Università degli Studi di Ferrara

DOCTORAL COURSE IN

"Biomedical Sciences and Biotechnology"

CYCLE XXXVII

COORDINATOR Prof. Paolo Pinton

Examining Gene Editing and Promoter Optimization Approaches and Their Implications to Future Therapies

Scientific/Disciplinary Sector (SDS) Bio/11

Candidate

Dott. Silverberg, Jr Joseph

Supervisor

Prof. Mingozzi, Federico

Year 2021/2024

Acknowledgement

I want to thank Ferrara University and Prof. Paola Pinton for providing me the opportunity to pursue a PhD in the Biomedical Sciences and Biotechnology program for Cycle XXXVII. I would also like to thank my mentor, Federico Mingozzi, for offering me this opportunity to become a PhD candidate under his supervision and for his continued guidance throughout this process. I extend my gratitude to Spark Therapeutics Inc. for allowing me the ability to perform my doctoral research in their state-of-the-art laboratories. I also would like to thank my colleagues at Spark Therapeutics, Daniel Cohen, Laura Pereira, and countless other colleagues, all of whom I would not have been able to complete this research without their assistance and guidance. Finally, I thank my parents, Joe and Joanne, along with my sisters, Heather and Danielle, and partner, Nick, who have believed in my abilities and supported me throughout the years. This achievement would not have been possible without the support, mentorship and guidance from those I have mentioned.

Chapter I: Introduction

1.1 Historical Concept Overview of Gene Therapy p4

1.2 Adeno-Associated Viruses p8

1.3 Clinical Applications of Adeno-Associated Viruses p12

1.4 Limitations of Adeno-Associated Viruses p16

1.5 Limitations of Expression Cassette Design, Focusing on Promoters p23

1.6 Tailoring of the Therapeutic Expression p25

1.7 Future Directions of This Research Project p33

Chapter II: Evaluation of Six Ubiquitous Promoters for Targeted Engineered Miniaturization to Generate Potent, Compact, and Tissue-Specific Promoters Ocular Gene Therapies

2.1 Abstract p34

2.2 Introduction p35

2.3 Results p37

2.4 Discussion p49

2.5 Materials and Methods p53

2.6 Appendix p57

Chapter III: Tunable Expression of an AAV Delivered Transgene via Recruitment of Endogenous ADAR Enzymes

3.1 Abstract p75

3.2 Introduction p75

3.3 Results p77

3.4 Discussion p85

3.5 Materials and Methods p87

3.6 Appendix p92

Chapter IV: Conclusion p94

References p97

Chapter 1: Introduction

The field of *in vivo* gene therapy has experienced substantial evolution since first being studied a few decades ago. Advancements in the field have resulted in several drug approvals with products being brought to market to treat genetic diseases. However, further advancements to the field have been hindered by the limitations of the Adeno-Associated Virus (AAV) vector gene delivery platform. Limitations of AAV-based gene therapy include, but are not limited to, the following: packaging constraints, immunogenicity of the viral capsid, systemic toxicities, integration, and durability of expression. The ability to control therapeutic expression of a transgene, in particular, is critical for the continued success of gene therapies and is something that can be achieved through proper cassette design. To this achieve this aim, several different methodologies can be utilized to create truncated promoters. This dissertation will examine two approaches to drive potent transgene expression levels and to achieve tunable expression. The first series of experiments focuses on optimizing the properties of the promoter and other cis-regulatory elements (CREs) to enhance the design of AAV expression cassette. The second approach centers on creating tunable transgene expression utilizing an ADAR editing system to generate an on-switch for an AAV-delivered episome. Collectively, findings from this work will inform future directions to generate compact promoters with increased specificity while preserving potency. To provide context to these experiments, this introduction includes a brief overview of the history of gene therapy, current gene therapy products, challenges faced when developing therapies, and best strategies for tailoring current gene therapy delivery approaches to overcome existing barriers in utilizing AAV in gene therapy.

1.1 Historical Concept Overview of Gene Therapy

The conceptual idea of gene therapy can be traced to early research published in the 1940s with experiments conducted by Avery et al. on bacterial transformation. From these experiments, they concluded that DNA was successfully isolated via alcohol precipitation, which served as the transformative property that could alter the physiology of the bacterial

strain, thus achieving gene therapy in bacteria.[1] The isolation of DNA led to the theory of the central dogma of biology. This implies that the flow of genetic information is unidirectional, stating that DNA is eventually converted into functional products like proteins. The central dogma states that information flows from DNA to RNA, which makes protein, but genetic information cannot travel from protein to nucleic acid. Further building upon the central dogma of biology, Marshall Nirenberg, in 1961, discovered that protein is composed of a unique combination of 21 different amino acids encoded from DNA by three nucleotides.[1, 2] These early discoveries informed future directions for scientists in the field of gene therapy.

In the 1970s, researchers began to investigate the capabilities of viruses to transform cancerous tissues from a healthy state via the virus's ability to insert their DNA and RNA. DNA/RNA transfer mechanism from the viruses was investigated to deliver a therapeutic payload by removing the tumorigenesis part of viruses while maintaining the genome insertion capabilities.[1, 3] Some of the first engineered vectors were produced with gamma retroviruses as the delivery tool due to their ability to effectively insert their genome into targeted cells.[1, 4] Gamma retroviruses are RNA viruses that undergo reverse transcription upon entry into a cell to obtain the DNA sequence, ultimately responsible for integrating into the host cell's genome for viral replication, thus becoming a provirus.[1] To make gamma retroviruses an ideal tool for the delivery of gene therapy, the infectious elements of the virus must be removed. Three key elements of the gamma retroviruses were identified as critical for creating a viral vector to deliver therapeutic payload: envelope, gag protein and polymerase, and vector sequence. The envelope element would ensure virus binding to the host cell and facilitate entry into the cell. The gag-pol element is responsible for vector RNA packaging and reverse transcription of the RNA. The vector sequence elements incorporated the packaging signal elements, which are essential for the packaging process. Additionally, long terminal repeats (LTRs) serve as internal promoters for the payload production and contain elements required to join the two ends together in a circular formation.[1]

Advancements in gene therapy research continued to be observed throughout the 1980s and 1990s. The 1990s saw the start of the human genome project, one of the most ambitious and important scientific endeavors, with the goal of sequencing the entire human genome and the genome of some nonhuman organisms.[5] [6] The project official ended in 2003 lasting 13 years, at that time researchers were able to accurately sequence 92% of the human genome.[5, 7] There were still 400 areas where the DNA sequence was unknown. It was not until 2022 when researchers were able to fill in those gaps and truly complete the human genome sequence.[5, 7] The human genome project identified disease associated genes allowing researchers the ability to study the underling molecular elements of those diseases. Further, this work enabled researchers to develop gene-based treatments, such as production of natural substances to gene therapy strategies.[8] The mapping of the human genome also informed the development of genetic testing for disease causing mutations and to identify recessive carriers of a disease.[8, 9] The human genome project propelled advancements in sequencing technologies, genomics, and bioinformatics all of which have great implications in molecular based research.[10]

It was also during this decade of the 1990s that the limitations of utilizing retroviruses became more apparent. Gamma retroviruses were first used for clinical purposes between the 1980s and 1990s to deliver a normal copy of a defective gene into the genome of hematopoietic stem cells or T-lymphocyte cells from patients suffering from cancer or immunodeficiency, thereby becoming first-generation gene therapy.[11] The first successful clinical application of gene therapy occurred in 1990 in a pediatric patient with adenosine deaminase severe combined immunodeficiency (ADA SCID).[12] Scientists and doctors administered a normal functioning copy of the ADA gene into the patient's T cells via retroviral delivery, making this the first clinical application of gene transfer.[13] The patients were kept on the recombinant protein infusions of ADA even after the administration of gene therapy; this was an interesting clinical approach taken with the patient's best interests in mind, but one that made it difficult to assess the clinical benefit of the gene therapy.[12] Not all early clinical gene therapy trials were as successful as the first one in 1990. A clinical trial in Paris, France, enrolled patients from 1999 to 2002 with

a clinical diagnosis of SCID-X1 for treatment with gamma retrovirus to restore the missing IL-2 receptor γ gene.[13-15] In this clinical trial, 9 of the 10 participants were successfully treated via gene therapy; however, 4 participants later developed T cell leukemia after the treatment with gamma retroviral vector.[15] The onset of leukemia most likely occurred due to retroviral vector integration into host cellular oncogenes. There was also a possibility that the onset of leukemia was not only from viral vector integration but also in conjunction with overactivation of LMO2 gene caused by enhancer sequence of LTR region.[14] Natural limitations of gamma retrovirus, such as the inability to use alternative promoters instead of LTR, payload size restrictions, cell division needed for vector insertion, and integration into 5' untranslated region of genes, led scientists to search for another type of retrovirus to use for vector delivery.[1] Lentiviruses are viral vectors engineered from the human immunodeficiency virus, which, unlike gamma retrovirus, can transfect quiescent cells and have alternative tropisms along with an increased payload size of approximately eight kilobases.[1, 11, 16] However, some of the limitations of lentivirus vectors arise from the concerns that they are derived from the human immunodeficiency virus (HIV), of which there is a possibility of self-replication during the manufacturing process, and the risk of oncogenesis occurs in cells due to integration mutagenesis. Lentiviruses are often used in *ex vivo* genetic modification of cells and are the go-to choice for hematopoietic stem cell-based gene therapies today.[4, 11, 16] These limitations to retroviruses required additional approaches be examined.

Retroviruses are not the only viral vectors used in gene therapy; adeno-associated viral (AAV) vectors are another widely used tool for delivering gene therapies, in particular for *in vivo* gene transfer. In 1965, Bob Atchison, M. David Hoggan, and Wallace Rowe observed in their research that other virus-like particles were engaged in an adenovirus preparation they were working with.[17] This was the first identification of what would become the adeno-associated virus (AAV), and soon after their discovery, the characterization of AAV began. AAV is a member of the *Parvoviridae* family and is classified under the genus *Dependovirus* due to the need for a helper virus such as adenovirus or herpesvirus for viral infection. Still, it is not associated with any known human diseases.[18, 19]

1.2 Adeno-Associated Viruses

AAV is a nonenveloped virus with a single stranded DNA genome of approximately 4.7 kb, encapsulated via an icosahedral protein capsid, and its diameter ranges from 20 to 25 nm.[18-20] Wildtype AAV (wtAAV) comprises a 60-mer subunit for the capsid with the genomic ssDNA flanked by 145bp ITRs on each side. The genome of wtAAV has two main open reading frames (ORF) flanked by ITRs where three promoters, p5, p19, and p40, as well as different initiation codons and alternative splicing, generate replication (Rep) and capsid (Cap) proteins.[19] The left ORF encodes the nonstructural Rep genes required for viral replications, Rep78, Rep68, Rep52, and Rep40.[18, 21] The right ORF is the Cap gene, which encodes for three viral proteins(VP) subunits: VP1, VP2, and VP3 at a ratio of 1:1:10. The Cap gene also contains a third reading frame for assembling activating protein (AAP) needed for assembly of the virus.[18, 20-22] The VP3 protein comprises beta-strands linked by exposed surface variable loops, which aid in the morphology of the AAV and determine the AAV serotype-specific function.[20] The first 125 nucleotides (nt) of ITRs in wtAAV make two small palindromes flanked by a larger palindrome that fold onto each other to create a T-shaped hairpin structure, and the remaining 20nt is the D sequence and stay unpaired.[18, 19, 21] The ITRs are the only cis-acting elements of wtAAV that contain sequences required for viral origins for replication, packaging signals, and integration.[18, 19] The genome of wtAAV has been known to integrate into the human genome on chromosome 19 (q13.4), termed locus AAVS1, and established latency when not co-transfected with a helper virus. However, once co-infection occurs, wtAAV switches from a latent to a lytic infection. [18, 19, 21] Several serotypes of wtAAV have been identified from tissues of human and non-human primates (NHP), of which the ITRs of AAV2 are most often utilized in gene therapy for vector design.[20, 21]

Figure 1: Wildtype Adeno-Associated Virus (AAV)

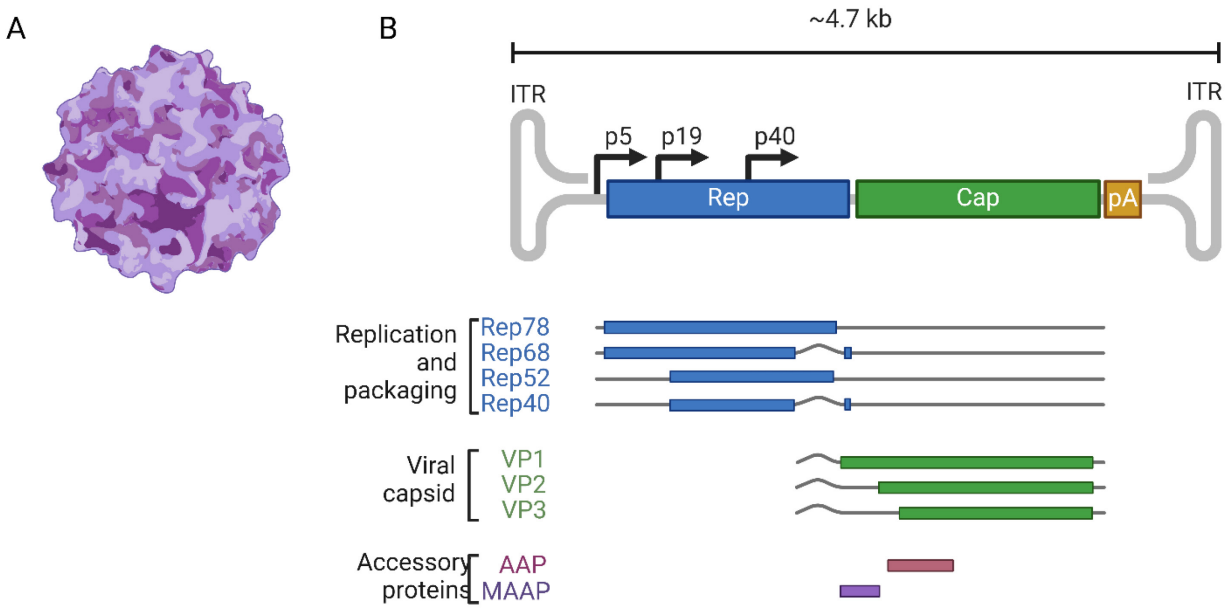


Figure 1 Caption: A. 3D representation of an AAV capsid. **B.** The genetic map of the wildtype AAV. The genome is roughly 4.7kb positioned between the ITR at 5' end and 3' end, which are essential for replication and packing of the viral genome, these are the only wildtype components used in rAAV. Showing the four Rep proteins (Rep 78, Rep68, Rep52, and Rep40) involved in replication, genome packing and regulation of gene expression with transcription driven by promoters p5, p19, and p40. The three Cap genes (VP1, VP2, and VP3) needed for the structural formation of the capsid. Accessory proteins AAP and MAAP are shown which are needed for viral capsid assembly and other functions. [19] (image created in BioRender)

Recombinant AAV (rAAV) vectors are derivatives of wtAAV used for gene transfer. rAAV has the same capsid sequence and structure as the wtAAV. Still, all the protein-coding elements, such as Rep and Cap genes, of the wtAAV are replaced with a therapeutic gene expression cassette flanked by the same ITRs from wtAAV in a rAAV vector.[20, 23] The ITRs from AAV2 are the ones that are most often utilized in rAAV vector designs since

the field has a greater understanding of AAV2 biology.[23] The rAAV vector therapeutic expression cassette components generally comprise a promoter, a therapeutic transgene, and a termination signal flanked by the ITRs. They are limited to a packing size of approximately 4.7kb.[20, 21, 23] The therapeutic transgene contains the DNA sequence of the gene of interest that you want to deliver via the rAAV platform. Examples of transgene utilized in gene therapy products or for research are human Factor IX (hFIX) cDNA for Hemophilia B or a fluorescent report gene such as green fluorescent protein (GFP) cDNA. In the expression cassette, a Kozak sequence is placed at the beginning of the transgene, prior to the methionine codon. This aids in ribosome assembly and translation initiation, ensuring the protein is accurately translated.[24] The termination signal in an expression cassette, which is often a polyadenylation signal (polyA) located at the 3' end of the transgene to indicate the termination of transcription on downstream DNA.[25] Additionally, the termination signal aids in stability and nuclear export through the addition of the polyA tail to the end of the mRNA.[25] Bovine growth hormone polyadenylation (BGH-PolyA) is a commonly used termination sequence in rAAV expression cassette design.

Figure 2: Recombinant Adeno-Associated Virus Expression Cassette

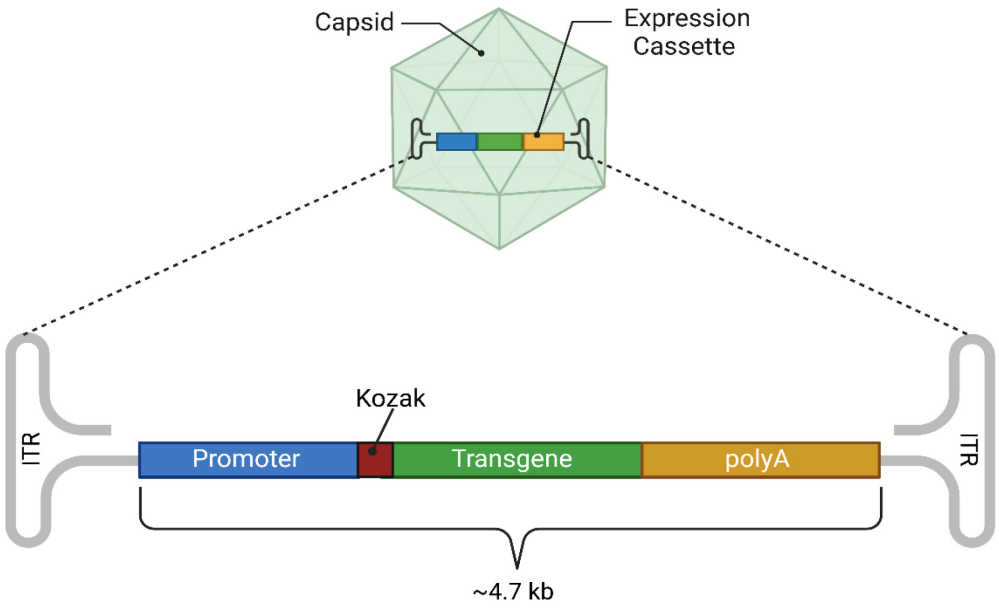


Figure 2 Caption: This is genome map of rAAV expression cassette. The components of the wildtype AAV between the ITRs are replaced with the expression cassette. The expression cassette is composed of a promoter, followed by the transgene with Kozak sequence at beginning, followed by the addition of the polyA tail. All of which are positioned between the 5' ITR and 3' ITR. [26] (image created in BioRender)

The promoter of the rAAV expression cassette is responsible for directing the coding sequence's transcription into mRNA for protein translation. It also indicates the direction in which transcription will occur and from what strand of the coding sequence. Within the DNA sequence of the promoter, there are strands of nucleotides termed transcription factor binding motifs (TFBM), which recruit different transcription factors that interact with RNA polymerase II enzymatic complex to initiate transcription and lead to the production of protein. [26, 27] Promoters in rAAVs can vary in size, with the stronger promoters often being larger, at approximately 1kb. Promoters can also vary in function and specificity. Ubiquitous promoters are expressed across various tissue types and are frequently utilized to maximize the transgene expression throughout the body. Examples of ubiquitous natural promoters are human translation elongation factor-1 alpha subunit (EF1α) and cytomegalovirus (CMV) promoter. Most often, promoters in rAAV expression cassettes are composed by combining different promoter elements to create a hybrid promoter, such as the CAG promoter. CAG comprises the CMV enhancer region, chicken beta-actin promoter, and first intron and splicing acceptor region of the rabbit beta-globin gene.[26] Promoters are not the only elements responsible for driving the transgene expression. Promoters are but one component within a system of different cis-regulatory elements (CREs) utilized in driving transgene expression.[28] Other CREs can include enhancers, introns, and untranslated regions (UTR) of DNA at the 5' or 3' ends.[29] CREs are sequences and structural motifs that regulate mRNA abundance through their interaction with transcriptional activity.[30]

1.3 Clinical Applications of Adeno-Associated Viruses

At the time of their discovery in 1965, AAVs research was not focused on clinical applications but rather to gain a better understanding of how AAVs function.[31] In 1983 the whole genome of AAV2 was sequenced allowing for cloning efforts of the AAV DNA to the engineering of rAAV2 vectors. These rAAV2 vectors demonstrated successful gene transfer in mammalian cells in experiments conducted in 1984, thus unlocking their potential as tool for gene therapy.[31-33] In the 1990s, efforts initiated by the Cystic Fibrosis Foundation, the first *in vivo* experiments conducted in rabbits occurred using rAAV2 vectors to deliver the transgene encoding the cystic fibrosis transmembrane regulator (CFTR), 4.4 kb, with expression driven by AAV p5 promoter interacting with rAAV2 ITR to form a 263bp cassette.[31, 34, 35] This study demonstrated successful gene transfer and protein expression from the rAAV2-CFTR vector following bronchoscopic delivery to a lobe of the lung in New Zealand White Rabbit. In this experiment, vector episomal DNA was observed for at least six months in the airway epithelial cells post-delivery of rAAV2-CFTR .[31, 35] An *in vivo* study with NHPs confirmed the findings seen in the rabbits of persistent vector episomal DNA in the airway epithelial cells post rAAV2-CFTR delivery.[31, 36] In 1995 the first human AAV based gene therapy was conducted where rAAV2-CFTR vector was administered into the maxillary sinuses of 10 pancreas insufficient Cystic Fibrosis (CF) patients.[37] In this study vector persistence was seen for 10 weeks post treatment with rAAV2-CFTR and restoration of CFTR function was observed.[37] Despite the transient detection of the rAAV2-CFTR vector in this study, the investigators did demonstrate successful transduction in the sinuses with rAAV2-CFTR vector. Furthermore, there were no evidence of toxicity adding to the safety profile of rAAV in humans.[37] These advances in AAV research and successful use in clinical research continued to drive advancements in gene therapies.

Different iterations of this clinical trial for AAV based treatment for CF have been conducted since the first one in 1995. However, most trials directly after the original were unable to demonstrate efficacy above placebo and all demonstrated transient infection

from rAAV2-CFTR vector only lasting 60 days, on average.[31, 37] Not long after the start of the AAV focused clinical trials for CF, researchers began the preclinical investigations of rAAV vectors for the treatment of genetic diseases such as Parkinson's, anemia, and hemophilia B.[31] In the treatment for hemophilia B, rAAV2-FIX vectors demonstrated rAAV persistence and sustained expression of FIX in the livers of both mice and canines.[31, 38, 39] These early *in vivo* studies of rAAV vectors for different diseases contributed to the mounting evidence supporting the safety profile of rAAV vectors.[31] This established safety profile allowed for the continued investigation of rAAV vectors in the early 2000s for these diseases and others, such as Leber's congenital amaurosis (LCA), aromatic amino acid decarboxylase (AADC), and lipoprotein lipase deficiency (LPLD). Studies targeting these conditions continued to demonstrate clinical efficacy using rAAV vectors.[31, 40-43] In 2006 the first clinical trial for hemophilia B utilizing rAAV2-FIX was conducted for liver-directed gene delivery by portal vein infusion.[44] In the study, patients that received a dose of 2.00E12 vg/kg had decreasing levels of FIX starting at week 4 and by week 8 FIX was undetectable.[44, 45] The decrease in FIX was accompanied by an elevation of liver transaminases from an immune response related to the capsid.[44]Despite this setback researchers still had hope for gene therapy but further investigation for advancing the field was still needed.

Continued research in AAV delivery for gene therapy progressed throughout the first two decades of the 2000s. Between the years of 2015 to 2018 the number of clinical trials initiated, and IND applications submitted to the Food and Drug Administration (FDA) have quadrupled compared to gene therapy applications submitted before 2003. [46, 47] In 2012, Glybera (alipogene tiparvovec) became the first European gene therapy product, gaining approval from the European Medicines Agency (EMA).[13] Glybera was approved for the treatment of LPLD, a disease that inhibits the ability for proper fat metabolism. In the therapeutic, a copy of the Ser447X gene, the functional variant of the LPL gene, is delivered to patients in the form of a rAAV1 vector. [48, 49] However, Glybera never gained FDA approval, and the product manufacturer did not renew the marketing authorization with the EMA, causing the Glybera to be removed from the market in

October 2017.[50] Despite this initial barrier to bringing gene therapies to US markets substantial improvements in the field continued.

As of March, 2023 there are 3,900 gene therapy clinical trials listed as either completed, ongoing or approved across 46 countries.[51] Of the 3,900 clinical trials from 1994 to 2023 only 350 involve AAV as the intervention or treatment.[51] The clinical landscape of AAV-mediated gene transfer throughout the years has made significant advancements in the clinical application of the treatment of genetic diseases such as congenital blindness, hemophilia, and muscular dystrophy.[52] Currently, 39 FDA-approved cellular and gene therapy drugs are on the market. Of those 39 therapies, 7 are AAV-mediated gene therapy products.[53] The first AAV-based gene therapy for treating disease caused by a specific gene, Luxturna (voretigene neparvovec-rzyl), gained FDA approval in 2017 for treating congenital blindness. Luxturna is an AAV-based gene therapy that delivers a functional copy of the gene encoding for the retinal pigment epithelium-specific 65kDa protein (RPE65) in patients with Leber's congenital amaurosis (LCA). LCA2 is a type of inheritable retinal dystrophies (IRD) due to defective biallelic RPE65 gene that results in loss of vision and even complete blindness.[54, 55] The administration of Luxturna requires the patients to undergo a surgical procedure where the AAV is administered via a subretinal injection into the eye. Treatment is done one eye at a time, 6 days or less before the treatment of the second eye.[56] The FDA's approval of Luxturna in 2017 was followed by the approval of 6 other AAV-based gene therapies for treating diseases such as Spinal Muscular Atrophy, Hemophilia B, Duchenne Muscular Dystrophy, Hemophilia A, and Aromatic L-amino Acid Decarboxylase (AADC) deficiency. The table below details the 7 FDA-approved AAV-based gene therapies, such as the disease, name, manufacture, dose, route of administration (ROA), AAV serotype, and cost.

Table 1: FDA- Approved AAV-Based Gene Therapies as of 11/25/2024

Disease	Therapeutic Name	Manufacture	AAV Serotype	ROA	Dose	Cost	Ref.
RPE65-mutation-associated retinal dystrophy	LUXTURNA® (voretigene neparvovec-rzyl), 2017	Spark Therapeutics, Inc.	AAV2	Subretinal	1.50E+11 vg	$850,000 ($425,000/eye)	[57]
Survival motor neuron 1 (SMN1) associated spinal muscle atrophy	ZOLGENSMA® (onasemnogene abeparvovec-xioi), 2019	Novartis Gene Therapies, Inc.	AAV9	Intravenous	1.10E+14 vg/kg	$2,125,000	[58, 59]
Hemophilia B (FIX)	HEMGENIX® (etranacogene dezaparvovec-drlb), 2022	CSL Behring LLC, uniQure	AAV5	Intravenous	2.00E+13 vg/kg	$3,500,000	[60, 61]
Certain types of Duchenne muscular dystrophy	ELEVIDYS® (delandistrogene moxeparvovec-rokl), 2023	Sarepta Therapeutics, Inc.	AAVrh74	Intravenous	1.33E+14 vg/kg	$3,200,000	[62, 63]
Hemophilia A (FVIII)	ROCTAVIAN™ (valoctocogene roxaparvovec-rvox), 2023	BioMarin Pharmaceutical Inc.	AAV5	Intravenous	6.00E+13 vg/kg	$2,900,000	[64, 65]
Hemophilia B (FIX)	BEQVEZ™ (fidanacogene elaparvovec-dzkt), 2024	Pfizer, Inc.	AAVRh74var	Intravenous	5.00E+11 vg/kg	$3,500,000	[66, 67]
AADC deficiency	EU&UK: Upstaza™ US: KEBILIDI™ (eladocagene exuparvovec-tneq), 2024	PTC Therapeutics, Inc.	AAV2	Intraputaminal infusion	1.80E+11 vg	US price not yet disclosed UK: $3,710,000	[68-70]

Table 1 Caption: Even though 7 FDA-approved AAV-based gene therapy products are available on the market, with over 100 clinical trials underway testing future products,

AAV-based gene therapies are not a perfect platform. There are a multitude of unique and specific factors that can be classified as limitations of AAV-based gene therapies.

1.4 Limitations of Adeno-Associated Viruses

Limitations of AAV-based gene therapy include but are not limited to the following: packaging constraints, immunological issues around the capsid, systemic toxicology, integration, and persistent expression. The genome packaging restraints of AAV are problematic due to its size of being approximately 4.7kb, significantly limiting the ability to incorporate larger transgenes since promoters and other regulatory elements need to be incorporated as well in the vector genome. The 4.7kb restriction not only complicates the vector design for diseases with larger transgenes, but also can be problematic for gene editing applications or when multiple therapeutic cassettes are needed to be expressed in one vector such as with CAR T-cell therapy.[71] Some genetic diseases with large transgenes include cystic fibrosis, Duchenne muscular dystrophy, some forms of Leber congenital amaurosis, and hemophilia A.[72] To try and accommodate larger-sized transgenes such as the ones mentioned above, most efforts focus on creating miniaturized expression cassettes, either on the transgene or the CREs.[73]. For the transgene, parts of the sequence for the gene of interest (GOI) deemed nonessential are often removed in attempts at decreasing the size of the transgene while still allowing for proper protein function. Efforts in transgene miniaturization were successful in the creation of gene therapies both for Duchenne muscular dystrophy (DMD) and hemophilia A. For DMD, a truncated version of the dystrophin gene was identified in individuals with Becker muscular dystrophy, and it was indicated that the truncated version of the dystrophin gene led to the restoration of the functional protein in DMD.[74] In hemophilia A, it was found that the B domain, which is ~44% of the whole FVIII gene, had no functionality in the coagulation function. This allowed for the miniaturization of the FVIII transgene by creating versions with parts or all of the B domain removed.[75] Despite this approach being successful for two approved gene therapies, it may not be applicable for most oversized transgenes. [73] An approach that can bypass the space restriction of AAV to incorporate oversized transgenes is the use of dual AAV vectors.[76] In the dual

vector approach, oversized transgenes are split into two and packaged separately into two different AAV vectors, with the goal of the transgene reassembling *in vivo* through co-infection of both AAV vectors in the same cell.[77] In this approach the transgene is reconstituted in cells infected by the dual AAV vector through concatamerization which is viral genome recombination and subsequent splicing, by homologous recombination of overlapping regions, by the hybrid strategy which combines both the concatamerization and homology overlapping approach into one system, or by protein trans-splicing where joining occurs at the protein level via joining of two polypeptides into one mediated by intervening protein segments (inteins).[77] This approach allows for the expression of the full-length protein but often requires higher dosing of vectors due to low efficiency of reconstitution of the full-length genome. [76]

Figure 3: Strategies for Expressing Transgenes that Exceed the Packing Capacity of AAVs

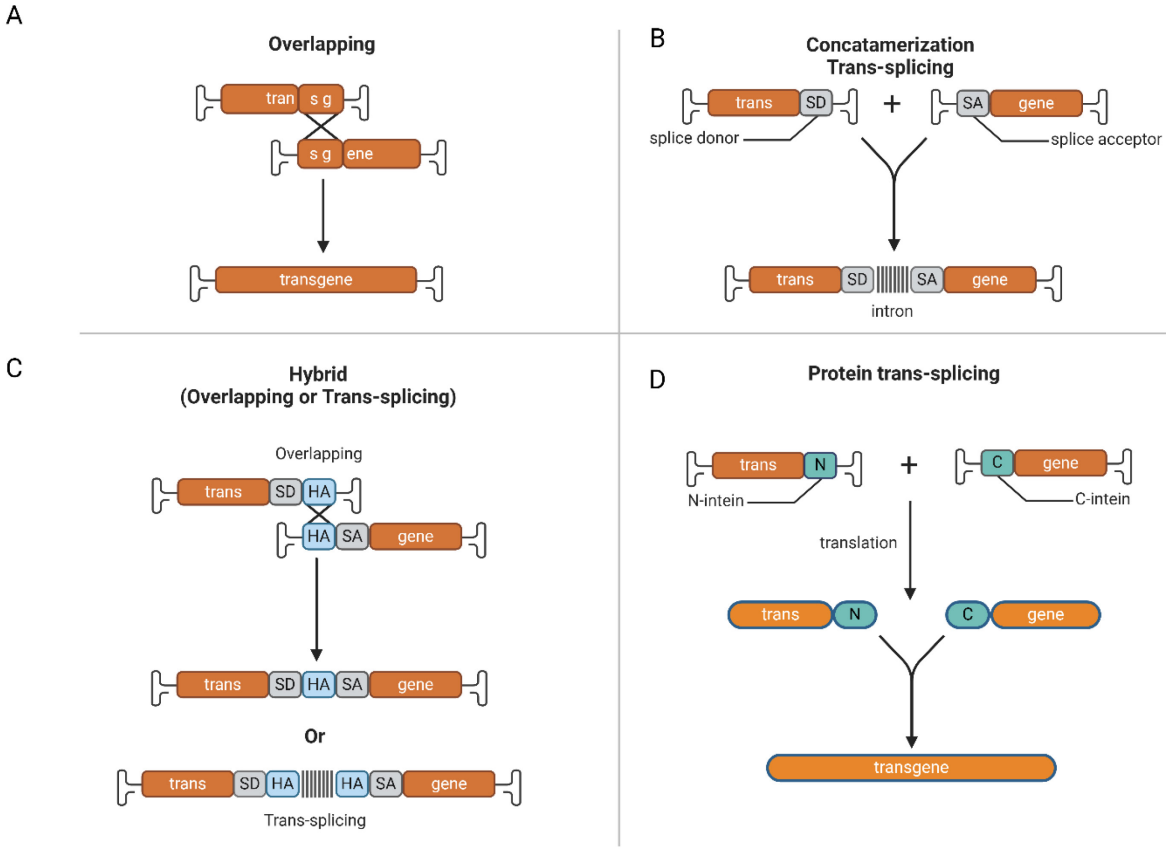


Figure 3 Caption: A. Overlapping strategy where the transgene is split between two AAV vectors with overlapping sequences. Upon delivery the overlapping sequences facilitate homologous recombination to reconstitute the whole transgene. **B.** Concatamerization trans-splicing approach where the transgene is divided into two vectors, one with a splice doner site and other with a splice acceptor site upon co-delivery the transgene is rejoined via splicing at the intron. **C.** Hybrid approach includes both overlapping regions and splicing sites to facilitate either recombination or splicing. **D.** Protein trans-splicing approach is where one vector encodes the N-terminal half the protein fused to intein while the half of the protein is fused to an intein at the C-terminal half, after translation the inteins auto splice reconstituting the protein.[77] (image created in BioRender)

Immune response to the AAV vectors is a considerable limitation of the AAV technology and a severe challenge to the field of gene therapy. The AAV vector capsid closely resembles the capsid of the wildtype virus, which means the capsids can cause immune responses that are similar to those activated by the wildtype virus.[78] When an AAV vector is administered, at least three different immune interactions can alter or inhibit the therapy's effects, including a T cell-mediated response, humoral immune response, and transgene immunogenicity. Studies have indicated the presence of cell-mediated immune responses carried out by capsid-specific CD8+ T cells, which leads to issues with the safety and efficacy of AAV gene transfer.[76] For example, clinical trials utilizing AAV2 and AAV8 vectors for the treatment of hemophilia B resulted in patients experiencing a transient and asymptomatic elevation of liver enzymes post vector administration.[42, 45, 76] This caused loss of FIX transgene expression due to expansion of and activation of capsid-specific CD8+ T cells, which ultimately caused the failure of the transduced hepatocytes.[42, 45, 76, 79] The onset of T cell-mediated immune response to the AAV vector is dose-dependent, with the likelihood of hepatoxicity occurring along the loss of transgene expression associated with activation of capsid-specific T cells due to capsid antigen load crossing a specific threshold.[76] Data suggest that only some patients could require treatment with immunosuppression regimens to manage the onset of the immunological response and mitigate the loss of transgene expression.[42, 76] However,

data from other clinical trials show the ineffectuality of mediating loss of transgene expression with an immunosuppression regimen once a T cell-mediated response is activated.[76, 80, 81] The immunological response initiated by the AAV vector is sometimes directly associated with the dose that is administered. Researchers have observed that high vector doses are often correlated to adverse events such as hepatotoxicity, complement activation, which can lead to more widespread systemic toxicity.[80, 81] When utilized in the clinic, they often need to be co-administered with immunosuppressants to be tolerated for successful gene transfer.[80, 82] However, some diseases, such as neuromuscular ones, require a therapeutic threshold that can only be achieved by very high doses of AAV vectors, often at 1.00E+14 vg/kg or higher.[80]

Humoral immunity against the AAV capsid is developed post-exposure to wildtype AAV. Immunity can be present at birth due to maternal anti-AAV antibodies; however, these will eventually go away a few months after birth until exposure to the virus happens later, usually around the second year of life. [76, 82, 83] Due to the natural occurrence of infections with various serotypes of AAV throughout a person's life, there is just a small window of time when most individuals are naïve to anti-AAV neutralizing antibodies (NAbs), making navigating NAbs a significant hurdle for AAV based gene therapy. NAbs bind to the capsid of the AAV vector, causing the vector to become neutralized, eliminating vector from the body before gene transfer can occur. The abundance of NAbs can significantly limit the efficacy of gene transfer in AAV-based gene therapies. Thus, levels of NAbs must be accurately measured before administration of the AAV vector and often have a cut-off ratio of <1:5.[76, 84] Mitigating the presence of the NAbs can usually be obtained via the route of administration that the AAV vector is given. The presence of NAbs and the titers in the body are not an issue as far as delivery and efficacy in AAV vectors that are delivered via subretinal injection, intramuscular injection, or focal delivery to the brain; it seems that the efficacy of gene transfer is disrupted by NAbs mainly when AAV vector is delivered systemically.[85] In the clinic, one strategy to overcome NAbs to AAV vectors includes transient immunosuppression of patients with corticosteroids such as prednisone.[86] Another strategy involves delivering an AAV vector that has both empty and full capsids in hopes of absorbing the anti-AAV antibodies with the empty

capsids to allow vector transduction with the full capsids.[85-87] Efforts to evade NAbs for possibly redosing with AAV are actively being investigated. Currently, there are no means to redose with the same AAV serotype due to the amount of NAbs produced after the first injection.[88] However, investigators have demonstrated in NHP that measurable levels of two different transgenes could be obtained from subsequential delivery of AAV5ch vector followed by AAV1 vector at clinically relevant dose. Thus demonstrating that redosing systemically could be a viable option via capsid switching.[76, 88] The use of Imlifidase (IdeS) has been investigated as a way to allow for vector readministration since it is an IgG degrading enzyme. IdeS degrades the IgG into F(ab')2 and Fc fragments.[89] Studies conducted in NHP demonstrated that treatment of IdeS to seropositive NHP lowered the presence of neutralizing antibodies to AAV allowing for successful readministration of an AAV vector.[89] An alternative IgG degrading enzyme IdeZ, which is similar to IdeS but utilizes a different IgG degrading enzyme originating from *S.equi ssp.*, is also being investigated as an approach for vector readminstration.[90] Similar to the immune responses associated with the AAV capsid, immune responses specific to the transgene protein can occur, causing transgene immunogenicity. CD8+ T cells can target the transgene, which can induce a humoral immune response and cytotoxicity, all of which would affect the stability of the transgene.[91]

It is apparent that the level of transgene immunogenicity encountered depends heavily on the organ targeted for gene transfer. Higher levels of transgene immunogenicity have been associated with systemic and intramuscular vector delivery with ubiquitous promoters to drive gene transfer but have not been an issue in organs with immune privilege or when targeting the liver via systemic administration and tissue-specific promoters.[91, 92] In a clinical trial for DMD, patients received a rAAV vector dose at 1.00E+11 vg/kg of a microdystrophin transgene driven by CMV promoter and failed to elicit a long term transgene expression. This led investigators to measure cell mediated immune responses in all patients.[93, 94] In one patient investigators observed a T-cell mediated response to non-self episomes that were encoded by the therapeutic transgene. This finding was one of the first instances where investigators saw immune response directed at the transgene itself.[93] It is theorized that the T-cell mediated response

against the transgene could be due to the large sequence difference between the defective self-gene and the therapeutic transgene.[93] In three separate clinical trials enrolling DMD patients, participants all had adverse reactions caused by cytotoxic T-cell immune response to dystrophin.[94] All three clinical trials utilized different microdystrophin transgene driven by different muscle specific promoters and all with different rAAV vectors all dosed intravenously between 1.00E+13 to 2.00E+14 vg/kg.[94] When investigators examined the epitope mapping of the patients and the therapeutic transgenes it was noted that all patients had deletions between exon 8 to exon 21, while all therapeutic transgenes contained hinge 1 and spectrin-like repeat domains encoded by exon 8 to exon 11. Investigators concluded these were causing the non-self episomes. As such, enrollment criteria has been updated to exclude any individuals with these deletions.[94]

Compared to using another virus, utilizing AAV as a delivery tool for gene therapies is generally associated with low levels of vector genome integration in the host cells and, therefore, a low risk for genotoxicity.[95] Despite the historical understanding of AAV vector integration, in terms of clinical applications, this is an issue due to the possible integration of the vector genome into the host's genome. This could result in the disruption of normal cellular homeostasis.[76] rAAV vectors lack the protein in the wildtype AAV needed for integration, indicating that integration of rAAV vectors into the host genome depends on the host itself.[96] AAV vector integration sites are in areas prone to DNA damage or regions of genomic instability. Favored integration sites include chromosomal regions prone to double-stranded breaks, actively transcribed genes, and CpG islands.[76, 96, 97] Unlike retroviruses, AAV vectors have not been associated clinically with malignancy due to the activation of oncogenes. Still, the possibility could arise from delivering a CRE that transactivates a nearby gene.[97] In one *in vivo* study where neonatal mice were injected with AAV, it was observed that the progression of hepatocellular carcinoma occurred due to integration into the Rian locus. However, to date no clinical observation of tumorigenesis due to AAV systemic delivery has been noted. This locus is not present in the human genome.[76, 98, 99] It has been shown that AAV vector dose and high transcriptional activity play a part in the preference of viral

genome integration into a specific locus.[76] The strength of the promoter included in the design of the expression cassette also plays a role in the upregulation of genes near the AAV vector's integration site.[98] It has been reported that strong promoters, like CAG and thyroxin-binding globulin (TBG), led to tumorigenesis via activation of nearby genes. Still, tissue-specific promoter human alpha-1 antitrypsin (hAAT) did not.[76, 98] AAV integration events could be mitigated via thoughtful designing of the transgene expression cassette by incorporating regulatory elements that account for the delicate balance between genotoxicity and expression potency.

Gene therapies delivered via an AAV are thought to have a persistent expression that does not wane over time. This is because AAVs transduce both dividing and nondividing cells, and the persistent expression comes from the formation of DNA payload forming circular episomes in transduced cells.[76, 100] AAV-based gene therapies are generally delivered to tissues and/or cell types with relatively low cellular turnover levels, leading to persistent and durable expression.[100] Durable and persistent expression has been noted in long-term follow-ups in animal models and clinical trials. Data from a hemophilia focused study where canines were administered with rAAV-FVIII demonstrated persistent and durable expression 10 years post-AAV delivery.[101] In 2018, it was reported that patients treated with scAAV2/8-LP1-hFIXco for hemophilia B had persistent expression up to 8 years after administration.[102] However, persistent expression of therapeutic transgene may not always be wanted. For example, persistent expression of gene editing modules such as CRISPR-Cas9 and DNA base editors can lead to increased risk of off target effects and possible triggering of an immune response from prolonged expression of non-human proteins like Cas9.[103, 104] Overexpression of therapeutic transgene, which is often associated with the effects of the promoter due to persistent expression, can lead to transgene toxicity.[105, 106] Therefore, it would not be ideal to have persistent expression when targeting diseases with dynamic disease states, such as inflammatory-related diseases.[107]

1.5 Limitations of Expression Cassette Design, Focusing on Promoters

The limitations previously observed in clinical trials also apply to pre-clinical research efforts focused on the design of the expression cassette, CREs, and promoters. Specifically, these challenges relate to the promoter since the promoter has implications in all the limitations outlined earlier. These limitations include limited capacity, immunogenicity, genotoxicity, and systemic toxicities. The importance of promoters is undeniable to advancing gene therapies. For example, the use of a small promoter within an expression cassette has direct impact by saving valuable space for a large transgene given the limited capacity of an AAV.[108, 109] The use of the right promoter has a direct impact on immunogenicity, if a potent and tissue specific promoter is utilized in an expression cassette it can significantly reduce the required dose of the vector needed to achieve a therapeutic effect and lower immunogenicity by efficiently targeting gene expression in the desired tissue.[110, 111] The use of a tissue specific promoters can reduce the chances of genotoxicity by limiting the transactivation activity.[112] An *in vivo* study in mice for hepatocellular carcinoma saw ubiquitous promoters CBA and TGB increased events of genotoxicity due increasing activity of nearby genes from the promoters. However, with hAAT, a liver specific promoter, this did not occur.[112] Systemic toxicity from AAV vectors is associated with a high vector dose. As such, incorporating a potent and tissue specific promoter into the expression cassette design could reduce the possibility of systemic toxicity by lowering the required vector dose to achieve therapeutic expression.[113] As mentioned previously, the packing constraints are tight within rAAVs, with packaging capacity being ~4.7kb.[18-20] When designing an expression cassette, it is optimal to design the expression cassette around the therapeutic transgene and optimize the promoter and other CREs for the specific needs of the therapy, such as tissue specificity, transgene size, and desired potency.[28, 114] It has been reported that approximately 50% of clinical applications for AAV gene therapies utilize large ubiquitous promoters like the CMV, CBA, and CAG promoters. These promoters are often associated with strong expression in various cell types.[110, 115, 116] Despite CMV, CBA, and CAG promoters being strong ubiquitous promoters, they are often utilized in AAV therapies for diseases requiring specific tissue and organ targeting. This is because they can achieve

a high expression level throughout various tissue types.[110, 117] These promoters are often utilized at the expense of the size of the therapeutic transgene used in an expression cassette since they range from 500bp to over 1kb.[110]

Balancing promoter size and potency can be a delicate process when designing an expression cassette. If investigators choose to utilize a small promoter, this should only be done if the transgene size is fairly large. A promoter that is too small when it is not needed can be detrimental to the overall design of the expression cassette and affect the stability and packaging efficiency of the AAV vector, especially when the total cassette is under 4.1kb.[118] The smaller the promoter is, the less potency it tends to possess compared to larger promoters. This can be problematic and lead to the use of higher therapeutic doses to achieve the same level of potency.[108, 119] The alternative option to utilizing smaller promoters is larger promoters, which tend to be ubiquitous and are often constitutively expressed. There are several benefits when transgene size allows for the use of larger ubiquitous promoters. Ubiquitous promoters allow for the expression across a variety of tissues, often with consistent levels of expression, and are often well characterized and have proven efficacy.[110, 120, 121] A drawback of ubiquitous promoters is their ability to express in various cell types. This can lead to overexpression of the transgene in unwanted tissues, causing off-target expression, and potential toxicities.[116] Dorsal root ganglia (DRG) toxicity is an example of non-specific transgene expression which can lead to neuroinflammation and neuronal cell death which was observed *in vivo* for a SMA gene therapy utilizing AAV9 as the vector.[122] Some ubiquitous promoters, like CMV and potentially CAG, can go through CpG methylation. This can cause gene silencing, essentially shutting down the expression of the transgene.[116, 123] Depending on the therapeutic transgene and the tissues associated with the disease, using a ubiquitous promoter might not always be favorable. Therefore, a tissue-specific promoter should be utilized. Tissue-specific promoters tend to be less potent than ubiquitous ones. Still, the lack in potency is not as dramatic since tissue-specific promoters can directly target specific cells and eliminate common off-target expression seen with ubiquitous promoters.[121, 124, 125] Even in systemic delivery, the mechanism of nontarget delivery of an AAV with a moderate dose and using a tissue-

specific promoter can help reduce the genotoxicity and toxicity associated with the transgene.[112] A potential reason why tissue-specific promoters have less toxicity associated with them than ubiquitous promoters is that they tend to drive lower levels of transgene expression than ubiquitous promoters.[112, 119, 126] The ability to alter and regulate therapeutic expression is critical for the continued success of gene therapies, something that can be achieved through proper cassette design.

1.6 Tailoring of the Therapeutic Expression

Promoter specificity is the ability of the promoter to drive gene expression in a selective base manner, such as towards a particular tissue, cell type, or condition. The specificity level depends upon the regulatory elements within the promoters and the transcription factors to which they recruit.[109] Obtaining promoter specificity is essential in the gene therapy sector since it allows precise control over transgene expression, which could help diminish off-target effects and improve efficacy.[127-129] The optimal approach by which to increase the specificity of the promoter is by selecting tissue-specific promoters since they already have intrinsic properties that limit the reach of their expression.[109, 130] Depending on the chosen promoter, sometimes further optimization may be needed to achieve the required specificity and potency while adhering to the size restrictions of AAV packaging capacity. Often, it is ideal to start with the core or minimal promoter region since the activity of RNA polymerase II and basal transcription machinery present in the area are sufficient for initiating gene expression; including other CRE that could enhance expression levels.[131, 132] Novel core promoters can be easily identified with bioinformatics tools by their upstream proximity to the transcription start site of 25 to 30bp and via the presence of the TATA box for tissue-specific promoters or with Initiator (Inr) element for ubiquitously expressed promoters.[132, 133] Once the core region of the promoters is identified, the other CREs can be easily located upstream or downstream of the core region if they are chosen to be included in the design of the expression cassette. As mentioned, using the core promoter alone might be insufficient to achieve the desired expression level unless other CREs are incorporated into the design.

One approach to achieve desired expression is to create hybrid promoters by combining a promoter region of one promoter with an enhancer of a different one to elicit more potent expression in the desired tissue or cell type.[131] In the literature, there are many examples where hybrid promoters are created to increase expression by adding the enhancer region of the human cytomegalovirus immediately early (CMV) promoter into the expression cassette, with CAG being the most well-known.[134, 135] Previously, CMV enhancer has also been added to the expression cassette of muscle-specific and neuronal-specific promoters to increase expression levels in their respective tissues with success.[136-138] Another potential approach to increase the specificity of a promoter is by incorporating silencers into the expression cassette. Yet, there is a lack of research examining the use of slicers in expression cassettes. However, this strategy could serve to increase specificity of promoters and deter off target expression. Silencers are a cis-regulatory sequence that causes a decrease in transgene expression when bound with transcription factors. However, this can depend upon the cell type in which they are located.[139] For example, the cone-rod homeobox (CRX) transcription factor has recognition sequences in enhancers and silencers for different cell types within the retina.[140] Within the retina, CRX binds to enhancer sequences within rod photoreceptors, while in cone photoreceptors, it represses expression by interacting with sequences in the silencer element.[140] Regardless of the cis-regulatory region CRX interacts with, it recognizes the same binding site in silencers and enhancers.[140] This could indicate whether CRE acts as a silencer or enhancer depending on the cell type in which they are located and that silencers are often enhancers in different cell types.[140, 141] This could also suggest that the transcription factors recruited by the CREs are some of the most important elements that could increase specificity in an expression cassette.

Transcription factors are critical elements that are essential in the biological process by exerting control over developmental patterning as well as tissue differentiation.[142, 143] Transcription factors are comprised generally of two domains: a DNA-binding domain and a trans-activating domain.[144] The DNA-binding domain recognizes specific sequences in DNA that will direct the transcription factor to bind, and the trans-activating domain imposes regulatory effects, which often involve binding RNA polymerase.[144, 145]

Sometimes, the DNA binding domain can form protein-protein interaction with coregulators, which allows activity to be influenced by the recruitment of other transcription factors.[145] In a mammalian system, over 1,500 transcription factors are encoded into the genome, and they are often grouped together in families according to the structure of their DNA binding domain.[146] One analysis of transcription factors in mouse tissues noted that half of the transcription factors detected demonstrated properties of enrichment in certain tissues, indicating they could be tissue-specific transcription factors.[146] Investigators also noted that certain transcription factors ubiquitously interacted with other transcription factors to regulate expression, while also detecting transcription factors that showed properties of being tissue-specific did not demonstrate this phenomenon.[146] These findings indicate that the incorporation of tissue-specific transcription factors or, more specifically, transcription factor binding motifs (TFBM) into an expression cassette could potentially increase the likelihood of specificity. This is thought to be due to some tissue-specific transcription factors can function independently to activate gene expression in select targeted tissues and not rely on multiple tissue-specific transcription factors.[147] If increased specificity is desired for targeting ocular tissues, specifically photoreceptors incorporating TFBM for transcription factors like CRX, RX2, NR2E3, these could assist in increasing specificity since these transcription factors are involved in driving differentiation of photoreceptors during development.[148] The relationship between CREs and transcription factors is pivotal for both the level of expression and the specificity that is obtained. Therefore, methodologies that are used for promoter truncation must try to preserve the necessary TFBM in order to maintain expression.

The need for compact promoters in AAV-based gene therapies is critical for incorporating oversized transgenes into an expression cassette since the packaging capacity is limited to 4.7kb.[17, 20, 21, 23] Several different methodologies can be utilized to create truncated promoters. Some of the most common techniques are sequential deletion analysis and bioinformatics-based approaches to generate synthetic or hybrid promoters. Sequential deletion analysis is the most widely used technique to identify critical regions of the promoters to preserve expression. In sequential deletion analysis, small promoter

regions are deleted from the sequence one at a time until regions that control transcription are identified using a reporter gene, and expression is evaluated in assays like a luciferase-based one.[149] Sequential deletion analysis is a very direct and simplistic method to truncate a promoter. Still, it can often be labor-intensive and time-consuming since each deletion needs to be cloned and evaluated individually. Bioinformatic tools have been utilized heavily within the field to create compact promoters and to increase cell specificity through synthetic promoter design or the creation of hybrid promoters. One example of utilizing bioinformatics for creating compact synthetic promoters was conducted by Johnson et al., where investigators identified *in silico* transcription factor response elements specific to dendritic cells utilizing publicly available transcriptomic data sets. The researchers combined these elements to generate synthetic promoters all under 500bp and evaluated expression.[150] Promoters that are compact and tissue-specific are not the only elements that might be desirable from a promoter. For example, the use of a regulatable promoter can sometime be beneficial for gene therapies.

A regulatable promoter has many benefits for gene therapies, including control of expression levels, temporal regulation, and increased safety. In theory, a regulatable promoter should allow for the controlled expression of a transgene. This is desirable for gene therapies because having the ability to fine-tune transgene expression would ensure the therapeutic effect is reached without any added toxicities, especially when overproduction of the transgene could be toxic.[87, 151] Regulatable promoters that provide dynamic control of transgene expression could greatly benefit future gene therapy development. These promoters enable therapies for diseases where consistent expression is not desirable. Instead, these promoters allow activation of gene expression during specific stages of disease progression or in response to a triggering event. This approach is particularly relevant for chronic conditions like type 1 diabetes mellitus or rheumatoid arthritis.[152, 153] The safety aspect of gene therapies could be increased significantly through regulatable promoters by minimizing potential adverse events. Safety could be enhanced by fine-tuning the expression to appropriate levels or by incorporating a kill switch to halt transgene expression if severe adverse events occur.[152] Regulatable promoters are utilized not only for the fine-tuned control of the

transgenes, but also for the controlled use of gene editing tools, such as base editors and CRISPR-Cas9 technologies.[154-156]

Similar methods mentioned earlier as ways to increase promoter specificity, such as incorporating transcription factors binding sites, can be utilized to generate tunable promoters. Another form of tunable promoters includes inducible promoters.[157] Inducible promoters have regions within their sequence that can be controlled by external stimuli. Stimuli, such as heat, metal ions, interferons, and steroids, can regulate the expression of the promoter by turning them on or off at the transcriptional level.[152] One of the most widely utilized tunable systems is the Tet system, where gene expression is controlled through tetracycline or doxycycline to create a Tet-Off or Tet-On system.[158, 159] This system is based upon the binding of the Tet repressor protein with the DNA sequence of the Tet operator where the introduction of tetracycline or doxycycline inhibits the binding of Tet repressor protein, causing conformation change that prohibits binding to the Tet operator creating the Tet-Off system.[159] Meanwhile, the Tet-On system is activated only in the presence of tetracycline or doxycycline, while in its absence, it is in a steady state.[159] The Tet system enables researchers to understand the temporal effects of gene expression since the system allows for specific and reversible control of expression with minimal off-target effects due to the high-affinity binding of tetracycline or doxycycline with the transcriptional activator. [158-160] Some limitations of the Tet system could be the complexity of cassette design since the tetracycline-controlled transactivator protein sequence and several repeats of the Tet operator need to be incorporated into the design, taking away from the already limited cargo space.[157] Early designs of the Tet system had issues with leakiness, but the most recent designs have addressed the issue with limited leakiness observed now. Even though the Tet system has been utilized in rodents and non-human primates, there are still concerns about translating into human. This is due to the potential risk of immunogenicity issues since the system utilizes bacterial-derived proteins.[157, 161, 162] Another tunable system that works on the transcriptional level is the rapamycin-inducible system. This system works through the interaction of two inactive transcription factors: a DNA binding domain and a DNA transcriptional activation domain. In this system each domain is fused with a heterologous

binding domain that recognizes rapamycin.[163] When rapamycin is added to the system, the binding and activation domains form a heterodimeric transcription factor that allows transgene expression.[163, 164] The advantage of this system is that rapamycin is an approved drug that can cross the blood-brain barrier and has a dose-dependent response allowing for controlled gene expression.[164, 165] Key disadvantages of the rapamycin system include that adding rapamycin can disrupt the metabolism and impact the growth of cells if it interacts with the endogenous mammalian target of rapamycin complex 1.[164, 165] Another drawback of the system is the slow kinetics for turning gene expression off once rapamycin is removed.[164]

Figure 4: Tet On/Off Inducible Promoter

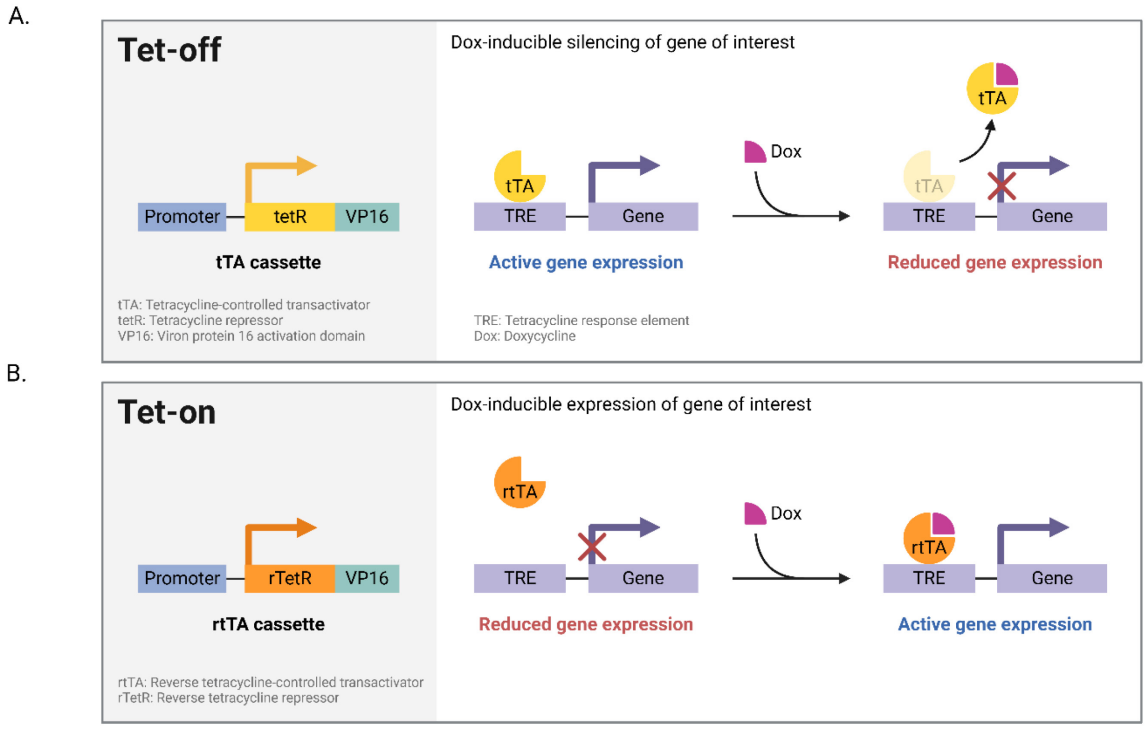


Figure 4 Caption: A. The Tet-Off system the tetracycline-controlled transactivator (tTA) binds to the tetracycline response element (TRE) allowing for transcription of the gene. In the presence of tetracycline or doxycycline (Dox) in the system it binds with the tTA causing it to dissociate from the TRE preventing transcription resulting

in an off state. **B.** The Tet-On system the reverse tetracycline-controlled transactivator (rtTA) cannot bind to TRE. In the presence of Dox, Dox binds to the rtTA allowing it to bind to the TRE thus activating transcription of the gene resulting in gene expression being on.[158-160] (image created in BioRender)

There are also tunable platforms that work on the post-transcriptional level. Some examples of these platforms are RNA interference, riboswitches and aptamers, and RNA editing tools like adenosine deaminase acting on RNA (ADAR).[157, 166, 167] Post-transcriptional gene regulation is the ability to control gene expression at the RNA level directly after transcription has occurred but before translation into protein begins. [168] RNA interference (RNAi) is a regulatory mechanism in eukaryotic cells that can initiate gene silencing.[169] These can be utilized in gene therapies through the incorporation of small RNAs, such as small interfering RNAs and microRNAs, to interact with the enzymatic machinery of RNAi. This can inhibit translation of specific mRNAs through homology-dependent degradation, allowing for the fine tune control of gene expression.[163, 169, 170] The Advantages of RNAi are the precise targeting of specific mRNAs, which can successfully distinguish between target and nontarget genes, and high potency. This means lower levels of effector molecule concentrations are needed to tune expression when using RNAi. [169, 171] Some limitations of RNAi include potential competition with endogenous RNAs by saturating microRNA machinery, which could cause disruptions in natural gene processes.[169, 172] Another limitation is the possibility of initiating an immune response to siRNA sequence motifs, especially if they have GU-rich regions that interact with Toll-like receptors, leading to the release of inflammatory cytokines.[169, 173] Riboswitches are small naturally occurring RNA structures with ligand-sensing aptamer domains and expression platforms that regulate gene expression through aptamer ligand binding.[174-176] Riboswitches can regulate gene expression through interactions with the polyA signal, stopping mRNA processing and regulating splicing sites, impacting the presence or absence of exons in mRNA transcripts.[175] Riboswitches have been utilized to treat anemia *in vivo* with an AAV-delivered erythropoietin transgene, where expression is tuned post-delivery by small-molecule drugs.[175, 177] Some key advantages of using riboswitches to regulate expression

include their small size for packing constraints, they are non-immunogenic, and they are modular, allowing for swapping of aptamer domain and expression platform for new riboswitches that interact with different ligands.[175] A major drawback of riboswitches is the poor performance observed in mammalian cells. Efforts are underway to try and improve their performance and increase their use in therapeutic applications.[175, 178]

Even though ADAR is an RNA editing technology, this aspect of the system can be utilized to control transgenes in gene therapies. The mammalian system of the ADAR enzyme family is composed of ADAR1, ADAR2, and ADAR3, all of which are RNA editing enzymes that deaminate adenosine to inosine in double-stranded RNA substrates in both coding and non-coding regions of the codon.[107, 179] All three members of the ADAR gene have the same structural domain components consisting of double-strained RNA binding domains that are 65 amino acids (AA) in length and a separate catalytic domain. [175, 180] ADAR1 and ADAR2 have catalytically active enzymes, while ADAR3 has yet to be shown as catalytically active.[181] ADAR2 is primarily active in the brain but has demonstrated expression elsewhere, while ADAR3 is exclusively found to have expression only in the brain, while ADAR1 is ubiquitously expressed.[181] ADAR1 has two isoforms p110 and p150. ADAR1p110 is mainly localized in the nucleus, while ADAR1p150 is primarily active within the cytoplasm.[175, 180] All ADAR family members play an essential function during developmental stages and are associated with RNA editing, which is required for normal central nervous system functioning. In contrast, editing deficiency has been linked to neurological disorders, such as epilepsy.[175, 180, 182] Autoimmune diseases such as Aicardi-Goutières syndrome have also been associated with mutated ADAR1-encoded genes.[175] ADAR can mediate various cellular processes that can aid in regulating expression, such as altering splicing patterns, changing mRNA isoforms, and ultimately altering protein products.[183] Another way ADAR can be utilized for regulating expression is by changing the coding region of the mRNA through the A to I edits, thus altering the functional amino acid.[179, 183] Key advantages of using ADAR are the transient nature of the effects of the editing, which reduces the risk of lasting off-target effects, and the safety aspect of the ADAR system since it is an endogenously expressed enzyme in mammals.[179, 183] Some limitations

of the ADAR system include the risk of off-target editing, the limited scope of edits to A-to-G only, and the instability of RNA, which requires methods to improve the stability of guide RNAs.[107, 183] Utilizing circular guide RNA (cadRNA) in an AAV expression cassette has been shown to improve stability and recruit endogenous ADAR1 enzyme.[184]

1.7 Future Directions of This Research Project

The overarching goal presented in this work demonstrates two ways in which transgene expression levels can be controlled for tunable expression for use in AAV-based gene therapies. The first approach is through the proper design of an AAV expression cassette, focusing mainly on the properties of the promoter and other CREs. These efforts identify ways to generate compact promoters with increased specificity without huge sacrifices to potency to be incorporated in AAV expression cassettes with oversized transgenes primarily for use in ocular-related diseases. The second approach for creating tunable transgene expression utilizes an ADAR editing system to generate an on-switch for an AAV-delivered episome.

Chapter II: Evaluation of Six Ubiquitous Promoters for Targeted Engineered Miniaturization to Generate Potent, Compact, and Tissue-Specific Promoters Ocular Gene Therapies

2.1 Abstract

This study explores the feasibility of miniaturizing ubiquitous promoters to generate promoters that are compact, retain potency, and become more tissue-specific for potential use in ocular-related gene therapies where tissue-specific small promoters are needed to accommodate transgene expression cassettes into adeno-associated virus (AAV) vectors. We characterized six ubiquitous promoters (i.e., Ubiq-1, Ubiq-2, Ubiq-3, Ubiq-4, Ubiq-5, and Ubiq-6) *in vitro* across an array of cell types (i.e., ARPE19, Y79, Neuro2A, Huh7, and HEK293) utilizing either a fluorescent report gene dTomato or hFIX transgene to measure their expression and evaluate key components of each promoter. We performed a targeted deletion series on the Ubiq-6 promoter and generated variants of all the promoters where transcription factor binding motif (TFBM) were swapped to increase tissue specificity. Our results show that the presence of a 5'UTR and/or intron are cell-type dependent when it comes to the expression level achieved in the promoters evaluated. Our tiled ablation deletion series for Ubiq-6 identified key areas of the promoter, resulting in truncated promoters of 282bp to 682bp in length, named Comp-6 and Comp-7, respectively. We identified that tissue specificity increased as promoter size decreased. This showed that swapping TFBM with ocular-specific TFBM can increase expression in ocular cell lines, as seen with the Comp-7-Var1 promoter in Y79 cells ($p < 0.0001$). These findings suggest that ubiquitous promoters can be truncated to maintain potency and be engineered to have properties that increase tissue specificity. These hypothesis generating findings inform future research examining promoters that can be utilized in gene therapies with oversized transgenes to treat ocular-related diseases.

2.2 Introduction

The use of adeno-associated virus (AAV) has proven to be a helpful tool in the delivery of gene therapies, especially for ocular-associated diseases, with clinical trials utilizing AAV-directed gene augmentation for diseases such as choroideremia, XL retinoschisis, retinitis pigmentosa, and neovascular age-related macular degeneration.[185] Some clinical trials in these ocular diseases have utilized ubiquitous promoters, such as CMV, CAG, and CBA, to drive the transgene products.[185, 186] Although using ubiquitous promoters to drive the transgene expression is suitable for many ocular diseases, it is not an option for all. Most ubiquitous promoters are large and sometimes achieve greater than 1kb in length, which can be problematic due to the genome packaging constraints of AAV vectors being limited to approximately 4.7kb.[187] This limited size leaves little space, especially given that genes associated with ocular diseases are often larger or just under 4.7kb, to incorporate an appropriate promoter to drive transgene expression.[188] Additionally, it has been reported in the literature that expression driven by ubiquitous promoters can be associated with transgene toxicity.[126] Researchers have reported that when ubiquitous promoters such as CAG, CBA, and UbiC were tested with a reporter gene, GFP, they saw substantial toxicity in retinal pigment epithelial cells (RPE), but the toxicity was not present when tested with a photoreceptor-specific promoter.[126] This phenomenon of toxicity from ubiquitous promoters and none from tissue-specific promoters has been seen in other tissues, such as the heart and central nervous system.[126, 189-191] Therefore, subsequent research focused on optimizing promoter size while mitigating toxicity is needed to advance AAV research.

One of the major advantages of gene therapy over other therapeutic modalities like small biologics or antibodies is the ability to enable precise and cell specific activation of the transgene.[152] The ability to allow for spatial and temporal control of gene expression is a major differentiator from the other therapeutic modalities mentioned where there is an increased risk of off target activity often from systemic exposure.[20, 192] The development of cell specific promoters will further differentiate gene therapy from other therapeutic modalities by limiting the chances of off target expression.[109, 131] Cell

specific promoters will ensure the expression of transgenes in desired cell types through the binding of transcription factors that would regulate expression to the desired cell type only increasing the safety profile and ensuring there would be no activation of potential oncogenes.[109, 127, 131, 193]

The promoter is not only responsible for driving the expression of the transgene, but it also determines the cell specificity for which it requires activation by transcription factors present in the transduced cells.[117] Much research has gone into finding retina-specific promoters of various sizes, often larger than 1kb, to be utilized in ocular gene therapies.[117, 194-197] Efforts have been made to engineer native ocular tissue-specific promoters to be smaller; however, those promoters range in size from 784bp to 2484bp.[198] Tissue-specific promoters are not the only strategy to regulate transgene specificity but can be advantages when compared to other mechanism like the incorporation of miRNA. Tissue-specific promoters control specificity of the transgene at the transcriptional level and rely on endogenous regulatory elements leading to more direct regulate of gene expression, while miRNA control expression post transcriptionally through miRNA binding sites added in the transgene mRNA.[127, 199-201] Tissue-specific promoters often obtain maximum expression in the targeted tissues with minimal expression in unwanted tissues, while miRNA tend to have leaky expression due to incomplete suppression in non-targeted tissues.[200-202] Thus, enhancing the specificity of ubiquitous promoters capable of driving transgene expression is therefore needed.

Most AAV optimization around cis-regulatory element research focuses on the miniaturization of tissue-specific promoters for use in ocular gene therapies.[198] The present study focuses on enhancing miniaturization efforts for ubiquitous promoters. In this study, we aimed to characterize six ubiquitous promoters and test their potency across various cell types. We then tested the different functional elements of the promoters as a means to miniaturize them and identified that 5’UTR and intron are integral for expression depending on the cell type. Utilizing the Ubiq-6 promoter, we were able to generate miniaturized versions that ranged in size from 682bp to 282bp. We also noted that cell type specificity increased as the promoters decreased in size. This led us to swap

transcription factor binding motifs (TFBM) within our ubiquitous promoters with ocular-specific TFBM to generate more tissue-specific promoters.

2.3 Results

Six ubiquitous promoters (Figure 1) were identified and screened across various cell types and benchmarked against the control promoter EF1A to determine if they truly had ubiquitous characteristics. Specifically, promoters Ubiq-1, Ubiq-2, Ubiq-3, Ubiq-4, Ubiq-5, and Ubiq-6 were utilized and examine across ARPE19, Y79, Neuro2A, Huh7 and HEK293cell lines. To test functionality in each cell type, the promoters were transiently transfected, and 48 hours later, FACS analysis was performed to measure the mean fluorescence intensity of the report gene, dTomato. The promoters had the most similar expression when compared directly to the EF1α in ARPE19 (retinal pigment epithelium), Y79 (retina), and Neuro2A (neuroblastoma) cell lines (Figure 1A, 1B, 1C). In Huh7 (liver) and HEK293 (kidney) cell lines, expression varied across the different ubiquitous promoters when normalized to EF1α (Figure 1D, 1E), with Ubiq-5 having little to no expression in both cell lines.

Figure 1: Candidate Promoters Potency Observed Across Multiple Cell Types

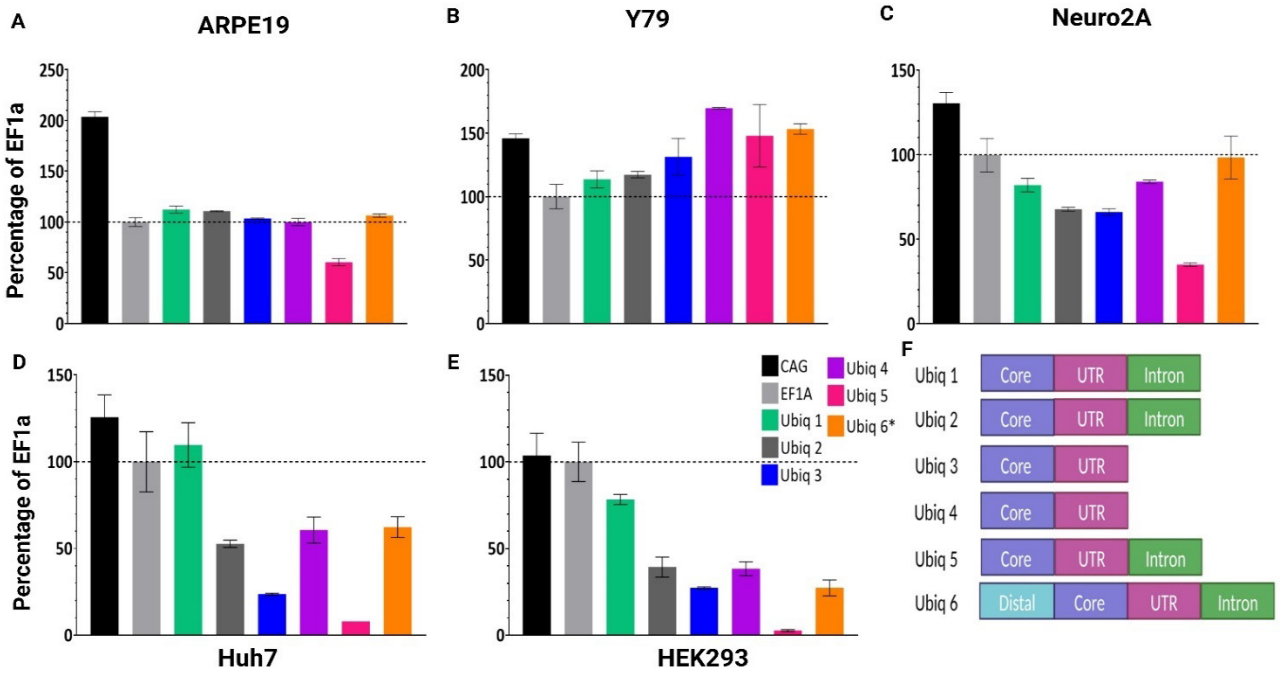


Figure 1 Caption: The dotted line in various cell lines indicates the six ubiquitous promoters benched marked against EF1α. dTomato reporter gene was measured via FACS following transient transfection. **A.** Expression in ARPE19 cells, which are an epithelial cell line. All the ubiquitous promoters except for Ubiq-5 had no statistically significant difference in expression compared to EF1α. EF1α expression was statistically significant compared to Ubiq-5 with a p-value of $p < 0.0001$. **B.** Expression in Y79 cells, which is a photoreceptor-like cell line. Ubiq-3 ($p = 0.0001$), Ubiq-4 ($p < 0.0001$), Ubiq-5 ($p < 0.0001$), and Ubiq-6 ($p < 0.0001$) all demonstrated statistical significance in expression when benchmarked against EF1α indicated by their p values. **C.** Expression in Neuro2A cells, which are a neuroblastoma cell line, showed that when benchmarked against EF1α, there was no significant statistical difference in expression obtained by Ubiq-1 ($p = 0.1256$), Ubiq-4 ($p = 0.2412$), or Ubiq-6 ($p > 0.9999$). However, Ubiq-2 ($p = 0.0001$), Ubiq-3 ($p < 0.0001$), and Ubiq-5 ($p < 0.0001$) had statistically significant less expression when compared to EF1α. **D.** Expression in Huh7 cells, a hepatocyte-like cell line, benchmarked against EF1α. Ubiq-1 ($p = 0.8053$) had no statistically significant expression compared to EF1α. Ubiq-2, Uniq-3,

Ubiq-4, Ubiq-5, and Ubiq-6 all had less expression than EF1α, which was statistically significant with a p-value of $p < 0.0001$ for Ubiq-2 through Ubiq-6. **E.** Expression in HEK293 cells, which are epithelial cell lines, when benchmarked again EF1α all ubiquitous promoters had statistically significant less expression indicated by their p values. Ubiq-1 ($p = 0.0250$) while Ubiq-2 through Ubiq-6 had p values of $p < 0.0001$. **F.** It is a cartoon depiction of the main components that comprise each ubiquitous promoter.

The first attempt at miniaturizing the ubiquitous promoters was by altering them so that any 5'UTR or intron was removed, leaving only the non-coding regulator region that we termed "core" (Figure 2F). All the Ubiq-Core promoters are plotted as a percentage of their parental sequence in the different cell lines. In the ARPE19, Y79, and Neuro2A (Figure 2A, 2B, 2C), most of the Ubiq-core promoters seem to achieve at least 50% expression of their parental sequence except for Ubiq-6-Core in the Neuro2A (Figure 2C). However, in the Huh7 and HEK293 cell lines (Figure 2D, 2E), the removal of the 5'UTR and/or intron appeared to have a dramatic impact on the expression of those Ubiq-Core promoters. The exception was Ubiq-5-Core, which had an expression similar to its parental sequence. A two-way ANOVA utilizing Tukey's multiple comparison tests (Table 1) was performed to compare the expression of the parental sequence to the Ubiq-Core variants when both expression values were normalized to EF1α. This showed that only Ubiq-5-Core in the Neuro2A cell line had greater expression than parental when normalized to EF1α ($p < 0.0001$).

Figure 2: Promoter Potency is Affected by 5' UTRs and/or Introns in a Cell-Type and Sequence Dependent Manner

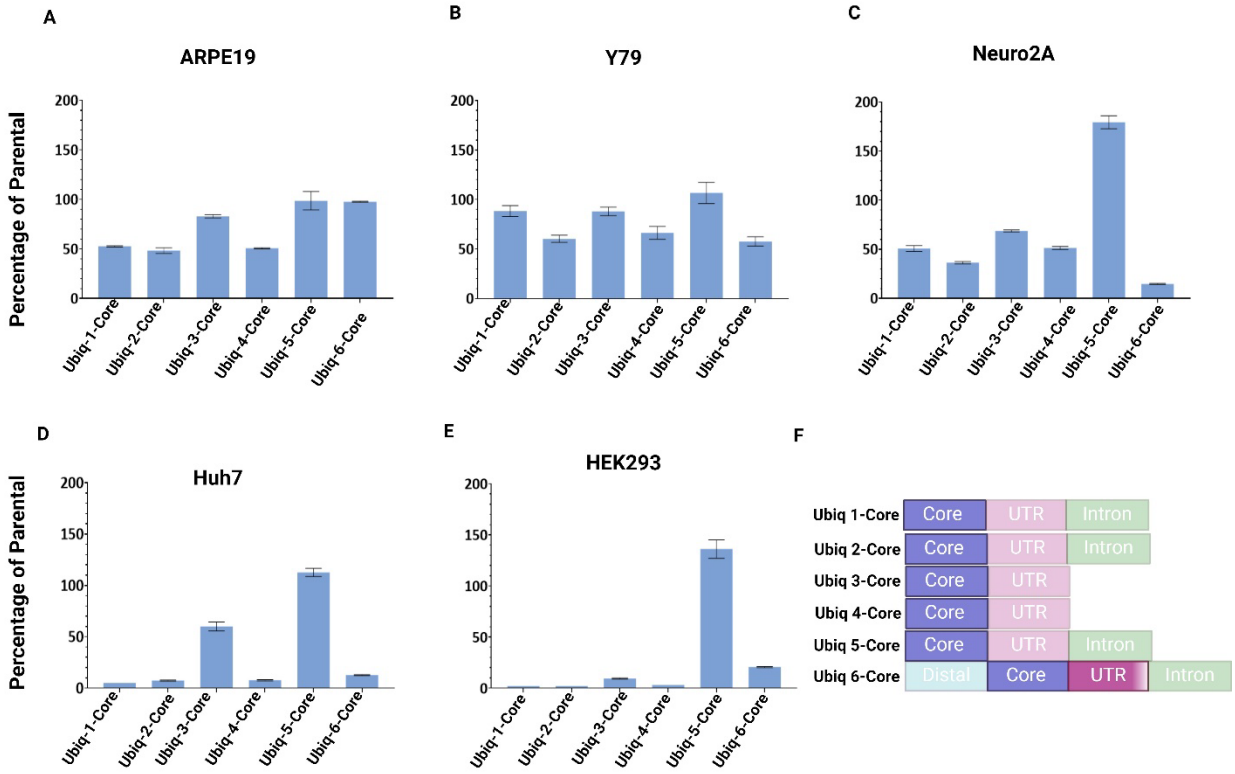


Figure 2 Caption: Shows the non-coding regulatory regions termed "core" of the ubiquitous promoters assayed for functionality and plotted as a percentage of their parental sequence in multiple cell types. **A.** Core promoters' expression in ARPE19 cells where all six core promoters demonstrate at least 50% of their parental sequence, if not more. There is statistically significant less expression in Ubiq-1-Core ($p < 0.0001$), Ubiq-2-Core ($p < 0.0001$), and Ubiq-4-Core ($p < 0.0001$) when compared to their parental sequence, all represented as a % of EF1α. **B.** Core promoters' expression in Y79 cells where all six core promoters demonstrate at least 50% of their parental sequence, if not more. When normalizing Ubiq core promoters to EF1α and comparing to parental sequences also normalized to EF1α, Ubiq-2-Core ($p < 0.0001$), Ubiq-4-Core ($p < 0.0001$), and Ubiq-6-Core ($p < 0.0001$) all had less expression that was statistically significant. **C.** In the Neuro2A expression, all but Ubiq-6-Core have approximately 50% expression compared to

the percentage of their parental. When normalized to EF1α for both Ubiq core promoters and the parental sequences, they all have statistically significant less expression except Ubiq-5-Core ($p < 0.0001$), which has statistically significant greater expression. **D.** In Huh7 cells, only Ubiq-3-Core and Ubiq-5-Core achieved at least 50% of their parental expression. Normalizing both parental and core promoters to EF1α shows statistically significant less expression in Ubiq-1-Core ($p < 0.0001$), Ubiq-2-Core ($p < 0.0001$), Ubiq-4-Core ($p < 0.0001$), and Ubiq-5-Core ($p < 0.0001$). **E.** In HEK293 cells, only Ubiq-5-Core had > 50% expression of parental. Normalizing both parental and core promoters to EF1α shows statistically significant less expression from Ubiq-1-Core ($p < 0.0001$), Ubiq-2-Core ($p < 0.0001$), Ubiq-3-Core ($p = < 0.0006$), Ubiq-4-Core ($p < 0.0001$), and Ubiq-5-Core ($p < 0.0001$). **F.** It is a cartoon depiction demonstrating which regions from the parental promoter were removed to create the core promoters.

Next, efforts focused on the Ubiq-6 promoter for miniaturization. The Ubiq-Core promoter work demonstrated the importance of the intron and its ability to achieve good expression in various cell types. In an attempt at miniaturizing the Ubiq-6 promoter, the native intron region of the Ubiq-6 promoter was swapped with smaller introns, such as IgH intron, VCL intron, and a synthetic intron. The functionality of these newly modified Ubiq-6 promoters was tested in the different cell lines, as shown in Figure 3. The promoters were transiently transfected into the cells and assayed for potency. In ARPE19, Y79, and Neuro2A (Figure 3A, 3B, 3C), potency was determined by reporter gene dTomato and analyzed via FACS. While in the Huh7 and HEK293 cells (Figure 3D, 3E), hFIX transgene was utilized, and expression was measured via an ELISA. All the data was normalized to the Ubiq-6 promoter and represented as a percentage of that promoter in all the cell types. The data shows minimal difference in expression with the different introns added onto the Ubiq-6 promoter in ARPE19, Y79, and Neuro2A cells (Table 2). Only Mod1 had a statistically significant difference in expression compared to Ubiq-6 in Y79 ($p = 0.0462$). In the Huh7 and HEK293, the addition of the Mod1 intron and Mod2 intron failed to achieve at least 50% expression when compared to Ubiq-6; this level of expression is only achieved with its native intron or partially with the synthetic intron. Regardless, all the Mod introns in

Huh7 and HEK293 had a decrease in expression when compared to Ubiq-6, which was statistically significant ($p < 0.0001$), except Mod3 in Huh7 ($p = 0.0002$) and HEK293 ($p = 0.0004$) (Table 2).

Figure 3: Intron Miniaturization Through Swapping Heterologous Sequences Yields Mixed Results

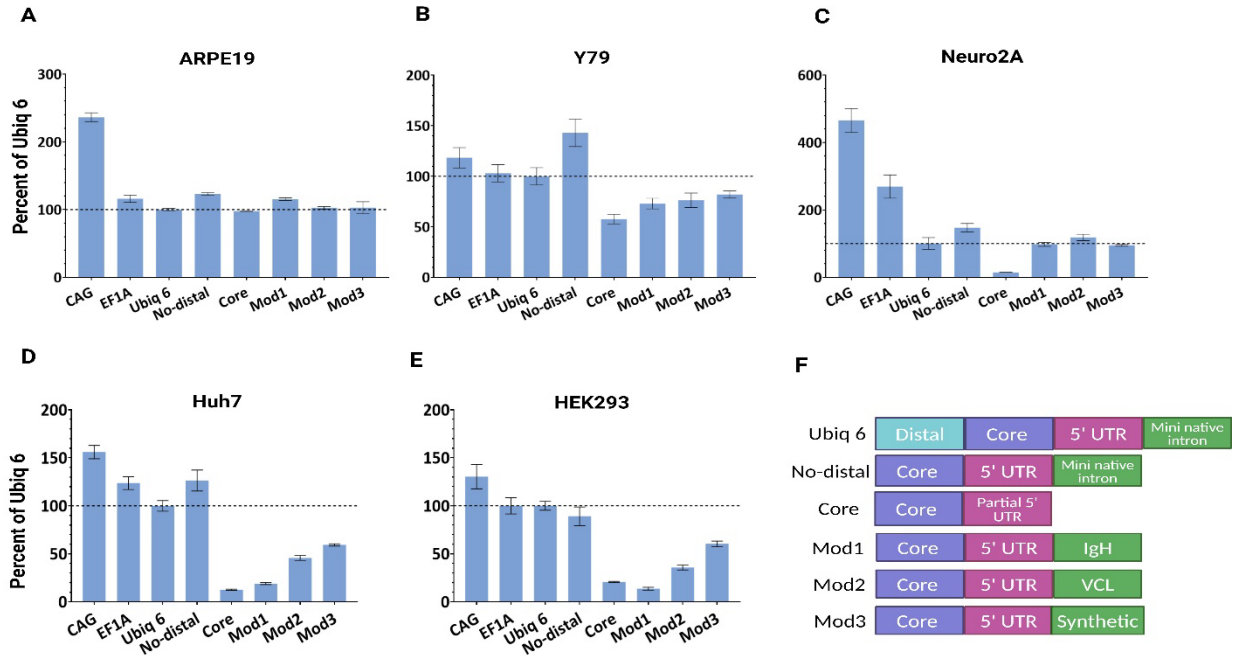


Figure 3 Caption: Demonstrates the effect the intronic region has on the Ubiq-6 promoter and how altering the intron can affect the potency of the promoter in a cell type-dependent manner. All data in each panel is normalized to the Ubiq-6 promoter **A.** Expression in ARPE19 cells where all constructs have expression comparable to Ubiq-6. There is no statistically significant difference in any of the altered Ubiq-6 promoters. **B.** Expression in Y79 cells shows a statistically significant increase in expression for the No-distal ($p < 0.0001$) compared to Ubiq-6. Meanwhile, Core ($p = 0.0001$) and Mod1 ($p = 0.0462$) have a statistically significant decrease in expression compared to Ubiq-6. **C.** Expression in Neuro2A cells shows Mod1, Mod2, and Mod3 have a comparable expression to Ubiq-6. At the same time, No-distal ($p < 0.0001$) has increased expression that is statistically significant, and Core ($p < 0.0001$) has a statistically significant decrease in expression from Ubiq-6. **D.** Huh7 expression shows a non-statistically significant

increase in expression with No-distal from Ubiq-6 and a statistically significant decrease in expression for Core ($p < 0.0001$), Mod1 ($p < 0.0001$), Mod2 ($p < 0.0001$), and Mod3 ($p = 0.0002$). **E.** Expression in HEK293 shows similar expression to Ubiq-6 from No-distal and a statistically significant decrease in expression from Core ($p < 0.0001$), Mod1 ($p < 0.0001$), Mod2 ($p < 0.0001$), and Mod3 ($p = 0.0004$). **F.** It is a cartoon depiction demonstrating the alterations in the intron of Ubiq-6 to generate a different version.

This led us to realize that the native intron might be the best to achieve maximum expression for the Ubiq-6 promoter. To miniaturize the Ubiq-6 promoter further, a tiled ablation deletion series was created (Figure 4A). This ablation series spanned the whole length of the Ubiq-6 promoter, with each ablation covering 30bp and a 10bp overlap with subsequent ablations. All 34 ablations were individually transiently transfected into HEK293 cells with hFIX transgene, and expression was measured via an ELISA and normalized to the Ubiq-6 promoter, with each ablation represented as a percentage of Ubiq-6. The data showed key areas that could be targeted for removal to miniaturize the promoter without sacrificing potency. This data was utilized to engineer nine miniaturized variants of the Ubiq-6 promoter. These combinatorial or composite ablations span approximately 700bp to 300bp in length (Figure 4B). They were assayed with dTomato measured via FACS for ARPE19, Y79, and Neuro2A cell lines. The hFIX transgene was used for Huh7 and HEK293 cell lines, and activity was determined via an ELISA. All the data was normalized to the Ubiq-6 promoter and presented as a heatmap (Figure 4B). As the Ubiq-6 promoter got smaller in size, it appeared that they were becoming less ubiquitous and slightly more tissue-specific, especially in the APRE19 and Y79 cell lines (Table 3).

Figure 4: Miniaturization Efforts for Ubiq-6 Promoter

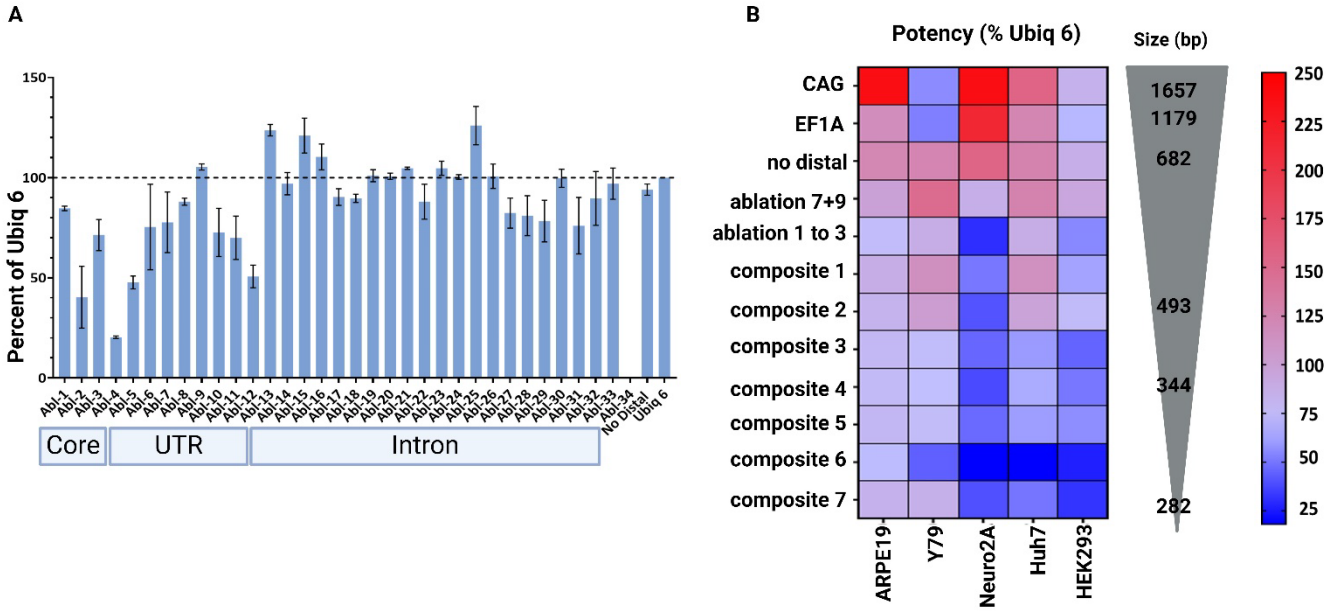


Figure 4 Caption: A. Ablation series for Ubiq-6 promoter where key regions are identified via a deletion series spanning the entire length of the promoter sequence. HEK293 cells were transiently transfected with the ablations series, expression was measured via FIX ELISA, and data was normalized to Ubiq-6 as a percentage. **B.** Data from the ablation series used to engineer miniaturized variants of the Ubiq-6 promoter spanning 682bp to 282bp in size is represented as a heatmap normalized to Ubiq-6 and represented as a percentage of it. The combinatorial or composite ablations were assayed utilizing the reporter gene dTomato for ARPE19, Y79, and Neuro2A and FIX for Huh7 and HEK293.

Engineering the Ubiq-6 promoter through miniaturization made them slightly more cell-specific, possibly due to the lower abundance of transcription factor binding motifs (TFBM) in the miniaturized promoters. This informed the process of swapping out two TFBM in the other Ubiq promoters with ocular-specific TFBM. We then examined if these other promoters resulted in increased expression in ocular cell lines ARPE19 and Y79 when compared to their wildtype versions. The new variants constructed from the five Ubiq promoters and two composite Ubiq-6 promoters all utilized dTomato as the reporter

gene. They were transiently transfected into ARPE19 and Y79 and analyzed via FACS cells. The mean fluorescent intensity of the 10 variants and wildtype were normalized to the CAG promoter and plotted as a scatter plot. The data is plotted as a matrix with Y79 data represented on the y-axis and ARPE19 data on the x-axis. In Ubiq-1 for Y79 cells, all but the Ubiq-1-Var10 (Figure 1A) had an expression that was either greater to or similar to wildtype. Meanwhile, in ARPE19 cells, all variants had an expression similar to, if not greater than, wildtype (Figure 5A). The negative values in Table 4 indicate that the mean difference in expression was greater than the referent group. Two-way ANOVA statistical test with multiple comparisons utilizing Dunnett's test revealed that in ARPE19, Var1 and Var5 promoters statistically differed, while in Y79, Var5 did not differ (Table 4). In Ubiq-2, no variants demonstrated expression similar to the wildtype in Y79 cells. Still, in ARPE19 Cells, Ubiq-2-Var10 (Figure 5B) had stronger expression than wildtype, which was statistically significant ($p = 0.0005$) (Table 5). When looking at Ubiq-3 (Figure 5C), only Var5 had an expression stronger than wildtype in Y79 cells, while there were none when tested in the ARPE19 cells (Table 6). Utilizing Ubiq-4 (Figure 5D) in the ARPE19 cells, all variants demonstrated similar expression, if not stronger than wildtype. However, in the Y79, all variants demonstrated statistically significant greater expression than wildtype, with the exception being Ubiq-4-Var8, which had a statistically significant decrease in expression ($p<.0001$) (Table 7). For Ubiq-6 (Figure 5E), only Var2 and Var9 showed statistically significant increases in expression in ARPE19 cells with $p <0.0001$ (Table 8). In the Y79, only Var1 had a statistically significant increase in expression (Table 8), while Var9 demonstrated similar levels of expression compared to the wildtype. For Comp-6 (Figure 5F), Var3 and Var8 had increased expression to the wildtype in ARPE19 cells, although it was not statistically significant (Table 9). In the Y79 cells, Var2, Var4, and Var8 demonstrated an increase in expression compared to the wildtype, but it was not statistically significant (Table 9). For the Comp-7 (Figure 5G) in the ARPE19 cells, Var2 demonstrated expression that slightly increased compared to the wildtype but was not statistically significant (Table 10). In the Y79, all variants except Var5 and Var10 had expressions similar to or greater than the wildtype.

Figure 5: Motif Swapped Promoters Individually Transfected in Ocular Specific Cell Type

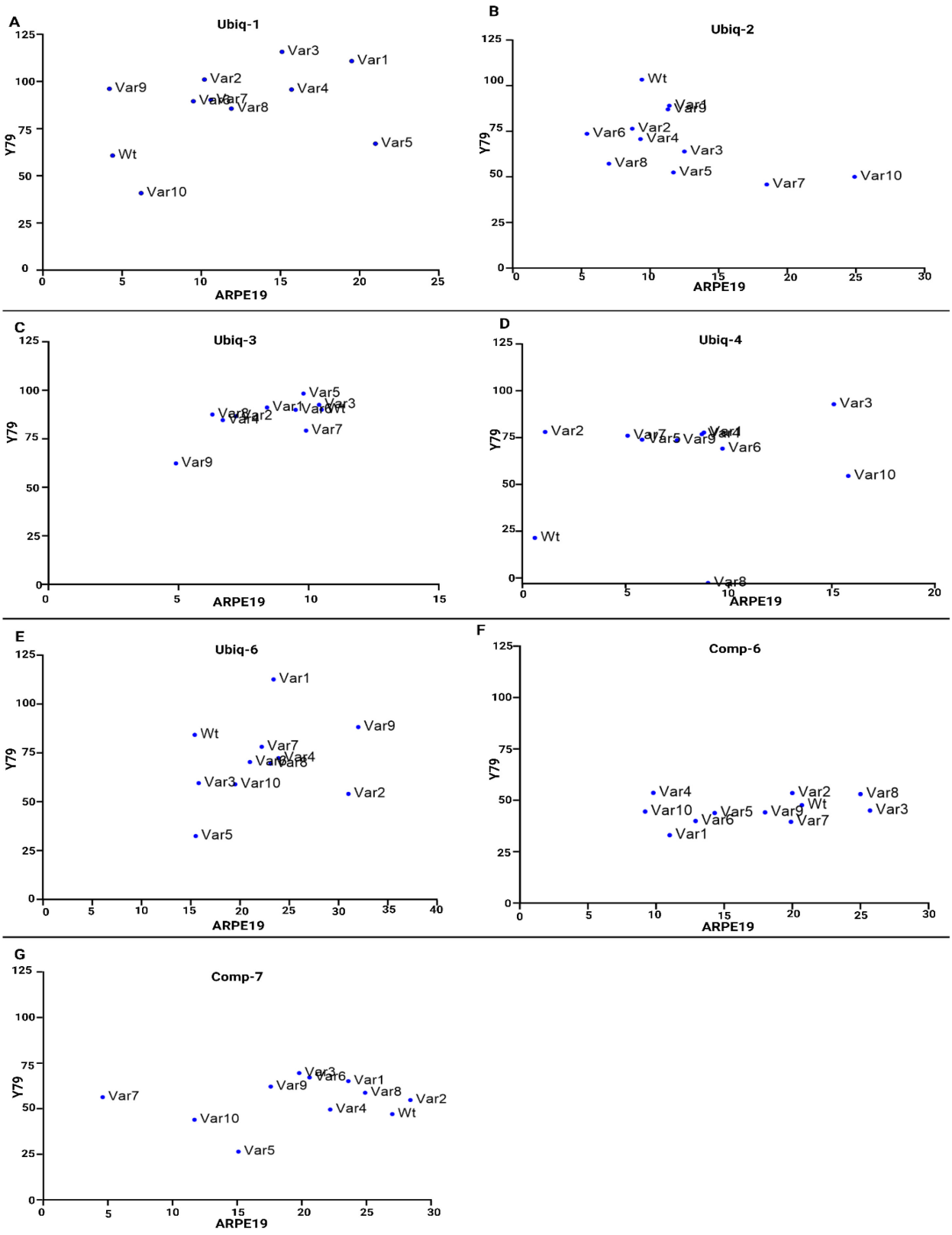


Figure 5 Caption: Shows Ubiquitous promoters Motif Swap variants transiently transfected in Y79 cells (y-axis) and ARPE19 cells (x-axis) and normalized to CAG and represented as a percentage of it. **A.** Ubiq-1 variants in ARPE19 show a statistically significant increase in expression compared to wildtype only in Var1 ($p = 0.016$) and Var5 ($p = 0.0066$). In Y79, Var10 showed a nonstatistical significant decrease in expression compared to wildtype. At the same time, all the variants expect Var5 ($p = 0.6648$) to have a statistically significant increase in expression indicated via a p-value of $p < 0.0001$. **B.** In Ubiq-2 variants, ARPE19 Var10 ($p = 0.0005$) has a statistically significant increase in expression over wildtype. In the Y79, all variants showed a statistically significant decrease in expression compared to wildtypes. **C.** Ubiq3 in ARPE19 cells has no variants with statistical significance in either direction compared to wildtype; Var10 data in ARPE19 is missing. In Y79 cells, only Var5 ($p = 0.0129$) has a statistically significant increase in expression when compared to wildtype, while Var7 ($p = 0.0007$), Var9 ($p < 0.0001$), and Var10 (not plotted) ($p < 0.0001$) all had a statistically significant decrease in expression. **D.** In Ubiq-4 for ARPE19 cells, Var3 ($p = 0.0006$), Var6 ($p = 0.0433$), and Var10 ($p = 0.0003$) all have an increase in expression compared to wildtype that is statistically significant. In Y79, Var8 ($p < 0.0001$) decreases expression compared to wildtype, which is statistically significant, while all the rest of the variants have a statistically significant increase in expression represented by p-value $p < 0.0001$. **E.** Ubiq-6 in ARPE19 cells Var2 ($p < 0.0001$) and Var9 ($p < 0.0001$) both have an increase in expression compared to wildtype, which is statistically significant. In Y79 cells, only Var1 ($p < 0.0001$) has a statistically significant increase in expression. **F.** Comp-6 for ARPE19 cells, no variants show an increase in expression compared to wildtype, which is statistically significant. In Y79, no variants show an increase in expression compared to wildtype, that are statistically significant. **G.** Comp-7 for ARPE19 cells there are no variants that have an increase in expression compared to wildtype that are statistically significant. In Y79, the following variants all have an increase in expression that is statistically

significant: Var1 ($p < 0.0001$), Var3 ($p < 0.0001$), Var6 ($p < 0.0001$), Var7 ($p = 0.0196$), Var8 ($p = 0.0015$), and Var9 ($p < 0.0001$).

The 70 different variants of the motif-swapped promoters were pooled together to create a plasmid library and were tested *ex vivo* in mouse retina explants for expression. Figure 6 shows that the dTomato reporter gene was detected in both group 1 and group 2, showing successful exploration with the pooled library with 77 components. Fluorescence were also detected in group 3, which contained only the wildtype versions of the promoters.

Figure 6: dTomato Expression in Retinal Explants Electroporated with Motif Swap Pooled Plasmid Library

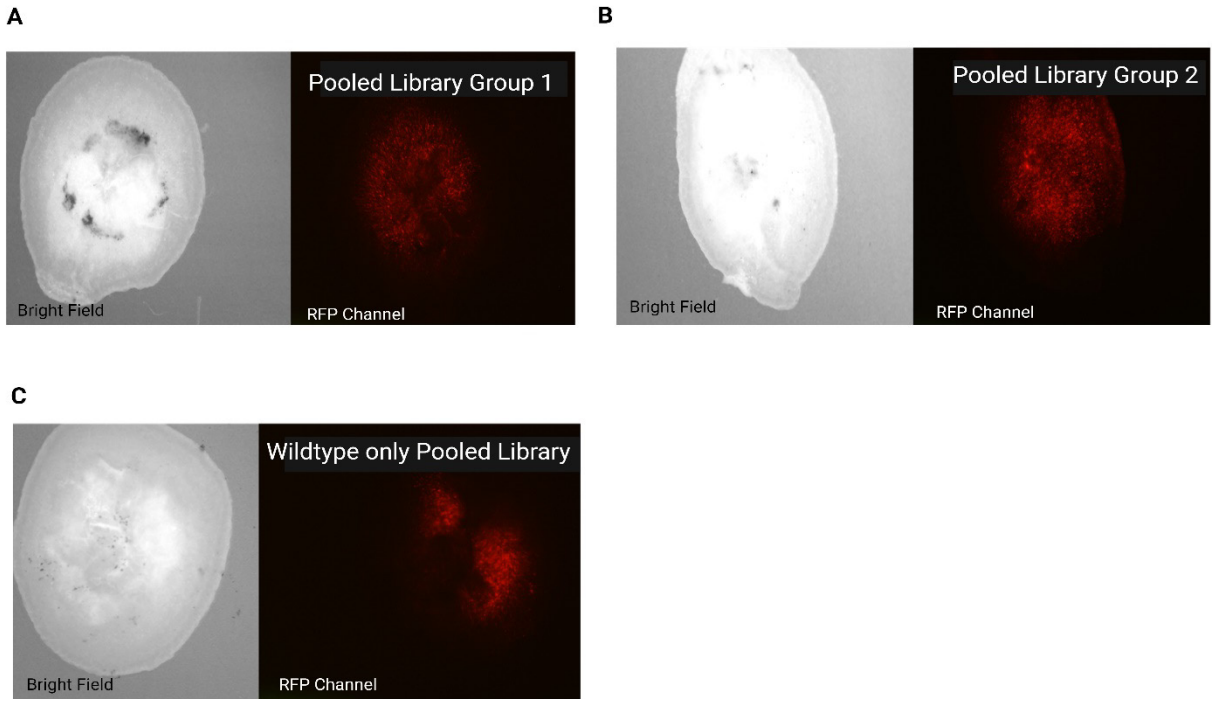


Figure 6 Caption: Images of retinal explants electroplated and viewed under Leica Fluorescent Dissecting Scope with both bright field of the retina and red fluorescent protein (RFP) channel. **A.** Motif swap library resuspended in PBS and electroporated at 1µg/µL. **B.** Motif swap library resuspended in HBBS and electroporated at 1µg/µL. **C.** A smaller pool of 7 promoters comprised only the

wildtypes from the motif swap library resuspended in PBS and electroporated at 1µg/µL.

2.4 Discussion

Gene therapy is a promising avenue for the treatment of ocular-related diseases that are due to genetic abnormalities. In 2017 the FDA approved the first AAV-based gene therapy, Luxturna, for the treatment of biallelic RPE65 mutation-associated retinal dystrophy.[203] There are over 40 active gene therapy clinical trials ongoing that use AAV vectors for the treatment of ocular-related diseases.[203, 204] The need to identify small promoters and regulatory elements is crucial, since in many cases genes do not fit in AAVs due to the limited packing capacity of 4.7kb.[23] An alternative method utilized for delivering therapies with oversize genes is trans-splicing where the gene is separated and packaged into different AAVs and administered together as a dual AAV vector.[77, 205, 206] In this approach the transgene is reconstituted in cells infected by the dual AAV vector through viral genome recombination and subsequent splicing or by homologous recombination of overlapping regions.[205] Despite increasing the delivery capacity of an AAV based gene therapy these dual AAV vectors often result in lower efficiency of reconstitution of the full length genome which could require a higher dose of vector thus increasing the need to identify small promoters and regulatory elements.[76, 207]

This study identified potential promoter candidates to be miniaturized while maintaining potency for use in AAVs for ocular-related diseases. This study utilized the use of existing ubiquitous promoters compared to known core/minimal promoters or fully synthetic promoters for miniaturization since minimal promoters could require external enhancers increasing the total size of the cassette and fully synthetic promoters often require extensive optimization efforts along with increased concerns with off target effects. The six ubiquitous promoters that were screened had a similar expression to that of the control promoter EF1α in certain cell types, specifically ARPE19, Y79, and Neuro2A. However, there was variation in the expression of the ubiquitous promoters in Huh7 and HEK293 cells when normalized to EF1α. A hindrance in expression was observed in most cell lines

tested with the Ubiq-5 promoter. A sharp decline in expression was seen in the Huh7 and HEK293 cell lines, but the opposite was observed in the Y79 cells, which showed expression greater than that of the EF1α promoter. This observation with the Ubiq-5 promoter could indicate that it is not as strong of a ubiquitous promoter or does not function well in tissues with epithelial-like properties, such as the Huh7, HEK293, and ARPE19 cell lines. Despite the findings with Ubiq-5, the rest of the ubiquitous promoters seemed to demonstrate a broad range of expression across the different cell types. This phenomenon of ubiquitous promoters having varied expression profiles in different cell types is not unique to our study. Other investigators have reported that some commonly used ubiquitous promoters, like CMV promoter, have significant variability across cell types, while EF1α tends to be more consistent.[120] The range of expression helps justify these promoters as ubiquitous and potentially viable candidates for miniaturization efforts to generate smaller promoters without sacrificing potency.

We utilized an iterative approach to advance previous research focused on miniaturization. In our first attempt at promoter miniaturization with the ubiquitous promoters, we removed any region of the native promoter other than the non-coding regulatory region, which we termed core. Removing the 5’ untranslated region (UTR) and any intron associated with the native promoter, leaving just the core region, significantly impacted the potency of the Ubiq-core promoters, depending on the cell type. We saw a hindrance in expression in the ARPE19, Y79, and Neuro2A cell lines indicated by achieving approximately 50% of their parental. Still, the expression was completely abolished in the Huh7 and HEK293 cells by removing the 5’UTR and intron regions. The removal of 5’UTR and intron is something that can alter the expression of the promoter. The intron associated with a promoter can contain an abundance of different TFBM, which are essential for driving transcription. They can also harbor other regulatory elements, such as enhancers.[208-210] The 5’UTR often contains sequences that attract and bind ribosomes, which are needed for translation to occur, and removing such regions could have an impact on the potency of the promoter.[208]

Our targeted approach to swapping native introns of promoters advanced previous research and potentially enables future researchers to target specific DNA sequences that do not serve functionality in the promoter to make the promoters more compact. Focusing on the Ubiq-6 promoter, we looked into swapping the native intron with heterologous sequences smaller than the current native intron to miniaturize the promoter. We saw that the addition of the different introns, although functional in their natural state, did not add to the expression of the Ubiq-6 promoter, depending on the cell line. This finding led us to believe that the best intron for the Ubiq-6 promoter was the native intron. We still needed to find an optimal way to try and miniaturize the promoter. A deletion ablation series is a common technique to identify key regions within a promoter that, when removed, do not affect their overall potency. Others in the literature have used this technique, often pairing the ablations with reporter genes like luciferase or a fluorescent reporter to assess the potency of the modified promoter in desired cell types.[127, 211, 212]

Our data demonstrates that miniaturization of ubiquitous promoters is feasible, as seen with Ubiq-6. However, there could be implications associated with miniaturization efforts, such as a loss in potency and a difference in tissue specificity. The increase in tissue specificity and reduced promoter size can be advantageous, especially for therapies with an oversized transgene that require targeting of specific tissues. The literature has reported that miniaturized versions of promoters tend to demonstrate different tissue specificity from nontruncated versions. This phenomenon was noted in the miniaturization efforts with the latency-associated promoter 2 (LAP2).[212] Other researchers have shown a loss of approximately 70% of promoter function in the spinal cord with the truncated LAP2 promoter, with expression primarily observed only in brain neurons.[212] Investigators also noted a decrease in off-target tissue expression with truncated LAP2 promoter compared to the wildtype version.[212] These findings validate the phenomenon seen with our truncated versions of the Ubiq-6 promoter, where tissue specificity increased as the size of the Ubiq-6 promoter decreased, as seen with Composite 6 and 7.

Our efforts to generate promoters with increased affinity for ocular cell types through swapping TFBM with ocular-specific motifs did generate some variants with increased expression compared to the wildtype with the different ubiquitous promoters tested. Our *in vitro* testing in ARPE19 and Y79 cell lines indicates that through swapping, TFBM promoters can be engineered to have increased activity in specific tissues, which can be advantageous for ocular-related gene therapies. The literature also indicates how altering TFBM within a promoter can dramatically affect where expression occurs and the potency of the promoter.[30, 127, 208, 212] A study by Zimmerman et al. demonstrated how adding the CTCF and NF-Y motifs can have drastically different outcomes depending on the promoter they were inserted into and the cell line they were tested in. For example, when CTCF and NF-Y motifs were added to the EF1α promoter and CMV promoter, an increase in expression was observed only in the EF1α promoter.[213] They also observed that the expression from the EF1α promoter with CTCF and NF-Y motifs depended on which cells they were tested in.[213] They reported that in the MCF-7 cell line, they saw an increase in expression, while the Jurkat line of leukemia T cells did not increase and had expression more similar to wildtype EF1α promoter.[213] A major limitation of our study is that the work presented is conducted solely as *in vitro* experiments. A challenge with *in vitro* experiments that use transient transfections is normalization between experiments due to factors such as transfection efficiency, cell viability, and plasmid copy number. To address this issue a common technique is to use a dual-reporter system like the GLUC-on system which helps to improve reproducibility. The transient transfections of the promoters provided valuable insights into the activity of the promoters across a variety of cell types; however, *in vitro* data does not always correlate well with data generated *in vivo*. This is especially true when developing promoters specific for ocular gene therapies since there are not good *in vitro* representative models to test.[214] Even *in vivo* between different animal models are not representative of each other. Future work is needed to explore the expression of the motif swap promoter library in ocular tissues, preferably in the context of an AAV-delivered *in vivo* study.

In conclusion, this study provided novel insights into the practicality of engineering ubiquitous promoters to become smaller without substantial sacrifices to their potency. It

also demonstrated the feasibility of swapping TFBM with tissue-specific motifs to increase expression in ocular cell types. This work also demonstrated the feasibility of swapping TFBM with tissue-specific motifs to increase expression in ocular cell types. In the future changing the number or the order of the ocular-specific TFBM could cause a positive or negative effect but either way it would further optimize the promoter properties by influencing the expression strength, onset and duration. This work adds to the effort of creating promoters that can be utilized in gene therapies where more compact and tissue-specific promoters are needed.

2.5 Materials and Methods

Construction of Ubiquitous Promoter Candidates

Promoters were designed *in silico* using bioinformatics software Geneious Prime (Dotmatics, Auckland, New Zealand), and DNA synthesis was cloned into a pUC57-Kan plasmid flanked by NotI and MluI restriction sites and performed at GenScript (Piscataway, NJ USA). Constructs were then double digested with enzymes NotI and MluI (New England Biolabs, Cat: R3189L & R3198L). Bands were excised from 0.8% agarose gel, made with 200mL of 1X UltraPure TAE buffer (Invitrogen, Cat: 1558026) and 1.6g of UltraPure Agarose (Invitrogen, Cat: 16500500), with SYBR Safe (Invitrogen, Cat: S33102). Excised bands were purified with the QIAquick Gel Extraction Kit (Qiagen, Cat: 28706). Promoter DNA was then cloned into either AAV backbone expression dTomato or hFIX as the reporter gene utilizing the Quick Ligation Kit (New England Biolabs, Cat: M2200L) and followed the protocol for restriction digest cloning. The ligation mixture was then transformed into One Shot™ Stbl3™ Chemically Competent E. coli (Invitrogen, Cat: C737303) following the protocol supplied. Colonies were picked and grown overnight in 5mL cultures of LB broth with kanamycin in a shaking incubator at 37ºC. DNA was isolated the following day using the QIAprep Spin Miniprep Kit (Qiagen, Cat: 27106) as outlined in the respective protocol.

Generation of motif swap promoter library

TFBM for the different ubiquitous promoters were mapped and annotated from a publicly available ENCODE (Stanford University, CA USA) data set. In Excel, a randomization generator was created where two motifs from each promoter were selected to be replaced by a different ocular-specific TFBM with degenerative sequencing of the motifs (Tables 11-17). Each ubiquitous parent sequence had 10 variants created with unique TFBM swapped. At GenScript (Piscataway, NJ USA), they cloned each variant promoter into a pAAV backbone plasmid containing dTomato as the reporter gene and with a unique 5' barcode. GenScript shipped each promoter along with a pooled plasmid library.

Cell Culture and Transfection

ARPE19, Neuro2A, Huh7, and HEK293 cells were grown in DMEM supplemented with 10% FBS, while Y79 were grown in RPMI 1640 supplemented with 20% FBS, all in an incubator at 37ºC with 5% CO2 atmosphere. All cells were seeded into a 24-well plate (Corning, Cat: 3527) at 1.00E+5 cells/well. Neuo2A, Huh7, and HEK293 cells were transiently transfected with 500ng of total DNA utilizing transfection reagents jetOptimus (Polyplus, Cat: 117-07) following manufacturer protocol. The ARPE19 and Y79 cells were transiently transfected with 500ng of total DNA utilizing transfection reagent Lipofectamine LTX Reagent with PLUS Reagent (Invitrogen, Cat: 15338100) following manufacturer protocol.

Retinal Explant and Electroporation

All animal procedures were performed in accordance with protocol IACUC007: 2020-1226: New Viral and Non-viral Gene Delivery Systems for the Eye approved by Institutional Animal Care and Use Committee of Spark Therapeutics, Inc. Explants of retina intact with lens were harvested from C57BL/6J neonatal mice at age P2 and placed into 2.0mL round bottom tube filled with Neurobasal-A Medium and sent to research lab on wet ice 2-6 hours after collection. Before explants arrive prepare enough 6-well plates 2mL of Neurobasal-A Medium supplemented with Pen/Strep, B-27, and N-2 and place membrane filter to each well and place in an incubator at 37ºC with a 5% CO2 atmosphere for at least 1 hour. Next, turn on NEPA21 Electro-Kinetic Transfection Systems (Bulldog Bio, Portsmouth, NH) and set voltage and time settings. Transfer retina explants once

they arrive from their tube into a petri dish with 1X PBS using a modified transfer pipette. Check for intact retina and discard any damaged ones. Add 30µL of the electroporation solution, plasmid DNA in 1xPBS at 1µg/µL concentration, to the 6-well electrode chamber. Next, transfer the retina explants into 6-well electrode chamber with a max of 2 retinas per chamber well. Align retina explants so the lens is facing the positive electrode. Then run electroporation protocol of 5 pluses, 30 V amplitude, 50 ms pulse width, and 950 ms between pulses. Next, transfer the electroplated retinas into multi well plate with explant media using modified transfer pipette and with fine forceps detach and remove the lens from the retina. Then transfer reach retina explants on top of the filter in the earlier prepared 6-well plates with supplemented media, and 1-2 drops of supplemented media directly on top of retinas and place into incubator at 37ºC with a 5% CO2 atmosphere. Change the media on post-natal date (PND) P3 by adding 2mL of fresh supplemented media with 1 drop on top of retinas and on PND P4 with 3mL of fresh supplemented media with 2 drops on top of retinas. On PND P7 image retinas with a Leica Fluorescent Dissecting Scope like Leica M205 FCA (Leica Microsystems Inc., Deerfield, IL). After images are taken retina explants can be processed for RNA extraction utilizing RNeasy Plus Micro Kit (Qiagen, Cat: 74034) following the manufacturer's protocol.

Quantification of Promoter Activity

For the analysis of promoters driving reporter gene dTomato, cells were collected 48hrs after transfection and resuspended in Fluidigm Corp MAXPAR CELL STAINNG BFFR 500ML (Fluidigm Corp, Cat: 201068), and FACS was performed on the cells. Data was analyzed via FlowJo Software (FlowJo, Ashland, OR USA) and graphed via Prism (GraphPad, San Jose, CA).

For the analysis of promoters driving reporter gene, hFIX supernatant was collected 48hrs after transfection, and ELISA was performed to detect hFIX antigen using Factor IX Paired Antibody Set (Affinity Biological, Cat: FIX-EIA) and following an in-house protocol. The ELISA was read via BioTek Synergy H1 Multimode Reader (Agilent, Santa Clara, CA USA) and analyzed via SoftMax Pro Software (Molecular Devices, LLC. San Jose, CA USA). Data was graphed via Prism (GraphPad, San Jose, CA, USA).

Statistical Analysis

All statistical analysis was conducted in Prism (GraphPad, San Jose, CA). ANOVAs were conducted to examine differences in expression across promoters within and between cell types. Post hoc multiple comparisons were performed using Tukey's and Dunnett's tests.

2.6 Appendix

Table 1: Two-way ANOVA with Tukey's Multiple Comparison Test

Two-way ANOVA with Tukey's Multiple Comparison Test					
ARPE19	**Tukey's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 1 vs. Ubiq1-Core	52.67	34.58 to 70.75	****	<0.0001
	Ubiq 2 vs. Ubiq2-Core	57	38.91 to 75.09	****	<0.0001
	Ubiq 3 vs. Ubiq3-Core	17.67	-0.4210 to 35.75	ns	0.0629
	Ubiq 4 vs. Ubiq4-Core	49.33	31.25 to 67.42	****	<0.0001
	Ubiq 5 vs. Ubiq5-Core	1	-17.09 to 19.09	ns	>0.9999
	Ubiq 6 vs. Ubiq6-Core	1.667	-16.42 to 19.75	ns	>0.9999
Y79	**Tukey's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 1 vs. Ubiq1-Core	13.33	-4.754 to 31.42	ns	0.4047
	Ubiq 2 vs. Ubiq2-Core	46.67	28.58 to 64.75	****	<0.0001
	Ubiq 3 vs. Ubiq3-Core	16	-2.088 to 34.09	ns	0.1439
	Ubiq 4 vs. Ubiq4-Core	56.67	38.58 to 74.75	****	<0.0001
	Ubiq 5 vs. Ubiq5-Core	-9.667	-27.75 to 8.421	ns	0.8613
	Ubiq 6 vs. Ubiq6-Core	40.67	22.58 to 58.75	****	<0.0001
Neuro2A	**Tukey's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 1 vs. Ubiq1-Core	40.33	22.25 to 58.42	****	<0.0001
	Ubiq 2 vs. Ubiq2-Core	43	24.91 to 61.09	****	<0.0001
	Ubiq 3 vs. Ubiq3-Core	20.67	2.579 to 38.75	*	0.0106
	Ubiq 4 vs. Ubiq4-Core	41	22.91 to 59.09	****	<0.0001
	Ubiq 5 vs. Ubiq5-Core	-28.33	-46.42 to -10.25	****	<0.0001
	Ubiq 6 vs. Ubiq6-Core	31.33	13.25 to 49.42	****	<0.0001
Huh7	**Tukey's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 1 vs. Ubiq1-Core	104.3	86.25 to 122.4	****	<0.0001
	Ubiq 2 vs. Ubiq2-Core	48.67	30.58 to 66.75	****	<0.0001
	Ubiq 3 vs. Ubiq3-Core	9.333	-8.754 to 27.42	ns	0.8891
	Ubiq 4 vs. Ubiq4-Core	55.67	37.58 to 73.75	****	<0.0001
	Ubiq 5 vs. Ubiq5-Core	-1	-19.09 to 17.09	ns	>0.9999
	Ubiq 6 vs. Ubiq6-Core	70.67	52.58 to 88.75	****	<0.0001
HEK293	**Tukey's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 1 vs. Ubiq1-Core	76.33	58.25 to 94.42	****	<0.0001
	Ubiq 2 vs. Ubiq2-Core	38.33	20.25 to 56.42	****	<0.0001
	Ubiq 3 vs. Ubiq3-Core	24.67	6.579 to 42.75	***	0.0006
	Ubiq 4 vs. Ubiq4-Core	37.33	19.25 to 55.42	****	<0.0001

	Ubiq 5 vs. Ubiq5-Core	-1	-19.09 to 17.09	ns	>0.9999
	Ubiq 6 vs. Ubiq6-Core	79	60.91 to 97.09	****	<0.0001

Table 2: Two-way ANOVA with Tukey's Multiple Comparison Test

Two-way ANOVA with Tukey's Multiple Comparison Test					
ARPE19	**Tukey's multiple comparisons test**	**Mean Diff**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. No-distal	-23	-49.75 to 3.745	ns	0.145
	Ubiq 6 vs. Core	2.333	-24.41 to 29.08	ns	>0.9999
	Ubiq 6 vs. Mod1	-15.33	-42.08 to 11.41	ns	0.6318
	Ubiq 6 vs. Mod2	-2.5	-32.40 to 27.40	ns	>0.9999
	Ubiq 6 vs. Mod3	-3	-29.75 to 23.75	ns	>0.9999
Y79	**Tukey's multiple comparisons test**	**Mean Diff**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. No-distal	-43	-69.75 to -16.25	****	<0.0001
	Ubiq 6 vs. Core	42.33	15.59 to 69.08	***	0.0001
	Ubiq 6 vs. Mod1	27	0.2549 to 53.75	*	0.0462
	Ubiq 6 vs. Mod2	23.67	-3.078 to 50.41	ns	0.1217
	Ubiq 6 vs. Mod3	18	-8.745 to 44.75	ns	0.4268
Neuro2A	**Tukey's multiple comparisons test**	**Mean Diff**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. No-distal	-47	-73.75 to -20.25	****	<0.0001
	Ubiq 6 vs. Core	85.67	58.92 to 112.4	****	<0.0001
	Ubiq 6 vs. Mod1	2.333	-24.41 to 29.08	ns	>0.9999
	Ubiq 6 vs. Mod2	-18	-44.75 to 8.745	ns	0.4268
	Ubiq 6 vs. Mod3	5.333	-21.41 to 32.08	ns	0.9985
Huh7	**Tukey's multiple comparisons test**	**Mean Diff**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. No-distal	-26.33	-53.08 to 0.4118	ns	0.0567
	Ubiq 6 vs. Core	87.33	60.59 to 114.1	****	<0.0001
	Ubiq 6 vs. Mod1	81	54.25 to 107.7	****	<0.0001
	Ubiq 6 vs. Mod2	54.33	27.59 to 81.08	****	<0.0001
	Ubiq 6 vs. Mod3	40.67	13.92 to 67.41	***	0.0002
HEK293	**Tukey's multiple comparisons test**	**Mean Diff**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. No-distal	11	-15.75 to 37.75	ns	0.9033
	Ubiq 6 vs. Core	79.33	52.59 to 106.1	****	<0.0001

	Ubiq 6 vs. Mod1	86.33	59.59 to 113.1	****	<0.0001
	Ubiq 6 vs. Mod2	64.33	37.59 to 91.08	****	<0.0001
	Ubiq 6 vs. Mod3	39.67	12.92 to 66.41	***	0.0004

Table 3: Two-way ANOVA with Dunnett's Multiple Comparison Test

Two-way ANOVA with Dunnett's Multiple Comparison Test					
	Dunnett's multiple comparisons test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
ARPE19	Ubiq 6 vs. CAG	-136.3	-151.4 to -121.2	****	<0.0001
	Ubiq 6 vs. EF1α	-16	-31.12 to -0.8843	*	0.032
	Ubiq 6 vs. No distal	-23	-38.12 to -7.884	***	0.0004
	Ubiq 6 vs. Ablations 7+9	1.667	-13.45 to 16.78	ns	>0.9999
	Ubiq 6 vs. Ablations 1 to 3	27	11.88 to 42.12	****	<0.0001
	Ubiq 6 vs. Composite 1	12.67	-2.449 to 27.78	ns	0.1498
	Ubiq 6 vs. Composite 2	18.33	3.218 to 33.45	**	0.0088
	Ubiq 6 vs. Composite 3	22.33	7.218 to 37.45	***	0.0007
	Ubiq 6 vs. Composite 4	25	9.884 to 40.12	****	<0.0001
	Ubiq 6 vs. Composite 5	22	6.884 to 37.12	***	0.0008
	Ubiq 6 vs. Composite 6	31.33	16.22 to 46.45	****	<0.0001
	Ubiq 6 vs. Composite 7	17	1.884 to 32.12	*	0.0188
Y79	**Dunnett's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. CAG	44.33	29.22 to 59.45	****	<0.0001
	Ubiq 6 vs. EF1α	47.33	32.22 to 62.45	****	<0.0001
	Ubiq 6 vs. No distal	-25	-40.12 to -9.884	****	<0.0001
	Ubiq 6 vs. Ablations 7+9	-48.67	-63.78 to -33.55	****	<0.0001
	Ubiq 6 vs. Ablations 1 to 3	13	-2.116 to 28.12	ns	0.1306
	Ubiq 6 vs. Composite 1	-15	-30.12 to 0.1157	ns	0.0529
	Ubiq 6 vs. Composite 2	-2.333	-17.45 to 12.78	ns	>0.9999
	Ubiq 6 vs. Composite 3	27.33	12.22 to 42.45	****	<0.0001
	Ubiq 6 vs. Composite 4	29.33	14.22 to 44.45	****	<0.0001
	Ubiq 6 vs. Composite 5	26.33	11.22 to 41.45	****	<0.0001
	Ubiq 6 vs. Composite 6	56	40.88 to 71.12	****	<0.0001
	Ubiq 6 vs. Composite 7	15.67	0.5510 to 30.78	*	0.038
Neuro2A	**Dunnett's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. CAG	-136	-151.1 to -120.9	****	<0.0001
	Ubiq 6 vs. EF1α	-112.7	-127.8 to -97.55	****	<0.0001
	Ubiq 6 vs. No distal	-55	-70.12 to -39.88	****	<0.0001
	Ubiq 6 vs. Ablations 7+9	14.33	-0.7823 to 29.45	ns	0.0726
	Ubiq 6 vs. Ablations 1 to 3	69.67	54.55 to 84.78	****	<0.0001
	Ubiq 6 vs. Composite 1	49.67	34.55 to 64.78	****	<0.0001

	Ubiq 6 vs. Composite 2	59.33	44.22 to 74.45	****	<0.0001
	Ubiq 6 vs. Composite 3	54.33	39.22 to 69.45	****	<0.0001
	Ubiq 6 vs. Composite 4	62	46.88 to 77.12	****	<0.0001
	Ubiq 6 vs. Composite 5	53.33	38.22 to 68.45	****	<0.0001
	Ubiq 6 vs. Composite 6	82	66.88 to 97.12	****	<0.0001
	Ubiq 6 vs. Composite 7	60.33	45.22 to 75.45	****	<0.0001
Huh 7	**Dunnett's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. CAG	-56	-71.12 to -40.88	****	<0.0001
	Ubiq 6 vs. EF1α	-23.67	-38.78 to -8.551	***	0.0003
	Ubiq 6 vs. No distal	-26.33	-41.45 to -11.22	****	<0.0001
	Ubiq 6 vs. Ablations 7+9	-27	-42.12 to -11.88	****	<0.0001
	Ubiq 6 vs. Ablations 1 to 3	12.33	-2.782 to 27.45	ns	0.1712
	Ubiq 6 vs. Composite 1	-13.33	-28.45 to 1.782	ns	0.1134
	Ubiq 6 vs. Composite 2	3	-12.12 to 18.12	ns	0.9996
	Ubiq 6 vs. Composite 3	40.33	25.22 to 55.45	****	<0.0001
	Ubiq 6 vs. Composite 4	35.67	20.55 to 50.78	****	<0.0001
	Ubiq 6 vs. Composite 5	39.33	24.22 to 54.45	****	<0.0001
	Ubiq 6 vs. Composite 6	82	66.88 to 97.12	****	<0.0001
	Ubiq 6 vs. Composite 7	50.33	35.22 to 65.45	****	<0.0001
HEK 293	**Dunnett's multiple comparisons test**	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
	Ubiq 6 vs. CAG	32.33	17.22 to 47.45	****	<0.0001
	Ubiq 6 vs. EF1α	32.33	17.22 to 47.45	****	<0.0001
	Ubiq 6 vs. No distal	14.33	-0.7823 to 29.45	ns	0.0726
	Ubiq 6 vs. Ablations 7+9	6	-9.116 to 21.12	ns	0.9039
	Ubiq 6 vs. Ablations 1 to 3	45.33	30.22 to 60.45	****	<0.0001
	Ubiq 6 vs. Composite 1	37.67	22.55 to 52.78	****	<0.0001
	Ubiq 6 vs. Composite 2	26	10.88 to 41.12	****	<0.0001
	Ubiq 6 vs. Composite 3	55	39.88 to 70.12	****	<0.0001
	Ubiq 6 vs. Composite 4	50	34.88 to 65.12	****	<0.0001
	Ubiq 6 vs. Composite 5	43.67	28.55 to 58.78	****	<0.0001
	Ubiq 6 vs. Composite 6	73.33	58.22 to 88.45	****	<0.0001
	Ubiq 6 vs. Composite 7	68	52.88 to 83.12	****	<0.0001

Table 4: Ubiq-1 2-way ANOVA Dunnett's Multiple Comparisons Test

Ubiq-1 2-way ANOVA Dunnett's Multiple Comparisons Test				
ARPE19 Cells				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-15.03	-28.01 to -2.055	*	0.016
Wt vs. Var2	-5.8	-18.78 to 7.179	ns	0.7731
Wt vs. Var3	-10.63	-23.61 to 2.345	ns	0.1559
Wt vs. Var4	-11.23	-24.21 to 1.745	ns	0.1187
Wt vs. Var5	-16.5	-29.48 to -3.521	**	0.0066
Wt vs. Var6	-5.067	-18.05 to 7.912	ns	0.871
Wt vs. Var7	-6.233	-19.21 to 6.745	ns	0.707
Wt vs. Var8	-7.433	-20.41 to 5.545	ns	0.5169
Wt vs. Var9	0.2	-12.78 to 13.18	ns	>0.9999
Wt vs. Var10	-1.8	-14.78 to 11.18	ns	>0.9999
Y79				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-50.13	-63.11 to -37.15	****	<0.0001
Wt vs. Var2	-40.37	-53.35 to -27.39	****	<0.0001
Wt vs. Var3	-55.03	-68.01 to -42.05	****	<0.0001
Wt vs. Var4	-34.97	-47.95 to -21.99	****	<0.0001
Wt vs. Var5	-6.5	-19.48 to 6.479	ns	0.6648
Wt vs. Var6	-28.77	-41.75 to -15.79	****	<0.0001
Wt vs. Var7	-29.47	-42.45 to -16.49	****	<0.0001
Wt vs. Var8	-24.93	-37.91 to -11.95	****	<0.0001
Wt vs. Var9	-35.4	-48.38 to -22.42	****	<0.0001
Wt vs. Var10	19.9	6.921 to 32.88	***	0.0007

Table 5: Ubiq-2 2-way ANOVA Dunnett's Multiple Comparisons Test

Ubiq-2 2-way ANOVA Dunnett's Multiple Comparisons Test				
ARPE19 Cells				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-2	-11.78 to 7.781	ns	0.9979
Wt vs. Var2	0.7	-9.081 to 10.48	ns	>0.9999
Wt vs. Var3	-3.067	-12.85 to 6.714	ns	0.9592
Wt vs. Var4	0.1	-9.681 to 9.881	ns	>0.9999
Wt vs. Var5	-2.333	-12.11 to 7.448	ns	0.9933
Wt vs. Var6	3.967	-5.814 to 13.75	ns	0.847
Wt vs. Var7	-9.133	-18.91 to 0.6478	ns	0.0775
Wt vs. Var8	2.367	-7.414 to 12.15	ns	0.9926
Wt vs. Var9	-1.9	-11.68 to 7.881	ns	0.9986
Wt vs. Var10	-15.47	-25.25 to -5.686	***	0.0005
Y79				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	14.27	4.486 to 24.05	**	0.0014
Wt vs. Var2	26.87	17.09 to 36.65	****	<0.0001
Wt vs. Var3	39.3	29.52 to 49.08	****	<0.0001
Wt vs. Var4	32.6	22.82 to 42.38	****	<0.0001
Wt vs. Var5	50.87	41.09 to 60.65	****	<0.0001
Wt vs. Var6	29.7	19.92 to 39.48	****	<0.0001
Wt vs. Var7	57.5	47.72 to 67.28	****	<0.0001
Wt vs. Var8	46.1	36.32 to 55.88	****	<0.0001
Wt vs. Var9	16.23	6.452 to 26.01	***	0.0002
Wt vs. Var10	53.3	43.52 to 63.08	****	<0.0001

Table 6: Ubiq-3 2-way ANOVA Dunnett's Multiple Comparisons Test

Ubiq-3 2-way ANOVA Dunnett's Multiple Comparisons Test				
ARPE19 Cells				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	2.133	-4.860 to 9.126	ns	0.9659
Wt vs. Var2	3.333	-3.660 to 10.33	ns	0.7163
Wt vs. Var3	0.1333	-6.860 to 7.126	ns	>0.9999
Wt vs. Var4	3.833	-3.160 to 10.83	ns	0.5685
Wt vs. Var5	0.7	-6.293 to 7.693	ns	>0.9999
Wt vs. Var6	1.067	-5.926 to 8.060	ns	0.9998
Wt vs. Var7	0.6	-6.393 to 7.593	ns	>0.9999
Wt vs. Var8	4.2	-2.793 to 11.19	ns	0.4641
Wt vs. Var9	5.6	-1.393 to 12.59	ns	0.1734
Y79				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-1.133	-8.126 to 5.860	ns	0.9997
Wt vs. Var2	3.4	-3.593 to 10.39	ns	0.6969
Wt vs. Var3	-2.533	-9.526 to 4.460	ns	0.9111
Wt vs. Var4	5.367	-1.626 to 12.36	ns	0.2085
Wt vs. Var5	-8.3	-15.29 to -1.307	*	0.0129
Wt vs. Var6	0.06667	-6.926 to 7.060	ns	>0.9999
Wt vs. Var7	10.8	3.807 to 17.79	***	0.0007
Wt vs. Var8	2.533	-4.460 to 9.526	ns	0.9111
Wt vs. Var9	27.67	20.67 to 34.66	****	<0.0001
Wt vs. Var10	53.85	46.03 to 61.67	****	<0.0001

Table 7: Ubiq-4 2-way ANOVA Dunnett's Multiple Comparisons Test

Ubiq-4 2-way ANOVA Dunnett's Multiple Comparisons Test				
ARPE19 Cells				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-9.367	-19.36 to 0.6232	ns	0.0756
Wt vs. Var2	-1.667	-11.66 to 8.323	ns	0.9996
Wt vs. Var3	-15.6	-25.59 to -5.610	***	0.0006
Wt vs. Var4	-9.267	-19.26 to 0.7232	ns	0.0806
Wt vs. Var5	-6.367	-16.36 to 3.623	ns	0.3954
Wt vs. Var6	-10.2	-20.19 to -0.2101	*	0.0433
Wt vs. Var7	-5.633	-15.62 to 4.357	ns	0.5346
Wt vs. Var8	-9.567	-19.56 to 0.4232	ns	0.0664
Wt vs. Var9	-8.067	-18.06 to 1.923	ns	0.1667
Wt vs. Var10	-16.33	-26.32 to -6.343	***	0.0003
Y79				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-56.33	-66.32 to -46.34	****	<0.0001
Wt vs. Var2	-56.67	-66.66 to -46.68	****	<0.0001
Wt vs. Var3	-71.43	-81.42 to -61.44	****	<0.0001
Wt vs. Var4	-55.5	-65.49 to -45.51	****	<0.0001
Wt vs. Var5	-52.6	-62.59 to -42.61	****	<0.0001
Wt vs. Var6	-47.7	-57.69 to -37.71	****	<0.0001
Wt vs. Var7	-54.63	-64.62 to -44.64	****	<0.0001
Wt vs. Var8	24.07	14.08 to 34.06	****	<0.0001
Wt vs. Var9	-52.2	-62.19 to -42.21	****	<0.0001
Wt vs. Var10	-33.2	-43.19 to -23.21	****	<0.0001

Table 8: Ubiq-6 2-way ANOVA Dunnett's Multiple Comparisons Test

Ubiq-6 2-way ANOVA Dunnett's Multiple Comparisons Test				
ARPE19 Cells				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-8.1	-16.71 to 0.5052	ns	0.0738
Wt vs. Var2	-16.2	-24.81 to -7.595	****	<0.0001
Wt vs. Var3	-0.4333	-9.039 to 8.172	ns	>0.9999
Wt vs. Var4	-8.5	-17.11 to 0.1052	ns	0.0543
Wt vs. Var5	-0.1333	-8.739 to 8.472	ns	>0.9999
Wt vs. Var6	-5.667	-14.27 to 2.939	ns	0.3593
Wt vs. Var7	-6.867	-15.47 to 1.739	ns	0.1759
Wt vs. Var8	-7.767	-16.37 to 0.8385	ns	0.0945
Wt vs. Var9	-16.6	-25.21 to -7.995	****	<0.0001
Wt vs. Var10	-4.167	-12.77 to 4.439	ns	0.699
Y79				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-28.4	-37.01 to -19.79	****	<0.0001
Wt vs. Var2	30.23	21.63 to 38.84	****	<0.0001
Wt vs. Var3	24.7	16.09 to 33.31	****	<0.0001
Wt vs. Var4	11.97	3.361 to 20.57	**	0.0025
Wt vs. Var5	51.8	43.19 to 60.41	****	<0.0001
Wt vs. Var6	13.97	5.361 to 22.57	***	0.0003
Wt vs. Var7	6.133	-2.472 to 14.74	ns	0.2768
Wt vs. Var8	14.73	6.128 to 23.34	***	0.0001
Wt vs. Var9	-4	-12.61 to 4.605	ns	0.7381
Wt vs. Var10	25.33	16.73 to 33.94	****	<0.0001

Table 9: Comp-6 2-way ANOVA Dunnett's Multiple Comparisons Test

Comp-6 2-way ANOVA Dunnett's Multiple Comparisons Test				
ARPE19 Cells				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	9.767	3.482 to 16.05	***	0.0006
Wt vs. Var2	0.7333	-5.552 to 7.018	ns	>0.9999
Wt vs. Var3	-4.967	-11.25 to 1.318	ns	0.1837
Wt vs. Var4	10.9	4.615 to 17.19	***	0.0001
Wt vs. Var5	6.433	0.1482 to 12.72	*	0.0425
Wt vs. Var6	7.833	1.548 to 14.12	**	0.0081
Wt vs. Var7	0.8667	-5.418 to 7.152	ns	>0.9999
Wt vs. Var8	-4.333	-10.62 to 1.952	ns	0.3104
Wt vs. Var9	2.733	-3.552 to 9.018	ns	0.7957
Wt vs. Var10	11.53	5.248 to 17.82	****	<0.0001
Y79				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	14.63	8.348 to 20.92	****	<0.0001
Wt vs. Var2	-5.933	-12.22 to 0.3518	ns	0.0725
Wt vs. Var3	2.6	-3.685 to 8.885	ns	0.8334
Wt vs. Var4	-5.967	-12.25 to 0.3185	ns	0.0701
Wt vs. Var5	3.8	-2.485 to 10.09	ns	0.455
Wt vs. Var6	7.667	1.382 to 13.95	**	0.01
Wt vs. Var7	8.1	1.815 to 14.39	**	0.0058
Wt vs. Var8	-5.4	-11.69 to 0.8851	ns	0.1233
Wt vs. Var9	3.433	-2.852 to 9.718	ns	0.5706
Wt vs. Var10	3.067	-3.218 to 9.352	ns	0.6914

Table 10: Comp-7 2-way ANOVA Dunnett's Multiple Comparisons Test

Comp-7 2-way ANOVA Dunnett's Multiple Comparisons Test				
ARPE19 Cells				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	3.867	-3.735 to 11.47	ns	0.6485
Wt vs. Var2	-0.9333	-8.535 to 6.668	ns	>0.9999
Wt vs. Var3	7.667	0.06503 to 15.27	*	0.0471
Wt vs. Var4	5.333	-2.268 to 12.93	ns	0.2926
Wt vs. Var5	12.4	4.798 to 20.00	***	0.0003
Wt vs. Var6	6.933	-0.6683 to 14.53	ns	0.089
Wt vs. Var7	22.9	15.30 to 30.50	****	<0.0001
Wt vs. Var8	2.6	-5.002 to 10.20	ns	0.9333
Wt vs. Var9	9.967	2.365 to 17.57	**	0.0048
Wt vs. Var10	15.77	8.165 to 23.37	****	<0.0001
Y79				
Test	**Mean Diff.**	**95.00% CI of diff.**	**Summary**	**Adjusted P Value**
Wt vs. Var1	-17.5	-25.10 to -9.898	****	<0.0001
Wt vs. Var2	-7.033	-14.63 to 0.5683	ns	0.0819
Wt vs. Var3	-21.83	-29.43 to -14.23	****	<0.0001
Wt vs. Var4	-1.867	-9.468 to 5.735	ns	0.9917
Wt vs. Var5	21.27	13.67 to 28.87	****	<0.0001
Wt vs. Var6	-19.47	-27.07 to -11.87	****	<0.0001
Wt vs. Var7	-8.6	-16.20 to -0.9984	*	0.0196
Wt vs. Var8	-11.03	-18.63 to -3.432	**	0.0015
Wt vs. Var9	-14.4	-22.00 to -6.798	****	<0.0001
Wt vs. Var10	3.767	-3.835 to 11.37	ns	0.6757

Table 11: Ubiq 1 Motif Swap for Each Variant

Ubiq 1 Motif Swap for Each Variant			
New Motif Sequence	**New Motif**	**Replaced Motif**	**Variant #**
AGGAGGATTAG	CRX_RC	NYF	Var1
CTCAGCACTAG	NRL	EBF1	
AAGAGGATTAA	CRX_RC	UAK42	Var2
ATCAGCAAACC	NRL	NYF	
GGGGGGATTAA	CRX_RC	ZNF143-ext	Var3
TATAATCCGATTACT	OTX2	SP1	
GTTAATCGGATTAGT	OTX2	GABP	Var4
ATCAGCAGACA	NRL	AP2	
CTAATCCCCCC	CRX	CTCF	Var5
TTTAATCCGATTATG	OTX2_RC	SP1	
GGGAGGATTAG	CRX_RC	NYF	Var6
CATAATCCGATTAGC	OTX2_RC	E2F4	
TCTAATCCGATTAGG	OTX2	GABP	Var7
GTCAGCGATCC	NRL_RC	AP2	
TCTAATCCGATTATG	OTX2_RC	ZNF143-ext	Var8
ATCAGCGGGCA	NRL	NYF	
CGAATGCTGAG	NRL_RC	EGR1	Var9
CTAATCCTCTC	CRX	CTCF	
CTAATCCCCCC	CRX	GABP	Var10
CATAATCGGATTACG	OTX2_RC	E2F4	

Table 12: Ubiq 2 Motif Swap for Each Variant

Ubiq 2 Motif Swap for Each Variant			
New Motif Sequence	**New Motif**	**Replaced Motif**	**Variant #**
GTCCTGCTGAG	NRL_RC	MAX	Var1
CTAATCCCCCT	CRX	E2F1	
AGTAATCGGATTAGC	OTX2_RC	CREB-ext	Var2
TTAATCCTCCC	CRX	UA3	
CATAATCGGATTACC	OTX2	CTCF	Var3
TTAATCCCCTT	CRX	TBP	
TATAATCCGATTATG	OTX2_RC	SRF	Var4
CTCAGCGAGTC	NRL	MYC	
CTAATCCCCCT	CRX	UA3	Var5
GAGGGGATTAA	CRX_RC	SRF	
GAGGGGATTAG	CRX_RC	CTCF	Var6
TCTAATCCGCTTATC	OTX2_RC	CREB-ext	
GCTAATCCGATTATT	OTX2_RC	E2F1	Var7
CTAATCCTCCC	CRX	MAX	
ATCAGCGCTAT	NRL	ELF1	Var8
CTAATCCTCCT	CRX	UAK42	
AGTAATCCGCTTATG	OTX2_RC	SP1	Var9
TTAATCCCCCC	CRX	CTCF	
GAGAGGATTAG	CRX_RC	CTCF	Var10
TACATGCTGAC	NRL_RC	CREB-ext	

Table 13: Ubiq 3 Motif Swap for Each Variant

Ubiq 3 Motif Swap for Each Variant			
New Motif Sequence	**New Motif**	**Replaced Motif**	**Variant #**
GGGAGGATTAA	CRX_RC	YY1	Var1
ACTAATCCGATTAAG	OTX2	PU1	
TGTAATCCGATTAGT	OTX2_RC	E2F4	Var2
GGGGGGATTAA	CRX_RC	ELF1	
CTCAGCGATCC	NRL	UAK42	Var3
TTAATCCTCCC	CRX	YY1	
AATAATCGGCTTAAC	OTX2_RC	UA2	Var4
TTAATCCCCTT	CRX	E2F1	
GAGAGGATTAA	CRX_RC	E2F4	Var5
ATTAATCCGATTACC	OTX2	E2F1	
CTAATCCTCTC	CRX	GABP and ELF1	Var6
ACTAATCCGATTATG	OTX2	UAK42	
CTAATCCTCTT	CRX	UAK42	Var7
GAGAGGATTAA	CRX_RC	ELF1	
CTTAAGCCGATTACA	OTX2	GABP and ELF1	Var8
GGGAGGATTAG	CRX_RC	ELF1	
ATCGTGCTGAG	NRL_RC	UA2	Var9
TTAATCCTCCT	CRX	YY1	
CACCTGCTGAC	NRL_RC	PU1	Var10
TTTAATCGGATTATC	OTX2	GABP and ELF1	

Table 14: Ubiq 4 Motif Swap for Each Variant

Ubiq 4 Motif Swap for Each Variant			
New Motif Sequence	**New Motif**	**Replaced Motif**	**Variant #**
ATTAATCGGATTACT	OTX2	PU1	Var1
GGACTGCTGAG	NRL_RC	ELK4	
CCTAATCCGCTTAGT	OTX2_RC	ELF1	Var2
CGCCCGCTGAC	NRL_RC	EGR1	
GAGGGGATTAG	CRX_RC	UAK42	Var3
CATAATCGGATTAGC	OTX2	CREB-ext	
ATCAGCGTACG	NRL	UAK42	Var4
AATAAGCGGATTAAT	OTX2	ELK4	
CCTAATCGGATTAAC	OTX2	ELK4	Var5
AACGCGCTGAG	NRL_RC	GABP	
GAGGGGATTAA	CRX_RC	E2F4	Var6
AGTAATCGGCTTAAT	OTX2_RC	GABP	
CTCAGCGTTTG	NRL	CREB-ext	Var7
AGTAATCCGATTAGA	OTX2_RC	EGR1	
CTCGCGCTGAC	NRL_RC	PU1	Var8
CTAATCCTCTT	CRX	ELK4	
AAGAGGATTAG	CRX_RC	EGR1	Var9
GGTAATCGGCTTAGT	OTX2_RC	GABP	
GATAATCCGATTATT	OTX2	YY1	Var10
CTCAGCAGTAT	NRL	CREB-ext	

Table 15: Ubiq 6 Motif Swap for Each Variant

Ubiq 6 Motif Swap for Each Variant			
New Motif Sequence	**New Motif**	**Replaced Motif**	**Variant #**
TGTAATCCGCTTAAG	OTX2_RC	CREB	Var1
CGCGTGCTGAC	NRL_RC	UAK42	
CTTAATCCGATTATT	OTX2_RC	NFIC	Var2
GAGGGGATTAG	CRX_RC	SP1	
AGGAGGATTAG	CRX_RC	GABP	Var3
ATCAGCATTTA	NRL	ELF1	
AAGGGGATTAG	CRX_RC	TCF7L2	Var4
CTCAGCGCAAA	NRL	GABP	
CTAATCCTCTT	CRX	CREB	Var5
GTCAGCACAAC	NRL	E2F4	
CTTAATCCGATTAGG	OTX2	GABP	Var6
TTAATCCCCTT	CRX	CREB	
ATTAATCGGATTAGT	OTX2	ZNF143-ext	Var7
TTAATCCTCCT	CRX	CREB	
TTCTCGCTGAT	NRL_RC	EGR1	Var8
AGTAATCCGATTAAT	OTX2_RC	NFIC	
AGGGGGATTAG	CRX_RC	AP2	Var9
TACTTGCTGAG	NRL_RC	NRF1	
TTTAATCGGATTATC	OTX2	GABP	Var10
AAGAGGATTAA	CRX_RC	UAK42	

Table 16: Comp-6 Motif Swap for Each Variant

Comp-6 Motif Swap for Each Variant			
New Motif Sequence	**New Motif**	**Replaced Motif**	**Variant #**
CTCAGCACAAT	NRL	E2F4	Var1
CATAATCGGATTATC	OTX2	SP1	
ACTAATCCGCTTAAT	OTX2_RC	E2F4	Var2
GTCAGCGGTCC	NRL	CEBPB	
TCTAAGCGGATTACT	OTX2_RC	TCF7L2	Var3
TTTTCGCTGAT	NRL_RC	EGR1	
ATCAGCGAGTC	NRL	SP1	Var4
GGGAGGATTAG	CRX_RC	TCF7L2	
CTCAGCATATT	NRL	GABP	Var5
TTAATCCTCTT	CRX	ELF1	
AAGAGGATTAG	CRX_RC	GABP	Var6
CTAATCCCCTC	CRX	ZNF143-ext	
CTCAGCACTAG	NRL	SP1	Var7
GATAATCCGATTAGA	OTX2_RC	NFIC	
GGTAATCCGCTTAAG	OTX2_RC	NFIC	Var8
TTAGCGCTGAG	NRL_RC	E2F4	
GGGGGGATTAG	CRX_RC	CEBPB	Var9
TAAGCGCTGAT	NRL_RC	E2F4	
ATTGCGCTGAT	NRL_RC	EGR1	Var10
GTCAGCAAGAA	NRL	TCF7L2	

Table 17: Comp-7 Motif Swap for Each Variant

Comp-7 Motif Swap for Each Variant			
New Motif Sequence	**New Motif**	**Replaced Motif**	**Variant #**
CTTGTGCTGAT	NRL_RC	ZNF143-ext	Var1
AAGAGGATTAG	CRX_RC	ELF1	
CGTGTGCTGAT	NRL_RC	YY1	Var2
CCTAATCCGATTAGG	OTX2_RC	EGR1	
TATAATCCGATTACT	OTX2	ELF1	Var3
TTAATCCTCTT	CRX	ZNF143-ext	
TATAATCCGATTAAT	OTX2	ZNF143-ext	Var4
CCTAATCCGATTAAC	OTX2_RC	EGR1	
GAGGGGATTAG	CRX_RC	GABP	Var5
GATAATCGGCTTATT	OTX2_RC	EGR1	
AAGAGGATTAG	CRX_RC	EGR1	Var6
AACACGCTGAT	NRL_RC	YY1	
AAGAGGATTAG	CRX_RC	EGR1	Var7
CTAATCCCCTC	CRX	ZNF143-ext	
AGGGGGATTAA	CRX_RC	GABP	Var8
GTCAGCACATT	NRL	ELF1	
AGGAGGATTAG	CRX_RC	GABP	Var9
GGCACGCTGAT	NRL_RC	YY1	
TTAATCCCCTC	CRX	ZNF143-ext	Var10
TTTAATCGGATTAGA	OTX2	ELF1	

Chapter III: Tunable Expression of an AAV Delivered Transgene via Recruitment of Endogenous ADAR Enzymes

3.1 Abstract

Gene therapies delivered via an AAV vector typically yield steady state gene expression with minimal options for post-dose regulation. This work provides proof of concept evidence of ADAR-mediated RNA editing for the controlled expression of an AAV-delivered transgene. Through the introduction of a premature termination codon (PTC) in the transgene, we established an OFF/ON switch where protein expression is triggered by ADAR-mediated mRNA editing. *In vitro* and *in vivo* studies utilizing the coagulation human factor IX (hFIX) transgene demonstrated successful rescue of protein expression, obtaining close to 50% of wildtype FIX expression levels post-delivery of the trigger RNA. *In vivo* studies showed a dose-dependent response between the amount of cadRNAis LNPs delivered and the levels of FIX expression detected in mouse plasma. They also demonstrated the feasibility to redose with cadRNAis LNPs and achieve similar levels of FIX expression seen in the initial dose of cadRNAis LNPs. The work presented demonstrates a methodology to achieve tunable expression of an AAV-delivered transgene via ADAR-mediated RNA editing.

3.2 Introduction

Adeno-associated viruses (AAVs) based gene therapies, typically for gene augmentation purposes, yield constitutive gene expression. Therapeutic transgene design involves the selection of a promoter that allows for the desired levels of expression in the appropriate cell type. Additionally, the introduction of other regulatory sequences, such as UTRs, introns, or miRNA binding sites, can further control transgene expression.[20, 215-217] However, in their current design, approved gene therapies are irreversible. Recently, approaches that involve RNAi sense and antisense molecules (REVERSIR) and RNA aptamers are emerging as valuable tools to tune the expression of a therapeutic transgene post-dosing.[20, 218, 219] Thus, the ability to have tunable transgene

expression post-AAV delivery would be beneficial in facilitating safer and broader applicability of gene therapies and enabling patient-centric care. Implementation of regulatable transgenes in a clinical sense could bring lots of potential for future gene therapies but there are also some challenges. Challenges could include the complexity of the regulatable system like those operated by drug-inducible or repressible promoters which can often require the incorporation of small molecule ligands or transcription regulators adding complexity for the design of the vector.[220] Another challenge for regulatable transgene expression could be the *in vivo* monitoring of transgene expression in real time and detection of safety signals before adverse events occur, most applications available to monitor transgene expression do not reflect transgene activity in specific tissues and there is a lack of ability to detect early biomarkers that could warn against overexpression and cell stress.[152]

Existing gene editing tools could be leveraged to control AAV-delivered transgene expression post-dosing. For example, base editors for DNA or RNA could generate precise point mutations in the episomal DNA or its mRNA transcript to control gene expression.[175] Adenosine deaminase acting on RNA (ADAR) enzyme is known to be ubiquitously expressed across mammalian species.[221] The ability to harness the ADAR mediated RNA editing could lead to the ability to direct the editing of genes throughout the body for therapeutic purposes. RNA editing mechanisms via adenosine deaminases acting on RNA (ADARs) have been shown to correct faulty endogenous genes via delivery of the trigger RNA in an AAV.[222] This work shows that endogenous ADARs can be recruited to modify a PTC to an amino acid codon in the AAV-delivered transcript to induce transgene expression. Transgene levels go from undetectable to peaking at 24hrs post-trigger RNA delivery. Interestingly, the efficacy of the ON switch was not decreased when re-dosing. The method presented here describes the possibility of tuning transgene expression post-AAV delivery.

3.3 Results

To control gene expression from an AAV-delivered transgene through ADAR-based RNA editing, we designed an expression switch system that relied on the adenosine (A) to inosine (I) editing of a premature TAG amber stop codon to the amino acid codon TIG (Figure 1A). Inosines are recognized as guanosines by the translational machinery based on their structural similarities.[223] In the OFF state, AAV-delivered transgene expression is abolished by the inability to generate protein from the transcript carrying a TAG PTC. Delivery of trigger RNA in the form of a circular RNA with interspersed loops (cadRNAis) complementary to the adenosine target recruits endogenous ADAR and promotes RNA editing.[184] Conversion of the TAG to the native TGG on the transgene mRNA allows for protein translation (ON state).

To assess our switch's dynamic range and efficiency, we took advantage of the robust readouts (Capillary Western Analysis ELISA, qPCRs) of human coagulation Factor IX (hFIX). TGG codons in the hFIX sequence were identified and mutated to TAGs to generate Tryptophan (W) to PTC mutants. In transfection assays, we found that mutating the TGG encoding Tryptophan in position 118 to a PTC completely abolished hFIX protein expression in Huh7 cells measured by Wes (Figure 1B) and confirmed by ELISA. Different W-to-PTC mutants yielded variable levels of hFIX expression in the same assay (Figure S1A). To recruit endogenous ADAR, we co-transfected a cadRNAis spanning the W118_STOP mutation (or G to A 353). 48 hours later, we detected hFIX protein expression via Wes (Figure 1B). An ELISA was used to quantify hFIX levels in the transfection assays. Data showed that after ADAR-based editing, transgene expression reached 50% of the wildtype FIX levels (Figure 1C). In the same cellular extracts, mRNA levels for hFIX showed that the introduction of the PTC significantly decreased mRNA expression. We hypothesize that this is a product of nonsense-mediated decay (NMD).[224, 225] When ADAR-based editing is triggered by the expression of the cadRNAis, mRNA levels reach approximately 50% of the wt hFIX (Figure 1D), similar to what we had observed at the protein level. Dose-response experiments for the cadRNAis delivered as a plasmid or as an RNA showed a linear relationship with hFIX expression measured by ELISA (Figure 1E and 1F). PCR products from the targeted site were Sanger

sequenced to detect the edit at the RNA level. The traces showed the presence of the unedited A and an edited G with predominantly more abundance of G (Figure 1G). No off-target edits were seen (not shown).

Figure1. *In vitro* data: Determining proof of concept of regulatable expression with endogenous ADAR from PTC

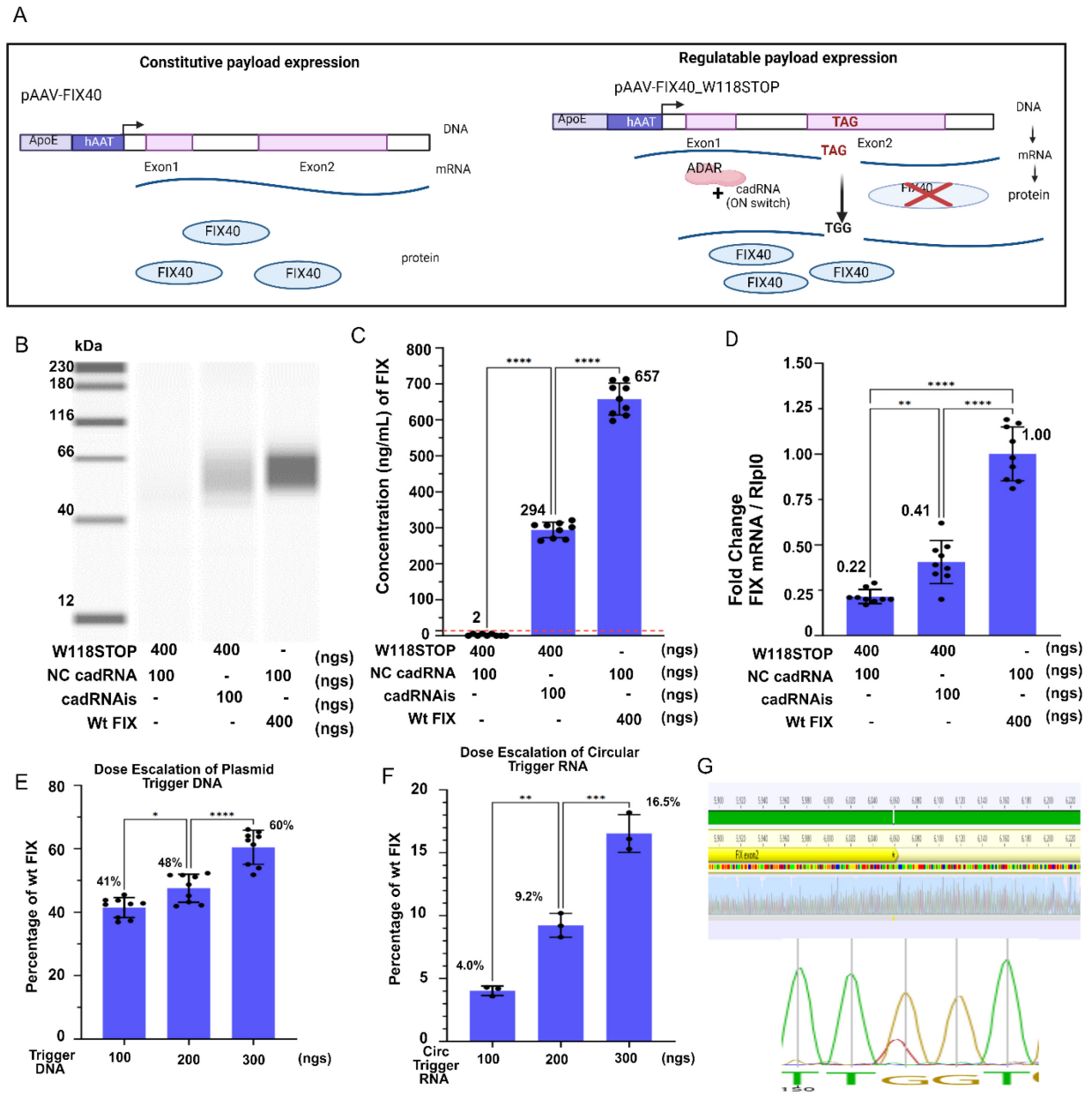


Figure 1 Caption: A. Cartoon depiction of a constitutive payload expression compared to a regulatable payload expression where protein is not produced until the addition of the cadRNAis recruits endogenous ADAR to edit the mRNA and

create the production of protein by targeting and editing the A in the TAG codon into a G for the native TGG codon. **B.** FIX protein measured via Protein Simple Wes capillary system in Huh7 cells 48hrs after transfection. NC: negative control cadRNA. **C.** ELISA measuring FIX expression from the experiment shown in B. Addition of cadRNAis to W118STOP demonstrated statistical signification p <0.0001 compared to W118STOP with NC cadRNA. Statistical significance exists in the abundance of protein expression from W118STOP with cadRNAis compared to wildtype FIX expression with a p <0.0001. **D.** RT-qPCR data for copy number of FIX mRNA normalized to Rlpl0 housekeeping gene represented as foldchange of wildtype FIX. There is statistical significance between W118STOP without cadRNAis compared to W118STOP with cadRNAis, as indicated by a p-value of p = 0.0039. Statistical significance in mRNA abundance of W118STOP with cadRNAis compared to wildtype with a p-value of p <0.0001. **E.** Represents dose escalation of plasmid Trigger DNA in Huh7 cells 48hrs post-transfection from 100ng to 300ng as a percentage of wildtype FIX expression. There is statistical significance from 100ng to 200ng, indicated by a p-value of p = 0.0196, and from 200ng to 300ng with a p-value of p = 0.0001. **F.** Represents dose escalation of trigger RNA in Huh7 cells 48hrs post-transfection from 100ng to 300ng as a percentage of wildtype FIX expression. There is statistical significance from 100ng to 200ng, indicated by a p-value of p = 0.0022, and from 200ng to 300ng with a p-value of p = 0.0003. **G.** Sanger Sequencing results demonstrate editing in Huh7 cells co-transfected with W118STOP and cadRNAis plasmids via detection of stop codon back to native tryptophan codon TGG.

Next, we focused on evaluating the behavior of the ADAR-based switch for hFIX *in vivo*. Aiming also to understand the therapeutic window where hFIX could be detected in plasma, mice were separated into groups with different terminal endpoints (Figure 2A). Mice were dosed with a liver-tropic AAV-hFIX (n=4) or AAV-W118STOP (n=24). As a precautionary measure, our cadRNAis LNPs were co-formulated with a Gaussia Luciferase mRNA to assess efficiency of transfection. Luminescence values were obtained and followed the expected half-life decay of Gaussia Luciferase (Figure 2B).

Two weeks after dosing, wt hFIX levels were at 70 ng/mL, as expected, and W118_PTC had no detectable levels of hFIX (Figure 2C). 1 day post-delivery of cadRNAis LNPs, hFIX levels were at an average of 48 ng/mL. A steady decline in levels of hFIX were observed from day 2 at 38 ng/mL until levels were below the quantification limit (BQL) of 14ng/mL by day 7. This data showed that hFIX expression peaked at 24hs after cadRNAis LNP dosing and that hFIX expression remained in plasma for 7 days. A cohort of mice received a second dose of cadRNAis LNPs at D15. This cohort of mice had BQL levels of hFIX Ag at D10, but by D16 1 day post the second dose of cadRNAis LNPs, their levels of hFIX were at 35 ng/mL. This result showed that the efficiency of the RNA editing is not decreased when redosing with the trigger. To understand the relationship between dosed vector genomes and FIX mRNA expression, we quantified the abundance of hFIX mRNA in the liver normalized to the vector genome copy number (VGCN) of what was dosed. Unlike *in vitro,* no statistically significant difference existed between the AAV-W118STOP treated mice and the control AAV-hFIX. PCR products were sent for Sanger Sequencing to detect the edit at the RNA level. In the sequence where the A to G edit would have occurred the chromatogram showed the presence of A (Figure S2A). Due to the undetectable A to G edit in the chromatogram we theorize that the mechanisms by which ADAR editing occurs is gone by 24hrs, hence why the A to G edit was not detect in the mRNA (Figure S2A). [226]

Figure 2. ***In vivo*** **data. Determining the therapeutic widow of trigger induced FIX expression via endogenous ADAR. I would move the** ***in vivo*** **study design to supplementary information.**

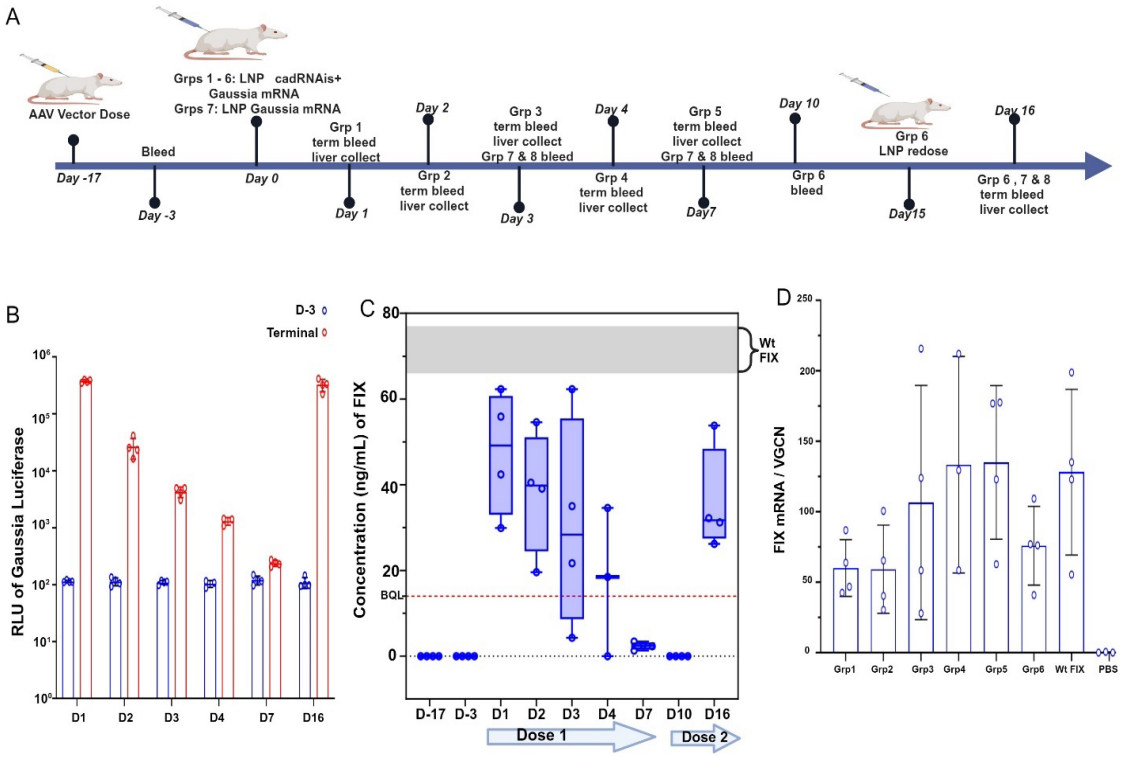


Figure 2 Caption: A. Represents a schematic timeline of an *in vivo* experiment where mice were injected with an AAV vector ~2 weeks prior to being dosed with Trigger LNP. Each day, a different group was sacrificed, and blood and livers were collected. D15 group 6 was redosed and sacrificed on D16 1 day post-dose. **B.** Relative luminescence values for Gaussia Luciferase expression from each group both at a bleed before LNP dosing and at the terminal timepoint bleed. The expression seen from the Gaussia Luciferase follows the Gaussia Luciferase half-life in plasma ~6hrs with R^2 = 0.996 from an exponential decay regression. **C.** FIX expression levels measured via an ELISA at timepoints just prior to trigger LNP delivery and at different terminal timepoints. Wildtype FIX is represented via the grey bar. D1 – D7 represents FIX expression post-dose 1 from different groups of mice. There was no statistical significance in FIX expression seen from D1 to D2, D2 to D3, D3 to D4, or D4 to D7. A group of mice was given LNP dose 2 at D15 and sacrificed 1 day post the 2nd dose. There was no statistically significant

difference between FIX expression seen at D1 or D16. **D.** RT-qPCR data for FIX mRNA abundance normalized to housekeeping gene Rplp0 represented as a fold-over change compared to wildtype FIX; the data is then further normalized to the vector genome copy numbers (VGCN) obtained via RT-qPCR for genomic DNA normalized to housekeeping gene Rlpl0. No statistically significant difference was seen in any groups when compared directly to wildtype abundance of FIX mRNA/VGCN.

Finally, we conducted two dose escalation studies *in* vivo to understand the limiting factor for our ADAR-based expression switch. Mice were dosed with increasing amounts of cadRNAis LNPs (Figure 3A) or dosed with a constant cadRNAis supplemented with exogenous hADAR mRNA (Figure 3B). We noted a dose-dependent response between the amount of cadRNAis LNPs delivered and the level of hFIX achieved, measured at 24hrs after LNP dosing. At the highest dose, 4 at 15 µg, hFIX levels were 400 ng/mL, while the lowest dose, 1 at 3.75 µg, had hFIX levels at 31 ng/mL. Statistical significance between hFIX levels from Dose 1 to Dose 4 is indicated by p-value ($p = 0.0092$) at D1 (Figure 3A). No dose-dependent response was seen when supplemented with exogenous hADAR mRNA, as with cadRNAis LNPs. No statistically significant difference existed between any dose groups with hADAR mRNA; the only statistically significant difference was between Dose 1 (control dose) and Dose 4, the highest dose ($p = 0.0136$) at D1 (Figure 3B). This led us to conclude that the amount of trigger RNA present is the rate-limited factor for achieving FIX expression and that endogenous levels of ADAR enzyme present are sufficient for gene augmentation. To understand the effect of LNP formulation choice on cadRNAis delivery, stability, and hFIX protein levels, we performed a study comparing three liver-directed LNPs (Figure 3C). There were no statistically significant differences in their expression level at any given timepoint (Figure 3C). With invivofectamine 3.0, we observed consistent expression up to D8 and minimal expression out to D20 (Figure 3C). We evaluated terminal timepoints closer to the dosing of LNPs to detect the occurrence of ADAR editing at the RNA level. As expected, there were no detectable levels of hFIX at 1hr and barely any at 4hr detected via an ELISA, but we did have expression 24hrs post LNP at 140 ng/mL (Figure 3D). We normalized the

abundance of hFIX mRNA to the VGCN. We saw no statistically significant difference in the treatment groups, but there was statistical significance from each treatment group to the wt FIX group, with all p-values p <0.0001 (Figure 3E). PCR products from the targeted site were sent for Sanger Sequencing to detect the edit at the RNA level (Figure S2B). The traces showed a majority presence of the unedited A and a minimal edited G (Figure S2B); however, as time increased, the traces of edited Gs became more prevalent. No off-target edits were seen (not shown).

Figure 3. Dose Escalation *in vivo* of Trigger RNA and ADAR mRNA independently and Evaluation of Different LNP Formulations along with Timepoints to Obtain Molecular Kinetics of Editing

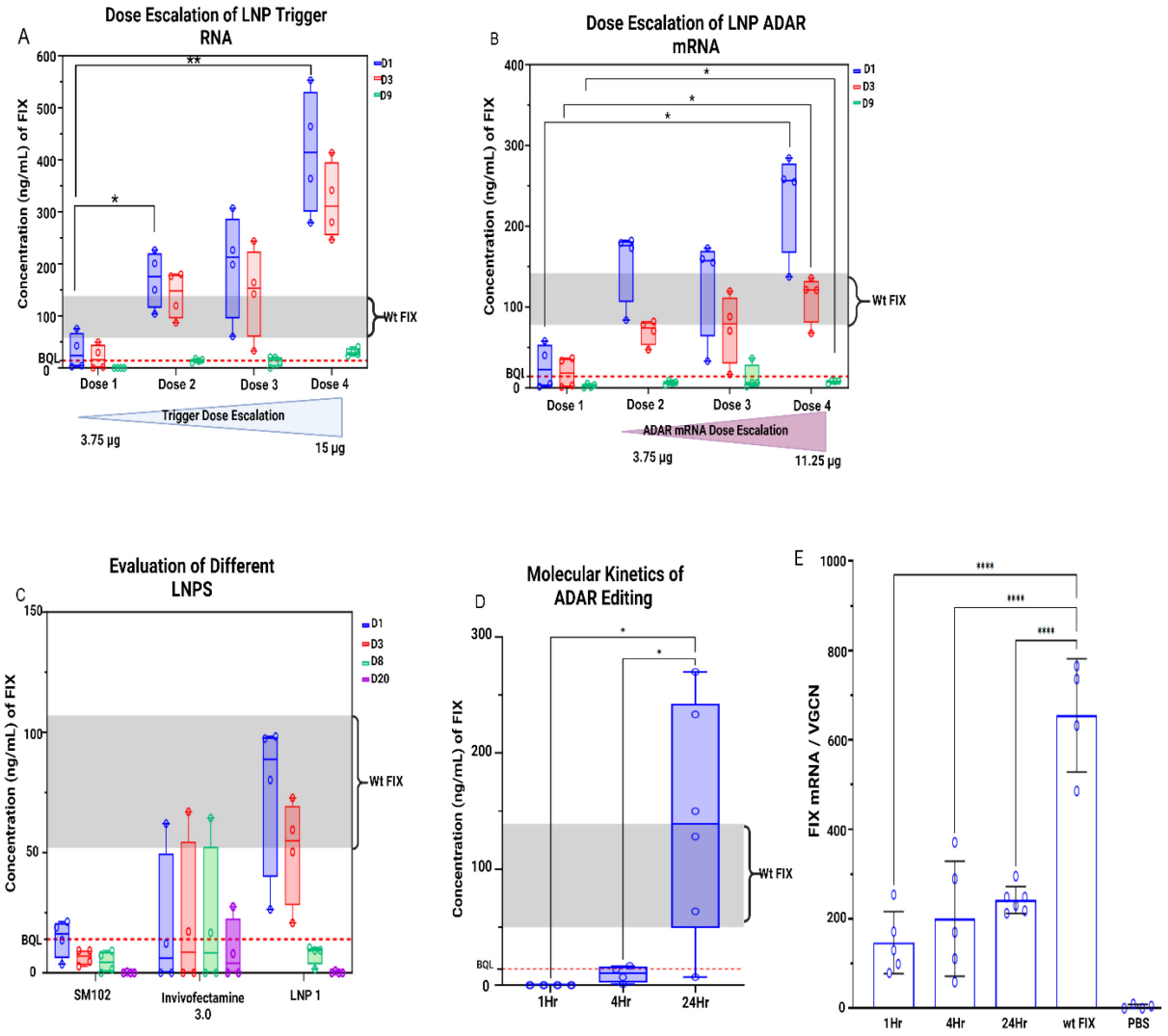


Figure 3 Caption: A. In an *in vivo* experiment where Trigger LNP was increased from 3.75μg to 15μg and followed for 9 days, FIX expression was obtained via an ELISA. There is a statistically significant difference in FIX expression levels seen in D1 from Dose 1 to both Dose 2 and Dose 4 with corresponding p values of $p = 0.0471$ and $p = 0.0092$, respectively. There is also statistical significance at D3 and D4 in the same dose groups with p values of $p = 0.030$ and $p = 0.0039$ for D3 and $p = 0.0106$ and $p = 0.013$ for D9. **B.** *In vivo* experiment where ADAR mRNA was increased from 3.75μg to 11.25μg and compared to the control group. Results were obtained via an ELISA measuring FIX expression. Dose 1, where no addition of ADAR mRNA was added. A statistically significant difference was indicated from dose 1 to dose 4 at all timepoints with a p-value of $p = 0.0136$ for D1, $p = 0.0109$ for D3, and $p = 0.0362$ for D9. **C.** *In vivo* experiment comparing different liver-directed LNPs for evaluation. Differences in expression were assessed utilizing an ELISA to detect FIX expression. There were no statistically significant findings between the different LNPs assessed or between the timepoints based on a Two-way ANOVA with alpha = 0.05 and Tukey's multiple comparison tests performed. **D.** *In vivo* experiment with terminal timepoints at 1hr, 4hr, and 24hr to try and catch molecular kinetics of editing. Results were captured via an ELISA measuring FIX expression. There is statistical significance in the 1hr vs 24hr timepoint with a p-value of $p = 0.0185$, and statistical significance is seen at the 4hr vs 24hr timepoint with a p-value of $p = 0.0271$. There is no statistical significance in expression at 1hr vs 4hr timepoint. **E.** RT-qPCR data for FIX mRNA abundance normalized to housekeeping gene Rplp0 represented as a fold-over change compared to wildtype FIX; the data is then further normalized to the VGCN obtained via RT-qPCR for genomic DNA normalized to housekeeping gene Rlpl0. For all experimental timepoints of 1hr ($p < 0.0001$), 4hr ($p < 0.0001$), and 24hr ($p < 0.0001$), there is statistical significance in the abundance of hFIX mRNA normalized to VGCN compared to the wildtype. There is no statistical significance between the experiment groups when compared to each other for the abundance of hFIX mRNA compared to the amount of VGCN present.

3.4 Discussion

In this study, we have successfully demonstrated an OFF/ON switch for the augmentation of a transcript mediated by an AAV episome utilizing ADAR-based RNA editing in both *in vitro* and *in vivo*. We identified a tryptophan codon in the hFIX transgene that halted protein production when changed to a PTC. There are reports in the literature on the use of ADAR editing in modifying endogenous genes either with overexpression of exogenous ADAR or via the recruitment of endogenous ADAR enzymes with guide RNA, which is often delivered via an AAV.[184, 226-228] Studies conducted by Katrekar et al. demonstrated the utilization of a circular gRNA in an AAV to recruit endogenous ADAR for editing.[184, 228] However, our study utilizes cadRNAis delivered in an LNP to recruit endogenous ADAR to edit our PTC from an amber stop codon TAG back to tryptophan TGG via A to I editing of the mRNA.

In vitro, we demonstrated that we could rescue at least 50% of wildtype FIX protein expression via ADAR-mediated editing of the TAG codon through recruitment by cadRNAis. We also demonstrated that the amount of protein rescued compared to wildtype FIX levels is dependent upon the amount of cadRNAis delivered and validated in both plasmid form and as an RNA. We also observed possible NMD of the mRNA due to the presence of PTC in the transgene, which is known in the literature to initiate mRNA degradation pathways.[225, 229] With the addition of the cadRNAis, there was more mRNA present compared to the PTC alone, indicating that the cadRNAis potentially mitigated the effects of NMD and allowed for mRNA stability for protein production, observed both *in vitro* and *in vivo*.[107] *In vivo,* we demonstrated that cadRNAis had a dose-dependent response and that the amount of cadRNAis delivered is the rate-limited factor in achieving high levels of ADAR editing and protein expression. This phenomenon has also been observed with data presented at ASGCT 2024 by Shape Therapeutics, where they reported increased editing with higher doses of AAV-gRNA that were delivered. We also noted that supplementation with exogenous ADAR had no increased benefit in the augmentation of the FIX transgene from AAV-W118STOP dosed mice. This indicates that the level of endogenous ADAR in a mammalian system is sufficient, and

efforts to try and increase ADAR editing should be focused on the dose of the cadRNAis delivered and not supplementation with exogenous ADAR.

Although we could obtain detectable levels of FIX *in vivo* through the delivery of cadRNAis LNPs, we did not detect high levels of A to G editing in liver tissues from our molecular kinetic-based study through target PCR and Sanger Sequencing. This could be due to the low number of cells targeted to be edited with the cadRNAis and ADAR, which was sufficient to produce detectable protein expression. A study conducted by Katrekar et al. reported 7-17% editing of PTC with an AAV-cadRNA in livers of Idua-W392X mice which is a mouse model for Hurler syndrome a form of mucopolysaccharidosis type 1.[184] Despite the low levels of editing, they did observe a 33% decrease in the accumulation of glycosaminoglycans in those mice, indicating successful partial restoration of α-L-iduronidase activity.[184] Another possibility is that the cadRNAis could be outcompeted by natural substrates of the ADAR enzyme.[13] However, this hypothesis is counterintuitive to the results generated with the supplementation of exogenous ADAR and how there was no significant increase in FIX expression as supplementation of exogenous ADAR increased. A theoretical limitation of the work presented is that it is mostly limited to the liver. Mammalian ADAR enzyme is known to be present in many different tissues types across the body.[230] Theoretically with a different AAV serotype, with a tropism other than liver, ADAR could be utilized in a similar fashion to target a different tissue. Future studies could validate ADAR mediated switch in a larger animal model and could use a different transgene to ensure this off to on switch is not transgene specific. Challenges involving the safety aspect of ADAR based editing system include off target editing, uncontrolled expression, and immunogenicity. To mitigate for off target editing it crucial to optimize the design of the trigger RNA to have high specificity for targeted area of editing.[226] To mitigate immunogenicity from ADAR protein it is best to try and recruit endogenous ADAR enzyme instead of supplementing with the biological system with additional amount of ADAR protein.[231]

In conclusion, we created an expression switch that can regulate an AAV-delivered transcriptome. This expression switch activates the production of protein from a dormant

state due to RNA editing occurring from the cadRNAis ability to recruit endogenous ADAR to edit the PTC to the native W codon. Our results demonstrated RNA editing *in vitro* and partial editing *in vivo* of the hFIX transgene delivered via an AAV. We also demonstrated that therapeutic levels of hFIX protein could be achieved with the administration of cadRNAis LNPs in the context of AAV-delivered transgene.

3.5 Materials and Methods:

PTC Construct Design

Tryptophan or Glutamine (Q) amino acids that could be changed to a stop codon such as TAG, TAA, or TGA and undergo editing either by ADAR or cytosine base editor (CBE) were identified within the coding sequence of the hFIX transgenes utilizing the Geneious Prime software (Dotmatics, Auckland, New Zealand). We identified six codons at amino acid (AA) 88AA, 118AA, 143AA, 167AA, 240AA, and 261AA, all of which could be changed to a stop codon. W at positions 118AA and 240AA were changed to TAG stop codon, while W at 88AA and 261AA were changed to TGA stop codon. All W AA changed can be targeted by ADAR editing. Q at positions 143AA and 167AA could be altered to TAA and TAG stop codon respectfully and target by CBE. Q5® Site-Directed Mutagenesis Kit (NEB, Cat: E0554S) was used to generate the stop codons in the FIX transgene. Primers to conduct the site-directed mutagenesis for each AA were designed using the NEBaseChanger Tool (Ipswich, MA), and oligos were synthesized by Integrated DNA Technologies (Coralville, Iowa). The manufacturer's protocol for the Q5® Site-Directed Mutagenesis Kit was followed to generate the pAAV-FIX with the different stop codons. The PCR product was then transformed into One Shot™ Stbl3™ Chemically Competent E. coli (Invitrogen, Cat: C737303) following the protocol supplied. Colonies were picked and grown overnight in 5mL cultures of LB broth with kanamycin in a shaking incubator at 30ºC. DNA was isolated the following day using the QIAprep Spin Miniprep Kit (Qiagen, Cat: 27106) as outlined in the respective protocol. Plasmids were confirmed correct by NGS sequencing performed by Plasmidsaurus (South San Francisco, CA).

cadRNAis and LNP Formulations

cadRNAis were designed utilizing the approach laid out by Katrekar et al. with the 200bp trigger RNA length, the edit at the 100bp mark, and interspersed loops with the trigger.[184, 228] The adRNAis were circularized by RNA ligase RtcB, which were recruited by flanking the adRNAis with twister ribozyme regions on the 5' and 3' end, all driven by pol3 from the human U6 promoter.[232] Sequences were designed with Geneious Prime software (Dotmatics, Auckland, New Zealand) and synthesized into pUC57 at Azenta (South Plainfield, NJ). Circular mRNA for the cadRNAis was synthesized at GenScript (Piscataway, NJ), and control Gaussia Luciferase mRNA was synthesized at TriLink BioTechnologies (San Diego, CA). Both the cadRNAis mRNA and Gaussia Luciferase mRNA were formulated into separate proprietary LNPs that are liver-targeted and then mixed. LNP formulation was conducted in-house following in-house generated encapsulation protocols utilizing proprietary components. A commercially purchased LNP, Invivofectamine 3.0 (Invitrogen), was used, and the manufacturer's protocol was followed for encapsulating the mRNA.

Cell Culture and Transfection

In vitro experiments were conducted in Huh7 cells grown in DMEM supplemented with 10% FBS in an incubator at 37ºC with a 5% CO_2 atmosphere. Huh7 cells were seeded into a 24-well plate (Corning, Cat: 3527) at a density of 1.00E+5 cells/well. Cells were transiently transfected with 500ng of total nucleic acids. Transfection reagent jetOptimus (Polyplus, Cat: 117-07) was used for all DNA-based transfections following the manufacturer's protocol. Transfection reagent TransIT-mRNA Transfection Kit (MirusBio, Cat: MIR 2225) was used for all RNA-based transfections following the manufacture protocol. Media and cells were harvested at 48h after transfection.

Detection of FIX Antigen and Protein

An ELISA to detect the hFIX antigen from cell supernatant or mouse plasma was conducted with Factor IX Paired Antibody Set (Affinity Biological, Cat: FIX-EIA), and an in-house protocol for a sandwich ELISA was followed. The ELISA was read via BioTek Synergy H1 Multimode Reader (Agilent, Santa Clara, CA USA) and analyzed via SoftMax

Pro Software (Molecular Devices, LLC. San Jose, CA USA). Data was graphed via GraphPad Prism 10 software (GraphPad, San Jose, CA).

A capillary-based western blot analysis (WES) (Protein Simple, San Jose, CA) was conducted to detect the FIX protein with 12-230 kDA Wes Separation Module (Protein Simple, Cat: SM-W004) following the manufacturer's protocol. The primary antibody for the detection of hFIX was the capture antibody in the Factor IX Paired Antibody Set (Affinity Biological, Cat: FIX-EIA) used at a 1:100 dilution. The secondary antibody was the detection antibody in the Factor IX Paired Antibody Set (Affinity Biological, Cat: FIX-EIA) used at a 1:100 dilution. Data was analyzed using the software Compass (Protein Simple, San Jose, CA).

AAV Vector

AAV-hFIX and AAV-W118STOP were prepared in house by Capsid Engineering Team at Spark Therapeutics with an in-house protocol for AAV production which utilized a triple transfection and purification via cesium chloride gradient. Titers for AAV vectors were determined with a RT-qPCR based assay using TaqMan™ Universal PCR Master Mix (Applied Biosystems, Cat: 4304437) and following the manufacturer's protocol. The vectors were aliquoted and stored at -80ºC before use.

Animal Experiment

All animal procedures were performed in accordance with protocol AUFSPK#1: Immunology of Gene Transfer in Mice approved by Institutional Animal Care and Use Committee of Spark Therapeutics, Inc. C57BL/6J male mice 8 weeks of age were acquired from Jackson Laboratories and housed in ambient temperature with ~55% humidity and 12h light/dark cycle. AAVs were injected intravenously at a dose of 5.00E+09 or 5.00E+10 vector genomes per mouse. LNPs were injected intravenously at doses of 0.3mpk, 0.45mpk, 0.6mpk, and 0.75mpk. There were 4 mice per group, experimental and control. Plasma and liver tissue samples were collected in a serial sampling of 1-day post-LNP dose, followed by 2-day post-LNP dose, and so on until day 16. In other studies, plasma samples were collected at timepoint intervals of D1, D3, and D9 post-LNP dose.

The final experiment collected plasma and liver tissue samples at 1h, 4h, and 24h post-LNP dose.

Luciferase Assay

Gaussia Luciferase was detected with the Renilla Luciferase Assay System (Promega, Cat: E2820) following the manufacturer's protocol. Mouse plasma was diluted 1:10 into control C57BL/6 plasma and plated into a 96-well black plate with a solid bottom. The plate was then loaded into GloMax Discover (Promega, Madison, WI), and 100µL of assay buffer was injected one well at a time at a speed of 200µl/s with a luminescence kinetics integration time of 1 second. Relative luminesce units are graphed via GraphPad Prism 10 software (GraphPad, San Jose, CA).

Cell and Tissue Extraction for RNA and Genomic DNA

RNA from cells and liver tissues were extracted using RNeasy Plus Mini Kit (Qiagen, Cat: 74134) following the manufacturer's protocol. cDNA was synthesized with a High-Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Applied Biosystems, Cat: 4374967) following the manufacturer's protocol. 500ng of cDNA was amplified by PCR by NEBNext® High-Fidelity 2X PCR Master Mix (NEB, Cat: M0541L) with primers specific to the target site, generating a PCR product of ~500bp following the manufacturer's protocol. Following the manufacturer's protocol, The PCR product was purified with DNA Clean & Concentrator-5 (Zymo Research, Cat: D4014). The purified PCR product was premixed with primers and sent for Sanger Sequencing at Azenta (South Plainfield, NJ). Trace files were analyzed in Geneious Prime software (Dotmatics, Auckland, New Zealand) for the presence of editing.

Genomic DNA was extracted from liver tissues using the QIAamp DNA Mini Kit (Qiagen, Cat: 51304) following the manufacturer's protocol.

Detection of hFIX mRNA and Vector Genome Copy Number (VGCN)

2µL of cDNA or gDNA were used to set up a 20µL RT-qPCR reaction using TaqMan™ Universal PCR Master Mix (Applied Biosystems, Cat: 4304437) and following the manufacturer's protocol. Primers and probe assays were designed to detect hFIX,

BGHpA, and housekeeping genes RLPL0 in humans and mice. They were ordered from IDT as a PrimeTime™ qPCR Probe Assay with 5' FAM reporter dye and 3'ZEN/IowaBlackFQ quencher. Three technical replicates were done for each sample. RT-qPCR was run on the QuantStudio 7 (ThermoFisher, Waltham, MA) with standard curves. Samples were analyzed in the QuantStudio software ExpressionSuite (ThermoFisher, Waltham, MA), where mRNA and VGCN were normalized to the housekeeping genes h/mRLPL0 and graphed in GraphPad Prism 10 software (GraphPad, San Jose, CA).

hFIX primer 1: 5'-CGCTTGTACCCATTGATGATG-3', hFIX primer 2: 5'-TTATAACCCTGAAGTGCTCGAC-3',

hFIX probe: 56-FAM/5'-AGGCCCTGC-3'/ZEN/5'-ACTCTCGCTT-3'/3IABkFQ

hRlpl0 primer 1: 5'-TGTCTGCTCCCACAATGAAAC-3', hRlpl0 primer 2: 5'-TCGTCTTTAAACCCTGCGTG-3', hRlpl0 probe: 56-FAM/5'-CCCTGTCTT-3'/ZEN/5'-CCCTGGGCATCAC-3'/3IABkFQ

mRlpl0 primer 1: 5'-CCTGCTGTCTGCTGAATGTA-3', mRlpl0 primer 2: 5'-TACACTCTCTCTCCAGATTCCC-3', mRlpl0 probe: 56-FAM/5'-ACCAAAGAG-3'/ZEN/5'-GTATAACTCTGGCAAGCT-3'/3IABkFQ

BGHpA primer 1: 5'-CTTGCCTTCCTTGACCCT-3', BGHpA primer 2: 5'-CCCAGAATAGAATGACACCTACT-3', BGHpA probe: 56-FAM/5'-TTAGGAAAG-3'/ZEN/5'-GACAGTGGGAGTGGC-3'/3IABkFQ

Statistical Analysis

All statistical analysis was conducted in GraphPad Prism 10 software (GraphPad, San Jose, CA). ANOVAs were conducted to examine differences in expression from *in vivo* and *in vitro* experiments. Post hoc multiple comparisons were performed using Tukey's and Dunnett's tests. For all figures * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

3.6 Appendix

Figure S1

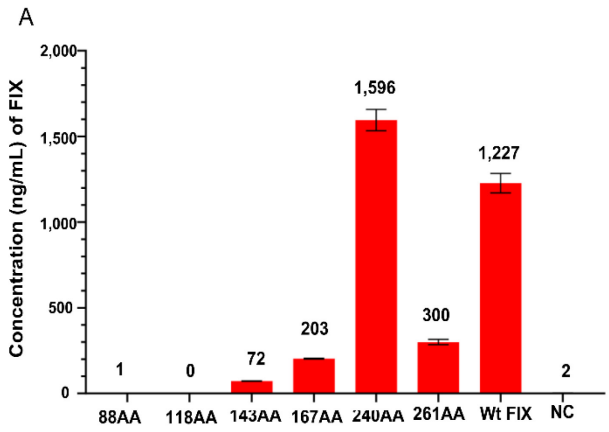


Figure S1 Caption: A. ELISA showing FIX expression levels obtained from Huh7 transiently transfected cells with plasmids of the PTC codons at different AA locations within the FIX transgene. PTC at positions 143AA, 167AA, 240AA, and 261AA all demonstrated leaky expression.

Figure S2

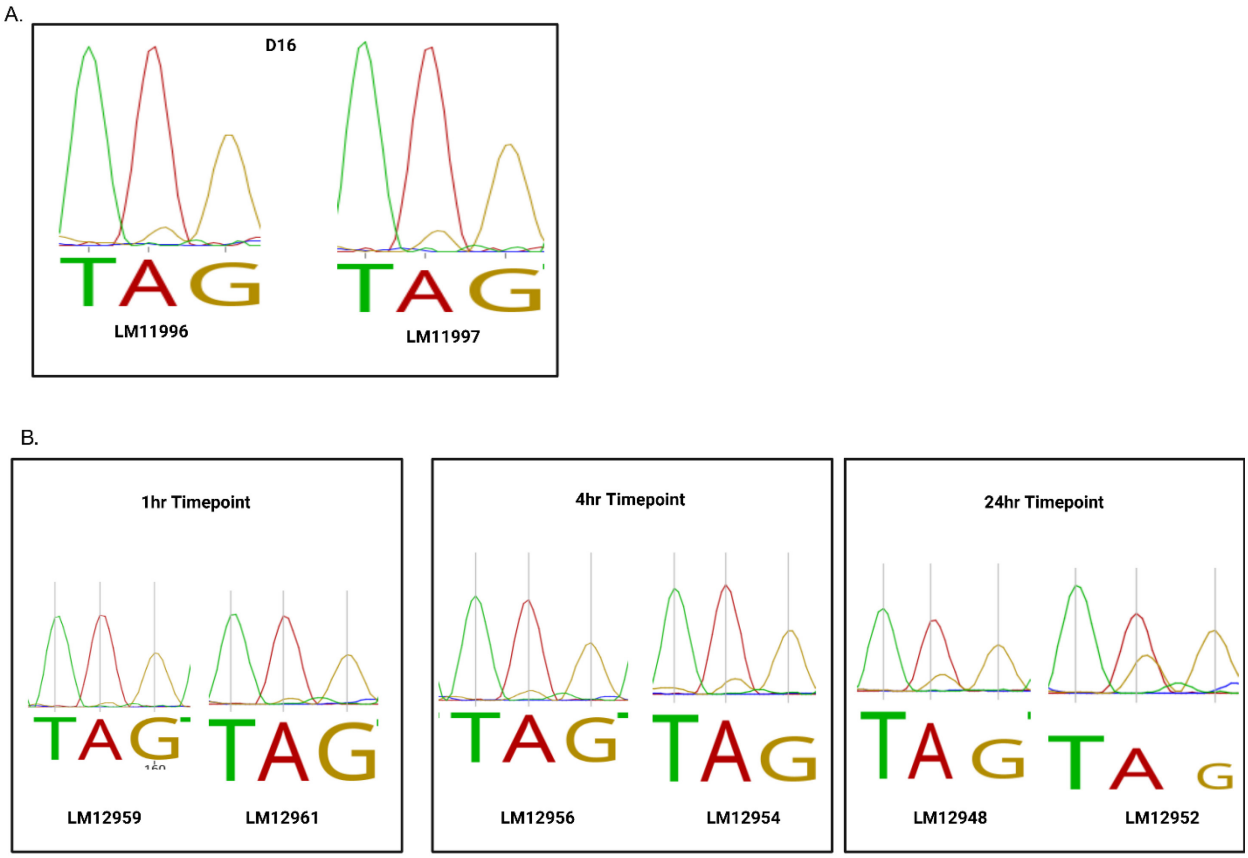


Figure S2 Caption: A. Sanger Sequencing results from mice livers at D16, were dosed with AAV-W118STOP-FIX and with cadRNAis LNP on D1 and D15, for the detection of A to G edit in chromatogram. LM# is the animal ID number. **B.** Sanger Sequencing results from mice livers dosed with AAV-W118STOP-FIX and cadRNAis LNP at timepoint intervals of 1 hour, 4 hours, and 24 hours post LNP dose for detection of A to G edit in the chromatogram. LM# is the animal ID number.

Chapter IV: Conclusion

My graduate research sought to address known limitations to transgene expression for AAV-delivered gene transfer. The hypothesis-generating results from this work can guide future efforts toward the development of safer and more effective gene therapies. Gene therapy has undergone a remarkable evolution over the past few decades.[1] Recent advances in the field have led to the development and approval of numerous therapies addressing genetic disorders.[51] Despite the progress in gene therapy, optimizing AAV-based gene therapies still face challenges.[52, 203] These challenges include limited DNA packaging capacity (<5kb), therapies triggering an immunological response to both the capsid and the transgene, and issues surrounding the precise control of transgene expression across cell types.[76, 85, 152, 203] The research presented in this dissertation investigated approaches to regulate transgene expression in the context of AAV-delivered gene therapies while addressing these known limitations. The ability to regulate transgene expression is essential in AAV-based gene therapies for a multitude of reasons, including enhancing tissue specificity, reducing the risk of immune-related responses and toxicities, and providing optimal therapeutic doses.[26, 127, 151] This project included two distinct aims focused on the enhancement of upstream promoter aspects and post-administration regulation of the transgene expression.

The first approach studied in this dissertation focused on enhancing promoters' properties within the AAV expression cassette through promoter truncation and increasing tissue specificity. This work investigated the feasibility of truncating ubiquitous promoters to generate compact variants that retained functionality and potency, enhanced specificity, and are suitable for ocular gene therapies, in which small and specific promoters are highly needed. Six ubiquitous promoters, Ubiq-1 through Ubiq-6, underwent *in vitro* characterization across a multitude of cell types. These characterization efforts utilized reporter genes to access expression levels to aid in identifying regions that were critical components to regulate promoter expression. The Ubiq-6 promoter underwent a highly targeted deletion series to assist in miniaturizing the promoter. This deletion series led to the generation of multiple truncated versions that ranged in length from 282 bp to 682 bp.

This led to over a 75% reduction in size of promoters. Promoters were further engineered by swapping TFBMs with ocular-specific ones to enhance tissue specificity. Our results demonstrated that swapping TFBMs with tissue-specific ones was able to increase specificity in ocular cell lines for promoter Comp-7-Var1. These findings indicate that ubiquitous promoters can be successfully shortened and tailored to enhance tissue specificity while maintaining functional potency. These findings build upon previous miniaturization efforts in gene therapy and provides a foundation for designing compact promoters optimized for gene therapies targeting ocular diseases, particularly those that require large transgenes.[198] The results offer valuable insights that will guide future efforts to refine promoter designs for therapeutic applications.

The second approach studied in this dissertation research focused on creating a tunable expression system for AAV-delivered episomes that allows for post-administration regulation of the transgene. This set of experiments demonstrated proof-of-concept of a tunable transgene expression by utilizing ADAR-mediated editing of mRNA at the post-transcriptional level to generate an Off-to-On switch through altering PTC in the transgene region. The results generated in vitro and *in vivo* utilizing the hFIX transgene demonstrated at least 50% successful protein recovery when compared directly to wildtype FIX. This successful protein recovery was mediated by the administration of cadRNAis and their ability to recruit endogenous ADAR enzymes to edit the PTC back to the tryptophan codon. Further *in vivo* testing demonstrated the ability to readminister cadRNAis LNPs and achieve similar levels of FIX expression as in the first dose. The work represents a proof of concept of an innovative method for regulating transgene expression in AAV-based gene therapies using ADAR-mediated RNA editing. Unlike previous approaches used by other investigators, utilizing ADAR-mediated RNA editing enables greater flexibility and precision in therapeutic gene control, opening new possibilities for more adaptable treatment strategies for gene therapies.[226] ADAR-mediated RNA editing could be beneficial in the other gene delivery systems like disease modeling where it could be beneficial to silence the transgene at different stages of disease progression to understand the gene function better.[107]

Future work is warranted to evaluate the TFBM swap variants in the context of an AAV-delivered system, preferably in an *in vivo* study, to properly understand the effect of swapping TFBM has on tissue specificity. The goal of the research in this area would be to incorporate generated compact, potent promoters in expression cassettes for ocular-related diseases with oversized transgenes. Future work for ADAR-mediated mRNA editing for the tunable transgene expression should be evaluated in a larger animal model, and a different transgene should be utilized to ensure the off-to-on switch is not transgene-specific. This work indicates this tunable approach can be used for other transgenes.[226] Future studies should aim to incorporate this ADAR-mediated switch to be used in clinical applications for disease conditions in which constant expression of a transgene is not desirable, for example in the case of expression genes modulating the immune system in autoimmunity. Collectively, the studies included in this dissertation research indicate that enhancing promoters' properties within the AAV expression cassette through extensive promoter truncation while increasing tissue specificity is feasible, and that post-administration regulation of the transgene is possible by utilizing a tunable expression system for AAV vectors.

References

1. Malech, H.L., E.K. Garabedian, and M.M. Hsieh, *Evolution of Gene Therapy, Historical Perspective.* Hematol Oncol Clin North Am, 2022. **36**(4): p. 627-645.
2. Matthaei, J.H., et al., *Characteristics and composition of RNA coding units.* Proc Natl Acad Sci U S A, 1962. **48**(4): p. 666-77.
3. Friedmann, T., *A brief history of gene therapy.* Nat Genet, 1992. **2**(2): p. 93-8.
4. Dunbar, C.E., et al., *Gene therapy comes of age.* Science, 2018. **359**(6372).
5. Institute, N.H.G.R. *Human Genome Project*. 2024 June 13, 2024; Available from: https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genome-project.
6. Wilson, B.J. and S.G. Nicholls, *The Human Genome Project, and recent advances in personalized genomics.* Risk Manag Healthc Policy, 2015. **8**: p. 9-20.
7. Taylor, D.J., et al., *Beyond the Human Genome Project: The Age of Complete Human Genome Sequences and Pangenome References.* Annu Rev Genomics Hum Genet, 2024. **25**(1): p. 77-104.
8. Collins, F.S. and L. Fink, *The Human Genome Project.* Alcohol Health Res World, 1995. **19**(3): p. 190-195.
9. Sivam, V., *Has the Human Genome Project Delivered for Healthcare?* Ann Med Surg (Lond), 2012. **1**: p. 19-20.
10. Hood, L. and L. Rowen, *The Human Genome Project: big science transforms biology and medicine.* Genome Med, 2013. **5**(9): p. 79.
11. Escors, D. and K. Breckpot, *Lentiviral vectors in gene therapy: their current status and future potential.* Arch Immunol Ther Exp (Warsz), 2010. **58**(2): p. 107-19.
12. King, N.M.P. and C.E. Bishop, *New treatments for serious conditions: ethical implications.* Gene Ther, 2017. **24**(9): p. 534-538.
13. Arabi, F., V. Mansouri, and N. Ahmadbeigi, *Gene therapy clinical trials, where do we go? An overview.* Biomed Pharmacother, 2022. **153**: p. 113324.
14. Howe, S.J., et al., *Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients.* J Clin Invest, 2008. **118**(9): p. 3143-50.

15. Hacein-Bey-Abina, S., et al., *Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1.* J Clin Invest, 2008. **118**(9): p. 3132-42.
16. Breckpot, K., J.L. Aerts, and K. Thielemans, *Lentiviral vectors for cancer immunotherapy: transforming infectious particles into therapeutics.* Gene Ther, 2007. **14**(11): p. 847-62.
17. Hastie, E. and R.J. Samulski, *Adeno-associated virus at 50: a golden anniversary of discovery, research, and gene therapy success--a personal perspective.* Hum Gene Ther, 2015. **26**(5): p. 257-65.
18. Daya, S. and K.I. Berns, *Gene therapy using adeno-associated virus vectors.* Clin Microbiol Rev, 2008. **21**(4): p. 583-93.
19. Lai, C.M., Y.K. Lai, and P.E. Rakoczy, *Adenovirus and adeno-associated virus vectors.* DNA Cell Biol, 2002. **21**(12): p. 895-913.
20. Wang, J.H., et al., *Adeno-associated virus as a delivery vector for gene therapy of human diseases.* Signal Transduct Target Ther, 2024. **9**(1): p. 78.
21. Wang, D., P.W.L. Tai, and G. Gao, *Adeno-associated virus vector as a platform for gene therapy delivery.* Nat Rev Drug Discov, 2019. **18**(5): p. 358-378.
22. Elmore, Z.C., et al., *The membrane associated accessory protein is an adeno-associated viral egress factor.* Nat Commun, 2021. **12**(1): p. 6239.
23. Naso, M.F., et al., *Adeno-Associated Virus (AAV) as a Vector for Gene Therapy.* BioDrugs, 2017. **31**(4): p. 317-334.
24. Khalid, K. and C.L. Poh, *The development of DNA vaccines against SARS-CoV-2.* Adv Med Sci, 2023. **68**(2): p. 213-226.
25. McFarland, T.J., et al., *Evaluation of a novel short polyadenylation signal as an alternative to the SV40 polyadenylation signal.* Plasmid, 2006. **56**(1): p. 62-7.
26. Keravala, A. and M. Gasmi, *Expression Cassette Design and Regulation of Transgene Expression.* Int Ophthalmol Clin, 2021. **61**(3): p. 17-33.
27. Haberle, V. and A. Stark, *Eukaryotic core promoters and the functional basis of transcription initiation.* Nat Rev Mol Cell Biol, 2018. **19**(10): p. 621-637.
28. Domenger, C. and D. Grimm, *Next-generation AAV vectors-do not judge a virus (only) by its cover.* Hum Mol Genet, 2019. **28**(R1): p. R3-R14.

29. Powell, S.K., R. Rivera-Soto, and S.J. Gray, *Viral expression cassette elements to enhance transgene target specificity and expression in gene therapy.* Discov Med, 2015. **19**(102): p. 49-57.
30. Fan, K., et al., *Genetic and epigenetic features of promoters with ubiquitous chromatin accessibility support ubiquitous transcription of cell-essential genes.* Nucleic Acids Res, 2021. **49**(10): p. 5705-5725.
31. Flotte, T.R., *Birth of a new therapeutic platform: 47 years of adeno-associated virus biology from virus discovery to licensed gene therapy.* Mol Ther, 2013. **21**(11): p. 1976-81.
32. Srivastava, A., E.W. Lusby, and K.I. Berns, *Nucleotide sequence and organization of the adeno-associated virus 2 genome.* J Virol, 1983. **45**(2): p. 555-64.
33. Tratschin, J.D., et al., *A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase.* Mol Cell Biol, 1984. **4**(10): p. 2072-81.
34. Flotte, T.R., et al., *Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter.* J Biol Chem, 1993. **268**(5): p. 3781-90.
35. Flotte, T.R., et al., *Stable in vivo expression of the cystic fibrosis transmembrane conductance regulator with an adeno-associated virus vector.* Proc Natl Acad Sci U S A, 1993. **90**(22): p. 10613-7.
36. Conrad, C.K., et al., *Safety of single-dose administration of an adeno-associated virus (AAV)-CFTR vector in the primate lung.* Gene Ther, 1996. **3**(8): p. 658-68.
37. Guggino, W.B. and L. Cebotaru, *Gene Therapy for Cystic Fibrosis Paved the Way for the Use of Adeno-Associated Virus in Gene Therapy.* Hum Gene Ther, 2020. **31**(9-10): p. 538-541.
38. Snyder, R.O., et al., *Persistent and therapeutic concentrations of human factor IX in mice after hepatic gene transfer of recombinant AAV vectors.* Nat Genet, 1997. **16**(3): p. 270-6.
39. Snyder, R.O., et al., *Correction of hemophilia B in canine and murine models using recombinant adeno-associated viral vectors.* Nat Med, 1999. **5**(1): p. 64-70.

40. 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.
41. Muramatsu, S., et al., *A phase I study of aromatic L-amino acid decarboxylase gene therapy for Parkinson's disease.* Mol Ther, 2010. **18**(9): p. 1731-5.
42. 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.
43. Burnett, J.R. and A.J. Hooper, *Alipogene tiparvovec, an adeno-associated virus encoding the Ser(447)X variant of the human lipoprotein lipase gene for the treatment of patients with lipoprotein lipase deficiency.* Curr Opin Mol Ther, 2009. **11**(6): p. 681-91.
44. Peyvandi, F. and I. Garagiola, *Clinical advances in gene therapy updates on clinical trials of gene therapy in haemophilia.* Haemophilia, 2019. **25**(5): p. 738-746.
45. Manno, C.S., et al., *Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response.* Nat Med, 2006. **12**(3): p. 342-7.
46. Lapteva, L., et al., *Clinical Development of Gene Therapies: The First Three Decades and Counting.* Mol Ther Methods Clin Dev, 2020. **19**: p. 387-397.
47. Kaemmerer, W.F., *How will the field of gene therapy survive its success?* Bioeng Transl Med, 2018. **3**(2): p. 166-177.
48. Gruber, K., *Europe gives gene therapy the green light.* Lancet, 2012. **380**(9855): p. e10.
49. Gaudet, D., et al., *Efficacy and long-term safety of alipogene tiparvovec (AAV1-LPLS447X) gene therapy for lipoprotein lipase deficiency: an open-label trial.* Gene Ther, 2013. **20**(4): p. 361-9.
50. Agency, E.M.; Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/glybera.
51. Ginn, S.L., et al., *Gene therapy clinical trials worldwide to 2023-an update.* J Gene Med, 2024. **26**(8): p. e3721.
52. Mingozzi, F. and K.A. High, *Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges.* Nat Rev Genet, 2011. **12**(5): p. 341-55.

53. Bio, M.; Available from: https://www.mirusbio.com/fda-approved-gene-cell-therapies/.
54. Padhy, S.K., et al., *Voretigene Neparvovec and Gene Therapy for Leber's Congenital Amaurosis: Review of Evidence to Date.* Appl Clin Genet, 2020. **13**: p. 179-208.
55. Petit, L. and C. Punzo, *Gene therapy approaches for the treatment of retinal disorders.* Discov Med, 2016. **22**(121): p. 221-229.
56. Darrow, J.J., *Luxturna: FDA documents reveal the value of a costly gene therapy.* Drug Discov Today, 2019. **24**(4): p. 949-954.
57. Gao, J., R.M. Hussain, and C.Y. Weng, *Voretigene Neparvovec in Retinal Diseases: A Review of the Current Clinical Evidence.* Clin Ophthalmol, 2020. **14**: p. 3855-3869.
58. Mahajan, R., *Onasemnogene Abeparvovec for Spinal Muscular Atrophy: The Costlier Drug Ever.* Int J Appl Basic Med Res, 2019. **9**(3): p. 127-128.
59. Bitetti, I., et al., *Onasemnogene abeparvovec gene replacement therapy for the treatment of spinal muscular atrophy: a real-world observational study.* Gene Ther, 2023. **30**(7-8): p. 592-597.
60. Tice, J.A., et al., *The effectiveness and value of gene therapy for hemophilia: A Summary from the Institute for Clinical and Economic Review's California Technology Assessment Forum.* J Manag Care Spec Pharm, 2023. **29**(5): p. 576-581.
61. Pipe, S.W., et al., *Gene Therapy with Etranacogene Dezaparvovec for Hemophilia B.* N Engl J Med, 2023. **388**(8): p. 706-718.
62. Leigh Ann Anderson, P.; Available from: https://www.drugs.com/medical-answers/how-elevidys-cost-3577480/.
63. Hoy, S.M., *Delandistrogene Moxeparvovec: First Approval.* Drugs, 2023. **83**(14): p. 1323-1329.
64. Madan, B., et al., *Three-year outcomes of valoctocogene roxaparvovec gene therapy for hemophilia A.* J Thromb Haemost, 2024. **22**(7): p. 1880-1893.

65. PERRONE, M. *$2.9 million gene therapy for severe hemophilia is approved by FDA*. 2023; Available from: https://apnews.com/article/hemophilia-gene-therapy-fda-approval-56ec64489b344e3345a49cb1390cfd10.
66. Dunleavy, K. *Pfizer scores FDA nod for hemophilia B gene therapy, will charge $3.5M per dose*. 2024; Available from: https://www.fiercepharma.com/pharma/pfizer-scores-fda-nod-hemophilia-b-gene-therapy-will-charge-35m-dose.
67. *CADTH Reimbursement Review Fidanacogene Elaparvovec (Beqvez).* Canadian Journal of Health Technologies, 2024. **4**.
68. PTC Therapeutics, I., *Prescribing Information for Kebilidi*. 2024.
69. *FDA Approves Gene Therapy for Ultra-Rare Nervous System Disorder*. 2024; Available from: https://www.navlindaily.com/article/23886/fda-approves-gene-therapy-for-ultra-rare-nervous-system-disorder.
70. PTC Therapeutics, I. *PTC Therapeutics Announces FDA Approval of AADC Deficiency Gene Therapy*. 2024; Available from: https://ir.ptcbio.com/news-releases/news-release-details/ptc-therapeutics-announces-fda-approval-aadc-deficiency-gene.
71. Bui, T.A., et al., *Advancements and challenges in developing in vivo CAR T cell therapies for cancer treatment.* EBioMedicine, 2024. **106**: p. 105266.
72. Wu, Z., H. Yang, and P. Colosi, *Effect of genome size on AAV vector packaging.* Mol Ther, 2010. **18**(1): p. 80-6.
73. Sun, L., J. Li, and X. Xiao, *Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization.* Nat Med, 2000. **6**(5): p. 599-602.
74. Kolesnik, V.V., et al., *Optimization strategies and advances in the research and development of AAV-based gene therapy to deliver large transgenes.* Clin Transl Med, 2024. **14**(3): p. e1607.
75. Hu, Z., et al., *Targeted B-domain deletion restores F8 function in human endothelial cells and mice.* Signal Transduct Target Ther, 2022. **7**(1): p. 189.
76. 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.

77. Marrone, L., P.M. Marchi, and M. Azzouz, *Circumventing the packaging limit of AAV-mediated gene replacement therapy for neurological disorders.* Expert Opin Biol Ther, 2022. **22**(9): p. 1163-1176.
78. Veron, P., et al., *Humoral and cellular capsid-specific immune responses to adeno-associated virus type 1 in randomized healthy donors.* J Immunol, 2012. **188**(12): p. 6418-24.
79. Mingozzi, F., et al., *CD8(+) T-cell responses to adeno-associated virus capsid in humans.* Nat Med, 2007. **13**(4): p. 419-22.
80. 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.
81. Monahan, P.E., *Gene therapy in an era of emerging treatment options for hemophilia B.* J Thromb Haemost, 2015. **13 Suppl 1**(0 1): p. S151-60.
82. Li, C., et al., *Neutralizing antibodies against adeno-associated virus examined prospectively in pediatric patients with hemophilia.* Gene Ther, 2012. **19**(3): p. 288-94.
83. Calcedo, R., et al., *Adeno-associated virus antibody profiles in newborns, children, and adolescents.* Clin Vaccine Immunol, 2011. **18**(9): p. 1586-8.
84. Calcedo, R., et al., *Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses.* J Infect Dis, 2009. **199**(3): p. 381-90.
85. Masat, E., G. Pavani, and F. Mingozzi, *Humoral immunity to AAV vectors in gene therapy: challenges and potential solutions.* Discov Med, 2013. **15**(85): p. 379-89.
86. Earley, J., et al., *Evading and overcoming AAV neutralization in gene therapy.* Trends Biotechnol, 2023. **41**(6): p. 836-845.
87. Mingozzi, F. and K.A. High, *Immune responses to AAV vectors: overcoming barriers to successful gene therapy.* Blood, 2013. **122**(1): p. 23-36.
88. Majowicz, A., et al., *Successful Repeated Hepatic Gene Delivery in Mice and Non-human Primates Achieved by Sequential Administration of AAV5(ch) and AAV1.* Mol Ther, 2017. **25**(8): p. 1831-1842.
89. Leborgne, C., et al., *IgG-cleaving endopeptidase enables in vivo gene therapy in the presence of anti-AAV neutralizing antibodies.* Nat Med, 2020. **26**(7): p. 1096-1101.

90. Elmore, Z.C., et al., *Rescuing AAV gene transfer from neutralizing antibodies with an IgG-degrading enzyme.* JCI Insight, 2020. **5**(19).
91. Costa Verdera, H., K. Kuranda, and F. Mingozzi, *AAV Vector Immunogenicity in Humans: A Long Journey to Successful Gene Transfer.* Mol Ther, 2020. **28**(3): p. 723-746.
92. Poupiot, J., et al., *Role of Regulatory T Cell and Effector T Cell Exhaustion in Liver-Mediated Transgene Tolerance in Muscle.* Mol Ther Methods Clin Dev, 2019. **15**: p. 83-100.
93. Mendell, J.R., et al., *Dystrophin immunity in Duchenne's muscular dystrophy.* N Engl J Med, 2010. **363**(15): p. 1429-37.
94. Bonnemann, C.G., et al., *Dystrophin Immunity after Gene Therapy for Duchenne's Muscular Dystrophy.* N Engl J Med, 2023. **388**(24): p. 2294-2296.
95. Kay, M.A., *AAV vectors and tumorigenicity.* Nat Biotechnol, 2007. **25**(10): p. 1111-3.
96. Deyle, D.R. and D.W. Russell, *Adeno-associated virus vector integration.* Curr Opin Mol Ther, 2009. **11**(4): p. 442-7.
97. Russell, D.W., *AAV loves an active genome.* Nat Genet, 2003. **34**(3): p. 241-2.
98. 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.
99. Donsante, A., et al., *AAV vector integration sites in mouse hepatocellular carcinoma.* Science, 2007. **317**(5837): p. 477.
100. Muhuri, M., et al., *Durability of transgene expression after rAAV gene therapy.* Mol Ther, 2022. **30**(4): p. 1364-1380.
101. Nguyen, G.N., et al., *A long-term study of AAV gene therapy in dogs with hemophilia A identifies clonal expansions of transduced liver cells.* Nat Biotechnol, 2021. **39**(1): p. 47-55.
102. Nathwani, A.C., et al., *Adeno-Associated Mediated Gene Transfer for Hemophilia B:8 Year Follow up and Impact of Removing "Empty Viral Particles" on Safety and Efficacy of Gene Transfer.* Blood, 2018. **132**(Supplement 1): p. 491-491.
103. Yamaguchi, T., et al., *Aspects of Gene Therapy Products Using Current Genome-Editing Technology in Japan.* Hum Gene Ther, 2020. **31**(19-20): p. 1043-1053.

104. Raguram, A., S. Banskota, and D.R. Liu, *Therapeutic in vivo delivery of gene editing agents.* Cell, 2022. **185**(15): p. 2806-2827.
105. Payne, R.M., *Gene therapy for Friedreich ataxia: Too much, too little, or just right?* Mol Ther Methods Clin Dev, 2022. **25**: p. 1-2.
106. Palmieri, M., D. Pozzer, and N. Landsberger, *Advanced genetic therapies for the treatment of Rett syndrome: state of the art and future perspectives.* Front Neurosci, 2023. **17**: p. 1172805.
107. Booth, B.J., et al., *RNA editing: Expanding the potential of RNA therapeutics.* Mol Ther, 2023. **31**(6): p. 1533-1549.
108. Gray, S.J., et al., *Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors.* Hum Gene Ther, 2011. **22**(9): p. 1143-53.
109. Papadakis, E.D., et al., *Promoters and control elements: designing expression cassettes for gene therapy.* Curr Gene Ther, 2004. **4**(1): p. 89-113.
110. Chai, S., et al., *Strong ubiquitous micro-promoters for recombinant adeno-associated viral vectors.* Mol Ther Methods Clin Dev, 2023. **29**: p. 504-512.
111. Fornes, O., et al., *OnTarget: in silico design of MiniPromoters for targeted delivery of expression.* Nucleic Acids Res, 2023. **51**(W1): p. W379-W386.
112. Suoranta, T., N. Laham-Karam, and S. Yla-Herttuala, *Strategies to improve safety profile of AAV vectors.* Front Mol Med, 2022. **2**: p. 1054069.
113. Assaf, B.T., *Systemic Toxicity of Recombinant Adeno-Associated Virus Gene Therapy Vectors.* Toxicol Pathol, 2024: p. 1926233241298892.
114. Choi, J.H., et al., *Optimization of AAV expression cassettes to improve packaging capacity and transgene expression in neurons.* Mol Brain, 2014. **7**(1): p. 17.
115. Kuzmin, D.A., et al., *The clinical landscape for AAV gene therapies.* Nat Rev Drug Discov, 2021. **20**(3): p. 173-174.
116. Au, H.K.E., M. Isalan, and M. Mielcarek, *Gene Therapy Advances: A Meta-Analysis of AAV Usage in Clinical Settings.* Front Med (Lausanne), 2021. **8**: p. 809118.
117. Nieuwenhuis, B., et al., *Improving adeno-associated viral (AAV) vector-mediated transgene expression in retinal ganglion cells: comparison of five promoters.* Gene Ther, 2023. **30**(6): p. 503-519.

118. Dong, J.Y., P.D. Fan, and R.A. Frizzell, *Quantitative analysis of the packaging capacity of recombinant adeno-associated virus.* Hum Gene Ther, 1996. **7**(17): p. 2101-12.
119. Khabou, H., et al., *Dosage Thresholds and Influence of Transgene Cassette in Adeno-Associated Virus-Related Toxicity.* Hum Gene Ther, 2018. **29**(11): p. 1235-1241.
120. Qin, J.Y., et al., *Systematic comparison of constitutive promoters and the doxycycline-inducible promoter.* PLoS One, 2010. **5**(5): p. e10611.
121. Wagner, H.J., W. Weber, and M. Fussenegger, *Synthetic Biology: Emerging Concepts to Design and Advance Adeno-Associated Viral Vectors for Gene Therapy.* Adv Sci (Weinh), 2021. **8**(9): p. 2004018.
122. Van Alstyne, M., et al., *Gain of toxic function by long-term AAV9-mediated SMN overexpression in the sensorimotor circuit.* Nat Neurosci, 2021. **24**(7): p. 930-940.
123. Brooks, A.R., et al., *Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle.* J Gene Med, 2004. **6**(4): p. 395-404.
124. de Leeuw, C.N., et al., *rAAV-compatible MiniPromoters for restricted expression in the brain and eye.* Mol Brain, 2016. **9**(1): p. 52.
125. Fujita, K., et al., *Spatially and Temporally Regulated NRF2 Gene Therapy Using Mcp-1 Promoter in Retinal Ganglion Cell Injury.* Mol Ther Methods Clin Dev, 2017. **5**: p. 130-141.
126. Xiong, W., et al., *AAV cis-regulatory sequences are correlated with ocular toxicity.* Proc Natl Acad Sci U S A, 2019. **116**(12): p. 5785-5794.
127. Zheng, C. and B.J. Baum, *Evaluation of promoters for use in tissue-specific gene delivery.* Methods Mol Biol, 2008. **434**: p. 205-19.
128. Zheng, C. and B.J. Baum, *Evaluation of viral and mammalian promoters for use in gene delivery to salivary glands.* Mol Ther, 2005. **12**(3): p. 528-36.
129. Greenshpan, Y., et al., *The Contribution of the Minimal Promoter Element to the Activity of Synthetic Promoters Mediating CAR Expression in the Tumor Microenvironment.* Int J Mol Sci, 2022. **23**(13): p. 7431.

130. Vilaboa, N. and R. Voellmy, *Regulatable gene expression systems for gene therapy.* Curr Gene Ther, 2006. **6**(4): p. 421-38.
131. Boulaire, J., P. Balani, and S. Wang, *Transcriptional targeting to brain cells: Engineering cell type-specific promoter containing cassettes for enhanced transgene expression.* Adv Drug Deliv Rev, 2009. **61**(7-8): p. 589-602.
132. Nikolov, D.B. and S.K. Burley, *RNA polymerase II transcription initiation: a structural view.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 15-22.
133. Roy, A.L. and D.S. Singer, *Core promoters in transcription: old problem, new insights.* Trends Biochem Sci, 2015. **40**(3): p. 165-71.
134. Niwa, H., K. Yamamura, and J. Miyazaki, *Efficient selection for high-expression transfectants with a novel eukaryotic vector.* Gene, 1991. **108**(2): p. 193-9.
135. Xu, L., et al., *CMV-beta-actin promoter directs higher expression from an adeno-associated viral vector in the liver than the cytomegalovirus or elongation factor 1 alpha promoter and results in therapeutic levels of human factor X in mice.* Hum Gene Ther, 2001. **12**(5): p. 563-73.
136. Barnhart, K.M., et al., *Enhancer and promoter chimeras in plasmids designed for intramuscular injection: a comparative in vivo and in vitro study.* Hum Gene Ther, 1998. **9**(17): p. 2545-53.
137. Hagstrom, J.N., et al., *Improved muscle-derived expression of human coagulation factor IX from a skeletal actin/CMV hybrid enhancer/promoter.* Blood, 2000. **95**(8): p. 2536-42.
138. Wang, C.Y., et al., *Improved neuronal transgene expression from an AAV-2 vector with a hybrid CMV enhancer/PDGF-beta promoter.* J Gene Med, 2005. **7**(7): p. 945-55.
139. Segert, J.A., S.S. Gisselbrecht, and M.L. Bulyk, *Transcriptional Silencers: Driving Gene Expression with the Brakes On.* Trends Genet, 2021. **37**(6): p. 514-527.
140. Friedman, R.Z., et al., *Information content differentiates enhancers from silencers in mouse photoreceptors.* Elife, 2021. **10**.
141. Gisselbrecht, S.S., et al., *Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts.* Mol Cell, 2020. **77**(2): p. 324-337 e8.

142. Lee, T.I. and R.A. Young, *Transcriptional regulation and its misregulation in disease.* Cell, 2013. **152**(6): p. 1237-51.
143. Lambert, S.A., et al., *The Human Transcription Factors.* Cell, 2018. **172**(4): p. 650-665.
144. Soto, L.F., et al., *Compendium of human transcription factor effector domains.* Mol Cell, 2022. **82**(3): p. 514-526.
145. Burdach, J., et al., *Regions outside the DNA-binding domain are critical for proper in vivo specificity of an archetypal zinc finger transcription factor.* Nucleic Acids Res, 2014. **42**(1): p. 276-89.
146. Zhou, Q., et al., *A mouse tissue transcription factor atlas.* Nat Commun, 2017. **8**: p. 15089.
147. Yu, X., et al., *Computational analysis of tissue-specific combinatorial gene regulation: predicting interaction between transcription factors in human tissues.* Nucleic Acids Res, 2006. **34**(17): p. 4925-36.
148. Laranjeiro, R. and D. Whitmore, *Transcription factors involved in retinogenesis are co-opted by the circadian clock following photoreceptor differentiation.* Development, 2014. **141**(13): p. 2644-56.
149. Xu, Y.Z., et al., *Promoter deletion analysis using a dual-luciferase reporter system.* Methods Mol Biol, 2013. **977**: p. 79-93.
150. Johnson, A.O., et al., *Bioinformatic Design of Dendritic Cell-Specific Synthetic Promoters.* ACS Synth Biol, 2022. **11**(4): p. 1613-1626.
151. Ertl, H.C.J., *Immunogenicity and toxicity of AAV gene therapy.* Front Immunol, 2022. **13**: p. 975803.
152. Goverdhana, S., et al., *Regulatable gene expression systems for gene therapy applications: progress and future challenges.* Mol Ther, 2005. **12**(2): p. 189-211.
153. Gould, D.J. and Y. Chernajovsky, *Application of regulatable vectors in gene therapy of experimental arthritis.* Arthritis Research & Therapy, 2004. **6**(Suppl 1): p. 83.
154. Byun, G., J. Yang, and S.W. Seo, *CRISPRi-mediated tunable control of gene expression level with engineered single-guide RNA in Escherichia coli.* Nucleic Acids Res, 2023. **51**(9): p. 4650-4659.

155. Li, X.T., et al., *tCRISPRi: tunable and reversible, one-step control of gene expression.* Sci Rep, 2016. **6**: p. 39076.
156. Fontana, J., et al., *Regulated Expression of sgRNAs Tunes CRISPRi in E. coli.* Biotechnol J, 2018. **13**(9): p. e1800069.
157. Lipinski, D.M., *A Comparison of Inducible Gene Expression Platforms: Implications for Recombinant Adeno-Associated Virus (rAAV) Vector-Mediated Ocular Gene Therapy.* Adv Exp Med Biol, 2019. **1185**: p. 79-83.
158. Baron, U. and H. Bujard, *Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances.* Methods Enzymol, 2000. **327**: p. 401-21.
159. Zhou, X., et al., *Optimization of the Tet-On system for regulated gene expression through viral evolution.* Gene Ther, 2006. **13**(19): p. 1382-90.
160. Das, A.T., L. Tenenbaum, and B. Berkhout, *Tet-On Systems For Doxycycline-inducible Gene Expression.* Curr Gene Ther, 2016. **16**(3): p. 156-67.
161. Le Guiner, C., et al., *Transgene regulation using the tetracycline-inducible TetR-KRAB system after AAV-mediated gene transfer in rodents and nonhuman primates.* PLoS One, 2014. **9**(9): p. e102538.
162. Chenuaud, P., et al., *Optimal design of a single recombinant adeno-associated virus derived from serotypes 1 and 2 to achieve more tightly regulated transgene expression from nonhuman primate muscle.* Mol Ther, 2004. **9**(3): p. 410-8.
163. Ingusci, S., et al., *Gene Therapy Tools for Brain Diseases.* Front Pharmacol, 2019. **10**: p. 724.
164. Kallunki, T., et al., *How to Choose the Right Inducible Gene Expression System for Mammalian Studies?* Cells, 2019. **8**(8).
165. Fussenegger, M., *The impact of mammalian gene regulation concepts on functional genomic research, metabolic engineering, and advanced gene therapies.* Biotechnol Prog, 2001. **17**(1): p. 1-51.
166. Chang, A.L., J.J. Wolf, and C.D. Smolke, *Synthetic RNA switches as a tool for temporal and spatial control over gene expression.* Curr Opin Biotechnol, 2012. **23**(5): p. 679-88.

167. Chi, Y., et al., *A compact, versatile drug-induced splicing switch system with minimal background expression.* Cell Rep Methods, 2024. **4**(9): p. 100842.
168. Licatalosi, D.D., *Roles of RNA-binding Proteins and Post-transcriptional Regulation in Driving Male Germ Cell Development in the Mouse.* Adv Exp Med Biol, 2016. **907**: p. 123-51.
169. Aagaard, L. and J.J. Rossi, *RNAi therapeutics: principles, prospects and challenges.* Adv Drug Deliv Rev, 2007. **59**(2-3): p. 75-86.
170. Mattick, J.S. and I.V. Makunin, *Non-coding RNA.* Hum Mol Genet, 2006. **15 Spec No 1**: p. R17-29.
171. Bertrand, J.R., et al., *Comparison of antisense oligonucleotides and siRNAs in cell culture and in vivo.* Biochem Biophys Res Commun, 2002. **296**(4): p. 1000-4.
172. Grimm, D., et al., *Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways.* Nature, 2006. **441**(7092): p. 537-41.
173. Marques, J.T. and B.R. Williams, *Activation of the mammalian immune system by siRNAs.* Nat Biotechnol, 2005. **23**(11): p. 1399-405.
174. Yokobayashi, Y., *Aptamer-based and aptazyme-based riboswitches in mammalian cells.* Curr Opin Chem Biol, 2019. **52**: p. 72-78.
175. Tickner, Z.J. and M. Farzan, *Riboswitches for Controlled Expression of Therapeutic Transgenes Delivered by Adeno-Associated Viral Vectors.* Pharmaceuticals (Basel), 2021. **14**(6).
176. Khorkova, O., et al., *Amplifying gene expression with RNA-targeted therapeutics.* Nat Rev Drug Discov, 2023. **22**(7): p. 539-561.
177. Zhong, G., et al., *A reversible RNA on-switch that controls gene expression of AAV-delivered therapeutics in vivo.* Nat Biotechnol, 2020. **38**(2): p. 169-175.
178. Hallberg, Z.F., et al., *Engineering and In Vivo Applications of Riboswitches.* Annu Rev Biochem, 2017. **86**: p. 515-539.
179. Aquino-Jarquin, G., *Novel Engineered Programmable Systems for ADAR-Mediated RNA Editing.* Mol Ther Nucleic Acids, 2020. **19**: p. 1065-1072.
180. Shiromoto, Y., et al., *ADAR1 RNA editing enzyme regulates R-loop formation and genome stability at telomeres in cancer cells.* Nat Commun, 2021. **12**(1): p. 1654.

181. Han, R.C. and R.E. MacLaren, *RNA gene editing in the eye and beyond: The neglected tool of the gene editing armatorium?* Int Rev Cell Mol Biol, 2022. **372**: p. 175-205.
182. Matthews, M.M., et al., *Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity.* Nat Struct Mol Biol, 2016. **23**(5): p. 426-33.
183. Khosravi, H.M. and M.F. Jantsch, *Site-directed RNA editing: recent advances and open challenges.* RNA Biol, 2021. **18**(sup1): p. 41-50.
184. Katrekar, D., et al., *Efficient in vitro and in vivo RNA editing via recruitment of endogenous ADARs using circular guide RNAs.* Nat Biotechnol, 2022. **40**(6): p. 938-945.
185. Bennett, J., *Taking Stock of Retinal Gene Therapy: Looking Back and Moving Forward.* Mol Ther, 2017. **25**(5): p. 1076-1094.
186. MacLaren, R.E., et al., *Retinal gene therapy in patients with choroideremia: initial findings from a phase 1/2 clinical trial.* Lancet, 2014. **383**(9923): p. 1129-37.
187. Corbo, J.C., *The role of cis-regulatory elements in the design of gene therapy vectors for inherited blindness.* Expert Opin Biol Ther, 2008. **8**(5): p. 599-608.
188. Lee, J., et al., *Quantitative fine-tuning of photoreceptor cis-regulatory elements through affinity modulation of transcription factor binding sites.* Gene Ther, 2010. **17**(11): p. 1390-9.
189. Merentie, M., et al., *Efficacy and safety of myocardial gene transfer of adenovirus, adeno-associated virus and lentivirus vectors in the mouse heart.* Gene Ther, 2016. **23**(3): p. 296-305.
190. Klein, R.L., et al., *Efficient neuronal gene transfer with AAV8 leads to neurotoxic levels of tau or green fluorescent proteins.* Mol Ther, 2006. **13**(3): p. 517-27.
191. Watakabe, A., et al., *Comparative analyses of adeno-associated viral vector serotypes 1, 2, 5, 8 and 9 in marmoset, mouse and macaque cerebral cortex.* Neurosci Res, 2015. **93**: p. 144-57.
192. Smith, A.J., *New horizons in therapeutic antibody discovery: opportunities and challenges versus small-molecule therapeutics.* J Biomol Screen, 2015. **20**(4): p. 437-53.

193. Cabrera, A., et al., *The sound of silence: Transgene silencing in mammalian cell engineering.* Cell Syst, 2022. **13**(12): p. 950-973.
194. Simpson, E.M., et al., *New MiniPromoter Ple345 (NEFL) Drives Strong and Specific Expression in Retinal Ganglion Cells of Mouse and Primate Retina.* Hum Gene Ther, 2019. **30**(3): p. 257-272.
195. Millington-Ward, S., et al., *Novel 199 base pair NEFH promoter drives expression in retinal ganglion cells.* Sci Rep, 2020. **10**(1): p. 16515.
196. Hanlon, K.S., et al., *A Novel Retinal Ganglion Cell Promoter for Utility in AAV Vectors.* Front Neurosci, 2017. **11**: p. 521.
197. Hulliger, E.C., S.M. Hostettler, and S. Kleinlogel, *Empowering Retinal Gene Therapy with a Specific Promoter for Human Rod and Cone ON-Bipolar Cells.* Mol Ther Methods Clin Dev, 2020. **17**: p. 505-519.
198. Korecki, A.J., et al., *Human MiniPromoters for ocular-rAAV expression in ON bipolar, cone, corneal, endothelial, Muller glial, and PAX6 cells.* Gene Ther, 2021. **28**(6): p. 351-372.
199. Merlin, S. and A. Follenzi, *Transcriptional Targeting and MicroRNA Regulation of Lentiviral Vectors.* Mol Ther Methods Clin Dev, 2019. **12**: p. 223-232.
200. Valinezhad Orang, A., R. Safaralizadeh, and M. Kazemzadeh-Bavili, *Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation.* Int J Genomics, 2014. **2014**: p. 970607.
201. Salva, M.Z., et al., *Design of tissue-specific regulatory cassettes for high-level rAAV-mediated expression in skeletal and cardiac muscle.* Mol Ther, 2007. **15**(2): p. 320-9.
202. Datta, N., et al., *MicroRNA-based therapeutics for inflammatory disorders of the microbiota-gut-brain axis.* Pharmacol Res, 2023. **194**: p. 106870.
203. Drag, S., F. Dotiwala, and A.K. Upadhyay, *Gene Therapy for Retinal Degenerative Diseases: Progress, Challenges, and Future Directions.* Invest Ophthalmol Vis Sci, 2023. **64**(7): p. 39.
204. Tian, B., et al., *Ocular Drug Delivery: Advancements and Innovations.* Pharmaceutics, 2022. **14**(9).

205. Xu, Z., et al., *Trans-splicing adeno-associated viral vector-mediated gene therapy is limited by the accumulation of spliced mRNA but not by dual vector coinfection efficiency.* Hum Gene Ther, 2004. **15**(9): p. 896-905.
206. Riedmayr, L.M., et al., *mRNA trans-splicing dual AAV vectors for (epi)genome editing and gene therapy.* Nat Commun, 2023. **14**(1): p. 6578.
207. Trapani, I., *Adeno-Associated Viral Vectors as a Tool for Large Gene Delivery to the Retina.* Genes (Basel), 2019. **10**(4).
208. Barrett, L.W., S. Fletcher, and S.D. Wilton, *Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements.* Cell Mol Life Sci, 2012. **69**(21): p. 3613-34.
209. Rose, A.B., *Introns as Gene Regulators: A Brick on the Accelerator.* Front Genet, 2018. **9**: p. 672.
210. Chorev, M. and L. Carmel, *The function of introns.* Front Genet, 2012. **3**: p. 55.
211. Lee, T., M.E. Bradley, and J.L. Walowitz, *Influence of promoter potency on the transcriptional effects of YY1, SRF and Msx-1 in transient transfection analysis.* Nucleic Acids Res, 1998. **26**(13): p. 3215-20.
212. Maturana, C.J., *Engineered compact pan-neuronal promoter from Alphaherpesvirus LAP2 enhances target gene expression in the mouse brain and reduces tropism in the liver.* Gene Ther, 2024. **31**(5-6): p. 335-344.
213. Zimmerman, D., et al., *Enhancement of transgene expression by nuclear transcription factor Y and CCCTC-binding factor.* Biotechnol Prog, 2018. **34**(6): p. 1581-1588.
214. Buck, T.M. and J. Wijnholds, *Recombinant Adeno-Associated Viral Vectors (rAAV)-Vector Elements in Ocular Gene Therapy Clinical Trials and Transgene Expression and Bioactivity Assays.* Int J Mol Sci, 2020. **21**(12): p. 4197.
215. Steri, M., et al., *Genetic variants in mRNA untranslated regions.* Wiley Interdiscip Rev RNA, 2018. **9**(4): p. e1474.
216. Mignone, F., et al., *Untranslated regions of mRNAs.* Genome Biol, 2002. **3**(3): p. REVIEWS0004.
217. Pichon, X., et al., *RNA binding protein/RNA element interactions and the control of translation.* Curr Protein Pept Sci, 2012. **13**(4): p. 294-304.

218. Subramanian, M., et al., *RNAi-mediated rheostat for dynamic control of AAV-delivered transgenes.* Nat Commun, 2023. **14**(1): p. 1970.
219. Di Ruscio, A. and V. de Franciscis, *Minding the gap: Unlockingthe therapeutic potentialof aptamers and making up for lost time.* Mol Ther Nucleic Acids, 2022. **29**: p. 384-386.
220. Doshi, A., et al., *Small-molecule inducible transcriptional control in mammalian cells.* Crit Rev Biotechnol, 2020. **40**(8): p. 1131-1150.
221. Walkley, C.R. and J.B. Li, *Rewriting the transcriptome: adenosine-to-inosine RNA editing by ADARs.* Genome Biol, 2017. **18**(1): p. 205.
222. Sinnamon, J.R., et al., *Site-directed RNA repair of endogenous Mecp2 RNA in neurons.* Proc Natl Acad Sci U S A, 2017. **114**(44): p. E9395-E9402.
223. Zinshteyn, B. and K. Nishikura, *Adenosine-to-inosine RNA editing.* Wiley Interdiscip Rev Syst Biol Med, 2009. **1**(2): p. 202-209.
224. Sato, H. and R.H. Singer, *Cellular variability of nonsense-mediated mRNA decay.* Nat Commun, 2021. **12**(1): p. 7203.
225. Kurosaki, T., M.W. Popp, and L.E. Maquat, *Quality and quantity control of gene expression by nonsense-mediated mRNA decay.* Nat Rev Mol Cell Biol, 2019. **20**(7): p. 406-420.
226. Bellingrath, J.S., et al., *Programmable RNA editing with endogenous ADAR enzymes - a feasible option for the treatment of inherited retinal disease?* Front Mol Neurosci, 2023. **16**: p. 1092913.
227. Wettengel, J., et al., *Harnessing human ADAR2 for RNA repair - Recoding a PINK1 mutation rescues mitophagy.* Nucleic Acids Res, 2017. **45**(5): p. 2797-2808.
228. Katrekar, D., et al., *Robust RNA editing via recruitment of endogenous ADARs using circular guide RNAs.* bioRxiv, 2021: p. 2021.01.12.426286.
229. Kim, Y.K. and L.E. Maquat, *UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond.* RNA, 2019. **25**(4): p. 407-422.
230. Nishikura, K., *Functions and regulation of RNA editing by ADAR deaminases.* Annu Rev Biochem, 2010. **79**: p. 321-49.
231. Ai, X., et al., *Enhancing RNA editing efficiency and specificity with engineered ADAR2 guide RNAs.* Mol Ther Nucleic Acids, 2025. **36**(1): p. 102447.

232. Litke, J.L. and S.R. Jaffrey, *Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts.* Nat Biotechnol, 2019. **37**(6): p. 667-675.