## Universiteit Leiden

The handle http://hdl.handle.net/1887/37412 holds various files of this Leiden University dissertation.

**Author**: Holkers, Maarten
**Title**: The roles of adenoviral vectors and donor DNA structures on genome editing
**Issue Date**: 2016-01-26

# Differential integrity of *TALE* nuclease genes following adenoviral and lentiviral vector gene transfer into human cells

Maarten Holkers[1], Ignazio Maggio[1], Jin Liu[1], Josephine M. Janssen[1], Francesca Miselli[2], Claudio Mussolino[3], Alessandra Recchia[2], Toni Cathomen[3] and Manuel A.F.V. Gonçalves[1,]

*[1]Department of Molecular Cell Biology, Leiden University Medical Center, Eithovenweg 20, 2333 ZC Leiden, The Netherlands. [2]Department of Life Sciences, University of Modena and Reggio Emilia, Via Campi 287, 41125 Modena, Italy. [3]Laboratory of Cell and Gene Therapy, Center for Chronic Immunodeficiency, University Medical Center Freiburg, Engesserstr. 4, 79108 Freiburg, Germany.*

## ABSTRACT

The array of genome editing strategies based on targeted double-stranded DNA break (DSB) formation have recently been enriched through the introduction of transcription activator-like type III effector nucleases (TALENs). To advance the testing of TALE-based approaches, it will be crucial to deliver these custom-designed proteins not only into transformed cell types but also into more relevant, chromosomally stable, primary cells. Viral vectors are among the most effective gene transfer vehicles. Here, we investigated the capacity of lentiviral and adenoviral vectors to package and deliver functional *TALEN* genes into various human cell types. To this end, we attempted at assembling particles of these two vector classes, each of which encoding a monomer of a TALEN pair targeted to a bipartite sequence within the *AAVS1* "safe harbor" locus. Vector DNA analyses revealed that adenoviral vectors transferred intact *TALEN* genes, whilst lentiviral vectors failed to do so, as shown by their heterogeneously sized proviruses in target cells. Importantly, adenoviral vector-mediated *TALEN* gene delivery resulted in site-specific DSB formation at the intended *AAVS1* target site at similarly high levels in both transformed and nontransformed cells. In conclusion, we demonstrate that adenoviral, but not lentiviral, vectors constitute a valuable *TALEN* gene delivery platform.

## INTRODUCTION

The ability to design DNA-binding motifs that recognize pre-defined sequences of choice within complex eukaryotic genomes is opening a wide range of basic and applied research possibilities that involves the modulation of endogenous gene expression and chromosomal DNA editing. The latter aim has been mostly pursued by designing zinc-finger nucleases (ZFNs) comprising a sequence-specific array of synthetic zinc-finger motifs fused to the nuclease domain of the type IIS restriction endonuclease FokI (1,2). In addition, the recoding of natural homing endonucleases for heritable genome modification purposes has also been undergoing intense investigation (3,4). These artificial enzymes are tailored to generate double-stranded DNA breaks (DSBs) at pre-defined genomic sequences. As a consequence of their activity, targeted gene knock-out or chromosomal insertion of exogenous DNA can ensue following the activation and engagement of the error-prone non-homologous end-joining (NHEJ) or the error-free homologous recombination (HR) DNA repair pathways, respectively.

The recent discovery of the rules governing the recognition of specific DNA sequences by bacterial transcription activator-like type III effectors (TALEs) (5,6) has translated in a novel and straightforward methodology to construct proteins with customized DNA binding specificities. TALEs are proteins found in phytopathogenic bacteria of the genus *Xanthomonas*, whose main well-established function is to induce expression of specific host plant genes that enhance virulence. TALEs comprise an N-terminal translocation domain, a central region of tandem direct repeats that confers DNA binding specificity and a nuclear localization motif preceding a C-terminal transcriptional activation domain (7). The DNA binding domain of TALEs displays a unique architecture composed of a tandem array of typically 15.5 to 19.5 repeat units with *circa* 34 residues each (7). Often, the only distinguishing feature among the different repeats is a 2-amino acid hypervariable polymorphism at positions 12 and 13 dubbed "repeat-variable di-residue" (RVD). Crucially, it has been discovered that each individual RVD dictates the binding of the repeat in which it is embedded to a single nucleotide (5,6). Therefore, this one repeat to one nucleotide code directly establishes a simple rule governing TALE-DNA interactions. This insight combined with the modular nature of the DNA binding domain, is permitting researchers to harness TALE-derived repeats as custom-made scaffolds on which to assemble functional heterologous domains. Hitherto, designer TALEs tested *in vitro* and *in cellula* have included those with transcriptional activation or FokI nuclease domains (8,9). The latter, called TALE nucleases (TALENs) operate similarly to ZFNs in that a pair is assembled at a given DNA sequence consisting of two half-target sites separated by a spacer sequence (10). The directional binding of each TALEN monomer to its respective half-target site induces dimerization of the FokI portions resulting in site-specific DNA cleavage. Importantly, in contrast to zinc-finger modules (11), there are as of yet no

indications that binding of an individual TALE repeat to its cognate base pair is altered by neighboring sequences. This context-independent feature combined with the simple DNA binding code suggests that, when compared to ZFNs, TALENs can bind a wider range of DNA sequences and be programmed in an easier and more predictable manner. Moreover, the ability to construct functionally viable TALEN pairs with relatively large arrays of repeats has the potential to render them more specific than ZFNs.

Thus far, nucleic acids encoding TALENs have been introduced into cells by nucleic acid microinjection or by transfection methods based on chemical agents or electroporation (8,9,12-14). Unfortunately, these methods suffer from low throughput or are too cytotoxic or too inefficient in populations of non-transformed cells. Notably, the therapeutic application and the thorough assessment of the genome-wide TALEN specificity, for instance, will be dependent on highly efficient TALEN delivery systems into relevant karyotypically stable cells. In the realm of biomedical research this will include tissue-specific stem and progenitor cells. From the broad variety of methods to introduce foreign nucleic acids into cells, viral vectors have proven to be the most effective thus far. In particular, vectors derived from the double-stranded linear DNA human adenovirus serotype 5 and from the single-stranded positive-sense RNA lentivirus human immunodeficiency type 1 (HIV-1) have become broadly used (15,16). Due to the infection mechanisms evolved by their parental viruses (17,18), adenoviral and lentiviral vectors are particularly fit for the *in vitro* and *in vivo* transduction of dividing and non-dividing cells. In the current study, we investigated the suitability of these two major vector classes as delivery vehicles of *TALEN* sequences. We report that, in contrast to lentiviral vectors, adenoviral vectors can package and deliver functional TALEN-encoding expression units into transformed, immortalized and primary human cells.

## MATERIALS AND METHODS

### Cells

The 293T and HeLa cells (American Type Culture Collection) were cultured in Dulbecco's modified Eagle's medium (DMEM; Invitrogen). The medium used for the culturing of the former and latter cell types was supplemented with 10% and 5% fetal bovine serum (FBS; Invitrogen), respectively. The adenoviral vector packaging cell lines PER.C6 (19) and PER. E2A (20) were maintained in DMEM containing 10% FBS, 10 mM MgCl2 (Sigma-Aldrich) without and with 250 µg/ml Geneticin (Invitrogen), respectively. These cells were kept in a humidified-air 10% CO2 atmosphere. The human myoblasts were obtained from a Duchenne muscular dystrophy patient and were immortalized through constitutive expression of *Bmi-1* and *TERT*. The generation of and culture conditions for these immortalized human myoblasts have been specified elsewhere (21). The human bone marrow-derived primary

fetal mesenchymal stem cells (hMSCs) were cultured as previously specified (22). The human myoblasts and hMSCs were kept in a humidified-air 5% $CO_2$ atmosphere.

**Recombinant DNA**

A TALEN pair (Supplementary Figure S1) targeting intron 1 of the *PPP1R12C* locus (*AAVS1* target site: 5'-TCTGTCCCCTCCACCCCAC-spacer-GACAGGATTGGTGACAGAA-3') were generated by Golden Gate cloning, as previously described (9). To generate pAdShu.TALEN-L[S1].IRES.RFP and pAdShu.TALEN-R[S1].IRES.RFP, plasmids 1383.pVAX.AAVS1.TALEN.L-94 and 1384.pVAX.AAVS1.TALEN.R-95 were digested with *Eco*32I (Fermentas) and *Pme*I (New England Biolabs) and the resulting 3.0-kb *TALEN* open reading frame (ORF) fragment were isolated and ligated into the *Eco*32I-linearized 8.8-kb plasmid vector backbone pShuttle. IRES.DsRedEx2.1. Cells transduced with vectors based on this construct can be traced and quantified via the encoded red fluorescent protein reporter DsRedEx2.1 (RFP). The firstgeneration (i.e. early region 1 [*E1*]-deleted) fiber-modified adenoviral vector molecular clones pAd.ΔE1.TALEN-L[S1].F[50] and pAd.ΔE1.TALEN-L[S1].F[50] (Figure 2A) were assembled by HR following transformation of E. coli cells BJ5183[pAdEasy-1.50] (23) with shuttle plasmids pAdShu. TALEN-L[S1].IRES.RFP and pAdShu.TALEN-R[S1].IRES.RFP, respectively, whilst the second-generation adenoviral vector molecular clones pAd.ΔE1ΔE2A.TALEN-L[S1].F[50] and pAd. ΔE1ΔE2A.TALEN-R[S1].F[50] (Figure 2A) were assembled by HR following transformation of *E. coli* cells BJ5183[pAdEasy-2.50] (23) with shuttle plasmids pAdShu.TALEN-L[S1].IRES.RFP and pAdShu. TALEN-R[S1].IRES.RFP, respectively. Prior to transformation all shuttle plasmids were digested to completion with *Mss*I. The structure and composition of lentiviral vector constructs pLV. TALEN-L[S1].i.RFP, pLV.TALEN-R[S1].i.RFP, pLV.TALEN-L[S1] and pLV.TALEN-R[S1] are described in Figure 1A. The plasmid CMVPRES (24; herein referred to as pLV.CMV.eGFP) served as a negative control in the PCR amplifications of *TALEN* sequences.

**Viral Vectors**

The production of adenoviral vectors Ad.ΔE1.TALEN-L[S1].F[50] and Ad.ΔE1.TALEN-R[S1].F[50] took place in PER.C6 cells, whereas that of adenoviral vectors Ad.ΔE1ΔE2A.TALEN-L[S1].F[50] and Ad.ΔE1ΔE2A.TALEN-R[S1].F[50] occurred in PER.E2A cells. The rescue of vector particles was initiated by transfecting each of the *Pac*I-linearized adenoviral vector molecular clones into their respective packaging cells by using a polyethyleneimine-based protocol (23). The procedures to propagate and purify the resulting viral vectors have been described elsewhere (23). Vector titers were determined through limiting dilutions on HeLa indicator cells seeded at a density of $0.8×10^5$ cells per well of 24-well plates (Greiner). At three days post-transduction, frequencies of RFP-positive cells were determined through flow cytometry by using a BD LSR II FACS (Becton Dickinson). As a result, the titers are expressed in terms of HeLa cell-transducing units (HTUs) per ml. The HIV-1-based and vesicular stomatitis virus

glycoprotein-G (VSV-G)-pseudotyped lentiviral vectors were generated by transfecting 293T cells essentially as specified elsewhere (25) except for the use of $7.5\times10^6$ cells in T75-cm$^2$ flasks instead of $1.75\times10^7$ cells in T175-cm$^2$ flasks. A total amount of 12.85 µg of DNA was transfected corresponding to 2:1:1 mixtures of lentiviral vector transfer plasmid, packaging construct psPAX2 (Addgene plasmid 12260) or psPAX2$^{D116N}$ (25) and VSV-G pseudotyping construct pLP/VSVG (Invitrogen). Lentiviral vector titers were determined by using HIV-1 p24$^{gag}$ immunocapture (Perkin Elmer).

**Immunofluorescence microscopy**

Fifty-thousand HeLa cells were seeded in wells of 24-well plates and, after an overnight incubation period, they were either mock-transduced or were transduced with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or with Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at 3 HTU/cell. Three days post-transduction, the cells were subjected to immunostaining essentially as described elsewhere (26) except for the use of an influenza hemagglutinin (HA)-specific mouse monoclonal IgG1 antibody (HA.11 clone 16B12; Covance) together with a goat anti-mouse IgG (H+L) Alexa 488-conjugated secondary antibody (A-11001; Molecular Probes).

**Western blot analysis**

HeLa cells seeded at a density of $1\times10^5$ cells per well of 24-well plates (Greiner) were transduced with different amounts of Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$. At 72 hours post-transduction, the cell monolayers were rinsed with ice-cold phosphate-buffered saline (pH 7.4) and lysed in 100 µl of ice-cold RIPA buffer composed of 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% sodium dodecyl sulphate (SDS), 0.5% sodium deoxycholate and 1% Nonidet P-40 plus a cocktail of protease inhibitors (Complete Mini, Roche Applied Science). Next, the cell lysates were spun at 4°C for 8 min at 16,400×g. The proteins were subjected to electrophoresis through an SDS-7.5% polyacrylamide gel after which they were transferred onto an Immobilon-P membrane (Millipore) by electroblotting overnight. After incubation with TBST blocking solution consisting of 10 mM Tris-HCl (pH 8.0), 150 mM NaCl and 0.2% Tween-20 supplemented with 10% non-fat dry milk powder (Elk, Campina), the membranes were incubated overnight at 4°C with the HA-specific monoclonal antibody F-7 (SC-7392, Santa Cruz Biotechnology) diluted 1:1000. Subsequently, the membranes were washed once with TBST at room temperature and subjected to a 3-hour incubation period with a horseradish peroxidase conjugated goat anti-mouse secondary antibody (SC-2005; Santa Cruz Biotechnology) diluted 1:5000 in blocking solution. Finally, the membranes were rinsed and processed for protein detection by chemiluminescence as described elsewhere (27).

The Western blot analysis to detect *TALEN* expression following lentiviral vector transduction was performed as follows. Hundred-thousand HeLa cells were seeded in wells of 6- well plates and were transduced with 200 ng of p24$^{gag}$ of LV.TALEN-L$^{S1}$ or 400 ng of p24$^{gag}$

of IDLV.TALEN-L[S1] in the presence of polybrene (8 µg/ml). Forty-eight hours after transduction, the cells were harvested and total protein lysates were prepared as aforementioned. Fifty micrograms of protein were subjected to electrophoresis through a NuPAGE 4-12% Bis-Tris Gel (Invitrogen) and were transferred onto a nitrocellulose membrane (Hybond-ECL, Amersham) that was subsequently exposed overnight at 4°C to a mouse monoclonal antibody against the HA-tag (clone 12CA5 a gift from Valeria Marigo). A horseradish peroxidase-conjugated goat anti-mouse secondary antibody (Amersham) was diluted 1:2000 in blocking solution. The proteins were detected by using a chemiluminescence labelling detection reagent (ECL; GE Healthcare Life Sciences). For protein loading normalization, the same membrane was stripped and incubated with a 2000-fold diluted mouse monoclonal antibody directed against β-actin (sc-81178, Santa Cruz Biotechnology).

**Isolation and structural analysis of adenoviral vector DNA**

DNA was isolated from purified adenoviral vector particles as follows. To a 50-µl vector preparation aliquot it was added 12 µl of DNaseI buffer (130 mM Tris-HCl [pH 7.5], 1.2 mM CaCl$_2$ and 50 mM MgCl$_2$), 8 µl of DNaseI (10 mg/ml; Fermentas) and 50 µl of Milli-Q water (Millipore). This mixture was incubated for 30 min at 37°C, after which, the DNaseI was inactivated for 1 hour at 55°C after the addition of 2.4 µl of 0.5M EDTA (pH 8.0), 6 µl of 10% SDS and 1.5 µl of a Proteinase K (21.4 mg/ml; Fermentas). Next, the vector DNA was purified by using the QIAEX II Gel Extraction Kit (Qiagen) and was eluted in 30 µl of 10 mM Tris-HCl (pH 7.8). After measuring its concentration with the aid of a Nanodrop (Thermo Scientific), the recovered DNA was subjected to *Bcu*I (Fermentas) digestion and agarose gel electrophoreses (200 ng) and to PCR analyses (1 ng). The PCR amplifications were carried out with primers TALEN-Seq-F (5'-CCAGCTGCTGAAGATCGCCAAGC-3') and TALEN-Seq-R (5'-CTATCCTGAGTGGAATTTCTGGC-3') whose target sequences flank the TALEN DNAbinding motifs. The DNA samples were subjected to PCR with 0.4 µM of these primers (Eurofins MWG Operon), 0.1 mM of each dNTP (New England Biolabs), 1× Colorless GoTaq Flexi buffer, 1 mM MgCl$_2$ and 2.5 units of GoTaq DNA polymerase (all from Promega). The PCR cycles were performed in a DNA Engine Tetrad 2 Peltier Thermal Cycler (Bio-Rad) using the following cycling parameters. After a 3-min denaturation step at 95°C, the samples were subjected to 30 cycles consisting of 30 s at 95°C, 30 s at 60°C and 4 min at 72°C. The reactions were terminated by a final elongation period of 3 min at 72°C. One-fifth of these reactions were subjected to electrophoresis through 1% agarose gels in 1× TAE buffer. Moreover, to establish the integrity of *TALEN* sequences in adenoviral vector particles at the nucleotide level, the 2.4-kb *Bcu*I fragments from Ad.ΔE1.TALEN-L[S1].F[50] or from Ad.ΔE1.TALEN-R[S1].F[50] DNA were inserted into a *Xba*I-linearized pUC19 backbone and were sequenced by using M13-F (5'-GACGTTGTAAAACGACGGCCAGT-3') and M13-R (5'-CAGGAAACAGCTATGACCATGA-3') primers.

**Surveyor nuclease S mutation detection assay**

Genomic DNA was extracted from mock-transduced or adenoviral vector-transduced target cells as described before (28). Briefly, at 72 hours post-infection, the cells were collected by trypsinization and incubated overnight at 55°C in 500 µl of lysis buffer (100 mM Tris-HCl [pH 8.5], 5 mM EDTA, 0.2% SDS, 200 mM NaCl) supplemented with freshly-added Proteinase K at a final concentration of 100 ng/ml. Next, the cell lysates were extracted twice with a buffersaturated phenol:chloroform:isoamyl alcohol mixture (25:24:1) and once with chloroform. The DNA in the aqueous phase 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 pellet was mildly dried and dissolved in 100 µl of TE buffer (10 mM Tris-HCl [pH 8.0] and 1mM EDTA) supplemented with RNase A (Fermentas) at a final concentration of 100 µg/ml.

The recovered DNA was used for PCR analyses. The extent of *AAVS1* processing by the error-prone NHEJ DNA repair pathway was evaluated by using the Surveyor assay (SURVEYOR Mutation Detection Kit, Transgenomics) according to the manufacturer's instructions. In brief, genomic DNA purified from mock-transduced or adenoviral vectortransduced cells was used as a template to PCR amplify a 469-bp *AAVS1* sequence flanking the TALEN pair target site. PCR mixtures contained 0.4 µM of primers AAVS1-CEL1-F (5'-TTCGGGTCACCTCTCACTCC-3') and AAVS1-CEL1-R (5'-GGCTCCATCGTAAGCAAACC-3'), 0.1 mM of each dNTP, 1× Colorless GoTaq Flexi buffer, 1mM $MgCl_2$ and 2.5 units of GoTaq DNA polymerase. The PCR cycles were performed using the following cycling conditions. An initial denaturation step at 95°C for 5 min was followed by 40 cycles of 30 s at 95°C, 30 s at 61°C and 30 s at 72°C. The reactions were terminated by a final elongation period of 5 min at 72°C. The PCR products, after being heat-denaturated and re-annealed according to the manufacturer's specifications, were treated with the nucleotide mismatch-sensitive Surveyor nuclease S (SNS) enzyme in order to detect small insertions and deletions (indels) caused by NHEJ-dependent DNA repair. Samples were loaded on a 1.7% agarose gel and electrophoresis was performed in 1× TAE buffer. The DNA was visualized by in-gel staining with 2 µg/ml of ethidium bromide (Sigma-Aldrich) in 1× TAE buffer followed by stain removal in Milli-Q water. The frequencies of target DNA cleavage were determined by using the ImageJ software (National Institutes of Health, USA). Genomic DNA derived from hMSC cells co-transduced with a 1:1 mixture of Ad.ΔE1.TALEN-L[S1].F[50] and Ad.ΔE1.TALEN-R[S1].F[50] at a total multiplicity of infection (MOI) of 24 HTU/cell, served as template to PCR amplify the *AAVS1* target site region. The pGEM-T Easy TA cloning vector (Promega) was ligated to the resulting 469-bp amplicons and was transformed into *E. coli* strain DH5α. Plasmid DNA from 20 randomly selected colonies was extracted and the respective inserts were subjected to DNA sequencing by using the M13-F primer.

**PCR analyses of TALEN sequences following lentiviral vector-mediated gene transfer**

Total cellular DNA from HeLa cells incubated with integration-competent or integrase-defective lentiviral vectors (IDLVs) was isolated as specified elsewhere (28). The recovered DNA was subjected to three different PCR protocols. To determine the integrity of the cellular DNA, a PCR targeting the human housekeeping gene hypoxanthine phosphoribosyltransferase 1 (*HPRT1*) was performed. To this end, reactions containing 0.4 µM of primers hHPRT.1 and hHPRT.2 (29), 0.1 mM of each dNTP, 1× Colorless GoTaq Flexi buffer, 1mM $MgCl_2$ and 2.5 U of GoTaq DNA polymerase, were assembled. The PCR cycles were performed using the following cycling conditions. An initial denaturating step at 95°C for 5 min was followed by 30 cycles of 30 s at 95°C, 30 s at 60°C and 120 s at 72°C.

A second set of PCR mixtures were assembled to amplify a 458-bp *TALEN* sequence corresponding to the N-terminal, non-repetitive region, of each nuclease. These consisted of 0.4 µM of primers TAL-N.s (5'-CCTTACGACGTGCCTGACTACG-3') and TAL-N.as (5'-CGCTTGGCGATCTTCAGC-3'), 0.1 mM of each dNTP, 1× Colorless GoTaq Flexi buffer, 1mM $MgCl_2$ and 2.5 U of GoTaq DNA polymerase. The PCR cycles were performed using the following cycling conditions. An initial denaturating step at 95°C for 5 min was followed by 30 cycles of 30 s at 95°C, 30 s at 62°C and 30 s at 72°C.

Finally, the third PCR set was designed to amplify a 2.5-kb *TALEN* sequence encompassing the region encoding the repetitive DNA-binding domain of each nuclease. To this end, reactions containing 0.4 µM of primers CMV-prom.F (5'-AATGGGCGGTAGGCGTGTA-3') and TALE.as.Pr2 (5'-GTCGGGTCTGCTCAGCTG-3'), 0.1 mM of each dNTP, 1× Colorless GoTaq Flexi buffer, 1mM $MgCl_2$ and 2.5 U of GoTaq DNA polymerase, were assembled. The PCR cycling parameters consisted of an initial denaturating step at 95°C for 5 min followed by 40 cycles of 30 s at 95°C, 30 s at 60°C and 180 s at 72°C. In all three protocols, the PCR amplifications were terminated by a final elongation period of 5 min at 72°C. Control samples subjected to the same PCR conditions were carried out in parallel. These controls consisted of using, in place of DNA from lentiviral vector-transduced HeLa cells, DNA from mocktransduced HeLa cells, DNA from Ad.ΔE1.TALEN-R[S1].F[50] -transduced HeLa cells, pAd.ΔE1.TALEN-L[S1].F[50], pLV.CMV.eGFP or MilliQ water. Additional controls were provided by PCR amplifications on DNA isolated from HeLa cells exposed to an LV.TALEN-R[S1] preparation made with the aid of pUC19 instead of the lentiviral vector packaging construct psPAX2 (Addgene plasmid 12260) or on DNA extracted from cells transfected with pLV.TALEN-R[S1] transfer plasmid mixed or not mixed with ExGen500 (Fermentas).

**Southern blot analysis**

Total cellular DNA from 4×10[5] hMSCs exposed to three different MOIs of Ad.ΔE1.TALEN-L[S1].F[50] or of Ad.ΔE1.TALEN-R[S1].F[50] was extracted 72 hours post-transduction and purified as described before (28). After overnight digestion with *Bam*HI (Fermentas), 10 µg of

DNA per sample was resolved in a 1% agarose gel in 1× TAE buffer. Next, the DNA was transferred by capillary action onto an Amersham Hybond-XL membrane (GE Healthcare Life Sciences) using a standard Southern blot technique. A 397-bp probe specific for *TALEN* DNA corresponding to the N-terminal region of the nucleases was isolated by digestion of 1383. pVAX.AAVS1.TALEN.L-94 with *Xba*I and *Pst*I (both from Fermentas) and was purified with the aid of the QIAEX II gel extraction kit (Qiagen). This DNA probe was radiolabelled with EasyTide (α-$^{32}$P) dCTP (3000 Ci/mmol; Perkin Elmer) by using the DecaLabel DNA labelling system (Fermentas). Prior to its addition onto the membrane, the radiolabeled fragment was separated from unincorporated dNTPs through size-exclusion chromatography with the aid of Sephadex-50 columns (GE Healthcare Life Sciences). A Storm 820 PhosphorImager (Amersham) as used for the detection of the probe-hybridizing DNA. Finally, images were acquired and processed by using the Storm scanner control 5.03 and the ImageQuant Tools 3.0 software, respectively (both from Amersham).

HeLa cell clones were derived from an LV.TALEN-L[S1]-transduced population by limiting dilution using a seeding density of 0.3 cells per well in 96-well plates. Southern blot analysis of genomic DNA extracted from various HeLa cell clones was performed as follows. Genomic DNA was extracted from 5×10$^6$ cells from each clone using a QIAamp DNA Mini Kit according to the manufacturer's instructions (Qiagen). Ten micrograms of genomic DNA was digested overnight with *Afl*II (New England Biolabs), subjected to electrophoresis through a 0.8% agarose gel and was transferred onto a Hybond-XL nylon membrane. Prior to DNA detection, this membrane was incubated with a (α-$^{32}$P)-labelled 814-bp probe corresponding to the lentiviral vector LTR elements R and U5, the packaging signal and part of the Rev-responsive element (RRE).

## RESULTS

### Lentiviral vectors introduce rearranged *TALEN* genes into human cells

We started by investigating the feasibility of generating lentiviral vectors to deliver TALENencoding expression units into human cells. Although the long-term presence of artificial nucleases in cells is, clearly, undesirable in most experimental settings, lentiviral vectors can be made with in-built stringent drug-controllable transgene expression systems (30) or class I integrase mutant moieties (31,32). These integrase mutations are non-pleiotropic as they confer an integration-defective phenotype to the vectors by specifically disabling the chromosomal insertion step of their DNA. IDLVs are, therefore, being increasingly exploited to deliver transgenes whose expression is only required during a short timeframe in populations of cycling cells. Examples include IDLV-mediated delivery of transposases, homing endonucleases and zinc finger nucleases (33-36).

Therefore, we constructed the transfer plasmids pLV.TALEN-L$^{S1}$ and pLV.TALEN-R$^{S1}$ to package the lentiviral vector pairs LV.TALEN-L$^{S1}$/IDLV.TALEN-L$^{S1}$ and LV.TALEN-R$^{S1}$/IDLV.TALEN-R$^{S1}$, respectively. These constructs encode a TALEN enzyme (i.e., TALEN-L$^{S1}$ or TALEN-R$^{S1}$) specific for the *AAVS1* locus at the human chromosome 19 (19q13.42-qter). Moreover, we also constructed pLV.TALEN-L$^{S1}$.i.RFP and pLV.TALEN-R$^{S1}$.i.RFP to make lentiviral vector particles carrying bicistronic cassettes encoding TALEN-L$^{S1}$ and TALEN-R$^{S1}$, respectively, and through an IRES, the RFP reporter. To investigate the integrity of vector genomes in target cells we carried out transduction experiments on HeLa cells deploying the lentiviral vector set harboring the *TALEN-R$^{S1}$* ORF. Following DNA extraction from vectortransduced cells, we set-up a PCR assay (Figure 1A) based on the use of primer pairs framing the repeats of the *TALEN* ORF and its upstream region encoding the nuclease's N-terminus (CMVprom. F/TALE.as.Pr2) or targeting exclusively the latter segment (TAL-N.s/TAL-N.as). PCR amplifications with primers CMV-prom.F and TALE.as.Pr2 on DNA extracted from HeLa cells transduced with LV.TALEN-R$^{S1}$, IDLV.TALEN-R$^{S1}$ or LV.TALEN-R$^{S1}$.i.RFP did not yield detectable levels of 2.5-kb amplicons diagnostic for full-length *TALEN-R$^{S1}$* sequences (Figure 1B, upper panels, lanes 5, 6 and 11, respectively). In fact, in these samples, the only PCR products that could be discerned were those consistent with heterogeneously sized sub-genomic templates whose molecular-weight range spanned from about 0.8 to 1.5 kb. Significantly, PCR amplifications carried out on the same DNA templates with primers TAL-N.s and TAL-N.as, gave rise to properly-sized 0.46 kb PCR products indicating that the non-repetitive upstream portion of the *TALEN-R$^{S1}$* ORF did not suffer major rearrangements in transduced cells (Figure 4B, middle panels). Of note, equivalent results were obtained in complementary transduction experiments that made use of the lentiviral vector set carrying the *TALEN-L$^{S1}$* ORF (not shown). To validate these data, PCR amplifications were done in parallel on control samples. These negative and positive controls consisted of water (Figure 1B, lane 4) or DNA extracted from mock-transduced HeLa cells (Figure 1B, lane 1), Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$-transduced HeLa cells (Figure 1B, lane 10), plasmid pAd.ΔE1.TALEN-L$^{S1}$.F$^{50}$ (Figure 1B, lane 2) or plasmid pLV.CMV.eGFP (Figure 1B, lane 3). Additional controls consisted of PCR amplifications on DNA isolated from HeLa cells exposed to an LV.TALEN-R$^{S1}$ preparation made with pUC19 instead of the lentiviral vector packaging construct psPAX2 (Figure 1B, lane 7) or incubated with plasmid pLV.TALEN-R$^{S1}$ mixed or not mixed with the DNA transfection agent ExGen500 (Figure 1B, lanes 8 and 9, respectively). PCR amplifications targeting *HPRT1* served as a control for DNA template integrity (Figure 1B, lower panels).

To complement these PCR results on whole target cell populations, we carried out Southern blot analysis on genomic DNA from individual HeLa cell clones derived from a bulk population transduced with LV.TALEN-L$^{S1}$. To this end, the chromosomal DNA from 15 clones was digested with AflII, Southern blotted and incubated with the probe depicted in Figure 1A. To serve as a molecular-weight reference, the parental transfer plasmid pLV.TALEN-L$^{S1}$

was subjected to the same procedures. Results depicted in Figure 1C demonstrate that none of the tested clones yielded the *Afl*II 6.4-kb fragment diagnostic for full-length LV.TALEN-L$^{S1}$ proviruses.

Finally, we investigated through HA tag-directed Western blot analysis, *TALEN-L$^{S1}$* expression following addition of 200 ng of p24$^{gag}$ of LV.TALEN-L$^{S1}$ or 400 ng of IDLV.TALEN-L$^{S1}$ onto HeLa cultures. These experiments failed to show detectable levels of HA-specific signals in protein lysates corresponding to vector-transduced cells, whilst HA-specific signals at a molecular-weight consistent with full-length TALEN-L$^{S1}$ proteins could be readily detected in control lysates derived from pLV.TALEN-L$^{S1}$-transfected HeLa cells or 293T producer cells (Figure 1D). To serve as controls for assembly of LV.TALEN-L$^{S1}$ or IDLV.TALEN-L$^{S1}$ particles, we produced in parallel LV and IDLV vector preparations carrying a human *PGK1* promoter-driven eGFP reporter (LV.PGK.eGFP and IDLV.PGK.eGFP, respectively). Exposure of HeLa cell cultures to 200 ng of p24$^{gag}$ of LV.PGK.eGFP or to 400 ng of IDLV.PGK.eGFP led to the transduction of 97% and 83% of the target cells, respectively (not shown).

Taken together, these data indicate that *TALEN* sequences introduced into human cells through lentiviral vectors are prone to extensive rearrangements resulting presumably from deletions involving the *TALE* repeats.
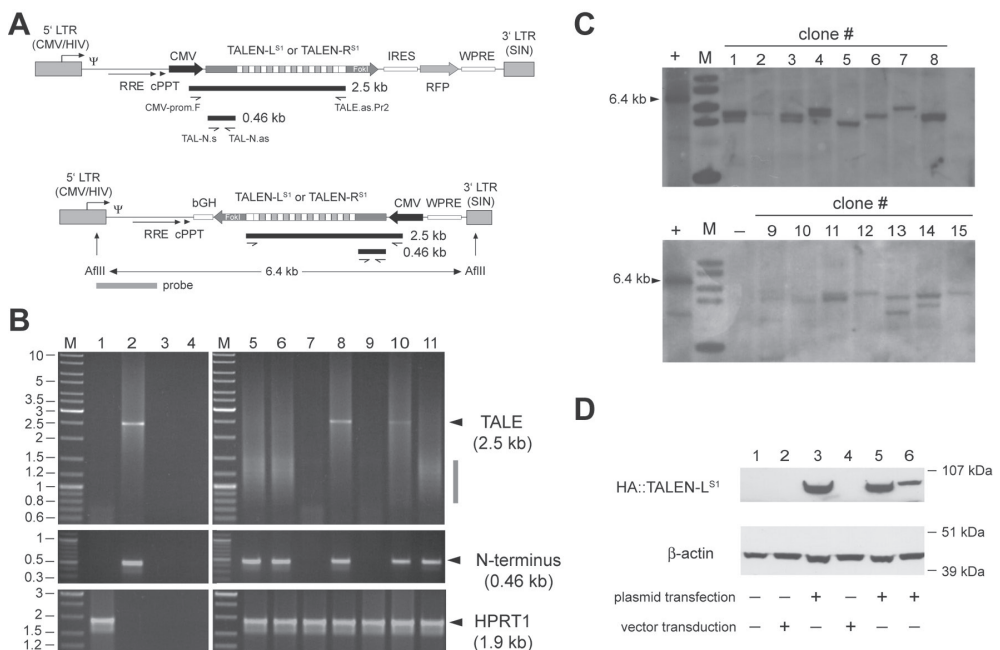
**Figure 1**. Structural analyses of lentiviral vectors harboring *TALEN* sequences in transduced cells. **(A)** Genetic organization of lentiviral vector constructs. Diagram of bicistronic pLV.TALEN-L$^{S1}$.i.RFP and pLV.TALEN-R$^{S1}$.i.RFP and of monocistronic pLV.TALEN-L$^{S1}$ and pLV.TALEN-R$^{S1}$ lentiviral vector transfer plasmids. The bicistronic constructs code for the *AAVS1*-specific custom-made nucleases TALEN-L$^{S1}$ or TALEN-R$^{S1}$ and, through an IRES sequence, a RFP reporter, whereas the monocistronic plasmids encode exclusively the *AAVS1*-specific designer nucleases. Each of the TALEN proteins are tagged by an HA antigen located close to their N-termini (not drawn). Gray boxes with broken arrow, hybrid 5' long-terminal repeat (LTR) containing CMV and HIV-1 sequences; gray boxes without broken arrow, selfinactivating (SIN) 3' LTR; Ψ, HIV-1 packaging signal; RRE, Rev-responsive element; cPPT, central polypurine tract; WPRE, Woodchuck hepatitis virus posttranscriptional regulatory element. In pLV.TALEN-L$^{S1}$.i.RFP and pLV.TALEN-R$^{S1}$.i.RFP the ORFs are under the transcriptional control of the CMV promoter and the polyadenylation signal present within the SIN 3' LTR, whilst the ORFs in pLV.TALEN-L$^{S1}$ and pLV.TALEN-R$^{S1}$ are under the control of the CMV promoter and the bovine *growth hormone* polyadenylation signal (bGH). Primers, PCR products and Southern blot probe used are depicted as half-arrows, black bars and gray bar, respectively, and are drawn in relation to their respective vector DNA sequences. For the sake of simplicity, plasmid backbone sequences are not show. The sizes of the PCR products expected from the amplification of full-length lentiviral vector genomes are indicated. The 6.4-kb *Afl*II restriction fragment corresponding to intact DNA from LV.TALEN-L$^{S1}$ or LV.TALEN-R$^{S1}$ is also indicated. **(B)** Expression analysis. PCR analyses of TALEN-encoding lentiviral vector genomes following transduction of HeLa cells. PCR amplifications with the aid of the primers shown in Figure 1A were carried out using, as template, total cellular DNA from mocktransduced HeLa cells, plasmid pAd.ΔE1.TALEN-L$^{S1}$.F$^{50}$, plasmid pLV.CMV.eGFP or nuclease-free H$_2$0 (lanes 1, 2, 3 and 4, respectively). PCR amplifications using the same primer pairs were performed in parallel on DNA extracted from HeLa cells transduced with integrationcompetent LV.TALEN-R$^{S1}$ (lane 5), integrase-defective IDLV.TALEN-R$^{S1}$ (lane 6), Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ (lane 10) or integration-competent LV.TALEN-R$^{S1}$.i.RFP (lane 11). Extra PCR controls supplementing those corresponding to lanes 1 through 4, were provided by using DNA isolated from HeLa cells exposed to an LV.TALEN-R$^{S1}$ preparation made in parallel but deploying pUC19 instead of lentiviral vector packaging construct psPAX2 (lane 7) or incubated with pLV.TALEN-R$^{S1}$ mixed or not mixed with the DNA transfection agent ExGen500 (lanes 8 and 9, respectively). Lane M, Gene Ruler DNA Ladder Mix. A primer pair targeting the human *HPRT1* served to control the integrity of the cellular DNA. **(C)** Clonal analysis. Southern blot analysis of *Afl*II-digested genomic DNA from HeLa cell clones 1 through 15 stably transduced with LV.TALEN-L$^{S1}$. The vector-specific probe that was used is shown in Figure 1A (horizontal gray bar). Lane M, 1 Kb molecular-weight marker. The parental transfer plasmid pLV.TALEN-L$^{S1}$ treated with *Afl*II served as an internal control for intact vector DNA. **(D)** TALEN protein detection. HA tag-directed Western blot analysis of full-length *TALEN-LS1* expression in 293T producer cells (lane 1), IDLV.TALEN-L$^{S1}$- and LV.TALEN-L$^{S1}$-transduced HeLa cells (lane 2 and lane 4, respectively) and in pLV.TALEN-L$^{S1}$-transfected 293T cells used to generate these IDLV.TALEN-L$^{S1}$ and LV.TALEN-L$^{S1}$ vector preparations (lanes 3 and 5, respectively). Lane 6, HeLa cells transfected with pLV.TALEN-L$^{S1}$. The β-actin served as loading control.
◄


## Adenoviral vectors deliver full-length TALEN genes into human cells inducing high-level DSB formation at an endogenous chromosomal target locus

In light of the previous results, we sought to investigate the feasibility of using adenoviral vectors as an alternative delivery platform for the introduction of intact *TALEN* genes into human cells. To this end, the full-length adenoviral vector molecular clones pAd.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and pAd.ΔE1.TALEN-R$^{S1}$.F$^{50}$ were assembled through HR in *E. coli* to produce the *E1*-deleted adenoviral vectors Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$, respectively (Figure 2A). These clones contain, in place of *E1* from the prototypic adenovirus serotype
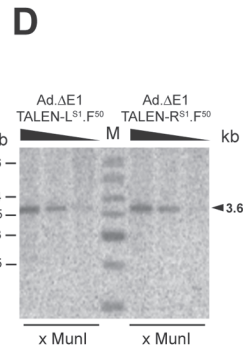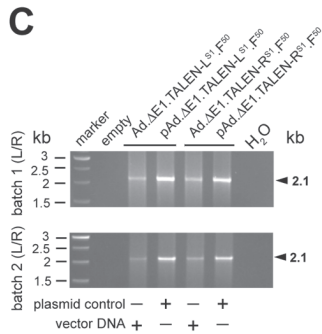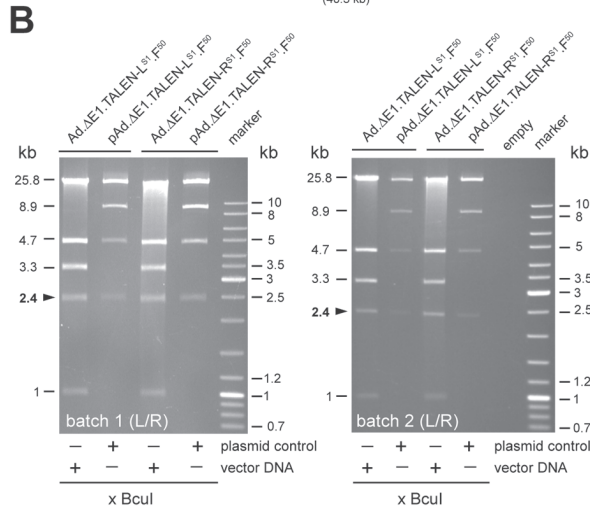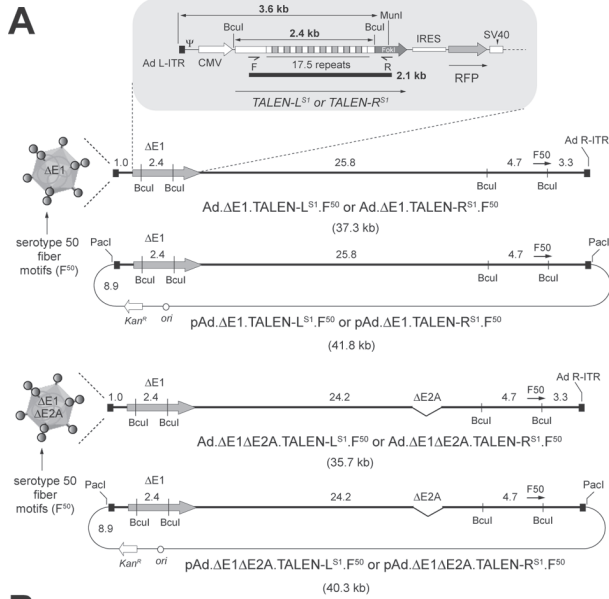
5, a bicistronic expression unit encoding a monomer of the *AAVS1*-specific TALEN pair and, for tracing and quantification purposes, the RFP reporter (Figure 2A). In addition, their adenoviral genomic sequences harbor an ORF coding for a vector retargeting chimeric fiber composed of serotype 5 basal shaft domains fused to apical shaft and knob domains derived from the species B adenovirus serotype 50 (F[50]). The latter CD46-interacting motif allows for entry of vectors displaying species B fiber domains into human cells lacking on their surface the Coxsackie B virus and adenovirus receptor, CAR. Of note, among these cells there are various therapeutically relevant tissue-specific normal and malignant stem and progenitor cell types (37-41). In addition, plasmids pAd.ΔE1ΔE2A.TALEN-L[S1].F[50] and pAd.ΔE1ΔE2A.TALEN-R[S1].F[50] were also constructed to assemble fiber-modified *E1*- plus *E2A*-deleted particles Ad.ΔE1ΔE2A.TALEN-L[S1]. F[50] and Ad.ΔE1ΔE2A.TALEN-R[S1].F[50], respectively (Figure 2A). By virtue of their lack of E2A, these so-called second-generation adenoviral vectors have been shown to display a muchreduced *de novo* replication and "leaky" viral gene expression profile in target cells when compared to their *E1*-deleted, first-generation, counterparts (see, for instance, ref. 20).
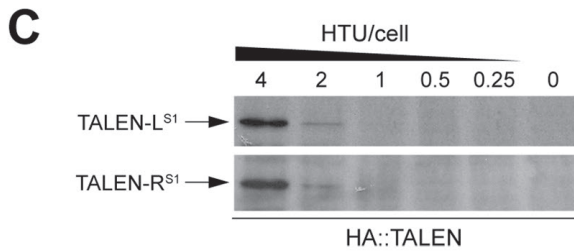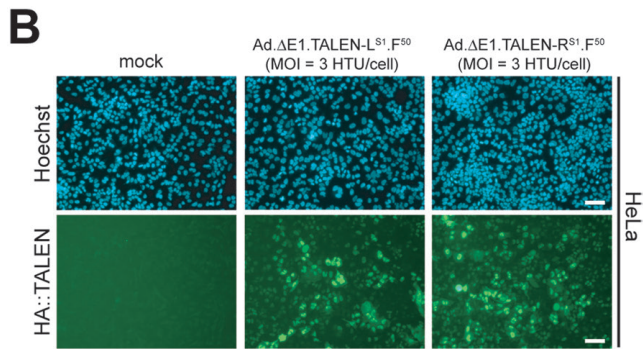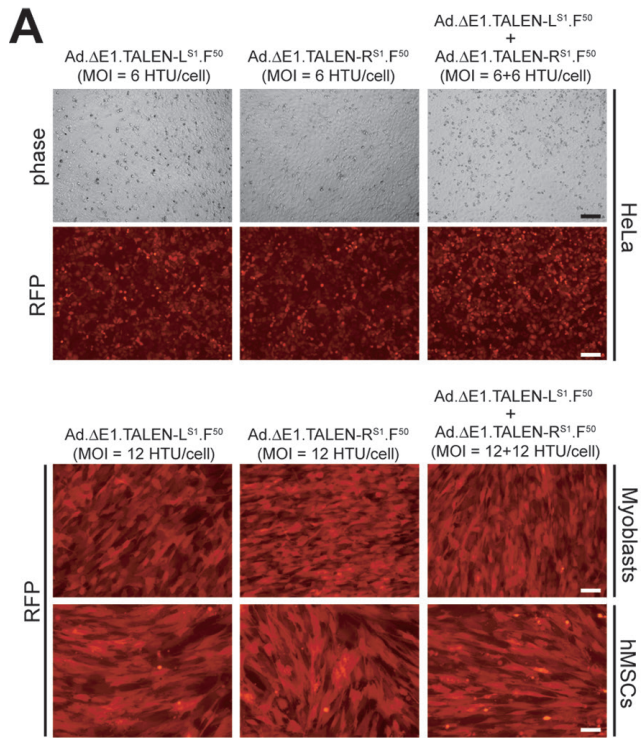
We started by testing whether *TALEN* sequences are stably maintained in first- and second-generation adenoviral vectors following their rescue and subsequent propagation on their respective packaging cell lines. Thus, these producer cells were transfected with *Pac*I-treated plasmids pAd.ΔE1.TALEN-L[S1].F[50], pAd.ΔE1.TALEN-R[S1].F[50], pAd.ΔE1ΔE2A.TALEN-L[S1].F[50] or pAd.ΔE1ΔE2A.TALEN-R[S1].F[50] to assemble the corresponding adenoviral vector particles (Figure 2A). The rescued particles were subsequently amplified following serial propagation on increasing amounts of packaging cells. The frequency of RFP-positive producer cells, as monitored by direct fluorescence microscopy, revealed a clear propagation round-dependent increase in vector yields (not shown). The Ad.ΔE1.TALEN-L[S1].F[50] and Ad.ΔE1.TALEN-R[S1].F[50] vectors, corresponding to two independent batches, were purified through CsCl buoyant density gradient ultracentrifugation, after which, the integrity of their genomes was probed by BcuI restriction fragment length analysis. To serve as an additional molecular-weight reference, the parental plasmids pAd.ΔE1.TALEN-L[S1].F[50] and pAd.ΔE1.TALEN-R[S1].F[50] were taken along in this analysis. Results depicted in Figure 2B show that the overall DNA restriction pattern is consistent with that expected from linear full-length adenoviral vector genomes (Figure 2A). Thus, importantly, the internal *TALE*-encompassing restriction fragments derived from Ad.ΔE1.TALEN-L[S1].F[50] and Ad.ΔE1.TALEN-R[S1].F[50] DNA did co-migrate with those corresponding to their respective 2.4-kb parental plasmid counterparts (Figure 2B, arrowheads). Results equivalent to these were obtained following the structural analyses of DNA isolated from Ad.ΔE1ΔE2A.TALEN-L[S1].F[50] and Ad.ΔE1ΔE2A. TALEN-R[S1].F[50] (Supplementary Figure S2). Next, we performed a more sensitive PCR-based assay using primers bracketing the 17.5 repeats of *TALEN-L[S1]* and *TALEN-R[S1]* to search for rearranged *TALEN* sequences in the purified vector preparations. Again, parental constructs

pAd.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and pAd.ΔE1.TALEN-R$^{S1}$.F$^{50}$ served as control DNA templates. This assay led to the detection of 2.1-kb amplicons whose size is diagnostic for vector DNA genomes harboring intact TALEN-encoding genes (Figure 2C). Moreover, Southern blot analysis of extrachromosomal DNA isolated from hMSCs exposed to different amounts of Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ also led to the sole detection of full-length *TALEN* sequences in the transduced cells (Figure 2D). Finally, we cloned into pUC19 the 2.4-kb DNA species from *Bcu*I-treated Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or *Bcu*I-treated Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ and sequenced the inserts from six individual clones. DNA sequence information corresponding to the 7 most downstream and to the 3 to 4 most upstream *TALEN* repeats could be retrieved (Supplementary Figure S3). The nucleotide sequences analyzed revealed to be isogenic to that in the parental constructs with no evidence for point mutations, local rearrangements and/or deletions.

Transduction experiments deploying Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ alone or mixed together were carried out on CAR-positive human cervix carcinoma HeLa cells (42) and on two CAR-negative cell types, namely, immortalized myoblasts from a Duchenne muscular dystrophy (DMD) patient and hMSCs (42,43). The use of relatively modest MOIs was sufficient to transduce most target cells in the various cultures (Figure 3A). Importantly, HA tag-specific immunofluorescence microscopy and Western blot analyses (Figure 3B and 3C, respectively) demonstrated the expression of *TALEN-L$^{S1}$* and *TALEN-R$^{S1}$* genes in target HeLa cells following their transduction with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$. F$^{50}$, respectively. Taken together, we conclude that first-generation adenoviral vectors can deliver expression units encoding full-length TALENs into human cells.

**Figure 2.** Structural analysis of TALEN-encoding adenoviral vector genomes. **(A)** DNA structures of the first-generation adenoviral vectors Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ and of the second-generation adenoviral vectors Ad.ΔE1ΔE2A.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1ΔE2A.TALEN-R$^{S1}$.F$^{50}$ drawn in relation to those of their respective plasmid clones. The adenoviral vector and shuttle plasmid DNA templates are depicted as *Bcu*I physical maps with numerals specifying the different restriction fragment sizes in kilobases (kb). These DNA molecules harbor a bi-cistronic expression unit coding for the *AAVS1*-specific artificial nucleases TALEN-L$^{S1}$ or TALEN-R$^{S1}$ and, through an IRES sequence, a RFP reporter. The open reading frames are under the transcriptional control of the cytomegalovirus immediate-early gene promoter (CMV) and the simian virus 40 (SV40) polyadenylation signal. The central region of *TALEN-L$^{S1}$* and *TALEN-R$^{S1}$* encode a tandem of 17.5 TALE repeats that define specificity to the *AAVS1* target sequence at the human chromosome 19 (19q13.42-qter). C-terminally both TALENs contain the non-specific cleavage domain of FokI, whose catalytic activity is contingent upon DNA binding-dependent dimerization. Each of the TALEN proteins are tagged by an HA antigen located close to their Ntermini (not drawn). The thick and thin lines represent adenoviral vector and plasmid backbone sequences, respectively. Ψ, human adenovirus serotype 5 packaging signal; Ad L-ITR and Ad RITR, "left" and "right", respectively, adenoviral inverted terminal repeat. F50, ORF encoding chimeric fiber composed of basal shaft domains from adenovirus serotype 5 fused to apical shaft and knob motifs from adenovirus serotype 50. The kanamycin-resistance gene (*Kan$^R$*) and the prokaryotic origin of replication (*ori*) are also indicated. **(B)** Restriction fragment length analysis of Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ genomes treated with *Bcu*I. Plasmids pAd.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ served as a reference. The sizes (in kb) of the DNA restriction fragments resulting from the BcuI treatment are indicated on the left. The arrowhead points to the 2.4 kb DNA fragments diagnostic for intact, non-rearranged, *TALE*-derived DNA sequences within TALEN-encoding genes. Marker, molecular-weight marker Gene Ruler DNA Ladder Mix. **(C)** PCR analysis on DNA isolated from Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ particles using primers TALEN-Seq-F and TALEN-Seq-R (F and R half arrows, respectively). These primers, drawn in relation to their respective target sequences, amplify most of the *TALEN-L$^{S1}$* and *TALEN-R$^{S1}$* DNA including the seventeen 102-bp tandem DNA repeats. The position (arrowheads) and size (in kb) of the 2.1-kb amplicon diagnostic for intact *TALEN* transgenes are indicated at the right. Marker, Gene Ruler DNA Ladder Mix. H$_2$0, PCR performed with nuclease-free water instead of vector DNA template. **(D)** Autoradiogram of *Mun*I-treated extrachromosomal DNA isolated from HeLa cells transduced with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or with Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at MOIs of 10, 5 and 1 HTU/cell. The DNA, extracted at three days post-transduction, was subjected to Southern blot analysis using a 397-bp probe hybridizing to a sequence encoding the NH2 terminus of the TALEN-L$^{S1}$ and TALEN-R$^{S1}$ proteins. ▶

**A**

Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$
(MOI = 6 HTU/cell)

Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$
(MOI = 6 HTU/cell)

Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$
+
Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$
(MOI = 6+6 HTU/cell)

phase

RFP

HeLa

Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$
(MOI = 12 HTU/cell)

Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$
(MOI = 12 HTU/cell)

Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$
+
Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$
(MOI = 12+12 HTU/cell)

RFP

Myoblasts

hMSCs

**B**

mock

Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$
(MOI = 3 HTU/cell)

Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$
(MOI = 3 HTU/cell)

Hoechst

HA::TALEN

HeLa

**C**

HTU/cell

4    2    1    0.5    0.25    0

TALEN-L$^{S1}$ →

TALEN-R$^{S1}$ →

HA::TALEN

Finally, to investigate the functionality of TALENs introduced into human cells through adenoviral vector-mediated gene transfer, we deployed the Surveyor mutation detection assay (Figure 4A). PCR amplification of *AAVS1* target DNA from mock-transduced cells or from cells transduced with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or with Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ yielded amplicons that, upon denaturation and reannealing, were not susceptible to site-specific cleavage by the mismatch-sensitive SNS enzyme (Figure 4B, solid arrowhead in the upper panel). This data validated the read-out system by demonstrating the absence of indels at the *AAVS1* locus under these experimental control conditions. However, co-transduction of HeLa cells and DMD myoblasts with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ resulted in efficient cleavage of the endogenous chromosomal target sequence, as evidenced by the detection of large amounts of SNS-digested products (Figure 4B, open arrowheads in the lower panels). This conclusion was substantiated through vector dose-response experiments on HeLa cells, human myoblasts and hMSCs (Figure 4C). Additional transduction experiments on hMSCs that made use of the second-generation adenoviral vectors Ad.ΔE1E2A.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1E2A.TALEN-R$^{S1}$.F$^{50}$ produced results similar to those shown in Figure 4C (Supplementary Figure S4). Finally, we sought to confirm the involvement of the NHEJ pathway in the repair of the TALEN-induced DSBs. To this end, we cloned the *AAVS1*-specific PCR products resulting from genomic DNA of hMSCs exposed to Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at an MOI of 12 HTU/cell each. DNA sequencing of twenty randomly selected clones revealed that 12 of these harbored indels characteristic of NHEJ-mediated repair of DSBs (Figure 4D).
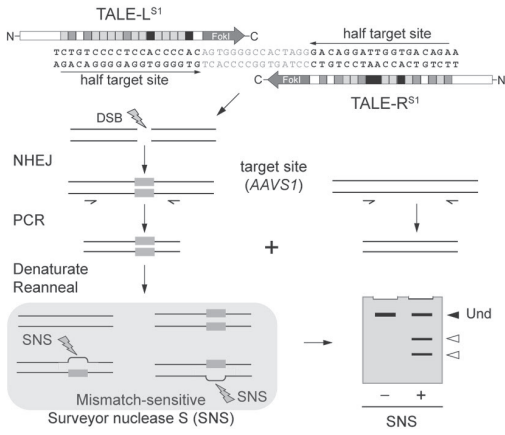
We conclude that delivery of TALENs by first- and second-generation adenoviral vectors is a highly effective approach to generate in a cell type-independent manner site-specific DSBs at an endogenous human target locus.

**Figure 3.** Adenoviral vector-mediated delivery of *TALEN* genes into transformed and nontransformed human cells. **(A)** Gene transfer efficiency. Upper panel, phase-contrast and live-cell RFP direct fluorescence microscopy on HeLa cells transduced with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or with Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at an MOI of 6 HTU/cell or with a mixture of the two adenoviral vectors at an MOI of 6 HTU/cell each. Lower panel, RFP direct fluorescence microscopy on immortalized DMD myoblasts and on primary hMSCs transduced with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or with Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at an MOI of 12 HTU/cell or co-transduced with a 1:1 mixture of Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at a total MOI of 24 HTU/cell. The fluorescence microscopy images were acquired at 48 hours post-infection. **(B)** HA tag-specific immunofluorescence microscopy on HeLa cells mock-transduced or transduced with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or with Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at an MOI of 3 HTU/cell. The direct fluorescence microscopy and the immunofluorescence microscopy were carried out at 48 and 72 hours post-infection, respectively. **(C)** HA tag-directed Western blot analysis of protein lysates derived from HeLa cell cultures incubated for 72 hours with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ or with Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$. The MOIs deployed are indicated.
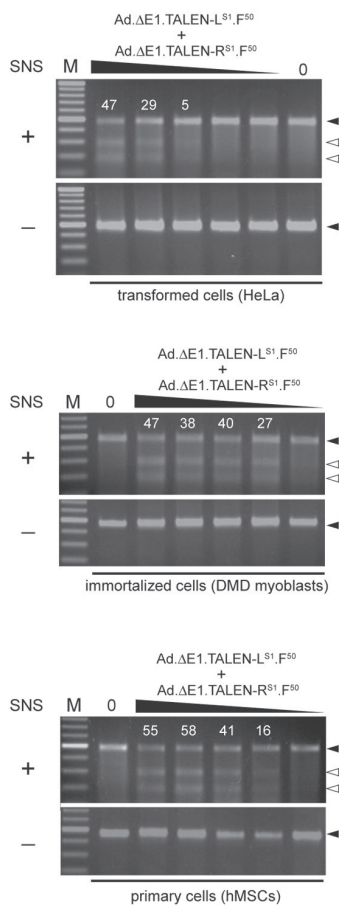◄

**Figure 4.** Targeted disruption of the chromosomal *AAVS1* locus. **(A)** Schematic representation of the Surveyor mutation detection assay used to probe TALEN activity in human cells. Assembly of the TALEN pair at the bipartite target sequence results in local DSB formation. The errorprone NHEJ DNA repair pathway links the broken chromosomal ends introducing in the process small insertions and deletions (indels) at the junction (grey boxes). PCR on chromosomal DNA with primers framing the target site yields a mixture of amplicons representing the heterogeneous population of mutated sequences and non-modified alleles. Denaturation and re-annealing of this mixture results in DNA species with internal base pairs that are mismatched. The Surveyor nuclease S (SNS) enzyme, by preferentially recognizing and cleaving at DNA sequences with such mismatched base pairs, yields low-molecular-weight DNA fragments (open arrowheads) that can be resolved from undigested DNA (solid arrowheads) through conventional agarose gel electrophoreses. Parallel samples not treated with SNS (-) served as a negative control. **(B)** TALEN activity at *AAVS1*. Upper panel, Surveyor assay on chromosomal DNA isolated from HeLa cells not exposed (mock) or exposed to Ad.ΔE1.TALEN-L$^{S1}$. F$^{50}$ or to Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at a high and a low dose (i.e. 6 and 2 HTU/cell, respectively). Lower panels, Surveyor assay on chromosomal DNA extracted from HeLa cells and DMD myoblasts (left and right panels, respectively). HeLa cells were mock-transduced (-) or were co-transduced with 6 HTU/cell of Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and 6 HTU/cell of Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ (+). Human myoblasts were mock-transduced (-) or were co-transduced with 12 HTU/cell of Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and 12 HTU/cell of Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ (+). Solid and open arrowheads indicate the positions of undigested and SNS-digested DNA molecules, respectively. White numerals correspond to the percentage of target site cleavage. **(C)** Dose-response of TALEN activity. Surveyor assay on cellular DNA from HeLa cells, DMD myoblasts and hMSCs untreated or treated with different doses of 1:1 mixtures of Ad.ΔE1.TALEN-L$^{S1}$. F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$. In these dose-response co-transduction experiments the total amounts of vector applied on HeLa cells were 12, 4, 1.3, 0.4 and 0.15 HTU/cell, whereas on DMD myoblasts and hMSCs HTU/cell were 24, 8, 2.7, 1.0 and 0.3. Solid and open arrowheads indicate the positions of undigested and SNS-digested DNA molecules, respectively. White numerals correspond to the percentage of target site cleavage. **(D)** Sequence analysis. DNA sequencing of cloned *AAVS1*-specific amplicons resulting from PCR amplifications on genomic DNA isolated from hMSCs co-transduced with Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$ and Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$ at 12 HTU/cell each. The presence of indels in 12 out of 20 clones, "footprint" for NHEJ processing of DSBs, demonstrates at the nucleotide level the TALEN-mediated cleavage of the chromosomal *AAVS1* target site.  ▶
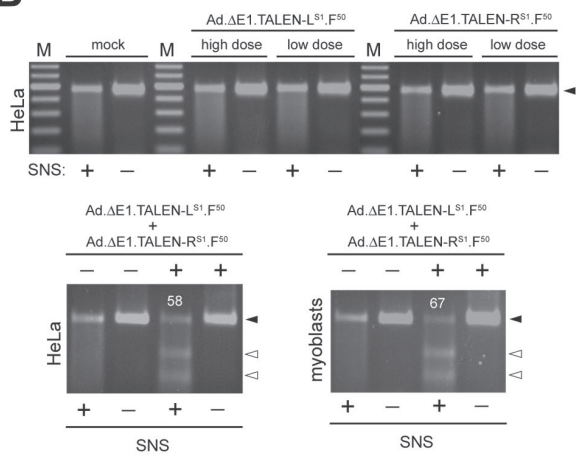
**A**

TALE-L$^{S1}$

half target site

half target site

TALE-R$^{S1}$

DSB

NHEJ

target site (AAVS1)

PCR

Denaturate
Reanneal

SNS

Mismatch-sensitive
Surveyor nuclease S (SNS)

Und

SNS

**B**

Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$

Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$

M    mock    M    high dose    low dose    M    high dose    low dose

HeLa

SNS:   +   −         +   −   +   −            +   −   +   −

Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$
+
Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$

Ad.ΔE1.TALEN-L$^{S1}$.F$^{50}$
+
Ad.ΔE1.TALEN-R$^{S1}$.F$^{50}$

HeLa          58

myoblasts          67

+   −   +   +              −   −   +   +

+   −   +   −              +   −   +   −

SNS                        SNS

**C**

Ad.ΔE1.TALEN-L$^{S1}$,F$^{50}$
+
Ad.ΔE1.TALEN-R$^{S1}$,F$^{50}$

SNS   M                                    0

+       47   29   5

−

transformed cells (HeLa)

Ad.ΔE1.TALEN-L$^{S1}$,F$^{50}$
+
Ad.ΔE1.TALEN-R$^{S1}$,F$^{50}$

SNS   M    0

+            47   38   40   27

−

immortalized cells (DMD myoblasts)

Ad.ΔE1.TALEN-L$^{S1}$,F$^{50}$
+
Ad.ΔE1.TALEN-R$^{S1}$,F$^{50}$

SNS   M    0

+            55   58   41   16

−

primary cells (hMSCs)

**3**

**D**

760    770    780    790    800    810    820    830

12/20                                                    edited

8/20                                                    original

half target site          spacer          half target site

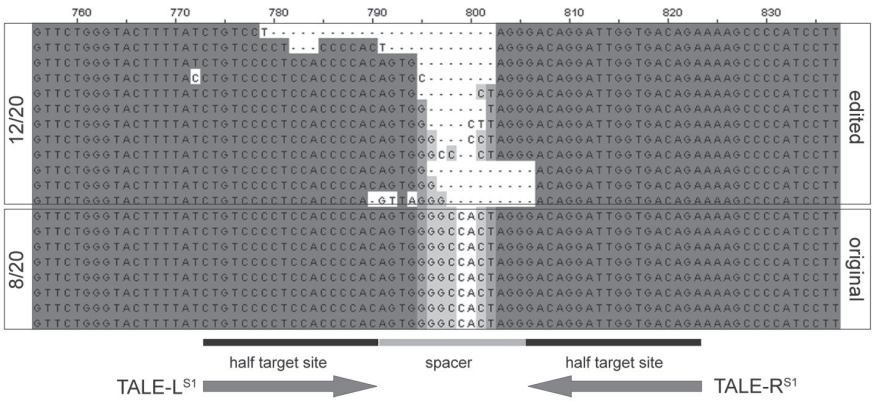TALE-L$^{S1}$                                    TALE-R$^{S1}$

## DISCUSSION

The approaches aiming at editing or modulating the expression of specific endogenous genes and targeting chromosomal insertion of exogenous DNA, have recently been enriched with the introduction of sequence-specific designer TALEs. To broaden the use of TALE-based proteins and to most optimally screen and evaluate their performance in many settings, it will be crucial to expand the range of cellular delivery tools for these molecules. Despite their promise and favorable set of attributes when compared to other custom-made sequence-specific proteins, TALEs are relatively large proteins whose DNA binding specificity is governed by an extensive tract of repetitive sequences. However, DNA repeats are known to be particularly unstable motifs. Indeed, examples of both physiological and function-impairing processes associated with instability of tandem repeats are copious and cover virtually all biological systems from viruses and prokaryotes to eukaryotes (44-46). Also in the context of gene delivery vehicles, such as those based on adenoviral and lentiviral vectors, examples exist pointing to certain repetitive sequence arrangements as the culprits behind DNA instability (47-50). Bearing this in mind, in the present study we sought to investigate the suitability of adenoviral and lentiviral vectors, two of the most commonly used viral vector systems, to serve as *TALEN* gene delivery platforms.

We report that lentiviral vector genomes bearing *TALEN* sequences are prone to rearrangements in target cells. In this regard, the structural analyses of these genomes in HeLa cell populations and in individual HeLa cell clones by PCR and Southern blot analyses, respectively, suggest that most of the rearrangements occurred through recombination events involving the *TALE* repeat array ultimately leading to deletions with various sizes. Conversely, we found that *TALEN* sequences are stably maintained in first- and second-generation adenoviral vectors following their serial propagation in producer cells. Indeed, the *TALEN* transgenes in the adenoviral vector genomes presented neither large rearrangements nor small-scale mutations as revealed by restriction fragment length analysis and DNA sequencing, respectively. Importantly, transduction experiments with the TALEN-encoding adenoviral vectors led to DSB formation at the intended target chromosomal locus in transformed and non-transformed human cells at similarly high frequencies.

Clearly, viruses from the *Adenoviridae* and *Retroviridae* families display diametrically different structural features, life cycles and replication strategies (15,51,52). The latter aspect in particular is likely to contribute to the disparate stability of *TALEN* ORFs in adenoviral versus lentiviral vectors. Adenoviruses replicate via a DNA strand-displacement mechanism using a virus-encoded high-fidelity DNA polymerase and a single-stranded DNA-binding protein that, by cooperatively coating the displaced single-stranded replicative intermediates, presumably inhibits spurious intramolecular hybridization (52,53). Retroviruses, on the other hand, encode error-prone reverse transcriptases (RTs), which convert their double-

copy RNA genomes into linear double-stranded DNA templates that serve as substrates for proviral chromosomal insertion (51). The RTs of simple and complex retroviruses also display a RNase H activity responsible for the digestion of the RNA component of hybrid RNA:DNA duplex intermediates. Moreover, they share a low template affinity and processivity. These features seem to have evolved to allow RT to undergo at least two obligate template switching events (i.e. minus-strand and plus-strand DNA transfer) in order to complete full-length viral DNA synthesis and, as "byproduct", generate genotypic diversity through recombination. It is known, however, that the presence of direct repeats in retroviral genomes increases the likelihood for additional inter- and intra-molecular template switching events, often resulting in the deletion of one of the repeats plus intervening sequences (47,48,54,55). In fact, the insertion of direct repeats in retroviral genomes provides for an elegant model system to study the RT template switching process in vivo or to induce precise rearrangements of *cis*- and *trans*-acting elements in retroviral vectors (54,56). Of note, the frequency of template switching increases with the size of and the distance between the direct repeat units. The latter finding suggests that spacing favors template "looping" which, in turn, aids in the intra-strand alignment of direct repeats, RT template switching and ensuing intra-molecular deletions (55). The models for these direct repeat-induced deletions (48,54,55) postulate that during reverse transcription, RT and the nascent DNA strand dissociates from the RNA template due to the aforementioned RNaseH activity and intrinsic low template affinity. Segments and, eventually, all nascent DNA strand hybridizes to the RNA template at the distal "acceptor" repeat serving as a novel "ectopic" initiation point for RNA-dependent DNA synthesis. The net result of these processes is often homogenously sized deletions of the retroviral genome. In contrast to simple direct repeat pairs, *TALE* sequences present a tract of numerous and contiguous repeats. Thus, each of these arrayed repeat units provide for multiple potential positions at which RT template switching may occur. This corollary is consistent with the heterogeneous and sub-genomic sized population of TALENcontaining lentiviral vector proviruses that were detected in stably transduced clones.

Notwithstanding these findings, one can envision strategies that might minimize or overcome the instability of *TALEN* sequences in lentiviral vectors. For instance, by exploiting the degeneracy of the genetic code, the individual TALE motifs within the array could be reengineered by the introduction of judiciously chosen silent substitutions to increase sequence divergence among repeats. Regarding this issue, it has been shown via a lentiviral vector direct repeat deletion assay that deletion frequencies can decrease steeply with incremental repeat sequence divergence (57).

In conclusion, *TALEN* sequences are prone to rearrangements following their introduction into human cells by lentiviral vectors. As a result, effective lentiviral vectormediated *TALEN* gene transfer will be dependent upon new optimization strategies. Conversely, we demonstrate that first- and second-generation adenoviral vectors are amenable to the

transfer of full-length *TALEN* genes into human cells. Importantly, the resulting TALEN protein yields suffice to induce high-level DSB formation at an endogenous chromosomal target locus.

## SUPPLEMENTARY DATA

Supplementary data are available at NAR Online: Supplementary Figures 1-4.

## FUNDING

## ACKNOWLEDGMENTS

# REFERENCES

1.  Kim,Y.G., Cha,J. and Chandrasegaran,S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA*, **93**, 1156-1160.

2.  Klug,A. (2010) The discovery of Zinc fingers and their applications in gene regulation and genome manipulation. *Annu. Rev. Biochem.*, **79**, 213-231.

3.  Redondo,P., Prieto,J., Muñoz,I.G., Alibés,A., Stricher,F., Serrano,L., Cabaniols,J.P., Daboussi,F., Arnould,S., Perez,C., Duchateau,P., Pâques,F., Blanco,F.J. and Montoya,G. (2008) Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature*, **456**, 107-111.

4.  Silva.G., Poirot,L., Galetto,R., Smith,J., Montoya,G., Duchateau,P. and Pâques,F. (2011) Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr. Gene Ther.*, **11**, 11-27.

5.  Boch,J., Scholze,H., Schornack,S., Landgraf,A., Hahn,S., Kay,S., Lahaye,T., Nickstadt,A. and Bonas,U. (2009) Breaking the code of DNA binding specificity of TALtype III effectors. *Science*, **326**, 1509-1512.

6.  Moscou,M.J. and Bogdanove,A.J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science*, **326**, 150.

7.  Boch,J. and Bonas,U. (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.*, **48**, 419-436.

8.  Miller,J.C., Tan,S., Qiao,G., Barlow,K.A., Wang,J., Xia,D.F., Meng,X., Paschon,D.E., Leung,E., Hinkley,S.J., Dulay,G.P., Hua,K.L., Ankoudinova,I., Cost,G.J., Urnov,F.D., Zhang,H.S., Holmes,M.C., Zhang,L., Gregory,P.D. and Rebar,E.J. (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.*, **29**, 143-148.

9.  Mussolino,C., Morbitzer,R., Lütge,F., Dannemann,N., Lahaye,T. and Cathomen,T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.*, **39**, 9283-9293.

10. Mussolino,C. and Cathomen,T. (2012) TALE nucleases: tailored genome engineering made easy. *Curr. Opin. in Biotech.*, **23**, 1-7.

11. Cathomen,T. and Joung,K.J. (2008) Zinc-finger nucleases: the next generation emerges. *Mol. Ther.*, **16**, 1200-1207.

12. Hockemeyer,D., Wang,H., Kiani,S., Lai,C.S., Gao,Q., Cassady,J.P., Cost,G.J., Zhang,L., Santiago,Y., Miller,J.C., Zeitler,B., Cherone,J.M., Meng,X., Hinkley,S.J., Rebar,E.J., Gregory,P.D., Urnov,F.D. and Jaenisch,R. (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.*, **29**, 731-734.

13. Tong,C., Huang,G., Ashton,C., Wu,H., Yan,H. and Ying,Q.L. (2012) Rapid and costeffective gene targeting in rat embryonic stem cells by TALENs. *J. Genet. Genomics*, **39**, 275-280.

14. Sun,N., Liang,J., Abil,Z. and Zhao,H. (2012) Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. *Mol. Biosyst.*, **8**, 1255-1263.

15. Gonçalves,M.A.F.V. and de Vries,A.A.F. (2006) Adenovirus: from foe to friend. *Rev. Med. Virol.*, **16**, 167-186.

16. Dropulić,B. (2011) Lentiviral vectors: their molecular design, safety, and use in laboratory and preclinical research. *Hum. Gene Ther.*, **22**, 649-657.

17. Yamashita,M. and Emerman,M. (2006) Retroviral infection of non-dividing cells: old and new perspectives. *Virology*, **344**, 88-93.

18. Henaff,D., Salinas,S. and Kremer,E.J. (2011) An adenovirus traffic update: from receptor engagement to the nuclear pore. *Future Microbiol.*, **6**, 179-192.

19.  Fallaux,F.J., Bout,A., van der Velde,I., van den Wollenberg,D.J., Hehir,K.M., Keegan,J., Auger,C., Cramer,S.J., van Ormondt,H., van der Eb,A.J., Valerio,D. and Hoeben,R.C. (1998) New helper cells and matched early region 1-deleted adenovirus vectors prevent generation of replication-competent adenoviruses. *Hum. Gene Ther.*, **9**, 1909-1917.

20.  Havenga,M.J., Holterman,L., Melis,I., Smits,S., Kaspers,J., Heemskerk,E., van der Vlugt,R., Koldijk,M., Schouten,G.J., Hateboer,G., Brouwer,K., Vogels,R. and Goudsmit,J. (2008) Serum-free transient protein production system based on adenoviral vector and PER.C6 technology: high yield and preserved bioactivity. *Biotechnol. Bioeng.*, **100**, 273-283.

21.  Cudré-Mauroux,C., Occhiodoro,T., König,S., Salmon,P., Bernheim,L. and Trono,D. (2003) 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.

22.  Gonçalves,M.A.F.V., Janssen,J.M., Nguyen,Q.G., Athanasopoulos,T., Hauschka,S.D., Dickson,G. and de Vries,A.A.F. (2011) Transcription factor rational design improves directed differentiation of human mesenchymal stem cells into skeletal myocytes. *Mol. Ther.*, **19**, 1331-1341.

23.  Janssen,J.M., Liu,J., Skokan,J., Gonçalves,M.A.F.G. and de Vries,A.A.F. (2012) Development of an AdEasy-based system to produce first- and second-generation adenoviral vectors with tropism for CAR- or CD46-positive cells. *Submitted*.

24.  Seppen,J., Rijnberg,M., Cooreman,M.P. and Oude Elferink,R.P. (2002) Lentiviral vectors for efficient transduction of isolated primary quiescent hepatocytes. *J. Hepatol.*, **36**, 459-465.

25.  Pelascini,P.L., Janssen,J.M. and Gonçalves,M.A.F.V. (2012) Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. Hum. Gene Ther. *In press*.

26.  Gonçalves,M.A.F.V., Swildens,J., Holkers,M., Narain,A., van Nierop,G.P., van de Watering,M.J., Knaän-Shanzer,S. and de Vries,A.A.F. (2008) Genetic complementation of human muscle cells via directed stem cell fusion. *Mol. Ther.*, **16**, 741-748.

27.  Gonçalves,M.A.F.G., Holkers,M., van Nierop,G.P., Wieringa,R., Pau,M.G. and de Vries,A.A.F. (2008) Targeted chromosomal insertion of large DNA into the human genome by a fiber-modified high-capacity adenovirus-based vector system. *PLoS One*, **29**, e3084.

28.  van Nierop,G.P., de Vries,A.A.F., Holkers,M., Vrijsen,K.R. and Gonçalves,M.A.F.V. (2009) Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease. *Nucleic Acids Res.*, **37**, 5725-5736.

29.  Gonçalves,M.A.F.V., van der Velde,I., Knaän-Shanzer,S., Valerio,D. and de Vries,A.A.F. (2004) Stable transduction of large DNA by high-capacity adeno-associated virus/adenovirus hybrid vectors. *Virology*, **321**, 287-296.

30.  Szulc,J., Wiznerowicz,M., Sauvain,M.O., Trono,D. and Aebischer,P. (2006) A versatile tool for conditional gene expression and knockdown. *Nat. Methods*, **3**, 109-116.

31.  Philpott,N.J. and Thrasher,A.J. (2007) Use of nonintegrating lentiviral vectors for gene therapy. *Hum. Gene Ther.*, **18**, 483-489.

32.  Wanisch,K. and Yáñez-Muñoz,R.J. (2009) Integration-deficient lentiviral vectors: a slow coming of age. *Mol. Ther.*, **17**, 1316-1332.

33.  Staunstrup,N.H. and Mikkelsen,J.G. (2011) Integrase-defective lentiviral vectors-a stage for nonviral integration machineries. *Curr. Gene Ther.*, **11**, 350-362.

34.  Cornu,T.I. and Cathomen,T. (2007) Targeted genome modifications using integrasedeficient lentiviral vectors. *Mol. Ther.*,**15**, 2107-2113.

35.  Lombardo,A., Genovese,P., Beausejour,C.M., Colleoni,S., Lee,Y.L., Kim,K.A., Ando,D., Urnov,F.D., Galli,C., Gregory,P.D., Holmes,M.C. and Naldini,L. (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.*, **25**, 1298-1306.

36. Izmiryan,A., Basmaciogullari,S., Henry,A., Paques,F. and Danos,O. (2011) Efficient gene targeting mediated by a lentiviral vector-associated meganuclease. *Nucleic Acids Res*., **39**, 7610-7619.

37. Shayakhmetov,D.M., Papayannopoulou,T., Stamatoyannopoulos,G. and Lieber,A. (2000) Efficient gene transfer into human CD34(+) cells by a retargeted adenovirus vector. *J. Virol*., **74**, 2567-2583.

38. Knaän-Shanzer,S., van der Velde,I., Havenga,M.J., Lemckert,A.A., de Vries,A.A.F. and Valerio,D. (2001) Highly efficient targeted transduction of undifferentiated human hematopoietic cells by adenoviral vectors displaying fiber knobs of subgroup B. *Hum. Gene Ther*., **12**, 1989-2005. Erratum in: *Hum. Gene Ther*. 2003 Aug 10;14.

39. Knaän-Shanzer,S., van de Watering,M.J., van der Velde,I., Gonçalves,M.A.F.V., Valerio,D. and de Vries,A.A.F. (2005) Endowing human adenovirus serotype 5 vectors with fiber domains of species B greatly enhances gene transfer into human mesenchymal stem cells. *Stem Cells*, **23**, 1598-1607.

40. Tuve,S., Wang,H., Ware,C., Liu,Y., Gaggar,A., Bernt,K., Shayakhmetov,D., Li,Z., Strauss,R., Stone,D. and Lieber,A. (2006) A new group B adenovirus receptor is expressed at high levels on human stem and tumor cells. *J. Virol*., **80**, 12109-12120.

41. Hoffmann,D., Meyer,B. and Wildner,O. (2007) Improved glioblastoma treatment with Ad5/35 fiber chimeric conditionally replicating adenoviruses. *J. Gene Med*., **9**, 764-778.

42. Gonçalves,M.A.F.V., Holkers,M., Cudré-Mauroux,C., van Nierop,G.P., Knaän-Shanzer,S., van der Velde,I., Valerio,D. and de Vries,A.A.F. (2006) Transduction of myogenic cells by retargeted dual high-capacity hybrid viral vectors: robust dystrophin synthesis in duchenne muscular dystrophy muscle cells. *Mol. Ther*., **13**, 976-986.

43. Gonçalves,M.A.F.V., de Vries,A.A.F., Holkers,M., van de Watering,M.J., van der Velde,I., van Nierop,G.P., Valerio,D. and Knaän-Shanzer,S. (2006) Human mesenchymal stem cells ectopically expressing full-length dystrophin can complement Duchenne muscular dystrophy myotubes by cell fusion. *Hum. Mol. Genet*., **15**, 213-221.

44. Pearson,C.E., Nichol Edamura,K. and Cleary,J.D. (2005) Repeat instability: mechanisms of dynamic mutations. *Nat. Rev. Genet*., **6**, 729-742.

45. Zhao,J., Bacolla,A., Wang,G. and Vasquez,K.M. (2010) Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci*., **67**, 43-62.

46. Holkers,M., de Vries,A.A.F. and Gonçalves,M.A.F.V. (2012) Nonspaced inverted DNA repeats are preferential targets for homology-directed gene repair in mammalian cells. *Nucleic Acids Res.*, **40**, 1984-1999.

47. Rhode,B.W., Emerman,M. and Temin,H.M. (1987) Instability of large direct repeats in retrovirus vectors. *J. Virol*., **61**, 925-927.

48. Pathak,V.K. and Temin,H.M. (1990) Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: deletions and deletions with insertions. *Proc. Natl. Acad. Sci. USA*, **87**, 6024-6028.

49. Steinwaerder,D.S., Carlson,C.A. and Lieber,A. (1999) Generation of adenovirus vectors devoid of all viral genes by recombination between inverted repeats. *J. Virol*., **73**, 9303-9313.

50. Belousova,N., Harris,R., Zinn,K., Rhodes-Selser,M.A., Kotov,A., Kotova,O., Wang,M., Aurigemma,R., Zhu,Z.B., Curiel,D.T. and Alvarez,R.D. (2006) Circumventing recombination events encountered with production of a clinical-grade adenoviral vector with a double-expression cassette. *Mol. Pharmacol*., **70**, 1488-1493.

51. Coffin,J.M., Hughes,S.H. and Varmus,H.E. (1997) Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

52. de Jong,R.N., van der Vliet,P.C. and Brenkman,A.B. (2003) Adenovirus DNA replication: protein priming, jumping back and the role of the DNA binding protein DBP. *Curr Top Microbiol Immunol.*, **272**, 187-211.

53.   Zijderveld,D.C., Stuiver,M.H. and van der Vliet,P.C. (1993) The adenovirus DNA binding protein enhances intermolecular DNA renaturation but inhibits intramolecular DNA renaturation. *Nucleic Acids Res*., **21**, 2591-2598.

54.   Julias,J.G., Hash,D. and Pathak,V.K. (1995) E- vectors: development of novel selfinactivating and self-activating retroviral vectors for safer gene therapy. *J. Virol*., **69**, 6839-6846.

55.   Delviks,K.A. and Pathak,V.K. (1999) Effect of distance between homologous sequences and 3' homology on the frequency of retroviral reverse transcriptase template switching. *J. Virol*., **73**, 7923-7932.

56.   Delviks,K.A., Hu,W.S. and Pathak,V.K. (1997) Psi- vectors: murine leukemia virusbased self-inactivating and self-activating retroviral vectors. *J. Virol*., **71**, 6218-6224.

57.   An,W. and Telesnitsky,A. (2002) Effects of varying sequence similarity on the frequency of repeat deletion during reverse transcription of a human immunodeficiency virus type 1 vector. *J. Virol*., **76**, 7897-7902.