UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

GENERATION AND CHARACTERIZATION OF NOVEL ADENOVIRAL VECTORS FOR HYBRID NUCLEASE-MEDIATED GENE TARGETING

SARA FILIPA DIAS HENRIQUES

DISSERTATION

(PUBLIC VERSION)

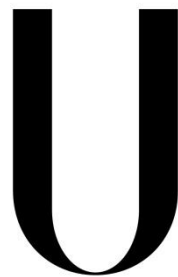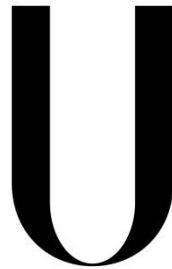### MESTRADO EM BIOLOGIA HUMANA E AMBIENTE

2013

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

U

LISBOA

UNIVERSIDADE
DE LISBOA

GENERATION AND CHARACTERIZATION OF NOVEL ADENOVIRAL VECTORS FOR HYBRID NUCLEASE-MEDIATED GENE TARGETING

SARA FILIPA DIAS HENRIQUES

External supervisor: Doutor Manuel A. F. V. Gonçalves (Leiden University Medical Center, The Netherlands)
Internal supervisor: Professora Doutora Deodália Dias (Faculdade de Ciências da Universidade de Lisboa, DBA, Portugal)

DISSERTATION

(PUBLIC VERSION)

MESTRADO EM BIOLOGIA HUMANA E AMBIENTE

2013

Foreword

Items (tables, figures and references) in this thesis are presented according to the format of the journal *Nature Methods.*

Acknowledgements

Abstract

Adenovirus-based vectors are among the most efficient and disseminated gene transfer vehicles currently in use in a broad range of basic and applied research applications including the testing of gene therapy. As a gene therapy modality, gene targeting relies on the site-specific genome modification based on "integrases" or on the error-free homologous recombination pathway of the cell.

Here, by deploying engineered nucleases together with targeting donor DNA delivered by adenoviral vectors, was tested the influence of the homologous donor topology on the specificity and the accuracy of gene targeting. Was shown clear evidence for the susceptibility of adenoviral vector DNA to the catalytic activities of FLPe recombinase and TALEN proteins. Also, the excision of donor DNA from the context of AdV genomes leads to an increase in stable transduction levels with the majority of integrants having precise HR-derived junctions at both ends, indicating that the vector topology greatly impacts the outcome of the genomic editing process.

Keywords: adenoviral-vectors; TALENs; gene targeting; homologous-donor; homologous-recombination

Resumo

O domínio científico da terapia genética tem, nos últimos anos, sido alvo de grande expansão e desenvolvimento, sendo descrita como a introdução de nova informação genética, por meio de vetores, em células de um indivíduo com a finalidade de corrigir um defeito no genoma. Muitas dessas correções têm como alvo genes implicados em doenças, como na distrofia muscular de Duchenne, uma doença ligada ao cromossoma X. A ambição de uma cura genética tem impulsionado a descoberta e aperfeiçoamento de ferramentas úteis a essa ciência, tanto na área dos vetores utilizados (nomeadamente os sistemas víricos), como em endonucleases de restrição como as nucleases ZFNs e TALENs.

Vetores baseados em adenovírus estão entre os veículos de transferência de genes mais eficientes e são, atualmente, de utilização disseminada. São utilizados numa grande variedade de aplicações de pesquisa básica e aplicada e apresentam características que os distinguem dos restantes sistemas utilizados, como uma eficiente transdução da maioria dos tecidos, uma grande capacidade de acomodação de transgenes e a forma episomal em que o seu genoma permanece na célula, impedindo indesejáveis integrações aleatórias no genoma. Com o consecutivo melhoramento da ciência, também os vetores baseados em adenovírus foram sucessivamente aperfeiçoados, tendo a remoção total ou parcial de genes virais não essenciais, aumentado o tamanho de acomodação do transgene e diminuindo as respostas imunitárias ao vector. Atualmente são descritos essencialmente três tipos de vetores adenovíricos: primeira-geração, segunda- e terceira-geração e *high-capacity*, com deleções nos genes virais *E1, E2/E4* e em todos os genes, respetivamente.

Uma das modalidades da terapia genética, a terapia genética *targeted*, baseia-se na modificação, através de "integrases" ou na via de recombinação homóloga da célula, do genoma numa localização específica, como o locus humano *AAVS1*, no cromossoma 19, um local largamente utilizado para integração de genes exógenos.

A recombinação homóloga é uma via, isenta de erros, de reparação de quebras de ADN utilizada pela célula e é dependente da existência de uma sequência dadora (ADN dador), homóloga em parte com a sequência onde a quebra se deu. Por recombinação entre as duas sequências dá-se troca de informação genética complementar e a correção da quebra. Este sistema é muito utilizado em terapia genética, nomeadamente em sistemas de correção de mutações, mutagénese ou de inserção genética específica. Apesar de ser uma via de reparação eficiente, a célula

utiliza mais prontamente uma outra via de reparação: a "*non-homologous end joining*", em que, como o nome indica, liga as duas pontas livres, produzindo erros no processo. Para uma terapia genética eficiente e livre de erros, a recombinação homóloga deve ser induzida preferencialmente. Sabe-se que este sistema é recrutado preferencialmente em quebras duplas de ADN e a introdução exógena destas lesões por endonucleases de restrição, como as ZFNs e TALENs, pode aumentar a frequência de ocorrência da recombinação homóloga.

Transcription activator-like effectors (TALEs) são proteínas encontradas na bactéria fitopatogénica do género *Xanthomonas* e que funcionam essencialmente como activadores transcripcionais. A estrutura modular e a sua capacidade de ligação ao ADN numa relação de um-para-um (resíduo-nucleótido) altamente específica fazem com que, associadas à enzima de restrição FokI, tenham dado origem às nucleases TALENs que, após dimerização dos domínios FokI, produzem uma quebra dupla na sequência desejada. As proteínas TALEN podem ser introduzidas na célula eficientemente através de vetores víricos como os vetores adenovíricos.

No presente trabalho, através da transferência de TALENs (específicas para o locus *AAVS1* humano) e de ADN dador, por vetores adenovíricos de segunda-geração, foi testada a influência da topologia do ADN dador homólogo sobre a especificidade e a precisão da terapia genética *targeted*.

Os dadores homólogos foram construídos numa base genómica adenovírica, com duas sequências homólogas ao locus *AAVS1*, flanqueando o transgene *eGFP* direcionado por um promotor PGK. As topologias finais dos dadores foram produzidas a partir de reacções em sequências clonadas: *flipase recognitition target* (FRT), TALEN-*target site* (T-TS) e *adeno-associated virus inverted terminal repeat* (AAV-ITR). Pela recombinação das sequências FRT, mediada pela recombinase flipase, resultou um dador circular; pela libertação do dador do contexto adenovírico, mediado pela quebra dupla por TALENs nas sequências T-TS, foi produzido um dador linear e pelas sequências AAV-ITR, foi produzido um dador com estruturas secundárias terminais com recrutamento de fatores de reparação. Após aquisição da topologia final, e por quebra dupla específica no locus *AAVS1* mediada pelas TALENs, a via de recombinação homóloga será recrutada, levando à recombinação entre as sequências homólogas no locus e no ADN dador e, consequentemente, à integração do transgene.

Após a caracterização dos vetores por análise enzimática de restrição, demonstrou-se evidências da suscetibilidade do ADN vetor dador para as atividades catalíticas da flipase e das proteínas TALEN através de técnicas de PCR. A atividade

de transferência dos dadores pelos vetores adenovirais foi determinada por titulações em células Hela, após as quais se procedeu às experiências de transdução, tanto em células HeLa como em mioblastos derivados de pacientes com distrofia muscular de Duchenne.

No sentido de investigar a especificidade e precisão da recombinação, mediada por TALENS, dos diferentes dadores no locus-alvo *AAVS1*, a frequência de células positivas para *eGFP* foi monitorizada por citometria de fluxo ao longo do período de tempo em que as células transduzidas foram mantidas em cultura. Observou-se na primeira experiência de transdução, em células HeLa, que apesar dos elevados níveis iniciais de células positivas para *eGFP*, essas frequências rapidamente diminuiram e, após 28 dias, foram consideradas como níveis basais em todas as condições à excepção das células transduzidas com o dador linear (com sequências T-TS) em combinação com as TALEN. Esta condição apresentou 1.7 vezes maior frequência de células positivas para *eGFP* que o dador controlo. O período da experiência foi suficiente para diluir as formas episomais do vetor, sendo este valor final resultado da integração genómica do transgene.

Uma segunda experiência de transdução em células HeLa, neste caso utilizando somente o dador linear (T-TS) e as TALENs, confirmou os resultados da experiência inicial.

Uma última experiência de transdução foi feita, utilizando mioblastos derivados de pacientes de Duchenne e, para além do dador linear T-TS, um dador normal para comparação. Igualmente nesta experiência, o dador linear em combinação com as TALEN revelou uma frequência de células positivas para *eGFP* mais elevada que o dador controlo.

Porque um aumento nos níveis de eficiência de integração do transgene não demonstra directamente um equivalente aumento na especificidade, procedeu-se a uma avaliação desse parâmetro por experiências de PCR. Para isso, isolaram-se clones de células HeLa da primeira experiência, positivos para *eGFP*, da condição do dador linear T-TS em combinação com as TALEN. Primers foram desenhados de forma a ser avaliada a integração do transgene, tanto na extremidade 5' como na 3' do locus-alvo *AAVS1*, assim como primers para o gene *eGFP* como controlo interno. Em 30 clones isolados, somente 22 clones (73%) revelaram junções do locus provenientes de recombinação homóloga.

Com base nos resultados deste trabalho pode-se afirmar que o contexto do vetor influencia a especificidade e a precisão do processo de edição genético.

Comparação com outros sistemas víricos poderá levar ao desenvolvimento de ferramentas e sistemas mais eficientes e seguros para o campo da terapia genética.

Palavras-chave: vetores adenovíricos; TALENs; terapia genética *targeted*; dador homólogo; recombinação homóloga

Table of contents

Figures index

Tables index

Abbreviations

AdV: adenoviral vector

HAdV: human adenoviral vector

AAV: adeno-associated viruses

AAV-ITR: adeno-associated virus inverted terminal repeat

CAR: coxsackie B virus and adenovirus receptor

CPE: cytopathic effect

CTL: cytotoxic T lymphocite

DBP: DNA-binding protein

dsDNA: double-stranded DNA

DSB: double-strand breaks

eGFP: enhanced green fluorescent protein

FLPe: flipase

FRT: flipase recombinase target

HAdV: human adenoviruses

HC-AdV: high-capacity adenoviral vector

HR: homologous recombination

ITR: inverted terminal repeat

MOI: multiplicity of infection

NHEJ: non-homologous end-joining

OTC: ornithine transcarbamylase deficiency

PGK: phosphoglycerate kinase

Pol: DNA polymerase

Pr: protease

pTP: precursor of the terminal protein

RCA: replication-competent AdV

RVD: repeat variable di-residue

ssDNA: single-stranded DNA

TALE: transcription activator-like effector

TALEN: TALE nuclease

TP: terminal protein

T-TS: TALEN target site

ZFN: zinc-finger nucleases

Chapter 1. Introduction

1.1 Genetic Therapy

Gene or genetic therapy is typically defined as the delivery of new genetic information, through a gene delivery vehicle or vector, into cells of an individual, with the intention of functionally correcting a defect in the genome[1, 2]. This genetic alteration is made in order to restore health, halt, slow down or reverse the progression of a disease[3].
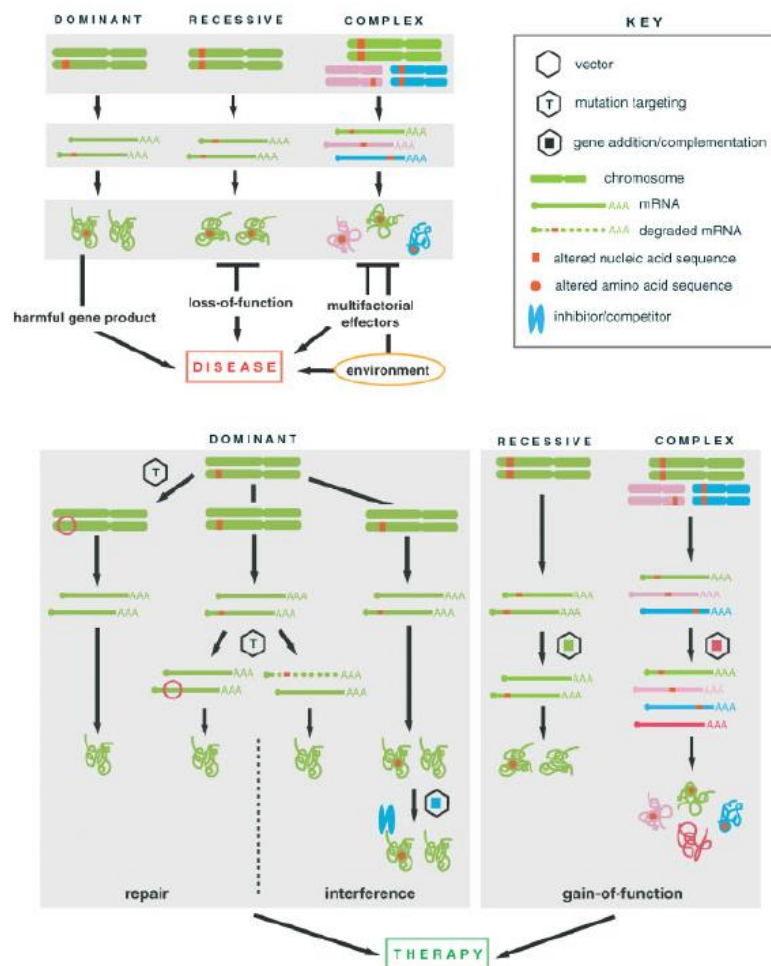
Figure 1 | Overview of the main gene therapy approaches directed towards treating the different classes of genetic defects. Mutations in the genome of an individual can lead, ultimately, to disease. Gene therapy aim is, through vectors, restore that genetic defect[2].

First designed as a therapy for hereditary monogenetic recessive disorders[4], gene therapy has been adapted, and today is being tested in multiple areas, including the treatment of complex diseases such as cardiovascular diseases[5], cancer[6], neurodegenerative disorders[7], as well as infectious diseases[8,9]. The ground-breaking field of gene therapy had to surpass the initial clinical failures attributed to the inefficacy of early recombinant viral vectors and reduced transgene expression levels. It became clear that the understanding of the fundamentals of gene therapy science were lacking, motivating the scientific community on an interdisciplinary quest to establish specific, efficient and safe gene delivery systems.
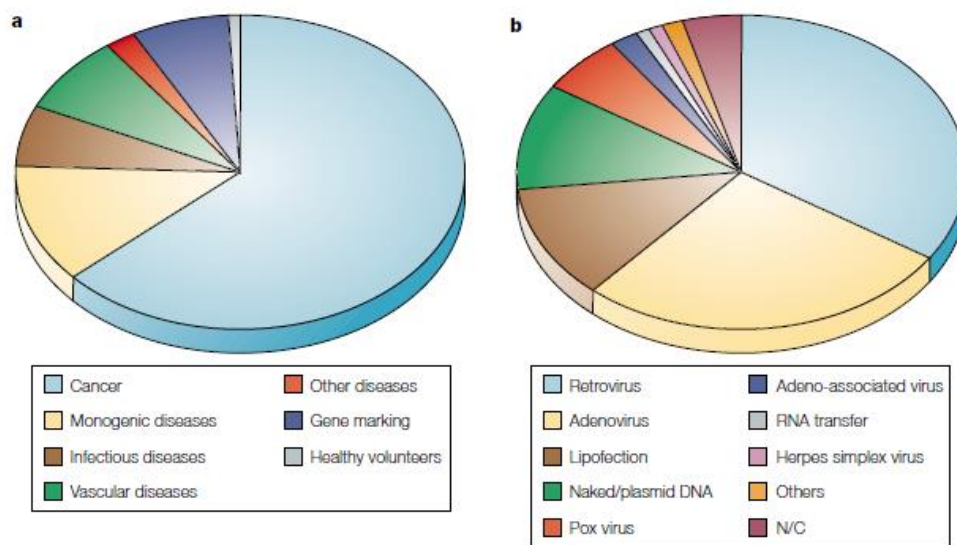
Figure 2 | Gene therapy clinical trials. The clinical trials are categorized by (**a**) types of diseases and (**b**) types of vectors used[10].

Gene therapy is typically designed with three alternative modes of action: gene addition, gene knockout or gene correction/alteration[1]. Gene addition is necessary to correct, by complementation, the phenotype of a disorder in which a gene is functionally disabled as a result of a mutation (deletion or otherwise) such as in the case of the dystrophin-encoding *DMD* gene causing Duchenne muscular dystrophy[11]. Gene knockdown or, ideally, gene knockout, is required if the mutation in a gene results in the expression of a gene product with a dominant damaging action in the cell. However, correcting a genetic mutation through gene targeting is regarded to be the ideal approach. Importantly, the development of tools that might allow a therapeutically successful use of gene targeting strategies is currently well underway. These tools include designer nucleases such as zinc-finger nucleases (ZFNs)[12] and transcription

activator-like effector nucleases (TALENs)[13]. These artificial proteins can be exploited together with the endogenous DNA repair mechanisms of the cell, that is, non-homologous end-joining (NHEJ) and homologous recombination (HR)[14], to bring about a specific genetic modification (**Fig. 3**).
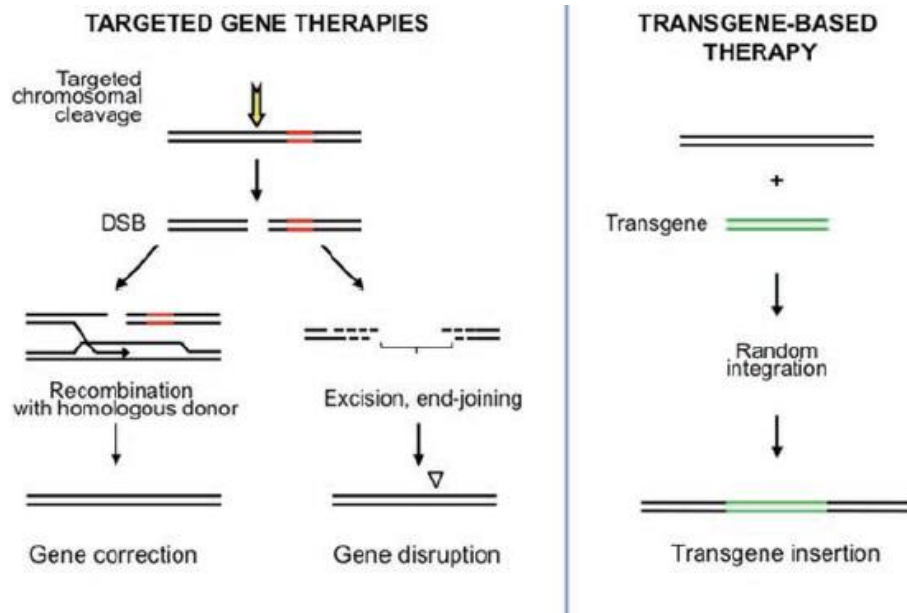
Figure 3 | Targeted gene therapies versus conventional gene therapy. Targeted gene therapy uses specific and targeted tools to, either restore the mutated essential gene, or disrupt a deleterious one. Transgene-based therapy aims to integrate a transgene with the required genes to a random location in the host cell genome[15].

1.1.1 Vectors

The transduction of the target cells is the first challenge in gene therapy and is typically achieved by using viral or non-viral vectors[3]. These can be directly administered *in vivo* or via the *ex vivo* transduction of autologous cells derived from the patient. Depending on the type of vector, the transgene can be designed to either integrate into the host cell genome or remain in an episomal state in the nucleus of target cells[1]. The selection of which type of vector to use (i.e. episomal or chromosomally-integrating) depends on the specific gene therapy setting.

1.1.1.1 Non-viral vectors

Typically, these vectors consist of "naked" plasmid DNA packaged within liposomes[16] or complexed with polycations. They are widely used owing especially to the following attributes: simplicity, absence of immunogenic viral components, no transgene size limit and ease of large scale production[1, 3]. The main drawback is their limited gene transfer efficiency and only transient transgene expression[17].

1.1.1.2 Viral vectors

Viruses are for long time known to efficiently transduce their host cells and, with the rise of the gene therapy science, they were envisioned, by Tatum and Rogers independently (1966)[18, 19], as potentially good vectors for gene delivery. The proof-of-concept for the use of viruses as gene delivery vehicles was demonstrated by Rogers, in 1968, in a Nature article titled: "Use of viruses as carriers of added genetic information"[20]. The RNA-containing murine oncoretroviruses were the first viruses to be modified into viral vectors to be used as gene delivery systems for gene therapy. Subsequently, DNA viruses, such as adenoviruses and adeno-associated viruses, started also, in the mid-1980s, to be exploited as gene delivery systems. Together, these studies demonstrated the clear advantage of using viral vectors for high-level transgene expression[3].

Viral vectors are derived from their wild-type parental viruses by the replacement of several viral genes for a foreign or exogenous transgene. The gene products corresponding to the deleted viral genes are provided *in trans* by complementary producer cell lines[3] ensuring the production of the replication-defective viral vector. Viral vectors can be divided/categorized according to their persistence in the target cell as chromosomally integrating or non-integrating vectors. Thus, the former vector types, by integrating into the host cell genome, have the potential for providing life-long transgene expression, whereas the latter, by remaining as an extrachromosomal episomal form, are diluted out with the consecutive cell divisions providing, as a result, short-term transgene expression[10]. Due to these features, integrating viral vectors, like oncoretroviral and lentiviral vectors, are used for stable

and long-lasting gene expression, whereas the non-integrating viral vectors such as adenoviral vectors are used for transient transgene expression[21].

As aforementioned, murine oncoretroviral vectors were among the first to be developed into gene delivery vehicles. These recombinant RNA viral vectors retain important parental virus-derived features as well as some major limitations (**Table 1**), namely their inability to infect non-dividing cells and their semi-random DNA integration profile that might cause host gene disruption and/or proto-oncogene deregulation[10]. Unlike the previous ones, lentiviral vectors possess the feature of transducing both dividing and non-dividing cells, making them good vectors for gene therapy clinical trials[10, 21]. Adeno-associated viruses (AAV) are small single-stranded DNA viruses that require the help of another virus, such as adenovirus or herpesvirus, to complete their life cycle. They possess some key features that make them attractive viral vectors but also have some disadvantages[10].

Integrating viral vectors insert the exogenous DNA in a semi-randomly fashion into the host cell genome and, as a result, can cause the disruption of important cellular genes and/or deregulate the expression of proto-oncogenes potentially leading to tumorogenesis as the induction of Murine Leukemia[22]. So there is currently a pressing need to develop viral vector systems that insert the foreign DNA into specific and safe genomic regions.

Table 1 | Characteristics of the main viral vectors used in gene therapy[10].

Vector	Packaging capacity	Tropism	Immune response	Vector genome form	Main advantages	Main limitations
Enveloped, RNA genome						
Retrovirus	8 kb	Only dividing cells	Low	Integrated	Persistent gene transfer in dividing cells	Only transduces dividing cells Integration might induce oncogenesis
Lentivirus		Broad			Persistent gene transfer in most tissues	Integration might induce oncogenesis Unknown safety risk
Non-enveloped, DNA genome						
AAV	< 5 kb	Broad	Low	Episomal (>90%) Integrated (<10%)	Non-pathogenic	Small packaging space
Adenovirus	8 -30 kb		High	Episomal	Efficient transduction of most tissues	Potent Immune response

1.2 Adenovirus

Adenoviruses were isolated from adenoid tissue in 1953[23]. Adenoviruses are included in the genus Mastadenovirus of the *Adenoviridae* family and are currently divided into 5 genera and sub-divided into species from A through F[24]. The human adenoviruses (HAdVs) are divided in more than 50 different serotypes that display various degrees of tissue specificity such as for skeletal muscle, lung and brain[24, 25].

They infect a wide-range of hosts like non-human primates, mouse, dog, pig, chicken and humans[26]. With the exception of immune-compromised individuals[27],

human adenoviruses are mostly only mildly pathogenic, targeting, amongst other tissues, the respiratory, gastrointestinal and urinary tracts as well as the eye[27].

1.2.1 Virion structure

The 90-nm non-enveloped icosahedral HAdV particle consists of a viral capsid comprising assemblies of the three major proteins: 240 trimers of hexon protein (pII), 12 pentamers of penton base polypeptide (pIII) and trimers of knobbed fiber proteins (pIV)[28]. They also have other minor proteins such as pIIIa, pVIII, pIX, pIVa2, pVI (the core protein that binds tightly to the viral genome forming a nucleoprotein complex[29]) and pV linking the viral chromosome to the virion shell[30]. Moreover, the virion contains a virus-encoded protease (Pr), necessary for the processing of some of the structural proteins needed for the maturing of the viral particle[31]. The fibers (trimers of pIV) protrude at the vertices of the adenoviral capsid. The fibers consist of a N-terminal structure that interacts with the penton base, a rod-like shaft, and a C-terminal globular domain or "knob" motif which is responsible, necessary and sufficient, for the primary attachment of the virus to host cellular receptors[28].
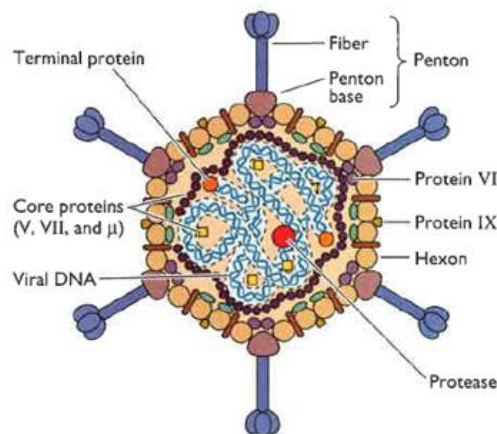
Figure 4 | Adenovirus virion structure[32].

Adenoviruses have a lytic life cycle and the early infection phase comprises the selective attachment, entry of the virus into the host cell and passage of the viral genome into the cell nucleus[25].

1.2.2 Cellular receptors and life cycle

The primary cellular receptor for the HAdVs from species C serotypes, such as type-2 and type-5[33], as well as from those of species A, D, E and F[32], is the Coxsackie B virus and adenovirus receptor (CAR)[34], a 46 kDa plasma membrane protein of the immunoglobulin superfamily. HAdVs from other species utilized other cellular receptors. For instance, species B HAdVs use the transmembranar protein CD46 as their primary attachment and entry moiety[24].

After cell receptor binding, the HAdV particles start to shed their fibers and are internalized through a clathrin coated pit/vesicle endocytosis pathway[35, 36]. As the pH decreases in the route from the endosome to lysosome, it triggers conformational changes in the penton base proteins, exposing components that are lytic to the vesicle membrane, which in turn leads to the release of the virion into the cytosol[24]. Next, the virus binds to motor proteins and is transported along microtubules into the nuclear pore complex after which the viral DNA is injected into the nucleoplasm. Once in the nucleus, the HAdV terminal protein attaches tightly to components of the nuclear matrix[37] located in nuclear subdomains enriched in cellular factor necessary for the transcription, replication and assembly of the mature virions[24].

1.2.3 Viral genome organization and expression

The HAdV genome consists of a linear double-stranded DNA (dsDNA) molecule of about 37 kb[27, 37]. The 5' HAdV DNA termini is covalently attached to a terminal protein (TP)[38] and contains 103 nucleotide-long inverted terminal repeat sequences (ITRs) in which the viral origins of replication are embedded[38]. The HAdV genome is transcribed by the host RNA polymerase II and its expression profile can be divided in early, intermediate and late phases. The early and late transcriptional units, are transcribed before and after the onset of viral DNA replication respectively[28] and, together with the intermediate units, encode for over 50 polypeptides[21].
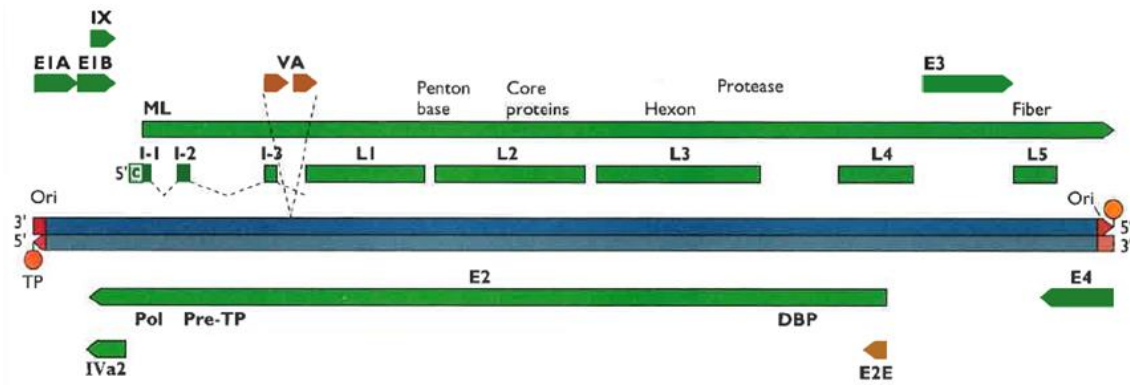
Figure 5 | The HAdV genome organization. Early, intermediate and late transcriptional units encodes for a large amount of polypeptides involved in the viral replication and assembly[32].

The early expression units are composed of six components (*E1A*, *E1B*, *E2A*, *E2B*, *E3* and *E4*) that, by differential splicing and alternative start codon usage, give rise to multiple mRNAs[24]. The first transcripts to be synthesized are those encoded by the *E1* region[39]. These transcripts encode two major regulatory components: 289R (or 13S) and 243R (or 12S)[40] that modulate cellular processes such as cell division and tumor suppressor *p53* gene expression, to make the cell more susceptible to viral replication[25]. The *E1*-derived gene products cooperate in the trans-activation of the other early transcriptional units, initiating the viral transcriptional program[30].

The *E2* region encodes components for the viral replication, like the precursor of the terminal protein (pTP), the adenoviral DNA polymerase (Pol) and the DNA-binding protein (DBP). The latter protein is involved in the stabilization of single-stranded DNA (ssDNA) forms that are formed during viral DNA replication, in the activation of late viral gene expression[25] and virion assembly[23, 41].

The gene products of *E3*, have as primary function the modulation of the host immune responses to the viral infection by inhibiting inflammatory responses[23], preventing transportation of viral antigens to the cell surface by MHC class I proteins[42] and inhibiting pro-apoptotic pathways[25]. Of note, the *E3* gene products are non-essential for viral replication in tissue culture[30, 40].

The *E4* region encodes a group of polypeptides (encoded by *E4* ORFs 1 through 7)[23] that are involved in multiple tasks such as preferential translation of viral mRNA over that of the host cell, viral DNA replication, host-cell protein synthesis shut-off and blockage of apoptotic pathways[43].

After differential splicing events, five late ORFs (L1 through L5) are transcribed from a common major late promoter[40]. These ORFs encode viral structural components essentials for viral DNA encapsidation and viral maturation in the nucleus. These proteins include the core and capsid proteins as well as a cysteine protease, important for proteolytic maturation of TP[23,47] from the precursor protein pTP. The HAdV genome also transcribes non-translated RNAs (i.e. VA RNAs I and II) that, amongst other functions, play a role in modulating cellular defense mechanisms[25].

1.2.4 Virion assembly and release

The virion assembly process, like replication, also occurs in the host cell nucleus[24], starting about at 8 hours after infection, resulting in the production of around 10^4 to 10^5 progeny particles per cell[44], released by cell lysis at 30 to 40 hours post-infection. The encapsidation of the viral genome is governed by a viral packaging signal (ψ) located at the "left" end of the viral genome and by other *cis*-acting sequences[45], giving rise to a selective incorporation of the double-stranded HAdV genome into the HAdV capsids. These capsids are assembled by a series of polypeptides and protein maturations by the HAdV DNA-encoded protease[46]. Accompanying these, there are major changes in nuclear infrastructure and permeabilization, releasing the mature virus into the cytosol and, by cell lysis, out from the infected host cell[25, 47].

1.2.5 Adenoviral-based vectors for gene therapy

Adenoviral vectors had a turbulent beginning as a tool for gene therapy, with the adverse reaction and ultimately death of a patient undergoing gene therapy for ornithine transcarbamylase deficiency (OTC)[10], caused by the interaction of the viral vector with the host immune system. Despite this, HAdVs are currently being tested in *in vivo* therapeutic settings as recombinant vaccines for infectious diseases such as HIV/AIDS[48] and Ebola[49].

HAdVs are used as gene therapy viral vectors for clinical trials, usually Ad2 or Ad5[44], exploiting some of their key advantages such as their high systemic transfer efficiency[1] (although given intravenously most of the vectors ends up in the liver, making direct injection more reliable[50]), ample transgene space[1], episomal genome form[51], high titer production[1], broad tissue tropism[10], low integration efficiency[51] and the

ability to transduce both quiescent and actively dividing cells[10, 52]. They are used as recombinant vaccines, for delivery of genes into cells (i.e. gene therapy)[25], enhance immune system ability to fight infections or cancer therapy for tumor suppression and elimination[25, 48, 53]. They are used to transduce tissues such as the liver, skeletal muscle, heart, brain, lung and tumors[54].

Despite their many advantages, adenoviral-based vectors have been shown to induce severe toxicity and immune response. For this reason, the most frequent adenoviral vector manipulations are aimed at transforming them into less immunogenic vectors for *in vivo* gene therapy applications[10].

1.2.5.1 First-generation adenoviral vectors

First-generation adenoviral vectors have the *E1* or the *E1* and *E3* regions deleted[55, 56] (**Fig. 6**). The former and latter vector types can accommodate up to 5 and 8 kb kb of foreign DNA, respectively[24, 55]. In addition, the removal of the *E1* region disrupts the normal activation of the viral gene expression program, rendering the vector replication-defective[40]. Thus, to generate these vectors, the *E1* functions have to be provided *in trans* by a complementing producer cell line such as HEK293[57], 911[58] or PER.C6[59].

First-generation adenoviral vectors are easy to construct and can be produced to high titers (up to 10^{12} to 10^{13} vector particles per milliliter)[44, 56]. Moreover, they achieve efficient transgene delivery and high-level transient transgene expression in a broad array of mammalian cell types.

Despite their positive attributes, it became apparent that the deletion of the *E1* functions alone didn't completely abolish the vector's ability to produce viral proteins due to the presence of E1-like proteins in transduced cells[24, 60]. These *de novo*-synthesized viral proteins can lead to immune responses, often cytotoxic T lymphocite (CTL)-mediated[61], that in turn are responsible for transgene expression extinction and vector-transduced cell clearance within 2-3 weeks of after vector administration[24, 25]. Moreover, some degree of homologous recombination between the HAdV sequences of the vector backbone and those inserted in some of the producer cells, such as HEK293, can regenerate *E1* and produce contaminating replication-competent HAdVs (RCAs)[62]. The production of newer producer cell lines (like PER.C6) with *E1* sequences trimmed in such a way as to avoid homologous recombination-mediated generation of RCAs was required for the production of clinical-grade vector batches[59].

Therefore, further manipulations (i.e. deletions) were deemed necessary to further cripple the adenoviral-based vectors.

1.2.5.2 Second-generation adenoviral vectors

Second-generation adenoviral vectors combine the *E1/E3* deletion(s) with deletions in other early regions such as *E2* (*E2A* or *E2B*) or *E4*[63, 64]. As a result, second-generation adenoviral vectors are propagated in complementing cell lines that express, in addition to *E1*, the other viral early region that is absent from the vector backbone. Importantly, these deletions not only reduce the likelihood for the generation of contaminating RCAs[59, 65] but also have shown reduced cytotoxicity in target tissues. Yet, even second-generation adenoviral vectors were shown to lead exclusively to transient transgene expression after their *in vivo* administration into immunocompetent animal models[10, 24, 40].

1.2.5.3 High-capacity adenoviral vectors

In an attempt to overcome the immunological issues of first- and second-generation adenoviral vectors, one strategy pursued was the elimination of all coding sequences from the vector, giving rise to so-called "gutless", helper-dependent or high-capacity adenoviral vectors (HC-AdVs).

This new vector type only retains the *cis*-acting elements involved in viral DNA replication and encapsidation (i.e. the HAdV ITR and packaging sequence, respectively) requiring, as a consequence, a so-called helper vector to provide *in trans* all the missing adenoviral gene products[66-68]. HC-AdVs were the first viral vector system to transfer the full-length 11-kb dystrophin-encoding DNA sequence into a mouse model of the muscle-wasting disorder Duchenne muscular dystrophy[69]. In addition to their significantly larger transgene capacity (~37 kb)[68, 70], this new type of adenoviral vector is superior to their earlier generation counterparts in what safety (reduced immune response)[66] and long-term transgene expression profiles are concerned (up to 1-2 years in the liver of mice and baboons)[68, 70]. However, in part due to the aforementioned requirement for a helper adenoviral vector, the production of HC-AdVs is more complex and more difficult to standardize than that of their earlier generation counterparts.
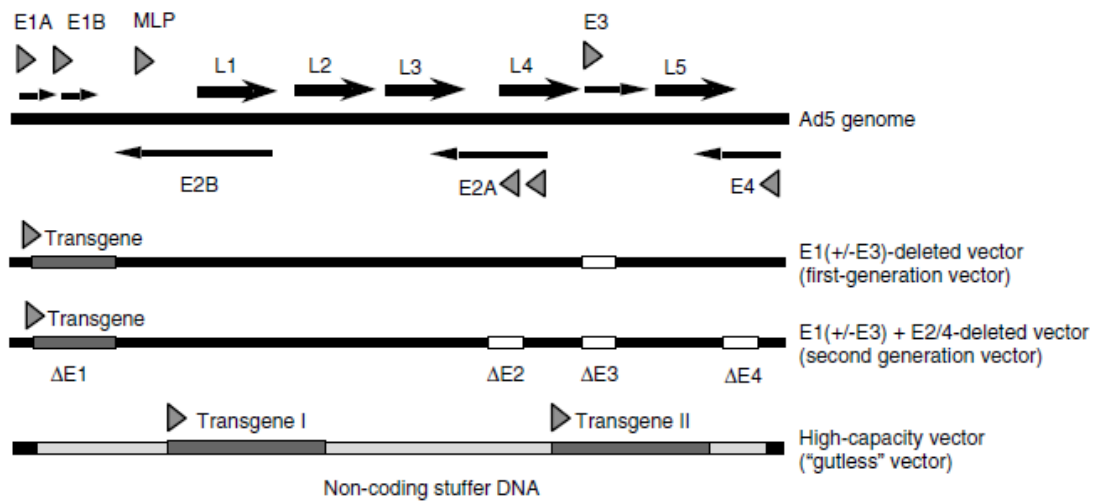
Figure 6 | Diagrammatic representation of the different types of adenoviral vectors and corresponding deletions. The deletion of non-essential viral genes can increase transgene size limits, overcome toxic effects and render the viral vector safe in clinical trials[44].

1.2.5.4 Adenoviral vector modifications

Fiber modification (pseudotyping): Many therapeutically relevant cell types like those of the endothelium, the hematopoietic system as well as many tumor cells, lack the CAR receptor which is the receptor utilized by vectors based on the prototypic adenoviral serotypes 2 and 5. The receptor recognition of the virus can be altered by switching the head domains of the fiber from other subgroup[71] [72, 73].

Long-term expression in proliferating cells: as aforementioned, the HAdV genome remains in an episomal form in the nucleus of the host cell, therefore, the DNA of conventional adenoviral vectors in transduced proliferating cells, such as hematopoietic stem and progenitor cells is lost over time due to target cell division. For this reason, many efforts are underway to try to endow adenoviral vectors with elements conferring episomal replication or genome integration. Strategies aiming at turning adenoviral vectors into gene delivery systems for long-term transgene expression in dividing cell populations include the development of adenovirus/retrovirus chimeric vectors[74], adenovirus/AAV chimeric vectors[75] and adenoviral vectors that exploit transposase systems or the error-free homologous recombination DNA repair pathway of the cell[76].

1.3 Homologous recombination

Double-strand breaks (DSBs) are constantly occurring in cells, either due to the normal DNA replication process or due to the exposure of cells to genotoxic agents[77]. Cells have evolved DNA repair systems of which the error-prone NHEJ[78] and the error-free HR[79] pathways are the two major ones. The DSB DNA repair pathway chosen is regulated by several cellular factors that, to a great extent, depend on the phase of the cell cycle phase, with the HR occurring specially in late S and G2 phases[80].

HR is as a process that, by interaction between two similar or identical DNA sequences, exchange of genetic information takes place with the recombinant products acquiring genetic information from the parental DNA molecules[87]. Gene targeting[81, 82, 89] exploits the endogenous HR pathway of the cell to bring about precise genomic modifications at a target sequence of choice. To this end, an exogenous DNA molecule (donor DNA) with the desired genetic alteration needs to be introduced into the target cells[90]. This HR-based gene targeting principle is currently being investigated for gene therapy and can also be used in 'knock in' and 'knock out' strategies[82]. Although HR is an highly accurate process under normal conditions its efficiency is very low most mammalian cell types, occurring at frequencies ranging from 10^{-7} to 10^{-8} [83].

The HR machinery is recruited at the sites of DSBs. The exogenous induction of these lesions at specific/targeted loci by designed restrictions endonucleases[84] like meganucleases such as I-SceI[85], ZFNs[86, 87], that increase HR frequency by 3-4 orders of magnitude or, more recently, TALENs that greatly enhances HR and, subsequently gene targeting by 10 to 10,000-fold. Also nicks or single-strand breaks can somewhat induce the same pathway, but to a less extent[88, 89]. Co-delivery of a donor transgene bearing locus-specific homology arms[90] or a single-stranded DNA oligonucleotide[91] is required in HR-based gene targeting approaches (**Fig. 7**).
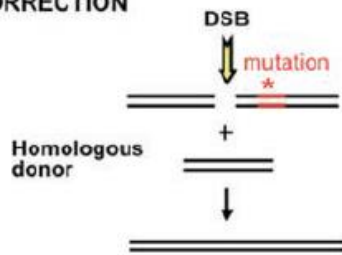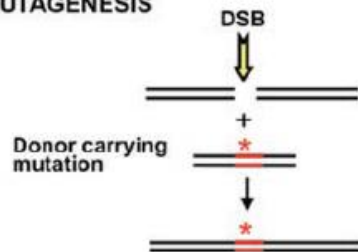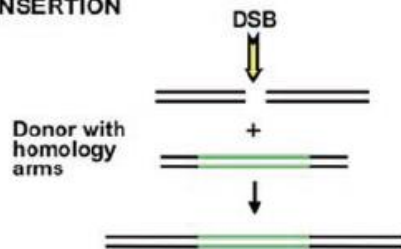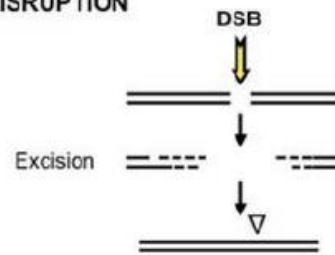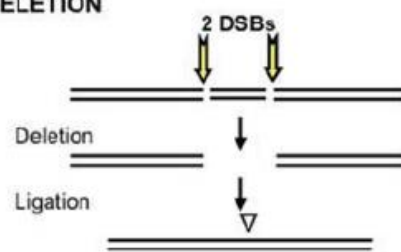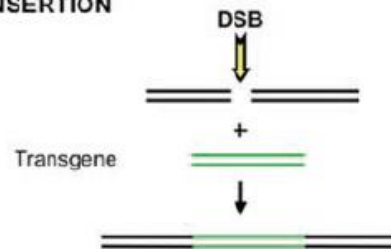
Figure 7 | Targeted genome engineering. Through the cells HR and NHEJ repair systems, genome alterations can be made using designed restriction endonucleases to produce targeted DSBs. A donor transgene is also required for the desired sequence modification[15].

1.4 Transcription activator-like effector nucleases

Transcription activator-like effectors (TALEs) are bacterial proteins found in the phytopathogenic bacteria of the genus *Xanthomonas.* They are injected and delivered to plant cells via a type III secretion system and imported to the nucleus[92, 93]. Their main established function is to target specific gene promoters and to activate their transcription, thus, functioning as transcriptional activators, they regulate host plant target genes to increase pathogen colonization and dissemination in the plant host[94, 95, 92].

Their structure comprises typically a N-terminal translocation domain, a central tandem direct repeats region (between 13 and 28 repeats), referred to as the DNA binding domain of the protein and a nuclear localization sequence preceding a C-terminal transcriptional activation domain[13, 92, 94] (**Fig. 8**).
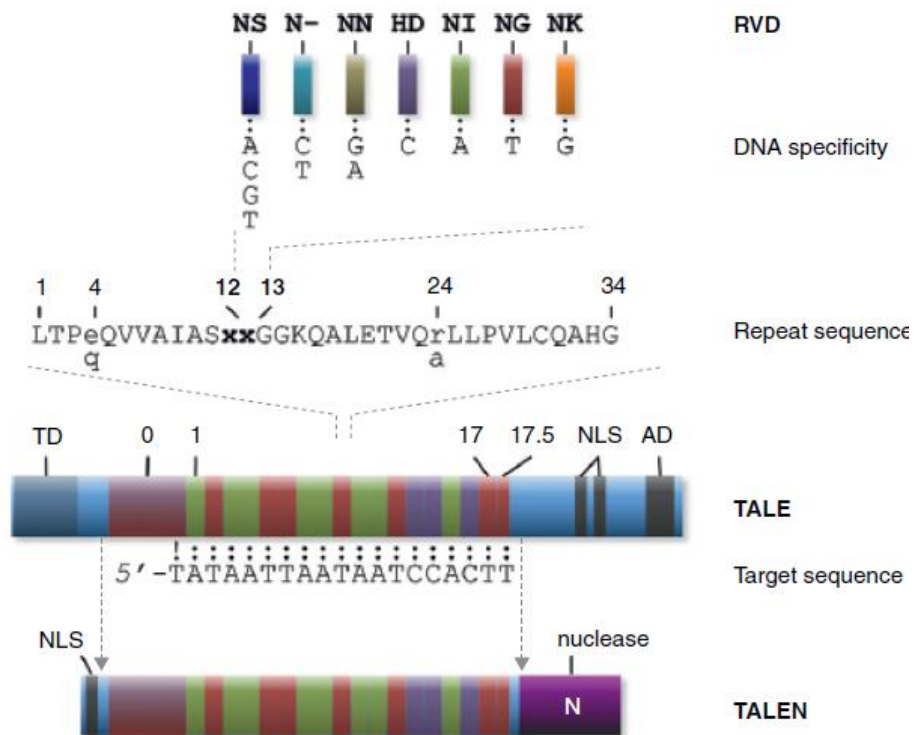
Figure 8 | Prototypical structure of an artificial transcription activator-like effector nuclease (TALEN). The modularity and sequence specificity nature of the TALE protein can be used to engineering designed nucleases for gene therapy purposes[96].

The DNA binding domain of the protein is composed of a tandem array of repeat units (of about 34 highly conserved amino acids each[94]), with the last repeat unit referred to as a 'half-repeat' as it only contains 20 amino acids. The modules commonly target sequences initiated by a thymidine (position 0) which is required for the effector activity[94]. The amino acid sequence is mostly invariant with the exception of two hypervariable amino acids at positions 12 and 13 of the repeat. These are called "repeat variable di-residue" (RVD) and, together with the other subunits, are responsible for the target DNA sequence binding activity[97], as different RVDs (HD, NG, NI and NN) associate preferentially with different nucleotides (C, T, A and G, respectively), with some being less selective than others[98]. This one-to-one (protein-to-DNA) correspondence for each RVD and target nucleotide, combined with the programmable modular nature of TALEs[98], makes possible to design TALEs to target specific sequences in the genome, using systems such as the 'Golden Gate' cloning strategy.

For the goal of producing DSBs in the genome, artificial TALE enzymes were generated by fusing the TALE scaffold to the non-specific DNA cleavage domain of the FokI restriction enzyme. This strategy results in the assembly of TALE nucleases (TALENs)[96]. These operate in pairs and are designed to bind opposing DNA target sites, separated by a spacer sequence. Upon binding of each TALEN monomer to its target sequence, the FokI domains dimerize[99] becoming catalytically active to induce a targeted DSB[100] (**Fig. 9**). TALENs have been demonstrated to be powerful genetic editing tools in a variety of experimental systems such as fruit fly[101], zebrafish[102], rat[103], cultured human cells[104] and induced pluripotent stem (iPS) cells[105].

For gene therapy purposes, TALENs can be used for several types of genome modifications (**Fig. 10**) such as targeted gene mutagenesis (disruption of gene function by NHEJ), targeted gene insertion (HR-dependent integration of exogenous DNA into host cells at a specific target location for gene addition or gene repair) and targeted chromosomal rearrangements like deletions, inversions and translocations[93].

TALENs can be delivered in the cell through nucleic acid microinjection or DNA transfection methods based on chemical agents or electroporation[105-107] but these methods suffer from low efficiency especially in primary cells. Adenoviral vectors, on the contrary, allow for highly efficient gene delivery of functional TALEN genes into human cells[108].
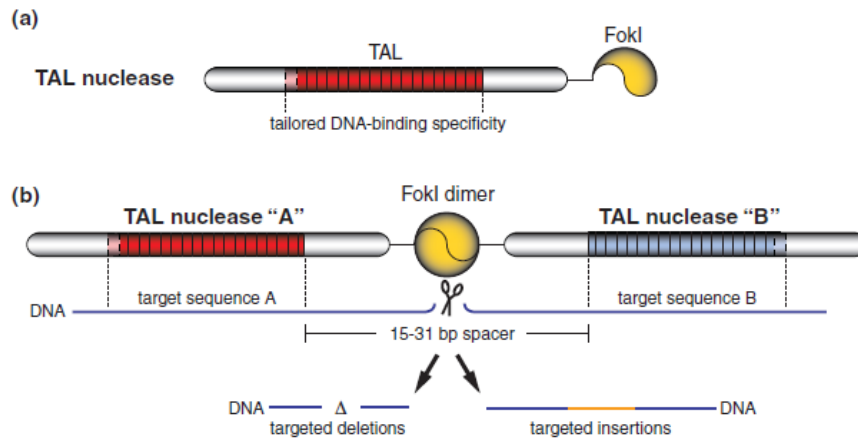
Figure 9 | TALE-based nucleases (TALENs). These tailored nucleases have a (**a**) TALE protein domain fused to a FokI nuclease domain that (**b**) dimerizes and produce a DSB in the DNA sequence used in gene therapy approaches such as targeted deletions and/or insertions[93].
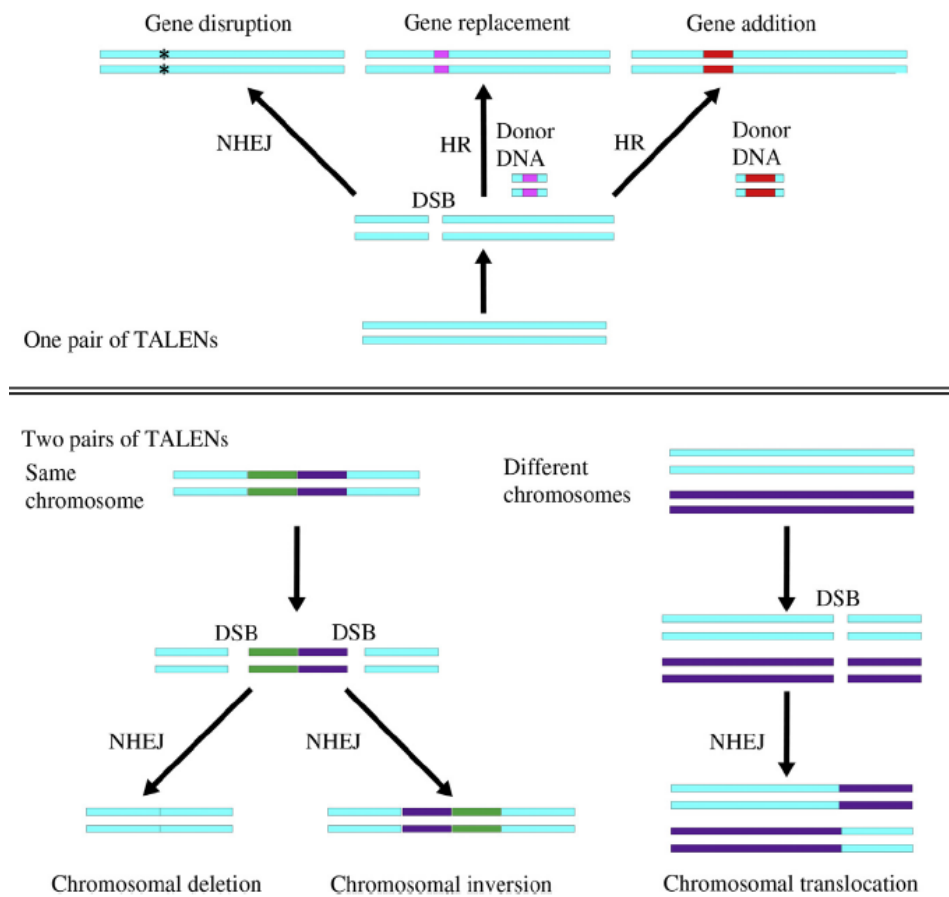
Figure 10 | Genome modifications with TALENs. By TALEN-mediated DSB production on a target sequence, various gene modifications types can be achieved[109].

1.5 Objective of the project

The objective of this project is to construct and test donor DNA (*eGFP* reporter gene with two sequences of homology to the human *AAVS1* locus sequence) with different topologies by means of terminal structures and determine their impact on exogenous DNA integration efficiency and accuracy following targeted DSB formation by *AAVS1*-specific TALENs in human HeLa cells and DMD myoblasts.

The different donor DNA molecules will be delivered by second-generation fiber 50-modified adenoviral vectors and the efficiency of transduction and transgene integration into the human genome will be determined.

Chapter 2. Materials and methods

2.1 Construction of pAd.shuttle donor plasmids

The three adenoviral shuttle donor plasmids pAd.shuttle.AAVS1.PGK.eGFP.pA.T-TS, pAd.shuttle.AAVS1.PGK.eGFP.pA.FRT and pAd.shuttle.AAVS1.PGK.eGFP.pA.ITR were made on the basis of the plasmid pAd.shuttle.AAVS1.PGK.eGFP.pA by using recombinant DNA techniques. This plasmid, in addition to a kanamycin antibiotic marker gene, has a phosphoglycerate kinase-1 (*PGK*) promotor directing the expression of an enhanced green fluorescent protein (*eGFP*) reporter gene. This expression unit is flanked by DNA sequences identical to those of the human *AAVS1* target locus located at 19q13-qter. It also contains DNA sequences derived from the human adenovirus type 5, including the left and right ITR (L-ITR and R-ITR) and encapsidation signal (ψ). The adenoviral shuttle donor plasmids were derived from pAd.shuttle.AAVS1.PGK.eGFP.pA by inserting oligonucleotides representing the yeast site-specific Flipase recombinase target sequence (FRT), the *AAVS1*-specific TALEN target site (T-TS) or the adeno-associated virus inverted terminal repeat (AAV-ITR) sequence (**Fig. 11**). Each of these inserts were endowed with the following discriminating restriction enzyme recognition sites: XbaI (Fermentas) for 5' and 3' FRT, BspLI (Fermentas) for 5' and 3'-T-TS and Eam1105I (Fermentas) for AAV-ITR.

The sense and the anti-sense oligonucleotides representing the T-TS and the FRT sites (Eurofins MWG Operon) were annealed to each other and were ligated into the vector plasmid pAd.shuttle.AAVS1.PGK.eGFP.pA. The insertion of the annealed oligonucleotides at the 5' end of the donor DNA (i.e. upstream of the *AAVS1* homologous "arm 1") was done by digesting the vector plasmid with Acc651 (Fermentas), whereas at the 3' end XhoI (Fermentas) was used. The AAV-ITR was isolated from plasmid pDD2[110] after its digestion with PvuII (Invitrogen) and BspLI (Fermentas). Next, the resulting cohesive termini were blunt-ended by using the Klenow fragment of *E. coli* DNA polymerase I (Klenow, Fermentas) and were ligated into the vector plasmid, also blunt-ended (Klenow, Fermentas), either at the Acc651 or the XhoI sites.
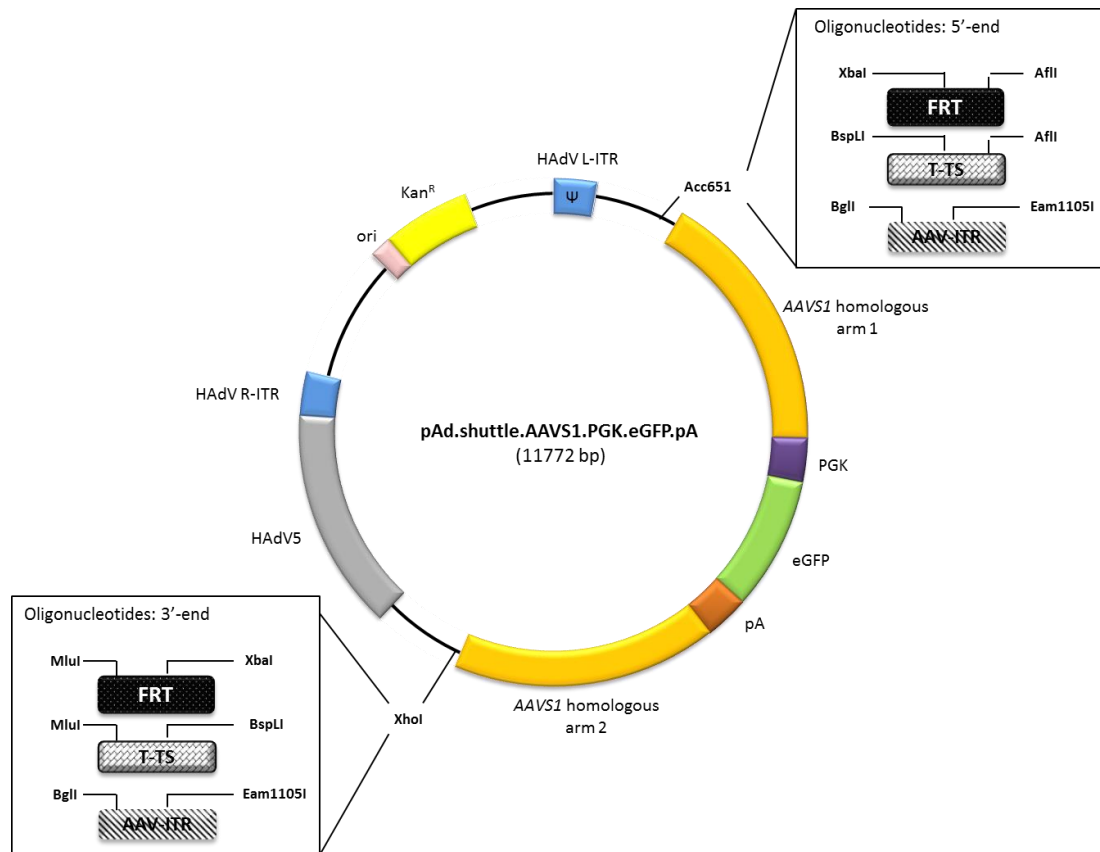
Figure 11 | Diagramatic representation of the adenoviral shuttle plasmids containing the AAV-ITR, the FRT sites or the T-TS recognition sequences. The plasmid contains a kanamycin antibiotic marker gene (Kan[R]), a PGK promoter, an *eGFP* reporter gene, DNA sequences identical to those of the human *AAVS1* target locus, DNA sequences derived from the human adenovirus type 5, L-ITR, R-ITR and encapsidation signal (ψ).

After the ligated DNA fragments, the plasmids were used to transform chemically-competent GT115 *E. coli* cells (InvivoGen). Next, the transformed cells were incubated for an overnight incubation period at 37°C, after which the plasmid DNA was extracted by using a standard small-scale DNA purification procedure ("miniprep"). In brief, the cell suspensions were transferred to 1.5 mL tubes, centrifuged briefly at 14.000 rpm for 15 seconds and the supernatant was removed apart from ~100 μL which was used to resuspend the cells. A lysis mix (lysis buffer and lysozyme) was added and the samples incubated at 100°C for 1 min. Through centrifugations and washing steps, the DNA was re-suspendend in TE with RNAseA. The recovered DNA was subjected to restriction fragment length analysis to access the correct incorporation and orientation of the different inserts. Positive *E. coli* clones were propagated and up-scaled for preparative DNA extraction by using the large-scale DNA purification system JETSTAR kit (Genomed) according to the instructions provided by the manufacturer.

2.2 Generation of full-length adenoviral vector molecular clones

The assembly of the full-length second-generation adenoviral vector molecular clones pAd.Easy.AdV.dE2A.E3+.F$^{50.}$AAVS1.PGK.eGFP.pA.FRT2+ (abbreviated to pAdEasy.AAVS1.eGFP.FRT),
pAd.Easy.AdV.dE2A.E3+.F$^{50.}$AAVS1.PGK.eGFP.pA.TALEN-TS2+
(pAdEasy.AAVS1.eGFP.T-TS) and
pAd.Easy.AdV.dE2A.E3+.F$^{50.}$AAVS1.PGK.eGFP.pA.AAV-ITR (pAdEasy.AAVS1.eGFP. ITR) was performed by deploying the 1, 2 and 3 adenoviral shuttle donor plasmids, respectively linearized with MssI (Fermentas). The MssI-treated shuttle donor plasmids were transformed into electro-competent BJ5183$^{pAdEasy-2.50}$ *E. coli* cells of the AdEasy system[111] harboring the remaining of the adenoviral vector backbone. The full-length adenoviral vector molecular clones were assembled by HR between the adenoviral shuttle and backbone constructs (**Fig. 12**). The DNA was extracted by using the above-described "miniprep" protocol and the correct clones were identified by restriction fragment length analysis. Next, the plasmids were transformed into DH5α *E. coli* cells (Invitrogen) for large-scale DNA preparation by using the DNA purification system JETSTAR kit (Genomed) according to the instructions provided by the manufacturer.
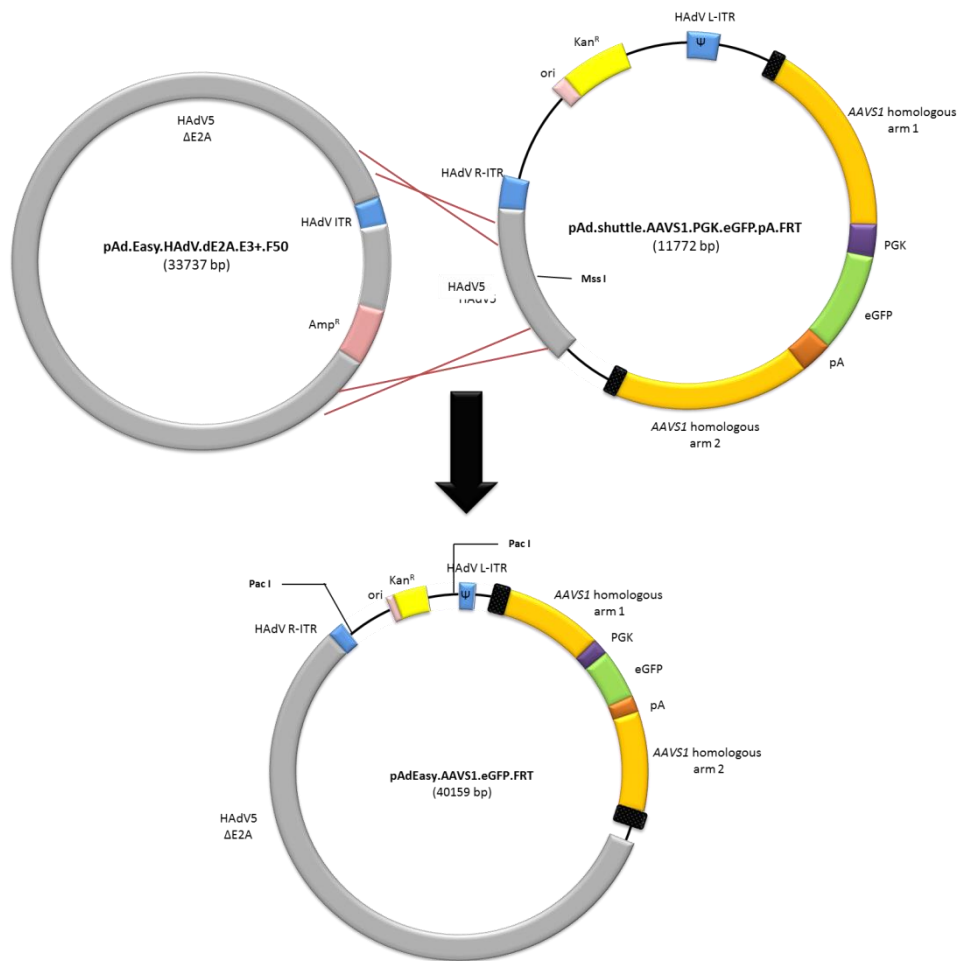
Figure 12 | Schematic representation of the assembly of full-length AdV molecular clones by HR in *E. coli* cells. The MssI-treated pAd.shuttle plasmids are individually transformed into *E. coli* BJ5183 cells containing the pAd.Easy-derived adenoviral vector plasmid backbone for the generation of fiber-modified second-generation AdVs (BJ5183[pAd-EasyΔE1.ΔE2A]). Exchange of genetic information through HR between the shuttle and backbone plasmids at the indicated regions (red crosses) yields AdV molecular clones containing full-length AdV genomes. Bacterial clones harboring these recombinant viral genomes can be selected for by growth on agar plates supplemented with the antibiotic kanamycin.

2.3 Rescue and large-scale production of adenoviral vectors encoding donor DNA templates

The adenoviral donor vectors were produced in PER.E2A cells[112] by linearizing each adenoviral vector molecular clone with PacI (Biolabs) and transfecting the resulting linear DNA into the *E1-* and *E2A*-complementing packaging cells seeded in 6-

well culture plates (Greiner) at a density of 1.5×10^6 cells per well. The DNA transfections were carried out by using a polyethyleneimine-based protocol[111].

The PER.E2A packaging cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) containing 10% fetal bovine serum (FBS; Invitrogen), 10 mM $MgCl_2$ (Sigma-Aldrich), and 250 µg/mL Geneticin (Invitrogen). During vector rescue and propagation, medium without Geneticin was used instead. They were kept in an atmosphere of humidified air with 10% CO_2 and were incubated at $39^{\circ}C$ during regular maintenance and at $34^{\circ}C$ during adenoviral vector production (note: these cells express mutant *E2A* gene encoding a thermo-sensitive version of the DBP of the HAdV5 virus, thus the need to incubate them at $34^{\circ}C$ during vector production).

Up-scaling of the vector productions to fifteen 175-cm^2 cell culture flasks (Greiner) was performed as follows. The transfected PER.E2A were harvested after the emergence of complete cytopathic effect (CPE). The CPE was visualized by rounding and detachment of cells from the cell culture surface at 13, 10 and 15 days post-transfection for pAdEasy.AAVS1.eGFP.FRT, pAdEasy.AAVS1.eGFP.ITR and pAdEasy.AAVS1.eGFP.T-TS respectively. Next, the rescued vector particles were released by subjecting the producer cells to three cycles of freezing (N_2) and thawing (in $37^{\circ}C$ water bath). The cellular debris was removed by centrifugation and the supernatants were filtered through 0.45-µm pore-sized filters. The vector particles present in the resulting clarified supernatants were serially propagated in increasing amount of producer cells with the last propagation round (passage 4) corresponding to fifteen 175-cm^2 cell culture flasks. After the emergence of CPE in these cultures, the cells were harvested and were taken for downstream processing of the vector preparations.

2.4 Purification of donor DNA-containing adenoviral vectors

The adenoviral vectors were purified following the DSP virus purification protocol[111]: the cells were collected by centrifugation and suspended in DMEM. They were lysed (incubation at $37^{\circ}C$, 30 minutes with 0.5% sodium deoxycholic acid). Addition of 40 mM $MgCl_2$ and 12.5 mg/ml DNaseI (grade II; Roche Applied Science) to digest cellular DNA. The mix was spun for 10 minutes at 800 g and then a second time for 20 minutes at 2000 g to remove cellular debris. The recovered supernatants were

loaded onto block gradients consisting of 1.4 and 1.24 g/ml CsCl in 30mM Tris–HCl (pH 7.5) and 5% (v/v) and a ultracentrifugation at 79 521 g for 2 hours at 10°C in an SW28 rotor (Beckman Coulter, Fullerton, CA, USA) was performed. After centrifugation, the vector particles concentrated in the interface of the two CsCl solutions (**Fig. 13**) were recovered and transferred to a 13.5 ml polyallomer Quick-Seal tube (Beckman Coulter) with 1.33 g/ml CsCl in 30 mM Tris–HCl (pH 7.5) and 5% (v/v) glycerol. They were spun overnight at 287 641 g in a VTi 65.1 rotor (Beckman Coulter) at 10°C. After, the vector particles were collected and desalted using an Amicon Ultra-4 Centrifugal Filter Devices (Millipore, Billerica, MA, USA) and storage buffer consisting of 20mM Tris–HCl (pH 8.0), 25mM NaCl and 5% (v/v) glycerol. Each viral vector was distributed in aliquots that were stored at -80°C.
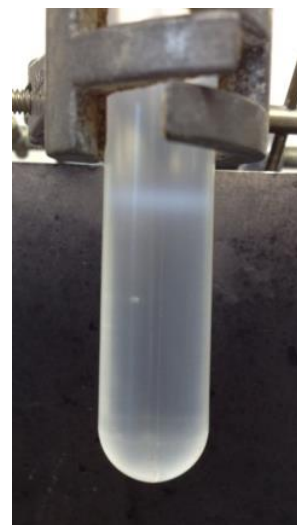
Figure 13 | Band of adenoviral vector particles after DSP virus purification.

The purified AdV donor vectors will be termed as AdV.Δ2.donor-T-TS.F50, AdV.Δ2.donor-FRT.F50 and AdV.Δ2.donor-ITR.F50, corresponding to each different inserted sequence.

2.5 Titration of donor DNA-containing adenoviral vector stocks in HeLa cells

Titers of adenoviral vector donor stocks were accessed through 5-fold limiting dilutions in HeLa indicator cells (American Type Culture Collection [ATCC]), seeded at a density of 0.8×10^5 cells per well of 24-well plates (Greiner). At 3 days post-transduction, frequencies of eGFP-positive cells were determined through flow cytometry by using a BD LSR II Flow cytometry machine (Becton Dickinson) and the titers were expressed in terms of HeLa cell-transducing units (HTUs) per milliliter.

The HeLa cells were cultured in DMEM (Invitrogen) supplemented with 5% FBS (Invitrogen). The cells were kept at 37°C in a humidified air atmosphere of 10% CO_2.

2.6 Transduction of HeLa cells with donor adenoviral vectors subjected to TALEN-mediated DSB formation

For this experiment, HeLa cells were transduced with the three different donors produced previously. They were also transduced with TALENs left and right, designed specifically for the *AAVS1* locus (TALEN[S1] L+R) for the induction of a DSB in that locus. The TALEN contains a DsRed reporter cassette and was delivered by a second-generation human adenoviral vector (AdV.Δ2.TALEN[S1] L+R).

The HeLa cells, in a density of 0.8×10^5 cells per well, were transduced, in triplo, with the AdV donors with or without the designer TALEN[S1] L+R and in combination with/without the Flipase (FLPe) in the case of AdV.Δ2.donor-FRT.F50. These sets of experiments were conducted in order to determine the frequencies of targeted transgene integration/expression in parallel with the proper controls.

HeLa cells were transduced with only the donor adenoviral vectors:

- Mock (control)
- AdV.Δ2.donor-FRT.F50
- AdV.Δ2.donor-FRT.F50 + FLPe
- AdV.Δ2.donor-T-TS.F50
- AdV.Δ2.donor-ITR.F50

or in combination with the TALEN[S1]:

- Mock (control)
- AdV.Δ2.TALEN[S1] L+R (control)
- AdV.Δ2.donor-FRT.F50 + AdV.Δ2.TALEN[S1] L+R
- AdV.Δ2.donor-FRT.F50 + FLPe + AdV.Δ2.TALEN[S1] L+R
- AdV.Δ2.donor-T-TS.F50 + AdV.Δ2.TALEN[S1] L+R
- AdV.Δ2.donor-ITR.F50 + AdV.Δ2.TALEN[S1] L+R

The multiplicity of infection (MOI), or the ratio of virus that infects target cells, chosen were 3 for the adenoviral donor vectors, 1.5 for the AdV.Δ2.TALEN[S1] L+R and 2 for the FLPe.

HeLa cells were kept in DMEM medium (Invitrogen) supplemented with 5% FBS (Invitrogen), at 37°C in an atmosphere of 10% CO_2 humidified air.

The final donor topologies will be achieved by 1) FLPe-mediated circularization of the AdV.Δ2.donor-FRT.F50 (circular donor), 2) TALEN-induced DSB of the

AdV.Δ2.donor-T-TS.F50 (linear free donor) and 3) secondary structure-ends with DNA binding repair factors of AdV.Δ2.donor-ITR.F50 (**Fig. 14**).
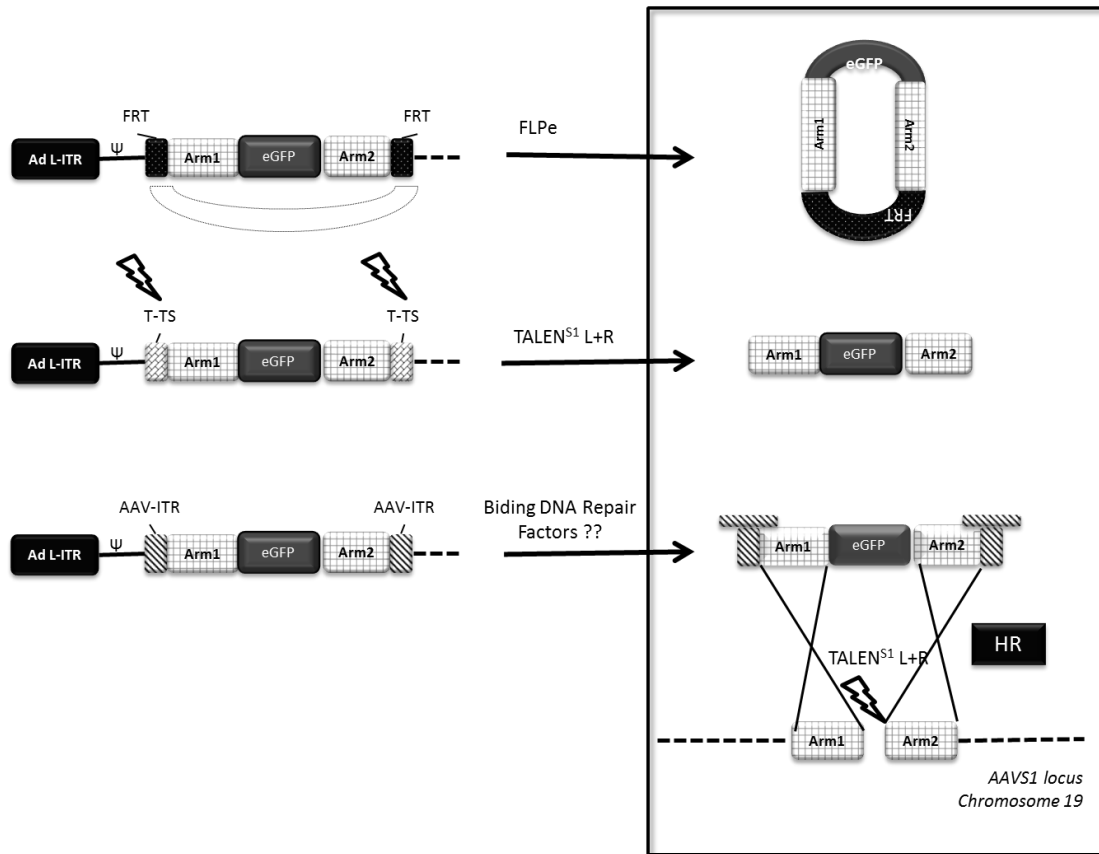
Figure 14 | Production of the final donor topologies and of the theoretical design of incorporation in the *AAVS1* locus. The AdV.Δ2.donor-FRT.F50, by FLPe-mediated recombination of the FRT sequences, will produce a circular form of donor. The AdV.Δ2.donor-T-TS.F50, by co-deliver with the AdV.Δ2.TALEN[S1] L+R, will produce a linear form, free from the rest of the vector backbone by TALEN-mediated DSB at the T-TS sequences. The AdV.Δ2.donor-ITR.F50, by means of the AAV-ITR sequences, will recruit unknown DNA repair factors that can potentially intervene with the HR gene insertion. Ψ –packaging sequence; thunder-bolts: TALEN-induced DSB.

2.7 HR and reporter gene integration/expression on HeLa cells

The transduced HeLa cells were maintained and kept in culture with DMEM (Invitrogen), 5% FBS (Invitrogen), 37°C and 10% CO_2 for nearly 4 weeks, time period

that is sufficient enough for viral clearance and episomal dilution. Measure of transgene expression parameters (i.e. frequencies of reporter-positive and reporter-negative target cells and mean fluorescence intensities) 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). Mock-transduced target cells were used to set background fluorescence levels. Typically, 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).

2.8 Molecular characterization of the viral donors and stably transduced cells

2.8.1 Isolation of the DNA from the purified AdV donors

A 50-µL aliquot of each purified AdV vector was used to extract the adenoviral donor genome through degradation of the capsid and the vector DNA molecules were extracted by using the JETSORB Gel Extraction Kit (GENOMED) following the instructions provided by the manufacturer.

2.8.2 Restriction enzyme analysis of AdV donor genomes

The adenoviral vector genomes of AdV.Δ2.donor-FRT.F50, AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.donor-ITR.F50 were digested with XbaI, MluI and BglI (all from Fermentas), respectively. Next, the restriction DNA patterns were analyzed to determine the integrity of the adenoviral vector genomes as a whole and the portions harboring the donor DNA templates in particular.

2.8.3 Isolation of the extrachromosomal DNA from the stably transduced cells

Extrachromosomal DNA was extracted from the stably transduced HeLa cells following a protocol described in reference ([113]) and a phenol:chloroform:isoamylalchohol (25:24:1) DNA extraction protocol. The purified DNA was used for polymerase chain reaction (PCR) experiments.

2.8.4 Analysis of *in vivo* circularization and excision of AdV donor DNA

The extracted extrachromosomal DNA was used as a template for PCR reactions using 2-µl samples. The PCR mixtures consisted of 0.4 µM of primer #001 and #002 (Eurofins MWG Operon), 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 95ºC, followed by 30 cycles of 30 sec at 95ºC, 30 sec at 62ºC and 45 sec at 72ºC. The reactions were terminated by a final extension period of 5 min at 72ºC. The detection of the resulting PCR products was performed by conventional agarose gel electrophoresis. The relevant controls were carried out in parallel.

Table 2 | Primers used in the PCR experiment for the analysis of *in vivo* circularization and excision of AdV donor DNA.

Primer	Sequence (5'-3')	Size (bp)	Amplicon (bp)
#001	-CGGCGTTGGTGGAGTCC-	17	619
#002	-GCACTGAAACCCTCAGTCCTAGG-	23	

2.8.5 Clonal sorting and amplification

Following several weeks of culture the stably transduced HeLa cells were sorted by flow cytometry-assisted cell sorting and the eGFP-positive cells were collected in 1:1 mixtures of regular medium containing 2× penicillin/streptomycin

(Invitrogen) and FBS (Invitrogen) in 5x 96-well plates (Greiner Bio-One), at a density of 0.3 cells per well and maintained in DMEM medium (Invitrogen) supplemented with 5% FBS (Invitrogen), 50 µM of α-thioglycerol (Sigma-Aldrich) and 20 nM of bathocuprione disulphonate (Sigma-Aldrich), at $37^{o}C$ in an atmosphere of 10% CO_2 humidified air. Over 250 individual clones were randomly selected for expansion and molecular analyses.

2.8.6 Isolation of the DNA from the stably transduced cells

Chromosomal DNA from 118 eGFP-positive HeLa clones was extracted using a phenol:chloroform:isoamylalchohol (25:24:1) DNA extraction protocol. The purified DNA was used for PCR reactions to access the targeting efficiency.

2.9 Transgene integration efficiency analysis

PCR was performed on the extracted DNA as follows: the PCR mixtures consisted of 0.4 µM of primer #331, #332, #551, #552, #111 and #112 (Eurofins MWG Operon), 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 95ºC, followed by 40 cycles of 30 sec at 95ºC, 30 sec at 61ºC, $63.5^{o}C$ and $62^{o}C$ for *AAVS1* 3'-, 5'-junction and *eGFP*, respectively, and 2 min at 72ºC. The reactions were terminated by a final extension period of 5 min at 72ºC. The detection of the resulting PCR products was performed by conventional agarose gel electrophoresis. The relevant controls were carried out in parallel.

Table 3 | Primers used in the PCR experiment for transgene integration efficiency analysis.

Primer	Sequence (5'-3')	Size (bp)	Amplicon (bp)
#331	-CGACAACCACTACCTGAGCA-	20	1709
#332	-GACCTGCCTGGAGAAGGAT-	19	
#551	-GCACCGTCCGCTTCGAG-	17	1665
#552	-AACCCCAACCCCGTGGAAG-	19	
#111	-ACCCCGACCACATGAAGCAGC-	21	417
#112	-CGTTGGGGTCTTTGCTCAGGG-	21	
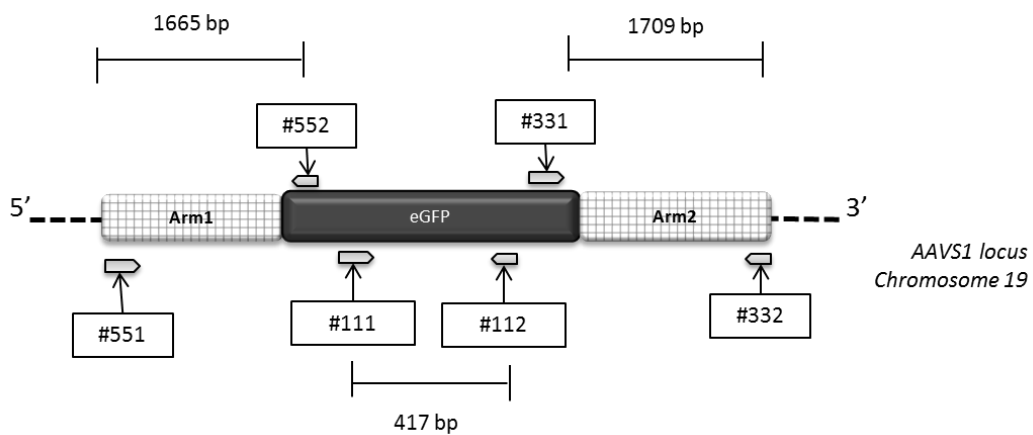
Figure 15 | Primers for detecting genomic transgene integration on stably transduced cells. The designed primers will amplify both *AAVS1* junctions, as well as *eGFP* reporter as internal control. The sizes (in bp) of the resulting amplicons are presented. Grey arrows, primers.

2.10 Transduction of HeLa cells with AdV.Δ2.donor-T-TS.F50 donor adenoviral vector subjected to TALEN-mediated DSB formation

A second transduction experiment was performed in HeLa cells, seeded at a density of 0.8 x 10^5 cells per well, in triplo in a 24-well plate (Greiner Bio-One), using only the AdV.Δ2.donor-T-TS.F50 with/without the AdV.Δ2.TALEN[S1] L+R in MOIs of 3 and 1.5 respectively.

HeLa cells were maintained in conditions previously described. After 4 weeks measure of transgene expression parameters was performed by flow cytometry as described previously.

2.11 Transduction of DMD myoblast with AdV.Δ2.donor-T-TS.F50 donor adenoviral vector subjected to TALEN-mediated DSB formation

The transduction experiment was performed in immortalized human myoblasts obtained from a Duchenne muscular dystrophy patient[114] (DMD myoblasts), seeded at a density of 2.0 x 10^5 cells per well, in triplo in a 24-well plate (Greiner Bio-One), using a AdV.Δ2.donor-eGFP.F50 (normal donor) and AdV.Δ2.donor-T-TS.F50 with/without the AdV.Δ2.TALEN[S1] L+R with MOIs of 10, 10 and 2.5 respectively.

The myoblasts were maintained in conditions specified in reference ([114]). After nearly 5 weeks measure of transgene expression parameters was performed by flow cytometry as described previously.

Chapter 3. Results

3.1 Generation and characterization of adenoviral vector shuttle plasmids encoding donor DNA templates targeting the human *AAVS1* locus

For the successful adenoviral vector production, a precise genome assembly is required, starting with the design and cloning the different inserts, composed of a PGK.eGFP.SV40.pA cassette flanked either by the T-TS, the FRT or the AAV-ITR sequences in the correct position and orientation, in a pAd.shuttle plasmid. The confirmation of a successful cloning was attained based on the restriction enzyme pattern analysis of randomly selected *E. coli* clones.
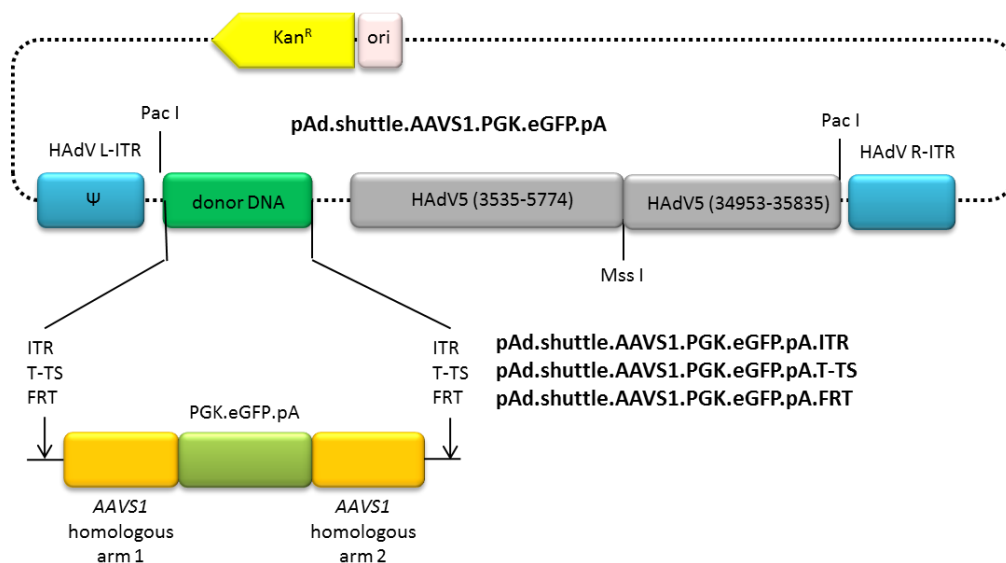
Figure 16 | Schematic representation of the three different pAd.shuttle donor plasmids designed and constructed. HAdV L-ITR, adenoviral left ITR containing the origins of replication; Ψ, adenoviral encapsidation signal; HAdV R-ITR, adenoviral right ITR; HAdV5, adenoviral genomic sequences necessary for HR-dependent assembly of full-length AdV genomes in *E. coli* strain BJ5183$^{pAd-EasyΔE1.ΔE2A}$; ITR, AAV ITRs; T-TS, *AAVS1*-specific TALEN target sites; FRT, site-specific FLPe recombinase target sites; *AAVS1*, DNA sequences identical to those framing the *AAVS1*-specific TALEN target site at the *AAVS1* locus on the human chromosome 19; KanR, Kanamicin resistence gene; ori, origin of replication.

3.1.1 Restriction enzyme analysis of pAd.shuttle donor plasmids encoding donor DNA templates

The cloning of the various inserts at an "upstream" and "downstream" position of the *AAVS1*-targeting donor DNA matrix gave rise to the pAd.shuttle donor plasmids pAd.shuttle.AAVS1.PGK.eGFP.pA.ITR, pAd.shuttle.AAVS1.PGK.eGFP.pA.T-TS and pAd.shuttle.AAVS1.PGK.eGFP.pA.FRT. The enzymes chosen for the restriction fragment length analysis of the different constructs gave rise to the expected DNA pattern following agarose gel electrophoresis (**Fig. 17**).
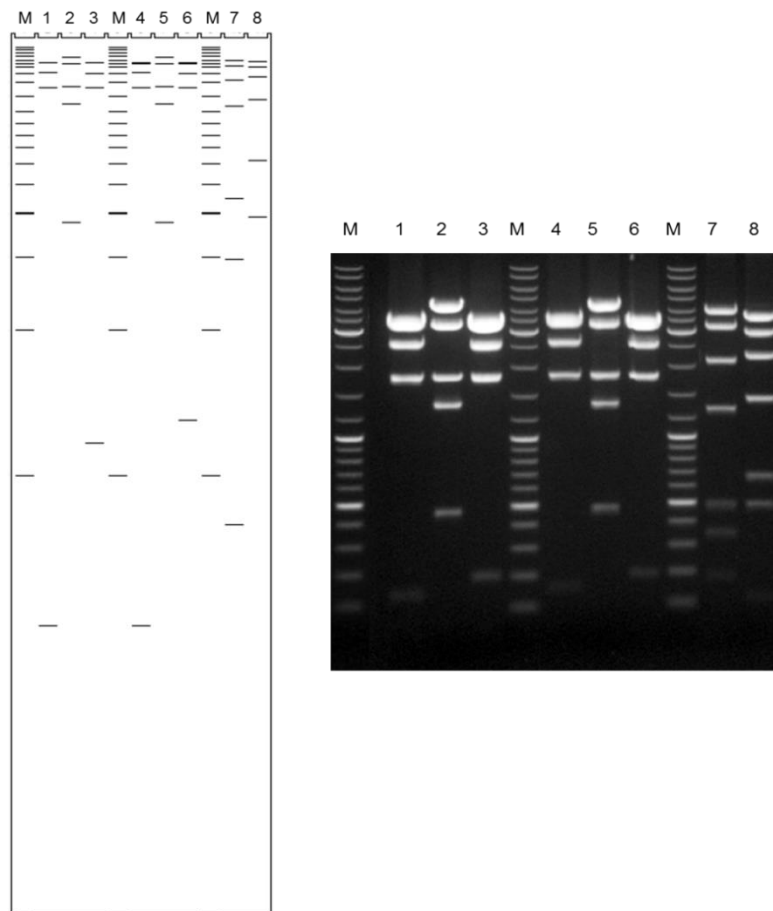
Figure 17 | Characterization of the pAd.shuttle donor plasmids by restriction fragment length analysis. Agarose gel electrophoresis of DNA fragments resulting from digestion of pAd.shuttle.AAVS1.PGK.eGFP.pA.T-TS and pAd.shuttle.AAVS1.PGK.eGFP.pA.FRT with NcoI and Bgl II (lanes 1 and 4, respectively), BamHI and MluI (lanes 2 and 5, respectively) and with NcoI and MluI (lanes 3 and 6, respectively). Digestion of pAd.shuttle.AAVS1.PGK.eGFP.pA.ITR with SmaI (lane 7) and with Eam1105I (lane 8). Lane M, GeneRuler DNA Ladder Mix molecular weight marker (Fermentas). The *in silico* restriction patterns were generated by using the Gene Construction Kit (version 2.5) software (Textco BioSoftware, Inc., West Lebanon, NH, USA).

After confirmation of the proper genetic structure of the pAd.shuttle plasmids, the constructs were digested with MssI and were used to transform recombinogenic *E. coli* strain BJ5183$^{pAd-Easy\Delta E1.\Delta 2A}$ containing the pAd.Easy-derived adenoviral vector plasmid backbone for the assembly of fiber-modified second-generation AdVs.

3.1.2 Restriction enzyme analysis of pAdEasy-derived molecular clones encoding donor DNA templates

After HR in the *E. coli* strain BJ5183$^{pAd-Easy\Delta E1.\Delta E2A}$, the resulting plasmids should contain the necessary sequences for adenoviral vector production, as well as the inserts. The proper genetic structure of the new full-length adenoviral vector molecular clones pAdEasy.AAVS1.eGFP.ITR, pAdEasy.AAVS1.eGFP.T-TS and pAdEasy.AAVS1.eGFP.FRT, was established by restriction fragment analysis following agarose gel electrophoresis (**Fig. 18**).
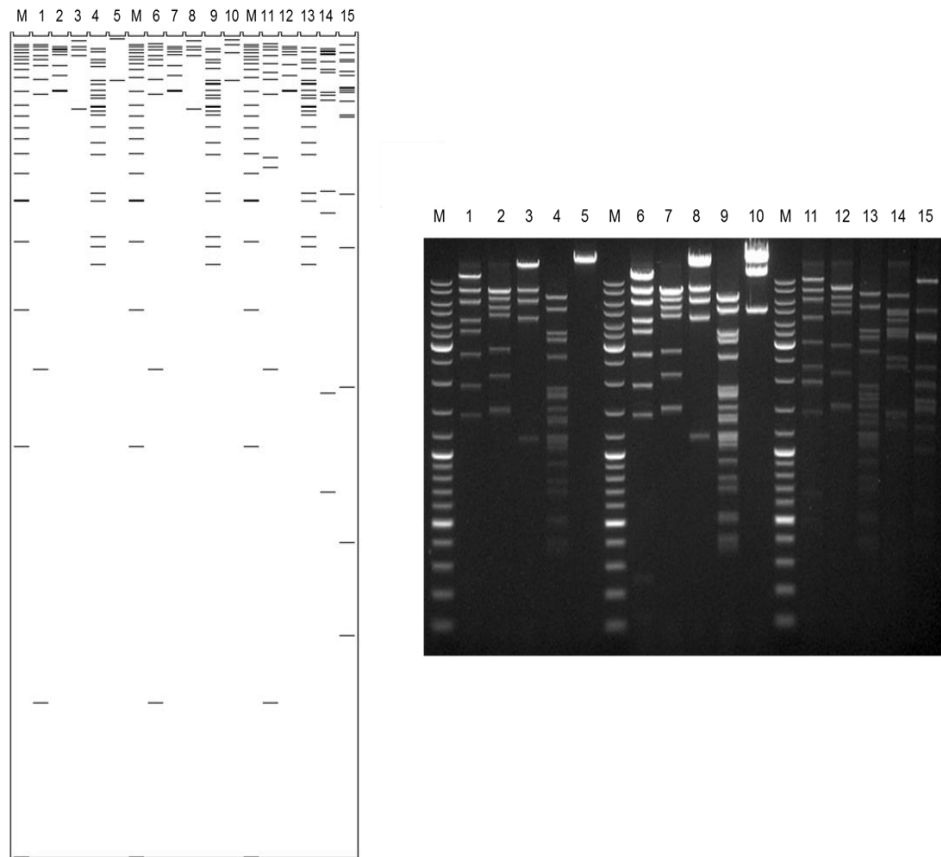
Figure 18 | Characterization of the full-length AdV molecular clones by restriction fragment length analysis. Agarose gel electrophoresis of DNA fragments resulting from digestion of pAdEasy.AAVS1.eGFP.T-TS and pAdEasy.AAVS1.eGFP.FRT with Eam1105I (lanes 1 and 6, respectively), HindIII (lanes 2 and 7, respectively), MluI (lanes 3 and 8, respectively), PvuII (lanes 4 and 9, respectively) and with XbaI (lanes 5 and 10, respectively). Digestion of pAdEasy.AAVS1.eGFP.ITR with Eam1105I (lane 11), HindIII (lane 12), PvuII (lane 13), SmaI (lane 14) and with BglI (lane 15). Lane M, GeneRuler DNA Ladder Mix molecular weight marker (Fermentas). The *in silico* restriction patterns were generated using Gene Construction Kit (version 2.5) software (Textco BioSoftware, Inc., West Lebanon, NH, USA).

Indeed, the DNA restriction pattern, although complex to analyze, it generally confirms the proper assembly of the three different types of AdV genomes in the selected molecules clones. The exception is a missing 1700 bp DNA band in the samples corresponding to pAdEasy.AAVS1.eGFP.T-TS and pAdEasy.AAVS1.eGFP.FRT DNA digested with XbaI (**Fig. 18**, **lane 5** and **10**) that can be the result of the sensitivity of the enzyme to cut at methylated sequences. It is also apparent an extra 200 bp in the plasmid that is assumed to be an integration event in the *E. coli* origin of recombination. These irregularities should not interfere with the viral production. From this constructs the corresponding adenoviral vectors were produced and characterized.

3.2 Production of second-generation fiber-modified AdVs encoding donor DNA templates

Next, it was performed the rescue, propagation and purification of the second-generation fiber-modified AdVs encoding the different donor DNA templates (i.e. AdV.Δ2.donor-ITR.F50, AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.donor-FRT.F50)

3.2.1 Characterization of recombinant AdV genomes by restriction enzyme analysis.

In order to investigate the structural integrity of vector genomes isolated from purified AdV preparations, restriction enzyme analysis was performed. In parallel, each of the corresponding full-length AdV molecular clones plasmids was taken along in this analysis to serve as molecular weight references. The restriction enzymes selected sites were XbaI, MluI and BglI for digesting AdV.Δ2.donor-FRT.F50, AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.donor-ITR.F50 genomes, respectively. The localization of the various restriction enzyme recognition sites along the various linear AdV DNA molecules and the respective restriction fragment patterns are presented (**Fig. 19**).

From the analysis of the various enzymatic restriction patterns (**Fig. 20**) one can observe that the restriction fragments diagnostic for the adenoviral genomic sequences and the transgenic DNA elements (i.e. the 4.9 kb, 5.3 kb and 5.0 kb bands for the AdV.Δ2.donor-FRT.F50, AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.donor-ITR.F50 vectors, respectively) match those of the expected patterns. Thus, we conclude that the three AdV-donor's contain intact full-length vector genomes with no evidence of deletions or other types of DNA rearrangements.
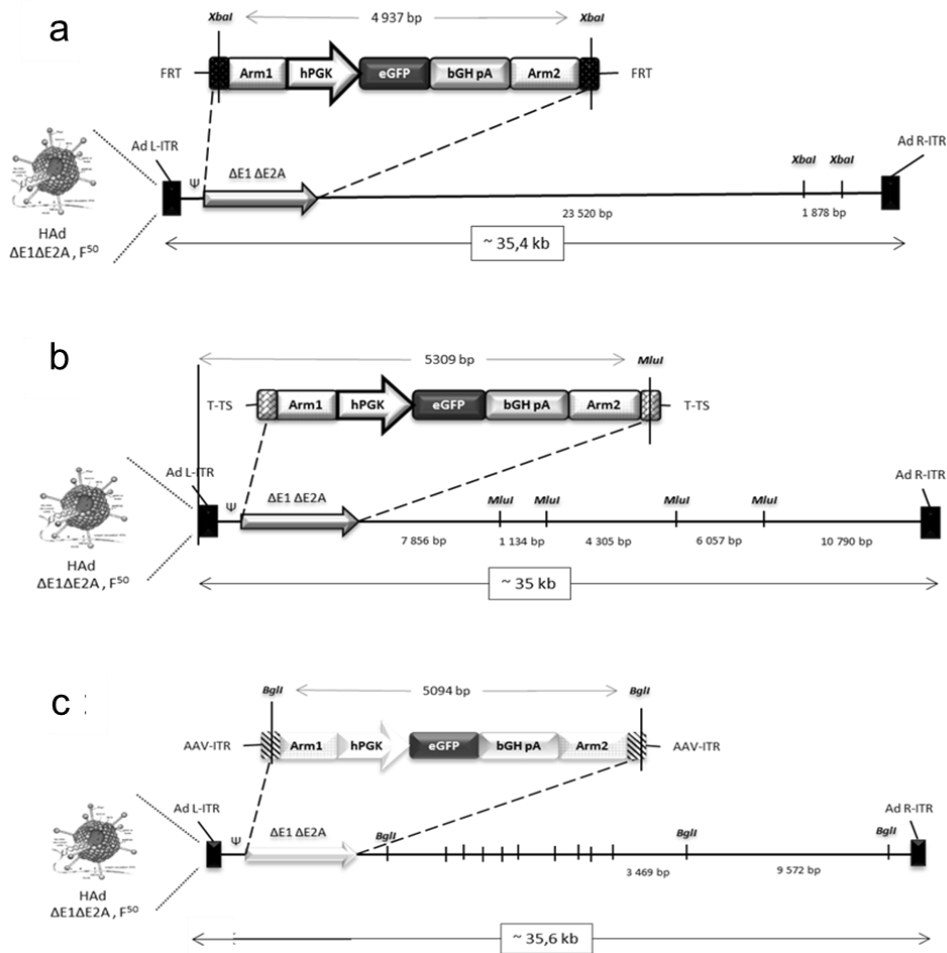
Figure 19 | Schematic representation of the physical maps of the AdV donor genomes. The restriction sites used to characterize and confirm the integrity of the donors genome of (**a**) AdV.Δ2.donor-FRT.F50 (**b**) AdV.Δ2..donor-T-TS.F50 and (**c**) AdV.Δ2.donor-ITR.F50 are represented as XbaI, MluI and BglI, respectively.
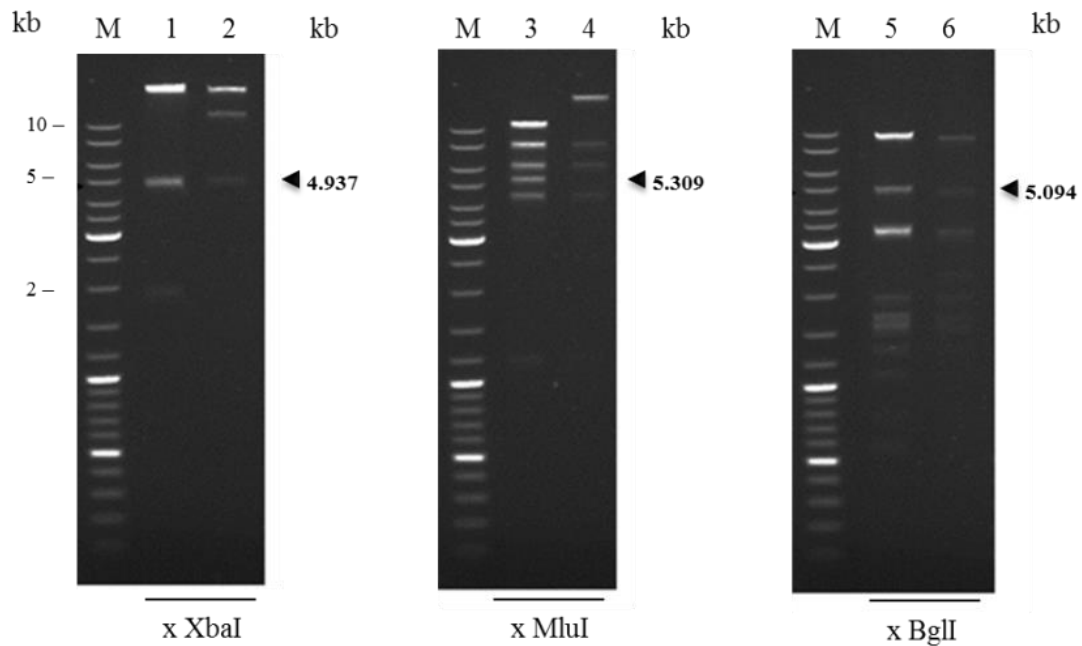
Figure 20 | Restriction enzyme analysis of AdV genomes to establish their structural integrity. Digestion of the AdVΔ2.donor-FRT.F50 and pAd.shuttle.AAVS1.PGK.eGFP.pA.FRT with XbaI (lane 1 and 2), AdV.Δ2.donor-T-TS.F50 and pAd.shuttle.AAVS1.PGK.eGFP.pA.T-TS with MluI (lane 3 and 4) and AdV.Δ2.donor.F50 and pAd.shuttle.AAVS1.PGK.eGFP.pA.ITR with BglI (lane 5 and 6). The arrowhead points to the DNA fragments diagnostic for intact transgenic DNA sequences in AdV genomes. Lane M, GeneRuler DNA Ladder Mix molecular weight marker (Fermentas).

3.2.2 Gene delivery activity of AdV preparations on target cell by end-point titrations

To determine the transduction efficiency of the newly produced AdV particles, we conducted an end-point titration assay in HeLa cells. The resulting functional titers, measured by eGFP-directed flow cytometry at three days post-transduction, are expressed in terms of HeLa cell-transducing units per milliliter (HTU/ml).

The AdV titers presented in Table 4 show that the three different AdV preparations display a very high and similar gene transfer activities in indicator target cells. These titers will be used to determine the MOI used in the subsequent experiments.

Table 4 | AdV donor titers in HeLa cells.

Vector donor	Target cells	Titer (HTU/ml)
AdV.Δ2.donor-FRT.F50		3.46×10^{10}
AdV.Δ2.donor-T-TS.F50	HeLa	2.44×10^{10}
AdV.Δ2.donor-ITR.F50		2.84×10^{10}

3.2.3 Testing FLPe- and TALEN-mediated topological change of donor DNA templates in target cells

The following experiments use FLPe and TALENs as mediators of circularization and linearization of donor DNA templates delivered by AdV.Δ2.donor-FRT.F50, and AdV.Δ2.donor-T-TS.F50 into target cells, respectively (**Fig. 21**). To this end, two primers were designed to target sequences within the two *AAVS1* arms of homology. If direct FLPe-dependent circularization and TALEN-dependent linearization followed by circularization of the donor DNA in AdV.Δ2.donor-FRT.F50, and AdV.Δ2.donor-T-TS.F50, respectively, does occur in target cells, it should yield a 619 bp PCR product. This product should not appear if the donor DNA remains embedded in the context of linear AdV genomes.

Consequently, cellular DNA was extracted from HeLa cells co-transduced with AdV.Δ2.donor-FRT.F50 and a FLPe-encoding vector hcAd.FLPe.F50, as well as from HeLa cells transduced with a combination of AdV.Δ2.donor-T-TS.F50, AdV.Δ2.TALEN-L[S1] and AdV.Δ2.TALEN-R[S1] (AdV.Δ2.TALEN[S1] L+R). PCR amplifications with "outward-facing" primers (half arrows) on linear templates will result in specific products only upon donor DNA excision and circularization.
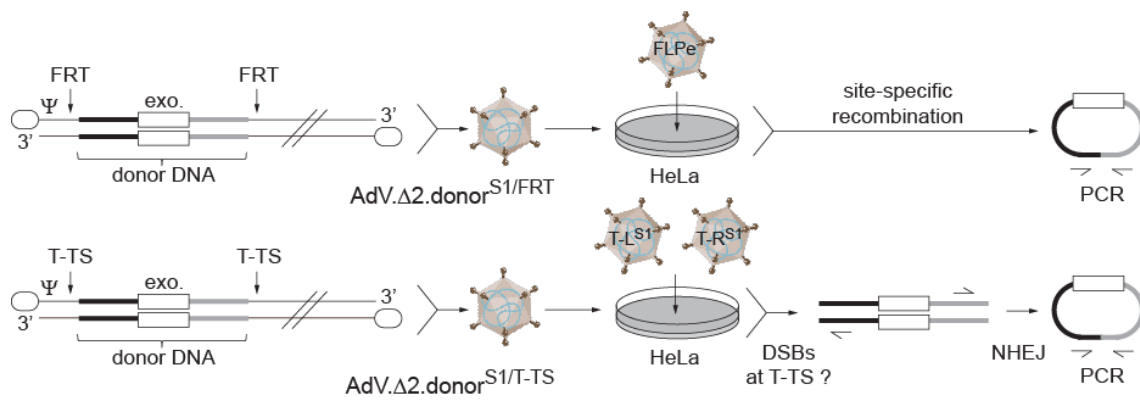
Figure 21 | Experimental design for testing FLPe- and TALEN-mediated release of donor DNA from AdV genomes in transduced cells. The generic structures of AdV.Δ2.donor-FRT.F50 and AdV.Δ2.donor-T-TS.F50 recombinant genomes display the FRT and T-TS target sites for the yeast site-specific FLPe recombinase and the *AAVS1*-specific TALEN pair, respectively. Open box, exogenous DNA flanked by *AAVS1*-targeting sequences (black and grey bars). Ψ, AdV packaging signal. Open oval, 5' covalently-attached AdV terminal protein.

In the combination between AdV.Δ2.donor-FRT.F50 and FLPe the expected 619 bp band appears, validating the activity of the FLPe and circularization of the donor (**Fig. 22**). Likewise, in the combination of AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN^{S1} L+R the discriminating appears, confirming the activity of the TALEN's to linearize and release the insert from the adenoviral genome context by DSB. These DSBs are recognized by the NHEJ cellular repair pathway and joined together, producing a circular form of the donor.
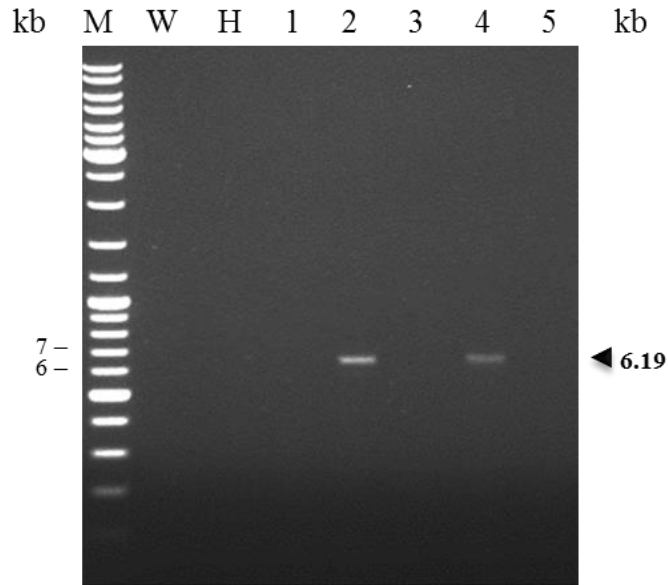
Figure 22 | Detection of FLPe-mediated circularization and TALEN-mediated linearization followed by circularization of the AdV.Δ2.donor-FRT.F50 and AdV.Δ2.donor-T-TS.F50 donor DNA templates, respectively. PCR products of cells transduced with AdV.Δ2.donor-FRT.F50, AdV.Δ2.donor-FRT.F50 + FLPe, AdV.Δ2.donor-T-TS.F50, AdVΔ2.donor-T-TS.F50 + AdV.Δ2.TALEN[S1] L+R and AdV.Δ2.donor-ITR.F50 (lanes 1, 2, 3, 4 and 5, respectively). The arrowhead points to the DNA fragments amplified by PCR. W, water control. H, mock control. Lane M, GeneRuler DNA Ladder Mix molecular weight marker (Fermentas).

After confirmation of the donors genome integrity, gene delivery activity and FLPe- and TALEN-mediated topological changes of donor DNA templates, these were used for the following transduction experiments.

3.3 Transduction experiments

3.3.1 Transduction experiment in HeLa cells

In the following experiment, HeLa cells were transduced with the three different AdV donors, in combination with AdV.Δ2.TALEN[S1] L+R and with FLPe in the case of the AdV.Δ2.donor-FRT.F50. The use of AdV.Δ2.donor-FRT.F50 in the absence of the the FLPe will serve to deliver into target cells donor DNA templates with a standard linear topology.

From the eGFP-directed fluorescence microscopic analysis (**Fig. 23**), it is observed similar initial gene transfer activities in the transductions corresponding to the various experimental set-ups. As expected, the eGFP signal is absent in the negative control and present in all the HeLa cell cultures exposed to the AdV donors, whereas the DsRed signal could only be detected in cultures incubated with the TALEN-encoding AdVs (**Fig. 24**). In addition, the cells looked healthy and didn't show overt vector-associated cytotoxicity.
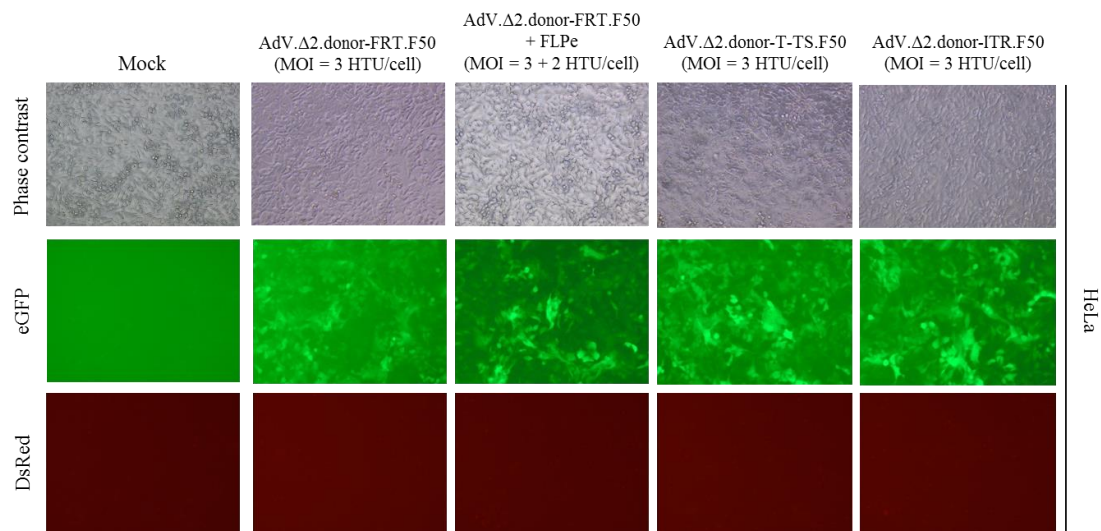
Figure 23 | Direct fluorescence microscopy of HeLa cell cultures exposed to the three different AdV donors at an early time point post-transduction. The HeLa cells were transduced at the indicated MOIs and were observed by phase contrast and by eGFP- and DsRed-directed fluorescence microscopy at 3 days post-infection. Original magnification: X40.
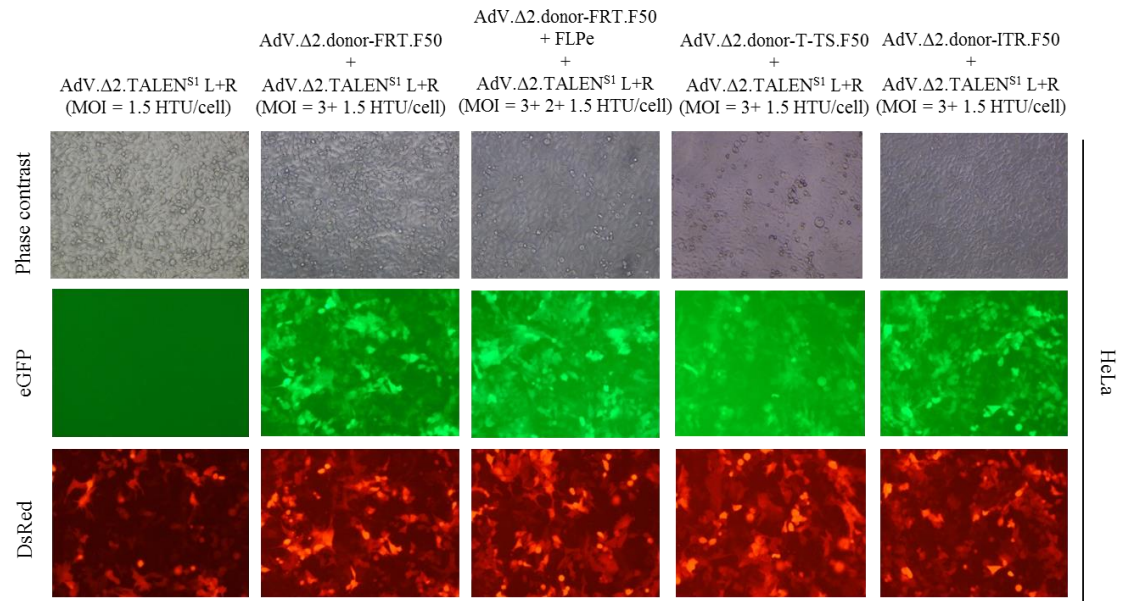
Figure 24 | Direct fluorescence microscopy of HeLa cell cultures exposed to the three different AdV donors in combination with the TALEN-encoding AdVs at an early time point post-transduction. The HeLa cells were transduced at the indicated MOIs and were observed by phase contrast and by eGFP- and DsRed-directed fluorescence microscopy at 3 days post-infection. Original magnification: X40.

The transduced HeLa cells were sub-cultured and the frequency of eGFP and DsRed positive cells was measured by flow cytometry in three samples for each condition, corresponding to 10^4 cells per sample; the mean from these three measurements was calculated. The samples corresponding to mock-transduced cells served to set the background fluorescence of the assay.

At 7 days post-transduction (**Fig. 25**), the cell cultures transduced with the AdV donors alone exhibited a high frequency of eGFP-positive cells (75-80%). The cells transduced with the AdV donors in combination with the DsRed-encoding AdV.Δ2.TALEN[S1] L+R held high frequencies of DsRed- (13-22%) as well as eGFP-positive cells (28-29%). The reduced number of eGFP-positive cells, when compared to the condition with the AdV donors alone, can perhaps be result of a vector "overload" of the cells with the combined two AdV vectors.

Additionally, it can also be observed that the mean percentages of reporter-positive cells are relatively similar between transductions and there are no evident initial experimental discrepancies in the transductions.

	Mock	FRT	FRT + FLPe	T-TS	AAV-ITR	TALEN L+R	FRT + TALEN L+R	FRT + FLPe + TALEN L+R	T-TS + TALEN L+R	AAV-ITR + TALEN L+R
■eGFP	0,08	76,13	74,64	80,44	75,36	0,02	27,98	28,05	27,54	28,59
☐DsRed	0,34	0,00	0,00	0,00	0,00	61,50	13,17	13,30	21,59	13,99

Figure 25 | Flow cytometric analysis of transduced HeLa cells at 7 days post-transduction. HeLa cells were transduced with AdV donors alone, in combination with the FLPe in the case of the AdV.Δ2.donor-FRT.F50 or co-transduced with AdV.Δ2.TALEN[S1] L+R. Mock-transduced HeLa cells (mock) and HeLa cells transduced with AdV.Δ2.TALEN[S1] L+R alone (TALEN L+R) served as negative and positive controls, respectively. Quantification of the number of eGFP-positive cells was carried out by flow cytometry at 7 days post-transfection with 10 000 events, corresponding to viable cells being measured per sample. FRT, AdV.Δ2.donor-FRT.F50; T-TS, AdV.Δ2.donor-T-TS.F50; AAV-ITR, AdV.Δ2.donor-ITR.F50; black bars, eGFP-positive cells; grey bars, DsRed-positive cells.

Some representative flow cytometry dot plots corresponding to these experiments (**Fig. 26**) also support the previous observations by visualization of the distribution of the single positive cells: the HeLa cells transduced with the AdV donors alone only exhibit eGFP-positive cells in a uniformly distributed population; HeLa cells co-transduced with the AdV donors and the TALEN encoding AdVs present a homogeneous distribution of eGFP- and DsRed-positive cells. Once more, there are no apparent dissimilarities between the AdV donors reporters frequencies.
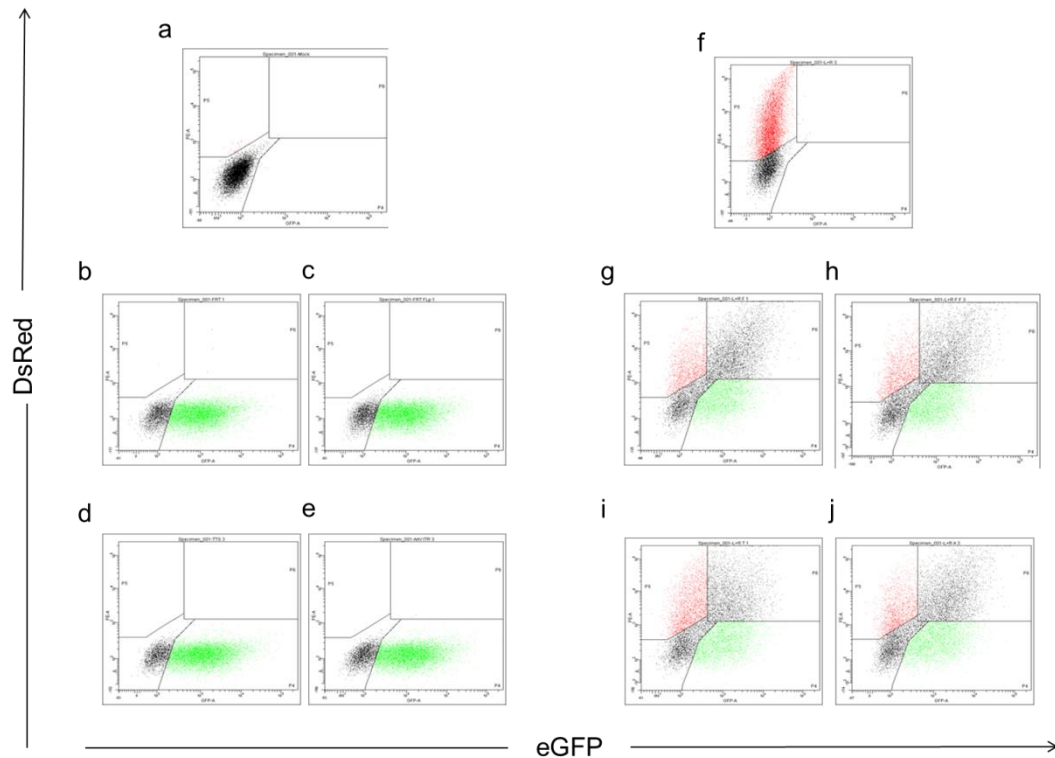
Figure 26 | Dot plot representation of reporter's expression in HeLa cells at 7 days post-transduction. Cells transduced with (**a**) mock, (**b**) AdV.Δ2.donor-FRT.F50, (**c**) AdV.Δ2.donor-FRT.F50+ FLPe, (**d**) AdV.Δ2.donor-T-TS.F50, (**e**) AdV.Δ2.donor-ITR.F50, (**f**) AdV.Δ2.TALEN[S1] L+R, **g**) AdV.Δ2.donor-FRT.F50 + AdV.Δ2.TALEN[S1] L+R, (**h**) AdV.Δ2.donor-FRT.F50+ FLPe + AdV.Δ2.TALEN[S1] L+R, (**i**) AdV.Δ2.donor-T-TS.F50 + AdV.Δ2.TALEN[S1] L+R and (**j**) AdV.Δ2.donor-ITR.F50 + AdV.Δ2.TALEN[S1] L+R. Flow cytometry was carried out 7 days post-transfection with 10 000 viable cells being analyzed per sample.

The HeLa cells were sub-cultured and, at 28 days post-transduction, each sample was mixed in a cell pool for each condition and the eGFP and DsRed reporter signals were measured, in a total of 100.000 events (**Fig. 27**). The DsRed-positive cells substantially decreased in all the conditions and reached nearly 0% in all of the conditions. Similarly, the average eGFP signal also decreased considerably and is much lower in the cells transduced with the AdV donors alone than in combination with AdV.Δ2.TALEN[S1] L+R (0.1-0.8% to 0.11-0.49%, respectively). There is a clear higher eGFP-positive cells frequency in cells co-transduced with the AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN[S1] L+R compared with the other donors.

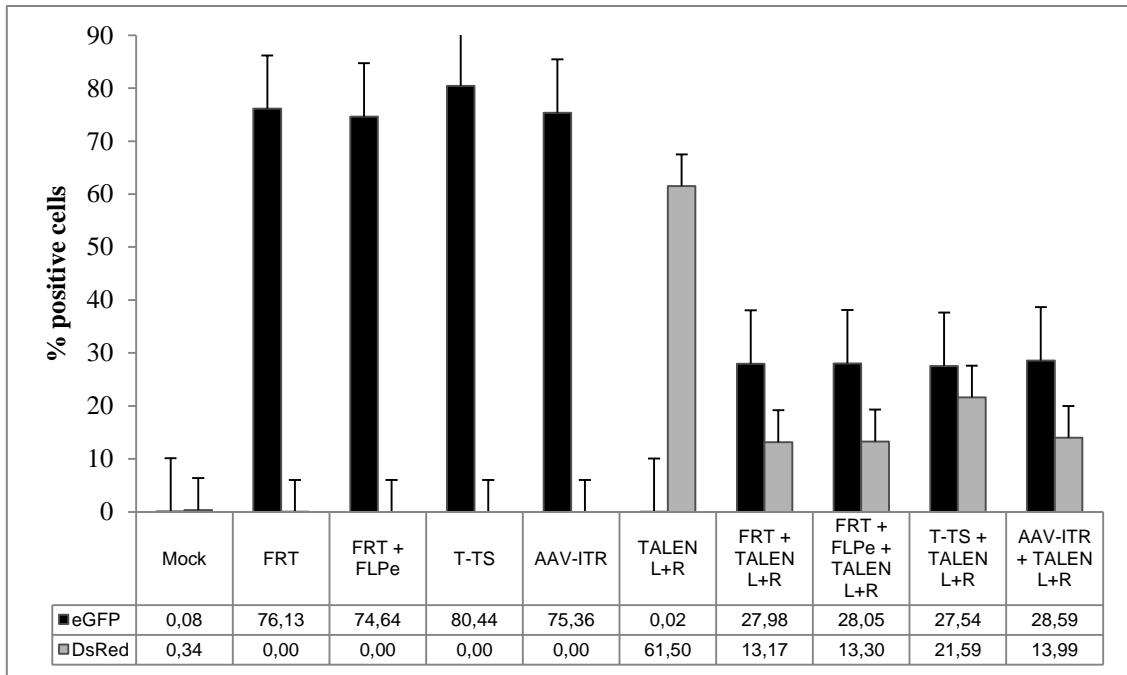|	Mock	FRT	FRT + FLPe	T-TS	AAV-ITR	TALEN L+R	FRT + TALEN L+R	FRT + FLPe + TALEN L+R	T-TS + TALEN L+R	AAV-ITR + TALEN L+R
■eGFP	0,00	0,07	0,01	0,03	0,08	0,00	0,26	0,30	0,49	0,11
▢DsRed	0,00	0,00	0,00	0,00	0,00	0,02	0,01	0,01	0,02	0,02
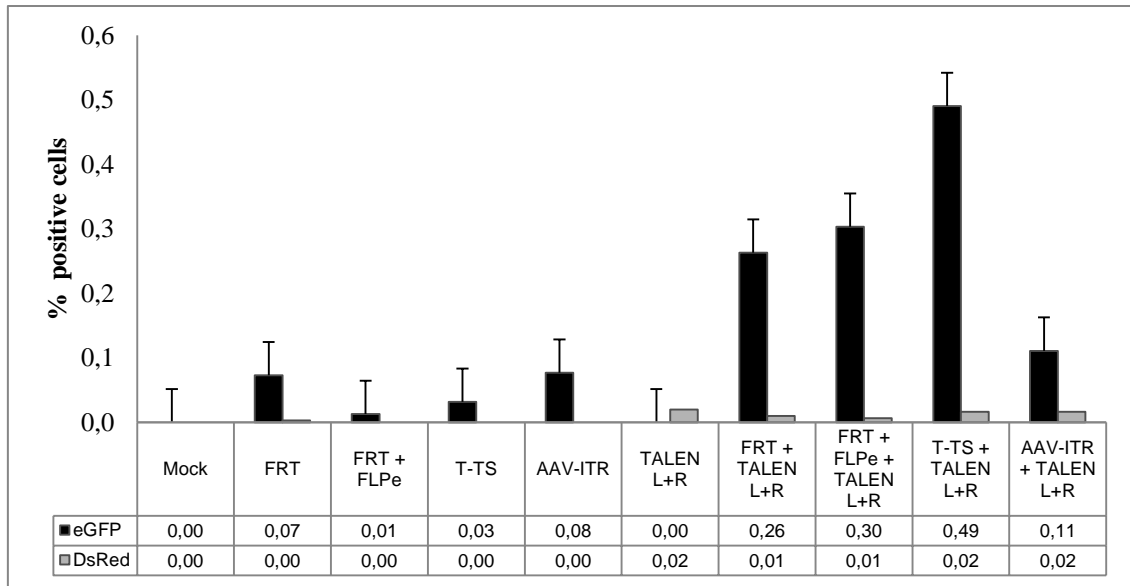
Figure 27 | Flow cytometric analysis of transduced HeLa cells at 28 days post-transduction. HeLa cells were transduced with AdV donors alone, in combination with the FLPe in the case of the AdV.Δ2.donor-FRT.F50 or co-transduced with AdV.Δ2.TALEN[S1] L+R. Mock-transduced HeLa cells (mock) and HeLa cells transduced with AdV.Δ2.TALEN[S1] L+R alone (TALEN L+R) served as negative and positive controls, respectively. Quantification of the number of eGFP-positive cells was carried out by flow cytometry at 28 days post-transfection with 100 000 events, corresponding to viable cells being measured per sample. FRT, AdV.Δ2.donor-FRT.F50; T-TS, AdV.Δ2.donor-T-TS.F50; AAV-ITR, AdV.Δ2.donor-ITR.F50; black bars, eGFP-positive cells; grey bars, DsRed-positive cells.

In summary, if we compare the measurements in the beginning and end of the transduction experiment (**Table 5**) we can observe high frequencies of eGFP-positive cells in the beginning, in the AdV donors alone, as well as in the combination of the AdV donors with the TALEN-encoding AdVs; those frequencies are similar between different donors. Throughout the time of the experiment there was a decrease of the eGFP signal, reaching nearly 0% in the AdV donors alone. In the combination of the AdV donors with the TALEN-encoding AdVs, the signal is higher, especially in the AdV.Δ2.donor-T-TS.F50 that presents a 1.7-fold increase compared with the standard linear topology.

Table 5 | Frequencies of eGFP-positive HeLa cells at 7 and 28 days post-transduction.

Condition		Mean % eGFP	
		Day 7 post-transduction	Day 28 post-transduction
Mock		0,08	0,00
AdV.Δ2.donor-FRT.F50		76,13	0,07
AdV.Δ2.donor-FRT.F50+ FLPe		74,64	0,01
AdV.Δ2.donor-T-TS.F50		80,44	0,03
AdV.Δ2.donor-ITR.F50		75,36	0,08
AdV.Δ2.TALEN[S1] L+R		0,02	0,00
AdV.Δ2.donor-FRT.F50	+ AdV.Δ2.TALEN[S1] L+R	27,98	0,26
AdV.Δ2.donor-FRT.F50 + FLPe		28,05	0,30
AdV.Δ2.donor-T-TS.F50		27,54	0,49
AdV.Δ2.donor-ITR.F50		28,59	0,11

3.3.2 Repetition of the transduction experiment in HeLa cells

To further substantiate these results, replication of the transduction experiment was made, this time only with the AdV.Δ2.donor-T-TS.F50 with or without the AdV.Δ2.TALEN[S1] L+R at the same MOIs. Although not presented, the transduction of the HeLa cells, observed by direct fluorescence microscopy, resembles the previous one.

At 3 days post-transduction (**Fig. 28**) frequencies of both reporter-positive cells are low in the mock control, as expected, and the percentage of eGFP-positive cells is higher in the AdV donor alone than in the combination with the AdV.Δ2.TALEN[S1] L+R, as seen in the first experiment.
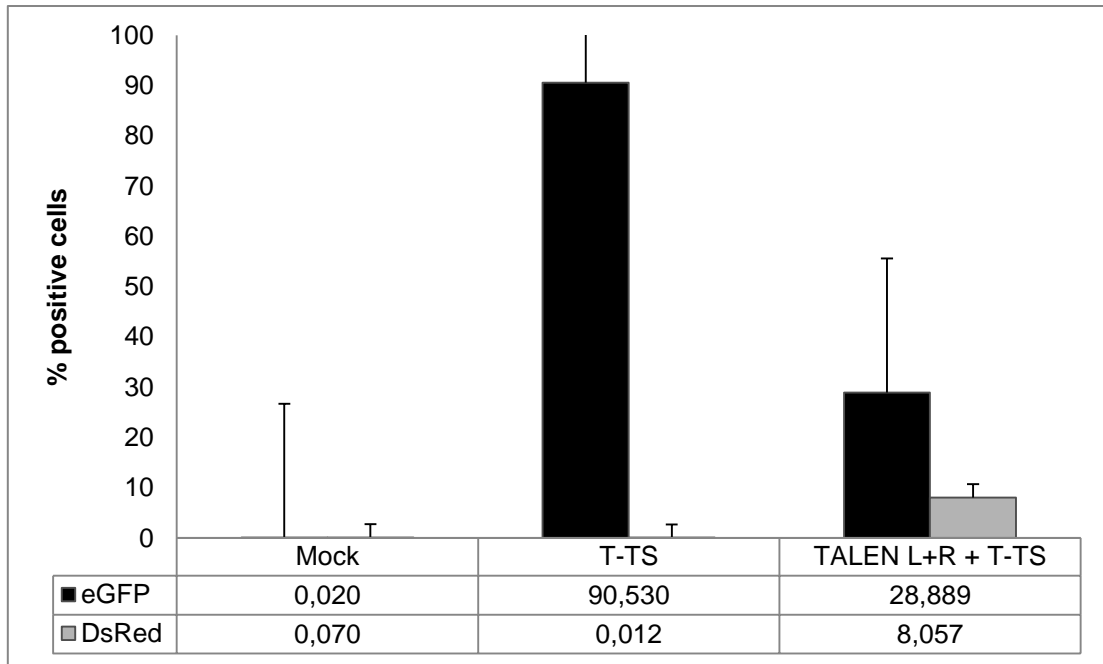
Figure 28 | Flow cytometric analysis of transduced HeLa cells at 3 days post-transduction in the second transduction experiment**.** HeLa cells were transduced with AdV.Δ2.donor-T-TS.F50 (T-TS) alone or co-transduced with AdV.Δ2.TALEN[S1] L+R (TALEN L+R). Mock-transduced HeLa cells (mock) served as negative control. Quantification of the number of eGFP-positive cells was carried out by flow cytometry at 3 days post-transfection with 10 000 events, corresponding to viable cells being measured per sample. Black bars, eGFP-positive cells; grey bars, DsRed-positive cells.

The flow cytometry dot blot (**Fig. 29**) shows the single events distribution and density and it is similar to the ones in the beginning of the first experiment for the conditions shown.
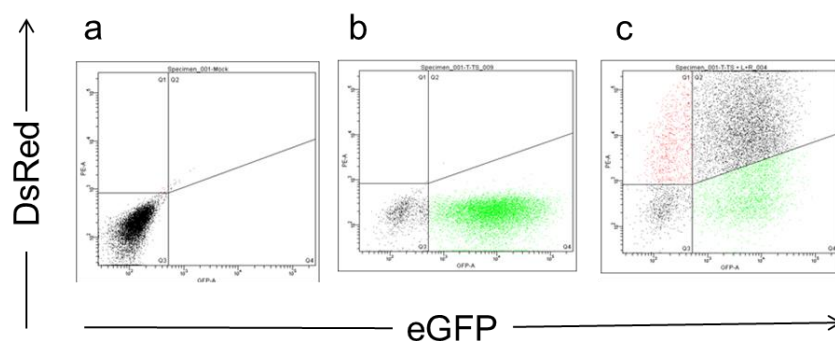
Figure 29 | Dot plot representation of reporter's expression in HeLa cells at 3 days post-transduction in the second transduction experiment. Cells transduced with (**a**) mock, (**b**) AdV.Δ2.donor-T-TS.F50 and (**c**) AdV.Δ2.donor-T-TS.F50 + AdV.Δ2.TALEN[S1] L+R. Flow cytometry was carried out 3 days post-transfection with 10 000 viable cells being analyzed per sample.

The last time point measured in this second experiment was at 27 days post-transduction (**Fig. 30**). The frequency of eGFP-positive cells in the combination of the AdV.Δ2.donor-T-TS.F50 with AdV.Δ2.TALEN[S1] L+R is higher than the AdV donor alone. Likewise, the flow cytometry dot blot (**Fig. 31**) displays a high density eGFP-positive population appearing in the cells co-transduced with AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN[S1] L+R.

	Mock	T-TS	TALEN L+R + T-TS
■ eGFP %	0,000	0,053	0,827
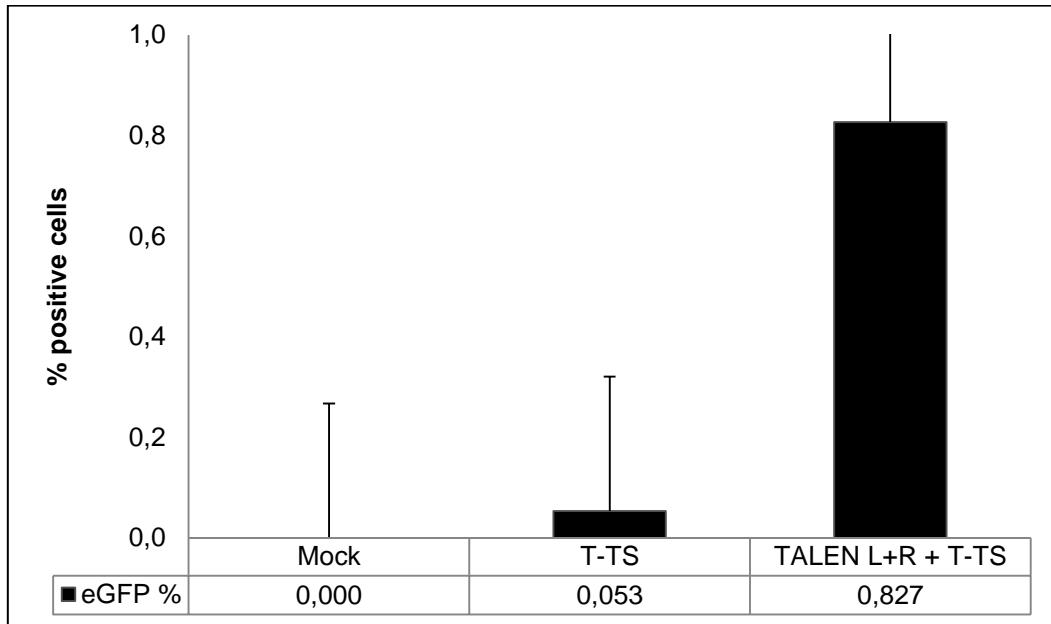
Figure 30 | Flow cytometric analysis of transduced HeLa cells at 27 days post-transduction in the second transduction experiment. HeLa cells were transduced with AdV.Δ2.donor-T-TS.F50 (T-TS) alone or co-transduced with AdV.Δ2.TALEN[S1] L+R (TALEN L+R). Mock-transduced HeLa cells (mock) served as negative control. Quantification of the number of eGFP-positive cells was carried out by flow cytometry at 27 days post-transfection with 10 000 events, corresponding to viable cells being measured per sample. Black bars, eGFP-positive cells.

In short, in this second experiment the percentage of eGFP-positive cells was, in the beginning of the experiment, higher in the AdV.Δ2.donor-T-TS.F50 than in combination with AdV.Δ2.TALEN[S1] L+R, but that shifted by the end of the experiment, being the latter around 16-fold higher than the AdV donor alone (**Table 6**).
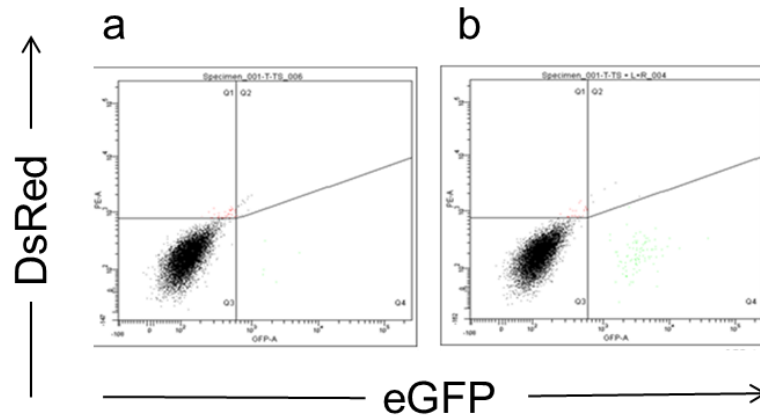
Figure 31 | Dot plot representation of reporter's expression in HeLa cells at 27 days post-transduction in the second transduction experiment. Cells transduced with (**a**) AdV.Δ2.donor-T-TS.F50 and (**b**) AdV.Δ2.donor-T-TS.F50 + AdV.Δ2.TALEN[S1] L+R. Flow cytometry was carried out 27 days post-transfection with 10 000 viable cells being analyzed per sample.

Table 6 | Frequencies of eGFP-positive HeLa cells at 3 and 27 days post-transduction in the second transduction experiment.

	Mean % eGFP	
Condition	**Day 3 post-transduction**	**Day 27 post-transduction**
Mock	0.02	0.00
AdV.Δ2.donor-T-TS.F50	90,53	0,05
AdV.Δ2.donor-T-TS.F50 + AdV.Δ2.TALEN[S1] L+R	28,89	0,83

From this second experiment we can validate the results from the first one and confirm that, indeed, the combination between AdV.Δ2.donor-T-TS.F50 and + AdV-TALEN[S1] L+R presents a significantly higher eGFP-positive cells frequency after 27-28 days post-transduction, i.e., in the stably transduced HeLa cells.

3.3.3 Transduction experiment in DMD myoblasts

The next transduction experiment was conducted in immortalized DMD myoblasts, as their genetic background is more stable than the HeLa cells, and should produce more sustainable results. Furthermore, in this transduction experiment, an AdV.Δ2.donor-eGFP.F50 was used as a positive control, as this donor contains the transgene devoid of the cloned terminal sequences.

From the observation of the transduced DMD myoblasts by direct fluorescence microscopy (**Fig. 32**), we can observe that the initial transductions were successful, with high numbers of eGFP-positive cells in the AdV donors, as well as of DsRed-positive cells in the co-transduction of the AdV donors with the AdV.Δ2.TALEN[S1] L+R.

The myoblasts were sub-cultured up until day 35 post-transduction, when the cells were visualized by direct fluorescence microscopy (**Fig. 33**). It was clear that the percentage of eGFP-positive cells in the combination of AdV.Δ2.donor-T-TS.F50 with AdV.Δ2.TALEN[S1] L+R was much higher than in the combination with the other AdV donor.
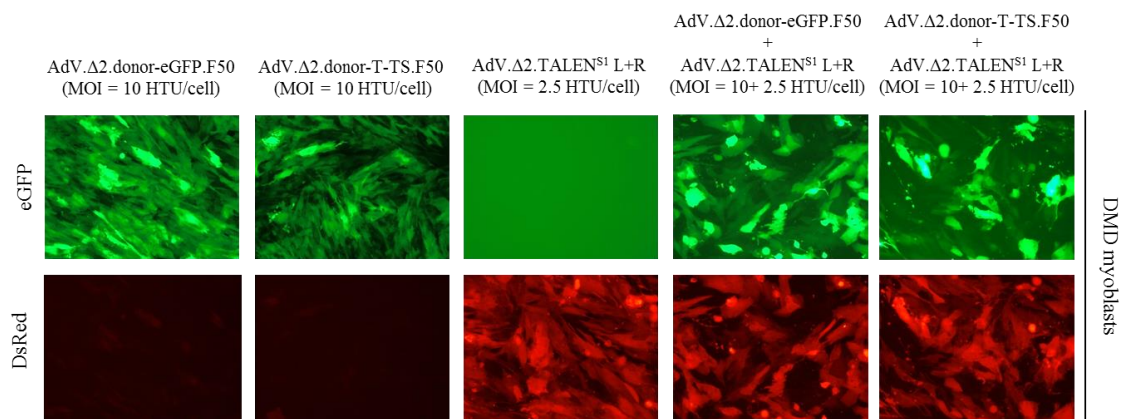
Figure 32 | Direct fluorescence microscopy of DMD myoblasts cell cultures exposed to the two different AdV donors in combination with the TALEN-encoding AdVs at an early time point post-transduction. The myoblasts were transduced with the AdV donors (AdV.Δ2.donor-eGFP.F50 and AdV.Δ2.donor-T-TS.F50) with or without AdV.Δ2.TALEN[S1] L+R at the indicated MOIs and were observed by phase contrast and by eGFP- and DsRed-directed fluorescence microscopy at 3 days post-infection. Original magnification: X40.
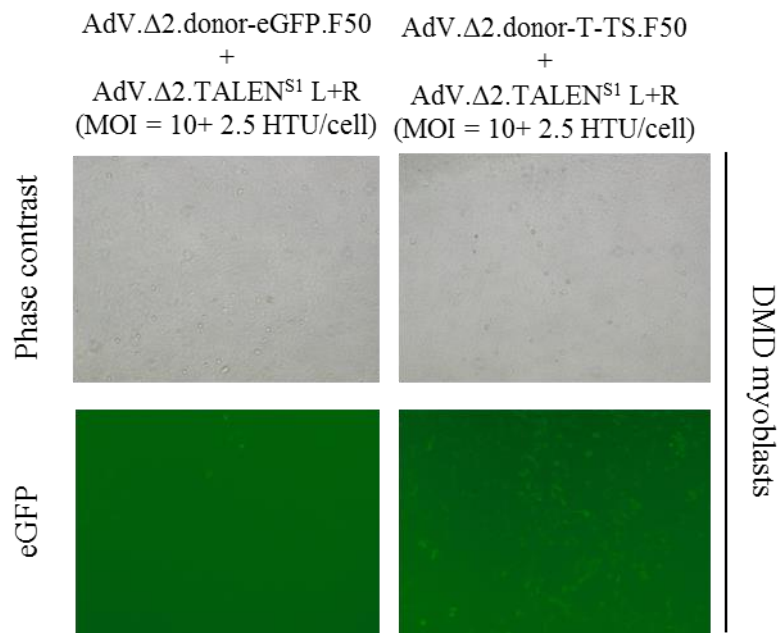
Figure 33 | Direct fluorescence microscopy of DMD myoblasts cell cultures exposed to the two different AdV donors in combination with the TALEN-encoding AdVs at 35 days post-transduction. The myoblasts transduced with the AdV donors (AdV.Δ2.donor-eGFP.F50 and AdV.Δ2.donor-T-TS.F50) with AdV.Δ2.TALEN[S1] L+R were observed by phase contrast and by eGFP-directed fluorescence microscopy at 35 days post-infection. Original magnification: X40.

Similarly, the frequencies of eGFP-positive cells were measured by flow cytometry at 35 days post-transduction (**Fig. 34** and **Table 7**), either in the different conditions or in the correspondingly cell pools. The measurements revealed that both the AdV donors, in combination with AdV-TALEN[S1] L+R, hold significantly higher percentages of eGFP-positive cells than the respective AdV donors alone. Additionally, and confirming the results of the two previous transduction experiments, the combination of AdV.Δ2.donor-T-TS.F50 with AdV-TALEN[S1] L+R displays a 2,9-fold increased eGFP-positive cell frequency than the other donor, taken as a positive control for integration.
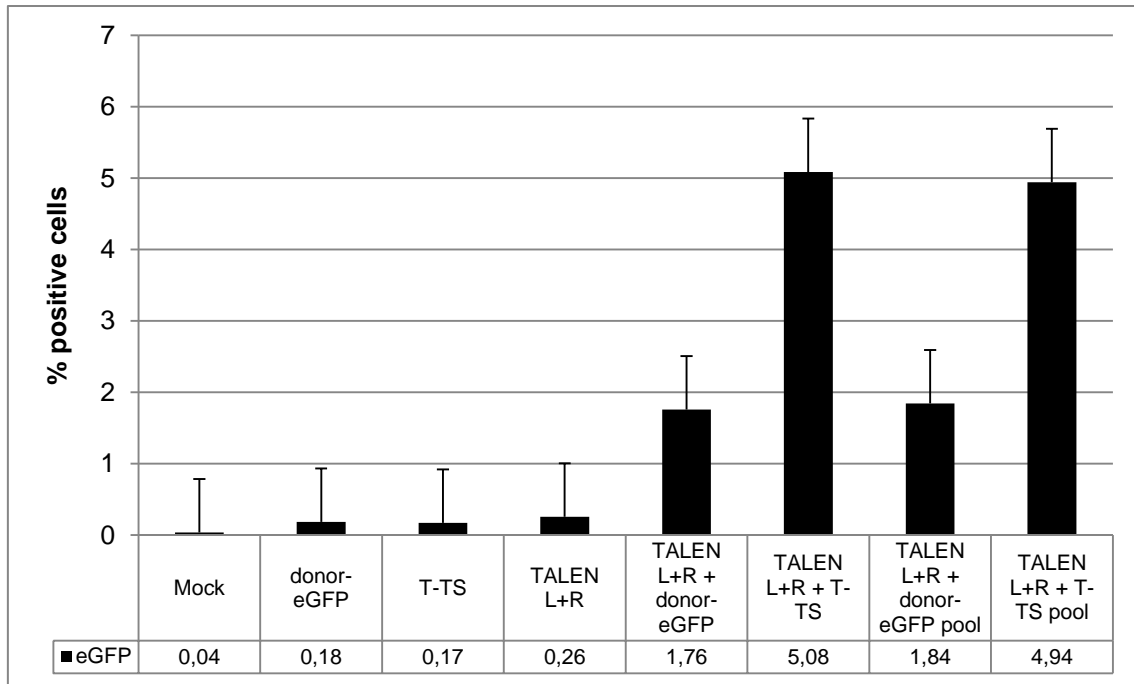
Figure 34 | Flow cytometric analysis of transduced DMD myoblasts at 35 days post-transduction. The myoblasts were transduced with AdV.Δ2.donor-T-TS.F50 (T-TS) or AdV.Δ2.donor-eGFP.F50 (donor-eGFP) alone or co-transduced with AdV-TALEN[S1] L+R (TALEN L+R). Mock-transduced myoblasts (mock) served as negative control. Quantification of the number of eGFP-positive cells was carried out by flow cytometry at 35 days post-transfection with 10 000 events or 100 000 events (in the cell pool), corresponding to viable cells being measured per sample. Black bars, eGFP-positive cells.

Table 7 | Frequencies of eGFP-positive DMD myoblasts at 35 days post-transduction.

Conditions	Mean % eGFP-positive cells
Mock	0,04
AdV.Δ2.donor-eGFP.F50	0,18
AdV.Δ2.donor-T-TS.F50	0,17
Ad AdV-TALEN[S1] L+R	0,26
AdV.Δ2.donor-eGFP.F50 + AdV-TALEN[S1] L+R	1,76
AdV.Δ2.donor-T-TS.F50 + AdV-TALEN[S1] L+R	5,08
AdV.Δ2.donor-eGFP.F50 + AdV-TALEN[S1] L+R (cell pool)	1,84
AdV.Δ2.donor-T-TS.F50 + AdV-TALEN[S1] L+R (cell pool)	4,94

3.4 Transgene integration analysis

We sought to investigate whether the stable genetic modification of the cells was brought about by targeted or by random DNA insertion. For that, the eGFP-positive HeLa cells co-transduced with AdV.Δ2.donor-T-TS.F50 and AdV-TALEN[S1] L+R from the first transduction experiment were FACS sorted at 28 days post-transduction and clonally expanded. Next, genomic DNA was extracted from selected cell clones (n=118) and was subjected to PCR analysis using primers designed to yield amplicons diagnostic for HR-derived junctions between exogenous and native target DNA: #331 and #332, #551 and #552, #111 and #112 specific for the 3'-junction of the *AAVS1* locus, 5'-junction of the *AAVS1* locus and the eGFP reporter gene, respectively. Mock-transduced HeLa cells and a shuttle plasmid encoding the transgene served as negative and positive control, respectively.

3.4.1 Integration in the *AAVS1* 3'-junction

HR-mediated transgene insertion in the sorted clones should give rise to a 1709 bp amplicon with primers #331 and #332. The PCR products consistent with this process were present in 78% (i.e. 92 out of 118) of the clones, i.e., are positive for *AAVS1* 3'-junction transgene integration (**Fig. 35** and **Supplementary Fig. 40**).
.
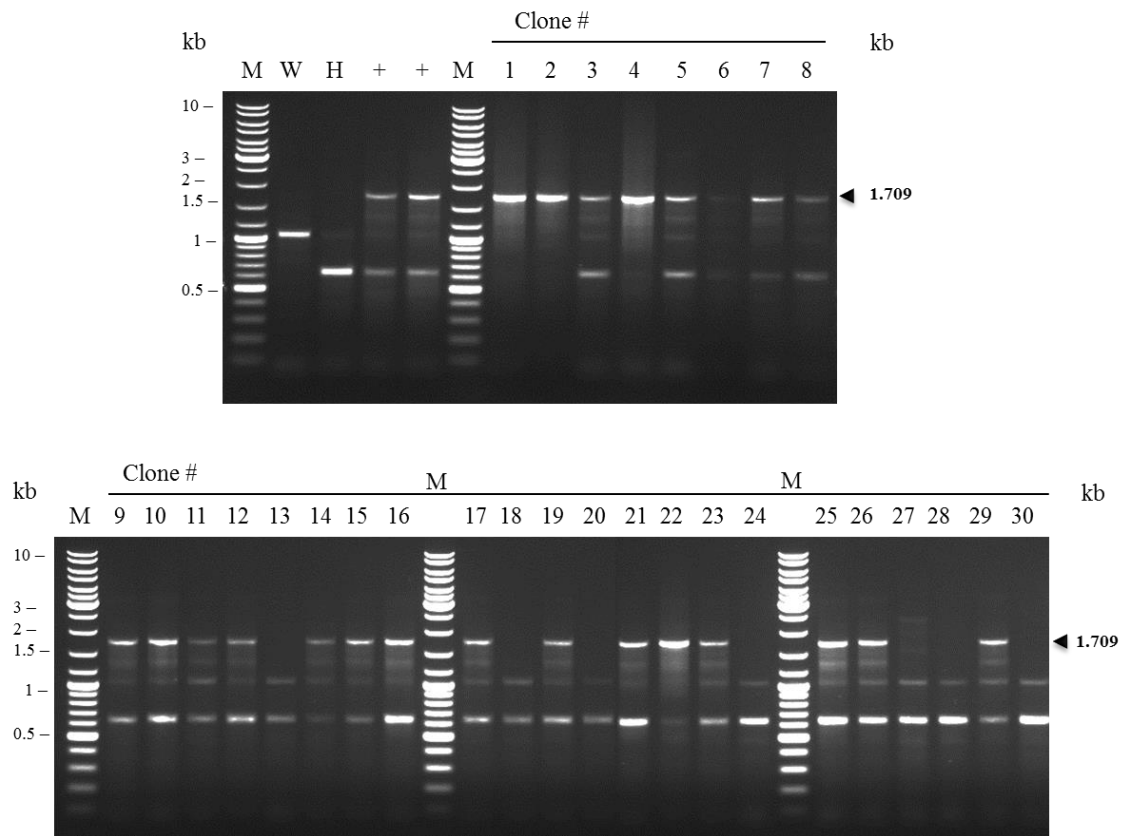
Figure 35 | PCR analysis for *AAVS1* 3'-junction transgene integration carried out on genomic DNA of eGFP-positive HeLa cell clones co-transduced with AdV.Δ2.donor-T-TS.F50 and AdV-TALEN[S1] L+R. The PCR reactions were performed with primers #331 and #332, specific for the 3'- junction of the *AAVS1* locus, for 30 clones. PCR mixtures using water (W) and mock-transduced HeLa cells (H) served as a negative control. Shuttle plasmid with the transgene served as a positive control (+). The positions and sizes (in kilo base pairs) of distinctive PCR products are indicated at the right. Water negative control displays a contamination band. M, Gene Ruler DNA Ladder Mix molecular weight marker (Fermentas).

3.4.2 Integration in the *AAVS1* 5'- junction

To confirm the full-length HR-mediated integration of the transgene, a PCR analysis for determining the integration in the *AAVS1* 5'- junction was conducted using primers #551 and #552 and should give rise to a 1665 bp amplicon. The experiment was conducted only for 30 clones.

The diagnostic PCR product appeared in 90% (i.e. 27 out of 30) of the clones, i.e., are positive for *AAVS1* 5'- junction transgene integration (**Fig. 36**). Of note, clones #6 and #8 appear not to possess any PCR product but they are indeed positive for integration confirmed by increased the DNA concentration (**Supplementary Fig. 41**).

56
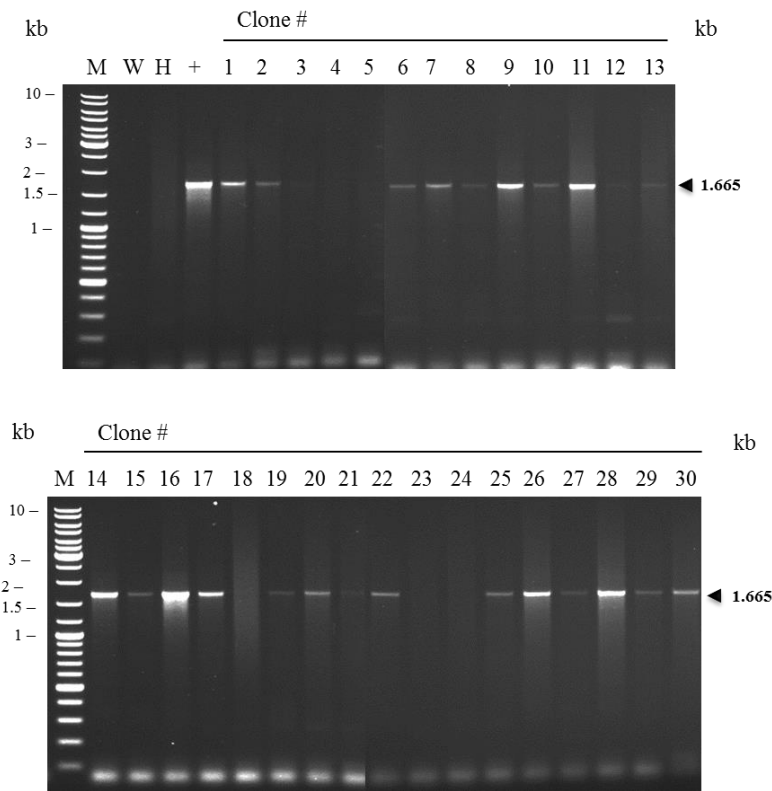
Figure 36 | PCR analysis for *AAVS1* 5'- junction transgene integration carried out on genomic DNA of eGFP-positive HeLa cell clones co-transduced with AdV.Δ2.donor-T-TS.F50 and AdV-TALEN[S1] L+R. The PCR reactions were performed with primers #551 and #552, specific for the 5'- junction of the *AAVS1* locus, for 30 clones. PCR mixtures using water (W) and mock-transduced HeLa cells (H) served as a negative control. Shuttle plasmid with the transgene served as a positive control (+). The positions and sizes (in kilo base pairs) of distinctive PCR products are indicated at the right. M, Gene Ruler DNA Ladder Mix molecular weight marker (Fermentas).

3.4.3 *eGFP* internal control

To serve as an internal control to monitor genomic DNA integrity and quantity, primers for *eGFP* reporter gene within the transgene were used (primers #111 and #112) and they should produce a 417 bp amplicon in eGFP-positive clones. A total of 118 clones were screened (**Fig. 37** and **Supplementary Fig. 42**) and the expected product was observed in 96.6% (i.e. 114 out of 118) clones.
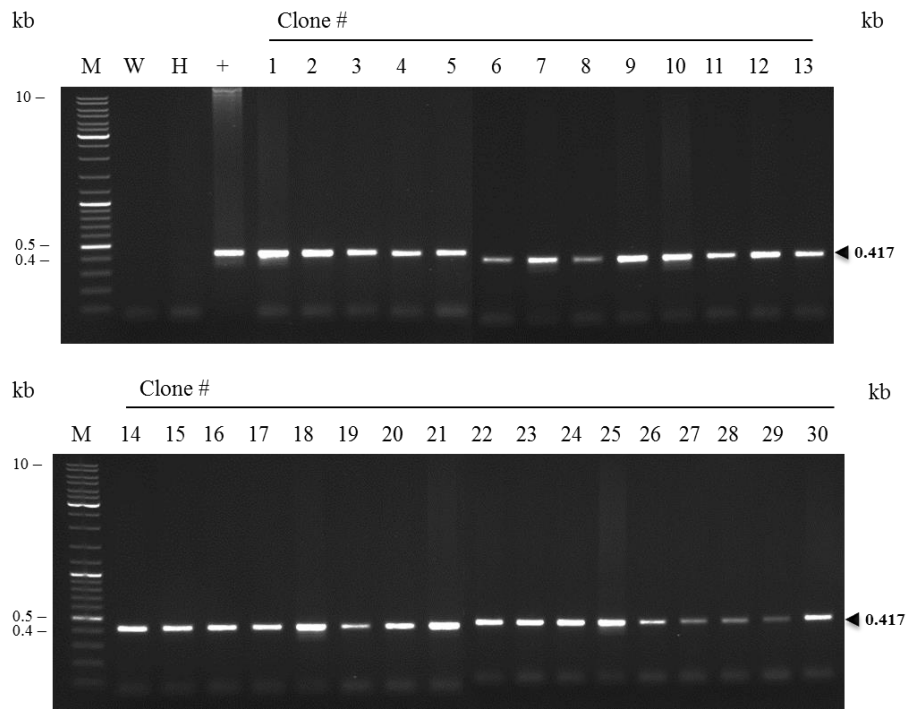
Figure 37 | PCR analysis for eGFP internal control carried out on genomic DNA of eGFP-positive HeLa cell clones co-transduced with AdV.Δ2.donor-T-TS.F50 and AdV-TALEN[S1] L+R. The PCR reactions were performed with primers #111 and #112, specific for the eGFP reporter gene, for 30 clones. PCR mixtures using water (W) and mock-transduced HeLa cells (H) served as a negative control. Shuttle plasmid with the transgene served as a positive control (+). The positions and sizes (in kilo base pairs) of distinctive PCR products are indicated at the right. M, Gene Ruler DNA Ladder Mix molecular weight marker (Fermentas).

Chapter 4. Discussion

Gene therapy, a relatively new and promising biomedical field, aims towards multiple goals such as the treatment of both inherited and acquired disease. Its concept is based on the principle that, with the introduction of an exogenous therapeutic gene into patient cells, its product will reverse, cure or slow down the progression of disease. With the continuing development of fields such as molecular biology, immunology and virology, gene therapy has gained the technological tools and insights needed to achieve its initial goals. A promising example of a valuable tool are designer endonucleases, such as the commonly used ZFNs, meganucleases and TALENs, with the latter showing the most promising attributes in terms of specificity and safety.

A relatively large number of human diseases have a well-identified genetic component and have been difficult to treat via standard medical interventions. The insertion or the correction of the missing or mutated genetic information holds the promise for a definite cure, for instance, via gene targeting approaches which involve the generation of a targeted DSB in the target cell genome and the delivery of a donor homologous DNA sequence.

TALENs are engineered phytopathogenic proteins merged to a FokI nuclease domain that, upon dimerization, produce a DSB in a specific DNA region in the genome. The high sequence specificity is given by the array of repeats in the TALE protein and these nucleases can be efficiently delivered to the target cell through viral vectors, namely by adenoviral vectors[108]. Adenoviral vectors show improved efficiency compared to the commonly used lentiviral vectors. Upon TALEN-mediated DSB, the cell's own repair systems are activated and promptly try to repair the chromosomal break, specifically via the NHEJ repair system, that joins the free ends in an error-prone fashion, and HR that, given a homologous exogenous template, will recombine the two molecules within the homology regions, fixing the break without major mutations or rearrangements.

Since the establishment of custom-designed TALENs, further research has been made in order to enhance its robustness, specificity, safety and also broaden the disease spectrum that can be effectively targeted by these nucleases. Investigations aiming to enhance the recombination with the homologous donor in the target cells have also been made. The donor sequence has been studied in order of its homology

content, but not much mechanistic studies have been done in order to enhance the recombination process.

The work of this project focused on the hypothesis that the topology of the homologous donor template affects the recombination and, by that, integration of the transgene into the targeted genome location. To test this hypothesis, three different donor DNA molecule topologies were designed and investigated: a circular topology, a linear topology and a secondary structure topology which recruits cellular DNA repair factors. To this end, donor DNA templates were flanked by (i) FRT sites to generate circular DNA molecules in target cells upon FLPe recombinase expression, (ii) T-TS site to generate linear DNA molecules in target cells in the presence of T-TS-specific TALENs or (iii) secondary structure-forming AAV-ITR sequences. The donors were delivered via second-generation fiber 50-modified human adenoviral vectors to HeLa cells and DMD myoblasts.

4.1 Molecular characterization of the AdV donor vectors

The different donors were constructed based on shuttle plasmids and the final full-length donor vectors were assembled by using a modified AdEasy system. In each step, the confirmation of the molecular integrity of the adenoviral vector DNA donor constructions were confirmed by restriction fragment length analysis.

The FLPe-mediated circularization of donor DNA templates delivered by AdV.Δ2.donor-FRT.F50 was investigated by deploying a PCR experimental set-up. These experiments established FLPe-mediated donor DNA circularization in target cells and confirmed by the presence of the discriminative amplicon. By using the same PCR-based experimental setup, the TALEN-dependent release of the donor DNA from the context of AdV.Δ2.donor-T-TS.F50 genomes could also be confirmed.

4.2 Vector transductions experiments

Adenoviral vectors are among the most efficient gene delivery vehicles for transgenes in gene therapy protocols, hence the experiment started by investigating the specificity and the accuracy of TALEN-mediated transgene recombination of the

different AdV donor DNA formats with the *AAVS1* target locus in human cells. To this end, HeLa cells were transduced with the viral vectors, as this cells display a high degree of genetic instability and, therefore, spontaneous chromosomal DSBs, being a more stringent cellular model system on which to evaluate HR-mediated gene targeting. The transduction of the donors alone established background levels of exogenous DNA genomic integration.

The initial transduction efficiency, determined by direct fluorescence microscopy, was high in cultures exposed to the various donor DNA-carrying adenoviral vectors as well as in those co-transduced with TALEN-encoding adenoviral vectors at low MOIs. The initial transduction frequencies were determined by flow cytometry analysis of eGFP and DsRed reporter gene expression, which mark donor DNA- and TALEN-positive cells, respectively. At the beginning of the experiment (i.e. at three days post-transduction), the HeLa cell cultures transduced with the adenoviral vector donors alone had similar frequencies of reporter-positive cells frequencies, as well in the combination with the TALENs. A difference in percentage of positive cells was evident between these two conditions where there was an apparent lowering of these percentages, possibly attributed to an overload of viral particles in the target cells.

The percentage of eGFP-positive cells is, in these experiments, used as a measure of the frequencies of chromosomal integration events as, by TALEN-mediated recombination, the transgene containing the reporter eGFP gene is expected to be integrated on the *AAVS1* locus of the target cells genome.

At the beginning of the transduction, the percentage of eGFP positive cells was high in all of the conditions tested but quickly decreased following sub-culturing of the transduced cells. The rapid decline in the frequencies of donor DNA-positive cells in cultures not exposed to TALEN expression was to be expected as, without the TALEN-mediated recombination, the transgene will remain in an episomal form and will eventually will be lost with the consecutives cell divisions. In contrast to this, cultures incubated with donor DNA- and TALEN-containing vectors, some of the transgene molecules should become incorporated into the host cell chromosomal DNA either in a random or targeted fashion and, as a result, is expected to be passed to daughter cells leading to stable and long-term transgene persistence in these cells.

By the end of the transduction experiment monitoring (i.e. at twenty-eight days post-transduction), we could observe that the frequency of eGFP-positive cells in the cultures transduced with the adenoviral vector donors alone reached very low levels (0-

0.08%), which can be considered as basal or background levels of the assay under these experimental conditions. In contrast to this, the experimental conditions in which the cultures were co-transduced with donor DNA- and TALEN-containing adenoviral vector particles had higher percentages of eGFP-positive cells (up to 0.49%). Because at this time point the AdVs viral particles were cleared and the episomal donor forms should have been diluted out, the remaining reporter positive cells should contain chromosomally integrated transgenes. Interestingly, the percentage of eGFP-positive cells in long-term cultures that were initially exposed to AdV.Δ2.donor-T-TS.F50 in combination with the TALENs was 1.8-fold higher than when the control donor vector AdV.Δ2.donor-FRT.F50 was used together with TALENs. Based on the flow cytometry dot blot, showing a well-defined eGFP-positive cell population together with the fact that episomal DNA is not expected to be present after many cell divisions, this population of compact eGFP-positive cells should represent cells in which the transgene with the reporter eGFP is stably integrated in the cell genome in a random and *AAVS1*-targeted fashion. Also, because the frequency of DsRed-positive cells, initially transduced with the TALEN-encoding adenoviral vectors, is close to 0%,TALEN-mediated DSB activity is not expected in these long-term cultures being, therefore, the numbers/values measured the result of integration events occurring early after transduction. Importantly, the initial transduction levels were similar among the different experimental settings allowing drawing a relationship between each of these settings and the corresponding stable transduction levels. Indeed, the similar initial transduction levels permits concluding that TALEN-meditated excision of donor DNA templates from the context of AdV.Δ2.donor-T-TS.F50 genomes results in increased frequencies of stably transduced HeLa cells.

In order to strengthen these data, a second independent transduction experiment was conducted. In this experiment we focused on using the donor and TALEN-encoding vectors AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN[S1] L+R, respectively. The transductions were once again conducted in HeLa cell cultures after which the frequencies of eGFP-positive cells were monitored for twenty-seven days by flow cytometry. In addition, to have more statistical power, the number of cells transduced was increased.

The initial transduction levels were high, similarly to those corresponding to the first transduction experiment, with the AdV.Δ2.donor-T-TS.F50 in combination with the TALENs having a lower initial frequency of eGFP-positive cells. A similar decrease of the reporter gene was observed within the monitored time period, reaching close to 0% in the donor alone (basal levels). A similar relatively high frequency of eGFP-positive

stably transduced cells appeared in cultures initially incubated with donor DNA- and TALEN-containing vectors, with the flow cytometry analysis showing 0.83% of eGFP-positive cells in these long-term cultures. Although the donor T-TS-negative adenoviral vector control was not taken alongside in this second experiment, the results are consistent with the previous conclusion that the "free" linear donor DNA topology leads to more efficient chromosomal insertion of a transgene payload following site-specific DSB formation.

One of the main objectives of the research carried out at the MCB group is, by means of gene therapy tools, correct the dystrophin gene in a Duchenne muscular dystrophy model to, at a later stage, translate the gained knowledge into the context of an experimental DMD therapy based on autologous transplantation of *DMD* gene-corrected patient cells.

A third transduction experiment was conducted in DMD myoblasts. These cells were co-transduced with AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN[S1] L+R or with control, T-TS-negative, AdV.Δ2.donor-eGFP.F50 and AdV.Δ2.TALEN[S1] L+R . Negative controls were provided by transducing parallel DMD cultures with the donor DNA-containing vectors alone. As in the previous experiments, the initial transduction levels were high as revealed by flow cytometric measurements of reporter-positive cells. The myoblasts were sub-cultured and monitored for thirty-five days, at the end of which the percentage of eGFP-positive cells was accessed by direct fluorescence microscopy and by flow cytometry analysis.

At thirty-five days post-transduction, it was distinctively visible that the percentage of eGFP-positive cells in various microscopic fields was higher in the cultures initially exposed to the combination of AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN[S1] L+R than in those incubated with the control, T-TS-negative, vector AdV.Δ2.donor-eGFP.F50 and AdV.Δ2.TALEN[S1] L+R. This qualitative assessment was confirmed by flow cytometry analysis, which showed a 2.9-fold increase in the frequency of eGFP-positive cells in the experimental set-up in which linear "free-ended" donor DNA molecules are formed by TALEN-mediated excision. This value is somewhat higher than that the measured in the first HeLa cell transduction experiment indicating a higher transgene integration frequency in these cells.

4.3 Molecular analysis of targeted transgene integration events

Among the aims of the project was to determine, not only the impact of the different donor topologies on the transgene integration efficiencies, but also the percentage of targeted insertions as well. Therefore, we investigated whether the entire transgene is indeed incorporated at our target sequence of choice (the human *AAVS1* locus).To this end, a PCR screening was conducted to determine the integrity of the *AAVS1* 3'-junction and 5'-junction in the eGFP-positive sorted HeLa clones derived from cultures initially transduced with AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN[S1] L+R. In this analysis, we used different sets of primers, designed to yield amplicons diagnostic for the HR-derived junctions between the donor and genomic target DNA. A PCR targeting the eGFP reporter gene was used as an internal control for transgene DNA integrity and quantity. HR-derived targeted insertion events yield specific PCR products corresponding to the donor DNA-*AAVS1* 3' junction and the donor DNA-*AAVS1* 5' junction

Of the one hundred and eighteen eGFP-positive sorted HeLa cell clones, twenty-two (19%) presented no distinguishing band being *AAVS1* 3'-junction negative.

Figure 38 | Schematic overview of the HeLa cell clones integration profiles assayed by PCR screening. A total of 118 eGFP-positive clones were analyzed using primers designed to amplify the 3'-junction of the *AAVS1* locus and the eGFP gene. The analyzed clones were generated by HeLa cell culture transductions with AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN[S1] L+R.

Next, we analyzed thirty clones with the primer set targeting the donor DNA-*AAVS1* 5'-junction. Of those, seven clones (17%) had been shown to be *AAVS1* 3'-junction negative and three clones (3%) were *AAVS1* 5'-junction negative. In addition, two clones (7%) were negative for both extremities and, importantly. twenty two clones (73%) were positive for both donor DNA-*AAVS1* junction extremities indicating that the majority of stably transduced cells contain the transgene accurately targeted at the *AAVS1* locus.

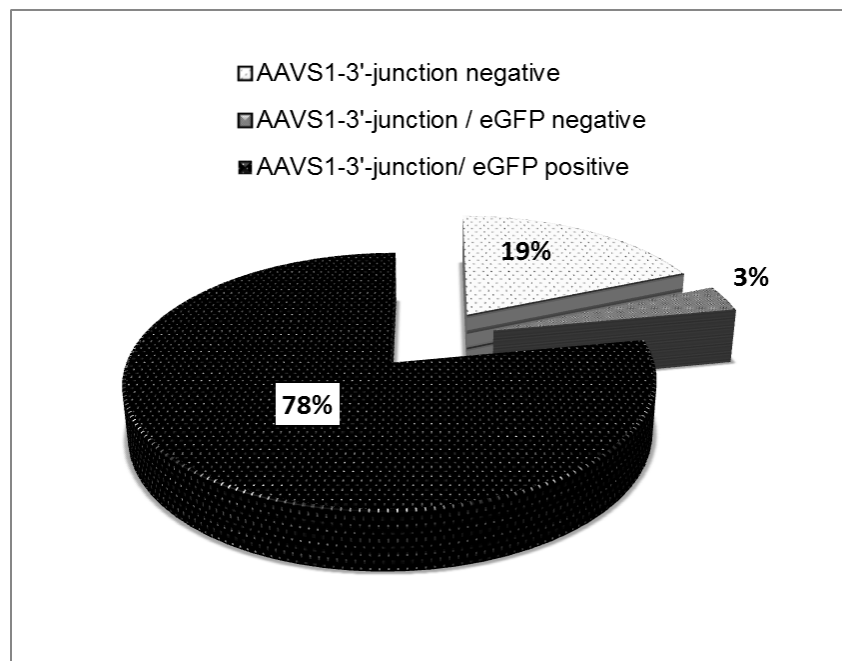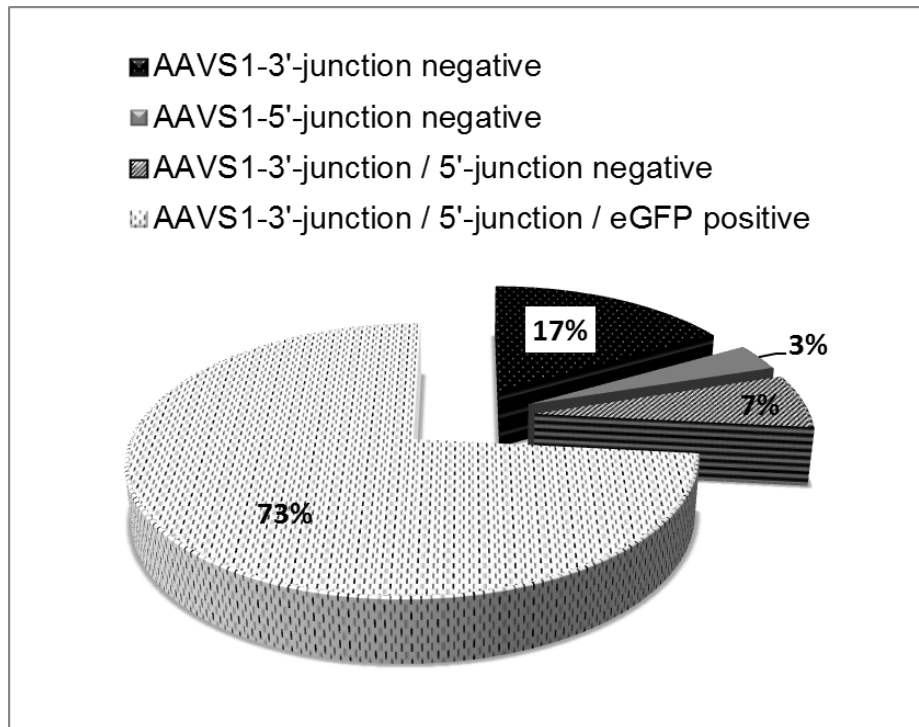**Figure 39** | Schematic overview of the HeLa cell clones integration profiles assayed by PCR screening on both junctions. A total of 30 eGFP-positive clones were analyzed by using primers designed to amplify the 3'-junction, the 5'-junction of the *AAVS1* locus and the eGFP gene. The analyzed clones were generated by HeLa cell culture transductions with AdV.Δ2.donor-T-TS.F50 and AdV.Δ2.TALEN[S1] L+R.

Chapter 5. Conclusion and final remarks

Although the ability to stably transduce efficiently different target cell types derives in part from the efficiency with which the viral vector coating allows target cell entry, the genome structure might also affect the efficiency of the chromosomal DNA integration process. This integration process is, perhaps, the most important aspect to improve towards a genetic therapy based on new gene targeting principles. This view is supported by the fact that efficient vector DNA entry into target cells can nowadays be easily achieved by modifying viral capsids and coats to allow the interaction of vector particles with specific cell surface receptors. Indeed, chromosomal integration of exogenous DNA and, more specifically, targeted integration into the genomic DNA of therapeutically relevant primary cells is currently a rather inefficient process.

The report presented in this thesis constitutes a mechanistic study on the influence of the homologous donor DNA template topology on the specificity and the accuracy of gene targeting following adenoviral vector-mediated gene transfer. We designed and constructed three different donor DNA topologies: a linear-, a circular- and a secondary structure-forming one that were delivered into human cells by second-generation adenoviral vectors.

We successfully produced high-titer adenoviral vector preparations with the donor DNA templates containing intact FRT, T-TS and AAV-ITR sites and used them subsequently in transduction experiments. The resulting data provided experimental evidence for the susceptibility of adenoviral vector DNA to the catalytic activities of site-specific FLPe recombinase and TALEN proteins. The latter finding opens news research possibilities based on the manipulation of adenoviral vector genomes in living cells. Most importantly, through the present work, it was found that TALEN-mediated excision of donor DNA from the context of AdV genomes leads to increased stable transduction levels with the majority of DNA integrants having precise HR-derived junctions at both ends. Independent transduction experiments as well as the testing of this new approach in different cell types strengthen this conclusion

Taken together, the results presented herein established that the donor DNA topology can indeed impact the outcome of genome editing based on engineered sequence-specific nucleases.

Chapter 6. References

1. Kay, M.A. State-of-the-art gene-based therapies: the road ahead. *Nature Reviews Genetics* **12**, 316-328 (2011).
2. Goncalves, M.A. A concise peer into the background, initial thoughts and practices of human gene therapy. *BioEssays : news and reviews in molecular, cellular and developmental biology* **27**, 506-517 (2005).
3. Somia, N. & Verma, I.M. Gene therapy: trials and tribulations. *Nature Reviews Genetics* **1**, 91-99 (2000).
4. Kay, M.A. & Woo, S.L.C. Gene-therapy for metabolic disorders *Trends Genet.* **10**, 253-257 (1994).
5. Isner, J.M. Myocardial gene therapy. *Nature* **415**, 234-239 (2002).
6. Lowenstein, P.R. Why are we doing so much cancer gene therapy? Disentangling the scientific basis from the origins of gene therapy. *Gene Ther.* **4**, 755-756 (1997).
7. Baekelandt, V., De Strooper, B., Nuttin, B. & Debyser, Z. Gene therapeutic strategies for neurodegenerative diseases. *Curr. Opin. Mol. Ther.* **2**, 540-554 (2000).
8. Bunnell, B.A. & Morgan, R.A. Gene therapy for infectious diseases. *Clin. Microbiol. Rev.* **11**, 42-+ (1998).
9. Lasaro, M.O. & Ertl, H.C.J. New insights on adenovirus as vaccine vectors. *Mol. Ther.* **17**, 1333-1339 (2009).
10. Thomas, C.E., Ehrhardt, A. & Kay, M.A. Progress and problems with the use of viral vectors for gene therapy. *Nature Reviews Genetics* **4**, 346-358 (2003).
11. Odom, G.L., Gregorevic, P. & Chamberlain, J.S. Viral-mediated gene therapy for the muscular dystrophies: successes, limitations and recent advances. *Biochim Biophys Acta.* **1772**, 243-262 (2007).
12. Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics* **11**, 636-646 (2010).
13. Bogdanove, A.J. & Voytas, D.F. TAL effectors: customizable proteins for DNA targeting. *Science* **333**, 1843-1846 (2011).
14. Vasquez, K.M., Marburger, K., Intody, Z. & Wilson, J.H. Manipulating the mammalian genome by homologous recombination. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8403-8410 (2001).
15. Humbert, O., Davis, L. & Maizels, N. Targeted gene therapies: tools, applications, optimization. *Crit. Rev. Biochem. Mol. Biol.* **47**, 264-281 (2012).
16. Li, S. & Huang, L. Nonviral gene therapy: promises and challenges. *Gene Ther.* **7**, 31-34 (2000).
17. Al-Dosari, M.S. & Gao, X. Nonviral gene delivery: principle, limitations, and recent progress. *Aaps J.* **11**, 671-681 (2009).
18. Tatum, E.L. Molecular biology nucleic acids and future of medicine. *Perspect. Biol. Med.* **10**, 19-& (1966).
19. Rogers, S. Shope papilloma virus -a passenger in man and its significance to potential control of host genome *Nature* **212**, 1220-& (1966).
20. Rogers, S. & Pfuderer, P. Use of viruses as carriers of added genetic information. *Nature* **219**, 749-& (1968).
21. Kay, M.K., Glorioso, J. & Naldini, L. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nature Medicine* **7**, 33-40 (2001).
22. Li, Z.X. et al. Murine leukemia induced by retroviral gene marking. *Science* **296**, 497-497 (2002).

23. Rowe, W.P., Huebner, R.J., Gilmore, L.K., Parrott, R.H. & Ward, T.G. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc. Soc. Exp. Biol. Med.* **84**, 570-573 (1953).

24. Goncalves, M.A. & de Vries, A.A. Adenovirus: from foe to friend. *Reviews in medical virology* **16**, 167-186 (2006).

25. Russell, W.C. Update an adenovirus and its vectors. *Journal of General Virology* **81**, 2573-2604 (2000).

26. Vogels, R. et al. Replication-deficient human adenovirus type 35 vectors for gene transfer and vaccination: Efficient human cell infection and bypass of preexisting adenovirus immunity. *Journal of Virology* **77**, 8263-8271 (2003).

27. Kojaoghlanian, T., Flomenberg, P. & Horwitz, M.S. The impact of adenovirus infection on the immunocompromised host. *Reviews in medical virology* **13**, 155-171 (2003).

28. San Martin, C. & Burnett, R.M. Structural studies on adenoviruses. *Current topics in microbiology and immunology* **272**, 57-94 (2003).

29. Vellinga, J., Van der Heijdt, S. & Hoeben, R.C. The adenovirus capsid: major progress in minor proteins. *Journal of General Virology* **86**, 1581-1588 (2005).

30. Matthews, D.A. & Russell, W.C. Adenovirus core protein V is delivered by the invading virus to the nucleus of the infected cell and later in infection is associated with nucleoli. *Journal of General Virology* **79**, 1671-1675 (1998).

31. Webster, A., Russell, S., Talbot, P., Russell, W.C. & Kemp, G.D. Characterization of the adenovirus proteinase - substrate-specific. *Journal of General Virology* **70**, 3225-3234 (1989).

32. Flint, S.J. Principles of virology: molecular biology, pathogenesis, and control of animal viruses, Edn. 1st edition. (American Society Microbiology, 1994).

33. Bergelson, J.M. et al. Isolation of a common receptor for coxsackie B viruses and adenoviruses 2 and 5. *Science* **275**, 1320-1323 (1997).

34. Zhang, Y.M. & Bergelson, J.M. Adenovirus receptors. *Journal of Virology* **79**, 12125-12131 (2005).

35. Svensson, U. Role of vesicles during adenovirus-2 internalization into HeLa-cells. *Journal of Virology* **55**, 442-449 (1985).

36. Wang, K.N., Huang, S., Kapoor-Munshi, A. & Nemerow, G. Adenovirus internalization and infection require dynamin. *Journal of Virology* **72**, 3455-3458 (1998).

37. Trotman, L.C., Mosberger, N., Fornerod, M., Stidwill, R.P. & Greber, U.F. Import of adenovirus DNA involves the nuclear pore complex receptor CAN/Nup214 and histone H1. *Nat. Cell Biol.* **3**, 1092-1100 (2001).

38. Rekosh, D.M.K., Russell, W.C., Bellet, A.J.D. & Robinson, A.J. Identification of a protein linked to ends of adenovirus DNA. *Cell* **11**, 283-295 (1977).

39. Flint, J. & Shenk, T. Adenovirus E1A protein paradigm viral transactivator. *Annu. Rev. Genet.* **23**, 141-161 (1989).

40. Wang, M. et al. Adenoviral vector systems for gene therapy. *Gene Therapy and Molecular Biology* **9**, 291-300 (2005).

41. Hay, R.T., Freeman, A., Leith, I., Monaghan, A. & Webster, A. Molecular interactions during adenovirus DNA replication. *Current topics in microbiology and immunology* **199 (Pt 2)**, 31-48 (1995).

42. Bennett, E.M., Bennink, J.R., Yewdell, J.W. & Brodsky, F.M. Cutting edge: adenovirus E19 has two mechanisms for affecting class I MHC expression. *J. Immunol.* **162**, 5049-5052 (1999).

43. Leppard, K.N. E4 gene function in adenovirus, adenovirus vector and adeno-associated virus infections. *Journal of General Virology* **78**, 2131-2138 (1997).

44. Volpers, C. & Kochanek, S. Adenoviral vectors for gene transfer and therapy. *J. Gene. Med.* **6**, S164-S171 (2004).

45. Hearing, P., Samulski, R.J., Wishart, W.L. & Shenk, T. Identification of a repeated sequence element required for efficient encapsidation of the adenovirus type-5 chromosome *Journal of Virology* **61**, 2555-2558 (1987).

46. D'Halluin, J.C. Virus assembly. *Current topics in microbiology and immunology* **199 (Pt 1)**, 47-66 (1995).

47. Tollefson, A.E., Ryerse, J.S., Scaria, A., Hermiston, T.W. & Wold, W.S.M. The E3-11.6-kDa adenovirus death protein (ADP) is required for efficient cell death: characterization of cells infected with adp mutants. *Virology* **220**, 152-162 (1996).

48. Shiver, J.W. & Emini, E.A. Recent advances in the development of HIV-1 vaccines using replication-incompetent adenovirus vectors. *Annu. Rev. Med.* **55**, 355-372 (2004).

49. Sullivan, N.J. et al. Accelerated vaccination for Ebola virus haemorrhagic fever in non-human primates. *Nature* **424**, 681-684 (2003).

50. Vrancken Peeters, M.J., Perkins, A.L. & Kay, M.A. Method for multiple portal vein infusions in mice: quantitation of adenovirus-mediated hepatic gene transfer. *BioTechniques* **20**, 278-285 (1996).

51. McConnell, M.J. & Imperiale, M.J. Biology of adenovirus and its use as a vector for gene therapy. *Hum. Gene Ther.* **15**, 1022-1033 (2004).

52. Halbert, D.N., Cutt, J.R. & Shenk, T. Adenovirus early region-4 encodes functions required for efficient DNA-reoplication, late gene-expression, and host-cell shutoff. *Journal of Virology* **56**, 250-257 (1985).

53. Dobbelstein, M. Replicating adenoviruses in cancer therapy. *Curr.Top.Microbiol.Immunol.* **273**, 291-334 (2004).

54. Bramson, J.L., Graham, F.L. & Gauldie, J. The use of adenoviral vectors for gene-therapy and gene-transfer in-vivo *Current opinion in biotechnology* **6**, 590-595 (1995).

55. Graham, F.L. & Prevec, L. Manipulation of adenovirus vectors. *Methods in molecular biology* **7**, 109-128 (1991).

56. Graham, F.L. & Prevec, L. Methods for construction of adenovirus vectors. *Mol. Biotechnol.* **3**, 207-220 (1995).

57. Graham, F.L., Smiley, J., Russell, W.C. & Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type-5 *Journal of General Virology* **36**, 59-72 (1977).

58. Fallaux, F.J. et al. Characterization of 911: a new helper cell line for the titration and propagation of early region 1-deleted adenoviral vectors. *Hum. Gene Ther.* **7**, 215-222 (1996).

59. 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).

60. Morral, N., Oneal, W., Zhou, H.S., Langston, C. & Beaudet, A. Immune responses to reporter proteins and high viral dose limit duration of expression with adenoviral vectors: comparison of E2a wild type and E2a deleted vectors. *Hum. Gene Ther.* **8**, 1275-1286 (1997).

61. Yang, Y.P. et al. Cellular-immunity to viral-antigens limits E1-deleted adenoviruses for gene-therapy. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4407-4411 (1994).

62. Lochmuller, H. et al. Emergence of early region-1-containing replication-competent adenovirus in stocks of replication-defective adenovirus recombinants (delta-E1 plus delta-E3) during multiple passages in 293-cells   *Hum. Gene Ther.* **5**, 1485-1491 (1994).

63. Amalfitano, A. et al. Production and characterization of improved adenovirus vectors with the E1, E2b, and E3 genes deleted. *Journal of Virology* **72**, 926-933 (1998).

64. Lusky, M. et al. In vitro and in vivo biology of recombinant adenovirus vectors with E1, E1/E2A, or E1/E4 deleted. *Journal of Virology* **72**, 2022-2032 (1998).

65. Fallaux, F.J., van der Eb, A.J. & Hoeben, R.C. Who's afraid of replication-competent adenoviruses? *Gene Ther.* **6**, 709-712 (1999).

66. Morsy, M.A. et al. An adenoviral vector deleted for all viral coding sequences results in enhanced safety and extended expression of a leptin transgene. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7866-7871 (1998).

67. Steinwaerder, D.S., Carlson, C.A. & Lieber, A. Generation of adenovirus vectors devoid of all viral genes by recombination between inverted repeats. *Journal of Virology* **73**, 9303-9313 (1999).

68. Morsy, M.A. & Caskey, C.T. Expanded-capacity adenoviral vectors - the helper-dependent vectors. *Mol. Med. Today* **5**, 18-24 (1999).

69. Clemens, P.R. et al. In vivo muscle gene transfer of full-length dystrophin with an adenoviral vector that lacks all viral genes. *Gene Ther.* **3**, 965-972 (1996).

70. Palmer, D.J. & Ng, P. Helper-dependent adenoviral vectors for gene therapy. *Hum. Gene Ther.* **16**, 1-16 (2005).

71. Miyazawa, N. et al. Fiber swap between adenovirus subgroups B and C alters intracellular trafficking of adenovirus gene transfer vectors. *Journal of Virology* **73**, 6056-6065 (1999).

72. Kirby, I. et al. Mutations in the DG loop of adenovirus type 5 fiber knob protein abolish high-affinity binding to its cellular receptor CAR. *Journal of Virology* **73**, 9508-9514 (1999).

73. Magnusson, M.K., Hong, S.S., Boulanger, P. & Lindholm, L. Genetic retargeting of adenovirus: novel strategy employing "deknobbing" of the fiber. *Journal of Virology* **75**, 7280-7289 (2001).

74. Reynolds, P.N., Feng, M.Z. & Curiel, D.T. Chimeric viral vectors - the best of both worlds? *Mol. Med. Today* **5**, 25-31 (1999).

75. Fisher, K.J., Kelley, W.M., Burda, J.F. & Wilson, J.M. A novel adenovirus-adeno-associated virus hybrid vector that displays efficient rescue and delivery of the AAV genome. *Hum. Gene Ther.* **7**, 2079-2087 (1996).

76. Yant, S.R. et al. Transposition from a gutless adeno-transposon vector stabilizes transgene expression in vivo. *Nature biotechnology* **20**, 999-1005 (2002).

77. Jackson, S.P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071-1078 (2009).

78. Lieber, M.R. The mechanism of human nonhomologous DNA end joining. *J. Biol. Chem.* **283**, 1-5 (2008).

79. Liang, F., Han, M.G., Romanienko, P.J. & Jasin, M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 5172-5177 (1998).

80. Kass, E.M. & Jasin, M. Collaboration and competition between DNA double-strand break repair pathways. *FEBS letters* **584**, 3703-3708 (2010).

81. Waldman, A.S. Targeted homologous recombination in mammalian cells. *Critical Reviews in Oncology|Hematology* **12**, 49-64 (1992).

82. Capecchi, M.R. Altering the genome by homologous recombination. *Science* **244**, 1288-1292 (1989).

83. Itzhaki, J.E. & Porter, A.C.G. Targeted disruption of a human interferon-inducible gene detected by secretion of human growth-hormone. *Nucleic acids research* **19**, 3835-3842 (1991).

84. Rouet, P., Smih, F. & Jasin, M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian-cells *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6064-6068 (1994).

85. Smih, F., Rouet, P., Romanienko, P.J. & Jasin, M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic acids research* **23**, 5012-5019 (1995).

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

87.   Bibikova, M. et al. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* **21**, 289-297 (2001).

88.   Holliday, R. A mechanism for gene conversion in fungi. *Genet. Res.* **89**, 285-307 (2007).

89.   Meselson, M.S. & Radding, C.M. General model for genetic recombination *Proc. Natl. Acad. Sci. U. S. A.* **72**, 358-361 (1975).

90.   Moehle, E.A. et al. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3055-3060 (2007).

91.   Chen, F.Q. et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* **8**, 753-U796 (2011).

92.   Bogdanove, A.J., Schornack, S. & Lahaye, T. TAL effectors: finding plant genes for disease and defense. *Curr. Opin. Plant Biol.* **13**, 394-401 (2010).

93.   Scholze, H. & Boch, J. TAL effectors are remote controls for gene activation. *Curr. Opin. Microbiol.* **14**, 47-53 (2011).

94.   Boch, J. & Bonas, U. in Annual Review of Phytopathology, Vol. 48. (eds. N.K. VanAlfen, G. Bruening & J.E. Leach) 419-436 (Annual Reviews, Palo Alto; 2010).

95.   Kay, S., Hahn, S., Marois, E., Hause, G. & Bonas, U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**, 648-651 (2007).

96.   Mussolino, C. & Cathomen, T. TALE nucleases: tailored genome engineering made easy. *Current opinion in biotechnology* **23**, 644-650 (2012).

97.   Moscou, M.J. & Bogdanove, A.J. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science* **326**, 1501-1501 (2009).

98.   Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-1512 (2009).

99.   Bitinaite, J., Wah, D.A., Aggarwal, A.K. & Schildkraut, I. FokI dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10570-10575 (1998).

100.   Smith, J. et al. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic acids research* **28**, 3361-3369 (2000).

101.   Liu, J.Y. et al. Efficient and specific modifications of the drosophila genome by means of an easy TALEN strategy. *J. Genet. Genomics* **39**, 209-215 (2012).

102.   Huang, P. et al. Heritable gene targeting in zebrafish using customized TALENs. *Nature biotechnology* **29**, 699-700 (2011).

103.   Tesson, L. et al. Knockout rats generated by embryo microinjection of TALENs. *Nature biotechnology* **29**, 695-696 (2011).

104.   Osborn, M.J. et al. TALEN-based gene correction for epidermolysis bullosa. *Mol. Ther.* **21**, 1151-1159 (2013).

105.   Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nature biotechnology* **29**, 731-734 (2011).

106.   Miller, J.C. et al. A TALE nuclease architecture for efficient genome editing. *Nature biotechnology* **29**, 143-148 (2011).

107.   Mussolino, C. et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic acids research* **39**, 9283-9293 (2011).

108.   Holkers, M. et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic acids research* **41**, e63 (2013).

109.   Chen, K. & Gao, C. TALENs: customizable molecular DNA scissors for genome engineering of plants. *J. Genet. Genomics* **40**, 271-279 (2013).

110.   Yue, Y.P. & Duan, D.S. Development of multiple cloning site cis-vectors for recombinant adeno-associated virus production. *BioTechniques* **33**, 672-+ (2002).

111.   Janssen, J.M., Liu, J., Skokan, J., Goncalves, M. & de Vries, A.A.F. Development of an AdEasy-based system to produce first- and second-generation adenoviral vectors with tropism for CAR- or CD46-positive cells. *J. Gene. Med.* **15**, 1-11 (2013).

112. Havenga, M.J.E. 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).

113. Goncalves, M., Pau, M.G., de Vries, A.A.F. & Valerio, D. Generation of a high-capacity hybrid vector: packaging of recombinant adenoassociated virus replicative intermediates in adenovirus capsids overcomes the limited cloning capacity of adenoassociated virus vectors. *Virology* **288**, 236-246 (2001).

114. Cudre-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).
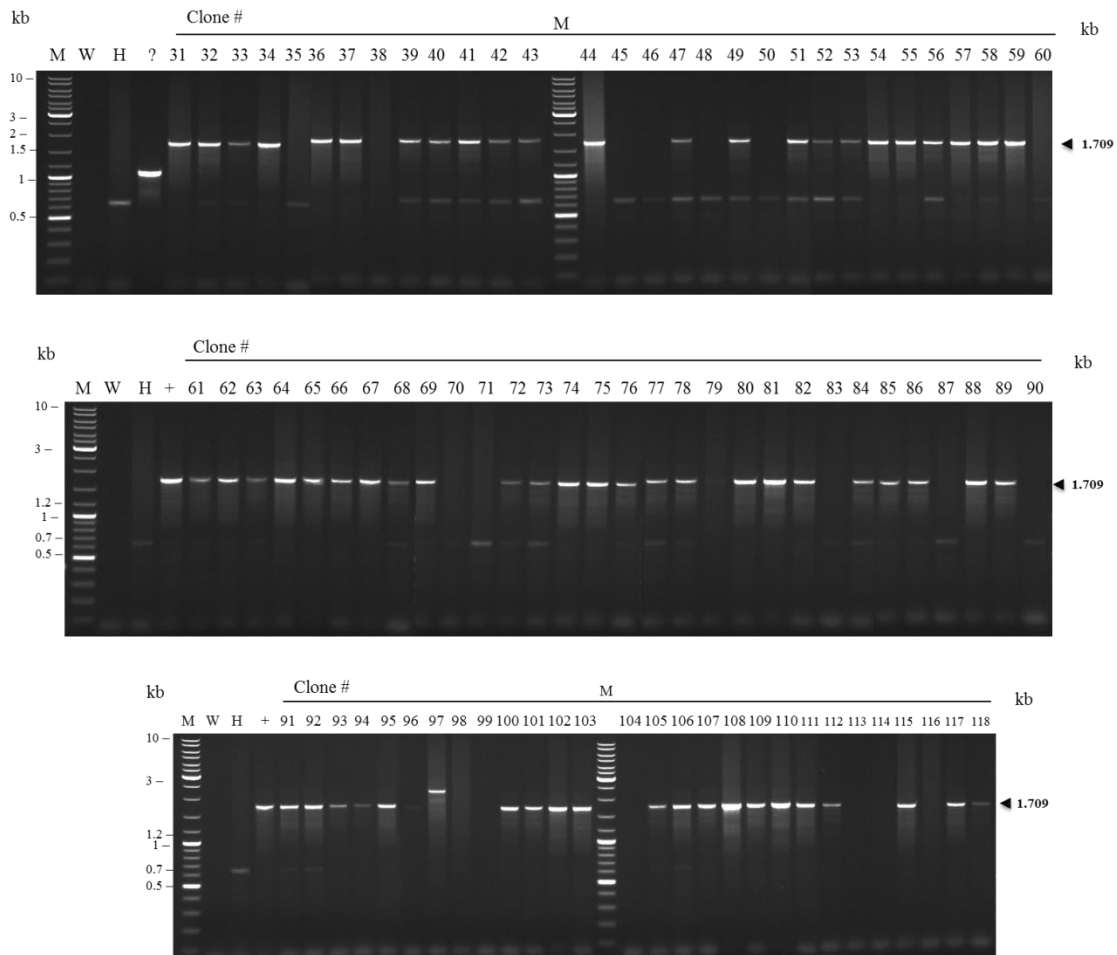
Chapter 7. Supplementary data

Figure 40 | PCR analysis for *AAVS1* 3'-junction transgene integration carried out on genomic DNA of eGFP-positive HeLa cell clones co-transduced with AdV.Δ2.donor-T-TS.F50 and AdV-TALEN[S1] L+R in the remaining clones. The PCR reactions were performed with primers #331 and #332, specific for the 3'- junction of the *AAVS1* locus, for the remaining 88 clones. PCR mixtures using water (W) and mock-transduced HeLa cells (H) served as a negative control. Shuttle plasmid with the transgene served as a positive control (+). The positions and sizes (in kilo base pairs) of distinctive PCR products are indicated at the right. Water negative control displays a contamination band. M, Gene Ruler DNA Ladder Mix molecular weight marker (Fermentas).
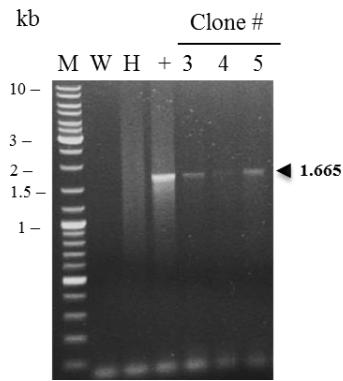
Figure 41 | PCR experiment for *AAVS1* 5'-end targeted integration with 4µL of DNA. Repetition of the PCR experiment for clone #6 through 8. Water (lane 1); mock (lane 2); positive control (lane 3).
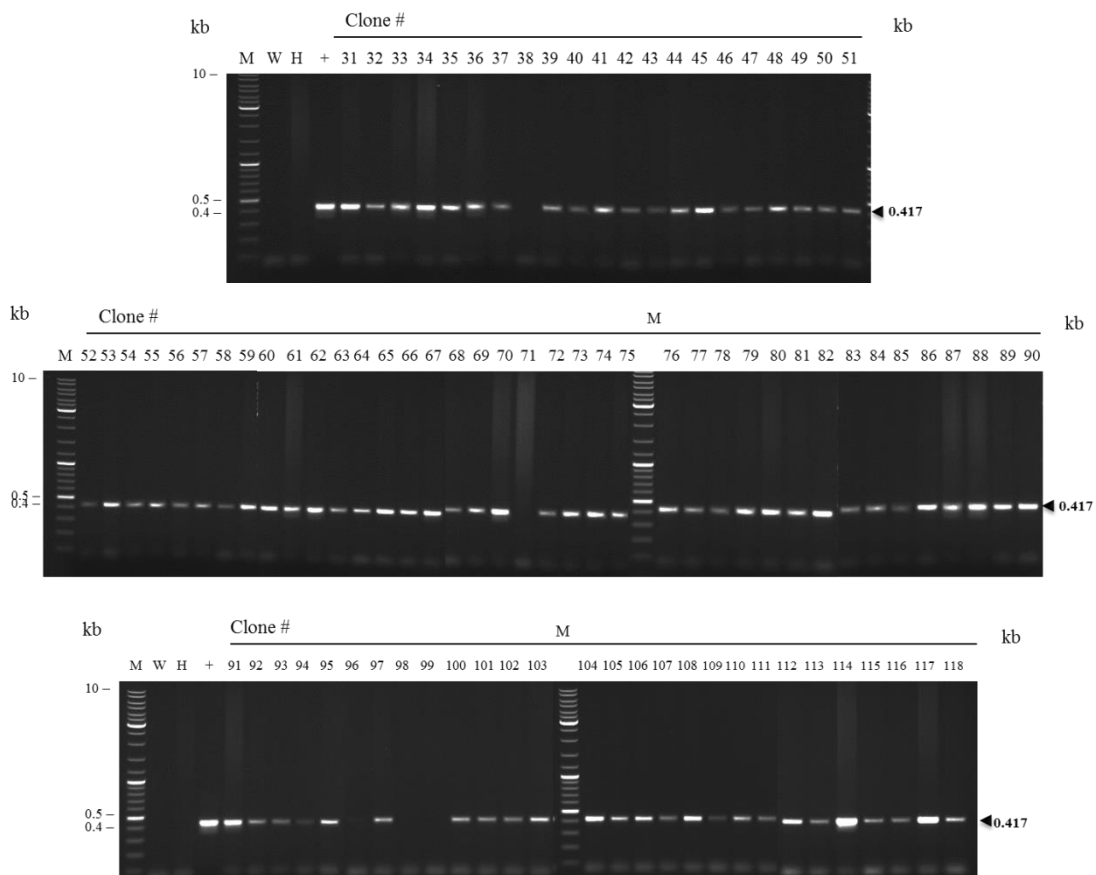

Figure 42 | PCR analysis for eGFP internal control carried out on genomic DNA of eGFP-positive HeLa cell clones co-transduced with AdV.Δ2.donor-T-TS.F50 and AdV-TALEN[S1] L+R in the remaining clones. The PCR reactions were performed with primers #111 and #112, specific for the eGFP reporter gene, for the remaining 88 clones. PCR mixtures using water (W) and mock-transduced HeLa cells (H) served as a negative control. Shuttle plasmid with the transgene served as a positive control (+). The positions and sizes (in kilo base pairs) of distinctive PCR products are indicated at the right. M, Gene Ruler DNA Ladder Mix molecular weight marker (Fermentas).