Universiteit Leiden
The Netherlands

Adenoviral vectors as genome editing tools : repairing defective DMD alleles

Maggio, Ignazio

Citation
Maggio, I. (2016, November 17). *Adenoviral vectors as genome editing tools : repairing defective DMD alleles*. Retrieved from https://hdl.handle.net/1887/44288

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/44288

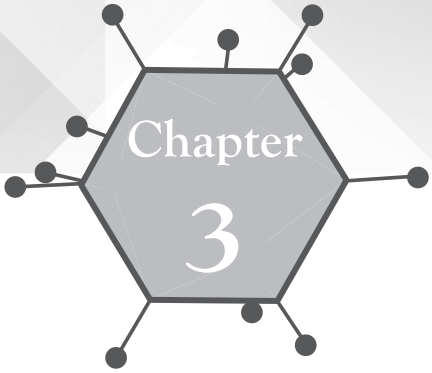
Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/44288 holds various files of this Leiden University dissertation

Author: Maggio, Ignazio
Title: Adenoviral vectors as genome editing tools : repairing defective DMD alleles
Issue Date: 2016-11-17



Adenoviral vector DNA for accurate genome editing with engineered nucleases

Maarten Holkers[1*], **Ignazio Maggio**[1*], Sara F.D. Henriques[1,2], Josephine M. Janssen[1], Toni Cathomen[3,4] and Manuel A.F.V. Gonçalves[1]

Nature Methods, 2014, **11**:1051-1057

* These authors contributed equally to this work.

[1.] Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, the Netherlands.
[2.] Departamento de Biologia Animal, Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal.
[3.] Institute for Cell and Gene Therapy, University Medical Center Freiburg, Freiburg, Germany.
[4.] Center for Chronic Immunodeficiency, University Medical Center Freiburg, Freiburg, Germany.

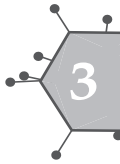
Abstract

Engineered sequence-specific nucleases and donor DNA templates can be customized to edit mammalian genomes via the homologous recombination (HR) pathway. Here we report that the nature of the donor DNA greatly affects the specificity and accuracy of the editing process following site-specific genomic cleavage by transcription activator–like effector nucleases (TALENs) and clustered, regularly interspaced, short palindromic repeats (CRISPR)-Cas9 nucleases. By applying these designer nucleases together with donor DNA delivered as protein-capped adenoviral vector (AdV), free-ended integrase-defective lentiviral vector or nonviral vector templates, we found that the vast majority of AdV-modified human cells underwent scarless homology-directed genome editing. In contrast, a significant proportion of cells exposed to free-ended or to covalently closed HR substrates were subjected to random and illegitimate recombination events. These findings are particularly relevant for genome engineering approaches aiming at high-fidelity genetic modification of human cells.

Contents chapter 3

Introduction

The exchange of genetic information between native acceptor loci and exogenous donor DNA through error-free HR is an established strategy to manipulate prokaryote and eukaryote genomes with nucleotide-level precision[1]. However, in mammalian somatic cells, typical frequencies of spontaneous HR-mediated gene-targeting range from 10^{-8} to 10^{-6} events per transfected cell, with most exogenous DNA being found randomly integrated throughout host-cell chromosomes[2-4]. Importantly, the deployment of sequence-specific nucleases greatly increases the odds of retrieving cells with specific allelic alterations[5]. Generation of double-stranded DNA breaks (DSBs) at predefined chromosomal positions together with the introduction of donor DNA containing sequences identical to those bracketing the genomic lesion can increase gene targeting by several orders of magnitude. In gene therapy, for instance, inserting transcriptional units into specific genomic positions (i.e., so-called safe harbors) or directly repairing faulty genes within their native chromosomal context is a highly desirable goal[6].



Viral vectors constitute attractive gene delivery vehicles owing to their efficient transduction of a wide range of cell types. In addition, in contrast to bulk nucleic acid transfections[7], viral transductions permit fine control over the number of DNA copies that reach the nucleus. We have investigated the role of the donor-delivering vectors on the specificity and accuracy of the DNA editing process. Specificity refers to the relative frequencies of on-target versus off-target insertions; accuracy refers to the structure or arrangement of site-specifically integrated exogenous DNA. Low-fidelity gene targeting includes the chromosomal integration of exogenous DNA copies in tandem (i.e., concatemers) as well as the incorporation of virus-derived sequences at the target site. These events may result from the generation of vector-vector or vector-host DNA junctions through error-prone nonhomologous end joining. Clearly, accurate site-specific gene addition and repair should yield, respectively, homogenous transgene activities and restored open reading frames (ORFs) in target cell populations.

In this study, we compared the specificity and the accuracy of nuclease-induced gene targeting upon the delivery of various types of HR substrates (i.e., protein-capped, free-ended and covalently closed DNA) and found that the nature of the donor DNA greatly affects these key parameters. We demonstrate that donor DNA transferred by protein-capped AdVs is amenable to homology-directed gene targeting after sequence-specific genomic cleavage by TALENs and CRISPR-Cas9 nucleases. Further, donor delivery by AdVs results in diminished off-target chromosomal insertion, concatemeric 'footprint' formation and prokaryotic DNA

incorporation. In consequence, genetically modified cell populations generated via AdV donor DNA transfer show homogenous transgene activity.

Results

Designer nuclease–induced IDLV DNA targeting is inaccurate

Integrase-defective lentiviral vectors (IDLVs)[8] are one of the most commonly used viral vectors for the delivery of HR substrates into human cells[9-14]. Because TALENs display a particularly favorable specificity profile[15], we started by investigating the specificity and fidelity of TALEN-induced chromosomal insertion of IDLV donor DNA. We introduced TALENs specific for the safe harbor *AAVS1* locus into human myoblasts through the early region 1 (*E1*)-deleted adenoviral vectors AdV.TALEN-L^{S1} and AdV.TALEN-R^{S1} (ref. 16) together with the target site-matched HR substrate delivered by IDLV.donor^{S1} particles (**Fig. 1a**). The AdV-delivered TALEN pair TALEN-L^{S1} and TALEN-R^{S1} forms a dimeric nuclease complex by binding, respectively, at the 'left' (L) and 'right' (R) end of the *AAVS1* target sequence (S1). The IDLV.donor^{S1}, generated on the basis of plasmid AQ25_pLV.donor^{S1} (**Supplementary Note**), contains a transgene encoding the fluorescent reporter EGFP flanked by sequences sharing identity to DNA bracketing the TALEN target site. Cells exposed to the latter vector alone or mixed with AdV.TALEN-L^{S1} served as negative controls. After subculturing to eliminate episomal vector DNA, flow cytometry (**Fig. 1b**) and live-cell fluorescence microscopy (**Supplementary Fig. 1a**) showed a significant ($P < 0.0001$) nuclease-dependent increase in the frequencies of stably transduced cells.

Next, to gauge the relative frequencies of on-target versus off-target IDLV donor DNA integration, we randomly selected EGFP$^+$ myoblast clones (n = 104) from cultures exposed to the *AAVS1*-specific TALENs. We observed, using PCR screening with primers designed to yield amplicons diagnostic for HR-derived junctions between foreign and native target DNA (**Fig. 1a** and **Supplementary Fig. 1b**), that 86.6% of the EGFP$^+$ cells underwent homology-directed chromosomal integration of the exogenous DNA (**Fig. 1c** and **Supplementary Fig. 1b**). The resulting *AAVS1*-donor DNA junctions represented events involving the telomeric (6.7%), centromeric (3.9%) or both ends of the targeting template (76.0%) (**Fig. 1c**). Further characterization of IDLV integrants revealed high frequencies of head-to-tail (H-T) concatemeric forms not only in the nontargeted but also in the three *AAVS1*-targeted clonal fractions (38.5%) (**Fig. 1c** and **Supplementary Fig. 1b**). Of note, PCR analysis of EGFP$^+$ cells sorted from cultures that were not exposed to TALENs did not yield amplicons diagnostic for homology-directed gene addition (**Supplementary Fig. 1c**).

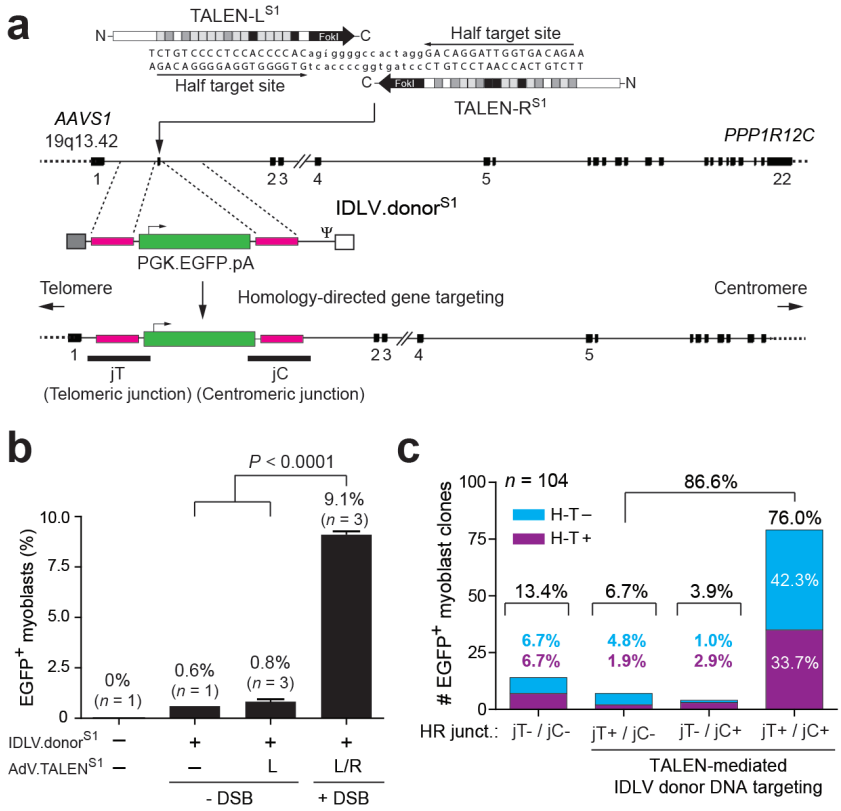


Figure 1. Gene targeting with IDLV donor DNA and TALEN-induced DSBs.

(**a**) The schematic shows TALENs drawn in relation to their target sites in the *AAVS1* locus. These bipartite target sites frame spacer sequences, which are cleaved upon local nuclease pair dimerization. IDLV.donor^{S1} vector genomes contain HR substrates consisting of a reporter expression unit flanked by sequences identical to those bracketing the TALEN target sequences. The transgene in IDLV.donor^{S1} comprises the human *PGK1* promoter, the *EGFP* ORF and the bovine *GH1* polyadenylation signal. The structure of the integrated exogenous DNA resulting from error-free HR events, forming 'telomeric' and 'centromeric' junctions (jT and jC), is depicted. Ψ, HIV-1 packaging signal; white and gray boxes, 5′ and 3′ long terminal repeat, respectively. (**b**) Flow cytometric analysis on myoblasts cotransduced with the indicated viral vector constructs. Flow cytometry was performed at 27 d post transduction; 10,000 events, each corresponding to a single viable cell, were measured per sample. Error bars, s.d. *P* value (by two-tailed *t*-test) and the number of independent experiments (*n*) are shown. DSB, double-stranded DNA break. (**c**) Cumulative molecular characterization of EGFP$^+$ myoblast clones generated by TALEN-L^{S1}:TALEN-R^{S1} and IDLV.donor^{S1} DNA delivery (**Supplementary Fig. 1b**). The frequencies of EGFP$^+$ myoblast clones with random insertions (jT−/jC−), HR-derived telomeric junctions (jT+/jC−), HR-derived centromeric junctions (jT−/jC+) and HR-derived telomeric and centromeric junctions (jT+/jC+) are plotted. The percentages of myoblast clones without (−) or with (+) head-to-tail IDLV concatemers (H-T) are also plotted.

These data indicate that a substantial fraction of incoming IDLV donor templates integrates randomly into host cell chromosomes, presumably at spontaneous DSBs through a noncanonical, i.e., integrase-independent, process[17,18]. Moreover, owing to their concatemeric structure, a sizable proportion of IDLV-exposed cells harbors unwanted HIV-derived *cis*-acting elements (**Fig. 1c** and **Supplementary Figs. 1b** and **2**). These tandem repeats are expected to neither restore endogenous ORFs nor yield homogeneous transgene expression levels in the context of gene repair and gene addition strategies, respectively. These results are in line with those of other experiments carried out by us (**Supplementary Results** and **Supplementary Fig. 3**) as well as by others[11,12,14] deploying IDLV donor DNA and zinc-finger nuclease (ZFN) technology.

Designer nuclease–induced AdV DNA targeting is accurate

Together with IDLVs, adeno-associated viral vectors (rAAVs) constitute the most commonly used viral vectors for the delivery of HR substrates into mammalian cells[4,19-22]. Like IDLVs, rAAV genomes have free ends and can become inserted at sporadic genomic DSBs after being co-opted by illegitimate recombination pathways involved in chromosomal DNA break repair[11,23]. In contrast, linear double-stranded AdV DNA has a terminal protein (TP) covalently attached at its 5′ ends. This led us to postulate that this capped DNA structure reduces the chance for interactions between the donor and off-target DSBs, such that pairing between acceptor and donor DNA through shared sequences would favor HR-dependent insertions at cleaved target sites. We therefore asked whether delivering donor DNA in the context of protein-capped AdV genomes[24] displays a less promiscuous chromosomal integration pattern and a more precise insertion profile than that resulting from using free-ended DNA. We generated the *E1*- and early region 2A (*E2A*)-deleted AdV.Δ2.donor^{S1} to introduce the *AAVS1*-matched HR substrate into myoblasts exposed to the *AAVS1*-specific TALENs, as well as to unexposed cells. This AdV-delivered donor DNA harbors the same EGFP-encoding transcriptional unit present in the aforementioned IDLV.donor^{S1} (**Fig. 1a**). Flow cytometry of the resulting long-term cultures showed a clear nuclease-dependent increase in the frequencies of stably transduced cells (**Fig. 2a**) with not only TALEN-induced but also residual DNA integration rates being lower than those measured in their IDLV.donor^{S1}-transduced counterparts (**Fig. 1b**). The degree of the TALEN-dependent stimulatory effect was nonetheless similar to that observed in IDLV.donor^{S1}-transduced myoblasts (**Fig. 1b**). Interestingly, cotransducing myoblasts with the AdV.Δ2.donor^{S1} and TALEN-encoding AdVs resulted in EGFP^{+} populations with a distribution of transgene expression levels almost as narrow as those of clones harboring *AAVS1*-targeted

donor^{S1} DNA and much narrower than those of IDLV.donor^{S1}-modified populations (**Fig. 2b**). We observed similar results at the clonal level by comparing the mean fluorescence intensity (MFI) of EGFP$^+$ myoblasts randomly selected from TALEN-treated cultures transduced either with IDLV.donor^{S1} or with AdV.Δ2.donor^{S1} (**Fig. 2c**). Collectively, these data indicate scarce chromosomal positional effects on transgene activity in AdV.Δ2.donor^{S1}-modified populations, possibly resulting from a preponderance of site-specific over random genomic DNA insertions. Indeed, all of the EGFP$^+$ myoblast clones isolated from cultures cotransduced with AdV.Δ2.donor^{S1} and TALEN-encoding AdVs (n = 110) had *AAVS1*–foreign DNA junctions resulting from HR events at both termini (**Fig. 2d** and **Supplementary Fig. 4a**).

To further probe the precision of AdV gene targeting, we set up a PCR assay to detect head-to-tail exogenous DNA concatemers. This assay failed to produce any discernible head-to-tail-specific PCR species from genomic DNA of EGFP$^+$ myoblasts sorted from cultures cotransduced with AdV.Δ2.donor^{S1} and TALEN-encoding AdVs (**Supplementary Fig. 4b**). Southern blot analysis of AdV.Δ2.donor^{S1}-modified clones confirmed target-site specificity. We note that we did identify a clone that, in addition to the typical *AAVS1*-targeted donor DNA, contained an integrant whose origin is consistent with HR-independent integration (**Supplementary Fig. 5a**). Nonetheless, in line with its low prevalence, we did not detect the latter type of integrant in the parental EGFP$^+$ myoblast population (**Supplementary Fig. 5b**). Southern blot analysis also identified a clone that underwent biallelic targeting.

Next we performed AdV gene-targeting experiments in HeLa cells. These cells display a high degree of genetic instability (**Supplementary Fig. 6**), providing, as a result, a more stringent model system in which to evaluate HR-mediated genome editing amidst a presumably high frequency of spontaneous chromosomal DSBs. Consistent with the experiments carried out in myoblasts (**Fig. 2a**), we observed a significant ($P < 0.0001$) TALEN-dependent increase in the frequency of stably transduced cells (**Fig. 3a,b**). We did not detect head-to-tail AdV DNA concatemers in the EGFP$^+$ populations resulting from TALEN-induced chromosomal insertion of Adv.Δ1.donor^{S1} DNA (**Fig. 3c**). Notably, all randomly selected EGFP$^+$ HeLa cell clones (n = 83) were genetically modified through homology-directed gene targeting at *AAVS1*, confirming a high level of target-site specificity following AdV-mediated delivery of donor DNA (**Fig. 3d** and **Supplementary Fig. 7**).

AdV-based gene targeting with CRISPR-Cas9

We investigated the compatibility of nuclease-induced AdV gene targeting with the versatile CRISPR-Cas9 RNA-guided nuclease system[25]. We deployed the vector pair AdV.Cas9 and AdV.gRNA^{S1}, which encode Cas9 and a single guide RNA

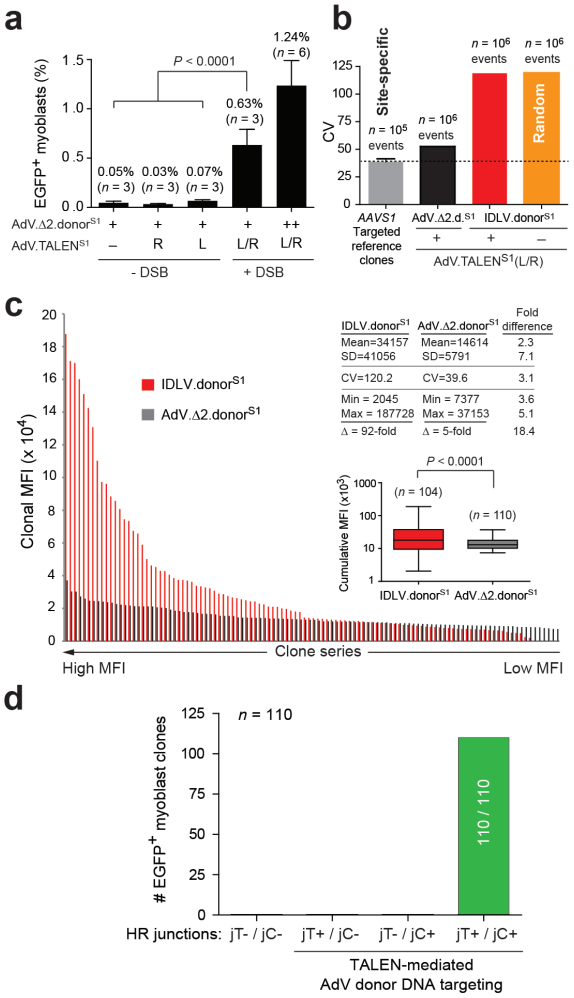


Figure 2. Gene targeting with AdV donor DNA and TALEN-induced DSBs.

(**a**) Flow cytometric analysis on myoblasts cotransduced with the indicated viral vector constructs. The AdV.Δ2.donor^{S1} was applied at multiplicity of infection values of 5 transducing units (TU)/cell (+) and 10 TU/cell (++). Cultures exposed to AdV.Δ2.donor^{S1} at 5 TU/cell alone (+, −) or mixed with AdV.TALEN-L^{S1} (+, L) or with AdV.TALEN-R^{S1} (+, R) served as negative controls. Flow cytometry was performed at 45 d post transduction; 10,000 events corresponding to single viable cells were measured per sample. Error bars, s.e.m. *P* value (by two-tailed *t*-test) and the number of independent experiments (*n*) are shown. DSB, doubled-stranded DNA break. (**b**) Coefficient of variation (CV) of EGFP$^+$ populations exposed to *AAVS1*-specific TALENs and HR substrates delivered via the indicated donors. Clones containing an *AAVS1*-targeted donor^{S1} DNA copy (*n* = 10 clones) and EGFP$^+$ populations with randomly inserted IDLV.donor^{S1} DNA served as controls (gray and orange bars, respectively). (**c**) Mean fluorescence intensity (MFI) values of EGFP$^+$ myoblast clones targeted with TALENs and either IDLV.donor^{S1} (red bars) or AdV.Δ2.donor^{S1} (black bars). Inset, box plot of the ▶

◀ cumulative MFI values corresponding to both series of myoblast clones analyzed; *n* indicates number of clones. Whiskers, minimum and maximum. *P* was calculated by two-tailed *t*-test. (**d**) Cumulative data on randomly selected EGFP$^+$ myoblast clones (n = 110) isolated from cultures cotransduced with AdV.Δ2.donor^{S1}, AdV.TALEN-L^{S1} and AdV.TALEN-R^{S1} (**Supplementary Fig. 4a**). The frequencies of EGFP$^+$ myoblast clones representing the different integrant classes are plotted as in **Figure 1c**.

(gRNA^{S1}) addressing the Cas9 nuclease to a genomic position overlapping with the target site for the *AAVS1*-specific TALENs (**Fig. 4a**). Exposing HeLa cells to AdV.Cas9 and AdV.gRNA^{S1} resulted in robust and dose-dependent DSB formation at *AAVS1* (**Supplementary Fig. 8**).

We initiated gene-targeting experiments by cotransducing HeLa cells with AdV.Δ2.donor^{S1}, AdV.Cas9 and AdV.gRNA^{S1}. HeLa cells cotransduced exclusively with AdV.Δ2.donor^{S1} and AdV.Cas9 served as negative controls. In these experiments, we also tested conventional nonviral plasmid donors with both circular and linear topologies deploying, in this case, TALENs as the designer nuclease system. Thus, HeLa cells were cotransfected with constructs encoding the *AAVS1*-specific TALENs mixed with supercoiled pAdV.donor^{S1}, PacI-linearized pAdV.donor^{S1} or TALEN-cleavable pAdV.donor$^{S1/T\text{-}TS}$ (**Supplementary Fig. 9**). The free-ended plasmid donor templates were generated *in vitro* and *in cellula* by restriction enzyme– and TALEN–induced DNA cleavage, respectively. Of note, the nonprokaryotic DNA portions of these plasmids are isogenic to those present in protein-capped AdV.Δ2.donor^{S1} genomes. HeLa cells transfected with the TALEN-L^{S1}-expressing construct plus each donor plasmid type served as negative controls.

The percentages of EGFP$^+$ cells in cultures exposed to Cas9:gRNA^{S1} complexes and AdV.Δ2.donor^{S1} (**Supplementary Fig. 9**) were comparable to those measured in populations treated with TALEN-L^{S1}:TALEN-R^{S1} dimers and AdV.Δ2.donor^{S1} (**Fig. 3a,b**). Moreover, the frequencies of genetically modified cells were clearly higher in cultures that had been subjected to site-specific DSBs. Crucially, stably transduced HeLa cells generated by co-delivering CRISPR-Cas9 complexes and AdV.Δ2.donor^{S1} DNA displayed a remarkably narrow range of transgene activities, as determined by flow cytometric screening of randomly selected EGFP$^+$ clones (**Fig. 4b,c**), reminiscent of data in myoblasts modified with TALENs and AdV donor DNA (**Fig. 2c**). In contrast, cells exposed to TALENs and plasmid donors led to EGFP$^+$ clones displaying significantly ($P < 0.0001$) broader distributions of transgene expression levels, independently of the topology of the donor DNA (**Fig. 4b,c**). The higher homogeneity of transgene expression among AdV-modified cells is also apparent in the correlation between the coefficient of variation (CV) and MFI values for each individual clone (**Supplementary Fig. 10a**), as well as by comparing the range of transgene expression levels in EGFP$^+$ populations resulting from

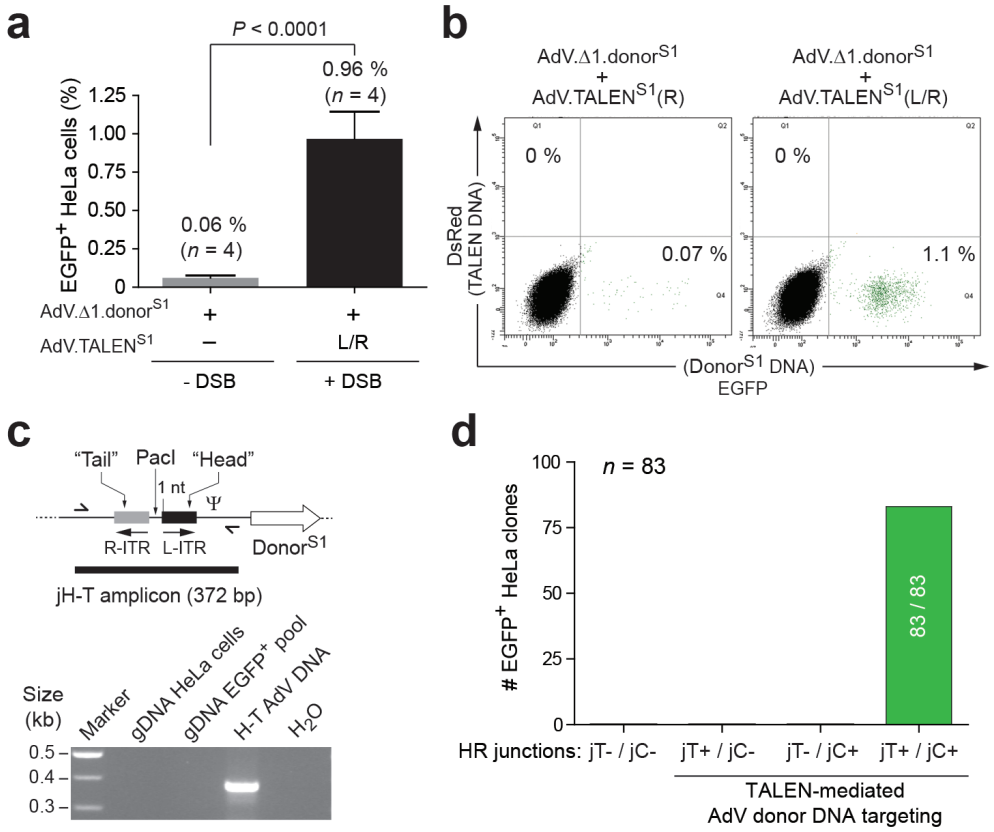


Figure 3. Nuclease-mediated gene targeting of AdV donor DNA in genetically unstable HeLa cells.

(**a**) Flow cytometric analysis of HeLa cells 24 d after transduction with the indicated viral vector constructs. Error bars, s.d. *P* value (by two-tailed *t*-test) and the number of independent experiments (*n*) are shown. (**b**) Representative flow cytometry dot plots corresponding to HeLa cell cultures exposed to the indicated viral vector constructs at 24 d post transduction. Owing to a bicistronic expression unit, cells transduced with TALEN-encoding AdVs become tagged with DsRedEx2.1. (**c**) PCR analysis probing for head-to-tail AdV DNA concatemers in transduced EGFP$^+$ HeLa cells. Top, schematics of *in vitro*–assembled head-to-tail AdV DNA junctions (jH-T). L-ITR and R-ITR, 'left' and 'right' AdV ITR, respectively; half arrows, primers; horizontal bar, amplicon diagnostic for the presence of head-to-tail AdV DNA concatemers. Bottom, PCR analysis of genomic DNA from EGFP$^+$ HeLa cells exposed to *AAVS1*-specific TALENs and AdV.Δ1.donor^{S1} (gDNA EGFP$^+$). DNA from parental HeLa cells (gDNA HeLa cells) and nuclease-free water are negative controls; *in vitro*–generated head-to-tail AdV DNA (H-T AdV DNA) is the positive control. (**d**) Cumulative data on randomly selected EGFP$^+$ HeLa cells isolated from cultures co-transduced with AdV.Δ1.donor^{S1}, AdV.TALEN-L^{S1} and AdV.TALEN-R^{S1} (**Supplementary Fig. 7**). The frequencies of EGFP$^+$ clones representing the different integrant classes are plotted as in **Figure 1c**.

CRISPR-Cas9–induced AdV gene targeting to those obtained using plasmid donors as templates (**Supplementary Fig. 10b**).

PCR analysis of donor DNA-*AAVS1* junctions demonstrated higher target-site specificity of AdV over plasmid HR substrates even when using the presumably more promiscuous[26-28], yet more versatile, CRISPR-Cas9 system[25] (**Fig. 4d** and **Supplementary Fig. 11**). Furthermore, as previously observed for IDLV donor DNA, plasmid donor DNA resulted in intermolecular illegitimate recombination events (**Supplementary Fig. 11a**). Of note, these data represent, to the best of our knowledge, the first demonstration of the utility of the CRISPR-Cas9 system for viral vector–mediated exogenous DNA targeting in human cells.

There are precedents for the delivery and persistence of bacterial DNA in mammalian cells exposed to both nonviral and viral vectors[29,30]. Besides their unpredictable structures, these prokaryotic DNA footprints are also undesirable owing to their immunostimulatory and methylation-prone nucleotide patterns (for example, CpG motifs). We probed for bacterial DNA in cell populations genetically modified by AdV- and plasmid-based gene targeting, and we detected Kan^R DNA exclusively in the latter case (**Supplementary Fig. 12**).

Capped AdV donor DNA reduces illegitimate recombination

Finally, we tested whether the protein cap of AdV genomes is a determinant of the specificity and fidelity of the AdV gene-targeting process. To this end, we generated AdV.Δ2.donor$^{S1/T\text{-}TS}$ and used it together with AdV.Δ2.TALEN-L^{S1} and AdV.Δ2.TALEN-R^{S1} to transduce HeLa cells. The AdV.Δ2.donor$^{S1/T\text{-}TS}$ has the same genetic makeup as the AdV.Δ2.donor^{S1} except that its donor DNA payload is flanked by recognition sequences for the *AAVS1*-specific TALENs (T-TS) (**Fig. 5a**). The AdV.Δ2.TALEN-L^{S1} and AdV.Δ2.TALEN-R^{S1} differ from AdV.TALEN-L^{S1} and AdV.TALEN-R^{S1}, respectively, in that they are deleted simultaneously in *E1* and *E2A*. In striking difference with the results obtained with the T-TS–negative AdV (**Fig. 3d** and **Supplementary Fig. 7**), PCR screening of thirty EGFP$^+$ HeLa cell clones expanded from cultures exposed to TALENs and AdV.Δ2.donor$^{S1/T\text{-}TS}$ revealed that seven of these lines lacked HR-derived *AAVS1*–exogenous DNA centromeric junctions (**Fig. 5a**). Control experiments established the release of donor$^{S1/T\text{-}TS}$ DNA from the AdV.Δ2.donor$^{S1/T\text{-}TS}$ genome in transduced cells (**Supplementary Fig. 13**). These results confirm that HR substrates delivered in the context of protein-capped AdV genomes lead to more precise gene targeting when compared to HR substrates transferred in free-ended linear templates (**Fig. 5b**). We obtained similar results following AdV transduction experiments in reporter cells (**Supplementary Results** and **Supplementary Fig. 14**) and differentiating myoblasts (**Supplementary Results** and **Supplementary Fig. 15**).

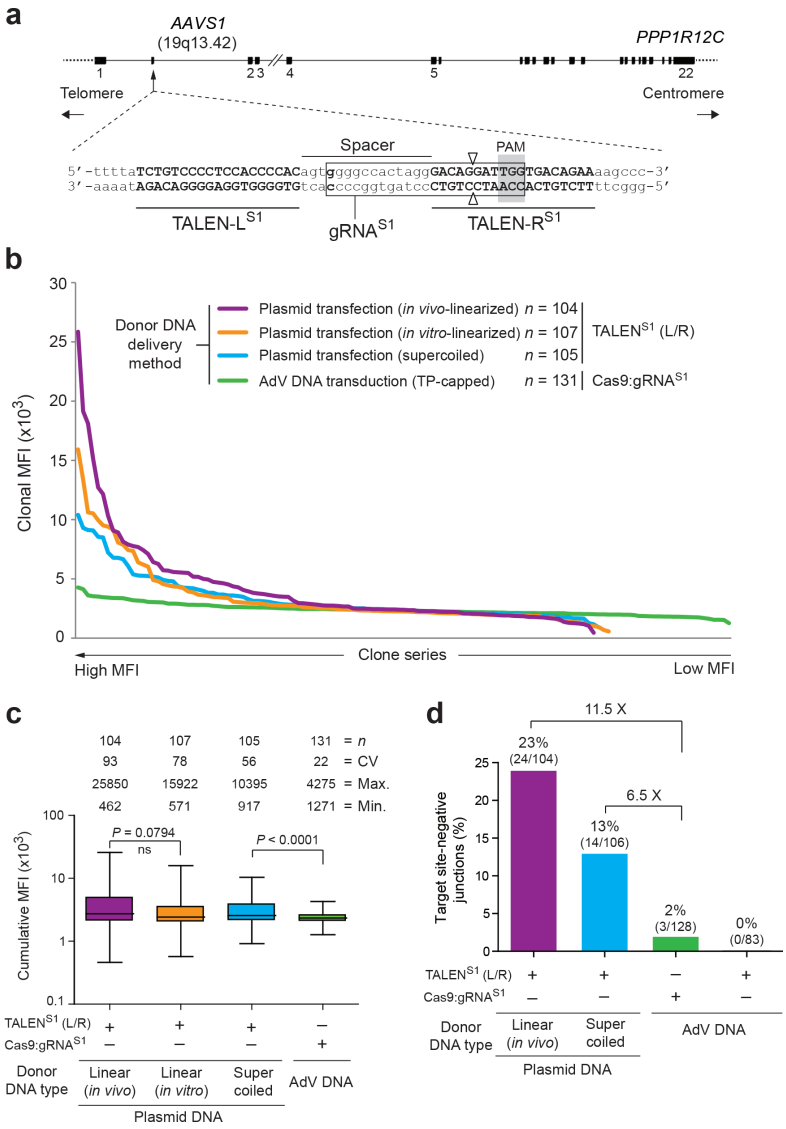


Figure 4. Nuclease-induced gene targeting with AdV- versus plasmid-mediated delivery of HR substrates.

(**a**) *AAVS1* recognition sequences for the TALEN and Cas9:gRNA complexes are drawn in relation to the *PPP1R12C* locus in which they are embedded. The target sites of the TALEN pair and the RNA-guided CRISPR-Cas9 nuclease are shown in upper case and boxed, respectively. The protospacer-adjacent motif (PAM) is shaded. Open vertical arrowheads indicate the Cas9 cleavage site. (**b**) MFI distribution of randomly selected EGFP$^+$ HeLa cell clones genetically modified by Cas9:gRNA^{S1} and AdV.Δ2.donor^{S1} (TP capped) or by TALEN-L^{S1}:TALEN-R^{S1} (L/R) and either pAdV.donor^{S1} (supercoiled), PacI linearized pAdV.donor^{S1} (*in vitro* linearized) or pAdV.donor$^{S1/T\text{-}TS}$ (*in vivo* linearized). (**c**) Box plot of cumulative MFI ▶

◀ values corresponding to panel **b**. Whiskers, minimum and maximum; ns, not significant (two-tailed *t*-test), *n* indicates number of clones. (**d**) Frequencies of EGFP^{+} clones lacking the *AAVS1*-donor DNA 'telomeric' junction, as identified by PCR screening of HeLa cell clones (**Supplementary Figs. 7** and **11**). The data correspond to the clones analyzed in **b**,**c** and **Supplementary Figure 7**.

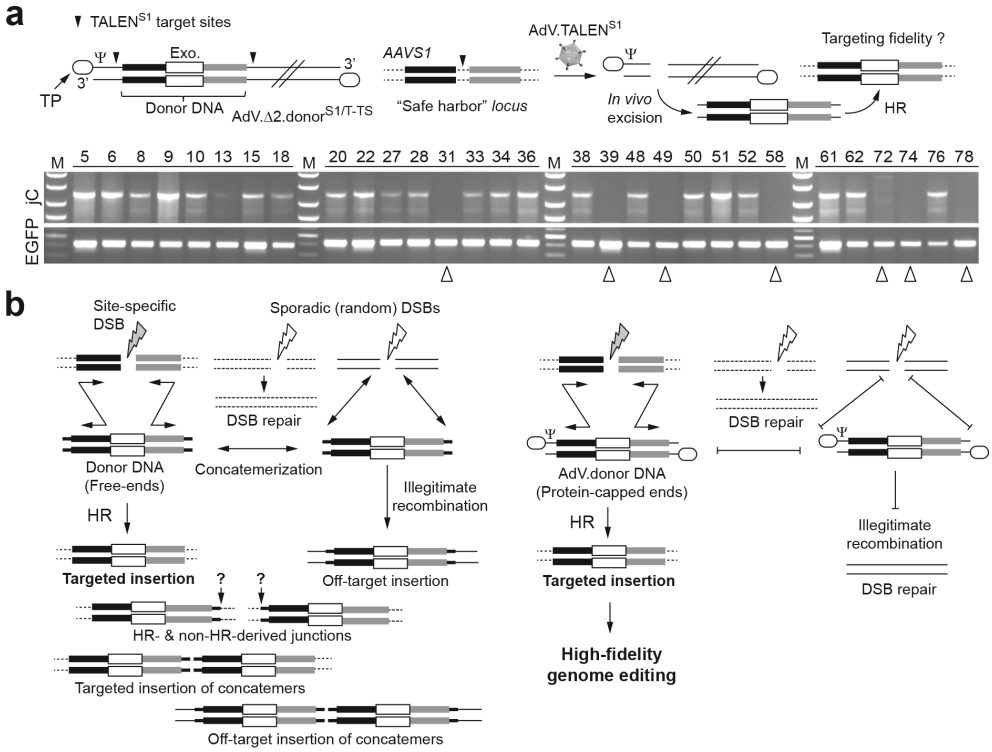


Figure 5. The protein cap of AdVs contributes to gene-targeting specificity and accuracy.

(**a**) The effect of donor DNA excision from the protein-capped AdV genomes on gene targeting frequency. Top, experimental strategy for the TALEN-mediated excision of HR templates from AdV DNA in transduced cells. AdV.Δ2.donor$^{S1/T-TS}$, AdV carrying donor^{S1} DNA framed by target sequences for the *AAVS1*-specific TALENs (solid vertical arrowheads); Open oval, terminal protein (TP) covalently attached to the 5′ termini of AdV DNA; Ψ, AdV packaging signal; Exo., exogenous DNA. Bottom, PCR screening of EGFP^{+} HeLa cell clones to detect DNA junctions formed by HR events between AdV.Δ2.donor$^{S1/T-TS}$ DNA and the 'centromeric' side of the *AAVS1* locus. Open arrowheads, clones harboring integrants lacking HR-derived centromeric junctions. (**b**) Model for high-fidelity genome editing based on site-specific DSBs and AdV-delivered donor DNA. Left, donor DNA with free ends are sensed as broken DNA and rerouted to off-target or sporadic DSBs. Precision can be further compounded by the chromosomal insertion of concatemeric vector DNA forms. Right, protein-capped AdV genomes make interactions between donor templates and off-target or sporadic DSBs less probable.

Discussion

The overall performance of genome-editing technology depends on its efficiency as well as on its specificity and accuracy. On-target specificity is being actively investigated, with significant efforts devoted to not only examining but also reducing off-target DNA cleaving activities of artificial nucleases[5,26-28,31-34]. The contribution of the donor DNA component to the specificity and accuracy with which different exogenous HR substrates are inserted at a target site has, comparatively, received scarcer scrutiny. It is known, however, that commonly used viral vector–borne donor templates lead to off-targeted events, presumably owing to their capture by random chromosomal DSBs[11,23,29]. In fact, IDLV and rAAV genomes are being exploited as tags for mapping the distribution of DSBs at a genome-wide scale[11,35,36]. Experimental evidence indicates that linear DNA molecules, by resembling DSB repair intermediates, are targets for illegitimate recombination mechanisms regardless of their viral or nonviral origin[37,38]. The ensuing processing of these molecules by DNA repair factors can result in their off-target chromosomal insertion as single or multiple copies. These outcomes reduce the predictability of transgene expression and gene repair in target cell populations. In principle, viral vector–mediated genome editing might be refined by adding, 'outside' the targeting HR module, a heterologous cassette conferring negative selection against cells with insertions at off-target sites. However, purging cells with targeted insertions formed by a combination of nonhomologous end joining and HR and/or with irregular concatemeric footprints would be less certain. Besides being complex and measurably 'leaky', negative selection strategies also reduce the effective packaging capacity of vector particles. The overall fidelity of the IDLV-based genome-editing process may also suffer from the intrinsic genetic 'plasticity' of lentiviruses. For instance, the error rates of retroviral reverse transcriptases are orders of magnitude higher than those of dsDNA viral polymerases[39].

The frequencies of genetic modification achieved by DSB-induced AdV gene targeting were lower than those obtained with the IDLV platform but were within the range of or above those reached with plasmid donors co-delivered with nucleases or reported for single-stranded rAAV templates[40]. Regardless, the high packaging capacity of AdV particles coupled to the availability of robust positive selection strategies can, if necessary, be exploited to rapidly enrich for the AdV-modified population. Notably, the precision of the AdV gene-targeting procedure is expected to reduce the dependency on time-consuming screening of genetically modified populations to identify properly targeted cells.

On the basis of our findings, we put forward the view that the numerous efforts

devoted to minimizing off-target activity of sequence-specific nucleases[5,31-34] should be complemented with those aiming at identifying HR substrates whose features maximize on-target and accurate insertion of foreign DNA. The development of these optimized HR templates is expected to promote fundamental and applied research activities dependent on the precise manipulation of mammalian genomes.

Methods

Cells

The HeLa cells (American Type Culture Collection) and EGFP-positive H27 clone derivative[41] were cultured in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 5% fetal bovine serum (FBS; Invitrogen). The HEK 293T lentiviral vector producer cells were maintained in DMEM containing 10% FBS, whereas the AdV packaging cell lines PER.C6 (ref. 42) and PER.E2A[43] were cultured in DMEM supplemented with 10% FBS and 10 mM $MgCl_2$ (Sigma-Aldrich) in the absence and in the presence of 250 μg/ml of Geneticin (Invitrogen), respectively. These cell types were kept in a humidified atmosphere containing 10% CO_2. The origin of and the culture conditions for the human myoblasts have been detailed elsewhere[44,45]. The cell batches used to generate the viral vector preparations and to carry out the experiments were mycoplasma free.

Recombinant DNA

The complete and annotated DNA sequences of lentiviral vector shuttle plasmids AP45_pLV.donor^{EGFP} and AQ25_pLV.donor^{S1} can be retrieved via GenBank accession numbers KF419293 and KF419294, respectively. The AdV shuttle plasmid AQ60_pAdV.donor^{EGFP} contains as foreign DNA the same targeting module as that in AP45_pLV.donor^{EGFP}. The AdV shuttle plasmid pAdV.donor$^{S1/T\text{-}TS}$ was constructed by inserting 'upstream' and 'downstream' of the donor^{S1} DNA module in pSh.AAVS1.eGFP[14] (herein named pAdV.donor^{S1}) two annealed oligodeoxyribonucleotides containing bipartite target sequences for the *AAVS1*-specific TALENs (T-TS). A similar approach based on inserting into pAdV.donor^{S1} a direct repeat of *FRT* sites in place of the T-TS sequences was pursued in parallel. These maneuvers resulted in pAdV.donor$^{S1/FRT}$. The AdV molecular clones AL25_pAdV.ΔE1.donor^{S1}.F^{50} and AL27_pAdV.ΔE1ΔE2A.donor^{S1}.F^{50} were assembled by HR in *Escherichia coli* strains[46] BJ5183$^{pAdEasy\text{-}1.50}$ and BJ5183$^{pAdEasy\text{-}2.50}$, respectively, transformed with MssI-treated pAdV.donor^{S1}. The AdV molecular clones pAdV.ΔE1ΔE2A.donor$^{S1/T\text{-}TS}$.F^{50} and pAdV.ΔE1ΔE2A.donor$^{S1/FRT}$.F^{50} were built by HR following the transformation of the latter cells with MssI-digested plasmids pAdV.donor$^{S1/T\text{-}TS}$ and pAdV.donor$^{S1/FRT}$,

respectively. The FP635-encoding AdV molecular clone AR02_pAdV.ΔE1.ΔE2A.donor^{EGFP}.F^{50} was generated by HR in BJ5183$^{pAdEasy-2.50}$ cells[46] transformed with MssI-digested AQ60_pAdV.donor^{EGFP}. The AdV shuttle plasmid AG03_pAdV.CMV.TALEN-L^{EGFP} and AF59_pAdV.CMV.TALEN-R^{EGFP} were constructed by inserting into the multiple cloning site of AQ17_pAdV.MCS.SV40pA the CMV promoter linked to the ORFs of TALEN-L^{EGFP} and TALEN-R^{EGFP}, respectively. The TALEN-L^{EGFP} and TALEN-R^{EGFP} proteins (**Supplementary Fig. 14a**) were custom designed (GeneArt) to recognize the EGFP half target sites 5′-TGAACTTCAAGATCCGCCA-3′ and 5′-TCGGCGAGCTGCACGCTGC-3′, respectively, and cleave within their 14-bp intervening spacer sequence (**Supplementary Fig. 14a**). Next, the full-length AdV molecular clones AF50_pAdV.ΔE1.TALEN-L^{EGFP}.F^{50} and AF52_pAdV.ΔE1.TALEN-R^{EGFP}.F^{50} were assembled by HR in *E. coli* by transforming strain BJ5183$^{pAdEasy-1.50}$ with MssI-treated AdV shuttle plasmids AG03_pAdV.CMV.TALEN-L^{EGFP} and AF59_pAdV.CMV.TALEN-R^{EGFP}, respectively.

The AdV shuttle plasmid pAdSh.PGK.Cas9 contains a humanized ORF encoding the *Streptococcus pyogenes* nuclease Cas9 under the transcriptional control of the *PGK1* promoter and the SV40 polyadenylation signal, whereas the AdV shuttle plasmid pAdSh.U6.gRNA^{S1} encodes a *U6* promoter–driven single guide RNA targeting the Cas9 protein to the human *AAVS1* locus[47]. The human codon-optimized *cas9* ORF and the RNA Pol III-dependent *gRNA^{S1}* expression unit have been published elsewhere[48] and were isolated from constructs hCas9 (Addgene plasmid 41815) and gRNA_AAVS1-T2 (Addgene plasmid 41818), respectively. Next, the AdV molecular clones pAdV^{Δ2}P.Cas9.F^{50} and pAdV^{Δ2}U6.gRNA^{S1}.F^{50} were generated by HR in BJ5183$^{pAdEasy-2.50}$ cells[46] following their transformation with MssI-treated pAdSh.PGK.Cas9 and pAdSh.U6.gRNA^{S1}, respectively[47]. Further details about DNA constructs generated for this study can be found in the **Supplementary Note**.

Production and titration of AdV vectors

The generation of the fiber-modified *E1*-deleted AdVs Ad.ΔE1.TALEN-L^{S1}.F^{50} and Ad.ΔE1.TALEN-R^{S1}.F^{50} (herein referred to as AdV.TALEN-L^{S1} and AdV.TALEN-R^{S1}, respectively) has been detailed elsewhere[46]. The same applies to the fiber-modified *E1*- and *E2A*-deleted AdVs Ad.ΔE1ΔE2A.TALEN-L^{S1}.F^{50} and Ad.ΔE1ΔE2A.TALEN-R^{S1}.F^{50} (herein named AdV.Δ2.TALEN-L^{S1} and AdV.Δ2.TALEN-R^{S1}, respectively)[16]. The productions of the fiber-modified *E1*-deleted AdV AdV.Δ1.donor^{S1} and of its *E1*- plus *E2A*-deleted derivative AdV.Δ2.donor^{S1}, were initiated by transfecting PER.C6 and PER.E2A cells with PacI-linearized AL25_pAdV.ΔE1.donor^{S1}.F^{50} and AL27_pAdV.ΔE1ΔE2A.donor^{S1}.F^{50}, respectively. The productions of the fiber-modified, EGFP-encoding, donor AdVs AdV.Δ2.donor$^{S1/T-TS}$ and AdV.Δ2.donor$^{S1/FRT}$ as well as

that of the fiber-modified, *E1*- and *E2A*-deleted, FP635-encoding, donor AdV AdV.Δ2.donor^{EGFP} were started by transfecting PER.E2A cells with PacI-treated pAdV.ΔE1ΔE2A.donor$^{S1/T\text{-}TS}$.F^{50}, pAdV.ΔE1ΔE2A.donor$^{S1/FRT}$.F^{50} and AR02_pAdV.ΔE1.ΔE2A.donor^{EGFP}.F^{50}, respectively. The *E1*- plus *E2A*-complementing PER.E2A cells were also used for rescuing and propagating the fiber-modified AdVs AdV.Cas9 and AdV.gRNA^{S1} following their transfection with the PacI-linearized molecular clones pAdV$^{\Delta 2}$P.Cas9.F^{50} and pAdV$^{\Delta 2}$U6.gRNA^{S1}.F^{50}, respectively. Finally, the generation of the first-generation and fiber-modified AdVs AdV.TALEN-L^{EGFP} and AdV.TALEN-R^{EGFP}, was carried out in *E1*-only complementing cells transfected with PacI-digested AF50_pAdV.ΔE1.TALEN-L^{EGFP}.F^{50} and AF52_pAdV.ΔE1.TALEN-R^{EGFP}.F^{50}, respectively.

The DNA transfection–mediated rescue of AdV particles in packaging cell lines and their subsequent propagation and purification were performed essentially as described previously[46, 49]. The isolation and restriction fragment length analysis of AdV.Δ2.donor$^{S1/T\text{-}TS}$ and AdV.Δ2.donor$^{S1/FRT}$ DNA was carried out as detailed elsewhere[46, 49]. The titers of the various reporter-encoding AdV stocks, expressed in terms of transducing units (TU) per ml, were determined through limiting dilutions on HeLa indicator cells seeded at a density of 8×10^4 cells per well of 24-well plates. At 3 d post transduction, frequencies of reporter-positive cells were measured by reporter-directed flow cytometry. The titers of the reporter-negative AdV preparations were established by $TCID_{50}$ assays in complementing cells and by fluorometric quantification of genome-containing vector particles (VP) per milliliter as described elsewhere[46,49].

Production and titration of lentiviral vectors

The generation of the vesicular stomatitis virus glycoprotein G (VSV-G)-pseudotyped lentiviral vectors LV.ZFN-1^{EGFP} and LV.ZFN-2^{EGFP} has been described before[50]. The generation of the VSV-G-pseudotyped vectors IDLV.donor^{EGFP} and IDLV.donor^{S1} was carried out by transient transfections of 293T cells with shuttle plasmids AP45_pLV.donor^{EGFP} and AQ25_pLV.donor^{S1}, respectively, together with packaging construct[51] AM16_psPAX2.IN^{D116N} and pseudotyping construct pLP/VSVG (Invitrogen) as detailed elsewhere[51,52]. The physical particle titers of lentiviral vector preparations were determined by using the RETRO-TEK HIV-1 p24 ELISA kit as specified by the manufacturer (Gentaur Molecular Products). Titers of these vector stocks expressed in terms of TU/ml were derived by using a conversion factor of 2,500 TU/ng of HIV-1 p24^{gag} protein.

Transduction experiments

The IDLV.donor^{EGFP} transductions on cultures of EGFP$^+$ H27 indicator cells

were carried out as follows. Eighty-thousand cells were seeded in wells of 24-well plates (Greiner Bio-One). The next day, IDLV.donor^{EGFP} was added at a multiplicity of infection (MOI) of 45 TU/cell together with the LV.ZFN-1^{EGFP} and LV.ZFN-2^{EGFP} vectors each applied at an MOI of 8 TU/cell. Parallel H27 cultures that were either untreated or were incubated exclusively with IDLV.donor^{EGFP} at an MOI of 45 TU/cell, served as controls. Next, after an extensive 5-week subculturing period, the frequencies of reporter-positive and reporter-negative H27 cell populations were monitored and quantified, respectively, by two-color fluorescence microscopy and flow cytometry.

Long-term transduction experiments on human myoblasts were initiated by seeding 2×10^5 cells per well in 24-well plates. The next day, the cells were incubated with AdV.TALEN-L^{S1} (2.5 TU/cell) and AdV.TALEN-R^{S1} (2.5 TU/cell) mixed with IDLV.donor^{S1} (10 TU/cell) or with AdV.Δ2.donor^{S1} (10 TU/cell). The frequencies of IDLV.donor^{S1}- and AdV.Δ2.donor^{S1}-modified myoblasts were determined at 27 and 45 d post transduction, respectively.

Long-term transduction experiments on HeLa cells were started by seeding 8×10^4 cells per well in 24-well plates. After an overnight incubation period, the cells were exposed to AdV.TALEN-L^{S1} (3 TU/cell) and AdV.TALEN-R^{S1} (3 TU/cell) together with AdV.Δ1.donor^{S1} (6 TU/cell). The frequencies of stably transduced cells were measured by flow cytometry at 24 d post transduction.

To investigate the effect of TALEN-mediated donor DNA excision on the rate of gene targeting, 8×10^4 HeLa cells, seeded 1 d before, were transduced with AdV.Δ2.TALEN-L^{S1} (1.5 TU/cell) and AdV.Δ2.TALEN-R^{S1} (1.5 TU/cell) together with AdV.Δ2.donor$^{S1/T\text{-}TS}$ (3 TU/cell). The frequencies of stably transduced cells were established at 24 d post transduction by flow cytometry. The EGFP$^+$ populations present in the various long-term cultures were sorted, after which over 250 individual clones were randomly selected for expansion and molecular analyses.

The gene-targeting experiments on EGFP$^+$ H27 cells with the FP635-encoding vector AdV.Δ2.donor^{EGFP} were carried out as follows. Two-hundred thousand cells were seeded in wells of 24-well plates. The next day, AdV.Δ2.donor^{EGFP} was added at an MOI of 4 TU/cell together with 1:1 mixtures of AdV.TALEN-L^{EGFP} and AdV.TALEN-R^{EGFP} applied at total doses of 6.0×10^3, 2.4×10^3 and 9.6×10^2 VP/cell. Controls consisted of H27 cells, HeLa cells exposed for 3 days to AdV.Δ2.donor^{EGFP} (0.5 TU/cell) and H27 cells transduced with AdV.Δ2.donor^{EGFP} (4 TU/cell) plus AdV.TALEN-L^{EGFP} at 6.0×10^3, 2.4×10^3 and 9.6×10^2 VP/cell. At 48 h post transduction, the various H27 cultures received fresh regular medium, and 1 d later they started to be subcultured at a rate of about twice per week in order to remove the episomal vector DNA. The frequencies of the various types of H27 cell populations were determined

by two-color flow cytometry at 37 d post transduction.

Flow cytometry and light microcopy

The measurement of transgene expression parameters (i.e., frequencies of reporter-positive and reporter-negative target cells, mean fluorescence intensities and coefficients of variation) were determined by using a BD LSR II flow cytometer (BD Biosciences). Data were analyzed with the aid of BD FACSDiva 6.1.3 software (BD Biosciences) or FlowJo 7.2.2 (Tree Star). Mock-transduced target cells were used to set background fluorescence levels. At least 10,000 viable single cells were analyzed per sample. The light microscopic analyses were carried out with an IX51 inverse fluorescence microscope equipped with a XC30 Peltier-cooled digital color camera (Olympus). The images were processed with the aid of Cell^{F} 3.4 imaging software (Olympus).

Functional validation of CRISPR-Cas9 complexes delivered into target cells by AdVs

Fifty-thousand HeLa cells were transduced with AdV.Cas9 alone (300 $TCID_{50}$/cell) or with 1:1 mixtures of AdV.Cas9 and AdV.gRNA^{S1} applied at total vector doses of 60, 120, 180, 240 and 300 $TCID_{50}$/cell. At 3 d post transduction, chromosomal DNA was extracted from the target cells by using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's instructions. Next, *AAVS1*-specific PCR amplifications and T7 endonuclease I–based genotyping assays were carried out essentially as previously described[49].

Designer nuclease–mediated gene targeting of linear and supercoiled plasmid donor DNA

HeLa cells were seeded at a density of 6.5×10^4 cells per well of 24-well plates. The next day, the cells were transfected with DNA mixtures consisting of 100 ng of 1383.pVAX.AAVS1.TALEN.L-94 (ref. 16), 100 ng of 1384.pVAX.AAVS1.TALEN.R-95 (ref. 16) and 200 ng of *AAVS1*-targeting AdV donor DNA plasmids. The targeting constructs were pAdV.donor^{S1}, PacI-linearized pAdV.donor^{S1} and pAdV.donor$^{S1/T\text{-}TS}$. Controls were provided by transfecting HeLa cells with 200 ng of 1383.pVAX.AAVS1.TALEN.L-94 mixed together with 200 ng of pAdV.donor^{S1}, PacI-linearized pAdV.donor^{S1} or pAdV.donor$^{S1/T\text{-}TS}$. The completion of the PacI digestions was confirmed by agarose gel electrophoreses and ethidium bromide staining. Each of the plasmid mixtures were diluted in 50 µl of 150 mM NaCl and received 1.32 µl of a 1 mg/ml polyethylenimine solution under vigorous shaking for about 10 s. After a 20-min incubation period at room temperature, the resulting polycation-DNA complexes

were directly added into the culture medium. After 7 h, the transfection mixtures were removed and fresh culture medium was added. The resulting cell populations were subsequently subjected to subculturing for 3 weeks, after which cells stably expressing EGFP in these populations were individually sorted by flow cytometry into wells of 96-well plates. Viable single cell-derived clones corresponding to the various experimental settings were randomly selected for transgene expression and integration status analysis.

Gene targeting of AdV donor DNA by using the RNA-guided nuclease Cas9

HeLa cells were seeded at a density of 8.0×10^4 cells per well in 24-well plates. The following day, the cells were transduced with AdV.Cas9 (150 $TCID_{50}$/cell), AdV.gRNA^{S1} (50 $TCID_{50}$/cell) and AdV.Δ2.donor^{S1} (10 TU/cell). To serve as negative controls, HeLa cells were either mock-transduced or were transduced with AdV. Cas9 (150 $TCID_{50}$/cell) and AdV.Δ2.donor^{S1} (10 TU/cell). At 3 d post transduction, mock- and vector-transduced HeLa cells started to be subcultured. At 17 d post transduction, EGFP stably expressing cells were individually sorted by flow cytometry into wells of 96-well plates. Viable single cell-derived clones isolated from cultures initially exposed to AdV.Cas9, AdV.gRNA^{S1} and AdV.Δ2.donor^{S1} were randomly selected for transgene expression and integration status analysis.

Cell sorting and clonal expansion

Flow cytometry–assisted cell sorting was done after the removal of donor DNA-containing episomes from long-term HeLa and human myoblast cultures. The EGFP$^+$ cells were collected in 1:1 mixtures of regular medium containing 2× penicillin-streptomycin (Invitrogen) and FBS. Next, the sorted cells were individually seeded in wells of 96-well plates (Greiner Bio-One) at a density of 0.3 cells per well in their respective medium supplemented with 50 μM of α-thioglycerol (Sigma-Aldrich) and 20 nM of bathocuprione disulphonate (Sigma-Aldrich) to increase cloning efficiency[53]. Finally, over 250 individual clones were randomly selected for expansion and molecular analyses.

Genomic DNA extraction

Genomic DNA was extracted from cell populations and clones essentially as described before[54]. In brief, the cells were collected and incubated overnight at 55 °C in 500 μl of lysis buffer (100 mM Tris-HCl, pH 8.5, 5 mM EDTA, 0.2% sodium dodecyl sulfate and 200 mM NaCl) supplemented with freshly added proteinase K (Thermo Scientific) at a final concentration of 100 ng/ml. The cell lysates were extracted twice with a buffer-saturated phenol:chloroform:isoamyl alcohol mixture (25:24:1) and

once with chloroform. Next, the genomic DNA was precipitated by the addition of 2.5 volumes of absolute ethanol and 0.5 volumes of 7.5 M ammonium acetate (pH 5.5). After washing with 70% ethanol, the DNA pellets were air dried and dissolved in 100 µl of Tris-EDTA buffer (10 mM Tris, pH 8.0, and 1 mM EDTA) supplemented with RNase A (Thermo Scientific) at a final concentration of 100 µg/ml. The genomic DNA of $EGFP^+$ clones derived from cultures transfected with plasmid DNA as well as that derived from cultures cotransduced with AdV.Cas9, $AdV.gRNA^{S1}$ and $AdV.\Delta 2.donor^{S1}$ was extracted by using the DNeasy Blood & Tissue Kit according to the protocol provided by the manufacturer. The genomic DNA samples of $FP635^+$ H27 clones were also isolated by deploying the DNeasy Blood & Tissue Kit.

PCR analyses of gene-targeting experiments with plasmid, IDLV and AdV donor DNA

The composition of the PCR mixtures and cycling parameters used for the analyses of genome-modification events are specified in the **Supplementary Tables 1** and **2**, respectively.

Statistics

Data sets were analyzed by using the GraphPad Prism 5 software package and evaluated for significance by applying unpaired two-tailed Student's *t*-tests ($P < 0.05$ considered significant).

Southern blot analyses

Genomic DNA was isolated from individual human myoblast clones and from parental human myoblast populations according to aforementioned organic solvent-based protocol. Next, DNA samples (10 µg or 20 µg) were digested overnight with NcoI (Thermo Scientific) and were resolved through a 0.8% agarose gel in 1× Tris-acetate-EDTA buffer. The DNA was transferred by capillary action onto an Amersham Hybond-XL membrane (GE Healthcare Life Sciences) using standard Southern blot techniques. The 635- and 790-bp DNA probes specific for the centromeric *AAVS1* arm and for the transgene in donor DNA templates, respectively, were isolated from source plasmids following standard restriction enzymes digestions and agarose gel electrophoreses. The purified DNA probes were subsequently radiolabeled with [α-^{32}P]dATP (GE Healthcare Life Sciences) by using the DecaLabel DNA labeling Kit following the manufacturer's instructions (Thermo Scientific). Prior to its deployment, the radiolabeled probe were separated from unincorporated dNTPs through size-exclusion chromatography in Sephadex-50 columns (GE Healthcare Life Sciences). A Storm 820 Phosphoimager (Amersham Biosciences) was used for

the detection of the probe-hybridized DNA. The images were acquired by using the Storm Scanner Control 5.03 software and were processed with the aid of ImageQuant Tools 3.0 software (both from Amersham Biosciences).

COBRA-FISH karyotyping

The COBRA-FISH karyotyping of HeLa target cells was done according to a published protocol[55].

Analysis of *in vivo* excision of AdV donor DNA

The experiments designed to assess TALEN-mediated excision of donor^{S1} DNA from AdV backbones were performed as follows. Eighty-thousand HeLa cells were seeded in wells of 24-well plates, and the following day they were cotransduced with AdV.Δ2.TALEN-L^{S1} (1.5 TU/cell), AdV.Δ2.TALEN-R^{S1} (1.5 TU/cell) and AdV.Δ2.donor$^{S1/T\text{-}TS}$ (3 TU/cell) or were cotransduced with 2 gene switch–activating units per cell of hcAd.FLPe.F50 (ref. 56) and AdV.Δ2.donor$^{S1/FRT}$ (3 TU/cell). Control samples were provided by parallel HeLa cell cultures exposed exclusively to AdV.Δ2.donor$^{S1/T\text{-}TS}$ (3 TU/cell) or to AdV.Δ2.donor$^{S1/FRT}$ (3 TU/cell). At 72 h post transduction, extrachromosomal DNA was isolated essentially as described previously[57], after which 2-μl DNA samples were subjected to PCR. The PCR mixtures consisted of 0.4 μM of primer #997 (5′-GCACTGAAACCCTCAGTCCTAGG-3′), 0.4 μM of primer #998 (5′-CGGCGTTGGTGGAGTCC-3′), 0.1 mM of each dNTP (Invitrogen), 1 mM $MgCl_2$ (Promega), 1× Colorless GoTaq Flexi Buffer (Promega) and 2.5 U of GoTaq Flexi DNA polymerase (Promega). Next, 50-μl PCR mixtures were subjected to an initial 2-min denaturation period at 95 °C, followed by 30 cycles of 30 s at 95 °C, 30 s at 62 °C and 45 s at 72 °C. The reactions were terminated by a final extension period of 5 min at 72 °C. The detection of the resulting PCR products was performed by conventional agarose gel electrophoresis.

Investigation of vector-vector illegitimate recombination in a myoblast-to-myotube cellular differentiation model

Two-hundred thousand human myoblasts were seeded in wells of 24-well plates. Two days later, the cells were transduced with 10 TU/cell of AdV.Δ2.donor^{S1} or with 10 TU/cell of IDLV.donor^{S1} mixed with AdV.TALEN-L^{S1} plus AdV.TALEN-R^{S1} at MOIs of 2.5 TU/cell each or mixed with AdV.TALEN-R^{S1} alone at an MOI of 5 TU/cell. Mock-transduced myoblasts served as negative controls. At 3 d post transduction, the growth medium[44] was removed, the cells were washed twice with 2 ml of PBS, and mitogen-free differentiation medium[44] was added onto the different myoblast cultures. Five days later, cellular DNA was extracted from the various

cultures containing post-mitotic syncytial myotubes. Next, PCR analysis of the cellular DNA was carried out for detecting head-to-tail vector DNA forms using as negative controls nuclease-free water and DNA extracted from the mock-transduced myotube cultures. Positive controls were obtained by using *in vitro*–assembled head-to-tail AdV DNA templates (**Fig. 3c** and **Supplementary Fig. 4b**) and DNA isolated from IDLV.donor^{S1}-transduced myoblast clone 10. The integrity of the various DNA templates was controlled for by parallel DMD-specific PCR amplifications.

Acknowledgments

The authors thank J. Liu for her technical assistance and L.P. Pelascini for the generation of the lentiviral vector stocks encoding ZFNs. We also thank A. Recchia (University of Modena and Reggio Emilia) for providing us with plasmid pSh.AAVS1.eGFP; R. Hoeben and D. Baker for critically reading the manuscript; K. Szuhai and D. de Jong for the COBRA-FISH karyotyping; and M. Rabelink for the p24^{gag} ELISA measurements. This work was supported by the Prinses Beatrix Spierfonds (grant W.OR11-18 to M.A.F.V.G.), the European Community's 7th Framework Programme for Research and Technological Development (PERSIST grant 222878 to T.C. and M.A.F.V.G.) and the European Community's ERASMUS Programme (grant PLISBOA02 to S.F.D.H.).

Additional information

Supplementary information

Supplementary text and figures are available at the online version of the paper (http://www.nature.com/nmeth/journal/v11/n10/extref/nmeth.3075-S1.pdf).

Accession codes

GenBank/EMBL/DDBJ: AP45_pLV.donorEGFP, KF419293; AQ25_pLV.donorS1, KF419294. Plasmids available at Addgene: hCas9, 41815; gRNA_AAVS1-T2, 41818.

Author contributions

M.H. and I.M. contributed equally to this work. M.H. and I.M. generated reagents and performed most of the experiments with the help of S.F.D.H., J.M.J. and M.A.F.V.G. T.C. generated and validated *AAVS1*-specific TALENs and *EGFP*-specific ZFNs. M.H., I.M., J.M.J. and M.A.F.V.G. designed the experiments and analyzed the data. M.A.F.V.G. conceived of and initiated the research. M.A.F.V.G. wrote the manuscript with the help from all authors.

Competing financial interests

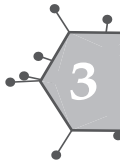
The authors declare competing financial interests: details are available in the online version of the paper.

References

1. Capecchi, M.R. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet.* **6**, 507–512 (2005).
2. Itzhaki, J.E. & Porter, A.C. Targeted disruption of a human interferoninducible gene detected by secretion of human growth hormone. *Nucleic Acids Res*. **19**, 3835–3842 (1991).
3. Brown, J.P., Wei, W. & Sedivy, J.M. Bypass of senescence after disruption of p21CIP1/WAF1 gene in normal diploid human fibroblasts. *Science* **277**, 831–834 (1997).
4. Russell, D.W. & Hirata, R.K. Human gene targeting by viral vectors. *Nat. Genet.* **18**, 325–330 (1998).
5. Segal, D.J. & Meckler, J.F. Genome engineering at the dawn of the golden age. *Annu. Rev. Genomics Hum. Genet.* **14**, 135–158 (2013).
6. Biasco, L., Baricordi, C. & Aiuti, A. Retroviral integrations in gene therapy trials. *Mol. Ther.* **20**, 709–716 (2012).
7. Varga, C.M. *et al.* Quantitative comparison of polyethylenimine formulations and adenoviral vectors in terms of intracellular gene delivery processes. *Gene Ther.* **12**, 1023–1032 (2005).
8. Wanisch, K. & Yáñez-Muñoz, R.J. Integration-deficient lentiviral vectors: a slow coming of age. *Mol. Ther.* **17**, 1316–1332 (2009).
9. Cornu, T.I. & Cathomen, T. Targeted genome modifications using integrase-deficient lentiviral vectors. *Mol. Ther.* **15**, 2107–2113 (2007).
10. Lombardo, A. *et al.* Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.* **25**, 1298–1306 (2007).
11. Gabriel, R. *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* **29**, 816–823 (2011).
12. Lombardo, A. *et al.* Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* **8**, 861–869 (2011).
13. Benabdallah, B.F. *et al.* Targeted gene addition of microdystrophin in mice skeletal muscle via human myoblast transplantation. *Mol. Ther. Nucleic Acids* **2**, e68 (2013).
14. Coluccio, A. *et al.* Targeted gene addition in human epithelial stem cells by zinc-finger nuclease-mediated homologous recombination. *Mol. Ther.* **21**, 1695–1704 (2013).
15. Mussolino, C. *et al.* TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res*. **42**, 6762–6773 (2014).
16. Holkers, M. *et al.* Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res.* **41**, e63 (2013).
17. Nightingale, S.J. *et al.* Transient gene expression by nonintegrating lentiviral vectors. *Mol. Ther.* **13**, 1121–1132 (2006).
18. Mátrai, J. *et al.* Hepatocyte-targeted expression by integrase-defective lentiviral vectors induces antigen-specific tolerance in mice with low genotoxic risk. *Hepatology* **53**, 1696–1707 (2011).
19. Miller, D.G., Petek, L.M. & Russell, D.W. Human gene targeting by adeno-associated virus vectors is enhanced by DNA double-strand breaks. *Mol. Cell. Biol.* **23**, 3550–3557 (2003).
20. Li, H. *et al*. *In vivo* genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, 217–221 (2011).
21. Ellis, B.L., Hirsch, M.L., Porter, S.N., Samulski, R.J. & Porteus, M.H. Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by

Food and Drug Administration-approved drugs. *Gene Ther.* **20**, 35–42 (2013).
22. Rahman, S.H. *et al.* The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. *Hum. Gene Ther.* **24**, 67–77 (2013).
23. Miller, D.G., Petek, L.M. & Russell, D.W. Adeno-associated virus vectors integrate at chromosome breakage sites. *Nat. Genet.* **36**, 767–773 (2004).
24. Gonçalves, M.A.F.V. & de Vries, A.A.F. Adenovirus: from foe to friend. *Rev. Med. Virol.* **16**, 167–186 (2006).
25. Terns, R.M. & Terns, M.P. CRISPR-based technologies: prokaryotic defense weapons repurposed. *Trends Genet.* **30**, 111–118 (2014).
26. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).
27. Hsu, P.D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
28. Cradick, T.J., Fine, E.J., Antico, C.J. & Bao, G. CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* **41**, 9584–9592 (2013).
29. Miller, D.G., Rutledge, E.A. & Russell, D.W. Chromosomal effects of adeno-associated virus vector integration. *Nat. Genet.* **30**, 147–148 (2002).
30. Chadeuf, G., Ciron, C., Moullier, P. & Salvetti, A. Evidence for encapsidation of prokaryotic sequences during recombinant adenoassociated virus production and their in vivo persistence after vector delivery. *Mol. Ther.* **12**, 744–753 (2005).
31. Miller, J.C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* **25**, 778–785 (2007).
32. Szczepek, M. *et al.* Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.* **25**, 786–793 (2007).
33. Doyon, Y. *et al.* Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods* **8**, 74–79 (2011).
34. Pattanayak, V., Ramirez, C.L., Joung, J.K. & Liu, D.R. Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. *Nat. Methods* **8**, 765–770 (2011).
35. Petek, L.M., Russell, D.W. & Miller, D.G. Frequent endonuclease cleavage at off-target locations *in vivo*. *Mol. Ther.* **18**, 983–986 (2010).
36. Miller, D.G. *et al.* Large-scale analysis of adeno-associated virus vector integration sites in normal human cells. *J. Virol*. **79**, 11434–11442 (2005).
37. Nakai, H. *et al.* Helper-independent and AAV-ITR-independent chromosomal integration of double-stranded linear DNA vectors in mice. *Mol. Ther*. **7**,101–111 (2003).
38. Choi, V.W., McCarty, D.M. & Samulski, R.J. Host cell DNA repair pathways in adeno-associated viral genome processing. *J. Virol.* **80**, 10346–10356 (2006).
39. Duffy, S., Shackelton, L.A. & Holmes, E.C. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet*. **9**, 267–276 (2008).
40. Khan, I.F., Hirata, R.K. & Russell, D.W. AAV-mediated gene targeting methods for human cells. *Nat. Protoc.* **6**, 482–501 (2011).
41. Gonçalves, M.A.F.V., van der Velde, I., Knaän-Shanzer, S., Valerio, D. & de Vries, A.A.F. Stable transduction of large DNA by high-capacity adenoassociated virus/adenovirus hybrid vectors. *Virology* **321**, 287–296 (2004).
42. Fallaux, F.J. *et al.* New helper cells and matched early region 1-deleted adenovirus vectors prevent generation of replication-competent adenoviruses. *Hum. Gene Ther.*

9, 1909–1917 (1998).
43. Havenga, M.J. *et al.* Serum-free transient protein production system based on adenoviral vector and PER.C6 technology: high yield and preserved bioactivity. *Biotechnol. Bioeng.* **100**, 273–283 (2008).
44. Cudré-Mauroux, C. *et al.* Lentivector-mediated transfer of Bmi-1 and telomerase in muscle satellite cells yields a Duchenne myoblast cell line with long-term genotypic and phenotypic stability. *Hum. Gene Ther.* **14**, 1525–1533 (2003).
45. Gonçalves, M.A.F.V. *et al.* Transcription factor rational design improves directed differentiation of human mesenchymal stem cells into skeletal myocytes. *Mol. Ther.* **19**, 1331–1341 (2011).
46. Janssen, J.M., Liu, J., Skokan, J., Gonçalves, M.A.F.V. & de Vries, A.A.F. Development of an AdEasy-based system to produce first- and secondgeneration adenoviral vectors with tropism for CAR- or CD46-positive cells. *J. Gene Med*. **15**, 1–11 (2013).
47. Maggio, I. *et al.* Adenoviral vector delivery of RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells. *Sci. Rep* **4**, 5105 (2014).
48. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
49. Holkers, M., Cathomen, T. & Gonçalves, M.A.F.V. Construction and characterization of adenoviral vectors for the delivery of TALENs into human cells. *Methods* doi:10.1016/j.ymeth.2014.02.017 (20 February 2014).
50. Pelascini, L.P.L. *et al.* Histone deacetylase inhibition rescues gene knockout levels achieved with integrase-defective lentiviral vectors encoding zincfinger nucleases. *Hum. Gene Ther. Methods* **24**, 399–411 (2013).
51. Pelascini, L.P.L., Janssen, J.M. & Gonçalves, M.A.F.V. Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. *Hum. Gene Ther.* **24**, 78–96 (2013).
52. Pelascini, L.P.L. & Gonçalves, M.A.F.V. Lentiviral vectors encoding zinc-finger nucleases specific for the model target locus *HPRT1*. *Methods Mol. Biol.* **1114**, 181–199 (2014).
53. Brielmeier, M. *et al.* Improving stable transfection efficiency: antioxidants dramatically improve the outgrowth of clones under dominant marker selection. *Nucleic Acids Res.* **26**, 2082–2085 (1998).
54. van Nierop, G.P., de Vries, A.A.F., Holkers, M., Vrijsen, K.R. & Gonçalves, M.A.F.V. Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease. *Nucleic Acids Res.* **37**, 5725–5736 (2009).
55. Szuhai, K. & Tanke, H.J. COBRA: combined binary ratio labeling of nucleic-acid probes for multi-color fluorescence in situ hybridization karyotyping. *Nat. Protoc.* **1**, 264–275 (2006).
56. Gonçalves, M.A.F.V. *et al.* Targeted chromosomal insertion of large DNA into the human genome by a fiber-modified high-capacity adenovirusbased vector system. *PLoS ONE* **3**, e3084 (2008).
57. Holkers, M., de Vries, A.A.F. & Gonçalves, M.A.F.V. Nonspaced inverted DNA repeats are preferential targets for homology-directed gene repair in mammalian cells. *Nucleic Acids Res.* **40**, 1984–1999 (2012).


3