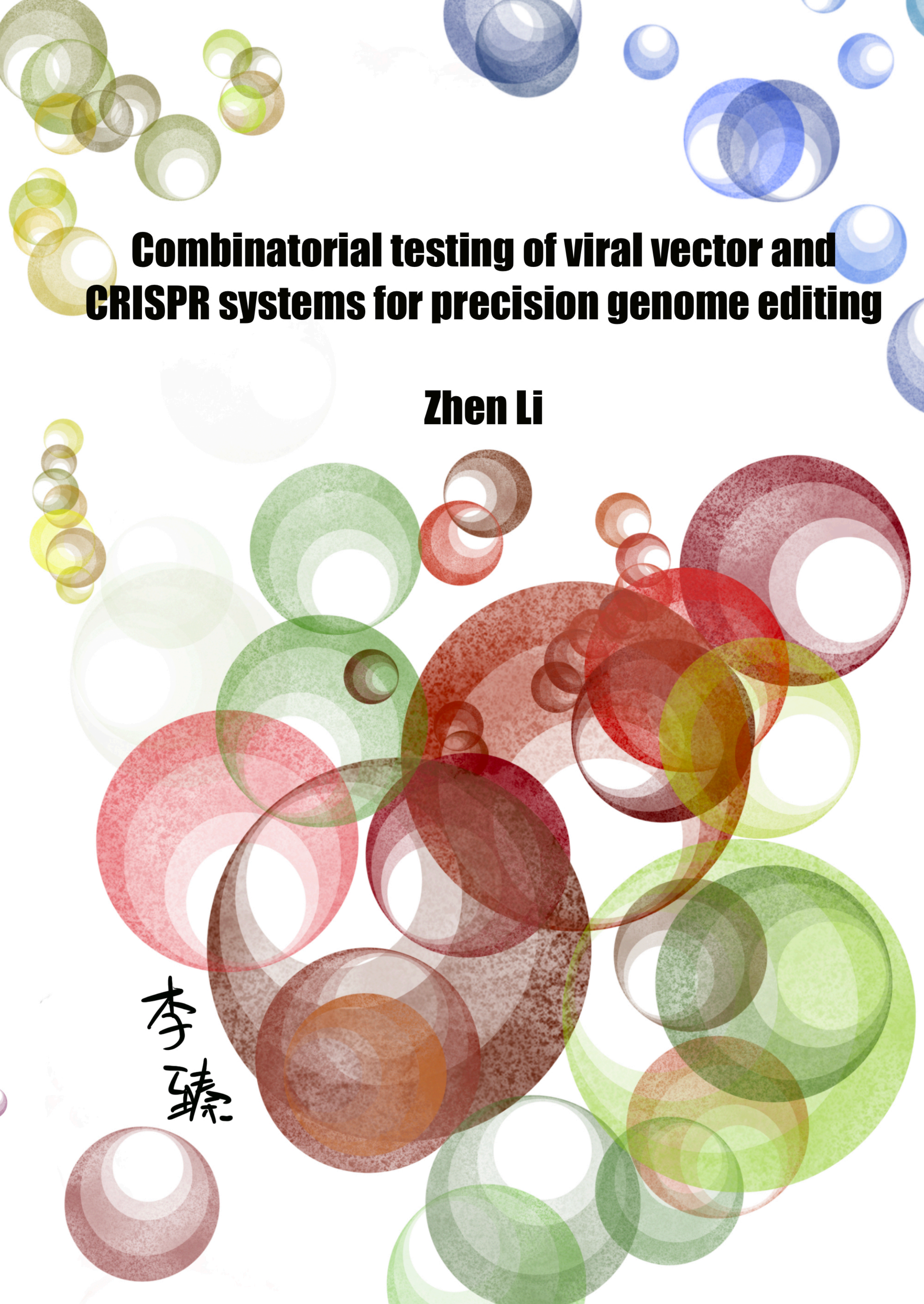
Combinatorial testing of viral vector and CRISPR systems for precision genome editing

Zhen Li



Combinatorial testing of viral vector and CRISPR systems for precision genome editing

Zhen Li

Combinatorial testing of viral vector and CRISPR systems for precision genome editing

Zhen Li

ISBN: 978-94-6522-568-5

The research presented in this thesis was supported by the China Scholarship Council - Leiden University joint program, the Prinses Beatrix Spierfonds, the Duchenne Parent Project NL, the Dutch Research Council (NWO) - Open Technology Program, and EU Marie Skłodowska-Curie Doctoral Network Actions.

Cover: concept and design by Zhen Li

Layout: Zhen Li

Printing: Ridderprint, www.ridderprint.nl

Copyright © Zhen Li, 2025, Leiden, the Netherlands. All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means without written permission of the copyright owner.

Combinatorial testing of viral vector and CRISPR systems for precision genome editing

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 woensdag 8 oktober 2025
klokke 10:00 uur

door

Zhen Li

geboren te Sanya, Hainan, China
in 1993

Promotor:

Prof. Dr. R.C. Hoeben

Co-promotor:

Dr. M.A.F.V. Gonçalves

Leden Promotiecommissie:

Prof. Dr. S.M. Chuva de Sousa Lopes

Prof. Dr. W.W.M. Pijnappel (Erasmus MC, Rotterdam)

Prof. Dr. M.H.M. Heemskerk

Prof. Dr. P. ten Dijke

Dr. Z.Y. Lei (University Medical Center Utrecht)

敬谢家人长久相伴

愿春祺夏安，秋绥冬禧

To my family for their love and trust

Contents

Chapter 1 General introduction 9

Chapter 2 "Soft" genome editing using CRISPR nickases as a potential source of safer cell products 15

Chapter 3 AAV-vectored base editor *trans*-splicing delivers dystrophin repair 31

Chapter 4 Precision genome editing using combinatorial viral vector delivery of CRISPR-Cas9 nucleases and donor DNA constructs 39

Chapter 5 Selector AAV-CRISPR vectors purge off-target chromosomal insertions and promote precise genome editing 97

Chapter 6 Summary and general discussion 151

Chapter 7 Nederlandse Samenvatting 161
- *Curriculum Vitae*
- *List of Publications*
- *Acknowledgements*

Chapter 1

General introduction

General introduction and outline of this thesis

Nowadays, advanced tools deployed for genome editing (GE) mainly derive from the CRISPR/Cas9 system found to be in 2012 as a prokaryotic RNA-guide antiviral system. Typically, Cas9 proteins or engineered Cas9 proteins, *e.g.*, nucleases, can make endogenously site-specific double-stranded DNA breaks (DSBs) which are predominantly repaired via the non-homologous end joining (NHEJ) DNA repair pathway leading to the introduction of mutagenic insertions/deletions (indels). Hence, NHEJ-mediated GE can lead to the removal of pre-existing genetic information, *i.e.*, knock-out (KO). Alternatively, in the presence of exogenous donor DNA templates, *e.g.*, a transgene cassette flanked by DNA sequences identical to those present around the targeted DSB made by engineered RNA-guided nucleases, can result in gene knock-in (KI) through the homology-directed repair (HDR) DNA repair pathway. Differently from DSBs made by RNA-guided Cas9 nucleases, there are Cas9 protein variants that induce instead single-strand DNA breaks (SSBs), or nicks, by virtue of having the D10A or H840A mutations that result in the catalytic inactivation of their RuvC or HNH nuclease domains, respectively. Significantly, in contrast to DSBs, nicks are not substrates for error-prone NHEJ processes. Furthermore, nicks caused by nicking Cas9 proteins, *i.e.*, nickases, can also serve as stimuli for HDR in the presence of exogenous donor DNA templates for the purpose of achieving KI at target chromosomal sites. Ideally, SSB-based KI strategies can alleviate or even avoid the buildup of indels caused by DSBs, and hence it can be regarded as a preferable GE strategy choice in the regards to therapeutical applications.

In **Chapter 2**, the pros and cons of GE involving DSB-based homology-directed repair are elaborated. We discuss the pressing need for the development and application of less mutagenic GE procedures, namely, via using DSB-independent nickases which, as aforementioned, lead only to residual amounts of mutagenic indels and can be tailored together with matched donor DNA templates for precise SSB-mediated HDR. As such, SSB-mediated HDR can serve as a valuable approach for editing therapeutically relevant cells, *e.g*., induced pluripotent stem cells (iPSCs) and T cells, leading to the manufacturing of potentially safer cell products for transplantation purposes. In **Chapter 3**, we further elaborate on the use of nickase variants in the form of base editors directed at therapeutic GE applications. In particular, we commented on the work conducted by Chai *et al.* (Mol. Ther. Nucleic Acids. 2023 32:522-535), investigating repair of defective *DMD* alleles causing Duchenne muscular dystrophy (DMD) via adeno-associated viral (AAV) vector delivery of *trans*-splicing base editors *in vitro* and *in vivo* settings. Instead of testing AAV/CRISPR-Cas9-based *DMD* reading frame restoration, associated with the potentially prevalent capture of Cas9-encoding AAV at on- and off-target nuclease sites, the authors investigated dual AAV delivery of *trans*-splicing adenine-base editor constructs yielding a Cas9 nickase fused to an adenine deaminase that, together with a cognate gRNA, mediates adenine (A) to guanine (G) substitutions. The resulting substitutions disrupt splice site motifs leading to exon skipping of *DMD* exon sequences bearing premature stop codons. On the top of overcoming the limited cargo capacity of AAV vectors (*i.e.*, 4.7 kb), such dual AAV method provides for a DSB-free *DMD* reading frame correction option for treating DMD patients whose disease is caused by out-of-frame deletions.

Even though, conceptually, one expects installing desired therapeutic gene edits with the aid of different GE tools, it has become patently clear the challenge of achieving efficient and safe delivery of the necessary tools into target cells, especially those cells resistant to transfection or relevant to therapeutical applications. Moreover, one needs to contend with the fact with the increasing precision of the most advanced GE tools there is a concomitant trend towards larger size increase in the associated machineries, which further hampers their delivery efficiency. Taken these delivery and safety aspects together, we reasoned that combining distinct engineered viral vectors lacking all viral genes could serve as sources of GE tools upon robust transduction of target cells. In particular, based on the complementary characteristics of adenoviral (AdV) and adeno-associated viral (AAV) vectors we postulated that a robust and precise GE system could be assembled by combining the former and latter platforms for delivering CRISPR/Cas9 reagents and donor DNA templates into human cells, respectively.

Indeed, contrary to linear uncapped AAV genomes, terminally capped double-stranded AdV genomes are not prone to off-target DNA insertions making them suitable for the delivery of Cas9 enzyme constructs since it is especially important to prevent chromosomal integration of these constructs to minimize the buildup of off-target effects. In addition, also in contrast to AAV vectors, AdV vectors can accommodate full-length Cas9 constructs together with single or multiple gRNAs. Conversely, AAV genomes are proficient substrates for HDR possibly owing to the AAV DNA structure featuring secondary-structured inverted terminal repeats flanking single-stranded DNA. Besides allocating AdV and AAV systems for the delivery of, respectively, Cas9 constructs and distinct types of donor DNA templates, the role of distinct AAV donor DNA structures on the efficiency and accuracy of genome editing was investigated. Hence, in **Chapter 4**, we put forward and systematically evaluate different iterations of this dual viral vector system in HeLa cells, human mesenchymal stem cells (hMSCs) and skeletal muscle progenitor cells (myoblasts) in terms of their efficiencies, specificities, and accuracies. In this context, we extend earlier observations showing that 3rd generation, fully viral gene-deleted AdV vectors have a notoriously dampened cytotoxicity profile when compared to that of their 2nd generation counterparts.

In **Chapter 5**, we build on the dual viral vector platform introduced in **Chapter 4** by investigating its compatibility with a marker-free co-selection system for selecting gene-edited cells and, simultaneously, purging imprecisely edited cells via ouabain selection. The main sub-unit of the ubiquitously expressed sodium/potassium pump (Na^+/K^+ ATPase) is encoded by *ATP1A1* whose ATPase product is responsible for many physiological functions including osmotic regulation. Interestingly, specific point mutations in *ATP1A1* conferring ouabain resistance are naturally found in the human population without disrupting the regular physiological functions of the Na^+/K^+ ATPase. Hence, aiming at improving the performance of AAV-based gene editing procedures, we sought to investigate AAV donor constructs harboring marker-free co-selection components (selector AAV vectors) permitting ouabain-dependent enrichment for genome-edited cells. We demonstrate that combining selector AAV vectors with ouabain treatments, in addition to enriching for genome-edited cell populations, eliminates imprecise on-target edits and off-target and/or random donor DNA insertions from said populations. Importantly, selector AAV vector titration experiments showed that the highest fold-enrichment factors for genome-edited cell fractions are associated with the lowest vector input amounts. This finding is expected to be beneficial for alleviating AAV vector production costs, off-target donor insertions and P53-dependent activation of the DNA damage response linked to AAV DNA, which is known to be particularly deleterious in stem cells with scientific and therapeutic relevance, *e.g.*, induced pluripotent stem cells and hematopoietic stem cells.

Chapter 2

“Soft” genome editing using CRISPR nickases as a potential source of safer cell products

Zhen Li[1] and Manuel A.F.V. Gonçalves[1]

Cell & Gene Therapy Insights 9:1201-1210 (2023)

[1]Leiden University Medical Centre, Department of Cell and Chemical Biology, Einthovenweg 20, 2333 ZC, Leiden, the Netherlands.

Abstract

The integration of the gene and cell therapy fields through the application of genome editing principles permits generating *ex vivo* transplantable grafts from stem cells or from their differentiated progenies (*e.g.*, T and NK cells) with novel genetically-engineered function(s). As such, these technologies are offering new therapeutic avenues to previously intractable inherited and acquired disorders (*e.g.*, malignant and infectious diseases). In this article, we discuss the main characteristics, advantages and limitations of genome editing involving the targeted chromosomal insertion of transgenes upon site-specific double-stranded DNA break (DSB) formation by programmable nucleases, namely, RNA-programmable CRISPR nucleases. Subsequently, building on this information and recent findings, we put forward the view that targeted transgene insertion strategies based on CRISPR nickases, as opposed to nucleases, address important limitations of conventional DSB-dependent genome editing approaches. In particular, the cytotoxicity and high genotoxicity resulting from DSBs especially in cell types highly sensitive to DNA damage, including pluripotent and hematopoietic stem cells.

Background

Genome editing or genome engineering is a fast-evolving field with growing impact on basic science, biotechnology, and medicine [1]. Particularly versatile genome editing strategies consist of inserting exogenous donor DNA constructs into specific genomic loci (knock-in) subjected to double-stranded DNA breaks (DSBs) made by engineered nucleases derived from class 2 type II CRISPR systems consisting of single guide RNA (gRNA) and Cas9 ribonucleoprotein complexes (CRISPR-Cas9 nucleases) [2]. This versatility stems from the robust activity and straightforward designing of these RNA-programmable nucleases and the amenability of gene knock-in strategies to genomic modifications spanning entire transgenes, including those encoding chimeric antigen receptors (CARs) and T-cell receptors (TCRs) alone or together with auxiliary factors, such as positive-selection markers and safety genetic switches [3,4]. Indeed, notwithstanding the growing mining for and adaption of CRISPR and CRISPR-like systems for genome editing purposes, engineered CRISPR-Cas9 nucleases based on the prototypic *Streptococcus pyogenes* CRISPR system and their molecularly evolved or structurally-guided designed variants (*e.g*., high-specificity and targeting range-expanded variants), remain leading tools for a wide variety of genome engineering applications [5,6].

Main attributes of CRISPR nuclease-assisted genome editing

Chromosomal gene knock-in procedures often entail the delivery of CRISPR-Cas9 nucleases together with donor DNA constructs designed as substrates for either homology-independent or homology-dependent repair (HDR) pathways [7]. Generally, HDR-mediated transgene knock-ins are more precise than those resulting from homology-independent processes in that they are naturally inserted at the chromosomal target site in a predefined orientation and present neither multiple-copy insertions nor imprecise 'foot-prints' at the junctions between genomic and exogenous DNA sequences [8,9]. Importantly, as HDR takes place during the late G2 and S phases of the cell cycle, therapeutically relevant dividing cell types, such as induced pluripotent stem cells (iPSCs), natural killer (NK) cells and T lymphocytes, are amenable to precise HDR-mediated genome editing. For instance, in what valuable target cells is concerned, genetically engineered CAR-T cells, serving as personalized 'living drugs', are yielding impressive results in terms of treating CD19-positive hematological malignancies [3,10]. This is so despite their high costs that stem in part from the difficulties in generating the large amounts of the respective engineered cell products. Since 2017, a growing number of these CAR-T cell products have in fact started to be approved by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) [11]. Building on the resulting CD19-targeted cancer therapy datasets, over 500 CAR-T cell therapies directed at different antigens in liquid and solid tumors are currently undergoing clinical testing worldwide [11,12]. Significantly, CAR-NK cells are also entering the adoptive immunotherapy field as a potential alternative to CAR-T cells owing to their intrinsic tumor-cell killing activity and fewer adverse effects in patients [13]. Yet, regardless of the target cell type, in what the genetic modification procedures are concerned, the adoptive immunotherapy field is moving from randomly integrating retroviral vector and DNA transposon systems to targeted transgene insertion approaches using programmable nucleases [3,10]. In contrast to unpredictable *CAR* or *TCR* donor construct integration, programmable nuclease-assisted genome editing assures stable and homogeneous transgene expression while minimizing insertional mutagenesis risks inherent to randomly integrating vehicles. In fact, in contrast to random, targeted *TCR* transgene insertion leads to predictable T-cell function *in vivo* [14]. In this context, genomic loci generically dubbed 'safe harbor' are particularly appealing endogenous landing pads for *CAR* and *TCR* transgenes as insertions at these sites minimize the chances for gene silencing or variegated transgene expression while preserving the endogenous transcriptome of engineered cells [15,16].

Main limitations of CRISPR nuclease-assisted genome editing

As aforementioned, programmable nucleases and HDR-tailored donor DNA constructs yield precise gene knock-ins. However, a major limitation regarding the use of programmable nucleases is the fact that, in mammalian cells, DSBs are prevalently repaired via mutagenic non-homologous end joining (NHEJ) or microhomology-mediated end joining (MMEJ) processes instead of accurate HDR [17,18]. Moreover, in contrast to HDR, end-joining processes take place throughout the cell cycle. As a result, amongst cells exposed to donor constructs and programmable nucleases, the vast majority contains one or both target alleles disrupted by NHEJ- or MMEJ-derived small insertions and deletions (indels). This mutagenic burden, in the form of indel 'footprints', can lead to target protein imbalances and cell fitness losses [19]. In addition, on-target DSB formation can also yield translocations and gross chromosomal rearrangements [19–22]. Recent studies have further uncovered that on-target DSBs are capable of triggering extensive chromosome fragmentation followed by haphazard reassembly (chromothripsis) [23,24] and the partial or entire loss of chromosomes (aneuploidy) [25]. Notably, the chromothripsis and aneuploidy phenomena were readily detected in T cells and hematopoietic progenitor cells subjected to CRISPR-Cas9 nuclease reagents used in clinical trials [23,25]. Notwithstanding these phenomena, recent findings are more reassuring in that, contingent upon gRNA target site selection, chromosomal losses in particular can be substantially minimized by inducing DSB formation before, as opposed to after, the activation/stimulation of the primary T-cell populations [26].

Finally, on-target DSBs trigger P53-dependent cell cycle arrest and apoptosis which limits the efficacy of HDR-mediated genome editing in regular P53-positive cells [27,28], and creates selective pressure for the emergence of mutations associated with tumorigenesis. Related to the latter matter, during sub-culturing, pluripotent stem cells can acquire ‘spontaneous’ tumor-associated P53 mutations in a recurrent fashion [29] which, by virtue of being more resistant to DSBs, are in principle more prone to expansion than their wild-type counterparts once exposed to programmable nucleases. Indeed, CRISPR-Cas9 nuclease activation of certain signaling pathways can lead to the selection of cells with potentially harmful loss-of-function or dominant-negative mutations in the tumor-suppressor P53 transcription factor or gain-of-function mutations in the KRAS oncoprotein [27,30]. Furthermore, recent mouse modelling experiments indicate that p53 mutant cells, rather than proceeding to malignancy via a haphazard route, are instead subjected to an unexpectedly more deterministic set of genetic instability events [31]. Together, these cytotoxic and genotoxic effects raise tangible concerns on the use of programmable nucleases for the optimal generation of autologous genetically-corrected cell products.

Rationale for "soft" genome editing based on CRISPR nickases

Although emerging high-specificity programmable nucleases can greatly minimize off-target DNA cleavage, *e.g.*, eSpCas9(1.1) [32] and Cas9-HF1 [33], they are inherently incapable of eliminating the potentially deleterious effects resulting from on-target DSB formation. Therefore, the substantial genotoxicity and cytotoxicity profiles associated with conventional nuclease-assisted genome editing create a pressing need for the development of alternative genetic engineering systems that reliably generate safer and functionally robust cell products. Indeed, DSB-dependent genome editing is expected to be particularly risky in the context to cell therapies based on the transplantation of populations of genetically engineered pluripotent stem cells, T lymphocytes and NK cells. The reasons are twofold. Firstly, in the context of extensive *ex vivo* cell amplification protocols underpinning the generation of these cell transplantation products, DSB-derived mutations and/or chromosomal rearrangements can cooperate in cell transformation and clonal expansion. Secondly, in instances where targeting multiple genes is needed for achieving a robust anti-tumor effect, *e.g.*, via combinatorial exogenous *CAR* transgene knock-ins and endogenous *TCR* or programmed cell death protein 1 (*PD1*) gene knockouts, simultaneous induction of the attendant DSBs at different genomic positions is expected to exacerbate the levels of undesirable genome editing by-products in the form of translocations and chromosomal rearrangements. In this context, investigations exploring alternative HDR-mediated gene knock-in approaches that rely on sequence- and strand-specific nucleases ('nickases') are valuable in that the resulting single-stranded DNA breaks (SSBs), or nicks, are substrates for neither NHEJ nor MMEJ. As a corollary, the balance between precise HDR to undesired end-joining events are dramatically biased towards the former. Moreover, although genomic SSBs are, *per se*, poor HDR stimuli, earlier experiments from our laboratory using the native adeno-associated virus Rep68/78 nickases demonstrated that concomitant SSB formation at acceptor sequences and donor DNA constructs fosters HDR-mediated gene knock-in at an endogenous human locus, namely, the prototypic safe harbour locus *AAVS1* [34]. The application of this generic *in trans* paired nicking (ITPN) principle was subsequently expanded to other genomic sequences through the use of more versatile RNA-programmable CRISPR-Cas9 nickases [35,36] that are simply obtained through site-directed mutagenesis of one of the two nucleases domains of the parental Cas9 protein (*i.e.*, HNH or RuvC) (**Figure 1**) [37]. Indeed, by stimulating otherwise inefficient SSB-dependent HDR, ITPN approaches based on the delivery of nicking CRISPR-Cas9 complexes and matched nickase-susceptible HDR donor constructs, are valuable for seamless and scarless chromosomal editing, including at multiple-copy or essential genomic tracts [19,38]. Additional examples regarding the application of ITPN methodologies in various mammalian cell types, *e.g.*, iPSCs, keratinocytes and organoids featuring normal or cancer traits, encompass: (i) repairing or installing predefined gene mutations [35,38-41], (ii) maximizing the integrity of unmodified alleles during allele-specific gene editing [42,43], and (iii) streamlining one-step biallelic gene editing or one-step multiplexing gene knock-in or tagging [35,44,45]. It is equally worth mentioning that, in contrast to regular and high-specificity CRISPR-Cas9 nucleases, CRISPR-Cas9 nickases constitute poor P53-dependent signalling triggers in human cells, including in DNA damage-sensitive iPSCs [38,40]. Hence, it is expected that the aforementioned growing mining for CRISPR systems buried in large genomic and metagenomic databases, will start unearthing enzymes that, via either their intrinsic or engineered nicking activities, enlarge the toolset for DSB-free genome editing. Examples include

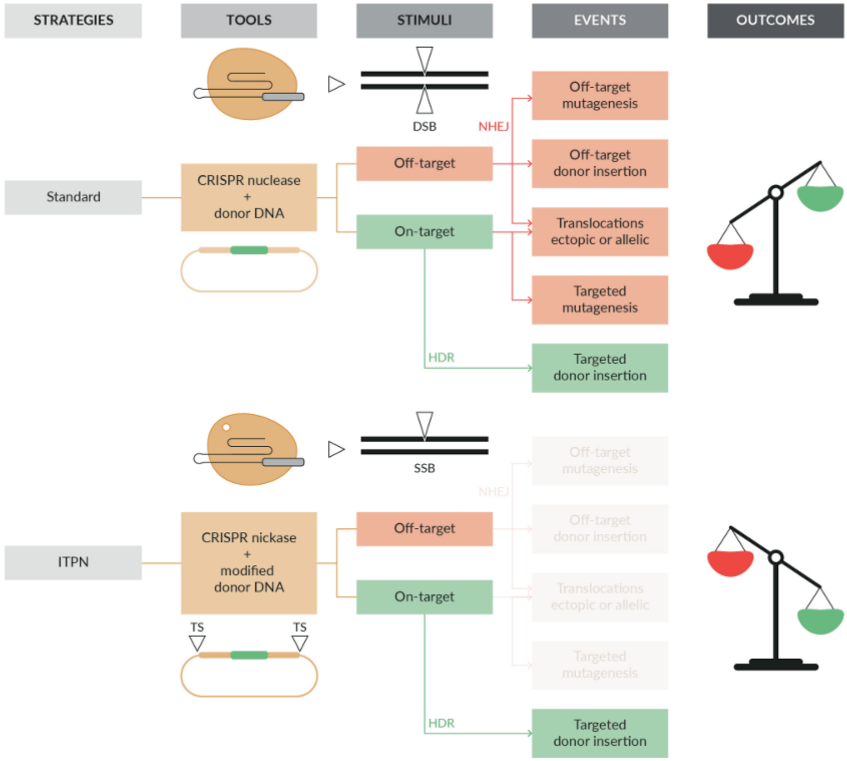


Figure1. Standard versus *in trans* paired nicking genome editing. The relative weights of desired and undesired genome editing outcomes derived from, respectively, homology-directed repair (HDR) and imprecise events caused by competing end-joining DNA repair pathways, *e.g.*, non-homologous end joining (NHEJ), are illustrated. DSB and SSB, double-stranded and single-stranded chromosomal breaks, respectively; 'nickases', sequence- and strand-specific nucleases. In contrast to standard donor constructs, modified donor constructs have nickase-susceptible target sites (TS) framing their targeting modules consisting of exogenous DNA (green bar) flanked by sequences homologous to the genomic target region ('homology arms').

the HNH-negative IsrB nickase derived from the ancestral CRISPR-like system OMEGA and the RuvC-only CRISPR class 2 type V Cas12i nuclease that nick and preferentially nicks, respectively, double-stranded DNA substrates [46–48]. Moreover, often, newly discovered CRISPR systems also yield genome editing components whose small sizes renders them more fitting for delivery through commonly used adeno-associated viral (AAV) vectors [49].

Finally, another recent ‘soft’ HDR-mediated genome editing concept that might be particularly suited for the repair of heterozygous or dominant mutations involves allele-specific chromosome nicking for the stimulation of interhomolog recombination (IHR) in somatic cells [50,51]. Through this process of allelic conversion, a pathogenic mutation in one allele can, in principle, be corrected using as donor template the endogenous ‘healthy’ allele (**Figure 2**). This elegant concept of using CRISPR-Cas9 nickases and endogenous homologous chromosomal DNA as repairing templates has been demonstrated in *Drosophila* models [51] and human cell lines [50,52]. Regarding the application of such exogenous donor DNA-free genome editing principles in human cells, recent investigations argue for multiplexing approaches in which primary allelic-specific gRNAs act in concert with secondary gRNAs to direct *in trans* paired nicking of homologous chromosomes and ensuing allelic conversion via IHR (**Figure 2**) [52]. Further research will be instrumental to advance CRISPR-Cas9 nickase-induced IHR from enticing proof-of-concept studies in cell lines to its application in human stem and progenitor cells.

Translation Insight & Outlook

There is a pressing need for investigating and validating alternative DSB-free and precise genome editing tools and strategies in various stem and progenitor cell types, *e.g.*, *bona fide* T and NK cells as well as precursor iPSCs from which different effector cells can be differentiated, including immunotherapeutic T and NK cell candidates. Genome engineering strategies covering targeted and precise chromosomal incorporation of genetic payloads with varying sizes will become ever-more relevant. In this regard, CRISPR nickases *per se* and fused to reverse transcriptases offer a complementary

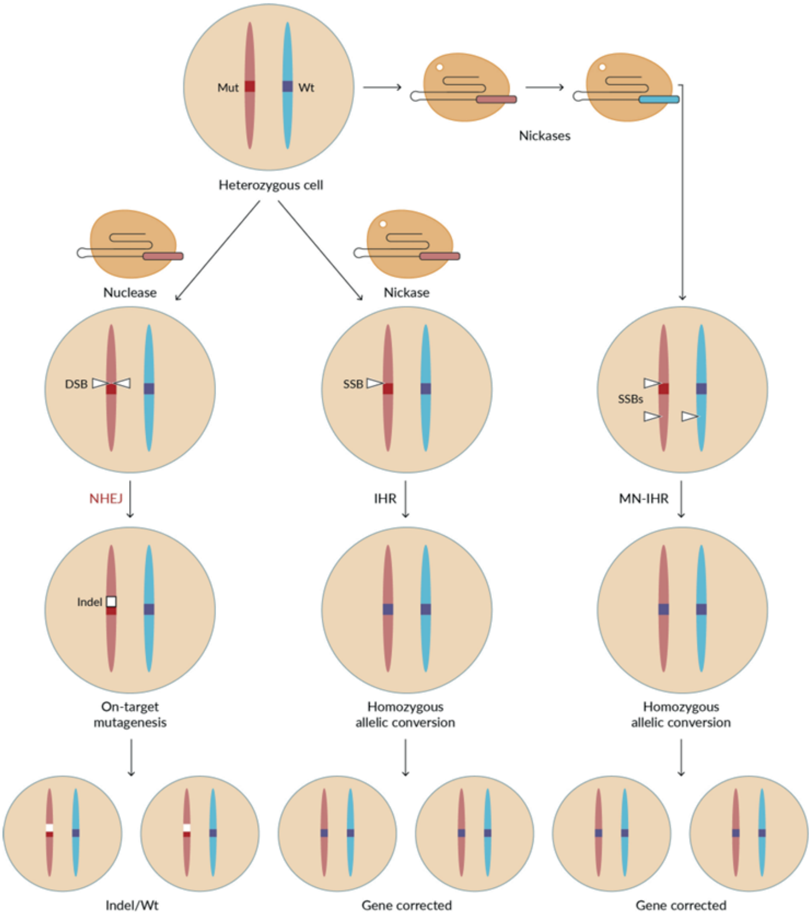


Figure2. Gene correction via interhomolog recombination between heterozygous allelic sequences. Interhomolog recombination (IHR) characteristic of meiosis in germ cells can be fostered in somatic cells subjected to allele-specific double-stranded DNA breaks (DSB), yet the major products are on-target mutagenesis in the form of NHEJ-derived small insertions and deletions (indels). In contrast, allele-specific single-stranded DNA breaks (SSB) can equally foster IHR in somatic cells especially when using multiplexing CRISPR-Cas9 nickases for *in trans* multiple nicking IHR (MN-IHR). In somatic cells with heterozygous mutations or compound heterozygous mutations (not shown) underlying genetic disorders, CRISPR-Cas9 nickase-induced IHR offers the prospect for new genetic therapy interventions via wild-type allele-templated gene repair.

toolbox for ‘soft’ genome editing involving HDR and prime editing, respectively. Contrary to HDR, prime editing does not require the transfer of donor DNA substrates and allows for genomic insertion of up to ~44-bp of foreign DNA despite the need for substantial optimization of extended prime-editing gRNAs (pegRNAs) [53]. Moreover, in contrast to HDR-based genome editing, prime editing can take place in post-mitotic cells albeit to lower efficiencies than in cycling cells [54]. Recent prime editing developments include the combinatorial use of dual pegRNAs and site-specific recombinases designed for replacing genomic sequences with up 250-bp of foreign DNA and inserting entire transgenes at prime editor-defined recombinase target sites, respectively [53]. Despite powerful and versatile, such combinatorial strategies require the delivery of large and multicomponent reagents into target cells. An aspect warranting attention when considering multiplexing approaches concerns the importance of introducing balanced amounts of the attendant individual components to maximize the performance and precision of genome editing interventions [55]. In addition, prime editing involving the delivery of dual pegRNAs is not compatible with large edits whereas sequential prime editing and site-specific recombination is not amenable to subtle genomic edits underlying endogenous gene repair due to discontinuous ‘footprint’ installation in the form of recombinase target sites.

In conclusion, considering the herein discussed findings and matters, one can submit that cell therapy products derived from the use of RNA-programmable nickases as such or with heterologous domains, will offer a complementary set of ‘soft’ genome engineering options whose safety profiles are potentially higher than those associated with the exposure of cells to programmable nucleases.

Contributions

The named author takes responsibility for the integrity of the work as a whole, and has given his approval for this version to be published.

Acknowledgements

None

Disclosure and potential conflicts of interest

Dr. Gonçalves is a member of the European Reference Network - Neuromuscular Diseases (ERN EURO-NMD).

Funding declaration

Research in the authors' laboratory is supported by the Prinses Beatrix Spierfonds, the Duchenne Parent Project NL, the Dutch Research Council (NWO) - Open Technology Program, and EU Marie Skłodowska-Curie Doctoral Network Actions. Dr. Li holds a PhD fellowship from the China Scholarship Council - Leiden University Joint Scholarship Program

References

1. Cornu, T.I.; Mussolino, C.; Cathomen, T.. Refining strategies to translate genome editing to the clinic. *Nat. Med*. **2017**, 23, 415–423.
2. Wang, J.Y.; Doudna, J.A.. CRISPR technology: a decade of genome editing is only the beginning. *Science* **2023**, 379, eadd8643.
3. Finck, A.V.; Blanchard, T.; Roselle, C.P.; Golinelli, G.; June, C.H.. Engineered cellular immunotherapies in cancer and beyond. *Nat. Med*. **2022**, 28, 678–689.
4. Sahillioglu, A.C.; Schumacher, T.N.. Safety switches for adoptive cell therapy. *Curr. Opin. Immunol.* **2022**, 74, 190–198.
5. 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.
6. Huang, X.; Yang, D.; Zhang, J.; Xu, J.; Chen, Y.E.. Recent advances in improving gene-editing specificity through CRISPR-Cas9 nuclease engineering. *Cells* **2022**, 11, 2186.
7. 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.
8. 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.
9. He, X.; Tan, C.; Wang, F.; *et al*. 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.
10. Globerson Levin, A.; Riviere, I.; Eshhar, Z.; Sadelain. M.. CAR T cells: building on the CD19 paradigm. *Eur. J. Immunol*. **2021**, 51, 2151–2163.
11. Detela, G.; Lodge, A.. EU regulatory pathways for ATMPs: standard, accelerated and adaptive pathways to marketing authorisation. *Mol. Ther. Methods Clin. Dev*. **2019**, 13, 205–232.
12. Ivica, N.A.; Young C.M.. Tracking the CAR-T revolution: analysis of clinical trials of CAR-T and TCR-T therapies for the treatment of cancer (1997–2020). *Healthcare (Basel)* **2021**, 9, 1062.
13. Vogler, M.; Shanmugalingam, S.; Sarchen. V.; *et al*. Unleashing the power of NK cells in anticancer immunotherapy. *J. Mol. Med. (Berl.)* **2022**, 100, 337–349.
14. Muller, T.R.; Jarosch, S.; Hammel, M.; *et al*. Targeted T cell receptor gene editing provides predictable T cell product function for immunotherapy. *Cell Rep. Med*. **2021**, 2, 100374.
15. Pavani, G.; Amendola, M.. Targeted gene delivery: where to land. Front. Genome Ed.

2021; 2, 609650. Corrigendum: *Front. Genome Ed.* **2021**, 3, 682171.

16. Odak, A.; Yuan, H.; Feucht, J.; *et al.* Novel extragenic genomic safe harbors for precise therapeutic T-cell engineering. *Blood* **2023**, 141, 2698–2712.
17. Lieber, M.R.; Ma, Y.; Pannicke, U.; Schwarz, K.. Mechanism and regulation of human non-homologous DNA end-joining. *Nat. Rev. Mol. Cell Biol.* **2003**, 4, 712–720.
18. Xue, C.; Greene, E.C.. DNA repair pathway choices in CRISPR-Cas9-mediated genome editing. *Trends Genet.* **2021**, 37(7), 639–656.
19. Chen, X.; Tasca, F.; Wang, Q.; *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.
20. 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.
21. 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.
22. Turchiano, G.; Andrieux, G.; Klermund, J.; *et al.* Quantitative evaluation of chromosomal rearrangements in gene-edited human stem cells by CAST- Seq. *Cell Stem Cell* **2021**, 28, 1136–1147.
23. Leibowitz, M.L.; Papathanasiou, S.; Doerfler, P.A.; *et al.* Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat. Genet.* **2021**, 53, 895–905.
24. Amendola, M.; Brusson, M.; Miccio, A.. CRISPRthripsis: the risk of CRISPR/ Cas9-induced chromothripsis in gene therapy. *Stem Cells Transl Med.* **2022**, 11, 1003–1009.
25. Nahmad, A.D.; Reuveni, E.; Goldschmidt, E.; *et al.* Frequent aneuploidy in primary human T cells after CRISPR-Cas9 cleavage. *Nat Biotechnol.* **2022**, 40, 1807–1813.
26. Tsuchida, C.A.; Brandes, N.; Bueno, R.; *et al.* Mitigation of chromosome loss in clinical CRISPR-Cas9-engineered T cells. *bioRxiv* **2023**, 36993359.
27. Ihry, R.J.; Worringer, K.A.; Salick, M.R.; *et al.* p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **2018**, 24, 939–946.
28. 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.
29. Merkle, F.T.; Ghosh, S.; Kamitaki, N.; *et al.* Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. *Nature* **2017**, 545, 229–233.
30. Sinha, S.; Barbosa, K.; Cheng, K.; *et al.* A systematic genome-wide mapping of oncogenic mutation selection during CRISPR-Cas9 genome editing. *Nat. Commun.*

2021, 12, 6512.

31. Baslan, T.; Morris, J.P.4th; Zhao, Z.; *et al.* Ordered and deterministic cancer genome evolution after p53 loss. *Nature* **2022**, 608, 795–802.
32. 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.
33. Kleinstiver, B.P.; Pattanayak, V.; Prew, M.S.; *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **2016**, 529, 490-495.
34. Gonçalves, M.A.; van Nierop, G.P.; Holkers, M.; de Vries, A.A.. Concerted nicking of donor and chromosomal acceptor DNA promotes homology-directed gene targeting in human cells. *Nucleic Acids Res*. **2012**, 40, 3443–3455.
35. Chen, X.; Janssen, J.M.; Liu, J.; *et al.* In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat. Commun*. **2017**, 8, 657.
36. 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.
37. 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.
38. Wang, Q.; Liu, J.; Janssen, J.M.; Gonçalves, M.A.F.V.. Precise homology-directed installation of large genomic edits in human cells with cleaving and nicking high-specificity Cas9 variants. *Nucleic Acids Res*. **2023**, 51, 3465–3484.
39. Kocher, T.; Wagner, R.N.; Klausegger, A.; *et al.* Improved double-nicking strategies for COL7A1-editing by homologous recombination. *Mol. Ther. Nucleic Acids* **2019**, 18, 496–507.
40. Hyodo, T.; Rahman, M.L.; Karnan, S.; *et al.* Tandem paired nicking promotes precise genome editing with scarce interference by p53. *Cell Rep*. **2020**, 30, 1195–1207.
41. Rahman, M.L.; Hyodo, T.; Karnan, S.; *et al.* Experimental strategies to achieve efficient targeted knock-in via tandem paired nicking. *Sci. Rep*. **2021**, 11, 22627.
42. Bollen, Y.; Stelloo, E.; van Leenen, P.; *et al.* Reconstructing single-cell karyotype alterations in colorectal cancer identifies punctuated and gradual diversification patterns. *Nat. Genet*. **2021**, 53, 1187–1195.
43. Fortschegger, K.; Husa, A.M.; Schinnerl, D.; Nebral, K.; Strehl, S.. Expression of RUNX1-JAK2 in Human induced pluripotent stem cell-derived hematopoietic cells activates the JAK-STAT and MYC pathways. *Int. J. Mol. Sci.* **2021**, 22, 1.
44. Bollen, Y.; Hageman, J.H.; van Leenen, P.; *et al.* Efficient and error-free fluorescent

gene tagging in human organoids without double-strand DNA cleavage. *PLoS Biol.* **2022**, 20, e3001527.

45. Hamaker, N.K.; Lee, K.H.. High-efficiency and multilocus targeted integration in CHO cells using CRISPR-mediated donor nicking and DNA repair inhibitors. *Biotechnol. Bioeng*. **2023**, 120, 2419–2440.
46. Hirano, S.; Kappel, K.; Altae-Tran, H.; *et al.* Structure of the OMEGA nickase IsrB in complex with ωRNA and target DNA. *Nature* **2022**; 610, 575–581.
47. Yan, W.X.; Hunnewell, P.; Alfonse, L.E.; *et al.* Functionally diverse type V CRISPR-Cas systems. *Science* **2019**, 363, 88–91.
48. Zhang, H.; Li, Z.; Xiao, R.; Chang, L.. Mechanisms for target recognition and cleavage by the Cas12i RNA-guided endonuclease. *Nat. Struct. Mol. Biol*. **2020**, 27, 1069–1076.
49. Chen, X.; Gonçalves, M.A.. Engineered viruses as genome editing devices. *Mol. Ther.* **2016**, 24, 447–457.
50. Davis, L.; Khoo, K.J.; Zhang, Y.; Maizels, N.. POLQ suppresses interhomolog recombination and loss of heterozygosity at targeted DNA breaks. *Proc. Natl Acad. Sci. USA* **2020**, 117, 22900–22909.
51. Roy, S.; Juste, S.S.; Sneider, M.; *et al*. Cas9/Nickase-induced allelic conversion by homologous chromosome-templated repair in *Drosophila* somatic cells. *Sci. Adv*. **2022**, 8, eabo0721.
52. Tomita, A.; Sasanuma, H.; Owa, T.; *et al.* Inducing multiple nicks promotes interhomolog homologous recombination to correct heterozygous mutations in somatic cells. *Nat. Commun*. **2023**, 14, 5607.
53. Chen, P.J.; Liu, D.R.. Prime editing for precise and highly versatile genome manipulation. *Nat. Rev. Genet.* **2023**, 24, 161–177.
54. 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.
55. 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. *Mol. Ther. Nucleic Acids* **2023**, 31, 746–762.

Chapter 3

AAV-vectored base editor *trans*-splicing delivers dystrophin repair

Zhen Li[1] and Manuel A.F.V. Gonçalves[1]

Cell & Gene Therapy Insights 9:1201-1210 (2023)

[1]Leiden University Medical Centre, Department of Cell and Chemical Biology, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands.

Duchenne muscular dystrophy (DMD) (MIM: 310200) is a severe and frequent neuromuscular disorder (incidence of ~1 in 5,500 boys) caused by mutations in the vast X-linked *DMD* gene (~2.2 Mb), whose largest product, the long rod-shaped 427-kDa dystrophin isoform, is a key structural component of the striated musculature [1]. Contributing to the urgency in the development of currently inexistent DMD therapies is the observation that ~1/3 of cases arise *de novo* through germline mutations, often intragenic deletions that disrupt the mRNA reading frame. Critically, naturally occurring *DMD* gene deletions resulting in in-frame transcripts coding for internally truncated, yet partially functional, dystrophins cause the milder Becker muscular dystrophy (MIM: 300376). Hence, *DMD* gene manipulations yielding Becker-like dystrophins via direct coding sequence reframing or exon skipping have the potential of offering long-lasting therapeutic effects [1]. Toward this end, among other technologies, CRISPR-Cas9 nucleases and adeno-associated viral (AAV) vectors are being investigated for rescuing dystrophin expression upon double-strand DNA break (DSB) formation and ensuing chromosomal end-joining [1]. These experiments demonstrate that AAV/CRISPR-Cas9-based dystrophin restoration can improve striated muscle function in mice; however, a potentially insidious outcome is the identification of prevalent capture of Cas9-encoding AAV at nuclease target sites, including at *Dmd* exons 51 and 53 [2,3]. Moreover, programmable nucleases can trigger other untoward effects, *e.g.*, locus- or chromosome-wide rearrangements [4]. There is, therefore, a pressing need to expand candidate DMD genetic therapies to those based on DSB-independent genome editing systems. In a timely study published in *Molecular Therapy – Nucleic Acids*, Chai and coworkers [5] identify adenine base editors (ABEs) and guide RNAs (gRNAs) (**Figure 1**) that, after implementing single base-pair substitutions (*i.e.*, A·T-to-G·C transitions) at splicing motifs, a process that they name "single-swap" editing, lead to genotype-specific *DMD* repair through exon skipping. Next, the authors assemble a dual AAV ABE *trans*-splicing system to demonstrate in dystrophin-defective mice the amelioration of dystrophic traits at the cellular and organismal levels upon intramuscular or systemic administrations. This study identifies ABE:gRNA complexes compatible with ~30% of DMD-causing genotypes and, notwithstanding its inherent complexity, establishes dual AAV ABE *trans*-splicing as a DSB-free *DMD* gene correction option.

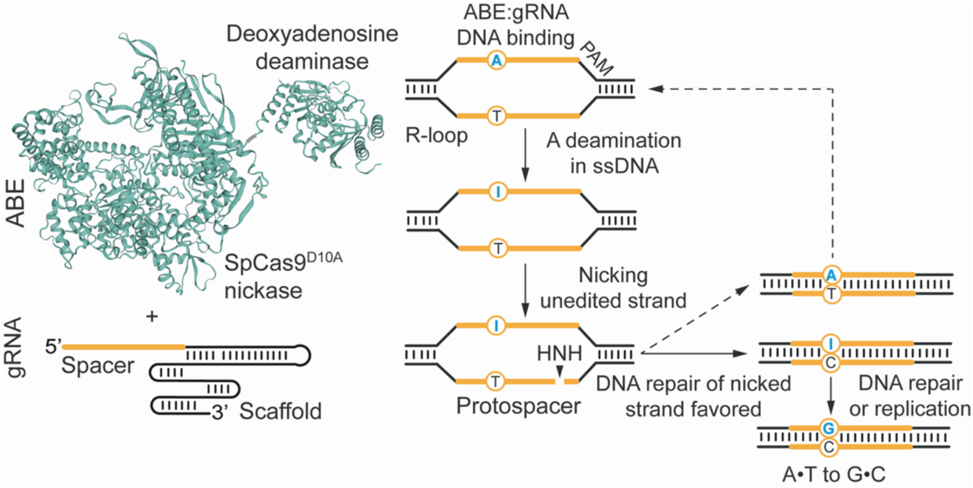


Figure 1. Adenine base editing. Adenine base editors (ABEs) catalyze A·T-to-G·C substitutions and consist of a fusion product between a disabled or nicking Cas9, or ortholog protein, and an evolved deoxyadenosine deaminase, *e.g.*, *Escherichia coli* tRNA adenosine deaminase (TadA) derivatives. Upon PAM binding, ABE:gRNA complexes form an R-loop at a gRNA-defined target sequence exposing a region of single-stranded DNA. A nucleotides in this single-stranded protospacer "bubble" become targets for the ABE effector domain that converts A nucleotides to inosine (I) intermediates preferentially within an "activity window." Subsequently, nicking of the unedited strand induces DNA repair that installs C nucleotides opposite I intermediates with additional DNA repair events (or replication) establishing the final A·T-to-G·C transitions.

Chai and colleagues start by testing in HEK293T cells and DMD iPSCs, ABE:gRNA complexes that, depending on their ABE component, *i.e.*, ABE8e [6] or ABEe-NG, recognize, respectively, canonical NGG or NG PAMs. DNA sequencing assays in DMD iPSCs identified ABE:gRNA complexes yielding high-frequency target-base editing at *DMD* exons 51 and 45 splice acceptor (SA) motifs (71.6% and 79.3%–83.3%, respectively). As a result, robust expression of Becker-like dystrophins was detected in cardiomyocytes differentiated from base-edited DMD iPSCs. Interestingly, it was found that *DMD* exon 51 targeting serendipitously established an in-frame 11-nucleotide deletion instead of the intended exon skipping, presumably due to internal cryptic splice site usage. This finding *per se* stresses the importance of carefully assessing gene-edited products even when using subtle DSB-independent systems. Moreover, various amounts of bystander A·T-to-G·C transitions were also detected (range: 3.3%–91%). These bystander changes might have limited consequences as they map to either spliced-out intron or, if accompanied with the intended SA edits, to skipped exon sequences. Further investigations will be, nonetheless, necessary to probe for slight (or otherwise) splicing alterations in different cell types or contexts. Regardless, base editors with narrower "editing windows" should facilitate more favorable target-to-bystander ratios.

AAV vectors are attractive *in vivo* gene-editing tool delivery vehicles owing to their lack of viral genes and serotype diversity of their parental viruses. Indeed, packaging vector genomes in AAV serotype capsids with a strong tropism for certain cell types (pseudotyping) facilitates tissue-directed transductions. However, the limited AAV packaging capacity (<4.7 kb) permits delivering neither base editing nor other large constructs. To obviate this limitation, researchers are developing base editors with compact architectures [7], testing alternative delivery systems or applying dual AAV strategies in which split constructs linked to N- and C-terminal intein domains are packaged in different AAV vectors (**Figure 2**). Upon target cell co-transductions, intein-mediated protein *trans*-splicing results in the *in situ* assembly of full-length proteins. Indeed, dual AAV-vectored base editor *trans*-splicing is currently undergoing intense investigation for tackling various disease-causing mutations, including *DMD* mutations [8,9]. In Chai *et al*. [5], a dual AAV ABE *trans*-splicing system is assembled to address frequent *DMD* deletions through exon skipping modulation. By exploiting high human-murine conservation over intron 44 to exon 45 junctions and guided by their earlier *in vitro* experiments, the authors apply dual AAVs to deliver a split version of an ABE8e variant (*i.e.*, ABE.TadA-8e^{V106W}), selected for its reduced off-target nucleic acid deaminase activities [6], together with a gRNA targeting the exon 45 SA region (**Figure 2**).

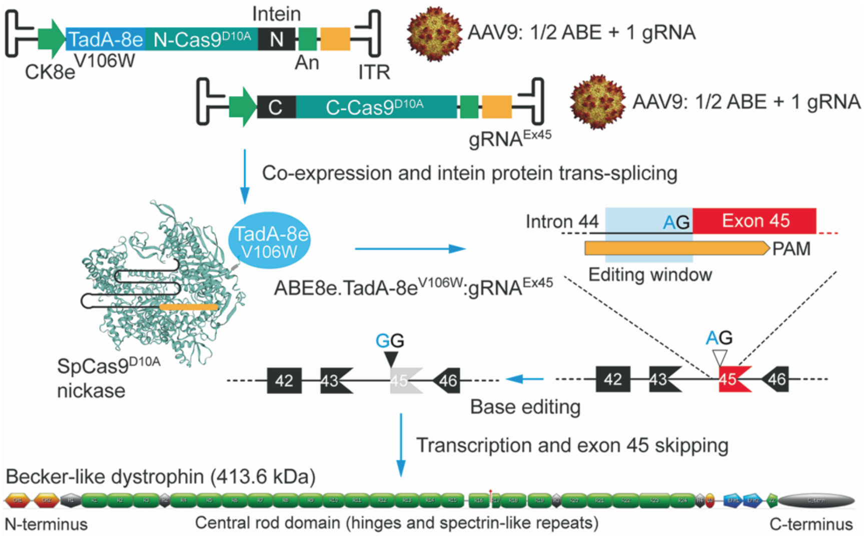


Figure 2. Dual AAV ABE *trans*-splicing system for dystrophin repair. Two AAV serotype-9 vectors expressing separated portions of an adenine base editor (ABE.TadA-8e^{V106W}) and a gRNA (gRNA^{Ex45}) lead to intein-mediated assembly of complete ABE:gRNA complexes. Base editing involving A·T-to-G·C transitions at the splice acceptor site of exon 45 establishes permanent exon 45 skipping in striated muscle cells. In dystrophin-defective muscle cells lacking exon 44, exon 45 skipping restores the reading of mature transcripts that code for a truncated Becker-like dystrophin with therapeutic potential for DMD patients. ITR, T-shaped hairpin-structured AAV serotype-2 inverted terminal repeats (*cis*-acting elements needed for vector DNA replication and packaging in producer cells). CK8e and An, synthetic striated muscle-specific promoter and polyadenylation signal, respectively. The dystrophin diagram was generated via: http://edystrophin.genouest.org/index.php?page=home.

To favor base editing in striated muscles over non-target organs of *Dmd* exon 44-deleted mice, the vector constructs were packaged in AAV serotype-9 capsids with the split ABE.TadA-8e^{V106W} moieties being expressed through the tissue-specific CK8e promoter. Intramuscular dual AAV co-administrations of 1 x 10^{11} total vector genomes (VGs) per tibialis anterior

(TA) led to 29.5% ± 2.7% A·T-to-G·C edits with minimal indel formation (*i.e.*, 0.2% ± 0.1%), as determined by deep sequencing at 3 weeks post injection. Systemic dual AAV co-administrations of 1.5 x 10^{14} VG kg^{-1} and 3 x 10^{14} VG kg^{-1} via the temporal facial veins of postnatal day 2 (P2) mice yielded, in TA muscles, 5.5% ± 1.2% and 8.1% ± 3.0% edits and, in hearts, 22.0% ± 2.2% and 26.2% ± 4.4% edits, respectively, with <0.1% indels detected at 8 weeks post injection. Critically, a general dose-dependent improvement of disease-associated molecular, cellular, and functional endpoints is reported. In this regard, up to 31% and 60% of wild-type dystrophin levels estimated in treated TA and heart muscles corresponded to over 76% and 95% of dystrophin-positive myofibers and cardiomyocytes, respectively. Partial dystrophin rescue translated, in turn, in noticeable reduction of myofiber central nucleation, diameter distribution, and fibrosis, all hallmarks of muscle degeneration. Finally, *Dmd* exon 44-deleted mice systemically treated at P2 with low and high doses of dual AAV ABE *trans*-splicing particles registered 31% and 41% grip strength augmentation, respectively, when compared with their untreated counterparts. Follow-up studies will be instructive to determine the long-term effects of the local and systemic gene-editing procedures in the treated animals. Of notice, despite the aforementioned transductional and transcriptional targeting measures, significant base editing was detected in the liver (*i.e.*, 11.1% ± 5.9%), which correlated with high VG copy numbers present specifically in this organ. These data confirm the importance of developing liver de-targeting protocols and strictly myotropic AAV capsids [10]. Indeed, as dose-dependent toxicity and immunological constrains have emerged during AAV clinical applications, optimization of tissue tropism and expression will be particularly important for dual AAV *trans*-splicing approaches due to the necessarily higher particle amounts required for maximizing co-expression and full-length protein assembly. Moreover, besides seeking the elimination of the observed editing at two of five top-ranked *in silico*-predicted candidate off-target sites [5], unbiased genome- and transcriptome-wide assessments of off-target effects, including gRNA-independent deamination, will complement the safety profile of *DMD*-targeting ABE:gRNA complexes. Concluding, Chai and colleagues demonstrate that AAV-vectored ABE *trans*-splicing can induce robust synthesis of Becker-like dystrophins in striated muscles of dystrophic mice upon DSB-free splice site knockout and exon skipping [5], The resulting improvement in pathological traits measured at the molecular, tissue, and functional levels validates the use of this platform for the efficacious testing and optimization of the growing number of ABE reagents *in vivo* and expands the range of potential treatment modalities for DMD patients.

Acknowledgements

Research in our laboratory is supported by the Prinses Beatrix Spierfonds (W.OR21-01), the Duchenne Parent Project NL, the Dutch Research Council (NWO) – Open Technology Program, and EU Marie Skłodowska-Curie Doctoral Network Actions. Z.L. holds a Ph.D. fellowship from the China Scholarship Council – Leiden University Joint Scholarship Program. Authors of this work are members of the European Reference Network – Neuromuscular diseases (ERN EURO-NMD).

Declaration of interest

The authors declare no competing interests.

References

1. Choi, E.; Koo, T.. CRISPR technologies for the treatment of Duchenne muscular dystrophy. *Mol. Ther*. **2021**, 29, 3179–3191.
2. Nelson, C.E.; Wu, Y.; Gemberling, M.P.; Oliver, M.L.; Waller, M.A.; Bohning, J.D.; Robinson-Hamm, J.N.; Bulaklak, K.; Castellanos Rivera, R.M.; Collier, J.H.; *et al*. Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. *Nat. Med.* **2019**, 25, 427–432.
3. 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.; *et al*. High levels of AAV vector integration into CRISPR-induced DNA breaks. *Nat. Commun*. **2019**, 10, 4439.
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. Chai, A.C.; Chemello, F.; Li, H.; Nishiyama, T.; Chen, K.; Zhang, Y.; Sánchez-Ortiz, E.; Alomar, A.; Xu, L.; Liu, N.; *et al*. Single-swap editing for the correction of common Duchenne muscular dystrophy mutations. *Mol. Ther. Nucleic Acids* **2023**, 32, 522–535.
6. Richter, M.F.; Zhao, K.T.; Eton, E.; Lapinaite, A.; Newby, G.A.; Thuronyi, B.W.; Wilson, C.; Koblan, L.W.; Zeng, J.; Bauer, D.E.; *et al.* Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **2020**, 38, 883–891.
7. Davis, J.R.; Wang, X.; Witte, I.P.; Huang, T.P.; Levy, J.M.; Raguram, A.; Banskota, S.; Seidah, N.G.; Musunuru, K.; Liu, D.R. Efficient in vivo base editing via single adeno-associated viruses with size-optimized genomes encoding compact adenine base editors. *Nat. Biomed. Eng.* **2022**, 6, 1272–1283.
8. Chemello, F.; Chai, A.C.; Li, H.; Rodriguez-Caycedo, C.; Sanchez-Ortiz, E.; Atmanli, A.; Mireault, A.A.; Liu, N.; Bassel-Duby, R.; Olson, E.N. Precise correction of Duchenne muscular dystrophy exon deletion mutations by base and prime editing. *Sci. Adv*. **2021**, 7, eabg4910.
9. Xu, L.; Zhang, C.; Li, H.; Wang, P.; Gao, Y.; Mokadam, N.A.; Ma, J.; Arnold, W.D.; Han, R. Efficient precise in vivo base editing in adult dystrophic mice. *Nat. Commun.* **2021**, 12, 3719.
10. Liu, J.; Koay, T.W.; Maiakovska, O.; Zayas, M.L.; Grimm, D. Progress in bioengineering of myotropic Adeno-associated viral (AAV) gene therapy vectors. *Hum. Gene Ther.* **2023**, 34(9-10), 350–364.

Chapter 4

Precision genome editing using combinatorial viral vector delivery of CRISPR-Cas9 nucleases and donor DNA constructs

Zhen Li[1], Xiaoling Wang[1], Josephine M. Janssen[1], Jin Liu[1],

Francesca Tasca[1], Rob C. Hoeben[1] and Manuel A.F.V. Gonçalves[1]

Nucleic Acids Research 53(2):gkae1213. (2025)

[1]Leiden University Medical Centre, Department of Cell and Chemical
Biology, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands.

Abstract

Genome editing based on programmable nucleases and donor DNA constructs permits introducing specific base-pair changes and complete transgenes or live-cell reporter tags at predefined chromosomal positions. A crucial requirement for such versatile genome editing approaches is, however, the need to co-deliver in an effective, coordinated and non-cytotoxic manner all the required components into target cells. Here, adenoviral (AdV) and adeno-associated viral (AAV) vectors are investigated as delivery agents for, respectively, engineered CRISPR-Cas9 nucleases and donor DNA constructs prone to homologous recombination (HR) or homology-mediated end joining (HMEJ) processes. Specifically, canonical single-stranded and self-complementary double-stranded AAVs served as sources of ectopic HR and HMEJ substrates, whilst second- and third-generation AdVs provided for matched CRISPR-Cas9 nucleases. We report that combining single-stranded AAV delivery of HR donors with third-generation AdV transfer of CRISPR-Cas9 nucleases results in selection-free and precise whole transgene insertion in large fractions of target-cell populations (*i.e.* up to 93%) and disclose that programmable nuclease-induced chromosomal breaks promote AAV transduction. Finally, besides investigating relationships between distinct AAV structures and genome-editing performance endpoints, we further report that high-fidelity CRISPR-Cas9 nucleases are critical for mitigating off-target chromosomal insertion of defective AAV genomes known to be packaged in vector particles.

Introduction

Genome editing technologies are advancing at a fast pace with their reach continuing to expand in science, biotechnology, and medicine [1]. Gene targeting involving insertion of exogenous donor DNA constructs into specific chromosomal loci (knock-in) subjected to double-stranded DNA breaks (DSBs) generated by engineered CRISPR-Cas9 nucleases is a common and remarkably versatile genome editing approach. The versatility of this generic approach results from its compatibility with large-scale genomic edits and the facile assembly of robust RNA-programmable CRISPR-Cas9 nucleases. Indeed, despite the discovery of a growing number of CRISPR systems, engineered CRISPR-Cas9 nucleases based on the prototypic *Streptococcus pyogenes* CRISPR system remain leading tools for genome editing purposes [1]. These designer nucleases consist of a single guide RNA (gRNA) and a Cas9 enzyme complex that binds to protospacer adjacent motif (PAM) sequences reading NGG. Targeted DSB formation, catalyzed by the HNH and RuvC nuclease domains of Cas9, follows whenever next to a PAM locates a *circa* 20-bp DNA stretch (protospacer) complementary to the 5′ end of the gRNA (spacer). Critically, directed evolution and structure-guided protein engineering efforts have yielded high-specificity Cas9 variants that, depending on the selected gRNA, have either undetectable or notably reduced off-target DNA cleaving activities [2].

Normally, gene targeting is achieved through the delivery of donor DNA substrates for ectopic homology-directed repair (HDR) [3] or non-homologous end joining (NHEJ) [4,5] at site-specific DSBs created by RNA-programmable nucleases. The former pathways engage donor DNA constructs whose designs favor homologous recombination (HR) [3], microhomology-mediated end joining (MMEJ) [6] or, more recently, homology-mediated end joining (HMEJ) [7-9]. MMEJ, HMEJ and HR donors contain sequences homologous to genomic sequences framing the targeted DSB (‘homology arms’) whilst NHEJ-prone donors lack ‘homology arms’ altogether [4,5]. As such, homology-bearing donors achieve directional single-step gene knock-ins in contrast to NHEJ donors including those with a homology-independent targeted integration (HITI) design [5].

Moreover, in contrast to HR donors, NHEJ, MMEJ and HMEJ donors have their targeting matrixes flanked by programmable nuclease cleaving sites. This ‘double-cut’ design assures the freeing of donor DNA substrates from delivery vehicle backbones which favors gene knock-ins via the processing and alignment between donor and target sequences. Of notice, when compared to MMEJ and NHEJ donors, HR and HMEJ donors mitigate indels at chromosomal-recombinant DNA junctions yielding more accurate and directional gene knock-ins [4,6,8].

Multiple physical and chemical transfection methods such as those based on electroporation and polycations, respectively, permit introducing genome editing reagents (*e.g.* engineered CRISPR nucleases and donor DNA constructs) into human cells. Yet, achieving optimal transfection efficiencies without evident cytotoxic effects is challenging. Moreover, effective transfection of multiple cell types often requires different reagents whose assembly or compositions are sometimes unknown due to proprietary reasons. Transfections are also dependent on systematic optimization of cell type-specific protocols whose ultimate performance often varies as a function of subtle experimental conditions, *e.g.* distribution of cell-cycle stages in cell populations at the time of transfection. Instead, viral vector transductions display high reproducibility and can be applied in a straightforward manner to different cell types. These features derive from the exquisite finely tuned mechanisms evolved by the vector parental viruses for the delivery of their nuclei acid genomes into host-cell nuclei. Adenoviral (AdV) and adeno-associated viral (AAV) vectors are particularly effective delivery vehicles in a wide variety of mammalian cell types independently of their cell-cycle statuses [10]. AdVs contain a linear protein-capped double-stranded DNA genome (up to 36 kb) [11]; AAVs have a linear single-stranded DNA genome (up to 4.7 kb) with T-shaped hairpin inverted terminal repeats (ITRs) [12]. Previous experiments have demonstrated that, in contrast to free-ended linear DNA, capped double-stranded DNA molecules, including AdV genomes, are refractory to concatemerization and off-target chromosomal insertion [13,14]. Amongst constructs encoding edits-of-interest or programmable nucleases it is especially critical to avoid chromosomal integration of the latter to minimize the potential build-up of off-target effects. Moreover, diversely from AAVs, AdVs can package complete CRISPR-Cas9 constructs encoding

Cas9 and single or multiple gRNAs. Conversely, recombinant AAV genomes serve as proficient substrates for HR, including in difficult-to-transfect stem cell populations [15], and for NHEJ-mediated HITI *in vitro* and *in vivo* [5]. Possibly, the peculiar structure of AAV vector genomes featuring secondary-structured ITRs flanking single-stranded DNA contributes to their recombinogenic character [16,17]. Hence, here, inspired by the complementary attributes of these viral vector platforms, we sought to investigate genome editing strategies based on allocating AdV and AAV systems for the delivery of, respectively, CRISPR-Cas9 nucleases and distinct types of donor DNA templates. In particular, second- and third-generation AdVs provided for Cas9 nucleases, whilst AAVs served as sources of HR and HMEJ donor substrates presented as single-stranded or double-stranded DNA molecules. In addition, to further study the role of AAV donor DNA structures on the efficiency and accuracy of genome editing, standard and self-complementary (sc) AAV vectors were assembled to present in target cell nuclei donor templates in single-stranded and double-stranded DNA formats, respectively.

Similarly to AAVs, third-generation AdVs, hereafter named high-capacity AdV particles (AdVPs), consist of viral gene-free recombinant DNA packaged in non-enveloped protein capsids [18,19]. In contrast to high-capacity AdVPs, first- and second-generation AdVs lack only a limited number of coding regions, *e.g.* early (E) region 1 (*E1*) alone or together with *E2A* or *E4*. These viral gene containing AdV vector genotypes can further possess deletions in immunomodulatory *E3* sequences whose functions are dispensable during vector amplification in complementing cell lines [10,11]. Co-transduction experiments using early-generation AdVs encoding CRISPR-Cas9 complexes and matched AAV donors tailored for HR or HMEJ yielded limited gene targeting due to strict dose dependent cytotoxicity likely associated with ‘leaky’ AdV gene expression in target cells. Co-transduction experiments using fully viral gene-deleted AdVPs and the aforementioned AAV donors led instead to robust gene targeting without noticeable cytotoxicity. In particular, combining single-stranded AAV delivery of HR donors with AdVP transfer of optimized high-specificity CRISPR-Cas9 nucleases yielded robust and precise gene targeting in human cells. We further disclose that productive AAV vector transduction directly correlates with the frequency of programmable nuclease-induced DSBs which may have a bearing on the ultimate performance and accuracy of AAV-based genome editing procedures. Finally, we demonstrate the importance of selecting high-specificity instead of regular Cas9 nucleases to avoid off-target chromosomal insertion of defective AAV DNA species in human cells.

Results and discussion

We started by investigating genome editing based on the delivery of donor DNA and CRISPR-Cas9 constructs in, respectively, AAV and second-generation AdVs lacking *E1* and *E2A*. Gene knock-in into 'safe harbor' loci is, amongst the range of genome editing strategies, a particularly versatile approach in that it offers the prospect for correcting recessive disorders independently of their underlying mutations in a predictable manner. The predictability aspect stems from a minimization of insertional mutagenesis, transgene silencing and variegated expression, all valuable features associated with gene knock-ins at 'safe harbor' loci [38]. Hence, AAV vectors carrying HR or HMEJ donors were tailored for targeted knock-in of *EGFP* expression units at the commonly used human safe harbor locus *AAVS1* at 19q13.4-qter (*i.e.* AAV-HR^{S1} or AAV-HMEJ^{S1}, respectively) (**Figure 1A**). Moreover, regardless of their genotypes, all adenovectors assembled for this study displayed adenovirus type-50 fibers. In contrast to prototypic type-5 fibers that recognize the coxsackievirus and adenovirus receptor (CAR), type-50 fibers bind CD46 instead. As such, these tropism-modified vectors facilitate transduction of otherwise refractory cell types with potential or established therapeutic relevance, such as, CAR-negative human mesenchymal stem cells, muscle progenitor cells (myoblasts) and hematopoietic stem cells [39–41]. Relatedly, AAV donors consisted of recombinant AAV genomes based on type-2 packaged in AAV type-6 capsids to facilitate delivery into a range of human cell types broader than that engaged by their prototypic AAV type-2 capsid counterparts [42]. In this regard, AAV transduction experiments in hMSCs revealed efficient transfer of donor constructs by these pseudotyped AAV particles, *i.e.* ~80% EGFP-positive hMSCs (**Figure 1A**). Interestingly, co-transduction of hMSCs with AAV and AdV vectors delivering *AAVS1*-targeting donor templates and CRISPR-Cas9 complexes, respectively, led to a clear AdV-dependent increase in AAV transduction as assessed through flow cytometry-based quantification of reporter-positive cells (*i.e.* 90–100%) and expression levels per cell (**Figure 1A**, top and bottom graphs, respectively) as well as via direct visualization of reporter-positive hMSCs using fluorescence microscopy (**Figure 1B**). Potential cause for this AdV-dependent AAV transduction enhancement is the expression of AAV helper functions from AdV genomes. Related to this aspect, wild-type AAVs (*Dependoparvovirus* genus) rely on unrelated viruses (*e.g.* adenoviruses) as helpers for the completion of their lytic infection cycle [12]. A well-established AAV helper function encoded in *E1*- and *E2A*-deleted AdV genomes is that of the E4ORF6 protein whose 'leaky' expression is known to buildup in an AdV dose-dependent manner [43]. The E4ORF6 helper function [44] has been implicated in converting incoming single-stranded AAV genomes into transcriptionally active double-stranded DNA [45,46]. Hence, it is likely that the AdV-dependent AAV transduction enhancing effect is contributed by the introduction of AAV helper functions into target cells. Regardless, upon episomal AAV DNA elimination through hMSC sub-culturing, CRISPR-Cas9-dependent stable transduction levels were variable and low (**Figure 1C**). In fact, these levels were not significantly higher than those detected in cultures initially exposed exclusively to AAV-HR^{S1} or AAV-HMEJ^{S1}, respectively (**Figure 1C**).

Based on the previous data, we next sought to improve the dual viral vector genome-editing system by assembling and testing high-capacity AdVPs encoding the high-specificity Cas9 nuclease SpCas9^{KARA} [35] alone (*i.e.* AdVP.C9^{KARA}) or the optimized high-specificity eCas9^{4NLS} nuclease [30] together with the *AAVS1*-targeting gRNA G^{S1} (*i.e.* AdVP.eC$9^{4NLS}$$G^{S1}$). The eCas$9^{4NLS}$ protein is a variant of eSpCas9(1.1) [29] whose enhanced performance results from having two extra nuclear localization signals [30]. Importantly, both AdVP.C9^{KARA} and AdVP.eC$9^{4NLS}$$G^{S1}$ contained intact vector genomes (**Supplementary Figure S1**). Dose-response transduction experiments in hMSCs confirmed that second-generation AdVs present a strikingly higher cytotoxicity profile than viral gene-free AdVPs (**Supplementary Figure S2** and **Supplementary Files** https://figshare.com/s/a999fdb4400aaf1828c5). These data are similar to earlier results from our laboratory showing that, amongst second-generation AdVs and viral gene-free AdVPs containing the same transgene cassettes, the former induces notably higher and dose-dependent cytotoxic effects in transduced cells [20].

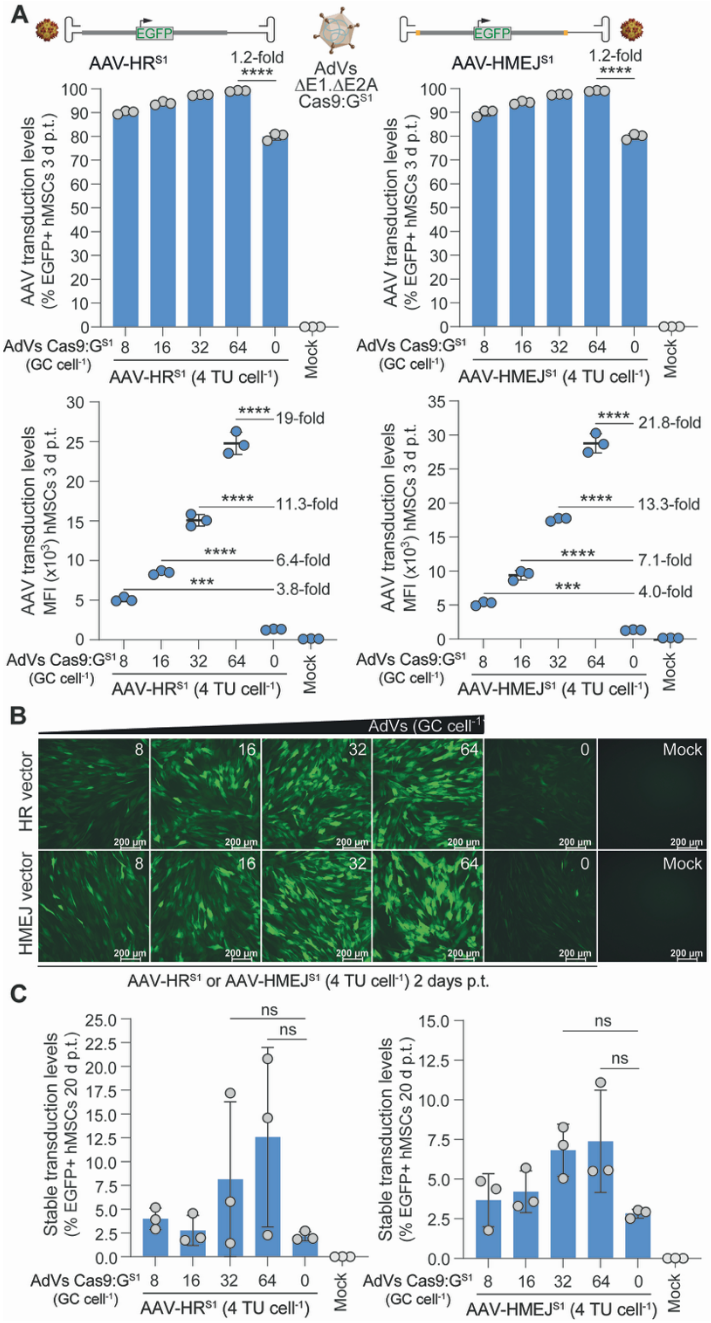


Figure 1. Genome editing using AAV delivery of HR or HMEJ donors and AdV transfer of CRISPR-Cas9 complexes. (**A**) AAV HR and HMEJ donor transduction efficiencies. Schematics of AAV donor structures are depicted. Horizontal gray bars, DNA

homologous to sequences flanking the *AAVS1* target site of gRNA G^{S1} ('homology arms'); gray box, *EGFP* transgene; orange boxes, G^{S1} target sequence. hMSCs were exposed to AAV-HR^{S1} or to AAV-$HMEJ^{S1}$ together with AdVs encoding *AAVS1*-targeting CRISPR-Cas9 complexes (Cas9:G^{S1}) at the indicated multiplicity-of-infection (MOI). TU and GC, transducing units and genome copies, respectively. AAV donor transduction was determined by quantifying EGFP-positive cells and corresponding mean fluorescence intensity (MFI) values by flow cytometry at 3 days post-transduction (top and bottom graphs, respectively). Mock-transduced hMSCs and hMSCs exposed exclusively to AAV donors served as controls. (**B**) AAV donor expression imaging. Representative direct fluorescence microscopy images of hMSCs transduced with AAV-HR^{S1} or AAV-$HMEJ^{S1}$ at a single MOI and different amounts of AdVs encoding Cas9:G^{S1} complexes or with AAV donors alone. Images were acquired at 2 days post-transduction (**C**) AAV stable transduction levels. Stable transduction frequencies were assessed by EGFP-directed flow cytometry after sub-culturing for 20 days hMSCs initially exposed to the indicated viral vector MOI. Data are presented as mean ± standard deviation (SD) of three biological replicates. Significant differences between the defined datasets were determined with one-way analysis of variance (ANOVA) followed by Tukey's test for multiple comparisons; $***P < 0.001$; $****P < 0.0001$; $P > 0.05$ was considered non-significant (ns).

These cumulative datasets are consistent with the fact that although *E1*- and *E2A*-deleted AdVs encode less viral regulatory functions than their first-generation, *E1*-deleted, counterparts, 'leaky' viral gene expression is nonetheless detected at high vector doses [11,43]. Importantly, transduction of hMSCs with viral gene-free AdVPs did not impair their differentiation capacity into osteoblasts as assessed by the detection of calcium deposits upon incubation in osteogenic differentiation medium and alizarin red S staining (**Supplementary Figure S3**).

Gene targeting experiments using AdVP.$C9^{KARA}$ and AAV-$HR^{L.S1}G^{S1}$, an AAV co-delivering gRNA G^{S1} and matching *AAVS1*-tailored HR templates, led to robust transduction of hMSCs (**Figure 2A** and **B**, left column) as well as HeLa cells (**Figure 2D**). Crucially, this co-transduction scheme yielded CRISPR-Cas9-dependent stable transduction frequencies of up to ~80% and 90% at end-point hMSC and HeLa cell cultures (**Figure 2C** and **E**, respectively). Indeed, in the absence of AdVP.C9^{KARA}, the frequencies of hMSCs and HeLa cells stably transduced with AAV donor DNA were barely above background levels (**Figure 2C** and **E**, respectively) as was also grasped via live-cell fluorescence microscopy of hMSC cultures at 19 days post-transduction (**Figure 2B**, right column). Additional control experiments consisting of exposing cells to AAV-HR$^{L.S1}$G^{S1} with and without AdVP.C9^{KARA} confirmed over 90% and background stable transduction frequencies, respectively; whilst those comprising the AAV donor alone or together with a control AdVP.mCherry vector yielded exclusively background stable transduction levels (**Supplementary Figure S4**). Analogous experiments in hMSCs using AdVP.C9^{KARA} together with AAV-HR^{C5}G^{C5}, an AAV co-delivering gRNA G^{C5} and HR substrates for gene knock-in at the alternative 'safe harbor' locus *CCR5*, also yielded efficient transduction (**Figure 2F**) and high CRISPR-Cas9-dependent stable transduction levels, *i.e.* up to 54% of initially transduced cells (**Figure 2G**). Of note, CRISPR-Cas9-dependent stable transduction levels in primary hMSCs with AdVP.C9^{KARA} and AAV-HR^{C5}G^{C5} (**Figure 2G**) compared favorably to those measured in HeLa cells (**Supplementary Figure S5**, bottom graph). Moreover, albeit by degrees lower than those previously observed using AAV and second-generation AdV vectors (**Figure 1A**), AAV and viral gene-free AdVP vectors also led to increased AAV transduction levels as measured by reporter-directed flow cytometry at 3 days post-transduction (**Figure 2A**, **D** and **F**, bottom graphs). Based on these data, we next sought to investigate the role of AdVPs *per se* versus AdVPs combined, or not, with targeted DSB formation on the proficiency of AAV vector transduction as determined by tracing vector DNA amounts and transgene expression levels.

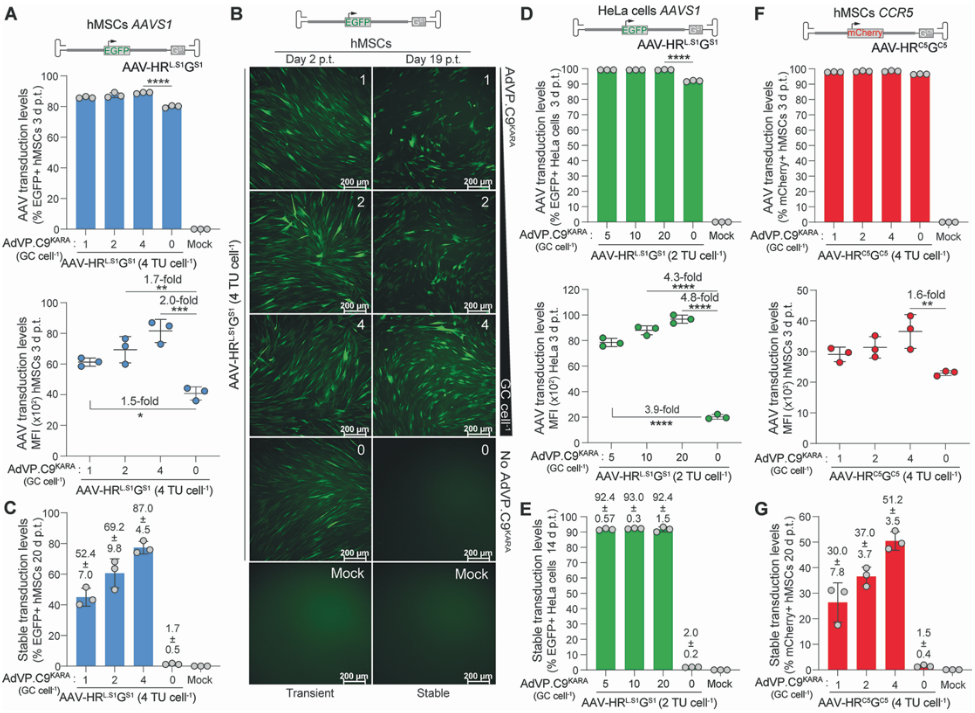


Figure 2. Genome editing combining AAV delivery of donor and gRNA components with high-capacity AdVP transfer of Cas9. (**A**) AAV donor transductions targeting *AAVS1* safe harbor loci in hMSCs. The structure of AAV-HR$^{L.S1}$G^{S1} donor is shown. Horizontal gray bars, DNA homologous to sequences flanking the *AAVS1* target site of gRNA G^{S1} ('homology arms'). hMSCs were transduced with AAV-HR$^{L.S1}$G^{S1} and AdVP.C9^{KARA} at the specified MOI. The former vector encodes the *AAVS1*-targeting gRNA G^{S1}; the latter vector encodes the high-specificity Cas9 nuclease SpCas9^{KARA}. AAV donor transduction was assessed by measuring EGFP-positive cells and respective MFI values by flow cytometry at 3 days post-transduction (top and bottom graphs, respectively). Mock-transduced hMSCs and hMSCs transduced solely with the AAV donor provided for controls. Data are shown as mean ± SD of three biological replicates. Significant differences amongst the marked datasets were calculated with one-way ANOVA followed by Tukey's test for multiple comparisons; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$ (**B**) Monitoring transient and stable transduction of hMSCs upon AAV-HR$^{L.S1}$G^{S1} delivery. Representative live-cell fluorescence microscopy images of hMSCs co-transduced with AAV-HR$^{L.S1}$G^{S1} and AdVP.C9^{KARA} at the indicated MOI. Images showing transiently and stably transduced cells were acquired at 2- and 19-days post-transduction, respectively. (**C**) Quantification of hMSCs stably transduced with AAV-HR$^{L.S1}$G^{S1}. Stable transduction frequencies were measured via EGFP-directed flow

cytometry after sub-culturing for 20 days hMSCs initially exposed to the indicated viral vector MOI. (**D**) AAV donor transductions targeting *AAVS1* safe harbor loci in HeLa cells. AAV-HR$^{L.S1}$G^{S1} and AdVP.C9^{KARA} were applied at the indicated MOI and, at 3 days post-transduction, EGFP-directed flow cytometry was used to measure the frequency of transgene expressing cells and respective MFI (top and bottom graphs, respectively). Mock-transduced HeLa cells and HeLa cells transduced exclusively with AAV-HR$^{L.S1}$G^{S1} provided for controls. Significant differences amongst the marked datasets were assessed with one-way ANOVA followed by Tukey's test for multiple comparisons; $^{****}P < 0.005$ ($n = 3$ biological replicates). (**E**) Quantification of HeLa cells stably transduced with AAV-HR$^{L.S1}$G^{S1}. Stable transduction frequencies were measured via EGFP-directed flow cytometry after sub-culturing for 14 days HeLa cells initially exposed to the indicated viral vector MOI. Data are shown as mean ± SD of three biological replicates. (**F**) AAV donor transductions targeting *CCR5* safe harbor loci in hMSCs. The structure of AAV-HR^{C5}G^{C5} donor is depicted. Horizontal gray bars, DNA homologous to sequences flanking the *CCR5* target site of gRNA G^{C5} ('homology arms'); gray box, *mCherry* transgene. AAV-HR^{C5}G^{C5} and AdVP.C9^{KARA} were applied at the specified MOI and, at 3 days post-transduction, mCherry-directed flow cytometry was used to determine the frequency of transgene expressing cells and corresponding MFI (top and bottom graphs, respectively). Mock-transduced hMSCs cells and hMSCs transduced with AAV-HR^{C5}G^{C5} alone served as controls. Significant differences between the marked datasets were assessed with one-way ANOVA followed by Tukey's test for multiple comparisons; $^{**}P < 0.01$ ($n = 3$ biological replicates). (**G**) Quantification of HeLa cells stably transduced with AAV-HR^{C5}G^{C5}. Stable transduction frequencies were measured via mCherry-directed flow cytometry after sub-culturing for 20 days hMSCs initially treated with indicated viral vector MOI. Numerals above the graph bars correspond to AAV stable transduction frequencies (mean ± SD) normalized to the initial transduction levels.

Initial quantification of vector DNA buildup in transduced cells in the presence and absence of AdVPs ruled out an effect of AdVPs *per se* on AAV vector cell entry (**Supplementary Figure S6**). Of note, previous experiments have demonstrated that DNA damage response (DDR) proteins (*e.g.* MRN complex factors MRE11-RAD50-NBS1) associate with incoming AAV vector DNA that presumably hinder transgene expression [47]. Indeed, the dampening of these proteins correlates with increased permissiveness to AAV vector expression *in vitro* and i*n vivo* [48] and exposing cells to pleiotropic genotoxic drugs (*e.g.* hydroxyurea and etoposide) equally fosters productive AAV vector transduction [49]. Moreover, RNA interference screens retrieved an overrepresentation of hits whose down-regulation stimulates both DNA damage and AAV transduction, although top-scoring hits act on DDR pathways only indirectly [50]. These cumulative findings support the hypothesis that genomic lesions recruit DNA damage-sensing factors that would otherwise inhibit productive AAV transduction via vector DNA binding [50].

To investigate in a more direct manner a causal relationship between productive AAV transduction and specific DNA lesions, namely programmable nuclease-induced DSBs, HeLa cells were exposed to CRISPR-Cas9 complexes endowed with either highly specific or promiscuous gRNAs, *i.e.* gRNA^{CALM2} and gRNA^{VEGFA}, respectively [29,51,35,36]. Real-time live-cell imaging of these cells transduced with an EGFP-expressing AAV vector or with its isogenic scAAV counterpart revealed a clear DSB-dependent increase in transgene expression regardless of the AAV vector structure (**Figure 3A** – **C**). scAAV vectors, assembled by using AAV transfer constructs in which one of the two ITR terminal resolution sites is mutated, promptly hybridize upon uncoating to form transcription-competent double-stranded genomes owing to their internal inverted repeat arrangement (**Supplementary Figure S7**) [12,52,53]. Monitoring of AAV vector transduction via real-time live-cell imaging established the expected higher transgene expression kinetics of scAAV-HR^{S1} over that of conventional single-stranded AAV-HR^{S1} (**Figure 3D** and **E**). Notably, regardless of AAV genome arrangement, productive transduction was highest in cells initially subjected to the highest number of DSBs induced by the promiscuous Cas9:gRNA^{VEGFA} complexes (**Figure 3A** – **C**). This data is consistent with the aforementioned hypothesis whereby DNA damage, in this case CRISPR-Cas9-induced DSBs, recruits DDR proteins known to bind palindrome-structured AAV genomes and, in doing so, leave said genomes unhindered for transgene expression.

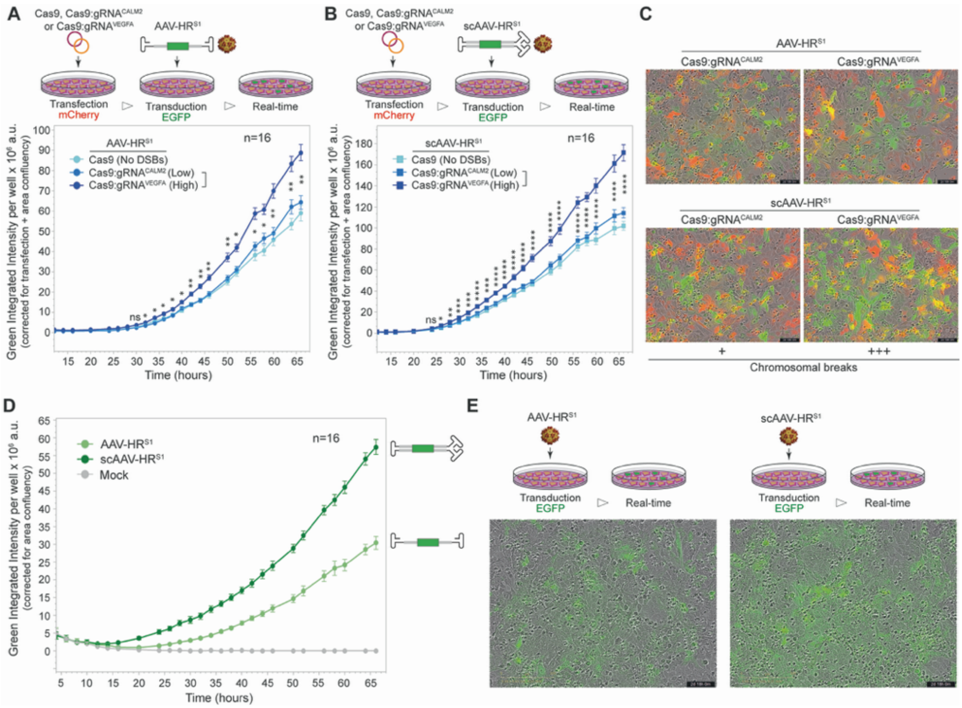


Figure 3. Assessing the role of chromosomal DNA breaks on AAV vector transduction. (**A**,**B**) Single-stranded AAV and double-stranded scAAV transduction kinetics in cells subjected to CRISPR-induced DSBs. HeLa cells exposed to Cas9:gRNA^{CALM2} or Cas9:gRNA^{VEGFA} complexes, inducing few or plenty off-target DSBs, respectively, were transduced with vectors AAV-HR^{S1} or scAAV-HR^{S1}. AAV transgene expression was traced by real-time live-cell imaging for up to 66 h in an Incucyte apparatus. Each datapoint represents mean and SD values of a total of 16 fluorescence intensity measurements derived from 16 independent microscopy fields per well. Significant differences between datapoints corresponding to cells receiving gRNA^{CALM2} versus gRNA^{VEGFA} were calculated with Student's t-tests; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; $P > 0.05$ was considered ns. (**C**) Representative microscopy fields corresponding to the indicated experimental settings at the 66-h timepoint. (**D**) Establishing faster transduction kinetics of scAAV over regular AAV vectors. HeLa cells were transduced with vectors AAV-HR^{S1} or scAAV-HR^{S1} and transgene expression levels were monitored through real-time live-cell imaging for up to 66 h in an Incucyte apparatus. Each datapoint represents mean and SD values of a total of 16 fluorescence intensity measurements derived from 16 independent microscopy fields per well. (**E**) Representative microscopy fields corresponding to HeLa cells transduced with AAV-HR^{S1} or scAAV-HR^{S1} at the 66-h timepoint.

Besides DDR proteins, episomes of both non-viral and viral origins, including AAV vector genomes, are prone to a vast array of cellular restriction factors that dampen or silence transgene expression, *e.g.* via the acquisition of heterochromatin epigenetic marks [33,54]. In this context, additional real-time monitoring of AAV vector transduction kinetics in the presence and absence of AdVP.C9$^{\mathrm{KARA}}$ was consistent with HDR-mediated AAV gene targeting occurring soon after site-specific DSB formation, which presumably results in heightened and earlier onset of transgene expression from integrated sequences over non-integrated episomes (**Supplementary Figure S8**). Such differential expression levels between chromosomally targeted and episomal AAV donor DNA have been in fact successfully exploited to enrich for gene knock-ins amongst transduced cell populations [15].

Directly editing specific endogenous genes for the purpose of tagging, modifying or knocking-out encoded gene product(s) is an alternative approach to inserting transgenes at safe harbor loci. A distinctive advantage of endogenous gene modification is maximizing appropriate physiologically regulated expression and subcellular positioning of the resulting gene product(s). Moreover, most genetic disorders are caused by numerous and diverse types of mutations scattered along gene bodies. Consequently, knocking-in recombinant coding sequences downstream of *cis*-acting regulatory elements offers an ‘universal’ strategy for rescuing disease phenotypes or properly tracing gene products. A broad range of mutations in the *DMD* gene, that normally codes for the long sarcolemma-stabilizing protein dystrophin (427 kDa), causes the lethal X-linked muscle-wasting disorder Duchenne muscular dystrophy (DMD; MIM no. 310200). Due to its prevalence (~1:4700 boys) and severity, DMD is a primary target for genetic therapies despite the challenges posed by the vast expanse of the affected striated musculature and of the *DMD* gene itself (>2.3 Mb) [55]. Candidate *in vivo* and *ex vivo* DMD genetic therapies have defined sets of pros and cons [56–58]. For instance, although autologous transplantation of *ex vivo* corrected myogenic stem/progenitor cells currently presents notable challenges (*e.g.* reduced cell survival and tissue engraftment), it offers nonetheless a controlled gene correction setting and, simultaneously, minimizes immune responses directed against vector and gene-editing tool components [56–58]. Hence, to further assess the performance of the dual viral vector genome-editing system, we assembled the AAV donor construct AAV-HR^{DMD}G^{IN4} for *DMD* gene knock-in using human muscle progenitors (myoblasts) as target cells (**Figure 4A**). This AAV donor encodes gRNA G^{IN4} for site-specific DNA cleavage within *DMD* intron 4 and it bears a matched HR template for directional HDR-mediated fusion of *EGFP* to the first four exons of *DMD* (**Figure 4A**). The gRNA G^{IN4} was selected via functional assays in HeLa cells and human myoblasts (**Supplementary Figure S9**). In addition to *EGFP*, the donor template in AAV-HR^{DMD}G^{IN4} also has a constitutively active *mCherry* expression unit for tracing purposes (**Figure 4A**). *DMD* gene editing experiments using AdVP.C9^{KARA} and AAV-HR^{DMD}G^{IN4} resulted in efficient AAV transduction of human myoblasts (**Figure 4B**, left graph), with the presence of AdVP.C9^{KARA} having a significant enhancing effect on AAV transduction levels (**Figure 4B**, right graph), as previously observed in HeLa cells and hMSCs (**Figure 2**). Importantly, CRISPR-Cas9-dependent stable transduction frequencies of up to 52.6 ± 6.1% were reached in initially co-transduced myoblast populations (**Figure 4C**). A multiplex RT-PCR assay demonstrated the co-expression of *EGFP*-tagged and unmodified *DMD* transcripts in myotubes differentiated from myoblasts engineered via AAV-HR^{DMD}G^{IN4} and AdVP.C9^{KARA} co-transduction (**Supplementary Figure S10A**). Moreover, Sanger sequencing confirmed the accurate assembly of mRNA fusion transcripts in edited myotubes corresponding to the *EGFP* sequence linked to the first four *DMD* exons (**Supplementary Figure S10B**).

In addition to the efficiency, the specificity and precision of donor DNA insertion are other parameters of paramount importance associated with programmable nuclease-assisted gene knock-in procedures. The specificity is established by detecting donor DNA sequences at the target site whilst the precision or fidelity can be ascertained by demonstrating that telomeric-sided and centromeric-sided junctions between exogenous and endogenous DNA sequences (jT and jC, respectively), result from HDR instead of imprecise end-joining processes (non-HDR). Hence, to assess the specificity and fidelity of *DMD* gene editing in human myoblasts co-transduced with AdVP.C9^{KARA} and AAV-HR^{DMD}G^{IN4}, junction PCR screens were carried out on randomly selected mCherry-positive myoblast clones, where each clone represents an individual genome-modifying event. Of notice, in contrast to splice acceptor gene-trapping constructs, cell isolation based on constitutively active transgenes, such as the *mCherry* in AAV-HR^{DMD} G^{IN4}, avoids biased selection for on-target chromosomal DNA insertions. Hence, it is significant that strictly off-target chromosomal donor DNA insertions were not detected (*i.e.* jT-/jC-) in any of the randomly isolated myoblast lines and, amongst the forty lines analyzed, only two were not properly targeted as they lacked HDR-derived chromosomal junctions at one of the two ends (*i.e.* jT+/jC-) (**Supplementary Figure S11**). Interestingly, similar to three myoblast lines that presented imprecise HDR-derived junctions, these two lines contained high molecular-weight products whose origins are consistent with AAV donor concatemers (**Supplementary Figure S11**).

To test endogenous gene tagging via the dual viral vector genome-editing system at an independent locus, the AAV-HR^{LMNA}G^{EX1} donor was assembled (**Figure 4D**). This AAV donor expresses gRNA G^{EX1} for site-specific DNA cleavage at exon 1 of *LMNA* and it contains a matched HR template for directional HDR-mediated fusion of the mScarlet-I reporter to the N-terminus of LMNA (**Figure 4D**). The LMNA protein is a component of the nuclear lamina matrix that locates underneath the inner nuclear membrane and whose roles include nuclear stability, chromatin structure and gene expression. Critically, *LMNA* mutations cause various disorders, *e.g.* Emery–Dreifuss muscular dystrophy, familial partial lipodystrophy, limb girdle muscular dystrophy, dilated cardiomyopathy and Hutchinson–Gilford progeria syndrome. Co-transduction experiments with AdVP.C9^{KARA} and AAV-

HR^{LMNA}G^{EX1} led to high and DSB-dependent LMNA protein tagging frequencies in HeLa cell populations, reaching at the uppermost AAV dose 86.5 ± 0.8% and 89.6 ± 2.5% as determined by flow cytometry at 3- and 14-days post-transduction, respectively (**Figure 4E**). The frequencies of LMNA tagging registered in the absence of AdVP.C9^{KARA} were at background levels (**Figure 4E**). Importantly, the LMNA-tagged cell fractions in populations exposed to the same transduction conditions were similar at early and late timepoints indicating a lack of negative selection against these gene-edited cell populations. Similar results were obtained upon LMNA tagging experiments in hMSCs (**Supplementary Figure S12A**). These data are, therefore, consistent with the low cytotoxicity profile of the dual viral vector genome-editing system based on viral gene-free AdVP and AAV particles in the herein tested cells as indicated by the stability of LMNA-tagged cell frequencies before and after sub-culturing. Finally, the synthesis and proper location of the fusion mScarlet-I::LMNA protein in cell nuclei was confirmed by direct live-cell fluorescence microscopy (**Figure 4F** and **Supplementary Figure S12B**). Taken together, these data indicate that the combined delivery of HDR-based genome editing components by AdVP and AAV particles is a robust and versatile genomic engineering approach that, in principle, can be tailored for generating disease models as well as cellular substrates for drug screens and, eventually, autologous cell therapies.

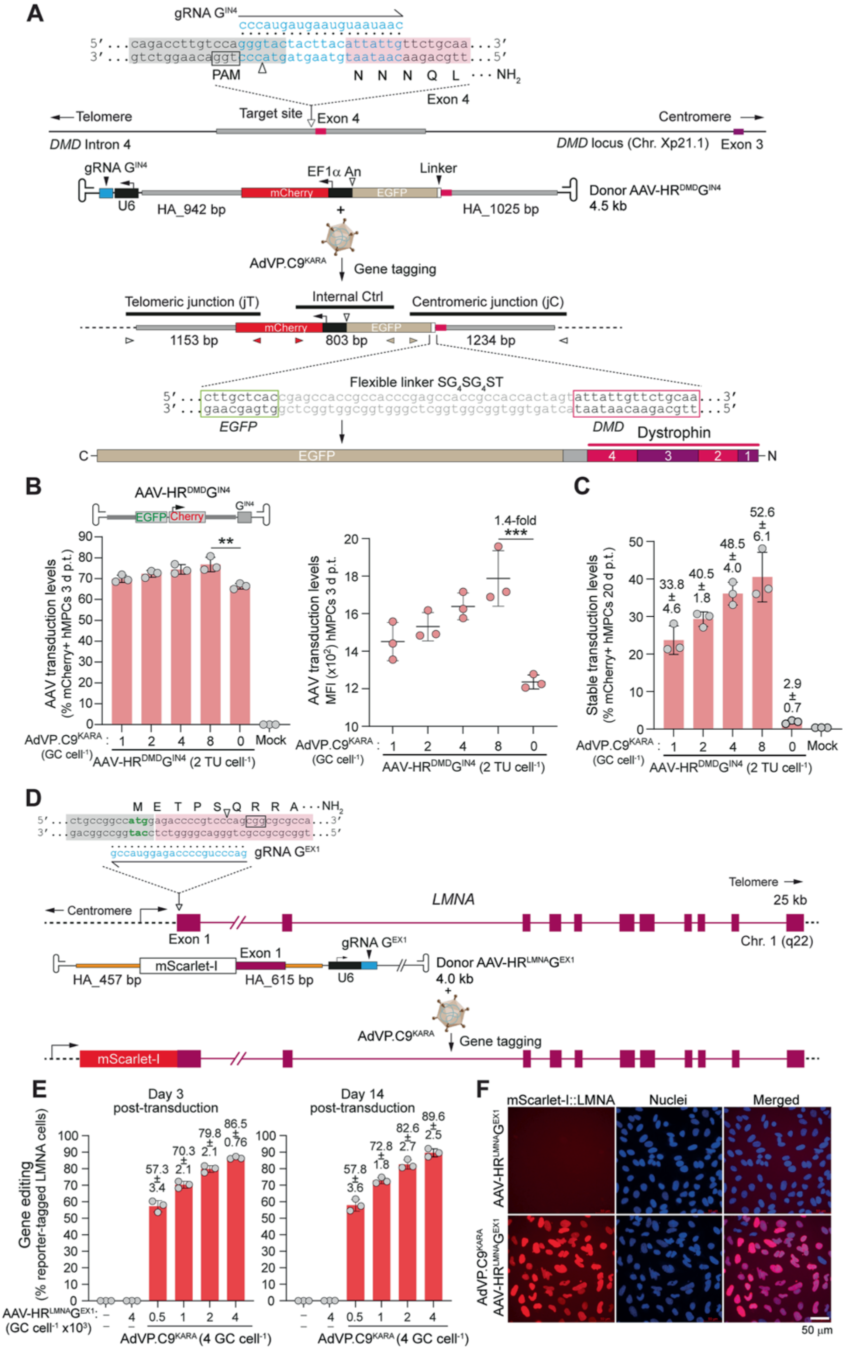


Figure 4. Endogenous gene tagging using dual viral vector delivery. (**A**) Diagrammatic representation of the *DMD* gene tagging strategy. Horizontal gray bars, AAV-HR^{DMD}G^{IN4} DNA homologous to sequences flanking the *DMD* target site of gRNA G^{IN4} ('homology

arms'); black boxes, *EF1α* and *U6* gene promoters driving mCherry and gRNA G^{IN4} expression, respectively. The elements of the junction PCR assays diagnostic for HDR-mediated gene tagging are also depicted. (**B, C**) Testing *DMD* gene tagging in human myoblasts. Human myoblasts were co-transduced with AdVP.$C9^{KARA}$ and AAV-$HR^{DMD}G^{IN4}$ at the indicated MOI. The initial and stable transduction levels, determined by flow cytometry at 3 days and 20 days post-transduction, are plotted in panel (**B**) and (**C**), respectively. Mock-transduced myoblasts and myoblasts transduced with AAV-$HR^{DMD}G^{IN4}$ alone served as controls. Significant differences between the marked datasets were assessed with one-way ANOVA followed by Tukey's test for multiple comparisons; $^{**}P < 0.01$; $^{***}P < 0.001$. Bars and error bars represent mean and SD values, respectively, of three biological replicates. Numerals above the graph bars correspond to AAV stable transduction frequencies (mean ± SD) normalized to the initial transduction levels. (**D**) Diagrammatic representation of the *LMNA* gene tagging strategy. (**E**) Assessing *LMNA* gene tagging in HeLa cells. HeLa cells were co-transduced with AdVP.$C9^{KARA}$ and AAV-$HR^{LMNA}G^{EX1}$ at the indicated MOI. The gene tagging frequencies were determined by flow cytometry at 3- and 14-days post-transduction. Bars and error bars correspond to mean and SD values, respectively, of three biological replicates. (**F**) Direct fluorescence microscopy analysis of LMNA-tagged cells. Representative fluorescence microscopy images of HeLa cells exposed only to AAV-$HR^{LMNA}G^{EX1}$ at 4000 GC $cell^{-1}$ or to this donor and AdVP.$C9^{KARA}$ at 4000 GC $cell^{-1}$ and 4 GC $cell^{-1}$, respectively. HeLa cell nuclei containing reporter-tagged LMNA proteins were visualized via direct fluorescence microscopy for mScarlet-I and the DNA dye Hoechst 33342.

We next sought to exploit the dual viral vector genome-editing platform to investigate the performance of different AAV donor designs (*i.e.* HR- versus HMEJ-prone) and structures (*i.e.* single-stranded versus double-stranded). To this end, AAV-HR^{S1} and AAV-$HMEJ^{S1}$ were first combined with AdVP.$eC9^{4NLS}G^{S1}$ for introducing *AAVS1*-tailored HR and HMEJ donors into hMSCs, respectively. CRISPR-Cas9-dependent stable transduction frequencies measured upon sub-culturing of co-transduced hMSCs were similar in cultures initially exposed to AAV-HR^{S1} or AAV-$HMEJ^{S1}$ (**Figure 5A**). In contrast, it is well-established that amongst HR and HMEJ substrates placed in plasmid or adenovector double-stranded DNA, genome editing frequencies are typically higher when using the latter 'double-cut' donor design [7–9,13,26]. Of notice, incoming single-stranded AAV genomes normally undergo conversion to a double-stranded format via either the engagement of host cell DNA polymerases at the priming 3′ ITR or after the hybridization of genomes with plus and minus polarity whose packaging in AAV capsids seems to occur at similar rates [12,45,59,60]. Hence, the single-to-double strand AAV DNA conversion kinetics is likely to be contingent upon particular cell types and experimental conditions that, together, influence the build-up of double-stranded AAV HMEJ substrates susceptible to targeted DNA cleavage (**Supplementary Figure S6**). For instance, inhibitory cellular factors binding AAV ITR *cis*-acting elements, whose amounts vary in different cell types, can render second-strand AAV DNA synthesis or transgene expression rate-limiting to different extents [46–48]. Therefore, to readily present HMEJ-prone AAV donor substrates in transduced cells and, simultaneously, probe the performance of single-stranded versus double-stranded AAV donor structures, scAAV-HR^{S1} and scAAV-HMEJ^{S1} were assembled and tested side-by-side with their respective AAV-HR^{S1} and AAV-HMEJ^{S1} counterparts (**Figure 5B**). Similar to the previous experiments in hMSCs (**Figure 5A**), combining AdVP.eC9^{4NLS}G^{S1} with AAV-HR^{S1} or AAV-HMEJ^{S1} yielded comparable CRISPR-Cas9-dependent stable transduction frequencies in HeLa cells (**Figure 5B**, left graphs). Notably, comparable CRISPR-Cas9-dependent genome editing frequencies were also observed in HeLa cells initially co-transduced with AdVP.eC9^{4NLS}G^{S1} and each of the scAAV-HR^{S1} and scAAV-HMEJ^{S1} vectors (**Figure 5B**, right graphs). Therefore, these data suggest that in contrast to HMEJ templates placed in the context of donor plasmids or adenovector genomes [7–9,13,26], HMEJ templates in AAV donor genomes do not necessarily outperform their HR counterparts during DSB-dependent genome editing.

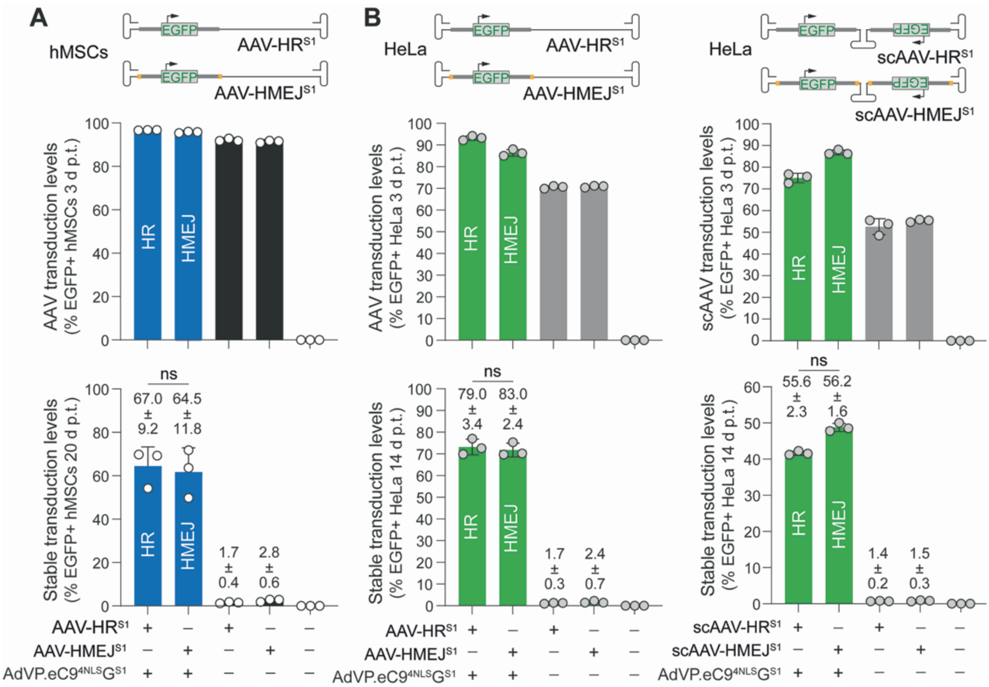


Figure 5. Investigating the role of AAV donor structures on genome editing. (A) Testing genome editing in hMSCs using single-stranded AAV HR and HMEJ donor constructs. Schematics of AAV donor structures are depicted. Horizontal gray bars, short tracts homologous to DNA flanking the *AAVS1* target site of gRNA G^{S1} ('short homology arms'); gray box, *EGFP* transgene; orange boxes, G^{S1} target sequence. hMSCs were transduced with AdVP.eC9^{4NLS}G^{S1} (4 GC cell^{-1}), expressing high-specificity eCas94^{NLS}:G^{S1} complexes, together with AAV-HR^{S1} or AAV-HMEJ^{S1} donors (4 TU cell^{-1} each). The frequencies of transient and stable transduction levels are plotted in the top and bottom graphs, respectively. Datapoints correspond to mean ± SD values from three biological replicates. (**B**) Testing genome editing in HeLa cells using single-stranded versus double-stranded AAV donor constructs. Diagrams of regular and sc AAV HR and HMEJ donors, are shown. HeLa cells were transduced with AdVP.eC9^{4NLS}G^{S1} (20 GC cell^{-1}) mixed with single-stranded AAV (2 TU cell^{-1}) or scAAV donors (0.5 TU cell^{-1}. The frequencies of transient and stable transduction levels are plotted in the top and bottom graphs, respectively. Datapoints correspond to mean ± SD values from three biological replicates. Numerals above the graph bars correspond to AAV stable transduction frequencies (mean ± SD) normalized to the initial transduction levels.

Knowledge on the relative contribution of different AAV donor designs (*i.e.* regular versus 'double-cut') and structures (*i.e.* single-stranded versus double-stranded) to genome editing endpoints is scant. To fill this knowledge gap, we capitalized on the isogenic set of structurally diverse AAV donors to systematically investigate the role of such donor substrates on the precision of genome editing. In particular, AAV donors with HR or HMEJ templates in single-stranded AAV or double-stranded scAAV genomes with targeting modules matching the commonly used *AAVS1* safe harbor locus. After co-transducing HeLa cells with AdVP.eC9^{4NLS}G^{S1} and each of the AAV donors, we performed junction PCR screens on randomly selected EGFP-expressing cell clones, each of which corresponding to individual genome-modifying events (**Figure 6A**). In the set of clones modified through the transfer of AdVP.eC9^{4NLS}G^{S1} and AAV-HR^{S1}, the *AAVS1*-targeted fraction was 97.7%, with 88.6% of the total DNA-modifying events representing precise genome editing (*i.e.* jT+/jC+) (**Figure 6B** and **Supplementary Figure S13**). The scAAV vectors scAAV-HR^{S1} and scAAV-HMEJ^{S1} yielded similar and lower numbers of precise genome-editing events, respectively, when compared to those obtained with AAV-HR^{S1} (**Figure 6B** and **Supplementary Figure S13**). Indeed, although the *AAVS1*-targeted cell fraction resulting from using scAAV-HMEJ^{S1} was 97.7%, only 71.7% of the total DNA-modifying events were precisely targeted (**Figure 6B** and **Supplementary Figure S13**). These data are consistent with previous studies showing that 'double-cut' donors placed in circular plasmids or protein-capped linear adenovector genomes normally yield less accurate genome-editing products than their respective HR donor counterparts [6,7,9,13,26,36]. Hence, it is also feasible that in the context of scAAV genomes with closed palindromic structures, generation of free-ended HMEJ donor products created by targeted DNA cleavage exacerbates the engagement of less precise NHEJ and/or MMEJ processes linking exogenous to endogenous DNA. Consequently, as single-stranded AAV-HR^{S1} and scAAV-HMEJ^{S1} yielded similar genome editing frequencies but the former presented higher HDR precision, we next focused on further testing single-stranded AAV constructs as sources of HR donor templates. To this end, gene targeting experiments were carried out in HeLa cells with AdVP.eC9^{4NLS}G^{S1} and AAV-HR$^{L.S1}$. The AAV-HR$^{L.S1}$ donor construct has 'homology arms' larger than those in the previous single-stranded and sc AAV vectors (*i.e.* 600-bp instead of 200-bp per arm). Moreover, in these experiments, various doses of AAV-HR$^{L.S1}$ were applied to assess the efficiency and precision of gene targeting as a function of AAV donor amounts (**Figure 6C**). As expected, the CRISPR-Cas9-dependent stable transduction frequencies determined at 14 days after transduction, were directly proportional to AAV donor amounts (**Figure 6C**). Interestingly, independently of AAV donor amounts, genome editing frequencies relative to the initial transduction levels were comparable (range: 84%–87%) (**Figure 6C**). Importantly, randomly isolated EGFP-expressing cell clones (n = 80) generated through AAV-HR$^{L.S1}$ transductions at 2 and 0.1 TU per cell were all shown to be targeted (**Figure 6D**) with, respectively, 97.6% and 100% of these being precisely targeted (**Figure 6D** and **E**).

Taken together, these data suggest that AAV HR donors engage target sequences more frequently through precise HDR than double-stranded AAV HMEJ donors and, reminiscent of double-stranded HR templates, extending the 'homology arms' in single-stranded AAV HR templates can improve the performance of gene targeting protocols.

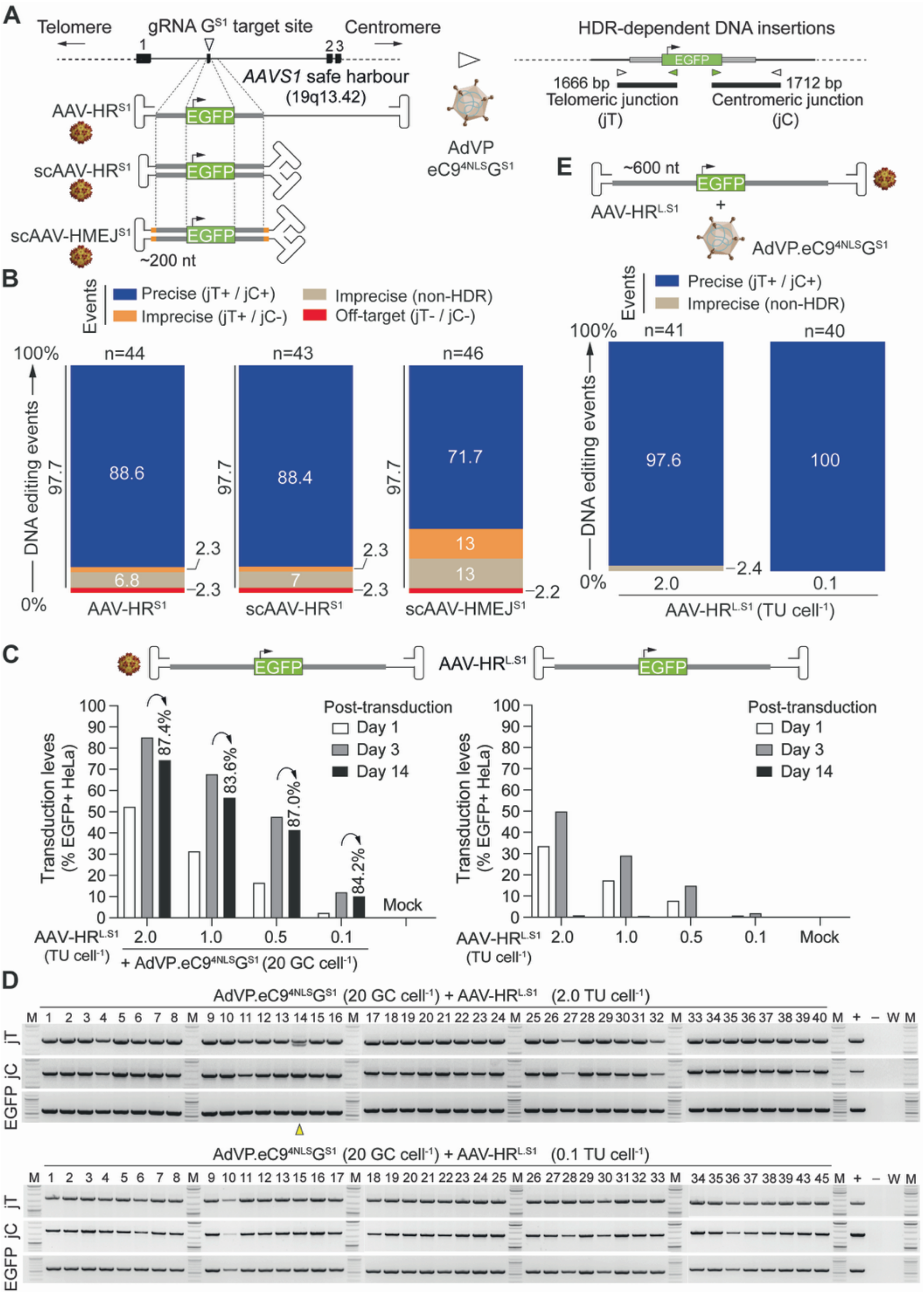


Figure 6. Investigating the role of AAV donor structures on the specificity and fidelity of genome editing. (**A**) Experimental setup. Homology arms in regular and sc AAV donors are drawn in relation to the *AAVS1* acceptor sequence before and after gene targeting. PCR amplicons diagnostic for HDR-mediated formation of exogenous-endogenous chromosomal junctions from the telomere and centromere side are also

depicted (jT and jC, respectively). (**B**) Cumulative characterization of genome editing events resulting from AAV donors with different structures and designs. Individual EGFP-positive HeLa cell clones genetically modified via transduction with AdVP.eCas9^{4NLS}G^{S1} and AAV-HR^{S1}, scAAV-HR^{S1} or scAAV-HMEJ^{S1} were subjected to junction PCR screens targeting HDR-derived telomeric and centromeric junctions (jT and jC, respectively). The frequencies of gene targeting events resulting from HDR at both termini (jT+/jC+) or involving imprecise HDR-independent processes (non-HDR and jT+/jC-), are plotted. The frequencies of clones with off-target insertions (jT-/jC-) are equally plotted. The respective PCR screening data are presented in **Supplementary Figure S13**. (**C**) Transient and stable transduction levels as a function of AAV donor input. HeLa cells transduced with a fixed MOI of AdVP.eC9^{4NLS}G^{S1} together with different MOI of AAV-HR$^{L.S1}$ were monitored through flow cytometry at early and late timepoints post-transduction to quantify initial and stable transduction levels corresponding to DSB-dependent genome editing frequencies. Controls consisted of HeLa cells exposed exclusively to the AAV-HR$^{L.S1}$ dose-range. (**D**) Characterization of genome editing outcomes using AAV-HR$^{L.S1}$. Screening of individual EGFP-positive HeLa cell clones genetically modified via transduction with AdVP.eCas9^{4NLS}G^{S1} and AAV-HR$^{L.S1}$ at 2 TU cell^{-1} (top panel) or 0.1 TU cell^{-1} (bottom panel). All randomly selected clones yielded amplicons diagnostic for HDR-derived telomeric and centromeric junctions (jT and jC, respectively). One clone yielded, in addition, an amplicon consistent with imprecise HDR-independent targeted insertion (yellow arrowhead). PCR mixtures with DNA from EGFP-positive cell populations and water provided for positive and negative controls, respectively. *EGFP* served as an internal control template. Lanes M, GeneRuler DNA Ladder Mix molecular weight marker. (**E**) Plots summarizing the PCR screening data presented in panel (**D**).

Earlier and recent research has established that AAV vectors yield imprecise genome-editing events at appreciable levels *in vitro* and *in vivo* whose origins are consistent with end-joining of vector DNA at on-target genomic positions (‘capture’), including fragment derivatives containing ITR sequences [61–63]. Related to this, it is well-established that, in addition to full-length genomes, AAV particles package heterogeneous species, such as, truncated genomes with ITR sequences [64,65]. Notably, the previous clonal analyses revealed that combining high-specificity CRISPR-Cas9 complexes with single-stranded AAV HR donor templates mostly yields precise HDR-mediated gene targeting at *AAVS1* (**Figure 6B**, **D** and **E**). Indeed, these analyses disclosed that most genome-modifying events consisted of on-target exogenous DNA insertions resulting from HDR at both termini, with such outcomes being especially prevalent when implementing longer ‘homology arms’ in the vector design (**Figure 6D** and **E**). Hence, we set out to further explore the genome-editing specificity and fidelity achieved with the dual viral vector system and, in parallel, probe the role that high-specificity CRISPR-Cas9 complexes have on these parameters. To this end, after exposing hMSCs to AAV-HR^{S1} and high-specificity or regular CRISPR-Cas9 complexes (*i.e.* eC9^{4NLS}:G^{S1} or Cas9:G^{S1}, respectively), stably transduced cell populations were sorted by EGFP-directed flow cytometry and subjected to junction PCR analyses diagnostic for on-target and off-target insertions at *AAVS1* and *CPNE5*, respectively (**Figure 7A**). Regarding the latter locus, previous work from our laboratory tracing off-target DNA cleavage in HEK293T cells through orthogonal high-throughput genome-wide translocation sequencing validated *CPNE5* as the top-ranked off-target site for gRNA G^{S1} [34,35]. Junction PCR analysis detected HDR-mediated *AAVS1* gene targeting in hMSCs exposed to AAV-HR^{S1} together with either eC9^{4NLS}:G^{S1} or regular Cas9:G^{S1} complexes (**Figure 7B**). Yet, *CPNE5* off-target insertions were detected exclusively in hMSCs initially exposed to regular Cas9:G^{S1} complexes (**Figure 7C**). Moreover, these insertions displayed heterogenous sizes consistent with end-ligation of truncated and/or rearranged AAV vector DNA. Related to the emergence of these byproduct species, as forementioned, it is known that AAV capsids package heterogeneously sized vector DNA that can integrate as such or as multimers. In fact, recent experiments have demonstrated that fragments encompassing AAV ITR sequences contribute to a significant fraction of off-target insertions [63]. Besides insertional mutagenesis, AAV ITR insertions create additional safety concerns in the form of potential deregulation of cellular

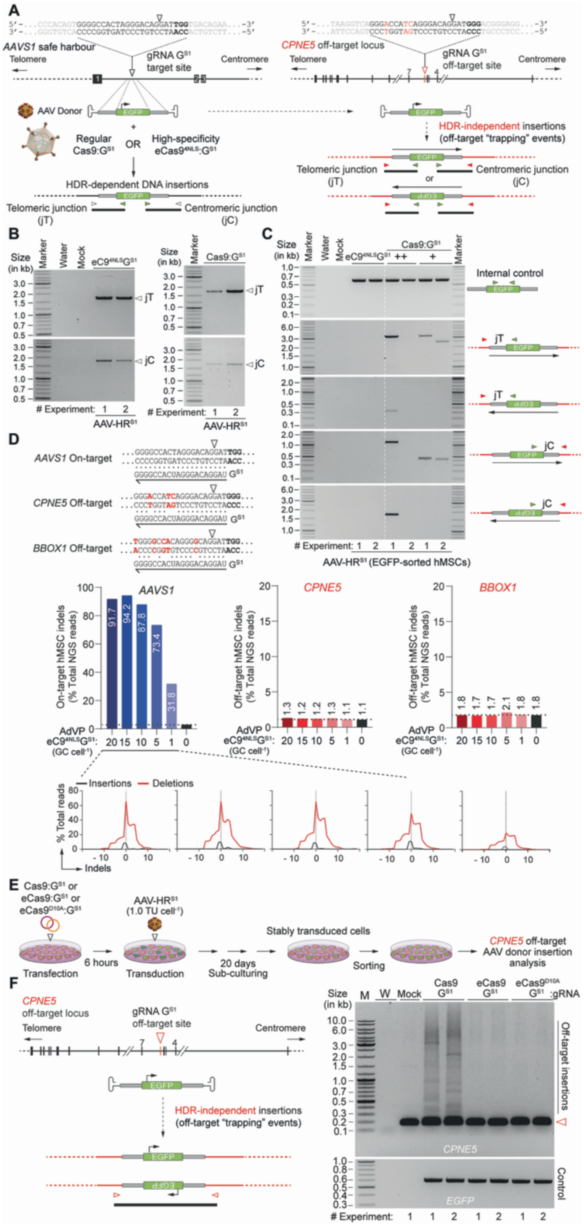


Figure 7. Investigating off-target AAV donor insertions resulting from using regular versus high-specificity CRISPR-Cas9 complexes. (A) Experimental setup for experiments in hMSCs. AAV HR donor sequences before and after chromosomal insertion at *AAVS1* and off-target *CPNE5* loci through HDR-dependent and HDR-independent processes, respectively, are shown. Likewise for the respective amplicons specific for on-target and off-target insertions. In addition to AAV-HR^{S1}, hMSCs were exposed to high-

specificity eCas9^{4NLS}:G^{S1} or regular Cas9:G^{S1} nucleases. (**B**) Establishing AAV gene targeting. EGFP-sorted hMSCs genetically modified via the delivery of AAV-HR^{S1} together with high-specificity eCas9^{4NLS}:G^{S1} or conventional Cas9:G^{S1} nucleases, were subjected to junction PCR assays detecting HDR-derived telomeric and centromeric junctions (jT and jC, respectively) (n = 2 biological replicates). (**C**) Assessing AAV gene targeting specificity. EGFP-sorted hMSCs genetically modified through the transfer of AAV-HR^{S1} together with high-specificity eCas9^{4NLS}:G^{S1} or regular Cas9:G^{S1} nucleases, were subjected to junction PCR assays specific for HDR-independent insertions at the validated top-ranked gRNA G^{S1} off-target site at *CPNE5*. The regular Cas9:G^{S1} complexes were applied using two total doses of AdVs, *i.e.* 32 and 64 GC cell^{-1}; the high-specificity eCas9^{4NLS}:G^{S1} complexes were applied at 20 GC cell^{-1}. (**D**) Assessing eCas9^{4NLS}:G^{S1} nuclease-induced DNA cleaving specificity. hMSCs were transduced with AdVP.eC9^{4NLS}G^{S1} at the indicated MOI and DNA cleaving at the *AAVS1* target site and at the highest and intermediate-ranked validated off-target sites in *CPNE5* and *BBOX1* loci, respectively, were quantified by next-generation deep sequencing at 3 days post-transduction (~50000 paired-end reads per sample). Mismatches between the spacer of gRNA G^{S1} and off-target DNA sequences within *CPNE5* and *BBOX1* alleles are highlighted in red. Histograms corresponding to the types and distributions of insertions and deletions (indels) at *AAVS1* established after NHEJ repair of DSBs induced by eC9^{4NLS}:G^{S1} complexes are also plotted. (**E**) Experimental setup for experiments in HeLa cells. HeLa cells transfected with plasmids encoding regular Cas9:G^{S1} or high-specificity eCas9^{4NLS}:G^{S1} nucleases or eCas9^{4NLS}:G^{S1} nickase, were transduced with donor AAV-HR^{S1} at 1.0 TU cell^{-1}. After sub-culturing and EGFP-directed cell sorting, stably transduced HeLa cell populations were subjected to off-target AAV donor insertion analysis by *CPNE5*-targeted PCR. (**F**) Off-target AAV donor DNA insertion analysis. Schematics of AAV donor-derived sequences inserted through non-homologous end-joining processes at the major off-target site of gRNA G^{S1} at *CPNE5*, are depicted (left panel). PCR analysis using a primer pair flanking the major off-target site of gRNA G^{S1} at *CPNE5* (right panel). Open arrowhead mark *CPNE5* sequences without donor insertions; *EGFP* served as an internal control template for availability and integrity of transgenic DNA (n = 2 biological replicates).

gene expression due to their intrinsic promoter activity [66,67]. Notably, high-throughput deep sequencing established a direct correlation between the detection of donor DNA insertions at *CPNE5* (**Figure 7C**) and DSB formation at this major gRNA G^{S1} off-target locus (**Supplementary figure**

S14). In line with this outcome, no exogenous DNA insertions were detected at *CPNE5* (**Figure 7C**) in the absence of DSB formation at this locus (**Figure 7D** and **Supplementary figure S14**). Finally, additional AAV-HR^{S1} transduction experiments in cells exposed to *AAVS1*-targeting CRISPR-Cas9 complexes containing gRNA G^{S1} and either regular or high-specificity Cas9 nucleases or a Cas9^{D10A} nickase (**Figure 7E**), confirmed the insertion of AAV donor DNA species at *CPNE5* in cells receiving regular Cas9:G^{S1} nuclease complexes (**Figure 7F**).

Taken together, these data stress the importance of selecting high-specificity CRISPR-Cas9 nucleases as opposed to parental Cas9 nucleases for achieving targeted and precise genome editing when using AAV donors.

In this study, we introduce a dual viral vector genome-editing system based on the transfer of engineered CRISPR-Cas9 nucleases and donor DNA templates via AdVP and AAV particles, respectively. In common, these delivery agents lack the entire viral-gene complement of their parental viruses allowing for broader dose-ranges than those permitted when using viral gene-containing AdVs. Well-defined complementary attributes of AAV and AdVP systems (*e.g.* HDR-prone and strictly episomal genomes, respectively), were exploited for achieving selection-free and precise engineering of various human cell types. These experiments involved targeting commonly used safe harbor loci (*i.e.* *AAVS1* and *CCR5*) and tagging *DMD* and *LMNA* alleles whose mutations cause notable human disorders. Specifically, mutations in former gene underpin *DMD* whilst in latter cause Emery–Dreifuss muscular dystrophy, limb girdle muscular dystrophy and Hutchinson–Gilford progeria syndrome, amongst others. We further exploited this dual viral vector platform to address the knowledge gap regarding the relationship between genome editing endpoints and different AAV donor designs (*i.e.* conventional HR versus 'double-cut' HMEJ templates) and structures (*i.e.* single-stranded versus double-stranded). These experiments revealed that, in contrast to HMEJ donors delivered in the context of circular plasmid DNA [7–9] or linear protein-capped adenovector genomes [13,26], AAV HMEJ donors do not necessarily outperform their AAV HR donor counterparts in terms of global gene knock-in proficiencies, which is consistent with data released while this work was under review [68]. It is conceivable, nonetheless, that in different cell types or genomic positions, differences in genome editing efficiencies between AAV vectors presenting HMEJ versus HR templates will arise. Possibly, incoming AAV genomes harboring intrinsically recombinogenic secondary-structured ITRs [16,17] preclude, to a large extent, the need for DSB-mediated HMEJ donor excision in order to facilitate target DNA engagement. On the other hand, earlier studies demonstrated that incoming AAV vector genomes form ITR head-to-tail concatemers in transduced cells [69,70]. Recent experiments established that such multiple copy genomes can be readily detected at CRISPR-Cas9-induced chromosomal breaks when derived from AAV HR donors [68]. These byproducts, alike head-to-tail integrase-defective lentivector donor concatemers [13], are expected to be substantially more disruptive when editing coding sequences than when targeting safe harbor loci introns or other non-coding sequences to, for instance, complement a genetic defect. Interestingly, supporting the view that the AAV ITRs drive end-to-end DNA recombination, targeted AAV donor concatemer insertions seem to be significantly diminished when using ‘double-cut’ AAV HMEJ donors [68]. However, the fate of the ITRs released upon DSB-processing of AAV HMEJ donors warrants further investigation. This is especially so in therapeutic contexts where off-target insertion of these transcriptionally competent elements can lead to insertional mutagenesis and, potentially, proto-oncogene deregulation and oncogenesis [63,66,67]. Moreover, determining whether additional DSB formation upon AAV HMEJ donor cleavage exacerbates DDR activation merits further investigation.

Consistent with earlier experiments using HMEJ and other ‘double-cut’ templates embedded in plasmid DNA [7–9] or protein-capped adenovector genomes [13,26], herein tested scAAV HMEJ donors yield lower amounts of precise HDR-derived events than their single-stranded or scAAV HR counterparts. Presumably, as for the other aforementioned types of vehicles, generation of free-ended double-stranded DNA from scAAV HMEJ substrates favors end-joining products at DSBs (targeted or otherwise). On the other hand, it will be valuable determining whether, similarly to 'double-cut' AAV HITI donors, AAV HMEJ donors also outperform AAV HR designs in non-dividing or post-mitotic cells. Ultimately, the selection of individual AAV donor substrates will depend on a judicious assessment of their performance taking into account the goal and specific settings in which they are meant to operate. Regardless, in the context of AAV-based genome editing procedures the use of high-specificity CRISPR-Cas9 complexes is critical to mitigate HDR-independent insertions of defective AAV genomic species at off-target chromosomal positions.

In conclusion, the dual viral vector platform introduced herein permits investigating different genome-editing reagents and strategies and offers the prospect for the effective and accurate chromosomal edition of hard-to-transfect cell types with scientific and therapeutic relevance. Applications may include establishing human disease-in-a-dish models and engineering cellular substrates for drug screens and, eventually, autologous cell therapies.

Materials and methods

Cells

The human embryonic kidney 293T (HEK293T) cells (ATCC) and the human cervix carcinoma (HeLa) cells (ATCC) were maintained in high-glucose Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific; cat. no.: 41966–029) containing, respectively, 10% and 5% fetal bovine serum (FBS; Biowest; cat. no.: S1810-500). Both cell lines were cultured at 37°C in a 10% CO_2 atmosphere. The PEC3.30 packaging cell line used for the assembly of AdVPs [20], were cultured in high-glucose DMEM with 10% FBS, 10 mM $MgCl_2$ and 0.4 $\mu g\ ml^{-1}$ puromycin (Thermo Fisher Scientific; cat. no.: A11138-03) at 39°C in a 10% CO_2 atmosphere. The bone marrow-derived primary human mesenchymal stem cells (hMSCs) were kept in Minimum Essential Medium α (MEM-α; Thermo Fisher Scientific; cat. no.: 22561–021) with 10% FBS, 100 U ml^{-1} penicillin/streptomycin (Thermo Fisher Scientific; cat. no.: 15140–122), 1 × non-essential amino acids (Thermo Fisher Scientific; cat. no.: 11140–050), 1 × GlutaMax supplement (Thermo Fisher Scientific; cat. no.: 35050–061) and 5 ng ml^{-1} Recombinant Human Fibroblast Growth Factor-basic (FGF-2; Peprotech; cat. no.: 100–18B), at 37°C in a 5% CO_2 atmosphere. The collection of human primary cells from bone marrow was, in accordance with the Best Practices Code of the Dutch Federation of Biomedical Scientific Societies, carried out from anonymous ‘left-over’ surgery material. No informed consent was demanded for the usage of such anonymous and non-traceable body materials and the institutional ethics committee of the LUMC waived the permission from the donor. The human muscle progenitor myoblasts, a gift from Vincent Mouly (Sorbonne University, Paris, France), were maintained at 37°C in a 5% CO_2 atmosphere in Ham’s F-10 Nutrient Mix containing 1 × GlutaMAX supplement (Thermo Fisher Scientific; cat. no.: 41550–021), 20% heat-inactivated FBS (HI FBS; Thermo Fisher Scientific; cat. no.: 10500–064), 10 ng ml^{-1} FGF-2, 1 μM Dexamethasone (Sigma-Aldrich; cat. no.: D2915) and 100 U ml^{-1} penicillin/streptomycin.

Recombinant DNA

The generation of all the plasmids used herein followed standard recombinant DNA techniques. The annotated maps and nucleotide sequences of the constructs used for the assembly of the nuclease-encoding high-capacity AdVP.C9^{KARA} and AdVP.eC9^{4NLS}G^{S1} particles (*i.e.* U67_pHC-AdVP.C9^{KARA} and S90_pHC-AdVP.eC9^{4NLS}G^{S1}, respectively); for the assembly of the single-stranded AAV donor vectors without gRNA expression units AAV-HR^{S1}, AAV-HMEJ^{S1}, AAV-HR$^{L.S1}$ and AAV-HMEJ$^{L.S1}$ (*i.e.* BG14_pAAV-HR^{S1}, BG13_pAAV-HMEJ^{S1}, BE22_pAAV-HR$^{L.S1}$ and AG68_pAAV-HMEJ$^{L.S1}$, respectively), for the assembly of the single-stranded AAV donor vectors with gRNA expression units AAV-

HR$^{L.S1}$G^{S1}, AAV-HR^{C5}G^{C5} and AAV-HR^{DMD}G^{IN4} (*i.e.* AK35_pAAV-HR$^{L.S1}$G^{S1}, BF12_pAAV-HR^{C5}G^{C5} and Y37_pAAV-HR^{DMD}G^{IN4}, respectively); and for the assembly of scAAV donor vectors pscAAV-HR^{S1} and pscAAV-HMEJ^{S1} (*i.e.* BG12_pscAAV-HR^{S1} and BG11_pscAAV-HMEJ^{S1}, respectively), are available in the **Supplementary information**. The latter plasmids were generated on the basis of the scAAV construct pscAAV-GFP (Addgene plasmid #32396) [21]. The construct AAV-HR^{LMNA}G^{EX1} contains the gRNA unit G^{EX1} for site-specific DNA cleavage at *LMNA* and a matched HR template for LMNA tagging with a mScarlet-I reporter at the N-terminus. The HR template was retrieved from plasmid LMNA_mScarlet-I_Donor (Addgene plasmid #178092) [22]. Finally, the plasmid AT51_pDG6.RSV.DsRed.SV40pA, used as the AAV packaging plasmid, expresses AAV serotype-2 *rep* and AAV serotype-6 *cap* proteins together with adenovirus helper functions for AAV production [23].

AdVP vector production, purification and characterization

The generation, purification, and characterization of the second-generation AdVs encoding the *AAVS1*-targeting gRNA G^{S1} and the regular *S. pyogenes* Cas9 endonuclease have been detailed elsewhere [24]. The generation of the high-capacity AdVP.C9^{KARA} and AdVP.eC9^{4NLS}G^{S1} particles was done based on the molecular clones U67_pHC-AdVP.C9^{KARA} (**Supplementary information**) and S90_pHC-AdVP.eC9^{4NLS}G^{S1} (**Supplementary information**), respectively, using procedures previously described [20,25,26]. In brief, the initial packaging of AdVP genomes into vector capsids was done by transfecting Cre recombinase-expressing and adenovirus type-5 *E1*-complementing PEC3.30 producer cells [20,25,26] seeded at a density of ~1.6 × 10^{6} cells per well of six-well plates (Greiner Bio-One). The next day, each AdVP molecular clone was first digested with MssI (6.25 μg per clone), to remove the plasmid backbone, and then diluted in 200 μl of a 150 mM NaCl solution. Subsequently, 20.6 μl of a 25-kDa linear polyethylenimine (PEI; Polysciences) solution at a concentration of 1 mg ml^{-1} was added to the DNA and, after an immediate *circa* 10-s homogenization in a vortex, was incubated at room temperature (RT) for 15 min to allow for DNA-PEI complex formation. The formed complexes were then directly added in a dropwise fashion to the producer cells medium. At 6 h post-transfection, the transfection medium was replaced with fresh medium containing the *E1*-deleted helper AdV vector AdV.SRα.LacZ.1.50 [27] at a multiplicity of infection (MOI) of 20 infectious units (IU) per cell for providing *in trans* all the proteins necessary for the replication and packaging of AdVP genomes into vector capsids. Importantly, AdV.SRα.LacZ.1.50 has its packaging signal flanked by loxP sites for selective Cre-mediated excision to hinder the assembly of helper vector particles. In addition, the PEC3.30 cells were transferred from 39°C to 34°C to allow for the proper folding of a thermosensitive version of the adenovirus DNA-binding protein expressed in PEC3.30 cells that further contributes to AdVP complementation. Upon the emergence of full cytopathic effect, the producer cells were collected and subjected to freezing and thawing for three cycles using liquid N_2 and 37°C water baths. After a centrifugation step at 2000 × *g* for 10 min, the clarified supernatants with rescued AdVPs were used for serial propagation in increasing numbers of producer cells co-transduced with the helper vector AdV.SRα.LacZ.1.50. For the rescue step and first two passages the helper was applied at 20 IU cell^{-1}; whilst for the third final passage the helper was used at 40 IU cell^{-1}. The AdVPs assembled at the last propagation round, involving 20 T175-cm^2 culture flasks, were purified by sequential block and continuous CsCl buoyant density ultracentrifugation gradients. Finally, the retrieved AdVPs were de-salted by ultrafiltration through Amicon Ultra-15 100K MWCO filters (MerckMillipore, cat. no. UFC910024) against storage buffer A195 [28] consisting of 10 mM Tris (pH 7.4), 75 mM NaCl, 5% (*w/v*) sucrose, 0.02% Tween-80 (Sigma-Aldrich; cat. no.: BCBV8843), 1 mM $MgCl_2$, 100 µm ethylenediaminetetraacetic acid (EDTA), 0.5% (*v/v*) ethanol and 10 mM histidine (Merck; cat. no.: 4350). The vector stocks were kept at –80°C until usage.

The AdVP transducing genome copy (GC) titers were established via quantitative PCR (qPCR) assays using a previously published protocol [25] based on the iQ SYBR Green Supermix (Bio-Rad, cat. no. L010171C) and a primer pair specific for the AdV packaging signal. (*i.e.*5′-CGGTGTACACAGGAAGTGACA-3′ and 5′-CAGATTTCACTTCCTCTT ATTCAG-3′). In brief, HeLa cells seeded one day before at a density of 8 × 10^4 cells per well of 24-well plates (Greinder Bio-One), were exposed to seven 3-fold serial dilutions of each purified AdVP stock. Genomic DNA extracted at 24 h post-transduction with the DNeasy Blood & Tissue kit (QIAGEN; cat. no.: 69506) was then subjected to qPCR analysis. In parallel, eight serial 10-fold dilutions of linearized parental AdVP plasmid stocks containing 1 × 10^7 GC ml^{-1} were used to generate a qPCR standard curve. Data analysis was carried out by using the Bio-Rad CFX Manager 3.1 software (Bio-Rad Laboratories) with the titers being subsequently calculated on the basis of the cycle threshold (Ct) values corresponding to the AdVP genome dilutions and plasmid DNA standard curves [25].

The structural integrity of AdVP genomes was established by restriction fragment length analysis assays as follows. First, 50 µl of purified AdVPs were incubated with 8 µl of 10 mg ml^{-1} DNase I (Sigma-Aldrich, cat. no. 10104159001) for 30 min at 37°C and, subsequently, the DNase I activity was inactivated by adding 1.5 µl of 20 mg ml^{-1} proteinase K (Thermo Fisher Scientific, cat. no. EO0491), 6 µl of 10% (*w/v*) sodium dodecyl sulfate (SDS) and 2.4 µl of 0.5 M EDTA (pH 8.0). After a 1-h incubation period at 55°C, AdVP genomes were isolated by using the QIAEX II Gel Extraction Kit (QIAGEN, cat. no. 20021) according to the manufacturer's recommendations. The isolated AdVP genomes were then subjected to the indicated restriction enzyme digestions and analyzed after agarose gel electrophoresis by using the Gel-Doc XR + system (Bio-Rad Laboratories) and the Image Lab 6.0.1 software (Bio-Rad Laboratories). The AdVP parental plasmids U67_pHC-AdVP.$C9^{KARA}$ and S90_pHC-AdVP.$eC9^{4NLS}G^{S1}$ digested with the same restriction enzymes applied to AdVP genomes provided for molecular weight references. The *in-silico* restriction patterns corresponding to vector and reference plasmid DNA were generated by using the SnapGene software (version 6.0.7).

AAV vector production, purification and characterization

The AAV vectors were generated by transfecting HEK293T cells seeded one day before in T175-cm^2 culture flasks at a density of 2 × 10^7 cells per flask (up to 18 flasks per AAV stock). The transfected DNA consisted of each AAV transfer plasmid (**Supplementary information**) and the packaging plasmid AT51_pDG6.RSV.DsRed.SV40pA mixed at 1:1 molar ratio (30 μg total DNA per flask). As aforementioned, this packaging construct expresses AAV serotype-2 *rep* and AAV serotype-6 *cap* together with adenovirus helper functions [23]. For each T175-cm^2, mixtures consisting of DNA and 99 μl of PEI (1 mg ml^{-1}) each diluted in 1 ml of 150 mM NaCl. Transfection mixtures were generated by dropwise addition of the PEI to the DNA solution followed by immediate homogenization in a vortex for 10 s. After incubation at room temperature for 16–18 min, the formed DNA-PEI complexes were directly added to the HEK293T cells and, 24 h later, the transfection medium was replaced by 20 ml of fresh medium. At 5 days post-transfection, the producer cells were detached with a cell scrapper and transferred together with the conditioned medium into 50-ml tubes and then centrifuged at 1000 × *g* for 10 min at 4°C. The resulting supernatants and cell pellets were separately recovered and stored at –80°C overnight. After thawing, each 100 ml of supernatant received 25 ml of a 40% (*w/v*) poly(ethylene) glycol 8000 solution (PEG 8000; Sigma-Aldrich; cat. no.: P2139) with the resulting mixture being subsequently slowly stirred for 1 h at 4°C followed by an overnight incubation at 4°C without stirring for full precipitation. Next, the supernatant-PEG 8000 mixtures were centrifuged at 2820 × *g* for 15 min at 4°C in 50-ml tubes and the resulting pellets were resuspended in 7 ml of phosphate-buffered saline (PBS) (pH 7.4) and mixed with 10 ml of clarified cell lysates to yield 17 ml of vector suspensions. The clarified cell lysates were obtained after resuspending the producer cell pellets in 10 ml of PBS (pH 7.4), subjecting the resuspended cells to three cycles of freezing and thawing in liquid N_2 and 37°C water baths, respectively, and finally removing the resulting cell debris by centrifugation at 3220 × *g* for 15 min at 4°C. The 17-ml AAV vector suspensions were subsequently treated with 50 U ml^{-1} of Benzonase (Millipore; cat. no.: E1014-25KU) for 1 h at 37°C and then centrifuged at 2420 × *g* for 10 min at 4°C. Next, clarified supernatants containing AAV particles were loaded onto Iodixanol-OptiPrep (Progen; cat. no.: 1114542) cushions of 15%, 25%, 40% and 60% placed in Quick-Seal round-top polypropylene tubes (Beckman; cat. no.: 342414). The AAV particles were then purified and concentrated via iodixanol gradient ultracentrifugation at 69 000 revolutions per minute in a 70Ti rotor (Beckman Coulter) at 16°C in a Beckman Coulter Optima XE-90 centrifuge. By piercing the ultracentrifuge tubes with a needle (18G needle BD Microlance^{TM}; cat. no.: 304622), the majority of AAV particles located within the 40% iodixanol cushion, were collected and subjected to buffer exchange using Amicon Ultra-15 100K MWCO filters (Millipore; cat. no.: UFC910024) and Dulbecco's phosphate-buffered saline (Thermo Fisher Scientific; cat. no.: 14040–091) containing 0.001% Poloxamer 188 (Sigma-Aldrich; cat. no.: P5556). Purified AAV batches were stored at –80°C and the respective transducing unit (TU) titers were determined by limiting dilution assays on HeLa cells as follows. Firstly, 5×10^4 cells were seeded in wells of 24-well plates (Greiner Bio-One) and 16–18 h later, the cells were exposed to 3-fold serial dilutions of each vector preparation. At 3 days post-transduction, the frequencies of transduced cells were determined by flow cytometry with the corresponding functional AAV vector titers corresponding to HeLa-cell TU per ml being calculated as the percentage of transduced cells × cells seeded × dilution factor x 1000/µl. The AAV vector batches produced, and their respective titers are listed in **Supplementary Table S1**.

The transducing GC titers of AAV-$HR^{LMNA}G^{EX1}$ were established via qPCR assays using the iQ SYBR Green Supermix (Bio-Rad, cat. no. L010171C) and a primer pair specific for the *mScarlet-I* reporter (*i.e.* 5′-CTACCTGGCGGACTTCAAGA-3′ and 5′-ACGGTGTAGTCCTCGTTGTG-3′). In brief, HeLa cells seeded one day before at a density of 8.5 × 10^4 cells per well of 24-well plates (Greinder Bio-One), were incubated with seven 3-fold serial dilutions of the purified AAV stock. Genomic DNA isolated at 24 h post-transduction with the DNeasy Blood & Tissue kit was subsequently used for the qPCR analysis. In parallel, eight serial 10-fold dilutions of the linearized parental AAV construct containing 1 × 10^7 GC ml^{-1} were used for qPCR standard curve generation. Data analysis was performed with the Bio-Rad CFX Manager 3.1 software (Bio-Rad Laboratories) and the titer determined based on the AAV DNA and plasmid standard curve Ct values.

Selection of *DMD*-targeting gRNA

The screens to identify a gRNA for AAV donor DNA targeting *DMD* in human myoblasts was done via transient transfections and viral vector transductions on HeLa cells and human myoblasts, respectively. The transient transfections started by seeding 4 × 10^4 HeLa cells per well of 24-well plates and, the next day, the cells were exposed to complexes formed by incubating 150 mM NaCl solutions containing 1.1 µl of PEI (1 mg ml^{-1}) and 300 ng of plasmid mixtures at 1:1 ratio. The later mixtures consisted of 218.5 ng of AV50_pCAG.eSpCas9.bGHpA encoding the high-specificity eSpCas9(1.1) nuclease [29] and 81.5 ng of each of the test gRNA constructs, *i.e.* BC52_pgRNA$^{DMD.IN3.1}$, BC53_pgRNA$^{DMD.IN3.2}$, BC54_pgRNA$^{DMD.IN3.3}$, BC55_pgRNA$^{DMD.IN4.1}$ and BC56_pgRNA$^{DMD.IN4.2}$ (**Supplementary information**). As controls, previously validated and non-targeting gRNA constructs, *i.e.* AY60_pgRNA$^{AAVS1-T2}$ [30] and AM51_pgNT [31], respectively, provided for positive and negative controls, respectively. All gRNAs contain an optimized scaffold [32] derived from the gRNA acceptor construct AY56_pUC.U6.opt-sgRNA.BveI-stuffer [30]. At 2 days post-transfection, genomic DNA was isolated with the DNeasy Blood & Tissue Kit (Qiagen; cat. no.: 69506) and targeted DSB formation was assessed via

T7 endonuclease I (T7EI)-based genotyping assays on amplicons spanning the target sites of the *DMD*-specific gRNAs.

Lentiviral vector production and characterization

The lentiviral vector transfer plasmids pLV.gRNA$^{DMD.IN3.1}$, pLV.gRNA$^{DMD.IN3.3}$ and pLV.gRNA$^{DMD.IN4.2}$ were assembled by inserting the gRNA expression units of, respectively, BC52_pgRNA$^{DMD.IN3.1}$, BC54_pgRNA$^{DMD.IN3.3}$ and BC56_pgRNA$^{DMD.IN4.2}$ (**Supplementary information**) into a lentiviral vector acceptor construct encoding the EGFP reporter. The corresponding vector particles were generated according to previously specified protocols [33]. In brief, 17 × 10^6 HEK293T cells were seeded in 175-cm^2 culture flasks (Greiner Bio-One). The next day, 30 µg of a DNA cocktail diluted in 1 ml of NaCl and 90 µl of a 1 mg ml^{-1} PEI solution diluted in 1 ml of NaCl were mixed by dropwise addition of the PEI to the DNA under gentle agitation followed by vigorous homogenization with a vortex for 10 s. After an incubation for 15 min at room temperature, the DNA-PEI complexes were directly added to the HEK293T cell culture medium. The DNA cocktail consisted of mixtures at 2:1:1 molar ratio of each lentiviral transfer plasmid, a packaging construct (psPAX2; Addgene #12260, a gift from Didier Trono) and a VSV-G-pseudotyping construct (pLP/VSVG; Invitrogen). At 2 days post-transfection, the conditioned medium was harvested, and the cellular debris was eliminated by centrifugation. The supernatants were then filtrated through 0.45-µm pore-sized cellulose acetate filters. Finally, the resulting clarified vector preparations were aliquoted and titrated by reporter-directed flow cytometry essentially as above-described for the titration of AAV vector batches.

Viral vector transductions

HeLa cells, hMSCs and human myoblasts were seeded at densities of, respectively, 5 × 10^4, 6 × 10^4 and 5 × 10^4 cells per well of 24-well plates (Greinder Bio-One) and, 16–18 h later, these cells were either mock-transduced or were transduced for 24 h with single or dual viral vector combinations at the indicated MOIs. Gene editing experiments were assessed by reporter-directed flow cytometry at early and late timepoints post-transduction for quantifying transient and stable transduction levels, respectively. Gene editing experiments in HeLa cells were assessed at 3 days and 14 days post-transduction and in hMSCs and human myoblasts were examined at 3 days and 20 days post-transduction. Additional gene editing control experiments in HeLa cells were assessed at 3-, 7-, 10-, 14- and 17-days post-transduction. These extra control experiments included the use of AAV-HR^{S1} and AAV-HR^{S1}G^{S1} at 2 TU cell^{-1} and AdVP.C9^{KARA} and AdVP.mCherry [25] at 0, 5 or 20 GC cell^{-1}.

Live-cell microscopy analyses

The live-cell imaging of cells expressing reporter proteins at the indicated early and late timepoints post-transduction was done with the aid of an AF6000 LX inverted fluorescence microscope (Leica). The acquired images were examined with the aid of LAS X (Leica Microsystems). The monitoring of hMSCs transduced with second-generation AdV versus AdVP vectors was done in an Incucyte S3 live-cell imaging analysis system (Sartorius) and real-time analyzed by using the Incucyte software (Essen BioScience).

Flow cytometry

The quantification of cell transduction efficiencies and stable transduction efficiencies as well as mean fluorescence intensities was performed on a per sample basis by using a BD LSR II flow cytometer (BD Biosciences). In brief, cells were washed with PBS (pH 7.4) and, after treatments with a 0.05% trypsin-EDTA solution (Thermo Fisher Scientific; cat. no.: 15400–054), cells present in the resulting cell suspensions were collected by a brief centrifugation and resuspended in FACS buffer consisting of PBS (pH 7.4) supplemented with 0.5% (*w/v*) bovine serum albumin and 2 mM EDTA (pH 8.0). Mock-transduced cells treated in parallel, were used to establish the background fluorescence threshold levels. For each sample, at least 10 000 viable single cells were acquired and subsequently analyzed with the aid of the FlowJo 10.9.0 software (BD Biosciences).

Cell viability assays

hMSCs were seeded at a density of 4 × 10^3 cells per well of 96-well plates (Greiner-BioOne) and, 16–18 h later, they were exposed for 24 h to second-generation vector AdV.Δ2.Cas9 or to high-capacity vector AdVP.C9^{KARA} at different MOIs. Mock- and vector-transduced hMSCs were then kept in an Incucyte S3 apparatus during 4 days for live-cell real-time monitoring of cell confluency and division rates. The Incucyte software (Essen BioScience) was used for this purpose (https://figshare.com/s/a999fdb4400aaf1828c5). Moreover, at consecutive 4 days post-transduction, the metabolic activities of hMSCs were determined by colorimetric quantification of the conversion of the tetrazolium salt (WST-8) to formazan using the CCK-8 kit (Boster Biological Technology; cat. no.: AR1160). In brief, after substituting regular culture medium by 100 µl of medium containing 10 µl of CCK-8 solution, hMSCs were incubated for 1 h at 37°C. Next, the absorbance of the solution in each well, directly proportional to the amounts of viable metabolically active cells, was measured at 450 nm in a multimode plate reader (PerkinElmer VICTOR X3).

Cell differentiation assays

Myogenic differentiation was initiated by exposing to low-mitogen medium confluent cultures of unmodified myoblasts and of *DMD* edited myoblasts seeded in plates pre-coated with 0.1% (*w/v*) gelatin. This differentiation medium consisted of phenol red-free DMEM (Thermo Fisher Scientific; cat. no.: 11880–028) supplemented with 100 U ml^{-1} penicillin/streptomycin, 10 µg ml^{-1} human insulin (Sigma-Aldrich; cat. no.: 19278) and 100 µg ml^{-1} human holo-transferrin (Sigma-Aldrich; cat. no.: T0665). After 4–5 days under differentiation conditions, total RNA was isolated from the resulting syncytial myotubes for multiplex reverse transcriptase-polymerase chain reaction (RT-PCR) analysis targeting endogenous and tagged *DMD* transcripts.

The capacity of mock- and AdVP-transduced hMSCs to differentiate into osteoblasts was tested as follows. After seeding hMSCs at 5 × 10^4 cells per well of 24-well plates, the cells were either treated or not treated with AdVP.C9^{KARA} at 2 TU cell^{-1}. After a 10-day sub-culturing period, the hMSCs were seeded at 1 × 10^4 cells in wells of 48-well plates and cultured for 2 weeks in osteogenic differentiation medium consisting of MEM-α containing 10% FBS, 100 U ml^{-1} penicillin/streptomycin, 1 × non-essencial amino acids, 1 × GlutaMax, 0.2 mM l-ascorbic acid 2-phosphate (Sigma-Aldrich; cat. no.: A8960), 10 mM β-glycerophosphate (Sigma-Aldrich; cat. no.: G6251), 2 μM dexamethasone (Sigma-Aldrich; cat. no.: D4902) and 100 ng ml^{-1} of recombinant human bone morphogenetic protein 6 (BMP6; PeproTech; cat. no.: 120–06). The differentiation medium was refreshed every 3–4 days and alizarin red S staining was performed for detecting calcium deposits. In brief, after fixing hMSCs in 4% paraformaldehyde in phosphate-buffered saline (PBS; pH 7.4) for 10 min, the cells were incubated at room temperature in the dark for 5 min in a 2% Alizarin Red S (pH 4.25) solution (Sigma-Aldrich; cat. no.: A5533). Finally, after sequential large-volume washes with PBS, bright-field microscopy pictures were acquired for the visualization of calcium deposits.

Gene-edited cell sorting, clonal expansion and screening

The sorting of EGFP-positive HeLa cells and of mCherry-positive human myoblasts was performed through a BD FACSAria III flow cytometer (BD Biosciences) or a CytoFLEX SRT Cell Sorter (Beckman), after subculturing the transduced cells for >25 days. The individually sorted HeLa cells were plated in wells of 96-well plates containing a 1:1 mixture of regular culture medium and FBS with penicillin/streptomycin at 100 U ml^{-1}. In addition, to increase the cloning efficiency, this mixture was supplemented with 50 μM α-thioglycerol and 20 nM bathocuprione disulphonate (both from Sigma-

Aldrich). The procedure for isolating individual mCherry-positive human myoblast clones was similar to that applied to HeLa cells except that heat-inactivated FBS was used instead. After 2–3 weeks, single cell-derived clones expressing reporter proteins were randomly collected for genomic DNA analysis via junction PCR assays using the Phire Tissue Direct PCR Master mix following the manufacturer's recommendations (Thermo Fisher Scientific; cat. no.: F170L). The compositions of the PCR mixtures and thermocycling parameters used for the clonal screening of HeLa cells are described in the **Supplementary Tables S2** and **S3**; the compositions of the PCR mixtures and thermocycling parameters applied for the clonal screening of human myoblasts are indicated in **Supplementary Tables S4** and **S5**.

T7EI genotyping assays

Targeted DSB formation in HeLa cells and human myoblasts driven by CRISPR-Cas9 complexes containing *DMD*-targeting gRNAs was assessed by T7EI assays. In brief, 10 μl of PCR mixtures containing amplicons spanning the *DMD* target sites were incubated in 15-μl reactions consisting of 5 U of the mismatch-sensing T7EI enzyme (New England Biolabs; cat. no.: M0302) and 1 × NEBuffer 2 (New England Biolabs; cat. no.: B7002S). After incubation at 37°C for 17 min, the DNA samples underwent electrophoresis through 2% (*w/v*) agarose gels in 1 × Tris-acetate-EDTA buffer. Ethidium bromide-stained DNA species were subsequently imaged by using a Gel-DocTM XR + apparatus (Bio-Rad) and the Image Lab 6.0.1 software (Bio-Rad).

Amplicon deep sequencing

To quantify and characterize DNA cleavage at *AAVS1* target sequences and at the *CPNE5* and *BBOX1* off-target sites of gRNA G^{S1} [34], hMSCs were transduced with different MOIs of adenovectors encoding high-specificity $eC9^{4NLS}$:G^{S1} or regular Cas9:G^{S1} complexes. At 3 days post-transduction, genomic DNA was extracted by using the protocol and reagents from the DNeasy Blood & Tissue Kit (Qiagen; cat. no.: 69506). The resulting DNA was then subjected to Illumina MiSeq next-generation amplicon deep sequencing for obtaining 50000 paired-end reads covering the aforementioned genomic regions. The pipeline and reagents used for this amplicon deep sequencing analysis have been previously specified in detail [25,35,36].

Gene-editing specificity assays

EGFP-expressing hMSCs generated via co-transduction with AAV-HR^{S1} donor and adenovectors encoding high-specificity eCas9^{4NLS}:G^{S1} or regular Cas9:G^{S1} complexes, were sorted through a CytoFLEX SRT Cell Sorter (Beckman) and, subsequently, were subjected to genomic DNA extraction by using the DNeasy Blood & Tissue Kit according to the manufacturer's instructions. The AAV vector DNA insertions at on-target *AAVS1* and at off-target *CPNE5* sequences were captured by junction PCR assays with the aid of Platinum SuperFi II DNA Polymerase (Thermo Fisher Scientific; cat. no.: 12361010). Amplicons specific for *EGFP* served as internal controls. The PCR primers and cycling conditions used in these junction PCR assays are listed in **Supplementary Tables S6** – **S9**, respectively.

EGFP-expressing HeLa cells generated by combining transfections of plasmids encoding Cas9:G^{S1}, eCas9^{4NLS}:G^{S1} or eCas9.D10A^{4NLS}:G^{S1} with transductions with the AAV-HR^{S1} donor (1.0 TU cell^{-1}), were sorted and subjected to genomic DNA isolation by using the DNeasy Blood & Tissue Kit following the manufacturer's instructions. The AAV-HR^{S1} transductions took place at 6 h post-transfection with the transfections having been performed with the reagents and conditions indicated in **Supplementary Table S10** on 1×10^5 HeLa cells seeded one day before in wells of 24-well plates. Prior to EGFP-directed cell sorting and genomic DNA isolation, the transfected and transduced HeLa cells were sub-cultured for 20 days. Finally, assessing AAV donor DNA insertions at the *CPNE5* off-target site was conducted through PCR with the primers #2142 and #2145 with sequences 5′-CCTTGGATTCCTCATCCCAG-3′) and 5′-CCCAGGGCTCTACTCAC

ATAG-3′, respectively. PCR products specific for *EGFP* served as internal controls for the availability and integrity of transgenic DNA and were amplified with primers #978 and #979 with sequences 5′-GAGCTGGACGGCGACGTAAACG-3′ and 5′-CGCTTCTCGTTGGGGTC TTTGCT-3′, respectively.

RT-PCR analyses of gene-editing experiments

Total RNA from myotubes differentiated from unmodified and *DMD* edited myoblasts was extracted with TRIzol Reagent (Thermo Fisher Scientific; cat. no.: 15596018) according to the manufacturer's recommendations. Next, 500 ng of the resulting RNA was reversely transcribed into cDNA by using the SuperScript II Reverse Transcriptase kit (Thermo Fisher Scientific; cat. no.: 18064022) and Random Hexamer Primers (Thermo Fisher Scientific; cat. no.: SO142). Next, Platinum SuperFi II DNA Polymerase (Thermo Fisher Scientific; cat. no.: 12361010) was used for multiplexing PCR assays targeting edited and endogenous *DMD* transcripts. The compositions of the RT-PCR mixtures and thermocycling parameters are described in the **Supplementary Tables S11** and **S12**, respectively.

Quantification of AAV vector DNA cell entry

HeLa were seeded at a density of 8 × 10^4 cells per well of 24-well plates. The next day, the cells were transduced with AAV-HR^{S1} at 0.5 or 2.0 TU cell^{-1} in the absence or in the presence of AdVP.C9^{KARA} at 5, 10 or 20 GC cell^{-1}. At 3-, 6-, 12- and 24-h post-transduction, the cells were thoroughly washed by four sequential washes with large volumes of PBS, after which total DNA was isolated by using the DNeasy Blood & Tissue Kit protocol (Qiagen; cat. no.: 69506). AAV vector DNA levels were determined by qPCR using iQ SYBR Green Supermix (Bio-Rad, cat. no. L010171C) together with the *EGFP*-specific primers #1243 (5′-CCATCCTGGTCGAGCTGG-3′) and #1246 (5′-CGGTGGTGCAGATGAACTTC-3′) at 0.3 μm each in a total volume of 15 μl. The PCR mixtures were subjected to an initial denaturation step at 95°C for 5 min followed by 40 cycles at 95°C for 10 s and 60°C for

30-s. A melting curve was generated by subjecting the samples to 95°C for 10 s and subsequently to 65°C to 95°C with 5-s 0.5°C increments. The relative amounts of AAV vector genomes were expressed in PCR Ct values corresponding to individual transduction conditions.

Tracking of AAV versus scAAV transduction kinetics

HeLa cells were seeded at 8×10^4 cells per well of 24-well plates. The next day, the cells were transduced with AAV-HR^{S1} or scAAV-HR^{S1} at 0.5 TU cell^{-1} and placed in an Incucyte S3 live-cell imaging analysis system (Sartorius) at 37°C in a 5% CO_2 atmosphere. Real-time fluorescence intensity measurements derived from 16 independent microscopy fields per well were recorded at 2-h intervals for a period of 3 days using the software G/R Optical Module. Representative timelapses of experiments establishing faster transduction kinetics of scAAV over regular AAV vectors are available via: https://figshare.com/s/8e3e6c8ef5438d9c3064

Tracking AAV vector transduction in cells with chromosomal breaks

HeLa cells were seeded at 8×10^4 cells per well of 24-well plates. The next day, the cells were exposed to complexes formed by incubating 150 mM NaCl solutions containing 1.54 µl of PEI (1 mg ml^{-1}) and plasmid mixtures consisting of 300 ng of AV62_pCAG.Cas9.rBGpA [37] and 100 ng AG66_pgRNA^{CALM2} [36] or 300 ng of AV62_pCAG.Cas9.rBGpA and 100 ng of AG65_pgRNA^{VEGFA} [36] or, as negative control, 300 ng of AV62_pCAG.Cas9.rBGpA and 100 ng of pMoluc (Addgene plasmid #1251). To control for transfection efficiency, all transfection mixtures were spiked with 50 ng of plasmid BE08_pCAG.mCherry.bGHpA expressing a red fluorescence-emitting reporter. After a 6-h incubation period, the transfection medium was removed, and the cells were transduced in regular culture medium containing AAV-HR^{S1} or scAAV-HR^{S1} at 0.5 TU cell^{-1}. The time-dependent AAV and scAAV transduction levels were followed in an Incucyte apparatus at 37°C in a 5% CO_2 atmosphere. Real-time fluorescence intensity measurements derived from 16 independent microscopy fields per well were recorded at 2-h intervals for a period of 3 days using the software G/R Optical Module. Representative timelapses of experiments assessing the role of chromosomal DNA breaks on AAV vector transduction are available via: https://figshare.com/s/7efc29e8d0d34a99142a

Tracking AAV donor knock-in kinetics

HeLa cells were seeded at 8×10^4 cells per well of 24-well plates. The next day, the cells were transduced with AdVP.C9^{KARA} (20 GC cell^{-1}) together with AAV-HR^{S1} (0.5 TU cell^{-1}) or with AAV-HR^{S1}G^{S1} (0.5 TU cell^{-1}). HeLa cells transduced exclusively with AAV-HR^{S1}G^{S1} (0.5 TU cell^{-1}) provided for controls. The tracing of AAV donor stable transduction levels in the presence and absence of targeted DSBs was performed by real-time fluorescence intensity measurements in an Incucyte S3 live-cell imaging analysis system (Sartorius) at 37°C in a 5% CO_2 atmosphere. Hence, expression of the live-cell EGFP reporter served as a proxy for productive AAV transduction monitoring. Sixteen independent microscopy fields per well were recorded at 2-h intervals for a period of 3 days using the G/R Optical Module software. Representative timelapses of experiments monitoring AAV transduction kinetics under conditions favoring or not favoring gene knock-ins are available via https://figshare.com/s/3c09e1031a48e7df84c1

Statistical analyses

The statistical parameters and analyses corresponding to the various experimental datasets are indicated in the respective figure legends whenever applicable.

Acknowledgements

The authors are thankful to Thilo M. Buck and Jan Wijnholds (Department of Ophthalmology, Leiden University Medical Centre, The Netherlands) for their advice during the setting-up of AAV vector production procedures in our laboratory and to Maarten van Dinther for his support with real-time live-cell imaging experiments in the Incucyte live-cell imaging analysis system. The authors also thank the personnel of the Flow Cytometry Core Facility of the LUMC for aiding with the cell sorting procedures and Vincent Mouly (Sorbonne University, Paris, France) for making available the human myoblasts used in this study. Authors in this study are members of the European Reference Network—Neuromuscular Diseases (ERN EURO-NMD).

Author contributions: Z.L. generated and characterized reagents, designed and performed the experiments, examined the datasets and wrote the paper together with M.A.F.V.G.; X.W. performed gene targeting experiments and clonal screens and analyzed the resulting datasets; J.L. generated, characterized and tested reagents; J.M.J. generated, characterized and tested reagents; F.T. setup the *DMD* gene targeting system; R.H. supervised the research and analyzed the results; M.A.F.V.G. designed and supervised the research, analyzed the data and wrote the paper together with Z.L.

Additional Information

Data availability: The results supporting this work are available in the article and accompanying supplementary files. The high-throughput deep sequencing library reads are deposited at the NCBI Sequence Read Archive (SRA) database under BioProject ID number: PRJNA1047301. The movies illustrating the viability of human mesenchymal stem cells transduced with second-generation versus high-capacity adenovector particles, the movies illustrating the faster productive transduction kinetics of self-complementary AAV over regular single-stranded AAV vectors, the movies illustrating the impact of CRISPR-induced chromosomal DNA breaks on productive AAV vector transduction, and the movies monitoring the productive AAV transduction kinetics under conditions favoring and not favoring gene knock-ins are all available through Figshare links indicated in the respective Materials and methods sections.

Supplementary data: Supplementary Data are available at doi: 10.1093/nar/gkae1213

Funding: China Scholarship Council - Leiden University Scholarship (to Z.L.

and X.W.). Research in the author’s laboratory is supported by the Prinses Beatrix Spierfonds, the Dutch Duchenne Parent Project, the Horizon Europe Programme, and the Dutch Research Council (NWO)—Open Technology Programme.

Conflict of interest statement: None declared.

References

1. Pacesa, M.; Pelea, O.; Jinek, M.. Past, present, and future of CRISPR genome editing technologies. *Cell* **2024**, 187, 1076–1100.
2. 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.
3. Liao, H.; Wu, J.; VanDusen, N.J.; Li,Y.; Zheng, Y.. CRISPR/Cas9-mediated homology-directed repair for precise gene editing. *Mol. Ther. Nucleic Acids* **2024**, 35, 102344.
4. 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.
5. 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.
6. Nakade, S.; Tsubota, T.; Sakane, Y.; Kume, S.; Sakamoto, N.; Obara, M.; Daimon, T.; Sezutsu, H.; Yamamoto, T.; Sakuma, T.; *et al.* Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat. Commun.* **2014**, 5, 5560.
7. Chen, X.; Janssen, J.M.; Liu, J.; Maggio, I.; 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.
8. 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.
9. 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.
10. Chen, X.; Gonçalves, M.A.. Engineered viruses as genome editing devices. *Mol. Ther.* **2016**, 24, 447–457.

11. Gonçalves, M.A.; Vries,A.A.. Adenovirus: from foe to friend. *Rev. Med. Virol.* **2006**, 16, 167–186.
12. Gonçalves, M.A.. Adeno-associated virus: from defective virus to effective vector. *Virol. J.* **2005**, 2, 43.
13. 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.
14. Medert, R.; Thumberger, T.; Tavhelidse-Suck, T.; Hub, T.; Kellner, T.; Oguchi, Y.; Dlugosz, S.; Zimmermann, F.; Wittbrodt, J.; Freichel,M. Efficient single copy integration via homology-directed repair (scHDR) by 5′ modification of large DNA donor fragments in mice. *Nucleic Acids Res.* **2023**, 51, e14.
15. Dever, D.P.; Bak, R.O.; Reinisch, A.; Camarena, J.; Washington, G.; Nicolas, C.E.; Pavel-Dinu, M.; Saxena, N.; Wilkens, A.B.; Mantri, S.; *et al.* CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **2016**, 539, 384–389.
16. Holkers, M.; de Vries, A.A.; Gonçalves, M.A. Nonspaced inverted DNA repeats are preferential targets for homology-directed gene repair in mammalian cells. *Nucleic Acids Res.* **2012**, 40, 1984–1999.
17. Gaj, T.; Staahl, B.T.; Rodrigues, G.M.C.; Limsirichai, P.; Ekman, F.K.; Doudna, J.A.; Schaffer, D.V. Targeted gene knock-in by homology-directed genome editing using Cas9 ribonucleoprotein and AAV donor delivery. *Nucleic Acids Res.* **2017**, 45, e98.
18. Tasca, F.; Wang, Q.; Goncalves, M.. Adenoviral vectors meet gene editing: a rising partnership for the genomic engineering of human stem cells and their progeny. *Cells* **2020**, 9, 953.
19. 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.
20. Brescia, M.; *et al.* High-capacity adenoviral vectors permit robust and versatile testing of DMD gene repair tools and strategies in Human cells. *Cells* **2020**, 9, 869.
21. Gray, J.T.; Zolotukhin, S.. Design and construction of functional AAV vectors. *Methods Mol. Biol.* **2011**, 807, 25–46.
22. Levesque, S.; Mayorga, D.; Fiset, J.P.; Goupil, C.; Duringer, A.; Loiselle, A.; Bouchard, E.; Agudelo, D.; Doyon, Y.. Marker-free co-selection for successive rounds of prime editing in human cells. *Nat. Commun.* **2022**, 13, 5909.
23. Grimm, D.; Kay, M.A.; Kleinschmidt, J.A.. Helper virus-free, optically controllable, and two-plasmid-based production of adeno-associated virus vectors of serotypes 1 to 6. *Mol. Ther.* **2003**, 7, 839–850.

24. 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.
25. 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.
26. 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.
27. Janssen, J.M.; Liu, J.; Skokan, J.; Gonçalves, M.A.; 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.
28. Evans, R.K.; Nawrocki, D.K.; Isopi, L.A.; Williams, D.M.; Casimiro, D.R.; Chin, S.; Chen, M.; Zhu, D.M.; Shiver, J.W.; Volkin, D.B.. Development of stable liquid formulations for adenovirus-based vaccines. *J. Pharm. Sci.* **2004**, 93, 2458–2475.
29. 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.
30. 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.
31. Janssen, J.M.; Chen, X.; Liu, J.; Gonçalves, M.A.F.V.. The chromatin structure of CRISPR-Cas9 target DNA controls the balance between mutagenic and homology-directed gene-editing events. *Mol. Ther. Nucleic Acids* **2019**, 16, 141–154.
32. 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.
33. 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.
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. 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.

36. Wang, Q.; Liu, J.; Janssen, J.M.; Gonçalves, M.A.F.V.. Precise homology-directed installation of large genomic edits in human cells with cleaving and nicking high-specificity Cas9 variants. *Nucleic Acids Res.* **2023**, 51, 3465–3484.
37. Chen, X.; Rinsma, M.; Janssen, J.M.; Liu, J.; Maggio, I.; Gonçalves, M.A.. Probing the impact of chromatin conformation on genome editing tools. *Nucleic Acids Res.* **2016**, 44, 6482–6492.
38. Pavani, G.; Amendola, M.. Targeted gene delivery: where to land. *Front. Genome Ed.* **2021**, 3, 682171.
39. Knaän-Shanzer, S.; Van Der Velde, I.; Havenga, M.J.; Lemckert, A.A.; 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.
40. Gonçalves, M.A.F.V.; de Vries, A.A.F.; 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.* **2006**, 15, 213–221.
41. 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.F.. 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.
42. Ellis, B.L.; Hirsch, M.L.; Barker, J.C.; Connelly, J.P.; Steininger, R.J. 3rd; Porteus, M.H.. A survey of ex vivo/in vitro transduction efficiency of mammalian primary cells and cell lines with nine natural adeno-associated virus (AAV1-9) and one engineered adeno-associated virus serotype. *Virol. J.* **2013**, 10, 74.
43. 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.
44. Allen, J.M.; Halbert, C.L.; Miller, A.D.. Improved adeno-associated virus vector production with transfection of a single helper adenovirus gene, E4orf6. *Mol. Ther.* **2000**, 1, 88–95.
45. Ferrari, F.K.; Samulski, T.; Shenk, T.; Samulski, R.J.. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J. Virol.* **1996**, 70, 3227–3234.

46. Qing, K.; Wang, X.S.; Kube, D.M.; Ponnazhagan, S.; Bajpai, A.; Srivastava, A.. Role of tyrosine phosphorylation of a cellular protein in adeno-associated virus 2-mediated transgene expression. *Proc. Natl Acad. Sci. U.S.A.* **1997**, 94, 10879–10884.
47. Cervelli, T.; Palacios, J.A.; Zentilin, L.; Mano, M.; Schwartz, R.A.; Weitzman, M.D.; Giacca, M.. Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins. *J. Cell Sci.* **2008**, 121, 349–357.
48. Lovric, J.; Mano, M.; Zentilin, L.; Eulalio, A.; Zacchigna, S.; Giacca, M.. Terminal differentiation of cardiac and skeletal myocytes induces permissivity to AAV transduction by relieving inhibition imposed by DNA damage response proteins. *Mol. Ther.* **2012**, 20, 2087–2097.
49. Russell, D.W.; Alexander, I.E.; Miller, A.D.. DNA synthesis and topoisomerase inhibitors increase transduction by adeno-associated virus vectors. *Proc. Natl Acad. Sci. U.S.A.* **1995**, 92, 5719–5723.
50. Mano, M.; Ippodrino, R.; Zentilin, L.; Zacchigna, S.; Giacca, M.. Genome-wide RNAi screening identifies host restriction factors critical for in vivo AAV transduction. *Proc. Natl Acad. Sci. U.S.A.* **2015**, 112, 11276–11281.
51. 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.
52. McCarty, D.M.; Monahan, P.E.; Samulski, R.J.. Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther.* **2001**, 8, 1248–1254.
53. McCarty, D.M.; Fu, H.; Monahan, P.E.; Toulson, C.E.; Naik,P.; Samulski, R.J.. Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther.* **2003**, 10, 2112–2118.
54. Ngo, A.M.; Puschnik, A.S.. Genome-scale analysis of cellular restriction factors that inhibit transgene expression from adeno-associated virus vectors. *J. Virol.* **2023**, 97, e0194822.
55. Duan, D.; Goemans, N.; Takeda, S.; Mercuri, E.; Aartsma-Rus, A.. Duchenne muscular dystrophy. *Nat. Rev. Dis. Primers* **2021**, 7, 13.
56. Roberts, T.C.; Wood, M.J.A.; Davies, K.E.. Therapeutic approaches for Duchenne muscular dystrophy. *Nat. Rev. Drug Discov.* **2023**, 22, 917–934.
57. Maggio, I.; Chen, X.; Gonçalves, M.A.. The emerging role of viral vectors as vehicles for DMD gene editing. *Genome Med.* **2016**, 8, 59.
58. Biressi, S.; Filareto, A.; Rando,T.A.. Stem cell therapy for muscular dystrophies. *J. Clin.*

Invest. **2020**, 130, 5652–5664.

59. Mayor, H.D.; Torikai, K.; Melnick, J.L.; Mandel, M.. Plus and minus single-stranded DNA separately encapsidated in adeno-associated satellite virions. *Science* **1969**, 166, 1280–1282.
60. Berns, K.I.; Adler, S.. Separation of two types of adeno-associated virus particles containing complementary polynucleotide chains. *J. Virol.* **1972**, 9, 394–396.
61. 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.; *et al.* High levels of AAV vector integration into CRISPR-induced DNA breaks. *Nat. Commun.* **2019**, 10, 4439.
62. Nelson, C.E.; Wu, Y.; Gemberling, M.P.; Oliver, M.L.; Waller, M.A.; Bohning, J.D.; Robinson-Hamm, J.N.; Bulaklak, K.; Castellanos Rivera, R.M.; Collier, J.H.; *et al.* Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. *Nat. Med.* **2019**, 25, 427–432.
63. Ferrari, S.; Jacob, A.; Cesana, D.; Laugel, M.; Beretta, S.; Varesi, A.; Unali, G.; Conti, A.; Canarutto, D.; Albano, L.; *et al.* Choice of template delivery mitigates the genotoxic risk and adverse impact of editing in human hematopoietic stem cells. *Cell Stem Cell.* **2022,** 29, 1428–1444.
64. Wada, M.; Uchida, N.; Posadas-Herrera, G.; Hayashita-Kinoh, H.; Tsunekawa, Y.; Hirai, Y.; Okada, T.. Large-scale purification of functional AAV particles packaging the full genome using short-term ultracentrifugation with a zonal rotor. *Gene Ther.* **2023**, 30, 641–648.
65. McColl-Carboni, A.; Dollive, S.; Laughlin, S.; Lushi, R.; MacArthur, M.; Zhou, S.; Gagnon, J.; Smith, C.A.; Burnham, B.; Horton, R.; *et al.* Analytical characterization of full, intermediate, and empty AAV capsids. *Gene Ther.* **2024**, 31, 285–294.
66. Flotte, T.R.; Solow, R.; Owens, R.A.; Afione, S.; Zeitlin, P.L.; Carter, B.J.. Gene expression from adeno-associated virus vectors in airway epithelial cells. *Am. J. Respir. Cell. Mol. Biol.* **1992**, 7, 349–356.
67. Haberman, R.P.; McCown, T.J.; Samulski, R.J.. Novel transcriptional regulatory signals in the adeno-associated virus terminal repeat A/D junction element. *J. Virol.* **2000**, 74, 8732–8739.
68. Suchy, F.P.; Karigane, D.; Nakauchi, Y.; Higuchi, M.; Zhang, J.; Pekrun, K.; Hsu, I.; Fan, A.C.; Nishimura, T.; Charlesworth, C.T.; *et al.* Genome engineering with Cas9 and AAV repair templates generates frequent concatemeric insertions of viral vectors. *Nat Biotechnol.* **2025**, 43(2), 204-213.
69. Duan, D.; Sharma, P.; Yang, J.; Yue, Y.; Dudus, L.; Zhang, Y.; Fisher, K.J.; Engelhardt, J.F.. Circular intermediates of recombinant adeno-associated virus have defined

structural characteristics responsible for long-term episomal persistence in muscle tissue. *J. Virol.* **1998**, 72, 8568–8577.

70. Yang, J.; Zhou, W.; Zhang, Y.; Zidon, T.; Ritchie, T.; Engelhardt, J.F.. Concatamerization of adeno-associated virus circular genomes occurs through intermolecular recombination. *J. Virol.* **1999**, 73, 9468–9477.

Chapter 5

Selector AAV-CRISPR vectors purge off-target chromosomal insertions and promote precise genome editing

Zhen Li[1,*], Xiaoling Wang[1,*], Jin Liu[1], Josephine M. Janssen[1],
Rob C. Hoeben[1] and Manuel A.F.V. Gonçalves[1]

In preparation

* These authors contribute equally to this work

[1]Leiden University Medical Centre, Department of Cell and Chemical Biology, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands.

Abstract

Adeno-associated viral (AAV) vectors are commonly used for genome editing owing to the proclivity with which their single-stranded genomes serve as homologous recombination (HR) substrates during programmable nuclease-assisted gene targeting. However, the high recombinogenic nature of recombinant AAV genomes also facilitates their non-homologous end joining at off-target chromosomal breaks ("capture") created by said nucleases, mutagens, or DNA metabolic processes. The collateral build-up of off-target and random insertions occurs in an AAV dose-dependent manner and greatly diminishes the overall genome-editing accuracy. Moreover, AAV donor constructs can equally yield imprecise on-target edits resulting from non-homologous recombination pathways. Here, we demonstrate that endowing AAV donors with marker-free selectable sequences permits enriching for cells precisely co-edited at target and endogenous *ATP1A1* alleles. These selector AAV donors instal *ATP1A1* polymorphisms conferring resistance to the small-molecule ouabain and, in the process, yield high frequencies of on-target and precisely edited cell populations independently of the initially applied vector dose (up to 99.4%). Crucially, we further report that next to marker-free enrichment for precisely edited cell populations, selector AAV donors achieve a thorough removal of cells with off-target DNA insertions heightening, therefore, the ultimate precision of AAV-based genome editing.

Introduction

Genome editing technologies are emerging at a fast pace with their application in scientific and biotechnological realms continuing to expand (Pacesa *et al.* 2024). Insertion of exogenous (donor) DNA at predefined chromosomal positions (gene targeting or knock-in) subjected to double-strand DNA breaks (DSBs) made by clustered regularly interspaced short palindromic repeat (CRISPR)-derived nucleases, forms a set of commonly used and highly versatile genome editing principles. This results from the amenability of these gene targeting approaches to large genomic edits (*e.g.*, whole transgene knock-ins) and the straightforward programmability of RNA-guided CRISPR nucleases, such as those built on the prototypic CRISPR-Cas9 adaptive immune system from *Streptococcus pyogenes* (Pacesa et a. 2024). Engineered CRISPR-Cas9 nucleases, consisting of a sequence-tailored guide RNA (gRNA) and a Cas9 enzyme induce DSBs at DNA sequences that, next to a protospacer adjacent motif (*i.e*., NGG), have a *circa* 20-bp nucleotide tract (protospacer) complementary to the 5' end of the gRNA (spacer). Subsequent DSB repair by donor DNA substrates tailored for homology-directed repair mechanisms, *e.g*., homologous recombination (HR) (Liao *et al*. 2024) or for alternative DNA end-joining processes (He *et al*. 2016; Suzuki *et al*. 2016), results in targeted genome editing. Critically, when compared to donor constructs tailored for DNA end-joining processes, HR donors yield directional and more accurate gene knock-ins by mitigating insertions and mutations at, respectively, off-target positions and endogenous-exogenous DNA junctions (Liao *et al*. 2024; He *et al*. 2016). Unfortunately, HR-mediated genome editing is an inefficient process that often requires auxiliary measurers, *e.g*., addition of inhibitors of competing and dominant error-prone DNA repair pathways, like non-homologous end joining (NHEJ) and microhomology end joining (MMEJ) or, more typically, incorporation of selectable marker expression units in donor constructs to enrich for gene-edited cell fractions (Chu *et al*. 2015; Wimberger *et al*. 2023; Schimmel *et al*. 2023). However, interfering with endogenous DNA repair processes raises genomic instability concerns (Bischoff *et al*. 2020); whilst chromosomal insertion of heterologous marker genes limits the applicability and increases the complexity of gene editing protocols (Mikkelsen and Bak, 2023). Interestingly, the co-transfection of donor HR plasmids with selectable markers and genes-of-interest targeting independent loci permits drug-dependent enrichment for cells edited simultaneously at both loci, indicating the preferential isolation of HR-proficient cells amongst heterogeneous cell populations (Shy *et al*. 2016; Mitzelfelt *et al*. 2017). Building on this phenomenon, marker-free co-selection strategies have been devised where a donor plasmid is co-transfected with a secondary donor construct or oligonucleotide designed for creating drug- or toxin-selectable dominant alleles with specific polymorphism(s) (Agudelo *et al.* 2017; Wiebking *et al.* 2020; Li *et al.* 2021). A marker-free co-selection strategy based on the acquisition of gain-of-function resistance to a highly potent and specific inhibitor of the sodium/potassium (Na^+/K^+) ATPase pump, namely the plant-derived cardiotonic steroid ouabain, constitutes a particularly powerful co-selection approach (Agudelo *et al.* 2017). The main attributes of this approach are two-fold. Firstly, it targets the essential and ubiquitously expressed *ATP1A1* gene yielding, as a consequence, a robust and universal selectable phenotype; and, secondly, it is based on a commercially available and cheap small-molecule that, over decades, has been administered for congestive heart failure (Wu *et al.* 2015). Moreover, distinct *ATP1A1* polymorphisms confer cellular resistance to a broad range of ouabain concentrations which can be exploited for reiterative implementation of distinct genomic edits within individual cells (Levesque *et al.* 2022).

Plentiful physical and chemical transfection methods allow for introducing genome editing reagents into human cells including donor DNA substrates in the form of plasmids or synthetic oligonucleotides. However, achieving optimal transfection efficiencies without the build-up of cytotoxic effects is demanding as it often requires systematic cell type-specific protocol optimizations. Moreover, the ultimate performance of these optimized protocols, whose reagents are sometimes unknown due to proprietary reasons, typically depends on subtle experimental conditions, *e.g.*, cell-cycle stage distributions during transfection. In contrast to transfections, viral vector transductions present higher reproducibility and can be directly applied to different cell types independently of their cell cycle statuses. Valuable viral vector delivery properties stem from the fine-tuned mechanisms evolved by their wild-type counterparts in delivering nuclei acids into the cytoplasm or nucleus of the host cell. In this regard, commonly used adeno-associated viral (AAV) vectors with regular, pseudotyped or engineered capsids are particularly effective sources of donor DNA in a broad range of mammalian cell types (Epstein and Schaffer, 2017). Moreover, AAV vector genomes consisting of single-stranded DNA flanked by hairpin-forming inverted terminal repeats (ITRs), are prone to HR when harbouring sequences identical to those framing programmable nuclease target sites. Indeed, AAV HR donors yield high-efficiency gene targeting including in human cells with potential and established therapeutic relevance (Epstein and Schaffer, 2017). Unfortunately, their recombinogenic nature also contributes to off-target and imprecise on-target chromosomal donor DNA insertions involving non-homologous recombination processes (Miller *et al*. 2004; Hanlon *et al*. 2019; Ferrari *et al*. 2022; Li *et al*. 2024). Other insidious byproducts include on-target and off-target chromosomal insertion of, respectively, concatemeric structures and heterogeneous AAV DNA species that often, harbour ITR sequences (Hanlon et a. 2019; Ferrari *et al*. 2022; Li *et al*. 2024; Suchy *et al*. 2024). The latter events raise transcriptome deregulation and insertional oncogenesis risks due to the known transcriptional competency of ITR elements (Ferrari *et al*. 2022; Flotte *et al*. 1992; Haberman *et al*. 2000; Bazick *et al*. 2024). Equally of concern, AAV genomes, possibly due to mimicking DNA lesions or repair intermediates, can impair cell viability through P53-dependent DNA damage response (DDR) activation whose consequences are particularly deleterious during stem-cell genomic engineering (Schiroli *et al*. 2019; Allen *et al*. 2022). Critically, imprecise and off-target byproducts as well as cytotoxic effects are strictly proportional to AAV vector amounts (Schiroli *et al*. 2019; Allen *et al*. 2022).

Besides the efficiency, additional key parameters of genome editing procedures include their specificity and accuracy or fidelity (Maggio and Gonçalves, 2015). The former corresponds to the relative levels of on-target to off-target donor DNA insertions; the latter relates to the proportions between precise and imprecise on-target editing events.

The performance of marker-free co-selection systems in the context of viral vector delivery is presently unknown. Moreover, their utility for purging genome-edited cell populations from off-target as well as imprecise on-target chromosomal insertions is also underexplored. To fill these knowledge gaps, in this study, we set out to investigate AAV donor constructs harbouring marker-free co-selection components (selector AAV vectors) allowing for ouabain-dependent enrichment for genome-edited cells (Agudelo *et al*. 2017). We demonstrate that combining selector AAV vectors with ouabain treatments, next to enriching for genome-edited cell populations, achieves concomitant elimination of imprecise on-target edits and off-target and/or random donor DNA insertions from said populations. Interestingly, selector AAV vector titration experiments revealed that the highest fold-enrichment factors of genome-edited cell fractions are associated with the lowest vector input amounts which are expected to be beneficial for alleviating AAV vector production costs, off-target donor insertions and P53-dependent DDR activation.

Results

To start investigating AAV-based genome editing involving marker-free ouabain co-selection (**Figure 1A**), we assembled the selector vector AAV-HR$^{S1.A1}$. This vector contains HR donor templates and matched gRNA units designed for CRISPR-Cas9-induced transgene insertion at the human *AAVS1* safer harbour locus (19q13.4-qter); and generation of *ATP1A1* alleles with the Q118R and N129D (RD) polymorphisms conferring resistance to ouabain (**Figure 1B**). The *AAVS1*- and *ATP1A1*-specific gRNAs are complementary to intronic sequences to mitigate NHEJ-mediated mutagenesis of target alleles and both HR templates and cognate gRNA units are packaged in single AAV particles to guarantee their co-delivery into individual cells. Amongst the increasing range of genome editing strategies, gene knock-in into genomic safe harbour loci remains a particularly flexible approach as it permits to, for instance, correct in a predictable and safe manner the phenotype(s) of recessive disorders regardless of their causative mutations (Pavani and

Amendola, 2021). The predictability and safety attributes result from a mitigation of insertional mutagenesis, transgene silencing and/or variegated expression normally associated with genome engineering systems yielding semi-random and random integration profiles, *e.g.*, retroviral vectors and transposons, respectively.

Mock-transfected HeLa cells and HeLa cells transfected with a plasmid expressing Cas9 were transduced with AAV-HR$^{S1.A1}$ and, after sub-culturing in the presence or absence of ouabain, each population was subjected to clonal analysis for the characterization of genome editing events at *AAVS1* and *ATP1A1*. The former characterization involved junction PCR analysis; the latter entailed restriction fragment polymorphism (RFLP) assays (**Figure 1C**). EGFP-directed flow cytometry at 20 days post-transduction revealed that incubating in ouabain cells initially exposed to Cas9 led to a 2.8-fold increase in the frequency of stably transduced cells, hence, genetically modified cells (**Figure 1D**). Of notice, amongst the HeLa cell cultures not exposed to Cas9, those untreated with ouabain contained a measurable amount of stably transduced cells (2.3%), whilst those treated with ouabain, as expected, died (**Figure 1D**). Albeit at low frequencies, there are precedents for programmable nuclease-free gene targeting using AAV HR donors (Spector *et al*. 2021; Biijani *et al*. 2022) as well as for a role of inverted repeats, including the AAV ITR, in HR stimulation (Holkers *et al*. 2012). Yet, the exclusive AAV-HR$^{S1.A1}$ delivery setup suggests, nonetheless, that the vast majority of stably transduced cells contained random or off-target donor DNA insertions, hence wild-type *ATP1A1* alleles, in that they were readily eliminated by ouabain (**Figure 1D**). Importantly, combining AAV-HR$^{S1.A1}$ with Cas9 delivery resulted in an ouabain-dependent 2.8-fold increase in the frequency of stably transduced cells, suggesting selection for *AAVS1*-targeted cells (**Figure 1D**). Independent experiments involving HeLa cells initially transfected with plasmids expressing Cas9 or an inactive dCas9 protein and, subsequently, equally transduced with AAV-HR$^{S1.A1}$, yielded similar results, *i.e.*, ouabain-dependent elimination and enrichment of stably transduced cells generated by AAV-HR$^{S1.A1}$ delivery alone and together with Cas9, respectively (**Supplementary Figure S1**).

As aforementioned, next to the efficiency, the specificity and accuracy of donor DNA integration are key parameters of genome editing procedures based on HR or otherwise. The specificity is defined by the presence of donor DNA at the target site whilst the accuracy results from generating seamless telomeric-sided and centromeric-sided junctions between exogenous and endogenous DNA (jT and jC, respectively). Therefore, to determine the specificity and accuracy of selector AAV-based genome editing in the presence and absence of ouabain, junction PCR screens were performed on isolated EGFP-positive cell clones, each of which representing individual genome-modifying events. Of notice, differently from splice acceptor gene-trapping and protein-tagging constructs, clonal isolation based on transcriptionally autonomous transgenes, such as that in AAV-HR$^{S1.A1}$, prevents the biased selection for on-target events. The screening of 160 randomly isolated clones (80 expanded with ouabain and 80 without), identified genome-modifying events corresponding to random or off-target donor DNA insertions (jT-/jC-), precise HR-mediated *AAVS1* gene targeting (jT+/jC+), partial gene targeting (jT+/jC- or jT-/jC+), and non-homologous recombination, namely, containing differently-sized amplicon(s) diagnostic for imprecise DNA end-joining events (**Figure 1E**, left panel, red, cyan, orange and yellow arrowheads, respectively; and **Supplementary Figure S2**). Moreover, ouabain selection led to a remarkable increase in the number of *ATP1A1* gene editing as traced through RFLP assays (**Figure 1E**, right panel and **Supplementary Figure S2**).

Significantly, when compared with untreated cultures, cultures treated with ouabain contained substantially lower amounts of imprecisely edited cells (**Figure 1F**, top panel). Indeed, all the three categories representing imprecisely edited cells were smaller in ouabain-treated cultures (**Figure 1F**, top panel; and **Supplementary Figure S2**). The substantial enrichment for

AAVS1-targeted and precisely edited cells in the presence of ouabain was paralleled by a remarkable expansion of *ATP1A1*-edited cells as assessed by RFLP assays (**Figure 1F**, bottom panel; and **Supplementary Figure S2**). This data demonstrates that besides enriching for gene-targeted cell fractions in an ouabain-dependent manner, selector AAV-HR$^{S1.A1}$ construct deployment is valuable for purging said fractions from off-target and/or imprecise genome editing byproducts. These results further support the robust ouabain-resistance phenotype conferred by installing the RD polymorphisms at *ATP1A1*.

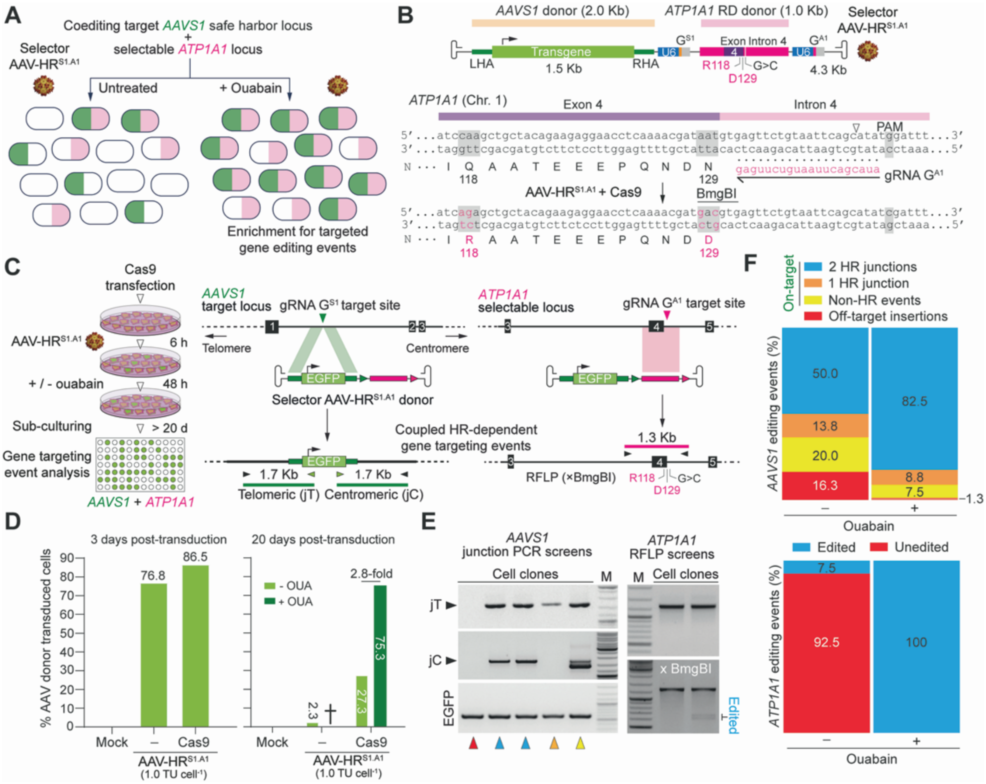


Figure 1. Testing and characterizing selector AAV genome editing at *AAVS1* and *ATP1A1*. (**A**) The marker-free co-selection principle. The cell population fraction whose intracellular milieu is conducive for HR (*e.g.*, cells undergoing the late G2/S phases of the cell cycle), are prone to simultaneous HR-mediated editing at two independent loci. As corollary, cells co-edited at a target sequence of interest and at a secondary selectable locus

can be enriched for when the latter locus acquires an edit(s) conferring a dominant resistance to a small-molecule drug. In this study, AAV-based genome editing of target and selectable *ATP1A1* alleles is assessed in the presence and absence of ouabain, a highly potent and specific inhibitor of the essential Na^+/K^+ ATPase pump. (**B**) Schematics of selector AAV donor construct and selectable *ATP1A1* site. The vector AAV-$HR^{S1.A1}$ contains *AAVS1* and *ATP1A1* donor templates and cognate matched gRNA units that, in the presence of Cas9, trigger HR-mediated chromosomal insertion of a transgene and polymorphisms at the former and latter loci, respectively. The *ATP1A1* polymorphisms Q118R and N129D (RD) confer resistance to ouabain. *ATP1A1* editing can be probed via restriction fragment length polymorphism (RFLP) assays using BmgBI in untreated and ouabain-treated cell populations. (**C**) Schematics of the experimental setup. Cas9-transfected HeLa cells are transduced with selector AAV-$HR^{S1.A1}$ donor and, after sub-culturing in the presence or absence of ouabain, are subjected to EGFP-directed flow cytometry and to clonal screens using junction PCR analysis and RFLP assays at *AAVS1* and *ATP1A1*, respectively. (**D**) Quantification of selector AAV donor delivery and DNA editing. HeLa cells subjected or not to Cas9 plasmid transfections were transduced with AAV-$H^{S1.A1}$ at 1 TU $cell^{-1}$. Donor delivery was assessed 3 days later by flow cytometry (left graph). After 20 days of sub-culturing in the presence and absence of ouabain, flow cytometry established AAV stable transduction frequencies (right graph). The cross indicates complete cell death in ouabain-treated cultures exposed exclusively to AAV-$HR^{S1.A1}$. (**E**) Characterization of genome editing outcomes. Left panel, representative clones yielding *AAVS1* amplicons diagnostic for gene targeting involving precise HR events at the telomeric and centromeric side of the target sequence are marked by cyan arrowheads (jT and jC, respectively). Representative clones lacking *AAVS1*-specific insertions (off-target), containing HR-independent targeted insertions and with only one HR-derived junction between transgenic and *AAVS1* sequences are marked by red, yellow, and orange arrowheads, respectively. Right panel, representative *ATP1A1* amplicons resistant and susceptible to BmgBI digestion diagnostic for unedited and edited *ATP1A1* alleles, respectively, are also depicted. (**F**) Cumulative characterization of genome editing outcomes. The frequencies of the different types of genome-modifying events detected in cell clones randomly isolated from HeLa cell populations stably transduced with AAV donor DNA and expanded in the presence or absence of ouabain are plotted (**Supplementary Figure S2**).

Recently, we introduced and characterized a dual viral vector genome-editing system based on the delivery of CRISPR-Cas9 nucleases and donor DNA templates via high-capacity adenovector particles (AdVPs) and AAV vectors, respectively (Li *et al.*, 2024). Earlier experiments from our laboratory and those of others have shown that, contrary to linear free-ended DNA, capped double-stranded DNA, including adenovector genomes, are refractory to end-joining processes underpinning off-target and random chromosomal DNA insertions (Holkers *et al.* 2014; Medert *et al.* 2023). Moreover, besides their vast packaging capacity (*i.e.*, up to 36 kb), viral gene-free AdVPs display a remarkably lower cytotoxicity profile when compared to that of their viral gene-containing, earlier-generation, counterparts (Li *et al.* 2024; Brescia *et al.* 2020; Tasca *et al.* 2020; Ricobaraza *et al.* 2020). Hence, to expand and streamline the testing of selector AAV vectors in established cell lines as well as in difficult-to-transfect primary cells, we used this dual viral vector platform in HeLa cells and human mesenchymal stem cells (hMSCs). Initial co-transduction experiments in HeLa cells using AdVP.C9^{KARA}, a vector encoding the high-specificity nuclease SpCas9^{KARA} (**Figure 2A**) (Wang *et al.* 2021), and different amounts of AAV-HR$^{S1.A1}$, led to a clear dose-dependent increase in productive AAV transduction as assessed by EGFP-directed flow cytometry at 3 days post-transduction (**Figure 2B**). The significant AdVP.C9^{KARA}-dependent enhancement on productive AAV-HR$^{S1.A1}$ transduction (**Figure 2B**) is mostly caused by the higher transgene expression levels resulting from the buildup of chromosomally targeted templates over non-integrated episomes known to be prone to cellular restriction factors (Li *et al.*, 2024; Dever *et al.* 2016). Moreover, earlier experiments have also established a causal relationship between CRISPR-Cas9-induced DSBs and productive AAV transduction (Li *et al.* 2024). Most importantly, after a 20-day sub-culturing period in the presence and absence of ouabain, EGFP-directed flow cytometry revealed significantly higher frequencies of stably transduced cells in the presence of ouabain (**Figure 2C**). Of notice, the lowest and highest stably transduced cell fold-enrichment factors were associated with the highest and lowest AAV-HR$^{S1.A1}$ dosages, respectively (**Figure 2C**). Junction PCR and RFLP analyses (**Figure 2A**) of genomic DNA from stably transduced cell populations established ouabain-dependent co-selection of cells edited at *AAVS1* and *ATP1A1* (**Figure 2D**).

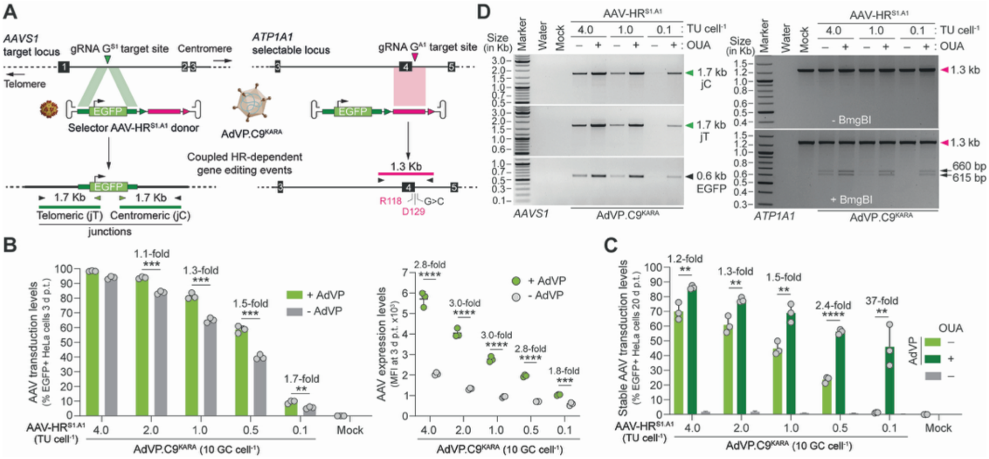


Figure 2. Selector AAV gene targeting at *AAVS1* and *ATP1A1* using combinatorial viral vector delivery. (**A**) Schematics of the experimental setup. Selector AAV-based gene targeting upon co-transduction of HeLa cells with vector AAV-HR$^{S1.A1}$, encoding donor and gRNA sequences, and vector AdVP.C9^{KARA}, encoding a Cas9 nuclease, was assessed via EGFP-directed flow cytometry and genotyping assays based on junction PCR and RFLP assays as depicted. (**B**) Quantification of selector AAV donor delivery. HeLa cells were transduced with AAV-HR$^{S1.A1}$ alone or together with AdVP.C9^{KARA} at the specified multiplicities of infection (MOI). Transduction levels were determined by flow cytometric quantification of EGFP-positive cell frequencies and respective mean fluorescence intensity (MFI) values at 3 days post-transduction (left and right graphs, respectively). Mock-transduced cells provided for negative controls. (**C**) Quantification of selector AAV-based DNA editing with and without ouabain. AAV stable transduction frequencies were measured via EGFP-directed flow cytometry after sub-culturing HeLa cells initially exposed to the indicated vector doses for 20 days. During sub-culturing, the cells were incubated or not with ouabain. Mock-transduced cells and cells transduced exclusively with AAV-HR$^{S1.A1}$ served as controls. The results are presented as mean ± SD of three biological replicates. Significant differences amongst the marked datasets were calculated by Student's t-tests; **$P<0.05$. (**D**) Genotyping of *AAVS1* and *ATP1A1* in co-

transduced cell populations. *AAVS1* and *ATP1A1* gene editing in HeLa cells co-transduced with the indicated vectors and incubated or not with ouabain was assessed through junction PCR and RFLP assays (left and right panels, respectively). Mock-transduced cells provided for negative controls.

To test genome editing at an independent target locus, the selector AAV-HR$^{LMN.A1}$ donor was applied to HeLa cells and to primary hMSCs together with AdVP.C9^{KARA} (**Figure 3** and **Supplementary Figure S3**, respectively). The AAV-HR$^{LMN.A1}$ vector contains the same selectable sequence of AAV-HR$^{S1.A1}$ together with a *LMNA*-specific gRNA unit and a matched HR template for tagging LMNA at its N-terminus with the live-cell reporter mScarlet-I (**Figure 3A**). Of note, *LMNA* mutations have been linked to, amongst others, Emery-Dreifuss muscular dystrophy, limb girdle muscular dystrophy, dilated cardiomyopathy, and Hutchinson-Gilford progeria syndrome. In contrast to the use of autonomous transgene expression units, gene tagging setups are contingent on precise gene knock-in to guarantee expression from endogenous *cis*-acting regulatory elements. As such, this setup directly traces and quantifies HR-mediated gene editing events. Co-transductions targeting *LMNA* alleles broadly recapitulated the results obtained through experiments targeting the *AAVS1* locus (**Figure 2**). In particular, reporter-directed flow cytometry (**Figure 3B** and **Supplementary Figure S3A**), together with junction PCR and RFLP assays (**Figure 3C**, left and right panel, respectively) demonstrated an ouabain-dependent co-selection for cells edited at target *LMNA* and *ATP1A1* alleles. Again, the lowest and highest gene editing fold-enrichment factors resulted from applying the highest and lowest selector AAV doses, respectively (**Figure 3B** and **Supplementary Figure S3A**). The highest fold-enrichment factor (30-fold) was, in fact, observed in hMSCs initially co-transduced with AdVP.C9^{KARA} and AAV-HR$^{LMN.A1}$ at 500 GC cell^{-1} (**Supplementary Figure S3A**). These results indicate the feasibility in generating high frequencies of genome-edited cell populations while using low amounts of AAV vectors known to trigger dose-dependent cytotoxic effects in cell types with therapeutic relevance, *e.g.*, stem cells. Further consistent with precise gene editing, direct fluorescence microscopy revealed that mScarlert::LMNA fusion products were present and properly located in cell nuclei exclusively in cultures exposed simultaneously to CRISPR-Cas9 and donor DNA reagents (**Figure 3D** and **Supplementary Figure S3B**). Finally, RFLP assays on unsorted cells and on mScarlet-positive and mScarlet-negative cells, further confirmed the strict ouabain-dependent selection of the *ATP1A1* RD variant. The strong positive selection for this endogenous marker gene is in fact particularly evident in cultures of mScarlet::LMNA-negative cells treated with ouabain (**Figure 3E**).

Taken together, the above-described experiments demonstrate that combining ouabain with tailored selector AAV vectors achieves a strong elimination of cells with imprecise on-target edits and off-target exogenous DNA insertions from CRISPR-edited cell populations. However, although the co-delivery of selectable *ATP1A1* and target donor sequences in single AAV vectors guarantees the presence of both donor templates in individual cells, at low vector doses, bipartite donor availability at primary and secondary loci is expected to become limiting. This consideration is supported by the decreasing ouabain-dependent genome editing frequencies at primary target loci as a function of diminishing AAV vector amounts, *i.e.*, at *AAVS1* safe harbour and *LMNA* loci (**Figure 2C** and **Figure 3B**).

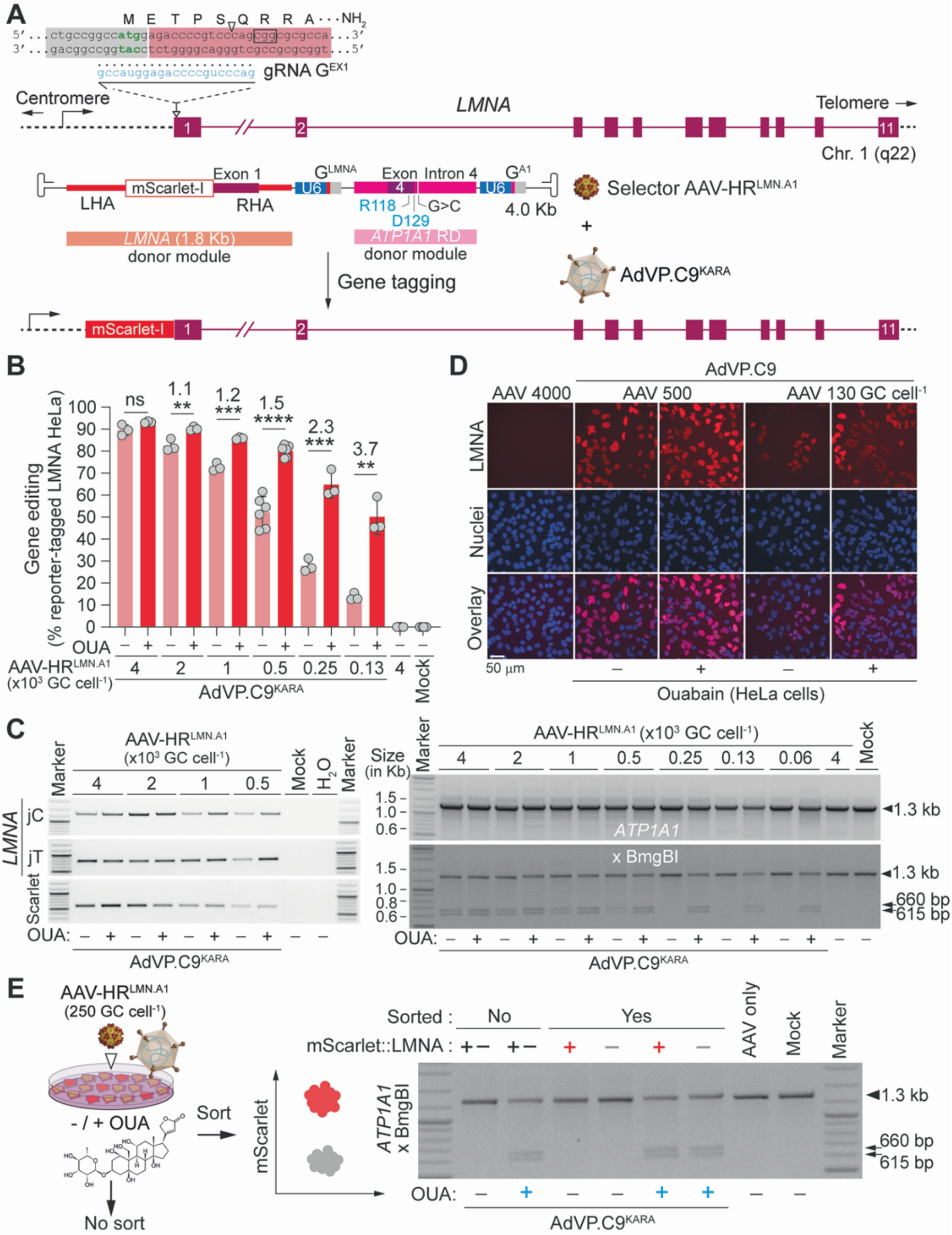


Figure 3. Selector AAV gene editing at *LMNA* and *ATP1A1* using combinatorial viral vector delivery. (**A**) Schematics of the experimental setup. Selector AAV-based gene tagging upon co-transduction of HeLa cells with AAV-HR$^{LMN.A1}$ donor and adenovector

AdVP.C9^{KARA} was monitored through mScarlet-directed flow cytometry and genotyping assays based on junction PCR and RFLP assays (not shown). (**B**) Quantification of selector AAV gene tagging with and without ouabain. Gene tagging was determined in HeLa cell populations initially exposed to the indicated vector doses via mScarlet-directed flow cytometry at 14 days post-transduction. Mock-transduced cells and cells transduced exclusively with AAV-HR$^{S1.A1}$ at 4,000 GC cell^{-1}, served as negative controls. The data are shown as mean ± SD of three biological replicates. Significant differences amongst the marked datasets were calculated by Student's t-tests; $^{**}P<0.05$. (**C**) Genotyping of *LMNA* and *ATP1A1* in co-transduced cell populations. *LMNA* and *ATP1A1* gene editing in HeLa cells co-transduced with the indicated vectors and incubated with or without ouabain was assessed through junction PCR and RFLP assays (left and right panels, respectively). Mock-transduced cells served negative controls. (**D**) Characterization of LMNA protein tagging. LMNA tagging and nuclear localization was monitored by combining direct fluorescence microscopy for mScarlet expression and nuclei labelling using the DNA dye Hoechst 33342. (**E**) Assessing ouabain-dependent selection of *ATP1A1* edited cells. HeLa cells co-transduced with AAV-HR$^{LMN.A1}$ and AdVP.C9^{KARA} and then cultured with or without ouabain were either sorted or not sorted to isolated LMNA-tagged positive and negative cells. *ATP1A1* editing in each cell population was probed through RFLP assays. Mock-transduced cells and cells exposed only to AAV-HR$^{LMN.A1}$ at 4,000 GC cel l^{-1}, served as negative controls.

Hence, we sought to investigate whether a more strict enrichment for genome-edited cells with sustained purging of random and/or off-target insertions is achievable by using selector AAV donors with transgenic DNA juxtaposed to a selectable polymorphism (**Figure 4A**). To this end, the vector AAV-HR$^{A1.IN17}$ was assembled (**Figure 4B**). This vector contains a gRNA unit for directing cleavage at intron 17 of *ATP1A1* and a matched donor DNA template designed for (i) *mScarlet-I* transgene knock-in at this intron; and (ii) installing the ouabain-selectable polymorphism T480N at the contiguous exon 17 (Agudelo *et al.* 2017). We started testing this selector AAV-HR$^{A1.IN17}$ vector with in-linkage transgene and T840N by transducing HeLa cells transfected with plasmids expressing Cas9 nuclease or Cas9^{D10A} nickase proteins. After sub-culturing in the presence or absence of ouabain, stably transduced cells were quantified by mScarlet-I-directed flow cytometry and subjected to clonal analysis for characterizing genome-modifying events at the single-cell level (**Figure 4C**). At 3 days post-transduction, AAV-HR$^{A1.IN17}$ donor delivery into virtually all HeLa cells was achieved (**Figure 4D**, left panel). At 20 days post-transduction, ouabain untreated cultures initially exposed to Cas9 and Cas9^{D10A} had *circa* 60% and 7.4% of genome-modified cells, respectively (**Figure 4D**, right panel). This difference is consistent with single-strand DNA breaks (SSBs), or nicks, being generally weaker HR stimuli than DSBs (Chen *et al*. 2017), including when using AAV HR donors (Pavani *et al*. 2021). Despite this, nickase-based genome editing offers notable advantages that include a striking reduction in off-target effects and on-target allelic mutagenesis as, in contrast to DSBs, SSBs are typically not engaged by error-prone end joining repair pathways (Chen *et al*. 2017). Importantly, addition of ouabain to cultures transduced with AAV-HR$^{A1.IN17}$ and exposed to Cas9 or Cas9^{D10A} led to a significant increase in the frequencies of genome-modified cells, *i.e*., 1.6- and 12-fold, respectively (**Figure 4D**, right panel). Moreover, as previously observed when using the bipartite donor AAV-HR$^{S1.A1}$ (**Figure 1D** and **Supplementary Figure S1**), HeLa cell cultures exposed exclusively to AAV-HR$^{A1.IN17}$ contained a low, yet clearly measurable, proportion of stably transduced cells (3.5%). The selective abolishment of this cell fraction in the presence of ouabain (**Figure 4D**, right panel), indicates that it results from the random chromosomal insertion of vector DNA. In line with this, data from the subsequent junction PCR screening of 160 arbitrarily isolated mScarlet-positive clones (80 expanded from cultures exposed to Cas9 and 80 expanded from cultures exposed to Cas9^{D10A}), was consistent with the absence of randomly inserted donor DNA in cells grown in the presence of ouabain (**Figure 4E** and **Supplementary Figure S4**). Equally reminiscent of the results obtained with the bipartite donor AAV-HR$^{S1.A1}$ (**Figure 1F** and **Supplementary Figure S2**), next to off-target DNA insertion purging, combining ouabain incubation with in-linkage donor AAV-HR$^{A1.IN17}$ delivery also yielded a substantial reduction of imprecise non-homologous recombination (**Figure 4E** and **Supplementary Figure S4**).

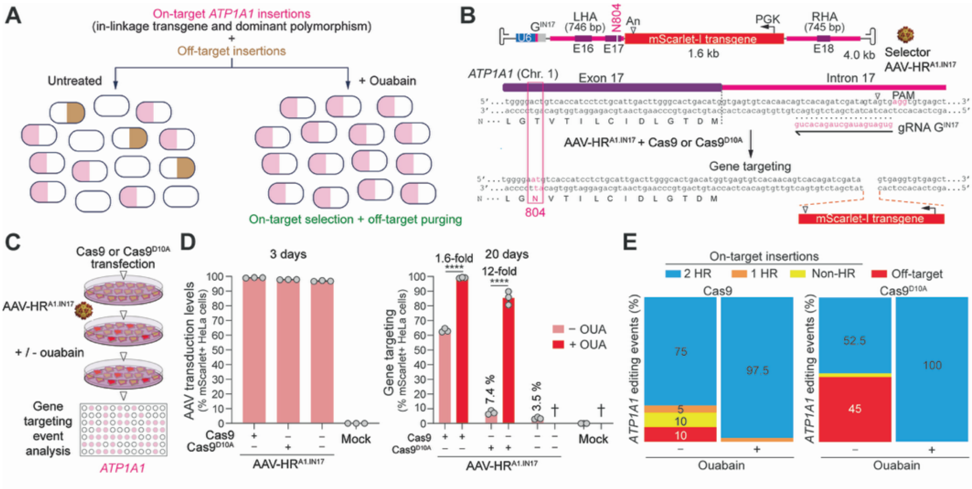


Figure 4. Testing and characterizing selector AAV genome editing based on gene knock-in and marker-free selection linkage. (**A**) Gene targeting with off-target insertions purging via gene knock-in and marker-free selection linkage. HR donors designed for concomitant knock-in of transgenes and dominant ouabain-resistance polymorphisms are postulated to achieve a strict enrichment for gene targeted cells with the simultaneous thorough eradication of random and/or off-target donor DNA insertions. (**B**) Selector AAV donor with in-linkage transgene and dominant polymorphism. The selector AAV-HR$^{A1.IN17}$ vector harbours a gRNA unit specific for intron 17 of *ATP1A1* and a matched HR donor whose region homologous to *ATP1A1* encodes the ouabain-resistance polymorphism T804N. LHA and RHA, "left" and "right" homology arms, respectively. (**C**) Schematics of the experimental setup. HeLa cells transfected with Cas9 nuclease or Cas9^{D10A} nickase constructs are transduced with AAV-HR$^{A1.IN17}$ and cultured in the presence or absence of ouabain. The frequencies of genome-modified cells and the characterization of genome editing events are subsequently assessed trough mScarlet-directed flow cytometry and clonal screens using junction PCR analysis at *ATP1A1*, respectively. (**D**) Quantification of selector AAV donor delivery and DNA editing. HeLa cells subjected to Cas9 and Cas9^{D10A} plasmid transfections were transduced with AAV-HR$^{A1.IN17}$ at 1 TU cell^{-1}. Donor delivery was assessed 3 days later by flow cytometry (left graph). After 20 days of sub-culturing in the presence and absence of ouabain, flow cytometry determined gene targeting frequencies (right graph). The crosses denote full cell death in ouabain-treated cultures exposed exclusively to AAV-HR$^{A1.IN17}$. (**E**) Cumulative

characterization of genome editing outcomes. The frequencies of the different types of genome-modifying events detected in cell clones randomly isolated from HeLa cell populations stably transduced with AAV-HR$^{A1.IN17}$ donor DNA and expanded in the presence and absence of ouabain are plotted. Gene targeting events derived from precise HR at the telomeric and centromeric side of the target sequence are marked in cyan. Gene targeting events involving partial HR or no HR are labelled in orange and yellow, respectively. Off-target donor DNA insertion events are coloured in red (**Supplementary Figure S4**).

To test the performance of the selector AAV in-linkage design in terms of its robustness for ouabain-dependent selection of gene targeted cells (**Figure 5A**), hMSCs and HeLa cells were co-transduced with AdVP.C9^{KARA} and AAV-HR$^{A1.IN17}$ (**Figure 5B** and **5C**, respectively). As aforementioned, the observed AdVP.C9^{KARA}-dependent enhancement on productive AAV transduction (**Figure 5B** and **5C**) is primarily caused by the higher transgene expression resulting from the accumulation of chromosomally targeted exogenous DNA, known to be more refractory to cellular restriction factors than non-integrated episomal DNA (Dever *et al*. 2016; Li *et al*., 2024). Critically, after a sub-culturing period in the presence and absence of ouabain, mScarlet-directed flow cytometry disclosed remarkably strong ouabain-dependent positive selection that consistently yielded over 90% of gene targeting frequencies independently of selector AAV vector doses and transduced cell type (**Figure 5D** and **5E**).

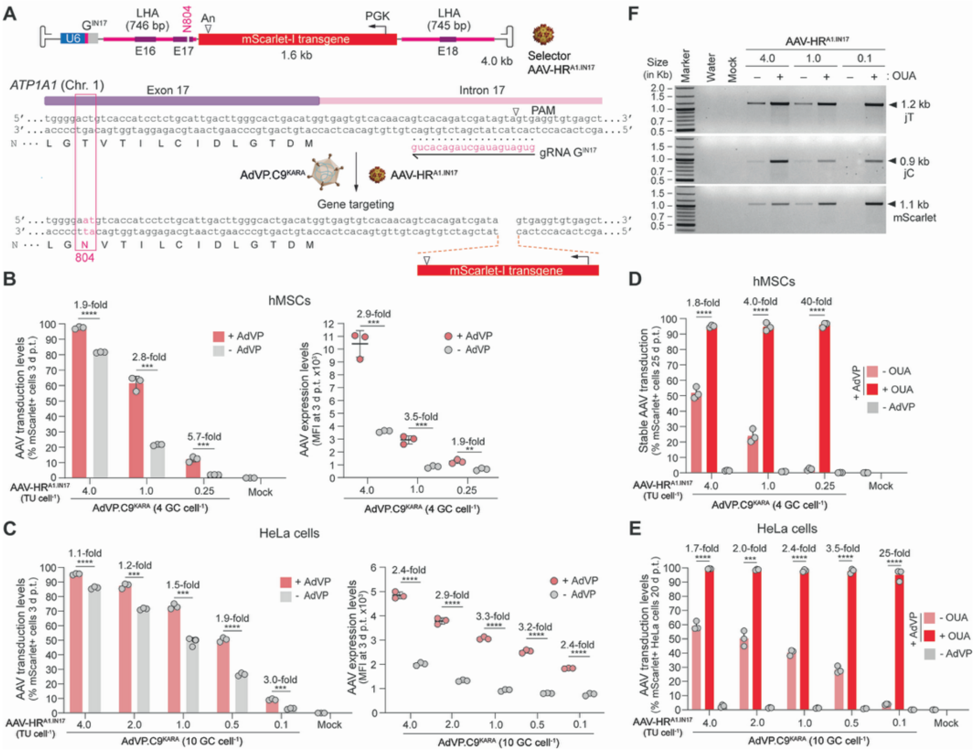


Figure 5. Selector in-linkage AAV gene targeting using combinatorial viral vector delivery. (**A**) Diagram of the experimental setup. The selector AAV-HR$^{\text{A1.IN17}}$ vector contains a gRNA unit specific for intron 17 of *ATP1A1* and a matched HR donor with the ouabain-resistance polymorphism T804N in its *ATP1A1* homologous region. LHA and RHA, "left" and "right" homology arms, respectively. (**B** and **C**) Quantification of selector AAV donor delivery. hMSCs and HeLa cells (panel B and C, respectively) were transduced with AAV-HR$^{\text{A1.IN17}}$ alone or together with AdVP.C9$^{\text{KARA}}$ at the specified MOI. Transduction levels were determined by flow cytometric quantification of mScarlet-positive cell frequencies and respective MFI values at 3 days post-transduction (left and right graphs, respectively). Mock-transduced cells provided for negative controls. (**D** and **E**) Quantification of selector in-linkage AAV gene targeting with and without ouabain. Gene targeting frequencies in hMSC and HeLa cell cultures (panel D and E, respectively) were measured via mScarlet-directed flow cytometry after sub-culturing these cultures initially exposed to the indicated vector doses in the presence and absence of ouabain. Mock-transduced cells and cells transduced exclusively with AAV-HR$^{\text{S1.A1}}$ served as controls. The results are presented as mean ± SD of three biological replicates. Significant differences amongst the marked datasets were calculated by Student's t-tests; **P<0.05.

(**F**) Genotyping of *ATP1A1* in co-transduced cell populations. *ATP1A1* gene targeting in HeLa cells co-transduced with the indicated vectors and incubated or not with ouabain was assessed through junction PCR. Mock-transduced cells provided for negative controls.

Similarly to the bipartite donor AAV vectors co-targeting *ATP1A1* and *AAVS1* or *LMNA*, the lowest and highest stably transduced cell fold-enrichment factors were associated with the highest and lowest AAV-HR$^{A1.IN17}$ dosages, respectively (**Figure 5D** and **5E**). However, clearly, the in-linkage donor vector AAV-HR$^{A1.IN17}$ achieved a more thorough and homogeneous positive selection of gene targeted cells than that obtained with the bipartite donor vectors AAV-HR$^{S1.A1}$ and AAV-HR$^{LMN.A1}$. In fact, the highest fold-enrichment for gene targeted cells was obtained in hMSC cultures co-transduced with AdVP.C9^{KARA} and AAV-HR$^{A1.IN17}$ at 0.25 TU cell^{-1} where ouabain selection resulted in a 40-fold increase in the frequency of gene knock-ins (*i.e*., 96±1.6% and 2.4±0.7% with and without selection, respectively) (**Figure 5D**). The purity levels for gene targeted cells in hMSC and HeLa cell cultures attained upon AdVP.C9^{KARA} and AAV-HR$^{A1.IN17}$ co-transductions and ouabain selection were consistently high, varying from a minimum of 94.6% to a maximum of 99.4% regardless of donor vector input amounts (**Figure 5D** and **5E**). Finally, junction PCR analysis of genomic DNA from stably transduced cell populations confirmed strict ouabain-dependent enrichment for *ATP1A1*-targeted cells (**Figure 5F**).

Taken together, these data establish the selector AAV in-linkage design as a robust strategy for achieving a strict selection for gene targeted cells through precise HR and, simultaneously, the purging of random and/or off-target donor DNA insertions from engineered cell populations.

Discussion

AAV vectors are commonly used in genome editing protocols as sources of donor HR substrates. However, the recombinogenic character of AAV vector genomes, that bear these substrates, fosters their “capture” at on-target and off-target or random chromosomal breaks through non-homologous end joining processes (Miller *et al*. 2004; Hanlon *et al*. 2019; Ferrari *et al*. 2022; Li *et al*. 2024). Moreover, recent studies have demonstrated the pervasiveness of additional genomic DNA byproducts consisting of heterogeneous AAV vector-derived fragments and concatemeric species at off-target and on-target sites (Hanlon *et al*. 2019; Ferrari *et al*. 2022; Suchy *et al*. 2024; Li *et al*. 2024). The former AAV fragment intermediates, known to be packaged in vector particles (McColl-Carboni *et al*. 2024), typically contain transcriptionally-competent ITR elements (Flotte *et al*. 1992; Haberman *et al*. 2000), raising transcriptome deregulation and insertional oncogenesis concerns (Ferrari *et al*. 2022; Bazick *et al*. 2024). Finally, when combined with programmable nucleases, AAV vector genomes exacerbate P53 build-up and ensuing DDR activation that impairs cell viability in a strict vector dose-dependent manner (Schiroli *et al*. 2019; Allen *et al*. 2022).

We hypothesized that marker-free co-selection systems can be co-opted for addressing the aforementioned AAV-based genome editing shortcomings. These systems require neither chromosomal integration of exogenous selectable markers nor cell isolation reagents and equipment (*e.g*., FACS and MACS) (Mikkelsen and Bak, 2023). For this study, we selected a strategy based on ouabain, a highly potent and specific inhibitor of the ubiquitously expressed Na^+/K^+ ATPase pump (Agudelo *et al*. 2017). In contrast to other marker-free co-selection systems, such as that based on the potent inhibitor of mammalian protein synthesis diphtheria toxin protein, ouabain-dependent systems require a cheap small-molecule drug that has been used for congestive heart failure (Wu *et al*. 2015). Moreover, distinct *ATP1A1* polymorphisms confer resistance to a broad range of ouabain concentrations that can be exploited for sequential installation of distinct genomic edits (Levesque *et al*. 2022), including those underpinning regulatory systems, complex gene circuits and other synthetic biology devices. For instance, the herein tested polymorphisms T804N, located in the third extracellular loop, and Q118R/N129D located in the first extracellular loop of Na^+/K^+ ATPase, create variants resistant to ouabain inhibition at 10 μM and over 1000 μM, respectively, in K562 cells (Levesque *et al*. 2022). Hence, after assembling AAV donor constructs endowed with matched gRNA units and ouabain-selectable sequences, we demonstrate that these selector AAV vectors, in addition to enriching for gene-targeted cell populations, achieve concomitant removal of imprecise HR-independent edits and off-target and/or random AAV donor DNA insertions. Interestingly, selector AAV vector titration experiments revealed that the highest fold-enrichment factors of gene-targeted cell fractions are associated with the lowest vector input amounts which are expected to be beneficial for alleviating both AAV production costs and detrimental P53-dependent DDR activation.

In this study, two types of selector AAV vector designs were investigated, namely, vector particles containing bipartite donor modules for *ATP1A1* and target gene co-editing, and vector particles bearing an in-linkage donor DNA module for direct *ATP1A1* targeting and selection. Selector AAV vectors with bipartite donors can be customized to implement distinct types of genomic edits (*e.g*., gene knock-ins, gene-tagging, or gene-repairing) at different loci, like the herein targeted *AAVS1* safe harbour locus and *LMNA* alleles. The former locus is a commonly used genomic landing pad for achieving homogenous and stable transgene expression (Lombardo *et al*. 2011; Pavani *et al*. 2021); the latter encodes lamin, a product found in the nuclear lamina matrix of proteins located underneath the inner nuclear membrane and whose mutations underpin, for instance, Emery-Dreifuss muscular dystrophy, limb girdle muscular dystrophy, dilated cardiomyopathy, Charcot-Marie-Tooth disease, and Hutchinson-Gilford progeria syndrome. However, although the co-delivery of selectable *ATP1A1* and donor sequences in single AAV vectors assures their presence in individual cells, at low vector doses, bipartite donor availability at primary and secondary loci should become limiting. This point is consistent with the observed gradual decrease in ouabain-dependent gene editing levels at primary target sequences as a function of diminishing AAV vector amounts.

Selector AAV vectors with in-linkage selecting and targeting donor templates are, on the other hand, restricted to creating gene knock-ins at *ATP1A1* alleles. Yet, *ATP1A1* can in principle serve as a suitable transcription-favourable genomic landing pad to, for instance, overexpress proteins in producer cells, control cell behaviour with synthetic gene circuits, or complement genetic defects in autologous patient-derived cells. Moreover, albeit less versatile than bipartite donors, selector AAV vectors with in-linkage selecting and targeting HR templates yield engineered cell populations with substantially higher degrees of purity for gene targeted cells independently of vector doses and transduced cell types (range: 94.6%-99.4%). Indeed, the ability to generate such high frequencies of gene targeted cells using low AAV vector doses should allow creating genome engineered cells with minimal risks of harboring off-target and/or random chromosomal insertion of exogenous DNA (intact or otherwise). These favorable selector AAV performance features might permit streamlining cell engineering efforts via bypassing the need for time-consuming cell line isolation and screening and, in addition, expand said efforts to cell types refractory to single-cell isolation and expansion, such as most primary cells whose proliferation is restricted by their Hayflick limit and ensuing senescence. Equally of note, selector AAV vectors with in-linkage donor designs require single instead of dual CRISPR-Cas9 nucleases, therefore reducing genomic instability risks. There is nonetheless a growing realization that especially in DNA damage sensitive cells, like stem cells, programmable nuclease-induced DSBs can be detrimental to target locus stability (Frock *et al.* 2015; Kosicki *et al.* 2018) and cell viability (Chen *et al.* 2017; Ihry *et al.* 2018; Schiroli *et al.* 2019). Significantly, research from our laboratory has demonstrated that, when compared with Cas9 nucleases, Cas9^{D10A} nickases are substantially less disruptive to on-target and off-target sequences (Wang *et al.* 2021; Chen *et al.* 2020) and present greatly dampened P53-dependent DDR activation levels (Wang *et al.* 2023). Hence, the herein provided proof-of-principle that selector AAV-based genome editing is transportable to protocols involving

Cas9^{D10A}-induced HR is relevant for further refining and applying marker-free co-selection approaches.

In conclusion, in the present study, we integrate AAV HR donor delivery with a marker-free co-selection principle based on a commercially available and cheap small molecule, ouabain. We demonstrate that combining these selector AAV vectors with ouabain treatments, in addition to selecting for precisely edited cell populations, eradicates otherwise prevalent off-target and/or random AAV donor DNA insertions. Moreover, through selector AAV vector titration experiments, we report that the highest fold-enrichment levels for genome-edited cells is associated with the lowest vector inputs which is expected to be beneficial for mitigating AAV vector production costs and P53-dependent DDR activation processes known to limit AAV-based genome editing in DNA damage-sensitive cells. Selector AAV vectors are, therefore, expected to become useful in a broad array of basic and applied research contexts, such as for expediting cell engineering endeavours and generating well-defined populations of hard-to-transfect cell types including those that are refractory to clonal expansion but that hold nonetheless therapeutic relevance.

Materials and Methods

Cells

The human cervix carcinoma HeLa cells (ATCC) were cultured in high-glucose Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific; Cat. No.: 41966-029) containing 5% fetal bovine serum (FBS; Biowest; Cat. No.: S1810-500). The HeLa cells were kept at 37°C in a humidified-air 10% CO_2 atmosphere. The primary human mesenchymal stem cells (hMSCs) were isolated from bone marrow and cultured in Minimum Essential Medium α (MEM-α; Thermo Fisher Scientific; Cat. No.: 22561-021) supplemented with 10% FBS, 100 U ml^{-1} penicillin/streptomycin (Thermo Fisher Scientific; Cat. No.: 15140-122), 1× non-essential amino acids (NEAA; Thermo Fisher Scientific; Cat. No.: 11140-050), 1× GlutaMax supplement

(Thermo Fisher Scientific; Cat. No.: 35050-061) and 5 ng ml^{-1} Recombinant Human Fibroblast Growth Factor-basic (FGF-2; Peprotech; Cat. No.: 100-18B). The hMSCs were kept at 37°C in a humidified-air 5% CO_2 atmosphere. The harvesting of these human primary cells was done following the Best Practices Code of the Dutch Federation of Biomedical Scientific Societies on anonymous surgery material remnants.

Recombinant DNA plasmids

The AAV transfer plasmids BI17_pAAV-HR$^{S1.A1}$, BI19_pAAV-HR$^{LMN.A1}$, and BI38_pAAV-HR$^{A1.IN17}$ were assembled by using standard recombinant DNA techniques. The complete nucleotide sequences and respective annotated maps of these constructs are available in the **Supplementary information.**

Recombinant AAV productions

Recombinant AAV particles were assembled on the basis of BI17_pAAV-HR$^{S1.A1}$, BI19_pAAV-HR$^{LMN.A1}$, and BI38_pAAV-HR$^{A1.IN17}$ as follows. HEK293T cells were seeded in T175-cm^2 culture flasks at a density of 2×10^7 cells per flask (up to 18 flasks per AAV vector stock) and, the next day, they were transfected with each AAV transfer plasmid (**Supplementary information**) together with the packaging plasmid AT51_pDG6.RSV.DsRed. SV40pA mixed at 1:1 molar ratios (30 µg total DNA per T175-cm^2 flask). This packaging plasmid expresses the AAV serotype-2 *rep* and AAV serotype-6 *cap* genes together with adenovirus helper functions, *i.e.*, VA RNAs I and II, E4ORF6, and E2A (Grimm *et al.* 2003).

Each T175-cm^2 culture flask received 99 µl of a 25-kDa linear polyethylenimine (PEI) solution (Polysciences) at 1 mg ml^{-1} and DNA mixtures, each diluted in 1 ml of 150 mM NaCl. These transfection mixtures were made by dropwise addition of the PEI to the DNA followed by direct homogenization in a vortex for 10 seconds. After 16-18 minutes at room temperature, the resulting DNA-PEI complexes were added to the HEK293T cells with the transfection medium being replaced 24 hours later by 20 ml of culture medium. The HEK293T cell were detached at 5 days post-transfection by using a cell scrapper and collected into 50-ml tubes together with the conditioned medium. This material was then centrifuged at 1,000 ×*g* for 10 min at 4°C and the resulting supernatants and cell pellets were separately recovered and stored at -80°C until further processing. After thawing, 25 ml of a 40% (*w/v*) polyethylene glycol 8000 solution (PEG 8000; Sigma-Aldrich; Cat. No.: P2139) was added per 100 ml of supernatant with this mixture being first gently stirred for 1 hour at 4°C and subsequently stayed overnight at 4°C without stirring for particle precipitation. Next, the supernatant-PEG8000 mixtures were subjected to 2,820 ×*g* for 15 min at 4°C in 50-ml tubes after which the pellets were resuspended in 7 ml of PBS (pH 7.4) and mixed with 10 ml of clarified cell lysates to yield 17 ml of vector suspensions. The clarified cell lysates were generated by resuspending the producer-cell pellets in 10 ml of PBS (pH 7.4), subjecting the resuspended cells to three rounds of freezing and thawing using liquid N_2 and 37°C water baths, respectively, and eliminating cell debris via centrifugation at 3,220 ×*g* for 15 min at 4°C. The 17-ml AAV vector suspensions were then exposed to 50 U ml^{-1} of Benzonase (Millipore; Cat. No.: UFC910024) for 1 hour at 37°C and subsequently centrifuged at 2,420 ×*g* for 10 min at 4°C.

Recombinant AAV purification and characterization

The clarified supernatants harboring the AAV particles were then placed onto Iodixanol-OptiPrep (Progen; Cat. No.: 1114542) cushions of 15%, 25%, 40% and 60% in Quick-Seal round-top polypropylene tubes (Beckman; Cat. No.: 342414). The AAV vectors were purified through iodixanol gradient ultracentrifugation at 69,000 RPM in a 70Ti rotor (Beckman Coulter) at 16°C in a Beckman Coulter Optima XE-90 centrifuge. The ultracentrifuge tubes were pierced with a needle (18G needle BD Microlance™; Cat. No.: 304622) for recovering AAV vector particles in the 40% iodixanol cushion. The collected material was then subjected to buffer exchange using Amicon Ultra-15 100K MWCO filters (Millipore; Cat. No.: UFC910024) and Dulbecco's Phosphate-Buffered Saline (DPBS; Thermo Fisher Scientific; Cat. No.:

14040-091) containing 0.001% Poloxamer 188 (Sigma-Aldrich; Cat. No.: P5556). The purified batches of AAV-HR$^{S1.A1}$, AAV-HR$^{LMN.A1}$, and AAV-HR$^{A1.IN17}$ were stored at -80°C and their transducing unit (TU) titers were determined by end-point titrations on HeLa cells using flow cytometry or qPCR assays as readouts. For AAV-HR$^{S1.A1}$ and AAV-HR$^{A1.IN17}$ titrations, HeLa cells were seeded at a density of 5×10^4 cells per well of 24-well plates (Greiner Bio-One) and approximately 18 hours later, they were incubated with 3-fold serial dilutions of each vector batch. The frequencies of transduced cells were determined at 3 days post-transduction by EGFP- or mScarlet-directed flow cytometry with the functional AAV vector titers corresponding to TU per ml being determined as the percentage of transduced cells × number of cells seeded × dilution factor x 1000 μl^{-1}. The transducing titers of AAV-HR$^{LMN.A1}$, expressed in GC per ml, were determined via a qPCR assay based on the iQ SYBR Green Supermix (Bio-Rad, cat. No. L010171C) and the *mScarlet-I*-specific primers 5'-CTACCTGGCGGACTT CAAGA-3' and 5'-ACGGTGTAGTCCTCGTTGTG-3'. In brief, HeLa cells were seeded at 8.5×10^4 cells per well of 24-well plates (Greinder Bio-One) and, the next day, they were exposed to seven 3-fold serial dilutions of purified vector stock. Next, qPCR analysis was performed on genomic DNA isolated via the DNeasy Blood & Tissue kit at 24 hours post-transduction. In parallel, eight serial 10-fold dilutions of linearized parental AAV vector DNA containing 1×10^7 GC ml^{-1} served to setup a qPCR standard curve. Data analysis was done with the Bio-Rad CFX Manager 3.1 software (Bio-Rad Laboratories) and the titer was determined on the basis of the AAV vector DNA and plasmid Ct value standard curve.

Testing selector AAV genome editing at *AAVS1* and *ATP1A1*

HeLa cells were seeded at the density of 5×10^4 cells per well of 24-well plates. The next day, the cells were exposed or not to complexes formed by incubating 150 mM NaCl solutions containing 2.19 μl of PEI (1 mg ml^{-1}) and plasmid mixtures consisting of 420 ng of AV62_pU.CAG.Cas9.rBGpA or 420 ng AB66_pU.CAG.dCas9.rBGpA as negative control. To control for transfection efficiency, all transfection mixtures were spiked with 80 ng of plasmid BE08_pCAG.mCherry.bGHpA expressing a red fluorescence-emitting reporter. After a 6-h incubation period, the medium of cells transfection-treated was removed, and the cells were transduced in regular culture medium containing AAV-HR$^{S1.A1}$ particles at 1 TU cell^{-1}. Ouabain selection of HeLa cells was initiated at 48 h post-trasdution with the concentration of 0.2 µM. HeLa cells were gone through flow cytometry at day 3 and day 20 to determine the transient and stable editing efficiencies, respectively, in the condition of ouabain selection or not.

Testing selector AAV in-linkage donor

HeLa cells were seeded 5 × 10^4 cells per well of 24-well plates. The next day, the cells were exposed or not to complexes formed by incubating 150 mM NaCl solutions containing 2.19 µl of PEI (1 mg ml^{-1}) and plasmid mixtures consisting of 420 ng of AV62_pU.CAG.Cas9.rBGpA, AB65_pU.CAG.Cas9 D10A.rBGpA, or AB66_pU.CAG.dCas9.rBGpA as negative control. To control for transfection efficiency, all transfection mixtures were spiked with 80 ng of plasmid AZ15_pU.CAG.eGFP.rBGpA expressing a green fluorescence-emitting reporter. After a 6-h incubation period, the medium of cells transfection-treated was removed, and the cells were transduced in regular culture medium containing AAV-HR$^{A1.IN17}$ particles at 1 TU cell^{-1}. Ouabain selection of HeLa cells was initiated at day 4 post-trasdution with the concentration of 0.2 µM. HeLa cells were went through flow cytometry at day 3 and day 20 to determine the transient and stable editing efficiencies, respectively, in the condition of Ouabain selection or not.

Viral vector transductions and DNA editing assays

HeLa cells and hMSCs were seeded at, respectively, 2×10^4 and 5×10^4 cells per well of 48- and 24-well plates (Greinder Bio-One) and, after overnight incubation, they were mock-transduced or were transduced for 24 hours with combinations of AdVP.C9^{KARA} (Li *et al.* 2024) and selector AAV vectors or with selector AAV vectors alone at the indicated MOIs. At 5- and 6-days post-transduction, respectively, HeLa cells and hMSCs were cultured in the absence and in the presence of ouabain at final concentrations of 0.2 µM and 1.0 µM, respectively. Parallel cultures of mock-transduced cells were also exposed and not exposed to ouabain. Subsequently, genome modification endpoints were assessed at the indicated timepoints post-transduction by a combination of reporter-directed flow cytometry and *AAVS1*, *ATP1A1*, and *LMNA* genotyping analyses. The latter analysis involved RFLP and junction PCR assays whose details are specified in **Supplementary Tables S6 - S9**, **S12** and **S13**.

Microscopy analysis

The direct fluorescence microscopy analysis of reporter-tagged LMNA at early and late timepoints post-transduction was performed with an AF6000 LX inverted fluorescence microscope (Leica) and the resulting images were examined with the aid of the LAS X software (Leica Microsystems).

Flow cytometry analysis

Selector AAV vector transduction efficiencies and corresponding mean fluorescence intensities per cell were determined through reporter-directed flow cytometry using a BD LSR II FACS (BD Biosciences). The same apparatus was also used to quantify AAV stable transduction levels and *LMNA*-tagging frequencies. In brief, mock- and vector-transduced cells were rinsed with PBS (pH 7.4) and incubated in 0.05% trypsin-EDTA (Thermo Fisher Scientific; Cat. No.: 15400-054). The resulting cell suspensions were then collected in cultured medium, briefly centrifuged and resuspended in FACS buffer composed of PBS (pH 7.4) containing 0.5% (*w/v*) BSA and 2 mM EDTA (pH 8.0). The mock-transduced cells served to set the background fluorescence threshold cutoff. At least 10,000 viable single cells were acquired per sample. The resulting datasets were analyzed with the aid of the FlowJo 10.9.0 software (BD Biosciences).

Characterization of genome editing events

After sub-culturing selector AAV transduced cells for more than two weeks, reporter-positive cells were sorted by using either a BD FACSAria III flow cytometer (BD Biosciences) or a CytoFLEX SRT Cell Sorter (Beckman). Next, individual reporter-positive cells were seeded in wells of 96-well plates in 1:1 mixtures of culture medium and FBS supplemented with penicillin/streptomycin at 100 U ml^{-1} with or without 0.2 μM ouabain. Moreover, to increase cell cloning efficiencies, α-thioglycerol and bathocuprione disulphonate (both from Sigma-Aldrich) were added at final concentrations of 50 μM and 20 nM, respectively. Single cell-derived clones expressing reporter proteins were arbitrarily collected after 2-3 weeks for genotyping through junction PCR analyses using the Phire Tissue Direct PCR Master mix according to the manufacturer's instructions (Thermo Fisher Scientific; Cat. No.: F170L). The PCR mixture reagents and thermocycling parameters used in the clonal screens are indicated in the **Supplementary Tables S2 - S5**, **S10** and **S11**.

Statistical analyses

Statistical analyses were performed with the aid of the GraphPad Prism software (version 9.3.1) on datasets derived from a minimum of three biological replicates. Two-tailed unpaired Student's t tests were performed to assess statistical significance amongst two independent experimental groups. Details on statistical parameters are also indicated in the figure legends where applicable. P values inferior to 0.05 were considered to be statistically significant.

Acknowledgements

The authors are thankful to Thilo M. Buck and Jan Wijnholds (Department of Ophthalmology, Leiden University Medical Centre, The Netherlands) for their advice during the setting-up of AAV vector production procedures in our laboratory. The authors also thank the personnel of the Flow Cytometry Core Facility of the LUMC for aiding with the cell sorting procedures.

Authors in this study are members of the European Reference Network—Neuromuscular Diseases (ERN EURO-NMD).

Author contributions: Z.L. and X.W. generated and characterized reagents, designed and performed the experiments, examined the datasets and wrote the paper together with M.A.F.V.G.; J.L. generated, characterized and tested reagents; J.M.J. generated, characterized and tested reagents; R.H. supervised the research and analyzed the results; M.A.F.V.G. designed and supervised the research, analyzed the data and wrote the paper together with Z.L and X.W..

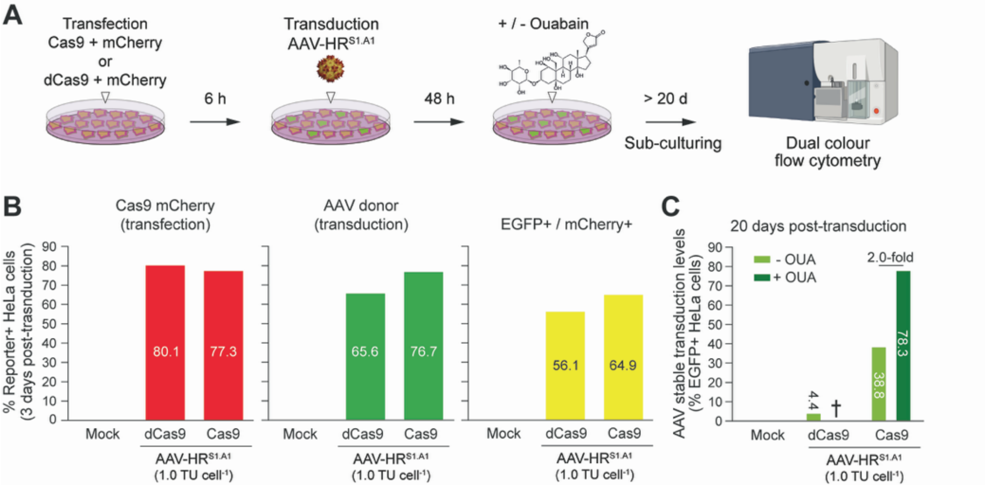
Funding

China Scholarship Council - Leiden University Scholarship to Z.L. and X.W.. Research in the author's laboratory is supported by the Prinses Beatrix Spierfonds, the Dutch Duchenne Parent Project, the Horizon Europe Programme, and the Dutch Research Council (NWO)—Open Technology Programme.

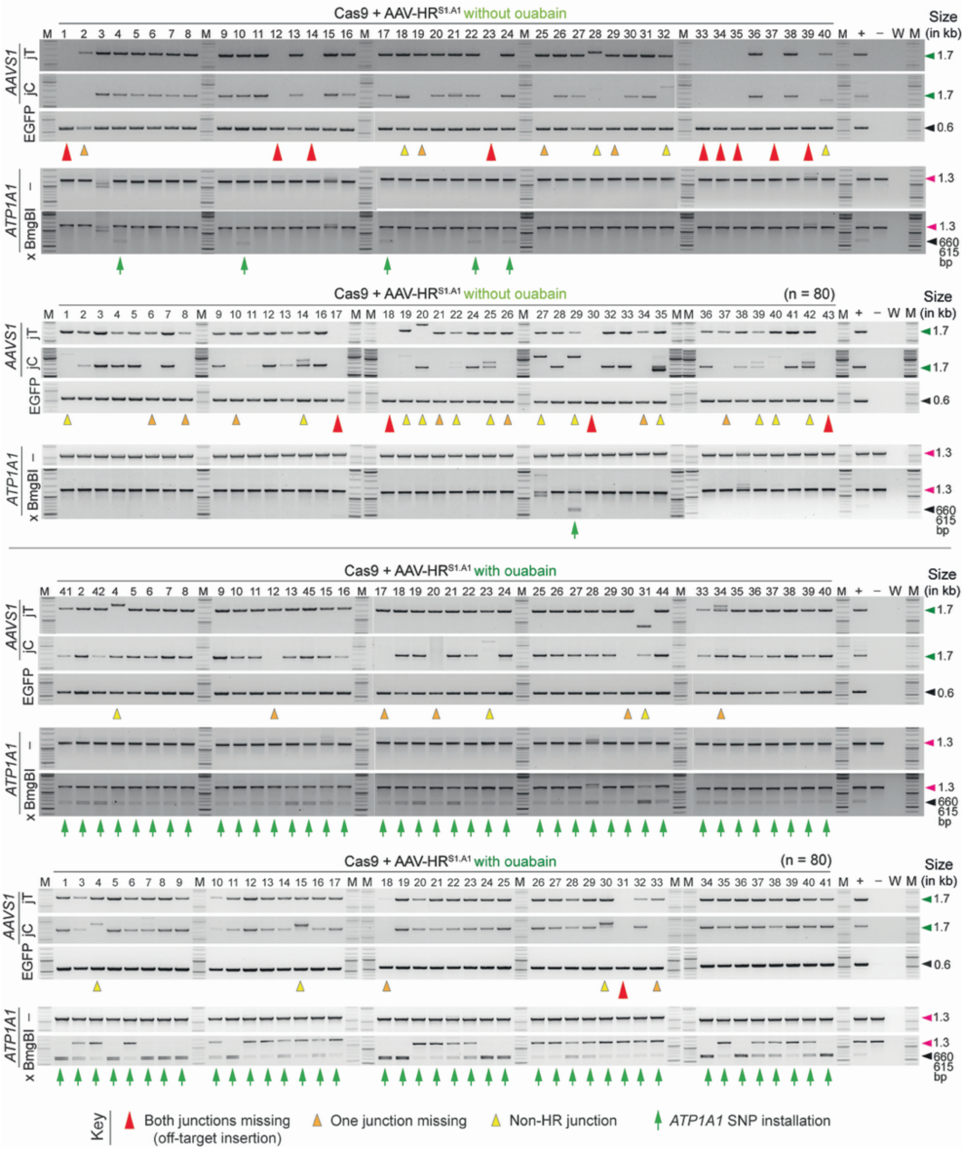
Conflict of interest statement

None declared.

Supplementary Figures

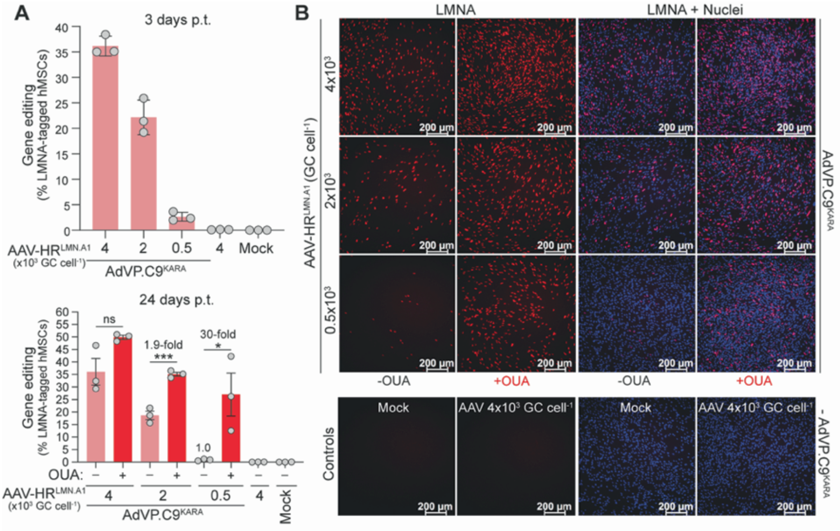


Supplementary Figure S1. Testing selector AAV-based genome editing at *AAVS1*. (**A**) Diagram of the experimental setup. HeLa cells transfected with plasmids expressing Cas9 and mCherry or inactive dCas9 and mCherry are transduced with the selector AAV-HR$^{S1.A1}$ donor and, after sub-culturing in the presence or absence of ouabain, are subjected to EGFP- and mCherry-directed flow cytometry. HeLa cells only exposed to AAV-HR$^{S1.A1}$ served as a control. (**B**) Quantification of transfection efficiency and selector AAV donor delivery. The transfection efficiency and AAV-HR$^{S1.A1}$ donor delivery were determined by mCherry- and EGFP-directed flow cytometry at 3 days post-transduction (left and central graphs, respectively). The frequencies of transfected and transduced cells are also plotted (right graph). (**C**) Quantification of selector AAV-based DNA editing. The AAV stable transduction levels, serving as a proxy for genome editing frequencies, were determined at 20 days post-transduction by EGFP-directed flow cytometry. The cross indicates complete cell death in ouabain-treated cultures exposed exclusively to AAV-HR$^{S1.A1}$.

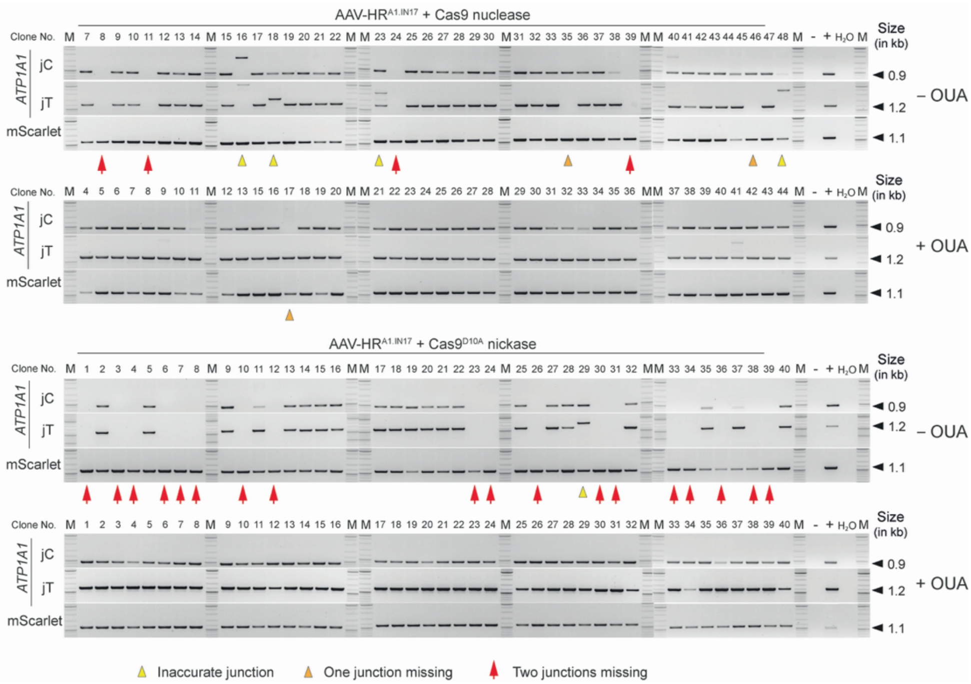

Cas9 + AAV-HR^S1.A1 without ouabain
Cas9 + AAV-HR^S1.A1 without ouabain (n = 80)
Cas9 + AAV-HR^S1.A1 with ouabain
Cas9 + AAV-HR^S1.A1 with ouabain (n = 80)
M
+
−
W
Size (in kb)
AAVS1
jT
jC
EGFP
ATP1A1
−
x BmgBI
1.7
1.7
0.6
1.3
1.3
660
615
bp
Key
Both junctions missing (off-target insertion)
One junction missing
Non-HR junction
ATP1A1 SNP installation

Supplementary Figure S2. Characterization of *AAVS1* and *ATP1A1* co-targeting resulting from selectable AAV donor delivery. PCR genotyping of individual genetically-modified cell clones generated via AAV-HR$^{S1.A1}$ transduction of HeLa cells transfected with a Cas9-encoding plasmid. Prior to clone isolation, the HeLa cell populations were expanded either in the presence or absence of ouabain. Nuclease-free water (W) and DNA from unmodified HeLa cells (-) provided for negative controls; genomic DNA from sorted EGFP-positive HeLa cells exposed to Cas9 and AAV-HR$^{S1.A1}$, served as positive control (+). The *AAVS1* and *ATP1A1* genotype screens involved junction

PCR (jC and jT) and RFLP (×BmgBI) assays, respectively. Clones lacking *AAVS1* targeted insertions or with only one HR-derived junction between transgenic and *AAVS1* sequences are marked by red and orange arrowheads, respectively; whilst clones containing HR-independent donor DNA insertions are highlighted by yellow arrowheads. The remaining clones, modified through precise HR events involving the telomeric and centromeric side of the target sequence (jT and jC, respectively), are not highlighted. Clones with edited *ATP1A1* alleles, scored via BmgBI amplicon digestion, are marked by green arrows. The cumulative datasets corresponding to these *AAVS1* and *ATP1A1* genotype screens (n=160 clones representing independent genome-modifying events), are plotted in **Figure 1F**.



Supplementary Figure S3. Selector AAV gene editing of *LMNA* in hMSCs using combinatorial viral vector delivery. (**A**) Quantification of AAV gene editing with and without ouabain. hMSCs were co-transduced with AAV-HR$^{LMN.A1}$ and AdVP.C9^{KARA} at the specified doses. Gene editing frequencies were measured through mScarlet-directed flow cytometry at 3- and 24- days post-transduction (top and bottom graph, respectively). Prior to flow cytometry analysis at the latter timepoint, the hMSCs were cultured in the presence or absence of ouabain. Mock-transduced hMSCs and hMSCs transduced only with the highest dose of AAV-HR$^{LMN.A1}$ provided for negative controls. The results are presented as mean ± SD of three biological replicates. Significant differences amongst the marked datasets were calculated by Student's t-tests; **$P<0.05$. (**B**) Characterization of LMNA protein tagging in hMSCs. LMNA tagging and nuclear localization was monitored by combining direct fluorescence microscopy for the reporter mScarlet and the DNA dye Hoechst 33342, respectively.


AAV-HR$^{A1.IN17}$ + Cas9 nuclease
Clone No. M 7 8 9 10 11 12 13 14 M 15 16 17 18 19 20 21 22 M M 23 24 25 26 27 28 29 30 M 31 32 33 35 36 37 38 39 M M 40 41 42 43 44 45 46 47 48 M - + H$_2$O M
Size (in kb)
ATP1A1
jC
jT
mScarlet
0.9
1.2
1.1
– OUA
Clone No. M 4 5 6 7 8 9 10 11 M 12 13 15 16 17 18 19 20 M M 21 22 23 24 25 26 27 28 M 29 30 31 32 33 34 35 36 MM 37 38 39 40 41 42 43 44 M - + H$_2$O M
+ OUA
AAV-HR$^{A1.IN17}$ + Cas9^{D10A} nickase
Clone No. M 1 2 3 4 5 6 7 8 M 9 10 11 12 13 14 15 16 MM 17 18 19 20 21 22 23 24 M 25 26 27 28 29 30 31 32 MM 33 34 35 36 37 38 39 40 M - + H$_2$O M
– OUA
Clone No. M 1 2 3 4 5 6 7 8 M 9 10 11 12 13 14 15 16 M M 17 18 19 20 21 22 23 24 M 25 26 27 28 29 30 31 32M M 33 34 35 36 37 38 39 40 M - + H$_2$O M
+ OUA
Inaccurate junction
One junction missing
Two junctions missing

Supplementary Figure S4. Characterization of *ATP1A1* gene targeting resulting from selectable in-linkage AAV donor delivery. PCR genotyping of individual genetically-modified cell clones generated via AAV-HR$^{A1.IN17}$ transduction of HeLa cells transfected with a Cas9- or Cas9^{D10A}-encoding plasmids. Prior to clone isolation, the HeLa cell populations were expanded either in the presence or absence of ouabain. Nuclease-free water (W) and DNA from unmodified HeLa cells (-) provided for negative controls; genomic DNA from sorted mScarlet-positive HeLa cells exposed to Cas9 or Cas9^{D10A} and AAV-HR$^{A1.IN17}$, served as positive controls (+). The *ATP1A1* genotype screens involved junction PCR (jC and jT) analysis. Clones lacking *ATP1A1* targeted insertions or with only one HR-derived junction between transgenic and *ATP1A1* sequences are marked by red and orange arrowheads, respectively; whilst clones containing HR-independent donor DNA insertions are highlighted by yellow arrowheads. The remaining clones, modified through precise HR events involving the telomeric and centromeric side of the target sequence (jT and jC, respectively), are not highlighted. The cumulative datasets corresponding to these *ATP1A1* genotype screens (n=160 clones representing independent genome-modifying events), are plotted in **Figure 4E**.

Supplementary Tables

Supplementary Table S1. Titres of AAV vectors generated for this study.

Plasmid code	**AAV/AdVP vector**	**Titre**
BI17	AAV-HR$^{S1.A1}$	1.86E+08 TU/mL
BI19	AAV-HR$^{LMN.A1}$	1.14E+11 GC/mL
BI38	AAV-HR$^{A1.IN17}$	8.61E+07 TU/mL
U67	AdVP.C9^{KARA}	5.43E+07 GC/mL

Supplementary Table S2. Primers and PCR mixtures used for clonal screening of *AAVS1* (**Figures 1E**, **1F** and **Supplementary figure S2**)

Target	Primer code	Primers (5’ → 3’) / final concentrations (µM)	2X Phire Tissue Direct PCR Master mix µl	DMSO	Amplicon size (bp)
GFP	#978	GAGCTGGACGGCGACGTAAACG / 0,5	10		596
	#979	CGCTTCTCGTTGGGGTCTTTGCT / 0,5			
jT.*AAVS1*	#986	AACCCCAACCCCGTGGAAG / 0,5	10	2%	1666
	#1004	GCACCGTCCGCTTCGAG / 0,5			
jC.*AAVS1*	#1046	CGACAACCACTACCTGAGCA / 0,5	10		1712
	#1047	GACCTGCCTGGAGAAGGAT / 0,5			

Supplementary Table S3. PCR cycling parameters used for clonal screening of *AAVS1* (**Figures 1E**, **1F** and **Supplementary figure S2**)

Target	Initial denaturation	Denaturation	Annealing	Elongation	Cycles	Final elongation
GFP	98 °C	98 °C	72 °C		35	72 °C
	5 min	5 sec	20 sec			2 min
jT.*AAVS1*	98 °C	98 °C	72 °C		35	72 °C
	5 min	7 sec	30 sec			2 min
jC.*AAVS1*	98 °C	98 °C	72 °C	72 °C	TD 0,5 °C decrease/cycle*10	72 °C
	5 min	7 sec	7 sec	32 sec		
		98 °C	67 °C	72 °C	25	2 min
		7 sec	7 sec	32 sec		

Supplementary Table S4. Primers and PCR mixtures used for clonal screening of *ATP1A1* (**Figures 1E**, **1F** and **Supplementary figure S2**)

Target	Primer code	Primers (5' → 3') / final concentrations (μM)	dNTP (mM)	$MgCl_2$ (mM)	GoTaq Flexi Buffer	GoTaq (Units)	Amplicon size (bp)
ATP1A1	#2225	CCCCTCCCGACAAAATCAATAC/0,4	0.4	1	1x	1.25	1275
	#2228	TAGCACCACACCCAGGTACA/0,4					

Supplementary Table S5. PCR cycling parameters used for clonal screening of *ATP1A1* (**Figures 1E**, **1F** and **Supplementary figure S2**)

Target	Initial denaturation	Denaturation	Annealing	Elongation	Cycles	Final elongation
ATP1A1	95 ℃	95 ℃	64 ℃	72 ℃	40	72 ℃
	5 min	30 sec	30 sec	1 min 20 sec		5 min

Supplementary Table S6. Primers and PCR mixtures used for genotyping of *AAVS1* and *ATP1A1* (**Figures 2D**)

Target	Primer code	Primers (5’ → 3’) / final concentrations (µM)	dNTP (mM)	MgCl$_2$ (mM)	GoTaq Flexi Buffer	GoTaq (Units)	DMSO	Amplicon size (bp)
EGFP	#978	GAGCTGGACGGCGA CGTAAACG / 0,4	0.4	1	1×	1.25	/	596
	#979	CGCTTCTCGTTGGG GTCTTTGCT / 0,4						
jT.*AAVS1*	#986	AACCCCAACCCCGT GGAAG / 0,4	0.4	1	1×	1.25	2%	1666
	#1004	GCACCGTCCGCTTC GAG / 0,4						
jC.*AAVS1*	#1046	CGACAACCACTACC TGAGCA / 0,4	0.4	1	1×	1.25	/	1712
	#1047	GACCTGCCTGGAGA AGGAT / 0,4						
ATP1A1	#2225	CCCCTCCCGACAAA ATCAATAC / 0.4	0.4	1	1×	1.25	/	1275
	#2228	TAGCACCACACCCA GGTACA / 0.4						

Supplementary Table S7. PCR cycling parameters used for genotyping of *AAVS1* and *ATP1A1* (**Figures 2D**)

Target	Initial denaturation	Denaturation	Annealing	Elongation	Cycles	Final elongation
EGFP	95 °C	95 °C	72 °C		35	72 °C
	5 min	30 sec	20 sec			5 min
jT.*AAVS1*	95 °C	95 °C	61 °C	72 °C	35	72 °C
	5 min	30 sec	30 sec	1 min 30 sec		1 min 30 sec
jC.*AAVS1*	95 °C	95 °C	67 °C	72 °C	35	72 °C
	5 min	30 sec	30 sec	1 min 30 sec		1 min 30 sec
ATP1A1	95 °C	95 °C	64 °C	72 °C	40	72 °C
	5 min	30 sec	30 sec	1 min 20 sec		5 min

Supplementary Table S8. Primers and PCR mixtures used for genotyping of *LMNA* and *ATP1A1* (**Figures 3C**, **3E**)

Target	Primer code	Primers (5’ → 3’) / final concentrations (µM)	dNTP (mM)	$MgCl_2$ (mM)	GoTaq Flexi Buffer	GoTaq (Units)	Amplicon size (bp)
mScarlet	#2328	CACGAGTTCGAGATCGAGGG / 0.4	0.2	1	1x	1.25	573
	#2329	TTCGTACTGTTCCACCACGG / 0.4					
jT.*LMNA*	#2332	GAACAGTACGAACGCTCCGA / 0.4	0.2	1	1x	1.25	781
	#2333	CTGGGTGCCCAGAGTTCTTC / 0.4					
jC.*LMNA*	#2330	TGAGTCACACTGATGGGCAC / 0.4	0.2	1	1x	1.25	1263
	#2331	GGTGTAGTCCTCGTTGTGGG / 0.4					
ATP1A1	#2225	CCCCTCCCGACAAAATCAATAC / 0.4	0.2	1	1x	1.25	1275
	#2228	TAGCACCACACCCAGGTACA / 0.4					

Supplementary Table S9. PCR cycling parameters used genotyping of *LMNA* and *ATP1A1* (**Figures 3C**, **3E**)

Target	Initial denaturation	Denaturation	Annealing	Elongation	Cycles	Final elongation
mScarlet	95 °C	95 °C	62.9 °C	72 °C	30	72 °C
	5 min	30 sec	30 sec	30 sec		5 min
jT.*LMNA*	95 °C	95 °C	62.9 °C	72 °C	30	72 °C
	5 min	30 sec	30 sec	40 sec		5 min
jC.*LMNA*	95 °C	95 °C	62.9 °C	72 °C	30	72 °C
	5 min	30 sec	30 sec	1 min		5 min
ATP1A1	95 °C	95 °C	64 °C	72 °C	30	72 °C
	5 min	30 sec	30 sec	1 min		5 min

Supplementary Table S10. Primers and PCR mixtures used for clonal screening of *ATP1A1::mScarlet* (**Figures 4E** and **Supplementary figure S4**)

Target	Primer code	Primers (5’ → 3’) / final concentrations (µM)	2X Phire Tissue Direct PCR Master mix µl	Amplicon size (bp)
mScarlet	#2120	ACGGTGTAGTCCTCGTTGTG / 0.5	10	1095
	#1648	TCTCGCACATTCTTCACGTC / 0.5		
jT.*ATP1A1*	#2229	ACTACAGGGCGTGCATACAG / 0.5	10	1199
	#2230	CCCACAACGAGGACTACACC / 0.5		
jC.*ATP1A1*	#2234	GGTGACCTACCAGCCAAACT / 0.5	10	945
	#2235	CTTGGAAAAGGCGCAACCC / 0.5		

Supplementary Table S11. PCR cycling parameters used for clonal screening of *ATP1A1::mScarlet* (**Figures 4E** and **Supplementary figure S4**)

Target	Initial denaturation	Denaturation	Annealing	Elongation	Cycles	Final elongation
mScarlet	98 °C	98 °C	72 °C		30	72 °C
	5 min	7 sec	32 sec			2 min
jT.*ATP1A1*	98 °C	98 °C	67 °C	72 °C	30	72 °C
	5 min	7 sec	5 sec	20 sec		2 min
jC.*ATP1A1*	98 °C	98 °C	72 °C		35	72 °C
	5 min	7 sec	20 sec			2 min

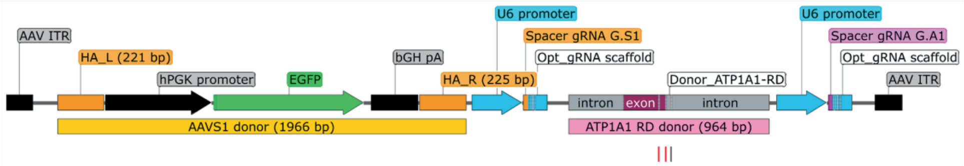
Supplementary Table S12. Primers and PCR mixtures used for junction PCR analysis of *ATP1A1::mScarlet* (**Figures 5F**)

Target	Primer code	Primers (5’ → 3’) / final concentrations (µM)	dNTP (mM)	$MgCl_2$ (mM)	GoTaq Flexi Buffer	GoTaq (Units)	Amplicon size (bp)
mScarlet	#2120	ACGGTGTAGTCCTCGTTGTG / 0.4	0.4	1	1x	1.25	1095
	#1648	TCTCGCACATTCTTCACGTC / 0.4					
jT.*ATP1A1*	#2229	ACTACAGGGCGTGCATACAG / 0.4	0.4	1	1x	1.25	1199
	#2230	CCCACAACGAGGACTACACC / 0.4					
jC.*ATP1A1*	#2234	GGTGACCTACCAGCCAAACT / 0.4	0.4	1	1x	1.25	945
	#2235	CTTGGAAAAGGCGCAACCC / 0.4					

Supplementary Table S13. PCR cycling parameters used for junction PCR analysis of *ATP1A1::mScarlet* (**Figures 5F**)

Target	Initial denaturation	Denaturation	Annealing	Elongation	Cycles	Final elongation
mScarlet	95 ℃	95 ℃	63 ℃	72 ℃	35	72 ℃
	5 min	30 sec	30 sec	1 min		5 min
jT.*ATP1A1*	95 ℃	95 ℃	65 ℃	72 ℃	35	72 ℃
	5 min	30 sec	30 sec	1 min		1 min 30 sec
jC.*ATP1A1*	95 ℃	95 ℃	65 ℃	72 ℃	35	72 ℃
	5 min	30 sec	30 sec	1 min		1 min 30 sec

Supplementary Information

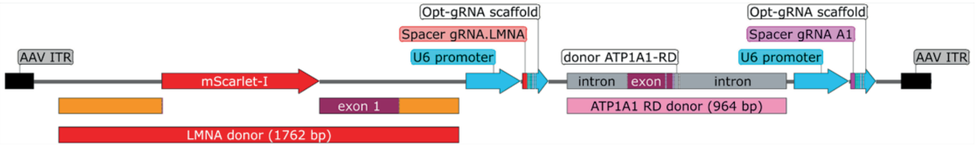


>BI17_pAAV-HR$^{S1.A1}$ (4304 bp)

```
ctggcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcct
cagtgagcgagcgagcgcgcagagagggagtggccaactccatcactaggggttccttgtagttaatgattaac
ccgccatgctacttatctacgtggccactagtacttctcgagctctgtacatgtccgcggtcgcgacgtacgcg
tatcgatggcgccagctgcagagctctagctcttccagcccctgtcatggcatcttccaggggtccgagagct
cagctagtcttcttcctccaacccgggcccctatgtccacttcaggacagcatgtttgctgcctccagggatcc
tgtgtccccgagctgggaccaccttatattcccagggccggttaatgtggctctggttctgggtacttttatct
gtcccctccaccccacagtggggcaagcttccacggggttggggttgcgccttttccaaggcagccctgggttt
gcgcagggacgcggctgctctgggcgtggttccgggaaacgcagcggcgccgaccctgggtctcgcacattctt
cacgtccgttcgcagcgtcacccggatcttcgccgctacccttgtgggccccccggcgacgcttcctgctccgc
ccctaagtcgggaaggttccttgcggttcgcggcgtgccggacgtgacaaacggaagccgcacgtctcactagt
accctcgcagacggacagcgccagggagcaatggcagcgcgccgaccgcgatgggctgtggccaatagcggctg
ctcagcagggcgcgccgagagcagcggccgggaaggggcggtgcgggaggcggggtgtggggcggtagtgtggg
ccctgttcctgcccgcgcggtgttccgcattctgcaagcctccggagcgcacgtcggcagtcggctccctcgtt
gaccgaatcaccgacctctctccccaccggtcgccaccatggtgagcaagggcgaggagctgttcaccggggtg
gtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcga
tgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcg
tgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgaccacatgaagcagcacgacttcttcaag
tccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgc
cgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggca
acatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaagaac
ggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagca
gaacacccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtccgccctgagca
aagaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatg
gacgagctgtacaagtaaagcggccggccgcgtcgagtctaggatcagcctcgactgtgccttctagttgccag
ccatctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaata
aaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctggggggtggggtggggcaggacagca
agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggaagctttactagggaca
ggattggtgacagaaaagccccatccttaggcctcctccttcctagtctcctgatattgggtctaacccccacc
tcctgttaggcagattccttatctggtgacacacccccatttcctggagccatctctctccttgccagaacctc
taaggtttgcttacgatggagccagagaggatcctgggagggagagcttggcagggggtgggagggagctagag
tccggccgccccttcaccgagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgtta
gagagataattggaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataa
tttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagta
tttcgatttcttggctttatatatcttgtggaaaggacgaaacaccggggccactagggacaggatgtttcaga
gctatgctggaaacagcatagcaagttgaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcg
gtgcttttttgaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaa
tcgccttgcagcacatccccctttcgccagaacaaccggtaccaaatttattgatggatcaatttaaagagttt
taatctgggtgttatgagttccttgggcctattgtttgcctgaaccctgtggggactggctcatcagcagaatt
attcatggaggaatttgctaggttttaccttggctctctagcttgggacattttgtttcttccttaaatcctta
ttgcaaccgtccagctaccaggtaggtatattgccttgtaagtgctggtacagtttgccttatttatattccac
tgcttctcagggattaacatctgctcgtgcagctgagatcctggcgcgagatggtcccaacgccctcactcccc
ctcccactactcctgaatggatcaagttttgtcggcagctctttggggggttctcaatgttactgtggattgga
gcgattctttgtttcttggcttatagcatcagagctgctacagaagaggaacctcaaaacgatgacgtgagttc
```

```
tgtaattcagcatatcgatttgtagtacacatcagatatcttctccgtctttgtctcccacttcttctcaatta
ccactcattacttaatggttatgaactcattacttaatggttatgaacagctgttgccttcaaggctcatccat
tcttccttcgtttccatttcctctctctaccacccacgttgtagatgctcttacaagtgggatgcccacctgca
tgtgctgctgtgagcaggcagctctgctcaggccccggccgccacccactggatggcagagcacagcgattcat
gttggcacatccacctgtccagaaatgcttggtggccattcttcaaaatccacaagttggttgaaaaacaatct
ttgttttataaagcaggagaaactgatgcatctagaacctttccaaacgtccagttagtgatcaagtgttggtg
tgcctgactccatttctgacccttcctgtgtggtctttagaagggaattggatcctagagtccggccgcccccct
tcaccgagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattgga
attaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtt
tgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttgg
ctttatatatcttgtggaaaggacgaaacaccgagttctgtaattcagcatagtttcagagctatgctggaaac
agcatagcaagttgaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttttttttgaa
ttgatccatatatagggcccgggttataattacctcaggtcgacgtcccatgtgcaggtgctgaattgtagata
agtagcatggcgggttaatcattaactacaaggaacccctagtgatggagttggccactccctctctgcgcgct
cgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcga
gcgagcgcgcag
```

Map and nucleotide sequence of selector AAV donor construct BI17_pAAV-HR$^{S1.A1}$. Diagram of AAV-HR$^{S1.A1}$ vector genome containing an *AAVS1*-targeting HR donor template (1966 bp). The selector donor DNA sequence (964 bp) is designed for installing the ouabain resistance gain-of-function Q118R and N129D SNPs (RD), vertical red dashes, within the exon 4 of *ATP1A1* upon gRNA G^{A1}-directed cleavage within intron 4. AAV ITR, adeno-associated virus type-2 inverted terminal repeat; orange bars, homology arms ("Left" and "Right") consisting of DNA homologous to sequences flanking the *AAVS1* target site of gRNA G^{S1}; hPGK, human phosphoglycerate kinase 1 gene promoter; EGFP, enhanced green fluorescence protein reporter; bGH pA, bovine growth hormone gene polyadenylation signal; U6 promoter, human U6 snRNA polymerase III promoter driving the expression of each gRNA. The vector plasmid backbone is not shown.

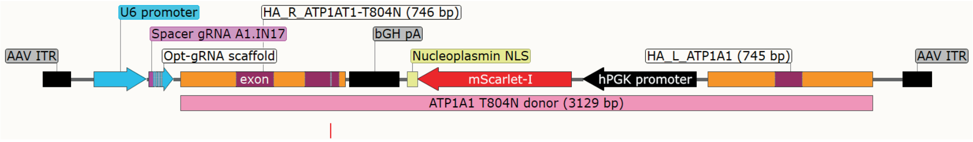


>BI19_pAAV-HR$^{LMN.A1}$ (4075 bp)

```
ctggcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcct
cagtgagcgagcgagcgcgcagagagggagtggccaactccatcactaggggttccttgtagttaatgattaac
ccgccatgctacttatctacgtggccactagtacttctcgagctctgtacatgtccgcggtcgatcgcatgcct
gcagagctctagtaccggcaagcttggagccgacagtgctgagcaggcaggagccaagagaggggaagcttgag
cctcacgcagttaggggtgcgctggagagggtggggcccgactccgccacaccccaacggtccttccccctcct
caccactcccgcccccaccccccaatggatctgggactgcccctttaagagtagtggcccctcctcccttcagag
gaggacctattagagcctttgccccggcgtcggtgactcagtgttcgcgggagcgccgcacctacaccagccaa
cccagatcccgaggtccgacagcgcccggcccagatccccacgcctgccaggagcaagccgagagccagccggc
cggcgcactccgactccgagcagtctctgtccttcgacccgagccccgcgccctttccgggacccctgccccgc
gggcagcgctgccaacctgccggccatggtgagcaagggcgaggcagtgatcaaggagttcatgcggttcaagg
tgcacatggagggctccatgaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacgagggc
acccagaccgccaagctgaaggtgaccaagggtggccccctgcccttctcctgggacatcctgtcccctcagtt
catgtacggctccagggccttcatcaagcaccccgccgacatccccgactactataagcagtccttccccgagg
gcttcaagtgggagcgcgtgatgaacttcgaggacggcggcgccgtgaccgtgacccaggacacctccctggag
gacggcaccctgatctacaaggtgaagctccgcggcaccaacttccctcctgacggccccgtaatgcagaagaa
gacaatgggctgggaagcatccaccgagcggttgtaccccgaggacggcgtgctgaagggcgacattaagatgg
ccctgcgcctgaaggacggcggccgctacctggcggacttcaagaccacctacaaggccaagaagcccgtgcag
atgcccggcgcctacaacgtcgaccgcaagttggacatcacctcccacaacgaggactacaccgtggtggaaca
gtacgaacgctccgagggccgccactccaccggcggcatggacgagctgtacaagaccccgtcccagcggcgcg
ccacccgcagcggggcgcaggccagctccactccgctgtcgcccacccgcatcacccggctgcaggagaaggag
gacctgcaggagctcaatgatcgcttggcggtctacatcgaccgtgtgcgctcgctggaaacggagaacgcagg
gctgcgccttcgcatcaccgagtctgaagaggtggtcagccgcgaggtgtccggcatcaaggccgcctacgagg
ccgagctcggggatgcccgcaagacccttgactcagtagccaaggagcgcgcccgcctgcagctggagctgagc
aaagtgcgtgaggagtttaaggagctgaaagcgcggtgagttcgcccaggtggctgcgtgcctggcggggagtg
gagagggcggcgggccggcgcccctggccggccgcaggaagggagtgagagggcctggaggccgataactttgc
catagtctcctccctccccggaactgcccccagcgggtgactggcagtgtcaaggggaattgtcaagacaggac
agagagggaagtggtggtctctgggagagggtcggggaggatataaggaatggtgggggtatcagggacaagtt
ggcgaattctagagtccggccgccccccttcaccgagggcctatttcccatgattccttcatatttgcatatacg
atacaaggctgttagagagataattggaattaatttgactgtaaacacaaagatattagtacaaaatacgtgac
gtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttacc
gtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccgccatggagacccc
gtcccaggtttcagagctatgctggaaacagcatagcaagttgaaataaggctagtccgttatcaacttgaaaa
agtggcaccgagtcggtgctttttttgaattcggtaccggcgcgcccgtacgactaggcctattaatattccgg
agtatacgtagccggctaacgttaacaaccggtaccaaatttattgatggatcaatttaaagagttttaatctg
ggtgttatgagttccttgggcctattgtttgcctgaaccctgtggggactggctcatcagcagaattattcatg
gaggaatttgctaggttttaccttggctctctagcttgggacattttgtttcttccttaaatccttattgcaac
cgtccagctaccaggtaggtatattgccttgtaagtgctggtacagtttgccttatttatattccactgcttct
cagggattaacatctgctcgtgcagctgagatcctggcgcgagatggtcccaacgccctcactccccctcccac
tactcctgaatggatcaagttttgtcggcagctctttgggggttctcaatgttactgtggattggagcgattc
tttgtttcttggcttatagcatcagagctgctacagaagaggaacctcaaaacgatgacgtgagttctgtaatt
cagcatatcgatttgtagtacacatcagatatcttctccgtctttgtctcccacttcttctcaattaccactca
ttacttaatggttatgaactcattacttaatggttatgaacagctgttgccttcaaggctcatccattcttcct
tcgtttccatttcctctctctaccacccacgttgtagatgctcttacaagtgggatgcccacctgcatgtgctg
ctgtgagcaggcagctctgctcaggccccggccgccacccactggatggcagagcacagcgattcatgttggca
catccacctgtccagaaatgcttggtggccattcttcaaaatccacaagttggttgaaaaacaatctttgtttt
ataaagcaggagaaactgatgcatctagaacctttccaaacgtccagttagtgatcaagtgttggtgtgcctga
```

```
ctccatttctgacccttcctgtgtggtctttagaagggaattggatcctagagtccggccgccccttcaccga
gggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattggaattaatt
tgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagtt
ttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttata
tatcttgtggaaaggacgaaacaccgagttctgtaattcagcatagtttcagagctatgctggaaacagcatag
caagttgaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgctttttttgaattgatcc
atatatagggcccgggttataattacctcaggtcgacgtcccatgtgcaggtgctgaattgtagataagtagca
tggcgggttaatcattaactacaaggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgc
tcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcg
cgcag
```

Map and nucleotide sequence of selector AAV donor construct BI19_pAAV-HR$^{LMN.A1}$. Diagram of AAV-HR$^{LMN.A1}$ vector genome containing an *LMNA*-targeting HR donor template (1762 bp). The selector donor DNA sequence (964 bp) is designed for installing the ouabain resistance gain-of-function Q118R and N129D SNPs (RD) within the exon 4 of *ATP1A1* upon gRNA G^{A1}-directed cleavage within intron 4. AAV ITR, adeno-associated virus type-2 inverted terminal repeat; DNA sequences homologous to target *LMNA* alleles (homology arms) flank the mScarlet-1 reporter coding sequence (red arrow). The gRNA^{LMNA} directs HR-mediated *LMNA* gene tagging upon targeted DNA cleavage at the N-terminus of *LMNA* alleles. U6 promoter, human U6 snRNA polymerase III promoter driving the expression of each gRNA. The vector plasmid backbone is not shown.



>BI38_pAAV-HR$^{A1.IN17}$ (4014 bp)

```
ctggcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcc
tcagtgagcgagcgagcgcgcagagagggagtggccaactccatcactaggggttccttgtagttaatgatta
acccgccatgctacttatctacgtggccactagtacttctcgagctctacgtagaattctctagagtccggcc
gccccttcaccgagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgttagagaga
taattggaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaataatttct
tgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttc
gatttcttggctttatatatcttgtggaaaggacgaaacaccgtcacagatcgatagtagtggtttcagagct
atgctggaaacagcatagcaagttgaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggt
gctttttttgaattcggtacgagcgctgctaggcaattccacccaccagagtcattactttgagtagctcttg
ttaatgctgggggctatgtttgttgtcacttctcagttctgttatttggtgtagggcctgtgtgaatacttgc
ctgtgacggttctcaggcttcataaatagtctcaataggaaaggagcagtgtctgtaatgagtgctcagtggg
ggcatgcatcgcactatttccatcgctaggaaaagtgattggtattaacccgtttttcccttttctgggtagg
gtgctatcgtggctgtgactggtgacggtgtgaatgactctccagctttgaagaaagcagacattggggttgc
tatggggattgctggctcagatgtgtccaagcaagctgctgacatgattcttctggatgacaactttgcctca
attgtgactggagtagaggaaggtgagagctatttaaggtgtacaccaagatcttattcagatactgcccatt
agcatccatttctgtatacttcttggatatgttcagtttccagtgtgcttgtctcataagctaacagtaaaaa
atcttggtttttcataggtcgtctgatctttgataacttgaagaaatccattgcttataccttaaccagtaaca
ttcccgagatcaccccgttcctgatatttattattgcaaacattccactaccactggggaatgtcaccatcct
ctgcattgacttgggcactgacatggtgagtgtcacaacagtcacagatcgatatcagacgtgatatgtcgac
ccatagagcccaccgcatccccagcatgcctgctattgtcttcccaatcctcccccttgctgtcctgccccac
cccacccccccagaatagaatgacacctactcagacaatgcgatgcaatttcctcattttattaggaaaggaca
gtgggagtggcaccttccagggtcaaggaaggcacgggggaggggcaaacaacagatggctggcaactagaag
gcacagcctgcagggtttaaacgcggccgctcgagttatcactttttctttttgcctggccggcctttttcg
tggccgccggccttttcttgtacagctcgtccatgccgccggtggagtggcggccctcggagcgttcgtactg
ttccaccacggtgtagtcctcgttgtgggaggtgatgtccaacttgcggtcgacgttgtaggcgccgggcatc
tgcacgggcttcttggccttgtaggtggtcttgaagtccgccaggtagcggccgccgtccttcaggcgcaggg
ccatcttaatgtcgcccttcagcacgccgtcctcggggtacaaccgctcggtggatgcttcccagcccattgt
cttcttctgcattacggggccgtcaggagggaagttggtgccgcggagcttcaccttgtagatcagggtgccg
tcctccagggaggtgtcctgggtcacggtcacggcgccgccgtcctcgaagttcatcacgcgctcccacttga
agccctcggggaaggactgcttatagtagtcggggatgtcggcggggtgcttgatgaaggccctggagccgta
catgaactgaggggacaggatgtcccaggagaagggcagggggccacccttggtcaccttcagcttggcggtc
tgggtgccctcgtagggggcggccctcgccctcgccctcgatctcgaactcgtggccgttcatggagccctcca
tgtgcaccttgaaccgcatgaactccttgatcactgcctcgcccttgctcaccatggctccctgaaaatacag
attctcggcggccgccttaaggctgagggtacccctggggagagaggtcggtgattcggtcaacgagggagcc
gactgccgacgtgcgctccggaggcttgcagaatgcggaacaccgcgcgggcaggaacagggcccacactacc
gccccacaccccgcctcccgcaccgccccttcccggccgctgctctcggcgcgccctgctgagcagccgctat
tggccacagcccatcgcggtcggcgcgctgccattgctccctggcgctgtccgtctgcgagggtactagtgag
acgtgcggcttccgtttgtcacgtccggcacgccgcgaaccgcaaggaaccttcccgacttaggggcggagca
ggaagcgtcgccgggggcccacaagggtagcggcgaagatccgggtgacgctgcgaacggacgtgaagaatg
tgcgagacccagggtcggcgccgctgcgtttcccggaaccacgcccagagcagccgcgtccctgcgcaaaccc
agggctgccttggaaaaggcgcaaccccaaccccgtggatgcattaaaaacctcccacacctcccccctgaac
ctgaaacaagcttgtgaggtgtgagctgtgtcttcattcactggcactatgctcccagcatccatgaggctgg
```

```
cgttgggaaaagaaattctcaggaccagtatccagtgtgtgtcccaatcccggcttcacagaatcagtattac
cactcttcaaggcagcagttaacatttgtttatcagatgggaatctttatttttgtatacttcacctaaaaga
tttttaaatggagggcaagaattttataacaaaaggttcacaatattagcttccttatttttagtaactaaa
ttccttctccccaccccttcccaggttcctgccatctccctggcttatgagcaggctgagagtgacatcatga
agagacagcccagaaatcccaaaacagacaaacttgtgaatgagcggctgatcagcatggcctatgggcagat
tggtaagctgcagcctggagtgggaagctggcacatctaaggcatctgaggtgatggtgtccacctcaggttg
aggttgatttcagagactgcaaatccaggcgactttcaggtctaggatgagccctaacggagtgagcctgtgg
agtttctctgaactctcttccatgtgagaatacatctgttgctgtttgcccagttaccttttgtggaaactgg
ttttctatactgagcaccttggaactgggctcaactctgcctgagaaatgaggatgtccttcctcactgagcc
ttgcggaacaggaggcatggttctaagggctagtcctagccacgtgaccggttcgcgaggatccatatatagg
gcccgggttataattacctcaggtcgacgtcccatgtgcaggtgctgaattgtagataagtagcatggcgggt
taatcattaactacaaggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactga
ggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcgcag
```

Map and nucleotide sequence of selector AAV donor construct BI38_pAAV-HR$^{\mathrm{A1.IN17}}$. Diagram of AAV-HR$^{\mathrm{A1.IN17}}$ vector genome containing an *ATP1A1*-targeting HR donor template. The selector donor DNA sequence (3129 bp bp) is designed for concomitant HR-mediated installation of a transgene and an ouabain resistance gain-of-function SPN (T804N) within exon 17 and intron 17 of *ATP1A1* alleles, respectively. The matched gRNA G$^{\mathrm{A1.IN17}}$ directs DNA cleavage at the ATP1A1 intron 17. U6 promoter, human U6 snRNA polymerase III promoter driving the expression of gRNA$^{\mathrm{A1.IN17}}$ AAV ITR, adeno-associated virus type-2 inverted terminal repeat; DNA sequences homologous to target *ATP1A1* alleles (homology arms) flank a transgene consisting of the human phosphoglycerate kinase 1 gene promoter (hPGK); the mScarlet-1 reporter coding sequence (red arrow), and the bovine growth hormone gene polyadenylation signal (bGH pA). The vector plasmid backbone is not shown.

References

1. Pacesa M, Pelea O, Jinek M. Past, present, and future of CRISPR genome editing technologies. Cell 2024 187:1076-1100.
2. Hongyu Liao, Jiahao Wu, Nathan J. VanDusen, Yifei Li, Yanjiang Zheng CRISPR/Cas9-mediated homology-directed repair for precise gene editing. Mol Ther Nucleic Acids 2024 DOI: 10.1016/j.omtn.2024.102344.
3. He X., Tan C., Wang F., Wang Y., Zhou R., Cui D., You W., Zhao H., Ren J., Feng B. Knockin of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. Nucleic Acids Res. 2016; 44:e85.
4. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, Hatanaka F, Yamamoto M, Araoka T, Li Z, Kurita M, Hishida T, Li M, Aizawa E, Guo S, Chen S, Goebl A, Soligalla RD, Qu J, Jiang T, Fu X, Jafari M, Esteban CR, Berggren WT, Lajara J, Nuñez-Delicado E, Guillen P, Campistol JM, Matsuzaki F, Liu GH, Magistretti P,

Zhang K, Callaway EM, Zhang K, Belmonte JC. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature 2016 540:144-149.

5. Bischoff N, Wimberger S, Maresca M, Brakebusch C. Improving Precise CRISPR Genome Editing by Small Molecules: Is there a Magic Potion? Cells 2020 9:1318.
6. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, Kühn R. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat. Biotechnol. 2015 33:543-548.
7. Wimberger S, Akrap N, Firth M, Brengdahl J, Engberg S, Schwinn MK, Slater MR, Lundin A, Hsieh PP, Li S, Cerboni S, Sumner J, Bestas B, Schiffthaler B, Magnusson B, Di Castro S, Iyer P, Bohlooly-Y M, Machleidt T, Rees S, Engkvist O, Norris T, Cadogan EB, Forment JV, Šviković S, Akcakaya P, Taheri-Ghahfarokhi A, Maresca M. Simultaneous inhibition of DNA-PK and PolΘ improves integration efficiency and precision of genome editing. Nat. Commun. 2023 14:4761.
8. Schimmel J, Muñoz-Subirana N, Kool H, van Schendel R, van der Vlies S, Kamp JA, de Vrij FMS, Kushner SA, Smith GCM, Boulton SJ, Tijsterman M. Modulating mutational outcomes and improving precise gene editing at CRISPR-Cas9-induced breaks by chemical inhibition of end-joining pathways. Cell Rep. 2023 42:112019.
9. Mikkelsen NS, Bak RO. Enrichment strategies to enhance genome editing. J Biomed Sci. 2023 30:51.
10. Shy BR, MacDougall MS, Clarke R, Merrill BJ. Co-incident insertion enables high efficiency genome engineering in mouse embryonic stem cells. Nucleic Acids Res. 2016 44:7997-8010.
11. Mitzelfelt KA, McDermott-Roe C, Grzybowski MN, Marquez M, Kuo CT, Riedel M, Lai S, Choi MJ, Kolander KD, Helbling D, Dimmock DP, Battle MA, Jou CJ, Tristani-Firouzi M, Verbsky JW, Benjamin IJ, Geurts AM. Efficient Precision Genome Editing in iPSCs via Genetic Co-targeting with Selection. Stem Cell Reports 2017 8:491-499.
12. Agudelo D, Duringer A, Bozoyan L, Huard CC, Carter S, Loehr J, Synodinou D, Drouin M, Salsman J, Dellaire G, Laganière J, Doyon Y. Marker-free coselection for CRISPR-driven genome editing in human cells. Nat Methods 2017 14:615-620.
13. Wiebking V, Patterson JO, Martin R, Chanda MK, Lee CM, Srifa W, Bao G, Porteus MH. Metabolic engineering generates a transgene-free safety switch for cell therapy. Nat. Biotechnol. 2020 38:1441-1450.
14. Li S, Akrap N, Cerboni S, Porritt MJ, Wimberger S, Lundin A, Möller C, Firth M, Gordon E, Lazovic B, Sieńska A, Pane LS, Coelho MA, Ciotta G, Pellegrini G, Sini M, Xu X, Mitra S, Bohlooly-Y M, Taylor BJM, Sienski G, Maresca M. Universal toxin-

based selection for precise genome engineering in human cells. Nat Commun. 2021 12:497.

15. Wu J, Li D, Du L, Baldawi M, Gable ME, Askari A, Liu L. Ouabain prevents pathological cardiac hypertrophy and heart failure through activation of phosphoinositide 3-kinase α in mouse. Cell Biosci. 2015 5:64.
16. Levesque S, Mayorga D, Fiset JP, Goupil C, Duringer A, Loiselle A, Bouchard E, Agudelo D, Doyon Y. Marker-free co-selection for successive rounds of prime editing in human cells. Nat Commun. 2022 13:5909.
17. Epstein BE, Schaffer DV. Combining Engineered Nucleases with Adeno-associated Viral Vectors for Therapeutic Gene Editing. Adv Exp Med Biol. 2017;1016:29-42.
18. Miller DG, Petek LM, Russell DW. Adeno-associated virus vectors integrate at chromosome breakage sites. Nat Genet. 2004 36:767-773.
19. Hanlon KS, Kleinstiver BP, Garcia SP, Zaborowski MP, Volak A, Spirig SE, Muller A, Sousa AA, Tsai SQ, Bengtsson NE, Lööv C, Ingelsson M, Chamberlain JS, Corey DP, Aryee MJ, Joung JK, Breakefield XO, Maguire CA, György B. High levels of AAV vector integration into CRISPR-induced DNA breaks. Nat Commun. 2019 10:4439.
20. Ferrari S, Jacob A, Cesana D, Laugel M, Beretta S, Varesi A, Unali G, Conti A, Canarutto D, Albano L, Calabria A, Vavassori V, Cipriani C, Castiello MC, Esposito S, Brombin C, Cugnata F, Adjali O, Ayuso E, Merelli I, Villa A, Di Micco R, Kajaste-Rudnitski A, Montini E, Penaud-Budloo M, Naldini L. Choice of template delivery mitigates the genotoxic risk and adverse impact of editing in human hematopoietic stem cells. Cell Stem Cell. 2022 29:1428-1444.e9.
21. Li Z, Wang X, Janssen JM, Liu J, Tasca F, Hoeben RC, Gonçalves MAFV. Precision genome editing using combinatorial viral vector delivery of CRISPR-Cas9 nucleases and donor DNA constructs. Under review.
22. Suchy FP, Karigane D, Nakauchi Y, Higuchi M, Zhang J, Pekrun K, Hsu I, Fan AC, Nishimura T, Charlesworth CT, Bhadury J, Nishimura T, Wilkinson AC, Kay MA, Majeti R, Nakauchi H. Nat Biotechnol. 2024. doi: 10.1038/s41587-024-02171-w. Online ahead of print.
23. Flotte TR, Solow R, Owens RA, Afione S, Zeitlin PL, Carter BJ. Gene expression from adeno-associated virus vectors in airway epithelial cells. Am. J. Respir. Cell. Mol. Biol. 1992 7:349-356.
24. Haberman RP, McCown TJ, Samulski RJ. Novel transcriptional regulatory signals in the adeno-associated virus terminal repeat A/D junction element. J. Virol. 2000 74:8732-8739.
25. Bazick HO, Mao H, Niehaus JK, Wolter JM, Zylka MJ. AAV vector-derived elements

integrate into Cas9-generated double-strand breaks and disrupt gene transcription. Mol. Ther. 2024 S1525-0016(24)00656-7. doi: 10.1016/j.ymthe.2024.09.032. Online ahead of print

26. Schiroli G, Conti A, Ferrari S, Della Volpe L, Jacob A, Albano L, Beretta S, Calabria A, Vavassori V, Gasparini P, Salataj E, Ndiaye-Lobry D, Brombin C, Chaumeil J, Montini E, Merelli I, Genovese P, Naldini L, Di Micco R. Precise Gene Editing Preserves Hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. Cell Stem Cell. 2019 24:551-565.e8.
27. Allen D, Weiss LE, Saguy A, Rosenberg M, Iancu O, Matalon O, Lee C, Beider K, Nagler A, Shechtman Y, Hendel A. High-Throughput Imaging of CRISPR- and Recombinant Adeno-Associated Virus-Induced DNA Damage Response in Human Hematopoietic Stem and Progenitor Cells. CRISPR J. 2022 5:80-94.
28. Maggio I, Gonçalves MA. Genome editing at the crossroads of delivery, specificity, and fidelity. Trends Biotechnol. 2015 33:280-291.
29. Pavani G, Amendola M. Targeted Gene Delivery: Where to Land. Front Genome Ed. 2021 2:609650. Corrigendum: Front Genome Ed. 2021 3:682171.
30. Spector LP, Tiffany M, Ferraro NM, Abell NS, Montgomery SB, Kay MA. Evaluating the Genomic Parameters Governing rAAV-Mediated Homologous Recombination. Mol Ther. 2021 29:1028-1046.
31. Bijlani S, Pang KM, Sivanandam V, Singh A, Chatterjee S. The Role of Recombinant AAV in Precise Genome Editing. Front Genome Ed. 2022 3:799722.
32. Holkers M, de Vries AA, Gonçalves MA. Nonspaced inverted DNA repeats are preferential targets for homology-directed gene repair in mammalian cells. Nucleic Acids Res. 2012 40:1984-1999.
33. Holkers M, Maggio I, Henriques SF, Janssen JM, Cathomen T, Gonçalves MA. Adenoviral vector DNA for accurate genome editing with engineered nucleases. Nat Methods. 2014 11:1051-1057.
34. Medert R, Thumberger T, Tavhelidse-Suck T, Hub T, Kellner T, Oguchi Y, Dlugosz S, Zimmermann F, Wittbrodt J, Freichel M. Efficient single copy integration via homology-directed repair (scHDR) by 5' modification of large DNA donor fragments in mice. Nucleic Acids Res. 2023 51:e14.
35. Brescia, M., *et al.*, High-Capacity Adenoviral Vectors Permit Robust and Versatile Testing of DMD Gene Repair Tools and Strategies in Human Cells. Cells, 2020 9: 869.
36. Tasca F., Wang Q., Goncalves M Adenoviral vectors meet gene editing: a rising partnership for the genomic engineering of human stem cells and their progeny. Cells. 2020; 9:953.

37. 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.
38. Wang Q, Liu J, Janssen JM, Le Bouteiller M, Frock RL, Gonçalves MAFV. Precise and broad scope genome editing based on high-specificity Cas9 nickases. Nucleic Acids Res. 2021 49:1173-1198.
39. Dever DP, Bak RO, Reinisch A, Camarena J, Washington G, Nicolas CE, Pavel-Dinu M, Saxena N, Wilkens AB, Mantri S, Uchida N, Hendel A, Narla A, Majeti R, Weinberg KI, Porteus MH. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. Nature. 2016 539:384-389.
40. Pavani G, Fabiano A, Laurent M, Amor F, Cantelli E, Chalumeau A, Maule G, Tachtsidi A, Concordet JP, Cereseto A, Mavilio F, Ferrari G, Miccio A, Amendola M. Correction of β-thalassemia by CRISPR/Cas9 editing of the α-globin locus in human hematopoietic stem cells. Blood Adv. 2021 5:1137-1153.
41. McColl-Carboni A, Dollive S, Laughlin S, Lushi R, MacArthur M, Zhou S, Gagnon J, Smith CA, Burnham B, Horton R, Lata D, Uga B, Natu K, Michel E, Slater C, DaSilva E, Bruccoleri R, Kelly T, McGivney JB 4th. Analytical characterization of full, intermediate, and empty AAV capsids. Gene Ther. 2024 31:285-294.
42. Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. Nat. Biotechnol. 2015 33:179-186.
43. 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.
44. Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF, Neri M, Magnani Z, Cantore A, Lo Riso P, Damo M, Pello OM, Holmes MC, Gregory PD, Gritti A, Broccoli V, Bonini C, Naldini L. Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat. Methods 2011 8:861-869.
45. Ihry RJ, Worringer KA, Salick MR, Frias E, Ho D, Theriault K, Kommineni S, Chen J, Sondey M, Ye C, Randhawa R, Kulkarni T, Yang Z, McAllister G, Russ C, Reece-Hoyes J, Forrester W, Hoffman GR, Dolmetsch R, Kaykas A. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. Nat Med. 2018 24:939-946.
46. Chen X, Tasca F, Wang Q, Liu J, Janssen JM, Brescia MD, Bellin M, Szuhai K, Kenrick J, Frock RL, Gonçalves MAFV. 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.

47. Wang Q, Liu J, Janssen JM, Gonçalves MAFV. Precise homology-directed installation of large genomic edits in human cells with cleaving and nicking high-specificity Cas9 variants. Nucleic Acids Res. 2023 51:3465-3484.

Chapter 6

Summary and general discussion

Genome editing (*aliases,* genomic engineering and gene editing) is a fast-paced field with increasing impact on fundamental science, biotechnology, and medicine [1]. With the aid of engineered RNA-guided nucleases (RGNs) derived from clustered regularly interspaced short palindromic repeat (CRISPR) systems initially discovered as prokaryotic antiviral machineries in bacteria and archaea [2], genome editing has become more versatile and customizable with regards to the options for achieving gene knock-out (KO), gene knock-in (KI), targeted DNA replacement and tagging, amongst other chromosomal DNA modification endpoints [1,2]. Despite the increasing discovery and adaption of new CRISPR and CRISPR-like systems for genome editing purposes [3], engineered CRISPR-derived RGNs based on the prototypic *Streptococcus pyogenes* CRISPR-Cas9 system, and their variants created by directed evolution or rational design (*e.g.*, high-specificity and targeting range-expanded variants), remain commonly used reagents for a broad range of genome engineering applications [4]. This stems in large part from their relative robustness in engaging eukaryotic chromatin [5].

Yet, although nuclease-induced double-stranded DNA break (DSB) formation yields robust and cell cycle-independent gene KO endpoints upon the installation of small insertions and deletions (indels) by canonical and alternative non-homologous end joining (NHEJ) pathways, these DNA repair pathways are disruptive to cell genotypes in the context of gene KI procedures based on homology-directed repair (HDR) [6]. Moreover, the possibility for RGN off-target and, more pervasively, on-target activities and ensuing mutagenic effects, which intrinsically stem from NHEJ processes, demands the exploration of more precise and less mutagenic ('soft') genome editing strategies. An emerging class of such DSB-independent strategies is reviewed in **Chapter 2** where sequence- and site-specific nucleases ('nickases') derived from CRISPR systems are exploited for precise HDR-mediated gene KI using tailored donor DNA substrates. Indeed, in contrast to DSBs, single-stranded DNA breaks (SSBs), or nicks, are not substrates for mutagenic NHEJ DNA repair pathways, canonical or otherwise. As corollary, genomic engineering using nickases is dramatically biased towards precise HDR over imprecise NHEJ events. Of note, although genomic SSBs are *per se* weak HDR stimuli, earlier experiments demonstrated that concomitant SSB formation at acceptor genomic sequences and donor DNA substrates by CRISPR/Cas9 nickases fosters HDR-mediated gene KI [7-9]. The application of this generic *in trans* paired nicking principle yields genomic engineering of stem cells with minimal allelic indels, translocations, and P53-dependent activation of the DNA damage response (DDR) known to trigger apoptosis or, alternatively, halt the cell cycle progression needed for HDR [7,9]. Therefore, as pointed out in **Chapter 2**, cell therapy products derived from using RNA-programmable nickases as such or coupled to heterologous effector domains, like prime editor reverse transcriptases or base editor deaminases, will start to increasingly offer a complementary set of 'soft' genome editing options whose performances and safety profiles are potentially higher than those resulting from exposing cells to programmable nucleases.

In **Chapter 3**, the relevance of DSB-independent genome editing efforts is further elaborated through a commentary to a study by Chai and colleagues [10]. In this study, base editors based on Cas9 nickases fused to deaminase effector domains are tested for the purpose of restoring dystrophin synthesis in *in vitro* and *in vivo* models of Duchenne muscular dystrophy (DMD; MIM: 310200). In particular, the authors identify adenine base editors and cognate guide RNAs that, by installing A·T-to-G·C transitions at splicing motifs, lead to defective *DMD* reading frame repair via exon skipping and subsequent removal of premature stop codons that arose due to out-of-frame deletions. Indeed, by using dual AAV vectors for the *trans*-splicing assembly of complete base editor proteins, the authors demonstrate the rescue of dystrophic traits at the cellular and organismal levels in dystrophin-defective mice upon dual AAV vector administrations. Specifically, base editor constructs linked to N- and C-terminal intein domains were split and packaged in two distinct AAV vectors that, upon target cell co-transductions, led to intein-mediated protein *trans*-splicing and *in situ* assembly of full-length base editor proteins. This work departs from earlier studies based on dual AAV delivery of RGNs where potentially deleterious outcomes in the form of prevalent end-joining ‘capture’ of Cas9-encoding AAV genomes at nuclease target sites were detected in various murine tissues, including skeletal muscle [11,12].

It is consensual the view that a key bottleneck regarding the application of genome editing technologies concerns the difficulty in delivering the attendant tools in an effective manner, especially so into regular, non-transformed and primary cells including therapeutically relevant cell types. In this context, physical and chemical transfection methods based on, for instance, electroporation, polyplexes and polycations, permit introducing RGNs and donor DNA constructs into mammalian cells. However, achieving optimal transfection efficiencies without noticeable cytotoxic effects in the aforementioned primary cell types, remains challenging. Moreover, neither electroporation nor transfections of RGNs and/or donor DNA substrates are easily amenable to *in vivo* settings. And, often, these gene-editing tool delivery methods depend on cell type-specific reagents whose compositions may be unknown due to proprietary reasons. Equally of note, electroporation and transfections are further reliant on time-consuming optimizations of cell type-specific parameters whose performances can vary due to subtle experimental conditions, *e.g.* cell-cycle stage profiles of target cell populations. Diversely, viral vector transductions are straightforward to perform and typically offer higher reproducibility independently of the target cell type of choice [13]. These characteristics stem from the fine-tuned mechanisms evolved by the parental viruses for nuclei acid transfer into host cell nuclei.

In **Chapter 4**, inspired by the complementary attributes of two distinct classes of viral gene-free viral vectors, namely, high-capacity adenoviral vectors (AdV) and AAV vectors (*e.g.*, large cargo compatibility and source of recombinogenic HDR substrates, respectively), a dual viral vector system is introduced, characterized, and tested for HDR-mediated gene KI purposes. In this dual viral vector genome-editing platform, RGN constructs are packaged in AdV capsids; whilst donor DNA templates are placed in AAV capsids.

Next to genome-editing tool delivery considerations linked to the ultimate efficiency of gene KI procedures, other parameters to consider regard the specificity and precision of donor DNA insertion. The specificity results from detecting donor DNA sequences at the target site; whilst the precision can be assessed by establishing the precedence of accurate HDR-dependent DNA insertions over imprecise NHEJ-derived events. In this regard, knowledge concerning the relative contributions of distinct AAV donor designs (*i.e.*, conventional versus ‘double-cut’) and structures (*i.e.*, single-stranded versus double-stranded) to the genome-editing specificity and precision parameters, is scant. Hence, besides achieving remarkably high gene KI efficiencies in transformed, non-transformed and primary cells, the dual viral vector platform studied in **Chapter 4** served as a probe to study the relationship between the aforementioned AAV DNA arrangements and the specificity and precision of gene KI and tagging procedures. This research disclosed that combining single-stranded AAV delivery of HDR donors with high-capacity AdV transfer of RGN constructs leads to precise genome editing in large fractions of target-cell populations. And, it revealed that RGN-induced chromosomal DNA breaks promote productive AAV vector transduction as scored by transgene expression presumably owing to fostering single-stranded to transcriptionally-competent double-stranded DNA conversion. Finally, in **Chapter 4**, it is further disclosed the critical importance of using high-fidelity RGNs for minimizing off-target donor AAV insertions in the form of defective vector genomes known to be packaged in vector particles [14,15].

The aforementioned recombinogenic character of AAV vector genomes, that makes them proficient HDR substrates, also contributes to their promiscuous “capture” at on-target and off-target or random chromosomal DNA breaks through imprecise NHEJ processes [11,12,16,17]. These end-joining processes can further yield chromosomal insertion of concatemeric and sub-genomic AAV species further compounding the range of HDR-independent bystander events [11,12,17,18]. Finally, possibly due to mimicking DNA lesions or repair intermediates, AAV genomes can impair cell viability through P53-dependent DDR activation whose consequences have been reported to be particularly deleterious in stem cells [19,20]. In conclusion, a growing amount of research indicates that there are distinctive genotoxic and cytotoxic effects linked to chromosomally integrated and episomal AAV DNA forms, respectively, with the build-up of these deleterious effects being strictly proportional to input AAV vector amounts.

Hence, with the aim of tackling the shortcomings associated with AAV-based genome editing, in **Chapter 5**, a marker-free co-selection system [21], dependent on the potent specific inhibitor of the Na^+/K^+ ATPase pump encoded by *ATP1A1*, ouabain, was co-opted and investigated. To this end, AAV donor vectors endowed with a secondary *ATP1A1*-selectable donor module and matched gRNA units were assembled and tested. Such AAV designs, dubbed selector AAV vectors, seek enriching for cells coedited through at a primary target locus and *ATP1A1* alleles that confer resistance to ouabain upon HDR-mediated acquisition of specific polymorphisms. Importantly, besides enriching for gene KI cell populations, combining selector AAV vectors with ouabain selection triggered the elimination of HDR-independent edits and off-target and/or random AAV donor DNA insertions from said populations. This selector AAV principle was successfully applied for inserting whole transgenes at safe harbor genomic loci as well as for tagging endogenous proteins. Further, selector AAV vector titration experiments revealed that the highest fold-enrichment factors of gene-targeted cell fractions were associated with the lowest vector input amounts. This feature might become beneficial for alleviating both AAV production costs and detrimental P53-dependent DDR activation. Of note, the dual viral vector platform described in **Chapter 4** was put to use for streamlining and expanding selector AAV vector testing into hard-to-transfect primary cells.

As covered in **Chapter 2** and **Chapter 3**, there is a growing realization that, especially in DNA damage sensitive stem cells, programmable nuclease-induced DSBs are detrimental to target loci stability [22,23], and cell viability [7,19]. Significantly, when compared with Cas9 nucleases, Cas9^{D10A} nickases are notoriously less mutagenic at both on-target and off-target genomic sequences as determined by reporter and unbiased genome-wide high-throughput assays, respectively [24,25]. Cas9^{D10A} nickases further display greatly reduced P53-dependent DDR activation when tested side-by-side with their Cas9 nuclease counterparts [9,26]. In this context, transduction experiments described in **Chapter 5** have shown that combining a particular selector AAV design, dubbed in-linkage selector AAV, with Cas9 or Cas9^{D10A} exposure leads to a significant increase in the frequencies of gene-targeted cells. The ouabain-dependent enrichment factor for gene-targeted cells was *circa* 7.5-fold higher when using Cas9^{D10A} nickase delivery. Importantly, combining in-linkage selector AAV transduction with Cas9 nuclease or Cas9^{D10A} nickase delivery led to the thorough elimination of donor DNA inserted at off-target genomic positions from genome-edited cell populations. Therefore, these proof-of-principle experiments indicating that selector AAV-based genome editing is transportable to protocols involving Cas9^{D10A}-induced HDR are relevant for furthering the refinement and application of DSB-independent cell engineering strategies.

In conclusion, through the investigation and integration of distinct viral vectors and RNA-programmable Cas9 proteins, the work presented in this thesis provides new insights and toolboxes for advancing genome editing through the improvement of performance aspects related to its efficiency, specificity and fidelity.

References

1. Pacesa M, Pelea O, Jinek M. Past, present, and future of CRISPR genome editing technologies. Cell 2024 187:1076-1100.
2. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 2014 346:1258096.
3. Koonin EV, Makarova KS, Zhang F. Diversity, classification and evolution of CRISPR-Cas systems. Curr Opin Microbiol. 2017 37:67-78.
4. Chen X, Gonçalves MAFV. DNA, RNA, and Protein Tools for Editing the Genetic Information in Human Cells. iScience 2018 6:247-263.
5. Chen X, Rinsma M, Janssen JM, Liu J, Maggio I, Gonçalves MA. Probing the impact of chromatin conformation on genome editing tools. Nucleic Acids Res. 2016 44:6482-92.
6. Liao H, Wu J, VanDusen NJ, Li Y, Zheng Y. CRISPR-Cas9-mediated homology-directed repair for precise gene editing. Mol Ther Nucleic Acids. 2024 35:102344.
7. Chen X, Janssen JM, Liu J, Maggio I, 't Jong AEJ, Mikkers HMM, Gonçalves MAFV. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. Nat Commun. 2017 8:657.
8. Nakajima K, Zhou Y, Tomita A, Hirade Y, Gurumurthy CB, Nakada S. Precise and efficient nucleotide substitution near genomic nick via noncanonical homology-directed repair. Genome Res. 2018 28:223-230.
9. Wang Q, Liu J, Janssen JM, Gonçalves MAFV. Precise homology-directed installation of large genomic edits in human cells with cleaving and nicking high-specificity Cas9 variants. Nucleic Acids Res. 2023 51:3465-3484.
10. Chai AC, Chemello F, Li H, Nishiyama T, Chen K, Zhang Y, Sánchez-Ortiz E, Alomar A, Xu L, Liu N, Bassel-Duby R, Olson EN. Single-swap editing for the correction of common Duchenne muscular dystrophy mutations. Mol Ther Nucleic Acids. 2023 32:522-535.
11. Hanlon KS, Kleinstiver BP, Garcia SP, Zaborowski MP, Volak A, Spirig SE, Muller A, Sousa AA, Tsai SQ, Bengtsson NE, Lööv C, Ingelsson M, Chamberlain JS, Corey DP, Aryee MJ, Joung JK, Breakefield XO, Maguire CA, György B. High levels of AAV vector integration into CRISPR-induced DNA breaks. Nat Commun. 2019 10:4439.
12. Nelson CE, Wu Y, Gemberling MP, Oliver ML, Waller MA, Bohning JD, Robinson-Hamm JN, Bulaklak K, Castellanos Rivera RM, Collier JH, Asokan A, Gersbach CA. Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. Nat Med. 2019 25:427-432.
13. Chen X, Gonçalves MA. Engineered Viruses as Genome Editing Devices. Mol Ther. 2016 24:447-457.

14. Wada M, Uchida N, Posadas-Herrera G, Hayashita-Kinoh H, Tsunekawa Y, Hirai Y, Okada T. Large-scale purification of functional AAV particles packaging the full genome using short-term ultracentrifugation with a zonal rotor. Gene Ther. 2023 30:641-648.
15. McColl-Carboni A, Dollive S, Laughlin S, Lushi R, MacArthur M, Zhou S, Gagnon J, Smith CA, Burnham B, Horton R, Lata D, Uga B, Natu K, Michel E, Slater C, DaSilva E, Bruccoleri R, Kelly T, McGivney JB 4th. Analytical characterization of full, intermediate, and empty AAV capsids. Gene Ther. 2024 31:285-294.
16. Miller DG, Petek LM, Russell DW. Adeno-associated virus vectors integrate at chromosome breakage sites. Nat Genet. 2004 36:767-773.
17. Ferrari S, Jacob A, Cesana D, Laugel M, Beretta S, Varesi A, Unali G, Conti A, Canarutto D, Albano L, Calabria A, Vavassori V, Cipriani C, Castiello MC, Esposito S, Brombin C, Cugnata F, Adjali O, Ayuso E, Merelli I, Villa A, Di Micco R, Kajaste-Rudnitski A, Montini E, Penaud-Budloo M, Naldini L. Choice of template delivery mitigates the genotoxic risk and adverse impact of editing in human hematopoietic stem cells. Cell Stem Cell 2022 29:1428-1444.
18. Suchy FP, Karigane D, Nakauchi Y, Higuchi M, Zhang J, Pekrun K, Hsu I, Fan AC, Nishimura T, Charlesworth CT, Bhadury J, Nishimura T, Wilkinson AC, Kay MA, Majeti R, Nakauchi H. Genome engineering with Cas9 and AAV repair templates generates frequent concatemeric insertions of viral vectors. Nat Biotechnol. 2025 43:204-213.
19. Schiroli G, Conti A, Ferrari S, Della Volpe L, Jacob A, Albano L, Beretta S, Calabria A, Vavassori V, Gasparini P, Salataj E, Ndiaye-Lobry D, Brombin C, Chaumeil J, Montini E, Merelli I, Genovese P, Naldini L, Di Micco R. Precise Gene Editing Preserves Hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. Cell Stem Cell. 2019 24:551-565.
20. Allen D, Weiss LE, Saguy A, Rosenberg M, Iancu O, Matalon O, Lee C, Beider K, Nagler A, Shechtman Y, Hendel A. High-Throughput Imaging of CRISPR- and Recombinant Adeno-Associated Virus-Induced DNA Damage Response in Human Hematopoietic Stem and Progenitor Cells. CRISPR J. 2022 5:80-94.
21. Agudelo D, Duringer A, Bozoyan L, Huard CC, Carter S, Loehr J, Synodinou D, Drouin M, Salsman J, Dellaire G, Laganière J, Doyon Y. Marker-free coselection for CRISPR-driven genome editing in human cells. Nat Methods. 2017 :615-620.
22. Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. Nat Biotechnol. 2015 33:179-186.

23. 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.
24. Chen X, Tasca F, Wang Q, Liu J, Janssen JM, Brescia MD, Bellin M, Szuhai K, Kenrick J, Frock RL, Gonçalves MAFV. 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.
25. Wang Q, Liu J, Janssen JM, Le Bouteiller M, Frock RL, Gonçalves MAFV. Precise and broad scope genome editing based on high-specificity Cas9 nickases. Nucleic Acids Res. 2021 49:1173-1198.
26. Hyodo T, Rahman ML, Karnan S, Ito T, Toyoda A, Ota A, Wahiduzzaman M, Tsuzuki S, Okada Y, Hosokawa Y, Konishi H. Tandem Paired Nicking Promotes Precise Genome Editing with Scarce Interference by p53. Cell Rep. 2020 30:1195-1207.

Chapter 7

Nederlandse Samenvatting

Nederlandse samenvatting

Genome editing (ook wel genoemd: genome engineering en gene editing) is een snel ontwikkelend vakgebied met een toenemende impact op fundamentele wetenschap, biotechnologie en geneeskunde [1]. Met behulp van gemodificeerde RNA-gestuurde nucleasen (RGN's) afgeleid van 'geclusterde regelmatig interspaced korte palindromische repeat (CRISPR)-systemen', die werden ontdekt als prokaryotische antivirale machines in bacteriën en archaea [2], is genoombewerking veelzijdiger geworden voor het bereiken van gen-knock-out (KO), gen-knock-in (KI), gerichte DNA-vervanging en tagging [1,2]. Ondanks de aanpassing van nieuwe CRISPR en CRISPR-achtige systemen voor genoom-bewerkingsdoeleinden [3], blijven CRISPR-afgeleide RGN's gebaseerd op het prototypische *Streptococcus pyogenes* CRISPR-Cas9-systeem, en door gerichte evolutie gecreëerde varianten, veelgebruikte reagentia voor vele toepassingen [4]. Dit komt grotendeels voort uit hun robuustheid bij het aangaan van eukaryotisch chromatine [5].

Hoewel nuclease-geïnduceerde dubbelstrengs DNA-breuk (DSB)-vorming robuuste en celcyclus-onafhankelijke gen-KO oplevert door het ontstaan van kleine inserties en deleties (indels) door niet-homologe end-joining (NHEJ)-paden, zijn deze DNA-herstelpaden verstorend in de context van gen-KI-procedures gebaseerd op homologie-gerichte reparatie (HDR) [6]. Bovendien vereist de mogelijkheid voor RGN off-target en on-target activiteiten en daaropvolgende mutagene effecten, die voortkomen uit NHEJ-processen, het ontwikkelen van nauwkeurigere en minder mutagene genoom-bewerkingsstrategieën. Een opkomende klasse van dergelijke DSB-onafhankelijke strategieën wordt besproken **in Hoofdstuk 2**, waar sequentie- en plaats-specifieke nucleasen ('nickases') afgeleid van CRISPR-systemen worden benut voor nauwkeurige HDR-gemedieerde gen-KI met behulp van op maat gemaakte donor-DNA-substraten. In tegenstelling tot DSB's zijn enkelstrengs DNA-breuken (SSB's) of nicks geen substraten voor mutagene

NHEJ DNA-herstelpaden, canoniek of anderszins. Als gevolg hiervan is genome engineering met behulp van nickases dramatisch bevoordeeld ten opzichte van precieze HDR boven onnauwkeurige NHEJ-gebeurtenissen. Opmerkelijk is dat hoewel genomische SSB's op zichzelf zwakke HDR-stimuli zijn, eerdere experimenten hebben aangetoond dat gelijktijdige SSB-vorming bij acceptor-genoomsequenties en donor-DNA-substraten door CRISPR/Cas9-nickases HDR-gemedieerde gen-KI bevordert [7-9]. De toepassing van dit generieke *in trans*-paired nicking-principe levert genome engineering op van stamcellen met minimale indel vorming, translocaties en P53-afhankelijke activering van de DNA-schaderespons (DDR) waarvan bekend is dat deze apoptose veroorzaakt of de celcyclusprogressie stopt die nodig is voor HDR [7,9]. Daarom zullen, zoals aangegeven in **Hoofdstuk 2**, celtherapieproducten die zijn afgeleid van het gebruik van RNA-programmeerbare nickases als zodanig of gekoppeld aan heterologe effectordomeinen, zoals prime editor reverse transcriptases of base editor deaminases, steeds vaker een complementaire set van 'zachte' genoombewerkingsopties gaan bieden waarvan de prestaties en veiligheidsprofielen mogelijk hoger zijn dan die voortvloeien uit het blootstellen van cellen aan programmeerbare nucleasen.

In **Hoofdstuk 3** wordt de relevantie van DSB-onafhankelijke genoombewerkingsinspanningen verder uitgewerkt door middel van een commentaar op een studie van Chai en collega's [10]. In deze studie worden base-editors op basis van Cas9-nickases die fuseren met deaminase-effectordomeinen getest met als doel de dystrofinesynthese te herstellen in *in vitro*- en *in vivo*-modellen van Duchenne-spierdystrofie (DMD; MIM: 310200). In het bijzonder identificeren de auteurs adenine-base-editors en cognate guide-RNA's die, door A·T-naar-G·C-overgangen te installeren bij splicing-motieven, leiden tot *DMD*-leesraamreparatie via exon-skipping en daaropvolgende verwijdering van voortijdige stopcodons die ontstonden door out-of-frame-deleties. Door gebruik te maken van dubbele AAV-vectoren voor trans-splicing-assemblage van complete base-editor-eiwitten, tonen de auteurs inderdaad de redding van dystrofische eigenschappen op cellulair en organismeniveau aan in dystrofine-defecte muizen bij toediening van dubbele AAV-vectoren. Specifiek werden base-editorconstructies gekoppeld aan N- en C-terminale inteïnedomeinen gesplitst en verpakt in twee afzonderlijke AAV-vectoren die, na doelcelcotransducties, leidden tot inteïne-gemedieerde eiwittrans-splicing en in situ-assemblage van volledige base-editoreiwitten. Dit werk wijkt af van eerdere studies gebaseerd op duale AAV-afgifte van RGN's waarbij potentieel schadelijke uitkomsten in de vorm van prevalente end-joining 'capture' van Cas9-coderende AAV-genomen op nuclease-doellocaties werden gedetecteerd in verschillende muizenweefsels, waaronder skeletspieren [11,12].

Het is de consensus dat een belangrijk knelpunt met betrekking tot de toepassing van genoom-bewerkingstechnologieën de moeilijkheid betreft om de bijbehorende hulpmiddelen op een effectieve manier af te leveren, met name in normale, niet-getransformeerde, primaire cellen, waaronder therapeutisch relevante celtypen. In deze context maken fysieke en chemische transfectiemethoden gebaseerd op bijvoorbeeld elektroporatie, polyplexen en polykationen het mogelijk om RGN's en donor-DNA-constructies in zoogdiercellen te introduceren. Het bereiken van optimale transfectie-efficiënties zonder merkbare cytotoxische effecten in de bovengenoemde primaire celtypen blijft echter een uitdaging. Bovendien zijn elektroporatie noch transfecties van RGN's en/of donor-DNA-substraten gemakkelijk geschikt voor *in vivo*-instellingen. En vaak zijn deze methoden voor het leveren van genbewerkingstools afhankelijk van celtypespecifieke reagentia waarvan de samenstellingen om bedrijfseconomische redenen onbekend kunnen zijn. Eveneens opmerkelijk is dat elektroporatie en transfecties verder afhankelijk zijn van tijdrovende optimalisaties van celtypespecifieke parameters waarvan de prestaties kunnen variëren vanwege subtiele experimentele omstandigheden, bijvoorbeeld celcyclusfaseprofielen van doelcelpopulaties. Virale vectortransducties zijn daarentegen eenvoudig uit te voeren en bieden doorgaans een hogere reproduceerbaarheid, onafhankelijk van het gekozen doelceltype [13]. Deze kenmerken komen voort uit de nauwkeurig afgestemde mechanismen die door de ouderlijke virussen zijn ontwikkeld voor de overdracht van nucleïnezuur naar de kernen van gastcellen.

In **Hoofdstuk 4**, geïnspireerd door de complementaire kenmerken van twee verschillende klassen van virale genvrije virale vectoren, namelijk adenovirale vectoren met hoge capaciteit (AdV) en AAV-vectoren (respectievelijk bijvoorbeeld grote vrachtcompatibiliteit en bron van recombinogene HDR-substraten), wordt een dubbel viraal vectorsysteem geïntroduceerd, gekarakteriseerd en getest voor HDR-gemedieerde gen-KI-doeleinden. In dit duale virale vector-genoombewerkingsplatform worden RGN-constructies verpakt in AdV-capsiden; terwijl donor-DNA-sjablonen in AAV-capsiden worden geplaatst.

Naast overwegingen over de levering van genoombewerkingstools die verband houden met de uiteindelijke efficiëntie van gen-KI-procedures, zijn er andere parameters om te overwegen met betrekking tot de specificiteit en precisie van donor-DNA-insertie. De specificiteit is het resultaat van het detecteren van donor-DNA-sequenties op de doellocatie; terwijl de precisie kan worden beoordeeld door de frequentie van nauwkeurige HDR-afhankelijke DNA-inserties boven onnauwkeurige NHEJ-afgeleide gebeurtenissen vast te stellen. In dit opzicht is de kennis over de relatieve bijdragen van verschillende AAV-donorontwerpen (d.w.z. conventioneel versus 'dubbel gesneden') en structuren (d.w.z. enkelstrengs versus dubbelstrengs) aan de specificiteit en precisieparameters van genoombewerking schaars. Daarom diende het duale virale vectorplatform dat in **hoofdstuk 4** werd bestudeerd, naast het bereiken van opmerkelijk hoge gen-KI-efficiënties in getransformeerde, niet-getransformeerde en primaire cellen, als een probe om de relatie te bestuderen tussen de eerdergenoemde AAV-DNA-arrangementen en de specificiteit en precisie van gen-KI en taggingprocedures. Dit onderzoek onthulde dat het combineren van enkelstrengs AAV-afgifte van HDR-donoren met AdV-overdracht met hoge capaciteit van RGN-constructies leidt tot nauwkeurige genoombewerking in grote fracties van doelcelpopulaties. En het onthulde dat RGN-geïnduceerde chromosomale DNA-breuken productieve AAV-vectortransductie bevorderen, zoals gescoord door transgeenexpressie, vermoedelijk als gevolg van het bevorderen van enkelstrengs naar transcriptioneel competente dubbelstrengs DNA-conversie. Ten slotte wordt in **Hoofdstuk 4** verder het cruciale belang onthuld van het gebruik van high-fidelity RGN's voor het minimaliseren van off-target donor AAV-inserties in de vorm van defecte vectorgenomen waarvan bekend is dat ze verpakt zijn in vectordeeltjes [14,15].

Het eerdergenoemde recombinogene karakter van AAV-vectorgenomen, dat ze tot efficiënte HDR-substraten maakt, draagt ook bij aan hun promiscue "vangst" bij on-target en off-target of willekeurige chromosomale DNA-breuken door onnauwkeurige NHEJ-processen [11,12,16,17]. Deze end-joining-processen kunnen verder leiden tot chromosomale insertie van concatemere en subgenomische AAV-soorten, waardoor het bereik van HDR-onafhankelijke omstandergebeurtenissen verder wordt vergroot [11,12,17,18]. Tot slot kunnen AAV-genomen, mogelijk door het nabootsen van DNA-laesies of reparatie-intermediairen, de levensvatbaarheid van cellen aantasten door P53-afhankelijke DDR-activering waarvan de gevolgen naar verluidt bijzonder schadelijk zijn in stamcellen [19,20]. Concluderend geeft een groeiend aantal onderzoeken aan dat er onderscheidende genotoxische en cytotoxische effecten zijn gekoppeld aan respectievelijk chromosomaal geïntegreerde en episomale AAV-DNA-vormen, waarbij de opbouw van deze schadelijke effecten strikt evenredig is met de hoeveelheid AAV-vectorinvoer.

Om de tekortkomingen aan te pakken die verband houden met AAV-gebaseerde genoombewerking, werd in **hoofdstuk 5** een merkervrij coselectiesysteem [21] onderzocht, dat afhankelijk is van de krachtige specifieke remmer van de Na^+/K^+ ATPase-pomp, ouabaïne. Hiervoor werden

AAV-donorvectoren met een secundaire ouabaïne-selecteerbare donormodule en bijpassende gRNA-eenheden samengesteld en getest. Dergelijke AAV-ontwerpen, selector-AAV-vectoren genoemd, kunnen zorgen voor verrijking voor cellen die co-bewerkt zijn via een primaire doellocus en *ATP1A1*-allelen die resistentie verlenen tegen ouabaïne bij HDR-gemedieerde verwerving van specifieke polymorfismen. Belangrijk is dat naast verrijking voor gen-KI-celpopulaties, het combineren van selector-AAV-vectoren met ouabaïneselectie de eliminatie van HDR-onafhankelijke bewerkingen en off-target en/of willekeurige AAV-donor-DNA-inserties uit genoemde populaties veroorzaakte. Dit selector-AAV-principe werd succesvol toegepast voor het invoegen van hele transgenen op veilige haven-genomische loci en voor het taggen van endogene eiwitten. Verder onthulden selector-AAV-vectortitratie-experimenten dat de hoogste verrijkingsfactoren van gen-gerichte cellen geassocieerd zijn met de laagste vector doses. Deze eigenschap zou gunstig kunnen zijn voor het verlichten van zowel AAV-productiekosten als schadelijke P53-afhankelijke DDR-activering. Opmerkelijk is dat het dubbele virale vectorplatform dat **in hoofdstuk 4** wordt beschreven, werd gebruikt voor het stroomlijnen en uitbreiden van selector-AAV-vectortesten in moeilijk te transfecteren primaire cellen.

Zoals besproken in **Hoofdstuk 2** en **Hoofdstuk 3**, is er een groeiend besef dat, met name in DNA-schadegevoelige stamcellen, programmeerbare nuclease-geïnduceerde DSB's schadelijk zijn voor de stabiliteit van doelloci [22,23], en de levensvatbaarheid van cellen [7,19]. Het is veelzeggend dat Cas9^{D10A}-nickases, vergeleken met Cas9-nucleasen, notoir minder mutageen zijn bij zowel on-target als off-target genomische sequenties, zoals bepaald door reporter en onbevooroordeelde genoombrede high-throughput-assays, respectievelijk [24,25]. Cas9^{D10A}-nickases vertonen verder een sterk verminderde P53-afhankelijke DDR-activering wanneer ze naast hun Cas9-nuclease-tegenhangers worden getest [9,26]. In deze context hebben transductie-experimenten beschreven in **Hoofdstuk 5** aangetoond dat het combineren van een bepaald selector AAV-ontwerp, genaamd in-linkage selector AAV, met Cas9- of Cas9^{D10A}-blootstelling leidt tot een significante toename van de frequenties van gen-gerichte cellen. De ouabaïne-afhankelijke verrijkingsfactor voor gen-gerichte cellen was ongeveer 7,5 keer hoger bij gebruik van Cas9^{D10A}-nickase-afgifte. Belangrijk is dat het combineren van in-linkage selector AAV-transductie met Cas9-nuclease of Cas9^{D10A}-nickase-afgifte leidde tot de grondige eliminatie van donor-DNA dat was geïnserteerd op off-target genomische posities van genoom-bewerkte celpopulaties. Daarom zijn deze proof-of-principle-experimenten die aangeven dat op selector AAV gebaseerde genoombewerking transleerbaar zijn naar protocollen met Cas9^{D10A}-geïnduceerde HDR relevant voor het verder verfijnen en toepassen van DSB-onafhankelijke celengineeringstrategieën.

Concluderend kan gesteld worden dat het werk dat in dit proefschrift wordt gepresenteerd, door het onderzoek en de integratie van verschillende virale vectoren en RNA-programmeerbare Cas9-eiwitten, nieuwe inzichten en hulpmiddelen biedt voor het bevorderen van genoombewerking door het verbeteren van prestatieaspecten met betrekking tot de efficiëntie, specificiteit en betrouwbaarheid ervan.

References

1. Pacesa M, Pelea O, Jinek M. Past, present, and future of CRISPR genome editing technologies. Cell 2024 187:1076-1100.
2. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 2014 346:1258096.
3. Koonin EV, Makarova KS, Zhang F. Diversity, classification and evolution of CRISPR-Cas systems. Curr Opin Microbiol. 2017 37:67-78.
4. Chen X, Gonçalves MAFV. DNA, RNA, and Protein Tools for Editing the Genetic Information in Human Cells. iScience 2018 6:247-263.
5. Chen X, Rinsma M, Janssen JM, Liu J, Maggio I, Gonçalves MA. Probing the impact of chromatin conformation on genome editing tools. Nucleic Acids Res. 2016 44:6482-92.
6. Liao H, Wu J, VanDusen NJ, Li Y, Zheng Y. CRISPR-Cas9-mediated homology-directed repair for precise gene editing. Mol Ther Nucleic Acids. 2024 35:102344.
7. Chen X, Janssen JM, Liu J, Maggio I, 't Jong AEJ, Mikkers HMM, Gonçalves MAFV.

In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. Nat Commun. 2017 8:657.

8. Nakajima K, Zhou Y, Tomita A, Hirade Y, Gurumurthy CB, Nakada S. Precise and efficient nucleotide substitution near genomic nick via noncanonical homology-directed repair. Genome Res. 2018 28:223-230.
9. Wang Q, Liu J, Janssen JM, Gonçalves MAFV. Precise homology-directed installation of large genomic edits in human cells with cleaving and nicking high-specificity Cas9 variants. Nucleic Acids Res. 2023 51:3465-3484.
10. Chai AC, Chemello F, Li H, Nishiyama T, Chen K, Zhang Y, Sánchez-Ortiz E, Alomar A, Xu L, Liu N, Bassel-Duby R, Olson EN. Single-swap editing for the correction of common Duchenne muscular dystrophy mutations. Mol Ther Nucleic Acids. 2023 32:522-535.
11. Hanlon KS, Kleinstiver BP, Garcia SP, Zaborowski MP, Volak A, Spirig SE, Muller A, Sousa AA, Tsai SQ, Bengtsson NE, Lööv C, Ingelsson M, Chamberlain JS, Corey DP, Aryee MJ, Joung JK, Breakefield XO, Maguire CA, György B. High levels of AAV vector integration into CRISPR-induced DNA breaks. Nat Commun. 2019 10:4439.
12. Nelson CE, Wu Y, Gemberling MP, Oliver ML, Waller MA, Bohning JD, Robinson-Hamm JN, Bulaklak K, Castellanos Rivera RM, Collier JH, Asokan A, Gersbach CA. Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. Nat Med. 2019 25:427-432.
13. Chen X, Gonçalves MA. Engineered Viruses as Genome Editing Devices. Mol Ther. 2016 24:447-457.
14. Wada M, Uchida N, Posadas-Herrera G, Hayashita-Kinoh H, Tsunekawa Y, Hirai Y, Okada T. Large-scale purification of functional AAV particles packaging the full genome using short-term ultracentrifugation with a zonal rotor. Gene Ther. 2023 30:641-648.
15. McColl-Carboni A, Dollive S, Laughlin S, Lushi R, MacArthur M, Zhou S, Gagnon J, Smith CA, Burnham B, Horton R, Lata D, Uga B, Natu K, Michel E, Slater C, DaSilva E, Bruccoleri R, Kelly T, McGivney JB 4th. Analytical characterization of full, intermediate, and empty AAV capsids. Gene Ther. 2024 31:285-294.
16. Miller DG, Petek LM, Russell DW. Adeno-associated virus vectors integrate at chromosome breakage sites. Nat Genet. 2004 36:767-773.
17. Ferrari S, Jacob A, Cesana D, Laugel M, Beretta S, Varesi A, Unali G, Conti A, Canarutto D, Albano L, Calabria A, Vavassori V, Cipriani C, Castiello MC, Esposito S, Brombin C, Cugnata F, Adjali O, Ayuso E, Merelli I, Villa A, Di Micco R, Kajaste-Rudnitski A, Montini E, Penaud-Budloo M, Naldini L. Choice of template delivery

mitigates the genotoxic risk and adverse impact of editing in human hematopoietic stem cells. Cell Stem Cell 2022 29:1428-1444.

18. Suchy FP, Karigane D, Nakauchi Y, Higuchi M, Zhang J, Pekrun K, Hsu I, Fan AC, Nishimura T, Charlesworth CT, Bhadury J, Nishimura T, Wilkinson AC, Kay MA, Majeti R, Nakauchi H. Genome engineering with Cas9 and AAV repair templates generates frequent concatemeric insertions of viral vectors. Nat Biotechnol. 2025 43:204-213.
19. Schiroli G, Conti A, Ferrari S, Della Volpe L, Jacob A, Albano L, Beretta S, Calabria A, Vavassori V, Gasparini P, Salataj E, Ndiaye-Lobry D, Brombin C, Chaumeil J, Montini E, Merelli I, Genovese P, Naldini L, Di Micco R. Precise Gene Editing Preserves Hematopoietic Stem Cell Function following Transient p53-Mediated DNA Damage Response. Cell Stem Cell. 2019 24:551-565.
20. Allen D, Weiss LE, Saguy A, Rosenberg M, Iancu O, Matalon O, Lee C, Beider K, Nagler A, Shechtman Y, Hendel A. High-Throughput Imaging of CRISPR- and Recombinant Adeno-Associated Virus-Induced DNA Damage Response in Human Hematopoietic Stem and Progenitor Cells. CRISPR J. 2022 5:80-94.
21. Agudelo D, Duringer A, Bozoyan L, Huard CC, Carter S, Loehr J, Synodinou D, Drouin M, Salsman J, Dellaire G, Laganière J, Doyon Y. Marker-free coselection for CRISPR-driven genome editing in human cells. Nat Methods. 2017 :615-620.
22. Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. Nat Biotechnol. 2015 33:179-186.
23. 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.
24. Chen X, Tasca F, Wang Q, Liu J, Janssen JM, Brescia MD, Bellin M, Szuhai K, Kenrick J, Frock RL, Gonçalves MAFV. 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.
25. Wang Q, Liu J, Janssen JM, Le Bouteiller M, Frock RL, Gonçalves MAFV. Precise and broad scope genome editing based on high-specificity Cas9 nickases. Nucleic Acids Res. 2021 49:1173-1198.
26. Hyodo T, Rahman ML, Karnan S, Ito T, Toyoda A, Ota A, Wahiduzzaman M, Tsuzuki S, Okada Y, Hosokawa Y, Konishi H. Tandem Paired Nicking Promotes Precise Genome Editing with Scarce Interference by p53. Cell Rep. 2020 30:1195-1207.

Curriculum Vitae

Zhen Li was born on April 1st, 1993 in the city of Sanya, Hainan Province, China. During 2011-2016, he was enrolled in the 5-year Bachelor study of clinical medicine consisted of one year of general college studies (2011-2012) and 4 years of fundamental and clinical medicine (2012-2016), which includes one year of clinical internship rotation in the Tongji Hospital affiliated to Huazhong University of Science and Technology, Wuhan, China.

From the year of 2016 to 2019, he, as a Master student of Surgery, joined the group of Prof. Dr. Liping Su in the Department of General Surgery at the Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, where he focused on the research investigating the impact of the long non-coding RNA *FMO6P* on the invasion and migration of gastric cancer cells. In the meantime, he also got the registered qualification certificate of practicing physician in 2017.

In the year of 2019, he received the Ph.D. scholarship from the China Scholarship Council - Leiden University joint program to chase the doctoral diploma in the Netherlands. With the supervision of Prof. Dr. Rob C. Hoeben and Dr. M.A.F.V. Gonçalves in the Department of Cell and Chemical Biology at the Leiden University Medical Center, he combined the CRISPR/Cas9 gene editing system and different viral vectors for achieving robust and versatile gene editing. Along with demonstrating the efficiencies of such viral vector-based system, he also probed the accuracy and specificity of the system of using different engineered Cas9 proteins delivered by viral vectors. Instead of being a full-time researcher in the laboratory, he prefers to engage with patients and choose to be a surgeon in a hospital.

List of Publications

1. **Li Z.**, Gonçalves M.A.F.V. "Soft" genome editing using CRISPR nickases as a potential source of safer cell products. Cell & Gene Therapy Insights 2023; 9:1201-1210. doi: 10.18609/cgti.2023.158.
2. **Li Z.**, Gonçalves M.A.F.V. AAV-vectored base editor *trans*-splicing delivers dystrophin repair. Molecular Therapy Nucleic Acids. 2023; 30;32:900-902. doi: 10.1016/j.omtn.2023.05.013.
3. **Li Z.**, Wang X., Janssen J.M., Liu J., Tasca F., Hoeben R.C., Gonçalves M.A.F.V. Precision genome editing using combinatorial viral vector delivery of CRISPR-Cas9 nucleases and donor DNA constructs. Nucleic Acids Res. 2025; 53(2):gkae1213. doi: 10.1093/nar/gkae1213.
4. **Li Z.***, Wang X.L.*, Liu J., Janssen J.M., Hoeben R.C., Gonçalves M.A.F.V. Selector AAV-CRISPR vectors purge off-target chromosomal insertions and promote precise genome editing. In preparation. * Shared first co-authorship

Acknowledgements

It’s a long and tough road to get here submitting the thesis for doctoral defense. Working with viruses is interesting and tricky, since we can use the power of our enemies to serve our own benefits but with amounts of intellectual and physical collaboration to achieve the goals by unpredictable cost. The explorative work presented in this thesis could be done by the financial supports of the China Scholarship Council - Leiden University joint program, the Prinses Beatrix Spierfonds, the Duchenne Parent Project NL, the Dutch Research Council (NWO) - Open Technology Program, and EU Marie Skłodowska-Curie Doctoral Network Actions. Beyonds the money, I also got the energy and confidence to finish the work from my family, friends, each member in Dr. Manuel Gonçalves’ and Prof. Dr. Rob Hoeben’s group, the personnel of the Flow Cytometry Core Facility of the LUMC and each people I met in CCB, LUMC, Leiden and the Netherlands. But at the very beginning, I am from Sanya in China, the route to Leiden is quite long, 5 years in Wuhan, 3 years in Shanghai and happily I am back to China pursuing the realization of the vow being a medical doctor. Till now I would like to thank all the people I met so far, not only the people I met in the journey to be a PhD, but each one I met in the journey of my life. No matter what experience the people I met brought to me, it makes who I am, therefore it’s also appreciated to be parts of the live of the people I met. Since there are too many names of the people who I would like to cherish and acknowledge in the path of life. Limited by the space and preferring to express in a simple way personally, my sincere gratitude is to all the people I met, meet and will.