PhD degree in Systems Medicine

curriculum in Human Genetics

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples "Federico II"

Settore disciplinare: MED/03

THERAPEUTIC GENOME EDITING IN RETINA AND LIVER

Manel Lladó Santaeularia

Tigem or Ceinge, Naples

Matricola n. R11477

Supervisor: *Prof. Alberto Auricchio*

Dept. Of Advanced Biomedicine, Federico II University,

TIGEM, Naples, Italy

Anno accademico 2019-2020

ACKNOWLEDGMENTS

To my family, for encouraging me to follow my dreams and for supporting me even when those dreams took me away from them. To my parents for giving me everything they could and for making me the man I am today. And to my sister Sandra, my touchstone and the one that always keeps me in check. To my grandparents Manel and Maria, for their unconditional love and for making me the happiest kid in the world while they helped my parents raise me.

To my mentor, Alberto Auricchio, for always believing I can do better, for being an inspiration both as a scientist and as a person, and for seeing the potential in me when sometimes I didn't. For trusting me to do things that were supposed to be beyond my capabilities, and always believing I had what it took to do anything he asked of me. For making me grow and teaching so much I can never thank him enough.

To Lele Castello, for his invaluable support as a teacher, a scientist, an advisor and more importantly a friend, and as literally anything I have ever needed. For finding any and all my mistakes in my presentations and teaching me about the importance of details and how to properly spell "photoreceptors".

To FDA, Fabio The King, my best friend, for all the amazing times we have spent together both in the lab and outside. For being my neapolitan brother forever.

To Renato Minopoli, the world's best technician, for being the heart and soul of the lab and for always finding new ways to make everyone else smile. For helping me believe in myself and for all the fun we have had together.

To Maria Sirakov, my neapolitan auntie, for her support and her advice in the beginning of this journey, and for being there until the end of it.

To Carolina Iodice, for her perseverance, her patience with my messiness and for taking the time to help in any way she can, and for making injection sessions fun and thrilling at the same time.

To all the past and present lab members who have contributed in any way to making me grow as a scientist and as a person, and who have left an imprint forever in me. Andrea, Patrizia, Ivana, Miriam, Rita, Margherita, Hristiana, Federica, Eugenio, Arjun and many more.

To Knut Stieger and Davide Cacchiarelli for their support and advice during my PhD and for being both a professional inspiration and great people.

To Graciana Diez-Roux for always being there before I even realized I need her and for always being available to give a hand.

INDEX

ABBREVIATIONS

AAP: assembly-activating protein

AAV: Adeno-associated virus

AdRP: Autosomal Dominant RP

Alb: murine albumin gene

ARSB: arylsulfatase B

bGH: bovine growth hormone polyA

bp: base pairs

CMV: cytomegalovirus promoter

Cbh: Chicken beta-actin hybrid promoter

CjCas9: Campylobacter jejunii Cas9

CNS: Central Nervous System

DEL: deletion

DNA: deoxyribonucleic acid

DSB: double-strand break

DsRed: Dicosoma red fluorescent protein

EDTA: etilendiaminotetraacetic acid

EGFP: Enhanced green fluorescent protein

ERG: electroretinogram

ERT: enzyme replacement therapy

FACS: fluorescence activated cell sorting

Fwd: forward

GAGs: glycosaminoglycans

GC: genome copies

GCL: gangliar cell layer

GE: Genome Editing

gRNA: guide RNA

GT: Gene Therapy

HEK293: human embryonic kidney 293

HDR: homology-directed repair

HITI: homology-independent targeted integration

HLP: hybrid liver promoter

INDELs: insertions and deletion

INL: inner nuclear layer

INS: insertion

IRBP: interphotoreceptor retinol-binding protein

IRES: internal ribosomal entry site

ITR: Inverted Terminal Repeats

Kozak: consensus kozak sequence

KO: knockout

LSD: lysosomal storage disease

MMEJ: microhomology-mediated end joining

MPS: Mucopolysaccharidosis

MPSVI: Mucopolysaccharidosis type VI

mRNA: messenger RNA

mRho: murine rhodopsin gene

NGS: next-generation sequencing

NHEJ: non-homologous end joining

NmCas9: Neisseria meningitidis Cas9

ONL: outer nuclear layer

ORF: open reading frame

PAM: protospacer-adjacent motif

PCR: polymerase chain reaction

PFA: paraformaldehyde

PLR: pupillary light response

polyA: polyadenylation signal

PR: photoreceptor

pRho: porcine rhodopsin gene

Rev: reverse

RHO: human rhodopsin gene

RHO: rhodopsin protein

RP: Retinitis Pigmentosa

RPE: retinal pigmented epithelium

RT-PCR: real-time PCR.

shpolyA: short synthetic polyA

ssDNA: single-stranded DNA

SpCas9: Steptococcus pyogenes Cas9

TALENs: Transcription Activator-Like Effector Nucleases

TBG: thyroxin-binding globulin promoter

TIDE: Tracking of INDELs by Decomposition

tracrRNA: trans-activating crRNA

T7E: T7 Endonuclease I

UTR: untranslated region

Vg: viral genome

VQR-Cas9-HF1: VQR SpCas9 High Fidelity

WT: wildtype

ZFNs: Zinc-Finger Nucleases

ABSTRACT

In vivo gene therapy with adeno-associated viral (AAV) vectors has been successful at treating several inherited diseases, specifically those caused by loss of function mutations which require transfer of a correct copy of a gene. This would not benefit dominant diseases due to gain of function mutations which produce toxic protein products. In addition, since AAV genomes persist as episomes in target cells, AAV mediated transgene expression might be short lived in tissues where cell proliferation occurs when newborn or after damage, like for example the liver. To overcome these challenges, I have developed AAV-based therapeutic approaches which use genome editing to introduce stable modifications at specific genomic loci. First, an allele-specific approach which targets the Rhodopsin P347S dominant mutation was developed and tested both *in vitro* and *in vivo*. I achieved allele-specific targeting of human P347S rhodopsin, which reduced mRNA levels and improved retinal electrical function in a mouse model of autosomal dominant retinitis pigmentosa. Second, I developed a mutation- and homology-independent targeted integration (HITI) approach for gene correction in photoreceptors. I demonstrated feasibility of this approach in mouse and pig photoreceptors using a reporter gene and characterized on-target precision of HITI in the murine rhodopsin locus. I then tested the therapeutic potential of this approach in a mouse model of autosomal dominant retinitis pigmentosa and observed mild and transient improvement of retinal function in treated eyes, which suggests that the levels of editing obtained need optimization. Third, I developed a HITI approach for expressing therapeutic genes from the liver by targeting the albumin locus, which is highly transcribed in hepatocytes. I demonstrated feasibility and efficiency of this approach using a reporter gene, and characterized on-target precision of HITI, as well as off-target integration due to Cas9 cleavage. I then tested the therapeutic potential of the integration of a copy of the human arylsulfatase B (ARSB) gene, which is mutated in a rare lysosomal storage disease, mucopolysaccharidosis type VI (MPS VI), in the albumin locus in the liver of newborn MPSVI mice. I demonstrated that this approach achieves stable expression of ARSB at levels that reduce glucosaminoglycan (GAG) urinary secretion, one of

the main readouts of MPSVI phenotype. This stable expression of ARSB is contrary to the decrease of transgene expression observed in neonatal MPSVI mice injected with the same dose of a conventional gene therapy vector, thus overcoming the potential loss of transgene expression caused by hepatocyte proliferation.

Overall, I have developed different genome editing approaches for conditions that are inherited as either dominant or recessive. I have tested these approaches in two relevant tissues for gene therapy like retina and liver and shown the potential to provide AAV with persistent transgene expression in proliferating tissues like the newborn liver.

INTRODUCTION

Gene Therapy:

Gene Therapy (GT) is the treatment or cure of a disease by expressing an exogenous gene or silencing a particular transcript in the target cells. More recently, GT has also used exogenous sequences to modify the endogenous DNA sequences of the target cells in order to achieve genomic correction. In all cases, the most important factor for successful GT is the efficient delivery of nucleic acids to the target cells, which is known as gene transfer [1]. Gene transfer can be performed *in vivo*, directly to the target tissues in the animal model or patient; or *ex vivo*, in cultured cells isolated from said animal or patient and re-implanted to it after gene transfer [2]. Gene transfer can be achieved using vectors from different origins. The main classification is between viral-derived vectors and non-viral vectors [3]. Viral-derived vectors exploit the ability of viral capsids to infect cells and transfer DNA to their nucleus, but avoid viral replication by eliminating the viral genes from the viral genome (vg) and leaving only those sequences that are important for the packaging of the vg into the viral capsid [1]. Several kinds of viral vectors have been used over the years. The first GT experiments used vectors derived from the Moloney murine leukemia retrovirus, which as all retrovirus has a linear single-stranded RNA 7-10Kb genome that can be integrated in the genome of the target cell. This conferred long-term expression of the transgene, but the affinity for integration in transcriptionally active genes, including proto-oncogenes [4, 5], was an important drawback for retroviral vectors, since insertional mutagenesis, as well as insertional activation of proximal genes, were reported to cause cancer in patients enrolled in a gene therapy clinical trial for X-linked severe combined immunodeficiency [1, 6]. In 2016, Strimvelis™, a retroviral vector-based gene therapy product for Severe Combined Immunodeficiency due to Adenosine Deaminase deficiency was approved for clinical use [7, 8]

In later years, lentiviral vectors were derived from the human immunodeficiency virus type-1 (HIV-1). Lentiviral vectors have a packaging capacity of 8-10 Kb and an RNA genome similar to retroviral vectors, but they have not been reported to cause insertional carcinogenesis, since their integration seems to favor different sequences compared to retroviral vectors. Additionally, lentiviral vectors can infect both dividing and non-dividing cells, while retroviral vectors can only infect dividing cells [2]. Lentiviral vectors are mainly used for ex vivo applications, particularly gene transfer into CD34 + hematopoietic stem cells, which has been used to treat several genetic diseases, including β-thalassemia [9], X-linked adrenoleukodystrophy [10], metachromatic leukodystrophy [11, 12], and Wiskott-Aldrich Syndrome [13]. Several clinical trials are ongoing, and no adverse events related to the vector have been reported in these trials [2, 14].

Adenoviral vectors were derived from adenovirus, which is a family of DNA viruses with a genome of around 36Kb that commonly infect humans, leading to inexistent or very mild symptoms. Several generations of adenoviral vectors were generated by successively deleting parts of the adenovirus genome, which allows for a very high packaging capacity. Adenoviruses are divided in serotypes, which have different membrane proteins that confer them with natural tropism for various tissues. However, most adenoviral vectors are derived from Adenovirus 5, which very efficiently transduces the liver after intravenous injection [15]. More than 500 gene therapy trials have been conducted using adenoviral vectors, most of them for cancer applications but also for treatment of monogenic diseases [8]. The main limitation of adenoviral vectors is their high immunogenicity, which induces a strong innate immune response [16], and the presence of pre-existing immunity against them, due to their very high prevalence in the normal population [15]. This normally limits the long-term persistence of transgene expression, which is an important drawback for GT with adenoviral vectors. In addition, the large size of adenoviral vectors reduces their effectivity in transducing some complex tissues like the retina.

Adeno-associated Viral (AAV) Vectors:

AAVs are small single-stranded DNA (ssDNA) viruses from the parvovirus family that need co-infection of an adenovirus in the same cell in order to complete their replicative cycle [17]. Their genome is comprised of two open reading frames, called Rep and Cap, and two terminal packaging signals called Inverted Terminal Repeats (ITR) (Figure 1). The Rep gene encodes 4 proteins (Rep 40, Rep 52, Rep 68 and Rep 78) necessary for the replication of the viral genome inside the nucleus of the host cell. The Cap gene encodes three proteins (VP1, VP2 and VP3) that assemble to constitute the viral capsid [18]. Vectors derived from AAVs maintain only the ITRs from the original AAV genome, which are crucial for viral genome integration, replication and packaging [19] and can package only 4,7Kb, which is the main limitation of AAV vectors [20]. AAV vectors are normally produced using the triple transfection method, in which usually human embryonic kidney 293 cells (HEK293) are transfected with 3 plasmids: one contains the DNA sequence that has to be encapsidated, flanked by the ITR sequences, normally from AAV2; a second plasmid encodes for the rep and cap genes; and a third plasmid encodes for the adenoviral helper genes that are necessary to complete the AAV cycle [21].

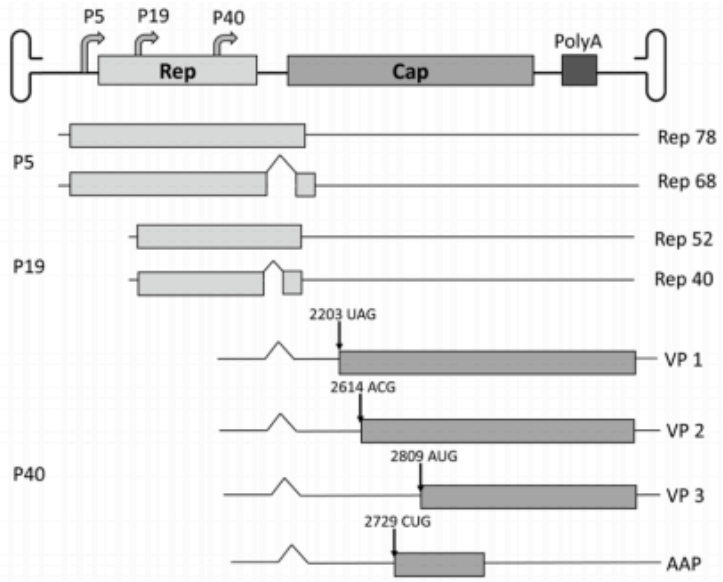


Figure 1: Schematic depiction of the AAV genome. The 4,7Kb ssDNA is flanked by 145bp inverted terminal repeats (ITRs). It contains two open reading frames (ORFs). Rep encodes for 4 proteins (Rep 40, Rep 52, Rep 68 and Rep 78) while Cap encodes for 3 (VP1, VP2 and VP3). An alternative ORF encodes for the assembly activating protein (AAP).

AAVs are some of the most frequently used vectors for in vivo applications of gene therapy, because of their safety profile, wide tropism and ability to provide long-term transgene expression while their genome remains in an episomal status, which avoids insertional mutagenesis [1]. Twelve different serotypes of AAV (AAV1-AAV12) have been identified so far from primate and human origin [18, 22]. Different serotypes can have very different tropisms depending on the proteins expressed in their capsid and their interaction with different cell surface receptors. This feature has been exploited for AAV vector production by using the Cap gene from a particular serotype in order to give a specific tropism to the vector. This has greatly broadened and diversified AAV applicability for transducing different tissues. Additionally, several authors have generated synthetic AAV capsids with different tropisms, either by rational design modification of natural serotypes or by gene reshuffling and high-throughput screening of generated variants [23, 24]. This has allowed to improve transduction efficiency and specificity while reducing immune recognition, which is crucial for systemic delivery approaches [25]. AAV-mediated gene therapy has been very successful in different animal models and has reached the clinical trial stage for several diseases. Glybera™, a gene therapy product for lipoprotein lipase deficiency was the first AAV-based gene therapy product approved for clinical use in Europe, in 2012 [8]. In December 2017, the Food and Drug Administration (FDA) approved Luxturna™, the first gene therapy product for an inherited disease in the United States [26], which is being used to treat patients affected by Leber's Congenital Amaurosis type 2 (LCA2) , an inherited retinal degeneration (IRD) due to mutations in the RPE65 gene [27]. In 2019, the FDA also approved Zolgensma, and AAV-based treatment for Spinal Muscular Atrophy. Several other clinical trials are using AAVs to target mainly the retina and the liver to treat a variety of genetic diseases [28]. In the retina the main target diseases apart from LCA2 have been achromatopsia, choroidoremia, macular degeneration and retinitis pigmentosa (RP) [28, 29]. In the liver, most efforts have focused in treating haemophilia A and B and diseases due to enzyme deficiencies like ornithine carbamylase deficiency, familial hypercholesterolemia and mucopolysaccharidosis type VI (MPSVI) [28].

Limitations of Conventional Gene Therapy with AAV:

Although GT using AAV vectors has been very successful for the treatment of cancer and several genetic diseases, it still presents some important limitations that have to be addressed:

The first limitation is the immune response against AAVs. Preexisting immunity against AAV capsids has been shown in several clinical trials to generate a CD8+ response against AAV2 and AAV8 serotypes, resulting in liver damage and loss of transgene expression [25, 30]. However, a clinical trial by Nathwani *et al*. showed that monitoring of liver damage markers like alanine aminotransferase could be used as an endpoint for applying an immunosuppression treatment, which resolved the immune response and avoided loss of transgene expression [31, 32]. On the other hand, the presence of preexisting neutralizing antibodies against common serotypes of AAV, like AAV2, greatly limits the transduction efficiency of those serotypes in some patients, and has been considered as an exclusion criterion for several GT clinical trials using AAVs [30].

Another important limitation of GT is its inapplicability to dominantly inherited diseases. While recessive diseases can be treated by achieving expression of a therapeutic gene in target cells, this is not true for most dominantly inherited diseases. This is particularly relevant in the retina, where mutations with gain-of-function or dominant-negative effects are responsible for 30-40% of cases of retinitis pigmentosa (RP)[33]. In those cases, gene addition therapies are ineffective. Instead, approaches aiming at reducing or eliminating the production of toxic protein products are being studied [29].

Last, the episomal nature of recombinant AAV genomes can be a limitation for gene transfer in dividing cells, since cell division causes dilution of viral genomes [34, 35]. This is particularly relevant in the liver, where neonate animals treated with AAV present progressive reduction of the expression of the therapeutic protein, which limits applicability of this approach in paediatric

patients [36]. Importantly, liver injury due to traumatic events or liver infection followed by regeneration can also lead to reduction of therapeutic protein expression in adult patients, as shown by Nakai et al., who performed partial hepatectomy 12 weeks after AAV transduction of the liver and observed significant reduction of transgene expression [35]. Similarly, Inagaki et al. showed reduction of AAV vector copy numbers in livers after partial hepatectomy [37].

The Genome Editing Revolution

Genome Editing (GE) has arisen in recent years as an exciting approach for therapy of inherited diseases. Its main objective is to correct the function of defective genes by modifying the genome of the affected cells with the highest possible precision [3]. The first GE publications reported use of the meganuclease I-SceI to generate a double-strand break (DSB) in the DNA, and demonstrated that the subsequent repair of the DNA generated modifications in the genome [38, 39]. However, I-SceI recognizes and cleaves a particular and very uncommon 18bp sequence, and thus its applicability in different loci was very low. In the following years, I-SceI and other meganucleases were modified in order to tailor them to target specific DNA sequences [40]. This was a very time and resource-intensive approach that required computational design, mutagenesis of the nuclease, high-throughput screening and in vivo validation. The most successfully used meganuclease was I-CreI, which was tailored to target different genes like XPC and RAG1, and also the HSV-1 virus [3, 40]

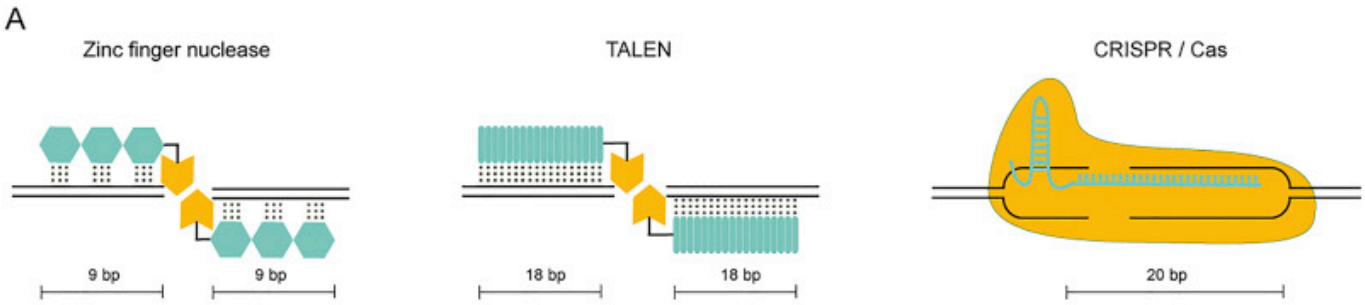


Figure 2: Scheme depicting the different target recognition characteristics of ZNs, TALENs and CRISPR/Cas9. Green structures represent DNA-binding mechanism. Yellow structures represent effector DNA-cleaving domains. The length of DNA target sequences is depicted below each drawing. (Modified from Stieger et al., 2017)

Due to the difficulty of tailoring meganucleases, they rapidly fell out of use and were replaced by the much more adaptable Zinc-Finger Nucleases (ZFNs). ZFNs were designer proteins generated by fusing a DNA-binding zinc-finger domain with a nuclease domain. Zinc-finger domains are composed of different 30-aminoacid fingers that are able to recognize 3-4 bases of the target DNA sequence. Each finger acted as a separate module, and protein engineering allowed generation of zinc-finger domains with the desired modules in order to target a specific 18-bp DNA sequence (Figure 2). The most widely used nuclease domain was derived from the restriction endonuclease Fok1, and was only able to generate a DSB when dimerized with another Fok1 domain bound in the opposite orientation at the other DNA strand at a distance of 5-7bp [41, 42]. Thus, two ZFNs had to be designed for each target sequence in order to allow DNA cleavage. One of the main limitations of ZFNs was the inability to generate new zinc-finger domains that recognized other DNA triplets, so only combinations of existing zinc-finger domains were possible, greatly reducing the availability of targets in the DNA [41]. However, despite this and the difficulty of generating ZFNs with low off-target activity, several publications showed successful DNA cleavage by ZFNs [43], and some approaches using ZFNs are currently being tested in clinical trials for Hemophilia B, Mucopolysaccharidosis I and II and HIV infections [44, 45].

Transcription Activator-Like Effector Nucleases (TALENs) were developed after ZFNs following a similar design. FokI was maintained as the nuclease domain, while the DNA-binding domain was derived from *Xanthomonas* transcription activator-like effector proteins. These proteins contain a DNA-binding region with 34-aa tandem repeats that differ in the aminoacids 12 and 13, which confers them differential recognition of a single base. In this way, each repeat is able to recognize a particular base (Figure 2). This supposed a clear advantage compared to ZFNs, because it allowed targeting of any DNA sequence by engineering the necessary modules together [45].

For this reason, FokI-TALENs and other versions with different endonuclease domains were more widely used [46-48]. However, they large size has been an important challenge for their in vivo delivery using viral vectors [49], which was achieved using adenoviral and lentiviral vectors [50].

CRISPR/Cas9

In 2010, the CRISPR system was described as a newly discovered prokaryotic adaptive immune system [51, 52]. This system is based in the ability of bacteria to cut small DNA fragments from the DNA of pathogens and store them in a DNA array called CRISPR. This CRISPR array is then transcribed into RNA molecules called crRNA corresponding to each DNA fragment from external origin. Each different crRNA is bound by a trans-activating crRNA (tracrRNA) to the endonuclease Cas9, generating a ribonucleoprotein complex. Cas9 is formed by 6 different domains. After binding of the tracrRNA and crRNA to the RecI domain, Cas9 suffers a large conformational rearrangement [53]. The formed ribunocleoprotein has the ability to recognize any DNA sequence that has Watson-Crick complementarity to the crRNA, as well as a particular 3-5bp sequence called a protospacer-adjacent motif (PAM) at the 3' end of the complementary region. The PAM is not part of the crRNA and is specifically recognized by the PAM-interacting domain of the Cas9 protein, while the rest of the sequence is recognized by the crRNA. This PAM sequence is absent in the CRISPR array in the bacterial genome, thus avoiding its cleavage by Cas9. [54]. When the Cas9 ribonucleoprotein recognizes the target sequence, the HNH and RuvC domains cleave the target and non-target strands respectively, generating a DSB (Figure 3) [53, 55, 56].

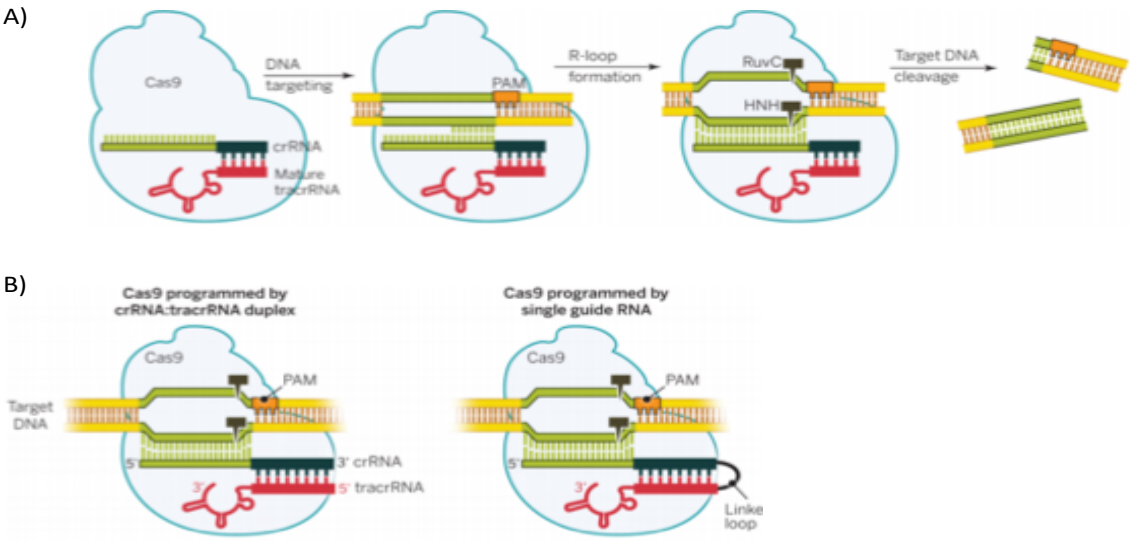


Figure 3: CRISPR/Cas9 recognition of target sites. (A) CRISPR/Cas9 binds to a tracrRNA and a crRNA and recognizes its target site, which is adjacent to a PAM sequence. RuvC and HNH domains cleave target and non-target strands respectively to generate a DSB. (B) sgRNA design for CRISPR/Cas9 targeting. Adapted from Doudna and Charpentier [56].

In 2012, the groups of Doudna and Charpentier adapted the CRISPR/Cas9 system for genome editing by combining the crRNA and tracrRNA molecules into what is now called a single guide RNA (sgRNA) (Figure 3). They showed that this sgRNA was able to bind to the Cas9 protein and target it to cleave a desired target sequence that is adjacent to a PAM motif, 3 bases upstream of the PAM sequence. The PAM was not part of the sgRNA, but immediately adjacent to it in the target sequence [55, 56] (Figure 4). In 2013 the Zhang group provided a toolkit that made CRISPR/Cas9 design widely available for other researchers [57]. This system is significantly more advantageous than either ZFNs or TALENs from a design standpoint, since it requires no protein engineering, but just the design and cloning of a 20bp nucleotide sequence into the gRNA scaffold to target Cas9 to any sequence in the genome (Figure 2). Interestingly, several gRNAs can be expressed together with a single Cas9 protein in order to target it to different loci in the same cell, thus generating precise large deletions [58, 59] and inversions [60]. In the last years, CRISPR/Cas9-mediated genome editing has been used in several pre-clinical models as a therapeutic approach, especially for Duchenne Muscular Dystrophy, Retinitis Pigmentosa and Hereditary Tyrosinemia type I [61].

Cas9 variants with altered PAM specificities:

In the last years, hundreds of publications have explored the capabilities of the canonical CRISPR/Cas9 system, which uses the *Streptococcus pyogenes* Cas9 (SpCas9) protein, as well as several other Cas9 proteins that have been derived from other bacteria (Table 1). The most relevant Cas9 proteins that have been described so far have been derived from *Staphylococcus aureus* (SaCas9) [62], *Campylobacter jejuni* (CjCas9) [63] and *Neisseria meningitidis* [64] (NmCas9). However, each Cas9 binds a particular PAM sequence, which limits the availability of its targets. Several authors have used directed evolution to generate Cas9 variants [65], mostly from SpCas9, that are able to recognize different PAMs (Table 1). The most widely used Cas9 variant has been the SpCas9-VQR developed by Kleinstiver et al., which recognizes a NGAN or NGNG PAM instead of the NGG PAM recognized by SpCas9 [66]. Alternatively, variants of the Cpf1 endonuclease [67], which is part of the CRISPR/Cas9 family, have also been isolated from different bacteria, and extensively modified to change its PAM specificity. Table 1 gives an overview of the most used Cas9 variants and their PAM specificities.

Species/Variant of Cas9	PAM Sequence
Streptococcus pyogenes (SP); SpCas9	3' NGG
SpCas9 D1135E variant	3' NGG (reduced NAG binding)
SpCas9 VRER variant	3' NGCG
SpCas9 EQR variant	3' NGAG
SpCas9 VQR variant	3' NGAN or NGNG
xCas9	3' NG, GAA, or GAT
SpCas9-NG	3' NG
Staphylococcus aureus (SA); SaCas9	3' NNGRRT or NNGRR(N)
Acidaminococcus sp. (AsCpf1) and Lachnospiraceae bacterium (LbCpf1)	5' TTTV
AsCpf1 RR variant	5' TYCV
LbCpf1 RR variant	5' TYCV
AsCpf1 RVR variant	5' TATV
Campylobacter jejuni (CJ)	3' NNNNRYAC
Neisseria meningitidis (NM)	3' NNNNGATT
Streptococcus thermophilus (ST)	3' NNAGAAW
Treponema denticola (TD)	3' NAAAAC

Table 1: Natural and modified Cas9 variants with different PAM specificity.
(https://www.addgene.org/crispr/guide/**)**

Cas9 variants that reduce off-target activity:

One of the main concerns with the use of Cas9 is its ability to target and cleave loci that are only partially homologous to the gRNA. Since Cas9 recognizes a 20bp DNA sequence, small mismatches are tolerated for cleavage, which represents an important risk for its application in mammalian cells. In order to avoid off-target activity, gRNA design tools include a prediction of potential off-target sites for Cas9, and gRNAs with low off-target scores are normally used. Several methods have been developed to identify the potential off-target sites, and deep sequencing of the sites with the highest scores is a routine analysis when using Cas9. However, these methods depend on software prediction of off-target sites. Recently, the Joung group developed GUIDE-Seq, which is an unbiased method of identifying off-target DSBs by Cas9. This system uses a double-stranded oligodeoxynucleotide (dsODN) that spontaneously integrates in DSBs at a high efficiency.

Probes specifically targeting this dsODN are then used for selective amplification and sequencing of the genomic loci where the dsODN has integrated [68]. A more recent tool, called CIRCLE-seq, improved on this method by generating circular DNA from the sheared genomic DNA. These circular DNA molecules are treated with Cas9 and a specific gRNA, and only those that are cleaved by Cas9 will be linearized. After ligation of adapters to the generated ends, linearized DNA molecules can be sequenced to detect the off-target sites [69]. These systems have proven to be more effective at identifying off-target sites with low efficiency of cleavage that were not detected by previous approaches.

Recently CRISPR/Cas9 development has focused in better characterizing and reducing the off-target activity of Cas9 in other genomic loci (Table 2). The first attempt to avoid off-target DNA cleavage was the generation of the Cas9 nickases. The most used Cas9 nickase carries an aspartate to alanin mutation in the 10th aminoacid (D10A) of the RuvC domain, which inactivates the RuvC cleavage of the non-target strand. Another variant carries a N863A mutation in the HNH domain which inactivates it. In these variants, Cas9 is only able to generate a single-strand break (nick) in the target strand. With this system, two gRNAs recognizing the two strands with 40-70bp distance are necessary in order to generate the DSB, greatly reducing the potential of off-target cleavage. However, this system is significantly more efficient in cleaving the target site when the PAMs of the target regions in both strands are facing the extremes of the target region, which reduces the availability of targets [57, 70]. Alternatively, a publication by Kleinstiver *et al*. showed that substitution of residues N497, R661, Q695, Q926 with four Alanine could be used in order to disrupt their hydrogen bonds with the target DNA. This reduced the binding energy due to non-specific DNA contacts by the DNA binding domains of Cas9, dramatically reducing the ability of Cas9 to bind to the DNA when the gRNA target recognition is not perfect. This Cas9 variant, named SpCas9-HF1, demonstrated almost non-existent off-target effects while maintaining high efficiency of on-target cleavage [71].

Slaymaker *et al*. generated another high-fidelity variant, which they called eSpCas9, by decreasing the positive charge in the groove formed by the Cas9 catalytic domains HNH and RuvC [72]. Alternatively, Casini *et al*. screened random mutations in the REC3 domain and were able to generate EvoCas9, which was demonstrated to be significantly more specific than either of the high-fidelity variants mentioned above [73]. A very different approach by Marienne *et al*. proposed KamiCas9, which uses a second gRNA targeting the start codon of the Cas9 gene in order to knock it out after Cas9 expression, thus limiting the time of expression of Cas9 and the probability of off-target cleavage [74]. Last, in 2016 the group of David Liu generated a fusion protein with Cas9 and a cytidine deaminase which uses a gRNA to recognize a target sequence, but does not generate a DSB and instead generates conversion from cytidine to uridine in a 5bp window. This allows for precise C→T or G→A mutation without generating DSBs [75]. This field was further developed by the Liu Lab in 2018 with the development of another base editor that is able to generate A→G or T→C mutations [76, 77]. All these variants and new ones being researched right now should solve the issue of off-target cleavage by Cas9, although it is still considered an issue for in vivo application of genome editing, and long-term studies of Cas9 toxicity are being conducted.

Cas9 variant	Mechanism for avoiding off-target	Author
SpCas9 D10A	Inactivation of RuvC domain cleavage activity	Ran et al.[57]
SpCas9 N863A	Inactivation of HNH domain cleavage activity	Ran et al. [57]
SpCas9-HF1	Reduction of Cas9 binding energy by protein engineering	Kleinstiver et al. [71]
eSpCas9	Reduction of positive charge of Cas9	Slaymaker et al. [72]
EvoCas9	Mutagenesis of REC3 domain	Casini et al. [73]
KamiCas9	Knockout of Cas9 after expression	Marienne et al. [74]

Table 2: Cas9 variants that reduce off-target activity and methods used for their generation.

DNA Repair Mechanisms after Double-Strand Breaks:

Double Strand Breaks (DSBs) are one of the most deleterious forms of DNA damage, because they promote chromosomal instability, such as translocations and inversions, if their repair is defective. For this reason, after detecting a DSB, a strong DNA repair response is activated and, in case it fails, cell death is promoted [78]. This has a clear anti-tumorigenic objective, since translocations are a common cause of activation of oncogenes. DSBs are recognized and bound by ATM (Ataxia telangiectasia mutated) and ATR (Ataxia telangiectasia and Rad3-related protein), which recruit and phosphorylate H2AX. This factor is key in recruiting several other factors and the main regulators of the two different pathways of DSB repair: 53BP1 is the main regulator of Non-Homologous End Joining (NHEJ) repair, while RAP80 regulates Homology-Directed Repair (HDR). The presence and activity of 53BP1 and RAP80 depend on the cell type and cell cycle phase, and configure the nature of DSB repair (Figure 4) [79].

Homology-Directed Repair (HDR):

This mechanism is mostly active in the G2 and S phases of the cell cycle, and is believed to be associated to DNA damage during chromosomal replication. HDR employs the sister chromatid as a template for homologous recombination in order to precisely repair the DSB. After activation of the HDR pathway by RAP80 and BRCA1, the MRN complex is recruited and associates with CtBP to generate single-stranded DNA 3' overhangs. These overhangs are stabilized by RPA, which is then replaced by RAD51. RAD51-coated ssDNA is then invaded by the homologous DNA region from the sister chromatid, which is then used as a template for DNA repair by polymerase-mediated extension of the ssDNA [78].

HDR has been exploited extensively for gene correction, by delivering an exogenous donor sequence with homology arms that can be used as a template for the DSB repair (Figure 4). For *ex vivo* applications, the most widely used templates are single-stranded oligodeoxynucleotides (ssODNs) with at least 40bp of homology on each side [57, 80].

This approach has been used extensively to modify hematopoietic stem cells and induced pluripotent stem cells (iPSCs). Different publications have shown feasibility of mutation correction in patient-derived iPSCs that can then be used to generate RPE cells or photoreceptors for retinal transplantation [81-83]. Although this approach is still far from clinical application, it could be very beneficial to treat inherited retinal diseases. In the last years, HDR has also been used for *in vivo* GE using a donor dsDNA with at least 800bp, with a significantly high efficiency [84, 85]. Interestingly, Barzel *et al*. showed in 2011 that delivery of an HDR template could generate HDR without need of DNA cleavage by endonucleases, albeit with extremely low efficiency. However, by inserting the Factor IX (FIX) gene in the albumin locus they were able to achieve high levels of expression of FIX from the liver, even with very low efficiency of integration [86, 87]. This approach using albumin as a "safe harbor" for integration has been further improved by endonuclease-mediated DSBs, and has achieved high levels of liver expression of several therapeutic proteins [84, 88-90]. Three clinical trials using ZFNs targeted to the albumin locus in order to produce therapeutic proteins from the liver are in progress for treatment of Hemophilia B and mucopolysaccharidosis type I and II respectively (trial numbers NCT02695160, NCT02702115 and NCT03041324, [8]).

Non-Homologous End Joining (NHEJ):

The NHEJ pathway is the most common DSB repair pathway in higher eukaryotes. It is active in all phases of the cell cycle, but it is most active during the G1 phase [91, 92]. NHEJ starts with the recognition of the two-ended DSB by the Ku70/Ku80 complex. This complex recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), which stabilizes and aligns the DNA ends. Both ends are then processed by the Artemis endonuclease, which eliminates the 5' and 3' single-stranded overhangs and reveals possible complementary nucleotide stretches. DNA polymerases μ and λ can also be recruited to fill in short gaps between nucleotides, which will then be ligated by ligase 4 [78]. In this process, microhomology regions of 5-25bp between the 5' and 3' ends can

direct the repair to a precise recombination between microhomologous regions, which is considered a particular kind of NHEJ called microhomology-mediated end joining (MMEJ) that generates small and predictable deletions [78, 93]. However, in most loci, no microhomology regions are present, and thus end resection and gap filling are the main mechanisms, which leads to the generation of small insertions and deletions (INDELs). These INDELs, in most cases, alter the reading frame of the gene, thus generating premature stop codons and the knock-out of the targeted gene (Figure 4). This has been used extensively to generate animal models of disease, as well as to correct disease-causing mutations. Specifically, allele-specific targeting of dominant mutant alleles is the most promising application of NHEJ for therapeutic purposes [3, 94].

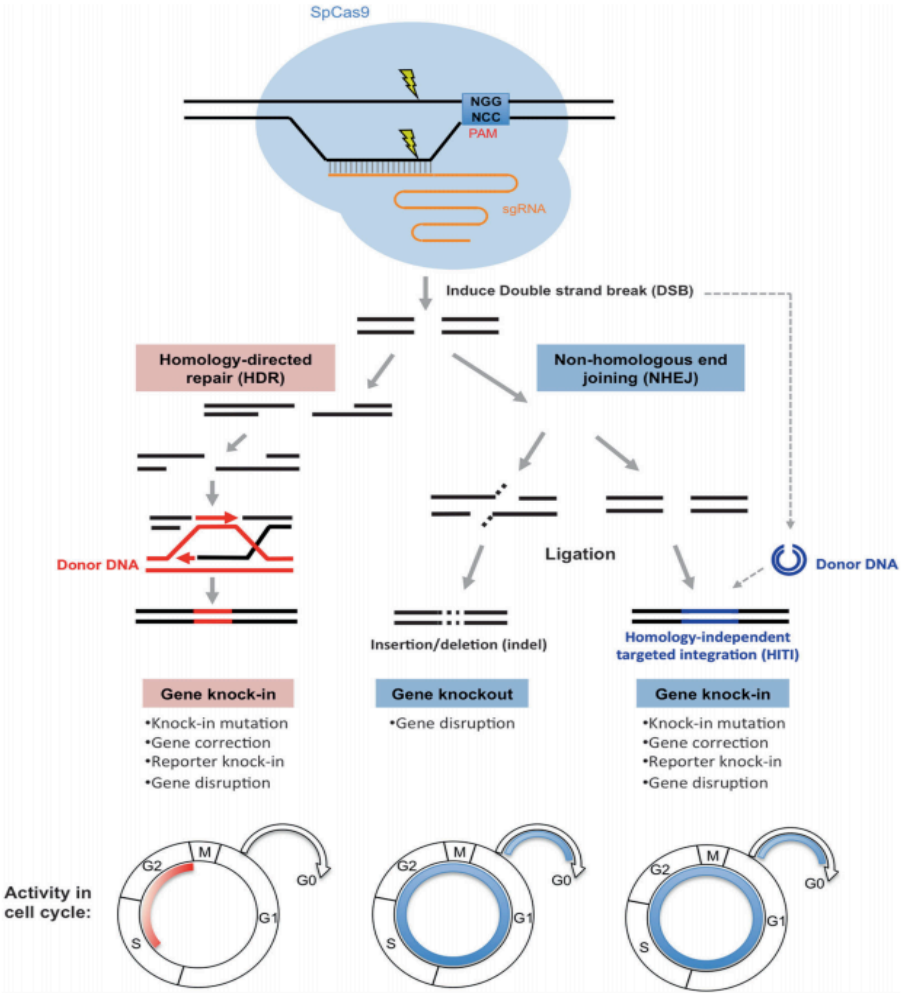


Figure 4: DSB repair mechanisms and their applications for genome editing. HDR uses a donor DNA with homology regions in order to correctly repair the DSB. NHEJ generates INDELs that cause gene disruption. HITI exploits NHEJ to integrate a donor DNA without homology regions into the DSB. HDR is only active in S and G2 phases, while NHEJ is active in all phases of the cell cycle. (Suzuki & Belmonte, 2018)

The detection of INDELs is the main readout that allows detection of DSBs caused by endonucleases. Several mechanisms can be used to detect INDELs, however the most widely used are direct sequencing of the cleaved DNA and the use of T7 Endonuclease I (T7E) assay. This assay takes advantage of the ability of T7E to cleave the DNA when there is a mismatched base. In this assay, an amplified PCR product containing the endonuclease target sequence is amplified from the genomic DNA, and then de-annealed and re-annealed using temperature gradients, so that some of the cleaved fragments will align with uncleaved fragments or fragments with different INDELs. These imperfect alignments are then recognized and cleaved again by T7E, generating two DNA fragments of predictable sizes. Another method for INDEL detection is the Tracking of Indels by DEcomposition (TIDE) analysis, which uses the chromatogram of a non-Cas9-cleaved PCR fragment as reference to decompose the frequency and types of INDELs present in the Cas9-cleaved PCR fragment. This gives a semiquantitative assessment of cleavage efficiency and of the kinds of INDELs generated after cleavage [95].

Recently, Shen *et al*. developed a machine learning software called inDelphi that accurately predicts the nature of insertions and deletions caused by a particular gRNA based on the sequence context around the cleavage site. This software is able to predict the frequency of 1-60bp deletions and the frequency and nature of 1bp insertions after CRISPR/Cas9 cleavage of a particular target sequence, and has allowed to identify loci that, after cleavage, yield more than 50% of major editing products with the same exact genotype, which is promising for therapeutic applications [96].

Homology- Independent Targeted Integration:

Since HDR is only active in dividing cells, several authors tried developing a mechanism for targeted transgene integration via NHEJ (Figure 4). The first approaches used ZFNs and TALENs *in vitro* to generate DSBs in the genomic locus as well as in a donor plasmid, facilitating the introduction of the exogenous DNA sequence via NHEJ [97, 98]. Other publications showed the feasibility of this approach in various organisms. Lackner *et al*. were the first to use CRISPR/Cas9

for NHEJ-mediated targeted integration, and proposed it as a mechanism for gene tagging in zebrafish [99]. However, the drawback of these approaches was the inability to control the direction of integration. Due to the lack of homology arms, NHEJ was able to integrate the donor DNA in both directions, lowering the efficiency of correct transgene integration and generating undesired integration products [100].

In 2016, Suzuki *et al*. developed Homology-Independent Targeted Integration (HITI). Their approach uses a donor DNA that is flanked by the same gRNA/Cas9 sequences targeted in the genome. The possible inverted integration of the donor DNA is avoided by inverting its gRNA target sequences, so that Cas9 can recognize and cut again the target sequence only if inverted integration occurs [101] (Figure 5).

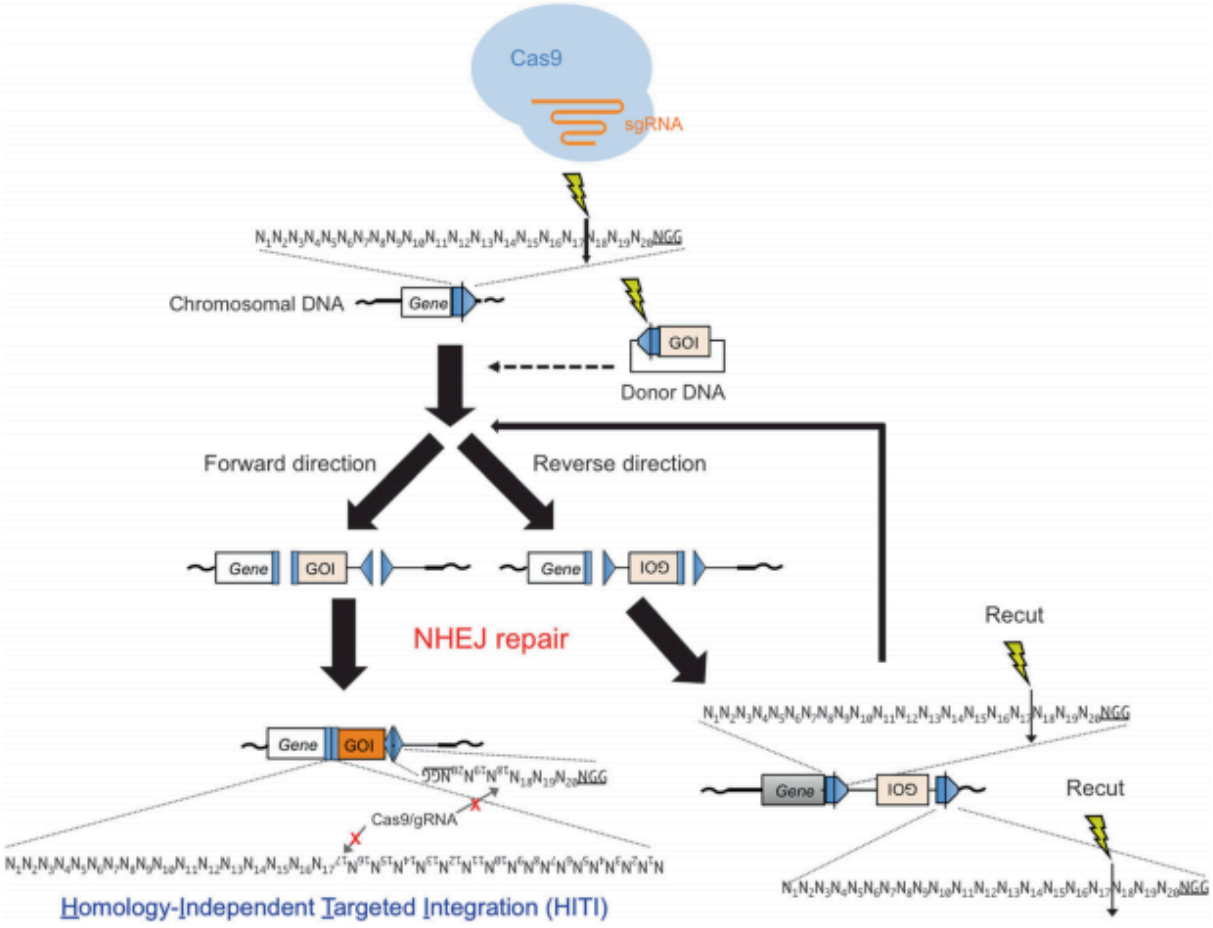


Figure 5: Schematic depiction of HITI design. Inverted target sites flanking the donor DNA allow directional integration after Cas9-mediated DSB. Inverted integration is avoided through Cas9 recut. Blue shapes represent target sequences for Cas9. GOI: Gene of interest. Lightning shapes represent CRISPR/Cas9-mediated DSBs. (Suzuki&Belmonte, 2018)

The main limitation of this approach is that it doesn't have a very high efficiency. However, in the cases where HITI occurs, a surprisingly high (60-80%) rate of INDEL-free directional integration was observed. Because HITI uses NHEJ, it is effective in terminally differentiated neurons, liver and skeletal and cardiac muscle. Interestingly, HITI was shown to be at least as effective as HDR in liver [101]. However, HITI precision and off-target integration need to be further characterized in order to establish it as a safe alternative to HDR. Moreover, its therapeutic efficacy has only been shown in the retinal pigmented epithelium, and further investigation of its therapeutic potential in differentiated neurons, like photoreceptors, and in the liver is necessary [100].

Inherited Retinal Degenerations:

Inherited Retinal Degenerations (IRDs), with an overall global prevalence of 1/2,000 [102], are a major cause of blindness worldwide. IRDs are mostly monogenic and are caused by mutations in genes mainly expressed in retinal photoreceptors (PRs; rods and cones) and to a lesser extent in the retinal pigmented epithelium (RPE) [103]. Retinitis pigmentosa (RP) is a heterogeneous group of IRDs affecting 1/3.000-5.000 people worldwide. RP is characterized by progressive degeneration and loss of PRs in the outer nuclear layer (ONL) of the retina. The main symptoms of the disease start with dark adaptation difficulties and night blindness, which progresses into loss of peripheral vision. In later stages of degeneration, further loss of peripheral vision causes tunnel vision, which progresses to loss of central vision [104]. RP is almost exclusively inherited as a Mendelian trait. With about 100 loci identified, RP is arguably one of the most genetically heterogeneous conditions in humans [33]. The majority of mutations causing RP occur in genes expressed in PRs. Most cases present an autosomal recessive inheritance. These cases tend to have an onset in early childhood, and most patients are declared legally blind between 40 and 60 years of age. Thirty-40% of cases of RP present an autosomal dominant inheritance (AdRP). These cases present a slower degeneration, frequently maintaining a relatively good visual acuity even in advanced ages [33, 104].

Rhodopsin and phototransduction:

The most common locus (*RHO*) involved in dominant RP encodes for the protein rhodopsin (RHO). RHO is a 398 aminoacid class A G-coupled receptor formed by seven transmembrane α-helices connected by three intra-discal and three cytoplasmic loops, as well as a C-terminal domain with cytoplasmic localization that acts as an effector. RHO is bound in the cytoplasmic part of its 7th helix to the chromophore 11-cis-retinal (Figure 6) [105].

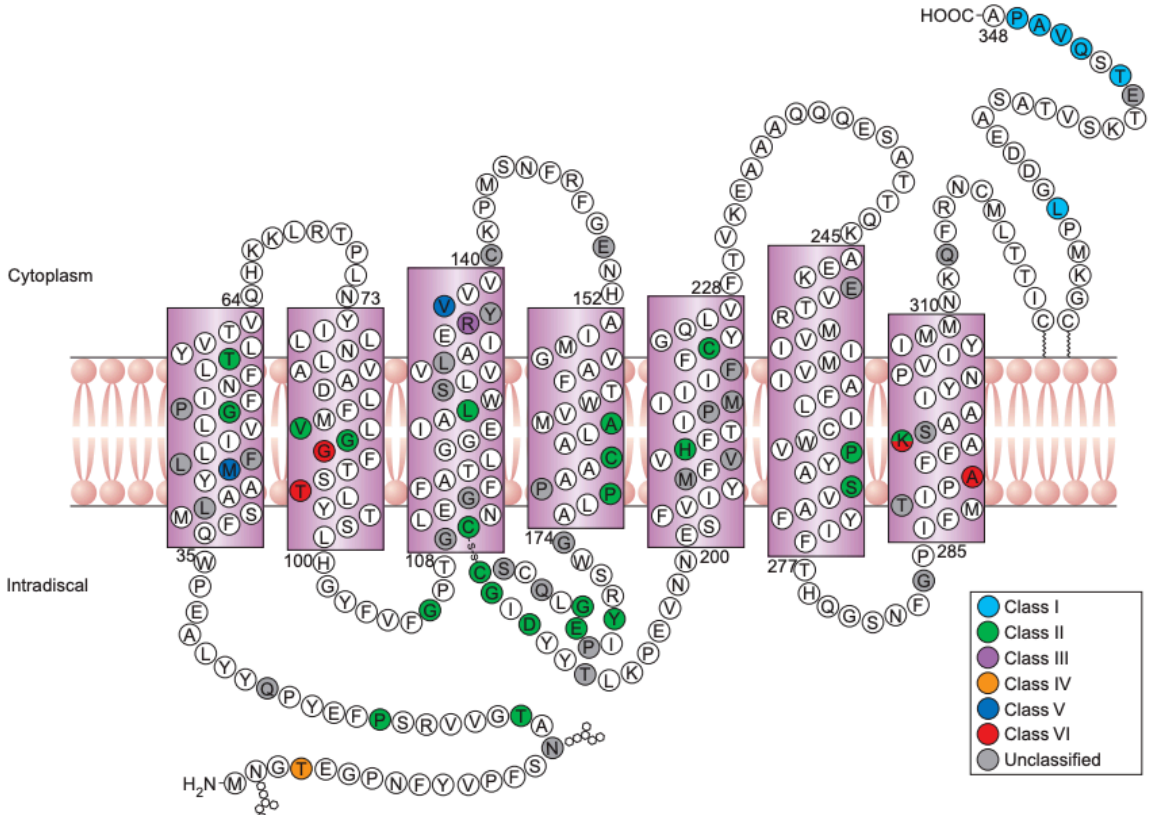


Figure 6: Secondary structure of rhodopsin showing all mutations causing RP4. Aminoacids are depicted as letters. Numbers represent positions of aminacids in the C-terminal and N-terminal ends of loops. Colors show the classification of mutations according to their pathologic mechanism. (Mendes et al, 2005).

RHO is found in the membranes of the disks present in the outer segments of rods, where it acts as a visual pigment. In dark conditions, cGMP-gated channels in the rod membrane are open, allowing entry of Na+ and Ca+2 ions, which maintain a depolarized state. When a photon is recognized by RHO, this causes isomerization of 11-cis-retinal to all-trans-retinal. This generates a conformational change in RHO, which then is able to bind and activate a photoreceptor-specific GTP-binding protein transducin.

Transducin then substitutes its bound GDP for GTP in its α domain. The activated α subunit causes disinhibition of the cGMP phosphodiesterase, which hydrolyzes cGMP to GMP. The resulting decrease in cGMP concentration causes the closure of the cGMP-gated Ca+2 channels in the PR membrane, leading to hyperpolarization of the PR and generating an electrical signal that is then transmitted to the bipolar cells in the inner nuclear layer (INL) (Figure 7). After this signal cascade, the kinase GRK1 phosphorylates various serine and threonine residues of RHO to inactivate it. Arrestin then binds to RHO and hydrolyses the all-trans-retinal. Then calmodulin reopens the Ca+2 channels causing de depolarization of the rod [106].

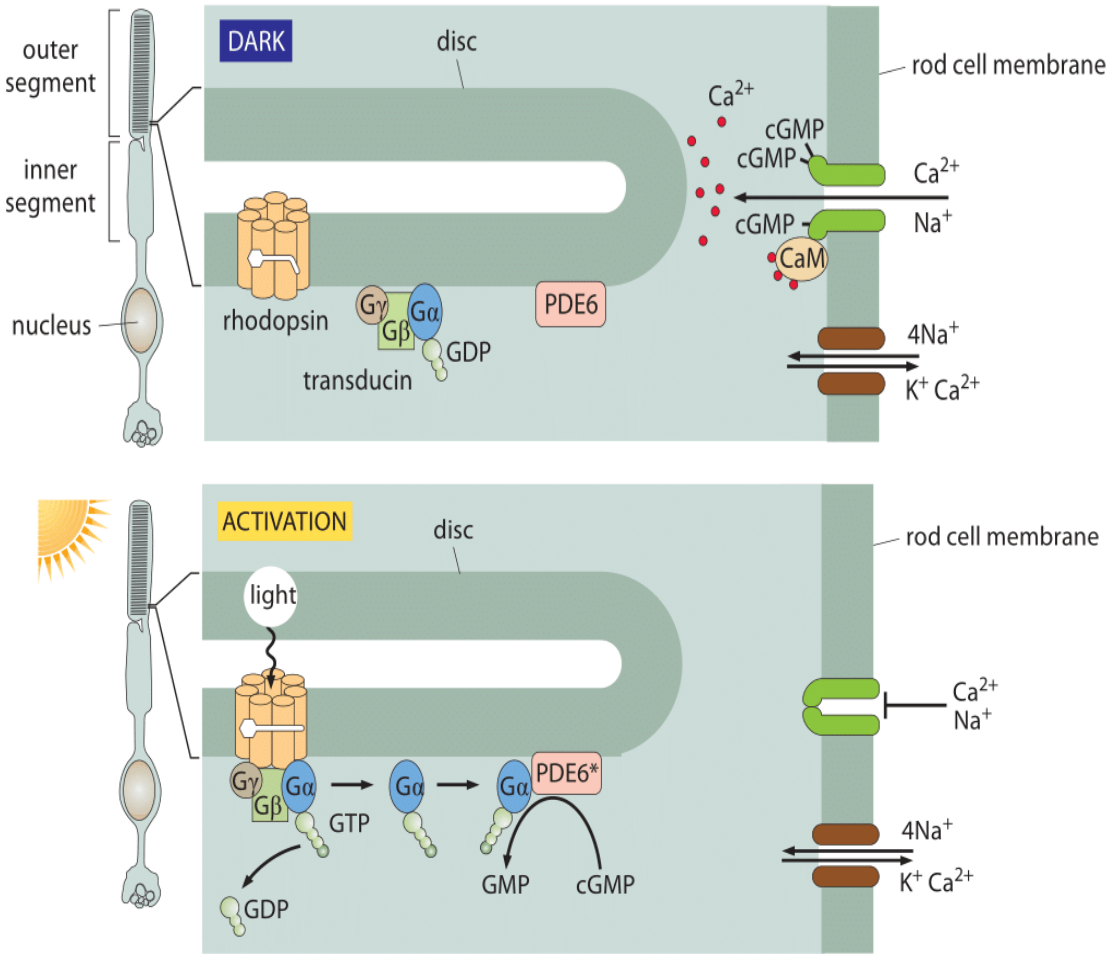


Figure 7: Phototransduction cascade. In dark conditions Ca+2 and Na+2 channels are open by cGMP, causing depolarization of the photoreceptor. Rhodopsin activates in response to a light stimulus, and activates the G-protein transducing, which causes cGMP hydrolysis and closure of Ca2+ and Na2+ channels, leading to photoreceptor hyperpolarization. Adapted from A. Stockman et al., Journal of Vision 8: 1, 2008

RHO Mutations as a cause of disease:

The *RHO* coding region is comprised of 5 exons that are highly conserved among vertebrates. Over 200 different point mutations have been identified in *RHO*. Most of these mutations are pathogenic. Although some of these mutations cause recessive RP, most *RHO* mutations cause autosomal-dominant RP (adRP) [107-109], being responsible for about 25% of all adRP cases in the United States and about 20% of cases elsewhere in the world [33]. RHO mutations causing adRP are classified in six classes depending of their pathogenic mechanism, with most classes depending of gain-of-function effects [107] (Figure 6). The Proline 23 to Histidine substitution (P23H) is the most common *RHO* mutation in North America, representing 9% of all cases of RP in the United States. RHO P23H does not fold properly and accumulates in the endoplasmic reticulum (ER) which results in ER stress and cytotoxicity ultimately leading to PR cell death [107, 110]. Mutations in the C-terminal end of the protein, like S334-Ter, P347S or P347L are considered to have a dominant-negative effect, and are more common in countries like Italy and Spain [111, 112].

Retinal Gene Therapy:

The retina is an ideal candidate for gene therapy for several reasons. First, it is a very accessible organ, so delivery of vector suspensions is easy and remains confined in the eye. Second, the eye is an immune-privileged site. This is due to several factors, namely the presence of the blood-retina barrier, the tight junctions formed between the cells in the RPE and also the local and systemic inhibition of immune responses against the eye [113]. Taken together, this reduces the risk of adverse systemic events after localized intraocular delivery of gene therapy vectors [114]. Another advantage of retinal gene therapy is the presence of an intraindividual control, which is the contralateral eye.

Furthermore, the retinal function and structure are easy to characterize, allowing for monitoring of disease progression and treatment success. Particularly, in vivo analyses like optical coherence tomography (OCT) and fundus autofluorescence allow monitoring of the structural integrity of the retina [29]. Separately, focal or multifocal electroretinogram (ERG), visual evoked potentials and measurements of afferent pupillary light responses are used to measure retinal function [115, 116]. All these techniques are available for use in patients but have also been adapted to the two most relevant animal models of retinal disease, which are rodents and the pig. Rodents, particularly mice and rats, have extensively been used as models of IRD and other kinds of retinal degeneration. Pig eyes, on the other hand, are used less often but better recapitulate the size and structure of the human eye, although they lack the macula, which is a cone-rich structure found only in higher primates. Instead, pigs possess a cone-rich visive streak that partially resembles the structure of the macula.

There are two main methods for intraocular delivery of vectors to the retina: Intravitreal injections consist in the delivery of the vector solution inside the vitreous, and is ideal for exposure of the anterior retina. However, the outer retina is not efficiently reached due to the barriers present between it and the vitreous. Instead, subretinal injection allows localized delivery between the neurosensory retina and the retinal pigmented epithelium (RPE), causing a small detachment between both structures, that is resolved in a period of hours or days as the vector solution is absorbed. Subretinal injection has been shown to expose vectors to PRs and the RPE, but is not able to reach all the retina (Figure 8) [29, 117].

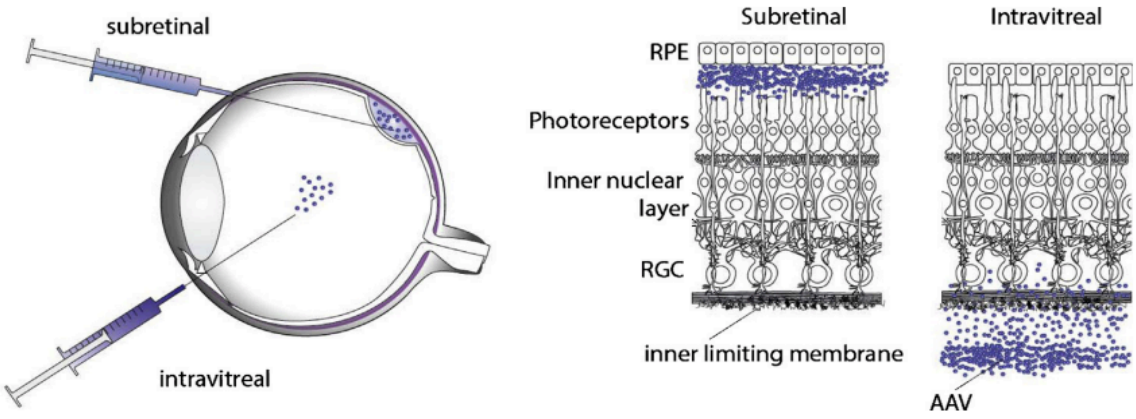


Figure 8: Scheme depicting architecture of the retina and differences between subretinal and intravitreal injections of AAV vectors. (Dalkara&Sahel 2014)

AAV Vectors for Retinal Gene Therapy:

The most used vectors for retinal GT are AAVs, with around 50 clinical trials performed or in progress in 2018 [29]. Importantly, one AAV-based gene therapy product, called Luxturna™, has been approved for clinical use in both Europe and the United States, for treatment of Leber's Congenital Amaurosis type 2 [118, 119]. While most serotypes are able to efficiently transduce the RPE after subretinal administration, PR transduction is limited to only some serotypes. The first and most widely used serotype has been AAV2, although AAV5, AAV7, AAV8 and AAV9 have all been shown to transduce PRs more efficiently than AAV2. AAV8 is the most efficient serotype for PR transduction in mice [120], pigs [121], dogs [122] and non-human primates [123] (Figure 9).

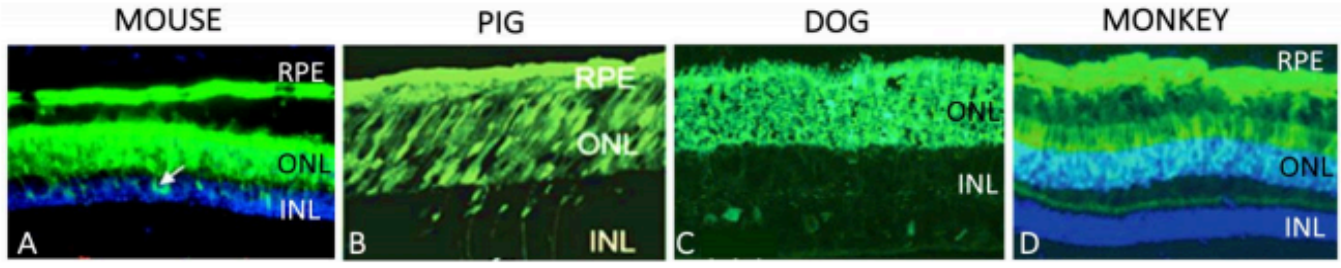


Figure 9: AAV 2/8 transduction in mouse, pig, dog and non-human primate retinas. Expression of EGFP after subretinal administration of AAV2/8 in C57BL/6 mice (A), pig (B), dog (C) and non-human primate (D). RPE: retinal pigmented epithelium, ONL: outer nuclear layer, INL: inner nuclear layer. (Modified from Allocca *et al*, 2007 (A), Mussolino *et al*, 2011 (B); Stieger *et al*, 2008 (C) and Vandenberghe *et al*, 2011 (D)).

Most efforts have been historically focused on using AAV to supply a correct copy of the otherwise mutated gene in recessive conditions due to a loss of function. However, in adRP due to RHO mutations (RP4), knockout of the *RHO* mutant allele rather than addition a correct *RHO* copy is required to achieve a therapeutic effect [29, 124]. Different strategies have been used to knock down mutant *RHO* in an allele-independent or allele-specific manner. Allele-independent strategies include *RHO* silencing with short hairpin RNAs coupled with AAV-mediated *RHO* expression [125-127] and *RHO* silencing using an inactive transcription factor [128]. Allele-specific strategies include allele-specific silencing and allele-specific genome editing. Genome editing with CRISPR/Cas9 (Cas9) has emerged in the last years as a versatile and efficient strategy for the treatment of dominant IRDs [3, 29]. Since PRs are terminally-differentiated cells, the efficiency of HDR correction after cleavage is low [100, 129]. NHEJ, on the other hand, is the main DNA repair mechanism in PRs. The resulting INDELs are normally used to knock out a targeted gene [94]. This has been used to induce allele-specific *RHO* knockout in the retina of various models of Retinitis Pigmentosa. Two different groups targeted the prevalent P23H mutation in the *mRho* and the *RHO* locus respectively. In both cases, genome editing in the retina of mouse models achieved partial correction of RP4 [130, 131], although Giannelli *et al.* used an allele-specific gRNA, while Latella *et al.* used 2 gRNAs targeted to the area close to the mutation but without allele specificity. Similarly, Bakondi *et al*. targeted the Ser334-Ter mutation in *mRho* in a transgenic rat model [132]. Yu *et al*. targeted the NRL locus and observed prevention of retinal degeneration in mice [133]. This method has also been used *ex vivo* by Burnright *et al*. to generate gene-corrected induced pluripotent stem cells, which could potentially be used for retinal therapy [81]. However, this approach is limited by both the availability of gRNA/PAM combinations at the GOF allele, and is mutation-specific, which limits its clinical applicability due to the genetic heterogeneity of RP4 [29, 107].

Lysosomal Storage Diseases:

The lysosome is a cellular organelle that contains a variety of hydrolytic enzymes in an acidic environment (pH=5)[134]. Lysosomal hydrolases are targeted to the lysosome by addition of a mannose 6-phosphate, which is recognized by receptors in the lysosomal membrane. However, lysosomal enzymes can also be secreted and then uptaken by other cells via mannose 6-phosphate receptors in the plasma membrane, followed by trafficking to the lysosome. Most lysosomal enzymes receive proteolytic and carbohydrate processing, which is not necessarily related to their catalytic activity. Lysosomal Storage Diseases (LSDs) are a group of more than 40 different inherited metabolic diseases caused by deficient activity of lysosomal hydrolases, which leads to accumulation of their substrates inside cells. This causes cellular and organ dysfunction. LSDs can be inherited as autosomal or X-linked recessive [135].

The mucopolysaccharidoses (MPS) are a group of LSDs caused by deficiency of the hydrolases that process the degradation of a class of substrate called glycosaminoglycans (GAGs) or mucopolysaccharides. Mutations in different genes cause accumulation of different GAGs, which has different effects, giving raise to 7 main kinds of MPS (MPSI-MPSVII). All MPS share a chronic and progressive evolution due to GAG accumulation in different tissues, causing multisystemic dysfunction. Common clinical features include organomegaly, dysostosis multiplex and abnormal facies. Depending on the kind of MPS, other systems like hearing, vision, the airways, the cardiovascular system, the joints or the central nervous system (CNS) can be affected [136].

Mucopolysaccharidosis type VI:

Mucopolysaccharidosis type VI (MPSVI), also known as Maroteaux-Lamy Syndrome was described in 1963. It is a very rare disease which affects 1-9/1.000.000 people worldwide (https://www.orpha.net/) and is caused by mutations in the Arylsulfatase B gene (*ARSB*) located in the 5th chromosome. The *ARSB* gene contains 8 exons that codify for the 502 aminoacid ARSB protein which is involved in the hydrolysis of the C4-sulfate ester link of dermatan sulfate during its lysosomal processing. Mutations in *ARSB* are very diverse and can affect enzyme synthesis, stability or maturation, which leads to lack of enzyme and the accumulation of dermatan sulfate in several organs [137]. This causes dysostosis multiplex, corneal clouding that can cause impaired vision, restriction of joint movement, hepatomegaly and cardiac valve dysfunction as the main symptoms. Differently to other MPS, MPSVI patients present no mental retardation [136]. A widely used biomarker for MPSVI is the elevated urinary presence of GAGs, which directly correlates with disease severity [137].

Current Treatment for MPSVI:

Enzyme Replacement Therapy (ERT) is the standard of care for several LSDs, including MPSVI. ERT for MPSVI consists in weekly intravenous administrations of the purified therapeutic protein product and achieves reduction of GAGs in tissues and urine, which leads to improvement of some symptoms [138]. However, the short half-life of the protein and its poor biodistribution make this treatment ineffective for the symptoms related to bone, heart and cartilage [139, 140]. Additionally, weekly infusions represent an extremely high cost for treating each patient, at between 150.000 and 450.000 euro per year [141], which is an issue specially in less developed countries. For these reasons, it is necessary to develop new therapeutic approaches for MPSVI.

New Therapeutic Approaches to treat LSDs:

Gene therapy has widely been used to treat LSDs, with two main approaches: *ex vivo* and *in vivo*. *Ex vivo* approaches have used lentivirus-engineered hematopoietic stem cells that constitutively overexpress the therapeutic enzyme. Using this approach, Biffi *et al*. were able to correct the phenotype in a mouse model metachromatic leucodystrophy [142]. After this, several others used similar approaches to treat murine models of different MPS, Pompe's disease and Fabry's Disease, among others [143].

Instead, *in vivo* approaches for LSDs have mainly focused on converting the liver into a factory for production and secretion of therapeutic proteins into the blood. Both adenovirus and AAV-mediated gene therapy has been used to treat different LSDs. Adenoviral vectors were effective but only achieved transient expression due to immune response against the vector, so they are not considered a viable option [143]. However, new helper-dependent adenoviral vectors have successfully been tested in the last years and have shown high levels of hepatic transduction, long-term expression and safety at low doses[144].

AAVs were first demonstrated to achieve long-term liver expression of the therapeutic transgene by Sferra *et al*., after liver parenchymal injection of an AAV2 in MPSVII mice [145]. Intravenous injection of AAV greatly increased the efficiency of liver transduction and thus the treatment effectivity [146]. AAV8 is the most frequently used serotype to transduce the liver, and has been proven to transduce up to 95% of hepatocytes after peripheral vein delivery (Figure 10A) [147]. The recent clinical success of liver gene therapy for the treatment of Hemophilia B [31, 32, 148] is increasing the interest in AAV-mediated GT in the liver. Specifically, GT has been shown to be at least as effective as ERT for treatment of MPSVI murine and feline models [149-152], which has led to a phase I/II clinical trial (Trial number NCT0317352).

However, as previously explained, liver-directed GT with AAVs has some limitations due to vector dilution during hepatocyte cell division (Figure 10B). This prevents neonatal or pediatric treatment, which could avoid or at least modulate some of the most severe symptoms of the disease, like dysostosis multiplex and joint stiffness.

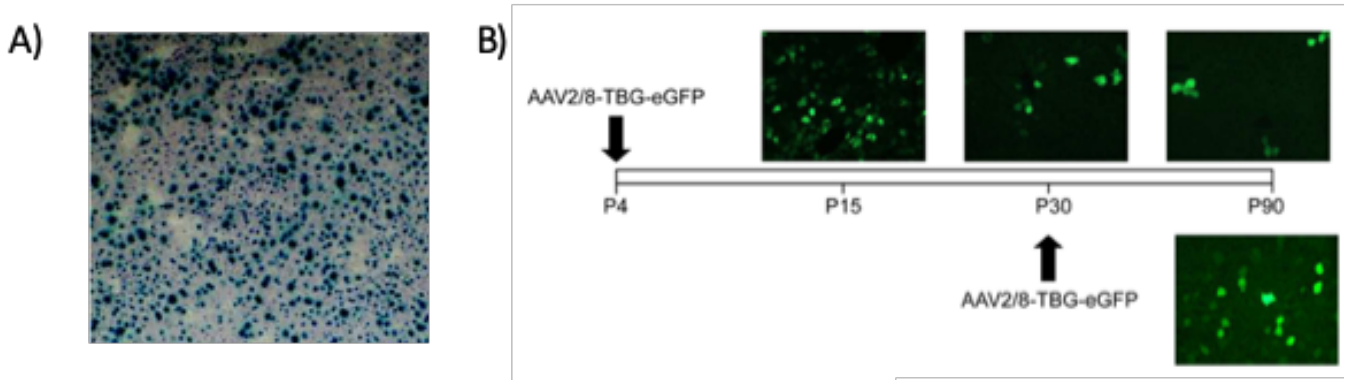


Figure 10: Liver transduction after intravenous delivery AAV2/8: A) Adult livers were transduced with 7*10^12gc/mouse of AAV8-EF1α-nlsLacZ, harvested after 6 weeks and stained with xGal and light hematoxylin. B) Neonatal (p4) mice were injected with 4*10^13vg/Kg of AAV8-TBG-EGFP. Liver sections were analyzed for fluorescence at p15, p30 and p90, observing dilution of EGFP signal. Mice injected at p30 and analyzed at p90 showed significantly higher presence of transduced hepatocytes. (Adapted from Nakai *et al.,* 2005 and Cotugno *et al*., 2011)

For this and other reasons, the field of genome editing has focused on LSDs as therapeutic targets. As explained previously, two clinical trials using ZFN-mediated HDR at the albumin locus for expression of therapeutic proteins have been approved for MPSI and MPSII respectively (Trial numbers NCT02702115 and NCT03041324). This approach targeting the albumin locus as a "safe harbor" for integration and expression of therapeutic proteins from the liver has been applied to animal models of several other diseases [87-90], and shows great promise as an alternative to conventional GT that could be used in pediatric patients.

AIMS

The overall aim of my PhD project was to develop AAV- CRISPR/Cas9 based genome editing approaches that could be used for treatment of diseases inherited either as autosomal dominant for which traditional gene replacement is ineffective, or recessive in which target cell proliferation may represent a limitation to persistence of transgene expression via episomal AAV vectors.

As a first model of disease I chose autosomal dominant Retinitis Pigmentosa. **The first aim of my project was to develop an allele-specific genome editing approach specific to the P347S mutation.**

A second parallel aim was to develop an allele-independent method for gene correction in AdRP, using a targeted integration system that could be efficient in mouse and pig photoreceptors, and to test the efficacy of this approach to correct the autosomal dominant Retinitis Pigmentosa phenotype in a mouse model.

After observing the high efficiency of the targeted integration system that I had developed, I decided to test its efficiency in dividing cells**.** Thus, **the third aim of my project was to use targeted integration in the albumin locus of neonatal mice to express a transgene**, and to assess whether this approach was enough to achieve therapeutic serum levels of arylsulfatase B in a mouse model of mucopolysaccharidosis type VI.

The development of new genome editing approaches that can be flexible and easy to adapt to other diseases would allow to expand the use of genome editing for the treatment of different kinds of inherited diseases.

MATERIALS AND METHODS

Plasmid Constructs:

Cas9 Effector Plasmids for gRNA screening

gRNAs were designed manually considering the position of the targeted P347S mutation. Effector plasmids were generated following the protocol by Zhang's lab [57]. In short, gRNAs were generated as a Fwd and Rev single-stranded DNA oligos, which were then phosphorylated and annealed using T4 Polynucleotide Kinase and T4 Ligation Buffer (NEB). Correct gRNA annealing was confirmed in a 50% acrylamide gel stained with SYBR Gold Nucleic Acid Stain (ThermoFisher). gRNAs were subsequently cloned into the px458 plasmid (#48138; Adgene) using BbsI sites. Oligonucleotides used are described in Table 3:

Oligo name	sequence
gRNA1 Fwd	5' CACCGCCTCGGCCTAAGACCTGCCT 3'
gRNA1 Rev	5' AAACAGGCAGGTCTTAGGCCGAGG 3'
gRNA2 Fwd	5' CACCGGAGACGAGCCAGGTGGCCT 3'
gRNA2 Rev	5' AAACAGGCCACCTGGCTCGTCTCC 3'
gRNA3 Fwd	5' CACCGCCTAGGCAGGTCTTAGGCCG 3'
gRNA3 Rev	5' AAACCGGCCTAAGACCTGCCTAGG 3'
gRNA4 Fwd	5' CACCGGTCTTAGGCCGAGGCCACC 3'
gRNA4 Rev	5' AAACGGTGGCCTCGGCCTAAGACC 3'
scramble Fwd	5' CACCGACTCGCGCGAGTCGAGGAG 3'
scramble Rev	5' AAACCTCCTCGACTCGCGCGAGTC 3'

Table 3: Oligos used for gRNA cloning. Each gRNA sequence was 20bp and included an overhang (in red) necessary for ligation in BbsI site.

Generation of AAV Vector Plasmids

The plasmids used for AAV vector production derived from the pAAV2.1[153] plasmid that contains the ITRs of AAV serotype 2. Specifically, I used a pAAV2.1 plasmid generated by our group for a previous publication.

The pAAV2.1-U6-gRNA1-RHO-GFP and pAAV2.1-U6-scramble-RHO-GFP plasmid were generated from an AAV2.1-RHO-GFP plasmid already present in the lab, in which the gRNA expression cassette was cloned upstream of the RHO promoter by ligation in an AflII restriction site. The pAAV2.1-gRNA5-RHO-EGFP plasmid was generated by Daniela Benati and Clarissa Patrizi (University of Modena and Regio Emilia) following a similar cloning strategy.

The kozak-DsRed and IRES-DsRed donor DNA vectors were generated by PCR-amplification of Discosoma red fluorescent protein (DsRed) CDS (675 bp) and the bovine growth hormone polyadenylation signal (bGHpA), from the plasmids generated in a previous publication from our group [154], adding the kozak or IRES [155] sequences, as well as the 5' and 3' gRNA target sites, as PCR primer overhangs. PCR fragments were subcloned in PCR-Blunt II-TOPO (Invitrogen, Carlsbad, California, United States) before cloning in the pAAV2.1 plasmid by Infusion (Takara, Kusatsu, Japan) using an AflII restriction site.

m*Rho*-gRNA , p*RHO*-gRNA, m*Alb*-gRNA (see Table 4) were designed using Benchling gRNA design tool (www.benchling.com), selecting the gRNAs with higher predicted on-target and off-target score targeting the first exon of each gene. gRNAs were then generated as Fwd and Rev oligonucleotides (Table 5), annealed and cloned in PX458 (pCbh-SpCas9-2A-GFP) as described by Zhang's lab [57]. gRNA expression cassettes (including scramble gRNA described above) were PCR-amplified and subcloned in PCR-Blunt II-TOPO (Invitrogen, Carlsbad, California, United States) before cloning in the pAAV2.1 plasmid by In-Fusion cloning (Takara, Kusatsu, Japan) using a NheI restriction site.

gRNA target	gRNA sequence	Endogenous PAM	Target Strand
murine *Rho* (m*Rho*)	GCAGCCGCAGTACTACCTGG	CGG	*Fwd*
porcine *RHO* (p*RHO*)	AGTACTGCGGATACTCAAAG	GGG	*Rev*
Murine *Alb* (m*Alb*)	ACAAGAGTGAGATCGCCCAT	CGG	*Fwd*
scramble (scr)	GACTCGCGCGAGTCGAGGAG	NGG	-

Table 4: gRNAs sequences used to target Cas9 to the rhodopsin and albumin locus in different species.

Oligo name	sequence
***mRho* Fwd**	5' CACCGCAGCCGCAGTACTACCTGG 3'
***mRho* Rev**	5' AAACCCAGGTAGTACTGCGGCTGC 3'
***pRho* Fwd**	5' CACCGAGTACTGCGGATACTCAAAG 3'
***pRho* Rev**	5' AAACCTTTGAGTATCCGCAGTACT 3'
***mAlb* Fwd**	5' CACCGACAAGAGTGAGATCGCCCAT 3'
***mAlb* Rev**	5' AAACARGGGCGATCTCACTCTTGT 3'

Table 5: Oligos used for HITI gRNA cloning. Each gRNA sequence was 20bp and included an overhang (in red) necessary for ligation in BbsI site.

pAAV2.1-IRBP-SpCas9-spA plasmid was generated by cloning the interphotoreceptor retinol-binding protein (IRBP) promoter to the commercial pAAV-pMecp2-SpCas9-spA (Addgene PX551[156]) using HindIII and AgeI restriction sites and conventional ligation. The pAAV2.1-IRBP-VQRCas9HF1-SpA plasmid was generated by Daniela Benati and Clarissa Patrizi (University of Modena and Regio Emilia) following a similar strategy.

pAAV2.1-HLP-SpCas9-spA plasmid was generated in collaboration with Hristiana Lyubenova from the Auricchio group by substituting the IRBP promoter in the previous primer with the hybrid liver promoter HLP) using AflII and AgeI restriction sites and In-Fusion cloning (Takara, Kusatsu, Japan).

Plasmids used as Cas9 templates

The pAAV2.1-CMV-hRHOP347S-3'UTR-bGH plasmid was generated by PCR-amplification of the 3' UTR region of the human rhodopsin using DNA extracted from HEK293 cells. PCR fragment was then cloned into a pAVV2.1-CMV-hRHOP347S-bGH present in our lab by ligation using BsmBI and NotI restriction sites. The pAAV2.1-CMV-m*Rho*P23H-bGH plasmid used was kindly provided by the group of Enrico Surace.

AAV Vector Production and Characterization

AAV vectors were produced by the TIGEM AAV Vector Core by triple transfection of HEK293 cells followed by two rounds of $CsCl_2$ purification [157]. For each viral preparation, physical titers (GC/mL) were determined by averaging the titer achieved by dot-blot analysis [158] and by PCR quantification using TaqMan (Applied Biosystems, Carlsbad,CA, USA) [157]. The probes used for dot-blot and PCR analyses were designed to anneal with the IRBP promoter for the pAAV2.1-IRBP-SpCas9-spA vector, the HLP promoter for the pAAV2.1-HLP-SpCas9-spA vector and the bGHpA region for the donor DNA vectors. The length of probes varied between 200 and 700 bp.

Culture and Transfection of HEK293 Cells

HEK293 cells were maintained in DMEM containing 10% fetal bovine serum (FBS) and 2 mM L-glutamine (Gibco, Thermo Fisher Scientific, Waltham, MA, USA). Cells were plated in 6-well plates ($1*10^6$cells/well), and transfected 16 hr later with the plasmids encoding for Cas9 and the different gRNAs and donor DNAs, using the calcium phosphate method (1 to 2mg/1*10*6cells); medium was replaced 4 hr later. Maximum material transfected was 3ug. In all cases, quantity of plasmid DNA was equilibrated between wells, using an empty vector when necessary.

Cytofluorimetric Analysis

HEK293 cells, plated in 6-well plates, were washed once with PBS, detached with trypsin 0.05% EDTA (Thermo Fisher Scientific, Waltham, MA USA), washed twice with PBS, and resuspended in sorting solution containing PBS, 5% FBS and 2.5 mM EDTA. Cells were analyzed on a BD FACS ARIA III (BD Biosciences, San Jose, CA, USA) equipped with BD FACSDiva software (BD Biosciences) using appropriate excitation and detection settings for EGFP and DsRed. Thresholds for fluorescence detection were set on untransfected cells, and a minimum of 10,000 cells/sample were analyzed. A minimum of 50,000 GFP+ or GFP+/DsRed+ cells/sample were sorted and used for DNA extraction.

Animal Models

Mice were housed at the TIGEM animal house (Pozzuoli, Italy) and maintained under a 12-hr light/dark cycle. C57BL/6J mice were purchased from Envigo Italy SRL (Udine, Italy). P347S mice were kindly provided by Enrico Surace. P347S transgenic mice were maintained as F0 by crossing them with themselves, and were crossed with C57BL/6 mice to generate experimental mice.

Rho-P23H knock-in [159] (referred as P23H) mice were imported from The Jackson Laboratory. The P23H mouse is a knock-in for the P23H mutation in the 1st exon of the murine *Rho* gene, inserted together with a Neomicin cassette. Mice were maintained by crossing homozygous females and males. Experimental heterozygous animals were generated by crossing homozygous P23H mice with C57BL/6 mice. The genotype of mice was confirmed by PCR analysis on genomic DNA (extracted from the mouse phalanx tip). Homozygous mice presented a 530bp PCR product, while heterozygous mice presented a 530bp and a 399bp product. Wildtype mice presented only a 399bp PCR product. The primers used for the PCR amplification are as follows:

Fwd: 5'- TGGAAGGTCAATGAGGCTCT-3'

Rev: 5'- GACCCCACAGAGACAAGCTC-3

The MPSVI mouse is a knock-out for the ARSB gene and a transgenic that expresses a truncated human ARSB, which generates tolerance for the therapeutic ARSB protein. This recapitulates the immunological state of MPSVI patients. MPSVI mice were maintained as heterozygotes and crossed to produce homozygote knock-out experimental mice. The genotype of mice was confirmed by PCR analysis on genomic DNA (extracted from the mouse phalanx tip). Knockout mice presented a 1400bp PCR product, while heterozygous mice presented both the 1400bp and another 234bp band. Wildtype mice presented only the 234bp band. The primers used for the PCR amplification are as follows:

Fwd: 5'- TGGGCAGACTAGGTCTGG -3'

Rev: 5'- TGTCTTCCACATGTTGAAGC -3'

The Large White female pigs (Azienda Agricola Pasotti, Imola, Italy) used in this study were registered as purebred in the LWHerd Book of the Italian National Pig Breeders' Association and were housed at the Centro di Biotecnologie A.O.R.N. Antonio Cardarelli (Naples, Italy) and maintained under a 12-hr light/dark cycle.

Subretinal Injection of AAV Vectors in Mice and Pigs

This study was carried out in accordance with the Association for Research in Vision and Ophthalmology Statement for the Use of Animals in Ophthalmic and Vision Research and with the Italian Ministry of Health regulation for animal procedures (Ministry of Health authorization number 147/2015-PR). Surgery was performed under general anesthesia, and all efforts were made to minimize animal suffering.

Mice (1 to 4 weeks old) were anesthetized with an intraperitoneal injection of 2 mL/100g of body weight of ketamine/xylazine, then AAV2/8 vectors were delivered subretinally via a trans-scleral trans-choroidal approach, as described by Liang et al [160]. Eyes were injected with 1uL of vector solution. The AAV2/8 dose (GC/eye) was between $1*10^9$ and $2,5*10^9$ GC of each vector/eye; thus, co-injection resulted in a maximum of $5*10^9$GC/eye.

Subretinal delivery of AAV2/8 vectors to the pig retina was performed as previously described [121]. Eyes (n = 3) were injected with two blebs of 100uL of AAV2/8 vector solution. The AAV2/8 dose was $2*10^{11}$ GC of each vector/bleb; thus, co-injection of triple AAV vectors resulted in a total of $4*10^{11}$ GC/bleb and $8*10^{11}$ GC/eye.

Neonatal Intravenous Injection via the Temporal Vein

p1-p2 neonatal mice were injected following the protocol published by Gombash Lampe et al. [161]. Total volume was 35uL. The dose injected was 4*10^13 GC/Kg or 6*10^13GC/Kg of each vector, for a total dose of 8*10^13 GC/Kg or 1,2*10^14GC/Kg.

Intravenous Injection via the Retroorbital Plexus

4-week old C57BL/6 mice were injected via the retroorbital plexus with a total volume of 360uL Injection dose was $1,3*10^{13}$ or $4*10^{13}$ GC/Kg for each vector.

Electrophysiological Recordings

For electroretinographic analyses, P347S or P23H mice were dark-adapted for 3 hr. Mice were anesthetized and positioned in a stereotaxic apparatus, under dim red light. Pupils were dilated with a drop of 0.5% tropicamide (Visufarma, Rome, Italy), and body temperature was maintained at 37.5 degrees

Light flashes were generated by a Ganzfeld stimulator (CSO, Costruzione Strumenti Oftalmici, Florence, Italy). The electrophysiological signals were recorded through gold-plate electrodes inserted under the lower eyelids in contact with the cornea. The electrodes in each eye were referred to a needle electrode inserted subcutaneously at the level of the corresponding frontal region. The different electrodes were connected to a two-channel amplifier. After completion of responses obtained in dark-adapted conditions (scotopic), the recording session continued with the purpose of dissecting the cone pathway mediating the light response (photopic). To minimize the noise, different responses evoked by light were averaged for each luminance step. The maximal scotopic response of rods and cones was measured in dark conditions (scotopic) with two flashes of 0.7 Hz and a light intensity of 20 cd s/m2, photopic cone responses were isolated in light conditions with a continuous background white light of 50 cd s/m2, with 10 flashes of 0.7 Hz and a light intensity of 20 cd s/m2.

Pupillary light response analysis

Pupillary light responses from P347S mice were recorded in dark condition using the TRC-50IX retinal camera connected to a charge-coupled device NikonD1H digital camera (Topcon Biomedical Systems). Mice were exposed to light-stimuli at different lux for approximately 10 seconds and one picture per eye was acquired using the IMAGEnet software (Topcon Biomedical Systems). For each eye, the pupil diameter was normalized to the eye diameter (from temporal to nasal side).

Assay for ARSB Enzymatic Activity Evaluation in Serum

Serum ARSB activity was measured by an immune capture assay based on the use of a specific anti-hARSB polyclonal antibody (Covalab), as previously described [150]. Briefly, 96-well plates (Nunclon) were coated with 5 µg/mL in 0.1 M $NaHCO_3$ (100 µL/well) and incubated overnight (O/N) at 4°C. The following day, plates were blocked with 1% milk; after 2 hr of incubation, 50 µL standard and unknown samples (diluted 1:10) was added to each well. Plates were incubated at 4°C O/N. The following day, 100 µL 5 mM 4-methylumbelliferylsulfate potassium salt (4-MUS; Sigma-Aldrich) substrate was added to each well and then incubated at 37°C for 4 hr. The reaction was stopped by the addition of 100 µL stop solution/well (0.2 M glycine). Plates were shaken for 10 min at room temperature and fluorescence was read (excitation of 365 nm/emission of 460 nm) on a multiplate fluorimeter (Infinite F200; TECAN). Serum ARSB was determined based on a rhARSB (Naglazyme; BioMarin Europe) standard curve and is expressed as picograms per milliliter.

Quantitative Analysis of GAG Accumulation in Urine

Urine samples were diluted 1:50 in water to measure GAG content. was used for the GAG assay, as previously described [162]. GAG concentrations were determined on the basis of a dermatan sulfate standard curve (Sigma-Aldrich). Tissue GAGs were expressed as micrograms of GAG per milligram of protein. Urinary GAGs were normalized to creatinine content, which was measured with a creatinine assay kit (Quidel, San Diego, USA). Thus, the units of urinary GAGs are given in micrograms of GAG per micromole of creatinine. Urinary GAGs are reported as the percentage of AF control mice.

Serum Albumin Quantification

Serum samples were analyzed with the mouse albumin ELISA kit (Abcam, Cambridge, UK) following manufacturer's instructions. Samples were diluted 30.000 times and albumin was determined on the basis of the competition of biotinylated albumin to bind to the plate. Serum albumin was expressed as milligrams of albumin per milliliter of serum.

Retinal Dissection

To isolate the temporal and nasal regions of the retina, after mice sacrifice, the temporal area of the eye was cauterized. After eye harvesting, eyes were dissected under a Leica M205FA Stereomicroscope (Leica, Wetzlar, Germany) to confirm GFP fluorescence of the temporal area. Temporal and nasal areas were dissected in two separate tubes.

Histology and Light and Fluorescence Microscopy

In vitro fluorescence imaging

To evaluate DsRed expression after HITI in vitro, HEK293 cells, plated in 6-wells at a density of $1*10^6$ were transfected as previously described. Forty-eight hours post-transfection, cells were washed once with PBS, fixed for 10 min with 4% paraformaldehyde (PFA) in PBS, washed three times with PBS, and mounted with Vectashield with DAPI (Vector Lab, Peterborough, UK). Cells were analyzed under an Axio Observer Z1 (Carl Zeiss, Oberkochen, Germany) equipped with ZEN software (Carl Zeiss) and using appropriate excitation and detection settings for EGFP, DsRed, and DAPI.

Retinal cryosections and fluorescence imaging

To evaluate DsRed expression in the retina after HITI in histological sections, C57BL/6J mice and large white pigs [121] were injected subretinally with IRBP-Cas9 and donor DNA AAV vectors. One month later, mice and pigs were sacrificed and eyes were fixed in 4% paraformaldehyde overnight and infiltrated with 30% sucrose overnight; the cornea and the lens were then dissected, and the eyecups were embedded in optimal cutting temperature compound (O.C.T. matrix; Kaltek, Padua, Italy). Ten-micrometer-thick serial retinal cryosections were cut along the horizontal meridian, progressively distributed on slides, and mounted with Vectashield with DAPI (Vector Lab, Peterborough, UK). Then, cryosections were analyzed under a confocal LSM-700 microscope (Carl Zeiss, Oberkochen, Germany), using appropriate excitation and detection settings. For assessment of HITI efficiency in mouse retinal cryosections following AAV administration, the highest transduced area of two sections/eye was selected and acquired at 40 magnification and then analyzed using ImageJ software (http://rsbweb.nih.gov/ij/). A minimum of 500 PRs, identified by DAPI staining, were counted for each image. PRs with signal compatible with DsRed expression were unequivocally identified based on their shape as observed in z-stacks of the analyzed sections, as well as the presence of DsRed+ outer segments.

Evaluation of Retinal Outer Nuclear Layer thickness

To evaluate retinal outer nuclear layer thickness, P347S or P23H mice were injected subretinally with IRBP-Cas9 and gRNA or donor DNA AAV vectors, respectively. Two or three months later, mice were sacrificed and eyes were fixed in Davidson's fixative (deionized water, 10% acetic acid, % 20% formalin, 35% ethanol) overnight dehydration in serial ethanols and then embedded in paraffin blocks. Ten-micrometer thick microsections were cut along the horizontal meridian, progressively distributed on slides and stained with hematoxylin-eosin. Then, the sections were analyzed under the microscope (Leica Microsystems GmbH; DM5000) and acquired at 20x magnification. For each eye one image from the temporal injected side of a slice in the central region of the eye was used for the analysis. Three measurements of the ONL thickness were taken in each image, masked to the genotype/treatment group, using the "freehand line" tool of the ImageJ software.

Liver cryosections and fluorescence imaging

To evaluate DsRed expression in the liver after HITI, C57BL/6J mice were injected either at p2 or at 4 weeks of age, and sacrificed one month after injection by cardiac perfusion. Liver was harvested and photographed in a Leica Stereomicroscope (Leica, Wetzlar, Germany) at 25 magnification. Then, a small piece of each lobe was dissected, and all pieces were fixed in 4% paraformaldehyde overnight and infiltrated with 15% sucrose overday and 30% sucrose overnight before being included in O.C.T. matrix (Kaltek, Padua, Italy) for cryosectioning. Five-micrometer-thick retinal cryosections were cut and distributed on slides and mounted with Vectashield with DAPI (Vector Lab, Peterborough, UK). Then, cryosections were analyzed under a confocal LSM-700 microscope (Carl Zeiss, Oberkochen, Germany), using appropriate excitation and detection settings. For assessment of HITI efficiency in mouse liver cryosections, 3 images of each liver were acquired ad 20 magnification and then analyzed using ImageJ software (http://rsbweb.nih.gov/ij/).

A minimum of 850 hepatocytes, identified by DAPI staining of the nucleus, was counted for each image. Hepatocytes with signal compatible with DsRed expression were unequivocally identified based on their shape.

DNA Cleavage Analysis

DNA extraction

Samples (GFP+ or GFP+/DsRed+ sorted HEK293 cells, retinal tissue or liver tissue) were lysed in commercial lysis buffer (GeneArt™ Genomic Cleavage Detection Kit, Invitrogen, Carlsbad, California, United States) or conventional lysis buffer for DNA extraction from tissue (400mM NaCl, 1% SDS, 20mM TRIS-CL (pH 8.0), 5mM EDTA (pH 8.0)) respectively. Lysis buffers were supplemented with proteinase K, which was inactivated after lysis for 15 minutes at 80 degrees.
50 to 200ng of DNA were used for PCR amplification of the region comprising the Cas9 target site (the first exon of RHO) from the pCMV-m*Rho*-P23H plasmid or from the mouse genome, respectively. Primers used are shown in Table 6:

Primer name	Sequence
P347S-Indel Fwd:	5' TGGATTTGAGTGGATGGGGC 3'
P347S-Indel Rev:	5' GTGGATGTCCCTTCTCAGGC 3'
pCMV-m*Rho*-Indel Fwd:	5' CCATGGTGATGCGGTTTTGG 3'
pCMV-m*Rho*-Indel Rev:	5' ATGTAGTTGAGGGGTGTGCG 3'
m*Rho*-HITI-Indel Fwd:	5' CAGTGCCTGGAGTTGCGCTG 3'
m*Rho*-HITI-Indel Rev:	5' GGGCCCAAAGACGAAGTAGCC 3'
p*RHO*-Indel-Fwd:	5' AGGCCTCAGCAGCATCCTTG 3'
p*RHO*-Indel-Rev:	5' GTGGTGGTGAAGCCTCCGAA 3'
m*Alb*-HITI-Indel Fwd:	5' ATTACGGTCTCATAGGGCCTGC 3'
m*Alb*-HITI-Indel Rev:	5' GCACACATTTCTACTGGACAGCA 3'

Table 6: Primers used to generate PCR fragments for Surveyor Assay and sequencing of INDELs.

P347S-Indel primers produced a 444bp PCR product. pCMV-m*Rho*-Indel primers produced a 634bp PCR product. m*Rho*-HITI-Indel primers produced a 426bp PCR product. p*RHO*-Indel primers produced a 341bp PCR product. m*Alb*-HITI-Indel primers produced a 592bp PCR product.

Surveyor Assay

1-3uL of the PCR products (according to PCR efficiency) was used for Surveyor Assay, following the GeneArt™ Genomic Cleavage Detection Kit manufacturer's recommendations. In short, DNA was de-annealed at 99ºC and re-annealed by a slow temperature gradient in a thermocycler. After re-annealing, 1uL of Detection Enzyme (T7Endonuclease) was added, and samples were incubated for an hour at 37 degrees. After incubation, samples were run in a 2% agarose gel in order to detect DNA cleavage products resulting from INDEL presence.

Tracking of INDELs by Decomposition

m*Rho*-HITI-Indel, p*RHO*-Indel and *mAlb*-HITI-Indel PCR products were also used for Sanger sequencing. Sequences were then used for TIDE software (https://tide.deskgen.com/) analysis of INDEL frequency.

HITI junction Characterization

Junction PCR Amplification

DNA extracted from retina or liver tissue was used for PCR amplification of HITI junctions. Both 5' and 3' junctions of integration were amplified. For the 5' junction, I used a forward primer recognizing the region upstream the first exon of the m*Rho* or *mAlb* gene and a reverse primer recognizing the DsRed coding sequence. For the 3' junction we designed a forward primer recognizing the bGH polyA sequence of the donor DNA, and a reverse primer recognizing the first Exon of m*Rho* or the second exon of *mAlb*, after the cleavage site. Table 7 shows the primers used.

Primer name	Sequence
***mRho* HITI 5' junction Fwd**	5' CAGTGCCTGGAGTTGCGCTG 3'
***mRho* HITI 5' junction Rev**	5' GGCTTGATGACGTTCTCAGTGC 3'
***mRho* HITI 3' junction Fwd**	5' CGACCTGCAGAAGCTTGGATCT 3'
***mRho* HITI 3' junction Rev**	5' GGGCCCAAAGACGAAGTAGCC 3'
***mAlb* HITI 5' junction Fwd**	5' GCCTGCTCGACCATGCTATACT 3'
***mAlb* HITI 5' junction Rev**	5' CCTTGGAGCCGTACTGGAACTG 3'
***mAlb* HITI 3' junction Fwd**	5' CGACCTGCAGAAGCTTGGATCT 3'
***mAlb* HITI 3' junction Rev**	5' TCTCTGGCTGCCACATTGCT 3'

Table 7: Primers used for HITI junction amplification.

5' junction primers produced a 663bp PCR product. 3' junction primers produced a 455bp PCR product. Both PCR products were cloned into PCR-Blunt II-TOPO (Invitrogen, Carlsbad, California, United States) and single clones were sequenced to confirm the identity of the PCR products before NGS analysis.

Library preparation and next-generation sequencing:

For library preparation a total of 47.5 ng of DNA from HITI junction PCR products was used as input for the synthesis of a DNA library with the SMART-Seq v4 Ultra Low Input RNA Kit for sequencing (Takara Bio USA, Mountain View, CA, USA). Manufacturer suggested protocol was followed, with minor modifications. Seventyfive pg of cDNA generated with SMART-Seq v4 Kit were used for library preparation using the NEXTERA XT DNA Library Preparation kit (Illumina, San Diego, CA, USA), following the suggested protocol. Quality of libraries was assessed by using Bioanalyzer DNA Analysis on a DNA high sensitivity chip (Agilent Technologies, Santa Clara, CA, USA), and quantified by using Qubit 4 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). Samples were sequenced using NextSeq 500/550 Mid Output v2 kit in a 150 + 150 paired-end run. The data were deposited in GEO:GSE10717. Illumina base call raw data were converted in fastq file through bcl2fastq software (version v2.20.0.422, Illumina, San Diego,USA).

Sequence reads were trimmed using Cutadapt software to reduce reads length from 600bp to 350bp. The custom reference sequence was built by predicting the NHEJ-mediated integration mediated by Cas9 in the locus of interest. The cleaved donor DNA sequence was inserted between the -3 and -4 bases of the gRNA sequence. All sequencing reads were aligned with their respective reference sequence using BWASW software and INDELs at each position respective to the cleavage site were quantified. Alignment was performed using BWA-SW software (MIT, Boston, USA). Total number of indel and specific nucleotide counts contained in aligned bam files were estimated with deepSNV [163] package. Length of INDELs was obtained from BAM CIGAR strings through an ad-hoc R algorithm. INDELs with a frequency lower than 0,5% of reads were not included in graphical representations of INDEL length.

HITI off-target Characterization

Library preparation and next-generation sequencing:

SureSelect custom probes designed to enrich the specific regions of interest (Donor DNA for mRho, donor DNA for m*Alb*, the m*Rho* locus, the m*Alb* locus and the AAV sequences) with supervision of Agilent technical support. Libraries were generated following the instructions from the manufacturer "SureSelect^{QXT} Target Enrichment for Illumina Multiplexed Sequencing, Version E0, April 2018" .All experiments were performed following the manufacturer's protocols, (Agilent Technologies, Santa Clara, CA, USA). 50 ng of genomic DNA was fragmented and adaptors were added in a single enzymatic step. The adaptor-tagged DNA library was purified and amplified. Next, 750 ng of each library was hybridized using SureSelect^{QXT} capture library for 90 minutes. The resulting libraries were recovered using streptavidin magnetics beads, and a post-capture PCR amplification was carried out.

Final libraries were quantified with a Qubit High Sensitivity kit (Invitrogen, Carlsbad, CA) and library quality was assessed on a Bioanalyzer High Sensitivity DNA chip (Agilent). Sixteen samples were included per run. Sequencing was carried out on a NextSeq instrument (Illumina) according to the manufacturer's protocol ("Denature and Dilute Libraries Guide", December 2018). Final library concentration was 2.2 pM, which was used to carry out cluster generation and sequencing on a NextSeq-Mid Output Flow Cell (2 × 150 cycles) with 10% of phiX.

Reads were aligned to the mouse genome using Integrative Genomcis Viewer (Broad Institute, Cambridge, USA) and reads containing the donor DNA were mapped using an ad hoc algorithm.

Statistical Analysis

All results are expressed as mean ± standard error (SE). Statistical comparisons were made using either *t*-test or one-way analysis of variance (ANOVA); the Tukey post hoc test was used. Statistical significance was considered if $P < 0.05$.

RESULTS:

Specific Aim 1: Development of an allele-specific genome editing approach specific to the P347S mutation.

gRNA design for allele-specific Cas9 targeting:

The aim of this part of the project was to identify one or several gRNAs that were able to target SpCas9 specifically to the P347S *RHO* allele without targeting the WT *RHO* allele. I theorized that, to achieve allele-specific targeting, the P347S mutation must be part of the gRNA target sequence or the PAM. Since SpCas9 has an NGG PAM, I was not able to include the C to T mutation as part of the PAM. For this reason, I designed 4 gRNAs that used all the available NGG PAMs next to the mutation. gRNA1, gRNA2 and gRNA4 were designed specifically to recognize the mutation, while gRNA3 was designed to target the area adjacent to the mutation but without including it in its sequence. This way, gRNA3 should not be able to discriminate between P347S and WT alleles and could be used as a positive control for DNA cleavage of both. As a negative control, I generated scramble-gRNA, which does not recognize any sequence in neither the murine nor the human genome (Figure 11).

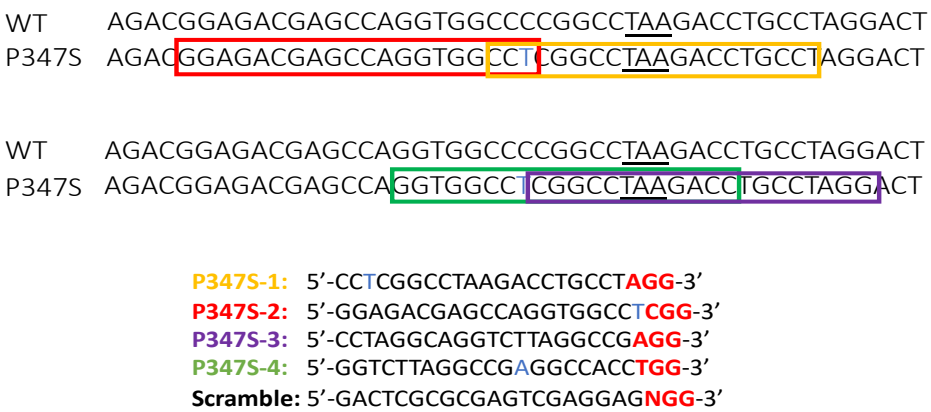


Figure 11. gRNA design for allele-specific targeting P347S *RHO*: *RHO* WT and P347S sequences are depicted. gRNAs 1, 2 and 4 were designed to have the mutation site as part of their sequence. gRNA 3 did not recognize the mutation site, and was used as a control of non-allele-specific targeting. Scramble gRNA was designed to not recognize any sequence in neither the murine nor the human genome. PAM sequences for each gRNA are depicted in red. The P347S mutation is depicted in blue. RHO STOP codon is underlined. PAM sequences are depicted in red.

In vitro gRNA testing reveals an allele-specific gRNA:

In order to test whether the designed gRNAs were able to perform allele-specific cleavage of P347S RHO, I decided to test them *in vitro*. Since I didn't have a cellular model with the P347S mutation, I generated a Cas9 template plasmid carrying *RHO* with the P347S mutation and the 3'UTR region of the *RHO* gene, which was necessary since some of the gRNAs recognized part of the 3'UTR (Figure 11). I then generated 5 different SpCas9-T2A-EGFP-U6-gRNA plasmids, each carrying one of the gRNAs I had designed. These plasmids encode for SpCas9 and EGFP, which are expressed together and equimolarly under control of the Chicken beta actin hybrid (Cbh) promoter, but remain as separate protein products because of a ribosomal skipping sequence called T2A. Additionally, they carry the gRNA expression cassette, comprised of the U6 promoter, the 20bp gRNA sequence and the rest of the gRNA scaffold necessary to bind the gRNA to Cas9 (Figure 12).

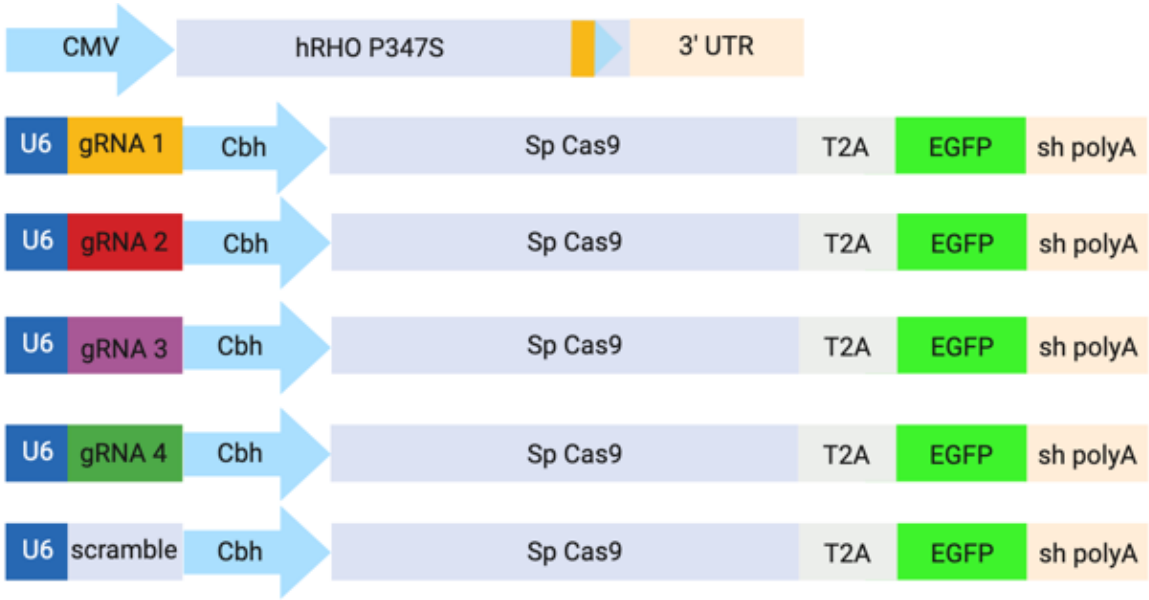


Figure 12. Schematic depiction of plasmids used for gRNA testing *in vitro*: For P347S targeting, a plasmid encoding *hRHO* P347S under control of the CMV promoter, as well as the 3'UTR, was generated. Plasmids carrying the gRNA expression cassette for each gRNA, as well as SpCas9-T2A-EGFP under control of the Cbh promoter, were also generated. CMV: cytomegalovirus promoter, 3'UTR: 3' untranslated region, Cbh: chicken beta actin hybrid promoter, T2A: ribosomal skipping sequence, sh polyA: short synthetic polyA.

For P347S-specific targeting, I transfected HEK293 cells with the *RHO*P347S-3'UTR and the SpCas9-T2A-EGFP-U6-gRNA plasmids. For WT-specific targeting, I transfected HEK293s only with the SpCas9-T2A-EGFP-U6-gRNA plasmid, to target WT *RHO* in the cell genome.

Forty-eight hours after transfection, I used FACS to sort EGFP+ cells and extracted their DNA. I then used Surveyor Assay to assess the ability of each gRNA to specifically cleave P347S and WT *RHO*. As expected, all gRNAs except for scramble were able to cleave P347S *RHO*. Instead, WT *RHO* was cleaved by gRNA2, gRNA3 and gRNA4, but not by gRNA1 or scramble (Figure 13). This shows that gRNA1 is able to specifically target SpCas9 to the P347S RHO without targeting WT *RHO*.

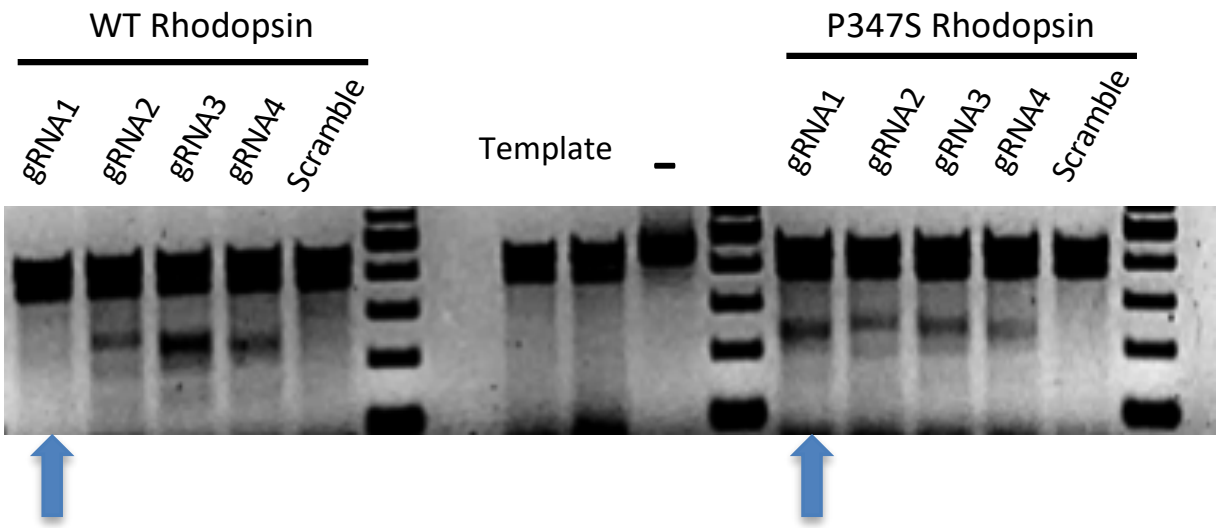


Figure 13. gRNA screening for P347S *RHO*: Cells were transfected with Cbh-Cas9-2A-EGFP plasmid carrying one of 4 different gRNAs designed to specifically recognize P347S *RHO*, or a scramble gRNA as control. For P347S targeting, a CMV-P347S*RHO*-3'UTR plasmid was transfected as well. EGFP+ cells were sorted and their DNA was extracted. The genomic region around the target sequence was amplified by PCR and used in Surveyor Assay. Expected sizes of cleavage products: gRNA1: 235+209bp, gRNA2: 252+192bp, gRNA: 241+203bp, gRNA4: 233+211bp. Blue arrows mark the allele-specific gRNA. 100bp ladder was used.

Thus, I selected gRNA1 as the only allele-specific gRNA. Our collaborators Clarissa Patrizi, Daniela Benati and Alessandra Recchia (Università di Modena e Reggio Emilia, UNIMORE) then performed a deep *in vitro* characterization of gRNA1 as well as another gRNA designed by them, which they called gRNA5. gRNA5 is specific for a variant of Cas9, called VQR-Cas9-HF1. This Cas9 variant recognizes a different PAM (NGAG) and has been shown to have lower off-target cleavage. Thus, I hypothesized that VQR-Cas9-HF1 and gRNA5 could have a better allele-specificity than SpCas9 and gRNA1. This was confirmed *in vitro* by collaborators in UNIMORE, as gRNA1 (6%) cleaved WT *RHO* more frequently than gRNA5 (1%) (data not shown). However, gRNA1 cleaved P347S *RHO* significantly more efficiently (61%) compared with gRNA5 (29%) (data not shown).

***RHO* P347S targeting *in vivo* is efficient and reduces mRNA levels in photoreceptors:**

Next, I decided to assess whether using gRNA1 and gRNA5 to target Cas9 to cleave P347S RHO in the retina of a mouse model could have a therapeutic effect. For that, AAV2/8 vectors were generated. The first two vectors encode for SpCas9 or VQR-Cas9-HF1 under control of the interphotoreceptor retinol-binding protein (IRBP) promoter. The other vectors carry gRNA expression cassettes for gRNA1, gRNA5 or scramble-gRNA, as well as EGFP under control of the *RHO* promoter (Figure 14). I combined AAV2/8-IRBP-SpCas9 with either AAV2/8-gRNA1-RHO-EGFP or AAV2/8-scramble-RHO-EGFP, and AAV2/8-IRBP-VQR-SpCas9-HF1 with either AAV2/8-gRNA5-RHO-EGFP or AAV2/8-scramble-RHO-EGFP (Figure 14).

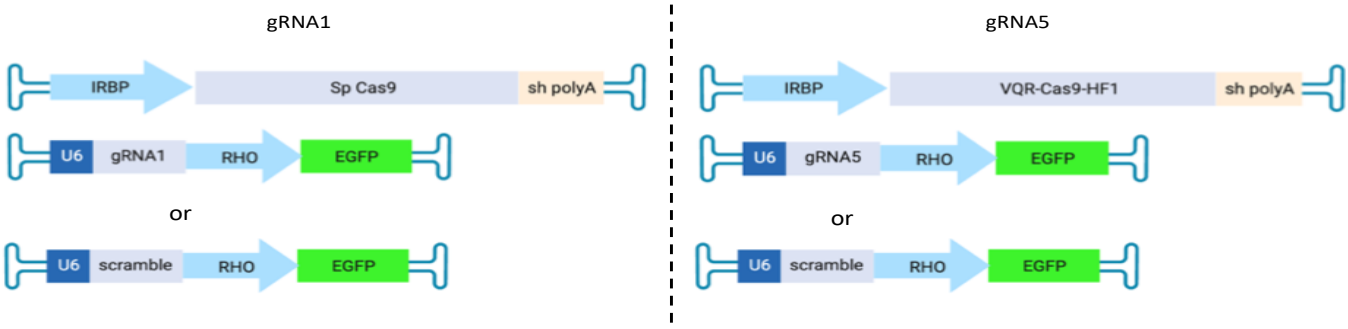


Figure 14: Schematic depiction of generated AAVs: AAV2/8-IRBP-SpCas9 was injected with either gRNA1 or scramble. AAV2/8-IRBP-VQR-Cas9-HF1 was injected with either gRNA5 or scramble. IRBP: interphotoreceptor retinol-binding protein promoter, shpolyA: short synthetic polyA, RHO: rhodopsin promoter.

P347S mice received subretinal injection of $2*10^9$ GC of each vector at p13 or p16. At p40, retinae were harvested and genomic DNA was extracted. TIDE analysis was performed by collaborators in UNIMORE to quantify presence of INDELs. INDELS ranging from 1,7% to 17% were detected (data not shown). No INDELs were detected in scramble-treated retinae. To further characterize INDEL frequency, the collaborators in UNIMORE performed next-generation sequencing (NGS) analysis on 5 retinae treated with gRNA1 and 7 retinae treated with gRNA5. With about 30.000 reads per retina, we observed INDELs ranging from 5% to 17% (Figure 15). A single scramble-treated retina was used as a control for NGS, with no INDELS being detected (data not shown).

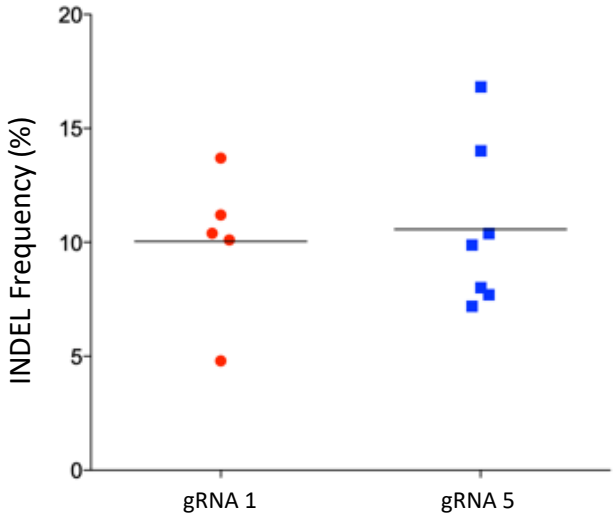


Figure 15: INDEL frequency in P347S retinae: Mice were injected with 2*10^{9}GC of AAV2/8-IRBP-Cas9/VQRCas9HF1 and AAV2/8 carrying gRNA1 (red dots) or gRNA5 (blue squares) respectively. INDEL frequency was quantified by NGS. Means are represented as bars. Performed in collaboration with Clarissa Patrizi, Daniela Benati and Alessandra Recchia, University of Modena and Reggio Emilia.

After this, collaborators at UNIMORE characterized whether the specific cleavage of the P347S *RHO* allele was enough to reduce the presence of mutant *RHO* in photoreceptors. First, they performed RT-PCR analysis of *RHO*P347S mRNA levels. For this analysis, P347S mice received subretinal injection of 1*10^{9}GC of each vector at p7. At p40, retinae were harvested and the nasal and temporal parts were isolated. mRNA levels were measured and normalized to the levels observed in the temporal portion of the contralateral retina, which was injected with the scramble gRNA. Clear reduction of mRNA levels was observed in the temporal (injected) region of gRNA-treated retinae when compared both to the untreated nasal region and to the contralateral scramble-treated retinae (Figure 16). This reduction was significant in 2 out of 3 analyzed eyes for gRNA1, and 3 out of 5 analyzed eyes for gRNA5. Interestingly, two eyes injected with gRNA5 showed reduction of RHOP347S mRNA also in the nasal region, which suggested a very efficient injection that had allowed sufficient vector diffusion to transduce also the nasal region of the retina (Figure 16).

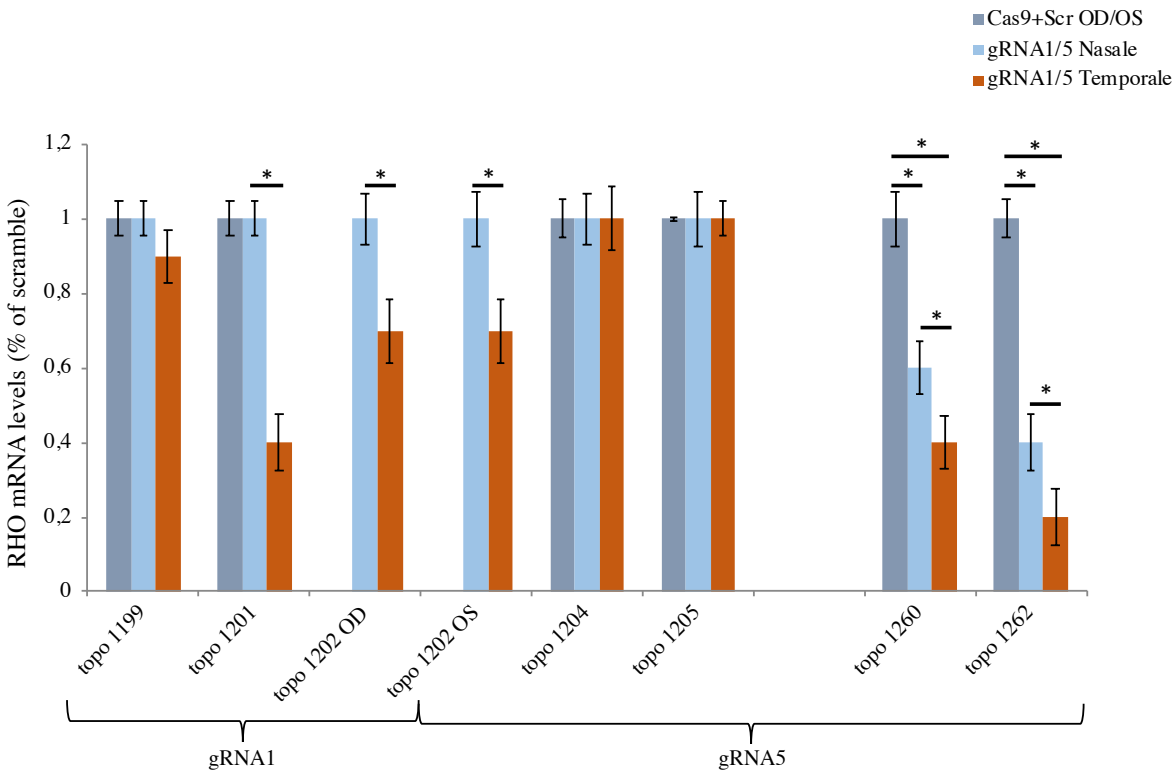


Figure 16. RHO P347S mRNA is reduced in mouse retinae treated with gRNA1 or gRNA5: P347S mice were injected at p7 with 1*10^{9}GC of each vector. Retinae were harvested at p60 and nasal and temporal regions were isolated. RT-PCR was used to quantify levels of RHO P347S mRNA. Values in gRNA treated eyes were normalized to mRNA levels in the temporal region of contralateral scramble eyes. In the case no contralateral eye was available, nasal region of the gRNA treated eye was used for normalization. *=p<0,05. Performed in collaboration with Clarissa Patrizi, Daniela Benati and Alessandra Recchia, University of Modena and Reggio Emilia.

These results show that targeting of P347S *RHO* can efficiently reduce its mRNA levels in the edited photoreceptors, even if cleavage occurs almost at the end of the coding sequence.

RHO P347S targeting partially rescues RP phenotype

To assess whether the reduction of P347S *RHO* mRNA observed was enough to reduce retinal degeneration, P347S mice were injected at p7 with 1*10^{9}GC of each vector, and retinal function was assessed at p40. Electroretinogram (ERG) analysis showed significant improvement of the B-wave amplitude at 20cd.s/m^2 in retinae treated with gRNA1 and gRNA5, compared to their respective scramble controls (Figure 17A). In total, 8 out of 14 eyes treated with gRNA1 or gRNA5 had B-wave values over 320μV, while no scramble-treated eyes did (Figure 17B,C).

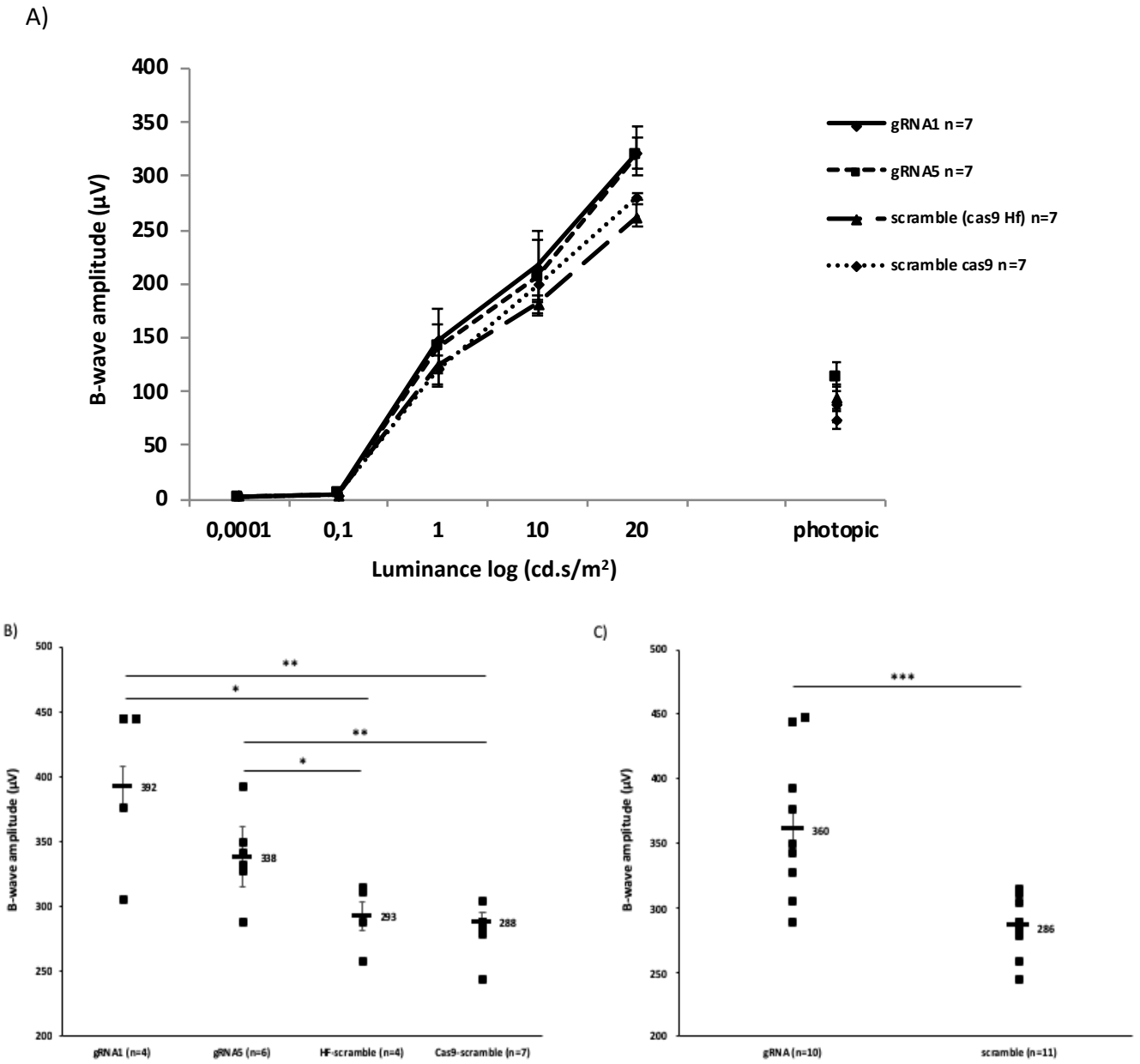


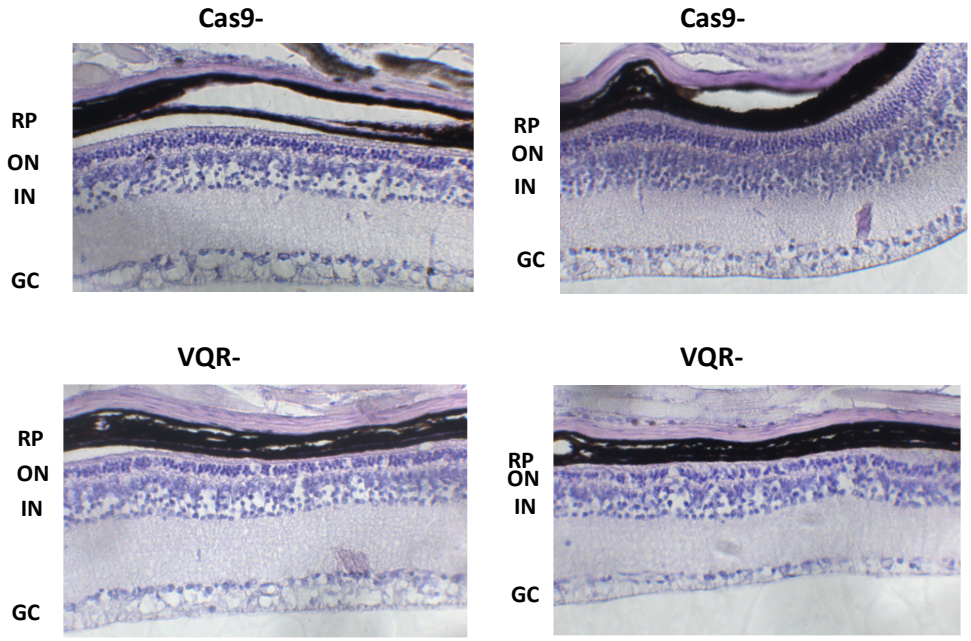
Figure 17. Retinal electrical function improvement after Cas9-mediated knockout of P347S *RHO*: Mice were injected at p7 with 10^9GC of each vector. Electroretinogram analysis was performed at p40. A) ERG B-wave amplitude at different luminance. cd.s/m^2: candles per square meter. B) Dot plot distribution of B-wave amplitude at 20cd.s/m^2. Eyes with an ERG response lower than 235uV were excluded. C) Dot plot depicting B-wave amplitude at 20cd.s/m^2 of treated and untreated eyes in single groups respectively. Black bars represent group averages. *=p<0,05; **=p<0,01; ***=p>0,001

For this analysis, all eyes with ERG response lower than 235uV at 20 cd.s/m^2 were excluded. This arbitrary threshold was set considering the distribution of the analyzed eyes, where only 4 eyes out of 14 in the treatment groups performed under 290 uV, and in all cases they performed under 235uV. Similarly, no eyes in the Cas9-scramble control group performed under 245uV.

Since we did not observe a significant difference in the outer nuclear layer thickness between eyes with higher or lower ERG values, we attributed the lowest values to artifacts in ERG analysis, which can be caused by relative opacity of the cornea, and thus decided to exclude all eyes performing under 235uV.

However, I was unable to see a difference in outer nuclear layer (ONL) thickness between treated and untreated eyes at p50 (Figure 18). Interestingly, the Cas9-scramble group seemed to have a thicker ONL than gRNA1 and gRNA5-treated groups, but also than the VQR-scramble group, which suggests that this difference was due to variability in the observed retinae.

A



B)

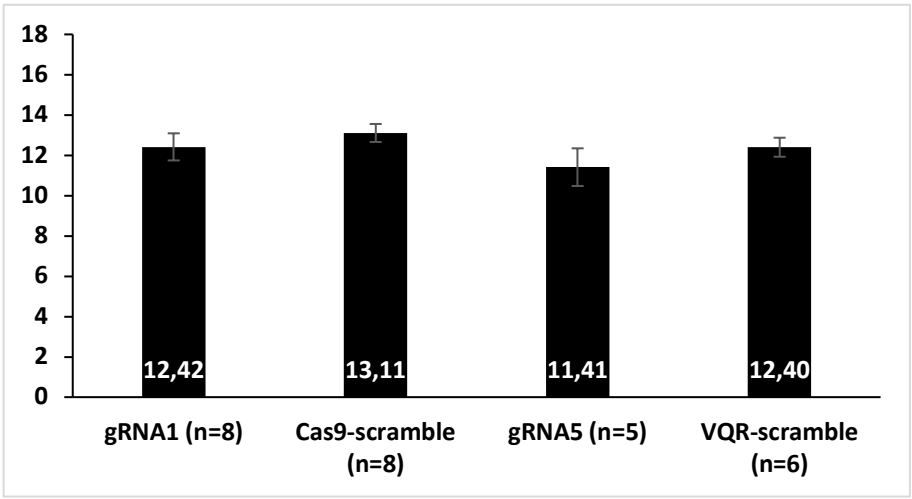


Figure 18. Histological analysis of ONL thickness: A) Representative light microscopy images of hematoxylin-eosin staining of retinas harvested at p50. B) Quantification of ONL thickness in the temporal region of analyzed retinas. *=p<0,05, **=p<0,01

Specific Aim 2

<u>HITI design to target the *mRho* locus</u>

The second aim of my project was to develop an allele-independent approach to treat RP4. For that, I decided to use HITI to integrate a donor DNA in the *mRho* locus. This gRNA has 4 nucleotides of difference with the *hRHO* locus, and a CGG PAM that is absent in *hRHO*, and thus should not generate cleavage in *hRHO*. First, I decided to perform a proof-of-concept experiment using the reporter gene DsRed. I designed a gRNA to target the first exon of the *mRho* gene (Figure 20) and a donor DNA carrying STOP codons in the 3 different frames, a translation START signal (kozak or a small 50bp synthetic IRES [155]), the reporter gene DsRed and the bGH polyA. This donor DNA was flanked by the same gRNA target site in the *mRho* gene but in an inverted orientation. After CRISPR-Cas9 cleavage, I expected that NHEJ would mediate the directional integration of the donor DNA in the DSB generated in *mRho* (Figure 19).

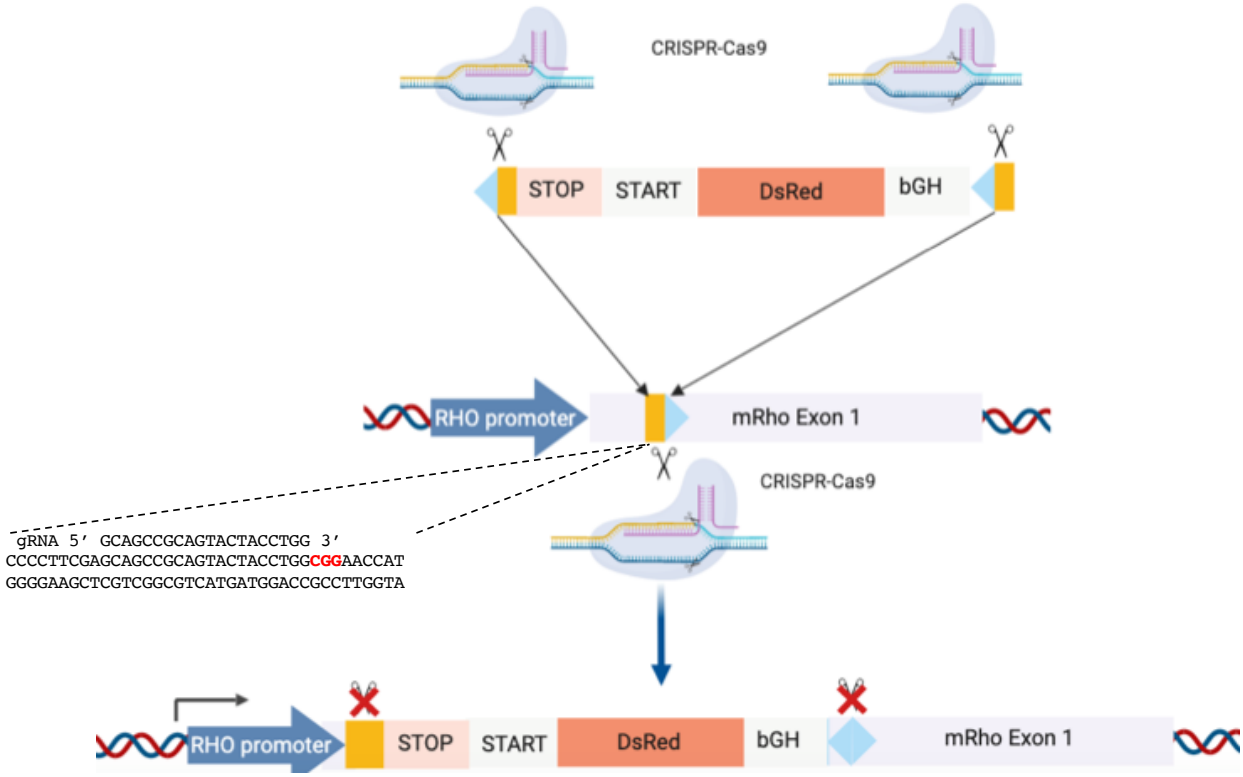


Figure 19. HITI design to target the *mRho* locus: yellow rectangle and blue triangle depict the two parts of the gRNA target sequence (upstream and downstream of Cas9 cleavage). Scissors represent Cas9-mediated DSBs. Crossed scissors represent inability of Cas9 to recognize and cleave the sequence. PAM sequence is depicted in red. STOP: STOP codons in 3 reading frames, START: translation start site, bGH: bovine Growth Hormone polyA.

HITI in the m*Rho* locus is efficient *in vitro*:

To test HITI in the m*Rho* locus *in vitro* I generated plasmids encoding for two versions of the donor DNA with different translation START sites: one carried a kozak sequence while the other carried a small synthetic 50bp IRES. Each donor DNA was cloned with the *Rho*-specific gRNA and a scramble gRNA that doesn't recognize any sequence in the mouse genome. This plasmid was transfected together with the CBh-SpCas9-2A-EGFP plasmid, as well as another plasmid encoding for m*Rho*P23H under control of the CMV promoter (Figure 20). This CMV-m*Rho*P23H plasmid was used as a template for Cas9 cleavage and donor DNA integration.

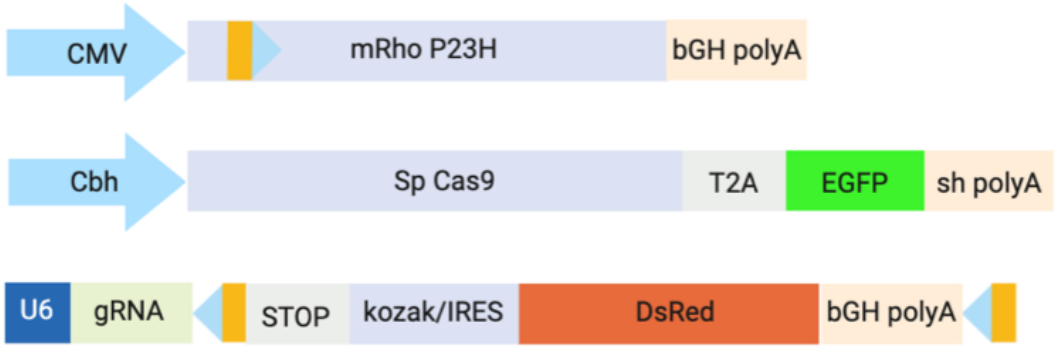


Figure 20. Plasmids used for transfection in HEK293 cells: yellow rectangle and blue triangle depict the two parts of the gRNA target sequence (upstream and downstream of Cas9 cleavage respectively). Cbh: chicken beta-actin hybrid promoter. STOP: STOP codons, START: translation start site, bGH: bovine Growth Hormone polyA.

I expected that, after HITI occurred, DsRed would be integrated into the template CMV-m*Rho*P23H sequence and expressed due to the transcriptional activity of the strong and ubiquitous CMV. Forty-eight hours after transfection in HEK-293 cells, fluorescence microscopy showed abundant presence of DsRed+ cells only in those cells treated with the gRNA and none in those treated with scramble (Figure 21).

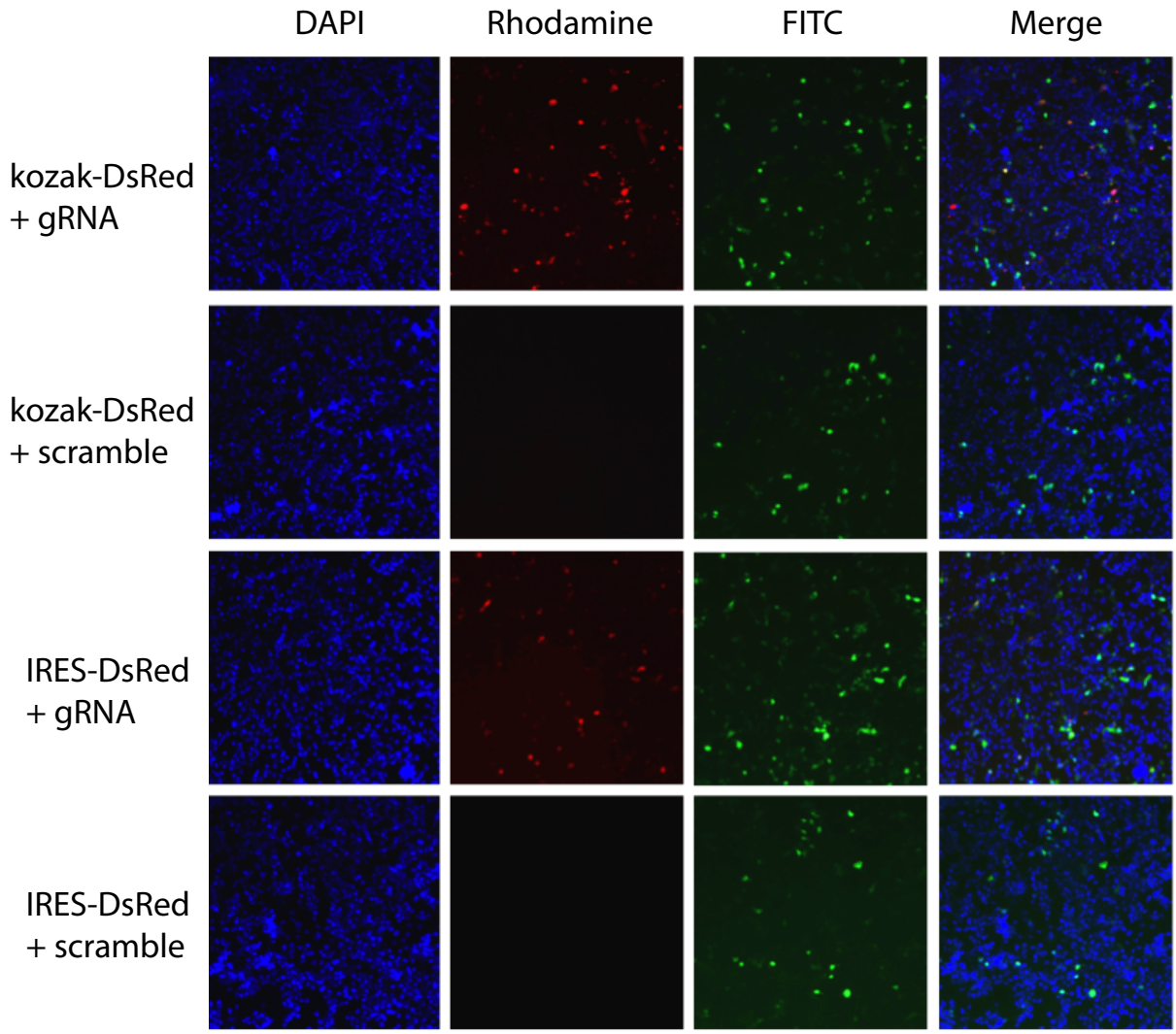


Figure 21. DsRed expression in HEK293 cells modified using HITI: Fluorescence microscopy imaging 48h after transfection. Cells were fixed for 15 minutes in PFA 4% and stained using a DAPI-containing mounting medium.

Since Cas9-transfected cells expressed EGFP, I used fluorescence-activated cell sorting (FACS) to determine the ratio between DsRed+ and EGFP+ cells in order to assess whether kozak or IRES achieved a higher expression of DsRed or a higher efficiency of integration (Figure 22A). No significant differences were found between kozak and IRES, although kozak seemed to perform slightly better (47% compared to 34%) (Figure 22B). Again, FACS demonstrated no presence of DsRed+ cells in those treated with scramble, confirming that DsRed positivity was dependent on gRNA presence.

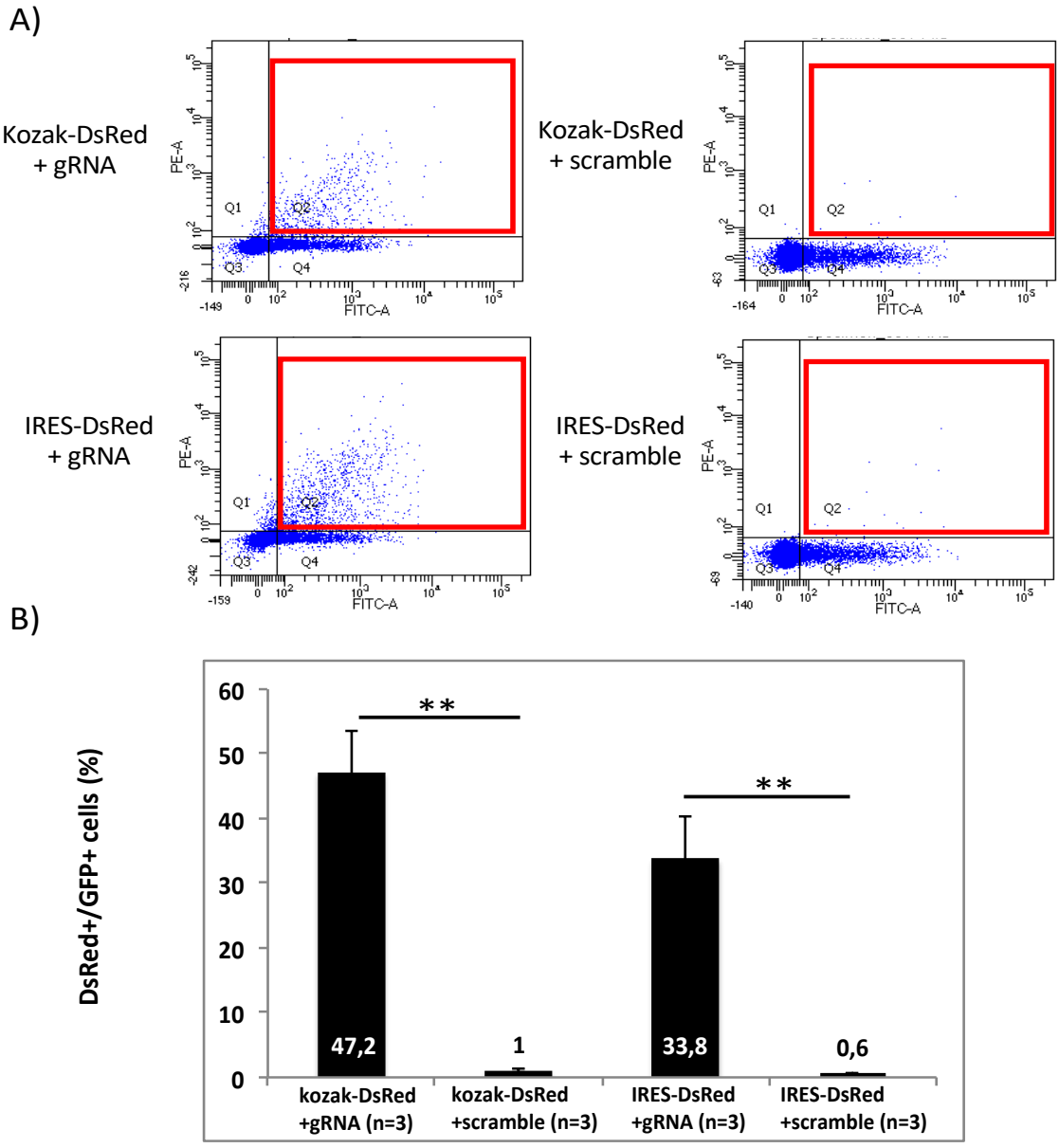


Figure 22. FACS of fluorescent HEK293 cells: Cells were detached 48h after transfection using trypsin 0.05% EDTA. 10.000 cells were counted for each sample. A) Representative FACS graphs showing presence or absence of EGFP+/DsRed+ cells. Q1: EGFP-/DsRed+, Q2: EGFP+/DsRed+, Q3: EGFP-/DsRed-, Q4: EGFP+/DsRed-. Red square depicts sorted cells. PE-A: DsRed fluorescence filter. FITC-A: EGFP fluorescence filter. B) Quantification of DsRed+ cells inside the EGFP+ population.

To confirm that DsRed expression was a result of integration following Cas9 cleavage in the template CMV-*Rho*P23H plasmid, I extracted DNA from EGFP+ sorted cells and PCR-amplified the region of the plasmid around the gRNA target site. Surveyor Assay showed presence of INDELs only in cells treated with gRNA (Figure 23), which correlated with DsRed expression.

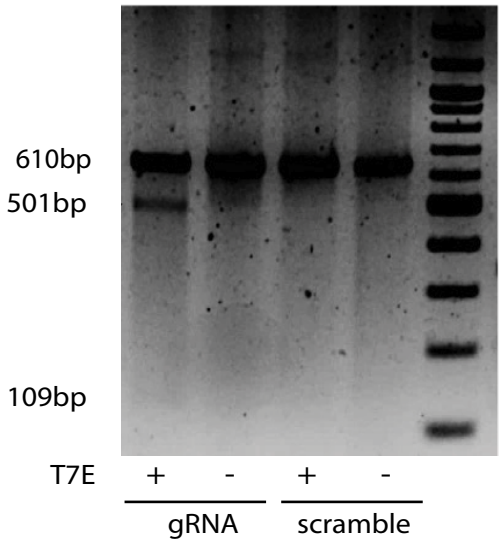


Figure 23: m*Rho*-specific gRNA cleaves the m*Rho* sequence: Surveyor Assay analysis of INDEL presence. Expected band sizes are depicted. T7E: T7 Endonuclease I treatment.

HITI is efficient in mouse photoreceptors:

In order to assess whether HITI was feasible in mouse photoreceptors I generated AAV2/8 vectors carrying the donor DNA with the kozak or the IRES sequence, as well as the gRNA expression cassette for the *mRho*-specific gRNA or the scramble gRNA (Figure 24). These vectors were injected together with the already described IRBP-SpCas9-shpolyA vector.

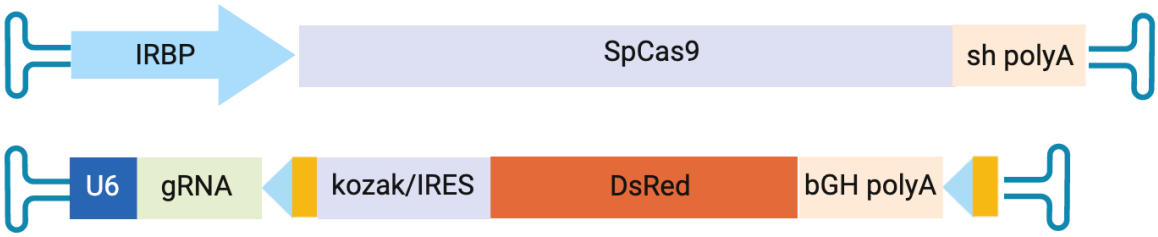
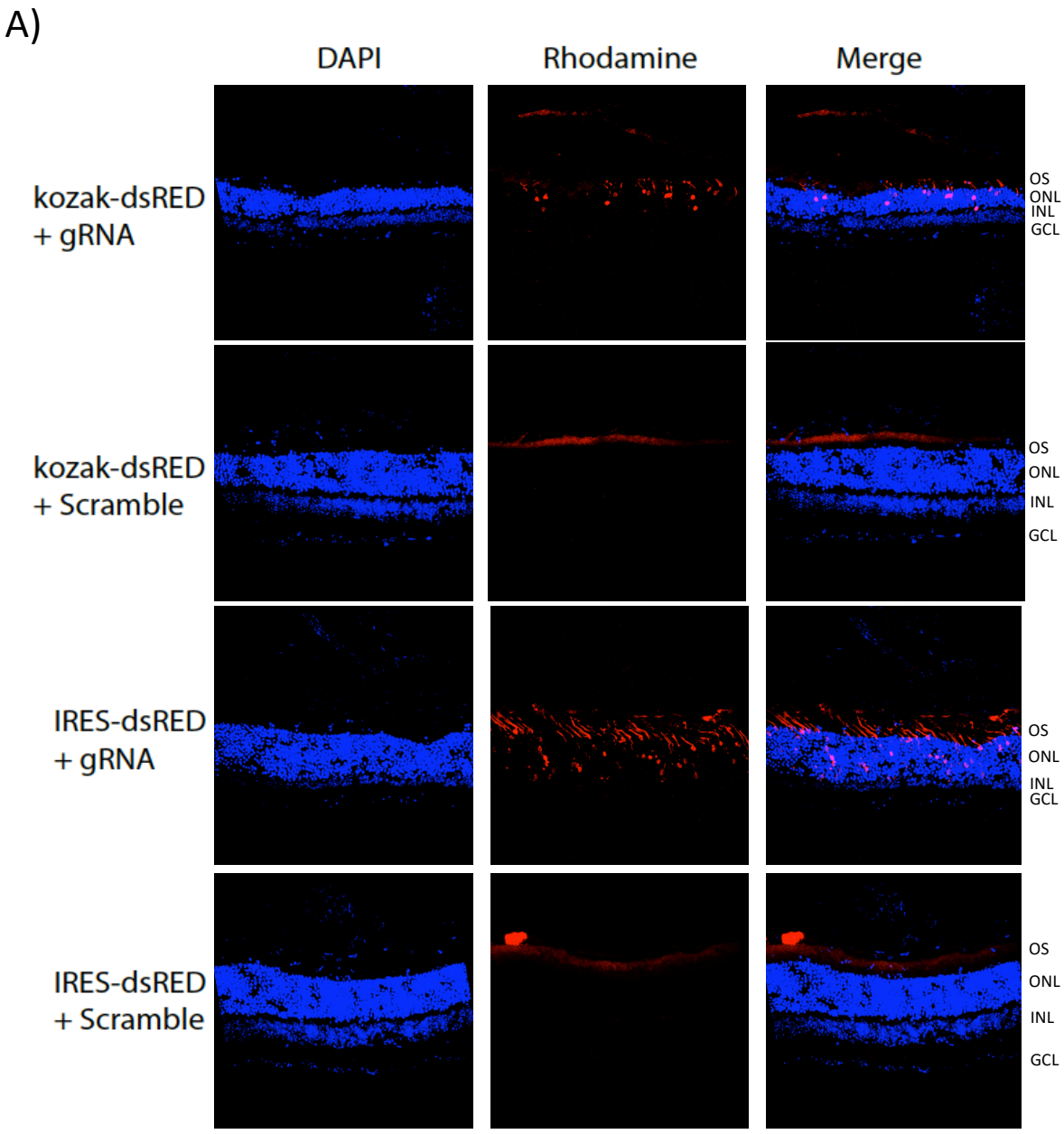


Figure 24. Schematic depiction of AAVs used for HITI in the mouse retina: IRBP: interphotoreceptor retinoid-binding protein promoter. shpolyA: short synthetic polyA. bGH: bovine growth hormone polyA. yellow rectangle and blue triangle depict the two parts of the gRNA target sequence (upstream and downstream of Cas9 cleavage). ITR sequences are depicted as blue loops.

Subretinal injection of 2,5*10^9 GC/eye of each vector was performed in 4-week old C57BL/6 mice. One month after injection, fluorescence microscopy of retinal cryosections showed presence of DsRed+ photoreceptors around the area of injection only in gRNA-treated retinae and not in scramble-treated retinae. Interestingly, IRES seemed to perform significantly better than kozak *in vivo* (9,2% compared to 3%) (Figure 25 A,B)).



B)

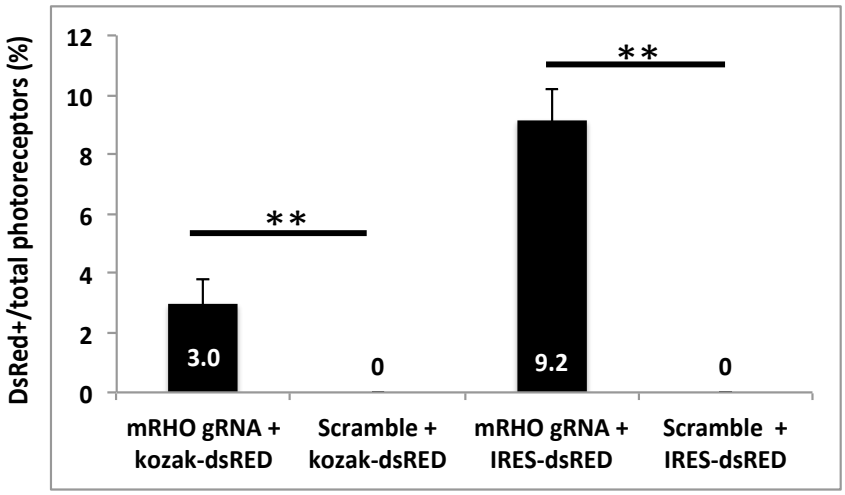


Figure 25. HITI is feasible in mouse photoreceptors. Eyes were harvested 30 days after injection and fixed overnight in 4% PFA. Eyes were treated with 30% sucrose for 6h and then included in optimal cooling medium. 10μm sections were made and stained with DAPI-containing mounting medium. A) Fluorescence microscopy of retinal cryosections: DsRed+ photoreceptors are only present in gRNA-treated retinae. B) Quantification of DsRed+ photoreceptors to assess HITI efficiency. **=p<0,01

Characterization of HITI in photoreceptors:

In order to better understand how precise HITI can be in photoreceptors, I designed two primer pairs to amplify the 5' and 3' junctions of integration (Figure 26A) respectively. Both primer pairs were able to amplify a PCR product of the expected size only in DNA extracted from a gRNA-treated retina, and not in the DNA from a scramble-treated retina. (Figure 26B). A faint 500bp band was detected in the 5' junction in both gRNA and scramble-treated livers, which could be compatible with viral genome integration. However, since this band was not dependent on Cas9 treatment, it was not sequenced. Off-target integration analysis could determine whether Cas9-independent AAV genome integration occurs in both treatment groups.

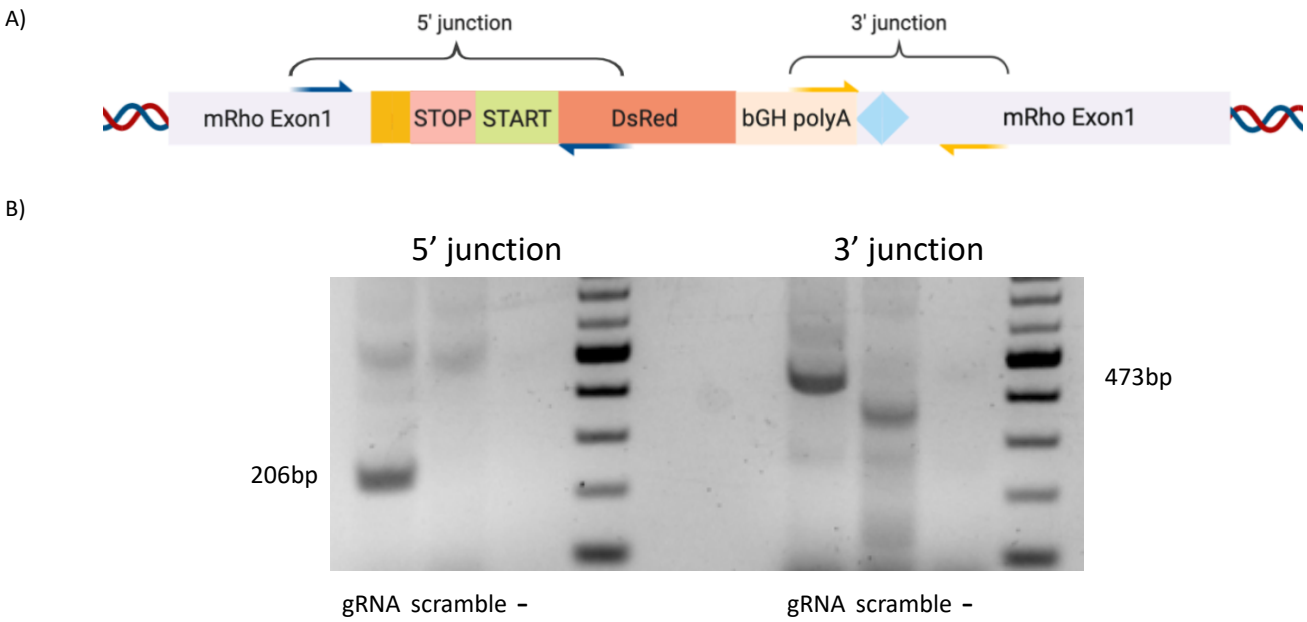


Fig 26. HITI Junction amplification. DNA was extracted from the isolated temporal side of the retina and used for PCR amplification of the HITI junctions. A) Scheme depicting primer design to amplify 5' and 3' junctions. B) 5' junction amplification. Expected fragment size was 206bp. PCR amplification was observed only in one gRNA-treated retina. C) 3' junction amplification. Expected fragment size was 473bp. PCR amplification was observed only in one gRNA-treated retina. Ladder used was 100bp.

PCR fragments were subsequently used for NGS analysis in collaboration with Davide Cacchiarelli (TIGEM). Between 1,5 and 3 million reads were obtained for each junction. Analysis of NGS reads showed low presence of insertions (5%) and deletions (15%) in the 5' junction, especially in the Cas9 cleavage site (position 1) (Figure 27). Most common deletions ranged between 1 and 9bp. Interestingly, I observed two common insertions of 40bp and 48bp sequences that I identified as being part of the AAV ITR sequences (Figure 27). Thus, INDEL-free HITI was quantified at 80% in the 5' junction. Differently, the 3' junction presented surprisingly abundant 1bp insertions (71%), which were mostly cytosine or thymine, some 2bp insertions and almost no presence of deletions (Figure 27). INDEL-free HITI was quantified at 12% in the 3' junction.

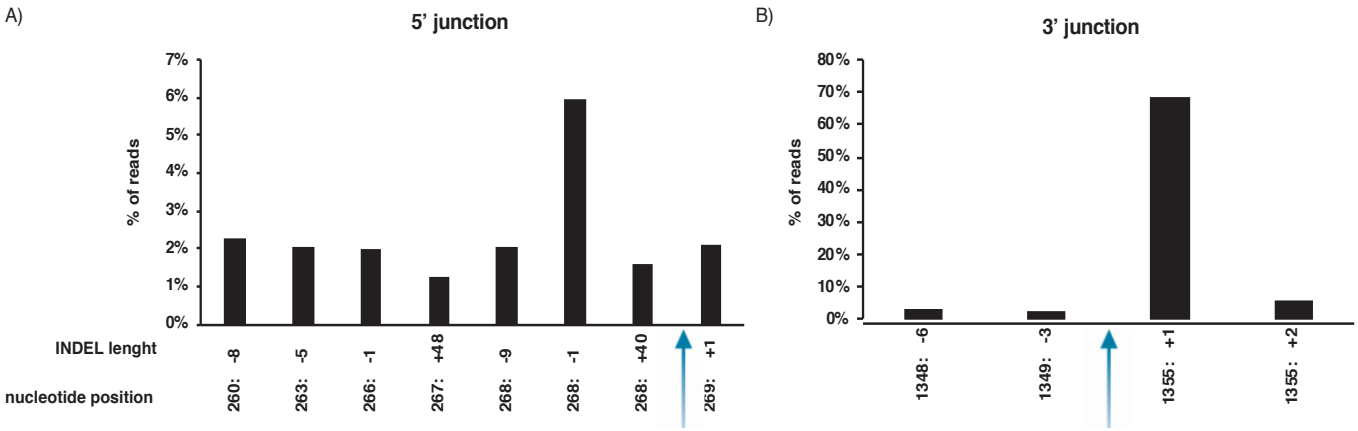
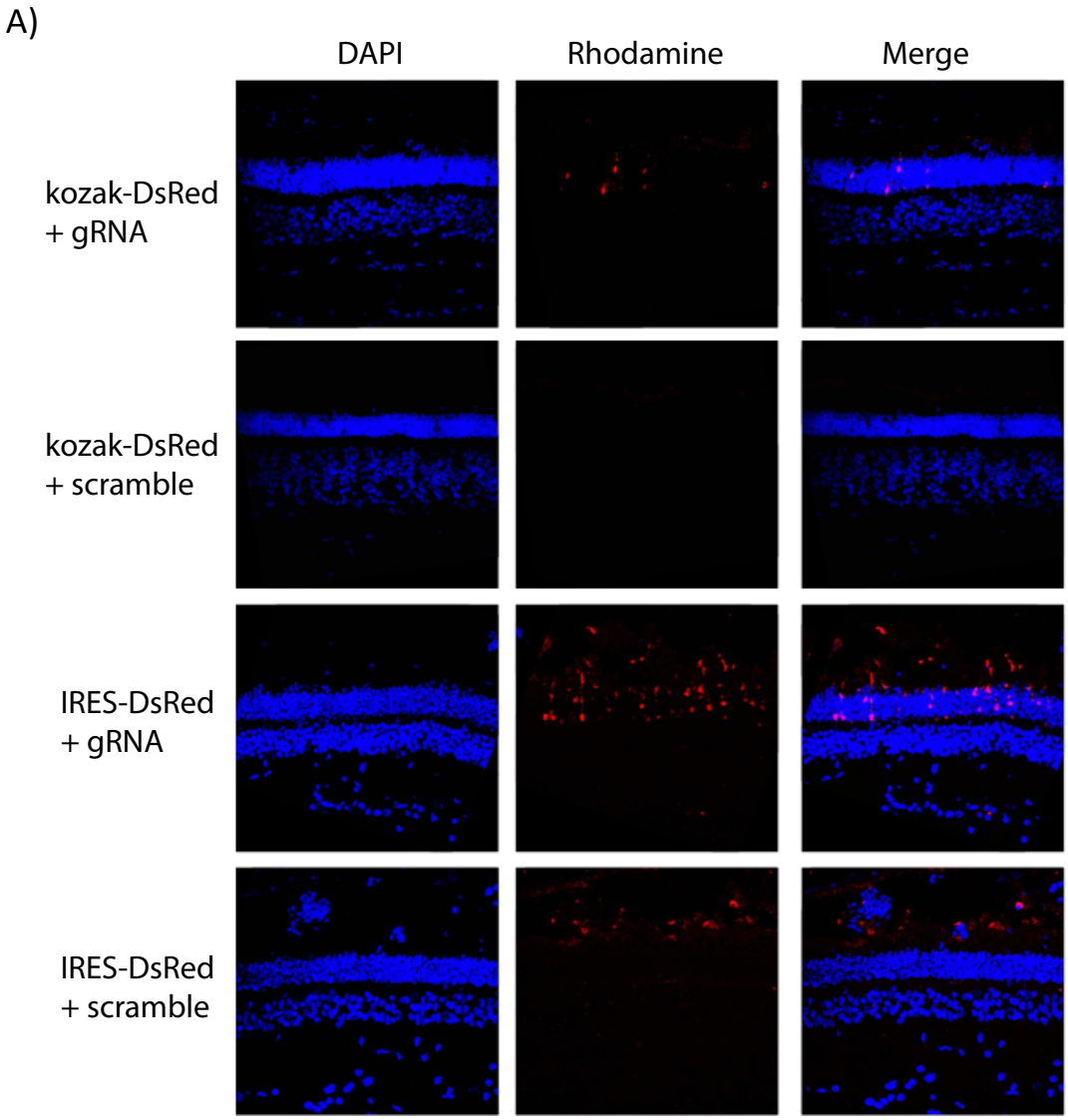


Figure 27. NGS characterization of HITI junctions in *mRho*: Relative INDEL frequency in each position surrounding the cleavage site. Negative and positive numbers represent deletions and insertions respectively. Cleavage site is shown with a blue arrow. Only INDELs with a relative frequency higher than 0,5% in their position are depicted.

Additionally, DNA extracted from mouse retinas was used for NGS analysis of potential off-target HITI integrations. Probes targeting DsRed were used to enrich for sequences in the genome that included the donor DNA. Surprisingly, no genomic sequences containing the donor DNA and compatible with Cas9 cleavage were found, even in the m*Rho* locus. This could be explained by the low number of reads (average of 59.000 reads recognizing the m*Rho* target site) and the low efficiency of integration observed, as well as the inability to select only DsRed+ photoreceptors for DNA extraction, thus diluting the edited genomes in the total DNA extracted.

HITI is efficient in pig photoreceptors:

In order to characterize HITI efficiency in an animal model that recapitulates the anatomy of the human eye better than the murine model, I decided to use the pig eye, which is a relevant preclinical model. For this, I designed a gRNA specific for the first exon of the pig rhodopsin (*pRho*) gene, and donor DNAs identical to the ones used in previous experiments but flanked by *pRho* target sites. With these, I generated the same types of vectors depicted in Figure 14. Subretinal injection of $2{,}5*10^{11}$GC/bleb of each vector was performed in 3-month-old large white pigs in 2 blebs per each retina. One month after injection, fluorescence microscopy of retinal cryosections showed presence of DsRed+ photoreceptors in the area of injection in gRNA-treated retinae and not in scramble-treated retinae (Figure 28A). Again, IRES performed significantly better than kozak (Figure 28A,B).



B)

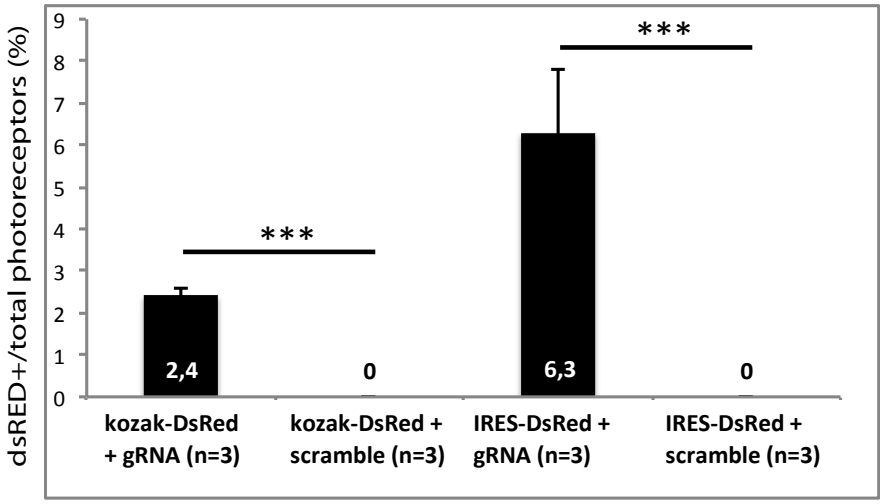


Figure 28. HITI is feasible in pig photoreceptors. A) Fluorescence microscopy of porcine retinal cryosections: DsRed+ photoreceptors were only present in gRNA-treated retinae. B) Quantification of DsRed+ photoreceptors to assess HITI efficiency in the pig retina. ***=p>0,001

In order to determine whether DsRed expression correlated with Cas9-mediated DSBs in the *pRho* gene, I extracted genomic DNA from pig retinae and RPE and performed Surveyor Assay. I observed presence of INDELs only in gRNA-treated retinae and not in scramble-treated retinae. As I expected, INDELs were present only in the retina and absent in the RPE, which is consistent with photoreceptor-specific expression of Cas9 by the IRBP promoter (Figure 29).

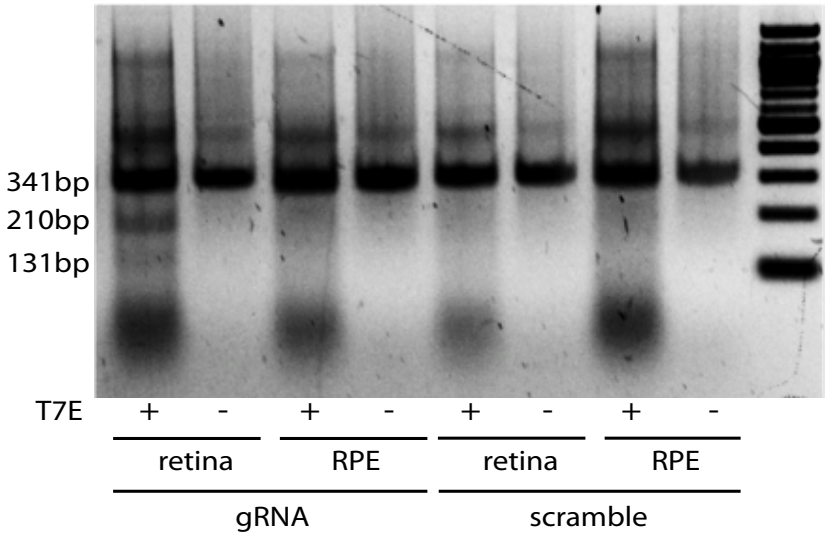


Figure 29. Surveyor Assay in the *pRho* locus using DNA extracted from retina or RPE: T7E: T7 Endonuclease I treatment. Sizes of PCR product and expected cleavage bands are depicted.

In order to better characterize the efficiency of cleavage, I used TIDE analysis, for which the same PCR fragments used for Surveyor Assay were used for SANGER sequencing. Resulting chromatograms were used as input for the TIDE software. The chromatogram from an RPE sample from a scramble-treated eye was used as a negative control, and other chromatograms were compared to it. TIDE showed 22% efficiency of cleavage in gRNA-treated retinae and negligible INDELs in scramble-treated retinae (Figure 30 A,B).

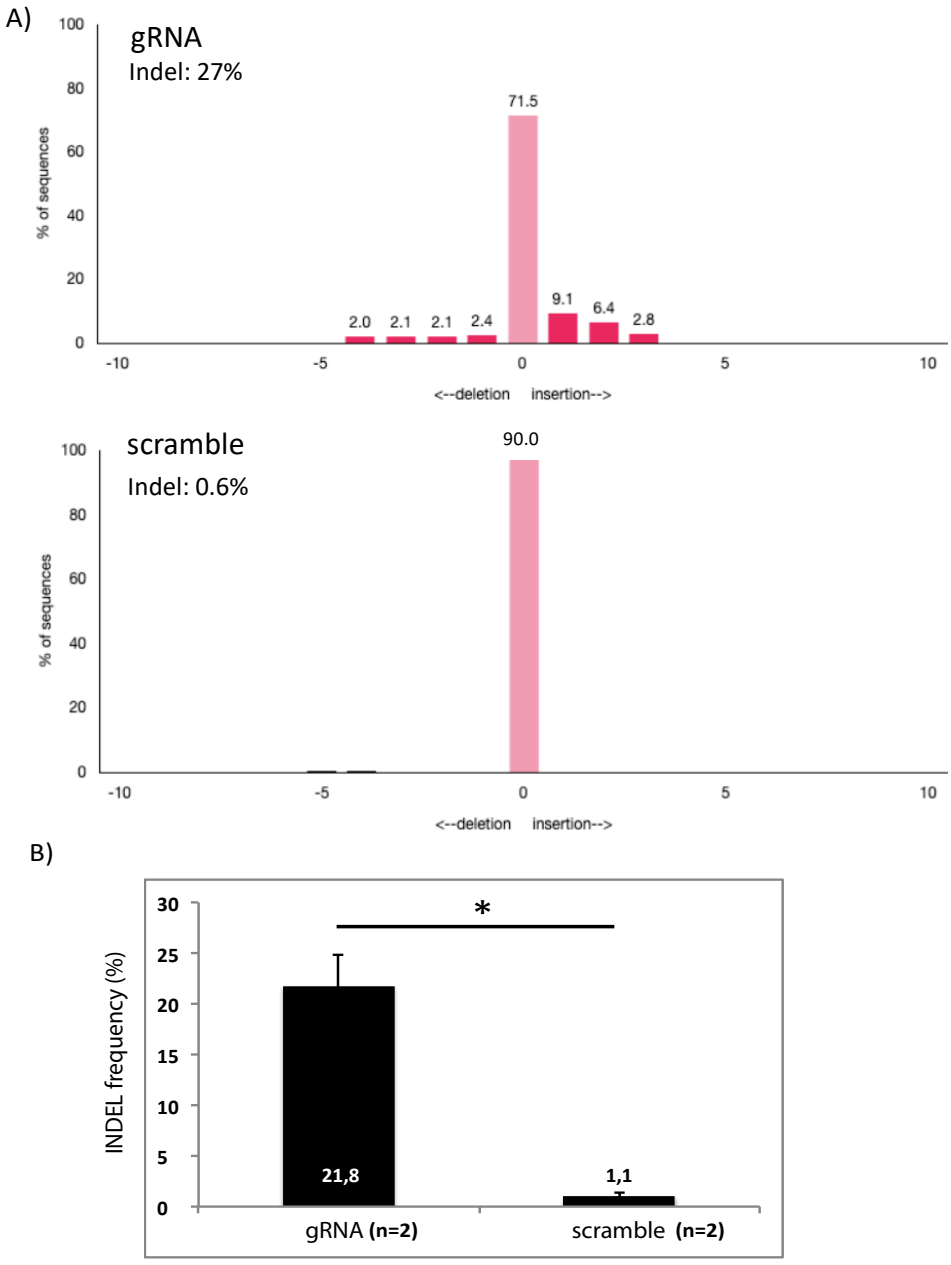


Figure 30. Cas9 generates INDELS only in the retina of gRNA-treated eyes: TIDE analysis of the p*Rho* locus: PCR products amplified from DNA extracted from retina or RPE were sequenced. RPE sequence was used as a template for TIDE. Chromatograms were used to deconstruct frequency and type of INDELs. A) Representative TIDE result. B) Quantification of INDEL frequency in pig retinae.

HITI at the *mRho* locus partially restores retinal function in a mouse model of AdRP:

To assess whether HITI efficiency was enough to correct the AdRP phenotype, I generated two AAV2/8 vectors carrying a donor DNA with the IRES sequence and the coding sequence of the *hRHO* gene (*hRHO* CDS), as well as the *mRho* or scramble gRNA (Figure 31). These vectors were injected together with the AAV2/8-IRBP-SpCas9 vector (Figure 31).

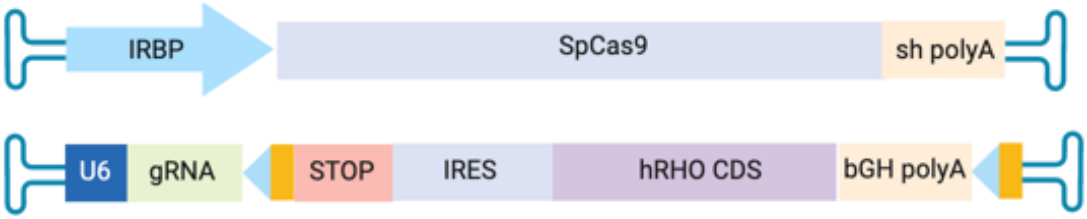


Figure 31. Schematic depiction of AAV vectors used for therapy in P23H+/- mice: ITRs are depicted in blue. IRBP: interphotoreceptor retinoid-binding protein promoter. shpolyA: short synthetic polyA. *hRHO* CDS: coding sequence of the human *RHO* gene, bGH: bovine growth hormone polyA. yellow rectangle and blue triangle depict the two parts of the gRNA target sequence (upstream and downstream of Cas9 cleavage).

Twenty-day old P23H+/- mice received subretinal injection of $2{,}5*10^9$GC of each vector. At p60, ERG analysis showed small but significant improvement at the two highest light stimuli analyzed (Figure 32).

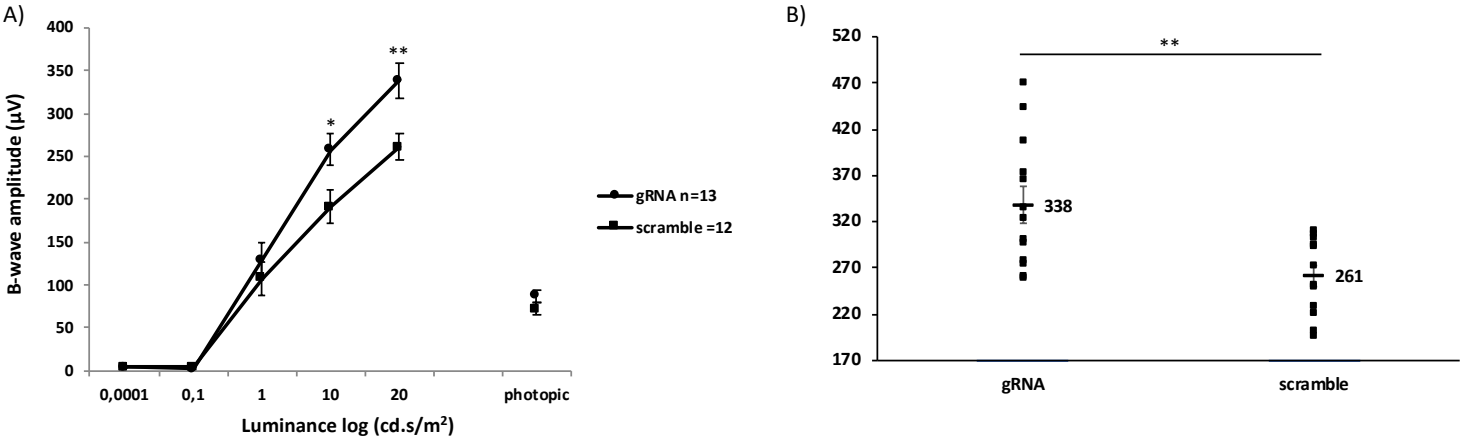


Figure 32. Improvement of ERG B-wave in P23H mice at p60. Cd.s/m² = candles per square meter. A) ERG measurements at different luminance. B) Dot plot of ERG measurements at 20cd.s/m². *=p<0,05, **=p<0,01

However, the observed improvement in ERG was only transient and disappeared at p90 and later timepoints (data not shown). Indeed, histological analysis at p120 showed no significant difference in retinal outer nuclear layer (ONL) thickness between gRNA and scramble groups (Figure 33).

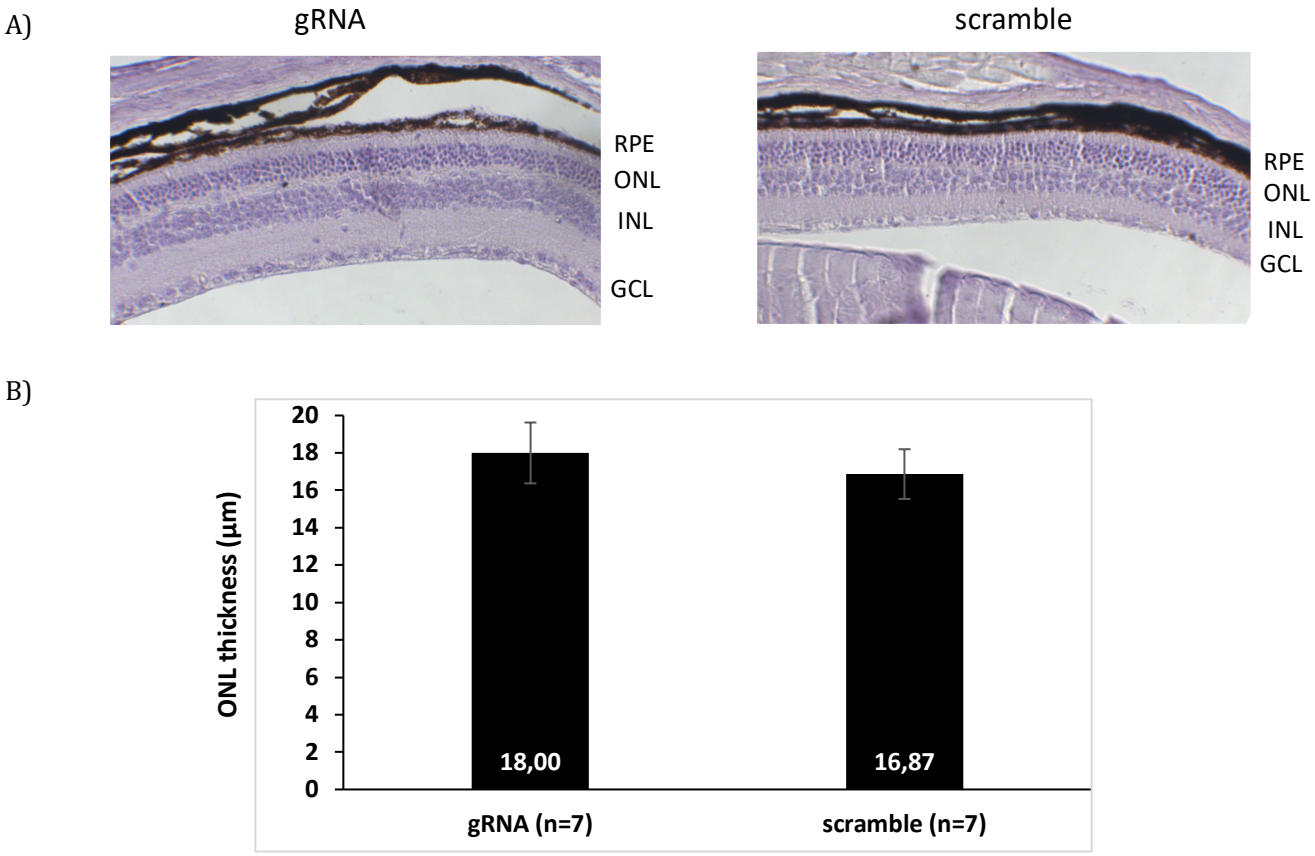


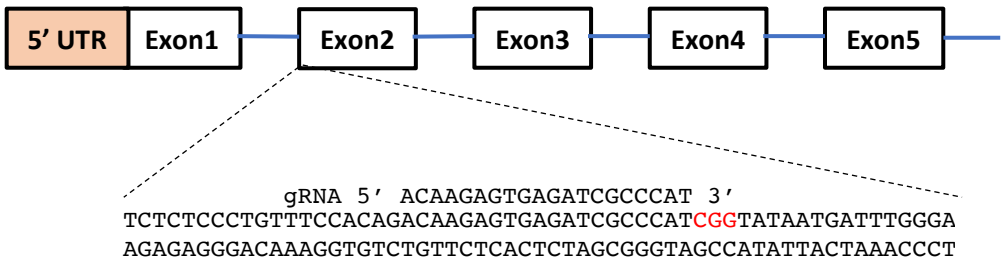
Figure 33. No differences in ONL thickness 4 months after injection. A) Histological analysis of ONL thickness: Microscopy images of hematoxylin-eosin staining of retinas harvested at p120. RPE: retinal pigmented epithelium; ONL: outer nuclear layer; INL: inner nuclear layer; GCL: gangliar cell layer. B) Quantification of ONL thickness in the temporal region of analyzed retinas.

Specific Aim 3

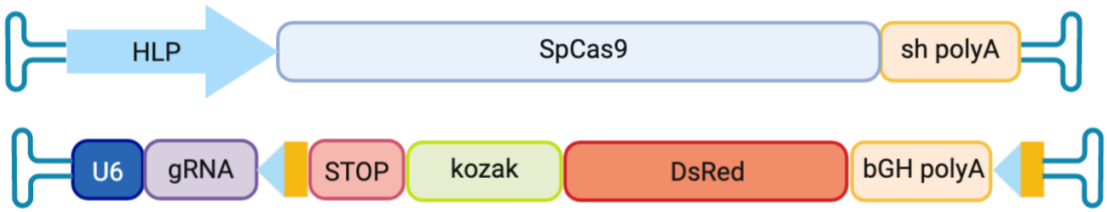
HITI design to target the Alb locus

The third aim of my project was to design a HITI strategy to perform targeted integration in the albumin locus in mouse hepatocytes. To do so, I designed a gRNA specific for the 2nd exon of the murine *Alb* gene (Figure 34A), which was the most upstream region where gRNAs with high predicted on-target score could be designed. Then I generated two AAV2/8 vectors: the first vector encoded for SpCas9 under control of the liver-specific hybrid liver promoter (HLP) and with a short synthetic polyA (sh polyA)[164]. The second vector encoded for a gRNA expression cassette and the donor DNA. The *Alb*-specific or scramble gRNA was under control of the U6 promoter. The donor DNA was flanked by the inverted albumin gRNA target sites, comprising the PAM. The donor DNA contained STOP codons in all 3 frames, a kozak signal to drive the start of translation, the transgene DsRed and the bGH polyA (Figure 34B). I expected that, after intravenous delivery, both vectors would reach the liver and, upon expression of Cas9, it would cleave both the *Alb* locus and the donor DNA, causing NHEJ-mediated integration of the donor DNA in the albumin locus (Figure 34C).

A)



B)



C)

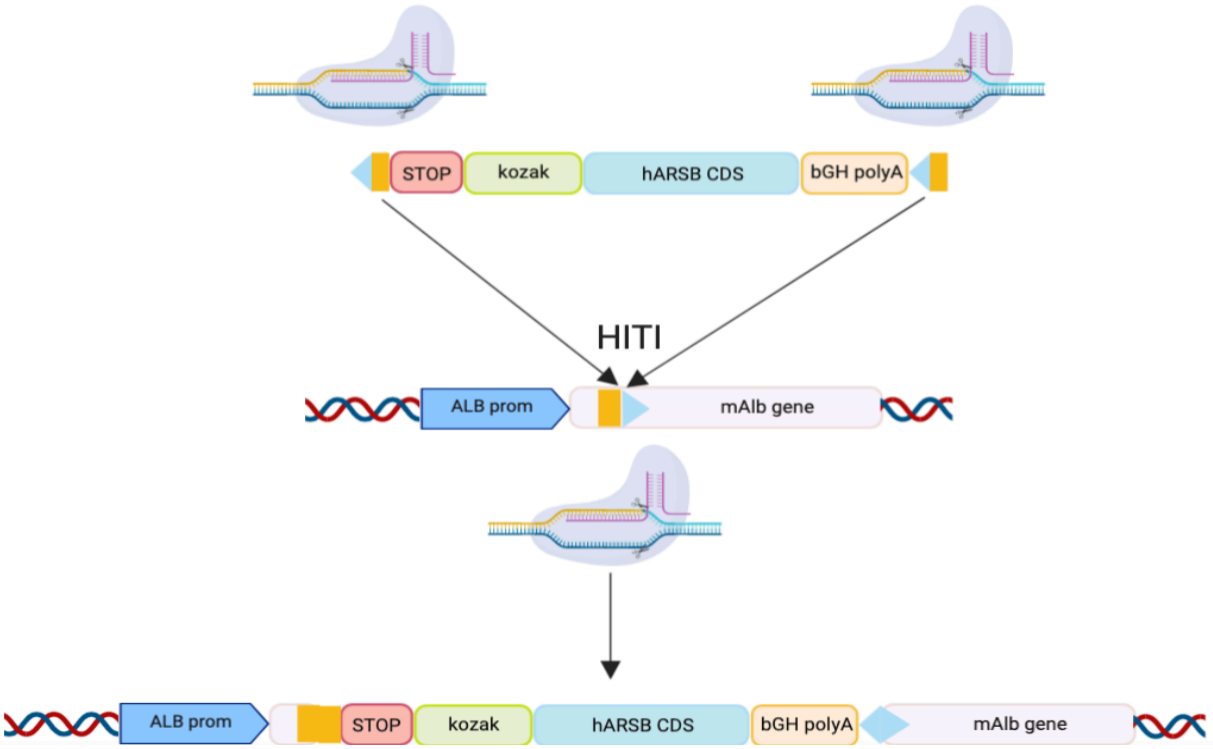


Figure 34. Schematic depiction of HITI design for integration in the Albumin locus and of gRNA and viral vectors used: A) Design of the gRNA specific for the 2nd intron of the murine Albumin gene. PAM sequence is depicted in red. B) Schematic depiction of the AAV2/8 vectors used for HITI in the mouse liver. Donor DNA is flanked by the same gRNA target sequences. Donor DNA contains STOP codons in the 3 frames, a translation START sequence (kozak), the reporter gene DsRed and the bGH poly-A. C) Depiction of the expected HITI after Cas9 and donor DNA delivery. Scissors represent Cas9-mediated DBSs. Crossed scissors represent inability of Cas9 to recognize target sites.

HITI is feasible and efficient in mouse hepatocytes

To assess whether HITI at the *Alb* locus in neonatal mice was feasible, 2-day old C57BL/6 mice received intravenous injection of AAV2/8 (dose of each vector: $4*10^{13}$ GC/Kg). Livers were harvested one month after injection and observed under a fluorescence stereomicroscope. gRNA-treated livers showed abundant presence of DsRed+ foci, which was consistent with the expected clonal expansion of hepatocytes after injection (Figure 35A). Scramble-treated livers showed no DsRed+ foci. Fluorescence microscopy of liver cryosections showed 3,26% of DsRed+ hepatocytes in livers treated with gRNA, and their absence in scramble-treated livers (Figure 35B, C). Interestingly, DsRed+ hepatocytes seemed to form clusters consistent with clonal expansion. (Figure 35B).

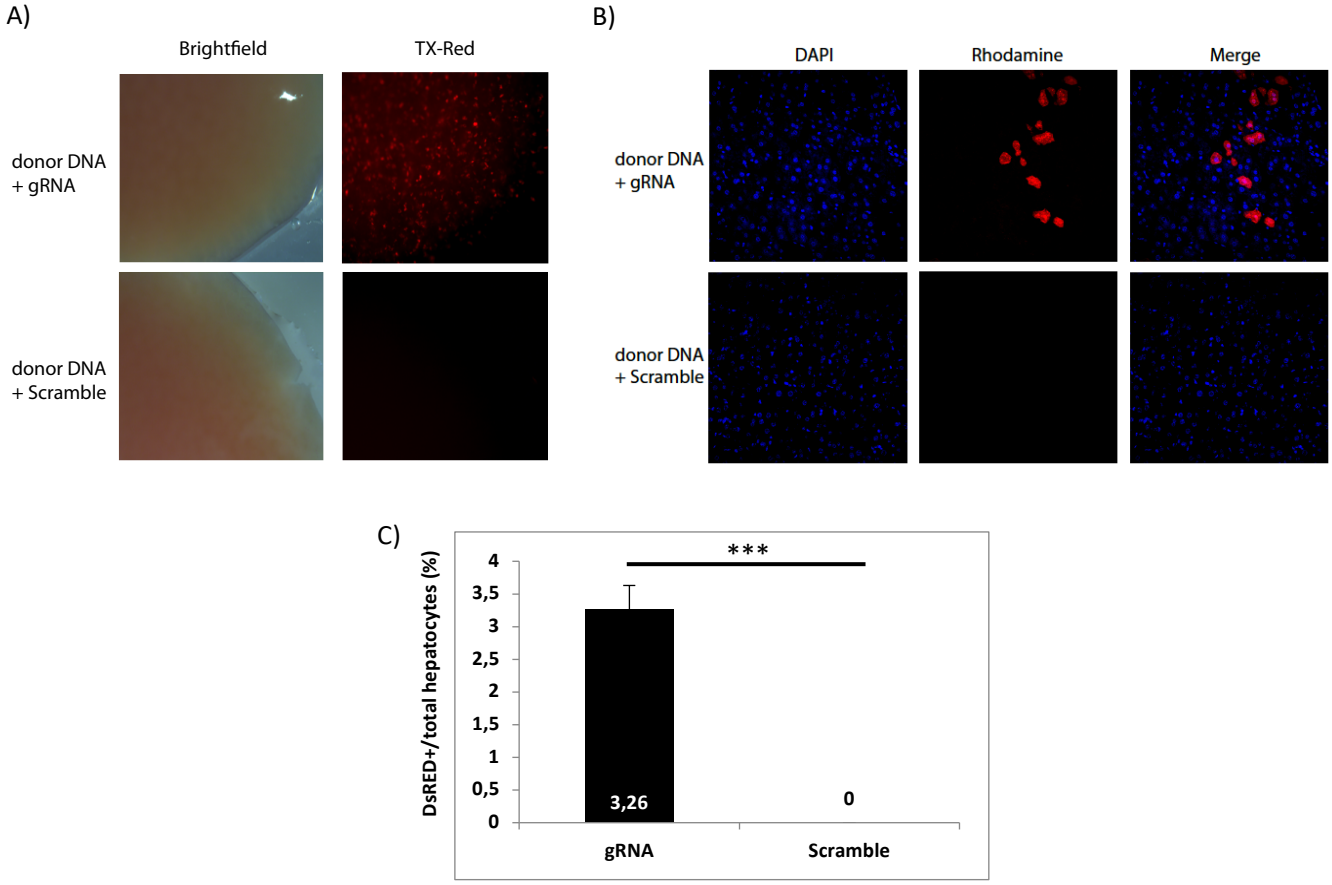


Figure 35. DsRed is expressed from mouse hepatocytes after neonatal HITI in the *Alb* locus. 4*10^{13}GC/Kg of each vector were administered intravenously through the temporal vein at p2. One month after injection livers were harvested and imaged. A) Stereomicroscope imaging of fresh livers shows widespread presence of DsRed+ foci only in gRNA-treated livers. B) Fluorescence microscopy of liver cryosections shows presence of DsRed+ hepatocyte foci. C) Quantification of DsRed+ hepatocytes. *** p<0,0001

I also generated AAV2/8 vectors carrying an IRES-DsRed donor DNA, which were injected following the same conditions. In this case, both the gRNA-treated and scramble-treated livers showed presence of DsRed+ hepatocytes, possibly due to hepatocyte-restricted promotorial activity of the IRES sequence, which was confirmed in Hepa1.6 cells (data not shown). For this reason, I only used kozak as a START signal for the following experiments.

Cas9 INDEL frequency in the albumin locus in mouse hepatocytes

In order to correlate the presence of DsRed+ hepatocytes to Cas9-mediated cleavage of the albumin locus, I extracted DNA from injected livers and performed PCR amplification of the region around the Cas9 target site. PCR fragments were subsequently used for Surveyor Assay, which showed presence of INDELs only in gRNA-treated livers and not in scramble-treated livers (Figure 36).

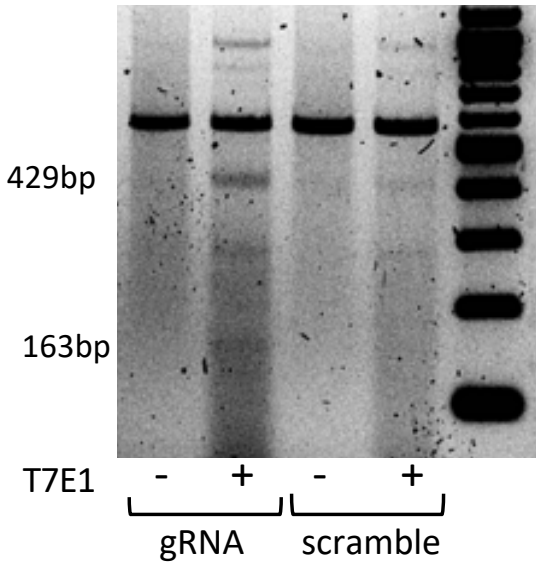


Fig 36. Only Alb-specific gRNA targets Cas9 to cleave the 2nd exon of the Alb gene: Surveyor Assay for INDEL detection. DNA was extracted from liver and used for PCR amplification of the genomic region surrounding the Cas9 target sequence. A 592bp fragment was amplified. Expected size of T7E1 digestion products is depicted. T7E1: T7 Endonuclease I treatment.

I then used TIDE analysis to assess the frequency and types of INDELs generated. I observed 9,9% INDEL efficiency in gRNA-treated livers, and negligible INDELs in scramble-treated or PBS-injected livers (Figure 37A, B). Interestingly, even if most deletions due to NHEJ repair are supposed to be between 1 and 6bp, in all gRNA-treated livers I observed a 7bp deletion which represented almost 50% of the observed INDELs (Figure 37A, blue arrow). Sequence analysis based in the models used by Shen et al. [96] for INDEL prediction allowed me to identify two small 4bp microhomology regions at both sides of the DSB which cause efficient MMEJ repair via this 7bp deletion (Figure 37C).

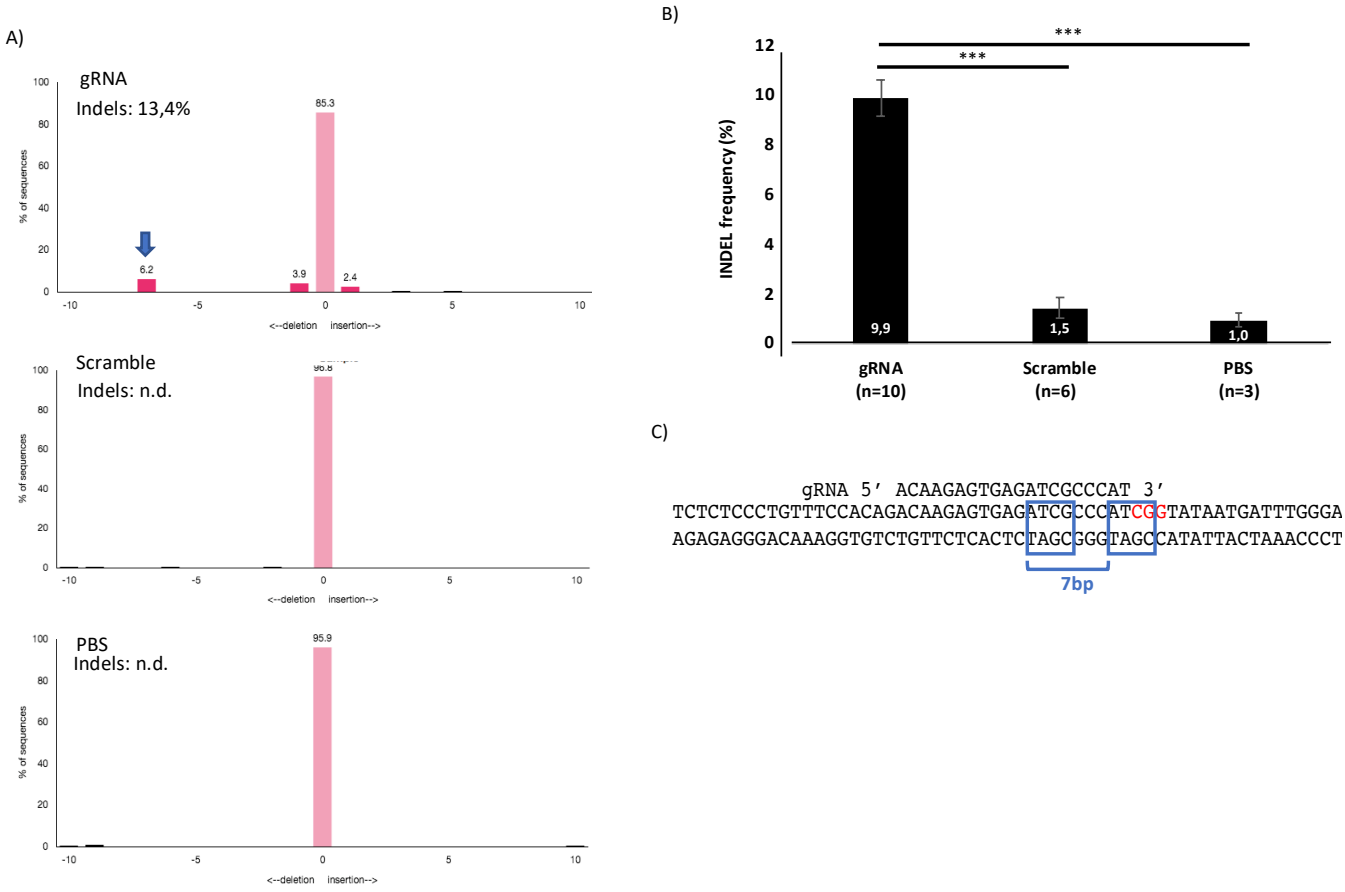


Figure 37. TIDE analysis of INDEL frequency. A) Representative TIDE graphs showing INDEL frequency and distribution for a sample from each treatment group. Pink color shows statistical significance. Black color shows that detected INDELs are not statistically significative. Blue arrow depicts common 7bp deletion observed in gRNA-treated livers. B) Quantification of INDEL frequency. Not significative INDELs observed in scramble and PBS were included. C) Schematic depiction of a common 7bp deletion due to microhomology-mediated end joining (MMEJ) DSB repair. Microhomolgy regions are depicted with blue squares. PAM is depicted in RED.

Characterization of HITI precision in the *Alb* locus of mouse hepatocytes:

Next, I designed PCR primers in order to amplify the 5' and 3' junctions after targeted integration. For the 5' junction, the *Fwd* primer recognized the 1st intron of albumin, while the *Rev* primer recognized the donor DNA. For the 3' junction, the *Fwd* primer recognized the donor DNA while the *Rev* primer recognized the 2nd exon of albumin (Figure 38A). In both cases, I observed amplification of a PCR product of the expected size only in gRNA-treated livers, consistent with presence of both INDELs and DsRed+ hepatocytes. No PCR products were amplified from DNA extracted from scramble-treated livers (Figure 38B).

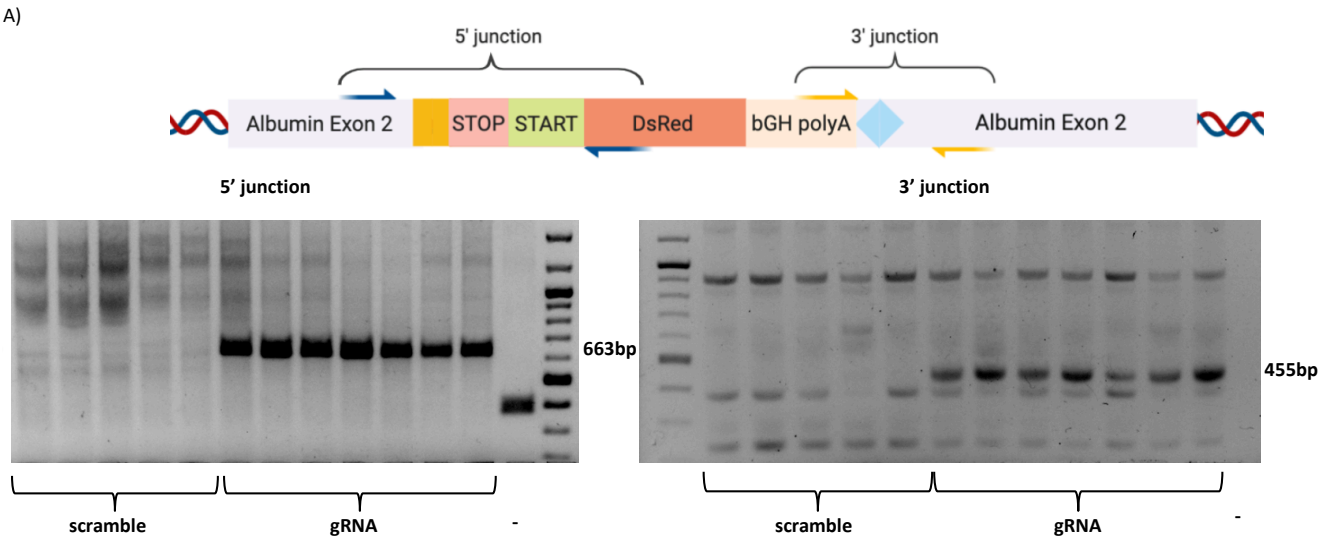


Fig 38. HITI Junction amplification. A) Scheme depicting primer design to amplify 5' and 3' junctions. B) 5' junction amplification. Expected fragment size was 663bp. PCR amplification was observed in all gRNA-treated and absent in all scramble-treated livers. C) 3' junction amplification. Expected fragment size was 455bp. PCR amplification was observed in all gRNA-treated and absent in all scramble-treated livers. 100bp ladder was used.

The amplified PCR products were then purified and cloned into bacteria. 26 clones were sequenced for each junction. I observed 38,5% of deletions and 3,8% of insertions in the 5' junction. In the 3' junction I observed 32% of deletions and 16% of insertions (Figure 39). All observed insertions were 1bp or 2bp. Deletions ranged from 1 to 55bp.

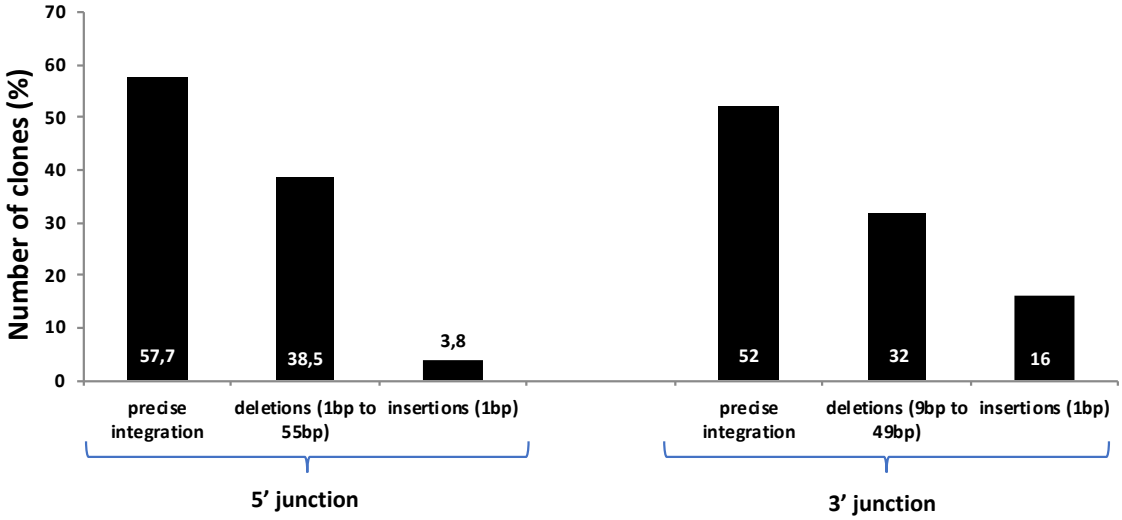


Fig 39. Frequency of INDEls in HITI junctions. HITI junction PCR fragments were cloned in bacteria and 26 clones for each junction were sequenced by SANGER sequencing. Presence of insertions, deletions and precise integrations was quantified.

To further characterize HITI precision, I used the same PCR fragments for NGS analysis in collaboration with Davide Cacchiarelli (TIGEM). Between 80.000 and 350.000 reads were obtained for each junction. Consistently with data from bacterial clones sequencing, in the 5' junction I observed low frequency of insertions and 38% frequency of deletions. Most insertions were of 1bp, while deletions mostly ranged from 1 to 47bp, with 1bp deletions being the most common. In the 3' junction, I observed a high frequency of 1bp insertions at the cleavage site, as well as some 2bp insertions. In total, deletions were more frequent (52%), but also more distributed around the cleavage site, ranging from 1bp to 38bp from the cleavage site. Most common deletions were between 1 and 18bp, and I also identified a common 42bp deletion (Figure 40).

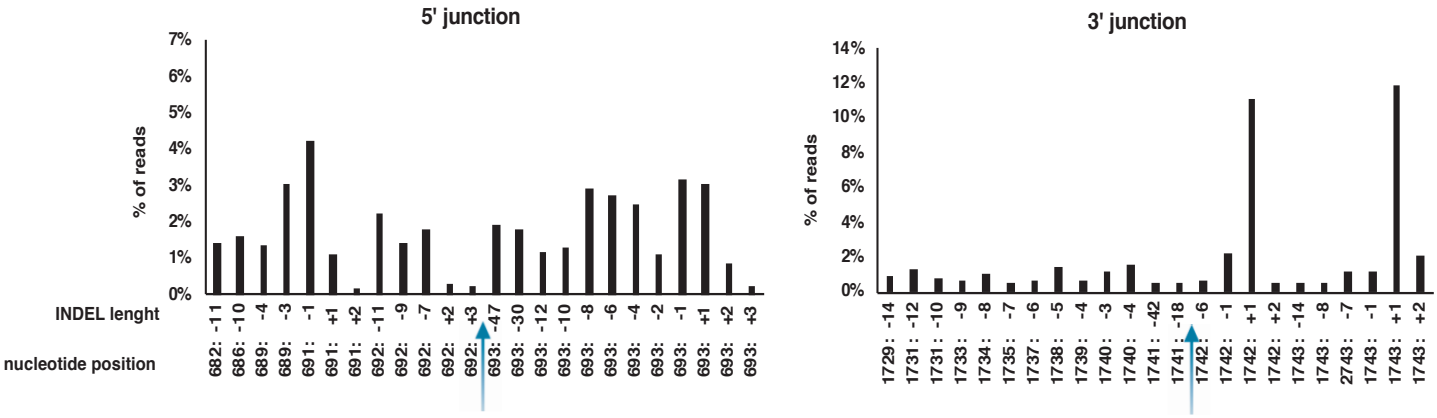


Fig 40. NGS of HITI junctions in the Alb locus. Distribution of Insertion and Deletion frequencies in NGS reads of HITI junctions. A) 5' junction: total reads = 80.000-100.000, B) 3' junction: total reads = 250.000-350.000 C) Relative INDEL frequency in each position surrounding the cleavage site. Negative and positive numbers represent deletions and insertions respectively. Blue arrows show the DSB.

We were not able to perform single-cell analysis of monoallelic or biallelic integration. In the future we expect to be able to perform single-cell isolation from treated livers and thus be able to perform this kind of analysis to determine the rate between monoallelic and biallelic integration.

HITI does not generate off-target integration of the donor DNA

To characterize whether potential off-target activity of Cas9 could lead to HITI of the donor DNA in other loci, I decided to use a genome-wide unbiased approach to detect potential off-target HITI. A library was generated using probes targeting DsRed to enrich for genomic sequences containing the donor DNA. This library was then used for NGS analysis. Only reads with a donor DNA sequence compatible with Cas9 cleavage were used for alignment with the mouse genome. In DNA samples extracted from gRNA-treated livers, integration of the donor DNA was observed only in the Alb locus with 11896 reads mapping to it, and no reads mapping elsewhere in the genome. This integration was not observed in scramble-treated livers. I then analyzed all reads containing the region of Alb close to the Cas9 cleavage site. I observed that 9,75% of reads contained the donor DNA, which is relatively consistent with the integration efficiency observed and proves that enrichment of DsRed sequences was not very efficient (Table 8).

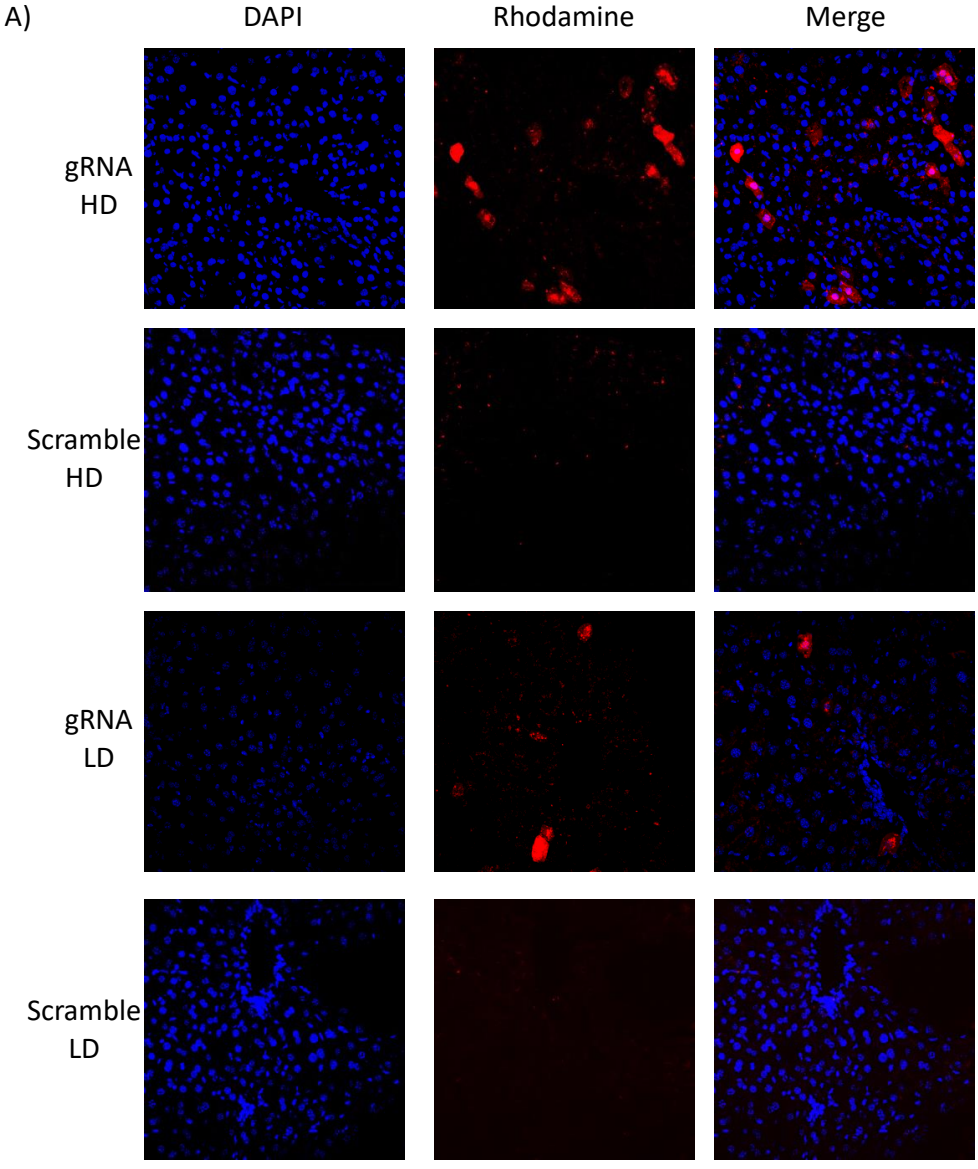
Sample	Treatment	Total reads in cleavage region	Total reads with donor DNA	Percentage of reads with donor DNA
1	gRNA	73469	4807	6,54
2	scramble	26822	0	0
3	scramble	32880	0	0
4	gRNA	31461	1214	3,86
5	gRNA	45960	2482	5,40
6	scramble	17689	0	0
7	gRNA	35806	1051	2,93
8	gRNA	122029	2342	1,91
TOTAL gRNA		**308725**	**11896**	**3,85**
TOTAL scramble		**77391**	**0**	**0**

Table 8: Albumin reads with donor DNA integrated in the targeted region.

Interestingly, I did observe integration of small parts of the AAV genome, normally close to the ITR sequences, in particular hotspots in the genome (data not shown). However, in this case integration was found at the same loci in DNA samples extracted from both gRNA-treated and scramble-treated livers, suggesting independence from Cas9-mediated DSBs.

HITI is efficient and dose-dependent in the adult mouse liver

Next, I wanted to assess whether HITI could also be performed in the liver of adult mice at a similar efficiency as in neonatal mice. For this reason, 4-week old C57BL/6 mice were injected intravenously with the same $4*10^{13}$GC/Kg dose used in neonatal mice (High Dose, HD) or a lower dose of $1,3*10^{13}$GC/kg (Low Dose, LD) of each vector. One month after injection I used fluorescence microscopy to assess presence of DsRed+ hepatocytes. I observed 2,76% of DsRed+ hepatocytes in the HD group and 1,25% in the LD group. DsRed+ hepatocytes were absent in scramble-treated livers independently of the vector dose used (Figure 41A, B). However, HITI efficiency was in both cases lower than achieved with neonatal injection. This could be explained by better transduction efficiency in the neonatal liver or by the clonal expansion of the modified hepatocytes, although theoretically HITI of DsRed in the albumin locus should confer no selective advantage.



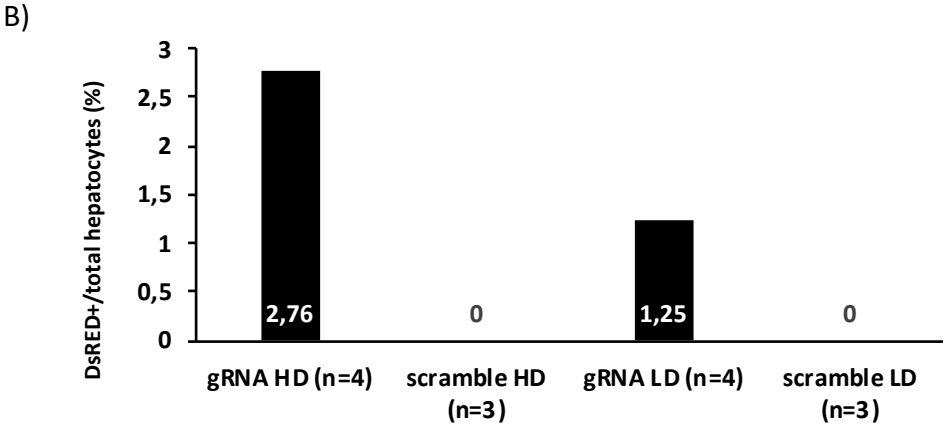


Fig 41. HITI is efficient and dose dependent in adult mouse liver. 4-week old C57BL/6 mice were injected with $4*10^{13}$GC/Kg (High Dose, HD) or $1,3*10^{13}$GC/Kg (Low Dose, LD) of each vector. A) Fluorescence microscopy of liver cryosections. B) Quantification of DsRed+ hepatocytes.

ARSB expression from the Albumin locus partially rescues the MPSVI phenotype

Next, I wanted to integrate the coding sequence of arylsulfatase B (ARSB) in the albumin locus of MPSVI mice. For that, I used the aforementioned AAV2/8-HLP-Cas9-shpolyA vector, and I generated another vector carrying the expression cassette for the *Alb*-specific or scramble gRNA, as well as a donor DNA encoding for STOP signals in the 3 frames, a kozak signal to start translation, the coding sequence of ARSB and the bGH polyA (Figure 42). Considering the deletions that I had observed in the 3' junction using the DsRed donor DNA, I added a 200bp Stuffer DNA between the bGH and the Cas9 target site in order to avoid unwanted deletions in the bGH sequence.

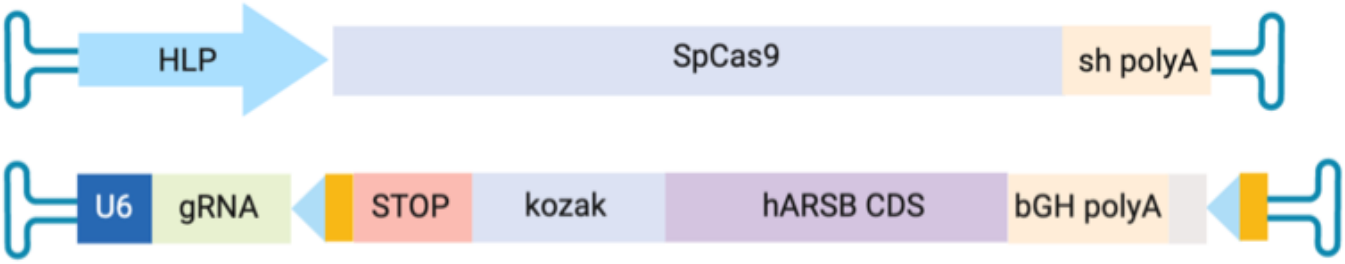


Fig 42. Design of AAV used to integrate the ARSB coding sequence in the albumin locus of mouse hepatocytes. HLP: Hybrid liver promoter, sh polyA: Short polyA, U6: U6 promoter for RNA polymerase 3, STOP: STOP codons in 3 different frames. hARSB CDS: coding sequence of human arylsulfatase B, bGH poly-A: bovine Growth Hormone poly-A. Stuffer DNA is depicted in grey. Yellow rectangle and blue triangle depict the two parts of the gRNA target sequence (upstream and downstream of Cas9 cleavage).

Two-day-old MPSVI-/- mice were injected intravenously with $6*10^{13}$GC/Kg of each vector. Monthly quantification of ARSB levels in the serum showed ARSB expression only in gRNA-treated mice, while ARSB was completely absent in the serum of scramble-treated mice. All gRNA-treated mice (n=5) presented ARSB expression at 1 month after injection (Figure 43). While 3 mice showed a relative decrease in ARSB levels between the first and second months, levels were stable after the second month, and in some cases, they increased by the fifth month after injection. The levels obtained ranged between 1/12 and 1/5 of the endogenous levels measured in unaffected mice using a different antibody for ARSB detection [150, 165], with the best performing mouse reaching 2572pg/mL (Figure 43). Scramble-treated mice (n=3) showed no serum ARSB at all timepoints of the analysis and were not included in Figure 43 for clarity.

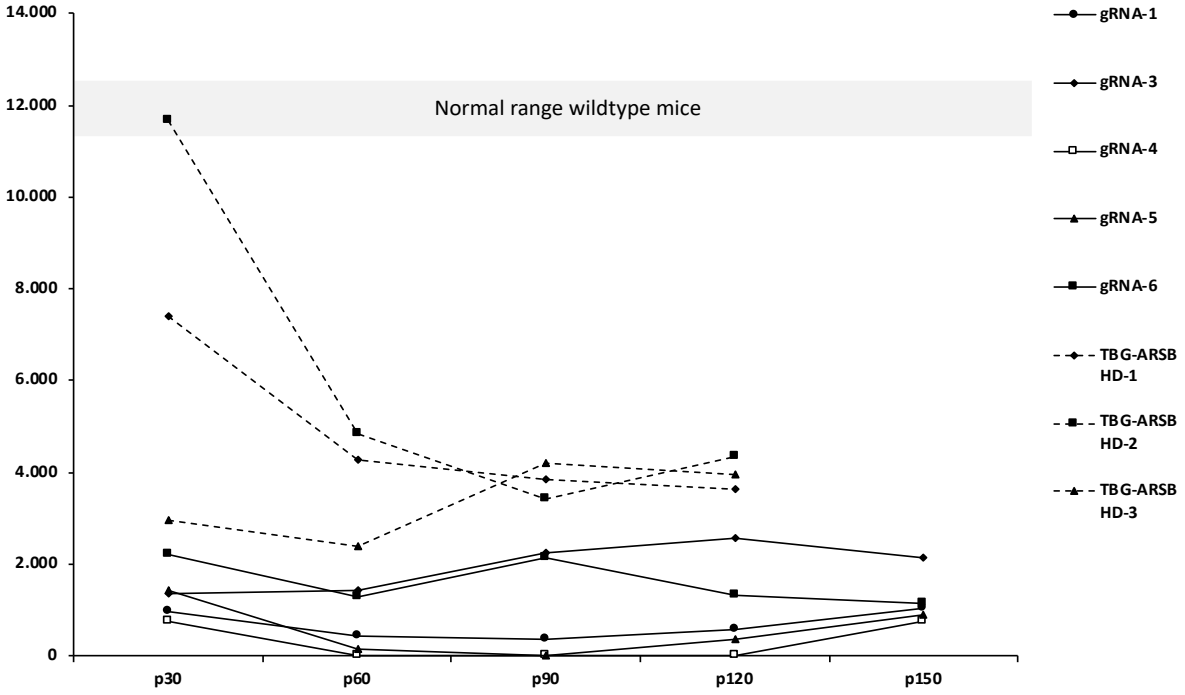


Fig 43. ARSB levels in mouse serum. Serum ARSB was measured monthly using an immunoassay with an antibody against human ARSB. Values observed at each timepoint are represented separately for each mouse. N=5 for HITI-treated mice. N=3 for scramble-treated mice. N=3 for TBG-ARSB treated mice.

Since the levels obtained after injection of a safe dose of $2*10^{12}$GC/Kg of an AAV2/8-TBG-ARSB vector in adult MPSVI mice were barely higher than what I observed with the HITI approach (3614±460pg/mL) [150], I decided to test whether neonatal injection of this same vector at the same dose used for HITI could stably achieve higher levels of ARSB expression. For this, neonatal MPSVI-/- mice (n=3) were injected with $6*10^{13}$GC/Kg of an AAV8-TBG-ARSB vector. As expected, ARSB levels at p30 were very high in the 3 analyzed mice but decreased at p60 (Figure 43), and stayed stable in later timepoints. To determine whether the serum ARSB levels reached were enough to correct the MPSVI phenotype, I quantified the urinary GAG levels 3 months after injection and observed a 46% reduction in GAG levels in gRNA-treated mice compared to scramble-treated mice (Figure 44), suggesting that the serum ARSB levels achieved are sufficient to restore GAG elimination in the tissues to levels comparable to the heterozygous MPSVI+/- mice. Similarly, mice injected with AAV8-TBG-ARSB showed 50% reduction in GAGs, and no statistically significant difference to gRNA-treated mice was observed. Further characterization of MPSVI phenotype correction at longer timepoints will be performed at sacrifice in order to quantify GAG accumulation in tissues like the cardiac valves.

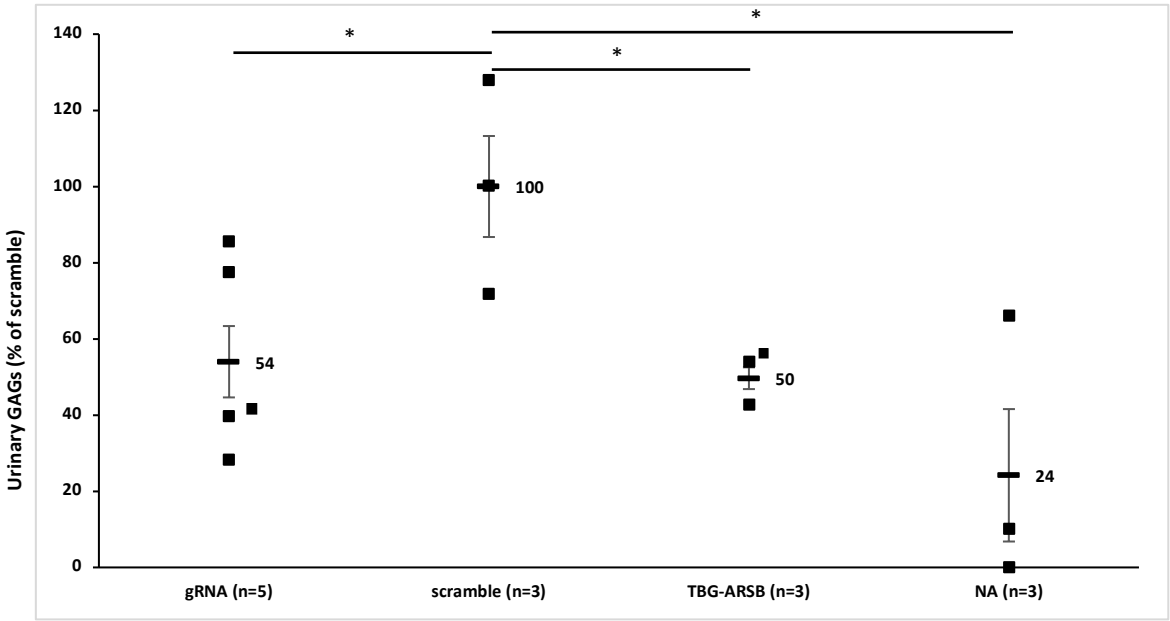


Fig 44. Urinary GAGs are similarly reduced 3 months after neonatal HITI or GT treatment. Urinary GAGs were measured from urine collected 3 months after injection. GAG levels were normalized with creatinine levels. Results are represented as percentage relative to GAG levels in control affected mice treated with scramble gRNA. Squares represent single analyzed mice. Bars represent group means and standard error. NA= not affected *=p<0,05

HITI at the albumin locus doesn't affect serum albumin levels

An important concern related to the strategy proposed to target Alb is the possible knock-out of Alb after genome editing. For this reason, I decided to assess whether treatment with Cas9 and the Alb-specific gRNA could reduce the serum albumin levels in treated MPSVI mice. For this, serum albumin levels were measured at p90. Although a relative decrease of albumin levels was observed in gRNA-treated mice when compared to scramble-treated or uninjected mice, this decrease was not significant and the levels observed were considered in the normal range for mice of this age (Figure 45).

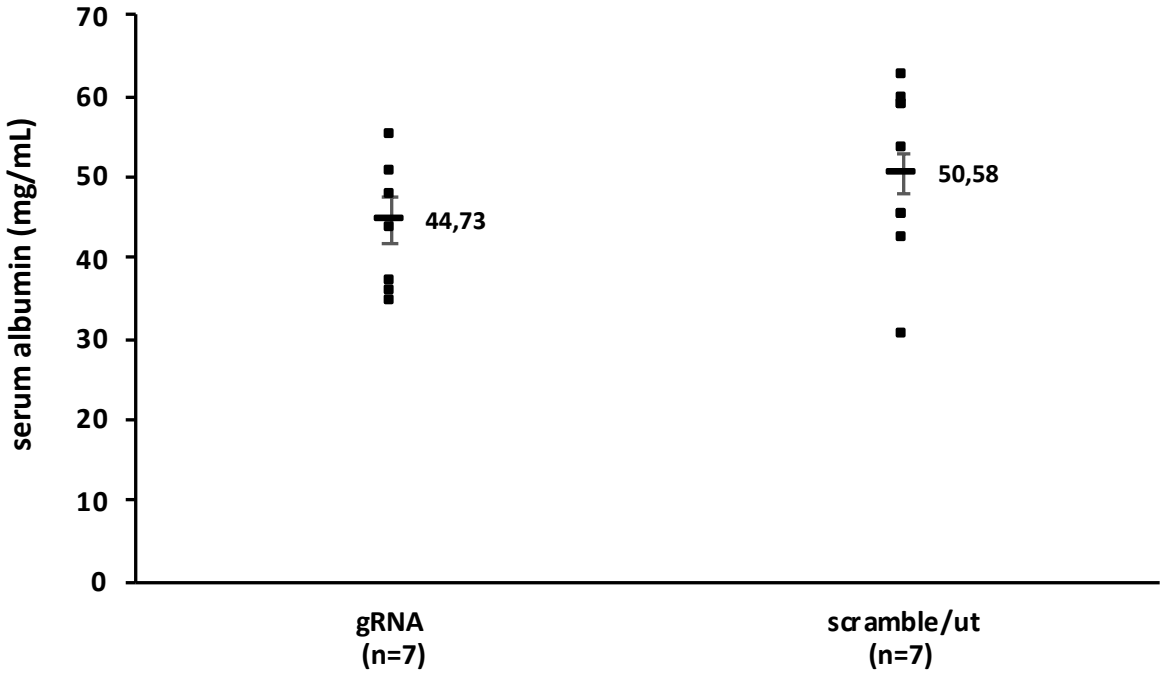


Fig 45. Serum albumin levels 3 months after treatment. Results are represented as absolute values for each analyzed mouse. Bars represent group means and standard error.

DISCUSSION

Gene therapy has been proven to be effective in animal models and in patients in several clinical trials. However, several important limitations still exist for the treatment of some diseases. In this thesis I have tried to solve two of these limitations: first is the inability to treat dominantly inherited diseases, in which knockout of the mutant allele, instead of gene supplementation, is needed to avoid gain of function effects [3, 94]. The other main limitation is the inability to deliver therapeutic genes to the neonatal liver using non-integrating vectors, due to vector dilution during hepatocyte division.

Allele-specific genome editing is feasible and effective in the retina:

Several authors have explored the possibility of using allele-specific genome editing to knockout gain of function alleles, with encouraging success [166, 167]. This is particularly relevant in the retina, where 30-40% of all cases of retinitis pigmentosa present a dominant inheritance [33]. While several approaches using CRISPR/Cas9 to perform allele-specific knockout in AdRP models have been developed [130-132], most of them have focused in very common mutations like the P23H mutation. Additionally, most approaches have used subretinal injection of DNA solutions *in utero* or at birth followed by electroporation as a method for delivering Cas9 in the retina, which is not a clinically relevant gene delivery approach. In Specific Aim 1 of this thesis I developed an AAV-based allele-specific genome editing approach targeting the less common P347S mutation, which is relevant in Spain and Italy [111, 112].

Gene knockout is normally performed by generating INDELs in the beginning of the coding sequence of a gene, thus changing the reading frame and causing truncation of protein translation [94]. However, due to the C-terminal position of the P347S mutation, gRNA design for allele-specific knockout was constricted to the last exon and the 3'UTR of the RHO gene. Bakondi *et al*. previously showed phenotypical improvement of AdRP after knockout of a C-terminal S334TER

allele, but didn't study the molecular mechanism that led to photoreceptor survival. In order to specifically target the P347S allele, I designed 3 gRNAs that contained the mutation in different positions. I also designed one gRNA that did not contain the mutation, as a control of allele-unspecific cleavage. Of those gRNAs, only gRNA1 was able to specifically target Cas9 to cleave P347S RHO without cleaving WT RHO. In this gRNA, the mutation was exactly in the -1 position from the PAM, which is in the 3' end of the gRNA sequence. Further research could confirm whether the presence of a mutation in this position is enough for allele-specific targeting in other loci. If this was confirmed, this kind of gRNA design could be applied to knockout other alleles.

Differently, the gRNA designed by our collaborators (gRNA5) placed the mutation in the PAM sequence, as first proposed by Bakondi *et al*. [132], so that when the mutation is not present, there is no PAM sequence and Cas9 cannot recognize the wildtype allele. This approach exploits the same system used by bacteria, in which the PAM sequence is absent in the CRISPR array in the bacterial genome, thus avoiding its cleavage by Cas9. Both targeting systems were shown to be viable and allele-specific, although gRNA1 showed significantly higher *in vitro* cleavage of P347S *RHO* and of WT *RHO*. For this reason, both gRNAs were tested in vivo.

The ability of gRNA1 and gRNA5 to cleave P347S *RHO* in vivo was confirmed in a mouse model of AdRP, and 5-17% INDEL efficiency was detected in treated retinae. Importantly, a significant reduction of the P347S *RHO* mRNA levels in the mouse retina was observed after cleavage, showing that generation of INDELs in terminal positions can be effective to knockout a gene. Furthermore, treatment with each gRNA significantly improved electrical retinal response to light stimuli in this animal model, as seen by ERG analysis. However, only treatment with gRNA1 achieved correction of pupillary light response, which was consistent with higher ERG B-wave values. The levels of ERG correction I observed were lower than those reported by Gianelli *et al.* [131], which was consistent with lower INDEL efficiency. However, when performing allele-specific genome editing, cleavage

efficiency is limited by the choice of gRNA, as more specific gRNAs are not necessarily more efficient. Thus, comparison of INDEL efficiencies in different mutations may not be informative due to different gRNA requirements. In the future, different Cas9 variants with different specificities could be used to target the P347S mutation and compared with the two approaches tested here. The results of this specific aim build on existing evidence of feasibility of allele-specific genome editing for the treatment of AdRP, and could lead to the development of a therapy for patients with the P347S mutation.

However, I didn't observe a clear difference in ONL thickness when comparing treated and untreated retinae. This could be due to the high variability in ONL thickness I observed in all analyzed P347S mice. Anyway, this suggests that the reduction of P347S mRNA observed may be enough to improve the function of surviving photoreceptors but not enough to avoid their death.

Homology-Independent Targeted Integration is feasible in photoreceptors:

Although I demonstrated that allele-specific knockout of a mutant allele was feasible and effective to treat an animal model of AdRP, clinical applicability of this approach is clearly limited by the large amount of RHO mutations that cause AdRP [102, 107]. Since developing an allele-specific approach for each different mutation is time and resource-consuming, in Specific Aim 2 I attempted to develop a genome editing approach that was not dependent on allele-specificity and could be used to treat all cases of AdRP due to RHO mutations. For this, I decided to use the HITI system, which allows targeted integration of an external donor DNA in the desired genomic locus [101]. This system was chosen preferentially to HDR because of the predominance of NHEJ-mediated DNA repair in photoreceptors and other non-dividing cells [78].

In this case, I decided to target the first exon of the murine Rho gene in an allele-independent manner, and to integrate a donor DNA carrying a correct copy of the human RHO gene that should replace the function of both the WT and mutant alleles after they have been cleaved by Cas9. First, as a proof of concept, I decided to use a donor DNA carrying the reporter transgene DsRed.

This allowed me to quantify the efficiency of HITI in mouse photoreceptors with two different donor DNA configurations previously tested *in vitro*. IRES worked significantly better than kozak as a translation start site, reaching 9% efficiency in the area of injection. This could be due to the distance between the integration site and the endogenous RHO promoter, which is larger than the one in the *in vitro* model. However, HITI efficiency in the retina was strongly restricted to the area immediately adjacent to the injection site. I hypothesized that this could be due to low efficiency of co-transduction with 2 different AAV vectors in the same photoreceptor, which has been reported to be of 24% [168]. In order to assess if HITI could be more efficient with higher AAV co-transduction, as well as whether HITI was feasible in the retina of a more clinically relevant animal model, I used HITI in the pig retina, which has been shown to have a much higher rate of co-transduction by 2 AAVs (73%) [168]. HITI was feasible also in the pig retina but showed a similar efficiency as in the mouse retina, albeit with a larger portion of the retina being edited. This suggests that while HITI efficiency could be limited by AAV co-transduction, the system could have intrinsic limits that may depend on the NHEJ repair machinery [100].

In order to assess whether the HITI efficiency observed was enough to rescue the AdRP phenotype, I generated a donor DNA carrying IRES and the coding sequence for the human RHO gene. Small but significant improvement of ERG response was observed 40 days after injection, although this improvement was only transient and disappeared one month later. This suggests that the HITI efficiency achieved was not enough to maintain the structure of the retina and thus the corrected photoreceptors were likely lost as a result of the degeneration of the neighboring photoreceptors,

which is a known trait of RP physiopathology [169]. This shows that, while HITI is feasible in photoreceptors, the complex structure and degeneration of the retina demand a more efficient correction that can involve an important number of photoreceptors in a larger area of the retina. For this reason, increasing HITI efficiency is crucial for clinical translation.

A way to increase HITI efficiency could be by increasing the ratio between Cas9 and donor DNA in order to have more available donor DNA inside the photoreceptor nucleus, thus possibly facilitating integration. However, this has not yet been tested in the retina. Additionally, using AAV serotypes that present a higher co-transduction efficiency, or other DNA or ribonucleoprotein delivery vectors, could help deliver Cas9 and the donor DNA more efficiently to photoreceptors, and thus increase HITI efficiency in the retina [29]. Alternatively, using other methods for targeted integration, like the homology-mediated end joining (HMEJ) approach proposed by Yao et al, which combines NHEJ and HDR [170], could increase the efficiency of targeted integration while maintaining the main features of the donor DNA I have developed.

Homology-Independent Targeted Integration in the Albumin locus achieves therapeutic correction of MPSVI:

In specific Aim 3, my objective was to develop a genome editing approach that could allow stable liver expression of a therapeutic gene after neonatal delivery. For that, I tried to assess whether the HITI platform I had developed for targeted integration into non-dividing cells like photoreceptors could be used for editing of neonatal hepatocytes, which are mainly in a dividing state. The albumin locus has widely been used as a safe harbor for integration of donor DNAs, leading to high and stable liver expression of therapeutic proteins [86-90]. For this reason, I decided to perform targeted integration of a donor DNA in the albumin locus of neonatal mouse hepatocytes. I used DsRed as a reporter gene in order to characterize HITI efficiency in the liver, as well as identify the best donor DNA configuration to use. One month after injection of neonatal

mice I observed widespread presence of edited hepatocytes, with around 3% of all hepatocytes being DsRed+. Cas9 cleavage efficiency in the albumin locus was measured at 11%. Both these numbers are very consistent with those reported by Suzuki *et al*. using similar vector doses to perform HITI in the liver of neonatal Ai14 mice [101]. This suggests that Cas9 cleavage and HITI efficiency may be limited by liver transduction and NHEJ efficiency.

However, the ratio between HITI efficiency and Cas9 cleavage efficiency shows that HITI is quite efficient in the liver, with a similar ratio to that reported by several authors when using HDR to target the factor IX locus. In short, Li *et al*. used ZFNs and HDR and observed 14% of INDELs and 2-5% of integration [85]. More recently, Wang *et al*. used Cas9 and HDR and observed 36% of INDELs and 11% of integration [171]. This shows that Cas9 cleavage efficiency could be crucial for HITI efficiency.

After this, I used the same HITI strategy in adult mice, and I demonstrated that HITI was feasible also in the adult liver, and that efficiency in the adult liver was dependent on the AAV dose. This is consistent with data published by Wang *et al*. [171]. In the adult liver, most hepatocytes are in a quiescent state. Accordingly, livers from mice treated at 4 weeks of age didn't show presence of foci of DsRed+ hepatocytes and presented more single edited hepatocytes than livers from neonatally-treated mice. This confirms that transgene expression is caused by integration of the DNA in the cell genome, and thus inherited by daughter cells after hepatocyte division.

One important point to consider is that integration was performed in the second exon of the albumin locus. This causes knockout of the endogenous albumin gene. Although levels of albumin were found to not be significantly reduced in gRNA-treated MPSVI mice, and were maintained within the normal range, other strategies that targeted HITI to the intronic regions or that could include the full albumin sequence in the final mRNA and protein product (for example using a 2A sequence between the transgene and the endogenous albumin), would be desirable to avoid potential detrimental effects due to albumin targeting.

After demonstrating that HITI in the albumin locus is effective, I tried to determine whether the observed efficiency was enough to achieve therapeutically relevant levels of liver expression of a therapeutic protein. For this, I used HITI to integrate a copy of the coding sequence of arylsulfatase B (ARSB) in the second exon of the albumin locus in the liver of a mouse model of MPSVI. Treated mice showed ARSB expression up to five months after treatment, and significant correction of the levels of urinary GAG excretion three months after injection, which is the main clinical marker for the disease [137], to levels comparable to non-affected mice. This improvement was comparable to the improvement reported after treating adult MPSVI mice with $2*10^{12}$GC/Kg of an AAV2/8-TBG-hARSB or ERT in [150], which is the standard of care for MPSVI patients worldwide. This is particularly relevant due to the fact that ERT is extremely expensive [137, 139, 141, 172, 173].

Additionally, while gene therapy with AAV has been shown to be efficient in murine [150, 165, 174] and feline [149, 151] animal models and is being tested in patients in a clinical trial, its main limitation is the inability to treat pediatric patients due to AAV genome dilution during hepatocyte division [34]. I demonstrated this by injecting $6*10^{13}$GC/Kg of AAV2/8-TBG-hARSB in neonatal mice and observing a significant decrease of serum ARSB levels from 1 to 2 months after injection. However, the serum ARSB levels reached with this dose were similar to what has been reported in adult mice injected with $2*10^{12}$GC/Kg [150], which is the dose being tested in the MPSVI gene therapy clinical trial. Although this shows that this approach might be feasible in pediatric patients and achieve improvement of the MPSVI phenotype, its major drawback is that random AAV integration in the genome of hepatocytes have been shown to cause hepatocellular carcinoma (HCC) after neonatal treatment. Indeed, Chandler *et al*. showed that neonatal intravenous injection of AAV8 with the TBG promoter, at similar doses as the ones used in this project, caused HCC in 60% of injected mice [175]. In contrast, the approach I have developed uses the weaker HLP promoter for Cas9 expression, and no promoter in the donor DNA. As Chandler *et al*. showed, weaker promoters do not cause genotoxicity after neonatal injection [175]. This approach achieved stable ARSB expression with one single AAV administration at a neonatal age, with ARSB

levels showing little variation over time. In the future, to further increase the safety of this approach, Cas9 could be delivered to the liver as a ribonucleoprotein using lipid nanoparticles or other protein delivery methods, thus avoiding the use of any promoters and further reducing the probability of AAV-induced carcinogenesis.

Moreover, although the levels of serum ARSB and of urinary GAGs were similar between mice treated with HITI and those treated with a conventional gene therapy approach, the gene therapy approach generates a very high initial expression of the therapeutic gene, which could potentially cause immune response against it. In contrast, HITI achieves lower but stable ARSB levels.

Importantly, the main limitation of conventional liver GT in adult patients is the possibility of loss of transgene expression after liver injury. Additional research could explore whether HITI of ARSB is therapeutically effective in adult MPSVI mice, and especially if ARSB expression is maintained after liver damage, which can be generated by partial hepatectomy.

Furthermore, the approach I have designed exploits the strong activity of the endogenous albumin promoter to achieve strong liver expression of the therapeutic protein. This approach could be adapted to treat other diseases by changing the therapeutic gene in the donor DNA without needing to change the gRNA or donor DNA conformation, which simplifies the applicability of this approach to a broad spectrum of diseases. This would be a viable alternative to the several approaches that use HDR to integrate therapeutic genes in the albumin locus.

Another important point to discuss is that AAV genome dilution during hepatocyte division would ensure that stable Cas9 expression would occur only in a very reduced number of hepatocytes. This could potentially reduce the probability of off-target cleavage and integration. Further

research could confirm this by quantifying the Cas9 vector copy numbers and Cas9 protein levels in hepatocytes at different timepoints after injection.

HITI is precise and limited to the targeted locus:

A part of this project that deserves to be discussed separately is the in-depth characterization of *in vivo* HITI that was performed using NGS in both the rhodopsin and albumin loci. In *Rho*, the 5' junction showed mainly small deletions at a low frequency, as well as two unexpected integrations of parts of the AAV ITR sequence. Instead, in the 3' junction, I observed a very frequent (70%) 1bp insertion in the 3' junction. This is contrary to what has been reported by Suzuki *et al*., who showed that HITI was quite precise in most loci both *in vitro* and *in vivo* [101]. However, they did describe imprecise HITI in the MERTK locus in RPE cells *in vivo*, but in the 5' junction and mostly due to deletions and not 1bp insertions.

In the *Alb* locus I observed around 40% INDELs in both the 5' and 3' junctions. Again, INDEL types were different between the two junctions, with insertions being significantly more represented in the 3' junction (30%). Furthermore, it has been reported that the nucleotide in position -4 from the PAM determines the nature of 1bp insertions [96]. In the Cas9 target sequence used in the albumin locus, the -4 position is a cytosine, which according to INDEL prediction should mostly generate cytosine insertions. Indeed, the most common insertion I observed in the albumin 3' junction was a single C. This was not consistent in RHO, where the base in the -4 position was a cytosine and the most frequent 1bp insertions observed were guanine and cytosine. Additionally, frequent 2bp GC insertions were observed. This suggests that, even though insertion nature after Cas9 cleavage can be predicted, this prediction is probably not valid for HITI, at least not in photoreceptors.

The deletions observed at the 5' junction in the Alb locus were not expected to affect the expression of our donor DNA. However, deletions at the 3' junction, which ranged from 1 to 42bp, affected part of the donor DNA. To avoid this, I generated a new donor DNA with a 200bp Stuffer DNA sequence that should serve as a buffer in case deletions occur, to avoid deletion of important sequences of the donor DNA. This kind of stuffer DNA could be used in donor DNA constructs for performing HITI in other loci, although characterization of HITI INDELs in each particular locus would give more information about the necessity of adding this stuffer DNA.

As a general conclusion, the NGS analysis of HITI junctions suggests that DNA repair is sequence-specific, as reported by Shen *et al*. [96], but without being perfectly consistent with expected repair. Interestingly, the endogenous DNA seems to be more important for the DNA repair pathway than the donor DNA, as suggested by the fact that deletions are more common at the 5' junction, which represents the sequence upstream of the Cas9 cleavage site, and 1bp-insertions are more common at the 3' junction, which represents the sequence downstream of the Cas9 cleavage site. Future studies should aim at confirming whether HITI precision can be different in cell types that present different levels and activity of certain factors involved in the DSB detection and repair [78]. Particularly, photoreceptors are quiescent cells where NHEJ is the main repair pathway, while neonatal hepatocytes are dividing cells, so HDR should be more active than in quiescent cells. It would be interesting to target the same loci in different tissues to better characterize the relevance of sequence and cell type in the repair pathway.

Additionally, both the *Rho* and *Alb* loci are the highest transcribed loci in photoreceptors and hepatocytes respectively, which confers them a very open chromatinic conformation. Since more active loci present higher presence of factors involved in DSB repair, it would be interesting to study whether the observed repair is similar in other less active loci and establish whether HITI precision can depend on chromatin structure and presence of DSB repair factors.

Since Cas9 can recognize and cleave off-target sites that are similar to its target site, undesired HITI could occur in those sites. This could have a detrimental effect, especially if integration occurs in coding regions or near oncogenes. To assess whether off-target HITI occurred, I decided to use an unbiased approach that could detect integration of the donor DNA anywhere in the genome. NGS analysis of potential off-target integration when using the *Rho*-specific gRNA was impossible due to low abundance of cells positive for the integration, which didn't allow detection of any products of HITI integration even at the *Rho* locus. Interestingly, I clearly observed random integration of partial AAV genome sequences in some loci, which was detected both in gRNA and scramble-treated retinas, and thus doesn't seem to be dependent on Cas9 cleavage. This is consistent with findings by several authors showing that AAV genomes or parts of these genomes can integrate spontaneously at DSBs [175-177].

When using the *Alb*-specific gRNA, precise integration of the donor DNA compatible with Cas9-mediated cleavage was detected only in the *Alb* locus and nowhere else in the entire mouse genome. This shows that if the right gRNAs are chosen, off-target HITI can be avoided. However, the assay used for this off-target characterization yielded a relatively low number of reads of the integrated donor DNA. Interestingly, I observed random integration of partial AAV genome sequences in several loci, of which most were shared with the retina. This analysis will be repeated with a higher sensitivity, expecting it to yield more significant results. In the future I also consider performing NGS directed to the predicted off-target sites with the highest score, in order to have a higher number of reads per each site and detect possible off-target HITI that our previous analysis was not sensitive enough to detect.

In conclusion I have demonstrated that HITI precision depends on the cell type and targeted locus, and that selection of a gRNA with low off-target activity avoids off-target HITI of the donor DNA.

Thus, HITI is a reliable method for targeted integration both in dividing and non-dividing cells *in vivo.*

Conclusion

In summary, I have developed three different genome editing approaches aimed at solving some of the remaining issues of gene therapy with AAVs. I have developed an allele-specific knockdown strategy for the P347S mutation of human *RHO*. I have also developed an allele-independent knockdown-and-replace strategy using HITI in the murine *Rho* locus. Both these genome editing strategies have been successful but not effective enough to achieve persistent correction of both retinal electrical response and retinal morphology. Last, I have developed a genome editing strategy that uses HITI to insert the *ARSB* coding sequence in the murine *Alb* locus, and achieved stable expression and secretion of a therapeutic protein from the liver at high enough levels to partially correct the MPSVI phenotype.

REFERENCES

1. Brunetti-Pierri N, A.A., *Gene Therapy of Human Inherited Diseases*, in *The Metabolic and Molecular Bases of Inherited Diseases*, S. R, Editor. 2010, McGraw Hill: New York.
2. Goswami, R., et al., *Gene Therapy Leaves a Vicious Cycle.* Front Oncol, 2019. **9**: p. 297.
3. Yanik, M., et al., *In vivo genome editing as a potential treatment strategy for inherited retinal dystrophies.* Prog Retin Eye Res, 2017. **56**: p. 1-18.
4. Cattoglio, C., et al., *Hot spots of retroviral integration in human CD34+ hematopoietic cells.* Blood, 2007. **110**(6): p. 1770-8.
5. Wu, X., et al., *Transcription start regions in the human genome are favored targets for MLV integration.* Science, 2003. **300**(5626): p. 1749-51.
6. Hacein-Bey-Abina, S., et al., *LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1.* Science, 2003. **302**(5644): p. 415-9.
7. Yla-Herttuala, S., *ADA-SCID Gene Therapy Endorsed By European Medicines Agency For Marketing Authorization.* Mol Ther, 2016. **24**(6): p. 1013-1014.
8. Ginn, S.L., et al., *Gene therapy clinical trials worldwide to 2017: An update.* J Gene Med, 2018. **20**(5): p. e3015.
9. Cavazzana-Calvo, M., et al., *Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia.* Nature, 2010. **467**(7313): p. 318-22.
10. Cartier, N., et al., *Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy.* Science, 2009. **326**(5954): p. 818-23.
11. Biffi, A., et al., *Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy.* Science, 2013. **341**(6148): p. 1233158.
12. Sessa, M., et al., *Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic leukodystrophy: an ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial.* Lancet, 2016. **388**(10043): p. 476-87.
13. Aiuti, A., et al., *Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome.* Science, 2013. **341**(6148): p. 1233151.
14. Milone, M.C. and U. O'Doherty, *Clinical use of lentiviral vectors.* Leukemia, 2018. **32**(7): p. 1529-1541.
15. Wold, W.S. and K. Toth, *Adenovirus vectors for gene therapy, vaccination and cancer gene therapy.* Curr Gene Ther, 2013. **13**(6): p. 421-33.
16. Muruve, D.A., *The innate immune response to adenovirus vectors.* Hum Gene Ther, 2004. **15**(12): p. 1157-66.
17. Atchison, R.W., B.C. Casto, and W.M. Hammon, *Adenovirus-Associated Defective Virus Particles.* Science, 1965. **149**(3685): p. 754-6.
18. Balakrishnan, B. and G.R. Jayandharan, *Basic biology of adeno-associated virus (AAV) vectors used in gene therapy.* Curr Gene Ther, 2014. **14**(2): p. 86-100.
19. Samulski, R.J. and N. Muzyczka, *AAV-Mediated Gene Therapy for Research and Therapeutic Purposes.* Annu Rev Virol, 2014. **1**(1): p. 427-51.
20. Hoggan, M.D., N.R. Blacklow, and W.P. Rowe, *Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics.* Proc Natl Acad Sci U S A, 1966. **55**(6): p. 1467-74.
21. Wright, J.F., *Manufacturing and characterizing AAV-based vectors for use in clinical studies.* Gene Ther, 2008. **15**(11): p. 840-8.
22. Gao, G., L.H. Vandenberghe, and J.M. Wilson, *New recombinant serotypes of AAV vectors.* Curr Gene Ther, 2005. **5**(3): p. 285-97.
23. Weinmann, J. and D. Grimm, *Next-generation AAV vectors for clinical use: an ever-accelerating race.* Virus Genes, 2017. **53**(5): p. 707-713.

24. Kotterman, M.A. and D.V. Schaffer, *Engineering adeno-associated viruses for clinical gene therapy.* Nat Rev Genet, 2014. **15**(7): p. 445-51.
25. Colella, P., G. Ronzitti, and F. Mingozzi, *Emerging Issues in AAV-Mediated In Vivo Gene Therapy.* Mol Ther Methods Clin Dev, 2018. **8**: p. 87-104.
26. *FDA approves hereditary blindness gene therapy.* Nat Biotechnol, 2018. **36**(1): p. 6.
27. Pierce, E.A. and J. Bennett, *The Status of RPE65 Gene Therapy Trials: Safety and Efficacy.* Cold Spring Harb Perspect Med, 2015. **5**(9): p. a017285.
28. Anguela, X.M. and K.A. High, *Entering the Modern Era of Gene Therapy.* Annu Rev Med, 2019. **70**: p. 273-288.
29. Trapani, I. and A. Auricchio, *Seeing the Light after 25 Years of Retinal Gene Therapy.* Trends Mol Med, 2018. **24**(8): p. 669-681.
30. Mingozzi, F. and K.A. High, *Immune responses to AAV vectors: overcoming barriers to successful gene therapy.* Blood, 2013. **122**(1): p. 23-36.
31. Nathwani, A.C., et al., *Adenovirus-associated virus vector-mediated gene transfer in hemophilia B.* N Engl J Med, 2011. **365**(25): p. 2357-65.
32. Nathwani, A.C., et al., *Long-term safety and efficacy of factor IX gene therapy in hemophilia B.* N Engl J Med, 2014. **371**(21): p. 1994-2004.
33. Dryja, T.P., *Retinitis Pigmentosa and Stationary Night Blindness*, in *Online Metabolic and Molecular Bases of Inherited Diseases*, A.L.B. David Valle, Bert Vogelstein, Kenneth W. Kinzler, Stylianos E. Antonarakis, Andrea Ballabio, Editor., McGraw-Hill: New York.
34. Cotugno, G., et al., *Impact of age at administration, lysosomal storage, and transgene regulatory elements on AAV2/8-mediated rat liver transduction.* PLoS One, 2012. **7**(3): p. e33286.
35. Nakai, H., et al., *Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction in vivo.* J Virol, 2001. **75**(15): p. 6969-76.
36. Arruda, V.R. and B.J. Samelson-Jones, *Obstacles and future of gene therapy for hemophilia.* Expert Opin Orphan Drugs, 2015. **3**(9): p. 997-1010.
37. Inagaki, K., et al., *Frequency and spectrum of genomic integration of recombinant adeno-associated virus serotype 8 vector in neonatal mouse liver.* J Virol, 2008. **82**(19): p. 9513-24.
38. Puchta, H., B. Dujon, and B. Hohn, *Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease.* Nucleic Acids Res, 1993. **21**(22): p. 5034-40.
39. Rouet, P., F. Smih, and M. Jasin, *Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease.* Mol Cell Biol, 1994. **14**(12): p. 8096-106.
40. Silva, G., et al., *Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy.* Curr Gene Ther, 2011. **11**(1): p. 11-27.
41. Chandrasegaran, S. and D. Carroll, *Origins of Programmable Nucleases for Genome Engineering.* J Mol Biol, 2016. **428**(5 Pt B): p. 963-89.
42. Kim, Y.G., J. Cha, and S. Chandrasegaran, *Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain.* Proc Natl Acad Sci U S A, 1996. **93**(3): p. 1156-60.
43. Chou, S.T., Q. Leng, and A.J. Mixson, *Zinc Finger Nucleases: Tailor-made for Gene Therapy.* Drugs Future, 2012. **37**(3): p. 183-196.
44. Tebas, P., et al., *Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV.* N Engl J Med, 2014. **370**(10): p. 901-10.
45. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.* Trends Biotechnol, 2013. **31**(7): p. 397-405.
46. Mussolino, C., et al., *A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity.* Nucleic Acids Res, 2011. **39**(21): p. 9283-93.

47. Christian, M., et al., *Targeting DNA double-strand breaks with TAL effector nucleases.* Genetics, 2010. **186**(2): p. 757-61.
48. Miller, J.C., et al., *A TALE nuclease architecture for efficient genome editing.* Nat Biotechnol, 2011. **29**(2): p. 143-8.
49. Holkers, M., et al., *Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells.* Nucleic Acids Res, 2013. **41**(5): p. e63.
50. Yang, L., et al., *Optimization of scarless human stem cell genome editing.* Nucleic Acids Res, 2013. **41**(19): p. 9049-61.
51. Horvath, P. and R. Barrangou, *CRISPR/Cas, the immune system of bacteria{Horvath, 2010 #115} and archaea.* Science, 2010. **327**(5962): p. 167-70.
52. Gasiunas, G., et al., *Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.* Proc Natl Acad Sci U S A, 2012. **109**(39): p. E2579-86.
53. Jinek, M., et al., *Structures of Cas9 endonucleases reveal RNA-mediated conformational activation.* Science, 2014. **343**(6176): p. 1247997.
54. Horvath, P. and R. Barrangou, *CRISPR/Cas, the immune system of bacteria and archaea.* Science, 2010. **327**(5962): p. 167-70.
55. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
56. Doudna, J.A. and E. Charpentier, *Genome editing. The new frontier of genome engineering with CRISPR-Cas9.* Science, 2014. **346**(6213): p. 1258096.
57. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.
58. Xiao, A., et al., *Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish.* Nucleic Acids Res, 2013. **41**(14): p. e141.
59. Dastidar, S., et al., *Efficient CRISPR/Cas9-mediated editing of trinucleotide repeat expansion in myotonic dystrophy patient-derived iPS and myogenic cells.* Nucleic Acids Res, 2018. **46**(16): p. 8275-8298.
60. Li, J., et al., *Efficient inversions and duplications of mammalian regulatory DNA elements and gene clusters by CRISPR/Cas9.* J Mol Cell Biol, 2015. **7**(4): p. 284-98.
61. Ho, B.X., et al., *In Vivo Genome Editing as a Therapeutic Approach.* Int J Mol Sci, 2018. **19**(9).
62. Ran, F.A., et al., *In vivo genome editing using Staphylococcus aureus Cas9.* Nature, 2015. **520**(7546): p. 186-91.
63. Kim, E., et al., *In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni.* Nat Commun, 2017. **8**: p. 14500.
64. Amrani, N., et al., *NmeCas9 is an intrinsically high-fidelity genome-editing platform.* Genome Biol, 2018. **19**(1): p. 214.
65. Haeussler, M. and J.P. Concordet, *Genome Editing with CRISPR-Cas9: Can It Get Any Better?* J Genet Genomics, 2016. **43**(5): p. 239-50.
66. Kleinstiver, B.P., et al., *Engineered CRISPR-Cas9 nucleases with altered PAM specificities.* Nature, 2015. **523**(7561): p. 481-5.
67. Zetsche, B., et al., *Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system.* Cell, 2015. **163**(3): p. 759-71.
68. Tsai, S.Q., et al., *GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases.* Nat Biotechnol, 2015. **33**(2): p. 187-197.
69. Tsai, S.Q., et al., *CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets.* Nat Methods, 2017. **14**(6): p. 607-614.
70. Alexandro E.Trevino, F.Z., *Chapter Eight - Genome Editing Using Cas9 Nickases*, in *Methods in Enzymology*, E.J.S. Jennifer A. Doudna, Editor. 2014, Elsevier Inc.
71. Kleinstiver, B.P., et al., *High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects.* Nature, 2016. **529**(7587): p. 490-5.

72. Slaymaker, I.M., et al., *Rationally engineered Cas9 nucleases with improved specificity.* Science, 2016. **351**(6268): p. 84-8.
73. Casini, A., et al., *A highly specific SpCas9 variant is identified by in vivo screening in yeast.* Nat Biotechnol, 2018. **36**(3): p. 265-271.
74. Merienne, N., et al., *The Self-Inactivating KamiCas9 System for the Editing of CNS Disease Genes.* Cell Rep, 2017. **20**(12): p. 2980-2991.
75. Komor, A.C., et al., *Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage.* Nature, 2016. **533**(7603): p. 420-4.
76. Gaudelli, N.M., et al., *Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage.* Nature, 2017. **551**(7681): p. 464-471.
77. Rees, H.A. and D.R. Liu, *Base editing: precision chemistry on the genome and transcriptome of living cells.* Nat Rev Genet, 2018. **19**(12): p. 770-788.
78. Iyama, T. and D.M. Wilson, 3rd, *DNA repair mechanisms in dividing and non-dividing cells.* DNA Repair (Amst), 2013. **12**(8): p. 620-36.
79. Stewart, G.S., et al., *The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage.* Cell, 2009. **136**(3): p. 420-34.
80. Chen, F., et al., *High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases.* Nat Methods, 2011. **8**(9): p. 753-5.
81. Burnight, E.R., et al., *Using CRISPR-Cas9 to Generate Gene-Corrected Autologous iPSCs for the Treatment of Inherited Retinal Degeneration.* Mol Ther, 2017. **25**(9): p. 1999-2013.
82. Wu, Z., et al., *A long-term efficacy study of gene replacement therapy for RPGR-associated retinal degeneration.* Hum Mol Genet, 2015. **24**(14): p. 3956-70.
83. Beltran, W.A., et al., *Optimization of Retinal Gene Therapy for X-Linked Retinitis Pigmentosa Due to RPGR Mutations.* Mol Ther, 2017. **25**(8): p. 1866-1880.
84. Anguela, X.M., et al., *Robust ZFN-mediated genome editing in adult hemophilic mice.* Blood, 2013. **122**(19): p. 3283-7.
85. Li, H., et al., *In vivo genome editing restores haemostasis in a mouse model of haemophilia.* Nature, 2011. **475**(7355): p. 217-21.
86. Barzel, A., et al., *Promoterless gene targeting without nucleases ameliorates haemophilia B in mice.* Nature, 2015. **517**(7534): p. 360-4.
87. Porro, F., et al., *Promoterless gene targeting without nucleases rescues lethality of a Crigler-Najjar syndrome mouse model.* EMBO Mol Med, 2017. **9**(10): p. 1346-1355.
88. Laoharawee, K., et al., *Dose-Dependent Prevention of Metabolic and Neurologic Disease in Murine MPS II by ZFN-Mediated In Vivo Genome Editing.* Mol Ther, 2018. **26**(4): p. 1127-1136.
89. Ou, L., et al., *ZFN-Mediated In Vivo Genome Editing Corrects Murine Hurler Syndrome.* Mol Ther, 2019. **27**(1): p. 178-187.
90. Sharma, R., et al., *In vivo genome editing of the albumin locus as a platform for protein replacement therapy.* Blood, 2015. **126**(15): p. 1777-84.
91. Takata, M., et al., *Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells.* EMBO J, 1998. **17**(18): p. 5497-508.
92. Delacote, F. and B.S. Lopez, *Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance: the trans-S double-strand break repair model.* Cell Cycle, 2008. **7**(1): p. 33-8.
93. McVey, M. and S.E. Lee, *MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings.* Trends Genet, 2008. **24**(11): p. 529-38.
94. Tu, Z., et al., *CRISPR/Cas9: a powerful genetic engineering tool for establishing large animal models of neurodegenerative diseases.* Mol Neurodegener, 2015. **10**: p. 35.
95. Brinkman, E.K., et al., *Easy quantitative assessment of genome editing by sequence trace decomposition.* Nucleic Acids Res, 2014. **42**(22): p. e168.

96. Shen, M.W., et al., *Predictable and precise template-free CRISPR editing of pathogenic variants.* Nature, 2018. **563**(7733): p. 646-651.
97. Maresca, M., et al., *Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining.* Genome Res, 2013. **23**(3): p. 539-46.
98. Cristea, S., et al., *In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration.* Biotechnol Bioeng, 2013. **110**(3): p. 871-80.
99. Lackner, D.H., et al., *A generic strategy for CRISPR-Cas9-mediated gene tagging.* Nat Commun, 2015. **6**: p. 10237.
100. Suzuki, K. and J.C. Izpisua Belmonte, *In vivo genome editing via the HITI method as a tool for gene therapy.* J Hum Genet, 2018. **63**(2): p. 157-164.
101. Suzuki, K., et al., *In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration.* Nature, 2016. **540**(7631): p. 144-149.
102. Sohocki, M.M., et al., *Prevalence of mutations causing retinitis pigmentosa and other inherited retinopathies.* Hum Mutat, 2001. **17**(1): p. 42-51.
103. Berger, W., B. Kloeckener-Gruissem, and J. Neidhardt, *The molecular basis of human retinal and vitreoretinal diseases.* Prog Retin Eye Res, 2010. **29**(5): p. 335-75.
104. Hartong, D.T., E.L. Berson, and T.P. Dryja, *Retinitis pigmentosa.* Lancet, 2006. **368**(9549): p. 1795-809.
105. Okada, T. and K. Palczewski, *Crystal structure of rhodopsin: implications for vision and beyond.* Curr Opin Struct Biol, 2001. **11**(4): p. 420-6.
106. Pepe, I.M., *Rhodopsin and phototransduction.* J Photochem Photobiol B, 1999. **48**(1): p. 1-10.
107. Mendes, H.F., et al., *Mechanisms of cell death in rhodopsin retinitis pigmentosa: implications for therapy.* Trends Mol Med, 2005. **11**(4): p. 177-85.
108. Athanasiou, D., et al., *The molecular and cellular basis of rhodopsin retinitis pigmentosa reveals potential strategies for therapy.* Prog Retin Eye Res, 2018. **62**: p. 1-23.
109. Rossmiller, B.P., R.C. Ryals, and A.S. Lewin, *Gene therapy to rescue retinal degeneration caused by mutations in rhodopsin.* Methods Mol Biol, 2015. **1271**: p. 391-410.
110. Lin, J.H. and M.M. Lavail, *Misfolded proteins and retinal dystrophies.* Adv Exp Med Biol, 2010. **664**: p. 115-21.
111. Ziviello, C., et al., *Molecular genetics of autosomal dominant retinitis pigmentosa (ADRP): a comprehensive study of 43 Italian families.* J Med Genet, 2005. **42**(7): p. e47.
112. Fernandez-San Jose, P., et al., *Prevalence of Rhodopsin mutations in autosomal dominant Retinitis Pigmentosa in Spain: clinical and analytical review in 200 families.* Acta Ophthalmol, 2015. **93**(1): p. e38-44.
113. Willett, K. and J. Bennett, *Immunology of AAV-Mediated Gene Transfer in the Eye.* Front Immunol, 2013. **4**: p. 261.
114. Masli, S. and J.L. Vega, *Ocular immune privilege sites.* Methods Mol Biol, 2011. **677**: p. 449-58.
115. Capozzi, M.E., et al., *Molecular imaging of retinal disease.* J Ocul Pharmacol Ther, 2013. **29**(2): p. 275-86.
116. Drexler, W. and J.G. Fujimoto, *State-of-the-art retinal optical coherence tomography.* Prog Retin Eye Res, 2008. **27**(1): p. 45-88.
117. Kumaran, N., et al., *Retinal gene therapy.* Br Med Bull, 2018. **126**(1): p. 13-25.
118. Maguire, A.M., et al., *Safety and efficacy of gene transfer for Leber's congenital amaurosis.* N Engl J Med, 2008. **358**(21): p. 2240-8.
119. Simonelli, F., et al., *Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration.* Mol Ther, 2010. **18**(3): p. 643-50.
120. Allocca, M., et al., *Novel adeno-associated virus serotypes efficiently transduce murine photoreceptors.* J Virol, 2007. **81**(20): p. 11372-80.

121. Mussolino, C., et al., *AAV-mediated photoreceptor transduction of the pig cone-enriched retina.* Gene Ther, 2011. **18**(7): p. 637-45.
122. Stieger, K., et al., *Subretinal delivery of recombinant AAV serotype 8 vector in dogs results in gene transfer to neurons in the brain.* Mol Ther, 2008. **16**(5): p. 916-23.
123. Vandenberghe, L.H., et al., *Dosage thresholds for AAV2 and AAV8 photoreceptor gene therapy in monkey.* Sci Transl Med, 2011. **3**(88): p. 88ra54.
124. Dalkara, D. and J.A. Sahel, *Gene therapy for inherited retinal degenerations.* C R Biol, 2014. **337**(3): p. 185-92.
125. O'Reilly, M., et al., *RNA interference-mediated suppression and replacement of human rhodopsin in vivo.* Am J Hum Genet, 2007. **81**(1): p. 127-35.
126. Chadderton, N., et al., *Improved retinal function in a mouse model of dominant retinitis pigmentosa following AAV-delivered gene therapy.* Mol Ther, 2009. **17**(4): p. 593-9.
127. Millington-Ward, S., et al., *Suppression and replacement gene therapy for autosomal dominant disease in a murine model of dominant retinitis pigmentosa.* Mol Ther, 2011. **19**(4): p. 642-9.
128. Botta, S., et al., *Rhodopsin targeted transcriptional silencing by DNA-binding.* Elife, 2016. **5**: p. e12242.
129. Bakondi, B., *In vivo versus ex vivo CRISPR therapies for retinal dystrophy.* Expert Rev Ophthalmol, 2016. **11**(6): p. 397-400.
130. Latella, M.C., et al., *In vivo Editing of the Human Mutant Rhodopsin Gene by Electroporation of Plasmid-based CRISPR/Cas9 in the Mouse Retina.* Mol Ther Nucleic Acids, 2016. **5**(11): p. e389.
131. Giannelli, S.G., et al., *Cas9/sgRNA selective targeting of the P23H Rhodopsin mutant allele for treating retinitis pigmentosa by intravitreal AAV9.PHP.B-based delivery.* Hum Mol Genet, 2018. **27**(5): p. 761-779.
132. Bakondi, B., et al., *In Vivo CRISPR/Cas9 Gene Editing Corrects Retinal Dystrophy in the S334ter-3 Rat Model of Autosomal Dominant Retinitis Pigmentosa.* Mol Ther, 2016. **24**(3): p. 556-63.
133. Yu, W., et al., *Nrl knockdown by AAV-delivered CRISPR/Cas9 prevents retinal degeneration in mice.* Nat Commun, 2017. **8**: p. 14716.
134. Ferreira, C.R. and W.A. Gahl, *Lysosomal storage diseases.* Transl Sci Rare Dis, 2017. **2**(1-2): p. 1-71.
135. Neufeld, E.F., *Lysosomal storage diseases.* Annu Rev Biochem, 1991. **60**: p. 257-80.
136. E Neufeld, J.M., *The mucopolysaccharidoses*, in *The mucopolysaccharidoses*, A.B. CR Scriver, WS Sly, DM Valle, Editor. 2001, McGraw-Hill: New York (2001). p. 3421-3452.
137. Harmatz, P. and R. Shediac, *Mucopolysaccharidosis VI: pathophysiology, diagnosis and treatment.* Front Biosci (Landmark Ed), 2017. **22**: p. 385-406.
138. Giugliani, R., P. Harmatz, and J.E. Wraith, *Management guidelines for mucopolysaccharidosis VI.* Pediatrics, 2007. **120**(2): p. 405-18.
139. Desnick, R.J. and E.H. Schuchman, *Enzyme replacement therapy for lysosomal diseases: lessons from 20 years of experience and remaining challenges.* Annu Rev Genomics Hum Genet, 2012. **13**: p. 307-35.
140. Harmatz, P., et al., *Pharmacokinetic profile of recombinant human N-acetylgalactosamine 4-sulphatase enzyme replacement therapy in patients with mucopolysaccharidosis VI (Maroteaux-Lamy syndrome): a phase I/II study.* Acta Paediatr Suppl, 2005. **94**(447): p. 61-8; discussion 57.
141. Schlander, M. and M. Beck, *Expensive drugs for rare disorders: to treat or not to treat? The case of enzyme replacement therapy for mucopolysaccharidosis VI.* Curr Med Res Opin, 2009. **25**(5): p. 1285-93.
142. Biffi, A., et al., *Correction of metachromatic leukodystrophy in the mouse model by transplantation of genetically modified hematopoietic stem cells.* J Clin Invest, 2004. **113**(8): p. 1118-29.

143. Biffi, A., *Gene therapy for lysosomal storage disorders: a good start.* Hum Mol Genet, 2016. **25**(R1): p. R65-75.
144. Piccolo, P. and N. Brunetti-Pierri, *Gene therapy for inherited diseases of liver metabolism.* Hum Gene Ther, 2015. **26**(4): p. 186-92.
145. Sferra, T.J., et al., *Widespread correction of lysosomal storage following intrahepatic injection of a recombinant adeno-associated virus in the adult MPS VII mouse.* Mol Ther, 2004. **10**(3): p. 478-91.
146. Sands, M.S. and B.L. Davidson, *Gene therapy for lysosomal storage diseases.* Mol Ther, 2006. **13**(5): p. 839-49.
147. Nakai, H., et al., *Unrestricted hepatocyte transduction with adeno-associated virus serotype 8 vectors in mice.* J Virol, 2005. **79**(1): p. 214-24.
148. Nienhuis, A.W., A.C. Nathwani, and A.M. Davidoff, *Gene Therapy for Hemophilia.* Mol Ther, 2017. **25**(5): p. 1163-1167.
149. Cotugno, G., et al., *Long-term amelioration of feline Mucopolysaccharidosis VI after AAV-mediated liver gene transfer.* Mol Ther, 2011. **19**(3): p. 461-9.
150. Ferla, R., et al., *Similar therapeutic efficacy between a single administration of gene therapy and multiple administrations of recombinant enzyme in a mouse model of lysosomal storage disease.* Hum Gene Ther, 2014. **25**(7): p. 609-18.
151. Ferla, R., et al., *Gene therapy for mucopolysaccharidosis type VI is effective in cats without pre-existing immunity to AAV8.* Hum Gene Ther, 2013. **24**(2): p. 163-9.
152. Tessitore, A., et al., *Biochemical, pathological, and skeletal improvement of mucopolysaccharidosis VI after gene transfer to liver but not to muscle.* Mol Ther, 2008. **16**(1): p. 30-7.
153. Auricchio, A., et al., *Isolation of highly infectious and pure adeno-associated virus type 2 vectors with a single-step gravity-flow column.* Hum Gene Ther, 2001. **12**(1): p. 71-6.
154. Maddalena, A., et al., *Triple Vectors Expand AAV Transfer Capacity in the Retina.* Mol Ther, 2018. **26**(2): p. 524-541.
155. Venkatesan, A. and A. Dasgupta, *Novel fluorescence-based screen to identify small synthetic internal ribosome entry site elements.* Mol Cell Biol, 2001. **21**(8): p. 2826-37.
156. Swiech, L., et al., *In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9.* Nat Biotechnol, 2015. **33**(1): p. 102-6.
157. Doria, M., A. Ferrara, and A. Auricchio, *AAV2/8 vectors purified from culture medium with a simple and rapid protocol transduce murine liver, muscle, and retina efficiently.* Hum Gene Ther Methods, 2013. **24**(6): p. 392-8.
158. Drittanti, L., et al., *High throughput production, screening and analysis of adeno-associated viral vectors.* Gene Ther, 2000. **7**(11): p. 924-9.
159. Sakami, S., et al., *Probing mechanisms of photoreceptor degeneration in a new mouse model of the common form of autosomal dominant retinitis pigmentosa due to P23H opsin mutations.* J Biol Chem, 2011. **286**(12): p. 10551-67.
160. Liang, F.Q., et al., *Intraocular delivery of recombinant virus.* Methods Mol Med, 2001. **47**: p. 125-39.
161. Gombash Lampe, S.E., B.K. Kaspar, and K.D. Foust, *Intravenous injections in neonatal mice.* J Vis Exp, 2014(93): p. e52037.
162. de Jong, J.G., et al., *Dimethylmethylene blue-based spectrophotometry of glycosaminoglycans in untreated urine: a rapid screening procedure for mucopolysaccharidoses.* Clin Chem, 1989. **35**(7): p. 1472-7.
163. Gerstung, M., E. Papaemmanuil, and P.J. Campbell, *Subclonal variant calling with multiple samples and prior knowledge.* Bioinformatics, 2014. **30**(9): p. 1198-204.
164. McIntosh, J., et al., *Therapeutic levels of FVIII following a single peripheral vein administration of rAAV vector encoding a novel human factor VIII variant.* Blood, 2013. **121**(17): p. 3335-44.

165. Alliegro, M., et al., *Low-dose Gene Therapy Reduces the Frequency of Enzyme Replacement Therapy in a Mouse Model of Lysosomal Storage Disease.* Mol Ther, 2016. **24**(12): p. 2054-2063.
166. Gyorgy, B., et al., *Allele-specific gene editing prevents deafness in a model of dominant progressive hearing loss.* Nat Med, 2019. **25**(7): p. 1123-1130.
167. Maule, G., et al., *Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing.* Nat Commun, 2019. **10**(1): p. 3556.
168. Colella, P., et al., *Efficient gene delivery to the cone-enriched pig retina by dual AAV vectors.* Gene Ther, 2014. **21**(4): p. 450-6.
169. Bovolenta, P. and E. Cisneros, *Retinitis pigmentosa: cone photoreceptors starving to death.* Nat Neurosci, 2009. **12**(1): p. 5-6.
170. Yao, X., et al., *Homology-mediated end joining-based targeted integration using CRISPR/Cas9.* Cell Res, 2017. **27**(6): p. 801-814.
171. Wang, L., et al., *CRISPR/Cas9-mediated in vivo gene targeting corrects hemostasis in newborn and adult factor IX-knockout mice.* Blood, 2019. **133**(26): p. 2745-2752.
172. Giugliani, R., et al., *Therapy for mucopolysaccharidosis VI: (Maroteaux-Lamy syndrome) present status and prospects.* Pediatr Endocrinol Rev, 2014. **12 Suppl 1**: p. 152-8.
173. Giugliani, R., et al., *Natural history and galsulfase treatment in mucopolysaccharidosis VI (MPS VI, Maroteaux-Lamy syndrome)--10-year follow-up of patients who previously participated in an MPS VI Survey Study.* Am J Med Genet A, 2014. **164A**(8): p. 1953-64.
174. Ferla, R., et al., *Non-clinical Safety and Efficacy of an AAV2/8 Vector Administered Intravenously for Treatment of Mucopolysaccharidosis Type VI.* Mol Ther Methods Clin Dev, 2017. **6**: p. 143-158.
175. Chandler, R.J., et al., *Vector design influences hepatic genotoxicity after adeno-associated virus gene therapy.* J Clin Invest, 2015. **125**(2): p. 870-80.
176. Chandler, R.J., M.S. Sands, and C.P. Venditti, *Recombinant Adeno-Associated Viral Integration and Genotoxicity: Insights from Animal Models.* Hum Gene Ther, 2017. **28**(4): p. 314-322.
177. Hanlon, K.S., et al., *High levels of AAV vector integration into CRISPR-induced DNA breaks.* Nat Commun, 2019. **10**(1): p. 4439.