was retrieved instead at 28-days post-transduction. These samples were analysed with the PCR mixtures and cycling parameters indicated in **Tables S7** and **S8**, respectively.

Target DNA cleaving assays

Targeted DSB formation in cells transduced with AdVP.eCas9^{4NLS}gRNA^{S1} was assessed by using assays based on the mismatch sensing T7EI enzyme and amplicon sequencing. To this end, genomic DNA samples from mock-transduced and vector-transduced cells were retrieved at 3 days post-transduction using the DNeasy Blood & Tissue kit (QIAGEN; Cat. No.: 69506) following the manufacturer's recommendations. The *AAVS1* target site-specific PCR amplifications were performed using Phusion High-Fidelity DNA Polymerase (ThermoFisher; Cat. No.: F-530). The primer sequences, PCR mixture compositions and cycling parameters are specified in **Tables S9** and **S10**, respectively. The resulting amplicons were denatured and reannealed by applying the program listed in **Table S11**. T7EI-based DNA cleaving assays were done as follows. First, 10 µl of each PCR mixture was incubated in 15-µl reactions consisting of 1× NEBuffer 2 (New England Biolabs; Cat. No.: B7002S) and 5 U of T7EI (New England Biolabs; Cat. No.: M0302). Next, after 17-min incubation at 37°C, the DNA samples were subjected to electrophoresis through 2% (w/v) agarose gels in 1× Tris–acetate–EDTA (TAE) buffer. The resulting ethidium bromide-stained DNA species were then detected by using a Molecular Imager Gel-DocTM XR+ system (Bio-Rad) and the proportions of T7EI-digested products were determined by densitometry using Image Lab 6.0.1 software (Bio-Rad). Target DNA cleaving activities were determined by uploading Sanger sequencing traces corresponding to the target site-specific PCR products into the Inference of CRISPR Edits (ICE) tool https://ice.synthego.com/#/ [45].

Next-generation sequencing for off-target DNA cleavage analysis

Wild-type myoblasts were transduced with various MOI of AdVP.eCas9^{4NLS}gRNA^{S1} and cultures of the same cell type were co-transduced in parallel with comparable MOI of the second-generation *E1*- and *E2A*-deleted adenoviral vectors AdV^{D2}P.Cas9.F^{50} and

AdV^{D2}U6.gRNA^{S1}.F^{50} (herein named AdV.Cas9 and AdV.gRNA^{S1}, respectively). AdV.Cas9 and AdV.gRNA^{S1} encode a wild-type Cas9 nuclease and the *AAVS1*-targeting gRNA gRNA^{S1}, respectively [46]. At three days post-transduction, genomic DNA was isolated with the DNeasy Blood & Tissue kit reagents and protocol. To assess the specificity profiles of Cas9:gRNA^{S1} versus eCas9^{4NLS}:gRNA^{S1} complexes, a previously described amplicon deep sequencing analyses pipeline was used [47, 48]. In brief, the *AAVS1* target site and two off-target sites of Cas9:gRNA^{S1} complexes experimentally validated by orthogonal high-throughput genome-wide translocation sequencing (i.e., *CPNE*5 and *BBOX1*) [34], were first amplified with primers containing adapter tag overhangs using Phusion High-Fidelity Polymerase (Thermo Fisher Scientific; Cat. No.: #F-530L). The primer sequences, PCR mixture compositions and cycling parameters are specified in **Tables S12** and **S13**, respectively. The resulting amplicons were subsequently purified with AMPure XP beads (Beckman Coulter; Cat. No.: A63881) and subjected to PCR barcoding using Illumina tag-specific primer pairs with unique sequence combinations for demultiplexing and sample identification. The cycling parameters, primer sequences and PCR mixtures used for the preparation of barcoded amplicons are indicated in **Tables S13**, **S14** and **S15**, respectively**.** Next, the samples were further purified with AMPure XP beads and the concentrations of barcoded amplicons were determined by using the Qubit dsDNA HS assay kit (Thermo Fisher Scientific; Cat. No.: Q32854) and a Qubit2.0 fluorometer. Finally, purified amplicons were pooled in equal molar ratios and then subjected to Illumina MiSeq deep sequencing for retrieving 50,000 paired-end reads. Finally, after demultiplexing and adapter trimming of the paired-end MiSeq raw reads (R1 and R2 fastq files) with Cutadapt version 2.10 [49], alignment of amplicon sequences to reference sequences was carried out by using the CRISPResso2 software [50]. The scripts applied in each CRISPResso2 analyses round are available as **Supplementary Information.**

Immunofluorescence microscopy

Cultures of undifferentiated myoblasts and differentiated myotubes were analysed through immunofluorescence analysis. Immunofluorescence analyses were also employed to detect the acquisition of differentiation markers (i.e., AFP, FOXA2, α-SMA, CD31 and TUBB3) in iPSCs-derived cells. Cells cultured on glass coverslips were fixed with 4% paraformaldehyde (PFA) for 10 min. Next, after three washes with PBS, the cells were permeabilized in 0.5% (v/v) Triton X-100 in TBS (50 mM Tris-HCl pH 7.5, 100mM NaCl) at RT for 5 min and were then washed three times for 10 min with 0.1% Triton X-100 in TBS. Subsequently, the cells were exposed for 2 h to a blocking Antibody Diluting Solution (Abdil) consisting of 0.1% Triton X-100, 2% bovine serum albumin and 0.1% sodium azide in TBS. The specimens were then incubated overnight at 4°C with the proper primary antibodies diluted in blocking solution (**Table S16**). After three 5-min washes with 0.1% Triton X-100 in TBS, the specimens were incubated with fluorochrome-conjugated secondary antibodies diluted in a blocking solution for 1 h in the dark at RT (**Table S16**). Finally, after three 10-min washes with 0.1% Triton X-100 in TBS, ProLong™ Gold Antifade Mounting reagent containing DAPI (ThermoFisher Scientific; Cat.

No.: P36931) was used for mounting the specimens. Immunofluorescence microscopy images were acquired by using an upright Leica SP8 confocal microscope equipped with Leica hybrid detectors HyD or a 3DHISTECH Pannoramic 250 digital slide scanner for the detection and quantification of eCas9^{4NLS} expressing cells. The acquisition of differentiation markers (i.e., cardiac troponin I cTnI and sarcomeric α-actinin) in iPSC-derived cardiomyocytes was also assessed through immunofluorescence microscopy. The iPSC-derived cardiomyocytes were first dissociated by incubation for 10 min at 37°C in 5× TrypLE Select (Thermo Fisher; Cat. No.: A1217701). The resulting cell suspensions were then seeded in wells of 96-well plates previously coated with Geltrex. After 4 to 6 days in culture medium, the cells were subjected to the same staining protocol described above except that (i) the cells were incubated overnight at 4°C in blocking solution, (ii) the cells were exposed to the appropriate primary antibody (**Table S16**) for 2 hours at RT and (iii) the nuclei were stained by incubation in Hoechst 33342 (Invitrogen; Cat. No.: H3570) diluted 1:1000 in PBS for 10 min at RT. Finally, microscopy images were analysed through the LAS X (Leica Microsystems), ImageJ (NIH, US National Institutes of Health) or CaseViewer (3DHISTECH) software packages whilst the cell image analysis software CellProfiler [51] at https://cellprofiler.org/#/, was employed for automated segmentation and quantification of eCas9^{4NLS}-positive nuclei.

Live-cell fluorescence microscopy

Reporter expression in HeLa cells, human myoblasts, iPSCs and iPSC-derived cardiomyocytes was monitored by direct fluorescence microscopy. mCherry- and EGFP-specific signals in HeLa cells, human myoblasts and iPSCs were detected by using an inverted DMi8 fluorescence microscope equipped with a DFC 450c camera. mCherry-specific signals in beating cardiomyocytes were recorded with a AF6000 LX system. The acquired images and videos were examined with the aid of LAS X (Leica Microsystems) and ImageJ software (NIH, US National Institutes of Health).

Western blotting

Unedited and vector-edited human myoblasts subjected to the differentiation conditions described previously were processed for western blot analysis as follows. After 4 to 6 days in differentiation medium, the myotube-containing cultures were lysed on ice for 30 min by incubation in 50 µl of RIPA buffer (ThermoFisher Scientific; Cat. No.: 89900) supplemented with a protease inhibitor cocktail (cOmplete Mini, Sigma-Aldrich; Cat. No.: 11836153001) and the resulting cell lysates were then passed through a 30-gauge syringe several times. Protein quantification was carried out by using the Pierce BCA Protein Assay Kit (ThermoFisher Scientific; Cat. No: 23225), following the manufacturer's instructions. Next, the indicated amounts of total protein were diluted in 4× sample buffer (Bio-Rad; Cat. No.: 161-0791) and 20× reducing agent (Bio-Rad; Cat. No.: 161-0792), were incubated at 95°C for 5 min. Protein samples and 15 µl of HiMark Prestained Protein Standard (ThermoFisher Scientific; Cat. No.: LC5699) were loaded in a 3–8% Criterion XT Tris-Acetate precast gel (Bio-Rad;

Cat. No.: 3450130). The polyacrylamide gel was then placed in a Criterion Cell containing XT Tricine running buffer (Bio-Rad; Cat. No.: 1610790) and run for 30 min at 75 V (0.07 A) and for 1.5 h at 150 V (0.12 A). Subsequently, the resolved proteins were transferred to polyvinylidene difluoride (PVDF) membranes with the aid of a Trans-Blot Turbo Midi PVDF pack (Bio-Rad; Cat. No.: 1704157) and a Trans-Blot Turbo system (Bio-Rad) according to the manufacturer's recommendations for high-molecular-weight proteins (2.5 A, 25 V, 10 min). The PVDF membranes were then blocked for 2 h at room temperature in 5% non-fat dry milk (Campina Elk; Cat. No.: 112349) dissolved in TBS with 0.1% (v/v) Tween 20 (TBST). Next, the membranes were incubated overnight at 4°C with primary antibodies (**Table S17**) diluted in 5% non-fat dry milk. After three 10-min washes in TBST, the membranes were incubated for 2 h at RT with the proper secondary antibodies (**Table S17**) conjugated to horseradish peroxidase (IgG-HRP) diluted in 5% non-fat dry milk. Proteins were detected by using horseradish peroxidase substrate Pierce ECL2 (ThermoFisher Scientific; Cat. No.: 80196) following the manufacturer's specifications. iPSCs, instead, were lysed in Laemmli buffer consisting of 8.0% glycerol, 3% SDS and 200 mM Tris-HCl (pH 6.8) and subsequently incubated for 5 minutes at 100°C. Protein concentrations in cell lysates were assessed by using a DC™ protein assay kit (Bio-Rad; Cat. No.: 5000111) according to the manufacturer's recommendations. Proteins were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Afterwards, the resolved proteins were transferred onto 45-µm PVDF membranes (Merck Millipore; Cat. No.: IPVH00010) that were subsequently blocked with 5% (w/v) non-fat dry milk dissolved in TBST at RT for 1 h. After the blocking step, the membranes were incubated overnight at 4°C with the proper primary antibody (**Table S17**) diluted in TBST supplemented with 5% BSA. Next, the membranes were washed with TBST thrice and probed at RT for 2 h with the proper secondary antibody (**Table S17**) diluted in TBST containing 1% (w/v) non-fat dry milk. Signal detection was performed with the Clarity™ Western ECL Substrate (Bio-Rad; Cat. No.: 1705060). All images were acquired using ChemiDoc Imaging System (Bio-Rad; Cat. No.: 17001402) and were analysed with the Image Lab 6.0.1 software (Bio-Rad).

Interphase FISH for integration mapping

Cells were grown on glass coverslips to 80-100% confluency and were fixed with 4% paraformaldehyde (PFA) for 10 min followed by three washes with PBS for 5 minutes each. FISH was performed as described before [52] with some modifications. In brief, glass coverslips were pre-treated at 37°C for 10 min with 100 µg ml^{-1} RNase A (Roche; Cat. No.: 10154105103) in 2× saline-sodium citrate (SSC; Sigma-Aldrich; Cat. No.: S0902) and were then incubated with 0.01% pepsin (Sigma-Aldrich; Cat. No.: P0525000) in 0.1 M HCl for 5 min at 37 °C. After a 10-min fixation at room temperature with 1% (v/v) formaldehyde (Merck; Cat. No.: 1.03999.1000) in PBS (pH 7.4), the specimens were dehydrated by three 3-min incubations in ethanol at increasing concentrations, i.e., 70%, 90% and 100% and airdried. For interphase FISH, 25 ng of Cy3-dUTP nick translation labelled BAC probe

(RP11-463M24) (Cy-dUTP Enzo Life Sciences; Cat. No: ENZ-42501) and 2.5 ng of Bio-11dUTP nick translation labelled transgene probe (Bio-11dUTP; Jena Biosciences; Cat. No.: NU-803-BIOX-S Nick translation) were used under a 12 mm round coverglass. The transgene probe sequences covering the mCherry and full-length dystrophin expression cassettes, were obtained by enzymatic digestion and gel extraction of BE14_pAdVP.DYS.mCherry DNA using established procedures. After DNA denaturation at 80 °C for 45 sec, hybridizations were done for 18 hours at 37 °C in a humidified chamber. Unbound probes were removed first by washing with 0.1% Tween-20 (Promega; Cat. No.: PRH5152) in 2× SSC, and then with 50% formamide (Merck; Cat. No.: 1.09684.1000) in 2× SSC (pH 7.0) at 44 °C followed by incubation at 60 °C in 0.1× SSC. Each washing step was done twice for 5 min. The biotin-labelled probe was detected by incubations in Streptavidin Alexa Fluor™ 488 Conjugate, diluted 1:250 (ThermoFisher; Cat. No.: S32354), followed by Biotinylated anti-Streptavidin diluted 1:100 (Jena Bioscienc; Cat. No.: BA-0500-5) and, finally, Streptavidin Alexa Fluor™ 488 Conjugate diluted 1:250 (ThermoFisher; Cat No: S32354). Each incubation took place at 37 °C for 30 minutes and was followed by three washing steps in PBS for 5 minutes each. After dehydration by exposure to the above-mentioned increasing concentrations of ethanol, the specimens were air-dried and embedded in Citifluor AF1/DAPI (400 ng ml^{-1}) solution (Aurion; Cat. No.: E17970). Stained chromosomes were visualized with a Leica DMRA fluorescence microscope (Leica, Wetzlar, Germany) and images were captured with the aid of a CoolSnap HQ2 camera (Photometrics, Tucson, USA).

Molecular Combing

Genome editing outcomes resulting from transducing human myoblasts with AdVPs designed for HR- or HMEJ-based gene targeting were assessed using the molecular combing gene editing quality control assay from Genomic Vision (Bagneux, France). Cells from control and AdVP-engineered cell lines were harvested and embedded in 1% low melting agarose plugs using the Genomic Vision FiberPrep® kit (Genomic Vision; Cat. No.: EXTR-001) at a concentration of 10×10^6 cells per plug. Subsequently, DNA extraction, combing, and immunostaining were performed according to the EasyComb procedure (Genomic Vision). Briefly, single long DNA molecules were extracted and stretched at a constant speed (~2 kb μm^{-1}) onto the surface of vinyl silane-treated glass coverslips (CombiCoverslips) employing the automated Molecular Combing System (MCS) instrument (Genomic Vision). The linearity and density of the combed DNA fibers were assessed by staining with the YOYO-1 dye to ensure precise and high-resolution measurements of hybridized DNA probes along the length of individually stretched DNA molecules (**Figure S1A**). DNA probes corresponding to the EGFP::DYS fusion construct (recombinant probes) and to the *AAVS1* "right" and "left" homology arms (RHA and LHA probes, respectively) were isolated after enzymatic digestion and gel extraction of AO75_pHC.Ad.EGFP::DYS DNA using established procedures. All the other DNA probes covering the *AAVS1* target locus (i.e., "right" and "left" flank anchoring probes) were generated through long-range PCR amplification using the appropriate primer

sets, reaction mixtures and cycling conditions (**Table S18** and **S19**). Next, FiberProbes® (Genomic Vision) were labelled and used as templates for FISH probe labelling by random priming. The correspondence between theoretical and experimental probe coverage patterns was validated by measuring probe hybridization lengths in control samples from unedited myoblasts (**Figure S1B**). Finally, coverslips from control and experimental samples were hybridized to the various labelled probes and fluorescent signals were detected by using the FiberVision® automated scanner (Genomic Vision). Image analysis and signal measurements were performed by using the FiberStudio® software (Genomic Vision).

Statistical analysis

Statistical analyses were performed with the aid of GraphPad Prism software (version 8.0.1) on datasets derived from independent biological replicates or technical replicates as defined in the figure legends. Statistical significances were calculated with the tests specified also in the figure legends. *P*-values lower than 0.05 were considered statistically significant.

Acknowledgements

Authors of this paper are members of the European Reference Network – Neuromuscular diseases (ERN EURO-NMD). The authors thank Ignazio Maggio and Aart Jochemsen (Department of Cell and Chemical Biology, LUMC, Leiden, The Netherlands) for assisting in the generation and characterization of AdVPs and for providing the short-hairpin RNA constructs used in this study, respectively, The authors are also thankful to Vincent Mouly (Institute of Myology, Sorbonne University, Paris, France), Jacques Tremblay (Laval University Hospital Center, Quebec, Canada) and Stephen Hauschka (Department of Biochemistry, University of Washington, Seattle, USA) for supplying the wild-type and DMD myoblasts, the construct encoding the human full-length dystrophin, and the synthetic striated muscle-restricted *CK8* promoter, respectively. Finally, the authors are grateful to all laboratory members for their support.

Additional Information

Data availability: All data gathered for and analysed in this study are included in the article and supplementary files. The flow cytometry datasets were deposited at FlowRepository under accession codes FR-FCM-Z4SL, FR-FCM-Z4SX, FR-FCM-Z4SG, FR-FCM-Z4SF, FR-FCM-Z4SM, FR-FCM-Z4SP and FR-FCM-Z4SN. The amplicon deep sequencing library reads generated for this study were deposited at the NCBI Sequence Read Archive (SRA) database under BioProject ID PRJNA834963.

Supplementary Data: Supplementary Data are available online at doi: 10.1093/nar/gkac567.

Funding: This project has received funding from the European Union's Horizon 2020

research and innovation program under the Marie Skłodowska-Curie grant agreement no. 765269 (F.T.). This study was also supported by the Prinses Beatrix Spierfonds (research grant W.OR16-13), the Dutch Duchenne Parent Project (research grant 17.012) and the China Scholarship Council–Leiden University Joint Scholarship programme (Q.W.).

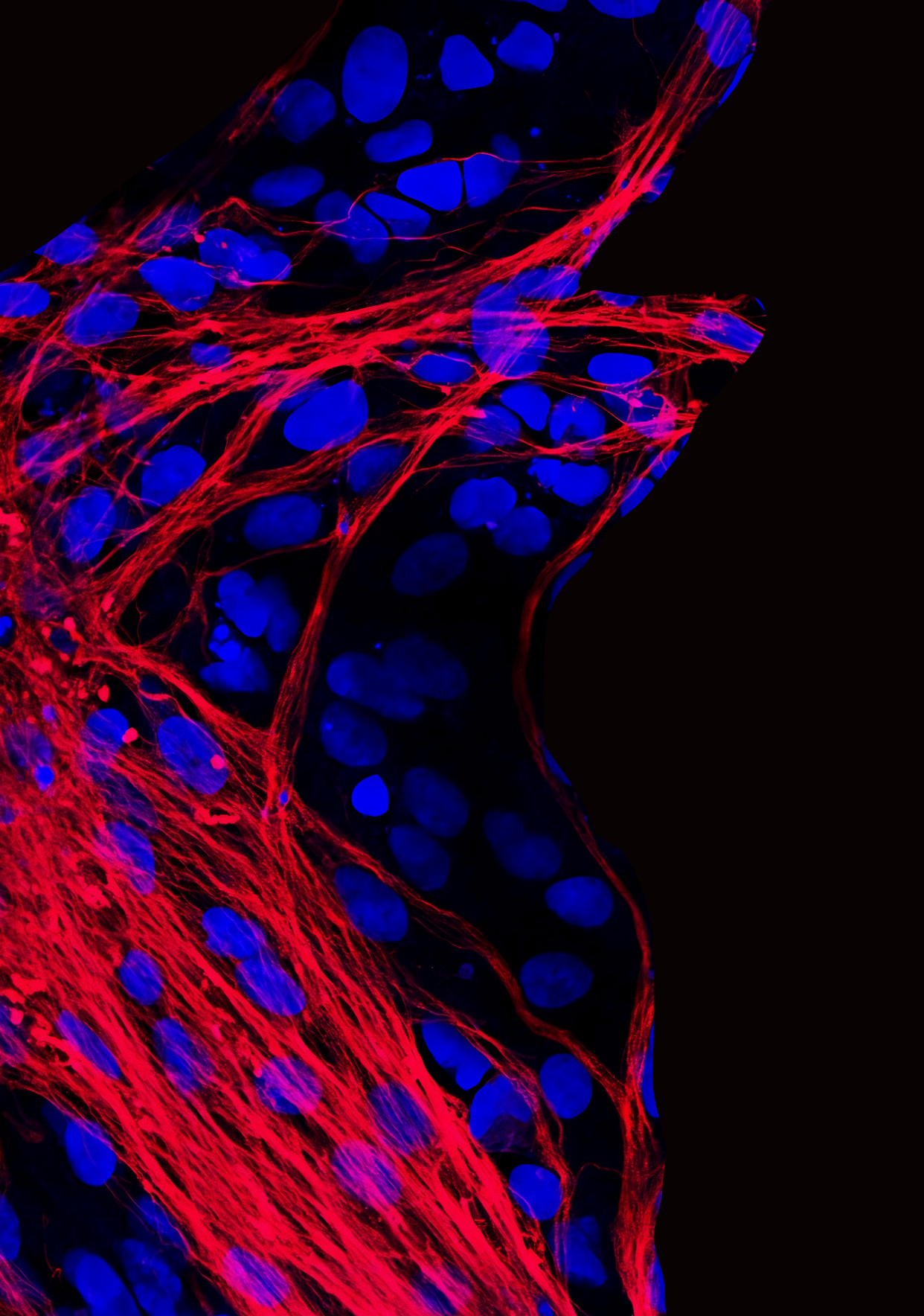
Conflict of interest statement: None declared.

References

1. Ernst, M.P.T.; Broeders, M.; Herrero-Hernandez, P.; Oussoren, E.; van der Ploeg, A.T.; Pijnappel, W. Ready for repair? gene editing enters the clinic for the treatment of human disease. *Mol Ther Methods Clin Dev.* **2020**, *18*, 532-557.
2. Sharma, G.; Sharma, A.R.; Bhattacharya, M.; Lee, S.S.; Chakraborty, C. CRISPR-Cas9: a preclinical and clinical perspective for the treatment of human diseases. *Mol Ther*. **2020**, *29*, 571-586.
3. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **2013**, *339*, 819-823.
4. Cho, S.W.; Kim, S.; Kim, J.M.; Kim, J.S.Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol.* **2013**, *31*, 230-232.
5. Jinek, M.; East, A.; Cheng, A.; Lin, S.; Ma, E.; Doudna, J. RNA-programmed genome editing in human cells. *Elife* **2013**, *2*, e00471.
6. Mali, P.; Yang, L.; Esvelt, K.M.; Aach, J.; Guell, M.; DiCarlo, J.E.; Norville, J.E.; Church, G.M. RNA-guided human genome engineering via Cas9. *Science* **2013**, *339*, 823-826.
7. Kleinstiver, B.P.; Pattanayak, V.; Prew, M.S.; Tsai, S.Q.; Nguyen, N.T.; Zheng, Z.; Joung, J.K. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **2016**, *529*, 490-495.
8. Slaymaker, I.M.; Gao, L.; Zetsche, B.; Scott, D.A.; Yan, W.X.; Zhang, F. Rationally engineered Cas9 nucleases with improved specificity. *Science* **2016**, *351*, 84-88.
9. Chen, X.; Gonçalves, M.A.F.V. DNA, RNA, and Protein Tools for Editing the Genetic Information in Human Cells. *iScience* **2018**, *6*, 247-263.
10. Chandrasegaran, S.; Carroll, D. Origins of Programmable Nucleases for Genome Engineering. J *Mol Biol.* **2016**, *428*, 963-989.
11. He, X.; Li, Y.X.; Feng, B. New turns for high efficiency knock-In of large DNA in human pluripotent stem cells. *Stem Cells Int*. **2018**, 9465028.
12. Jang, H.K.; Song, B.; Hwang, G.H.; Bae, S. Current trends in gene recovery mediated by the CRISPR-Cas system. *Exp Mol Med.* **2020**, *52*, 1016-1027.
13. Zhang, J.P.; Li, X.L.; Li, G.H.; Chen, W.; Arakaki, C.; Botimer, G.D.; Baylink, D.; Zhang, L.; Wen, W.; Fu, Y.W.; et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* **2017**, *18*, 35.
14. Yao, X.; Wang, X.; Hu, X.; Liu, Z.; Liu, J.; Zhou, H.; Shen, X.; Wei, Y.; Huang, Z.; Ying, W.; et al. Homology-mediated end joining-based targeted integration using CRISPR/Cas9. *Cell Res.* **2017**, *27*, 801-814.
15. He, X.; Tan, C.; Wang, F.; Wang, Y.; Zhou, R.; Cui, D.; You, W.; Zhao, H.; Ren, J.; Feng, B. Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res.* **2016**, *44*, e85.
16. Suzuki, K.; Tsunekawa, Y.; Hernandez-Benitez, R.; Wu, J.; Zhu, J.; Kim, E.J.; Hatanaka, F.; Yamamoto, M.; Araoka, T.; Li, Z.; et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **2016**, *540*, 144-149.
17. Byrne, S.M.; Ortiz, L.; Mali, P.; Aach, J.; Church, G.M. Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res*. **2015**, *43*, e21.
18. Palmer, D.J.; Grove, N.C.; Ing, J.; Crane, A.M.; Venken, K.; Davis, B.R.; Ng, P. Homology requirements for efficient, footprintless gene editing at the CFTR locus in human iPSCs with helper-dependent adenoviral vectors. *Mol Ther Nucleic Acids.* **2016**, *5*, e372.
19. Yao, X.; Wang, X.; Liu, J.; Shi, L.; Huang, P.; Yang, H. CRISPR/Cas9-mediated targeted integration in vivo using a homology-mediated end joining-based strategy. *J Vis Exp.* **2018**, *133*, e56844.
20. Holkers, M.; Maggio, I.; Henriques, S.F.; Janssen, J.M.; Cathomen, T.; Gonçalves, M.A.F.V. Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat Methods.* **2014**, *11*, 1051-1057.
21. Chen, X.; Gonçalves, M.A.F.V. Engineered viruses as genome editing devices. *Mol Ther.* **2016**, *24*, 447-457.
22. Gonçalves, M.A.F.V. Adeno-associated virus: from defective virus to effective vector. *Virol J*. **2005**, *2*, 43.

23. Gao, J.; Mese, K.; Bunz, O.; Ehrhardt, A. State-of-the-art human adenovirus vectorology for therapeutic approaches. *FEBS Lett.* **2019**, *593*, 3609-3622.
24. Tasca, F.; Wang, Q.; Gonçalves, M.A.F.V. Adenoviral vectors meet gene editing: a rising partnership for the genomic engineering of human stem cells and their progeny. *Cells* **2020**, *9*, e953.
25. Ricobaraza, A.; Gonzalez-Aparicio, M.; Mora-Jimenez, L.; Lumbreras, S.; Hernandez-Alcoceba, R. High-capacity adenoviral vectors: expanding the scope of gene therapy. *Int J Mol Sci.* **2020**, *21*, e3643.
26. Duan, D.; Goemans, N.; Takeda, S.; Mercuri, E.; Aartsma-Rus, A. Duchenne muscular dystrophy. *Nat Rev Dis Primers.* **2021**, *7*, 13.
27. Chemello, F.; Bassel-Duby, R.; Olson, E.N. Correction of muscular dystrophies by CRISPR gene editing. *J Clin Invest.* **2020**, *130*, 2766-2776.
28. Boyer, O.; Butler-Browne, G.; Chinoy, H.; Cossu, G.; Galli, F.; Lilleker, J.B.; Magli, A.; Mouly, V.; Perlingeiro, R.C.R.; Previtali, S.C.; et al. Myogenic cell transplantation in genetic and acquired diseases of skeletal muscle. *Front Genet.* **2021**, *12*, 702547.
29. Loperfido, M.; Steele-Stallard, H.B.; Tedesco, F.S.; VandenDriessche, T. Pluripotent stem cells for gene therapy of degenerative muscle diseases. *Curr Gene Ther.* **2015**, *15*, 364-380.
30. Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **2007**, *131*, 861-872.
31. Mamchaoui, K.; Trollet, C.; Bigot, A.; Negroni, E.; Chaouch, S.; Wolff, A.; Kandalla, P.K.; Marie, S.; Di Santo, J.; St Guily, J.L.; et al. Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. *Skelet Muscle.* **2011**, *1*, 34.
32. Cudrè-Mauroux, C.; Occhiodoro, T.; Konig, S.; Salmon, P.; Bernheim, L.; Trono, D. Lentivector-mediated transfer of Bmi-1 and telomerase in muscle satellite cells yields a duchenne myoblast cell line with long-term genotypic and phenotypic stability. *Hum Gene Ther.* **2003**, *14*, 1525-1533.
33. Brescia, M.; Janssen, J.M.; Liu, J.; Gonçalves, M.A.F.V. High-capacity adenoviral vectors permit robust and versatile testing of DMD gene repair tools and strategies in human cells. *Cells* **2020**, *9*, 869.
34. Chen, X.; Tasca, F.; Wang, Q.; Liu, J.; Janssen, J.M.; Brescia, M.D.; Bellin, M.; Szuhai, K.; Kenrick, J.; Frock, R.L.; et al. Expanding the editable genome and CRISPR-Cas9 versatility using DNA cutting-free gene targeting based on in trans paired nicking. *Nucleic Acids Res.* **2020**, *48*, 974-995.
35. Zhang, M.; D'Aniello, C.; Verkerk, A.O.; Wrobel, E.; Frank, S.; Ward-van Oostwaard, D.; Piccini, I.; Freund, C.; Rao, J.; Seebohm, G.; et al. Recessive cardiac phenotypes in induced pluripotent stem cell models of Jervell and Lange-Nielsen syndrome: disease mechanisms and pharmacological rescue. *Proc Natl Acad Sci U S A.* **2014**, *111*, E5383-E5392.
36. van den Berg, C.W.; Ritsma, L.; Avramut, M.C.; Wiersma, L.E.; van den Berg, B.M.; Leuning, D.G.; Lievers, E.; Koning, M.; Vanslambrouck, J.M.; Koster, A.J.; et al. Renal subcapsular transplantation of PSC-derived kidney organoids induces neo-vasculogenesis and significant glomerular and tubular maturation in vivo. *Stem Cell Reports* **2018**, *10*, 751-765.
37. Chapdelaine, P.; Moisset, P.A.; Campeau, P.; Asselin, I.; Vilquin, J.T.; Tremblay, J.P. Functional EGFP-dystrophin fusion proteins for gene therapy vector development. *Protein Eng.* **2000**, *13*, 611-615.
38. Gonçalves, M.A.F.V.; Janssen, J.M.; Nguyen, Q.G.; Athanasopoulos, T.; Hauschka, S.D.; Dickson, G.; de Vries, A.A. Transcription factor rational design improves directed differentiation of human mesenchymal stem cells into skeletal myocytes. *Mol Ther.* **2011**, *19*, 1331-1341.
39. Janssen, J.M.; Liu, J.; Skokan, J.; Gonçalves, M.A.F.V.; de Vries, A.A. Development of an AdEasy-based system to produce first- and second-generation adenoviral vectors with tropism for CAR- or CD46-positive cells. *J Gene Med.* **2013**, *15*, 1-11.
40. Herold, M.J.; van den Brandt, J.; Seibler, J.; Reichardt, H.M. Inducible and reversible gene silencing by stable integration of an shRNA-encoding lentivirus in transgenic rats. *Proc Natl Acad Sci USA* **2008**, *105*, 18507-18518.
41. Pelascini, L.P.; Janssen, J.M.; Gonçalves, M.A.F.V. Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. *Hum Gene Ther.* **2013**, *24*, 78-96.
42. Pelascini, L.P.; Gonçalves, M.A.F.V. Lentiviral vectors encoding zinc-finger nucleases specific for the model target locus HPRT1. *Methods Mol Biol.* **2014**, *1114*, 181-199.
43. Campostrini, G.; Meraviglia, V.; Giacomelli, E.; van Helden, R.W.J.; Yiangou, L.; Davis, R.P.; Bellin, M.; Orlova, V.V.; Mummery, C.L. Generation, functional analysis and applications of isogenic three-dimensional self-aggregating cardiac microtissues from human pluripotent stem cells. *Nat Protoc.* **2021**, *16*, 2213-2256.
44. Szuhai, K.; Tanke, H.J. COBRA: combined binary ratio labeling of nucleic acid probes for multi-color fluorescence in situ hybridization karyotyping. *Nat Protoc.* **2006**, *1*, 264–275.
45. Conant, D.; Hsiau, T.; Rossi, N.; Oki, J.; Maures, T.; Waite, K.; Yang, J.; Joshi, S.; Kelso, R.; Holden, K.; et al. Inference of CRISPR Edits from Sanger Trace Data. CRISPR J, 2022, *5*, 123-130.
46. Maggio, I.; Holkers, M.; Liu, J.; Janssen, J.M.; Chen, X.; Gonçalves, M.A.F.V. Adenoviral vector delivery of

RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells. *Sci Rep.* **2014**, *29*, 5105.
47. Wang, Q.; Liu, J.; Janssen, J.M.; Le Bouteiller, M.; Frock, R.L.; Gonçalves, M.A.F.V. Precise and broad scope genome editing based on high-specificity Cas9 nickases. *Nucleic Acids Res.* **2021**, *49*, 1173-1198.
48. Wang, Q.; Liu, J.; Janssen, J.M.; Tasca, F.; Mei, H.; Gonçalves, M.A.F.V. Broadening the reach and investigating the potential of prime editors through fully viral gene-deleted adenoviral vector delivery. *Nucleic Acids Res.* **2021**, *49*, 11986-12001.
49. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2011**, *17*, 10–12.
50. Clement, K.; Rees, H.; Canver, M.C.; Gehrke, J.M.; Farouni, R.; Hsu, J.Y.; Cole, M.A.; Liu, D.R.; Joung, J.K.; Bauer, D.E.; et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol.* **2019**, *37*, 224-226.
51. Stirling, D.R.; Swain-Bowden, M.J.; Lucas, A.M.; Carpenter, A.E.; Cimini, B.A.; Goodman, A. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics.* **2021**, *22*, 433.
52. Rossi, S.; Szuhai, K.; Ijszenga, M.; Tanke, H.J.; Zanatta, L.; Sciot, R.; Fletcher C.D.M.; Tos, A.P.D.; Hogendoorn, P.C.W. *EWSR1-CREB1* and *EWSR1-ATF1* fusion genes in angiomatoid fibrous histiocytoma. *Clin Cancer Res.* **2017**, *13*, 7322–7328.
53. Gonçalves, M.A.F.V.; de Vries, A.A. Adenovirus: from foe to friend. *Rev Med Virol.* **2006**, *16*, 167-186.
54. Gaggar, A.; Shayakhmetov, D.M.; Lieber, A. CD46 is a cellular receptor for group B adenoviruses. *Nat Med.* **2003**, *9*, 1408-1412.
55. Gonçalves, M.A.F.V.; Holkers, M.; Cudre-Mauroux, C.; van Nierop, G.P.; Knaan-Shanzer, S.; van der Velde, I.; Valerio, D.; de Vries, A.A. Transduction of myogenic cells by retargeted dual high-capacity hybrid viral vectors: robust dystrophin synthesis in duchenne muscular dystrophy muscle cells. *Mol Ther.* **2006**, *13*, 976-986.
56. Papapetrou, E.P.; Schambach, A. Gene insertion into genomic safe harbors for human gene therapy. *Mol Ther.* **2016**, *24*, 678-684.
57. Pavani, G.; Amendola, M. Targeted gene delivery: where to land. *Front Genome Ed.* **2020**, *2*, 609650.
58. Lombardo, A.; Cesana, D.; Genovese, P.; Di Stefano, B.; Provasi, E.; Colombo, D.F.; Neri, M.; Magnani, Z.; Cantore, A.; Lo Riso, P.; et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat Methods.* **2011**, *8*, 861-869.
59. Slaymaker, I.M.; Gao, L.; Zetsche, B.; Scott, D.A.; Yan, W.X.; Zhang, F. Rationally engineered Cas9 nucleases with improved specificity. *Science* **2016**, *351*, 84-88.
60. Maggio, I.; Zittersteijn, H.A.; Wang, Q.; Liu, J.; Janssen, J.M.; Ojeda, I.T.; van der Maarel, S.M.; Lankester, A.C.; Hoeben, R.C.; Gonçalves, M.A.F.V. Integrating gene delivery and gene-editing technologies by adenoviral vector transfer of optimized CRISPR-Cas9 components. *Gene Ther.* **2020**, *27*, 209-225.
61. Dang, Y.; Jia, G.; Choi, J.; Ma, H.; Anaya, E.; Ye, C.; Shankar, P.; Wu, H. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol.* **2015**, *16*, 280.
62. Bonini, C.; Grez, M.; Traversari, C.; Ciceri, F.; Marktel, S.; Ferrari, G.; Dinauer, M.; Sadat, M.; Aiuti, A.; Deola, S.; et al. Safety of retroviral gene marking with a truncated NGF receptor. *Nat Med.* **2003**, *9*, 367-369.
63. Tolle, F.; Stucheli, P.; Fussenegger, M. Genetic circuitry for personalized human cell therapy. *Curr Opin Biotechnol.* **2019**, *59*, 31-38.
64. Sharma, A.; Sances, S.; Workman, M.J.; Svendsen, C.N. Multi-lineage human iPSC-derived platforms for disease modeling and drug discovery. *Cell Stem Cell.* **2020**, *26*, 309-329.
65. Campostrini, G.; Windt, L.M.; van Meer, B.J.; Bellin, M.; Mummery, C.L. Cardiac tissues from stem cells: new routes to maturation and cardiac regeneration. *Circ Res.* **2021**, *128*, 775-801.
66. Ihry, R.J.; Worringer, K.A.; Salick, M.R.; Frias, E.; Ho, D.; Theriault, K.; Kommineni, S.; Chen, J.; Sondey, M.; Ye, C.; et al. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat Med.* **2018**, *24*, 939-946.
67. Schiroli, G.; Conti, A.; Ferrari, S.; Della Volpe, L.; Jacob, A.; Albano, L.; Beretta, S.; Calabria, A.; Vavassori, V.; Gasparini, P.; et al. Precise Gene Editing Preserves Hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. *Cell Stem Cell.* **2019**, *24*, 551-565.
68. Suzuki, K.; Yu, C.; Qu, J.; Li, M.; Yao, X.; Yuan, T.; Goebl, A.; Tang, S.; Ren, R.; Aizawa, E.; et al. Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell Stem Cell* **2014**, *15*, 31-36.
69. Palmer, D.J.; Turner, D.L.; Ng, P. A Single "All-in-One" Helper-Dependent Adenovirus to Deliver Donor DNA and CRISPR/Cas9 for Efficient Homology-Directed Repair. *Mol Ther Methods Clin Dev.* **2020**, *17*, 441-447.
70. Lee, B.C.; Lozano, R.J.; Dunbar, C.E. Understanding and overcoming adverse consequences of genome editing on hematopoietic stem and progenitor cells. *Mol Ther.* **2021**, *29*, 3205-3218.



Chapter 5

Expanding the editable genome and CRISPR-Cas9 versatility using DNA cutting-free gene targeting based on *in trans* paired nicking

Xiaoyu Chen[1,4,5], Francesca Tasca[1,5], Qian Wang[1], Jin Liu[1], Josephine M. Janssen[1], Marcella D. Brescia[1], Milena Bellin[2], Karoly Szuhai[1], Josefin Kenrick[3], Richard L. Frock[3], Manuel A.F.V. Gonçalves[1,*]

[1]Leiden University Medical Center, Department of Cell and Chemical Biology, Einthovenweg 20 2333 ZC, Leiden, the Netherlands, [2]Leiden University Medical Center, Department of Anatomy and Embryology, Einthovenweg 20 2333 ZC, Leiden, the Netherlands, [3]Stanford University School of Medicine, Division of Radiation and Cancer Biology, DNA Breakpoint Mapping Laboratory, 291 Campus DR, Stanford CA, 94305, USA, [4]Present address, Stanford University School of Medicine, Department of Psychiatry and Behavioral Sciences, 318 Campus DR, Stanford CA, 94305, USA.
[5]These authors contributed equally to this work.

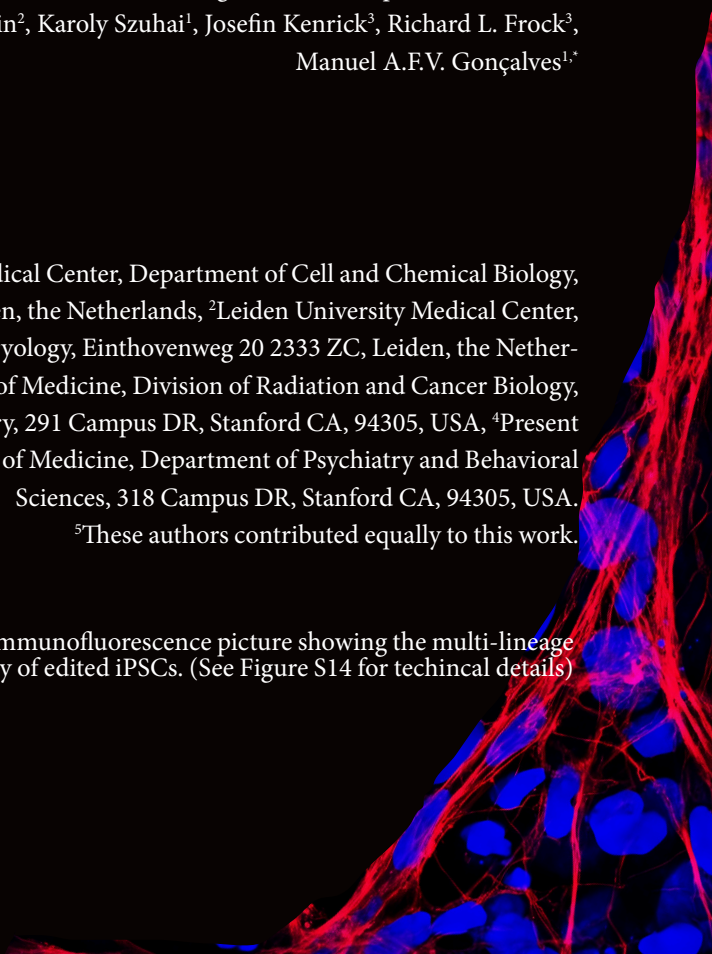


Illustration: Modified from immunofluorescence picture showing the multi-lineage differentiation capacity of edited iPSCs. (See Figure S14 for techincal details)

Abstract

Genome editing typically involves recombination between donor nucleic acids and acceptor genomic sequences subjected to double-stranded DNA breaks (DSBs) made by programmable nucleases (e.g. CRISPR–Cas9). Yet, nucleases yield off-target mutations and, most pervasively, unpredictable target allele disruptions. Remarkably, to date, the untoward phenotypic consequences of disrupting allelic and non-allelic (e.g. pseudogene) sequences have received scant scrutiny and, crucially, remain to be addressed. Here, we demonstrate that gene-edited cells can lose fitness as a result of DSBs at allelic and non-allelic target sites and report that simultaneous single-stranded DNA break formation at donor and acceptor DNA by CRISPR–Cas9 nickases (*in tran*s paired nicking) mostly overcomes such disruptive genotype-phenotype associations. Moreover, *in trans* paired nicking gene editing can efficiently and precisely add large DNA segments into essential and multiple-copy genomic sites. As shown herein by genotyping assays and high-throughput genome-wide sequencing of DNA translocations, this is achieved while circumventing most allelic and non-allelic mutations and chromosomal rearrangements characteristic of nuclease-dependent procedures. Our work demonstrates that *in trans* paired nicking retains target protein dosages in gene-edited cell populations and expands gene editing to chromosomal tracts previously not possible to modify seamlessly due to their recurrence in the genome or essentiality for cell function.

Introduction

Genome editing based on homology-dependent and homology-independent DNA repair pathways activated by programmable nucleases permits modifying specific chromosomal sequences in living cells [1]. Importantly, these genetic changes can span from single base pairs to whole transgenes [2]. However, the genomic double-stranded DNA breaks (DSBs) required for DNA repair activation inevitably yield complex and unpredictable genetic structural variants. These by-products result from the fact that DSBs (targeted or otherwise) are substrates for prevalent non-homologous end joining (NHEJ) pathways and other error-prone recombination processes [3]. These processes can trigger local [4] and genome-wide mutations and rearrangements, in the form of insertions and deletions (indels), duplications and/or translocations [5–10]. Likewise insidious, targeted DSBs at homologous allWWeles can result in the assembly of unstable dicentric chromosomes through head-to-head inversional translocations [10]. Finally, the engagement of donor DNA with target and off-target DSBs often leads to inaccurate and random chromosomal insertion events, respectively [2, 11]. This is especially so when donor DNA is presented in target cell nuclei as free-ended double-stranded recombination substrates [11–13].

The unpredictability of genome editing outcomes is naturally aggravated whenever nuclease target sites are located in (i) coding sequences, especially those associated with essentiality and haploinsufficiency, (ii) overlapping trans-acting or cis-acting sequences and (iii) multiple-copy sequences, such as those in paralogs and pseudogenes. To date, genotypic and phenotypic consequences resulting from editing these three types of genomic regions have received limited examination and remain to be addressed.

Single-stranded DNA breaks (SSBs) made by programmable sequence-specific and strand-specific nucleases (nickases) are intrinsically less disruptive than DSBs as they do not constitute canonical NHEJ substrates [14–17]. In this regard, CRISPR–Cas9 nickases consisting of guide RNAs (gRNAs) and Cas9 proteins with either their RuvC or HNH nuclease domains disabled (e.g. $Cas9^{D10A}$ and $Cas9^{H840A}$, respectively), are particularly appealing programmable nicking enzymes [18–20]. Indeed, similarly to their cleaving counterparts, CRISPR–Cas9 nickases target DNA consisting of a protospacer adjacent motif (PAM; NGG in *Streptococcus pyogenes* SpCas9) and a sequence complementary to the 5'-terminal 20 nucleotides (nts) of the gRNA (spacer) [18, 21]. Pairs of CRISPR–Cas9 nickases are commonly used to induce site-specific DSBs through coordinated nicking at opposite target DNA strands. This dual nicking strategy can significantly improve the specificity of DSB formation as SSBs made at off-target sites are, for the most part, faithfully repaired [22, 23]. However, genome editing based on paired CRISPR–Cas9 nickases remains prone to mutagenesis and chromosomal rearrangements due to the ultimate creation of DSBs [12, 22, 23].

The non-disruptive character of genome editing based on targeted chromosomal SSBs offers the possibility for seamlessly modifying a broad range of genomic sequences, including those that encode functional protein motifs or essential proteins or that are present in

genomic tracts with high similarity to DNA located elsewhere in the genome. Unfortunately, chromosomal SSBs are, *per se*, poor stimuli for genome editing via precise homology-directed DNA repair (HDR), even in instances in which single base pairs are due to be inserted at a target site [14–17, 24].

Here, we sought to determine whether chromosomal regions previously not possible to edit in an efficient and seamless manner could in fact be modified as such. In particular, we hypothesized that *in trans* paired nicking, comprising coordinated SSB formation at donor and acceptor HDR substrates by CRISPR–Cas9 nickases, permits expanding the 'editable genome', i.e. the genomic space amenable to operative DNA editing. Recently, it has been demonstrated that this genetic engineering principle achieves precise HDR-mediated genomic insertions, from a few base pairs [12, 25] to whole transgenes [12], without provoking the competing NHEJ pathway. However, the performance of *in trans* paired nicking at coding sequences of endogenous genes, in particular those associated with haploinsufficiency and essentiality, is unknown. To date, equally unknown is the performance of genome editing approaches based on repairing SSBs versus DSBs at these coding sequences using donor plasmids. By targeting exons in the *H2A.X variant histone* gene (*H2AX*) and the *POU class 5 homebox 1* gene (*POU5F1* or *OCT4*), whose products are essential for the DNA damage response and stem cell pluripotency, respectively, we demonstrate that in contrast to DSB-dependent strategies, *in trans* paired nicking achieves precise gene editing while disrupting neither functional motifs nor allelic or non-allelic homologous DNA. Moreover, after adapting linear amplification-mediated high-throughput genome-wide translocation sequencing (HTGTS) [10, 26] for the detection of SSB-initiated translocations, we found that CRISPR-SpCas9 nickases greatly reduce large-scale chromosomal rearrangements when compared to their nuclease counterparts. Finally, *PARP1* gene targeting experiments showed that, also in instances in which a target gene is not associated with haploinsufficiency or essentiality, *in trans* paired nicking achieves accurate HDR-mediated gene knock-ins without mutagenizing unmodified alleles, and hence, without reducing target protein dosages.

Results

Distinct prevalence of genome-wide rearrangements after SpCas9 versus SpCas9^{D10A} delivery

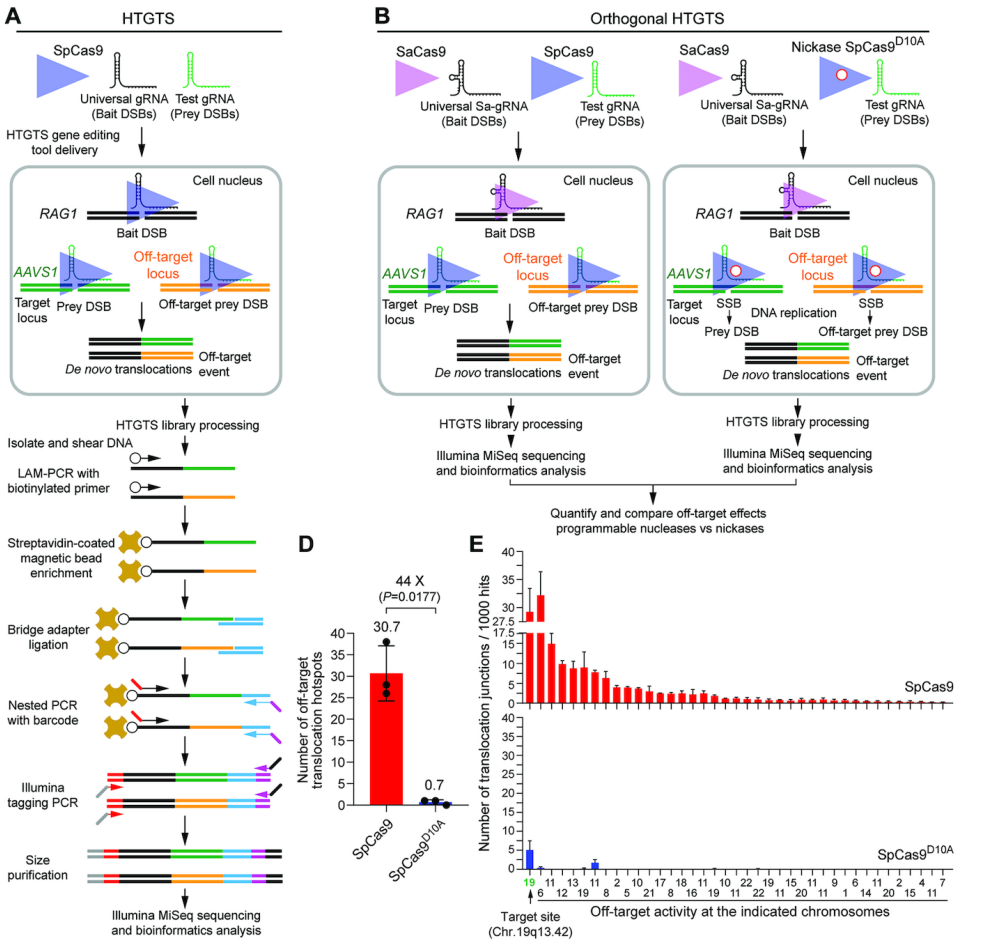
Genome-wide off-target effects of programmable nucleases are commonly assessed by high-throughput sequencing of exogenous DNA tags 'trapped' at two-ended DSB termini or, more recently, in situ detection of DSB repair factors [40, 41]. Although SSBs are mostly resolved through conservative repair processes they can in principle lead to DSBs if a replication fork advances through them and collapses [42]. However, the resulting one-ended chromosomal breaks are unlikely substrates for exogenous DNA 'trapping'. Therefore, to fulfil the lack of a sensitive and unbiased genome-wide assay for comparing off-target effects triggered by programmable nucleases versus programmable nickases, we have adapted the

HTGTS assay [10]. In contrast to other approaches, HTGTS detects off-target effects by deep sequencing of translocations joining bait and prey DSBs made by universal and test nucleases, respectively (**Figure 1A**). In addition to taking place at bona fide target sites, prey DSBs can also occur at off-target sites of a specific test nuclease under examination. In adapting the HTGTS assay for comparing off-target effects induced by nucleases versus nickases, we assured that bait DSBs are exclusively made by a universal nuclease whilst prey DSBs are instead generated by either test nucleases or test nickases. To this end, we combined S. pyogenes SpCas9 with its ortholog Staphylococcus aureus Cas9 (SaCas9). In particular, test S. pyogenes and universal S. aureus CRISPR complexes were designed for generating prey DNA lesions (i.e. SSBs or DSBs) and universal bait DSBs, respectively (orthogonal HTGTS). After selecting RAG1-targeting SaCas9:Sa-gRAG1.1 complexes as inducers of bait DSBs (**Figure S1**), HEK293T cells were exposed to these complexes together with SpCas9:gAAVS1 or SpCas9^{D10A}:gAAVS1, each cleaving or nicking, respectively, at the commonly used AAVS1 safe-harbour locus (**Figure 1B**). As expected, genotyping assays based on the mismatch-sensing T7EI enzyme, readily revealed indels at RAG1 and AAVS1 in cells subjected to SaCas9:Sa-gRAG1.1 and SpCas9:gAAVS1 (**Figure S2**). In contrast, indels were detected at RAG1 but not at AAVS1 in cells treated with SaCas9:Sa-gRAG1.1 and SpCas9^{D10A}:gAAVS1, confirming that the latter complex displays low mutagenicity at the target intron (**Figure S2**) [12]. Control orthogonal HTGTS read libraries generated by delivering SaCas9:Sa-gRAG1.1 alone, besides detecting a single poorly-enriched off-target site on chrmosome 1, revealed a genome-wide translocation pattern consistent with previously described S. pyogenes SpCas9:gRNA bait libraries (**Figure 1C**, **S3** and **S4**) [10]. Importantly, applying orthogonal HTGTS analyses to experimental DNA samples (**Figure 1C**, **S3** and **S4**), demonstrated that amidst cells exposed to SpCas9:gAAVS1 and SpCas9^{D10A}:gAAVS1, the former had significantly higher numbers of off-target translocation hotspots than the latter; i.e. 30.7 ± 6.4 versus 0.7 ± 0.6 recurrent hotspots, respectively (**Figure 1C** , **1D** and **S4**). In addition, SpCas9:gAAVS1 yielded higher frequencies of translocation junctions per hotspot than SpCas9^{D10A}:gAAVS1 (**Figure 1E** and **S4**). It is also noteworthy that, amongst the two translocation hotspots associated with SpCas9^{D10A}:gAAVS1 activity, was that involving RAG1 bait and AAVS1 prey target DNA (**Figure 1C**, **S3** and **S4**). This data suggests that individual SSBs can indeed be processed into chromosomal DSBs in living mammalian cells.

Together, these data establish orthogonal HTGTS as a sensitive method for the unbiased genome-wide detection of off-target effects elicited by genomic SSBs. Importantly, these results also lend support to SpCas9^{D10A} as a genome-editing tool that diminishes allelic and non-allelic chromosomal mutations and rearrangements.

In trans paired nicking minimizes disruptive genotype-phenotype associations

Earlier *AAVS1* gene targeting experiments in HeLa cells and human iPSCs demonstrated that DSB-dependent genome editing approaches yield more inaccurate and random donor


A
HTGTS
SpCas9
Universal gRNA (Bait DSBs)
Test gRNA (Prey DSBs)
HTGTS gene editing tool delivery
Cell nucleus
RAG1
Bait DSB
AAVS1
Off-target locus
Target locus
Prey DSB
Off-target prey DSB
De novo translocations
Off-target event
HTGTS library processing
Isolate and shear DNA
LAM-PCR with biotinylated primer
Streptavidin-coated magnetic bead enrichment
Bridge adapter ligation
Nested PCR with barcode
Illumina tagging PCR
Size purification
Illumina MiSeq sequencing and bioinformatics analysis
B
Orthogonal HTGTS
SaCas9
Universal Sa-gRNA (Bait DSBs)
Nickase SpCas9D10A
SSB
DNA replication
Quantify and compare off-target effects programmable nucleases vs nickases
D
44 X
(P=0.0177)
30.7
0.7
Number of off-target translocation hotspots
SpCas9
SpCas9D10A
E
Number of translocation junctions / 1000 hits
Target site (Chr.19q13.42)
Off-target activity at the indicated chromosomes
C
SaCas9:Sa-gRNA1.1 (DSB bait-inducing nuclease)
1 Translocation hotspot (20,515 Junctions)
SaCas9:Sa-gRNA1.1 + SpCas9:gAAVS1 nuclease
31 Translocation hotspots
30 Off-targets
(24,523 Junctions)
SaCas9:Sa-gRNA1.1 + SpCas9D10A:gAAVS1 nickase
3 Translocation hotspots
1 Off-target
(15,632 Junctions)

◄ **Figure 1. Comparing off-target effects triggered by cleaving versus nicking CRISPR complexes.** (**A**) Diagram of the HTGTS pipeline for detecting SpCas9-induced off-target effects. Cells are exposed to S. pyogenes CRISPR complexes containing universal and test gRNAs that induce bait and prey DSBs at RAG1 and target loci, respectively. The prevalence and distribution of off-target hotspots conferred by test gRNAs are determined by an HTGTS pipeline comprising next-generation sequencing of translocations between RAG1 and off-target DNA (black and orange lines, respectively). (**B**) Diagram of the orthogonal HTGTS pipeline for detecting SpCas9^{D10A}-induced off-target effects. Orthogonal HTGTS assays make use of S. aureus and S. pyogenes CRISPR complexes for generating bait DSBs at RAG1 and either prey DSBs or nicks at target loci, respectively. The orthogonality (i.e. lack of cross-talk) between gRNAs and Cas9 proteins from these CRISPR systems avoids nicking at RAG1 and cleaving at off-target sites of test SpCas9^{D10A}:gRNA complexes (right panel). Further, exchanging SpCas9^{D10A} by SpCas9 in parallel orthogonal HTGTS assays permits comparing side-by-side genomic disruptions inflicted by cleaving versus nicking CRISPR complexes (left panel). Original and orthogonal HTGTS assays share the same downstream library processing and bioinformatics analysis steps. (**C**) Cumulative orthogonal HTGTS analyses (i.e. Circos plots) from three biological replicates. Arrowheads on chromosome 11 indicate the location of the SaCas9:Sa-gRAG1.1 universal bait DSB for all sequence read libraries; stars on chromosome 19 mark the AAVS1 target site of test S. pyogenes CRISPR complexes. Blue-graded lines from bait DSBs at the RAG1 locus indicate bait-related off-targets whereas red-graded lines indicate test gAAVS1-related translocation hotspots from the activity of S. pyogenes CRISPR complexes at target and off-target sites. Hotspots are established only when significantly enriched translocation sites are present in the majority of independent HTGTS replicate experiments (n≥2). Black bars represent 5 Mb bins across each chromosome and enrichment is displayed on a custom color coded log scale by order of magnitude. (**D**) Number of gAAVS1 off-target translocation hotspots in SpCas9 and SpCas9^{D10A} sequence read libraries. Significance was calculated with paired two-tailed Student›s t tests. (**E**) Relative frequencies of junctions per gAAVS1 translocation hotspot in SpCas9 and SpCas9^{D10A} sequence read libraries. Individual experimental values and respective Circos plots are shown in **Figures S3** and **S4**, respectively. Bars and error bars in panels D and E indicate mean ± S.D., respectively (n = 3 independent biological replicates).

DNA insertions than *in trans* paired nicking [12]. Besides augmenting genotype-phenotype unpredictability, such as via insertional mutagenesis, random chromosomal DNA integration results in transgene expression variegation due to chromosomal positional effects [11, 12]. Similar AAVS1 gene targeting experiments performed in HEK293T cells support these previous findings [11, 12] by showing that heterogeneous transgene expression is prevalent in cell populations subjected to donor plasmids and DSB-forming nucleases (**Figure S5**).

Tagging endogenous proteins with fluorescent reporters is a frequent goal of genome editing endeavours, including for establishing live-cell screening systems or studying cellular processes in a spatiotemporal fashion. However, the need for targeting gene termini limits the availability of gRNAs with potentially high activities and/or specificities. The presence of functional motifs further limits gRNA design as, often, HDR-mediated knock-in of one allele is accompanied by NHEJ-induced knockout of the other allele creating functional gene-dose imbalances. The gRNA availability issue becomes extreme in cases where target sequences (coding or otherwise) are not unique in the genome. These sequences are in fact dubbed 'impossible to target' in the CRISPR tracks of the UCSC Genome Browser and are defined as having at least one identical copy in the genome [43]. Thus, as challenging targets for comparing the performance of SpCas9 versus SpCas9^{D10A}, we sought to tag housekeeping H2AX and cell type-specific OCT4 alleles with live-cell reporters. The difficulty in tagging these genes stems from the fact that H2AX function depends on a C-terminal SQ phosphorylation motif [44] that restricts gRNA selection in this coding region and OCT4 termini share 100% sequence identity with sequences found in four autosomal pseudogenes that prevents the identification of OCT4-specific gRNAs.

H2AX gene editing experiments were initiated by transfecting HeLa cells with plasmids

expressing cleaving SpCas9:gRNA or nicking SpCas9^{D10A}:gRNA complexes containing gRNA$^{H2AX.1}$ or gRNA$^{H2AX.2}$ (F**igure 2A)**. The transfection mixtures included donor constructs pDonor^{H2AX} or pDonor$^{H2AX.TS}$. The latter differs from the former in that it has the H2AX-specific gRNA target sites flanking the targeting module consisting of 'homology arms' and a mCherry reporter tag (**Figure 2A** and **2B**). After delivering these tools, we sought to access the efficiency and precision of gene editing involving (i) DSBs on target DNA (standard), (ii) DSBs on target and donor DNA (paired breaking; DSB^2), (iii) SSBs on target DNA (single nicking) and (iv) SSBs on target and donor DNA (*in trans* paired nicking; Nick^2) (**Figure 2B**). The efficiency and precision of H2AX gene editing was ascertained by combining flow cytometric quantification of mCherry$^+$ cells with molecular analysis of randomly isolated mCherry$^+$ clones, each of which, representing an individual genome-modifying event. Importantly, we exploited the fact that the mCherry-tagged intronless H2AX gene in donor plasmids behaves as an autonomous reporter unit (**Figure 2C**, top panel) to avoid biased selection of cells harbouring targeted exogenous DNA chromosomal insertions (targeted integrants). The frequencies of transiently and stably transfected cells were determined by flow cytometry before and after episomal templates had been eliminated through sub-culturing (**Figure 2C**, top and bottom panel, respectively). This analysis revealed that, for both gRNAs used, *in trans* paired nicking yielded ~4-fold higher percentages of stably transfected cells than those resulting from the single nicking approach (**Figure 2C**, bottom panel). The robust enhancement on the frequencies of genetically modified cells achieved by *in trans* paired nicking over those resulting from the single nicking strategy is consistent with previous experiments targeting introns [12]. Hence, in addition to supporting initial theoretical models postulating nicked DNA partners as homologous recombination substrates [45], these results further stress the limited utility of the single nicking approach. The paired breaking strategy led to the highest frequencies of stably transfected cells (**Figure 2C**, bottom panel). However, it is worth noting that the attendant free-ended donor DNA templates created in cellula by SpCas9-mediated cutting (paired breaking) are prone to yielding complex genome-modifying events, i.e., off-target and inaccurately targeted chromosomal insertions, including concatemeric and HDR-independent integrants [2, 11–13]. Indeed, although genetically modified cells expressed tagged H2AX transcripts independently of the gene editing procedure used (**Figure S6**), junction PCR screens of randomly selected mCherry$^+$ clones readily revealed that paired breaking yielded the least precisely targeted integrants when compared to standard and *in trans* paired nicking (**Figure 2D** and **S7**). Notably, untagged H2AX alleles in mCherry$^+$ clones exposed to SpCas9 and SpCas9^{D10A} had varying and uniform sizes, respectively (**Figure S7**). These results support recent findings indicating that, in addition to short indels, SpCas9 can induce gross structural variants at target sequences, such as, large insertions and deletions [4, 10]. To further characterize these collateral gene-editing events, nucleotide sequencing of H2AX alleles was done in mCherry$^+$ clones modified through either standard or *in trans* paired nicking procedures. This target site genotyping analysis confirmed the presence

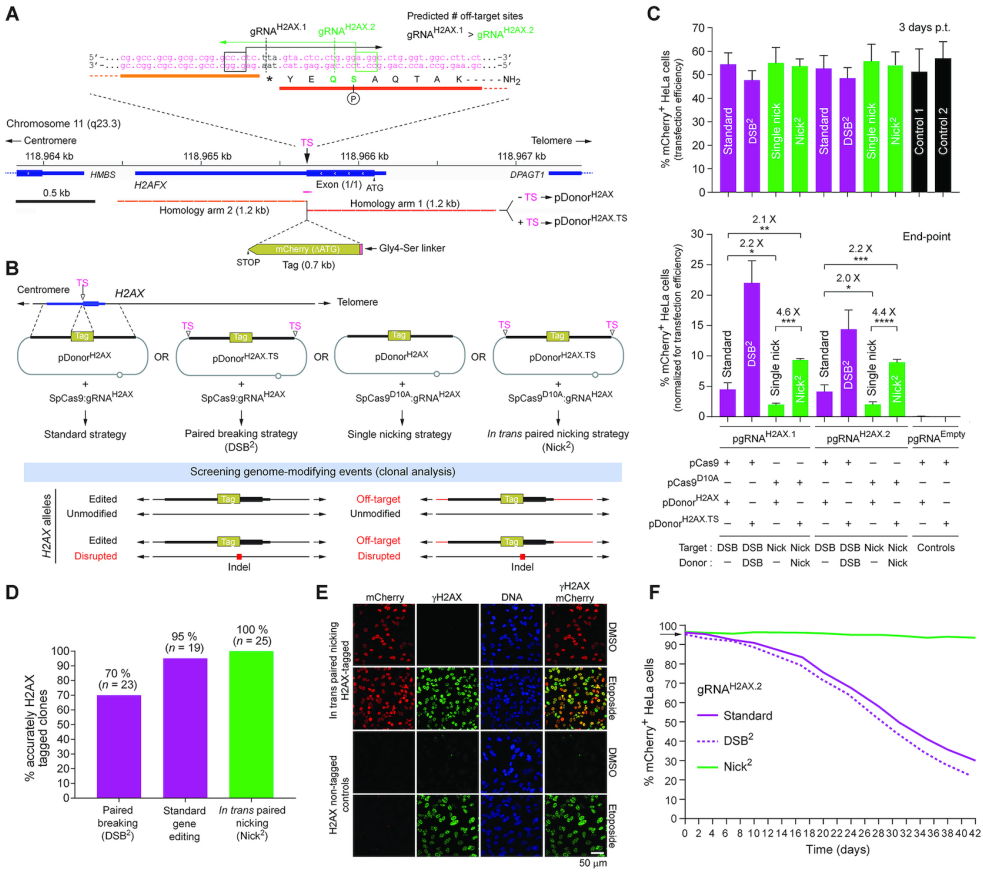


Figure 2. Homology-directed H2AX gene editing based on cleaving or nicking CRISPR complexes. (**A**) Diagram of the H2AX genomic region. The gRNA$^{H2AX.1}$ and gRNA$^{H2AX.2}$ target sites (TS) are highlighted by horizontal arrows and boxed nucleotides (PAMs). The H2AX post-translationally phosphorylated serine residue 129 is marked with a circled P. The donor plasmids pDonor^{H2AX} and pDonor$^{H2AX.TS}$ contain as targeting module H2AX sequences ('arms of homology') flanking a mCherry tag. (**B**) Schematics of H2AX gene editing strategies. Standard and paired breaking gene editing involve DSB formation at the genomic TS or at this TS and those in the donor DNA, respectively. Single nicking and *in trans* paired nicking gene editing comprise SSB formation at the genomic TS or at this TS and those in the donor DNA, respectively. Wanted and unwanted (red icons) genome-modifying events are depicted. (**C**) Quantification of transiently and stably transfected human cells. Flow cytometry was done on HeLa cell cultures co-transfected with the indicated plasmids. Top and bottom graphs, frequencies of mCherry^{+} cells at early and late time points after transfection (3 days and 2 weeks, respectively). Data are presented as mean ± s.e.m. of four independent biological replicates. Significance between the indicated datasets was calculated with one-way ANOVA followed by Tukey›s test for multiple comparisons; *P< 0.05; **P<0.01; ***P< 0.001; ****P< 0.0001. (**D**) Assessing H2AX gene editing accuracy. The frequencies of precisely targeted mCherry^{+} clones were determined through junction PCR screens (**Figure S7**). (**E**) Confocal microscopy analysis of H2AX gene-edited cells. HeLa cells genetically modified by *in trans* paired nicking were subjected to direct and indirect fluorescence microscopies for detecting, respectively, mCherry and H2AX phosphorylated at Ser-126 (γH2AX). Prior to microscopy, the cells were incubated with a DNA damaging antitumor agent (etoposide) or with vehicle (DMSO). Nuclei were stained with DAPI. (**F**) Competition experiments comprising unedited and H2AX edited cells. Long-term cultures of cells expressing H2AX::mCherry (95% at t = 0 days) mixed with unedited cells (5% at t = 0 days) were monitored by flow cytometry. H2AX tagging was done through standard, paired breaking (DSB²), or *in trans* paired nicking (Nick²) gene editing using gRNA$^{H2AX.2}$.

of a range of indel footprints in mCherry^{+} cells obtained via standard gene editing (**Figure S8**). In contrast, untagged H2AX alleles remained intact in mCherry^{+} cells generated through *in trans* paired nicking (**Figure S8**), with the respective tagged H2AX alleles expressing the H2AX::mCherry fusion protein in the nuclei of the respective cell populations (**Figure 2E**).

For further assessing the accuracy and mutagenicity of the different gene editing strategies (**Figure 2D** and **S7**, respectively), we randomly selected mCherry^{+} clones from cultures initially exposed to the

gRNA with the fewest predicted off-target sites, i.e., gRNA$^{H2AX.2}$ (**Figure S9**). Interestingly, gRNA$^{H2AX.2}$ directs SpCas9 and SpCas9^{D10A} to cut and nick, respectively, within the codons of the previously mentioned SQ phosphorylation motif whose integrity is crucial for H2AX function (**Figures 2A** and **S9**). In this regard, it is worth noting that reduced H2AX dosages in heterozygous H2AX$^{+/-}$ knockout mice have uncovered pleiotropic haploinsufficiency phenotypes [46]. For instance, embryonic fibroblasts from these H2AX$^{+/-}$ mice present growth kinetic curves that are in-between those of wild type and homozygous H2AX$^{-/-}$ mice [46]. Thus, we next compared the fitness of human cells whose H2AX loci had been edited by either *in trans* paired nicking or DSB-dependent gene editing approaches. To this end, populations of mCherry^{+} cells were mixed with a small fraction of unmodified cells (i.e. 5%) and were subsequently monitored by flow cytometry upon serial sub-culturing rounds. Such cell competition settings demonstrated a fitness loss (i.e. growth disadvantage) specifically in cells that had undergone standard and paired breaking gene editing after SpCas9:gRNA$^{H2AX.2}$ delivery (**Figure 2F**). This loss-of-fitness phenotype correlated with the time-dependent disappearance of cells harboring H2AX indels disabling the SQ phosphorylation target motif (**Figure S10**). We also performed competition experiments in which edited cells had initially been exposed to gRNA$^{H2AX.1}$ instead of gRNA$^{H2AX.2}$. Although displaying a higher potential for off-target effects than gRNA$^{H2AX.2}$, gRNA$^{H2AX.1}$ has a lower change of disrupting the SQ protein motif (**Figure 2A** and **S9**). In this case, we observed neither the replacement of edited cells by unedited cells (**Figure S11**) nor the elimination of cells with DSB-derived H2AX indels (**Figure S12**). Thus, in contrast to a process of 'purification' from mutations at the cost of gene-edited cell loss, there was instead, gene-edited cell maintenance at the cost of a 'fixation' of mutations in the populations subjected to SpCas9:gRNA$^{H2AX.1}$ complexes (**Figures S11** and **S12**). Importantly, reminiscent of the previous sequencing of H2AX alleles in individual clones (**Figure S8**), the population-level H2AX genotyping assays further confirmed the non-disruptive character of *in trans* paired nicking by revealing the striking dominance of gene edited cells lacking H2AX mutations at both time points analysed, independently of the gRNA used (**Figures S10** and **S12**, bottom **D** panels). Taken together, these data indicate that the loss-of-fitness phenotype seen in SpCas9:gRNA$^{H2AX.2}$-treated cells (**Figure 2F**) is attributable to functional H2AX haploinsufficiency caused by NHEJ-mediated disruption of the SQ post-translational modification motif (**Figures S8** and **S10**).

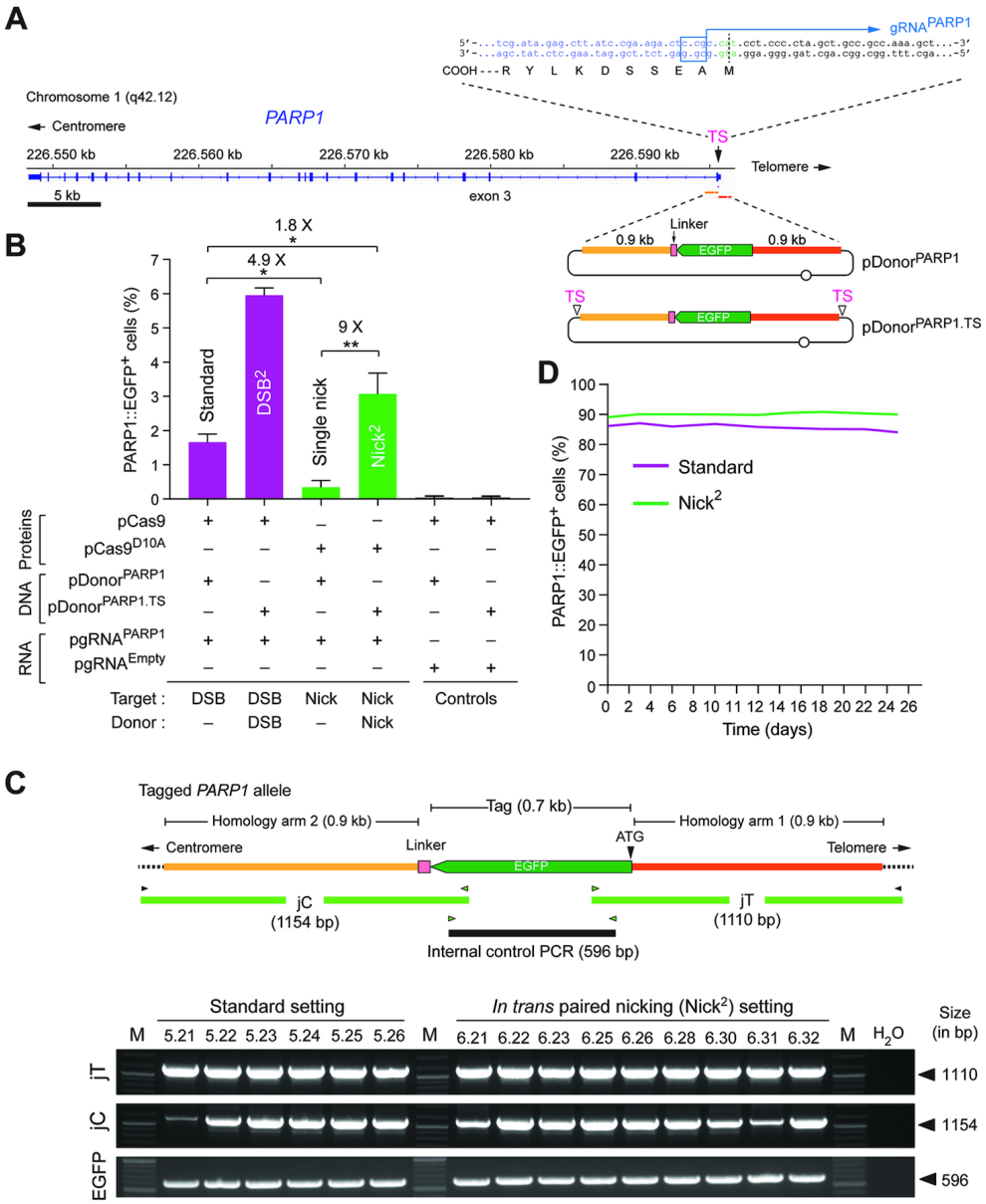


Figure 3. Homology-directed PARP1 gene editing based on cleaving or nicking CRISPR complexes. (**A**) Diagrams of PARP1 and PARP1-tailored gene editing tools. The gRNA^{PARP1} target site (TS) is indicated by the horizontal arrow and boxed nucleotides (PAM). The vertical dashed line marks the SpCas9:gRNA^{PARP1} cleaving position. The N-terminal PARP1 amino acids are drawn next to their respective codons. The donor constructs pDonor^{PARP1} and pDonor$^{PARP1.TS}$ have as targeting module PARP1 sequences ('arms of homology') flanking a EGFP tag. The latter construct has, in addition, TS sequences flanking the targeting module. (**B**) Quantification of genetically modified human cells. Flow cytometry of HeLa cell cultures co-transfected with the indicated plasmids. Data are presented as mean ± S.D. of three independent biological replicates. Significance between the indicated datasets was calculated

◂ with one-way ANOVA followed by Tukey's test for multiple comparisons; *P< 0.05; **P< 0.01. (**C**) Molecular characterization of human cells genetically modified through standard versus *in trans* paired nicking gene editing at PARP1. Top panel, Junction PCR assay for assessing PARP1 gene tagging. Amplicons diagnostic for HDR-derived centromeric and telomeric junctions between xogenous DNA and PARP1 (jC and jT, respectively) are depicted. Amplicons specific for EGFP served as internal controls (EGFP). Bottom panel, Junction PCR analysis on genomic DNA from EGFP$^+$ HeLa cell clones retrieved from cultures co-transfected with pCas9, pDonor^{PARP1} and pgRNA^{PARP1} (Standard setting) or with pCas9^{D10A}, pDonor$^{PARP1.TS}$ and pgRNA^{PARP1} (*In trans* paired nicking setting). H_2O, PCR sample containing nuclease-free water instead of genomic DNA. Lanes M, GeneRuler DNA Ladder Mix molecular weight marker. (**D**) Competition experiment involving unedited and PARP1 edited cells. Long-term cultures of HeLa cells expressing EGFP-tagged PARP1 mixed with unedited cells were monitored by flow cytometry. Green and magenta lines, EGFP$^+$ cells generated by *in trans* paired nicking and standard gene editing, respectively.

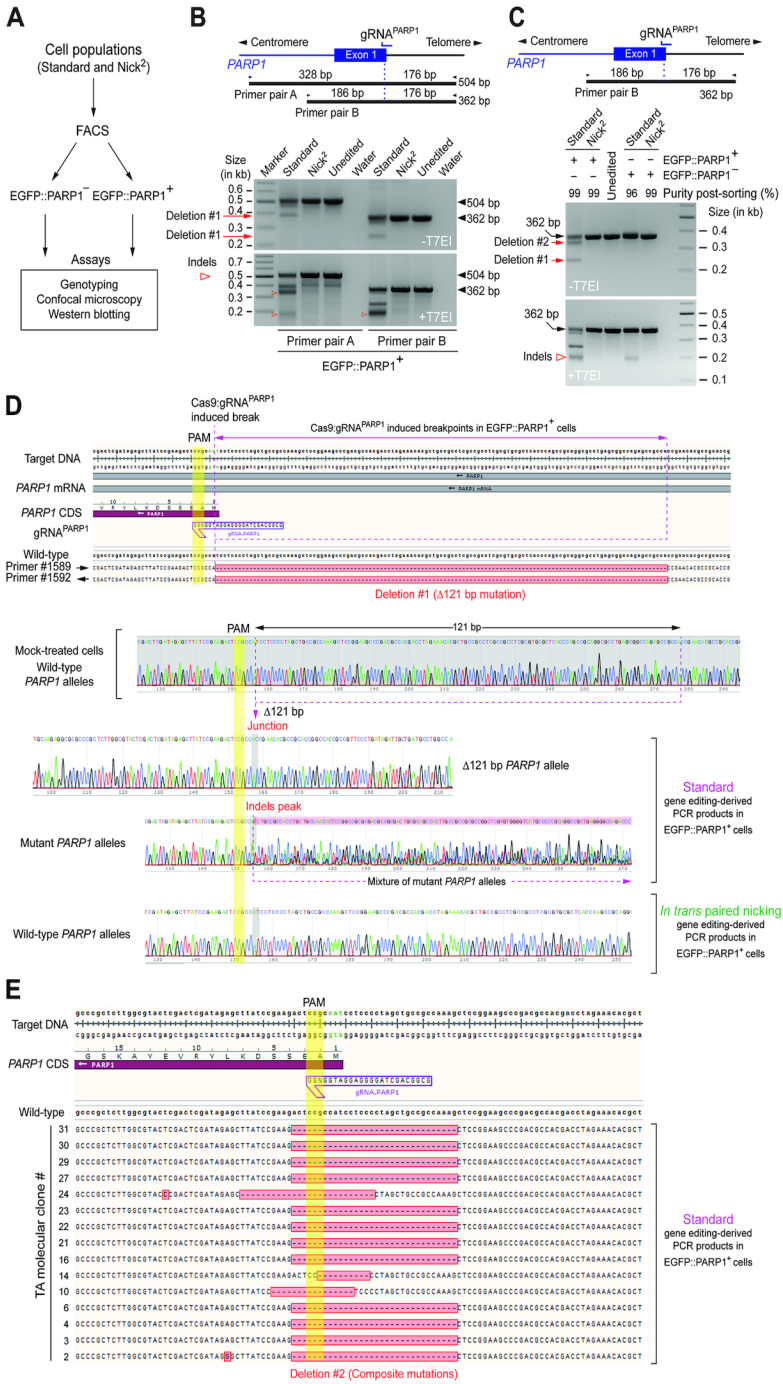
In trans paired nicking minimizes mutagenesis within coding sequences of target alleles

PARP1, like H2AX, is also involved in DNA repair; however, functional redundancies with other PARP family members are reported [47,48]. Tagging PARP1 with EGFP after delivering conventional pDonor^{PARP1} or target site-containing pDonor$^{PARP1.TS}$, together with cleaving SpCas9:gRNA^{PARP1} or nicking SpCas9^{D10A}:gRNA^{PARP1} complexes (**Figure 3A**), revealed that *in trans* paired nicking and standard gene editing led to higher frequencies of stably transfected cells than those reached by using the single nicking approach (**Figure 3B**). Importantly, junction PCR screens of randomly isolated EGFP$^+$ clones confirmed accurate DNA targeting events in cell populations subjected to *in trans* paired nicking and standard gene editing (**Figure 3C**). Moreover, cell competition experiments involving tracking mixtures of unedited and PARP1-edited cells provided no evidence for cell-fitness losses in each of the EGFP::PARP1-expressing populations (**Figure 3D**). Despite this, we sought to characterize EGFP::PARP1$^+$ and EGFP::PARP1$^-$ cell populations obtained through *in trans* paired nicking versus standard gene editing (Figure 4A). In addition to the typical small indels established after NHEJ-mediated DSB repair, the EGFP::PARP1$^+$ cell fraction generated through standard gene editing contained large PARP1 deletions (**Figure 4B** and **C**). Of note, small indels were even detected in the EGFP::PARP1$^-$ cell fraction isolated from cultures subjected to standard gene editing (**Figure 4C**). Sequence analysis of PARP1 target DNA in EGFP::PARP1$^+$ cells identified a 121-bp deletion mixed with shorter deletions of varying sizes (**Figure 4D** and **E**, respectively). These structural variants are reminiscent of those detected in

mCherry$^+$ cells that had been exposed to cleaving H2AX-specific CRISPR-SpCas9 complexes (**Figures S7** and **S8**), and further support the data indicating that targeted DSBs can trigger gross genomic alterations
[4, 10]. In contrast, PARP1 structural variants consisting of large deletions and small indels were detected neither in EGFP::PARP1$^+$ nor EGFP::PARP1$^-$ cell fractions generated through *in trans* paired nicking (**Figure 4B–D**).

Finally, dual-colour confocal microscopy showed that, regardless of the gene editing methodology, EGFP-tagged PARP1 localized properly in cell nuclei (**Figure 5A**). Tellingly, however, western blot analysis revealed that contrary to EGFP::PARP1$^+$ cells resulting from *in trans* paired nicking, EGFP::PARP1$^+$ cells derived from standard gene editing suffered a

substantial depletion of the endogenous, untagged, PARP1 protein (**Figure 5B**). This data is consistent with the high prevalence of PARP1 structural variants in EGFP::PARP1^{+} cells initially treated with pDonor^{PARP1} and SpCas9:gRNA^{PARP1} (**Figure 4B-E**).



◂ **Figure 4. Characterization of PARP1 alleles in cell populations subjected to standard versus *in trans* paired nicking gene editing.** (**A**) Overview of the experimental design. HeLa cell populations subjected to SSB-mediated *in trans* paired nicking and DSB-mediated standard gene editing were sorted in their respective EGFP::PARP1⁻ and EGFP::PARP1⁺ populations. Each of these cell fractions was next characterized at the DNA and protein levels by the indicated assays. (**B** and **C**) Examination of PARP1 mutagenesis after gene editing based on DSBs versus SSBs. Untreated and T7EI-treated PCR products spanning the gRNA^{PARP1} target site provided evidence for large deletions and small indels, respectively, in EGFP::PARP1⁺ cells generated by standard gene editing (panel B). Indels were equally detected in EGFP::PARP1⁻ cells exposed to standard gene editing (panel C). DNA species diagnostic for SpCas9:gRNA^{PARP1}-induced deletions and indels are marked with arrows and open arrowheads, respectively. (**D**) Sequence analysis of the PARP1 target region in gene edited cells. Top panel, Sanger sequencing of the low molecular weight amplicons shown in panel B (-T7EI, primer pair B) with forward and reverse primers revealing the presence of a 121-bp deletion at target sequences in EGFP::PARP1⁺ cells that underwent standard gene editing. The PARP1 proximal deletion breakpoint coincides with the predicted SpCas9:gRNA^{PARP1} cleaving position. Bottom panel, chromatograms corresponding to PARP1 alleles in EGFP::PARP1⁺ cells engineered by standard gene editing and *in trans* paired nicking. Chromatograms corresponding to wild-type PARP1 and to the 121-bp PARP1 deletion are also displayed. (**E**) Characterization of additional PARP1 deletion products. The PARP1 species with a molecular weight between unmodified and 121 bp-deleted alleles (Deletion #2) presented various mutations as determined by TA cloning and sequence analysis.

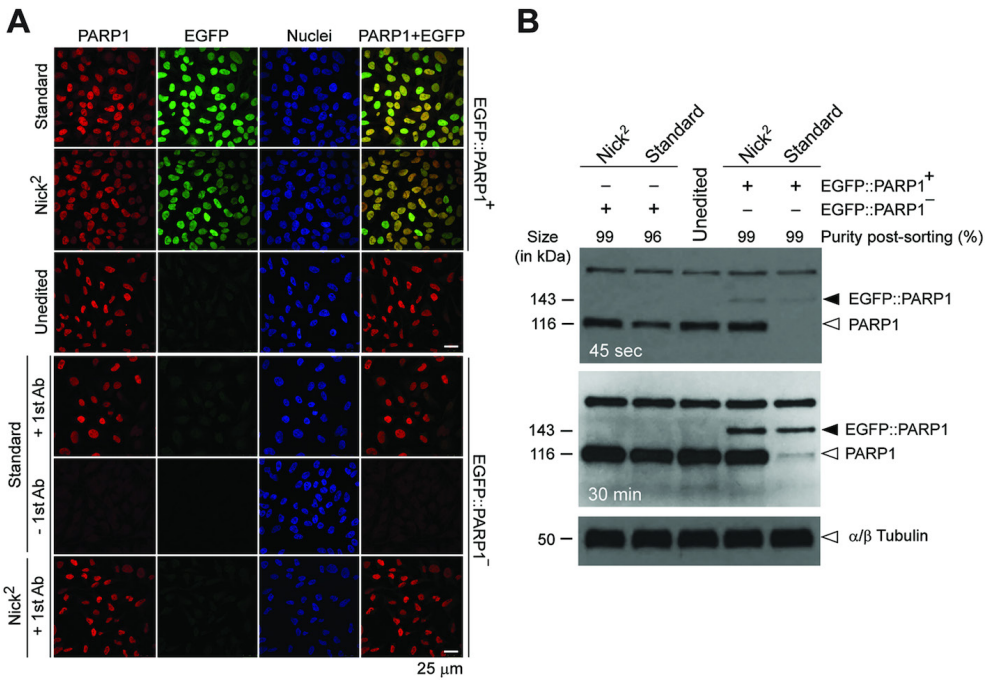


Figure 5. Examination of PARP1 protein status after gene editing triggered by DSBs versus SSBs. (**A**) Confocal microscopy analysis of HeLa cells expressing untagged and EGFP-tagged PARP1. Confocal microscopy of EGFP::PARP1⁺ and EGFP::PARP1⁻ cells confirming co-localization of PARP1 and EGFP in the nuclei of the former cell populations engineered by *in trans* paired nicking or standard gene editing. Nuclei were counterstained with DAPI. Unedited HeLa cells served as negative controls. Specimens of EGFP::PARP1⁻ cells not incubated with the primary PARP1-specific antibody (-1st Ab) provided for an additional staining control. (**B**) Western blot analysis of HeLa cells expressing untagged and EGFP-tagged PARP1. Western blotting of EGFP::PARP1⁺ and EGFP::PARP1⁻ cells exposing a striking reduction in the amounts of endogenous PARP1 antigen exclusively in EGFP::PARP1⁺ cells generated through standard DSB-dependent gene editing (open arrowhead). Properly sized EGFP::PARP1 fusion products were detected in both EGFP::PARP1⁺ cell populations (solid arrowhead). Unedited HeLa cells served as negative controls. α/β Tubulin antigens served as internal protein loading controls.

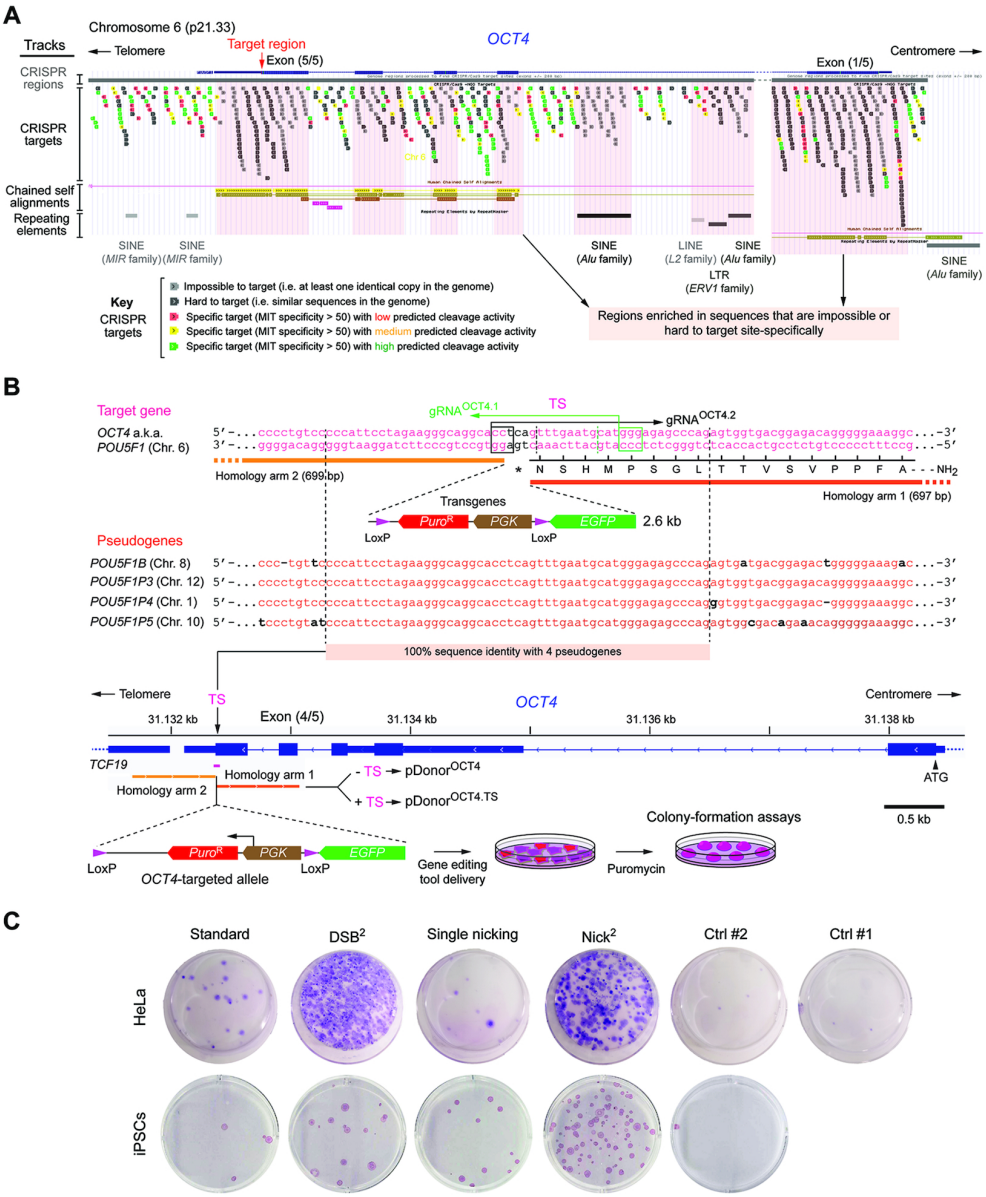


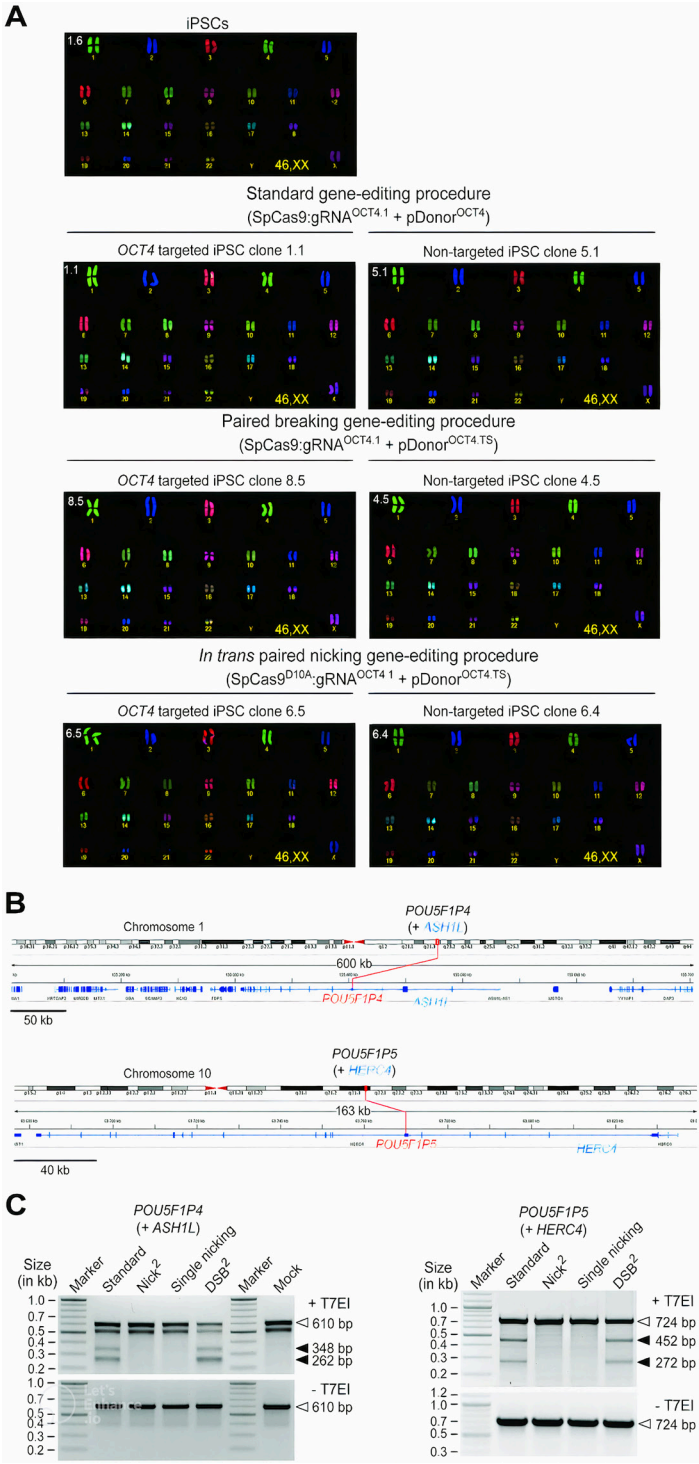
Figure 6. Homology-directed OCT4 gene editing based on cleaving or nicking CRISPR complexes. (**A**) The OCT4 genomic region. All potential S. pyogenes CRISPR-SpCas9 target sites, as defined by 20-mer spacers and canonical NGG PAMs, are colour-coded according to their predicted target site specificity and activity (CRISPR targets track). Genomic features sharing full or partial sequence identity with OCT4 are highlighted as duplications and repeats (chained self-alignments and repeating elements tracks, respectively). Tracks annotations were retrieved from the UCSC Genome Browser, Assembly GRCh38/hg38. (**B**) The OCT4 target region. The OCT4 terminal nucleotides are drawn in relation to similar sequences present in its pseudogenes and in donor plasmids pDonor^{OCT4} and pDonor$^{OCT4.TS}$. The former and latter constructs lack and contain, respectively, gRNA target sites (TS) flanking the targeting module. The target sites are indicated by horizontal arrows and boxed nucleotides (PAMs). Donor constructs are built to knock-in a floxed positive-selection cassette plus an EGFP reporter into OCT4 loci. The Cre-mediated excision of the selection cassette generates a traceable OCT4::EGFP fusion product exclusively in accurately targeted iPSCs. (**C**) OCT4 gene editing. Colony-

◄ formation assays for detecting stably transfected cells. iPSCs and HeLa cells were co-transfected with conventional pDonor^{OCT4} or target site-modified pDonor$^{OCT4.TS}$templates each mixed with constructs expressing SpCas9:gRNA$^{OCT4.1}$ or SpCas9^{D10A}:gRNA$^{OCT4.1}$. After puromycin selection, alkaline phosphatase and Giemsa staining identified genetically modified colonies of iPSCs and HeLa cells, respectively.

I*n trans* paired nicking achieves seamless editing of essential iPSC genomic sequences

The OCT4 transcription factor is essential for human embryogenesis [49] and for the genetic circuitry underpinning pluripotent stem cell states [50, 51]. For these reasons, it is a coveted gene-editing target. Yet, especially at its termini, OCT4 shares substantial homology with several of its pseudogenes (**Figure 6A** and **B**). These multiple-copy sequences make the identification of suitable gRNAs hard or impossible (**Figure 6A** and **S13**). Hence, we next sought to compare the performance of the different gene editing strategies in a challenging gene-editing model involving tagging OCT4 at its last exon using gRNAs that lack OCT4 specificity. To this end, HeLa cells and iPSCs were transfected with conventional pDonor^{OCT4} or target site-modified pDonor$^{OCT4.TS}$, each mixed with plasmids coding for SpCas9:gRNA$^{OCT4.1}$ or SpCas9^{D10A}:gRNA$^{OCT4.1}$ (**Figure 6B**). Colony-formation assays showed that, when compared to single nicking and standard gene editing approaches, *in trans* paired nicking comprising SSB formation at OCT4 and donor templates led to higher numbers of puromycin-resistant colonies regardless of the cell type (**Figure 6C**). Similar results were obtained in independent iPSC transfections in which an additional gRNA was included (**Figure S14**). Crucially, genomic DNA analysis of randomly isolated iPSC colonies readily revealed that *in trans* paired nicking achieved a much higher precision in OCT4 targeting than the DSB-dependent approaches (**Figure S15A** and **S15B**). Multicolour FISH-based molecular karyotyping (COBRA-FISH) revealed that neither iPSCs subjected to *in trans* paired nicking nor iPSCs exposed to the DSB-dependent protocols harboured overt chromosomal rearrangements (n = 6; **Figure 7A**). Possibly, this outcome is the result of a strong selection against iPSCs that had initially been exposed to multiple DSBs. Related with this, robust mutagenesis at gRNA$^{OCT4.1}$ target sites located in off-target chromosomal locations (**Figure 7B**) was readily detected in iPSC populations subjected to DSB-dependent gene editing (**Figure 7C**). The fact that gRNA target sequences in OCT4 pseudogenes overlap with coding cellular genes, further compounds the genotype of SpCas9:gRNA$^{OCT4.1}$-treated cells (**Figure 7B** and **C**).The generation of DSBs at OCT4 pseudogenes (**Figure 7C**) raises the possibility for the insertion of OCT4-targeting donor DNA at these off-target genomic positions due to the partial homology between them and donor DNA (**Figure S15C**). A junction PCR assay devised to investigate this possibility did not detect donor DNA insertions at OCT4 pseudogenes in puromycin-resistant iPSC clones (n = 22) randomly isolated from cultures subjected to *in trans* paired nicking (**Figure S15C** and **S15D**).

Previous experiments in pluripotent stem cells (i.e. human embryonic stem cells and iPSCs) revealed that *in trans* paired nicking using SpCas9^{D10A}:gAAVS1 complexes yields higher gene targeting frequencies than those achieved by standard gene-editing involving


A
iPSCs
1.6
46,XX
Standard gene-editing procedure
(SpCas9:gRNA^OCT4.1 + pDonor^OCT4)
OCT4 targeted iPSC clone 1.1
Non-targeted iPSC clone 5.1
1.1
5.1
46,XX
Paired breaking gene-editing procedure
(SpCas9:gRNA^OCT4.1 + pDonor^OCT4.TS)
OCT4 targeted iPSC clone 8.5
Non-targeted iPSC clone 4.5
8.5
4.5
In trans paired nicking gene-editing procedure
(SpCas9^D10A:gRNA^OCT4.1 + pDonor^OCT4.TS)
OCT4 targeted iPSC clone 6.5
Non-targeted iPSC clone 6.4
6.5
6.4
B
Chromosome 1
POU5F1P4
(+ ASH1L)
600 kb
50 kb
Chromosome 10
POU5F1P5
(+ HERC4)
163 kb
40 kb
C
POU5F1P4
(+ ASH1L)
Size
(in kb)
Marker
Standard
Nick2
Single nicking
DSB2
Marker
Mock
+ T7EI
610 bp
348 bp
262 bp
- T7EI
610 bp
POU5F1P5
(+ HERC4)
Size
(in kb)
Marker
Standard
Nick2
Single nicking
DSB2
+ T7EI
724 bp
452 bp
272 bp
- T7EI
724 bp

◂ **Figure 7. Characterization of iPSCs after OCT4 gene editing using cleaving versus nicking CRISPR complexes.** (**A**) Karyotyping of genetically modified iPSC clones. Overview of COBRA-FISH analysis of parental iPSCs and individual targeted and non-targeted clones showing a seemingly normal diploid karyotype (46,XX). Each clone was isolated after adding puromycin to iPSC populations subjected to the indicated gene editing strategies. (**B**) Chromosomal and genomic coordinates of POU5F1P4 and POU5F1P5. The former and latter OCT4 pseudogenes overlap with nucleotide sequences from ASH1L (ASH1-like histone lysine methyltransferase) and HERC4 (HECT and RLD domain containing E3 ubiquitin protein ligase 4), respectively. ASH1L codes for a member of the trithorax group of transcriptional activators and is ubiquitously expressed in over 25 tissues; HERC4 belongs to the HERC family of ubiquitin ligases and is ubiquitously expressed in over 25 tissues. As a result, indels generated at OCT4 pseudogenes inevitably create additional genotypic complexity in target cell populations whose, cell type-specific, phenotypic consequences are difficult to predict and assess. (**C**) Comparing genome-disrupting events at OCT4gRNA target sites located at off-target chromosomal positions. T7EI-based genotyping assays were performed on DNA from puromycin-resistant iPSC populations expanded after OCT4-targeting experiments involving the indicated gene editing procedures. T7EI-specific products diagnostic for mutant alleles generated by NHEJ-mediated DSB repair are pinpointed by closed arrowheads; products corresponding to intact alleles are instead indicated by open arrowheads in untreated and T7EI-treated samples. Marker, GeneRuler DNA Ladder Mix molecular weight marker.

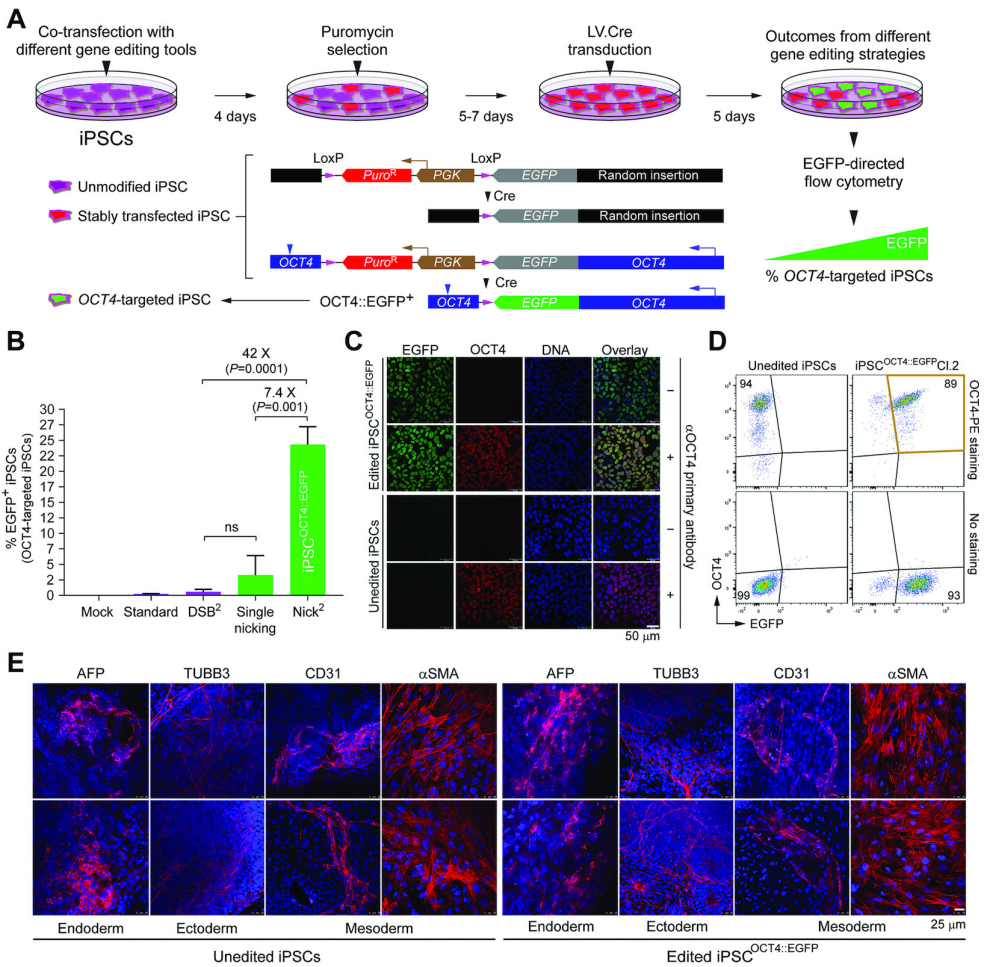


Figure 8. Comparing the accuracy of OCT4 gene editing after delivering cleaving versus nicking CRISPR complexes. (**A**) Genetic assay for determining OCT4 targeting frequencies. iPSCs co-transfected with plasmid combinations corresponding to each of the four different gene editing strategies, were sequentially exposed to puromycin and Cre recombinase. OCT4-targeted iPSCs

◄ expressing Cre-derived OCT4::EGFP fusion products report accurate genome-modifying events. The Cre recombinase was delivered by transducing iPSCs with lentiviral vector LV.Cre at a multiplicity-of-infection of 10 physical particles per cell. (**B**) Comparing the performance of OCT4 gene editing strategies in iPSCs. The frequencies of OCT4-targeted iPSCs expressing OCT4::EGFP were determined by EGFP-directed flow cytometry. Data are shown as mean ± S.D. of independent biological replicates. Significance was calculated with two-tailed Student's t tests (n = 3); ns, non-significant. (**C**) Confocal microscopy analysis of OCT4 edited iPSCs. OCT4::EGFP-expressing iPSCs engineered through in trans paired nicking and Cre delivery ($iPSC^{OCT4::EGFP}$) were subjected to indirect and direct fluorescence microscopies for detecting OCT4 and EGFP, respectively. Nuclei were stained with DAPI. Nuclear localization of OCT4::EGFP is highlighted by the merging of the three fluorescence signals. Unedited iPSCs served as negative controls. iPSC and $iPSC^{OCT4::EGFP}$ populations that were not incubated with the OCT4-specific primary antibody served as staining controls. (**D**) Flow cytometric analysis of OCT4 edited iPSCs. Flow cytometry of iPSC clone 2 isolated from an $iPSC^{OCT4::EGFP}$ population confirming OCT4 and EGFP co-labelling (coloured quadrant). Unedited iPSCs served as controls. Cultures of parental iPSCs and $iPSC^{OCT4::EGFP}$ clone 2 that were not exposed to the PE-conjugated OCT4 antibody were used as staining controls. (**E**) Testing multi-lineage differentiation capacity of iPSC populations expressing OCT4::EGFP. Immunofluorescence microscopy analysis of $iPSC^{OCT4::EGFP}$ cells differentiated into cellular lineages representative of endoderm, ectoderm and mesoderm. Unedited iPSCs served as differentiation controls. Markers for each germ layer are indicated. Nuclei were stained with DAPI.

SpCas9:gAAVS1 [12]. Similar AAVS1 gene targeting experiments performed in the iPSC line used in the current study were consistent with these earlier findings (**Figure S16**). To investigate whether chromosomal rearrangements were detectable in these iPSCs soon after their exposure to CRISPR complexes, we performed orthogonal HTGTS analysis on cell populations exposed to SaCas9:Sa–gRAG1.1 alone or together with SpCas9:gAAVS1 or $SpCas9^{D10A}$:gAAVS1 complexes (**Figure S17**). The orthogonal HTGTS assay detected translocations exclusively in iPSCs that had been co-treated with SaCas9:Sa-gRAG1.1 and SpCas9:gAAVS1 nucleases (**Figures S18** and **S19**). When compared with the orthogonal HTGTS experiments performed in aneuploid HEK293T cells (**Figure 1C**, **S3** and **S4**), the overall lower frequencies of translocations detected in iPSCs might have resulted from their diploid character and/or lower exposure to CRISPR complexes (compare **Figures S2** with **Figure S17**). Crucially, in line with the orthogonal HTGTS experiments in HEK293T cells, this data support that $SpCas9^{D10A}$ nickases trigger less chromosomal rearrangements than their SpCas9 counterparts, in this case, in diploid iPSCs (**Figures S18** and **S19**).

To complement the characterization of gene-edited iPSCs (**Figure 7** and **Figure S15**), we set-up a quantitative specificity assay in which Cre-mediated OCT4::EGFP assembly reports on precise gene editing in iPSCs (**Figure 8A**). The results from this functional genetic assay confirmed the strikingly different OCT4 targeting levels achieved by nicking versus cleaving CRISPR complexes. In particular, in contrast to the single nick-dependent and DSB-dependent approaches, induction of SSBs at acceptor and donor DNA results in efficient targeted gene editing in viable iPSCs (**Figure 8B**). Our results suggest that exposing iPSCs to nicking as opposed to cleaving CRISPR complexes overcomes a strong negative selection against OCT4-edited iPSCs. These results are in agreement with previous experiments showing that even very few DSBs, including those made by SpCas9 nucleases, can significantly reduce the division and survival rates of PSCs [12, 52–54].

Finally, dual-colour confocal microscopy and flow cytometry analyses confirmed proper EGFP tagging of the endogenous OCT4 protein in iPSCs subjected to *in trans* paired nicking, at both the population and clonal levels (**Figure 8C** and **D**, respectively). Importantly, these

OCT4::EGFP-expressing iPSCs were equally capable of differentiating along the three embryonic germ layers (**Figure 8E** and **Figure S20**).

In conclusion, unwarranted genotypes and deleterious phenotypic traits created by CRISPR-SpCas9 nucleases during gene knock-in procedures are mostly avoided by using *in trans* paired nicking genome editing.

Discussion

There are some concerns regarding the application of genome editing technologies. This is especially so when these applications are directed towards biotechnologies and genetic therapies [55]. In part these concerns stem from the fact that, regardless of their specificity, programmable nucleases generate DSBs that are prone to large-scale and small-scale mutagenic events [4–10]. In this regard, programmable nuclease-induced DSBs are particularly problematic, hence avoided, at multiple-copy sequences and/or at sequences needed for proper cell functioning or overall viability. As corollary, DSB-dependent genome editing substantially limits the editable genome. Moreover, in mammalian diploid cells, nuclease-induced homologous chromosome rearrangements [10] and allelic mutations potentiate cell transformation events and gene-dose unbalances, respectively. Equally insidious are the recent findings that DSB-induced nonsense mutations can trigger transcriptional compensatory mechanisms that further confound genotype-phenotype associations [56–58].

We report that concomitant SSB formation at target and donor DNA by CRISPR-SpCas9 nickases elicits accurate and non-disruptive gene editing, including at loci associated with haploinsufficiency and essentiality. This DSB-free *in trans* paired nicking approach prevented the loss of gene-edited cells due to the disruption of a functional protein motif or a pluripotency supporting gene in iPSCs. The observed difficulty in isolating iPSCs edited at OCT4 after CRISPR-SpCas9 delivery is in line with the essentiality of this gene in safeguarding stem cell phenotypes [49–51] and with earlier experiments showing that gene targeting frequencies at OCT4 are very low. Indeed, gene editing of iPSCs using TALENs and the herein used pDonor^{OCT4} construct, did not yield any correctly targeted clone (0/48) [28]. In another study, gene editing of human embryonic stem cells deploying SpCas9 and donor templates containing the same 'homology arms' of pDonor^{OCT4}, resulted in only 8 correctly targeted clones (8/288) [59]. In contrast to these studies, viable and correctly targeted iPSC clones were readily isolated after targeting OCT4 with pDonor$^{OCT4.TS}$ and SpCas9^{D10A} (21/22) (**Figure S15B**). Importantly, *in trans* paired nicking gene editing introduces a low mutagenic load into target cell populations by minimizing NHEJ-mediated chromosomal disruption of allelic and non-allelic target sequences, such as those in OCT4 and its pseudogenes, respectively. These multiple-copy gRNA target sites, are likely to have exacerbated the difficulty in isolating OCT4-targeted iPSCs after SpCas9 delivery (**Figure 8B** and **Figure S15B**) as pluripotent stem cells are particularly prone to DSB-induced cell cycle arrest and apoptosis [12, 52–54]. There are other experimental data linking detrimental genome editing

outcomes to target sequences associated with copy number variations. In particular, genome-wide CRISPR-SpCas9 library screens have demonstrated that DSBs mapping in amplified genomic regions create false-positive hits of gene essentiality in cancer cell lines [60, 61].

Notwithstanding the fact that nicking CRISPR complexes are significantly less mutagenic than their cleaving counterparts at both target and off-target sites, they can nonetheless trigger DNA disruptions if, for example, an advancing replication fork collapses after hitting the SSB product [42]. In the present work, by using orthogonal HTGTS assays, we have provided experimental evidence for such events in mammalian cells (**Figure 1C** and **Figures S3** and **S4**). These events should be most problematic at off-target sites. In this regard, it will be worth investigating whether *in trans* paired nicking is amenable to RNA-guided nickases built on high-specificity Cas9 scaffolds [62].

Although the OCT4 edited iPSC clones analysed lacked donor DNA insertions at SSB-susceptible OCT4 pseudogenes (**Figure S15D**), unwanted knock-ins at genomic regions exhibiting high homology with donor DNA constitute a possible limitation of *in trans* paired nicking. Therefore, whenever possible, this risk should be minimized by avoiding SSB formation at such potential off-target regions and/or reducing the extent of homology between them and donor DNA [63]. Conversely, assuring SSB formation at donor DNA and multiple-copy homologous sequences might offer the prospect for co-editing these recurrent regions in the genome without attendant large-scale chromosomal mutations and rearrangements.

In conclusion, HDR-mediated gene editing through *in trans* paired nicking offers high specificity and low mutagenicity, which, as a result, mostly preserves cellular genotypes and phenotypes. Moreover, the coordinated nicking of donor and acceptor HDR templates boosts the versatility of CRISPR-based gene editing by substantially enlarging the fraction of candidate gRNAs that can become operational, regardless of their a priori specificity profiles. The seamless and scarless character of *in trans* paired nicking should be particularly beneficial in instances in which precise and predictable genetic interventions are crucial. Examples include modelling or rescuing disease traits in stem cells [64] and functionally dissecting genomic sequences by multiplexed knock-in of donor DNA libraries [65]. Finally, *in trans* paired nicking might expand the 'editable genome' to different types of repetitive elements shedding light on this large and variegated portion of the functionally unknown genomic 'dark matter' [66].

Materials and Methods

Cells

Human cervix carcinoma HeLa cells and human embryonic kidney 293T (HEK293T) cells (both from American Type Culture Collection) were cultured in Dulbecco's modified Eagle's medium (DMEM; ThermoFisher Scientific; Cat. No.: 41966029) supplemented with 5% (v/v) and 10% (v/v), respectively, fetal bovine serum ultra-low endotoxin (FBS; biowest; Cat. No.: S1860500). The HeLa cells, authenticated before by karyotyping analysis [11], were

used for gene editing experiments. The HEK293T cells were used for assembling lentiviral vector LV.Cre particles and orthogonal HTGTS analyses. The generation and characterization of the human induced pluripotent stem cells (iPSCs) used in this work (LUMC0020iCTRL) were detailed elsewhere [27]. In the current study, these cells were further characterized by COBRA-FISH karyotyping. The iPSCs were cultured in feeder-free Essential 8 Medium (E8; ThermoFisher Scientific; Cat. No.: A1517001) supplemented with 25 U ml^{-1} penicillin and 25 μg ml^{-1} of streptomycin (ThermoFisher Scientific; Cat. No.: 15140122). The iPSCs were kept in wells of six-well plates (Greiner Bio-One; Cat. No.: 662160) coated for 1 h at room with Vitronectin Recombinant Human Protein (VTN-N; ThermoFisher Scientific; Cat. No.: A14700) diluted 1:100 to a final concentration of 5 ng ml^{-1} in Dulbecco›s phosphate-buffered saline, no calcium, no magnesium (DPBS; ThermoFisher Scientific; Cat. No.: 14190094). When ready for sub-culturing, to let cell-cell dissociation occur, the iPSCs were first washed with DPBS solution and then incubated with 0.5 mM ethylenediaminetetraacetic acid (EDTA; Invitrogen Cat. No.: 15575020) in DPBS at 37°C and room temperature for 4 and 1 min, respectively. After the removal of the EDTA solution, the cells were seeded in new wells of 24-well plates coated with VTN-N and containing E8 medium supplemented with a 1:200 dilution of RevitaCell (ThermoFisher Scientific; Cat. No.: A2644501). The cells used in this study were mycoplasma free and were kept at 37°C in a humidified-air atmosphere with 5% CO_2 (iPSCs) or 10% CO_2 (HeLa and HEK293T cells).

Recombinant DNA

The expression plasmids AU26_pCAG.Cas9 and AU28_pCAG.Cas9^{D10A} encoding cleaving SpCas9 and nicking SpCas9^{D10A} enzymes, respectively, have been described previously [12]. The control plasmid gRNA_Cloning Vector (Addgene #41824) and the OCT4-targeting donor construct eGFP-PGK-Puro (Addgene #31937), herein named pgRNA^{Empty} and pDonor^{OCT4}, respectively, were also described before [20, 28]. The annotated maps and nucleotide sequences of donor constructs AX74_pDonor$^{OCT4.TS}$, AX66_pDonor$^{OCT4.1TS}$, AZ44_pDonor^{H2AX}, AZ25_pDonor$^{H2AX.TS}$, AW77_pDonor^{PARP1} and AW69_pDonor$^{PARP1.TS}$ are available in pages 1–14 of the **Supplementary Information**. The annotated maps and nucleotide sequences of the S. pyogenes gRNA-expressing plasmids AZ34_pgRNA$^{H2AX.1}$, AZ35_pgRNA$^{H2AX.2}$, AM70_pgRNA^{PARP1}, AX33_pgRNA$^{OCT4.1}$, AX34_pgRNA$^{OCT4.2}$ are available in pages 15–24 of the **Supplementary Information**. The annotated map and nucleotide sequence of the Cre-expressing lentiviral vector construct BC17_pLV.Cre is available in pages 25–27 of the **Supplementary Information**.

The constructs used in the experiments for identifying CRISPR-SaCas9 nucleases inducing HTGTS bait DSBs at RAG1 were BA15_pCAG.SaCas9.rBGpA [29], AV85_pSa-gRAG1.1, AV86_pSa-gRAG1.2, AV87_pSa-gRAG1.3, AP65_pSa-gAAVS1. With the exception of BA15_pCAG.SaCas9.rBGpA [29], all these constructs are described in pages 28–33 of the **Supplementary Information**. The plasmid BPK2660 (Addgene #70709) served as a negative control as it encodes an irrelevant, non-targeting, Staphylococcus aureus gRNA,

herein named Sa-gNT [30]. Moreover, after BsmBI digestion, BPK2660 also served as an isogenic cloning vector for the insertion of annealed oligonucleotides corresponding to the spacers of S. aureus gRNAs; Sa-gRAG1.1, Sa-gRAG1.2, Sa-gRAG1.3 and Sa-gAAVS1.

Plasmids encoding S. aureus CRISPR components used for inducing universal HTGTS bait DSBs (i.e. BA15_pCAG.SaCas9.rBGpA and AV85_pSa-gRAG1.1), were combined with constructs AV62_pCAG.Cas9.rBGpA, AB65_pCAG.Cas9^{D10A}.rBGpA and gRNA_AAVS1-T2 [20] expressing Streptococcus pyogenes CRISPR elements for triggering test HTGTS prey DNA lesions in the form of AAVS1-targeted DSBs or SSBs. The latter plasmid (Addgene #41818) encodes an AAVS1-targeted gRNA, herein dubbed gAAVS1. The annotated maps and nucleotide sequences of AV62_pCAG.Cas9.rBGpA and AB65_pCAG.Cas9^{D10A}.rBGpA are described in pages 34–39 of the **Supplementary Information**. The full sequences and annotated maps of the plasmids applied in the AAVS1 gene targeting experiments; AV15_pCAG.Cas9.gRNA^{S1}, AV44_pCAG.Cas9^{D10A}.gRNA^{S1}, AV13_pCAG.Cas9.gRNA^{NT}, AV11_pDonor.EP^{S1} (Addgene #100296) and AV09_pDonor.EP$^{S1.TS}$ (Addgene #100297) are available elsewhere [12].

HeLa and HEK293T cell transfections

HeLa and HEK293T cells were seeded in the tissue culture vessels indicated in **Tables S1–S6**. The next day, transfections started by adding a 1 mg ml^{-1} 25 kDa linear polyethyleneimine (PEI, Polysciences) solution (pH 7.4) to each plasmid mixture diluted in 50 µl of 150 mM NaCl (Merck). The cell numbers, the amounts of PEI and DNA (in ng) as well as the compositions of each of the DNA mixtures corresponding to the different transfection reactions are specified in **Tables S1–S6**. After the addition of PEI, the transfection reactions were immediately and vigorously vortexed for 10 s, after which, DNA-PEI complexes were allowed to form for 15 min at room temperature. The resulting DNA-PEI complexes were subsequently added directly into the culture media of the target cells and, after 6 h, the transfection media were substituted by regular culture media. Whenever appropriate, reporter-directed flow cytometry was performed at 3 days post-transfection to determine the transfection efficiencies. In the gene targeting experiments, cell populations were then sub-cultured for at least 2 weeks to eliminate episomal donor DNA templates, after which, reporter-directed flow cytometry was used to quantify the frequencies of stably transfected cells.

Transfections of human iPSCs

The iPSCs were first seeded in wells of 24-well plates (Greiner Bio-One) that had been previously coated with VTN-N (ThermoFisher Scientific) as indicated above. The next day, the iPSC culture media were refreshed at least 2 h prior to transfection. Transfections were initiated by adding the appropriate plasmid mixtures together with Lipofectamine Stem Transfection Reagent (ThermoFisher Scientific, Cat. No.: STEM00003) to 50 µl of Opti-MEM medium (Gibco; Cat. No.: 31985-047) in 1.5-ml sterile Eppendorf tubes (**Tables S7** and **S8**). After mixing by pipetting, the transfection reactions were incubated at room temperature

for 10 min and were then added into the culture media of the target iPSCs (**Tables S7** and **S8**). The media were replaced 24 h later and, at 2–3 days post-transfection, the iPSCs were transferred into a new culture well and were subsequently expanded in wells of 6-well plates (Greiner Bio-One) for 5–7 days in the presence of 0.5 μg ml^{-1} puromycin in StemFlex Medium (ThermoFisher Scientific, Cat. No.: A3349401) containing 25 U ml^{-1} penicillin and 25 μg ml^{-1} of streptomycin. Parallel cultures of mock-transfected iPSCs served as negative controls. At the end of the selection period, puromycin-resistant iPSC colonies were identified by using the leukocyte alkaline phosphatase kit (Sigma-Aldrich; Cat. No.: 86R-1KT) for detecting enzymatic activity from the pluripotency marker alkaline phosphatase. Cultures of puromycin-resistant iPSC populations and individual randomly selected iPSC colonies were also expanded, collected and cryopreserved for further analyses.

The iPSC genomic DNA samples used for orthogonal HTGTS analyses were generated by nucleofecting iPSCs with constructs expressing SaCas9:Sa-gRAG1.1 and SpCas9:gAAVS1 or SaCas9:Sa-gRAG1.1 and $SpCas9^{D10A}$:gAAVS1. Nucleofection of iPSCs with plasmids expressing only the SaCas9:Sa-gRAG1.1 complexes needed for generating bait DSBs served as an orthogonal HTGTS assay control (**Table S9**). The iPSC nucleofections were performed in a Nucleofector 2b-device (Lonza) using Amaxa Human Stem Cell Nucleofector Kit 2 (Lonza; Cat. No.: VPH-5022). A total amount of 8 μg of DNA diluted in 10 μl of Milli-Q water were added to 100 μl of nucleofection buffer containing 2×10^6 iPSCs. After gentle mixing, the cell suspensions were transferred to the device-tailored cuvettes and immediately subjected to the nucleofection program B-016, selected for human embryonic stem cells. Next, the iPSCs were transferred to wells of 6-well plates (Greiner Bio-One) containing 2 ml of pre-warmed E8 medium (ThermoFisher Scientific; Cat. No.: A1517001) supplemented with a 1:100 dilution of RevitaCell (ThermoFisher Scientific; Cat. No.: A2644501). After an overnight incubation period, the culture medium was replenished and, at 3 days post-nucleofection, genomic DNA was extracted. Finally, genomic DNA samples were subjected to T7 endonuclease I (T7EI)-based genotyping assays directed at RAG1 and AAVS1 alleles and, subsequently, orthogonal HTGTS analyses was performed as described below.

Orthogonal HTGTS sample preparation

Transfections for generating genomic DNA samples for orthogonal HTGTS analyses were carried out in HEK293T cells and iPSCs (**Tables S1** and **S9**, respectively). The genomic DNA was isolated at 36 h post-transfection as described before [31]. In brief, the cells were collected by centrifugation and resuspended in lysis buffer consisting of 200 mM NaCl, 10 mM Tris–HCl pH 7.4, 2 mM EDTA pH 8.0, 0.2% (w/v), sodium dodecyl sulphate (SDS) and freshly added proteinase K (Thermo Fisher Scientific; Cat. No.: #EO0491) at a final concentration of 200 ng ml^{-1}. After an overnight incubation period at 56°C, the DNA was precipitated by adding isopropanol (1:1) and immediate mixing of the aqueous and organic phases. Next, the DNA was transferred to a new Eppendorf tube containing 1 ml of 70% (v/v) ethanol. The DNA was next pelleted by centrifugation at 13 000 × g for 5 min at 4°C, and

dissolved in TE buffer (10 mM Tris–HCl pH 8.0; 1 mM EDTA pH 8.0) for at least 2 h at 56°C.

Before orthogonal HTGTS analyses, genomic DNA samples were subjected to T7EI-based genotyping assays. These assays permitted assessing bait and prey chromosomal DNA breaks at RAG1 and AAVS1 alleles, respectively, in HEK293T and iPSC cell populations. To this end, the RAG1 and AAVS1 target regions were PCR-amplified with KOD Hot Start DNA Polymerase (Merck Millipore; Cat. No.: 71086–3) and GoTaq G2 Flexi DNA Polymerase (Promega; Cat. No.: M7805) using the PCR mixtures indicated in **Tables S10** and **S11**, respectively. The PCR primers and cycling parameters used to amplify RAG1 and AAVS1 DNA are specified in **Tables S12** and **S13**, respectively. Indels generated by NHEJ-mediated DSB repair were detected by exposing RAG1 and AAVS1 amplicons to T7EI (Biolabs; Cat. No.: M0302L) as below indicated.

Transfections for selecting Sa-gRNAs inducing universal HTGTS bait DSBs at RAG1 were performed on HeLa cells and HEK293T cells (**Table S2**). At 3 days post-transfection, indel formation at the target gene was assessed by T7EI-based genotyping assays as below indicated. To this end, genomic DNA was extracted by using the DNeasy Blood & Tissue kit (Qiagen; Cat. No.: 69506) according to the manufacturer's instructions. Next, the RAG1 target region in HeLa and HEK293T cells was PCR-amplified with KOD Hot Start DNA Polymerase (Merck Millipore). The PCR mixtures, primers and cycling parameters are indicated in **Tables S10**, S12 and S13, respectively. The construct expressing S. aureus gRNA Sa-gRAG1.1 was selected to induce bait DSBs at RAG1 in orthogonal HTGTS experiments in HEK293T cells and iPSCs (**Tables S1** and **S9**, respectively).

Gene targeting and gene tagging experiments

Transfections for AAVS1 gene targeting experiments were done in HEK293T cells and iPSCs (**Tables S3** and **S8**, respectively) using as donors plasmids AV11_pDonor.EP^{S1} (Addgene #100296) and AV09_pDonor.EP$^{S1.TS}$ (Addgene #100297) [12]. The former differs from the latter in that is has its targeting module flanked by gAAVS1 target sites. The targeting modules of these donors consist of sequences homologous to the AAVS1 locus framing expression units encoding both puromycin N-acetyltransferase and EGFP. In these experiments, these donors were combined with plasmids AV15_pCAG.Cas9.gRNA^{S1}, AV44_pCAG.Cas9^{D10A}.gRNA^{S1} and AV13_pCAG.Cas9.gRNA^{NT} which co-express SpCas9 proteins and gRNAs [12]. At 3 days post-transfection, the transfection efficiencies were determined by EGFP-directed flow cytometry. Subsequently, the cells were sub-cultured for 14 days, for the removal of episomal donor templates, after which stable transfection frequencies were established via EGFP-directed flow cytometry. In addition, stably transfected cells present in long-term HEK293T cell cultures were selected for by incubation with 3 μg ml^{-1} of puromycin (InvivoGen; Cat. No.: 58582) during 9 days. The distribution of EGFP expression levels in the resulting puromycin-resistant populations was assessed by EGFP-directed flow cytometry.

Transfections for tagging H2AX and PARP1 proteins were performed on HeLa cells (**Tables S4** and **S5**, respectively).

Transfections of HeLa cells for OCT4 gene targeting (**Table S6**), were assessed by colony-formation assays. To this end, at approximately 2 weeks post-transfection, the cells were counted and seeded at a density of 10^5 cells per 60 mm × 15 mm culture dishes (Greiner Bio-One; Cat. No.: 628160). After a 17-day exposure period to 1 μg ml^{-1} of puromycin (InvivoGen), HeLa cell colonies were identified by Giemsa staining.

Determining genome-wide off-target effects by orthogonal HTGTS analyses

The orthogonal HTGTS analyses were done in a blind fashion on genomic DNA samples isolated from HEK293T cells and iPSCs. Genomic DNA samples from the former and latter cell types were generated as described above using the transfection mixtures specified in **Tables S1** and **S9**, respectively. The reagents and procedures for HTGTS analysis have been detailed elsewhere [10, 31]. In brief, 25 μg of genomic DNA was used for each sample. Samples were sheared using a Bioruptor (Diagenode) with a circulating temperature of 4°C, on a low power setting: 2 × 30 s pulses interspaced by a 60 s cool down period. The biotinylated RAG1A/B – F1 primer [10] was used for LAM-PCR [31], and ssDNA products were enriched on streptavidin-coated magnetic beads (ThermoFisher Scientific; Cat. No.: 65002) prior to ligation of bridge adapters [10, 31]. Barcoded RAG1A/B – F2 I5 and AP2 I7 primers [10] were used for the nested PCR. P5–I5 and P7–I7 primers [31] were used in the final PCR. The resulting amplicons between 500 bp to 1 kb were separated and gel extracted (Qiagen; Cat. No.: 28706). Phusion polymerase (ThermoFisher Scientific; Cat. No.: F530L) was used in all PCR steps and the blocking enzyme step was omitted. HTGTS libraries were run on a Bioanalyzer (Agilent 2100) prior to MiSeq 2 × 250 bp sequencing (Illumina; Cat. No.: MS-102-2003). Pooled sequence reads were demultiplexed and trimmed according to predetermined molecular barcodes and adapter sequences; each library was subjected to bait/prey alignments (hg19), filtering, and post-pipeline analysis as described [31]. Significantly enriched translocation sites in sequence read libraries from individual experiments were identified using MACS2 as previously described [10]. Translocation hotspots were called if such enriched translocation sites were statistically significant in the majority of the independent replicate experiments.

Characterization of genome-modifying events by clonal analysis

$EGFP^+$ and $mCherry^+$ HeLa cells generated after PARP1 and H2AX gene editing, respectively, were sorted at 2–3 weeks post-transfection as single cells or as whole populations with the aid of a BD FACSAria III flow cytometer (BD Biosciences). The single cell-derived clones were seeded in wells of 96-well plates (Greiner Bio-One) and were grown in HeLa culture medium supplemented with 50 U ml^{-1} penicillin, 50 μg ml^{-1} of streptomycin and, to increase their cloning efficiency, 50 μM α-thioglycerol and 20 nM bathocuproine disulfonate (both from Sigma-Aldrich) [32]. Next, conventional and junction PCR analyses were performed on chromosomal DNA from individual clones, each of which representing

a specific genome-modifying event. The PCR screening of the mCherry^{+} HeLa cell clones was done with the GoTaq G2 Flexi DNA Polymerase system (Promega; Cat. No.: M7808) using the PCR mixtures and cycling parameters indicated in **Tables S14** and **S15**, respectively. The screening of the EGFP^{+} HeLa cell clones was performed with the reagents and protocol provided in the Phire Tissue Direct PCR Master Mix kit (ThermoFisher Scientific, Cat. No.: F170L). The PCR mixtures and cycling parameters used for these analyses are also indicated in the **Tables S14** and **S15**, respectively.

Characterization of genome-modifying events in iPSCs by clonal analysis

Puromycin-resistant iPSC colonies derived from OCT4 targeting experiments using pDonor^{OCT4} and pDonor$^{OCT4.TS}$, were picked from 6-well plates and transferred into wells of 96-well plates by applying a standard 'cut-and-paste' technique. The resulting iPSC clones, each of which representing an individual genome-modifying event, were first cultured in StemFlex Medium (ThermoFisher Scientific) containing 25 U ml^{-1} penicillin and 25 μg ml^{-1} of streptomycin supplemented with Revitacell (ThermoFisher Scientific). Next, the iPSC clones were expanded and adapted to E8 medium (ThermoFisher Scientific) in wells of 24-well plates (Greiner-BioOne). The junction PCR screening for detecting and characterizing genome-modifying events in iPSCs was done on total genomic DNA purified from iPSC clones using the reagents and protocol provided in the Phire Tissue Direct PCR Master Mix kit (ThermoFisher Scientific). The PCR mixtures and cycling parameters applied for these analyses are indicated in the **Tables S14** and **S15**, respectively.

Characterization of iPSC clones by COBRA-FISH analysis

Combined binary ratio labelling (COBRA) multicolour FISH-based molecular karyotyping (COBRA-FISH) was carried out on native and gene-edited iPSC lines essentially as detailed elsewhere [33]. In brief, glass coverslips containing metaphase spreads air-dried for at least 24 h were incubated with 100 μg ml^{-1} RNase A (Roche; Cat. No.: 10154105103) in 2× saline-sodium citrate (SSC; Sigma-Aldrich; Cat. No.: S0902) at 37 °C for 10 min, followed by incubation with 0.005% pepsin (Sigma-Aldrich; Cat. No.: P0525000) in 0.1 M HCl for 5 min at 37 °C and fixation with 1% formaldehyde (Merck; Cat. No.: 1.03999.1000) in PBS pH 7.4 at room temperature for 10 min. The specimens were dehydrated through a series of incubations in 70–90–100% ethanol solutions, 3 min each, followed by air drying. The probe mix containing the paint mixes covering all chromosomes was dissolved in hybridization mixture, denatured and let hybridize in a moist chamber for 72 h. After hybridization, the glass coverslips were washed in 2× SSC and 0.1% Tween-20 (Promega, Cat. No.: PRH5152), then in 50% formamide (Merck; Cat. No.: 1.09684.1000), 2× SSC pH 7.0 solution at 44°C followed by incubation in 0.1× SSC at 60°C. Each washing step was performed twice for 5 min. The specimens were then dehydrated through a series of incubations in 70–90–100% ethanol solutions, air-dried and embedded in Citifluor AF1/DAPI (400 ng ml^{-1}) solution (Aurion; Cat. No.: E17970). Stained chromosomes were visualised using a Leica DMRA fluorescence

microscope (Leica, Wetzlar, Germany) and images were captured with the aid of a CoolSnap HQ2 camera (Photometrics, Tucson, USA). For image processing and karyotyping ColorProc, an in-house developed software tool, was used. A detailed protocol of the whole procedure has been published elsewhere [33].

Reverse transcriptase PCR analysis

Analysis of H2AX transcripts in mCherry$^+$ cells subjected to standard, *in trans* paired nicking and paired breaking gene editing, using either gRNA$^{H2AX.1}$ or gRNA$^{H2AX.2}$, was done as follows. Total RNA was extracted with the aid of the NucleoSpin RNA kit (Macherey-Nagel) essentially as specified by the manufacturer after adding 350 µl of RA1 buffer and 3.5 µl of β-mercaptoethanol (Merck). Reverse transcription on 1 µg of total RNA was performed at 50°C for 1 h with 200 ng of random primers, 0.2 mM dNTPs, 1× First-Strand Buffer, 5 mM dithiothreitol, and 200 U of SuperScript III Reverse Transcriptase (all from ThermoFisher Scientific). Next, 1-µl cDNA aliquots were subjected to PCR amplifications with the GoTaq G2 Flexi DNA Polymerase system (Promega; Cat. No.: M7808) using 0.4 µM of primer #1444 (5'-CAACGACGAGGAGCTCAACA-3'), 0.4 µM of primer #1508 (5'-GGCGGTGGTGGCCCTTAAAA-3'), 1 mM $MgCl_2$, 0.4 mM dNTPs, 1× GoTaq Flexi buffer, 1.25 U GoTaq and Milli-Q H_2O to a final volume of 25 µl. Cycling parameters are specified in **Table S16**. To serve as internal controls, 1-µl cDNA aliquots were also subjected to GAPDH-directed PCR amplifications with the GoTaq G2 Flexi DNA Polymerase system (Promega; Cat. No.: M7808) using, in this case, 0.4 µM of primer #119 (5'-AGCCACATCGCTCAGACACC-3') and 0.4 µM of primer #120 (5'-GTACTCAGCGCCAGCATCG-3'). Cycling parameters are specified in **Table S16**. Finally, 10 µl PCR samples corresponding to H2AX and GAPDH transcripts were electrophoresed through a 2% (w/v) agarose gel in 1× TAE buffer.

Detection of indels by targeted amplicon sequencing

Target site genotyping of HeLa cell populations containing unmodified cells mixed with cells generated by gene editing involving standard, paired breaking or *in trans* paired nicking was performed as follows. PCR products spanning gRNA$^{H2AX.1}$ and gRNA$^{H2AX.2}$ target sites were amplified from total cellular DNA extracted from cells at two different timepoints by using the reagents and protocol provided in the DNeasy Blood & Tissue kit (Qiagen; Cat. No.: 69506). The cycling parameters and PCR mixture composition used for amplifying the H2AX target region are specified in **Tables S16** and **S17**, respectively. H2AX-specific PCR products amplified from unmodified HeLa cell populations served as controls. Next, the amplicons corresponding to untagged H2AX alleles were extracted following the QIAEX II Gel Extraction Kit (Qiagen Cat. No.: 20021) and were subjected to Sanger sequencing for determining indel frequencies and distributions with the aid of the ICE software https://ice.synthego.com/#/ [34].

Characterization of PARP1 alleles in gene-edited cell populations

EGFP+ HeLa cells resulting from PARP1 gene tagging experiments using *in trans* paired nicking and standard gene editing protocols, were sorted with the aid of a BD FACSAria III flow cytometer (BD Biosciences). Next, total genomic DNA from these EGFP+ populations and from unmodified HeLa cells was extracted by using the DNeasy Blood & Tissue Kit (Qiagen; Cat. No.: 69506), according to the manufacturer›s instructions. The various DNA samples were subsequently subjected to PCR amplifications with two different primer pairs (i.e. primer pair A and B). Milli-Q water served as negative controls. The cycling parameters and PCR mixture compositions that were applied are indicated in **Tables S16** and **S17**, respectively. Indels at PARP1 alleles were detected by exposing amplicons to the mismatching-sensing T7EI (Biolabs) as below indicated.

The presence of a 121-bp PARP1 deletion in EGFP+ HeLa cells generated through standard gene editing was established by direct Sanger sequencing of the low-molecular-weight species (241-bp) resulting from PCR with the primer pair B (**Table S17).** Finally, the amplicons spanning the SpCas9-induced composite mutations were cloned using the TA cloning protocol (ThermoFisher Scientific Cat. No.: K1214) and were subsequently subjected to Sanger sequencing.

Identification and in silico analyses of H2AX and OCT4 gRNAs

The number and distribution of candidate off-target sites for CRISPR complexes was probed by using publicly available algorithms [35, 36]. The UCSC Genome Browser (Assembly GRCh38/hg38) was used to display all canonical S. pyogenes CRISPR-SpCas9 gRNAs in and around the target sequences for tagging H2AX and OCT4. The tracks of the UCSC Genome Browser displayed in the present study are available through the links: https://genome.ucsc.edu/s/mafvg/hg38 Chen Tasca et al C-terminus H2AX CRISPR Zoom, https://genome.ucsc.edu/s/mafvg/hg38 Chen Tasca et al C-terminus CRISPR Zoom and https://genome.ucsc.edu/s/mafvg/hg38 Chen Tasca et al OCT4 CRISPR 1.5X. The computing of the predicted performance of each CRISPR-SpCas9 complex was made by a combination of algorithms in the crispor.org tool [36]. The tracks for chained self-alignments and repeating elements are presented in full mode with the former depicting alignments of the human genome with itself after filtering out the redundant chromosomal positions that map to each other. As specified in the UCSC Genome Browser (Assembly GRCh38/hg38] website, the chained self-alignments and repeating elements tracks were generated with the aid of Blastz [37] and RepeatMasker (http://www.repeatmasker.org/), respectively.

Production and purification of lentiviral vector particles

The vesicular stomatitis virus glycoprotein G (VSV-G)-pseudotyped lentiviral vector LV.Cre was generated according to previously detailed protocols [38, 39]. In brief, 17×10^6 HEK293T cells were seeded per 175-cm^2 culture flask (Greiner Bio-One). The next day, the cells were transfected by adding to 19 ml of regular HEK293T cell culture medium, 1 ml of a 150 mM NaCl solution containing a mixture of 30 μg of DNA composed of lentiviral vector

shuttle, packaging, and VSV-G-pseudotyping plasmids at a ratio of 2:1:1 (size-normalized for molecule copy number) and 90 µl of 1 mg ml^{-1} PEI solution (25 kDa PEI, Polysciences). The shuttle, packaging and pseudotyping constructs used were, BC17_pLV.Cre (**Supplementary Information**), psPAX2 (Addgene #12260) and pLP/VSVG (Invitrogen). The HEK293T cells were incubated overnight in a total 20-ml transfection mixture, after which, this transfection medium was removed and replaced by fresh DMEM supplemented with 5% FBS. At 3 days post-transfection, producer-cell conditioned media containing released vector particles were collected and the cellular debris were removed by centrifugation and filtration using 0.45-µm pore-sized HT Tuffryn membrane filter (Pall Life Sciences; Cat. No. PN4184). The resulting clarified supernatants were gently added onto 5-ml 20% (v/v) sucrose cushions in 35.8-ml polyallomer tubes (Beckman Coulter; Cat. No.: 326823). After ultracentrifugation (15,000 rpm for 2 h at 4°C) in an Optima LE-80K centrifuge (Beckman Coulter) using the SW28 rotor, vector-containing pellets were resuspended in 400 µl of ice-cold PBS pH 7.4 supplemented with 1% (w/v) bovine serum albumin. The vector particle titer of the purified LV.Cre stock was shown to be 31589 ng $p24^{gag}$ ml^{-1} after employing the RETROTEK HIV-1 p24 antigen ELISA kit reagents and protocol (ZeptoMetrix, Cat. No.: 0801111).

Quantification of OCT4 gene targeting frequencies in iPSCs

Puromycin-resistant iPSCs resulting from OCT4 gene editing via single nicking, *in trans* paired nicking, standard and paired breaking protocols, were seeded in wells of 24-well plates (Greiner Bio-One) at a density of 30 000 cells per well. The next day, LV.Cre was added to the iPSCs in a total volume of 500 µl of culture medium at a multiplicity-of-infection of 10 vector particles per cell. The frequency of iPSCs expressing OCT4::EGFP assembled via Cre-mediated recombination was determined by flow cytometry at 9 days and 18 days post-transduction.

Characterization of iPSCs with OCT4 gene-edited alleles

Gene edited iPSCs expressing OCT4::EGFP after coupling *in trans* paired nicking to Cre-mediated recombination, were sorted through a BD FACSAria III flow cytometer (BD Biosciences) as single cell-deposited clones or as polyclonal populations. Both the $OCT4::EGFP^+$ clones and the $OCT4::EGFP^+$ cell populations were deposited in StemFlex Medium (ThermoFisher Scientific; Cat. No.: A3349401) containing 25 U ml^{-1} penicillin and 25 µg ml^{-1} of streptomycin (ThermoFisher Scientific) supplemented with Revitacell (ThermoFisher Scientific). The medium of the iPSC clones was replenished every other day. The medium was refreshed every day when the wells of 96-well plates (Greiner Bio-One) contained visible clusters of viable cells. These cell colonies were further expanded into wells of 48-well plates (Greiner Bio-One) and subsequently into wells of 24-well plates (Greiner Bio-One). Finally, they were expanded and adapted to grow in E8 medium. The $OCT4::EGFP^+$ iPSC clones and iPSC polyclonal populations were subsequently subjected to OCT4/EGFP dual-colour confocal microscopy and flow cytometry assays. Finally, the

pluripotency of iPSCs was assessed after applying differentiation protocols and confocal microscopy analyses as detailed under the section 'Differentiation of iPSCs'.

Confocal microscopy analyses

Cells seeded in glass coverslips were fixed in 2% or 4% (v/v) paraformaldehyde (PFA) and were permeabilized in 0.5% (w/v) Triton X-100 in Tris-buffered saline (TBS) pH 7.6 (50 mM Tris–HCl pH 7.6, 150 mM NaCl) at room temperature for 5–10 min (**Table S18**). Subsequently, the cells were incubated for 1 h to 2 h with blocking Antibody Diluting Solution (Abdil) consisting of TBS, Triton X-100, 2% (w/v) bovine serum albumin and 0.1% sodium azide. In-between each fixation, permeabilization and blocking steps, the specimens were washed three times for 5 min at room temperature with 0.1% Triton X-100 in TBS. The primary antibodies were diluted in Abdil (**Table S18**) and were added to the cells for 1 h at room temperature. After three 5-min washes with 0.1% Triton X-100 in TBS, the cells were incubated with fluorochrome-conjugated secondary antibodies diluted in Abdil for 30 min to 1 h in the dark at room temperature (**Table S18).** Next, the specimens were subjected to three 5-min washes with 0.1% Triton X-100 in TBS and were mounted in ProLong Gold Antifade Mounting reagent containing DAPI (ThermoFisher Scientific; Cat. No.: P36931) or in VECTASHIELD Antifade Mounting Medium (VECTOR; Cat. No.: H-1000). Before the addition of the latter mounting medium, the specimens were incubated for 5 min in the dark with the DNA staining reagent DAPI (Invitrogen Cat. No.: R37606) diluted 1:1000 in TBS. Finally, fluorescence microscopy was carried out with an upright Leica SP8 confocal microscope (Leica Microsystems) equipped with Leica hybrid detectors, HyD (Leica Microsystems).

Differentiation of iPSCs

The culturing of clumps of iPSCs on glass coverslips coated with VTN-N triggered the spontaneous differentiation of iPSCs along the three embryonic germ layers. In brief, iPSCs were treated with PBS-EDTA for 1 min at 37°C and were subsequently gently dissociated into large cell clumps by scrapping. The resulting cell clumps were then cultured in suspension for 24 h on low-attachment plates at 37°C. Next, the iPSCs were seeded on coverslips coated with VTN-N in Essential 8 medium (ThermoFisher Scientific, #A1517001) supplemented with Revitacell (ThermoFisher Scientific, Cat. #A2644501). The day after, the medium was changed to DMEM/F12 growth medium (Gibco Cat. #31331–028) containing 20% FBS (Biowest Cat. #S1860–500). The DMEM/F12 medium was replenished every 2–3 days. After 3 weeks under differentiation conditions, the iPSCs were processed for immunofluorescence confocal microscopy for the detection of markers characteristic of the endoderm, mesoderm and ectoderm lineages (**Table S19**). The markers corresponding to the three embryonic germ layers that were tested were α-fetoprotein (AFP), forkhead box protein A2 (FOXA2), α-smooth muscle actin (α-SMA), endothelial cell adhesion molecule-1 (CD31), and tubulin β3 class III (TUBB3).

T7 endonuclease I-based genotyping assays

Genotyping assays based on the mismatch-sensing T7EI enzyme were performed for detecting indels at target sequences of CRISPR complexes located at human PARP1, RAG1 and AAVS1 alleles and at off-target chromosomal positions located in the human OCT4 pseudogenes POU5F1P4 and POU5F1P5. For the latter assays, the genomic DNA of puromycin-resistant iPSC populations grown after OCT4-targeting experiments was extracted by using the DNeasy Blood & Tissue Kit and protocol (Qiagen, Cat. No.: 69506). The GoTaq G2 Flexi DNA Polymerase system (Promega; Cat. No.: M7808) was subsequently applied to amplify the POU5F1P4 and POU5F1P5 genomic sequences. The cycling parameters and PCR mixture compositions are specified in **Tables S16** and **S17**, respectively. Next, the resulting amplicons were subjected to the thermocycling procedure indicated in **Table S20** after which, 10-µl samples were incubated at 37 °C for 17 min with 1.5 µl 10× NEBuffer 2, 0.5 µl (5U) T7EI (New England Biolabs; Cat. No.: M0302) and 3 µl of Milli-Q water. Samples that were not treated with T7EI provided for negative controls. Finally, after agarose gel electrophoresis, untreated and T7EI-treated amplicons were analysed by using the Gel-Doc XR+ system and the ImageLab 4.1 software (both from Bio-Rad).

Flow cytometry

The frequencies of cells expressing H2AX::mCherry, EGFP::PARP1, OCT4::EGFP and EGFP were determined by using a BD LSR II flow cytometer (BD Biosciences). Parental unmodified cells or cells corresponding to experimental negative controls were used to establish the thresholds corresponding to background fluorescence. At least 10 000 viable single cells were analysed per sample. Data were analysed with the aid of FlowJo 10.5.0 software (Tree Star).

Western blotting

After two washes with ice-cold PBS pH 7.4, sorted EGFP::PARP1^{+} and EGFP::PARP1^{-} HeLa cells that had been exposed to standard gene editing or *in trans* paired nicking procedures were collected from wells of six-well plates by adding 250 µl of lysis RIPA buffer (Pierce Cat. No.: 89900) supplemented with a protease inhibitor cocktail (cOmplete Mini, Sigma-Aldrich Cat. No.: 11836153001). Untreated HeLa cells were taken along as negative controls. The cell lysates were subsequently passed thrice through a 1 ml syringe with a 26 GA 3/8 0.45 × 10 needle (BD Plastipak Cat. No.: 300015) and spun at 14 000 RPM for 5 min at 4°C in an Eppendorf 5424 centrifuge. The protein concentrations in the resulting supernatants were determined by using the BCA Protein Assay Kit (ThermoFisher Scientific Cat. No.: 23225) according to the manufacturer›s instructions. Next, 15 µg of protein were diluted in 4× sample buffer and 20× reducing agent (both from Bio-Rad Cat. No. 161-0791 and 161-0792, respectively) and incubated at 95°C for 5 min. Protein samples were loaded in a 7% SDS-PAGE gel. After electrophoreses, the proteins were transferred to a PVDF membrane (Millipore Immobilon Cat. No.: IPVH00010) and were blocked overnight in TBS with 0.05%

(v/v) Tween 20 (TBST, ThermoFisher Scientifc Cat. No.: 28358) supplemented with 5% (w/v) Elk milk (Campina). Next, the membrane was incubated with PARP1 polyclonal antibody (Thermo Fisher, Cat. No.: PA5-34803) diluted 1:5,000 in blocking buffer or with α/β tubulin antibody (Cell Signalling Cat. No.: CST 2148) diluted 1:5000 in blocking buffer. After an overnight incubation period at 4°C, the membranes were washed in TBST and incubated for 4 h at 4°C with an anti-rabbit IgG secondary antibody conjugated to horseradish peroxidase (IgG-HRP; Santa Cruz Cat. No.: sc-2004) diluted 1:1,000 in TBST. Proteins were detected by using horseradish peroxidase substrate Pierce ECL2 (Pierce Cat. No.: 80196) following the manufacturer›s specifications and Super RX-N X-ray film (Fujifilm).

Statistical analyses

With the exception of genomic DNA samples used for assessing genome-wide off-target effects of CRISPR complexes by orthogonal HTGTS analyses, the researchers were not blinded to sample allocation. Statistical analyses were performed on data sets derived from a minimum of three biological replicates done on different days. These data were analyzed by using the GraphPad Prism 8.0.1 software. The statistical significances were calculated with the tests indicated in the figure legends. P values lower than 0.05 were considered to be statistically significant.

Acknowledgements

The authors thank Martijn Rabelink, Steve Cramer and Hans Vrolijk (Department of Cell and Chemical Biology, LUMC, Leiden, the Netherlands) for performing p24gag ELISA titrations of lentiviral vector particles, assistance with the research logistics and making available the ColorProc source code, respectively. The authors also thank Elisa Giacomelli and Dorien Ward-van Oostwaard for sharing their iPSC culture protocols and Christian Freund (Department of Anatomy and Embryology, LUMC, Leiden, the Netherlands) for making available the Nucleofector 2b-device.

Additional Information

Supplementary Data: Supplementary information can be found online at doi: 10.1093/nar/gkz1121.

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 765269; and from the Dutch Prinses Beatrix Spierfonds [W.OR16-13]; X.C. and Q.W. receive support from Ph.D. fellowships from the China Scholarship Council-Leiden University Joint Scholarship Programme.

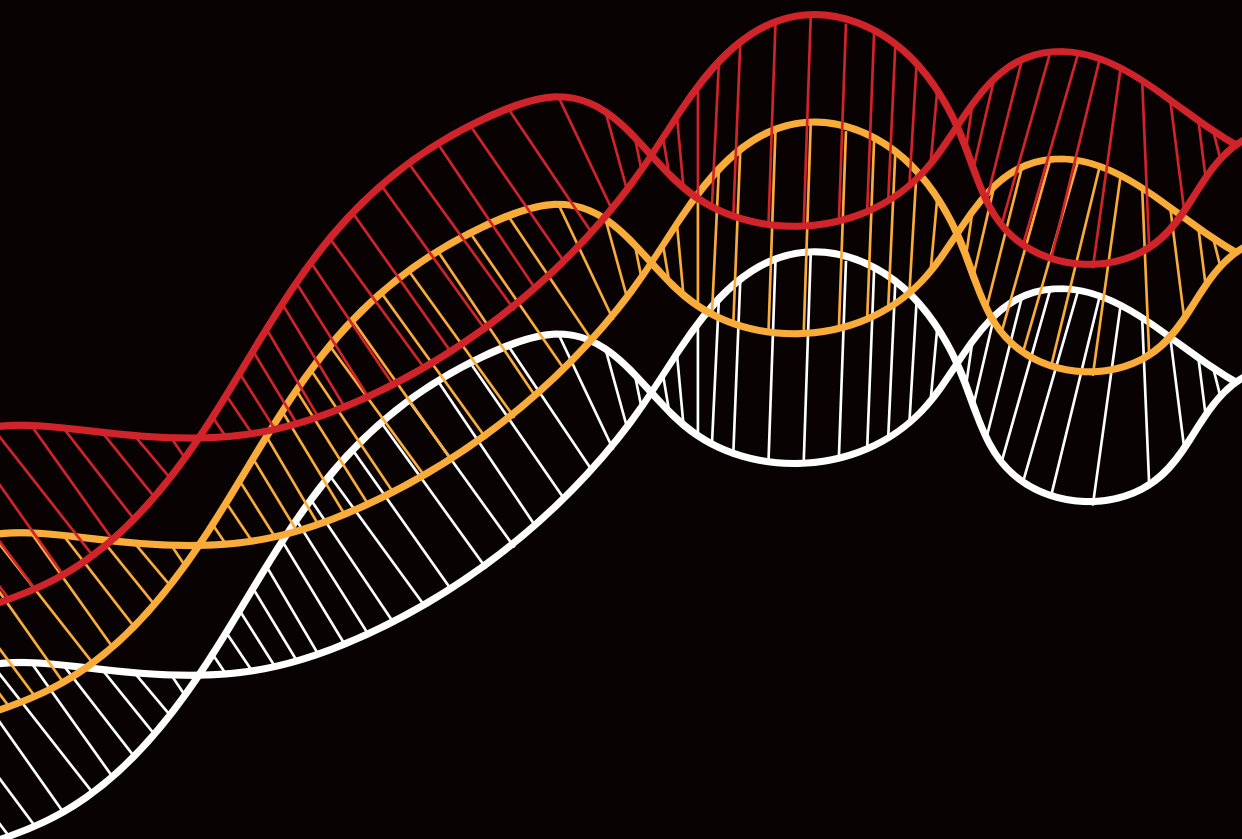
Conflict of interest statement: None declared.

References

1. Chandrasegaran, S.; Carroll, D. Origins of programmable nucleases for genome engineering. *J. Mol. Biol.* **2016**, 428, 963–989.
2. Maggio, I.; Gonçalves, M.A. Genome editing at the crossroads of delivery, specificity, and fidelity. *Trends Biotechnol.* **2015**, 33, 280–291.
3. Chang, H.H.Y.; Pannunzio, N.R.; Adachi, N., Lieber, M.R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **2017**, 18, 495–506.
4. Kosicki, M.; Tomberg, K.; Bradley A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **2018**, 36, 765–771.
5. Cradick, T.J.; Fine, E.J.; Antico, C.J.; Bao, G. CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* **2013**, 41, 9584–9592.
6. Fu, Y.; Foden, J.A.; Khayter, C.; Maeder, M.L.; Reyon, D.; Joung, J.K.; Sander, J.D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **2013**, 31, 822–826.
7. Cho, S.W.; Kim, S.; Kim, Y.; Kweon, J.; Kim, H.S.; Bae, S.; Kim, J.S. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **2014**, 24, 132–141.
8. Lin, Y.; Cradick, T.J.; Brown, M.T.; Deshmukh, H.; Ranjan, P.; Sarode, N.; Wile, B.M.; Vertino, P.M.; Stewart, F.J.; Bao, G. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res.* **2014**, 42, 7473–7485.
9. Kuscu, C.; Arslan, S.; Singh, R.; Thorpe, J.; Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* **2014**, 32, 677–683.
10. Frock, R.L.; Hu, J.; Meyers, R.M.; Ho, Y.J.; Kii, E.; Alt, F.W. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* **2015**, 33, 179–186.
11. Holkers, M.; Maggio, I.; Henriques, S.F.; Janssen, J.M.; Cathomen, T.; Gonçalves, MA. Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat. Methods.* **2014**, 11, 1051–1057.
12. Chen, X.; Janssen, J.M.; Liu, J.; Maggio, I.; 't Jong, A.E.; Mikkers, H.M.; Gonçalves, M.A. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat. Commun.* **2017**, 8, 657.
13. Zhang, J.P.; Li, X.L.; Li, G.H.; Chen, W.; Arakaki, C.; Botimer, G.D.; Baylink, D.; Zhang, L.; Wen, W.; Fu, Y.W.; et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* **2017**, 18, 35.
14. McConnell, Smith A.; Takeuchi, R.; Pellenz, S.; Davis, L.; Maizels, N.; Monnat, R.J. Jr.; Stoddard, B.L. Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106, 5099–5104.
15. Metzger, M.J.; McConnell-Smith, A.; Stoddard, B.L.; Miller, A.D. Single-strand nicks induce homologous recombination with less toxicity than double-strand breaks using an AAV vector template. *Nucleic Acids Res.* **2011**, 39, 926–935.
16. Ramirez, C.L.; Certo, M.T.; Mussolino, C.; Goodwin, M.J.; Cradick, T.J.; McCaffrey, A.P.; Cathomen, T.; Scharenberg, A.M.; Joung, J.K. Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res.* **2012**, 40, 5560–5568.
17. Wang, J.; Friedman, G.; Doyon, Y.; Wang, N.S.; Li, C.J.; Miller, J.C.; Hua, K.L.; Yan, J.J.; Babiarz, J.E.; Gregory, P.D.; et al. Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res.* **2012**, 22, 1316–1326.
18. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* **2012**, 337, 816–821.
19. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* **2013**, 339, 819–823.
20. Mali, P.; Yang, L.; Esvelt, K.M.; Aach, J.; Guell, M.; DiCarlo, J.E.; Norville, J.E.; Church, G.M. RNA-guided human genome engineering via Cas9. *Science.* **2013**, 339, 823–826.
21. Doudna, J.A.; Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR–Cas9. *Science.* **2014**, 346, 1258096.
22. Mali, P.; Aach, J.; Stranges, P.B.; Esvelt, K.M.; Moosburner, M.; Kosuri, S.; Yang, L.; Church, G.M. Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **2013**, 31, 833–838.
23. Ran, F.A.; Hsu, P.D.; Lin, C.Y.; Gootenberg, J.S.; Konermann, S.; Trevino, A.E.; Scott, D.A.; Inoue, A.; Matoba, S.; Zhang, Y.; et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell.* **2013**, 154, 1380–1389.
24. van Nierop, G.P.; de Vries, A.A.; Holkers, M.; Vrijsen, K.R.; Gonçalves, M.A. Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease. *Nucleic Acids Res.* **2009**, 37, 5725–5736.

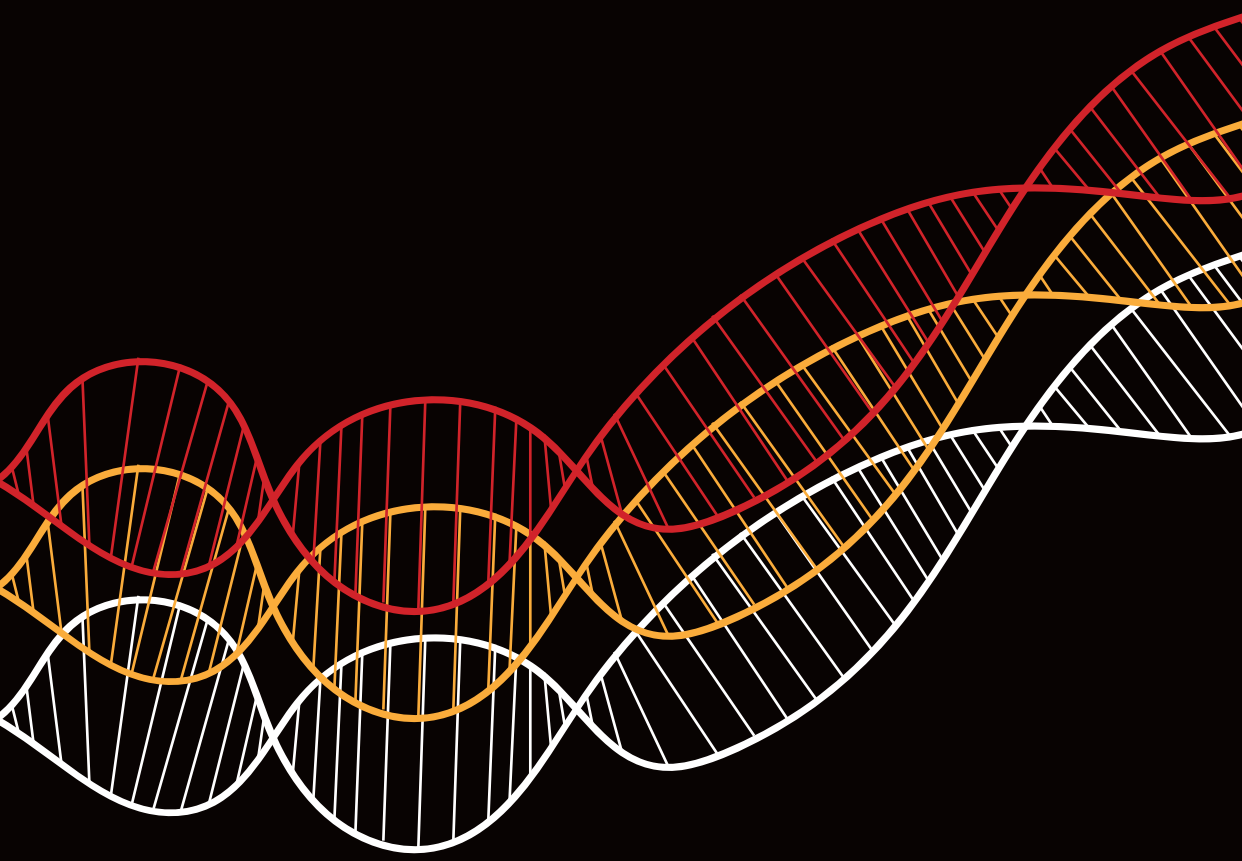
25. Nakajima, K.; Zhou, Y.; Tomita, A.; Hirade, Y.; Gurumurthy, C.B.; Nakada, S. Precise and efficient nucleotide substitution near genomic nick via noncanonical homology-directed repair. *Genome Res.* **2018**, 28, 223–230.
26. Paulsen, B.S.; Mandal, P.K.; Frock, R.L.; Boyraz, B.; Yadav, R.; Upadhyayula, S.; Gutierrez-Martinez, P.; Ebina, W.; Fasth, A.; Kirchhausen, T.; et al. Ectopic expression of RAD52 and dn53BP1 improves homology-directed repair during CRISPR–Cas9 genome editing. *Nat. Biomed. Eng.* **2017**, 1, 878–888.
27. Zhang, M.; D'Aniello, C.; Verkerk, A.O.; Wrobel, E.; Frank, S.; Ward-van Oostwaard, D.; Piccini, I.; Freund, C.; Rao, J.; Seebohm, G.; et al. Recessive cardiac phenotypes in induced pluripotent stem cell models of Jervell and Lange-Nielsen syndrome: disease mechanisms and pharmacological rescue. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, 111, E5383–E5392.
28. Hockemeyer, D.; Wang, H.; Kiani, S.; Lai, C.S.; Gao, Q.; Cassady, J.P.; Cost, G.J.; Zhang, L.; Santiago, Y.; Miller, J.C.; et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* 2011, 29, 731–734.
29. Chen, X.; Liu, J.; Janssen, J.M.; Gonçalves, M.A. The chromatin structure differentially impacts high-specificity CRISPR–Cas9 nuclease strategies. *Mol. Ther. Nucleic Acids.* 2017, 8, 558–563.
30. Kleinstiver, B.P.; Prew, M.S.; Tsai, S.Q.; Nguyen, N.T.; Topkar, V.V.; Zheng, Z.; Joung, J.K. Broadening the targeting range of Staphylococcus aureus CRISPR–Cas9 by modifying PAM recognition. *Nat. Biotechnol.* **2015**, 33, 1293–1298.
31. Hu, J.; Meyers, R.M.; Dong, J.; Panchakshari, R.A.; Alt, F.W.; Frock, R.L. Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc.* **2016**, 11, 853–871.
32. Brielmeier, M.; Béchet, J.M.; Falk, M.H.; Pawlita, M.; Polack, A.; Bornkamm, G.W. Improving stable transfection efficiency: antioxidants dramatically improve the outgrowth of clones under dominant marker selection. *Nucleic Acids. Res.* **1998**, 26, 2082–2085.
33. Szuhai, K.; Tanke, H.J. COBRA: combined binary ratio labeling of nucleic-acid probes for multi-color fluorescence in situ hybridization karyotyping. *Nat. Protoc.* **2006**, 1, 264–275.
34. Hsiau, T.; Maures, T.; Waite, K.; Yang, J.; Kelso, R.; Holden, K.; Stoner, R. Inference of CRISPR Edits from Sanger Trace Data. 2018; *bioRxiv* doi:10 August **2019**, preprint: not peer reviewedhttps://doi.org/10.1101/251082.
35. Bae, S.; Park, J.; Kim, J.S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics.* **2014**, 30, 1473–1475.
36. Doench, J.G.; Fusi, N.; Sullender, M.; Hegde, M.; Vaimberg, E.W.; Donovan, K.F.; Smith, I.; Tothova, Z.; Wilen, C.; Orchard, R.; et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR–Cas9. *Nat. Biotechnol.* **2016**, 34, 184–191.
37. Schwartz, S.; Kent, W.J.; Smit, A.; Zhang, Z.; Baertsch, R.; Hardison, R.C.; Haussler, D.; Miller, W. Human-mouse alignments with BLASTZ. *Genome Res.* **2003**, 13, 103–107.
38. Pelascini, L.P.; Janssen, J.M.; Gonçalves, M.A. Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. *Hum. Gene Ther.* **2013**, 24, 78–96.
39. Pelascini, L.P.; Gonçalves, M.A. Lentiviral vectors encoding zinc-finger nucleases specific for the model target locus HPRT1. *Methods Mol. Biol.* **2014**, 1114, 181–199.
40. Gabriel, R.; von Kalle, C.; Schmidt, M. Mapping the precision of genome editing. *Nat. Biotechnol.* **2015**; 33, 150–152.
41. Wienert, B.; Wyman, S.K.; Richardson, C.D.; Yeh, C.D.; Akcakaya, P.; Porritt, M.J.; Morlock, M.; Vu, J.T.; Kazane, K.R.; Watry, H.L.; et al. Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science.* **2019**, 364, 286–289.
42. Kuzminov, A. Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, 98, 8241–8246.
43. Casper, J.; Zweig, A.S.; Villarreal, C.; Tyner, C.; Speir, M.L.; Rosenbloom, K.R.; Raney, B.J.; Lee, C.M.; Lee, B.T.; Karolchik, D.; et al. The UCSC Genome Browser database: 2018 update. *Nucleic Acids. Res.* **2018**, 46, D762–D769.
44. Foster, E.R.; Downs, J.A. Histone H2A phosphorylation in DNA double-strand break repair. *FEBS J.* **2005**, 272, 3231–3240.
45. Holliday, R. A mechanism for gene conversion in fungi. *Genet. Res.* **1964**, 5, 283–304.
46. Celeste, A.; Difilippantonio, S.; Difilippantonio, M.J.; Fernandez-Capetillo, O.; Pilch, D.R.; Sedelnikova, O.A.; Eckhaus, M.; Ried, T.; Bonner, W.M.; Nussenzweig, A. H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. *Cell.* **2003**, 114, 371–383.
47. Hanzlikova, H.; Gittens, W.; Krejcikova, K.; Zeng, Z.; Caldecott, K.W. Overlapping roles for PARP1 and PARP2 in the recruitment of endogenous XRCC1 and PNKP into oxidized chromatin. *Nucleic Acids Res.* **2017**, 45, 2546–2557.
48. Ronson, G.E.; Piberger, A.L.; Higgs, M.R.; Olsen, A.L.; Stewart, G.S.; McHugh, P.J.; Petermann, E.; Lakin, ND. PARP1 and PARP2 stabilise replication forks at base excision repair intermediates through Fbh1-dependent

Rad51 regulation. *Nat. Commun.* **2018**, 9, 746.
49. Fogarty, N.M.E.; McCarthy, A.; Snijders, K.E.; Powell, B.E.; Kubikova, N.; Blakeley, P.; Lea, R.; Elder, K.; Wamaitha, S.E.; Kim, D.; et al. Genome editing reveals a role for OCT4 in human embryogenesis. *Nature*. **2017**, 550, 67–73.
50. Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. **2007**, 131, 861–872.
51. Yilmaz, A.; Peretz, M.; Aharony, A.; Sagi, I.; Benvenisty, N. Defining essential genes for human pluripotent stem cells by CRISPR–Cas9 screening in haploid cells. *Nat. Cell Biol.* **2018**, 20, 610–619.
52. Liu, J.C.; Lerou, P.H.; Lahav, G. Stem cells: balancing resistance and sensitivity to DNA damage. *Trends Cell Biol.* **2014**, 24, 268–274.
53. Ihry, R.J.; Worringer, K.A.; Salick, M.R.; Frias, E.; Ho, D.; Theriault, K.; Kommineni, S.; Chen, J.; Sondey, M.; Ye, C.; et al. p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **2018**, 24, 939–946.
54. Haapaniemi, E.; Botla, S.; Persson, J.; Schmierer, B.; Taipale, J. CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* **2018**, 24, 927–930.
55. Smalley, E. As CRISPR-Cas adoption soars, summit calls for genome editing oversight. *Nat. Biotechnol.* **2018**, 36, 4852018.
56. El-Brolosy, M.A.; Kontarakis, Z.; Rossi, A.; Kuenne, C.; Günther, S.; Fukuda, N.; Kikhi, K.; Boezio, G.L.M.; Takacs, C.M.; Lai, S.L.; et al. Genetic compensation triggered by mutant mRNA degradation. *Nature*. **2019**, 568, 193–197.
57. Ma, Z.; Zhu, P.; Shi, H.; Guo, L.; Zhang, Q.; Chen, Y.; Chen, S.; Zhang, Z.; Peng, J.; Chen, J. PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. *Nature*. **2019**, 568, 259–263.
58. Tuladhar, R.; Yeu, Y.; Piazza, J.T.; Tan, Z.; Clemenceau, J.R.; Wu, X.; Barrett, Q.; Herbert, J.; Mathews, D.H.; Kim, J.; et al. CRISPR–Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nat. Commun*. **2019**, 10, 4056.
59. Zhu, Z.; Verma, N.; González, F.; Shi, Z.D.; Huangfu, D. A CRISPR/Cas-mediated selection-free knockin strategy in human embryonic stem cells. *Stem Cell Rep.* **2015**, 4, 1103–1111.
60. Munoz, D.M.; Cassiani, P.J.; Li, L.; Billy, E.; Korn, J.M.; Jones, M.D.; Golji, J.; Ruddy, D.A.; Yu, K.; McAllister, G.; et al. CRISPR screens provide a comprehensive assessment of cancer vulnerabilities but generate false-positive hits for highly amplified genomic regions. *Cancer Discov.* **2016**, 6, 900–913.
61. Gonçalves, E.; Behan, F.M.; Louzada, S.; Arnol, D.; Stronach, E.A.; Yang, F.; Yusa, K.; Stegle, O.; Iorio, F.; Garnett, M.J. Structural rearrangements generate cell-specific, gene-independent CRISPR–Cas9 loss of fitness effects. *Genome Biol.* **2019**, 20, 27.
62. Chen, X.; Gonçalves, M.A. DNA, RNA, and protein tools for editing the genetic information in human cells. *iScience.* **2018**, 6, 247–263.
63. Deng, C.; Capecchi, M.R. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell Biol.* **1992**, 12, 3365–3371.
64. Bellin, M.; Marchetto, M.C.; Gage, F.H.; Mummery, C.L. Induced pluripotent stem cells: the new patient. *Nat. Rev. Mol. Cell Biol.* **2012**, 13, 713–726.
65. Findlay, G.M.; Boyle, E.A.; Hause, R.J.; Klein, J.C.; Shendure, J. Saturation editing of genomic regions by multiplex homology-directed repair. *Nature*. **2014**, 513, 120–123.
66. Sedlazeck, F.J.; Lee, H.; Darby, C.A.; Schatz, M.C. Piercing the dark matter: bioinformatics of long-range sequencing and mapping. *Nat. Rev. Genet*. **2018**, 19, 329–346.



Chapter 6

Conclusions and Final remarks



Conclusions and Final remarks

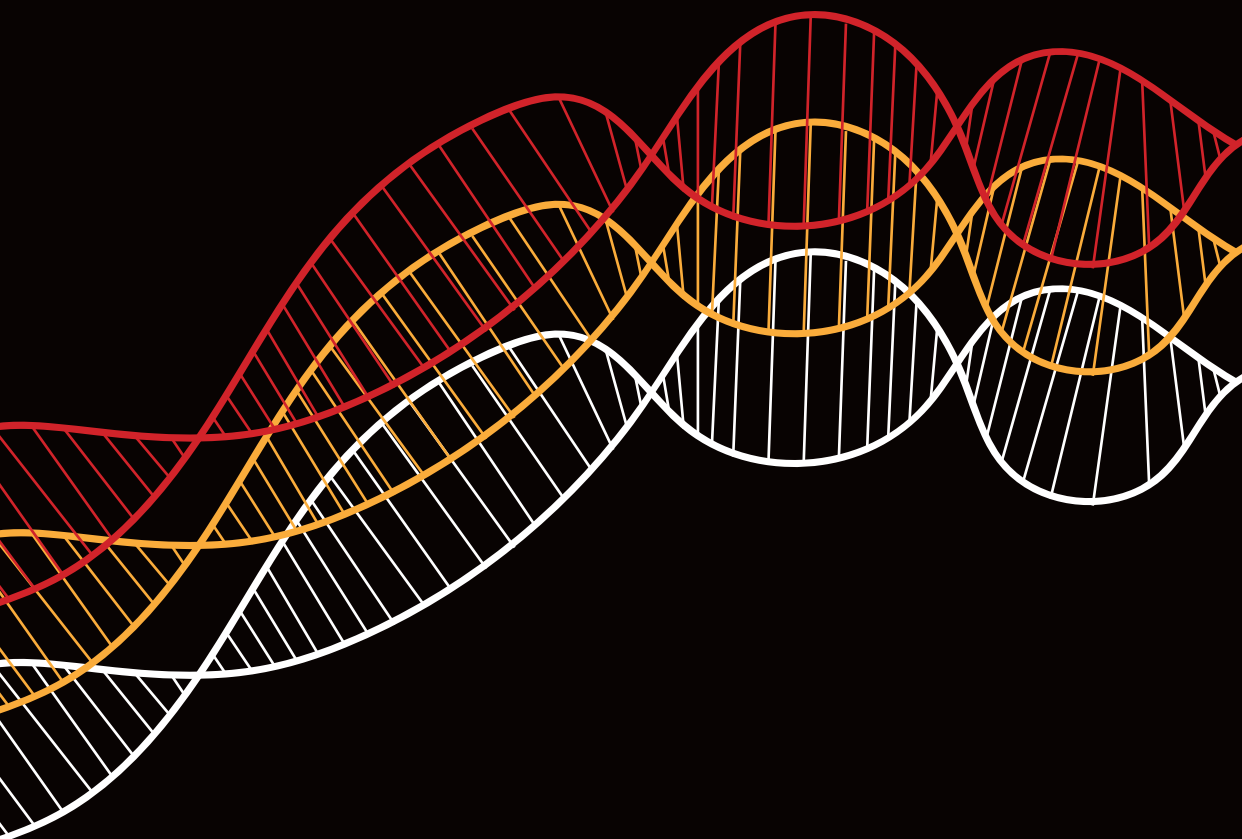
Genome editing (GE) strategies based on homology-dependent and homology-independent DNA repair pathways activated by programmable nucleases, e.g., RNA-guided nucleases (RGNs) from CRISPR-Cas systems, permit modifying specific chromosomal sequences in living cells, spanning from single base-pairs to whole genomic tracts. This set of techniques is already proving its outstanding potential in a variety of fields, ranging from basic research to applied biotechnology, biomedical research, and medicine. Nevertheless, constant efforts are in demand to further improve the efficiency and specificity of GE tools and procedures, especially those directed at clinical translation. As outlined in **Chapter 2**, key to broaden the possibilities of GE technologies to gene and cell therapies, is their application in stem cells. To this end, it is essential to implement delivery systems that permit introducing, in an efficient and non-cytotoxic manner, the latest-generation GE tools into hard-to-transfect target cell types (e.g., non-transformed somatic cells and human induced pluripotent stem cells, hiPSCs). It is in this context that, in **Chapter 3** and **Chapter 4**, high-capacity adenoviral vector particles (AdVPs) are investigated as carriers of GE tools consisting of conventional and novel programmable nucleases alone or together with donor DNA sequences prone to specific DNA repair pathways. The versatility of this delivery platform facilitated the testing of novel GE approaches in diverse stem and progenitor cells with therapeutic potential, including hiPSCs and muscle progenitor cells. Specifically, the strategies were tailored to correct the genetic defect underlying Duchenne Muscular Dystrophy (DMD). DMD is a fatal X-linked muscle wasting disorder caused by a broad range of loss-of-function mutations in the largest known protein-coding gene in the human genome, i.e., the ~2.4 Mb-long dystrophin-encoding *DMD* gene. To date, DMD-targeted gene therapies focus mainly on the overexpression of microdystrophins upon adeno-associated viral vector delivery or *in situ* assembly of Becker-like dystrophins after RGN transfer followed by non-homologous end joining (NHEJ)-mediated restoration of defective *DMD* reading frames.

In fact, **Chapter 3** explores the use of AdVPs for multiplexing GE approaches aiming at NHEJ-mediated repair of the *DMD* reading frame and ensuing expression of Becker-like dystrophins upon targeted DNA deletion. The multiplexing GE strategy investigated in this Chapter is based on the coordinated formation of DSBs by covalently joined RGN pairs (i.e., forced RGN heterodimers) designed for the excision of *DMD* reading frame-disrupting mutations. Uncoordinated activity of independently acting RGNs proved, in this study and elsewhere, to perform at diverse levels of efficiency and to generate substantial amounts unintended genomic modifications that compound the intended GE outcome in the form of precise DNA deletions. In **Chapter 3**, by employing AdVPs to deliver forced RGN heterodimers, it is uncovered that the frequency and accuracy of targeted DNA deletions are superior to that obtained when the various RGN components are delivered separately. This approach bodes well for GE applications in which generating targeted chromosomal deletions in a precise and efficient manner is necessary.

Chapter 4 expands on the research efforts described in **Chapter 3** by further leveraging the AdVP platform to, in this case, recruit homology-dependent DNA repair processes to achieve long-term complementation of DMD-causing mutations regardless of their type or location, namely, via site-specific chromosomal insertion of transgenes expressing the full-length striated muscle-specific isoform of dystrophin (427 kDa). The AdVP-based GE strategies investigated in **Chapter 4** involved the delivery of RGNs together with donor DNA templates prone to homologous recombination (HR) or homology-mediated end joining (HMEJ). HMEJ and HR donor constructs were both designed for targeted chromosomal integration of full-length dystrophin expression units at the commonly used "safe harbour" locus *AAVS1* at 19q13.42. Generally, it was found that the delivery of HMEJ donors led to higher frequencies of on-target integrations than those resulting from the transfer of HR donors. The efficiencies of the HMEJ- and HR-based GE strategies were instead found to be similar in hiPSCs yet, interestingly, DSB-induced gene targeting levels under p53 inhibiting conditions could be rescued specifically in hiPSCs transduced with HMEJ donors. In conclusion, the AdVP methodologies described in **Chapter 4** allow the investigation and application of different gene knock-in approaches in hard-to-transfect human cell types irrespective of RGN complex and transgene sizes. Altogether, the work described in **Chapter 3** and **Chapter 4** support the use of AdVPs for the development of effective and broadly applicable gene therapies based on CRISPR components, including those involving *ex vivo* correction and autologous transplantation of stem/progenitor cells. Nevertheless, both studies also highlight drawbacks of DSB-based GE strategies in the form of undesired on-target and off-target genomic modifications and, therefore, stress the need to carefully scrutinise GE outcomes when applying such strategies. Additionally, the data presented in **Chapter 4** further support the argument that DSB-based GE is particularly impaired in stem cells likely due to the induction of p53-dependent cell cycle arrest and apoptosis by the programmable nucleases. In light of these by-product events consistently detected in cells exposed to programmable nucleases , recent developments on genomic engineering comprise the progression from chromosomal cutting to chromosomal non-cutting approaches based on nicking CRISPR-Cas variants.

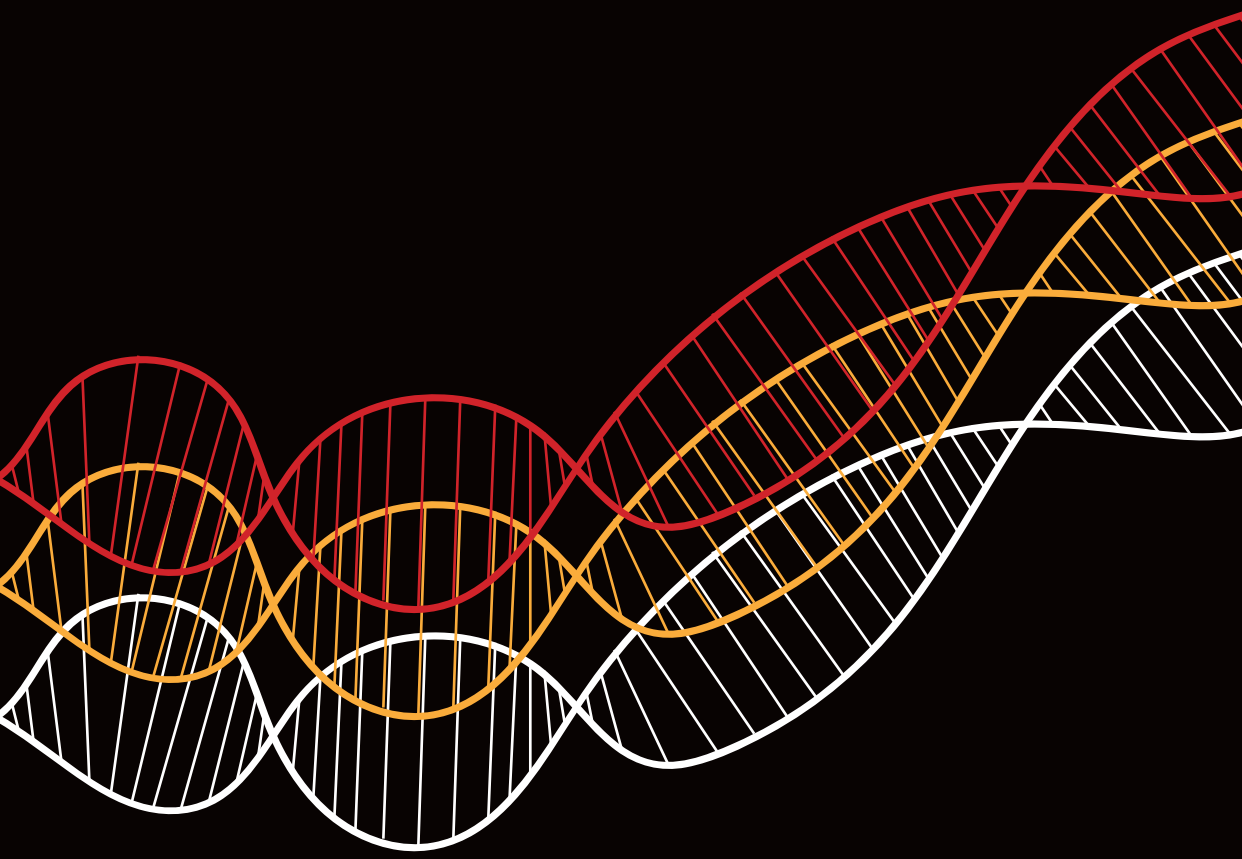
Contributing to this line of research, in **Chapter 5**, the benefits of employing a DSB-free GE strategy to modify particularly sensitive genomic regions and cells is instead investigated. This *in trans* paired nicking strategy, based on the simultaneous formation of single-strand DNA breaks (SSBs) at donor and acceptor DNA by CRISPR–Cas9 nickases, proved successful in knocking-in large DNA segments efficiently and precisely at loci associated with haploinsufficiency and essentiality in diverse human cell types, including hiPSCs. Moreover, this SSB-based GE strategy circumvents most large- and small-scale mutagenic events caused by DSBs maximizing the preservation of cellular genotypes and phenotypes. Hence, the seamless and scarless character of *in trans* paired nicking might be particularly beneficial in the editing of stem cells, especially when precise and predictable genetic interventions are

delivery of forced RGN heterodimer components which, in comparison to split conventional RGN multiplexes, engage target sequences in a more coordinated fashion. Moreover, besides serving as probes for investigating different gene targeting approaches independently of GE tool and donor DNA sizes, AdVPs constitute a robust platform for delivering and stably installing in a targeted homology-dependent manner large genetic payloads in human cells. Finally, closely monitoring GE procedures and further progressing towards DSB-free GE strategies will become ever-more crucial in aiding gene therapy research progressing to clinical application. In this context, tailoring AdVPs for the testing of such precision GE approaches in therapeutically relevant cell types and animal models can contribute to opening the doors to the development of safer therapeutic interventions that tackle the root cause of genetic disorders, such as DMD.



Chapter 7

Nederlandse Samenvatting



Nederlandse samenvatting

Genoom-aanpassing (GE) technieken gebaseerd op homologe-afhankelijk en homologe-onafhankelijk DNA-reparatie die geactiveerd kunnen worden door nucleasen, zoals bijvoorbeeld RNA-geleide nucleasen (RGNs) van het CRISPR-Cas systeem, laten specifieke modificaties in chromosomale sequenties in levende cellen toe, dit kan zowel een base-pair of gehele transgenen zijn. Deze technieken hebben al bewezen dat ze een uitstekende potentie hebben in veel verschillende velden, van basis onderzoek tot toegepaste biotechnologie, biomedische onderzoek, en medicijnen. Echter, het continue streven naar verdere verbetering van de effectiviteit en specificiteit van GE technieken is vereist, vooral als er gekeken wordt naar het gebruik in een klinische setting. Zoals beschreven in **Hoofdstuk 2**, is de sleutel tot het verbeteren van de mogelijke kandidaten voor gen en cel therapieën het gebruiken van de meest vergevorderde GE technieken en strategieën in stam cellen. Daarom is het essentieel om bezorgsystemen te implementeren die het toelaten om de laatste-generatie GE technieken in moeilijk om te transferten doel cellen (bijvoorbeeld humane somatische cellen en induceerbare pluripotente stam cellen, hiPSCs) te kunnen introduceren in een efficiënte, veelzijdige, en niet toxische manier.

In deze context wordt in **Hoofdstuk 3** en **Hoofdstuk 4** de hoge capaciteit adenovirale vectoren deeltjes (AdVPs), gebaseerd op gemodificeerde programmeerbare nucleasen en/of donor DNA sequenties gevoelig voor specifieke DNA-reparatiemechanismes, onderzocht als vervoerders voor GE technieken. De veelzijdigheid van dit levering platform helpt in het testen van nieuwe GE mogelijkheden in verschillende stam- en voorlopercellen met therapeutische potentiaal, met onder andere hiPSCs en spiervoorlopercellen. Om specifiek te zijn, de strategieën waren gemaakt om de genetische afwijking leidend tot Duchenne spierdystrofie (DMD) te corrigeren. DMD is een fatale X-chromosoom gebonden spier ziekte veroorzaakt door een grote range van verlies-van-functie mutaties in het grootste eiwit-coderende gen in het humane genoom (de ~2,4 Mb dystrofine-coderende *DMD*-gen). Tot op heden, focussen DMD-gerichte gen therapieën zich voornamelijk op de over expressie van micro-dystrofines na AdV bezorging of het *in situ* maken van Becker-like dystrofines na het inbrengen van RGN.

Hoofdstuk 3 draagt bij aan het GE-veld door het gebruik van AdVPs multiplexing GE strategieën gericht op het herstel van de DMD leesraam dat uiteindelijk leidt tot expressie van Becker-achtige dystrofines. De onderzochte multiplex strategie was gebaseerd op de gecoördineerde formatie van dubbelstrengs DNA-breuken (DSBs) door RGN-paren die waren ontworpen voor het verwijderen van verstorende sequenties in het DMD leesraam in Duchenne patiënten-afgeleide myoblasten. Ongecoördineerde activiteit van onafhankelijk werkende RGNs heeft aangetoond, hier en in eerder onderzoek, te werken op verschillende efficiëntie niveaus en om substantiële hoeveelheden onbedoelde genomische modificaties te genereren. Deze studie heeft ontdekt dat, door het gebruik van AdVPs om geforceerd RGN heterodimeren aan te leveren samen met hun gids-RNAs, dat de frequentie en nauwkeurigheid

van de bedoelde DNA deleties beter zijn in vergelijking met die zijn verkregen door verschillende componenten apart aan te leveren. Het gebruik van de geforceerde aanpak is een goed voorteken voor de verbetering van GE-toepassingen, waar het op een nauwkeurige en efficiënte manier genereren van gerichte DNA deleties noodzakelijk is.

Hoofdstuk 4 borduurt voor op het onderzoek dat beschreven is in **Hoofdstuk 3** door het AdVP platform verder te gebruiken om, in dit geval, homologie-afhankelijke DNA-herstelprocessen te rekruteren om langdurige complemetatie van DMD-veroorzakende mutaties te bereiken onafhankelijk van de type of locatie, namelijk via plaats specifieke chromosomale insertie van transgenen die de volledige lengte, dus volledig functioneel, dwarsgestreepte spier-specifieke isovorm van dystrofine (427 kDa) tot expressie brengt. De onderzochte AdVP-gebaseerde GE strategieën bevatten de levering van donor DNA-templates vatbaar voor homologe recombinatie (HR) of homologe-gemedieerde eindverbinding (HMEJ) en is gemaakt om de DNA integratie in de veel voorkomend gebruikte "veilige haven" locus *AAVS1* op 19q13.42 te plaatsen. Over het algemeen was er gevonden dat de HMEJ-gebaseerde GE strategieën leiden tot hogere frequentie van on-target integraties dan die van HR-gemedieerde GE in HeLa cellen en myoblasten. De efficiëntie van de HMEJ- en HR-gebaseerde GE strategieën waren vergelijkbaar in hiPSC, maar interessant genoeg kon de gen targeting niveaus onder p53-remmende omstandigheden gered worden specifiek in getransduceerde iPSC met HMEJ donors. In conclusie, de beschreven AdVP methodes maken het onderzoek en de toepassing van verschillende gen knock-in approaches mogelijk in moeilijk te transfereren cellen, ongeacht van het CRISPR complex en de grootte van de transgenen.

Samenvattend, het werk besproken in **Hoofdstuk 3** en **Hoofdstuk 4** ondersteund het gebruik van AdVPs voor de ontwikkeling van effectieve en breed toepasbare gen therapie gebaseerd op programmeerbare nucleasen, inclusief die met de *ex vivo* correctie en autologe transplantatie van stem-/voorlopercellen. Desalniettemin, hebben beide studies ook nadelen aan het licht gebracht van het gebruik van op DSB gebaseerde GE-strategieën in de vorm van on-target en off-target ongewenste aanpassingen en daarom is het noodzakelijk om GE-resultaten zorgvuldig te onderzoeken bij het toepassen van dergelijke strategieën. Bovendien ondersteunen de gegevens gepresenteerd in **Hoofdstuk 4** verder dat de programmeerbare nucleasen-gemedieerde GE strategieën voornamelijk minder efficiënt werkzaam zijn in stam cellen, waarschijnlijk door de DSB-getriggerde p53-afhankelijke cel cyclus stilstand en apoptose reactie. Door de bijproducten die ontstaan bij de programmeerbare nucleasen-gemedieerde GE, zorgen recente ontwikkelingen op het gebied van genomische technologie voor de vooruitgang van chromosomaal snijden naar chromosomaal niet-snijdende benadering op basis van "nicking" van Cas9-varianten.

Om verder te gaan met dit onderzoek, is in **Hoofdstuk 5** het voordeel van het gebruik van DSB-vrije GE strategieën onderzocht om vooral op gevoelige genomische regio's en cellen te richten. *In trans* gepaarde nicking, een strategie dat gebaseerd is op het gelijktijdig

vormen van enkelstrengs DNA-breuken (SSBs) op donor en acceptor DNA door CRISPR-Cas9 nickases, is succesvol in het introduceren van grote DNA segmenten efficiënt en precies op loci geassocieerd met haploinssuficienteit vooral in verschillende humane cellen, inclusief hiPSCs. Deze SSB-gebaseerde GE strategie zorgt ervoor dat grote- en kleine-schaal mutagene evenementen veroorzaakt door DSB gemeden zijn en zorgt vormelijk voor het behoud van cellulaire genotype en fenotypes. De accurate en specifieke manier van *in trans* gepaarde nicking kan voornamelijk behulpzaam zijn in editing van stam cellen, waar een precieze en voorspelbare genetische interventie essentieel is.

In conclusie, het nauwlettend volgen van op DSB-gebaseerde GE technieken en verdere vooruitgang naar DSB-vrije GE strategieën zal steeds belangrijker worden voor onderzoek naar translationele gen therapieën. In deze context, zal het afstemmen van AdVPs voor het testen van een precieze GE-benadering in therapeutisch relevante cellen ervoor kunnen zorgen dat deuren geopend worden naar de ontwikkeling van nieuwe therapeutische mogelijkheden die de grondoorzaken van (mono)genetische ziektes zoals DMD zullen aanpakken.

Curriculum Vitae

Francesca Tasca was born on February 23, 1993, in Marostica, Italy. She holds a Bachelor's degree in Biotechnology from the University of Padova, Italy, which she obtained between 2012 and 2015. She then pursued a Master's degree in Cellular and Molecular Biotechnology at the University of Trento, Italy, from 2015 to 2017. As part of her master's studies, Francesca also had the opportunity to attend an Erasmus exchange program at the University of Coimbra in Portugal. Following her Erasmus exchange, Francesca completed a six-month research internship at the Leiden University Medical Center in the Laboratory of Genome Editing, under the supervision of Dr. M.A.F.V. Gonçalves. In September 2017, Francesca successfully graduated with first-class honors (110/110L) from her Master's program.

In March 2018, Francesca started her PhD as part of the Marie Skłodowska-Curie Doctoral Network IMGENE - Improving Genome Editing Efficiency (European Union's Horizon 2020 Programme). She worked under the supervision of Dr. M.A.F.V. Gonçalves in the Genome Editing laboratory, part of the Virus and Stem Cell Biology group of Prof. Dr. Rob Hoeben. The results of her research are presented in this thesis.

List of Publications

1. **Tasca F.***, Wang Q.*, Gonçalves M.A.F.V. Adenoviral Vectors Meet Gene Editing: A Rising Partnership for the Genomic Engineering of Human Stem Cells and Their Progeny. Cells. 2020; 9:953. doi: 10.3390/cells9040953. * Shared first co-authorship.

2. **Tasca F.**, Brescia M., Liu J., Janssen J.M., Mamchaoui K., Gonçalves M.A.F.V. High-capacity adenovector delivery of forced CRISPR-Cas9 heterodimers fosters precise chromosomal deletions in human cells. Molecular Therapy Nucleic Acids. 2023; 31:746-762. doi:10.1016/j.omtn.2023.02.025

3. **Tasca F.**, Brescia M., Wang Q., Liu J., Janssen J.M., Szuhai K., Gonçalves M.A.F.V. Large-scale genome editing based on high-capacity adenovectors and CRISPR-Cas9 nucleases rescues full-length dystrophin synthesis in DMD muscle cells. Nucleic Acids Res. 2022; 50:7761-7782. doi: 10.1093/nar/gkac567.

4. Chen X.*, **Tasca F.***, Wang Q., Liu J., Janssen J.M., Brescia M.D., Bellin M., Szuhai K., Kenrick J., Frock R.L., Gonçalves M.A.F.V. Expanding the editable genome and CRISPR-Cas9 versatility using DNA cutting-free gene targeting based on in trans paired nicking. Nucleic Acids Res. 2020; 48:974-995. doi: 10.1093/nar/gkz1121. * Shared first co-authorship.

5. Wang Q., Liu J., Janssen J.M., **Tasca F.**, Mei H., Gonçalves M.A.F.V. Broadening the reach and investigating the potential of prime editors through fully viral gene-deleted adenoviral vector delivery. Nucleic Acids Res. 2021; 49:11986-12001. doi: 10.1093/nar/gkab938.

Acknowledgements

As I reach the end of my PhD journey, I reflect on the challenging and exciting moments that have passed, and the wonderful people I have met along the way. I am grateful for the path that has brought me here, and I couldn't have been happier with the people who were with me during this journey. I am deeply thankful for their support, encouragement, and friendship, which have made this experience truly unforgettable.

I am thankful to Rob Hoeben for his input and support during my PhD training. I would also like to express my deepest gratitude to my supervisor, Manuel Gonçalves, for his invaluable guidance and support throughout my PhD journey, for challenging me to think critically, and for always pushing me to do better. I am fortunate to have had the opportunity to work under their supervision and will always cherish the memories and lessons learned during this time.

I am grateful to my lab mates who have supported me day after day in my research. A special thank you to Marcella, my paranimf, who has always been there for me, from discussing scientific matters to offering moral support and sharing a good laugh. Your openness to mentor me and your positivity have made my PhD journey unique and have helped shape me into the scientist that I am today. I would also like to thank Igno, who welcomed me in Leiden, and showed me the ropes of both the lab and life in the Netherlands. Special thanks to Kim and Jin, who have always helped me in the lab and were always available for a chat in the office. I would like to give a thanks to Hidde for his camaraderie and thank you to my colleagues Chen, Qian, Li, and Xiaoling for their discussions, and kindness.

I am grateful to all the members of the IMGENE consortium who have inspired me to think outside the box and pushed me to challenge myself scientifically and personally.

Thank you to all the former and current members of the Department of Cell and Chemical Biology and in particular to the member of the Virus and Stem Cell Biology group (VSB) for the discussion and the help. Thank you to Timo, Vera, Sanne, Sophia, Willemijn, Selas, Diana, Steve, Martijn and Iris.

I got to give a huge shoutout to all the friends in Leiden. From the borrels in the terrace to the lunches, dinners and parties you made me forget about the stress and just enjoy life. Thanks to Dylan, Ale, David, Virginie, Frederick, Thilo, Ale, Ila, Sebastain, Robyn, Nicolette, Daniel, Rossella, Jesse, Mike and all the others wonderful people I had the opportunity to share this years with. I would also like to express my gratitude to Nanda, my paranimf, who has been a supportive friend since the beginning not only during my PhD but throughout my life in the Netherlands. I also want to thank Leo, who helped me through the quarantine period and was always ready to listen and offer advice. Special thanks go to the B.W. group Nanda, Leo, and Marghe, who shared the burden and reminded me of the importance of life beyond the lab.

Thanks to Pietro that from Italy to the Netherlands has been supportive in my adventures. Thanks to my lovely neighobours Silvia and Phil for being a certainty in always sticking to the

plan and to Tiso for always pointing out the bright side of every situation.

Thanks to Andy that more than anybody else hold my hand throughout this journey.

Thank you also to the M.U.G. group, that I can always count on. And thanks to my Italian friends Annalisa and Matilde for always being close despite the distance.

Finally thanks to my family that never doubt me and supported me throughout this PhD. Thank you to my mother and father for inspiring me to always explore outside my limits and thank you to my brother for being on my side while I navigated the water of the academic life.