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

Li, Z.

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.eC9^{4NLS}G^{S1}). The eCas9^{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.eC9^{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].

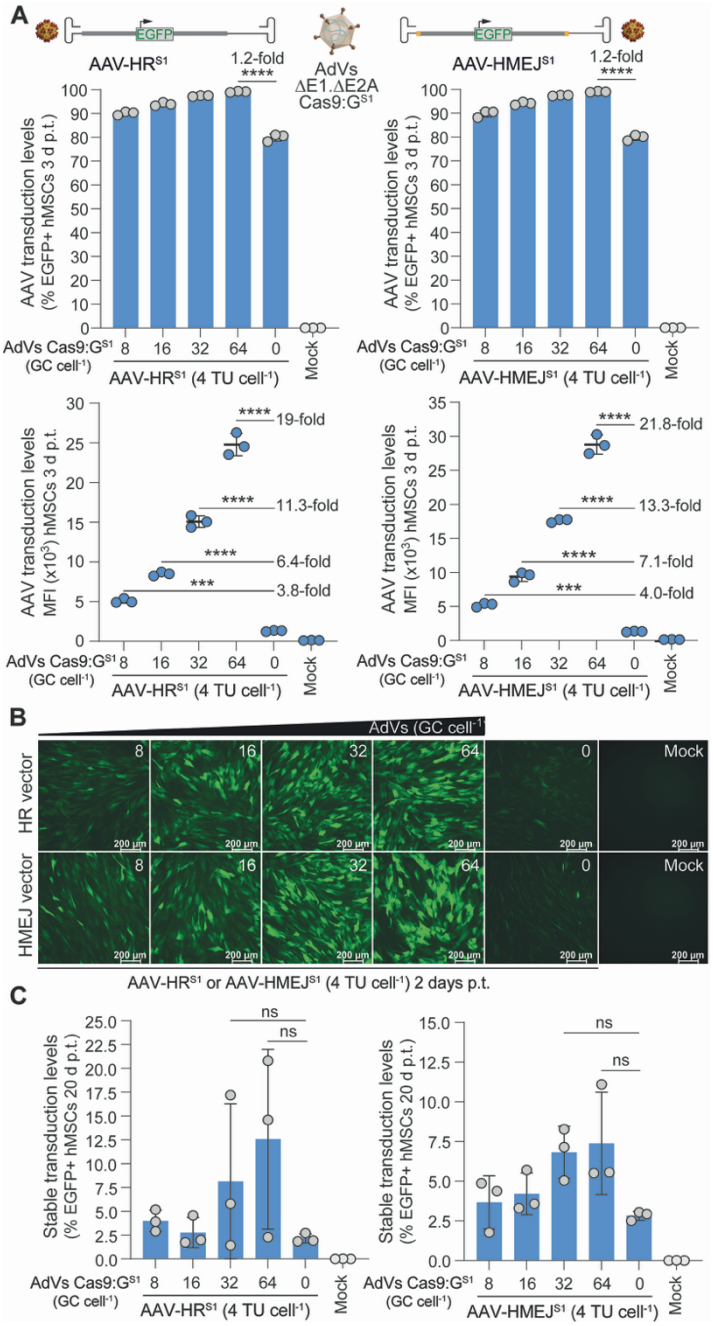

Figure 1. **Genome editing using AAV delivery of HR or HMEJ donors and AdV transfer of CRISPR-Cas9 complexes.** (**A**) AAV HR and HMEJ donor transduction efficiencies. Schematics of AAV donor structures are depicted. Horizontal gray bars, DNA

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.

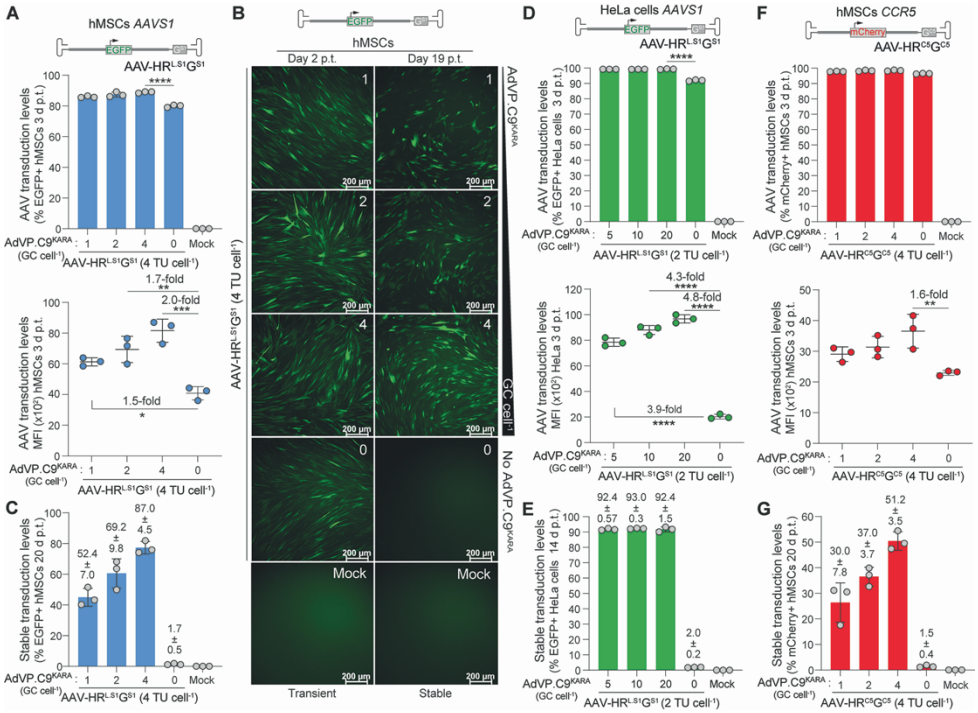

Figure 2. Genome editing combining AAV delivery of donor and gRNA components with high-capacity AdVP transfer of Cas9. (**A**) AAV donor transductions targeting *AAVS1* safe harbor loci in hMSCs. The structure of AAV-HR$^{L.S1}$G^{S1} donor is shown. Horizontal gray bars, DNA homologous to sequences 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.

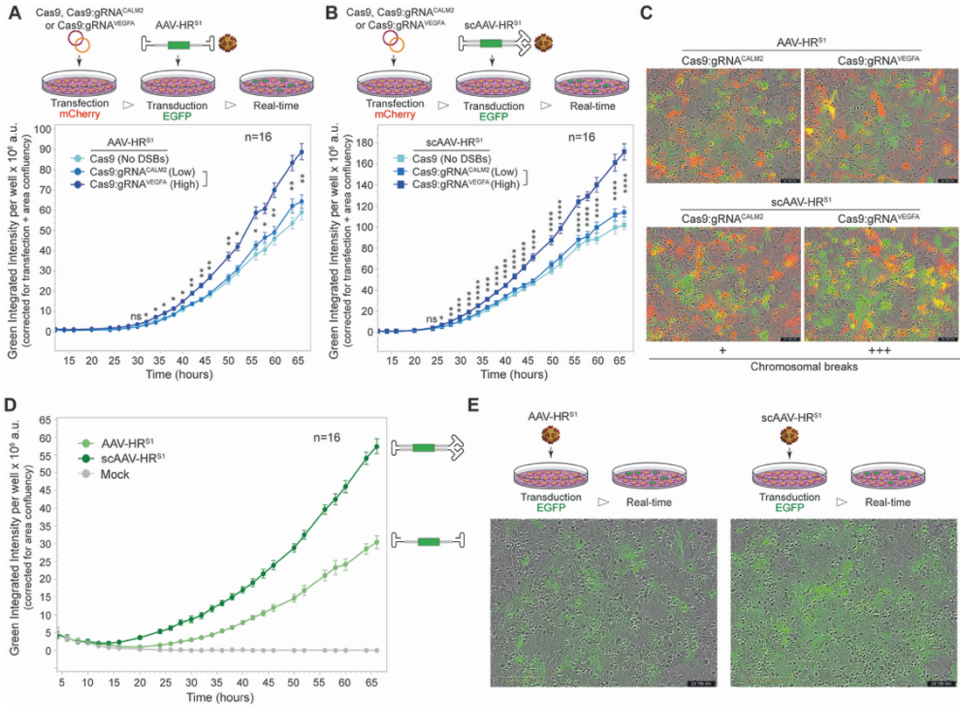

Figure 3. Assessing the role of chromosomal DNA breaks on AAV vector transduction. (**A**,**B**) Single-stranded AAV and double-stranded scAAV transduction kinetics in cells subjected to CRISPR-induced DSBs. HeLa cells exposed to Cas9:gRNA^{CALM2} or Cas9:gRNA^{VEGFA} complexes, inducing few or plenty 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^{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.

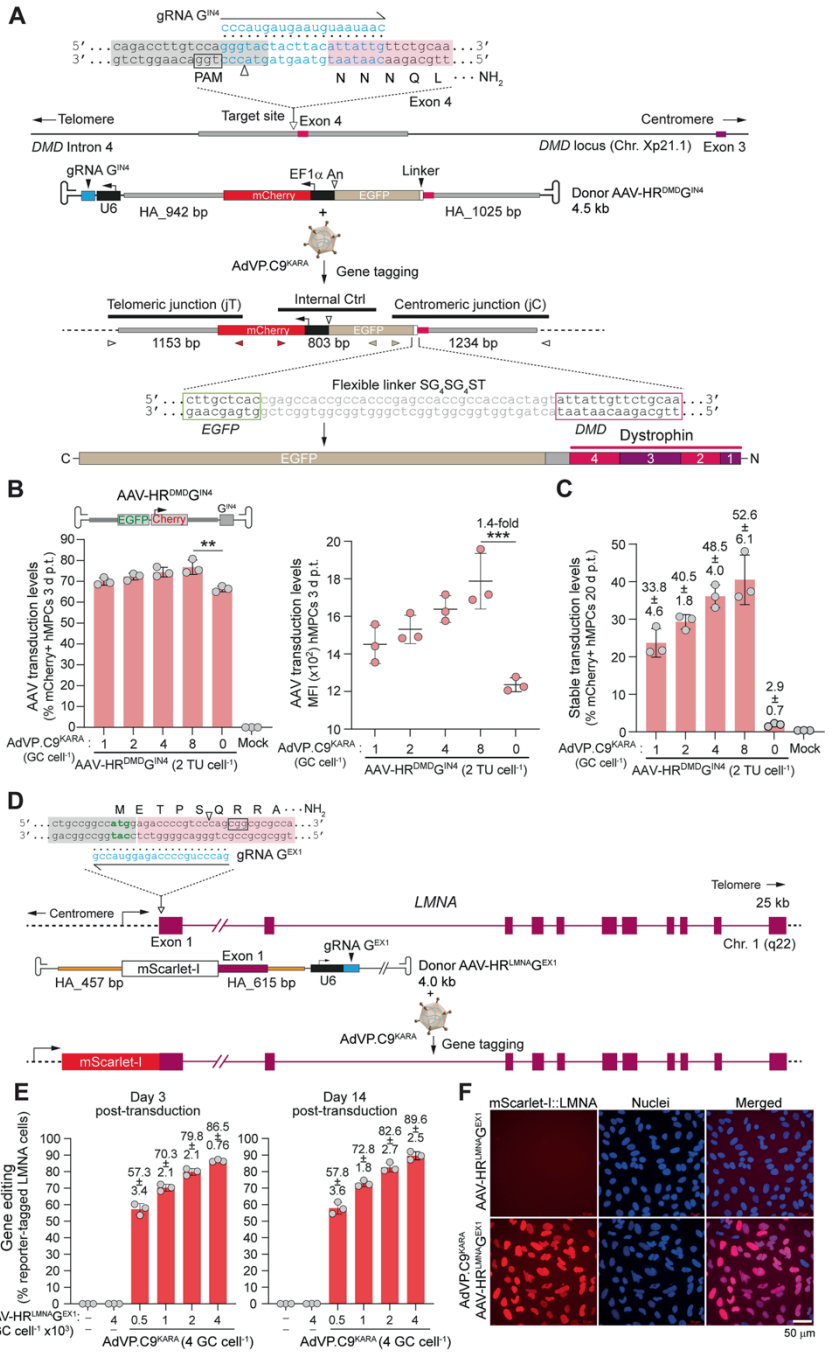

Figure 4. Endogenous gene tagging using dual viral vector delivery. (**A**) Diagrammatic representation of the *DMD* gene tagging strategy. Horizontal gray bars, AAV-HR^{DMD}G^{IN4} DNA homologous to sequences flanking the *DMD* target site of gRNA G^{IN4} ('homology

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.

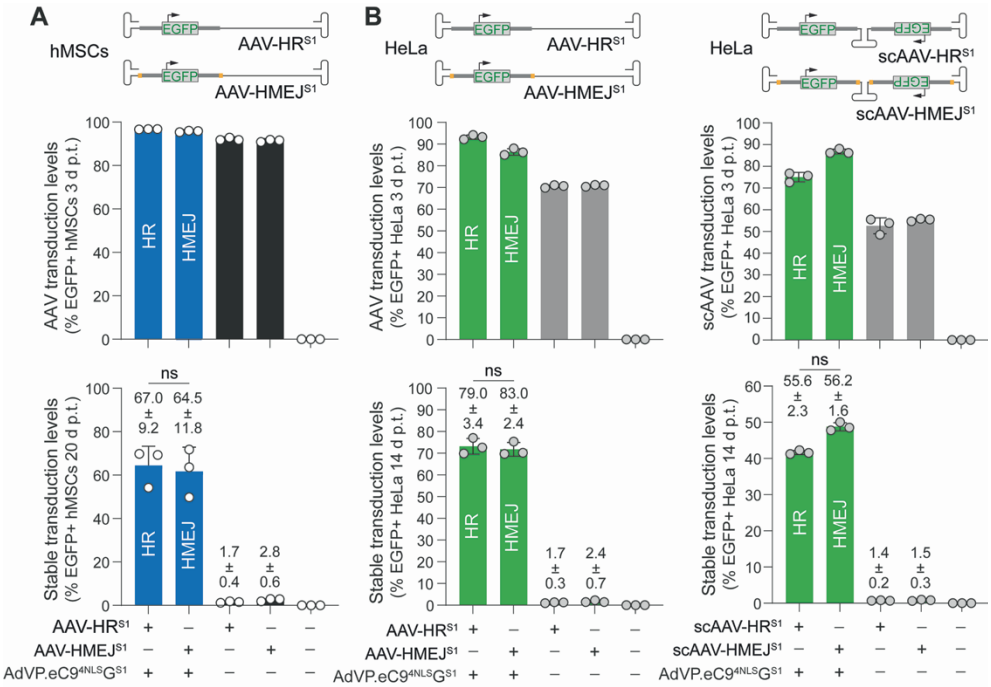

Figure 5. Investigating the role of AAV donor structures on genome editing. (**A**) Testing genome editing in hMSCs using single-stranded AAV HR and HMEJ donor constructs. Schematics of AAV donor structures are depicted. Horizontal gray bars, short tracts homologous to DNA flanking the *AAVS1* target 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.

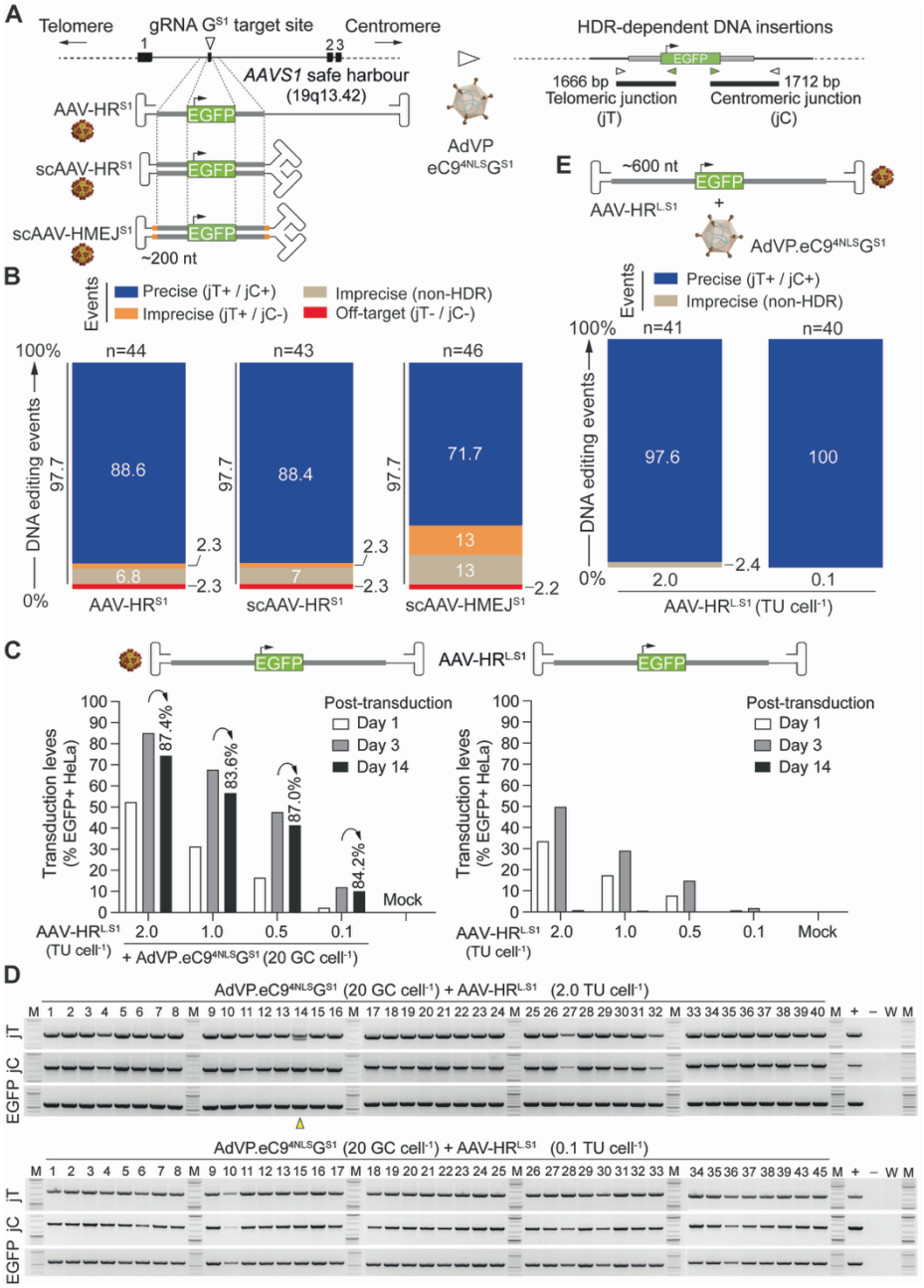

Figure 6. Investigating the role of AAV donor structures on the specificity and fidelity of genome editing. (**A**) Experimental setup. Homology arms in regular and sc AAV donors are drawn in relation to the *AAVS1* acceptor sequence before and after gene targeting. PCR amplicons diagnostic for HDR-mediated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

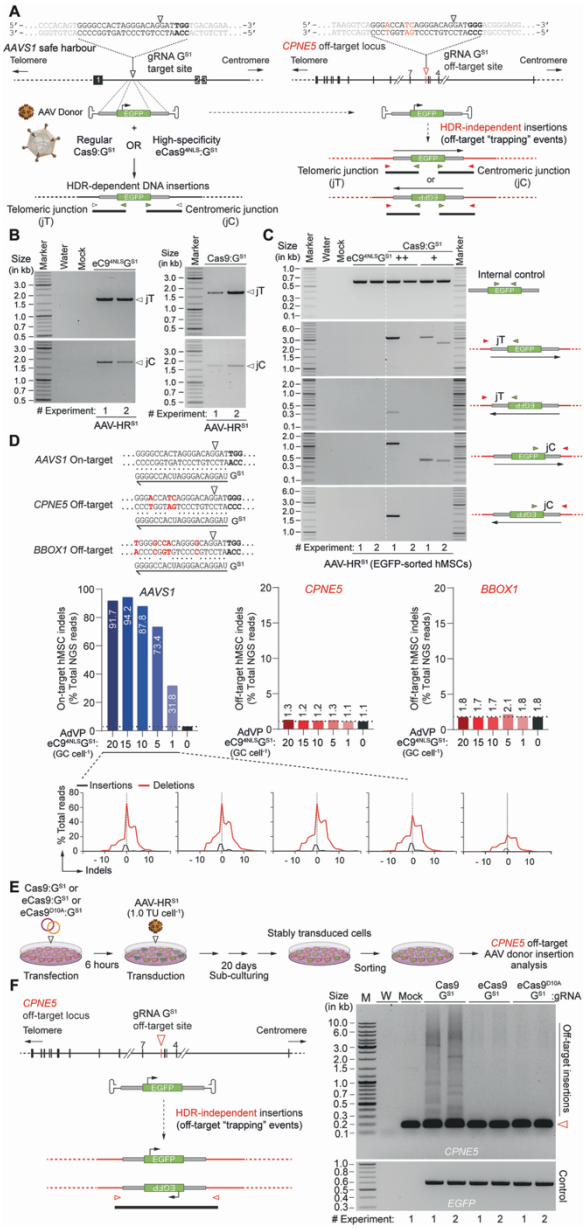

Figure 7. Investigating off-target AAV donor insertions resulting from using regular versus high-specificity CRISPR-Cas9 complexes. (**A**) Experimental setup for experiments in hMSCs. AAV HR donor sequences before and after chromosomal insertion at *AAVS1* and off-target *CPNE5* loci through HDR-dependent 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 μg 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™; 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℃ for 5 min followed by 40 cycles at 95℃ for 10 s and 60℃ 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.