Characterisation of adeno-associated virus vector preparation DNA contaminants and development of a high purity, P5 promoter replacement adeno-associated virus production system

Mr Mark Brimble

Submitted for the Degree of Doctor of Philosophy,

Department of Haematology

University College London Cancer Institute

Declaration:

I, Mark Brimble, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I have indicated this in the thesis.

Acknowledgements:

Firstly, I would like to thank my supervisors Professor Amit Nathwani and Dr Andrew Davidoff for their support, advice and trusting me to pursue a research topic that initially did not look likely to bear significant fruit. In addition, I would like to thank John Gray for his invaluable advice on project direction, and the members of the Davidoff lab for all their assistance in the course of these studies. I would like to thank my colleagues past and present at both UCL and St Jude Children's Research Hospital who have taught me what it means to be a scientist and the way to think about scientific problems.

I would especially like to thank my family and friends. My late Grandfather John Allen, whose embellished stories I will forever miss. My parents for raising me to have the mental wherewithal to deal with tough challenges. My brother and sister for their love and exquisite sense of humour. Rohan Bundell, Sammy Anzer and Mike Atkinson for their unending friendship, their moral fortitude only matched by their ridiculousness, and their much-needed help during my grandfather's passing and my father's illness.

Finally, the most important thanks go to my wife Laurin, the only person I would want by my side for such an undertaking. She has taught me far more than a lifetime of scientific research ever could, and I am a far better person for it. Her support is the only reason this thesis is here in its completed form.

Thesis Abstract

Gene therapy is the transfer of nucleic acids for therapeutic benefit, a process potentially achieved through several delivery methods. One of the most common gene therapy delivery modalities is adeno associated virus (AAV). The availability of a system to produce large quantities of AAV, and of diverse AAV serotypes, that can infect different cell types makes AAV very attractive for widespread clinical implementation. Viruses like AAV that are manipulated to become highly useful delivery tools rely on complex biological systems for production. It is imperative that these systems are well characterised and designed to be highly safe and efficacious.

This PhD thesis investigated nucleic acid contamination of AAV preparations, looking at the production system as a source, and the problems that this may pose for AAV infected cells. Specific regions of DNA from AAV producer plasmids were incorporated into virions more abundantly. These were found adjacent to the outside of the inverted terminal repeats (ITRs) that flank the expression cassette, and on the plasmid containing the replication and capsid genes, directly upstream of the P5 promoter which drives expression of the replication proteins, REP78 and REP68. Contaminant sequences were found to be transferred into transduced cells during AAV infection, persistent within cells and transcriptionally active. Investigation of P5 upstream contaminants showed potential for protein to be produced persistently from contaminant sequences both in vitro and in vivo.

Finally, precise sequence regions causing DNA incorporation upstream of the P5 promoter were examined. This information was used to produce alternative production

reagents that, in their first iteration, provided proof of principle of P5-related contaminant removal from the final AAV product, and in additional iterations resulted in a modified P5 promoter (P5-HS). This promoter limited upstream contamination thereby resulting in higher purity AAV, whilst retaining vector yields equivalent to the original AAV production system.

Impact Statement:

The finding that DNA contaminants within AAV preps can be transcribed and translated post infection highlights a previously undescribed safety issue with AAV gene therapy. The plasmid-based transfection system used to produce AAV is standard in the field. In the academic research setting this contamination could result in interference with experimental results through unintended downstream effects and interactions caused by any RNA or protein products resulting from the DNA contamination. In the clinical setting the effect could be more severe, as AAV is being administered directly to patients in an ever-growing number of clinical gene therapy trials. The findings in this thesis regarding the profile and activity of DNA contaminants in AAV illustrate why those using the virus should take measures to characterise the purity of their products in greater detail and implement all available methods to reduce production plasmid-derived impurities from AAV preparations.

The P5 spacer AAV production system (P5-HS) developed within this thesis to remove contamination upstream of the P5 promoter can be used to produce AAV of any serotype and protein of interest, thus providing a universal tool that can be used to improve the

purity and safety of any AAV gene therapy product. P5-HS also produces equivalent titers of AAV to standard production reagents. This retention of viral vector yield is important because very high doses of AAV are required to treat patients, and commercial entities will not wish to sacrifice production capabilities when making changes to their systems. P5-HS can be easily implemented in place of the current producer plasmids without the need to change any other aspect of the production system, keeping the cost of implementation to a minimum. P5-HS provides a means to take a significant step towards the FDA guidelines of reducing the amount of administered non-expression cassette DNA in an AAV gene therapy product to under 10ng per dose, the standard that has been in place for vaccines since 1997.[1]

Publications during PhD studies:

Brimble MA, Reiss UM, Nathwani AC, Davidoff AM. 2016. New and improved AAVenues: current status of hemophilia B gene therapy. *Expert Opin Biol Ther*. 16(1):79-92

Stefanelli G, Azam AB, Walters BJ, **Brimble MA**, Gettens CP, Bouchard-Cannon P, Cheng HM, Davidoff AM, Narkaj K, Day JJ, Kennedy AJ, Zovkic IB. 2018. Learning and Age-Related Changes in Genome-wide H2A.Z Binding in the Mouse Hippocampus. *Cell Rep*. 22(5):1124-1131.

Woods VMA, **Brimble MA**. 2018. Spring Forward: ESGCT Trains the Next Generation of Gene and Cell Therapists. *Hum Gene Ther.* 2018. 29(10):1074-1075.

Hu D, Jablonowski C, Cheng PH, AlTahan A, Li C, Wang Y, Palmer L, Lan C, Sun B, Abu-Zaid A, Fan Y, **Brimble M**, Gamboa NT, Kumbhar RC, Yanishevski D, Miller KM, Kang G, Zambetti GP, Chen T, Yan Q, Davidoff AM, Yang J. 2018. KDM5A Regulates a Translational Program that Controls p53 Protein Expression. *iScience*. *9*, 84-100.

Casey G, Askew C, **Brimble MA**, Samulski RJ, Davidoff AM, Li C, Walters BJ. 2020. Self-complementarity in adeno-associated virus enhances transduction and gene expression in mouse cochlear tissues. *PLoS One*. 15(11):e0242599.

*The histone variant macroH2A1, not macroH2A2, preferentially regulates learning-induced gene expression and memory formation in mice. Narkaj K, Singh G, **Brimble MA**, Stefanelli G, Hall M, Mitchnick KA, Creighton SD, Ianov L, Winters B, Walters B, Davidoff AM, Mitchell J, Zovkic IB. (Under revision at Nature communications)*

*Stefanelli G, Makowski C, **Brimble MA**, Hall, M, Reda A, Creighton SD, Baumbach J1, Greer CB, Davidoff AM, Walters BJ, Murphy PJ, Zovkic IB. The histone chaperone Anp32e regulates memory formation, transcription, and dendritic morphogenesis by regulating steady-state H2A.Z binding in neurons. (Under revision at Cell Reports)*

Altered P5 promoter vectors reduce the presence of transcriptionally active DNA contaminants in adeno Associated Virus – manuscript in preparation.

Awards during PhD studies:

Brimble MA, Zhou J, Morton C, Meagher M, Nathwani AC, Gray JT, Davidoff AM. (2016). 547. AAV Preparations Contain Contamination from DNA Sequences in Production Plasmids Directly Outside of the ITRs. Molecular Therapy. 24. 10.1016/S1525-0016(16)33355-X. – ***ASGCT 2016 outstanding poster presentation award.***

Patents during PhD studies:

Recombinant P5 promoter for use in reducing DNA contamination of AAV preparations. Filed provisionally to United States Patent Office. Serial# 69/029,839

Table of Contents

List of Figures:

List of Tables:

Abbreviations:

AAP	Adeno-associated virus assembly protein
AAV	Adeno-associated virus
ADA-SCID	Adenosine deaminase severe combined immune deficiency
BHK	Baby hamster kidney fibroblasts
CAP	Capsid gene/protein
CO_2	Carbon dioxide
DNA	Deoxyribonucleic acid
DMEM	Dulbecco's Modified Eagle's Medium
FACS	Fluorescence-activated cell Sorting
FBS	Foetal bovine serum
HEK293	Human embryonic kidney 293 cells
Huh7	Human hepatocellular carcinoma Huh7 cells
IHC	Immunohistochemistry
ITI	Immune tolerance induction
KDa	Kilodaltons
Kg	Kilogram
LM02	LIM domain only protein 2
MAAP	Membrane-associated accessory protein
M	Molar
mg	Milligram
ml	Millilitre
mM	Millimolar
ng	Nanogram
NHP	Non-human primate
NLS	Nuclear localisation sequence
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
PFA	Paraformaldehyde
rcAAV	Replication competent adeno-associated virus
REP	Replication gene/protein
REPBS	REP binding site
RNA	Ribonucleic acid
SC	Self-complementary
SS	Single-stranded
SF9	Spodoptera frugiperda clonal isolate 9
S/MAR	Scaffold/matrix attachment region
SV40	Simian virus 40
VG	Vector genome
VP	Capsid viral protein
WT	Wildtype
µg	Microgram
µL	Microlitre
µM	Micromolar

Chapter 1

Introduction

1. Introduction

1.1 Introduction to gene therapy:

Gene therapy is the therapeutic delivery or alteration of nucleic acids to correct a clinical phenotype caused by a genetic disorder. The origins of gene therapy date back to the late 1960s, when Renato Dulbecco's research group identified that the transformation of SV40 infected cells was due to the stable integration of the SV40 genome into the genome of the infected cells.[2] Four years later, the term "gene therapy" was coined by Theodore Friedmann and Richard Roblin in an editorial, as a possible therapeutic option for the 1500 genetic diseases identified at that time,[3] a number that sits around 10,000 in the present day. In the years since, there has been considerable progress; however, it was not until 1989 that the first gene transfer study was carried out in humans, in which the neomycin resistance gene was transferred by Steve Rosenberg's group into T cells that were subsequently infused into a patient.[4] A year later the first gene therapy patient was treated for adenosine deaminase deficiency (ADA-SCID) with retroviral modification of their T lymphocytes to express the adenosine deaminase gene. This proof of principle study, along with other preclinical research, led to a great enthusiasm at the potential for gene therapy approaches to usher in a new era of medicine.[5,6]

However, the field of gene therapy stalled with regards to clinical translation in 1997 after the death of Jesse Gelsinger, a patient with an ornithine transcarbamylase deficiency who suffered a cytokine storm after adenoviral gene therapy treatment. Furthermore, after an initially successful ex vivo gene therapy trial for X-linked SCID using a γ-retroviral vector,

leukemogenesis was observed in a patient with a proviral insertion at the LM0-2 locus, leading to altered expression of the LM0-2 transcript.[7,8] This same insertion was later seen in other patients treated with this vector who developed leukemia, and additional proviral insertions were detected near the proto-oncogenes BMI and CCND2.[9,10] Despite these considerable setbacks, gene therapy as a treatment modality has surged ahead in recent years. After initial successful trials using to treat Leber's congenital amaurosis and haemophilia B,[11–13] enthusiasm for genet approaches to treat genetic diseases has again gathered momentum.[5] Furthermore, after efforts to improve the safety profile of retroviral vectors,[14] curative clinical success has now been observed in gene therapy treatment for X-linked SCID.[15,16] There are now gene therapy products that have received market authorization by regulatory bodies in the United States, Europe, South Korea, China and Russia, with many more seeking market approval.[17] Whilst the current proportion of patients treated for genetic disorders with gene therapy approaches is exceedingly low, there have been over 3,000 gene therapy clinical trials conducted to date.[18] It is likely that nucleic acid based therapies will continue to increase significantly over time, providing a new and invaluable paradigm to the treatment of genetic disorders.

1.2 Gene therapy delivery and treatment methods:

The most fundamental aspect of a genetic treatment is how the intended DNA cargo will be delivered. Delivery strategies will vary depending on the target cell or tissue, and the type of genetic cargo being delivered. The most common delivery tool for gene therapy is viruses. Viruses are considered useful delivery vehicles for gene therapy due to their

innate ability to infect cells with high efficiency. Naturally occurring viruses have been co-opted by researchers to serve as systems to deliver nucleic acid constructs for a wide range of clinical applications (Table 1.1). The type of virus that will be used in each treatment setting depends on the requirements of the treatment. In this section a few of the most common gene therapy delivery strategies will be outlined.

1.3 Ex vivo gene therapy:

There are certain disorders in which the most appropriate method of delivery to the target cell is performed outside of the patient, in culture, with the idea that the modified cells will be transplanted back into the patient. This is known as ex vivo gene therapy. Some of the earliest gene therapy experiments were performed in an ex vivo setting on haematopoietic progenitor cells.[19,20] Other early gene therapy studies even used ex vivo gene transfer to hepatocytes for the potential treatment of disorders like hypercholesteremia,[21,22] and to the CNS for neurological diseases.[23,24] Ex vivo gene therapy has certain advantages over its in vivo counterpart. Firstly, any genetic manipulation that is made is confined to the targeted cells. This can both increase the precision and expand the possible genetic manipulation strategies that can be employed. Furthermore, ex vivo gene therapy can often enable sampling of the treated cells to ensure the safety and success of the gene transfer prior to treating the patient.

1.3.1 HSC gene therapy

For genetic disorders of the blood; ex vivo gene delivery requires the GOI to be transferred to the haematopoietic stem cell (HSC) population of CD34+ progenitor cells.

Virus	Genome type	Clinical Gene Therapy applications
Adenoviridae	dsDNA	Oncolytic Therapy; Gene delivery
Adeno-associated virus	ssDNA	Gene delivery; Gene editing nuclease delivery
Alphavirus	ssRNA (+ sense)	Cancer Gene therapy
Herpes Simplex Virus	dsDNA	Oncolytic Virotherapy; Gene delivery
Lentivirus	ssRNA (+ sense)	Ex-vivo gene delivery
Measles Virus	ssRNA (- sense)	Oncolytic virotherapy
Newcastle Disease Virus	ssRNA (- sense)	Oncolytic virotherapy
Picornaviridae	ssRNA (+ sense)	Oncolytic virotherapy
Reoviridae	dsRNA	Oncolytic virotherapy
Retroviridae	ssRNA (+ sense)	Ex-vivo gene delivery
Vaccinia Virus	dsDNA	Oncolytic virotherapy

Table 1.1 - Viruses used in clinical gene therapy

Viruses used in clinical gene therapy. Sourced from clinicaltrials.gov
Search terms: Adenovirus Gene Therapy; AAV gene therapy; Alphavirus cancer; HSV gene therapy; Lentivirus Gene Therapy; Measles Cancer; NDV Cancer; Picornavirus cancer; Reovirus cancer; Retrovirus Gene Therapy; Vaccinia Gene Therapy

To achieve this, integration of the genetic material is required, which can be achieved by infecting the cells with an integrating lentivirus containing the GOI cassette.[25] Lentivirus is an enveloped RNA virus with the desired properties for this application due to its high propensity to infect immune cells, and efficient semi-random integration into the genome in a pattern that whilst not without risk is deemed safe enough for therapeutic application [26,27] This genomic insertion allows the genetic alteration to become propagated as the cells are expanded before infusion into the patient, and in daughter cells post-engraftment. If a non-integrating virus is applied to these cells, expression of any transgene will be lost over time,[28] a less desirable outcome for most applications in this context. Lentiviral gene therapy of HSCs has provided clinical therapeutic benefit to diseases such as X-Linked SCID, Wiskott Aldrich syndrome, Sickle Cell anemia.[16,29,30]

1.3.2 Chimeric Antigen Receptor T cells.

A further way that gene therapy has been harnessed for cancer treatment is through the development of Chimeric Antigen Receptor T cells (CAR T cells). The purpose of CAR T cell therapy is to introduce an artificial T cell receptor into immune cells of either a cancer patient or healthy donor, with specificity to a tumour antigen, so that when modified cells are reinfused the cytotoxic T cells are able to recognise the tumour and mount an effective elimination of the tumour. The artificial construct must be introduced into T cells in a manner that will be retained upon T cell division, and this is most commonly achieved by the infection of the pre infused T cells with an integrating lentivirus,[31] although other methods to ensure construct retention that avoid integration such as the incorporation of

S/MAR elements into the construct and mRNA transfection of the chimeric receptor have been developed in an effort to reduce the genomic changes made to the infused T cells.[32,33] The best known and most commonly implemented example of this is for the CD19 antigen. CD19 is an antigen expressed specifically in B cells.[34] Remarkable clinical success has been observed through targeting this antigen in different B cell-derived cancers.[35–37] In addition to this success, CAR therapies have been developed and trialled for a plethora of tumour specific antigens such as CD22, GD2, BCMA, HER2 and others.[38–42] Most CAR-T research has been directed at the treatment of blood related i.e. liquid tumours; however, some groups are tackling the more difficult task of adapting this approach for the treatment of solid tumours.[43]

1.3.3 Epidermal replacement gene therapy:

One of the most remarkable successes of an ex vivo gene therapy is the treatment of a seven-year-old child with Junctional Epidermolysis Bullosa (JEB), a severe monogenic skin disorder with limited treatment options and high mortality.[44] Keratinocytes from this patient were cultured and infected with a retroviral vector encoding the *LAMB-3* component of the laminin-332 protein and expanded into epidermal sheets which were transplanted. Almost the entire epidermis was regenerated in this manner resulting in skin with normal functional characteristics.[45] The success of this proof of principle treatment provides considerable hope for the treatment of further patients with JEB, other forms of Epidermolysis Bullosa, and other genetic skin disorders such as Xeroderma Pigmentosum and Ichthyosis.

1.4 In vivo gene therapy:

If genetic cargo being used in a gene therapy treatment is to be administered directly into a patient, there are different considerations for an appropriate strategy. For instance, whilst vectors that rely on integration to yield expression such as lentivirus are appropriate for ex vivo gene therapy, they are less widely considered for in vivo patient strategies. Non integrating DNA viruses are one modality more suited for this purpose. This is partly due to the risk that a non-targeted integrating vector would pose to the downstream development of cancer. Early trials of in vivo gene therapy focused heavily on adenovirus, a large but immunogenic DNA virus and subsequently many other DNA based viruses have been explored as a possibility for in vivo gene delivery.

1.4.1 Oncolytic viruses:

A key application of in vivo viral gene therapy is for the treatment of cancer. There are many strategies that can use elements of gene therapy to approach this problem. A direct mode for this is oncolytic viruses. The concept that viruses could be harnessed to abrogate the growth of tumours was explored in animal models in the 1940s.[46,47] These early studies relied on the injection of wildtype virus into affected animals. Advances to the understanding of viral anticancer properties were made throughout the remainder of the 20th century. It wasn't until 1991 however, that the first well cited example of genetic modification being applied to a virus for gene therapy was seen in a treatment human glioma samples with herpes simplex virus, type 1 (HSV-1), engineered to be negative for the thymidine kinase protein to reduce neurovirulence.[48] The principle behind oncolytic viral therapies is that the engineered virus will replicate specifically in tumour cells, to an extent that cells become lysed, releasing cell debris and viral antigens into the tumour microenvironment, prompting an antitumour immune response from the host.

The first oncolytic virus to make it to the clinic was an engineered herpes simplex virus, type 1 (HSV-1) vector, designed to preferentially replicate in melanoma cells,[49] and currently there are several open clinical trials using oncolytic vectors, and a wide range of viruses in the clinic for this purpose.[50] Despite this, there is only one oncolytic virus that has received market approval, Talimogene laherparepvec (T-VEC), a herpes simplex virus engineered to encode the human GM-CSF gene, and injected into inoperable melanoma,[51] although this drug is under clinical investigation in other tumour types, with promising results.[52,53]

1.4.2 Adenovirus:

The adenovirus family comprise over 85 characterised double stranded DNA viruses that range from 26-45kb in length.[54] Adenoviral based Gene transfer was one of the earliest gene therapy tools to be used clinically and was implemented in several gene therapy trials in the 1990s.[55–57] There was a lot of excitement around the use of adenovirus as a gene therapy delivery tool; however, these early clinical trials revealed that expression of the provided transgene was transient, and that successful readministration was not usually feasible.[58] Another key problem with early generation adenovirus vectors as a treatment method was the potential for an immune response against the vector. This was seen acutely following vector administration to Jesse Gelisinger in an attempted treatment of ornithine transcarbamylase deficiency, leading to multiple organ failure concomitant with elevated IL-6 and IL-10 overproduction, resulting in death.[59,60]

Subsequent developments to the adenoviral recombinant system allowed for a larger space to be freed up for packaging an expression cassette of interest and reducing the

number of viral elements retained in the infused virus; however, retention of gene expression post infection has still been difficult to achieve with adenoviral vectors, and leaky expression of the adenoviral proteins can occur.[61] 'Gutless' adenovirurses contain none of the adenoviral genes in their packaged DNA, and only retain the inverted terminal DNA sequences from the adenovirus in the transferred DNA. These vectors are less immunogenic than the original adenoviral systems and may prove to be a useful gene therapy tool in future, due to the large packaging capacity conferred.[62] The uses of adenovirus in the clinic to date include as a vaccination agent, as an oncolytic virus and as a vector for delivery of transgene cassettes.[63]

1.4.3 Adeno-associated virus:

Adeno-associated virus (AAV) is a Dependovirus of the *Parvoviridae* family. It is a highly compact single stranded DNA (ssDNA) virus of approximately 4.7kb in length and approximately 25nm in size.[64,65] AAV was discovered in 1965 as a coinfection alongside adenovirus particles.[65] AAV is one of the most common gene therapy agents primarily because it has several properties conducive to safe, long term, gene expression. First, in contrast to lentivirus and other retroviral family members, it exists primarily in an episomal state.[66] AAV can also infect both dividing and non-dividing cells, and furthermore, it has not been definitively associated with any human disease. Whilst the earliest gene therapy trials with AAV did not yield long term transgene expression,[67] AAV has in recent years been used in the clinic as a successful gene therapy agent for genetic disorders such as haemophilia, spinal muscular atrophy and Leber's congenital amaurosis,[11,68,69] and there are currently over 100 active clinical trials using AAV as a gene therapy delivery vehicle for genetic diseases as varied as Alzheimer's, metabolic disorders, and diseases of the

retina (Appendix). Many more preclinical treatments are in development. Adeno-associated virus and its production methodology are discussed at length in the subsequent sections ‘The AAV genome’ and ‘recombinant AAV production.

1.4.4 Other in vivo gene therapy delivery tools:

Other DNA viruses are being developed for gene therapy purposes. For instance, Bocavirus is a single stranded DNA virus also from the *Parvoviridae* family, but with a packaging capacity of around 6kb. This could potentially mean that it can be used in certain gene delivery contexts where an expression cassette is too large to be accommodated by an AAV vector, for instance in delivery of the CFTR gene for treatment of cystic fibrosis.[70] There are also a number on non-viral delivery methods, such as Polymersomes, silica based delivery systems, and nanoparticles.[71–73] Use of these vehicles in clinical trials is lower than that of viral mediated-gene therapy, but some trials have been conducted; in particular studies using cationic liposomes as gene delivery vehicles.[74,75]

1.5 Gene therapy for haemophilia:

1.5.1 Haemophilia:

Haemophilia is an X linked monogenic bleeding disorder that occurs in approximately 1 in 5,000 males worldwide.[76] The two primary forms of haemophilia are haemophilia A and haemophilia B, caused by mutations in the clotting cascade factor eight (FVIII) and factor nine (FIX) genes respectively.[77] A deficiency in the factor XI protein is referred to as haemophilia C, however this is not X-linked and deficiency produces a milder bleeding diathesis.[78] There are additional coagulopathies, rarer and caused by deficiencies in other

proteins of the clotting cascade such as factor five (FV), factor seven (FVII) factor ten (FX), and factor thirteen (FXIII).[79] The severity of haemophilia within a given patient is dependent upon the type of mutation within these genes, with severe haemophilia characterised by circulating plasma levels of FVIII or FIX protein at <1% of normal.[80] Patients with circulating clotting factor levels under this threshold have a very severe bleeding diathesis and may exhibit regular spontaneous haemorrhage.[77] Of all diagnosed cases of haemophilia, severe cases constitute approximately 35% of haemophilia A cases and 29% of haemophilia B cases, with a severe incidence across both forms of around 7.1 per million people on a population basis.[81]

1.5.2 Treatment of Haemophilia:

The first references to haemophilia exist in texts from ancient Egypt and second century Talmudic writings.[82] If left entirely untreated, patients with severe haemophilia have a life expectancy of approximately 10 years.[83] The first known example of efficacious treatment for haemophilia was in 1840 when an 11 year old child was treated for postoperative bleeding via a blood transfusion.[84] Whilst modest amounts of plasma product were produced and administered to patients in the 1940s and 50s following a fractionation protocol developed in 1946,[85] considerable improvements in life expectancy were achieved after the discovery of cryoprecipitate, a precipitated isolate of slowly thawed blood plasma, in 1964.[86] In the early 1980s blood derived products contaminated with HIV and hepatitis C had devastating effects on recipients, with studies showing a large proportion of the haemophilia population of both haemophilia A and B patients receiving factor concentrate seropositive for HIV proteins.[87,88] Fortunately, the identification and recombinant cloning of the genes for normal clotting factors FVIII and FIX was completed

at around this time.[89–91] This allowed FVIII and FIX proteins to be produced in in vitro settings instead of being isolated from blood products. As such recombinant FVIII and FIX protein became rapidly implemented as the standard treatment option for haemophilia A and haemophilia B patients, respectively. In the present day, screening of blood products is much more stringent, and use of cryoprecipitate is only recommended by the World Federation of haemophilia in the absence of clotting factor concentrates and if the cryoprecipitate has undergone viral inactivation.[92]

Recombinant factor for haemophilia treatment can be administered either in response to bleeding events (on demand), or prophylactically. Whereas on demand treatment requires less recombinant factor, prophylactic administration is designed to retain a high enough circulating factor level to initiate clotting upon injury. The current standard of care for severe haemophilia in developed nations is the prophylactic administration of recombinant factor.[93] The benefits of prophylactic therapy versus as needed infusion are seen in the later life of the patient, with a lower rate and later onset of joint arthropathies.[94,95] However, prophylaxis is an expensive undertaking, and haemophilia treatment is estimated to cost around 200,000 euros per patient per year in Europe, and $150,000 per year in the United States.[96,97] This is despite many haemophilia patients in these countries only receiving intravenous infusions of recombinant factor in response to bleeding events instead of a prophylactic regimen.[98,99] The clinical standard in many developing nations is far from ideal, with only 30% of those diagnosed having access to clotting factor replacement therapy.[100] Furthermore, despite the effectiveness of recombinant protein infusions, around 30% of severe haemophilia A and 3% of haemophilia B patients develop inhibitors to the recombinant protein,[101] which requires

more complicated treatment strategies such as immune tolerance induction (ITI) and results in higher levels of morbidity and mortality.[102,103]

A significant issue regarding the treatment of haemophilia is the relatively short half-lives of the proteins involved, at approximately 12 hours in adults for FVIII and around 24 hours for FIX.[77] This means that to retain therapeutic levels of circulating protein, patients need to receive recombinant protein infusions at regular intervals that can be upwards of three times per week.[104] There have been efforts made to improve the efficiency of prophylactic therapy; fusion proteins have been generated to heighten the stability of the infused coagulation factors, resulting in products with extended half-lives.[105] One half-life extension strategy has been to fuse the biologically stable albumin protein to FIX and FVIIa.[106] Other innovations in haemophilia treatment include the use of an engineered antibody to mimic the activity of the FVIII protein,[107] and the use of RNAi technology to reduce blood antithrombin levels for treatment of both haemophilia A and B.[108]

1.5.3 AAV gene therapy for haemophilia:

Haemophilia was earmarked early as a target disease to treat with gene therapy. This is in part because it is a single gene disorder, and because only a relatively modest amount of protein activity (around 5%) would be required to convert the disease phenotype from severe to mild, which would alleviate many of the disorder-associated symptoms. Furthermore, because these proteins are secreted into the bloodstream, there would not be the need to deliver the gene to every cell in the target tissue to confer correction, or necessarily even target the cell type the protein is usually produced in. The first attempted

gene therapy trial for haemophilia B was published in 2000 and utilised intramuscular injection of AAV2 containing a FIX cassette.[109] Expression of the delivered transgene in this trial was detectable in one patient over 3 years post infection, however, circulating levels of FIX never surpassed 1% and no clinical benefit was seen.[110] A subsequent trial reported in 2006 showed transient correction after hepatic artery injection of an AAV2 FIX vector, with expression reaching 12% in one patient. However, in this case a rise in liver transaminases was seen, concomitant with a total loss of expression, and none of the patients on this trial sustained clinical benefit, or detectable circulating FIX protein.[111] Despite this setback, the liver remained a target organ for AAV gene transfer in the treatment of haemophilia, and after the discovery that the naturally occurring serotype AAV8 had high tropism for the liver after intravenous injection in a large animal model, a far safer route to the clinic was established.[112]

1.5.4 The St Jude/ University College London haemophilia B trial:

The first successful liver directed gene therapy trial with AAV, a collaboration between St Jude Children's Research Hospital and University College London, was published in the New England Journal of Medicine in 2011.[68] At the highest dose in this cohort, a rise in liver transaminases at approximately 2 months post gene transfer was concomitant with a decline in FIX levels, this decline was halted with the administration of the steroid prednisolone, a practice now standard in liver gene therapy trials. These original patients have retained stable expression of the FIX transgene long term,[113] with the earliest infused patients now over 10 years post AAV treatment.

1.5.4.1 The SJ/UCL AAV vector design:

The AAV treatment strategy used in the SJ/UCL trial for haemophilia B used components from extensive preclinical work carried out over the years in the field. The most commonly used serotype at the time was AAV2. However, the serotype chosen for this trial was AAV8. AAV8 was discovered alongside AAV7 in samples from rhesus macaques in 2002.[114] It's full length capsid gene is 3 amino acids longer than that of AAV2 and 83% of the residues are homologous.[115] It was observed that post intravenous injection into both mice and macaques, AAV8 had strong tropism for the liver,[116,117] aided in part by the rapid uncoating of the AAV8 capsid in comparison to other serotypes.[118] Furthermore, regarding the prevalence of innate immunity to the AAV capsid, a phenomena first noted in the 1960s,[119] the AAV8 capsid has consistently shown a lower prevalence of pre-existing immunity in humans than for AAV2[120–122] For use in a recombinant production system the AAV8 capsid was used in tandem with the AAV2 REP gene and AAV2 ITRs flanking the expression cassette.

The expression cassette packaged within the ITRs contained a series of features designed to maximise therapeutic potential (Fig. 1.1): The promoter sequence (referred to hence as LP1), comprises a combination of the human alpha-1-antitrypsin (hAAT) promoter and the human apolipoprotein hepatic control region (HCR)[123] This promoter limits expression of the transgene to liver cells. Other trials have used different promoter strategies to yield the same effect. An intron from the SV40 virus was included between the promoter and FIX sequence, as studies had shown the inclusion of this to boost protein expression levels. The FIX DNA sequence was codon optimised to allow for improved expression of the transgene and is substantially different from the naturally

occurring codon pattern (Fig 1.2).[123] The protein sequence remains unchanged aside from the incorporation of a naturally occurring Threonine to Alanine polymorphism known as the FIX Malmo variant (Fig 1.3).[124] Due to the short size of the construct (~2.35kb) one of the ITRs flanking the FIX expression cassette was mutated by deletion at the TRS site. This mutation prevents genome resolution at the terminal end containing the deletion, resulting in replication reading back across the genome before resolving at the other terminal repeat. Whereas a WT configuration will yield particles with DNA containing either the positive or negative strand of the expression cassette, the TRS deletion results in the expression cassette DNA of both the positive and negative stand to be present in a single virion. These 'self-complementary' AAV genomes intramolecularly anneal and do not need to undergo second strand synthesis in the infected nucleus to mediate transgene expression and can express protein products immediately and efficiently post AAV infection.[123,125]

This vector plasmid was transiently transfected into 293T cells in a two-plasmid transfection system in which the vector genome plasmid also contained the required adenoviral helper sequences. These include the E2A and E4orf6 genes and the VA RNA coding sequence,[126] but not E1A and E1b as these are already present within the 293T cells.[127] AAV production for this trial was carried out in adherent 293T cells at large scale in cell factories and purified by a three column chromatography method.[128] This yielded a high purity product with around 20% full particles that was approved for clinical trial infusion in both the United States and the United Kingdom.

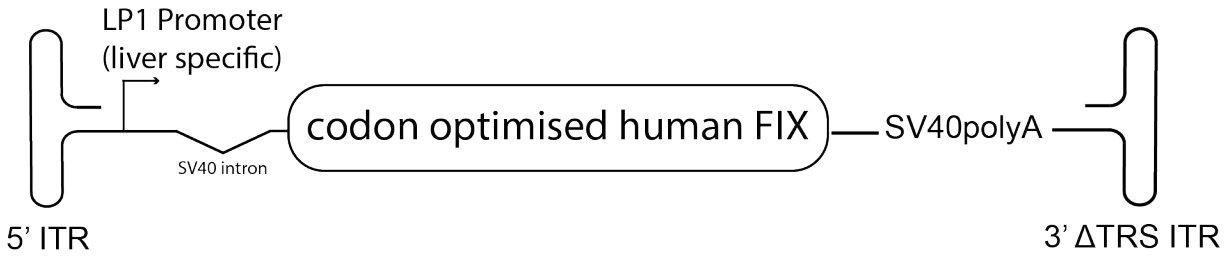


Figure 1.1: St Jude/UCL clinical trial expression cassette.

The expression cassette is flanked by inverted terminal repeats (ITRs) from the AAV2 virus. The 5' ITR sequence is from the WT virus and the 3' ITR contains a deletion in the TRS sequence that permits the generation of self-complementary genomes. The expression cassette contains a liver specific promoter with an intronic sequence from the SV40 virus that drives expression of a codon optimised form of the human FIX protein. A poly adenylation sequence from the SV40 virus is used.

Figure 1.2: DNA sequence of WT human and codon optimised FIX.

DNA sequence alignment (snapgene) of the WT human FIX coding sequence (GenBank: J00137.1) compared to the codon optimised sequence used in the SJ/UCL expression cassette. Matching bases shown in black. Altered bases shown in red.

Figure 1.3: Protein sequence of WT human and codon optimised FIX

Protein sequence alignment (snapgeneS) of the WT human FIX protein product compared to the codon optimised sequence used in the SJ/UCL expression cassette. Matching residues shown in black. Altered residues shown in red. Malmö mutation at amino acid 194 in full sequence (148 in mature protein)

1.5.5 Other AAV trials for haemophilia:

There are now several active and completed gene therapy trials for both haemophilia A (Table 1.2) and haemophilia B. (Table 1.3) Successful results have been published from an AAV mediated gene therapy for haemophilia A, using a, truncated form of the FVIII protein, showing durable efficacy over 2 years post treatment infusion.[129] Almost all haemophilia gene therapy trials in progress use gene addition via a single AAV cassette as the therapeutic approach, although one trial for haemophilia B from Sangamo therapeutics utilises a Zinc-finger based gene editing approach to attempt to insert the FIX gene into a safe harbour locus (NCT02695160). This requires 3 different AAV to enter the same cell, an AAV containing the ZFN left arm, and AAV containing the ZFN right arm and an AAV containing the FIX gene. So far, no results have been posted from this trial.

One trial for haemophilia B run by Baxalta therapeutics was halted after a rise in ALTs and concomitant loss of FIX expression that could not be brought under control by prednisolone administration.[130] This unsuccessful cassette was designed to have a high CpG content in order to boost protein expression levels.[131] It is likely that the high CpG content played a role in the robust immune response seen against this vector and it is now suggested that expression cassettes in AAV be depleted for CpG motifs.[132–134]

AAV serotype	Cassette	Dose (vg/kg)	Sponsor	Phase	Status as of Sept 2020	Clinical trial#
hu37	BDD-hFVIII (BAY2599023 (DTX201))	4 escalating - not listed	Bayer	I/II	Recruiting	NCT03588299
AAV6	BDD-hFVIII (SB-525)	adaptive' not listed	Pfizer	II	Recruiting	NCT03061201
AAV5	hFVIII-SQ (BMN-270)	6e13	BioMarin	I/II	Enrolling	NCT03520712
AAV6	hFVIII PF-07055480	not listed	Pfizer	III	Recruiting	NCT04370054
AAV8	HLP-hFVIII-V3	6e11 : 2e12 : 6e12	University College London	I	Recruiting	NCT03001830
AAV8	BDD-hFVIII (BAX888)	2e12 : 6e12 : 1.2e13	Baxalta (Shire)	I/II	Active, not recruiting	NCT03370172
AAV5	hFVIII-SQ (BMN-270)	not listed	BioMarin	I/II	not listed	NCT02576795
AAV5	hFVIII-SQ (BMN-270)	4e13	BioMarin	III	Active, not recruiting	NCT03392974
not listed	BDD-hFVIII (SPK-8016)	not listed	Spark Therapeutics	I/II	Active, not recruiting	NCT03734588
LK03	BDD-hFVIII (SPK-8011)	5e11 : 1e12 : 2e12	Spark Therapeutics	I/II	Recruiting	NCT03003533

Table 1.2 - AAV Haemophilia-A gene therapy trials

Sourced from clinicaltrials.gov. Search term: AAV Haemophilia

AAV serotype	Cassette	Dose (vg/kg)	Sponsor	Phase	Status	Clinical trial#
AAV5	hFIX codon optimised	5e12 : 2e13	Uniqure	I/II	Active, not recruiting	NCT02396342
AAV5	hFIXco-Padua (AMT-061)	2e13	Uniqure	II	Active, not recruiting	NCT03489291
AAV8	hFIX19	3 escalating - not listed	Spark Therapeutics	I	Terminated	NCT01620801
AAV5	hFIXco-Padua (AMT-061)	not listed	Uniqure	III	Active, not recruiting	NCT03569891
not listed	BBM-H901	5e12	Institute of Hematology & Blood Diseases Hospital, Tianjin	I	Recruiting	NCT04135300
AAV8	ss-3xCRM8-TTR-FIX_R338Lopt (SHP648)	3 escalating - not listed	Baxalta (Shire)	I/II	Recruiting	NCT04394286
AAV6	SB-FIX: ZFN1, ZFN2 and cDNA Donor.	3 escalating - not listed	Sangamo	I	Active, not recruiting	NCT02695160
rh10	hFIX (DTX101)	1.6e12 : 5e12	Ultragenyx Pharmaceutical Inc	I/II	Terminated	NCT02618915
Spark100	hFIX PF-06838435 (SPK-9001)	escalating - not listed	Spark Therapeutics	II	Completed	NCT02484092
AAV2	hFIX	escalating - not listed	Avigen	I/II	Terminated	NCT00076557
AAV8	scLP1hFIXco	2e11 : 6e11 : 2e12	St Jude Children's Research Hospital/ University College London	I	Active, not recruiting	NCT00979238
AAV8	hFIX Padua (AskBio009) (BAX 335)	2e11 : 1e12 : 3e12	Baxalta (Shire)	I/II	Active, not recruiting*	NCT01687608

Table 1.3 - AAV Haemophilia-B gene therapy trials

Sourced from clinicaltrials.gov. Search term: AAV Haemophilia
*BAX335 halted due to uncontrolled loss of expression

1.5.6 FIX Padua:

In 2009, a patient with thrombosis in a Padua clinic was identified as having a mutation in the FIX gene.[135] A single amino acid mutation (R338L) in the FIX gene conferred hyperactivity of the protein between 5 and 7-fold. This hyperactivity is due to a faster activation of Factor X facilitated by the cofactor FVIIIa.[136] This would mean that if applied in the same therapeutic context a lower dose could be given to achieve the same level of FIX activity. Preclinical work showed that incorporation of the FIX Padua mutation in the recombinant AAV setting could yield improved clotting activity at a given dose.[137–139] FIX Padua cassettes are now been used in several clinical trials and have yielded high clotting activity levels in the reported results from the first trial incorporating this mutation.[140] Other hyperactive mutations have since been identified and exhibited efficacy in preclinical models;[141,142] however, these have yet to be incorporated into clinical strategies.

1.5.7 Gene therapy for haemophilia with Inhibitors:

The potential of treating haemophilia by gene therapy for patients with inhibitors via liver gene therapy is an exciting prospect. To date no human patient treated for haemophilia with AAV gene therapy has developed inhibitors against the therapeutic protein product. However, this has been observed in at least one canine from the CHOP FIX dog cohort, albeit transiently, and in a dog from an inhibitor prone subcolony.[143] Indeed, there is evidence that the expression of the therapeutic transgene from the liver provides a tolerogenic environment, and AAV gene therapy is being investigated as a clinical strategy in haemophilia patients whom have developed inhibitors against recombinant protein.[144,145] Patients who develop inhibitors have the highest treatment costs, potentially upwards of a million euros per year;[96,97] therefore, the translation of gene therapy to the

clinic for this cohort of haemophilia patients could have both a substantial therapeutic and economic impact.

1.5.8 The future of haemophilia gene therapy:

Now that proof of principle has be attained, goals of the field will include attempting to lower the necessary dose of AAV to achieve normalisation of clotting factor activity. A diversification of gene delivery methods may also reach clinical potential. There has been extensive preclinical work for haemophilia treatment via ex vivo and in vivo lentiviral gene therapy,[146–149] and there are now active clinical trials in lentiviral transduction and infusion of CD34+ cells for both haemophilia A and B (Table 1.4). Additional changes to treatment strategy such as the delivery of engineered proteins and other variations such as the use of ancestral protein sequences have been shown to improve FVIII gene therapy in pre-clinical models,[150] and it is likely that in the future a range of recombinant gene products and strategies will be available for the gene therapy treatment of haemophilia, much as a range of recombinant protein products are available for infusion in the present day. It is incumbent upon the field to ensure this is done in as safe a manner as possible and therefore detailed study of these gene therapy products and the processes involved in their generation is essential.

Disease type	Cell target	Cassette	Sponsor	Phase	Status as of Sept 2020	Clinical trial#
Haemophilia-A	Autologous CD34+ PBSC	-889ITGA2B-BDDFVIII-WPTS (MUT6)(VSVg)	Medical College of Wisconsin	I	Recruiting	NCT03818763
Haemophilia-A	Autologous CD34+ haematopoetic stem cells	FVIII	Expression Therapeutics	I	Not yet recruiting	NCT04418414
Haemophilia-B	Autologous hematopoietic and mesenchymal stem cells	YUVA-GT-F901	Shenzhen Geno-Immune Medical Institute	I	Not yet recruiting	NCT03961243

Table 1.4 - Lentiviral treatment of CD34+ cells for Haemophilia gene therapy

Haematopoetic gene therapy clinical trials for Haemophilia. Sourced from clinicaltrials.gov
Search terms: Haematopoetic Haemophilia

1.6 The AAV genome:

The AAV genome comprises several overlapping reading frames encoding replication and capsid proteins that are situated between inverted terminal repeat sequences (Fig. 1.4 a). There are 4 overlapping REP proteins, named for their respective molecular weights in kilodaltons (KDa); REP78, REP68, REP52, and REP40. These and 3 overlapping capsid genes; VP1, VP2 and VP3 which, through alternative splicing and the use of alternate initiation codons, are produced at a 1:1:10 ratio that assembles into a 60mer (5 copies VP1, 5 copies VP2, 50 copies VP3) in the nucleus of an infected cell, resulting in the mature icosahedral virion. AAV contains 4 known promoters within its genome; P5, P19, P40, and P81. The P5 promoter is located at the 5' of the AAV genome, slightly downstream of the 5' ITR sequence. P5 drives expression of the large replication proteins, REP78 and REP68, and operates under a negative feedback loop to autoregulate its expression.[151] The P19 and P40 promoters are located within the coding sequence of the full-length REP gene; P19 drives the smaller REP proteins REP52 and REP40,[152] whereas the P40 promoter drives expression of the capsid proteins and the AAV Assembly protein (AAP),[153] produced from an ORF in an alternate reading frame within the capsid gene. In addition to this there are other identified ORFs such as the AAVX gene,[154] which produces a protein aiding in AAV replication, and the MAAP gene, which encodes a protein of unknown function that associates with the membrane of the host cell (Fig. 1.4.b).[155]

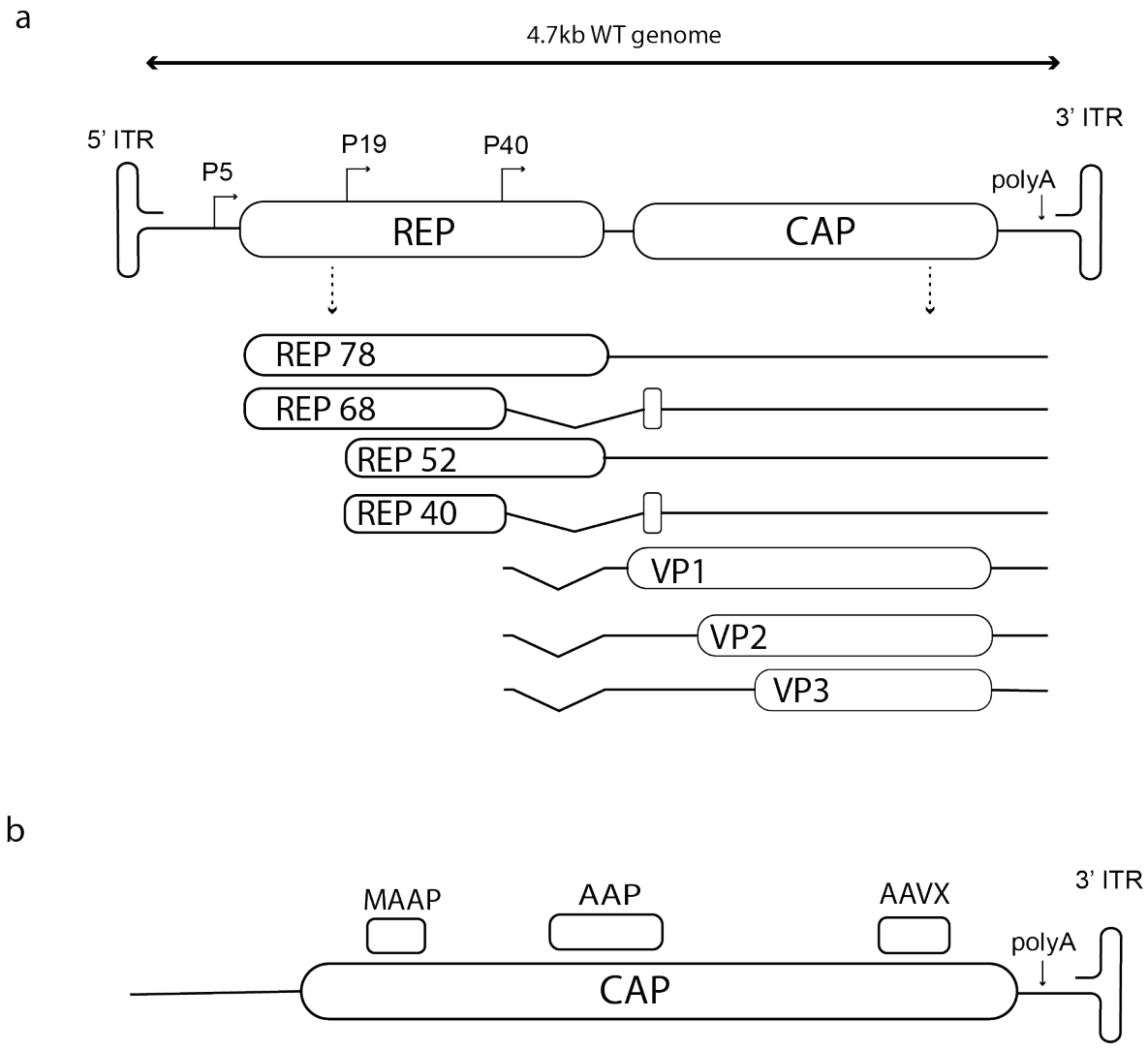


Figure 1.4: Schematic depicting the wildtype AAV genome and locations of overlapping genes.

(a) From top to bottom. The WT AAV genome with promoters and gene groups; The 78Kda replication protein; the 68KDa replication protein; the 52KDa replication protein; the 48KDa replication protein; viral protein 1; viral protein 2; viral protein 3. **(b)** positioning of alternate reading frame genes encoded within the AAV capsid sequence. From left to right. Membrane-associated accessory protein; Assembly activating protein; The AAV X protein.

1.6.1 The large form Replication genes: AAV REP78 and AAVREP68:

The overlapping AAV replication genes exist at the 5' half of the AAV genome. All 4 coding regions are in frame with each other and so their final protein products have matching domains. The large form REP genes REP78 and REP68 are driven from the P5 promoter and the encoded proteins exhibit DNA binding, helicase and endonuclease activity.[156] The genes are named for their molecular weight,[157] comprising 621 and 536 amino acids respectively (NCBI Reference Sequence: NC_001401.2). They are identical in their first 529 amino acids,[158] but differ in the c terminus due to alternate splicing at donor site nucleotide 1906 and acceptor nucleotide 2228.[159] This leads to a degree of redundancy in the REP proteins as both REP78 and REP68 can stimulate AAV DNA replication without the other protein present,[160] and high-level expression of either REP78 or REP68 can result in the production of infectious particles.[161] The DNA binding domain of REP78 and REP68 exists in the N-terminal region, a region not present in the lower MW REP proteins, with point mutation at amino acid 225 showing abrogated DNA binding in both REP78 and REP68.[162] This binding domain recognises and binds to a GAGC triplet at the inverted terminal repeats, and an imperfect GAGC repeat in the AAV2 P5 promoter $(GAGY)_2$, both of which can serve as the origination point for AAV replication.[163,164] When bound to the ITRs REP is able to form a phosphodiester bond with an exposed thymidine residue in the TRS loop in order to nick the DNA and facilitate genome replication.[165] REP78 and 68 can self-associate at these binding sites into complexes. Two studies suggested that such complexes are hexameric in nature,[166,167] although more recent work suggests that the actual nature, at least for the REP68 protein, is pentameric when in

contact with the dsDNA of the RBS and a double octamer in contact with ssDNA.[168] It is likely that these REP protein complexes are essential for mediating AAV integration during the REP dependent association of the AAV genome with the AAVS1 integration site on chromosome 19, that can be promoted by the ITR sequences and the P5 promoter.[169,170] The N terminal region of the large form REP proteins also harbour the site responsible for the endonuclease activity, a Y156F point mutation eliminates endonuclease activity on a ITR hairpin substrate.[171] The helicase activity of REP78 and REP68, necessary for the strand unwinding required for AAV replication is ATP dependent.[172] In AAV2 the consensus helicase motifs occur between amino acid 330 and 422, homologous to regions in replication proteins of other parvoviruses.[158] Furthermore, REP78 and REP68 contain basic residues in the shared region of their c-terminus which function as a nuclear localisation sequence (NLS).[173]

REP78 and REP68 have strong roles in the regulation of AAV genome transcription. In the absence of the helper genes from adenovirus or herpes simplex virus, the large forms of REP maintain a repressive role on AAV transcription from the P5 and P19 promoters.[174] In the presence of these helper genes the repressive function of the REP genes is lifted, and REP, acting in concert with other factors, functions as a transcriptional activator of the P19 promoter.[175] During the replication process, the REP proteins also interact with host cellular proteins. REP78 has been shown to bind to the TATA binding protein,[176] a general transcription factor shown to be necessary for initiation of replication at the P5 promoter.[177] The large form REP proteins are known to bind with other regulators of transcription like PC4 and SP1,[178,179] and to interact with other host proteins like B23/Nucleophosmin and KCTD5.[180,181] At the DNA level, consensus binding sites for the

GAGC repeat recognition sequences exist within the human genome, in and around genes like tyrosine kinase activator 1 and the breast cancer associated BRCA1 gene among others.[182] Downregulation of certain oncogenic promoters such as c-fos an c-myc by REP78 has been observed.[183] Conversely upregulation of the c-sis proto oncogene by REP68 has been described.[184] It has been reported the these proteins have an anti-proliferative effect on oncogenicity in certain contexts,[185–187] and that the large form REP proteins can exhibit toxicity to cells, inducing apoptosis through caspase-3 activation.[188] Moreover, when cells are exposed to other cytotoxic conditions, those expressing large form REP have been shown to exhibit greater levels of cell death than REP negative cells.[189]

In addition to their effects on human cells, the large form REP proteins have also been shown to affect the life cycles of both DNA and RNA viruses. REP78 inhibits papilloma family virus replication for both human and bovine variants,[190,191] hepatitis B virus through binding to the HBV core promoter,[192] and HIV.[193] Furthermore, REP78 and REP68 can inhibit the replication of the DNA viruses that provide AAV helper function, adenovirus and herpes simplex virus.[194,195]

1.6.2 The short form REP genes: AAV REP52 and AAV REP40:

The short form REP proteins are driven from the P19 promoter. The short form REP genes of AAV2 are in frame with the large form but translation of REP52 and 40 begins at the site of amino acid residue 225 in the large form REP proteins. REP52 and REP40 are 397 and 312 amino acids in length respectively (NCBI Reference Sequence:

NC_001401.2). They are also splice variants utilising the same c terminal splice donor and acceptor sites as the large form REP proteins.[159] Due to their translational start position both proteins lack the DNA binding domain present in REP78 and 68, and neither REP52 or REP40 can promote site specific integration of AAV into the host genome.[196] These shorter forms also lack the endonuclease domain required for trs strand nicking during replication, and no endonuclease activity for these proteins has been described.[156] Unlike the large forms of REP, REP52 and REP40 are also thought to only exist in monomeric form,[166,197] a hypothesis bolstered by the observation that at least for AAV5, the key oligomerisation domain exists in the N terminal region of the large form REPs that is absent from REP52 and REP40.[198] However, the short forms of REP are still required for efficient packaging of the AAV genome.[199] REP52 and REP40 retain helicase activity, which occurs in a 3'-5' direction,[197,200] and a major function of the shorter form Rep proteins appears to relate to the direct packaging with the viral genome through the preformed capsid pore, a process that requires the helicase activity of REP52 and 40, in which it is hypothesised that the helicase acts as a 'molecular motor' to load the ssDNA of AAV into the capsid through the open pore at the 5 fold symmetry axis.[199] and the presence of the short form REP proteins is also essential for the accumulation of single strand DNA in cells during replication.[201]

The short form REP proteins do not display the level of toxicity to cells as seen with REP78 and REP68. In one study that examined short form REP toxicity, REP52 did appear to have a moderately toxic effect on the cells tested, although much lower than that of REP78.[188] Also in contrast to the large form REP proteins, the short form do not appear to effect the transcription of the adenoviral late promoter significantly; only a

modest effect of adenoviral gene transcription is seen with REP52 and none is observed with REP40.[202] This is likely to be due to the lack of the REP78 and REP68 N terminal DNA binding domain, and it is likely that the shorter form REP proteins have a much diminished role, if any, in transcriptional regulation of viruses and cells, although wider effects of the shorter form REP proteins on cellular transcription have yet to be studied in any detail.

1.6.3 AAV capsid genes:

The AAV capsid is the outer shell of the AAV virion, essential for mediating successful delivery of the packaged genome. It is icosahedral in structure and coded for by 3 overlapping proteins designated VP1 VP2 and VP3. This overlapping capsid gene structure is common to other Parvovirus family members like *Bocaviridae* and *Densoviridae*.[203,204] Like the REP proteins, the VP proteins are all in frame with each other. The MW of the VP proteins for are approximately 87KDa 73KDa and 67KDa for VP1 VP2 and VP3 respectively.[205] The 3 capsid proteins are produced from a single mRNA by alternative splicing.[206] VP1 and VP3 of AAV2 are translated from a canonical AUG start codon whereas VP2 translation initiation occurs at the lower efficiency ACG codon.[207] The viral proteins assemble in a 1:1:10 ratio to yield a 60mer particle with a total MW (empty capsid) of 3.72MDa for AAV2.[208] Once assembled, the virion structure is remarkably stable. Long term storage at 4 degrees Celsius or exposure to temperatures up to at least 55 degrees Celsius appear to have no effect on the transducing capability of AAV.[209]

The AAV2 capsid undergoes assembly in the nucleolus.[210] Nucleolin has been shown to bind to the AAV2 capsid,[211] which may facilitate the nucleolar import of the capsid proteins. However, this interaction may be a feature unique to the AAV2 serotype as the VP proteins of other serotypes have not been shown to associate with the nucleolus.[212] The AAV capsid proteins can undergo a range of post translational modifications. Residues can be phosphorylated, ubiquitinated, glycosylated, SUMOylated and deamidated, with modifications like tyrosine phosphorylation and deamidation correlated with reduced functional transduction capabilities.[213–215] Recently it has also been shown that capsid can interact with the promoter sequence of its packaged expression cassette and influence tissue specificity of expression, at least within certain neurons.[216]

VP1 as the longest of the overlapping capsid proteins contains residues in its n-terminus that are not found in VP2 and VP3. This region is referred to as VP1u.[217] Capsids deficient in VP1 are unable to infect cells. VP1u contains basic amino acid residues that can influence the transduction capabilities, although basic residues are detected at the N-terminus of all the VPs.[218] VP1u also contains PDZ binding motifs at its n terminus. Mutation of these binding motifs prevents uptake of the viral particle into the nucleus, suggesting a role for these sites in the nuclear trafficking step post-endosomal escape.[219] Other consensus sequences in VP1u include SH2 and SH3 binding motifs, and endosomal sorting signal motifs.[219] One of the most prominent features of VP1u is a phospholipase A2 domain.[220] This region of the capsid undergoes a conformational change during the portion of AAV infection in which AAV containing endosomes are acidifying. This change exposes this protease region of VP1u and facilitates the escape of the AAV capsid from the endosome.[221]

VP2 also has basic residues in its N-terminus that are internalized in the fully formed particle but become exposed during the course of AAV infection.[222] Whilst VP1 and VP3 are essential to produce infectious AAV particles, virions that lack VP2 retain infectivity.[223] Furthermore, the N terminal region of VP2 can handle large insertions if provided separately from VP1 and VP3, a feature which has been utilised to try and improve specificity of AAV targeting.[224]

VP3 is the smallest and most abundant of the AAV capsid proteins. Whereas VP1 and VP2 localise to the nucleus, VP3 is seen in both the nucleus and cytoplasm of cells.[225] This could be due to absence of residues that serve as an NLS in the VP1 and VP2 proteins.[226] VP3 is the only essential capsid protein to form genome containing particles, but VP3 only AAV capsids are non-infectious.[223]

The full-length capsid sequence is what determines the AAV serotype. Naturally occurring AAV serotypes have been discovered in various tissues and organisms.[114,227–230] There is a relatively low sequence constraint within the capsid to produce functional AAV. This is extremely useful for gene therapy as the range of existing serotypes can be expanded artificially to engineer capsid variants for specialised therapeutic use. Whilst many current AAV gene therapy trials have used naturally occurring serotypes,[11,68,231] artificial capsids have been developed to improve gene delivery to cells of the inner ear, neurons, and muscle amongst others.[232–234] AAV capsids in which peptides have been inserted into regions of the capsid that can tolerate this without disruption have expanded the targeting capability of AAV.[235,236] More experimental designs have also shown that capsid proteins from heterologous serotypes can be combined in different ratios to yield functional ‘mosaic’ AAV particles with unique properties, such as altered transduction profiles and

capsids activatable only after protease cleavage [237,238] The continued research into identifying new capsids both natural and artificial in nature will further the applicability of AAV as a gene therapy agent.

1.6.4 The inverted terminal repeats:

Inverted repeat sequences are a common feature of functional significance across prokaryotes, eukaryotes and viruses, allowing the formation of complex secondary structures.[239] The terminal ends of viral genomes are often characterised by inverted repeat sequences. DNA viruses with linear genomes from families like adenoviruses, poxviruses, phycodnaviruses and parvoviruses all contain these features in some form.[204,240–242] The AAV genome is flanked at each end by a 145bp inverted terminal repeat (ITR). This is the only element of the viral genome retained in recombinant AAV, comprising distinct regions relative to its function. 125 of these nucleotides are an inverted repeat sequence, that results in a T shaped hairpin loop. This palindromic sequence contains within it additional shorter palindromes between nucleotides 42-62 and 64-84.[243] The remaining 20 nucleotides make up the 'D sequence'. The D sequence is necessary for replication and packaging of the AAV genome.[244] The D sequence is implicated in aiding AAV transduction and has been shown to bind the nuclear import proteins RFX1 and RFX3.[245] Within this D sequence is the trs site; of which the thymidine residue is nicked by the large form REP proteins in AAV replication. The site is conserved across AAV serotypes, aside from the most genetically divergent serotype AAV5 which harbours a unique TRS sequence within its ITRs.[246] Mutation of the TRS site leads to abolishment of replication at the ITR,[160] a facet used advantageously in the generation of self-

complementary AAV. Other efforts to substitute the D sequence in its entirety has yielded AAV with higher infectious potential, albeit with lower viral production yields.[247]

The region in between the D-sequence and the hairpin loops is designated the A-region.[248,249] The A region of AAV2 contains the 14bp REP binding site that directly interacts with the large form REP proteins. The hairpin loops of the ITRs are denoted as B', B C' and C. With each ITR contain a B and C loop. The loops are GC rich and assemble to form a T shaped cruciform structure.[250] The ITRs can either be oriented with the b loop or the c loop as the top of the cruciform. This is known as FLIP or FLOP orientation respectively.[248,249] During replication There are three thymidine residues within the B and C palindromes that elicit stronger binding of the REP proteins to the ITR, at the top of the B loop in the Flip orientation and the C loop in the FLOP orientation.[251] In either FLIP or FLOP orientation the CTTTG residue that contributes to REP binding is positioned in the same orientation with respect to the 14bp RBS and the TRS site.[252]

The ITR sequences are essential for the circularisation and concatemerisation of the AAV genome post infection.[253] ITR sequences are found to be present alongside integrated AAV genomes, although not necessarily in intact form.[254] It is felt that the presence of the ITRs enhances the integrative potential of the viral genome. This is certainly the case for targeted integration in the presence of REP. The ITRs of AAV are also known to have transcriptional activity. This is not a unique feature among viruses; the terminal repeat sequences of other DNA viruses like adenovirus and vaccinia virus are also known to exhibit transcriptional activity.[255,256] In early studies of recombinant AAV, the promoter activity of the ITR this was used advantageously to do away with the need for providing an additional promoter sequence, as was the case with early designs to deliver the CFTR

gene for cystic fibrosis.[257,258] A small region in the A/D junction of the ITR is enough to drive recombinant transgene expression.[259] It is now understood that the ITR promoter activity exists across serotypes and cell types.[260] In limited circumstances the ITRs have been implicated in contributing to cellular toxicity, although the scope of this has not been investigated in detail.[261]

In the recombinant setting, the ITRs from AAV2 are used almost ubiquitously, especially in the clinical context. However, work has been done to assess the use of ITRs from other serotypes to produce recombinant AAV. For instance, if producing recombinant AAV with the AAV5 REP protein, the most evolutionarily divergent of the natural serotypes, AAV5 ITRs must be used as AAV2 ITRs cannot undergo efficient TRS site nicking and will not package AAV2 ITR flanked cassettes.[246] Indeed, a recombinant system using AAV5 REP and AAV5 ITRs was generated in the late 1990s.[262] Additionally, AAV3 ITRs have been used in certain research contexts,[263] and the use of heterologous ITRs, i.e. the 5' from one serotype and the 3' from another has also been examined.[253] Much as directed evolution efforts have led to the production of a vast array of recombinant AAV capsids, it is likely that future research endeavours will lead to the generation of AAV REP protein variants with divergent and favourable properties.

1.6.5 The P5 promoter:

AAV contains 4 known promoters within its genome; P5, P19, P40 and P81. The P5 promoter is approximately 145bp in length and located at the 5' of the AAV genome, slightly downstream of the 5' ITR sequence. The P5 promoter is present in all catalogued

serotypes of AAV, although in AAV5 it is referred to as P7.[264] P5 drives expression of the large replication proteins REP78 and REP68 and operates under a negative feedback loop to autoregulate its expression.[265] The transcribed mRNA is polyadenylated at a site near the 3' end of the genome in the case of AAV2, and at a central intronic location for P7 of AAV5.[264,266] P5 expression is inhibited by all forms of the REP protein, the large forms (REP78/68) have a predominant inhibitory effect, with REP52 and REP40 reducing mRNA from P5 to a milder degree.[174,267] This autoregulatory effect on P5 is facilitated by the presence of a series of imperfect GCTC REP binding motifs within the TATA box that the large form REP proteins can bind to through an N terminal DNA binding domain.[158] In addition to the REP recognition sequence, the P5 promoter contains 2 identified sites for YY1 binding. These are denoted as YY1-60 and YY1+1, denoted relative to the start of transcription site at P5. The YY1-60 site has a repressive role in transcription from P5 in the absence of the adenoviral E1a, which switches to an activator role in the presence of E1a.[268] YY1-60 is also key to the transactivation of P19, with the REP and SP1 proteins forming a scaffold to position the P5 YY1-60 complex driving the timed production of REP52 and REP40.[175] The YY1+1 element exists downstream from the P5 REP binding site. This binding site for YY1 plays an activating role in transcription from P5 by aiding transcript initiation,[269] and limiting the auto-repressive effect by REP68 on P5.[265] P5 also contains a binding site at its 5' end for the major late transcription factor (MLTF). MLTF is a cellular transcription factor that upregulates transcription of the adenovirus late promoter.[270] The MLTF binding site is similar to that of YY1-60 in that it is repressive in the absence of E1a and enhancing in its presence.[271] P5 is also the only AAV promoter

to date to have any bidirectional activity characterised therein,[272] although no functional significance of any of its short, minus strand RNA products has been determined.

Outside of its role in transcription, the P5 promoter sequence plays a role in the retention of the AAV genome within a cell: A 138bp region within the P5 promoter has been designated the P5IEE.[273] This sequence is necessary and sufficient for REP mediated integration of AAV into the AAVS1 site on chromosome 19 of the human genome. Although AAVS1 is considered a safe harbour locus, other GCTC repeat containing loci in the human genome are not, and targeted integration mediated by P5 could pose a safety issue in the recombinant context. Fortunately, the P5 promoter as with the other AAV promoters is absent from the packaged expression cassette in recombinant AAV. Evidence suggests that this configuration reduces the proclivity for recombinant AAV to efficiently integrate at the AAVS1 locus in the presence of REP.[274] Despite not being present in the packaged recombinant genome, the P5 promoter is key to recombinant AAV production in the 293 based system. Previous efforts in the recombinant system to replace the P5 promoter with a constitutive promoter did not yield high titers of virus,[275] likely to be due to the inability of a constitutive promoter to retain optimal ratios of the different viral replication components. Generating high titers of recombinant AAV is essential for the practical application of gene therapy and, therefore, the P5 promoter is valuable to the recombinant AAV production system.

1.6.6 The P19 promoter:

The P19 promoter is a short promoter of approximately 180bp in length located within the REP gene of the AAV genome. In AAV2, P19 drives the expression of the smaller KDa

REP variants Rep52 and Rep40. Like the P5 promoter, under normal conditions the REP proteins have an inhibitory effect on the activity of the P19 promoter; in the presence of adenovirus, this repression is lifted.[276] This appears to be a downstream effect of P5 activation as the large form REP proteins cooperate with the SP1 transcription factor and the YY1-60 site in P5 to mediate this transactivation of P19 transcription by interacting with the DNA upstream of the P19 TSS.[175,277] An additional key element to transcription is a CArG like sequence towards the 5' of P19 that aids in this transactivation.[277,278] Direct binding of REP to the P19 promoter has been observed, although there is only 1 GCTC sequence present within the P19 sequence.[279] As with P5, in AAV2 the mRNA transcriptional products of P19 are polyadenylated near the inverted terminal repeat in the 3' of the AAV genome, around 2kb from the terminal end of their protein products.[266] Again this differs in AAV5, which contains polyadenylation for P19 product within an intronic sequence centrally located within the AAV genome.[264]

1.6.7 The P40 promoter:

The P40 promoter of AAV drives expression of the capsid VP proteins and the assembly activating protein from alternatively spliced mRNA. Like P19 it is located within the REP gene coding sequence, further to the 3' end of REP. There are two identified SP1 binding sites within the P40 promoter,[278] and like P19, the P40 promoter is also reliant on transactivation by the REP proteins in concert with SP1.[280] P40 also contains binding sites for the transcription factors EF1a, ATF and AP1, which when deleted yield a slight reduction in transcriptional activity.[278] The transcription factor binding sites are conserved across the P40 sequences of different AAV serotypes.[281] As with P5 an MLTF binding

site is also present at the 5' of P40, although no function of this site has been determined.[278] In the presence of adenovirus, P40 can be transactivated by both the REP binding element in the P5 promoter and within the inverted terminal repeats.[282] The AAV5 version of this promoter is denoted P41 but is functionally the same, and as with AAV2 drives capsid protein expression, although with this serotype there is a much reduced level of transactivation from the corresponding REP protein.[283] More recently short non coding RNAs transcribed from the P40 promoter have been identified, which appear to play a role in inhibition of adenovirus replication.[281,284]

1.6.8 The Assembly Activating Protein (AAP):

The size limitations imposed by the physical space within the AAV capsid require the viral genome to encode all essential DNA in a highly compact genome. The replication and capsid subunit genes contain multiple overlapping genes in the same frame, and it was thought for a long time that these were the only required elements for the AAV life cycle contained within the viral genome. In 2010, the Kleinschmidt group identified a +1 reading frame gene within the AAV2 capsid coding sequence.[210] This 204 amino acid protein product, designated the Assembly-Activating Protein (AAP), is approximately 23KDa and translated from the alternative initiation codon CTG. AAP was found in AAV2 to be essential for functional capsid formation and to aid in the transport of VP proteins to the cell nucleoli for assembly. AAP was subsequently identified across other naturally occurring AAV serotypes.[285] The AAP sequence shows strong homology across serotypes, with only AAP from AAV 4, 5, 11 and 12 exhibiting high divergence from the consensus coding sequences.[286] Interestingly the AAPs from serotypes 4, 5, and 11 were

later shown to not be essential for capsid assembly of those serotypes.[212] Conserved regions of AAPs include two hydrophobic domains in its N terminus, a proline rich region between amino acids followed by five serine threonine clusters, and a basic c terminus.[286] Elements in both the N and C terminus of AAP have been shown to promote AAV capsid assembly, and it has been posited that the surface charge of capsid lumen residues plays a role in dictating the interaction of AAP with the capsid proteins.[287] AAP can conformationally change and stabilise the VP proteins that otherwise are rapidly turned over.[288] Indeed, even the AAV serotypes that do not require AAP to form functional capsids yield higher titers of vector when their cognate AAP is present,[289] although exogenous overexpression of AAP does not improve titers further.[290]

There is high cross complementarity between AAPs. An early study found that AAP2 can facilitate capsid formation of AAV1, AAV8 and AAV9, and AAP1 can facilitate some degree of AAV2 capsid assembly, whereas AAP5 cannot.[285] A subsequent study confirmed this and showed that the vast majority of AAP/serotype combinations result in functional particles, with AAP4 and AAP5 having more restricted function, only rescuing the production of each other and not any of the other serotypes tested.[289] With regards to recombinant AAV production, AAP is necessary in both the mammalian cell line and baculoviral systems to generate functional virions,[289] and aids production yields of VP3 only virus like particles (VLPs) in *Escherichia Coli*.[291]

Synthetic AAVs are often produced by combinatorial shuffling and/or directed mutagenesis. A problem with this is that due to the nested nature of AAP an alteration to the AAP sequence is highly likely for any changes made to the capsid protein between amino acid residue 203 and 380. This may result in functional capsids for a given utility

not being selected for due to dysfunctionality of the resultant nested AAP sequence through the introduction of a stop codon or deleterious mutations. It is also possible with synthetic variants of AAV that the resultant nested AAP sequence may not produce the highest yields in all cases.

1.6.9 The Membrane-Associated Accessory protein:

The presence of an additional nested ORF in the +1 frame of the AAV2 capsid gene was identified by the Kleinschmidt group in 2010 in their studies that elucidated the AAP protein.[210] However, in this study, no effect upon vector yield was observed and the potential protein was not investigated further in the manuscript. Recently this protein was further characterized and found to associate with the cell membrane.[155] Further studies will be required to assess the importance of this protein in the AAV life cycle.

1.6.10 The AAVX protein:

An additional gene known as AAVX has been identified in the AAV2 genome.[292] This gene has not been well characterised but appears to have a function in AAV replication; removal of the AAVX gene reduces AAV2 replication in 293T cells.[154] It is posited that X expression is driven from a little studied promoter within the AAV capsid gene referred to as P81.[292]

Homologues to this gene do not exist in all serotypes. AAV6 for instance, has a disrupted ORF in the region where a putative X gene would exist. Exogenous addition of the AAV2X gene enhances AAV6 production.[293] When AAV serotypes 1-12 are examined the X gene appears to be intact in some but not all (Fig. 1.5). AAV8, the serotype used in the SJ/UCL

haemophilia trial does not contain a functional nested X gene, with several stop codons present in the putative open reading frame (Fig. 1.6).

1.7 AAV helper viruses:

One key characteristic of the AAV virus is its inability under normal cellular conditions to undergo efficient replication in the absence of proteins provided by a helper virus. AAV derives its name from its discovery alongside adenovirus, and it has since been demonstrated that adenovirus and other viruses express proteins that facilitate efficient replication of AAV. These are described herein.

1.7.1 Adenoviral helper genes:

AAV is unable to undergo replication in normal cellular conditions without the presence of a helper virus. The best characterised of these is adenovirus. Typically adenovirus type 5 (Ad5) genes are used to provide this helper function, given that the E1 gene insertion in 293 cells is from this adenovirus serotype.[294] However, genes from other adenovirus variants such as Ad2 have also been used in the recombinant setting.[295] A range of functions assisting with AAV replication are provided by the adenoviral proteins; some involve direct interaction with components of the AAV virus, and some interact with other factors in the cell to increase permissivity to AAV replication. An example of an adenoviral protein that directly interacts with AAV is the early phase gene E1a. E1a proteins are expressed early in the adenovirus life cycle and have a significant effect on transcription within infected cells.[296] Transcriptional regulation is the major role played by E1a with regards to AAV replication; E1a presence helps activate the P5 promoter by

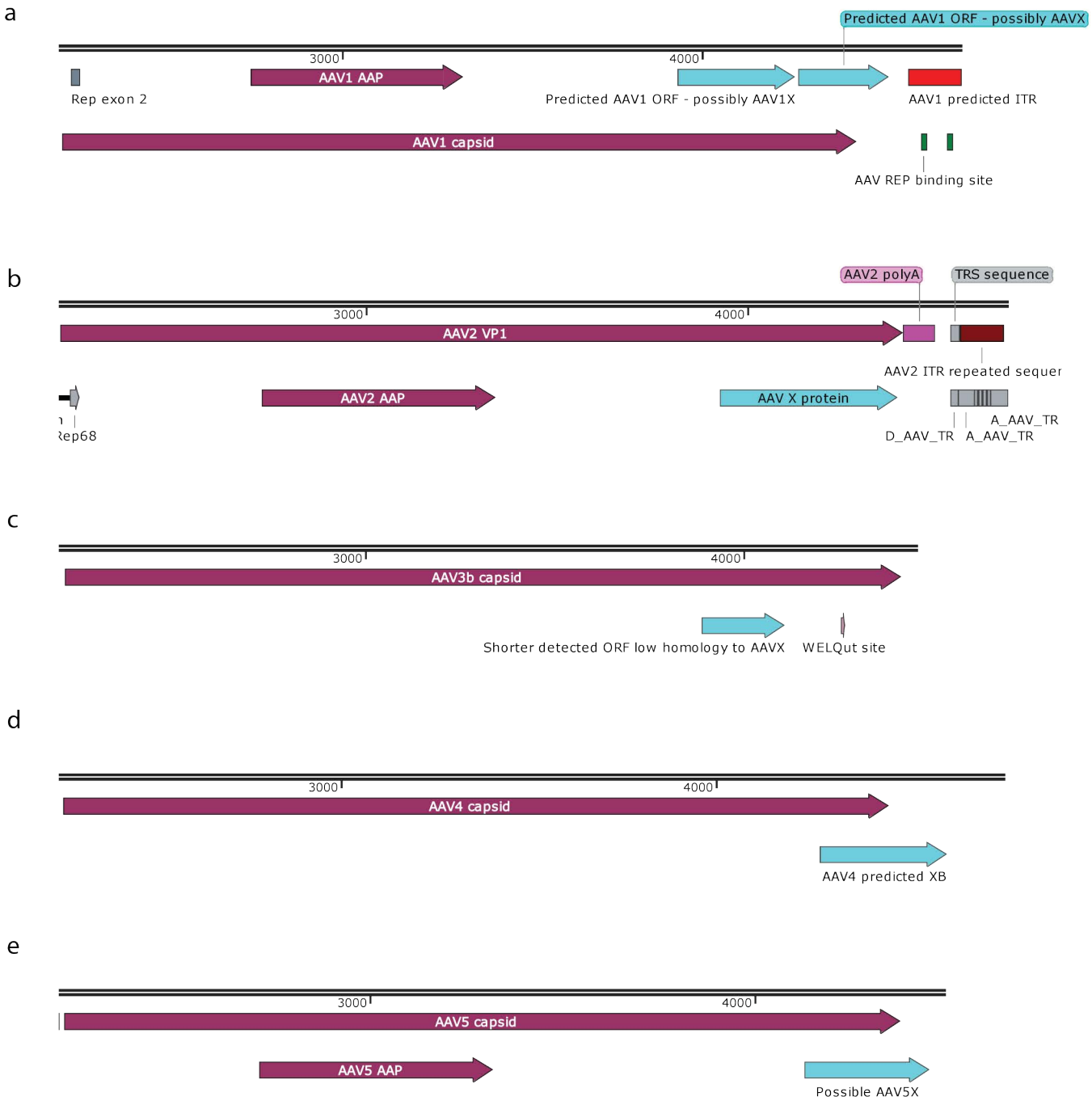
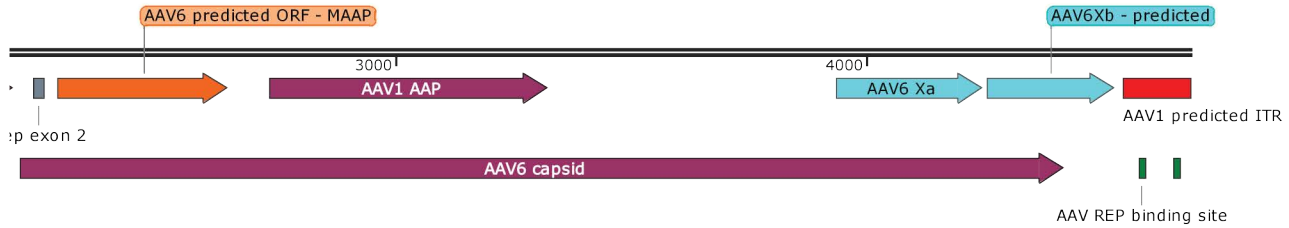


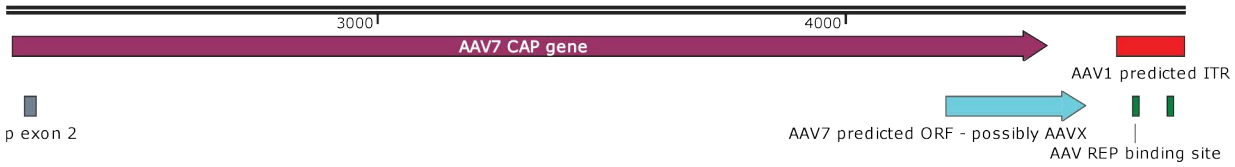
Figure 1.5: Schematic depicting AAVX gene homologues across naturally occuring sero-types

Mapped positions (snapgene) of potential AAVX homologues in **(a)** AAV1 - 2 Open reading frames **(b)** AAV2 (identified in literature) **(c)** AAV3b - low homology and short ORF **(d)** AAV4 **(e)** AAV5

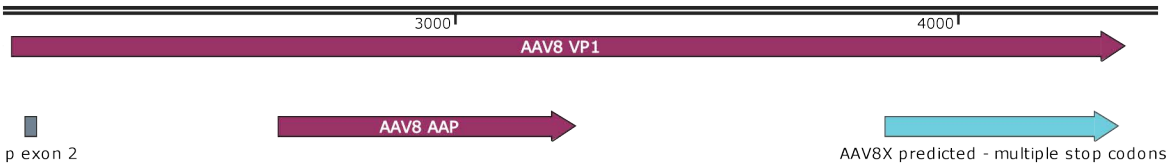
f



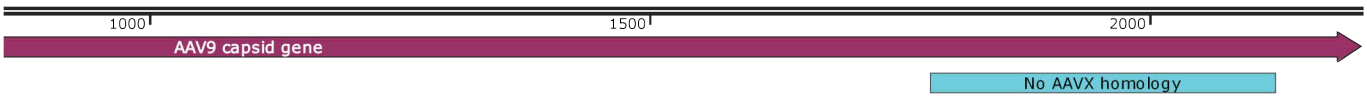
g



h



i



j

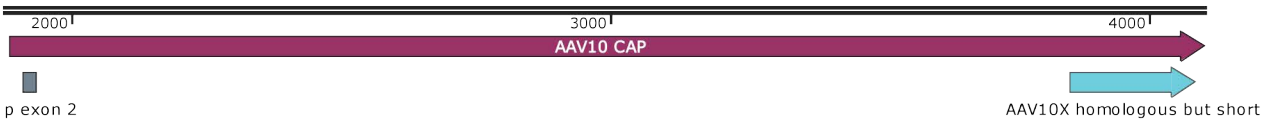
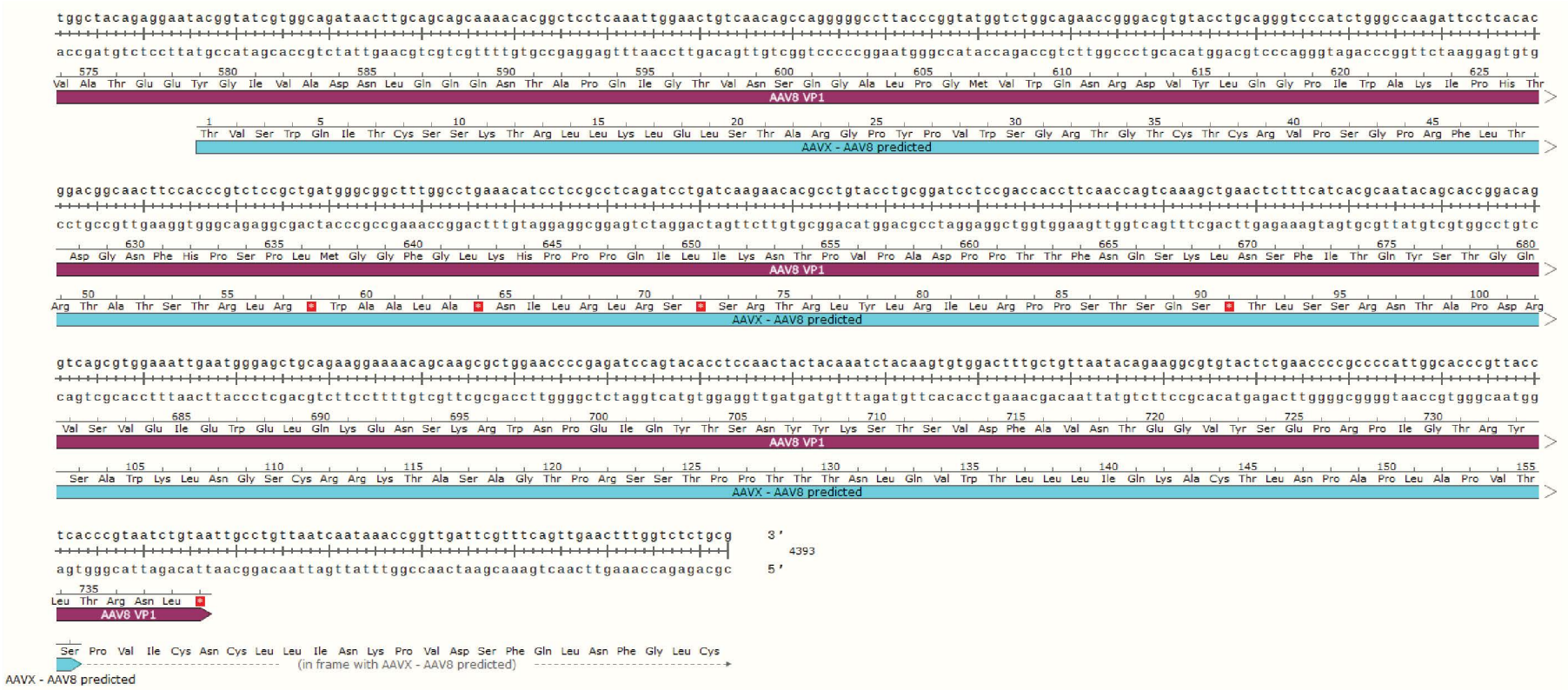


Figure 1.5 continued: Schematic depicting AAVX gene homologues across naturally occurring serotypes

Mapped positions (snapgene) of potential AAVX homologues in **(f)** AAV6 - 2 Open reading frames; identified in literature **(g)** AAV7 **(h)** AAV8 - multiple stop codons disrupt sequence **(i)** AAV9 - no detected homology **(j)** AAV10 - short homologous open reading frame

a



b

AAVX - AAV8 predicted	18	ELSTARGPYPVWSGRTGTCTCRVPSGPRFLTRTATSTRLR*WAALA*NILRLRS*SRTRLYLRILRPPSTSQS*TLSSRNTAPDRSAWKLNGSCRRKTASAGTPRSSTPP 127
		++ST + + WSGRT CT R PSG RF TRT T L W N L R SRT YLRILRPPS QS SS +T DRSAW+ +GSCRRKTA+AG P+ ST P
AAV X protein	18	QMSTHKAFFQAWSGRTEMCTFRGPSGQRFHTRTDIFTPLPSWVDSDLNTLLHRFSSRTPRYLRILRPPSVRQSLLPSSHSTPRDRSAWRSSGSCRRKTANAGIPKFSTLP 127
AAVX - AAV8 predicted	128	TTTNLQVWTLLLIQKACTLNPAPLAPVT 155
		TTT+L +WTLL AC + APLAP T
AAV X protein	128	TTTSLLMWTLLWTLMACIQSLAPLAPDT 155

Figure 1.6: AAV8 X predicted inactive gene homologue

(a) Map of AAV8X predicted protein homologue (snapgene) detailing positions of stop codons (red star) **(b)** Alignment of predicted translation product of putative AAV8X protein (top) with characterised AAV2X protein (bottom)

alleviating the transcriptional repression enacted on P5 by YY1 binding.[297] The E2a protein has also been shown to stimulate P5 transcriptional activity.[298] E2a , along with the VA RNAs, are involved in efficient translation of AAV capsid mRNAs.[299] Additional adenoviral proteins providing helper function to AAV include E1b55k and E4ORF6 proteins, which facilitate transport and accumulation of AAV mRNA into the cytoplasm[300] Interestingly in seeming contrast to this, E1b and E4ORF6 have been implicated in the degradation of Rep52 and preassembled capsid proteins, at least within the context of AAV5, an effect which is only overcome by the presence of VARNA.[301]

The interactions between adenovirus and AAV is not one way in terms of regulation; the REP proteins all inhibit the expression of the adenoviral E2a protein, with the larger REP proteins providing a greater inhibitory effect.[302] REP68 binding to the E2a promoter region has been observed and is thought to cause an inhibitory effect, although the observed binding site does not contain the GCTC motif recognised by the N-terminal region.[303] A similar inhibitory effect is seen at the adenovirus major late promoter, with all REP forms aside from REP40 capable of mediating inhibition.[202]

1.7.2 Herpes Simplex Virus helper genes:

The family *Herpesviridae* comprise large double stranded DNA viruses that infect a range of species. The HSV1 virus has a genome of approximately 150kb and encodes over 80 viral proteins and several transcribed regulatory RNAs.[304] HSV-1 has been shown to be able to provide helper function to assist AAV replication in the absence of adenoviral helper genes.[305] In fact the herpes simplex virus genes UL5, UL8, UL29, UL30, UL42, and UL52 are sufficient to facilitate AAV replication.[306] In a similar fashion to its effect on

adenoviral replication, AAV replication has an inhibitory effect on the process of HSV replication, mediated at least in part by the REP78 and REP68 proteins impairing the ability for HSV to form mature replication compartments in the nucleus.[307]

1.73 Other viral helper genes:

Aside from the well characterised AAV helper functions of adenovirus and herpes simplex virus, other viral genes have been shown to provide a helper effect to AAV replication. For instance, a number of proteins in the HPV16 genome have been shown to promote AAV replication and in particular the E1, E2 and E6 proteins of HPV16 have been shown to enhance recombinant AAV yields.[308–310] Vaccinia virus genes have also been implicated in promoting AAV replication,[311] although one study found it was insufficient to provide full helper function for AAV replication.[312] Furthermore, human Bocavirus 1 genes have shown to exhibit helper function for AAV replication, in the presence of adenovirus.[313] There is also evidence from an early study of the virus that in cells absent of other viruses that are placed under genotoxic stress, AAV replication can be facilitated.[314]

1.8 Recombinant AAV:

AAV was first incorporated into a recombinant system in 1982.[248] Soon after, the potential for transgene delivery using AAV was demonstrated by the successful transfer of the neomycin-resistance gene to human Detroit 6 cells.[315] These early production systems relied on the coinfection of adenovirus; however, in 1998 the Samulski group demonstrated a method of high titer viral production that used essential adenoviral genes

but did not require the use of live virus.[316] The utility of this recombinant system is that the REP and CAP genes are delivered to the production system in trans and the sequence between the ITRs is replaced with an expression cassette of interest of up to approximately 5kb in length, which is then packaged into the AAV capsid (Fig. 1.7). For certain monogenic disorders such as haemophilia and Leber's congenital amaurosis, the AAV system can be used to add a correct copy of the gene of interest. Additionally, approaches using silencing cassettes and the delivery of gene editing tools have proven successful in pre-clinical models for genetic disorders such as Huntington's, ALS, and others.[317–319] The typical production system uses the AAV2 REP gene and ITRs, however, the capsid gene is often substituted to produce different serotypes. There are several naturally occurring AAV serotypes that can be utilized in this system. Recently, however, groups have used methods such as DNA shuffling, peptide insertion and rational design to generate synthetic serotypes with desired properties, such as augmented infection of a target cell type, and evasion of immune response.[320–326] The production of AAV is discussed in detail in the section 'Recombinant AAV production'.

1.9 AAV infection.

Compared to other viral delivery types AAV infection is surprisingly inefficient. A very high multiplicity of infection (MOI) is used to deliver AAV to the cell type of interest, often upwards of $1X10^4$ particles dosed for every cell in an in vitro research setting,[327] and in a clinical setting this can be as high as $2x10^{14}$ vector particles administered per kilogram when administered systemically.[69] In contrast to this, the successful gene therapy treatment of X-linked SCID with ex vivo transduction of CD34+ cells with lentivirus uses an MOI of 20 in two successive exposures.[16] This relative inefficiency of AAV infectivity

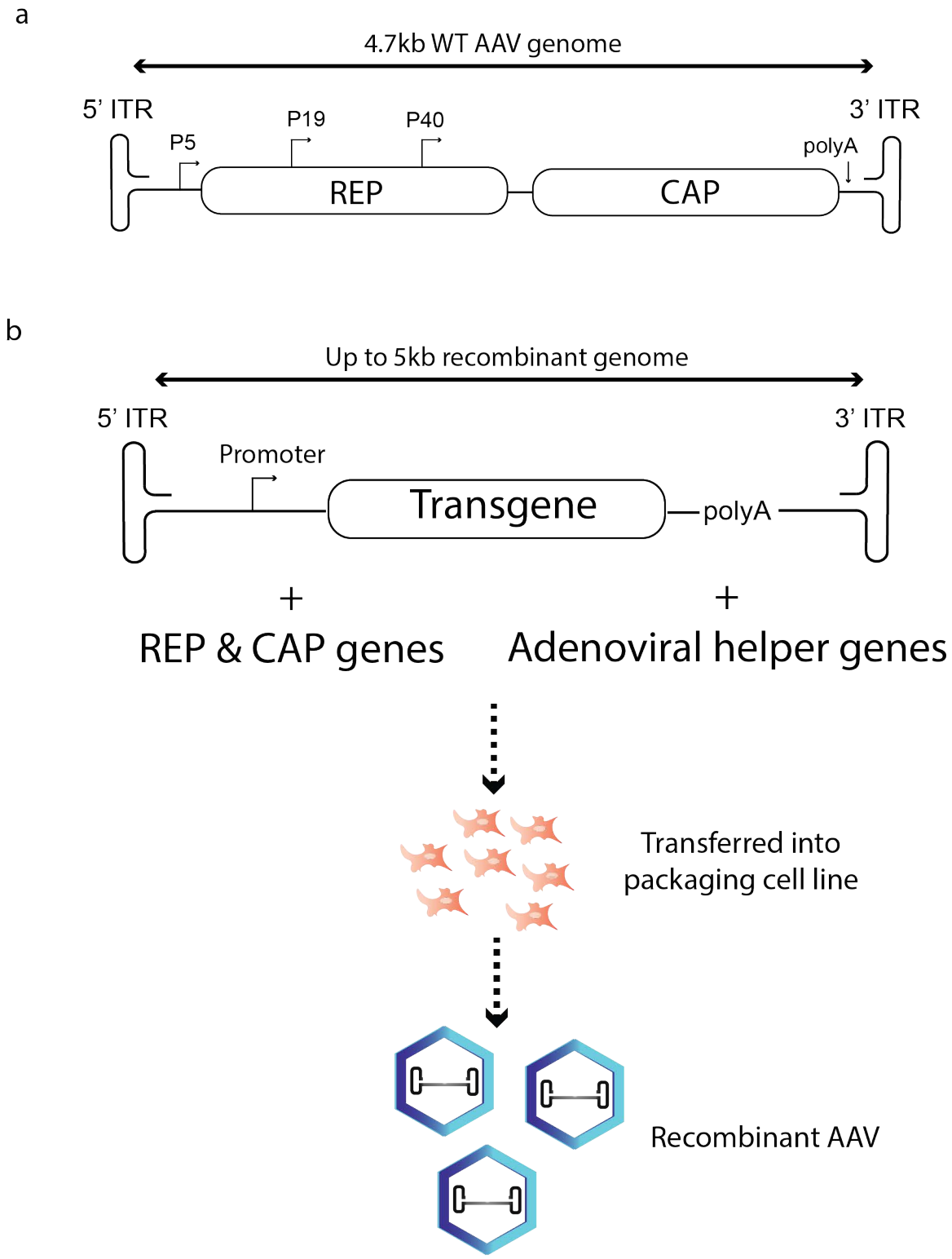


Figure 1.7 Recombinant AAV packaging system

(a) WT AAV genome without modification **(b)** Recombinant genome configuration with promoter, transgene and polyA sequence positioned between ITRs. Replication, capsid, and Adenoviral genes are introduced along with the recombinant AAV genome into the packaging cell line and recombinant AAV is subsequently generated in and harvested from those cells

is functionally mitigated by its ease of high yield production, exceeding $1X10^5$ viral particles per producing cell.[316] The high yield and a high tolerance for the virus allows for clinical dosing strategies to achieve these high infectivity requirements, circumventing the relative inefficiency of functional transduction on a per particle basis.

Whilst the precise mechanism of AAV infection is not fully elucidated, many aspects of the infection pathway are well characterised. AAV typically binds to a cell surface receptor via exposed residues on a threefold protruding external axis of the icosahedral capsid. The precise binding locations vary by serotype.[328–330] After binding, AAV particles are endocytosed, prevailing evidence suggesting this is largely a clathrin dependent process.[331,332] They are then transported within the endosome via the trans golgi network along microtubules in a unidirectional manner towards the nucleus.[333,334] Once the endosome reaches the perinuclear region, it undergoes acidification, leading to the exposure of the phospholipase A2 (PLA2) domain in the N terminus of the VP1 protein, a region that is otherwise conformationally hidden within the capsid. The PLA2 domain assists in the breakdown of the endosomal membrane allowing the escape of the AAV particle into the perinuclear cytoplasm.[222,335] The AAV virion then enters the nucleus, likely aided by a number of basic regions in the AAV capsid protein that have homology to NLSs and are required for AAV infectivity.[218,336] Once inside the nucleus the AAV capsid fully uncoats, and the single stranded AAV genome (WT or recombinant) utilises the cellular machinery to synthesise a second strand of DNA, converting it into a double stranded particle. This is an inefficient process and the rate-limiting step to functional transduction by single stranded AAV vectors.[337] Once the viral DNA is in a double stranded form transcription and translation of the AAV genome can occur. Alternatively,

genome transcription can also occur subsequent to intermolecular annealing of a positive and negative strand, bypassing the second strand synthesis step. Both second strand synthesis and intermolecular annealing are relatively inefficient. In recombinant self-complementary AAV genomes however, a positive and negative strand are contained within a single AAV particle, and expression of the transgene can be mediated far more rapidly in the target cell, due to intramolecular annealing of the two strands;[338] however, this is an artificial process that does not occur in WT virus.

A universal receptor for AAV cellular entry, KIAA0319L (also denoted as AAVR), has been proposed.[339] KIAA0319L is expressed broadly across cell types and knockout of this protein in cells severely abrogates AAV infection across many tested serotypes.[339] However, uninhibited infection occurs with a distinct lineage of AAV serotypes in a KIAA0319L deficient mouse model.[340] Furthermore the GPR108 and TM9SF2 proteins have been shown to be highly conserved entry factors, enhancing the transduction of serotypes across clades, with AAV5 the only tested natural serotype that does not utilise GPR108 as a cell entry factor.[341,342] There are still many aspects of AAV infection to be revealed and it is likely other host proteins critical to high efficiency AAV infection will become elucidated.

1.10 Recombinant AAV production and purification:

A primary concern of the AAV gene therapy field is the production of enough product to meet clinical need. Therapeutic doses of AAV depend on the target tissue. For retinal disorders, where AAV is delivered through sub-retinal or intravitreal injection, a far lower quantity of viral product is required than for larger organs or whole-body administration of

the AAV virus. A dose as low as 6e10vg per eye can achieve therapeutic transgene expression in human patients,[343] whereas for whole body directed therapies, viral loads of up to 2e14vg/kg have been administered.[69] To meet production needs for patients receiving high quantities of AAV, or indeed to expand the patient population for AAV treatments with lower dosing regimens, a highly efficient viral production system is essential. Fortunately, production systems can be scaled up to yield sufficient quantities, although challenges with producing high efficiency AAV still remain. There are two main production cell lines for generating recombinant AAV at clinical scale; HEK293 based and SF9 cells.[344]

1.10.1 HEK293 based AAV production:

HEK293 cells are derived from human embryonic kidney cells transformed with adenovirus serotype 5.[294] The cell line contains a stable insertion of the first approximately 4kb of the adenovirus genome into chromosome 19,[127] a sequence containing the E1A and E1b proteins (NCBI AC_000008.1). HEK293 cells are easy to culture and transfect and are therefore used in many applications requiring expression of exogenous proteins, such as protein production, for therapeutic injection; or recombinant adenoviral and retroviral production, for gene therapy.[345–348] Similarly, HEK293 cells can also be efficiently used to produce AAV for both research and clinical applications.

The HEK293 production method uses a 2 or 3 plasmid system; in which either adherent or suspension HEK293-derived cells are transfected with AAV viral production components alongside adenoviral helper genes. Transfected cells are then incubated, (most commonly at 37°C, 5-10% CO_2 neutral pH), in which time AAV replicates prolifically

in the cells, yielding in the order of 1e5 AAV particles per cell.[316] The time of production incubation prior to harvest ranges from 2-7 days depending on the production setup and research group.[349–351] A common derivative of HEK293 cells containing the SV40 T antigen (293T) is used routinely for AAV production at the research grade and has been used to manufacture clinical material.[128] The SV40 T antigen facilitates replication of DNA containing the SV40 origin of replication.[352]. Removal of the T antigen is now encouraged by the FDA for clinical gene therapy products, as it is a potentially transformative agent. Direct removal of the T antigen from these production cells generally results in dramatically reduced AAV titers.[353] However, many adaptations have been made to the 293 cell production system to facilitate large scale production capabilities. For instance, many early clinical trials in the field used adherent 293 cells to produce recombinant AAV.[354] This provides a 2 dimensional limitation to cell density, as the cells must be grown in a monolayer. Scale up with adherent 293 cells can be achieved by using cell stacks that contain multiple layers of culture surface for the cells to adhere to.[128,355] This is still monolayer culture and to overcome this on a larger scale, and maximise usage of production space, adapting the cells to grow in suspension gives a much higher cell density and ergo a much higher total yield, if efficiencies on the per cell basis can be maintained. 293SF cells for instance, which is a clonal cell line developed to operate in suspension, have been demonstrated to produce high yields of AAV.[356] Stable cell line methods of 293 based production also exist, in which the REP and/or CAP genes are inserted into the 293 genome and expressed after induction by transient transfection of the ITR flanked vector genome.[357]

Other 293 derivates adapted to various conditions have been demonstrated to provide useful properties for AAV production. Freestyle 293F cells (Thermo Fisher Cat# R79007) are an adapted cell line cultured to grow efficiently in serum free conditions. This is useful to AAV production for the clinic as raw animal derived biological products are advised to be kept to a minimum in the manufacturing of biological products.[358] Expi293F cells (Thermo Fisher Cat# A14527), are a further derivative of 293F cells and are designed to grow in denser culture.[359] Other 293 cell variants have been developed, like AAV-293 cells (Agilent Cat# 240073), and 293AAV cells (Cell biolabs Cat# AAV100); which both claim to improve AAV production yields. Given the benefits of improving AAV production it is likely that additional cell lines catered to aspects of AAV manufacturing will be developed.

1.10.2 Baculoviral AAV production:

Baculoviridae are a family of large double stranded DNA viruses of between 80 and 180kb.[360] Lepidoptera is the most well characterised baculoviral infection host;[361] however, baculoviridae have been found in several invertebrate species, including sawflies and crustaceans.[362,363] The baculovirus AAV production system uses the SF9 cell line derived from the *Spodoptera frugiperda* species of Lepidoptera.[364] Infection of SF9 cells with a baculovirus that contains recombinant expression cassettes has been used to produce proteins and viral particles since the late 1980s.[365] The first system to implement this system into high efficiency AAV production utilised a baculovirus derived from *Autographa californica* to deliver separate constructs containing the REP gene, the

CAP gene, and the ITR containing recombinant genome. These were delivered to SF9 cells by a triple coinfection.[366] All 3 of the key AAV promoters have activity in these cells, enabling recombinant production to occur.[367] However, a problem with this original system was genome instability, causing loss of the constructs with multiple passages of the virus, potentially limiting the scale of production.[368] A subsequent system used a dual coinfection configuration in which the REP and CAP genes were contained within the same baculovirus genome, a setup that produced higher titers and exhibited greater genetic stability. Other methods to improve the baculovirus system for AAV production have been developed. One group placed all of the AAV components within a single baculovirus construct to facilitate efficient production.[369] Another developed the 'OneBac' system, which contains the AAV REP and CAP genes stably inserted into the SF9 cell genome. In the OneBac system, expression of REP and CAP is induced only upon infection with the ITR flanked vector genome containing baculovirus.[370] The OneBac system is scalable due to the ability to easily expand SF9 cells, resulting in high titer AAV. A second version 'OneBac2.0' has been modified to improve the VP1 content of certain AAV serotypes and decrease baculoviral DNA impurities.[371] An additional issue with earlier generations of baculoviral production is the potential for the capsid proteins of certain serotypes to be digested by the baculoviral protease v-cathepsin, resulting in lower activity of the purified product, an issue that can be resolved by the addition of a protease inhibitor.[372] Furthermore, SF9 cells are known to contain an endogenous adventitious virus called sf-rhabdovirus.[373] Whilst unlikely to cause any negative downstream effects, an Sf9 based cell line with this virus removed (Sf-RVN) has been developed.[374]

AAV produced through baculoviral infection can be purified in a similar manner to that of the 293 cells system, through TFF and affinity chromatography, and AAV produced in this manner can efficiently transduce cells. Indeed, the first AAV gene therapy product to reach commercial authorisation (Glybera for treatment of familial hyperlipoproteinemia) was produced in a baculoviral cell production system.[375] Some groups have reported that AAV purified from baculoviral systems appears to be less infectious than that produced by the 293 cell based method. However, at least one group has reported that when optimisations are made to VP ratios, a greater transducing potency can be achieved using baculoviral production as compared to 293 cell based.[376] Further investigation will be required analysing the two production methods side by side to determine optimal conditions.

1.10.3 Other AAV production methods:

Other systems to produce recombinant AAV have been explored and developed. BHK cells that utilise herpes simplex virus for helper function, have been used to produce AAV for the treatment of A1AT deficiency in the clinic.[377–379] HeLa cells were an original production option for recombinant AAV, however concerns about clinical grade production with this cell line existed in part due to the presence of papilloma virus sequences contained within the HeLa cell genome.[380] However, Hela cell derivatives have been shown to be efficient cell line systems to produce AAV. Hela S3 cells, a derivate that can effectively grow in suspension culture, were used to produce AAV1 for a clinical trial of Serc2a delivery to heart failure patients.[354] Furthermore, before 293 derived cells became the standard production line, AAV was often produced in A549 cells[381,382] Additionally,

infectious AAV has also been successfully produced from Saccharomyces cerevisiae.[383] It is likely that with future research alternative cell lines will be generated to produce AAV from, both for academic and commercial purposes.

1.10.4 AAV purification:

Once AAV has been produced in the cell line of interest, it must be harvested and purified to a high concentration. This processing of the virus can be achieved in several ways. For research grade quality AAV, simplistic protocols designed to quickly and inexpensively purify infectious virus can be used. These include precipitation and ultra-centrifugation based methods,[384,385] and vary widely in whether input material is collected from just the supernatant, just the lysate, or a combination of the two.[355,386–389]

In the same manner that AAV production within cells must be performed at scale, purification methods must also be designed for processing of large volumes of virus containing media and cellular lysates. An initial step in AAV purification is often the reduction of material volume to an appropriate level for the processes further downstream. This can be achieved in a couple of ways; Tangential Flow through (TFF) continually passes material over a high MW filter to retain the virus within a smaller volume of media.[390] An alternative is the use of high MW polyethylene glycol (PEG). PEG is often used in protein purification and can precipitate AAV out of solution.[391] This precipitate can then either be processed further, by applying it to a gradient ultracentrifugation step,[392] or in the crudest of purification protocols, spun down and resuspended directly and used for infection.[385]

The primary ultracentriguation methods for purifying AAV are the use of sedimentation gradients. The two main types are the use of a caesium chloride step gradient and iodixanol. Caesium chloride gradients are typically a 2-stage process. Full particle density AAV in CsCl solution is approximately 1.4g/ml,[393] and so first a step gradient protocol where two phases, one above and one below the particle density of AAV are layered to form a single-phase interaction.[394] Full particles then sediment at the phase and can be extracted either for research grade use or for an additional CsCl continuous gradient sedimentation which improves the purity of the extracted virus.[395] The solution of extracted virus is then dialysed to remove Caesium Chloride, concentrated, filtered and resuspended. Iodixanol gradients operate on the same principle except the density of AAV particles in iodixanol is closer to 1.25g/ml and multiple density steps are typically layered to form the gradient.[396] Upon extraction of the virus from the gradient, the AAV can be concentrated for use without additional steps,[397,398] or further purified by chromatography.[399]

For clinical grade material, far more stringent and costly purification methodologies must be carried out to meet the standards of good manufacturing practice (GMP). These often involve multiple steps of column purification.[400] Affinity and anion exchange chromatography are the primary methods by which this can be achieved. Affinity chromatography in the context of AAV relates to the capsid proteins. A substance with binding affinity for the AAV capsid can be used in columns to bind and separate the capsid from other aspects of the cell lysate and media. This affinity agent can be a peptide, an antibody, or other agents with AAV capsid specificity.[401–404] One particular form of affinity chromatography useful to the field is the AAVX affinity resin, which is stated to elicit

serotype independent AAV column purification. The AAVX purification reagent is a camelid single chain antibody fragment with broad specificity across capsids of different AAV serotypes.[403]

The other class of column chromatography routinely used in AAV purification is anion exchange chromatography. This is a common purification process that yields separation of the intended product based upon charge. Columns are packed with positively charged resin that bind to negatively charged particles, the pH is then increased in a gradient to elute the bound particles from the column. Anion exchange chromatography can be used to purify many serotypes of AAV,[405] and strategies have been developed to use this method to differentiate between full and empty particles, to yield AAV preps with a higher proportion of expression cassette containing virions.[406,407] Size exclusion chromatography can also be implemented to isolate AAV particles.[408] Often this will be used as a second step in purification, and indeed two-step column purification of AAV is relatively common[408–411] With the ever-expanding range of AAV serotypes used in the clinic it is likely that a continued expansion of purification strategies that can either simplify or optimise purification of AAV will be developed.

1.11 Codon optimisation of AAV delivered transgenes:

Almost all life forms use DNA base triplicates to encode for amino acids. There are 64 codon combinations in total; 61 codon combinations that produce the 20 different amino acid, and 3 that result in translation termination. This provides some redundancy to the system, i.e. there are six different combinations of codons that code for leucine, and only

two amino acids (methionine and tryptophan) that are encoded by a single codon triplet.[412] It has long been apparent that a bias exists within genes and genomes for the codon triplets that are present.[413]

It has been shown that by codon optimising recombinant genes to match the highest codon usage for an organism, expression of recombinant transgenes can be increased as much as 100 fold.[414] Whilst it is easy to comprehend that different organisms have different usages of each codon, even within a single organism it has been shown that the expressed genes within a certain tissues have a unique codon bias signature.[415] This is an important consideration for gene therapy. Regarding therapeutic gene delivery, the benefit to codon optimisation will be situationally dependent, a gene with an already heavy bias toward preferred codon usage would have less to gain than one with a more divergent DNA sequence. For the FIX gene, codon optimisation has been demonstrated to produce a recombinant gene translated with 5 fold increased efficiency as compared to the WT sequence.[416] The benefit of this is that a greater therapeutic effect can be achieved at a given dose, or indeed in some cases it may end up being possible to lower the viral dose of a gene therapy vector to achieve a curative phenotype. Attempts to codon optimise proteins must be carefully considered depending on translational scenario, as the primary benefit from codon optimisation (increased translation speed), can result in altered protein conformation despite synonymous amino acid sequence.[416]

However, there are other parameters besides high expression that must be considered when codon optimising an expression cassette for an AAV construct. A trial launched by Baxalta therapeutics that incorporated a high percentage of CpG motifs in order to boost expression was halted after high dose treatment led to a considerable CD8+ T cell

mediated clearance of hepatocytes, and concomitant loss of FIX expression that could not be rescued by steroid administration.[417] Studies in pre-clinical models have shown that a lower CpG genome content in recombinant AAV results in a lower response from naïve CD8+ T cells, although, conversely, for any capsid specific CD8+ memory T cells present, a stronger activation is seen with a CpG depleted recombinant genome.[133] As such, the recommendations for codon optimisation of AAV expression cassettes now include minimisation of CpG motifs.[134] A further complicating factor is that simply matching codon usage to the target tissue may not provide the highest level of expression. Certain studies have shown that mRNA secondary structure can have a significant effect on translation efficiency. As such, great care needs to be taken to ascertain the ideal codon usage pattern of any expression cassette destined for the clinic.

1.12 Unresolved issues for AAV mediated gene therapy:

AAV mediated gene therapy has now been applied to several genetic diseases. Despite early clinical successes, there are still many issues to be resolved, of which some of the most prominent are discussed in this following section.

1.12.1 Serotype:

The ideal serotype to mediate AAV gene therapy for a given context is not a fully resolved issue. To use the liver as an example: AAV8 has been implemented successfully in clinical trials, due to its tropism and efficacy post intravenous injection in preclinical animal models, yet the transduction capability in human and NHP hepatocytes appears to be

sub-optimal. Comparative studies have been done with mixed results. A directed evolution study utilising a humanised mouse model showed that a capsid shuffled variant LK03 resulted in transduction of human hepatocytes approximately 10 fold that of AAV8.[321] Similar approaches have generated other artificial capsid variants that appear to exhibit enhanced liver transduction.[418,419] However, a separate study found that AAV3b, AAV8, LK03 and AAVrh10 all had similar human hepatocyte transduction efficacy in mice repopulated with human hepatocytes.[420] Other studies have reported conflicting data, for instance that AAV9 and AAVrh10 are superior at transducing human hepatocytes,[421] or that AAV7 has a higher transducing potential than AAV3b and AAV9.[422] Further research into the ideal serotype for human hepatocyte transduction is clearly required, although it may take several further independently reproduced studies before an answer can be settled upon. This same conundrum is present in the field of almost every target disease from delivery to the cochlea of the inner ear, to heart tissue, to retina. The iterative process of determining which capsid configuration is most suitable for specific disease applications is likely to continue for years into the future, and the production of synthetic serotypes to achieve this is a field unto itself.

1.12.2 AAV and the immune system:

A further unresolved issue within the field is the precise immune response observed after AAV administration. In several gene therapy trials with intravenous administration, a rise in liver transaminases (ALT/AST) has been observed.[12,69,423] Left untreated a concomitant loss of therapeutic protein expression can occur.[12] This phenotype has been successfully controlled with the administration of steroids upon the detectable rise in liver enzymes.[12,69,423] However, in at least one clinical trial, treatment with prednisolone was

insufficient to prevent the loss of therapeutic protein expression present alongside raised liver transaminases.[417] CD8+ T cell responses to both the capsid, and transgene encoded protein has been observed post infection in animal models and humans.[424–426] Additionally the DNA packaged within AAV can trigger the innate TLR9 immune pathway, stimulating interferon production.[427] Recombinant AAV genomes that are CpG rich or are engineered to be self-complementary (which contain three GC rich inverted repeat sequences instead of two) have been shown to elicit a greater TLR9 response.[428,429] Furthermore, whilst in the liver, the duration of the immune response against AAV appears to be transient, in a trial designed to treat Alpha-1 Antitrypsin deficiency by intramuscular injection, evidence of inflammation at the injection site and the presence of active T_{regs} and exhausted CD8+ T cells was observed over 5 years post infection, which the authors claim to be evidence of a tolerogenic effect post AAV treatment.[430]

Once a patient has been treated with AAV they will seroconvert, generating antibody-mediated immunity against AAV. In the clinical trial setting, the presence of high titer neutralizing antibodies against the AAV capsid has been seen long term post treatment infusion.[13] In the context of gene therapy, this leads to a ‘one and done’ situation; once an AAV gene therapy treatment has been given, a repeat administration of AAV of the same serotype is not possible under normal conditions. Groups have worked extensively to circumvent this problem; Multiple administrations of AAV in preclinical models have been achieved through immune modulation strategies.[431–433] One prominent example of this is the coadministration of rapamycin containing particles which allows for multiple administrations with the same AAV serotype in primates.[434]

The issue of pre-existing immunity against the AAV capsid is another considerable barrier to the applicability of gene therapy. The natural infection of humans by AAV yields a considerable portion of the human population exhibiting humoral immunity to AAV capsids. Pre-existing antibodies to AAV exist in humans worldwide and depending on serotype can be present in as much as 70% of the population.[120,121] Even artificially generated capsid serotypes that do not exist in natural settings have this issue, with one UK based study showing approximately 25% population seroprevalence against LK03, a capsid generated artificially by directed evolution.[321,435] Furthermore, cross reactivity across serotypes of certain anti-AAV antibodies could preclude merely treating with a different AAV serotype from being a viable option.[436] This would significantly reduce the eligible patient population to receive AAV based gene therapy. Work to circumvent pre-existing immunity to AAV has yielded the development of capsids designed to evade the humoral response,[437] and a method to remove AAV specific antibodies from the blood plasma via plasmapheresis have been developed.[438] Understanding the immune response against AAV in its entirety will be essential to determining the full safety profile of this virus as a mass market gene therapy agent.

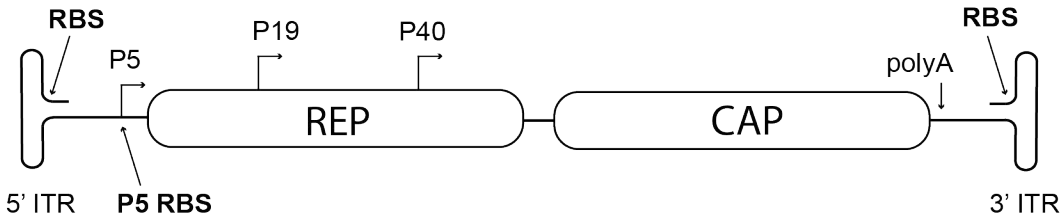
1.12.3 Dose limits and AAV toxicity:

Related but distinct from the issue of the immune response against AAV administration is the issue of the upper limit of AAV dosing. Successful systemic administration has been achieved at doses in excess of 1e14vg/kg[69]. Yet, preclinical studies have suggested that these high doses could yield significant toxicities in large animal models.[439] This has been compounded by the recent deaths in patients treated for X-linked myotubular myopathy.

Significant liver toxicity and subsequent sepsis was observed after a dose of 3e14vg/kg in multiple patients.[440] In this case, the livers already exhibited significant pathology prior to infusion. There is a need within the field to determine and set dose limits regarding AAV, perhaps dependent on the nature of particular genetic diseases, and to examine how to reduce dosing wherever possible without sacrificing therapeutic benefit.

1.12.4 Integration of AAV:

AAV is largely considered to be an episomal vector and has been detected in a wide variety of tissues in both animals and humans.[227,441] In the WT setting AAV has a distinct integration pattern. There are 3 characterised AAV preferential integration hotspots for WT AAV2 in the human genome. The best characterised site is AAVS1 positioned on chromosome 19q13.42,[442–445] a site only known to exist in humans and higher order primates.[446] However, additional preferential integration sites have been identified at chromosomes 5p13.3 and 3p24.3, denoted as AAVS2 and AAVS3 respectively,[447] and an ortholog to AAVS1 has been identified on chromosome 7 of the mouse genome.[448] These sites all contain GAGC repeat elements homologous to the REP binding sites present in the AAV genome (Fig. 1.8). The WT AAV5 virus, the most divergent of the naturally recovered AAV virus, however, has a characterized integration profile over 99% distinct from AAV2, and favours transcriptionally active GAGC repeats in intronic regions.[449] These targeted integration events are REP protein dependent and driven. Recombinant AAV does not contain the large form REP proteins to direct integration towards GAGC repeat sites. There have been numerous studies now looking at the profile



ITR REP binding sequence

GCGCTCGCTCGCTC

P5 REP binding sequence

GCGTGCTCACTCGGGC

AAVS1 REP binding sequence

GCTCGCTCGCTCGCTG

AAVS2 REP binding sequence

GCTCGCTCACTCACTC

AAVS3 REP binding sequence

GCTCGCTCACTCACTCACTC

Figure 1.8 AAV REP binding sites

(schematic) Positioning of REP binding sites in the WT AAV genome. **(Sequences - from top to bottom)** REP binding site within the inverted terminal repeat of an AAV2 ITR; REP binding site within the P5 promoter of the AAV2 genome; REP binding site of the AAVS1 integration site of the human genome; REP binding site of the AAVS2 integration site of the human genome; REP binding site of the AAVS3 integration site of the human genome.

of recombinant AAV integration in both preclinical and clinical samples that suggest a more randomised profile of integration events, although still favouring transcriptionally active regions.[450,451] A primary concern for the field, particularly after the fatal events in gene therapy trials using other delivery mediums, is whether AAV integration could contribute to a tumourigenic event. There have been a couple of studies in hepatocellular carcinoma patients linking WT AAV integration to oncogenesis.[452,453] In the recombinant setting, significant concern was generated after publication of a preferential integration in the RIAN locus on chromosome 12 resulting in HCC formation after neonatal injection of AAV in mice. However, whilst upregulation at the orthologous locus in human results in promotion of HCC, the precise signal sequence for this integration event in miR341 does not exist in the human ortholog of the locus.[454–456]

1.12.5 AAV vector purity:

Another significant challenge to the AAV gene therapy field is that of AAV prep purity. In the context of wildtype AAV, particles that did not contain full length viral genomes were noted in an early study of the virus.[457] When producing virus for clinical applications, the primary features of interest are viral yield and efficiency of transgene expression. However, purity of recombinant AAV is also a major concern due to how the virus is produced. Contamination of AAV preparations can arise in several ways. Firstly, protein contamination from the producer cell lines has been shown to impact therapeutic transgene activity post infection.[395] However, this concern is more applicable to research grade material, often performed by non-optimized ultracentrifugation methods. Clinical production of AAV to GMP standards requires more stringent purification methods, and host cell protein in preps is less. Secondly, AAV capsids do not require the packaging of

full viral genomes. In fact, the majority of AAV particles in a purified prep are non-expression cassette containing, i.e. empty capsids. These virions increase the viral antigen load of the administered virus but cannot provide a transducing genome. Often purified AAV will comprise up to 90% empty capsids, although certain purification methods can be employed to alter this ratio towards a higher level of full capsid particles.[458,459] There is disagreement within the literature about whether these empty capsids are detrimental to therapy administration or can be beneficial.[460] Some reports have shown that the presence of empty particles inhibits the transduction efficiency of the expression cassette.[406,461] One group however posited that empty capsids could be beneficial to gene therapy by developing a capsid mutant unable to enter host cells and adding these empty particles into an AAV prep to shield the transgene containing particles from humoral immunity.[462] Early studies of the virus also revealed heterogeneity within DNA containing particles, with DNA genome lengths ranging between 3% to 100% of the viral genome packaged, with shorter DNA fragments biased towards containing the inverted repeat sequences at the terminal end of the genome.[393] The issue of truncated vector genomes can also be exacerbated by the presence of secondary structure forming sequences within the ITR flanked expression cassette, including shRNA and sgRNA sequences.[463,464]

DNA contamination of the purified AAV can also occur; host cell DNA from the production cell line and DNA from the transfected producer plasmids can be found in purified AAV preps. Any contaminant DNA found outside of the viral particles can be readily removed in the purification process by enzymatic digestion. However, contaminant DNA fragments inside of the viral particle would be shielded from digestion, and therefore more difficult

to remove in the purification process. Previous studies have examined DNA contamination in AAV preps and have identified contaminating non expression cassette sequences from the producer plasmids and host cell DNA in purified AAV, in both 293 cell based and baculoviral AAV production methodologies.[465–468] It has been posited that at least a portion of this is due to ‘reverse packaging’ of the backbone DNA from the ITR containing expression cassette plasmid, and that this source of contamination can be reduced by increasing the backbone size of the expression cassette plasmid to over 5kb, as this exceeds the packaging capacity of AAV.[466] Another concern for the field is that the presence of contaminating sequence could yield to the production of replication competent AAV particles (rcAAV) post gene therapy. This concern was fuelled by the observation that REPCAP particles could be packaged in the absence of the inverted terminal repeats, albeit at low levels.[469] One approach to tackle this has been the development of ‘captron’ vectors. These are REPCAP plasmids in which the capsid gene of AAV is interrupted by a large intron, such that AAV can be produced normally but the entire capsid DNA sequence cannot be aberrantly incorporated into a single AAV particle.[470] Another is the physical separation of the REP and CAP genes onto different plasmids to form a “split packaging system” that can improve yields and yield AAV with rcAAV below detection limits.[471]

1.13 Aims of the project:

The research in this thesis will focus on contaminating DNA within AAV originating from the producer plasmids with two specific goals:

1. To characterise the profile of contaminant DNA sequences originating from AAV producer plasmids and assay their activity in cells post infection.
2. To develop a novel design of the AAV production plasmid system that will reduce producer plasmid related DNA impurities from AAV preparations.

Potential impact: The stated objectives could yield a greater understanding of determinants of AAV purity and key considerations with regards to vector design for clinical decision makers. Additionally, this work could further optimise the AAV vector toolkit to produce a higher purity product in a way that could be implemented to the AAV field in both the research and clinical setting.

Chapter 2

General Methods

2. General Methods:

2.1 Plasmid storage:

Plasmid constructs were kept in glycerol stocks. 100% Glycerol (Millipore Sigma Cat# G5516) was diluted to a working mixture of 50% glycerol with ddH20. Glycerol stocks were made with a 50:50 mixture of 50% glycerol and bacteria cultured in NZY media (Fisher Scientific Cat# BP24652) containing the desired plasmid. Plasmids were stored in stbl2 Max competency cells (ThermoFisher Cat# 10268019) and SURE2 Supercompetent cells. (Agilent Cat# 200152)

2.2 Plasmid purification:

Plasmids were grown in NZY media for 16 hours under either Kanamycin or Ampicillin selection. Plasmid purification from 6 well plate AAV production was performed by miniprep kit (Qiagen Cat# 27106) Plasmid for large scale AAV production was performed by Caesium Chloride gradient centrifugation. (See Appendix 1)

2.3 Construct Cloning:

Vector and insert gene fragments were digested overnight by restriction enzymes from New England Biolabs in CUTSMART buffer (New England Biolabs Cat# B7204S) at 37°C with shaking. Digested vector backbones were 5’ phosphate removed by Calf Alkaline phosphatase (New England Biolabs Cat# M0290) for 30 minutes. Digested plasmids were run on a 1% agarose gel containing 0.01% Ethidium bromide to separate the required fragments. Gel bands were extracted under UV light using a razor blade. Extracted fragments were gel extracted using a commercial kit (Qiagen Cat# 28706), and bands

were eluted in 35µL ddH_20. Vector:insert ligation reactions were prepared at a 1:3 ratio and constructs were ligated together using T4 ligase (Promega Cat# M1801) overnight at 4°C. Ligations reactions were transformed into stbl2 Max competency cells (ThermoFisher Cat# 10268019) and spread onto either Kanamycin resistant or Ampicillin resistant LB Agar plates. Colonies were picked and verified by digest for both size and, for AAV vector plasmids, ITR integrity. Successful ligations were validated by full plasmid illumina sequencing (Harvard biosequencing) prior to viral production.

P5 replacement promoter variants were cloned into the CR21AAV2_8 backbone by Not1 + Sal1 digest. P5-HS5 variants were cloned into CR21AAV2_8 backbone by Not1 + Sal1 digest. P5-HS5 variant was cloned into

2.4 PCR:

2.41 AAV particle PCR analysis:

Purified AAV2/8scLP1hFIXco was analysed by PCR to determine the presence of contaminant DNA in viral preparations. 50µL reactions were prepared in 0.2ml tubes. AAV2/8scLP1hFIXco virus or pAV2/8scLP1hFIXco plasmid at 5e8 vg or ds copies respectively were used as template DNA in each sample. 1µL of each primer (3µM), 1µL 10mM dNTPs (NEB cat# N0447S), and 5µL of *Taq* polymerase (NEB cat# MO273). Samples were made up to 50µL with DNase and RNase free ddH_20 (ThermoFisher cat# 10977015). The PCR machine was heated to 95°C for 5 minutes and samples were thermocycled 35X; 95°C for 15 seconds 55°C for 30 seconds; 72°C for 90 seconds, followed by a 120 second 72°C final extension phase and a 12°C indefinite hold. PCR

products were run on a 3% agarose gel and analysed in a Gel documentation system (Proteinsimple AlphaImager HP).

2.42 Genomic RNA extraction for PCR:

Mouse livers were harvested from mice infected with AAV, snap frozen in liquid nitrogen and stored at -80°C. To extract genomic RNA, frozen livers were pulverised with mortar and pestle, treated with RNASE free DNASE (Promega Cat# M6101) and purified by a commercial kit. (Qiagen Cat# 74104). Samples were quantitated for RNA by a NanoDrop OneC spectrophotometer (ThermoFisher). Whole liver RNA was converted to cDNA with random hexamer reverse transcription. 1µg of RNA was converted to cDNA using the Superscript III first strand synthesis system (Invitrogen). After cDNA generation 1µL of RNASE H, incubated at 37°C for 29 minutes and stored at -20°C.

2.43 Genomic RNA PCR:

For PCR analysis of genomic RNA 500ng of purified mouse liver genomic cDNA was loaded into 0.2ml tubes and subjected to 35 cycles of amplification at 95°C dissociation, 55°C annealing, 72°C extension. Plasmid DNA from the AAV2_8REPCAP was used as a positive control. Kanamycin resistance transcript primers - Forward: ATGATTGAACAAGATGGATTGCAC; Reverse: GAAGAACTCGTCAAGAAGGCG; Ampicillin resistance transcript primers – Forward: ATGAGTATTCAACATTTCCGTGTCGC; Reverse: CCAATGCTTAATCAGTGAGGCAC.

2.44 Quantitiative PCR:

2.44.1 DNA Standard preparation:

Standards were prepared from AAV production plasmids. Plasmids were linearised by restriction digest and purified by phenol chloroform extraction. Purified standards were resuspended in TE buffer (Promega Cat# V6231) to a concentration of 1e9 DNA copies per µL. Serial dilutions of 1:10 were made in ddH20 (Thermo Fischer Cat# 10977015) + 0.005% Pfu68 (Gibco Cat# 24040032). Standards of 5e8-5e4 were used for the assaying of purified AAV. Standards of 5e6-5e2 were used for the assaying of crude AAV from small scale (6 well) production. For vector genome plasmid analysis the plasmids pAVssHLPhFVIIIv3, pAVscLP1hFIXco+helpv3, and pAVssCMV_GFP were used as standards for FVII, FIX and GFP titer analysis respectively. For REPCAP contamination analysis AAV2_8REPCAP was used as a standard.

2.44.2 Master Mix Preparation:

Master mix for qPCR reactions were made up to contain (per sample) 12.5µL SYBR green (Applied Biosystems Cat# 4309155), 0.075µL of 100µM forward primer (IDT), 0.075µL of 100µM reverse primer, and 7.35µL ddH20 (Thermo Fischer Cat# 10977015).

2.44.3 Loading and analysis:

20µL of master mix and 5µL of sample or plasmid standard were loaded into individual wells of a 96 well plate (Applied Biosystems Cat# 4306737) and run on an Applied Biosystems 7500 Realtime PCR machine. (Cat# 4351107) Amplicon values were calculated against the curve generated by DNA standards to determine copy number of either expression cassette or contaminant amplicons.

2.5 Cell culture:

Cell lines were cultured in DMEM media (Corning Cat# 15-013-CV) supplemented with 10% FBS (GE Hyclone Cat# SH30071.03) and 2mM L-glutamine (Corning Cat#25-005-CI), referred to herein as D10 media. Cultured cells were incubated at 37°C 10% CO_2. Cells were passaged by the aspiration of media, followed by a wash with Phosphate buffered saline without calcium or magnesium (Lonza Biowhittaker – Cat# 17-516F) and addition of trypsin (Corning Cat#25-052-CI) for 5 minutes. Trypsinised cells were neutralised by D10 media and split to the required concentration into a D10 containing flask or dish. More sensitive cell lines were spun at 400g for 5 minutes and resuspended in D10 to remove trace trypsin. All cell lines used were identity confirmed by STR analysis and tested negative for mycoplasma.

2.6 AAV production:

Adherent 293T cells (ATCC CRL-3216) were used for all AAV production.

2.61 Large scale AAV production:

AAV8 FIX and FVIII viruses were produced by 2 plasmid transfections in CellSTACK culture chambers (Corning Cat #CLS3271) Prior to transfection adherent Human embryonic kidney (293T) cell line (ATCC CRL-3216) were cultured in DMEM (Lonza Bio Whittaker Catalog# 12-733Q) supplemented with 10% Fetal Bovine Serum (Fisher Scientific Catalog# SH3007103) and 2mM L-Glutamine (Corning Catalog# MT25005CI) at 37 ˚C 10% CO_2.

Transfection plasmids were as follows: A REP-CAP plasmid with AAV2_8 was used to provide the replication and capsid genes. scLP1hFIXco+helpV3 provided a self-complementary FIX expression cassette and adenoviral helper genes. ssHLPhFVIIIv3 provided a single stranded FVIII expression cassette with adenoviral helper genes. A full list of virus preparations and production plasmids is included in the supplemental data.

Transfection plasmids were resuspended in DMEM and passed through a 0.2μM filter into DMEM containing PEIpro (Polyplus Cat# 115). DNA-PEI mixture was inverted 5X, incubated for 15 minutes, and mixed with 800ml of media from CellSTACK culture chamber. Transfection mixture in media was then poured back into CellSTACK culture chamber and incubated at 37˚C 10% CO_2. Supernatants were harvested 7 days post transfection and concentrated to between 15-30ml using a multi-manifold of Pellicon XL 50 cassettes (Millipore-Sigma Cat# C1974). Concentrated supernatants were treated with Benzonase 25 units/ml (Millipore-Sigma Cat# E1014) in presence of 1mM $MgCl_2$ at 37°C for 1hr. Treated supernatants were passed through a 0.2μM filter, diluted in PBS and run on through a POROS CaptureSelect AAVX resin (Thermo Fisher Scientific Cat# A36740). Flow rate on column was set to 2ml/min. Column was equilibrated with 5 column volumes (CV) PBS pH7.4, AAV diluted supernatants applied to column. The column was then washed with 15CV PBS pH7.4 and eluted with 5CV of 0.1M Glycine-HCl, pH2.7 followed by 5CV of PBS pH7.4. Elution fractions of 4 ml were collected. Fractions were titered by qPCR and peak fractions were combined. Eluted AAV was concentrated by 100KDa Amicon filter (Millipore-Sigma Cat# C7715) and 0.25% rHA was added to the formulation. Virus was stored at 4°C short term and at -80°C long term.

2.62 Small scale AAV production:

293T cells were seeded in 6 well plates at a concentration of 800,000 cells per well. 24 hours after seeding cells were transfected with AAV production plasmids. In the context of FIX and FVIII 2 plasmid transfections were used.

2.7 AAV preparation for titer and contaminant analysis:

2.71 From small scale production tests:

2µL of Supernatant from 6 well plate small scale AAV production was pre-treated with 500U DNASE1 in NEBuffer3 (1X) in a reaction volume of 100µL. AAV were incubated at 37°C for 1 hour. Reactions were neutralized with 2µL 0.5M EDTA pH8 and incubated at 98°C for 10 minutes. Solution was cooled and capsid proteins were denatured via addition of 2µL 10% SDS, 2µL of 20mg/ml proteinase K (Ambion Cat# AM2546) and incubation at 55°C for 1 hour. Proteinase was inactivated at 98°C for 1 hour and 94µL of 0.01% Pfu68 (Gibco Catalog# 24040032) was added. Samples were then diluted 100-fold in ddH_20+0.01%Pfu68 to yield a total dilution of e-4 and analysed for production efficiency and contaminant amplicons by quantitative PCR on an Applied Biosytems 7500 machine.

2.72 From column purified AAV:

Large scale, column purified preps were serially diluted in ddh_20 to yield a total dilution of e-4 and analysed for production efficiency and contaminant amplicons by quantitative PCR on an Applied Biosytems 7500 machine.

2.8 In Vivo AAV injections:

For circulating protein experiments C57Bl/6 mice were injected by tail vein with 2e10vg of AAV in 100µL Phosphate buffered saline with 0.05% Pluronic acid. For high dose AAV contamination experiments C57Bl/6 mice were injected with 4.11e11vg in 100µL Phosphate buffered saline with 0.05% Pluronic acid.

2.9 ELISA:

20µL of plasma collected from CBL57/BL6 mice was used to assay for circulating protein levels. Circulating FIX protein was assayed by Asserachrom IX:Ag Enzyme immunoassay for FIX (Diagnostica stago Cat# 00943).

2.10 Western blot:

100µG protein was extracted from CBL57/BL6 mice or 40µG of protein extracted from Sure2 Competent *E.Coli* and separated by SDS-PAGE on 4-12% gels (Invitrogen NP0321PK2) in 1X running buffer (50ml 20X running buffer Invitrogen Cat# NP0002) + 950ml ddH_20). Separated protein was transferred onto nitrocellulose membrane in 1X transfer buffer (50 ml 20X transfer buffer (Invitrogen Cat# NP0006)) Membranes were blocked for 30 minutes in 5% milk in PBST (100ml 10XPBS (BioWhittaker Cat# 17-517Q) + 900ml ddH_20 +800µL Tween20 Cat# 85113). Blots were probed overnight at 4°C with either Anti-Neomycin Phosphotransferase 2 antibody (4B4D1) ab60018 or Anti-Beta Lactamase antibody (8A5.A10) (Abcam Cat# ab12251). Blots were washed 3X in PBST and incubated in Rabbit anti-mouse IgG secondary antibody (Abcam Cat# ab6728) for 1

hour. Blots were washed 3X in PBST, treated with enhanced chemiluminesence substrate (Thermo Scientific Cat# PI32106) and imaged on autoradiogaphy film (Thomas Scientific Cat# A8805).

2.11 FISH:

2.11.1 Probe preparation:

DNA for probes was prepared from plasmid DNA purified by a commercial kit. For samples that required specific sequence regions, plasmid DNA was digested overnight. Digested DNA was run on 1% agarose gel. Region band of correct size was extracted and purified by Qiagen kit. Purified DNA was then diluted, phenol chloroform extracted to increase purity, and resuspended in ddH20. Denatured probes were prepared by nick translation and resuspended in buffer consisting of 50% formamide, 2X saline-sodium citrate, and 10% dextran sulfate.[472]

2.11.2 Slide preparation – cell lines:

Cell line samples were media aspirated, washed in 1XPBS, treated with trypsin and incubated for 5 minutes at 37°C 10% CO_2. Trypsinised cell solution was neutralised with D10 media and spun down at 400g for 5 minutes. Cell sample was resuspended at between 50,000-100,000 cells total in 140µL 1XPBS. Cell resuspensions were transferred into a cytospin sample chamber and spun down at 400g onto glass cytospin slides (ThermoFisher Cat#5991059) using a cytospin cytocentrigfuge (ThermoFisher Cat# A78300003). Slides were incubated in 1% PFA in PBS + 0.05% NP-40 (MP Biomedicals Cat# RIST1315) for 5 minutes, then in 1% PFA in PBS for 5 minutes, and then in 70% EtOH for 5 minutes before transfer to storage in 70% EtOH at -20°C.

2.11.3 Slide preparation – Mouse liver:

Fresh liver samples were cut with a razor blade to expose a fresh edge and touched onto slides (Thermo-Scientific Cat# 5991051). Slides were incubated in 1% PFA in PBS + 0.05% NP-40 (MP Biomedicals Cat# RIST1315) for 5 minutes, then in 1% PFA in PBS for 5 minutes, and then in 70% EtOH for 5 minutes before transfer to storage in 70% EtOH at -20°C.

2.11.4 Slide treatment RNA FISH:

Fixed slides were RNA hybridized by dehydration in EtOH. Slides were sequentially treated for 2 minutes each in 70%, 80%, and 100% EtOH solutions. Slides were then dried and probed for either expression cassette RNA or contaminant RNA. Denatured probes were prepared by nick translation and resuspended in buffer consisting of 50% formamide, 2X saline-sodium citrate, and 10% dextran sulfate.[472] Probe was applied to slides and hybridized at 37ºC overnight. Slides were then washed in 50% formamide and 2X saline-sodium citrate at 37ºC for 5 minutes. Slides were mounted in Vectashield mounting medium. RNA FISH images were collected and analysed as previously described.[418]

For sequential DNA-FISH, slides were treated in 4% PFA, 0.5% tween 20, and 0.5% NP-40 for 10 minutes at RT, and then treated in 0.2N HCl, 0.5% triton X 100 for 10 minutes on ice. Following fixation, slides were denatured in 70% formamide 2X saline-sodium citrate at 80ºC for 10 minutes. Slides were then dehydrated in a graded alcohol series for

2 minutes each as for RNA FISH. Denatured probe (same labelled DNA as was used for RNA detection) was then applied to the slides and hybridized overnight at 37ºC.[472] Washing and mounting of slides is the same as for RNA FISH. DNA FISH images were collected and analysed as previously described.[418]

2.12 IHC:

Fresh CBL7/BL6 mouse liver was cut with a scalpel and placed into a plastic cassette. Liver tissue was fixed by submerging in formalin. Slides were cut into 4µm sections and stained for the presence of GFP protein.

2.13 Alkaline gel:

Virus samples were run on 0.8% alkaline gels. Alkaline running buffer (50X) was a solution of 2.5M NaOH and 0.5M EDTA. Gels were prepared by adding 1.2g Agarose (Invitrogen Cat# 16500100) to 147ml ddH_20. Solution was heated to dissolve agarose. 3ml of 50X Alkaline running buffer was added to dissolved solution and mixed. Solution was cooled to 50°C and poured into a gel tray with comb. Viral samples were prepared to required concentration by diluting in PBS to a volume of 62.5µL. 21.5µL of 4X Alkaline loading buffer was added to viral sample (20% glycerol (Millipore Sigma Cat# G5516) 1.2% SDS 4X Alkaline running buffer 1.75mM Alizarin Yellow (Millipore Sigma Cat# 206709) in 4X Alkaline running buffer.) Viral samples were run on gel for 18 hours at 15V. Gels were neutralised in neutralisation buffer (0.1M Tris-HCL pH8.0) for 1 hour. Neutralisation buffer was discarded, and gels were stained in staining buffer (1X GelRed (Biotium Cat# 41003) 0.1M NaCl) and imaged on a Gel documentation system (Proteinsimple AlphaImager HP).

2.14 FLOW:

2.14.1 Preparation of cells for GFP FLOW:

Media was aspirated from cells in a 24 well tissue culture plate and rinsed with 1XPBS (Lonza Biowhittaker – Cat# 17-516F). 150µL of 0.5% Trypsin (Corning Cat# 25-052-CI) was added to each required well. Plates were incubated at 37°C 10% CO_2 for 5 minutes. Trypsin was neutralised with 1ml D10 media. Neutralised cells were transferred to FACS tubes (Corning Cat# 352054). 2ml PBS was added to each tube and tubes were centrifugated at 400g for 5 minutes. Supernatant was aspirates and 3 ml of PBS added to each tube. Tubes were again centrifugated at 400g for 5 minutes. Supernatant was removed and cells were resuspended in PBS. Tubes of cells were placed on ice until being run on a BD LSR Fortessa FACS machine.

2.14.2 FLOW:

Cell samples for FLOW were run on a BD Fortessa. 10,000 events were collected per sample. FCS files were analysed by FLOWJO v10.

2.15 Next generation sequencing:

2.15.1 AAV prep DNA seq analysis:

1e11vg AAV genomes were diluted in PBS to a volume of 200uL and capsids were denatured at 95°C for 15 minutes. Samples were cooled on ice and 117µL of 100%ETOH was added to precipitate DNA. Viral DNA was extracted with a QIAamp MinElute Virus spin kit (QIAGEN Cat# 57704) and eluted in 25µL ddH_20. Purified DNA fragments were

then subjected to NexteraXT DNA sequencing using 10 million paired end reads 100bp in length. Sequencing reads were then mapped back to the AAV production plasmids.

2.15.2 Genomic DNAseq analysis:

Purified gDNA was library prepped and submitted to NexteraXT sequencing HiSeq 100 million reads. CBL57Bl/6 specific reads were removed, and remaining sequence reads were mapped back to producer plasmids and analysed by IGV and circos to form contamination plots.

2.15.3 Genomic RNAseq analysis:

Whole liver RNA from C57Bl/6 mice injected with AAV8scLP1_hFIXco or AAV8ssHLP_hFVIIIv3 were subjected to high throughput RNA sequencing. cDNA libraries were constructed from total liver RNA. Sequencing reads matching the C57Bl/6 mouse genome were removed and the remaining sequences were mapped against the production plasmid sequence constructs.

2.16 Schematics, Graphs and statistical analysis:

Schematics, image and figure assembly were performed in Adobe illustrator creative cloud. Image brightening was performed on Adobe Photoshop Creative cloud. Graphs and statistics were generated in Graphpad prism 8.0. All error bars represent Standard error of the mean (S.E.M.)

Chapter 3

Presence and post infection activity of adeno-associated virus DNA contaminants

3. Presence and post infection activity of adeno-associated virus DNA contaminants

3.1 Introduction:

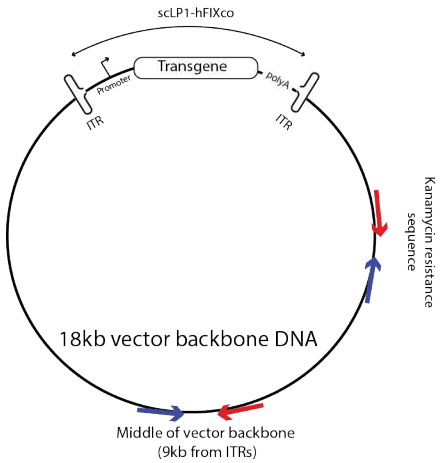
Adeno-associated virus is an invaluable tool for the clinical implementation of gene therapy. A contributing factor to the widespread use both of AAV in emerging clinical trials and in preclinical research worldwide is the ease of production using materials common to biological science laboratories. In a matter of weeks, a research team can produce, and purify recombinant virus to test the efficacy of gene delivery on a disease model of interest. For clinical grade production these same techniques are used, albeit with greater scale and stringency. This means that an exploration of the AAV production system at the research grade level can lead to insights that translate into the improvement of clinical production techniques. Factors influencing vector yield are always desirable to reducing production costs and are very important to the advancement to the field. Another facet of the AAV production system that is equally important is product characterisation. Gene therapy has had a storied history with unintended consequences, that at one point tempered wider enthusiasm for the modality altogether.[473,474] As the clinical expansion of AAV continues to greater patient numbers and types of diseases, it is imperative for the field to research potential downstream effects or unintended consequences of AAV delivery in a proactive manner. One such area of research is the packaging of contaminant nucleic acids that are not part of the designed expression cassette. This chapter aims to examine contaminant DNA sequences in AAV in greater detail than prior studies, focusing on the influence of the AAV producer plasmids and to examine any activity that contaminant sequences exhibit post AAV infection.

3.2 Results:

3.2.1 AAV vectors contain a distinct DNA contamination profile from producer plasmids.

Previous research into DNA contaminants of AAV has shown that sequences from the plasmid backbone are detectable in purified preps.[465–468] Despite this, relatively little research has been done to elucidate the causes and effects of this contamination. To assay whether contaminating sequences outside of the ITRs in the vector plasmid were present in purified AAV preps that utilised a large plasmid backbone, a PCR strategy was designed. Primers were designed to amplify regions in the vector plasmid backbone (Fig. 3.1). Amplicons corresponding to a sequence in the kanamycin resistance gene and a region in the plasmid 9kb from either ITR were detected in purified AAV8scLP1hFIXco (Fig. 3.1). To examine if these DNA contaminants were within the capsid, a forward primer was designed to span the ITR-backbone junction and read out into the plasmid DNA backbone. A series of reverse primers at increasing distances from this junction were also designed (Fig. 3.2). PCR amplification of the target amplicons was performed on purified AAV8scLP1hFIXco in the presence or absence of DNaseI in AAV with intact or denatured capsids (Fig. 3.2). Amplicon sequences reading out to 721bp were detected with or without DNaseI treatment when capsids were intact. The 1kb amplicon was not detected in the intact capsid samples, which could be due to either lack of presence or lack of template amplification under the reaction conditions. When capsids were denatured however, contaminant amplicons were only detected in the untreated sample, suggesting that contaminant DNA species of large size from the AAV production plasmids are present within the AAV capsid, and therefore at risk of transfer to infected cells.

a



b

Forward primer	Reverse primer	Amplicon length	Fragment
5'-TTAAGGACTAGTTTCGCGCC-3'	5'-CTGCGCCCCTGAGTTAG-3'	418bp	Vector backbone mid
5'-GCCAGTCCCTTCCCGCTTCAGTG-3'	5'-CCGCTTGGGTGGAGAGGCTATT-3'	147bp	Kan^r gene fragment

c

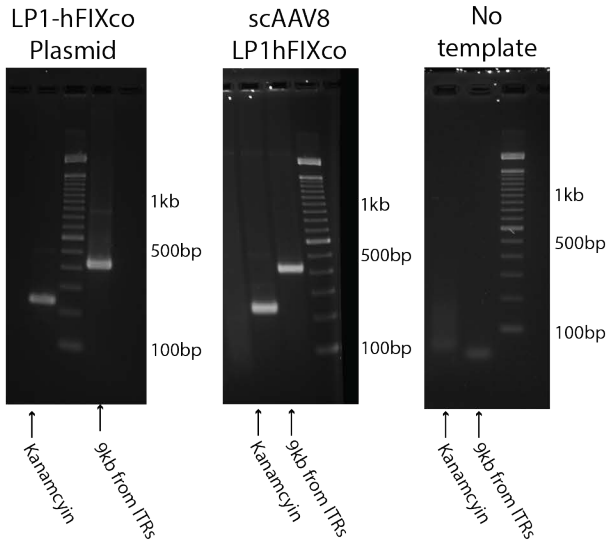
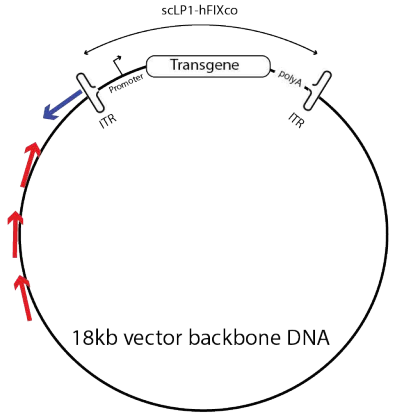


Figure 3.1 – Amplification of contaminant sequences in purified AAV.

(a) Schematic; primer locations for vector backbone amplification from AAV **(b)** table of primers for PCR amplification **(c)** Amplicons in the plasmid backbone sequence were amplified by PCR in samples containing the LP1hFIXco plasmid (left); the purified LP1hFIXco virus (middle) or no template (right).

a

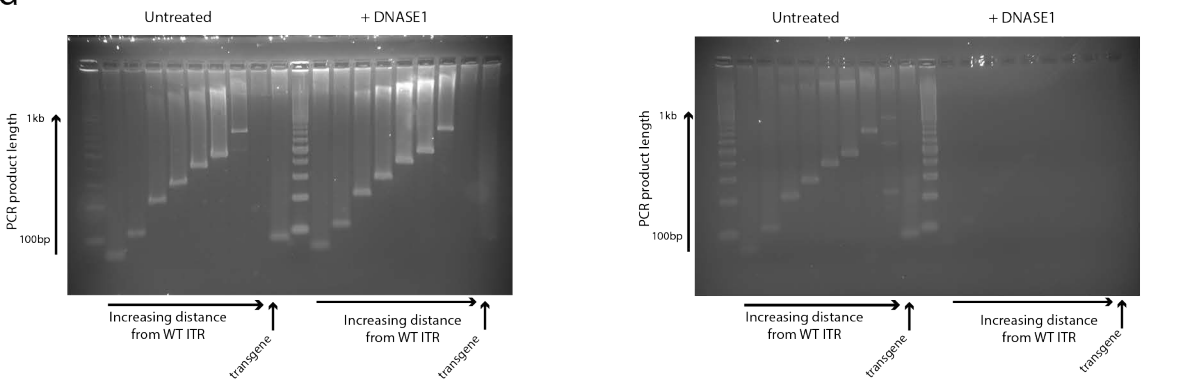


b

Forward primer	Reverse primer	Amplicon length
5'-GAGTGGCCCAGATCCC-3'	5'-CGGACCCAAGGACATGGT-3'	64bp
5'-GAGTGGCCCAGATCCC-3'	5'-CAAAACGAAGCCTGGGG-3'	111bp
5'-GAGTGGCCCAGATCCC-3'	5'-CCGCAGCGATAGATGCA-3'	212bp
5'-GAGTGGCCCAGATCCC-3'	5'-CTTCAGCCGATGAGGGG-3'	288bp
5'-GAGTGGCCCAGATCCC-3'	5'-GCGCATACCACCGCTTT-3'	389bp
5'-GAGTGGCCCAGATCCC-3'	5'-GTACACCGAGCTCTCGCAG-3'	469bp
5'-GAGTGGCCCAGATCCC-3'	5'-GCTGCGCAACATCTGGA-3'	721bp
5'-GAGTGGCCCAGATCCC-3'	5'-AAGCTAGCGCAGCAGC-3'	1010bp
5'-GGGCAAGTATGGCATCTACA-3'	5'-AAAGCATCGAGTCAGGTCAG-3'	83bp

c

d



Intact capsids

Heat denatured + Proteinase K treated capsids

Figure 3.2 – DNASE resistance of AAV contaminant amplicons

(a) Schematic; strategy to amplify increasingly large fragments of DNA outside of AAV **(b)** table of primers for PCR amplification **(c)** Schematic; AAV preps treated with/ without proteinase k and subsequently with/without DNASE 1 before vector plasmid backbone amplification. **(d)** Agarose Gels of PCR amplified vector plasmid backbone from Intact capsids treated with/without DNASE1 (left) and denatured capsids treated with/without DNASE1 (right).

Now that the presence of contaminant sequences in AAV had been identified, the next step was to determine whether certain sequences would be packaged at a higher rate than others. There are 3 REP binding sites in the WT AAV genome. In normal AAV replication these REP binding sites serve as initiation points for AAV genome replication, in each inverted terminal repeat sequence, and in the P5 promoter. We hypothesised that sequences near these REP binding sequences would be packaged into AAV at higher rates than sequences at greater distances. In the recombinant system used to produce this virus, these inverted terminal repeats remain flanking the expression cassette of interest, whereas the P5 promoter is used to drive REP78 and REP68 protein expression in trans. These elements were transfected along with the required adenoviral helper genes into the 293T producer cells to produce recombinant AAV (Fig. 3.3).

To assay the abundance of contaminant sequences in purified AAV, primers to generate amplicons reading into the vector backbone of the ITR containing expression cassette plasmid were designed (Fig 3.4.a-b). Purified AAV8scLP1hFIXco was analysed for the presence of contaminant backbone sequences. Amplicons close to the examined ITR were present at greater levels than those further from the ITR sequence in a distance dependent manner (Fig 3.4.e). The P5 promoter contains a REP binding site with a GCTC binding motif. In the context of AAV replication this site allows transactivation of the P19 promoter through positioning of the YY1-60 site in proximity to P19.[175] To examine if contamination was also initiated upstream of the P5 promoter, a series of amplicons at increasing distance from the P5 promoter were designed (Fig 3.4.c-d).

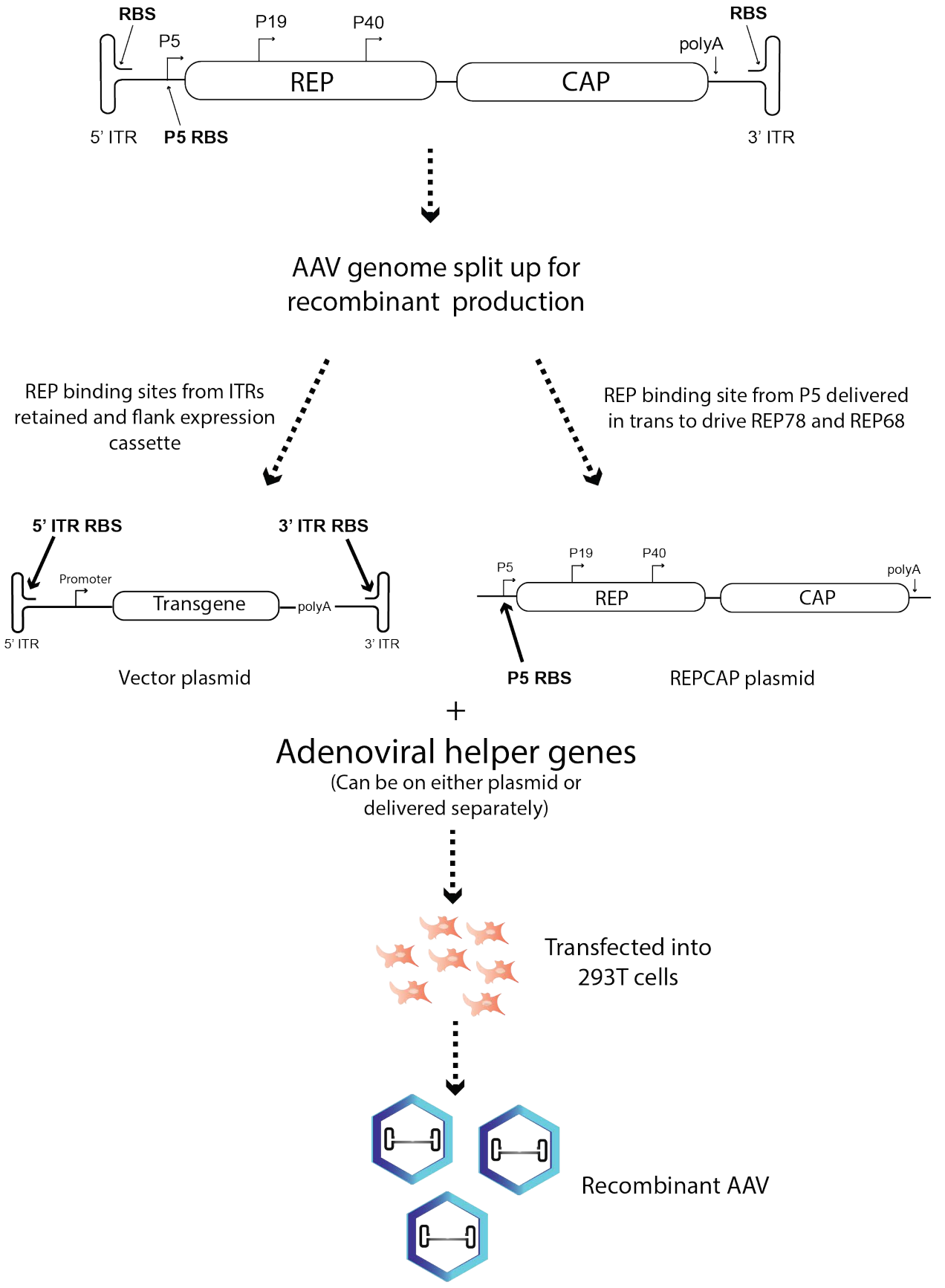


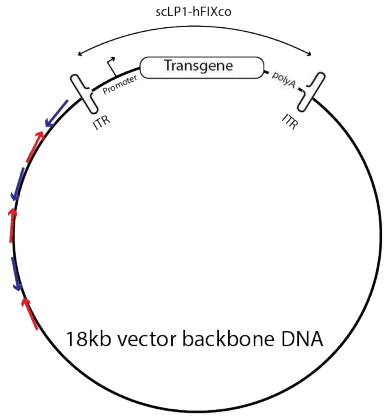
Figure 3.3 REP binding site position of recombinant AAV packaging system used in contaminant studies

Schematic; Positioning of REP binding sites in the WT AAV genome, found within each inverted terminal repeat and in the P5 promoter (top). Positioning of REP binding sites within recombinant packaging system (middle), 2 REP binding sites are present in the vector genome plasmid (1 at each ITR) and 1 REP binding site is present in the *in-trans* REPCAP plasmid (In the P5 promoter). Plasmids are cotransfected with Adenoviral genes into 293T cells to produce recombinant AAV

Contaminant presence was analysed by qPCR in AAV8ssHLPhFVIII and AAV8scLP1hFIXco. In both cases sequences close and upstream to the P5 promoter were present at greater levels than sequences at greater distance from P5 (Fig 3.4.e). To verify the contamination phenotype upstream of P5 and outside of the ITRs across different production runs; 3 individual preps of AAV8ssHLPhFVIII were assayed for 3 different contaminant amplicons, approximately equidistant from the hypothesised contamination initiation points. (Figure 3.5).

A further analysis of contaminants upstream of P5 was conducted on purified AAV8scLP1FIXco. When qPCR amplicons at different distances from P5 were analysed, the relative amount of incorporation seen decreased with distance from P5. (Figure 3.6.a-c). To ensure this was a phenotype related to the presence of the P5 promoter, a REPCAP configuration in which the P5 promoter was moved approximately 2kb upstream, so It was directly downstream of the CAP gene, was used to produce AAV8scLP1FIXco. In this 'P5 moved' configuration, the percentage of CAP gene contamination detected by qPCR went up, whereas the sequences that were no longer upstream of P5 declined (Figure 3.6.d-g). These amplicons were designed in such a way that the closer and further amplicon in each setup was equidistant to the closer and further amplicon in the converse setup. In both preps there is a visible decrease between the close and far amplicon in each setup, and that the 'P5 moved' configuration appears to have a lower level of contamination at the closer amplicon than the original P5 configuration.

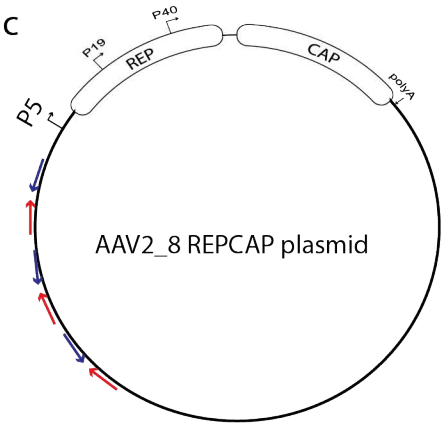
a



b

Forward primer	Reverse primer	Distance fromWT ITR (bp)
5'-GAGTGGCCCAGATCCC-3'	5'-CGGACCCAAGGACATGGT-3'	56
5'-TGCATCTATCGCTGCGG-3'	5'-CTTCAGCCGATGAGGGG-3'	280
5'-CCCCTCATCGGCTGAAG-3'	5'-GCGCATACCACCGCTTT-3'	381
5'-AAAGCGGTGGTATGCGC-3'	5'-GTACACCGAGCTCTCGCAG-3'	461
5'-GTATCCCACCAAAAAGTGCG-3'	5'-GCTGCGCAACATCTGGA-3'	713
5'-GTTCGAGCCCCGTATCC-3'	5'-AAGCTAGCGCAGCAGC-3'	1002
5'-AAACAGGTTCTGCCGCG-3'	5'-CTACCTGGACTTGGAGGAGG-3'	1532
5'-TTGCGGGCGTAAAACTTGC-3'	5'-ATAGTAGAGCCCGAGGGCC-3'	2080
5'-GATCAGAAACATCACCGCTTAGT-3'	5'-GCGGCGTGGAGGAATAT-3'	2501
5'-CTTTGCGGTGTCTCAGTCAC-3'	5'-GGTTGCACTAAACGCCTTC-3'	3036
5'-ATGGACATAAAGTTACTGGAGAATGG-3'	5'-ATTACAAGTGTCAGCCGCG-3'	3488

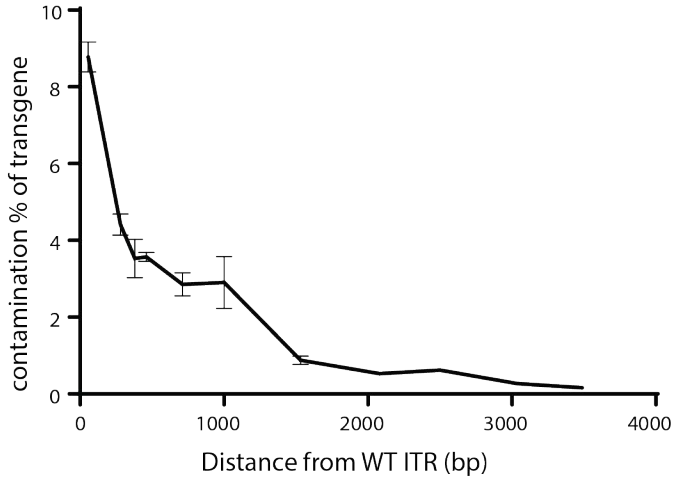
c



d

Forward primer	Reverse primer	Distance from P5 (bp)
5'-CCTTCAACCAGTCAAAGCTGAACT-3'	5'-CCAGCGCTTGCTGTTTTCC-3'	421
5'-GTATGGTCTGGCAGAACCGG-3'	5'-GCAGGTACAGGCGTGTTCTTG-3'	592
5'-CTCAAACAACAGGAGGCACGG-3'	5'-CAGCCAGTTCTTTGCCTGATTGG-3'	1057
5'-CCAACGACAACACCTACTTCGG-3'	5'-GATGAGTCGCTGCCAGTCAC-3'	1600
5'-GAGCGTCTGCAAGAAGATACGT-3'	5'-GCTCTACCGGTCTCTTCTTTCCAG-3'	2105
5'-GCAGCCATCGACGTCAGAC-3'	5'-CTTTCTGTCCGTGAGTGAAGCAG-3'	2752
5'-CTCAACGACCTTCGAACACCAG-3'	5'-GAATTCATGCTCCACCTCAACCAC-3'	2998
5'-GGAGCGTAAACGGTTGGTGG-3'	5'-GTTCTGCTCCTGCGTCTG-3'	3737

e



f

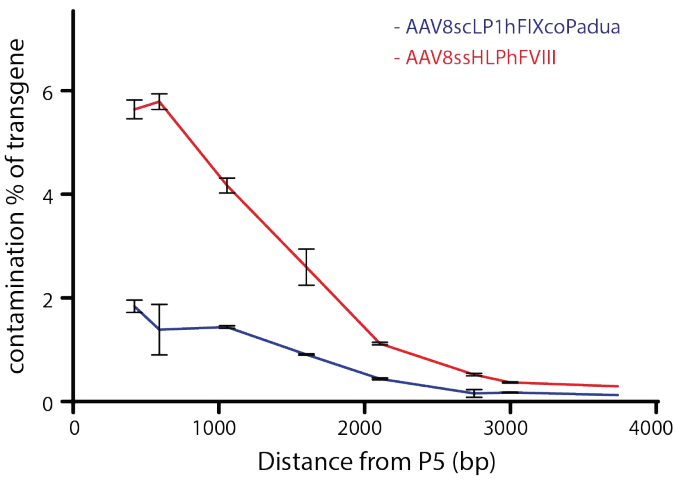
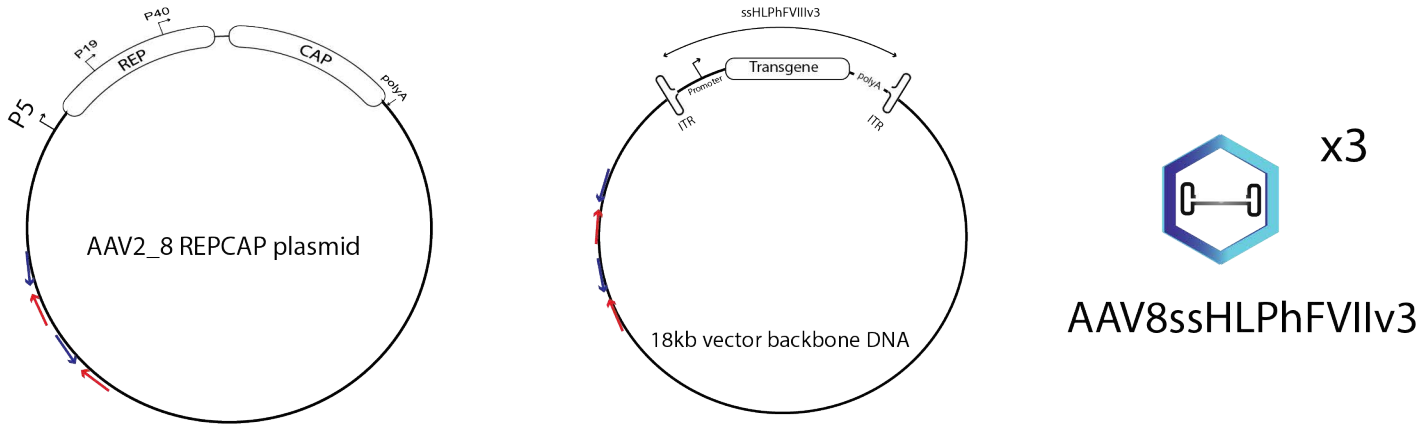


Figure 3.4 – AAV DNA contaminant sequence abundance depends on distance from REP binding sites

(a) Map of scLP1hFIXco vector plasmid used for assaying DNA contamination in AAV8scLP1hFIXco preparation. **(b)** Table of amplicon primers with distance (bp) from ITR sequence. **(c)** Map of AAV2_8 REPCAP plasmid used for assaying DNA contamination in AAV8 preparations. **(d)** Table of amplicon primers with distance (bp) from P5 sequence. **(e)** Graph showing results of qPCR of amplicons measuring abundance (*y axis* – expressed as a percentage of the expression cassette titer) plotted against distance from the ITRs (*x axis* – bp from outside end of ITR). **(f)** Graph showing results of qPCR of amplicons measuring abundance (*y axis* – expressed as a percentage of the expression cassette titer) plotted against distance from P5 (*x axis* – bp upstream of P5).

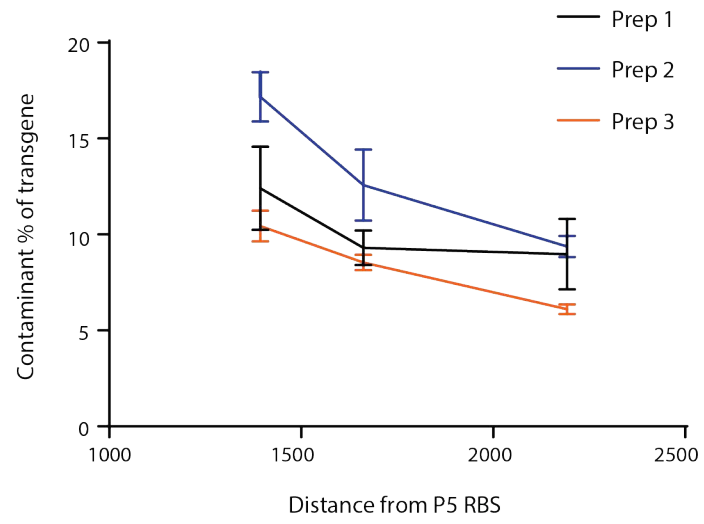
a



b

Forward primer	Reverse primer	Distance from P5 (bp)
5′-GCTGAGGTGGAGGACAACAT-3′	*5′-TCTTGGTTTCATTGGGCTTC-3′*	*transgene*
5′-GGGCGCCCGGTTCTTTTTGTC-3′	5′-GCCAGTCCCTTCCCGCTTCAGTG-3′	1394
5′-GGTCTTGTCGATCAGGATG-3′	5′-GAGCCTGGCGAACAGTTC-3′	1662
5′-TGAGAGTTTTCGCCCCG-3′	5′-GGGATAATACCGCGCCACATAGC-3′	2193
Forward primer	**Reverse primer**	**Distance from ITR (bp)**
5′-AAACAGGTTCTGCCGCG-3′	5′-CTACCTGGACTTGGAGGAGG-3′	1523
5′-TTGCGGGCGTAAAACTTGC-3′	5′-ATAGTAGAGCCCGAGGGCC-3′	2071
5′-GATCAGAAACATCACCGCTTAGT-3′	5′-GCGGCGTGGAGGAATAT-3′	2449

c



d

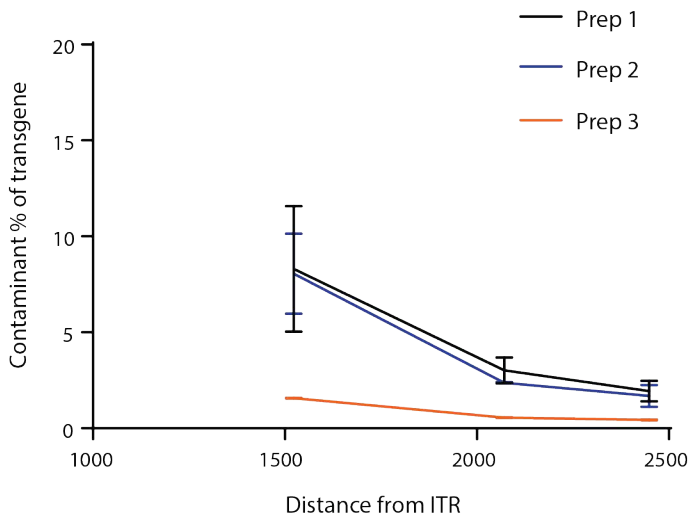
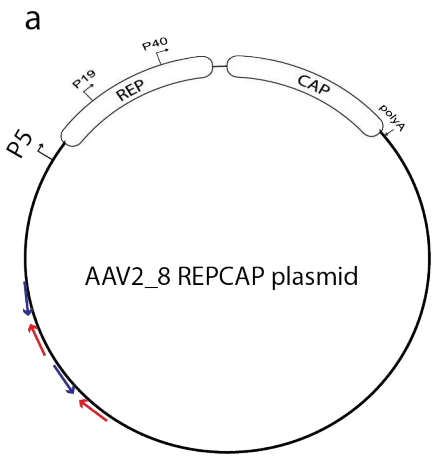


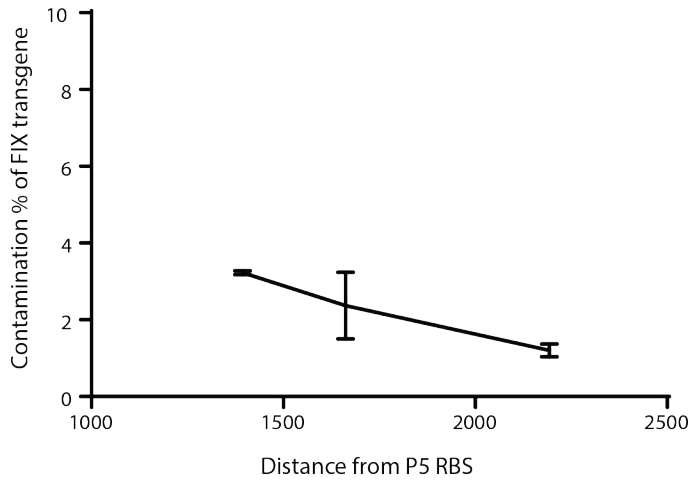
Figure 3.5 – AAV DNA contaminant sequence abundance depends on distance from REP binding sites

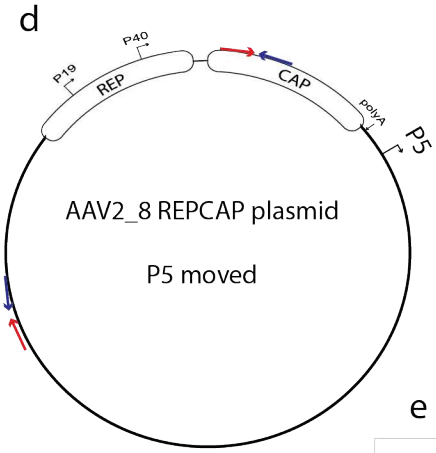
(a) Maps of AAV2_8 REPCAP plasmid (left) and ssHLPhFVIII vector plasmid (right) used for assaying DNA contamination in AAV8ssHLPhFVIII preparations. **(b)** Table of amplicon primers with distance (bp) from P5 and ITR sequence. **(c)** Graph showing results of qPCR of amplicons measuring abundance (*y axis* – expressed as a percentage of the expression cassette titer) plotted against distanceupstream from P5 (*x axis* – bp upstream of P5). **(d)** Graph showing results of qPCR of amplicons measuring abundance (*y axis* – expressed as a percentage of the expression cassette titer) plotted against distance from the ITRs (*x axis* – bp from outside end of ITR).



b

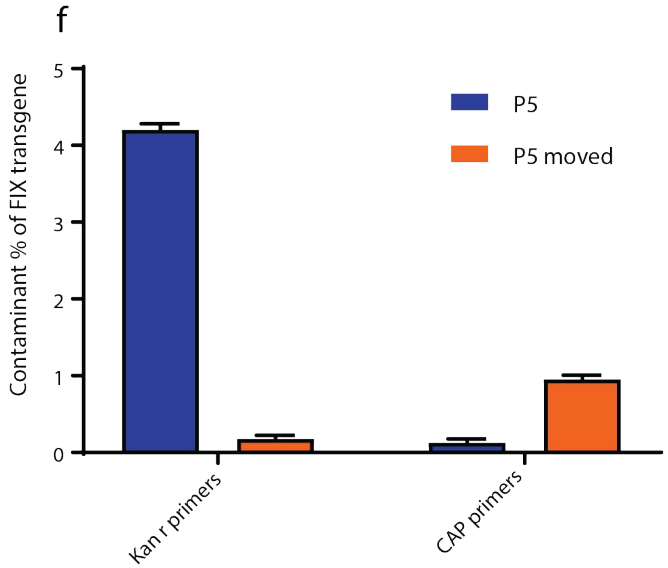
Forward primer	Reverse primer	Distance from P5 (bp)
5′-ATTTTATGTTTCAGGTTCAGGGGGAGGTG-3′	5′-GCGCAGAGAGGGAGTGGACTAGT-3′	transgene
5′-GGGCGCCCGGTTCTTTTTGTC-3′	5′-GCCAGTCCCTTCCCGCTTCAGTG-3′	1394bp
5′-GGTCTTGTCGATCAGGATG-3′	5′-GAGCCTGGCGAACAGTTC-3′	1662bp
5′-TGAGAGTTTTCGCCCCG-3′	5′-GGGATAATACCGCGCCACATAGC-3′	2193bp





e

Forward primer	Reverse primer	Distance from P5 (bp)	Distance from P5 moved (bp)
5′-ATTTTATGTTTCAGGTTCAGGGGGAGGTG-3′	5′-GCGCAGAGAGGGAGTGGACTAGT-3′	transgene	transgene
5′-GGGCGCCCGGTTCTTTTTGTC-3′	5′-GCCAGTCCCTTCCCGCTTCAGTG-3′	1394	5406
5′-CATCCAGGTCAAGGAGGTCACG-3′	5′-GGTGCTGGTGAGGTTATTGGCG-3′	5432	1419



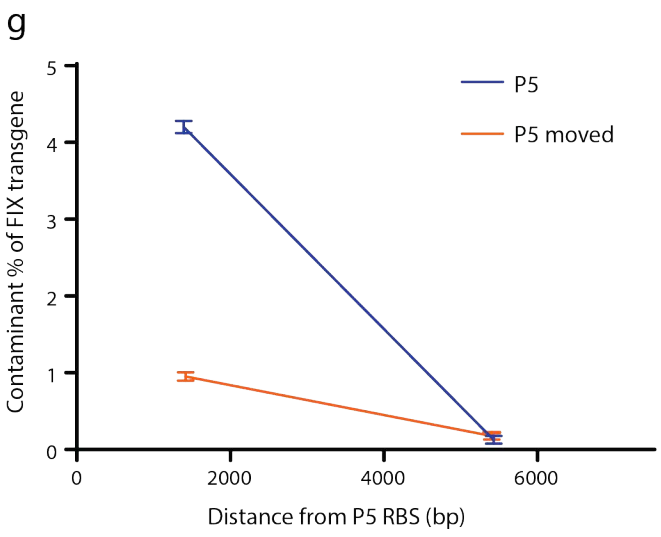


Figure 3.6 – Contaminant presence upstream of P5 is retained when P5 promoter is moved downstream of CAP gene

(a) Map of AAV2_8 REPCAP plasmid used for assaying DNA contamination in AAV8scLP1hFIXco preparations. **(b)** Table of amplicon primers with distance (bp) from P5 sequence. **(c)** Graph showing results of qPCR of contaminant amplicons from AAV8scLP1hFIXco measuring abundance (*y axis* – expressed as a percentage of the expression cassette titer) plotted against distance from P5 (*x axis* – bp upstream of P5). **(d)** Map of AAV2_8 REPCAP plasmid with moved P5 promoter. **(e)** Table of amplicon primers with distance (bp) from P5 sequence in the different AAV2_8 configurations. **(f)** Bar graph showing results of qPCR of amplicons measuring abundance (*y axis* – expressed as a percentage of the expression cassette titer) **(g)** Graph showing results of qPCR of amplicons measuring abundance (*y axis* – expressed as a percentage of the expression cassette titer) plotted against distance from P5 in standard (blue) and P5 moved (orange) configurations (*x axis* – bp upstream of P5).

To verify the DNA contamination profile identified by qPCR, AAV preps were DNA purified and submitted for next generation sequencing. Once sequencing was complete, reads were mapped back to the producer plasmid sequences (Fig. 3.7). As expected, most reads mapped to the ITR flanked expression cassette. A significant number of reads were also detected that mapped to non-expression cassette producer plasmid sequences. In the AAV8scLP1hFIXco prep 98.1% of the sequence reads mapped to the expression cassette and 1.9% of the reads were contaminants (Fig. 3.7.a). Of these contaminant sequences 1.56% mapped to the vector genome plasmid and 0.33% mapped to the REPCAP plasmid. From these 0.33% REPCAP plasmid reads 0.01% mapped to the REP gene of AAV2 and 0.32% mapped across the rest of the plasmid (REPCAP_Rest). In the AAV8ssHLPhFVIIIv3 prep 87.04% of the sequence reads mapped to the expression cassette and 12.96% of the reads mapped to contaminant sequences, with 8.70% of the reads mapping to the vector genome plasmid backbone and 4.26% of the contaminant reads mapping to the REPCAP plasmid, of which 0.03% of total sequences mapped to the REP gene. (Fig 3.7.b). Via circos plot mapping it was shown that sequences closely upstream of the P5 promoter were present at greater abundance than those further away (Fig 3.7.c-d). Interestingly sequences downstream of the P5 promoter did not appear to be actively incorporated into AAV, consistent with the described asymmetric nature of REP driven DNA replication.[475] The detected proportion of REP gene contamination was 0.01% and 0.03% in the FIX and FVIII preps respectively. As with the qPCR assay, to show that the contamination incorporation was linked to the P5 promoter and not another portion of the AAV genome the P5 promoter was moved approximately 2kb upstream, a configuration previously shown to still produce AAV titers effectively.[316] When AAV

produced using this method was deep sequenced, the map reads show that the contamination profile from this plasmid moves to match the repositioning of the P5 promoter (Fig. 3.7.e).

Next, to assess whether this contamination initiation was limited to WT ITR configurations, the self-complementary ITR of AAV8 FIX preparations, which contains a deletion in the TRS sequence required for DNA nicking by REP, was examined.[476,477] (Fig 3.8.a). qPCR analysis of amplicons at increasing distance from the self-complementary ITR showed that contamination still appeared to be present as a function of distance, but at a lower level than at a WT ITR (Fig 3.8.b). It is possible that differences in amplicon amplification efficiency could confound this result, and so a universal sequence was cloned upstream of the P5 promoter on the REPCAP plasmid, and an isolated single stranded or self-complementary ITR (Fig 3.8.c). In each instance adenoviral helper genes and REPCAP genes were provided to assist AAV production and then a single amplicon, universal and equidistant from each REP binding element was analysed by qPCR (Fig 3.8.d). When the amplicon was upstream of the P5 sequence it was most abundantly produced (2.31e9 copies), whereas the WT ITR produced over a log higher contaminant copies than the self-complementary ITR (WT ITR - 6.32e8; scITR – 2.65e7). To validate this difference; deep sequencing was conducted on 2 self-complementary AAV preps and the contamination profile directly outside of the ITRs was examined. Contamination read coverage outside of the self-complementary ITR was far lower than at the WT ITR (Fig 3.8.e). In the first prep, the read coverage at the 750bp outside the scITR was 43 times lower than the WT ITR in the first prep and 287 times lower in the second (Fig 3.8.f). When a FVIII prep with two WT ITRs was compared for read coverage a difference of

3.7x was seen between the two ITRs in read coverage (Fig 3.8.g). Combined, these results suggest there is a far lower contribution to overall contamination originating from self-complementary ITRs.

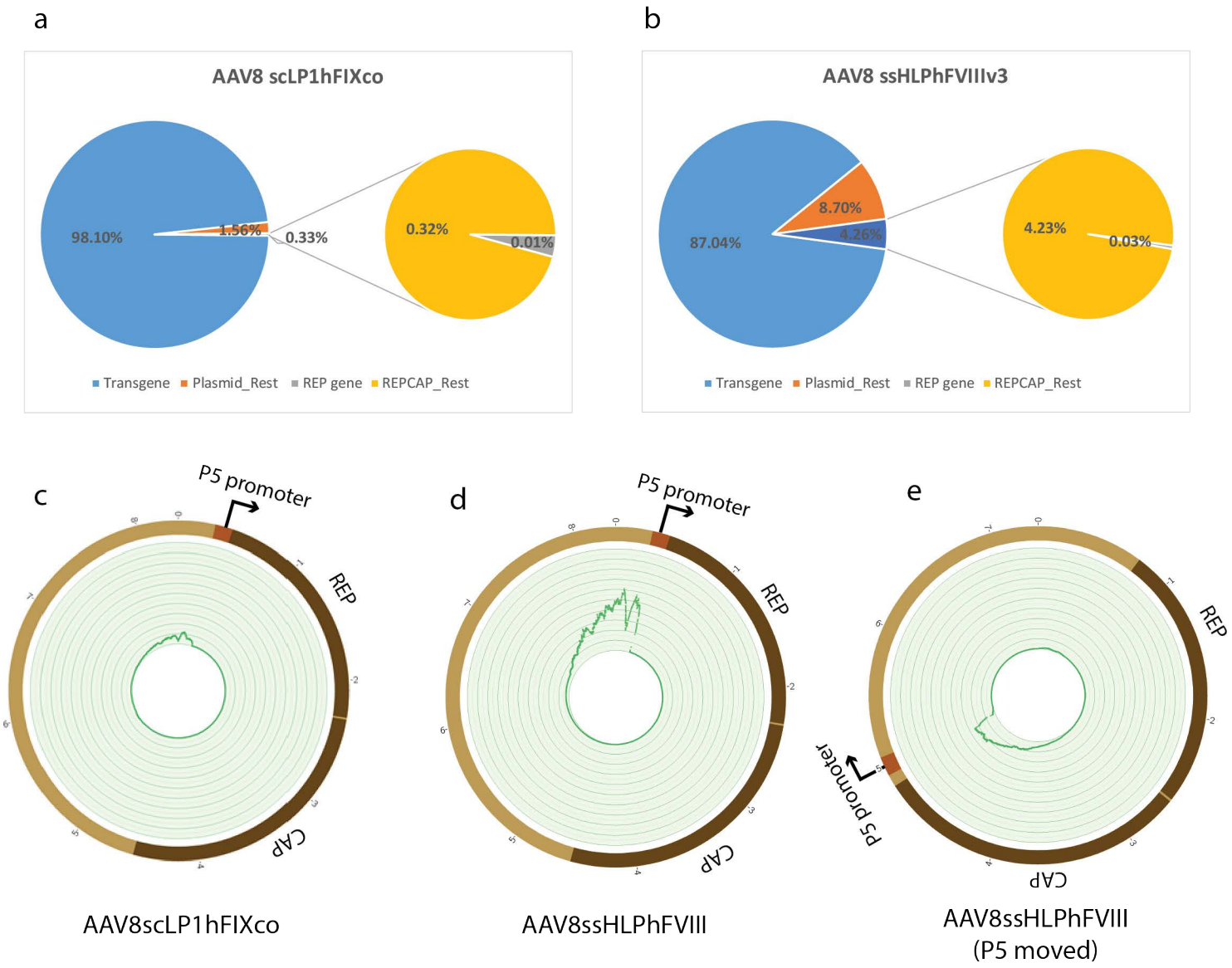


Figure 3.7 – Deep sequencing of AAV prep shows contaminant contribution from vector genome and REPCAP plasmid.

DNA purified from AAV FIX and FVIII purified preps and sequenced by Nextera XT, paired end reads, 100bp read length 10 million total reads. **(a+b)** pie charts of relative contributions of expression cassette (light blue) vector plasmid backbone (orange) and REPCAP plasmid (Dark blue expanded to yellow, REP gene mapping DNA in grey) in **(a)** AAV8scLP1hFIXco and **(b)** AAV9ssHLPhFVIIIv3 preparations. **(c-e)** Circos plots of REPCAP plasmid showing origination point of contamination from **(c)** AAV8scLP1hFIXco **(d)** AAV8ssHLPhFVIII and **(e)** AAV8ssHLPhFVIII with moved P5.

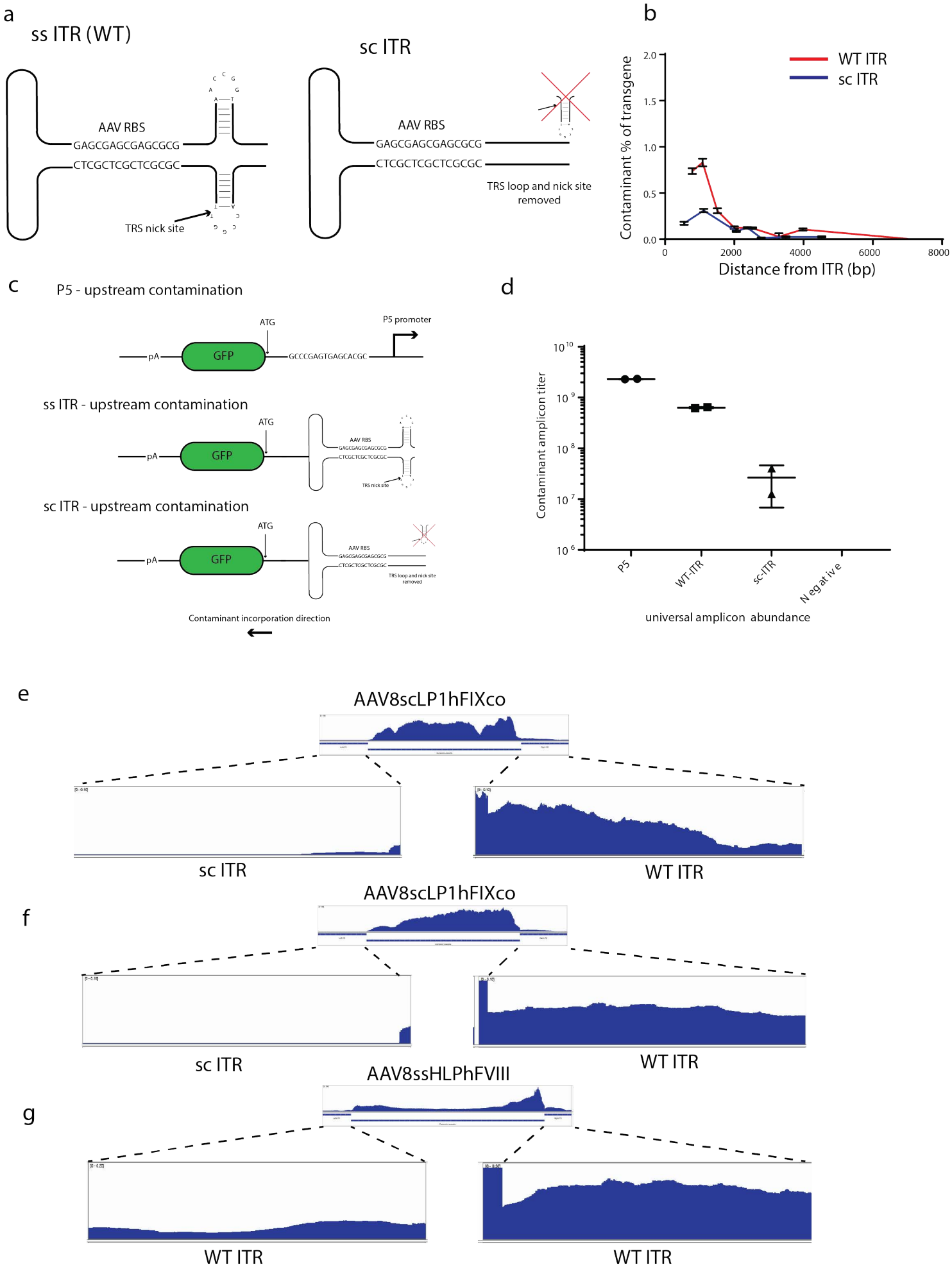

a
ss ITR (WT)
AAV RBS
GAGCGAGCGAGCGCG
CTCGCTCGCTCGCGC
TRS nick site
sc ITR
AAV RBS
GAGCGAGCGAGCGCG
CTCGCTCGCTCGCGC
TRS loop and nick site removed
b
WT ITR
sc ITR
Contaminant % of transgene
Distance from ITR (bp)
c
P5 - upstream contamination
ATG
P5 promoter
pA
GFP
GCCCGAGTGAGCACGC
ss ITR - upstream contamination
sc ITR - upstream contamination
Contaminant incorporation direction
d
Contaminant amplicon titer
P5
WT-ITR
sc-ITR
Negative
universal amplicon abundance
e
AAV8scLP1hFIXco
sc ITR
WT ITR
f
AAV8scLP1hFIXco
sc ITR
WT ITR
g
AAV8ssHLPhFVIII
WT ITR
WT ITR

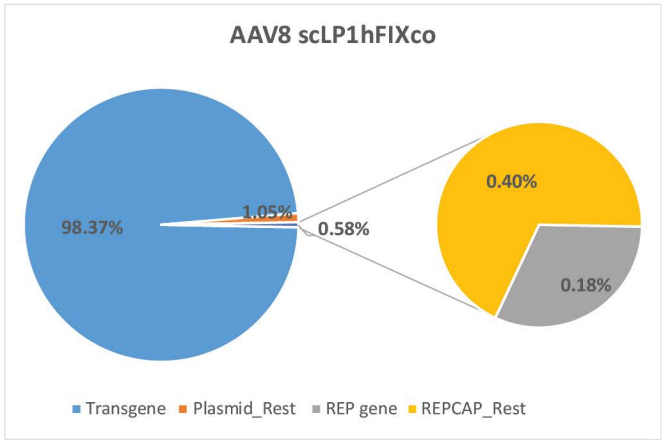
Figure 3.8 – WT ITRs initiate DNA backbone contamination at greater levels than self-complementary ITRs

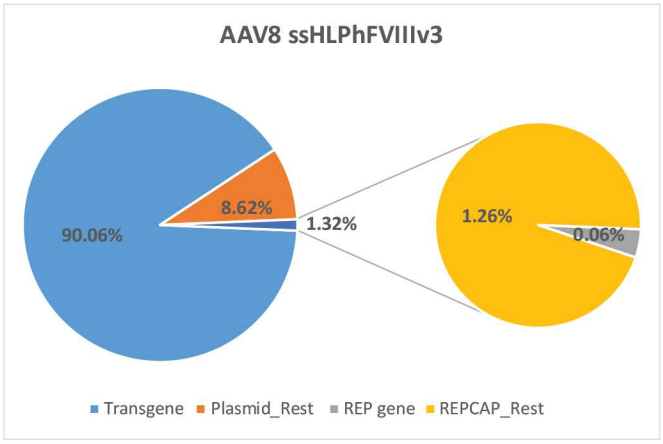
(a) Schematic; AAV ITRs showing a WT configuration containing REP binding site and TRS loop (left) and a self -complementary configuration in which the TRS loop is absent (right). **(b)** Graph showing results of qPCR of amplicons measuring abundance (*y axis* – expressed as a percentage of the expression cassette titer) plotted against distance from WT (red) or Self-Complementary (blue) ITRs (*x axis* – bp from outside end of ITR). **(c)** Schematic; Universal sequence placed outside/upstream of P5 promoter (top), WT ITR (middle), and Self-Complementary (bottom) **(d)** Bar graph showing results of qPCR of amplicons measuring abundance (*y axis* – expressed as copy number) of equidistant universal amplicon **(e-g)** Read mapping of next generation sequencing data *(viewed in IGV)* showing read abundance outside of ITRs from purified AAV from **(e+f)** AAV8scLP1hFIX-co and **(g)** AAV8ssHLPhFVIII.

3.2.2 AAV contaminant sequences are persistently transferred and transcriptionally active post infection:

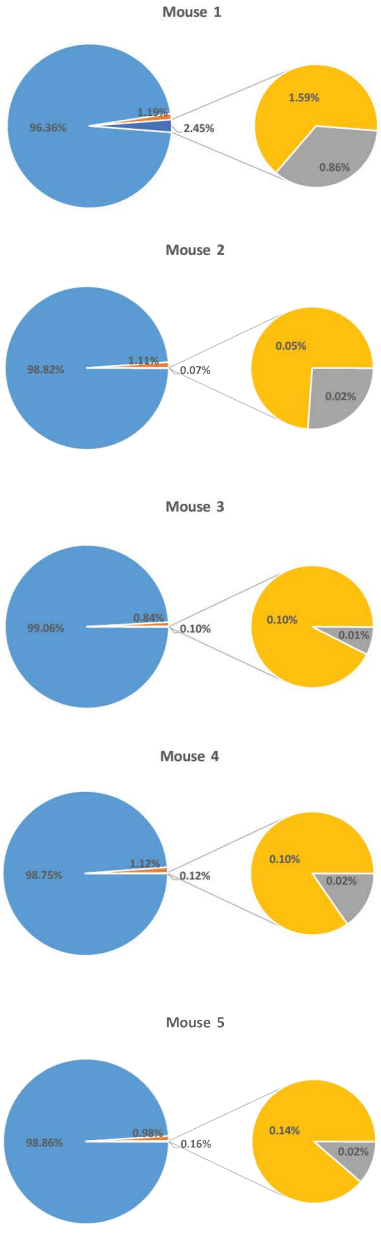
To examine whether these sequences were transferred and persistent post infection, CB57BL/6 mice were injected with 4.11e11vg of AAV8scLP1hFIXco or AAV8ssHLPhFVIIIv3. Four months post infection livers were harvested, DNA purified and deep sequenced. Reads that mapped to the CB57BL/6 genome were discarded and the remaining reads were mapped to the AAV production plasmids. The proportion of the reads that mapped to contaminant vector backbone sequences closely matched the contamination levels in the preps for both the FIX and FVIII constructs. In AAV8scLP1hFIXco infected mice 98.37% ± 1.01 of reads mapped to the FIX expression cassette. Producer plasmid contaminant reads made up 1.63% of the total with an average of 1.05% of reads mapping to the vector genome plasmid and 0.58% of reads mapping to the REPCAP plasmid (n=5). In the AAV8ssHLPhFVIII infected mice 90.06% ± 0.61 of sequence reads mapped to the FVIII expression cassette. Producer plasmid contaminant reads made up on average 9.94% of the total with 8.62% of the reads mapping to the vector genome plasmid backbone and 1.32% of the reads mapping to the REPCAP plasmid (n=5) (Fig. 3.9).

In AAV8ssHLPhFVIIIv3 infected mice in which liver DNA was assayed at 1-week post infection, 87.64% of reads mapped to the FVIII expression cassette and 12.36% of reads mapped to producer plasmid contaminant sequence (n=2) (Fig 3.10.a). Reads that mapped back to the REPCAP plasmids in these samples were combined and averaged across groups and showed an overrepresentation in the region upstream of the P5 promoter (Fig 3.10.b-d). Interestingly reads corresponding with the REP and CAP genes





n=5



n=5

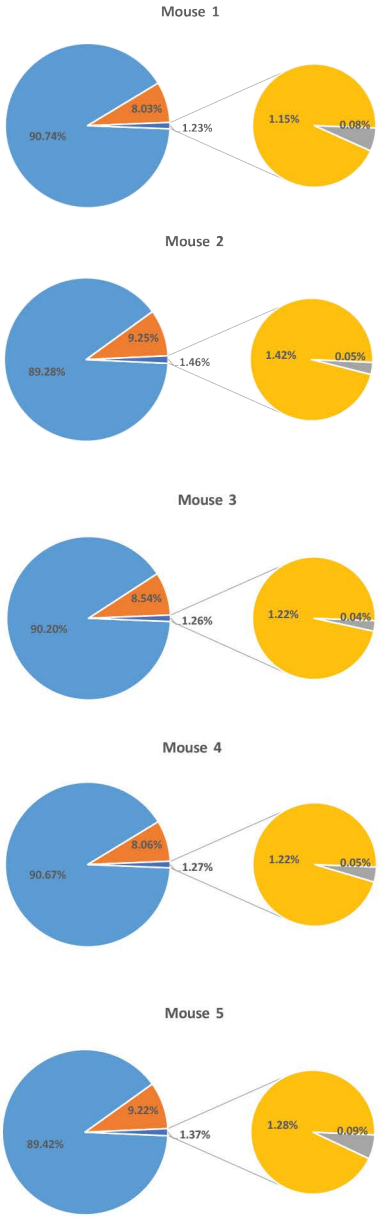


Figure 3.9 – AAV Contaminant DNA sequences are persistent in mouse livers post infection long term

DNA purified from livers 4 months post infection of CBL57/BL6 mice infected with 4.11e11vg of AAV8scLP1hFIXco (left) or AAV8ssHLPhFVIIIv3 (right) paired end reads, 100bp read length, 100 million total reads (n=5). Pie charts of relative contributions of expression cassette (light blue) vector plasmid backbone (orange) and REPCAP plasmid (Dark blue expanded to yellow, REP gene mapping DNA in grey). Shown as average read contribution (top) and individual (lower five) charts.

i.e. downstream of the P5 promoter were detected abundantly in one of the AAV8scLP1hFIXco infected mouse gDNA samples (Fig. 3.9 AAV8scLP1hFIXco mouse 1), which then appear in the plot of the average read counts (Fig 3.10.d).

To further verify the transfer of these contaminant sequences post AAV infection, Huh7 cells were infected with AAV8ssHLPhFVIIIv3. 48 hours post infection, cells were fixed on slides and probed by DNA FISH (Fig 3.11). DNA matching both the expression cassette and contaminants were detected in infected Huh7 cells (Fig 3.11.b). AAV8ssHLPhFVIIIv3 infected mouse hepatocytes were probed by DNA FISH and showed the presence of both expression cassette and contaminants at both 1 week and 4 months post infection, demonstrating both the transfer and persistence of contaminant sequences present in AAV preps (Fig 3.11.c).

A previous study had suggested that any contaminant sequences detectable post infection were transcriptionally inactive.[466] This study however, assayed preps that were 1. Of high purity 2. Using qPCR amplicons that were at distance from the regions of contaminant initiation that we have identified. We hypothesised that due to the sequence elements AAV contaminants were likely attached to, the ITRs and P5 sequence, that transferred DNA contaminants would exhibit transcriptional activity post infection. To assess the potential for contaminant sequences to be transcribed we first performed RT-PCR on two mice that had been infected with a high dose of AAV8ssHLPhFVIIIv3 (Fig 3.12). Whole transcript contaminant amplicons for the kanamycin and ampicillin resistance genes were successfully amplified by RT-PCR from these mice, suggesting the possibility that unintended protein products could be produced (Fig 3.12.a-b). To

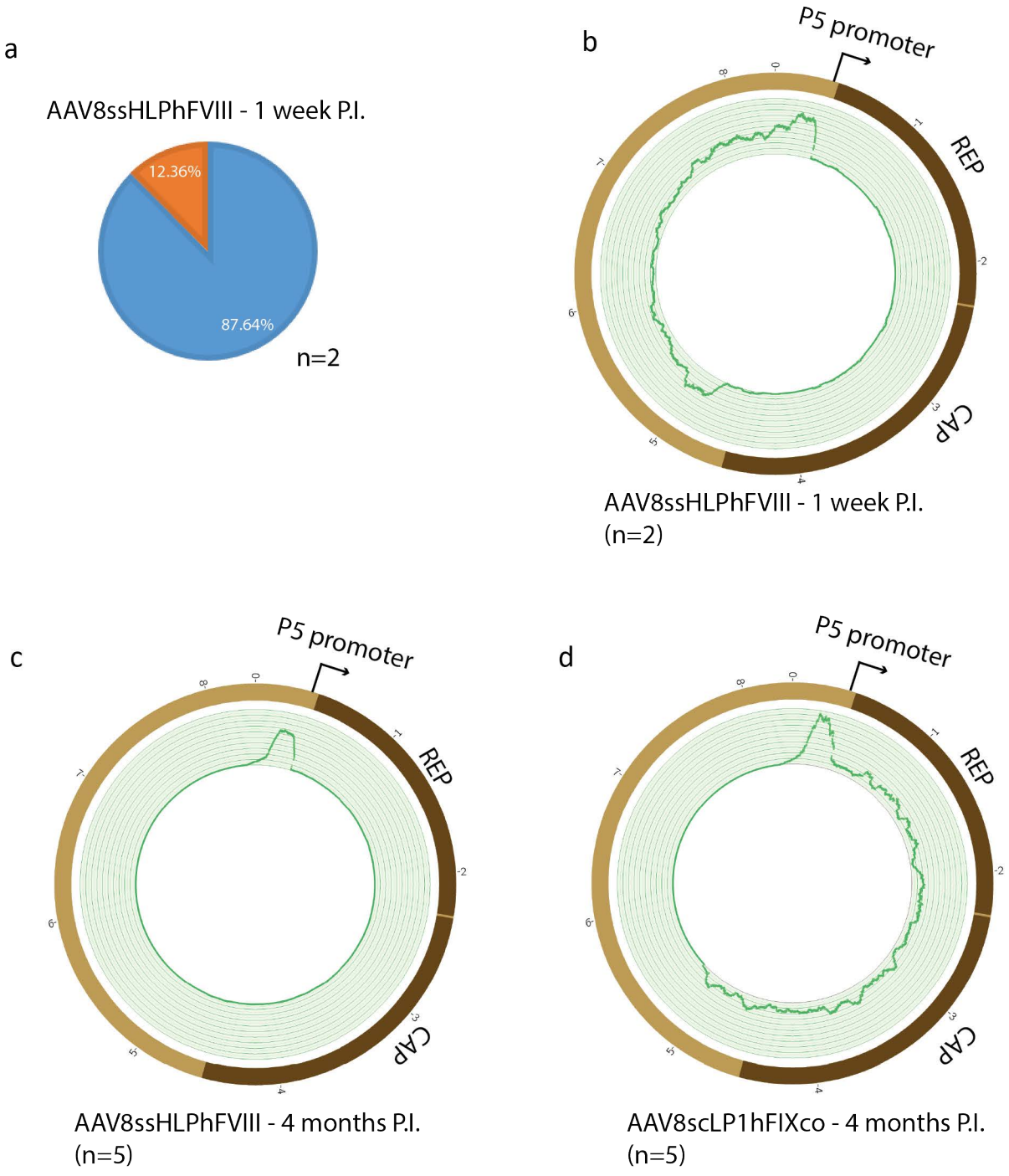


Figure 3.10 – AAV Contaminant DNA sequences upstream of P5 promoter persist long term post infection in mouse livers.

(a) Pie chart of next generation sequencing from CBL57/BL6 mice 1 week post infection with AAV8ssHLPh-FVIII showing relative contributions of expression cassette (blue) contaminants (orange)(n=2) **(b)** Circos plot of average reads mapping to AAV2_8 REPCAP plasmid from CBL57/BL6 mice 1 week post infection with AAV8ssHLPhFVIII (n=2) **(c)** Circos plot of average reads mapping to AAV2_8 REPCAP plasmid from CBL57/BL6 mice 4 months post infection with AAV8ssHLPhFVIII (n=5) **(d)** Circos plot of average reads mapping to AAV2_8 REPCAP plasmid from CBL57/BL6 mice 4 months post infection with AAV8scLP1hFIXco (n=5)

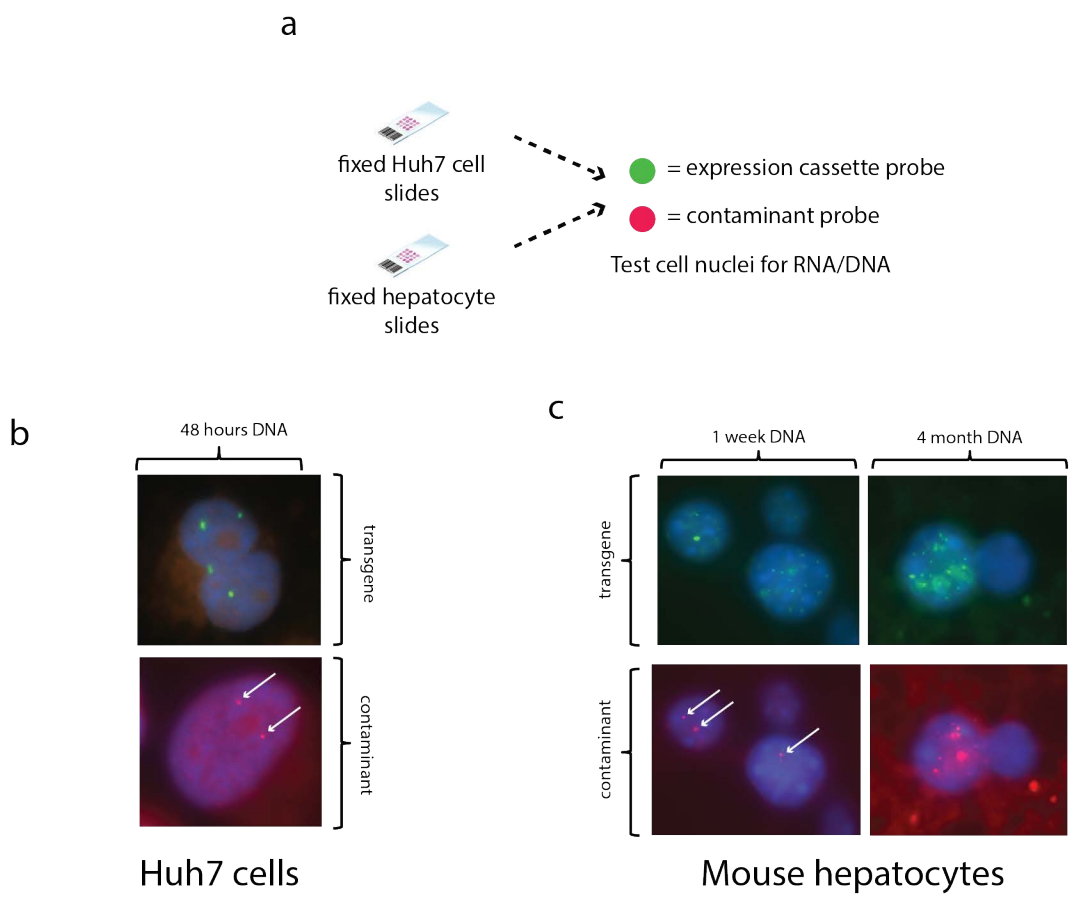


Figure 3.11 – AAV Contaminant DNA sequences detected in cells by FISH

(a) Schematic; Fixed cells on slides were probed for DNA sequences corresponding to the transgene (HLP-hFVIII) or contaminants (AAV2_8 REPCAP plasmid) **(b)** DNA FISH of fixed Huh7 cells probed for expression cassette (top - *green*) or contaminant (bottom - *red*) **(c)** DNA FISH of fixed CBL57/BL6 mouse hepatocytes 1 week (left) or 4 months (right) post infection, probed for expression cassette (top - *green*) or contaminant (bottom - *red*).

assay for the presence of unintended protein products within AAV infected cells, mouse livers from 1 week and 4 months post infection with AAV8ssHLPhFVIII were probed by western blot for the presence of Kanamycin and ampicillin resistance protein and was not detected in any of the tested mice (Fig 3.12.c-d).

Purified RNA from infected mouse hepatocytes of both AAV8scLP1hFIXco and AAV8ssHLPhFVIIIv3 treated mice was assayed by RNA sequencing. RNA sequence reads that mapped to mouse sequences were discarded. At 4 months post infection RNA reads from the FIX expression cassette made up 99.66% ±0.11 of detected RNA reads. 0.34% of the detected RNA reads mapped to the producer plasmids, of which 0.28% mapped the vector genome plasmid backbone and 0.06% mapped to the REPCAP plasmid (n=5). For AAV8ssHLPhFVIII infected mice, an average of 96.63% ±0.50 of RNA reads mapped to the FVIII expression cassette. On average 3.37% of the reads mapped to the producer plasmids of which 2.95% of the RNA reads mapped to the vector genome backbone and 0.48% of the reads mapped to the REPCAP plasmid (n=5) (Fig 3.13). RNAseq validation of contaminant transcription was also assessed at the 1-week timepoint for a separately produced prep of AAV8ssHLPhFVIIIv3 (n=2) showing similar a proportion of contaminant reads and transcript profile (Fig 3.14.a). Circos plot mapping of these RNA reads show transcripts localising to regions upstream of the P5 promoter region closely matching the DNA contaminant profile (Fig 3.14.b-d).

To further examine the potential for AAV contaminants to be transcribed, RNA FISH was performed on fixed, AAV8ssHLPhFVIIIv3 infected Huh7 cells and mouse hepatocytes

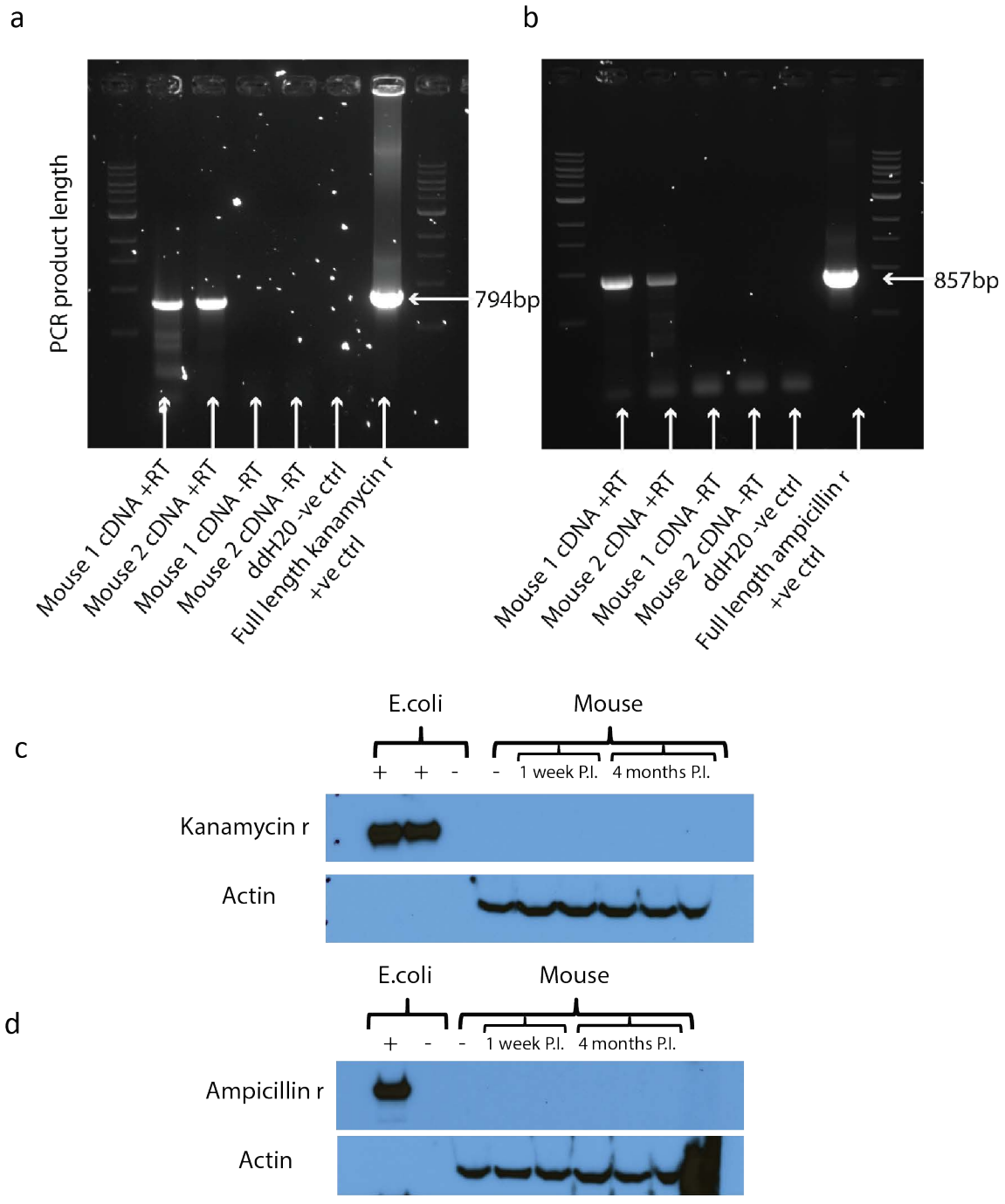
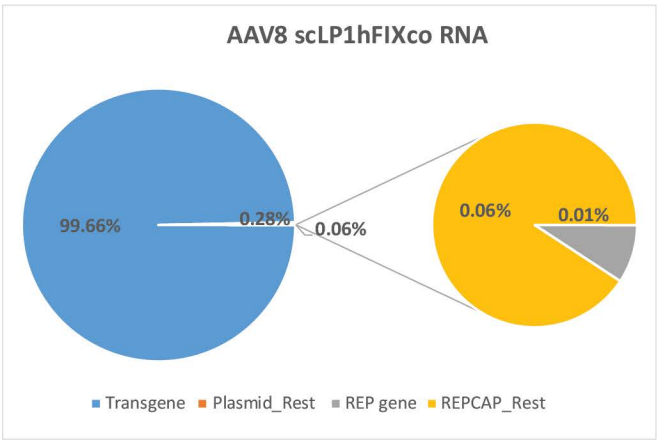
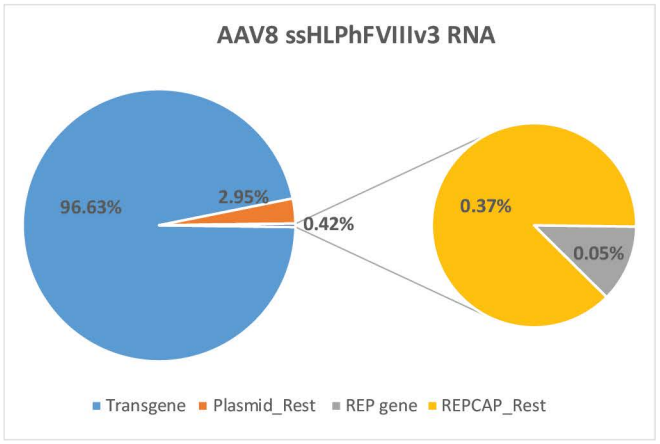


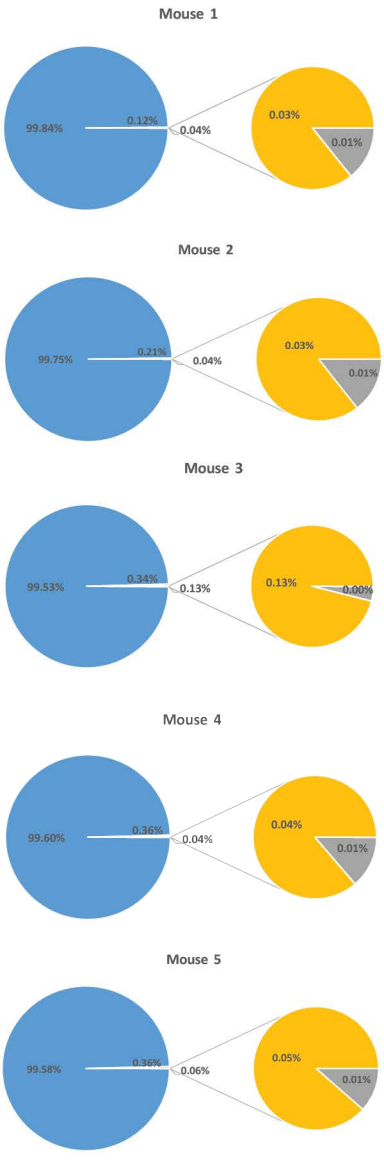
Figure 3.12 – Transcript of contaminant antibiotic resistance genes is detectable in mice post AAV infection

(a) Agarose gel of RT PCR reaction from CBL57/BL6 mouse liver RNA amplified with primers corresponding to the full length kanamycin gene sequence (794bp amplicon): *Lanes 1+2* - with reverse transcriptase step; *Lanes 3+4* - without reverse transcriptase step, *Lane 5* - ddH20 with reverse transcriptase step, *Lane 6* - plasmid DNA positive control. **(b)** Agarose gel of RT PCR reaction from CBL57/BL6 mouse liver RNA amplified with primers corresponding to the full length ampicillin gene sequence (857bp amplicon): *Lanes 1+2* - with reverse transcriptase step; *Lanes 3+4* - without reverse transcriptase step, *Lane 5* - ddH20 with reverse transcriptase step, *Lane 6* - plasmid DNA positive control. **(c)** Western blot for kanamycin resistance protein from: *Lanes1+2 E.Coli* expressing Kanamycin resitance protein; *Lane 3 E.Coli* - negative control; *Lane 4* - CBL57/BL6 mouse liver negative control; *Lanes 5+6* - CBL57/BL6 liver protein 1 week post infection; *Lanes 7-9* - CBL57/BL6 liver protein 4 months post infection. **(d)** Western blot for ampicillin resistance protein from: *Lanes1+2 E.Coli* expressing ampicillin resitance protein; *Lane 3 E.Coli* - negative control; *Lane 4* - CBL57/BL6 mouse liver negative control; *Lanes 5+6* - CBL57/BL6 liver protein 1 week post infection; *Lanes 7-9* - CBL57/BL6 liver protein 4 months post infection.





n=5



n=5

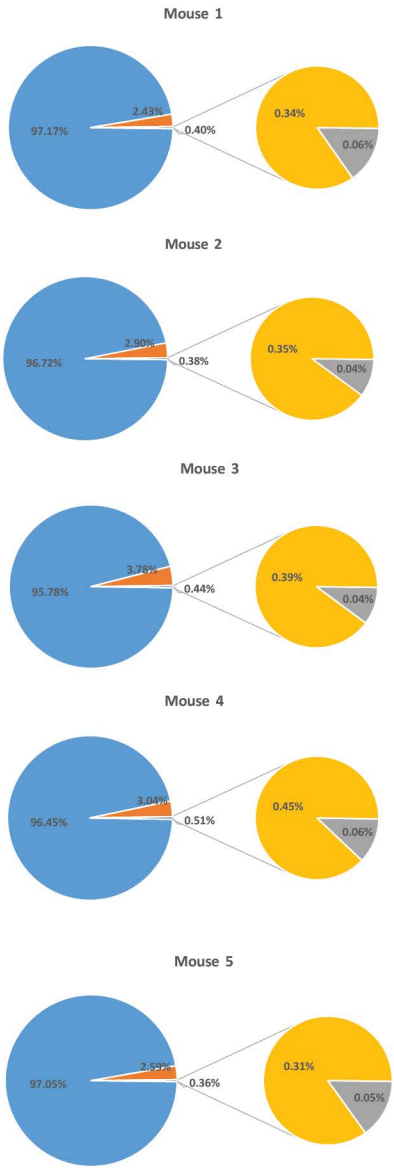
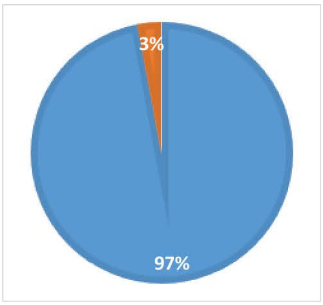


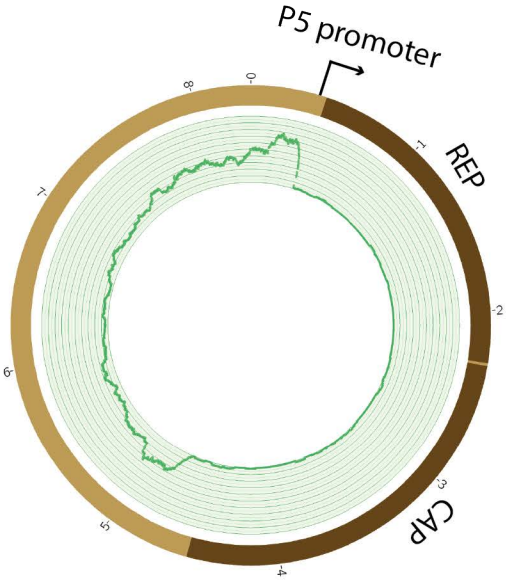
Figure 3.13– AAV Contaminant RNA sequences are persistently transcribed in mouse livers post infection long term

RNA purified from livers 4 months post infection of CBL57/BL6 mice infected with 4.11e11vg of AAV8scLP1hFIXco (left) or AAV8ssHLPhFVIIIv3 (right) paired end reads, 100bp read length, 100 million total reads (n=5). Pie charts of relative contributions of expression cassette (light blue) vector plasmid backbone (orange) and REPCAP plasmid (Dark blue expanded to yellow, REP gene mapping DNA in grey). Shown as average read contribution (top) and individual (lower five) charts.

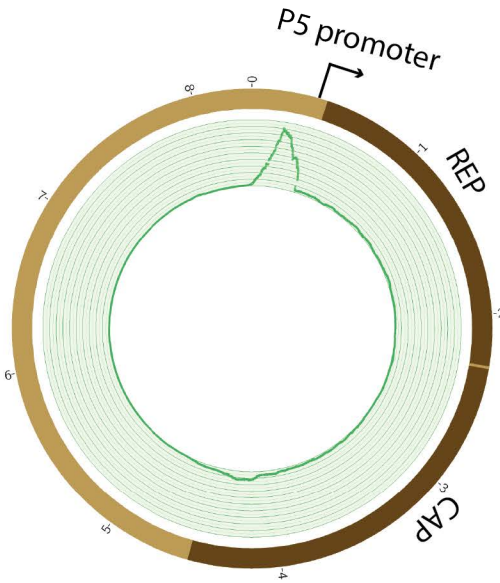
AAV8ssHLPhFVIII - 1 week P.I.



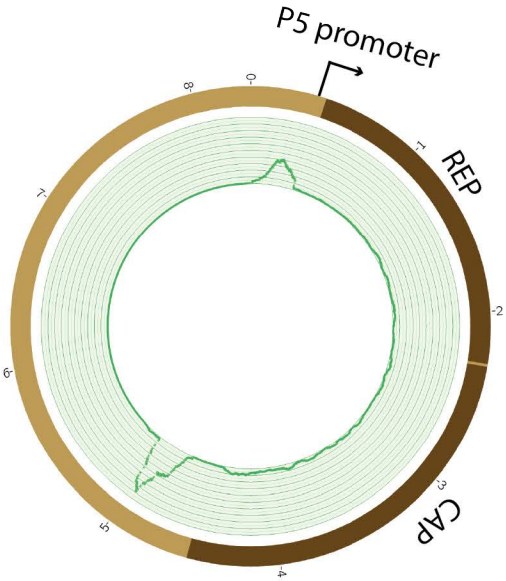
n=2



AAV8ssHLPhFVIII - 1 week P.I.
(n=2)



AAV8ssHLPhFVIII - 4 months P.I.
(n=5)



AAV8scLP1hFIXco - 4 months P.I.
(n=5)

Figure 3.14 – AAV contaminants upstream of the P5 promoter are transcriptionally active long term post infection in mouse livers.

(a) Pie chart of next generation RNA sequencing from CBL57/BL6 mice 1 week post infection with AAV8ssH-LPhFVIII showing relative contributions of expression cassette (blue) contaminants (orange)(n=2) **(b)** Circos plot of average reads mapping to AAV2_8 REPCAP plasmid from CBL57/BL6 mice 1 week post infection with AAV8ssHLPhFVIII (n=2) **(c)** Circos plot of average RNA reads mapping to AAV2_8 REPCAP plasmid from CBL57/BL6 mice 4 months post infection with AAV8ssHLPhFVIII (n=5) **(d)** Circos plot of average RNA reads mapping to AAV2_8 REPCAP plasmid from CBL57/BL6 mice 4 months post infection with AAV8scLP1hFIXco (n=5)

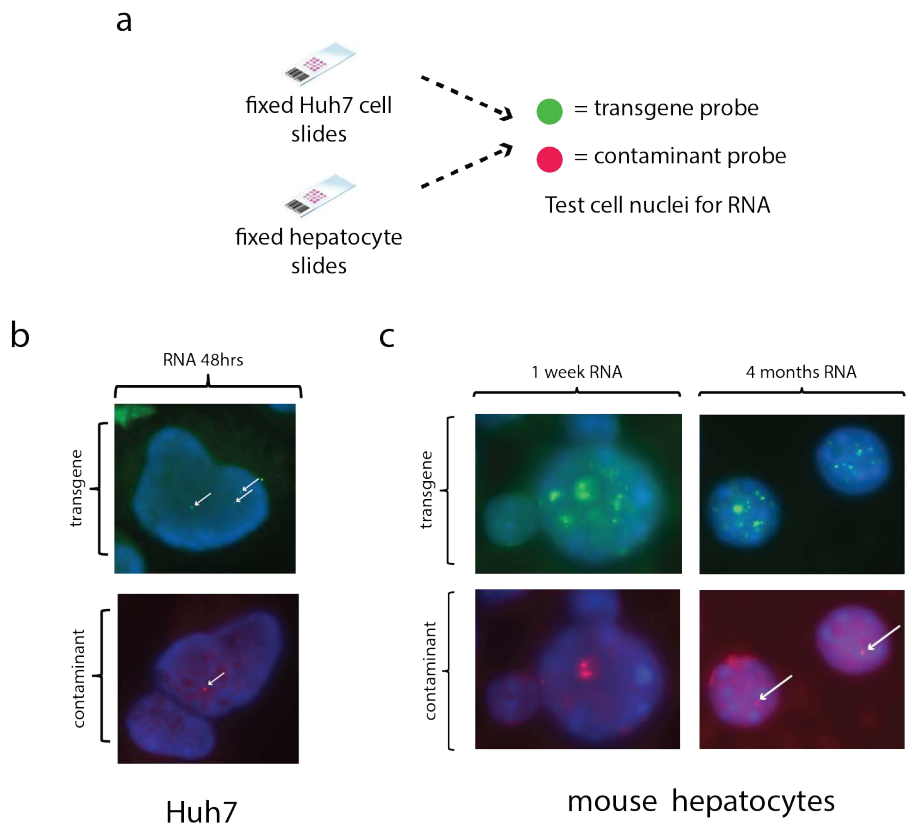


Figure 3.15 - Detection of contaminant RNA by FISH post AAV infection

(a) Schematic; Fixed cells on slides were probed for DNA sequences corresponding to the transgene (HLP-hFVIII) or contaminants (AAV2_8 REPCAP plasmid) **(b)** RNA FISH of fixed Huh7 cells probed for expression cassette (top - *green*) or contaminant (bottom - *red*) **(c)** RNA FISH of fixed CBL57/BL6 mouse hepatocytes 1 week (left) or 4 months (right) post infection, probed for expression cassette (top - *green*) or contaminant (bottom - *red*).

(Fig. 3.15). In the Huh7 cells at 48 hours post infection and in the mouse hepatocytes at both 1 week and 4 months post infection, RNA transcripts that matched the contaminant sequence was detected by fluorescent probes specific to sequences in the REPCAP plasmid. It is apparent from these studies that transcription of contaminant producer plasmid sequences is observed post infection in hepatic cells both in vitro and in vivo.

3.2.3 Contaminant sequences upstream of P5 can be translated post infection in vitro and in vivo:

Given the detectable transcription of AAV contaminant DNA sequences transferred to the cell post infection, the possibility for contaminant protein products from sequences present in the producer plasmids exists. Promoter activity from the ITRs has been observed.[257,478] Indeed, this promoter activity was utilised in early AAV constructs for the treatment of cystic fibrosis via the delivery of the CFTR gene.[257,258] The transcript results so far suggest the possibility for promoter activity driven by the reverse ITR sequence and the reverse direction of the P5 promoter, for the contaminants incorporated outside the ITRs and upstream of the P5 promoter respectively. We examined the inherent promoter activity of these sequences in 293T cells by transfecting plasmids containing either ITR sequence or the P5 promoter with a GFP cassette upstream, using the P5 promoter with a downstream cassette as a positive control (Fig. 3.16). Forty-eight hours post plasmid transfection, cells were assayed by FACS. Brightly GFP positive cells were detected when the gene encoding GFP was upstream of the P5 promoter but not in a control where GFP was placed in a reverse orientation (Fig. 3.16b). GFP positive cells were also detected when the gene was placed upstream of ITR sequences, at a lower

abundance than either forward or reverse direction of P5, in the absence of the REP protein. Self-complementary ITR appeared to have a significantly lower upstream transcriptional activity than the WT ITR (GFP MFI: WT-ITR - 17535±1063; scITR - P <0.0001), suggesting that the TRS sequence plays some role in transcription from this site (Fig. 2.16b). Surprisingly the MFI of GFP positive cells was around 75% of the level detected in the forward direction showing considerable activity (P5_GFP - 80519 ± 2124; GFP_P5 60165 ± 266) (Fig. 3.16b), suggesting that in the absence of REP both directions of the P5 promoter exhibit strong transcriptional activity. Furthermore, in a configuration where the REP proteins were present, the MFI of GFP positive cells dropped by over 75%, suggesting that transcriptional activity of P5 in the reverse direction is repressed by the presence of REP, a factor already known to repress P5 activity in the forward direction.[174]

Given the necessity to retain the inverted terminal repeats in AAV vector genomes in their location around the expression cassette, we focused on assaying the contaminants from upstream of the P5 promoter sequence. To do this we developed a reporter assay in which a GFP cassette was placed upstream of the P5 promoter in the AAV2_8 REPCAP plasmid in either the forward (GFP_P5) or reverse ($_{rev}$GFP_P5) orientation relative to the putative anti-P5 promoter, with the standard plasmid backbone as a negative control ($_{empty}$P5) (Fig. 3.17.a). As expected after transfection into 293T cells only the GFP_P5 plasmid produced GFP positive cells as observable by fluorescent microscope (Fig. 3.17.b).

AAV8ssHLPhFVIII was produced with the GFP_P5, $_{rev}$GFP_P5 and $_{empty}$P5 REPCAP plasmids. Post purification AAV was evaluated by deep sequencing and showed

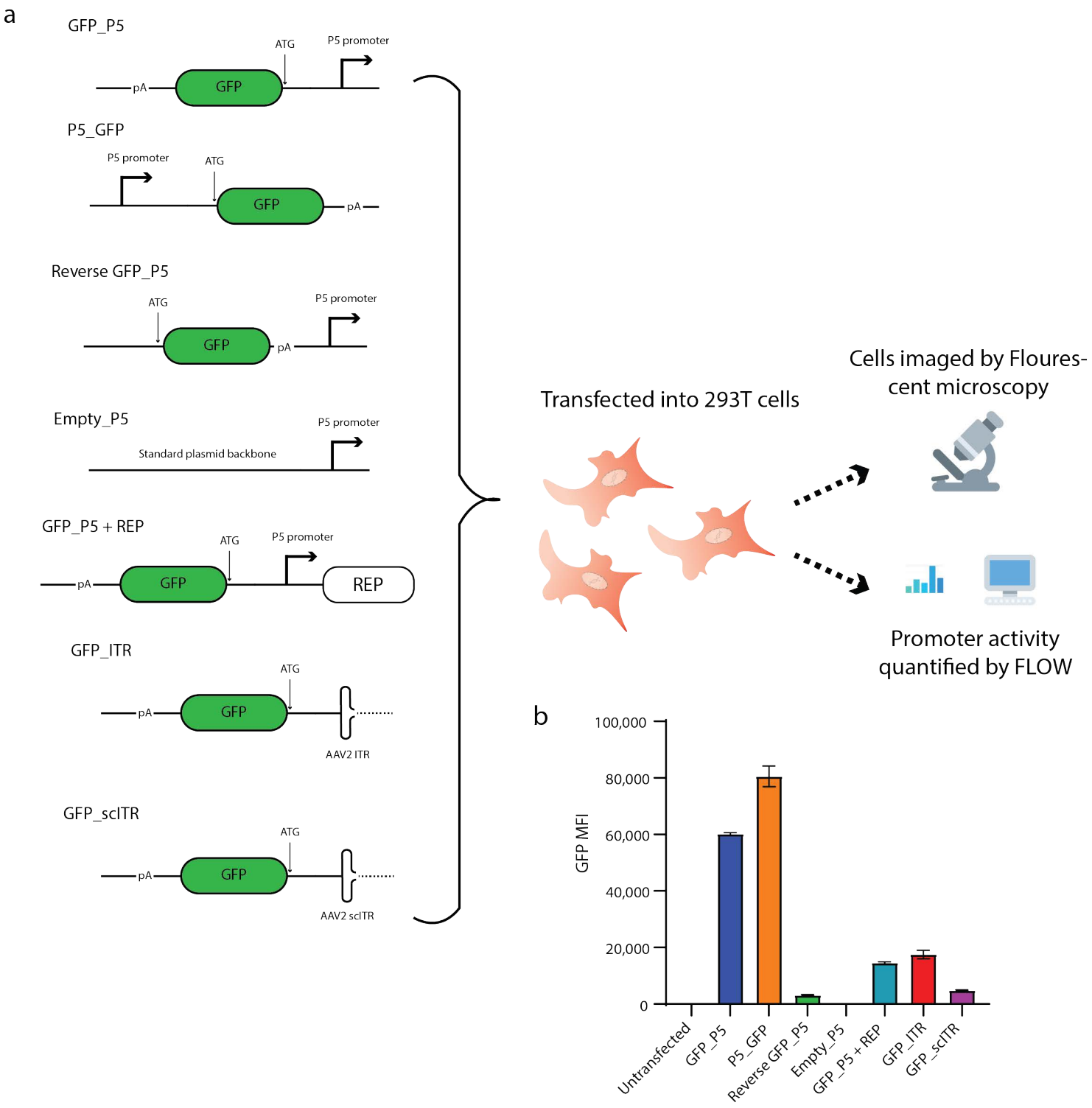


Figure 3.16 – Transcriptional activity of REP containing sequence elements from the AAV genome

(a) Schematic; Plasmids containing a GFP cassette upstream of REP containing elements of the AAV genome were transfected into 293T cells and analysed by FLOW. Constructs: (from top to bottom) P5 promoter with upstream GFP cassette; P5 promoter with upstream GFP cassette in reverse; P5 promoter without GFP cassette upstream; P5 promoter with GFP cassette downstream; P5 promoter with GFP cassette upstream in REPCAP plasmid context; WT ITR with GFP cassette upstream; Self-Complementary ITR with GFP cassette upstream. **(b)** Bar graph showing mean GFP flourescent intensity of 293 cells at 48 hours post transfection with REP element containing constructs. Analysed by One way ANOVA with Tukey's multiple comparison test.

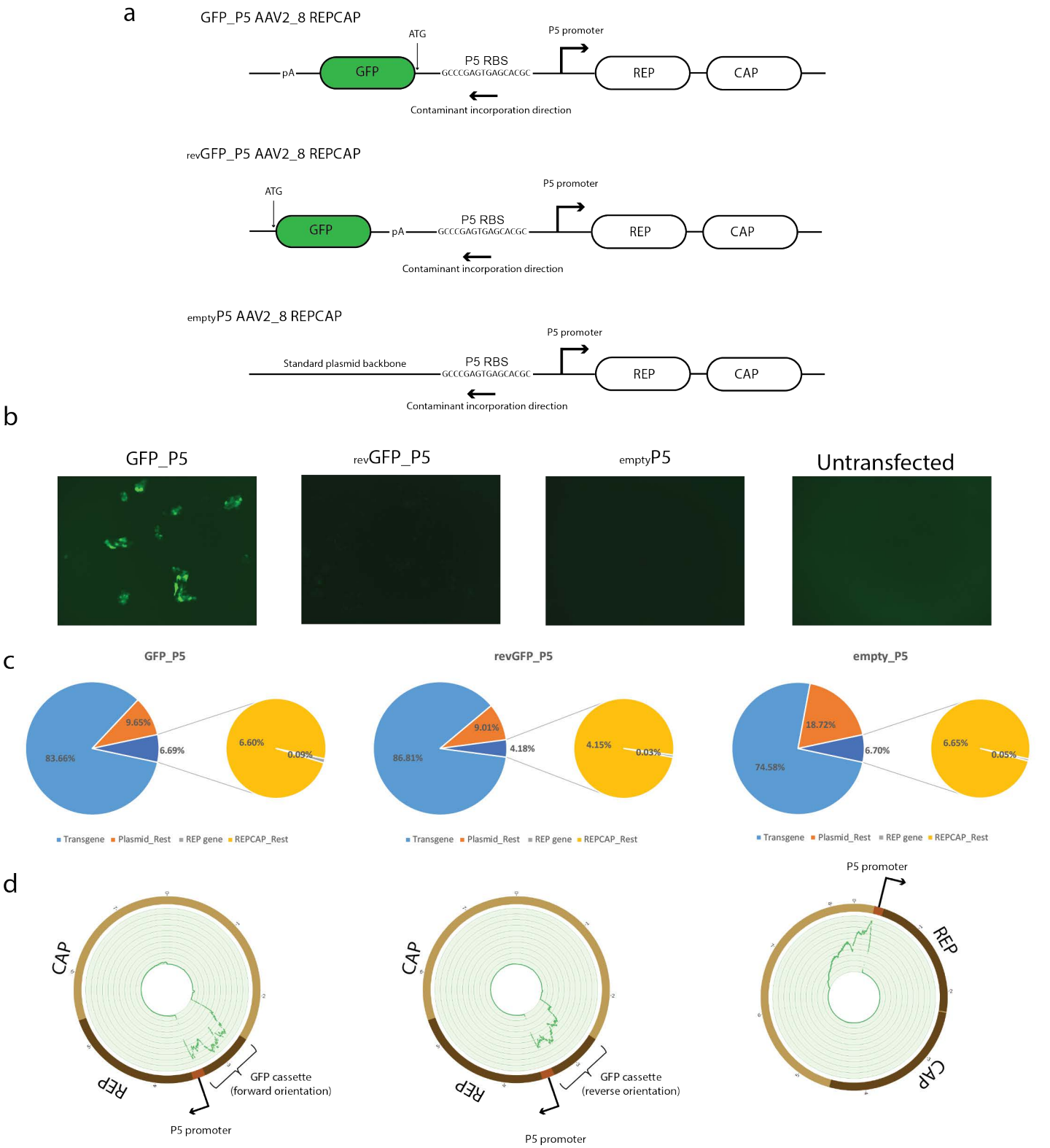


Figure 3.17 – GFP cassette flanked AAV preps produce AAV with P5 upstream contamination

(a) Schematic; AAV2_8 REPCAP constructs with (from top to bottom); GFP cassette upstream of P5 (GFP_P5 AAV2_8 REPCAP), GFP cassette in reverse upstream of P5 (revGFP_P5 AAV2_8 REPCAP), non-fluorescent reporter backbone (emptyP5 AAV2_8 REPCAP). **(b)** Flourescent microscopy of 293 cells 48 hours post plasmid transfection with (from left to right); GFP_P5 AAV2_8 REPCAP, revGFP_P5 AAV2_8 REPCAP, emptyP5 AAV2_8 REPCAP, untransfected. **(c)** Pie chart of DNA sequencing reads for AAV8ssHLPhFVIII preparations produced with (from left to right); GFP_P5 AAV2_8 REPCAP, revGFP_P5 AAV2_8 REPCAP, emptyP5 AAV2_8 REPCAP. Reads mapped to; expression cassette (light blue) vector plasmid backbone (orange) and REPCAP plasmid (Dark blue expanded to yellow, REP gene mapping DNA in grey). **(d)** circos plots of REPCAP plasmid from AAV8 produced with (from left to right); GFP_P5 AAV2_8 REPCAP, revGFP_P5 AAV2_8 REPCAP, emptyP5 AAV2_8 REPCAP.

abundant presence of producer plasmid contaminants including the GFP cassette upstream of P5 (Fig 3.17.c+d). Adherent 293T cells were infected at a range of MOIs of AAV8ssHLPhFVIII genomes produced with either GFP_P5, $_{rev}$GFP_P5 and $_{empty}$P5 AAV2_8 REPCAP plasmids (Fig. 3.18). 72 hours post infection, samples were gathered and analysed by FACS. AAV8 $_{rev}$GFP_P5 treated cells showed no shift in detected GFP intensity compared to $_{empty}$P5 cells. However, 293T cells treated with AAV8 GFP_P5 exhibited a shift, with 3.12% ±0.34 of cells exhibiting detectable GFP positivity at the 5e6MOI. (Fig 3.18.b+c). As expected, contaminant transduction rates were dose dependent. Certain initial studies of AAV infection have shown that in impure AAV preparations, protein can be transferred with the AAV capsid, leading to artefactual detection of a successful infection event referred to as pseudotransduction.[479] To ensure that the observed GFP positivity was not a result of transferred protein contaminants, 293T cells were assayed by FACS at subsequent timepoints; 48 and 72 hours post infection. GFP positivity increased significantly from 48 to 72 hours in both the number of detectable positive cells and the MFI of positive cells. (Fig 3.18.d+e). This is the converse of what would be expected from a pseudotransduction event. Together this shows that contaminant DNA incorporated upstream of the P5 promoter in AAV can be translated post infection in 293T cells.

To examine if protein expression from AAV contaminants could occur in vivo in the context of liver directed gene therapy, CB57BL/6 were administered 4.11e11vg (approximately 2e13vg/kg) of AAV8ssHLPhFVIII produced with either GFP_P5, $_{rev}$GFP_P5 or $_{empty}$P5 REPCAP plasmid. A low dose (2e10vg (1e12vg/kg)) of AAV8ssCMV_GFP was used as a positive control. Mouse livers were harvested at 1-week post infection and IHC was

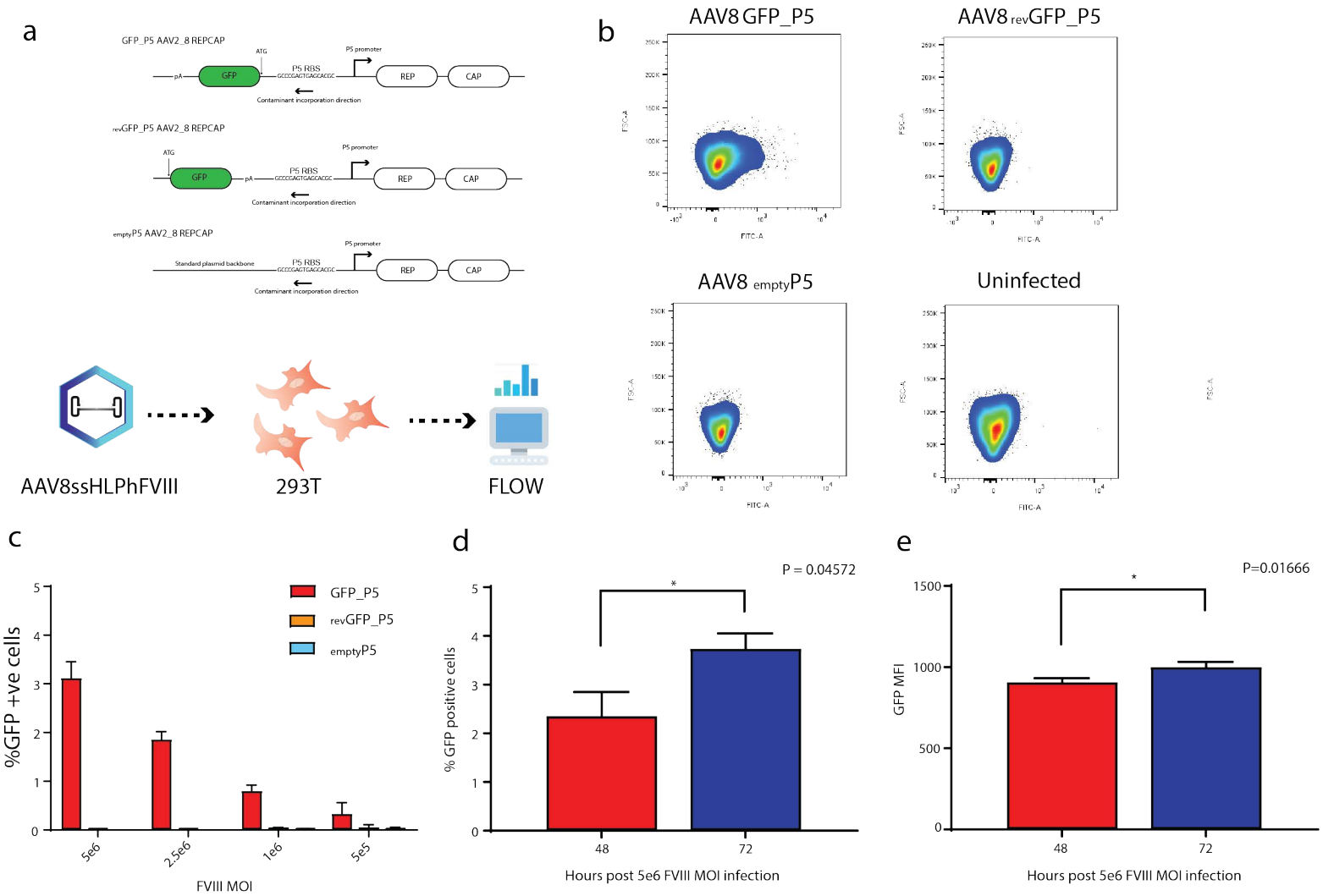


Figure 3.18 – Translation of a contaminant cassette in AAV infected cells

(a) Schematic; AAV2_8 REPCAP constructs used to produce AAV8ssHLPhFVIII infected into 293T cells and analysed post infection for flouresence by FLOW **(b)** FLOW panels showing GFP expression of 293T cells 72 hours post infection with AAV8ssHLPhFVIII produced with: (top left) GFP_P5; (top right) revGFP_P5; (bottom left) emptyP5; (bottom right) Uninfected 293T cells. **(c)** Bar graph of the % of GFP positive 293T cells detected 72 hours post infection at different MOIs of AAV8ssHLPhFVIII produced with: *(red)* GFP_P5; *(orange)* revGFP_P5; *(light blue)* emptyP5. **(d)** Bar graph of the % of GFP positive 293T cells detected at *(red)* 48 and *(blue)* 72 hours post infection with 5e6MOI of AAVAAV8ssHLPhFVIII produced with GFP_P5 AAV2_8 REPCAP **(d)** Bar graph of Mean Flourescent intensity of GFP positive 293T cells at *(red)* 48 and *(blue)* 72 hours post infection with 5e6MOI of AAVAAV8ssHLPhFVIII produced with GFP_P5 AAV2_8 REPCAP

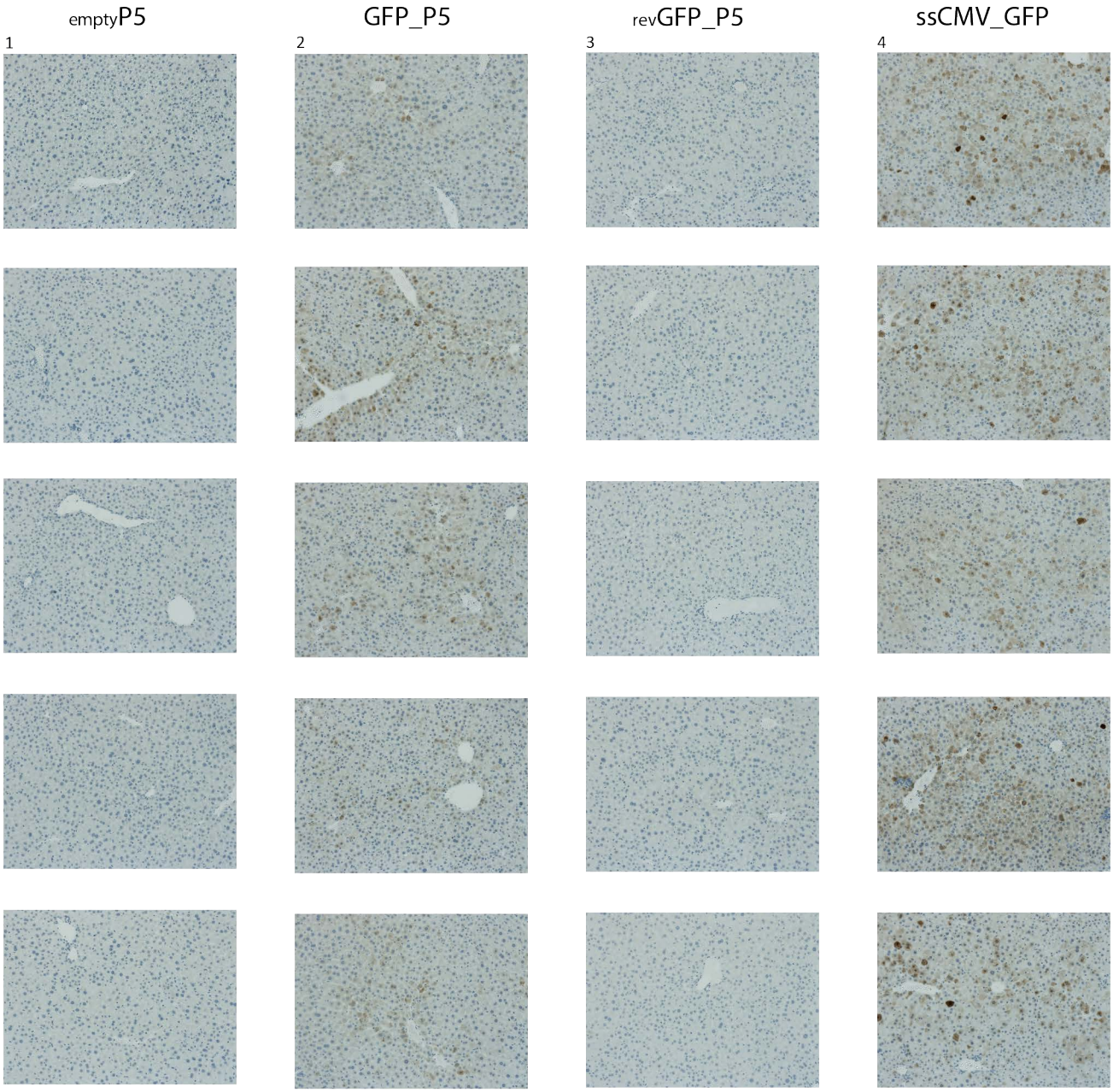


Figure 3.19 – Detectable P5 contaminant derived GFP protein in mouse hepatocytes post infection

IHC staining for GFP conducted on mouse hepatocytes from CBL57/BL6 mice 1 week post infection with AAV8. (Column 1) AAV8ssHLPhFVIII produced with emptyP5 AAV2_8 REPCAP - *4.11e11vg/kg* (Column 2) AAV8ssHLPhFVIII produced with GFP_P5 AAV2_8 REPCAP - *4.11e11vg/kg*; (Column 3) AAV8ssHLPhFVIII produced with revGFP_P5 AAV2_8 REPCAP - *4.11e11vg/kg* (Column 4) AAV8ssCMV_GFP produced with emptyP5 AAV2_8 REPCAP - *2e10vg/kg*

performed. GFP positive hepatocytes were seen in every mouse of both the ssCMV_GFP positive control and GFP_P5 injected mice but were not detected in the hepatocytes treated with the $_{rev}$GFP_P5 or $_{empty}$P5 vector (n=5) (Fig. 3.19). This is further evidence that not only does the anti-P5 sequence exhibit promoter activity, but that it can drive transcriptional and translational activity of incorporated contaminants in AAV post infection.

To understand whether these contaminants derived from the P5 sequence could be incorporated into AAV and act independently of the presence of an ITR flanked expression cassette, AAV8 was produced in the absence of an expression cassette using the GFP_P5 REPCAP plasmid hereby referred to as GFP_P5$_{noTG}$ (Fig. 3.20). After column purification, abundant GFP DNA was detected in the AAV8 GFP_P5$_{noTG}$ prep by qPCR (Fig. 3.20.b). When the DNA from these AAV was run on an alkaline gel, a smeared pattern was seen in both the GFP_P5$_{noTG}$ and GFP_P5 produced samples. This result was expected with GFP_P5 due to the presence of the oversized FVIII cassette, known to produce this pattern.[480] In the GFP_P5$_{noTG}$ sample however, there is no transgene and thus no outside of the ITR contaminants. As the contribution of host cell DNA is orders of magnitude lower than producer plasmid contaminants, it was deduced that this detected contaminant DNA was P5 derived and that these contaminants exhibit a range of sizes up to the packaging capacity of circa 5kb (Fig. 3.20.c). This demonstrates that the contaminants incorporated upstream of P5 are fragmented and are packaged into AAV in an ITR independent manner.

To examine the infection kinetics of P5 linked AAV contaminants, and whether these contaminants had the ability to remain stably present, 293T cells were infected at an MOI

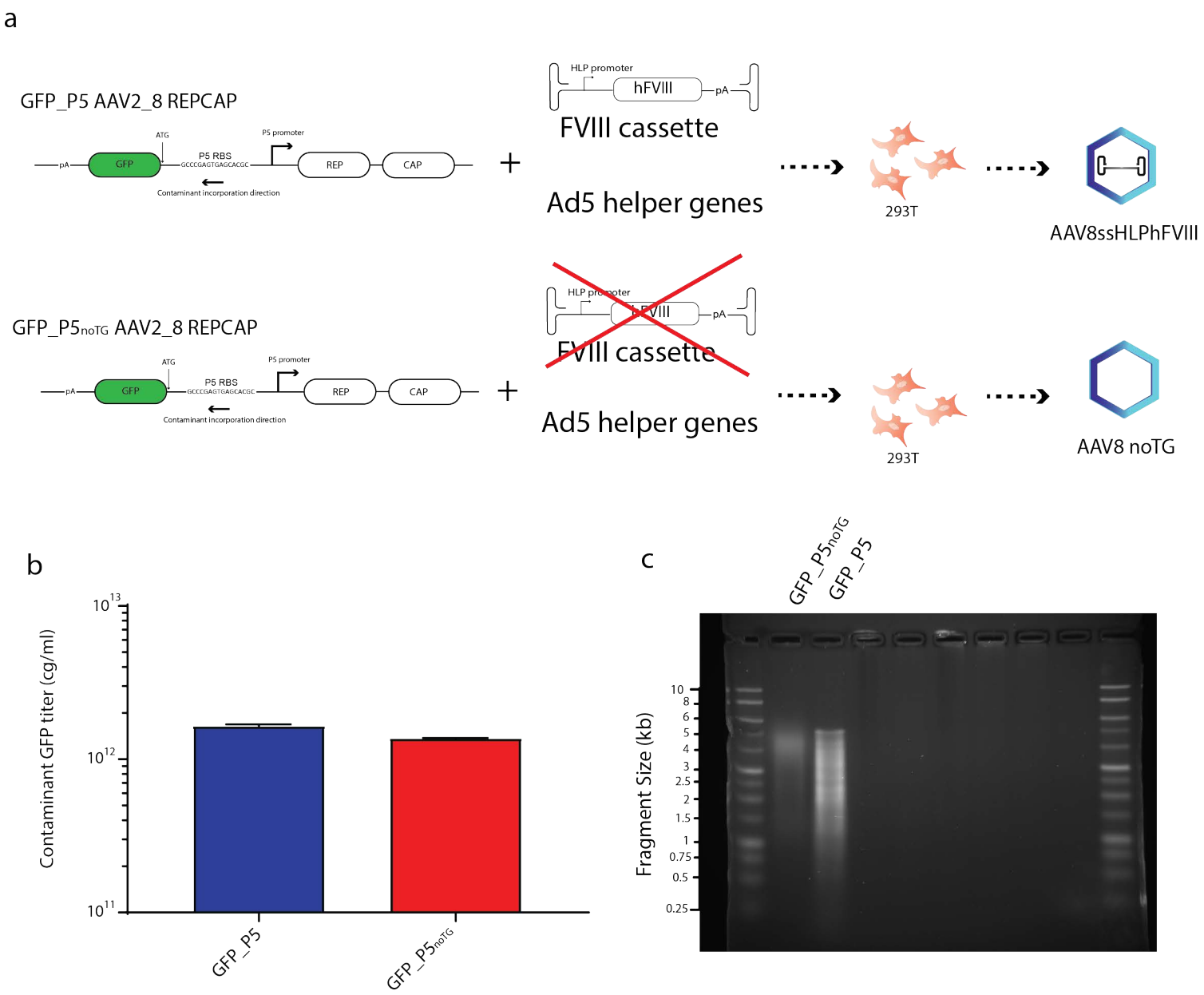


Figure 3.20 – P5 derived contaminants can be packaged independently of the expression cassette

(a) Schematic; GFP_P5 AAV2_8 REPCAP was used to produce AAV in the presence (GFP_P5) or absence (GFP_P5$_{noTG}$) of an ITR flanked FVIII cassette. **(b)** qPCR detection of GFP DNA in purified AAV8 produced with *(blue)* GFP_P5 or *(red)* GFP_P5$_{noTG}$ **(c)** Alkaline gel of DNA from AAV8 produced with (*Lane 1)* GFP_P5$_{noTG}$ and *(Lane 2)* GFP_P5

of 5e6 with AAV8ssHLPhFVIII GFP_P5, $_{rev}$GFP_P5, $_{empty}$P5, and an equivalent capsid load of AAV8 GFP_P5$_{noTG}$. A lower MOI (1e4) of AAV8ssCMV_GFP was used as a positive control. GFP positive cells were assayed by FACS at 1,2,3,7,10,14, and 21 days post infection (Fig. 3.21). Both GFP_P5 and GFP_P5$_{noTG}$ showed detectable GFP positive cells at multiple timepoints post infection, increasing from 24 to 72 hours and then decreasing to a low stable level of around 0.1% (GFP_P5 0.097% ± 0044; GFP_P5$_{noTG}$ 0.117% ± 0.051) as cells divided and were continually passaged. As in previous experiments, neither $_{rev}$GFP_P5 or $_{empty}$P5 infected cells showed expression of GFP protein. When compared with the ITR flanked ssCMV_GFP cassette the inefficiency of the contaminant cassette is apparent; a FVIII dose 500-fold higher than ssCMV_GFP was required to reach comparable levels of GFP positivity from the contaminant sequences (Fig 3.21.c). The detected expression from the GFP_P5$_{noTG}$ infected cells shows that at least for DNA contaminants in AAV originating from the P5 promoter, transcriptional activity of the contaminant DNA species is independent of productive infection from recombinant AAV genomes. Interestingly, the percentage of GFP expressing cells at the peak (72 hours) that were still expressing at the stable timepoints (10-21 days) timepoints was significantly greater (P=0.0058) in the GFP_P5$_{noTG}$ treated cells compared to the ssCMV_GFP treated cells (GFP_P5 4.75% ±0.66; GFP_P5$_{noTG}$ 9.89% ±2.18; ssCMV_GFP 0.63% ± 0.04). GFP_P5 did not reach significance above ssCMV_GFP in this regard (P=0.13), and the GFP_P5 and GFP_P5$_{noTG}$ were also not found to be significantly different to each other (P=0.07) (Fig. 2.21.d). This would suggest that P5 derived DNA contaminants in AAV could have a greater level of persistence than ITR flanked expression cassettes. However, this initial experiment did not use equivalent

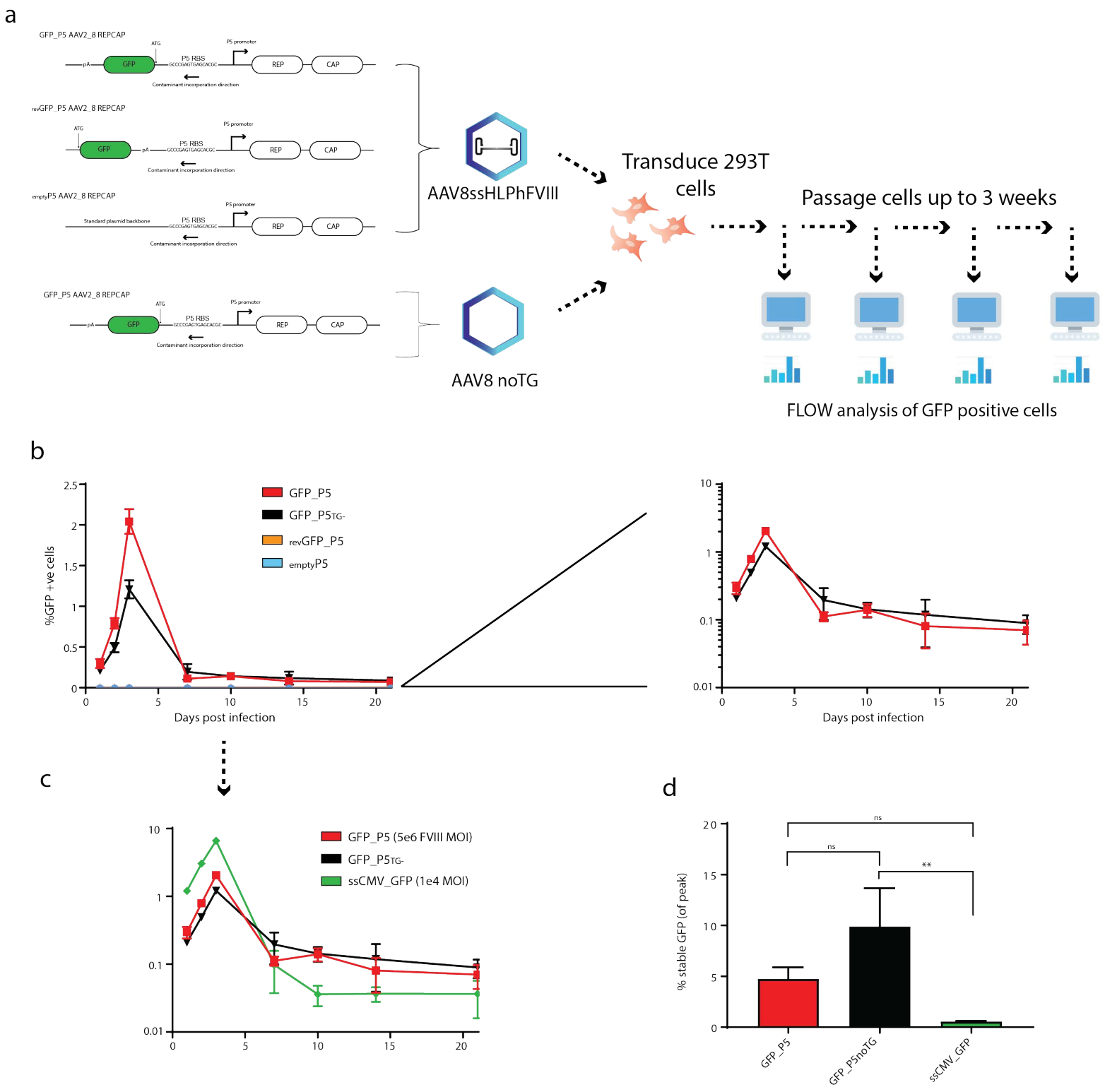


Figure 3.21 – A proportion of translated contaminants from P5 persist in dividing cells

(a) Schematic; AAV8 produced with GFP_P5, revGFP_P5 emptyP5 GFP_P5noTG were transduced at 5e6 MOI and AAV8ssCMV_GFP transduced at 1e4 MOI into 293T cells and continually passaged for 3 weeks, with FLOW analysis at 1,2,3,7,10,14, and 21 days post infection. **(b)** Graph of % GFP positive cells detected by FLOW in 293T cells transduced with (red) GFP_P5, (orange) revGFP_P5, (light blue) emptyP5, or (black) GFP_P5noTG. **(c)** Graph of % GFP positive cells detected by FLOW in 293T cells transduced with (red) GFP_P5 *(black)* GFP_P5noTG compared with a lower dose infection of AAV8ssCMV_GFP. **(d)** Bar graph of the percentage of GFP expressing cells at 21 days P.I. expressed as a percentage of GFP expressing cells at 72 hours P.I. (1 way Anova with Tukey's multiple comparisons test)

MOIs of expression cassette to the other samples. To investigate this further, 293T cells were infected with either AAV8 GFP_$P5_{noTG}$ or AAV8ssCMV_GFP a range of MOIs; 2e6, 1e6, 5e5, 1e5. FACS was conducted at 2,3,5,7,10,14, and 21 days post infection. (Fig 3.22.a). At equivalent MOIs, the number of GFP expressing cells was consistently higher in the ssCMV_GFP infected cells across timepoints, (Fig 3.22b) however the proportion of GFP positive cells at 72 hours that stably produced GFP was far greater in the GFP_$P5_{noTG}$ group (10.17% ± 1.983) than the ssCMV_GFP group (0.87% ± 0.012) (Fig. 3.22c). However, when gDNA of 293T cells was analysed for GFP copy number, both $P5_{GFP/TG-}$ and ssCMV_GFP infected cells had similar levels of stably present GFP copies relative to the detected copies 72 hours post infection (GFP_$P5_{noTG}$ 1.20% ± 0.28; ssCMV_GFP 0.88% ± 0.22) (P = 0.379). This suggests that these non ITR containing contaminants from P5 are stably present post infection, implying integration into the 293T cell genome at an equivalent rate to the ITR flanked expression cassette. Further qualitative evidence of this was found through sorting GFP_P5 contaminant infected 293T cells for GFP positive cells, and plating at low density in 96 well plates. Cells were cultured and imaged after 5 and 7 days of growth, yielding identifiable colonies that retained GFP positivity as well as colonies that were GFP negative after extended culture (Fig. 3.23).

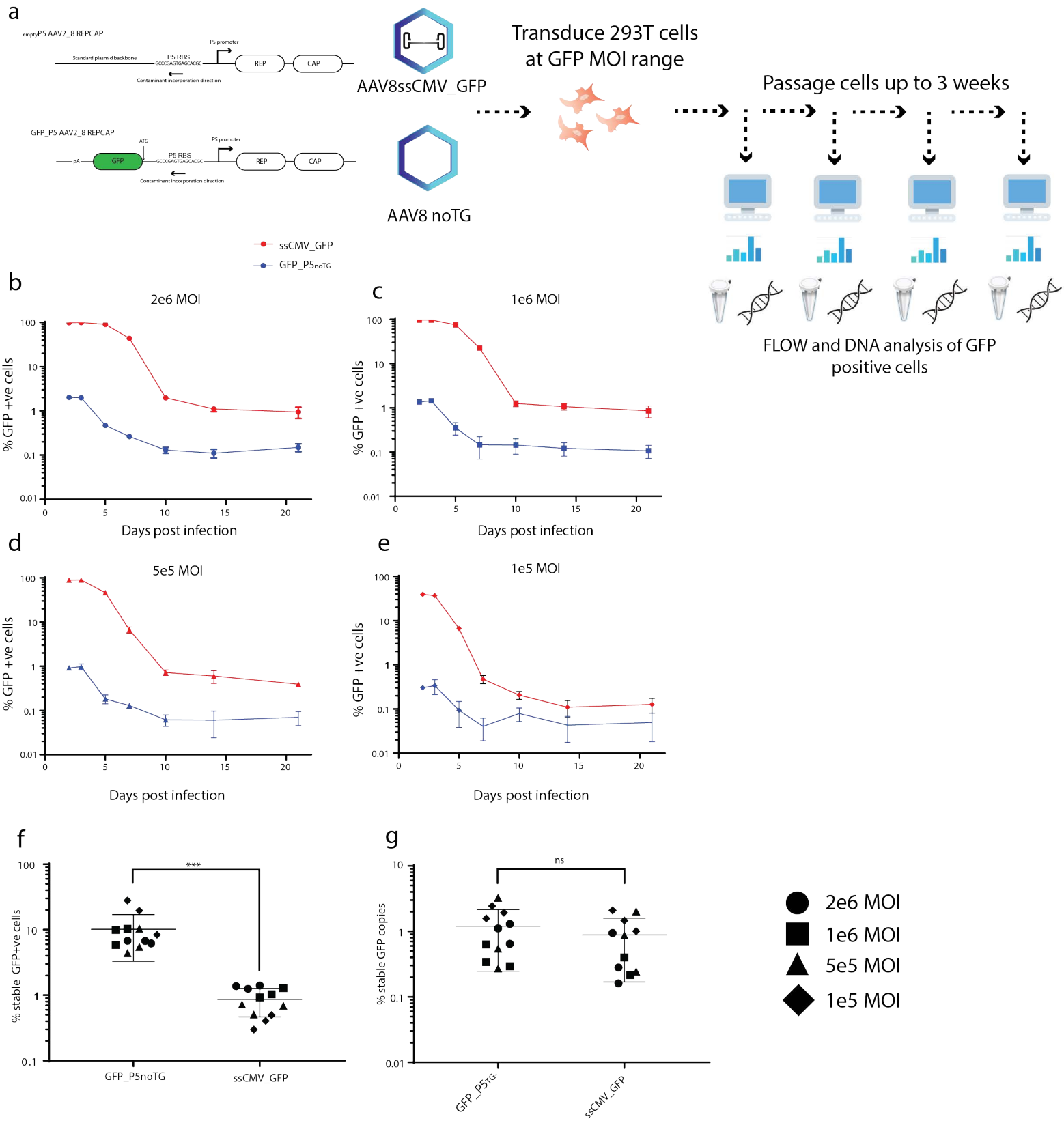


Figure 3.22 – P5 derived contaminants transferred by AAV persist in dividing cells at a similar frequency as ITR flanked vector genomes.

(a) Schematic; AAV8ssCMV_GFP or AAV8 GFP_P5$_{noTG}$ were transduced at different MOIs (2e6;1e6;5e5;1e5) into 293T cells and assayed at 2,3,5,7,10,14, and 21 days post infection by FLOW and at 72 hours and 21 days post infection by gDNA qPCR. **(b-e)** Graphs of % GFP positive cells detected by FLOW in 293T cells transduced with (red) AAV8ssCMV_GFP, or (blue) AAV8 GFP_P5$_{noTG}$ at an MOI of **(b)** 2e6 **(c)** 1e6 **(d)** 5e5 **(e)** 1e5. **(f)** Graph of % GFP expressing cells at 21 days P.I. with AAV8ssCMV_GFP or AAV8 GFP_P5$_{noTG}$ expressed as a percentage of GFP expressing cells at 72 hours P.I. (unpaired t-test) **(d)** Graph of stable DNA copies of GFP gene at 21 days P.I. expressed as a percentage of GFP copies detected at 72 hours P.I acros MOIs. (unpaired t-test)

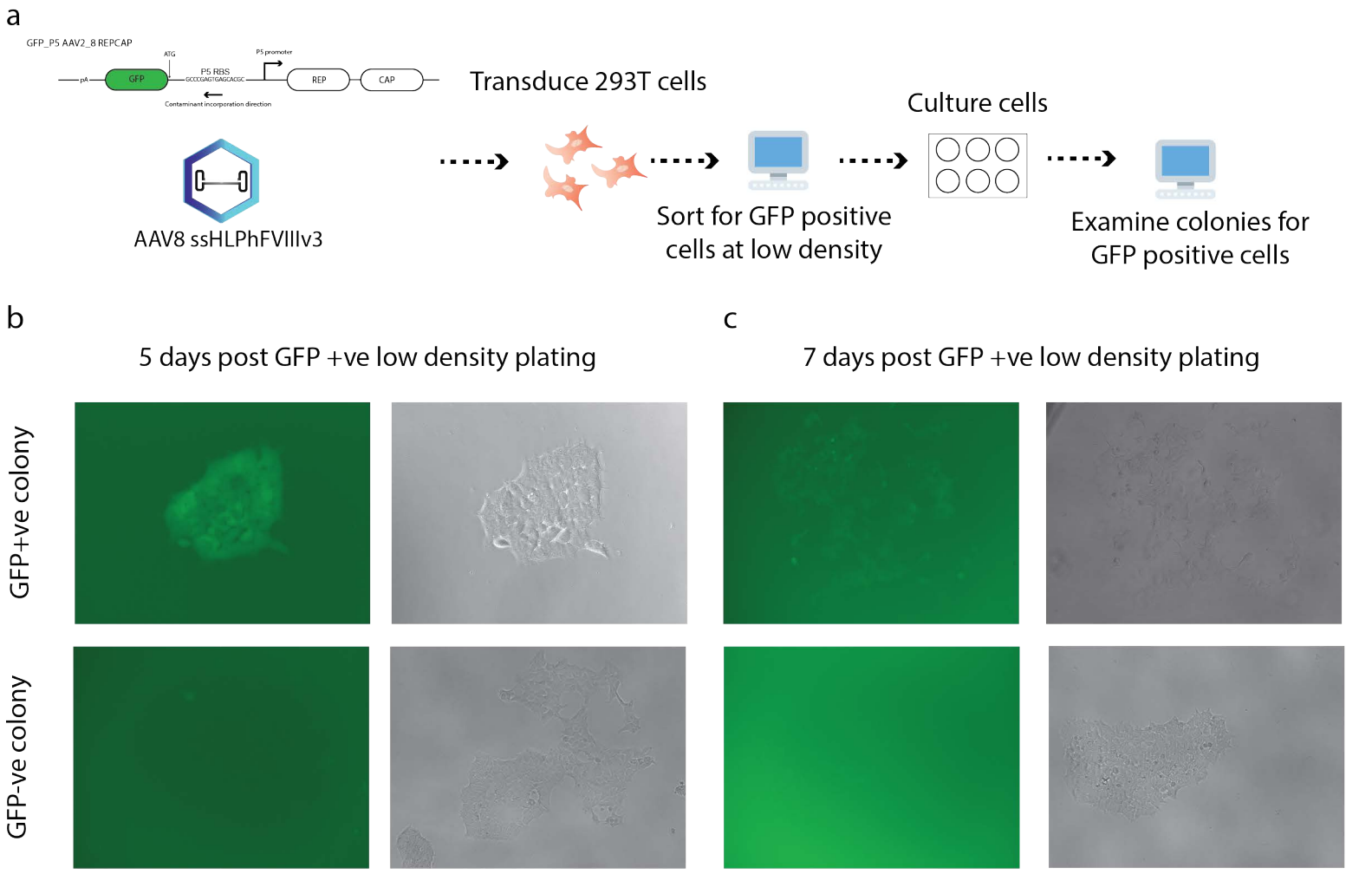


Figure 3.23 – GFP positive colonies grown from 293T cells infected with GFP_P5 contaminant AAV.

(a) Schematic; 293T cells were transduced with AAV8 made with GFP_P5 AAV2_8 REPCAP and sorted for GFP positive cells 72 hours post infection. Cells were plated at low density and allowed to form colonies. **(b)** Flourescent microscopy representative images of (top) GFP positive and (bottom) GFP negative colony at 5 days post plating. **(c)** Flourescent microscopy image of (top) GFP positive and (bottom) GFP negative colony at 7 days post plating.

3.3 Discussion:

This chapter set out to understand the nature of DNA contamination in AAV and its activity post AAV infection of the liver. Given the mapped profile of the DNA contaminants of multiple AAV vectors, it is now clear that most DNA contaminants derived from the producer plasmids are actively incorporated into the AAV capsid. If DNA contamination was a passive and random process, one would expect an equal contribution of contaminants across the length of the plasmid. It is likely that these incorporation events are REP dependent, as all initiation points of contamination have been found adjacent to REP binding sites in the AAV genome. The identified contaminants are not only present in the AAV prep, they are transferred to cells post AAV infection and can persist within the cells for an extended period. A previous study had indicated that any transferred contaminant DNA by AAV was transcriptionally inactive.[466] That study looked at a prep of relatively high purity for this and at amplicons a far distance from the points of contaminant initiation identified in this chapter. The experiments in this chapter have identified that contaminant sequences from AAV can be transcriptionally active post infection and, furthermore, under the appropriate conditions have the potential to be translated into protein. The evidence suggests that the anti-P5 promoter sequence is driving the expression of these contaminant sequences. In the wildtype AAV2 context, anti P5 promoter activity has been observed by RNA sequencing analysis, identifying both minus strand activity of the P5 promoter, and short RNA sequences derived therefrom.[272,284] In recombinant AAV the P5 promoter is retained to drive controlled expression of the large form REP proteins REP78 and REP68, coordinating the timed activation of the P19 promoter contained within the REP coding sequence.[175] What is likely occurring is that

the REP binding site on P5 is causing an incomplete replication event that drives a directional incorporation of DNA upstream of P5 in the producer plasmid. This results in contaminants of different sizes being inserted into AAV, as can be seen from the alkaline gel depicting virions that contain only P5 derived contaminant. Once AAV infection occurs, this DNA is then able, much like an ITR containing transgene, to undergo a productive infection process, perhaps through strand synthesis in an episomal context, or some degree of sequence integration into the genome. Having done this, the anti-P5 promoter is then able to drive transcriptional activity of the adjacent, upstream sequences, an effect that can persist up to 4 months in vivo, as shown by RNAseq analysis of infected livers.

One particularly interesting result is that the level of persistence of AAV contaminant DNA in rapidly overturning cells appears to be equivalent to that of the ITR containing transgene. Many previous studies have touted the ability of the AAV ITRs to facilitate integration into the genome at a high efficiency in a genome editing context, where AAV DNA is used to deliver a donor DNA template.[481–483] Our results from 293T cells at multiple infection doses appear to show no difference in the persistence of AAV DNA in rapidly dividing cells between an ITR flanked cassette and the P5 derived contaminants. In this case, persistence of AAV DNA is a surrogate assay for the general integration rate of these sequences. This would either suggest that the P5 sequence has a greater integrative potential than a standard DNA sequence, or that ITR flanked DNA contains no additional integrative advantage over any other ssDNA present in the cells. It would be interesting to see whether in the presence of the REP protein post AAV infection, this

profile of persistence would be increased for just the ITR flanked vector genome, or for the P5 derived contaminants as well.

The studies of the activity of these contaminants focused primarily on the liver. 293T cells are human embryonic kidney derived and, therefore, it is possible that this would also be seen in kidney transduced with AAV, a tissue that can now be efficiently targeted in vivo with the Anc80 AAV serotype.[484] Presumably, the potential for these contaminants to be transcribed relies on the activity of the anti-P5 sequence in the tissue. It would be interesting to understand whether the contaminant transcription/translation profile is limited just to the liver, or whether this will pose a problem in other key delivery tissues for gene therapy, such as muscle and the retina. There also remains the question of whether these contaminant sequences could contribute to any immune response against the AAV vector. It is already known that during AAV infection the DNA sensor TLR9 can trigger an immune response against the transgene DNA, and that the degree of such a response can be exacerbated by the presence of CpG motifs within the cassette.[133,428] It is therefore likely that since these backbone contaminant DNA fragments are contained within the AAV capsid, the same mechanism of DNA sensing and immune activation by TLR9 would occur, although this needs to be investigated further.

This study used an artificial cassette to study the potential for contaminant translation upstream of the P5 promoter. Whilst such an idealised setting would not occur in most AAV production plasmid setups, the backbone DNA is a little thought about feature of the AAV production system. As such, it is likely that these sequences will vary from plasmid to plasmid and system to system. There are antibiotic resistance genes and other coding sequences present in the backbone of many standard plasmids. If a resistance gene was

in a reverse orientation, perhaps it would be incorporated but no expression of that protein would be seen. If a setup had this in a forward orientation relative to P5, there could be the added immunogenic burden of producing a full-length protein additional to that of the intended therapeutic (Fig. 3.24). Furthermore, certain plasmid configurations used for clinical production may have a greater proportion of CpG sequences than others. The original system used in our study had 154 CpG sites in the 2kb upstream of P5. A strategy that could easily be implemented into reducing this potential risk could be to reduce the numbers of CpG motifs outside of the ITRs and upstream of P5, and also introduce a telomeric repeat sequence TTAGGG,[485] that has been shown when included inside the AAV expression cassette to bind to but not activate TLR9.[486]

This chapter has revealed an important feature and potential safety issue of recombinant AAV preps. It should be noted that there is no required or implemented uniformity to plasmid configuration or backbone sequences used in the production of AAV, and by the very nature of vector construction, different groups across the field will have used different plasmid setups. However, no matter the plasmid setup, in the 293 based production system, the P5 promoter and the inverted terminal repeats are retained. As such, this contamination phenotype and profile is likely to be widespread throughout preparations across the field and the issue of stringent characterisation of vector backbone sequences present in clinical AAV preps and the potential for transcriptional/ translational activity exhibited in on target or off target infected cells for a given gene therapy application should, in the future, be routinely assessed.

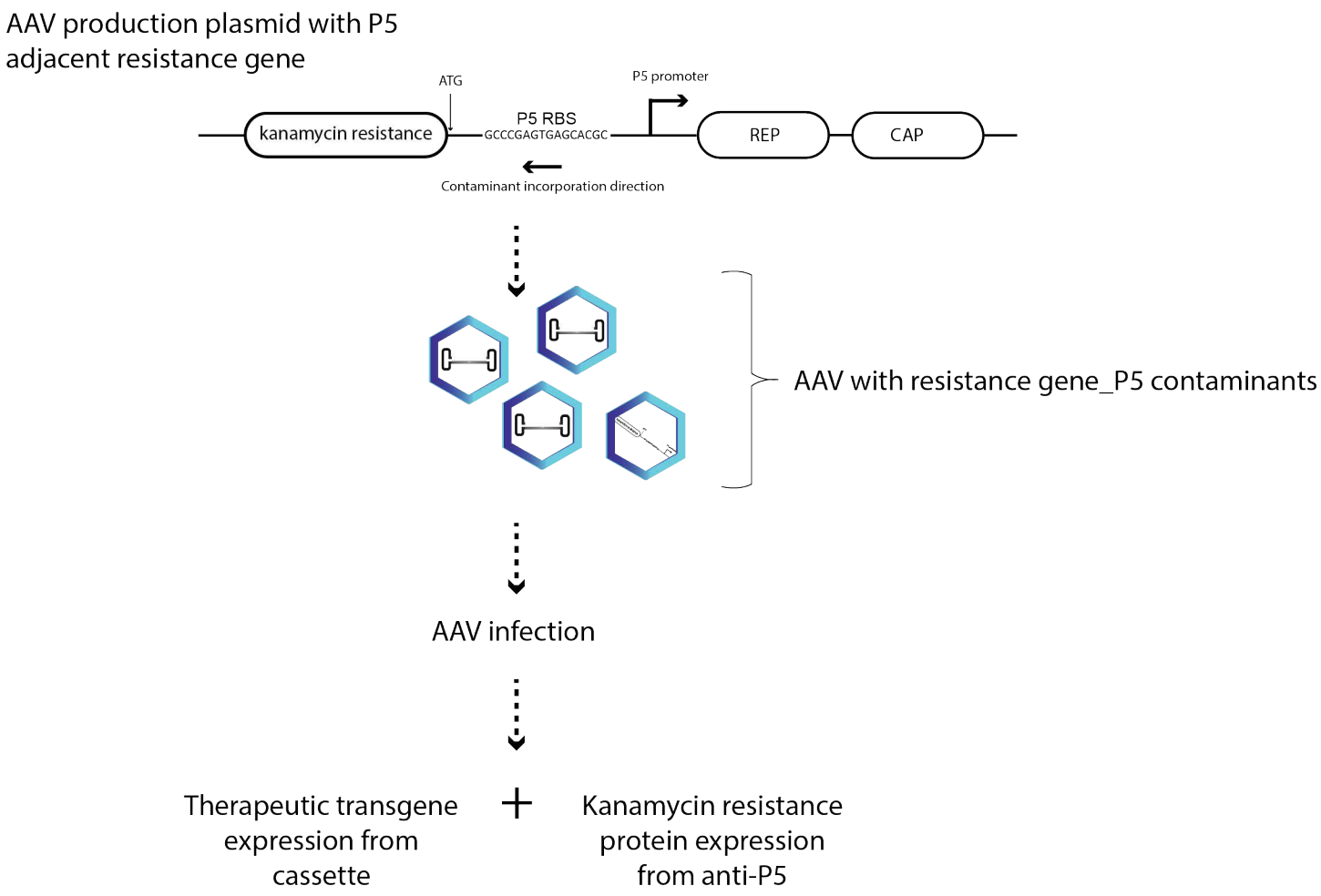


Figure 3.24 – Potential contaminant protein production in REPCAP plasmid backbone

If a gene contained within the plasmid backbone of a P5 promoter containing AAV production plasmid is positioned such that an ATG site is oritented in the orientation and proximity that would allow transcription and translation from the reverse P5 promoter, then the possibility exists for this to be abberantly incorporated into AAV and produced in cells post infection, potentially increasing the immunogenic load on the infected cell.

Chapter 4

Development of a P5 replacement promoter system for production of higher purity recombinant AAV

4. Development of a P5 replacement promoter system for production of higher purity recombinant AAV:

4.1 Introduction:

Having identified the presence and activity of DNA contaminants in AAV, there is an imperative to minimise the presence of these contaminants. Some clinical trials for extremely serious disorders require an AAV dose of up to 2e14vg/kg.[69] Even if a preparation for this purpose contained a low proportion of contaminants, for instance 0.5%, the vector administration would still yield the delivery of 1e12vg/kg of DNA contaminant containing particles to the subject. The contaminants identified have been observed from both outside of the ITRs and upstream of the P5 promoter. Due to the essential nature of the ITRs in packaging the viral genome, it was posited that these sequences could not be altered or disposed of without significant impact on viral titer. The focus of this chapter is, therefore, how contaminants in AAV from the P5 promoter origin on the REPCAP producer can be eliminated and a higher purity vector produced. We investigate the precise sequence motifs that result in incorporation of contaminant DNA from P5 and develop two systems to eliminate the contamination initiation; one based around the papilloma virus early promoter sequences, and one designed to retain all the sequence motifs of the P5 promoter.

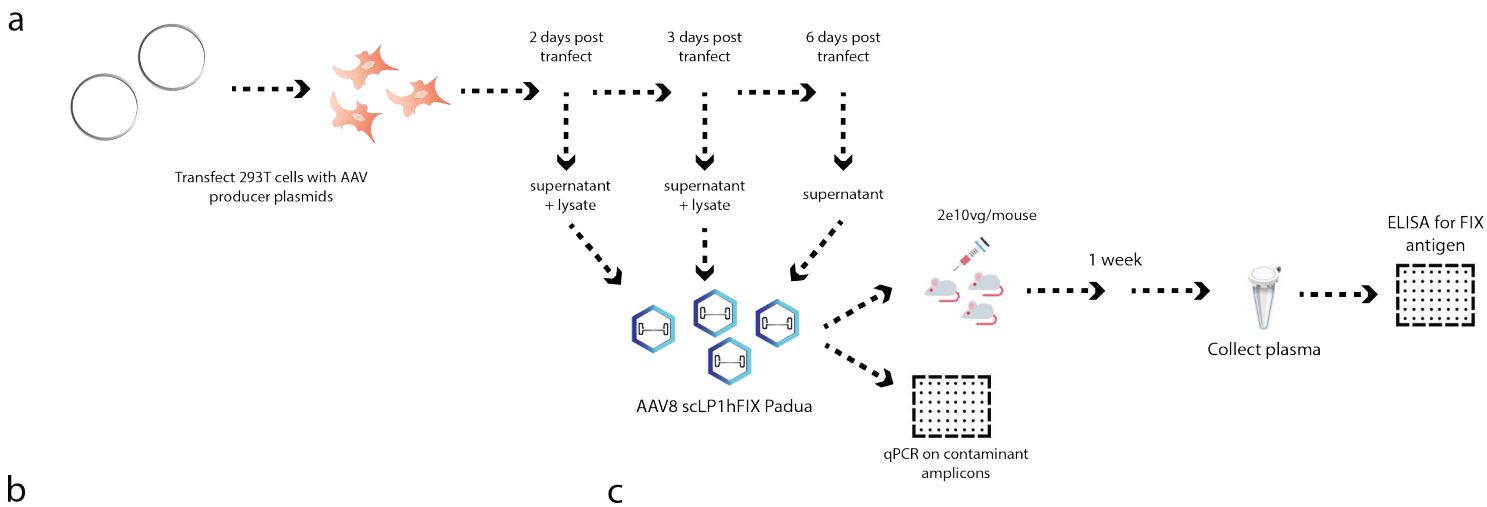
4.2 Results:

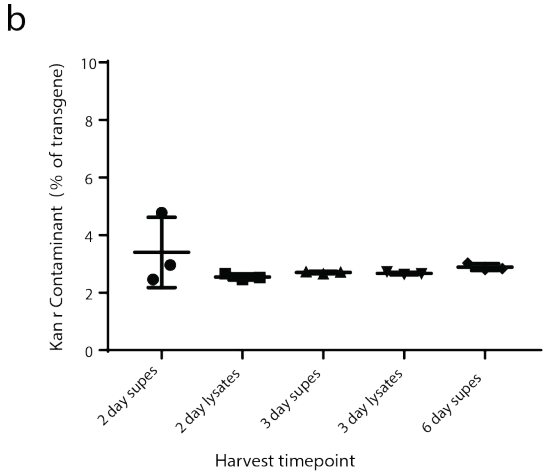
4.2.1: Use of papilloma early promoter sequences to drive REP78/68 expression:

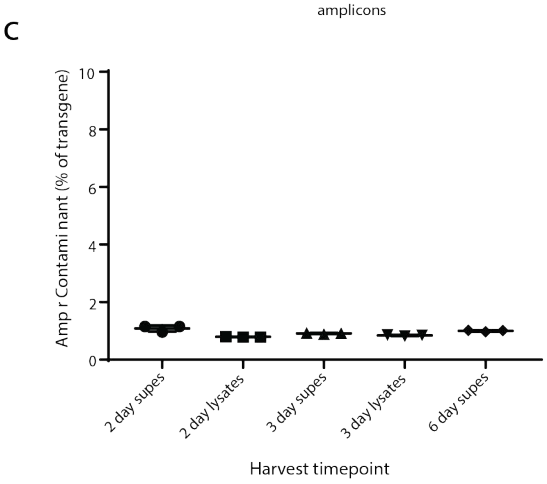
One variable in AAV production is the timepoint at which AAV is harvested. Certain protocols incubate production cells for 2 to 3 days post plasmid transfection.[350,487] Others suggest using a longer timepoint of up to 7 days to boost viral yields and ease purification procedures.[349] Some of these protocols suggest harvesting AAV from cell lysate, whereas others suggest from the supernatant. To determine whether these variables had an effect on the packaging of DNA contaminants into AAV, AAV8scLP1hFIXcoPadua was produced and then harvested at either 2, 3 or 6-days post plasmid transfection (Fig. 4.1). Supernatants and pellets were harvested separately at each timepoint and analysed. qPCR of two separate amplicons present in the production plasmid backbones showed no significant difference in contaminant DNA presence among any of the conditions tested (Fig 4.1. b+c). This suggests that the contaminating DNA phenotype is not altered by changing the incubation timepoint of AAV production.

Given that harvest timepoint appeared not to influence the incorporation of contaminant DNA, investigation was focused on the P5 containing REPCAP plasmid itself. Whilst the ITRs are required for genome packaging in the recombinant AAV system, the P5 promoter merely controls the expression of the large form REP proteins REP78 and REP68. As the REP and CAP proteins are delivered in trans, the ideal scenario to eliminate this anti-P5 contamination and subsequent activity would be to substitute the promoter sequence. However, due to the tight regulation of AAV replication, a simple promoter substitution is not likely to be successful. The P5 promoter efficacy in the context

of AAV replication relies on a negative feedback loop to inhibit its own activity during the AAV replication process (Fig. 4.2). [174,282] More specifically, the timing of P5 self-repression in concert with the later expression of the shorter forms of REP from the P19 promoter is essential to yield this effect.[175] As such, previous attempts to swap out this promoter for constitutively active and more classically 'active' promoters have failed to produce high titer recombinant virus.[275] Therefore, if a promoter substitution was to be effectively implemented in the recombinant AAV production system, it would have to fulfil the requirements of both eliminating the DNA contaminant incorporation phenotype upstream of this promoter region, and facilitating efficient replication of the virus to yield titers that could be implemented for successful gene transfer and thus treatment. To achieve this, we hypothesised that any REP78/68 promoter would 1. need to retain a negative feedback loop, and 2. Not contain a canonical GCTC REP binding site. Early work in AAV biology identified that the REP protein can have an inhibitory effect on non AAV genome associated promoters. This included the Long terminal repeat of the HIV-1 and MSV virus, and proto-oncogenes such as c-myc, c-fos, and H-ras.[183,488,489] However, these promoters all contain GCTC motifs (Fig. 4.3).[488,490–493] Another group of promoters that REP78 has been shown to inhibit are the human papilloma virus early promoter sequences,[191,267,494–496] and in the long control region of the bovine papilloma virus genome.[497] Intriguingly, these promoter sequences from the papilloma virus family are directly inhibited by the REP78 protein despite not containing the GCTC REP DNA binding sequence (Fig. 4.4). Furthermore, in a study of REP inhibition of the HPV18-URR when, in these studies, the c-terminal region of REP78 was mutated, inhibition of P5 was abrogated but inhibition of the HPV-18 P105 promoter was retained, suggesting that the







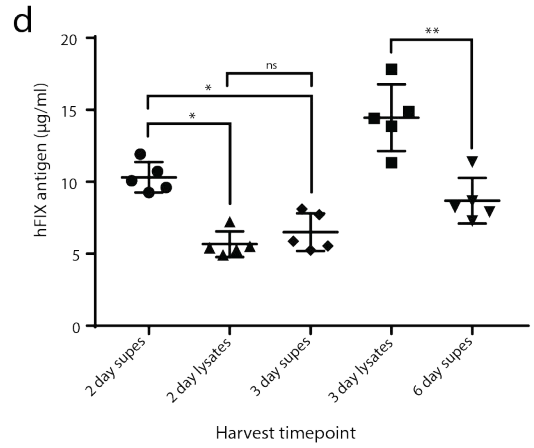


Figure 4.1 - DNA contamination of AAV is detected regardless of AAV harvest timepoint.

(a) Schematic; AAV Producer plasmids were transfected to produce AAV8scLP1hFIXcoPadua. Supernatants were harvested at 2, 3, and 6 days post transfection. Lysates were harvested at 2 and 3 days post transfection. Timepoint harvests were separately purified, analysed for contaminants by qPCR and injected into mice to assay circulating FIX antigen. **(b+c)** qPCR results of AAV timepoint harvest assayed for DNA contamination. *(y axis):* resistance gene DNA expressed as a percentage of the detected FIXPadua expression cassette. *(x axis):* Harvest condition. Analysed by 1 way Anova with Tukey's multiple comparison test. Assayed for **(b)** kanamycin resistance DNA **(c)** ampicillin resistance DNA **(d)** Circulating hFIX antigen detected in CBL57/BL6 mice 1 week post infection of 2e10vg AAV8scLP1hFIXcoPadua. *(y axis):* circulating FIX antigen (µg/ml). *(x axis):* Harvest condition. Analysed by 1 way Anova with Tukey's multiple comparison test.

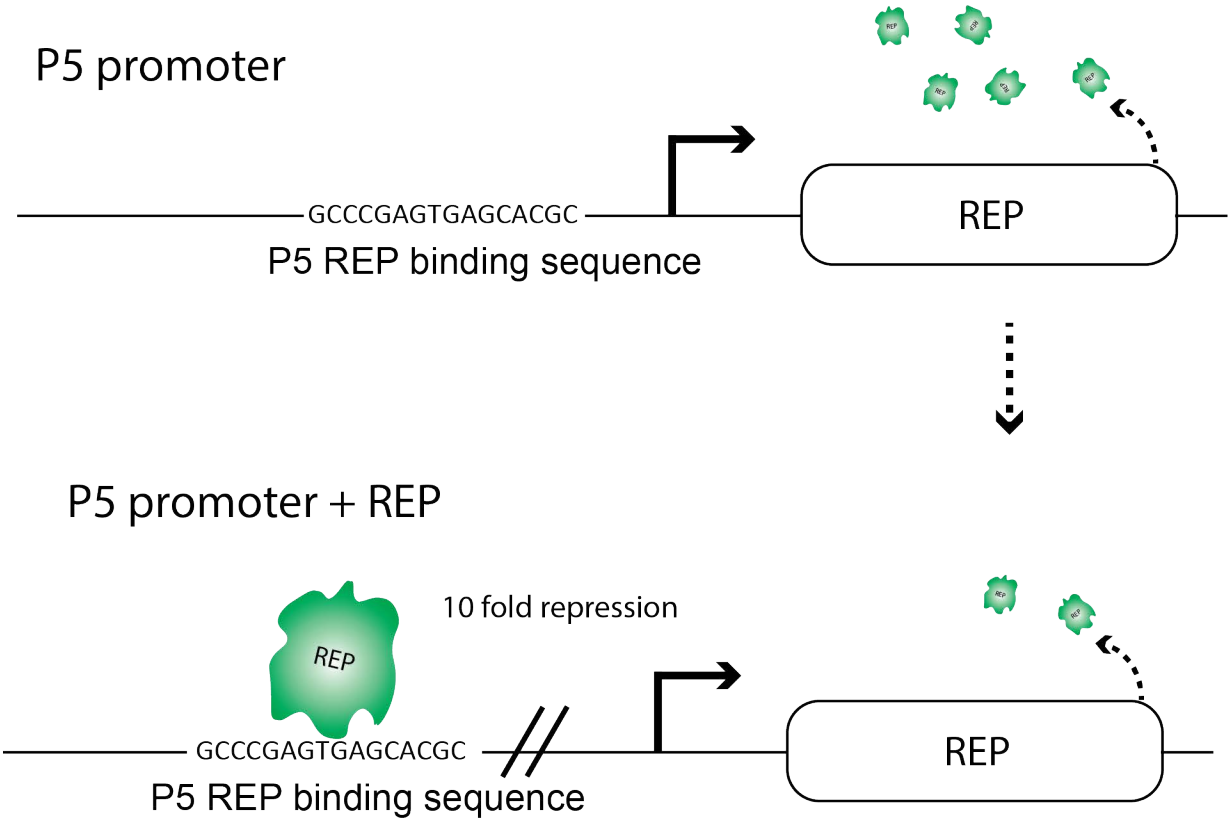


Figure 4.2 – Negative autoregulation of P5 promoter in AAV

Schematic; During AAV replication, the P5 promoter drives expression of the large for REP proteins (Green). The large form REP proteins interact with the P5 TATA box at a 16bp GCTC containing REP binding site, resulting in transcriptional repression and diminishing the further production of large form REP proteins (REP78 and REP68).

mechanism of inhibition may differ from that of P5 and the ITR REP binding sites.[267] In order to produce virus, any chosen promoter would need to have activity in the primary production cell line 293 based systems. There is evidence that the P97 promoter is active in these cells, as HPV16 shows low level replication in 293 cells.[498]

Designs were initially constructed using either the P97 promoter from the HPV16 virus, the P105 promoter from the HPV18 virus, and the upstream regulatory region from the bovine papilloma virus (BPV-URR) (Fig. 4.5). Some of these designs retained elements from the P5 promoter to keep some of the natural regulation intact; i.e. retention of YY1 box sequences, but all promoter sequences were devoid of GCTC motifs (P5+P97 combo and P5/P97 hybrid). Promoter sequences were cloned into an AAV8 capsid plasmid upstream of the REP protein of AAV2, the AAV REP protein most commonly used in AAV production.[499] Backbone plasmid sequences were kept identical to allow effective comparison of contaminant incorporation. REPCAP plasmids were transfected into 293T cells to make AAV8ssCMV_GFP. AAV was produced, purified and titered by qPCR for presence of the GFP transgene (Fig. 4.6). As expected, the P5 promoter positive control produced abundant AAV8. Designs that utilised either the P97 promoter from HPV16 or the URR of BPV were also able to produce AAV effectively, with three designs producing a titer almost half as much as the original P5 based setup (Fig. 4.6.b). These were the P5+P97 combo, HPV16-P97_ATG and P5/P97 hybrid configurations that produced 44.2%, 44.5% and 49.8% of the standard P5 production levels respectively. The HPV18 P105 configuration did not appear to package AAV with any useful efficiency, and detected transgene was approximately 3 logs lower than that of P5.

a P5 promoter sequence

TAGAGGTCCTGTATTAGAGGTCACGTGAGTGTTTTGCGACATTTTGCGACACCATGTGGTCACGCTGGGTATTTA
AGCCCGAGTGAGCACGCAGGGTCTCCATTTTGAAGCGGGAGGTTTGAACGCGCAGCCACCACG

b HIV-1 LTR

TGGAAGGGCTAATTCACTCCCACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCT
GATTAGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCA
GTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGG
GATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAG
AGCTGCATCGGGAGTACTTCAAGAACTGCTGATATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAG
GGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTTGCATATAAGCAGCTGCTTTTTGCCTG
TACT

c MSV LTR

CTGTACCCGCGCTTATTGCTGCCCAGCTCTATAAAAAGGGTAAGAACCCCACACTCGGCG

d HA-Ras promoter

GAGCAACCCCGAGCTCGGCTCCGGTCTCCAGCCAAGCCCAACCCCGAGAGGCCGCGGCCCTACTGGCTCCGCCTC
CCGCGTTGCTCCCGGAAGCCCCGCCCGACCGCGGCTCCTGACAGACGGGCCGCTCAGCCAACCGGGGTGGGGCG
GGGCCCGATGGCGCGCAGCCACAATGGTAGGCCGCGCCTGGCAGACGGACGGGCGCGGGGCGGGGCGTGCGC
AGGCCCGCCCGAGTCTCCGCCGCCCGTGCCC

e C-Fos promoter

ATCTCTGAGCCTCAGAACTGTCTTCAGTTTCCGTACAAGGGTAAAAAGGCGCTCTCTGCCCCATCCCCCCCGACCTC
GGGAACAAGGGTCCGCATTGAACCAGGTGCGAATGTTCTCTCTCATTCTGCGCCGTTCCCGCCTCCCCTCCCCCAG
CCGCGGCCCCCGCCTCCCCCCGCACTGCACCCTCGGTGTTGGCTGCAGCCCGCGAGCAGTTCCCGTCAATCCCTCC
CCCCTTACACAGGATGTCCATATTAGGACATCTGCGTCAGCAGGTTTCCACGGCCTTTCCCTGTAGCCCTGGGGGG
AGCCATCCCCGAAACCCCTCATCTTGGGGGGCCCACGAGACCTCTGAGACAGGAACTGCGAAATGCTCACGAGAT
TAGGACACGCGCCAAGGCGGGGGCAGGGAGCTGCGAGCCGTGGGACGCAGCCGGGCGGCCGCAGAAGCGCCC
AGGCCCGCGCGCCACCCCTCTGGCGCCACCGTGGTTGAGCCCGTGACGTTTACACTCATTCATAAAACGCTTGTT
ATAAAAGCAGTGGCTGCGGCGCCTCGTACT

f C-Myc promoter

GAGGGAGCAAAAGAAAATGGTAGGCGCGCGTAGTTAATTCATGCGGCTCTCTTACTCTGTTTACATCCTAGAGCTA
GAGTGCTCGGCTGCCCGGCTGAGTCTCCTCCCCACCTTCCCCACCCtCCCCACCCtCCCCATAAGCGCCCTCCCGGGT
TCCCAAAGCAGAGGGCGTGGGGGGAAAAGAAAAAAGATCCTCTCTCGCTAATCTCCGCCCACCGGCCCTTTATAAT
GCGAGGGTCTGGACGGCTGAGGACCCCCGAGCTGCTCGCGGCCGCCACCGCCGGGCCCCGGCCGTCCCTGGCTC
CCCTCCTGCCTCGAGAAGGGCAGGGCtTCTCAGAGGCTTGGCGGGAAAAAGAACGGAGGGAGGGATCGCGCTGA
GTATAAAAGCCGGTTTTCGGGGCTTTATCTAA

Figure 4.3 – REP interacting sequences containing GCTC binding motifs

REP interacting sequences containing GCTC/GAGC REP recogntion sequences highlighted in red. Sequences listed from **(a)** The AAV2 P5 promoter **(b)** Human Immunodeficiency Virus type 1 Long terminal repeat **(c)** Murine Sarcoma Virus Long terminal repeat **(d)** Ha-Ras (HRAS) protooncogene promoter **(e)** C-Fos protooncogene promoter **(f)** C-Myc protooncogene promoter.

a HPV16-P97 promoter sequence

ACTACAATAATTCATGTATAAAACTAAGGGCGTAACCGAAATCGGTTGAACCGAAACCGGTTAGTATAAAAGCA
GACATTTTATGCACCAAAAGAGAACTGCAACG

b HPV18-P105 promoter sequence:

ATTAATACTTTTAACAATTGTAGTATATAAAAAAGGGAGTAACCGAAAACGGTCGGGACCGAAAACGGTGTAT
ATAAAAGATGTGAGAAACACACCACAATACT ACG

c BPV upstream regulatory region sequence:

CGCCTGGCACCGAATCCTGCCTTCTCAGCGAAAATGAATAATTGCTTTGTTGGCAAGAAACTAAGCATCAATGG
GACGCGTGCAAAGC ACCGGCGGCGGTAGATGCGGGGTAAGTACTGAATTTTAATTCGACCTATCCCGGTAAAG
CGAAAGCGACACGCTTTTTTTTCACACATAGCGGGACCGAACACGTTATAAGTATCGATTAGGTCTATTTTTGTCT
CTCTGTCGGAACCAGAACTGGTAAAAGTTTCCATTGCGTCTGGGCTTGTCTATCATTGCGTCTCTATGGTTTTTGG
AGGATTAGACGGGGCCACCAGTAATGGTGCATAGCGGATGTCTGT ACCGCCATCGGTGCACCGATATAGGTTTG
GGGCTCCCCAAGGGACTGCTGGGATGACAGCTTCATATTATATTGAATGGGCGCATAATCAGCTTAATTGGTGA
GGACAAGCTACAAGTTGTAACCTGATCTCCACAAAGTACGTTGCCGGTCGGGGTCAAACCGTCTTCGGTGCTCG
AAACCGCCTTAAACTACAGACAGGTCCCAGCCAAGTAGGCGGATCAAAACCTCAAAAAGGCGGGAGCCAATCA
AAATGC AGCATTATATTTTAAGCTCACCGAAACCGGTAAGTAAAGACTATGTATTTTTTCCCAGTGAATAATTGT
TGTTAACAATAATCACACCATCACCGTTTTTTCAAGCGGGAAAAAATAGCCAGCTAACTATAAAAAGCTGCTGAC
AGACCCCGGTTTTCAC ACG

Figure 4.4 – REP interacting promoter sequences from Papilloma viruses lack GCTC REP binding sites

REP interacting sequences absent of GCTC/GAGC REP recogntion sequences. Sequences listed from **(a)** Human Papilloma Virus type 16 P97 promoter **(b)** Human Papilloma Virus type 18 P105 promoter **(c)** Bovine Papilloma Virus upstream regulatory region

a **P5 promoter sequence:**

TAGAGGTCCTGTATTAGAGGTCACGTGAGTGTTTTGCGACATTTTGCGACACCATGTGGTCACGCTGGGTATTTA
AGCCCGAGTGAGCACGCAGGGTCTCCATTTTGAAGCGGGAGGTTTGAACGCGCAGCCACCACG

b **HPV16-P97 promoter sequence:**

ACTACAATAATTCATGTATAAAACTAAGGGCGTAACCGAAATCGGTTGAACCGAAACCGGTTAGTATAAAAGCA
GACATTTTATGCACCAAAAGAGAACTGCAACG

c **HPV16-P97 + ATG promoter sequence:**

ACTACAATAATTCATGTATAAAACTAAGGGCGTAACCGAAATCGGTTGAACCGAAACCGGTTAGTATAAAAGCA
GACATTTTATGCACCAAAAGAGAACTGCAATG

d **HPV18-P105 promoter sequence:**

ATTAATACTTTTAACAATTGTAGTATATAAAAAAGGGAGTAACCGAAAACGGTCGGGACCGAAAACGGTGTAT
ATAAAAGATGTGAGAAACACACCACAATACTACG

e **BPV upstream regulatory region sequence:**

CGCCTGGCACCGAATCCTGCCTTCTCAGCGAAAATGAATAATTGCTTTGTTGGCAAGAAACTAAGCATCAATGG
GACGCGTGCAAAGCACCGGCGGCGGTAGATGCGGGGTAAGTACTGAATTTTAATTCGACCTATCCCGGTAAAG
CGAAAGCGACACGCTTTTTTTTCACACATAGCGGGACCGAACACGTTATAAGTATCGATTAGGTCTATTTTTGTCT
CTCTGTCGGAACCAGAACTGGTAAAAGTTTCCATTGCGTCTGGGCTTGTCTATCATTGCGTCTCTATGGTTTTTGG
AGGATTAGACGGGGCCACCAGTAATGGTGCATAGCGGATGTCTGTACCGCCATCGGTGCACCGATATAGGTTTG
GGGCTCCCCAAGGGACTGCTGGGATGACAGCTTCATATTATATTGAATGGGCGCATAATCAGCTTAATTGGTGA
GGACAAGCTACAAGTTGTAACCTGATCTCCACAAAGTACGTTGCCGGTCGGGGTCAAACCGTCTTCGGTGCTCG
AAACCGCCTTAAACTACAGACAGGTCCCAGCCAAGTAGGCGGATCAAAACCTCAAAAAGGCGGGAGCCAATCA
AAATGCAGCATTATATTTTAAGCTCACCGAAACCGGTAAGTAAAGACTATGTATTTTTTCCCAGTGAATAATTGT
TGTTAACAATAATCACACCATCACCGTTTTTTCAAGCGGGAAAAAATAGCCAGCTAACTATAAAAAGCTGCTGAC
AGACCCCGGTTTTCACACG

f **P5+P97 combination promoter sequence:**

TAGAGGTCCTGTATTAGAGGTCACGTGAGTGTTTTGCGACATTTTGCGACACCATGTGGTCACGCTGGGACTAC
AATAATTCATGTATAAAACTAAGGGCGTAACCGAAATCGGTTGAACCGAAACCGGTTAGTATAAAAGCAGACA
TTTTATGCACCAAAAGAGAACTGCAACG

g **P5/P97 hybrid promoter sequence:**

TAGAGGTCCTGTATTAGAGGTCACGTGAGTGTTTTGCGACATTTTGCGACACCATGTGGTCACGCTGGGATGTAT
AAAACTAAGGGCGTAACCGAAATCGGTTGAACCGAAACTCCATTTTGAAGCGGGAGGTTTGAACGCGCAGCCA
CCACG

Figure 4.5 – Papilloma based recombinant REP78/68 promoter designs for AAV production

Promoter sequences designed and placed directly upstream of REP78/68 start site to produce AAV8 **(a)** The AAV2 P5 promoter **(b)** Human Papilloma Virus type 16 P97 promoter **(c)** Human Papilloma Virus type 16 P97 promoter with optimised start codon **(d)** Human Papilloma Virus type 18 P105 promoter **(e)** Bovine Papilloma Virus upstream regulatory region **(f)** P5 sequence upstream of the TATA box followed by the HPV16 P97 promoter. **(g)** P5 sequence with P5 TATA box replaced with the P97 promoter REP binding site.

(red) GCTC REP binding site *(orange)* HPV16-P97 REP binding site *(green)* start codon optimisation *(blue)* YY1 box.

The setup using the BPV-URR and HPV16-P97 without and ATG start codon produced AAV successfully but also inefficiently yielding 5.4% and 21.4% of the P5 production titer respectively. To test whether these non GCTC motif containing promoters had eliminated the contamination incorporation upstream, amplicons in the kanamycin resistance genes, located closely upstream of the promoter in these REPCAP plasmids were assayed. As expected, the P5 promoter-produced AAV had detectable contamination upstream of P5. All the P5 replacement setups showed a dramatic decrease in the raw values of contaminant copies present (Fig. 4.6.c+d). With the level of P5 contamination normalised to 1, the mean levels of contamination in the P5 replacement AAV were as follows: Contaminant amplicon 1; HPV16-P97 - 0.06, HPV18-P105 - 0.04, BPV_URR 0.04, P5+P97 combo – 0.05, HPV16-P97_ATG – 0.03, P5/P97 hybrid – 0.07. Contaminant amplicon 2; HPV16-P97 - 0.06, HPV18-P105 - 0.05, BPV_URR 0.04, P5+P97 combo – 0.06, HPV16-P97_ATG – 0.04, P5/P97 hybrid – 0.13 (n=2 per promoter setup). Given the differences in production capability, these raw numbers were compared against the titer of each individual configuration to yield the overall percent contamination of contaminant amplicons for each P5 replacement construct (Fig. 4.6.e+f). For contaminant amplicon 1 the P5 promoter design yielded 1.2% contaminating sequence. The HPV-108 promoter, which produced a 3 log lower titer, yielded a higher ratio of contaminants (37.3% of the GFP transgene). All other promoter setups yielded lower levels. HPV16-P97 - 0.3%, BPV_URR 0.9%, P5+P97 combo – 0.1%, HPV16-P97_ATG – 0.1%, P5/P97 hybrid – 0.2%. For amplicon 2 the pattern was similar; the P5 promoter led to contamination of 1.1% of the transgene titer. Aside from HPV-108 (42.1%), all other promoter setups led to contaminants that amounted to a lower percentage of transgene titer. HPV16-P97 -

0.3%, BPV_URR 0.7%, P5+P97 combo – 0.1%, HPV16-P97_ATG – 0.1%, P5/P97 hybrid – 0.3%.

To ensure that infectivity was not compromised in this production setup, 293T cells were infected with a range of MOIs (5e5, 2e5, and 1e5) of AAV8ssCMV_GFP from the P5 replacement systems that produced a sufficient quantity of AAV (Fig. 4.7) All systems resulted in infectious AAV as detected by fluorescent microscopy (Fig. 4.7.b) and FACS analysis that appeared to transduce 293T cells as efficiently as the P5-produced system across the 3 MOIs (n=1) (Fig. 4.7.c). Following this preliminary experiment, the P5 replacement systems with the highest production yields were then used to produce AAV8ssHLPhFVIII to verify the contamination phenotype (Fig. 4.8) Again, AAV8 was produced to 28 - 50% of the P5 based system with the replacement REP promoters (Fig. 4.8.b). The observed titers with the P5 replacement constructs were still lower relative to that of P5; P97+P5 combo – 0.50 ± 0.02, HPV16-P97_ATG - 0.38 ± 0.06, P5/P97 hybrid – 0.28 ± 0.03 (n=4). When contaminant amplicons were measured in this context, a decrease was again observed both when this was expressed as a raw value of detected sequence copies ($P<0.0001$ for all constructs) and percentage of the transgene titer (<0.0005 for all constructs) (Fig. 4.8.c+d). To ensure these observed results for production capabilities were not influenced by purification procedures post-harvest we developed an assay to analyse the titer and contaminants in virus at the point of harvest from crude supernatants (Appendix 1). 293T cells were transfected in 6 well plates to increase throughput of construct testing and analysis. (Fig. 4.9) AAV8ssHLPhFVIII was produced using the 6 designed P5 replacement constructs. All promoter constructs produced significantly lower titers than P5 ($P<0.0001$).

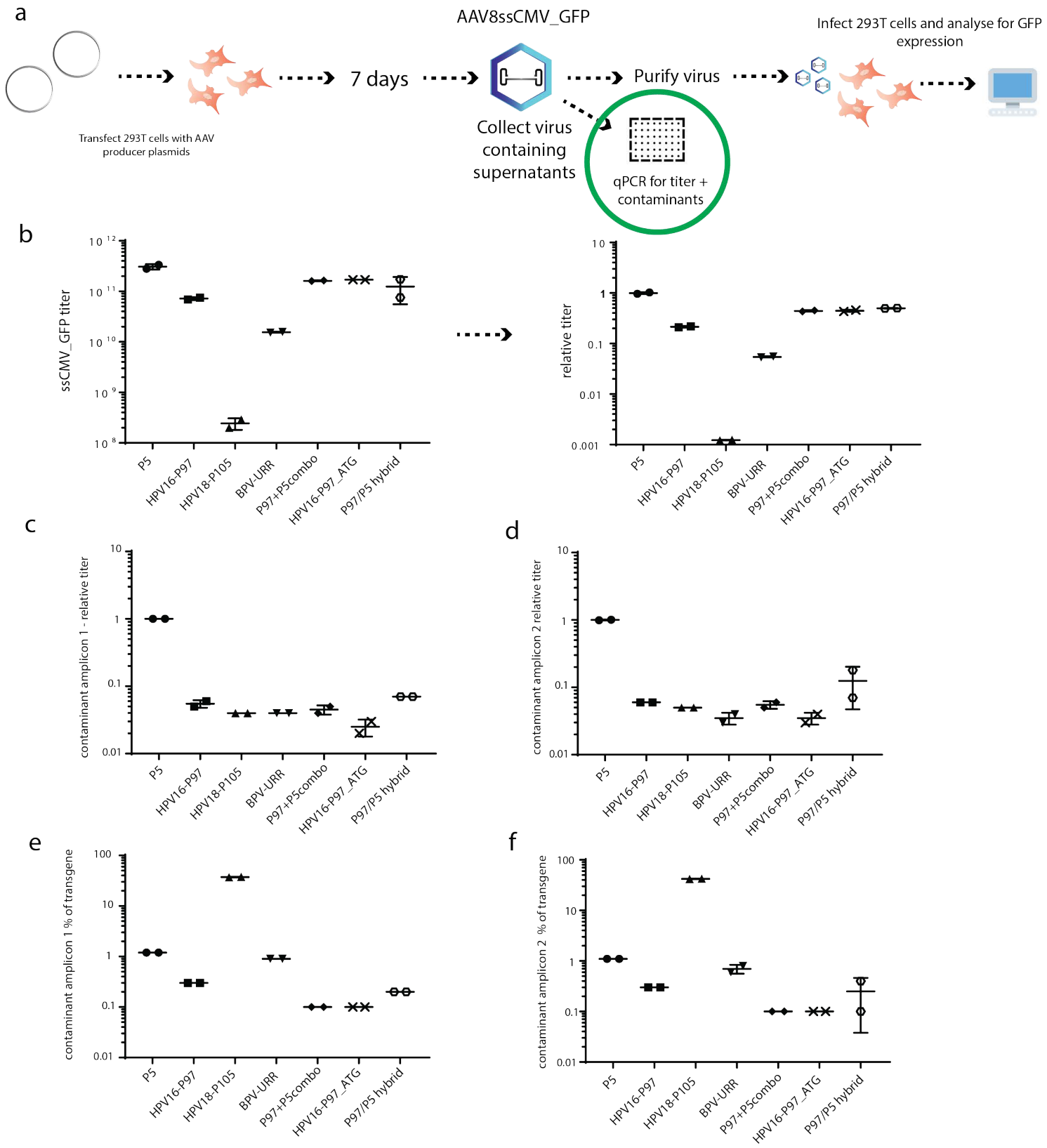


Figure 4.6 - Recombinant Papilloma based REP78/68 promoters produce AAV effectively and reduce upstream contamination

(a) Schematic; AAV8ssCMV_GFP was produced using AAV2_8 REPCAP plasmids with recombinant papilloma based designs. **(b)** qPCR results of AAV8ssCMV_GFP titer production efficiency with different REP78/68 promoter designs expressed as (left) CMV_GFP titer (vg/ml) (right) CMV_GFP titer normalised against P5 promoter production efficiency **(c-f)** qPCR results of AAV8ssCMV_GFP contaminant levels for two different upstream contaminant amplicons expressed as: **(c)** amplicon 1 abundance normalised against P5 promoter contamination **(d)** amplicon 2 abundance normalised against P5 promoter contamination **(e)** amplicon 1 abundance expressed as a percentage of the CMV_GFP expression cassette **(f)** amplicon 2 abundance expressed as a percentage of the CMV_GFP expression cassette.

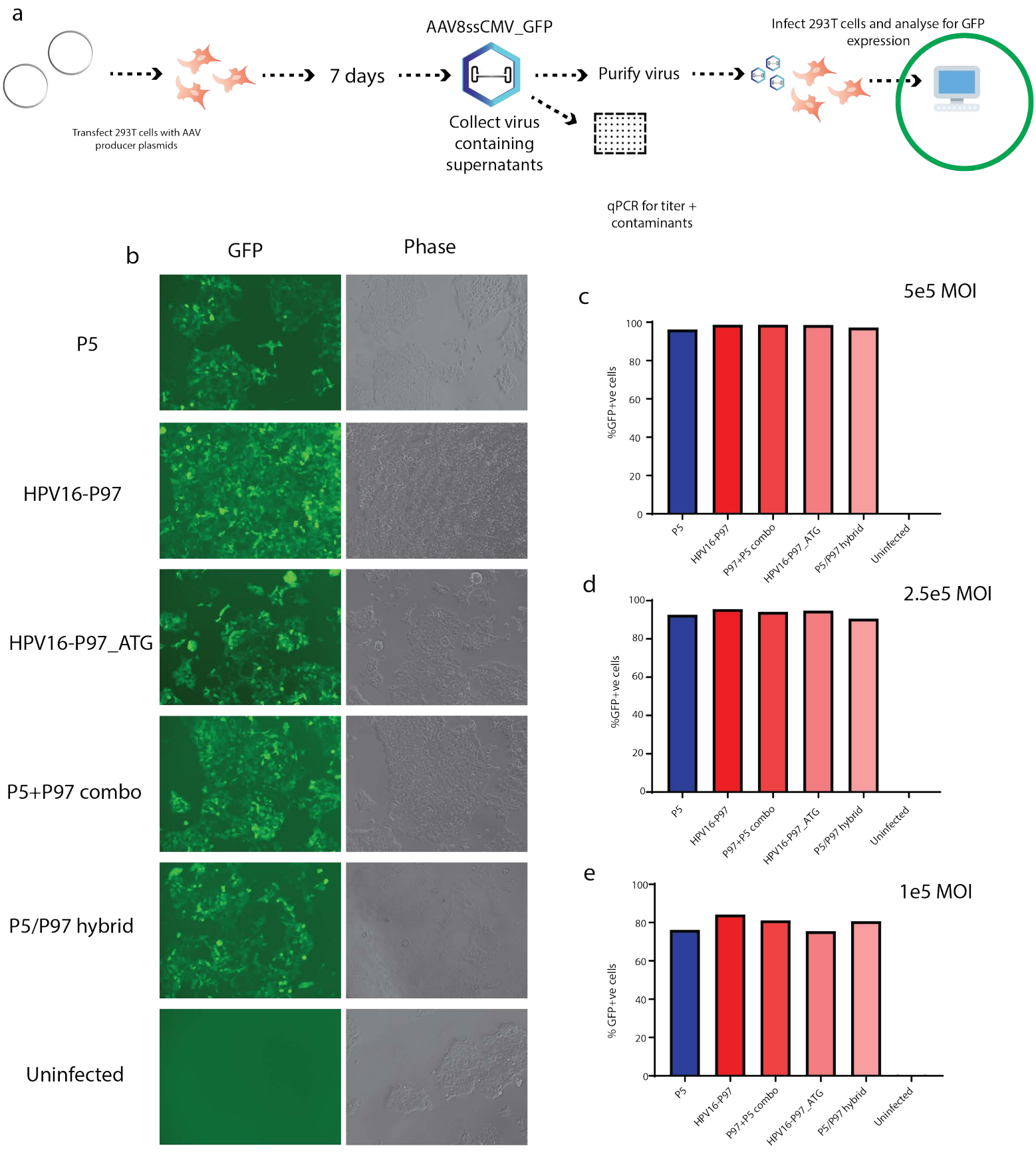


Figure 4.7 - AAV produced with Papilloma based promoter designs transduce cells efficiently.

(a) Schematic; AAV8ssCMV_GFP produced with recombinant papilloma based P5 replacement promoters was used to transduce 293T cells at a range of MOIS. **(b)** Flourescent microscopy of 293T cells 48 hours post infection with AAV8ssCMV_GFP produced with P5 replacement promoters. **(c-e)** Bar graph of FLOW results showing % of GFP expressing cells infected with AAV8ssCMV_GFP produced with *(blue)* P5 or *(red)* P5 replacement promoters at **(c)** 5e5 MOI **(d)** 2.5e5 MOI **(e)** 1e5 MOI.

The crude supernatant samples showed the same pattern as the purified AAV samples with respect to titer relative to that achieved by the P5 promoter (Fig. 4.9.b), with the P5+P97combo (0.23 ± 0.04), HPV16-P97_ATG (0.44 ± 0.06) and P5/P97 hybrid (0.18 ± 0.01) producing the highest amounts of virus relative to P5, and HPV18-P105 producing AAV 3 logs lower than P5 (0.001 ± 0.0003). When the samples were assayed for the contaminant amplicon upstream of P5, again all P5 replacement promoters produced AAV that had significantly lower levels of contamination than P5 (Fig. 4.9.c+d) ($P<0.001$). When this was expressed as a percentage of the observed titer, HPV18-P105 produced a contamination level that was higher than the transgene (177.3% ± 22.47). It can therefore be assumed that active viral packaging is not occurring when this promoter is used to drive REP78. In those samples that produced virus a relationship between the efficiency of the packaging and the percent contamination seen was again observed. With the tested contaminant amplicon, the P5 promoter production system resulted in contamination of 3.89% ± 0.29 of the expression cassette titer. Three constructs resulted in significantly lower levels of percent contamination (HPV16-P97_ATG – 0.48%; $P=0.0051$) (P5/P97hybrid – 1.09% ± 0.14; $P=0.0190$) (P97+P5combo – 0.88% ± 0.08; $P=0.0121$), whereas, two of the configurations did not produce significantly lower contaminant levels; the HPV16-P97 (1.77% ±0.07; $P=0.08$), and BPV-URR which yielded contamination significantly higher than P5 when expressed as a percentage of titer. (11.44% ± 1.36; $P<0.0001$).

To determine whether this phenotype held true across serotypes, an AAV2 capsid version of the P5+P97 combo promoter was constructed (Fig. 4.10). AAV2 produced with the P5+P97 combo promoter produced virus to the same log as the P5 promoter, although

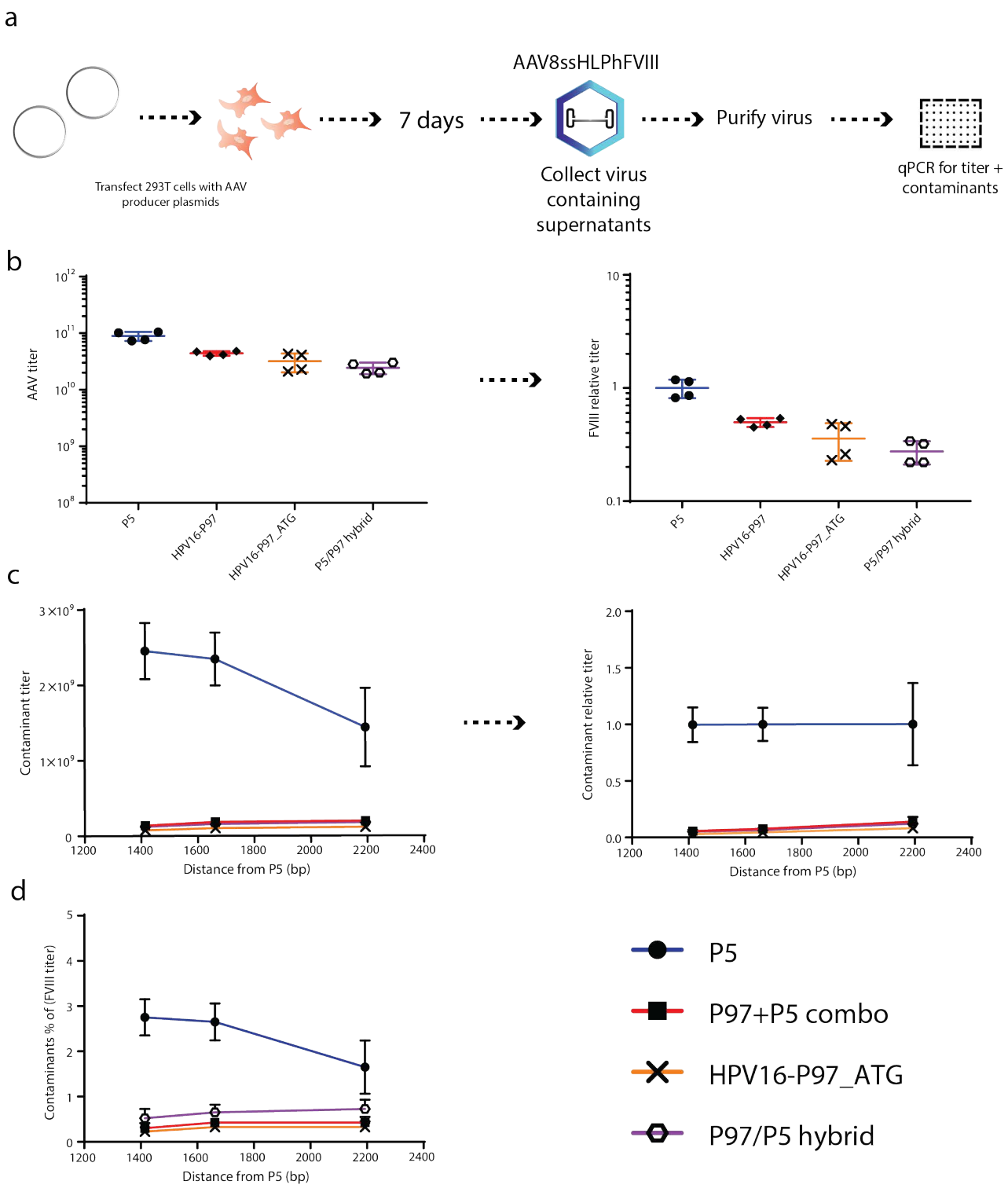


Figure 4.8 - Recombinant Papilloma based REP78/68 promoters significantly reduce upstream contamination

(a) Schematic; AAV8ssHLPhFVIII was produced using AAV2_8 REPCAP plasmids with HPV16-P97 based REP78/68 promoter designs and assayed for titer and contaminant abundance **(b)** qPCR results of AAV8ssHLPh-FVIII titer production efficiency with different REP78/68 promoter designs expressed as (left) HLPhFVIII titer (vg/ml) and (right) HLPhFVIII titer normalised against P5 promoter production efficiency **(c+d)** qPCR results of AAV8ssHLPhFVIII contaminant levels for three upstream contaminant amplicons at different distances from the promoter expressed as **(c)** (left) contaminant titer (vg/ml); (right) contaminant amplicon abundance normalised against P5 promoter contamination **(d)** contaminant abundance expressed as a percentage of HLPhFVIII titer. *Graphs b-d analysed by One Way Anova with Dunnet's multiple comparison test.*

as seen in the AAV8 setup the relative amount was significantly decreased as compared to P5. (0.30 ± 0.01; P = 0.016) (Fig. 4.10.b). A contaminant amplicon upstream of the P5 sequence was again seen in abundance in the P5 promoter version of AAV2ssHLPhFVIII and significantly reduced in the P5+P97combo promoter replacement (Fig. 4.10c). When this was analysed as a percentage of the transgene contaminants, the total contamination of P5_AAV2ssHLPhFVIII was 5.42 ± 0.83. In the P5+P97combo configuration, the contamination was decreased (2.84 ±0.84), but did not reach significance (P=0.06) (Fig. 4.10d).

4.2.2 P5 replacement configurations with lower AAV yields results in increased DNA contamination from outside of the ITRs:

To assess if the initiation of contamination at the REP78/68 promoter had been eliminated in these P5 replacement constructs, deep sequencing of AAV8ssHLPhFVIII viral particles was performed for vector produced with either a P5 driven or P5+P97combo setup (Fig. 4.11). When sequence reads were mapped back to producer plasmids, the P5-produced virus showed incorporation upstream of P5 as expected. For the P5+P97combo produced virus however, no contamination was detected from this site. (Fig. 4.11.b) As expected, both setups retained a contamination phenotype outside of the ITRs (Fig. 4.11.c). When the identity of all the reads was quantified, the reads that mapped to the REPCAP plasmid dropped from 6.7% of the total reads with P5 to 0.29% in the P5+P97 combo promoter. Surprisingly, when reads from the vector backbone plasmid, i.e. outside of the ITRs were quantified in this way, there was a dramatic increase; from 18.72% in P5 to 26.47% with P5+P97 combo. This could possibly be due to the lower expression cassette yield from

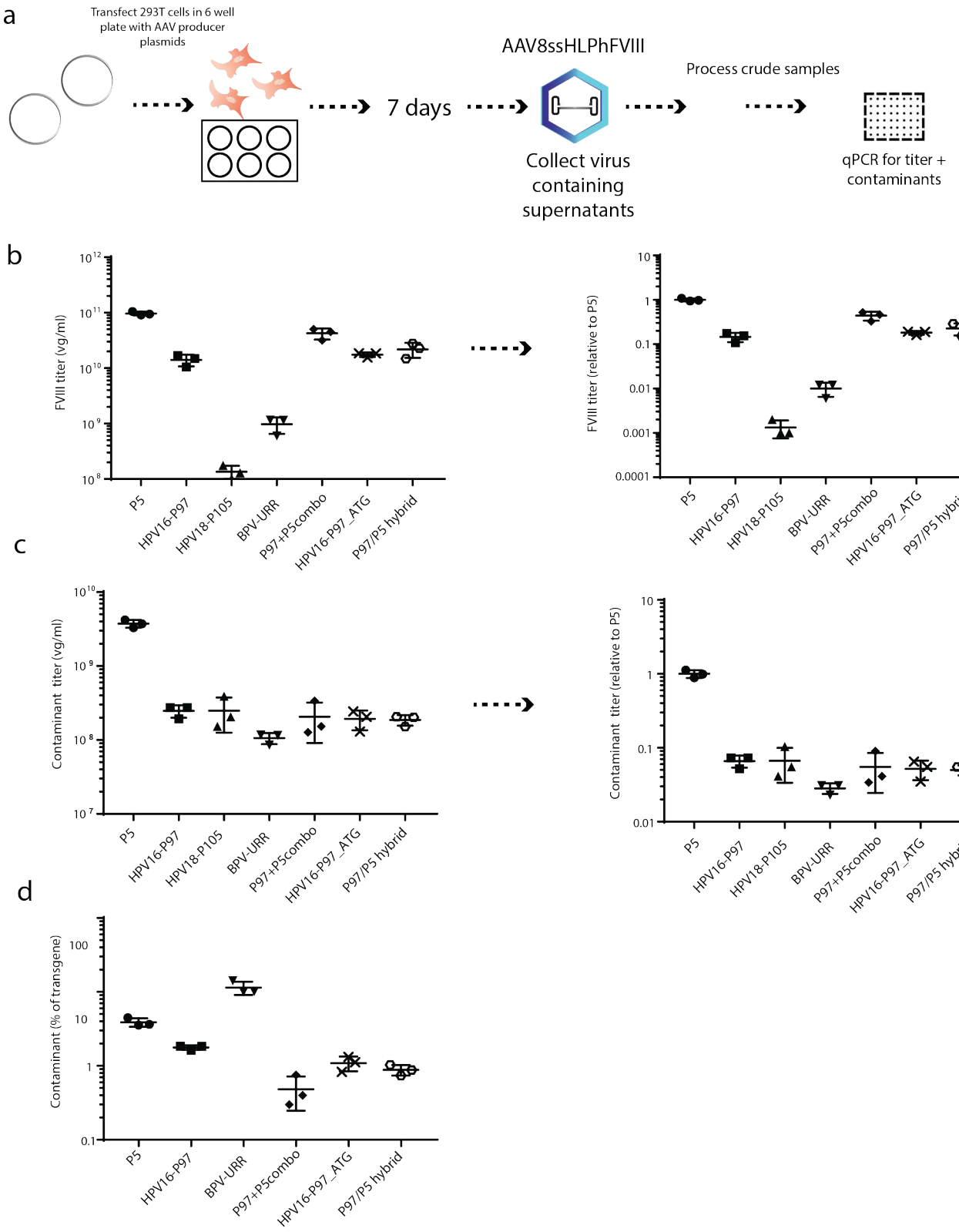


Figure 4.9 - Direct from harvest supernatant assay confirms P5 replacements lower upstream contamination.

(a) Schematic; AAV8ssHLPhFVIII was produced using AAV2_8 REPCAP plasmids with Papilloma based REP78/68 promoter designs and assayed for titer and contaminant abundance **(b)** qPCR results of AAV8ssHLPhFVIII titer production efficiency with different REP78/68 promoter designs expressed as (left) HLPhFVIII titer (vg/ml) and (right) HLPhFVIII titer normalised against P5 promoter production efficiency **(c+d)** qPCR results of AAV8ssHLPh-FVIII contaminant levels for an upstream contaminant amplicon expressed as **(c)** (left) contaminant titer (vg/ml); (right) contaminant amplicon abundance normalised against P5 promoter contamination **(d)** contaminant abundance expressed as a percentage of HLPhFVIII titer. *Graphs b-d analysed by One Way Anova with Dunnet's multiple comparison test.*

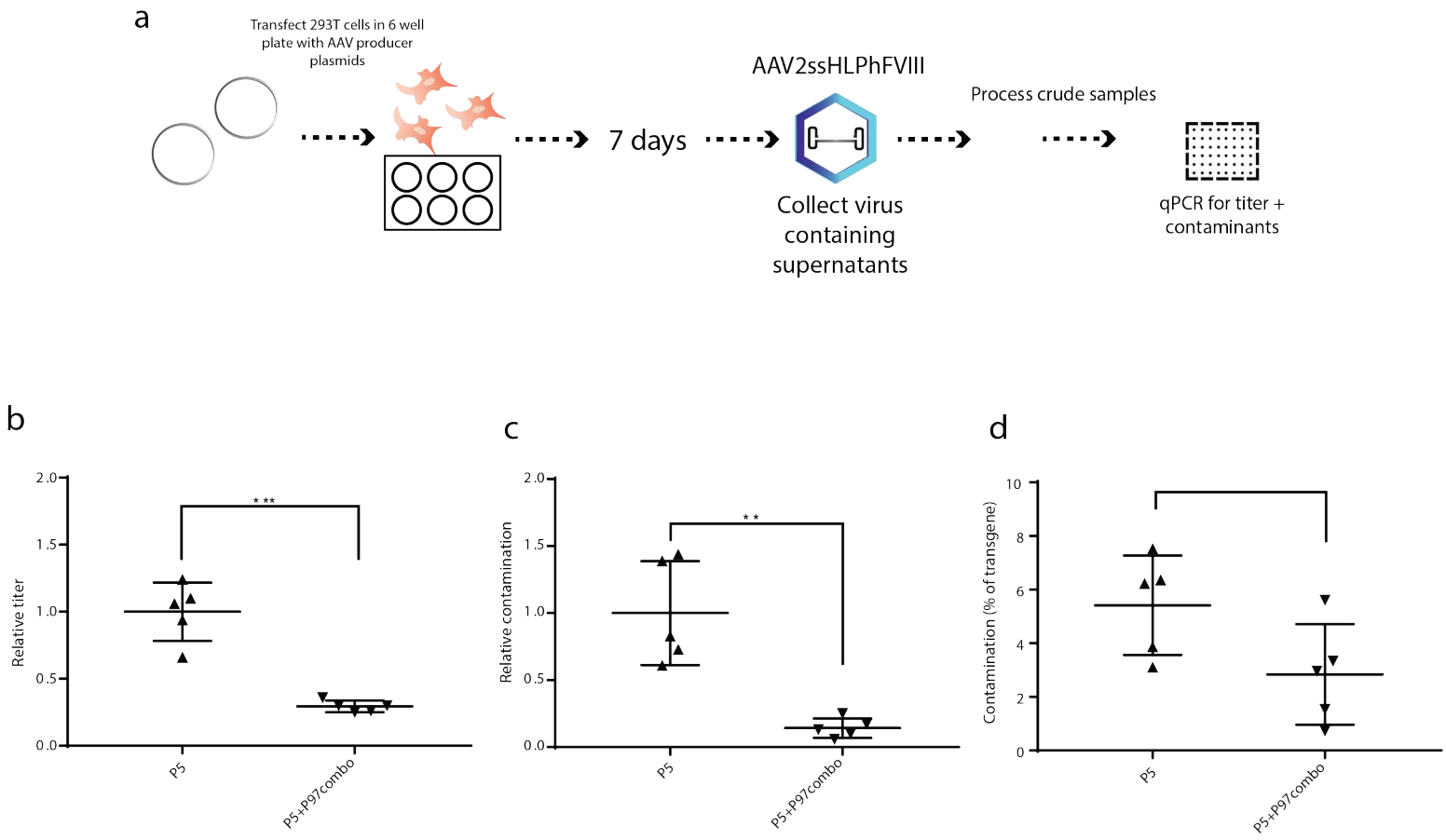


Figure 4.10 - P5 replacement contamination reduction of AAV is not serotype specific.

(a) Schematic; AAV2HLPhFVIII was produced with either P5 or P5+P97combination REP78/68 promoter. **(b)** qPCR results of AAV2ssHLPhFVIII titer production efficiency of P5+P97 combination promoter normalised against P5 titer. **(c)** qPCR results of AAV2ssHLPhFVIII contaminant levels for an upstream contaminant amplicon ormalised against contaminant amplicon abundance produced with P5 promoter. **(d)** contaminant abundance expressed as a percentage of HLPhFVIII titer. *Graphs b-d analysed by unpaired T test with Welch's correction.*

AAV produced with the P5+P97 combo promoter, which if contaminant incorporation remained relatively steady would result in a higher contribution of contaminants as a percentage of the expression cassette titer. Overall, the preps yielded a comparatively similar total level of contamination; 25.42% for P5 and 26.76% for P5+P97 combo (Fig. 4.11.d). To verify that this contamination outside of the ITR sequences was indeed raised in the P5 replacement setting, an amplicon 100bp outside of the ITR was assayed by qPCR (Fig. 4.12). Results showed in AAV8ssHLPhFVIII that whilst the raw values of contaminant sequences were lower in P5+P97combo (5.7e8 ± 1.52e8) than in P5 (1.26e9 ±7.97e7) (P=0.03), the P5+P97combo produced AAV had significantly higher contamination outside of the ITR when this value was expressed as a percentage of titer (P5 – 3.12% ± 0.13; P5+P97combo – 7.20% ± 0.17) (Fig. 4.12.b). This process was repeated in AAV2 and there was no significant difference in the raw values between P5 (1.01e8 ±2.37e7) and P5+P97combo (1.07e8 ± 2.26e7) (P=0.76) (Fig. 4.12.c). In concordance with the AAV8 data, the P5+P97combo had a significantly higher contamination percentage than the original P5 configuration (P5 – 3.12% ± 0.13; P5+P97combo – 7.20% ± 0.17) (P=0.006). These results show that whilst DNA contaminants upstream of the P5 promoter have been reduced with a P5 replacement strategy, the increase in representation of sequences from outside of the ITRs has negated this beneficial effect.. When analysing for vector purity it is important to consider the entirety of the prep, and with this current stock of promoters the overall purity is not improved as compared to when the original promoter, P5, is used to drive the large form REP78 and REP68 genes. It is unlikely that this strategy could be implemented into commercial or clinical production systems in its current form without improvements in

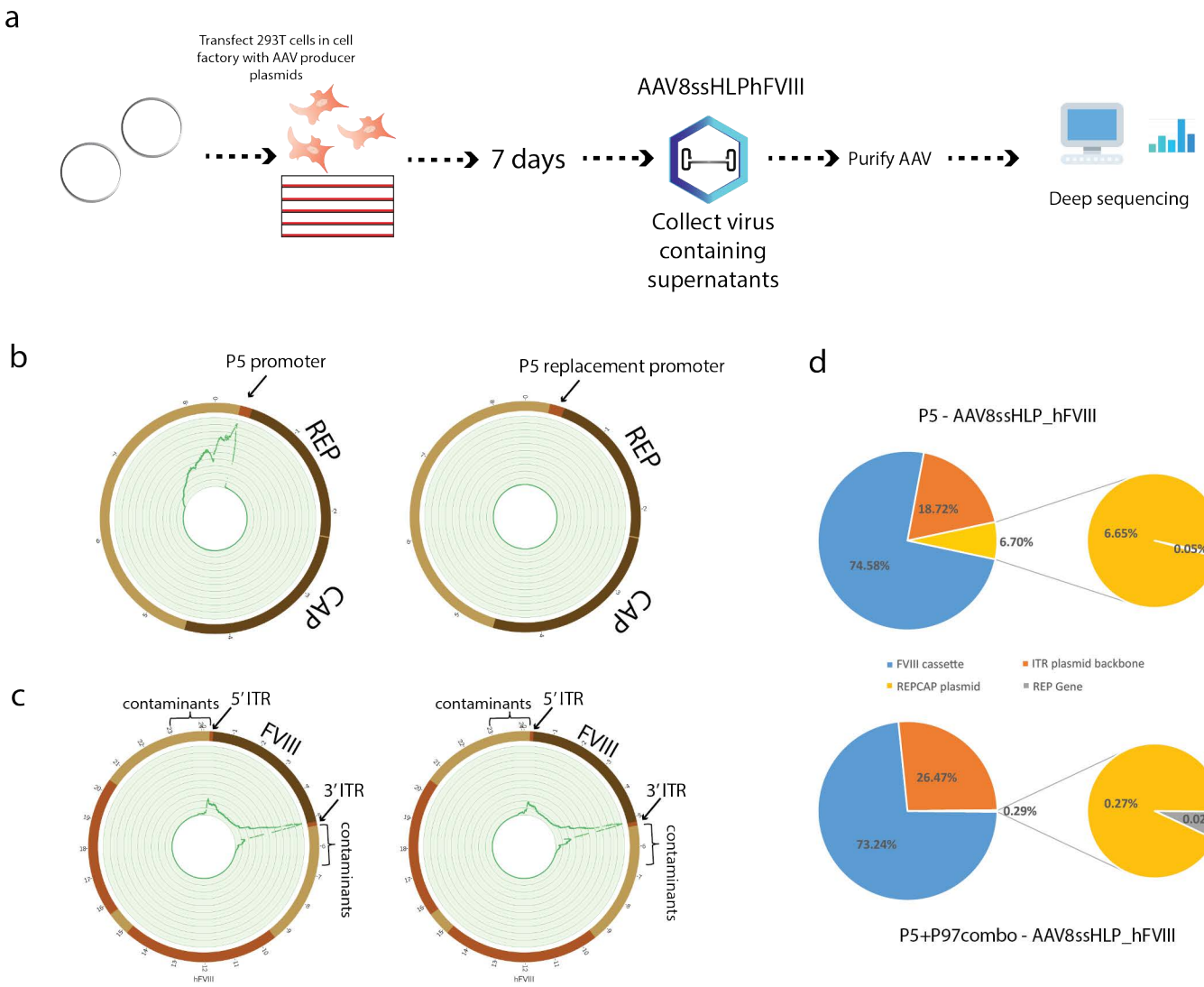


Figure 4.11 - P5 replacement vector does not result in improved AAV purity due to increased vector genome plasmid contamination

(a) Schematic; AAV8 HLPhFVIII was produced with either P5 or P5+P97combination REP78/68 promoter, purified and submitted for next generation sequencing *(NexteraXT, 10 million, paired end)* **(b)** Circos plot of AAV2_8 REPCAP plasmid from AAV8ssHLPhFVIII produced with (left) P5 and (right) P5+P97 combination promoter **(c)** Circos plot of ssHLPhFVIII vector genome plasmid from AAV8ssHLPhFVIII produced with (left) P5 and (right) P5+P97 combination promoter. **(d)** Pie chart of next generation sequencing reads mapped back to producer plasmids from AAV8ssHLPhFVIII produced with (top) P5 and (bottom) P5+P97 combination promoter.

either titer or total vector purity. Interestingly, whilst contamination outside of the ITRs when expressed as a percentage of the transgene was higher for the P5 replacement, the raw numbers were not, either yielding lower values as in the AAV8 samples or equal values as in the AAV2. This would suggest that if a means to improve the AAV titer to that of the P5 promoter could be achieved with a strategy that retains the contamination reduction of the P5 replacement promoter configurations, that the overall purity of the vector would be improved, and would thus be a more suitable product for clinical translation. In chapter 3 we identified the P5 promoter as the source of contamination incorporation via an experiment that moved this promoter upstream and observed a matching shift in the incorporating contaminant sequences. It is possible that through understanding the components of the P5 promoter that are yielding this incorporation phenotype, the principle of eliminating contamination initiation from the REPCAP plasmid could be retained without sacrificing AAV titer or total purity of the AAV vector. The previously tested variant designated P5/P97 hybrid utilised the REP binding site of the P97 promoter (nt14-56)[495] substituted into P5 in place of the P5 REP binding site yielding a result that showed a reduction in upstream contamination but a titer that did not exceed 50% of P5 (Figures 4.6 + 4.8). This suggested that the P5 REP binding site was a necessary element in the initiation of the P5 contamination phenotype. There are many GCTC containing sequences with high homology to the P5RBS in the human genome.[182] If mere REP binding was enough to cause DNA incorporation it is possible that certain genome sequences from the producer cell line genome would be overrepresented in the final AAV product. To ascertain if this binding element would be sufficient to initiate

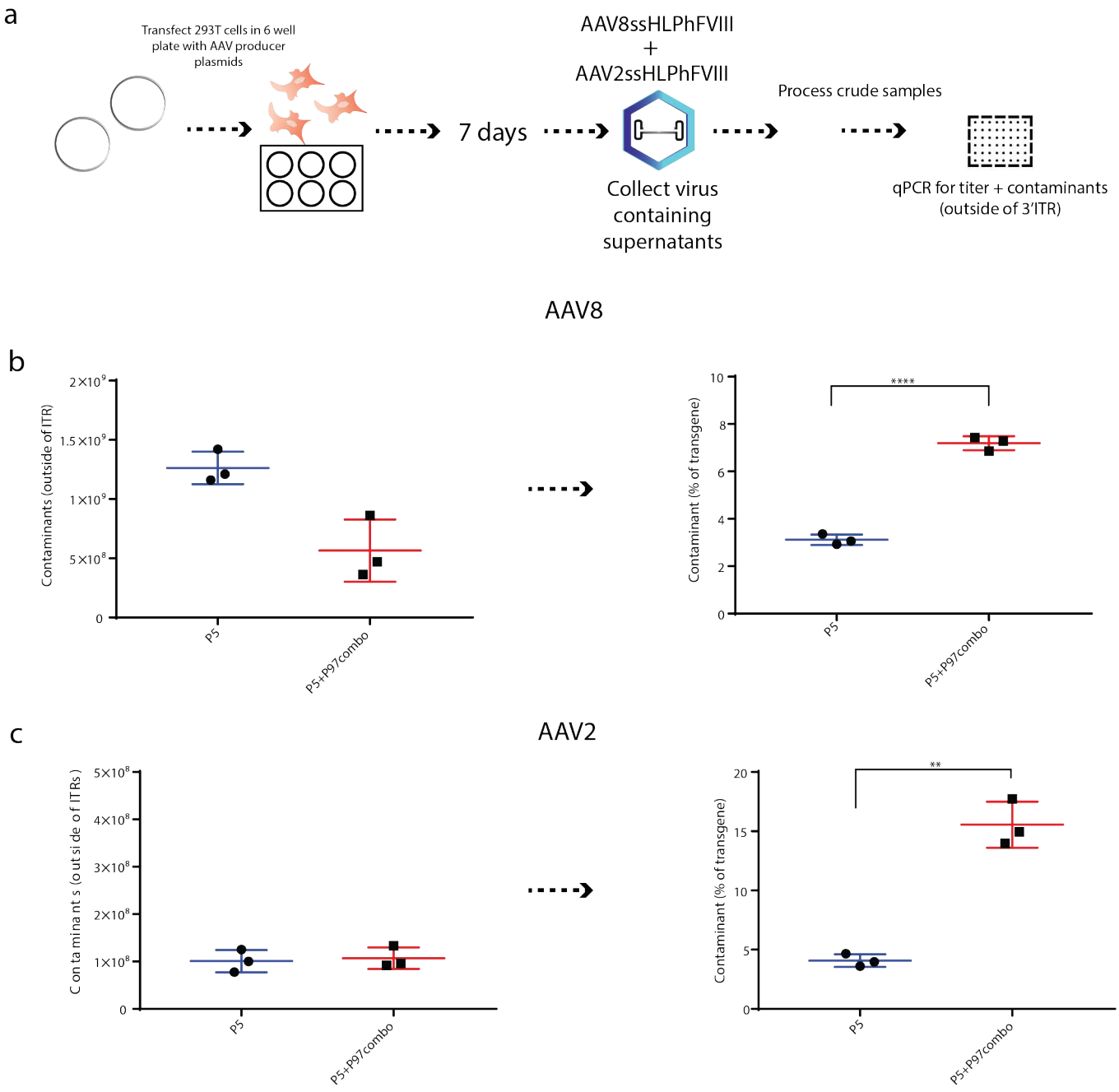


Figure 4.12 - P5 replacement configuration results in increased contamination outside of ITRs in both AAV8 and AAV2

(a) Schematic; AAV8HLPhFVIII and AAV2HLPhFVIII was produced with either P5 or P5+P97combination REP78/68 promoter. **(b)** qPCR results of AAV8ssHLPhFVIII (left) contaminant abundance of AAV produced with P5 and P5+P97 combination promoter and (right) contaminant abundance expressed as a percentage of HLPhFVIII titer. **(c)** qPCR results of AAV2ssHLPhFVIII (left) contaminant abundance of AAV produced with P5 and P5+P97 combination promoter and (right) contaminant abundance expressed as a percentage of HLPhFVIII titer. *Graphs b+c analysed by unpaired T test with Welch's correction.*

contamination in the recombinant AAV system, the reverse experiment was performed, in which the REP binding site from the P5 promoter was substituted into the HPV16-P97 promoter ($P97_{P5RBS}$) (Fig. 4.13). AAV8ssHLPhFVIII was produced with this configuration and compared to P5. $P97_{P5RBS}$ produced AAV effectively, although again significantly lower than the normalised P5 titer (0.38 ± 0.01) (P=0.0005). (Fig. 4.13.b). Surprisingly, when contaminant amplicons upstream of this promoter were assayed by qPCR and were significantly reduced in the $P5_{P97RBS}$ configuration relative to P5 (0.037 ± 0.003) (P=0.03), resulting in a lower contamination percentage (P5 – 2.02% ± 0.26; $P97_{P5RBS}$ 0.21% ± 0.02), despite the presence of the full P5 REP binding site (Fig. 4.13.c). When assayed for contaminants outside of the ITR however, the same pattern that emerged in the P5+P97combo setup emerged, with the raw values of contamination relative to P5 not differing in $P97_{P5RBS}$ (1.02 ± 0.04), however when expressed as a function of titer yielding a significantly more contaminated product (P5 – 3.12% ± 0.13; $P97_{P5RBS}$ 8.40% ± 0.09) (P<0.0001) (Fig.4.13.d). This shows that the P5RBS is necessary but not sufficient for REPCAP DNA contaminant incorporation into AAV to occur upstream of P5.

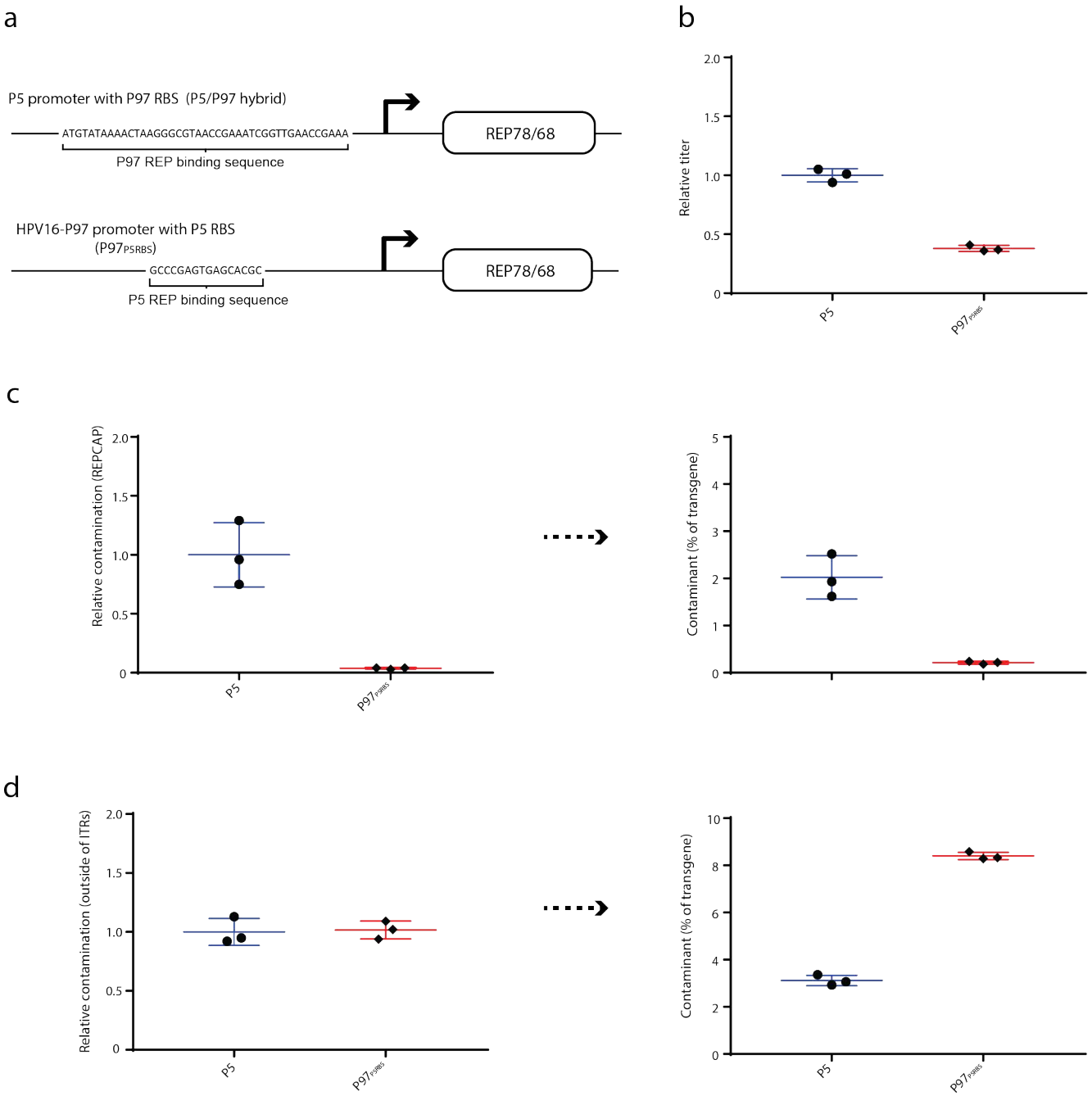


Figure 4.13 - The P5 REP binding site is not sufficient to initiate upstream contamination

(a) Schematic; AAV8HLPhFVIII awas produced with either P5 or HPV16-P97 containing the REP binding site of P5 (P97P5RBS). **(b)** qPCR results of AAV8ssHLPhFVIII titer normalised to production levels with the P5 promoter. **(c)** qPCR results of AAV8ssHLPhFVIII upstream contaminant abundance of AAV produced with P5 and P97P5RBS (left) normalised to P5 contamination and (right) contaminant abundance expressed as a percentage of HLPhFVIII titer. **(d)** qPCR results of AAV8ssHLPhFVIII outside of ITR contaminant abundance of AAV produced with P5 and P97P5RBS (left) normalised to P5 contamination and (right) contaminant abundance expressed as a percentage of HLPhFVIII titer. *Graphs b+d analysed by unpaired T test with Welch's correction.*

4.2.3. Development of P5-HS: A REP78/68 promoter modification to remove contamination and retain AAV production titer:

The function of REP in the AAV replication process is more than DNA binding. The REP proteins also have helicase activity and the large forms of REP exhibit site specific endonuclease activity.[172,197,200,500] In the ITRs during AAV replication, REP nicks the DNA at a site known as the TRS site (Fig. 4.14.a), facilitating AAV genome excision from surrounding DNA. Within the P5 promoter there is an AT dinucleotide that serves as a TRS mimic at the crest of a loop that forms adjacent to the P5 REP binding site between nucleotides 287 and 288 of the AAV2 genome (Fig. 4.14.b).[164] We hypothesised that removing this dinucleotide from the P5 promoter would also abrogate the upstream contamination derived from P5. A promoter that retained the REP binding site but was deleted for the AT dinucleotide ($P5_{\Delta AT}$) was designed (Fig. 4.14.c). AAV8ssHLPhFVIII was produced and examined for contaminant incorporation. $P5_{\Delta AT}$ produced virus efficiently, although again with a lower viral titer relative to that of intact P5 (0.38± 0.04. P=0.01) (Fig. 4.14.d). Contamination upstream of $P5_{\Delta AT}$ were significantly reduced in both relative terms ($P5_{\Delta AT}$ – 0.14 ± 0.01. P=0.008) and as a percentage of transgene (P5 2.027% ± 0.03; $P5_{\Delta AT}$ - 0.75% ± 0.02. P<0.0001) (Fig. 4.14.e+f). Amplicons found outside of the ITRs were then compared. As expected, the relative values from a given production run did not exceed that of P5 for either amplicon (Amplicon 1 0.66 ± 0.04 P=0.003; Amplicon 2 0.73 ± 0.01 P=0.08) (Fig. 4.14.g). However, when calculated as a percentage of the expression cassette titer the contamination outside of the ITRs was significantly increased (Amplicon 1: P5 - 2.4% ± 0.1; $P5_{\Delta AT}$ - 3.99% ± 0.05. P=0.0001) (Amplicon 2: P5 – 1.90% ± 0.1; $P5_{\Delta AT}$ - 3.66% ± 0.21. P=0.001) (Fig. 4.14.h). This again suggests that

the proportion of contaminants from outside the ITRs are related to the expression cassette packaging efficiency of the system.

To ascertain whether the loop portion of the TRS mimic site was essential for incorporation, a sequence was designed that kept the TRS dinucleotide intact, and the whole YY1box sequence, but removed its ability to form a loop with its surrounding sequence. This was positioned along with the P5 RBS in the context of the HPV-16 P97 promoter ($P97_{P5RBS\Delta loop}$) (Fig. 4.15.a). AAV was produced successfully, again lower relative to P5 driven production (0.59 ± 0.03) (Fig. 4.15.b), but relative DNA incorporation upstream of the promoter remained significantly lower than P5 incorporation levels with $P97_{P5RBS\Delta loop}$ when assayed by qPCR (0.14 ± 0.02. P=0.006) (Fig. 4.15.c). As in previous cases, the percentage contamination upstream of the promoter was also found to be significantly reduced (P5 - 2.03% ± 0.03; $P97_{P5RBS\Delta loop}$ - 0.50% ±0.07. P=0.0007) (Fig. 4.15.d). This suggests contaminant incorporation requires the same components for contaminant incorporation as it does for an AAV replication event - both an intact REP binding site and adjacent loop to position an available AT nicking site, and a GCTC containing binding site is not sufficient to mediate aberrant active contaminant packaging. $P97_{P5RBS\Delta loop}$ also showed that relative contamination outside of the ITR did not differ significantly (Amplicon 1 – 0.81 ± 0.02. P=0.12; Amplicon 2 – 0.91 ± 0.02. P=0.54) (Fig. 4.15.e). As expected, when titer was taken into consideration the $P97_{P5RBS\Delta loop}$ resulted in a higher percent contamination for both amplicons (Amplicon 1: P5 – 2.3% ± 0.10; $P97_{P5RBS\Delta loop}$ – 3.26% ± 0.19. P=0.002) (Amplicon 2: P5 – 1.90% ± 0.07; $P97_{P5RBS\Delta loop}$ –

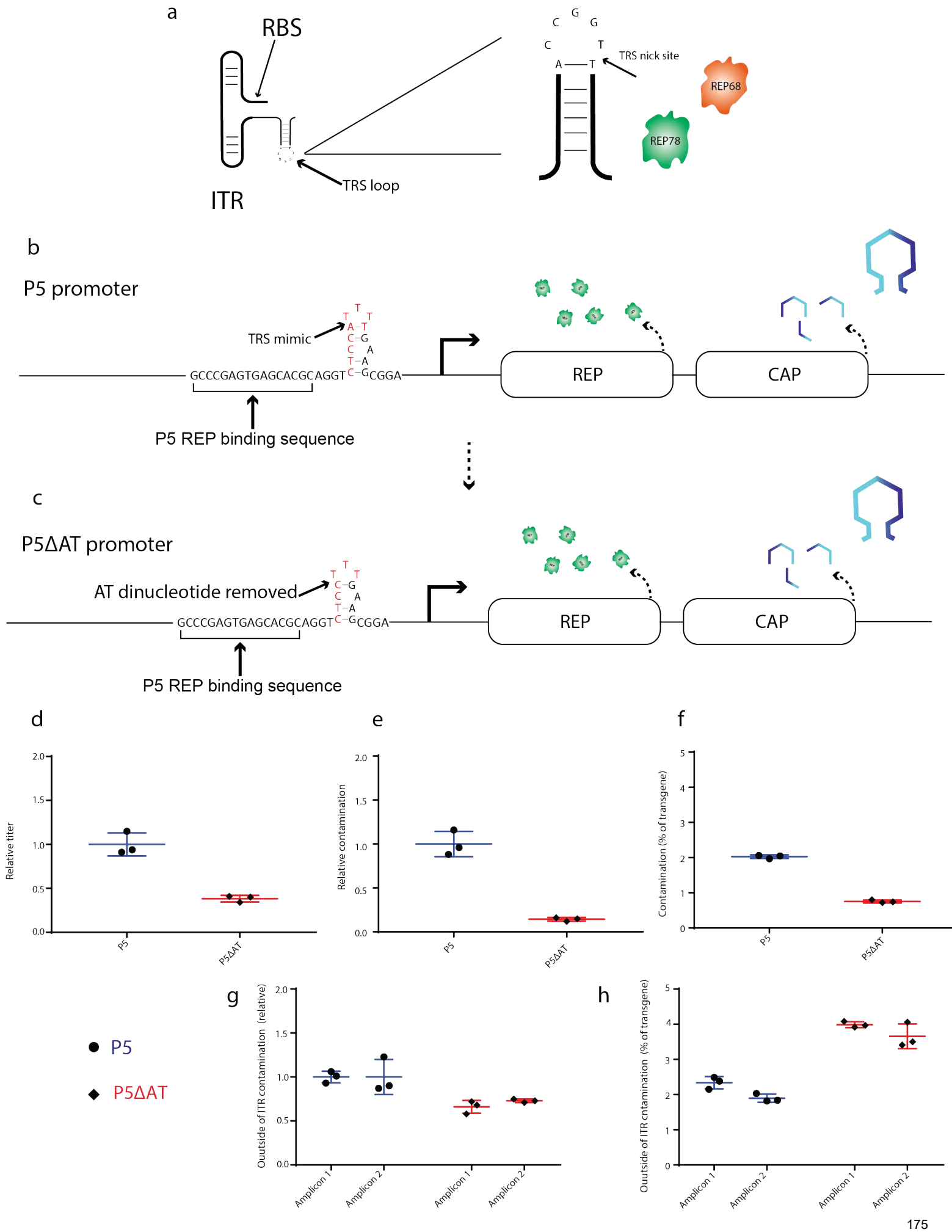

a
RBS
ITR
TRS loop
TRS nick site
REP68
REP78
b
P5 promoter
TRS mimic
GCCCGAGTGAGCACGCAGGTCGGA
P5 REP binding sequence
REP
CAP
c
P5ΔAT promoter
AT dinucleotide removed
GCCCGAGTGAGCACGCAGGTCGGA
P5 REP binding sequence
REP
CAP
d
Relative titer
P5
P5ΔAT
e
Relative contamination
P5
P5ΔAT
f
Contamination (% of transgene)
P5
P5ΔAT
P5
P5ΔAT
g
Ouutside of ITR contamination (relative)
Amplicon 1
Amplicon 2
Amplicon 1
Amplicon 2
h
Ouutside of ITR cntamination (% of transgene)
Amplicon 1
Amplicon 2
Amplicon 1
Amplicon 2

Figure 4.14 - Deletion of AT dinucleotide in P5 promoter removes upstream contamination but lowers titer

(a) Schematic; Positioning of the TRS loop witihin the AAVITR showing the canonical REP78/68 thymidine nicking site **(b)** Schematic; P5 promoter showing the positioning of the TRS mimic loop within the *(red)* YY1+1 box. **(c)** Schematic; P5 promoter design (P5ΔAT) showing the deletion of an AT dinucleotide at the nick site of the TRS mimic loop of P5. **(d)** qPCR results of AAV8ssHLPhFVIII titer produced with P5 and P5ΔAT normalised to the production of P5 **(e+f)** qPCR results of AAV8ssHLPhFVIII contaminant abundance upstream of the P5 and P5ΔAT promoter **(e)** normalised to contamination of P5 and **(f)** expressed as a percentage of HLPhFVIII titer. **(g+h)** qPCR results of AAV8ssHLPhFVIII contaminant abundance of two amplicons outside of the ITRs **(g)** normalised to outside ITR contamination of P5 and **(h)** expressed as a percentage of HLPhFVIII titer *Graphs d-f analysed by unpaired T test with Welch's correction. Graphs g+h analysed by one way Anova with Tukey's multiple comparison test.*

2.99% ± 0.14. P=0.0007) (Fig. 4.15.f). This indicates that this vector configuration would not result in an overall purer AAV product although further analysis of this promoter design was not undertaken to confirm this.

In order to achieve high titer AAV production whilst removing the contaminant phenotype, it was hypothesised that all characterised elements of the P5 promoter could be retained. This included the TATA box, REP binding site, YY1 boxes and the TRS dinucleotide with loop (Fig. 4.16.a). It was also hypothesised that these elements require tight positional requirements to achieve replication and thus contamination of the AAV prep. Considering this, a REP78/68 promoter design was constructed and cloned that contained the P5 promoter in its entirety, with a series of spacers of differing length inserted between the R5RBS and the YY1 box containing TRS loop (P5-HS) (Fig. 4.16.b). The rationale behind this placement of the spacer was to retain the integrity of the TATA box in P5, but create a physical separation between the two parts of the promoter that would interact with the REP protein; the P5 RBS and the TRS mimic sequence in YY1 box. Ideally this modification would prevent the incorporation of DNA without disrupting any of the natural promoter function of P5. Spacer sequences of 5bp and 100bp were chosen. A 5bp change would result in an approximate half helical turn in the DNA,[501] with the hypothesis both the lateral and rotational distance from the TRS mimic site in the YY1 box would be enough to prevent nicking and thus active incorporation. The 100bp spacer sequence was generated to provide an increased physical separation of the REP binding site from that of the YY1 box containing TRS loop mimic. The spacer sequences were generated with a random sequence generator and verified to not include GCTC to ensure no inadvertent dysregulation of REP. Spacer sequences were also verified to not contain an

AT dinucleotide in proximity to the REP binding site that could serve as a nicking site by the REP endonuclease domain. AAV8 ssHLPhFVIII was produced and titers were measured. P5-HS AAV titers did not differ significantly from the intact P5 production system (P5-HS5 – 0.87 ± 0.03. P=0.40; P5-HS100 0.96 ± 0.01. P=0.85) (Fig. 4.16.c). Furthermore, when assayed by qPCR for the presence of upstream contaminant amplicons, the presence of contaminant DNA was significantly reduced compared to P5 in both relative amounts (P5-HS5 – 0.28 ± 0.02. P=0.0004; P5-HS100 0.35 ± 0.07. P=0.0007), and percent contamination (P5 – 2.36% ± 0.22; P5-HS5 – 0.40% ± 0.02. P<0.0001; P5-HS100 0.60% ± 0.06. P=0.0002) (Fig. 4.16.d+e). With previous constructs, the proportion of contaminants from outside of the ITRs was greater than that of P5. With the P5HS vectors however, the proportion of these contaminants were similar to P5 when analysed as relative amounts for all but one amplicon: P5-HS5 amplicon 1 was significantly lower in relative terms (0.64 ± 0.04. P=0.01) (Fig. 4.16.f). When analysed as a percentage of the transgene, neither of the P5-HS produced AAV differed significantly from P5 with outside of ITR contaminants (Fig. 4.16.g). This means that in contrast to the previous promoter iterations that were examined, utilising the P5-HS system has reduced the overall contamination of AAV as tested by qPCR, whilst simultaneously retaining the production yield.

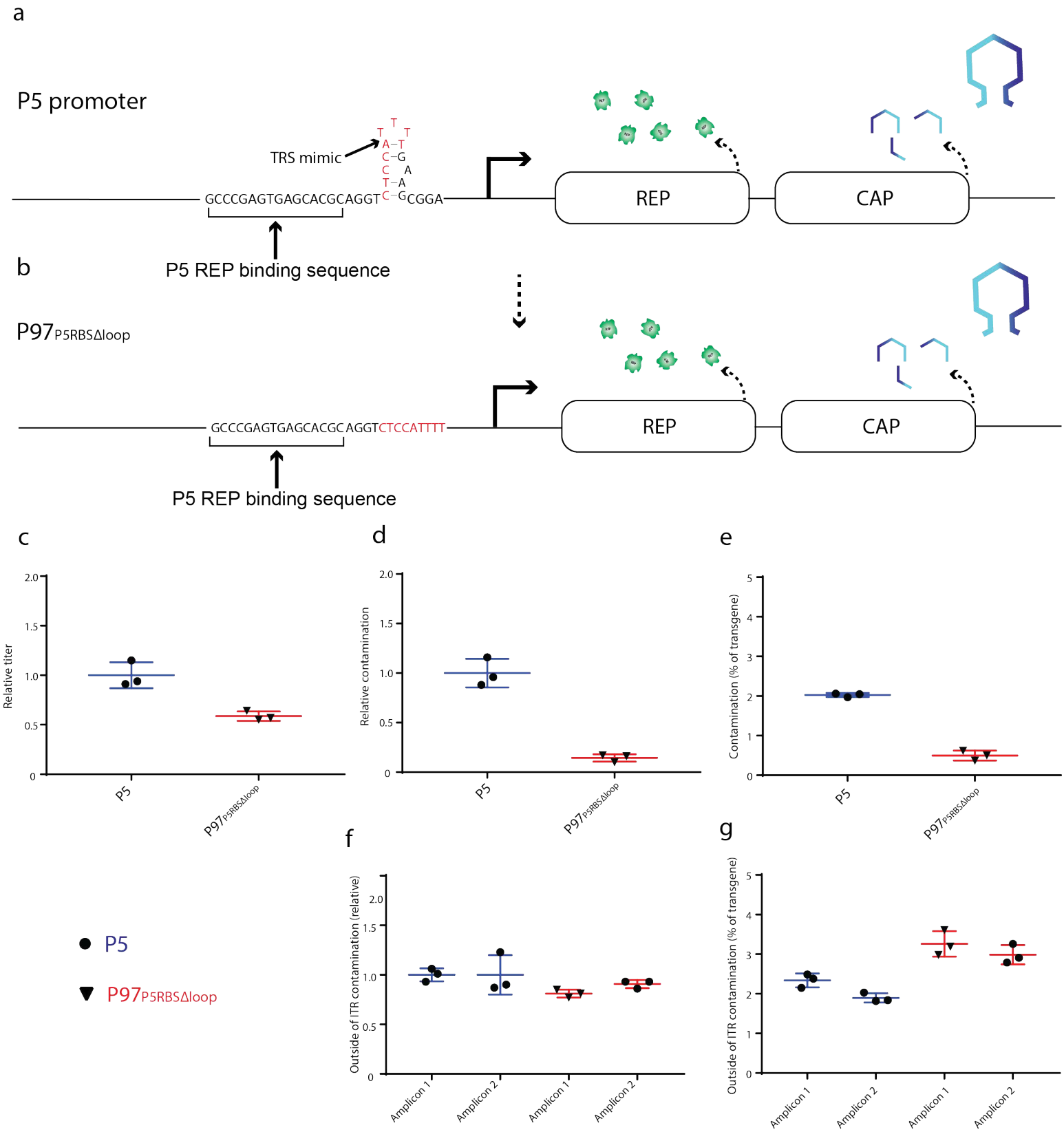

a
P5 promoter
TRS mimic
GCCCGAGTGAGCACGCAGGTCGGA
P5 REP binding sequence
REP
CAP
b
P97P5RBSΔloop
GCCCGAGTGAGCACGCAGGTCTCCATTTT
P5 REP binding sequence
REP
CAP
c
Relative titer
P5
P97P5RBSΔloop
d
Relative contamination
P5
P97P5RBSΔloop
e
Contamination (% of transgene)
P5
P97P5RBSΔloop
P5
P97P5RBSΔloop
f
Outside of ITR contamination (relative)
Amplicon 1
Amplicon 2
Amplicon 1
Amplicon 2
g
Outside of ITR contamination (% of transgene)
Amplicon 1
Amplicon 2
Amplicon 1
Amplicon 2

Figure 4.15 - Disruption of P5 TRS mimic loop formation reduces upstream contamination but reduces titer

(a) Schematic; P5 promoter showing the positioning of the TRS mimic loop within the *(red)* YY1+1 box. **(b)** Schematic; P5 promoter design (P97$_{P5RBS\Delta loop}$) showing the retention of the entire *(red)* YY1+1 sequence but removal of subsequent loop facilitating nucleotides **(c)** qPCR results of AAV8ssHLPhFVIII titer produced with P5 and P97$_{P5RBS\Delta loop}$ normalised to the production of P5 **(d+e)** qPCR results of AAV8ssHLPhFVIII contaminant abundance upstream of the P5 and P97$_{P5RBS\Delta loop}$ promoter **(d)** normalised to contamination of P5 and **(e)** expressed as a percentage of HLPhFVIII titer. **(f+g)** qPCR results of AAV8ssHLPhFVIII contaminant abundance of two amplicons outside of the ITRs **(f)** normalised to outside ITR contamination of P5 and **(g)** expressed as a percentage of HLPhFVIII titer *Graphs c-e analysed by unpaired T test with Welch's correction. Graphs f+g analysed by one way Anova with Tukey's multiple comparison test.*

To assess whether the contamination in AAV originating from outside of the ITRs was indeed related to the production efficiency of REP78/68 promoter, a regression analysis was performed comparing the titers observed relative to P5 and the relative contamination observed outside of the ITRs. Relative contamination from outside of the ITRs was negatively correlated with production titer (Amplicon 1 – R^2 = 0.75. P<0.0001; Amplicon 2 – R^2 = 0.89. P<0.0001) (Fig. 4.17.a+b). This result shows that the utilisation of a REP78/68 promoter that yields efficient AAV replication is essential for maximising the purity of the vector product. Higher vector purity should always be a desired goal of clinical viral production for gene therapy, as it is likely to improve vector safety and tolerance. However, there is also the possibility that vector purity could have an influence on the functioning of the prep itself. The numerous promoter variations created in order to produce a functioning construct give an opportunity to study this in a well-controlled manner. Large scale preps (1X cell factory) of AAV8ssHLPhFVIII were produced with either the P5 promoter, P5HS promoter, or a P5+P97 combo promoter to produce virus with quantifiably different levels of DNA contamination (Fig. 4.18). Virus was titered and contaminant amplicons were measured by qPCR to ensure the expected differences were present. (Fig. 4.18.b). Mice were injected with 2e10vg AAV8ssHLPhFVIII. Circulating FVIII levels were measured at 1-week post infection (n=5 per group) (Fig 4.18.c).

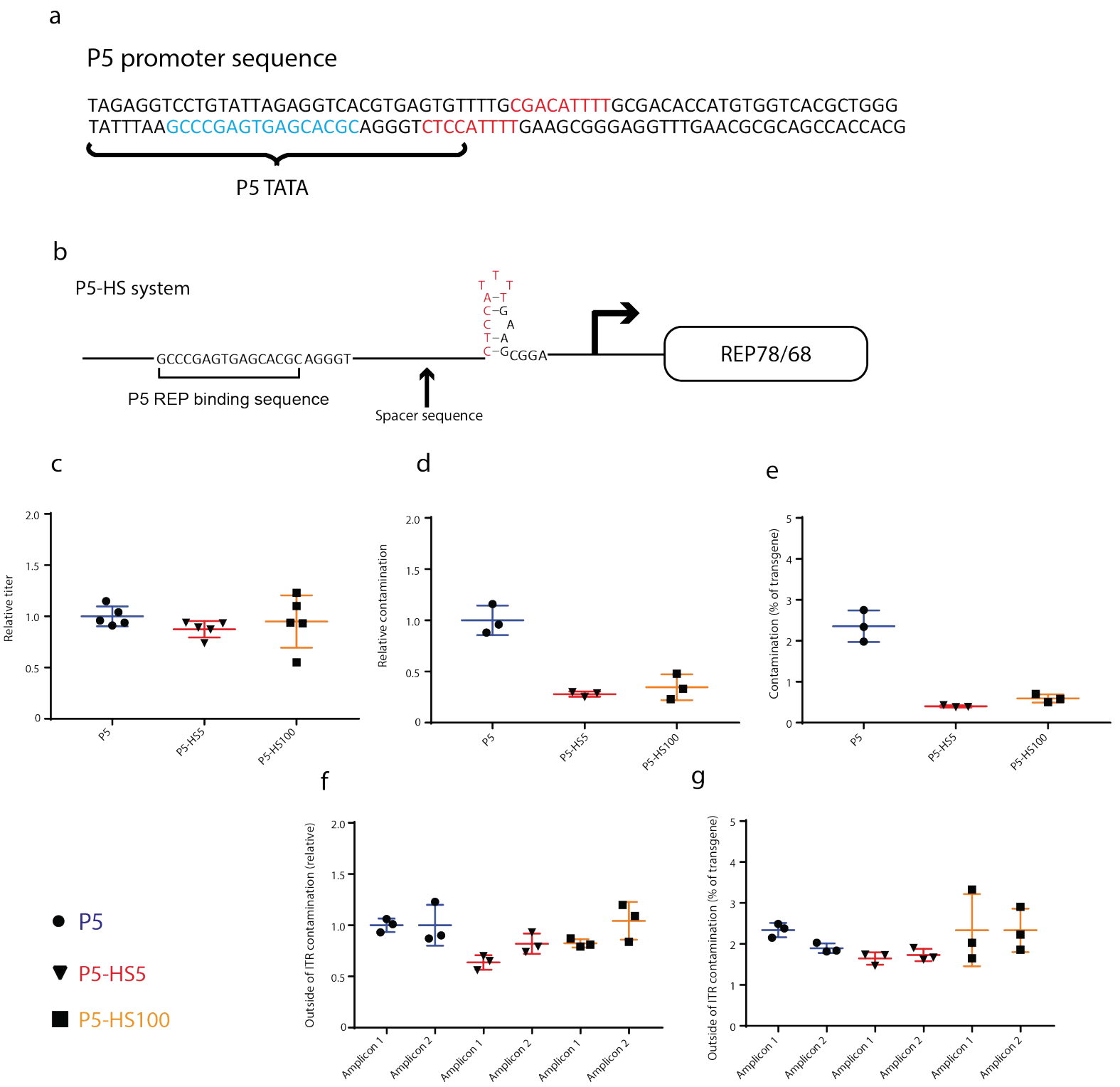


Figure 4.16 - Spacer insertion between P5 RBS and YY1+1 box reduces upstream contamination and retains titer.

(a) Schematic; P5 promoter showing the positioning of the P5 TATA box including the P5 REP binding site *(blue)* and the *(red)* YY1 box sites (YY1-60 and YY1+1). **(b)** Schematic; P5 promoter design (P5-HS) showing the insertion of a spacer between the TATA and YY1+1 elements of the P5 promoter **(c)** qPCR results of AAV8ssHLPhFVIII titer produced with P5 and P5-HS designs of either 5bp (P5-HS5) or 100bp (P5-HS100) normalised to the production of P5 **(d+e)** qPCR results of AAV8ssHLPhFVIII contaminant abundance upstream of the P5, P5-HS5 and P5-HS100 promoters **(d)** normalised to contamination of P5 and **(e)** expressed as a percentage of HLPhFVIII titer. **(f+g)** qPCR results of AAV8ssHLPhFVIII contaminant abundance of two amplicons outside of the ITRs **(f)** normalised to outside ITR contamination of P5 and **(g)** expressed as a percentage of HLPhFVIII titer *Graphs analysed by one way Anova with Dunnet's multiple comparison test.*

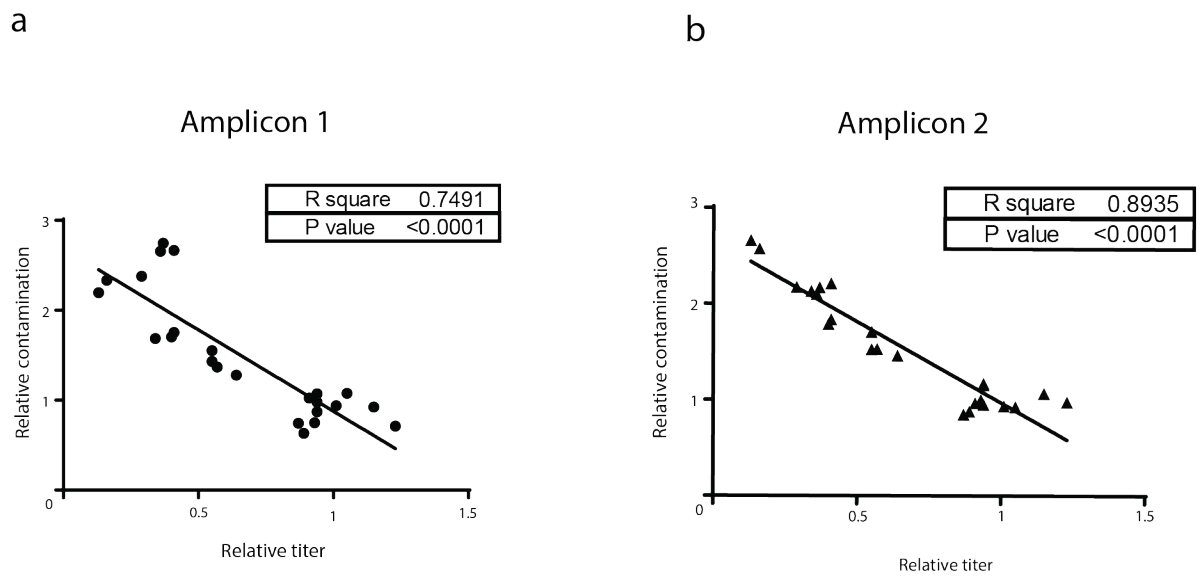


Figure 4.17 - Contamination levels outside of the ITR are correlated with titer when REP78/68 promoter is altered variable.

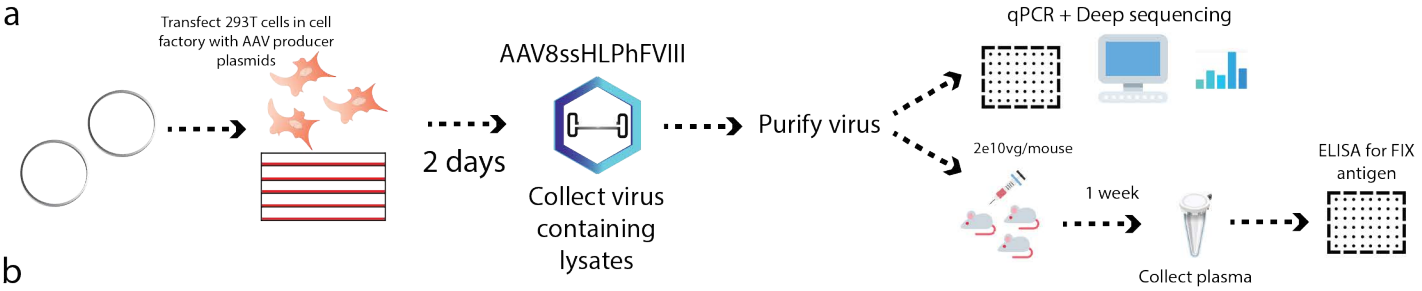
Linear regression of DNA contaminant abundance outside of the ITRs when REP78/68 promoter is altered. *(Y axis)* Relative contamination normalised to P5 contamination levels. *(x-axis)* Expression cassette titer normalised to P5 titer levels. **(a)** Outside of ITR contaminant amplicon 1 **(b)** Outside of ITR contaminant amplicon 2.

Interestingly, the P5-HS infected mice showed the highest level of circulating FVIII protein (286.7ng/ml ± 36.8), significantly higher than the AAV8hFVIII produced with the P5 promoter (111.5ng/ml ± 13.3. P=0.0003). There was no significant difference between circulating FVIII levels of virus produced with the P5 promoter and the P5+P97combo (104.0ng/ml ± 6.0. P=0.96) (Fig. 4.18.c). It is possible that the improvement in purity associated with the P5-HS promoter configuration allows the infected AAV to proceed through the steps to productive infection with less interference from the non-functional contaminating DNA. However, for any conclusion to be made on whether the P5-HS system does indeed result in the production of more efficacious product, this result would have to be replicated in further contexts with additional preparations of virus.

To test whether the P5-HS system could be applied to serotypes aside from AAV8, the 5bp spacer configuration (P5-HS5) was cloned into AAV1 and AAV2 REPCAP plasmids. AAV ssHLPhFVIIIv3 was then produced in either with either P5-HS5 or P5 driving REP78/68 expression (Fig. 4.19.a). Titers did not significantly differ between P5 and P5-HS across the tested serotypes (AAV1 P=0.99; AAV2 P=0.75) (Fig. 4.19.b). Given the differences in the plasmid backbones between REPCAP plasmids, a universal set of primers within the P5 promoter were used to assay contamination levels by qPCR. Analysis of the supernatants revealed that all serotypes showed a decrease in relative contamination levels of the tested amplicon when the P5-HS5 configuration was used for production (AAV1 P5-HS5 – 0.22 ± 0.03. P= 0001) (AAV2 P5-HS – 0.30 ± 0.11. P=0002), which when converted to a percentage of the transgene titer resulted in significantly lower contamination (AAV1_P5 – 15.10% ± 0.19; AAV1_P5-HS5 3.34% ± 0.54. P<0.0001) (AAV2_P5 – 16.37% ± 1.20; AAV2_P5-HS5 – 4.50% ± 1.04. P<0.0001) (Fig. 4.19.c +d).

Finally, to ensure that the reductions in contaminant presence detected by qPCR amplicons translated into a product with overall improved purity, FVIII and FIX preps produced side by side in 293T cells at the cell factory scale with either the original P5 promoter or the P5-HS5 were next generation sequenced, alongside GMP grade purified preps as a batch effect controls (Fig 4.20). In both FVIII and FIX contexts overall purity of the resultant AAV was improved by around 2%, (AAV8ssHLPhFVIII: P5 – 89.5% expression cassette DNA; P5-HS5 91.5% expression cassette DNA) (AAV8scLP1hFIXco: P5 – 95.7% expression cassette DNA; P5-HS5 97.7% expression cassette DNA) (Fig. 4.20.b). In both instances the P5-HS prep was a higher overall purity than previously produced GMP grade preps (FVIII – 76.1% expression cassette DNA; FIX – 97.0% expression cassette DNA). There was still a significant amount of contaminant DNA detectable from the vector genome plasmid backbone in both cases; however, unlike with the P5 replacement configurations that produced a lower yield, the contribution of vector genome plasmid contaminants did not differ between the P5 and P5-HS configurations in either FVIII or FIX. Interestingly unlike the papilloma-based designs, a slight representation of sequences upstream of the P5-HS promoter were detected, as can be seen on circos plots (Fig. 4.20.c+d). However, this is still considerably lower than in the unmodified P5 preps and in combination with the titer retention achieved the overall purity of the prep is improved with the P5-HS REPCAP plasmid configuration.

a



b

Promoter	Titer (vg/ml)	Upstream conamination %	outside of ITR contamination %	REP contamination %
P5	5.63E+12	5.98%	5.68%	0.48%
P5-HS5	5.30E+12	1.68%	7.23%	0.18%
P5+P97 combo	5.14E+11	1.68%	48.74%	0.23%

c

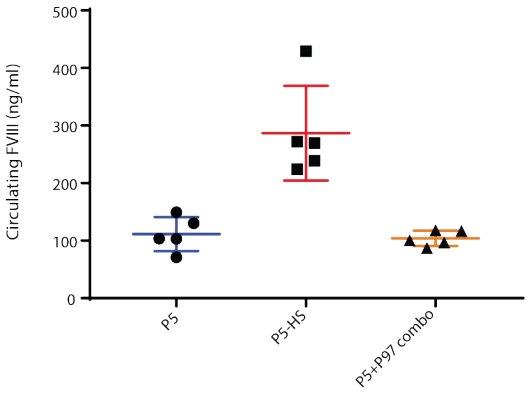


Figure 4.18 - Virus produced with P5-HS can efficiently deliver therapeutic cassettes in vivo.

(a) Schematic; AAV8ssHLPhFVIII was produced at cell factory scale with either P5, P5-HS5 or P5+P97 combination promoters, analysed for contaminants and injected into CBL57/BL6 mice at a dose of 2e10vg. **(b)** Table showing qPCR results of titer and contaminant amplicons in purified AAV8ssHLPhFVIII produced with P5, P5-HS5 or P5+P97 combination promoters. **(c)** Circulating FVIII detected in CBL57/BL6 mice 1-week post injection with AAV8ssHLPhFVIII produced with *(blue)* P5 *(red)* P5-HS5 and *(orange)* P5+P97 combination promoters. *Analysed by one way Anova with Tukey's multiple comparison test.*

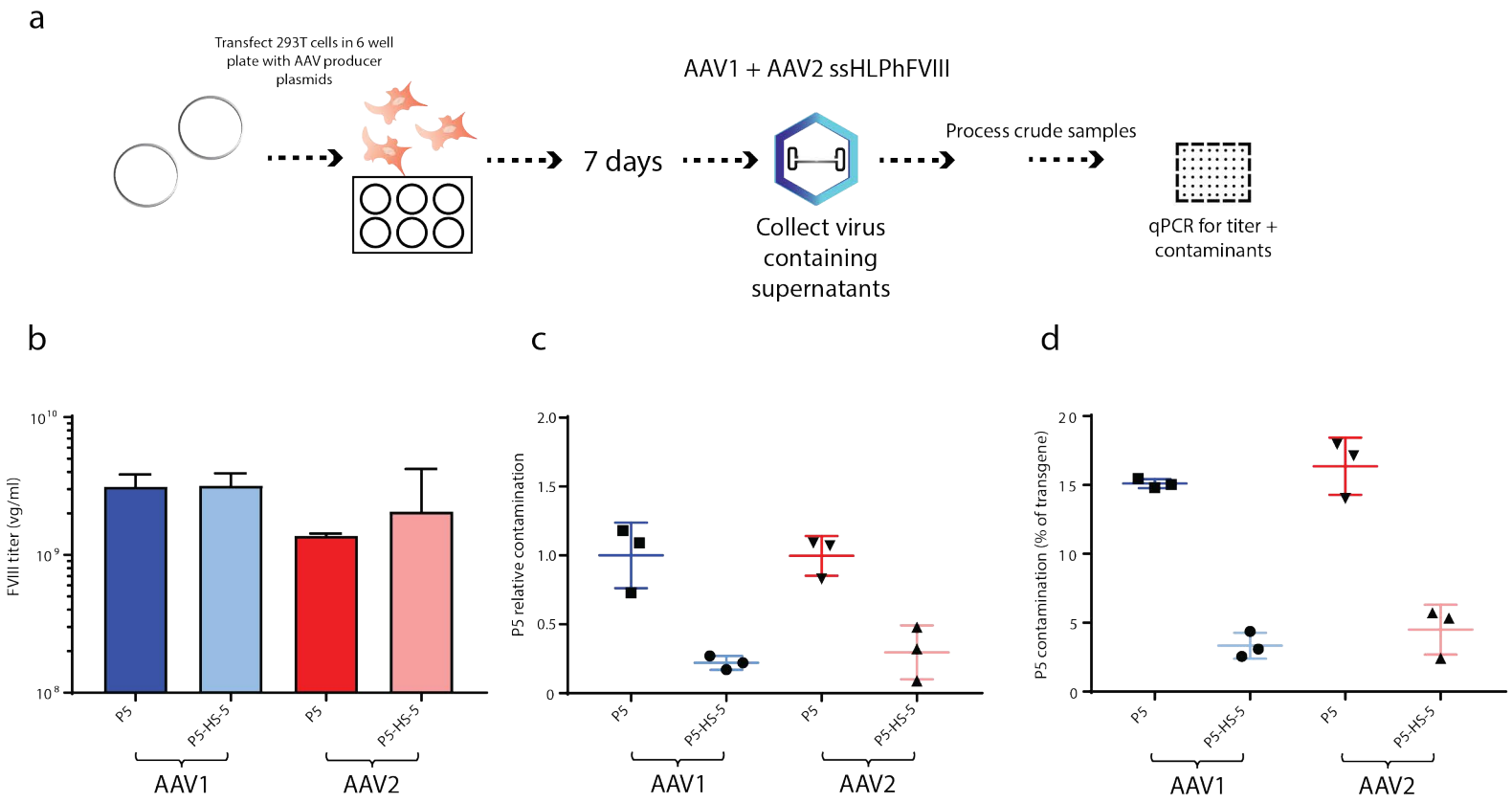
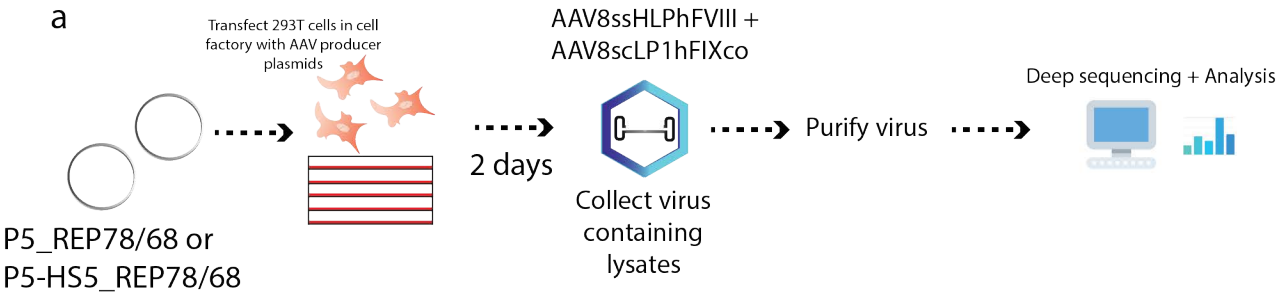


Figure 4.19 - P5-HS configuration reduces contamination in multiple serotypes.

(a) Schematic; AAV1 and AAV2 HLPhFVIII was produced with a REPCAP plasmid using either P5 or P5-HS5 and assayed for titer and contaminant presence **(b)** qPCR results of AAV8ssHLPhFVIII titer produced with *(dark)* P5 and *(light)* P5-HS5 of *(blue)* AAV1 and *(red)* AAV2. **(c)** qPCR results of upstream contaminant abundance of AAV8ssHLPhFVIII produced with *(dark)* P5 and *(light)* P5-HS5 of *(blue)* AAV1 and *(red)* AAV2 normalised to the production of P5 for each serotype **(d)** qPCR results of upstream contaminant abundance for AAV8ssH-LPhFVIII produced with *(dark)* P5 and *(light)* P5-HS5 of *(blue)* AAV1 and *(red)* AAV2 expressed as a percentage of the HLPhFVIII titer. *Analysis for b-d conducted by one way Anova with Sidak's multiple comparison test.*

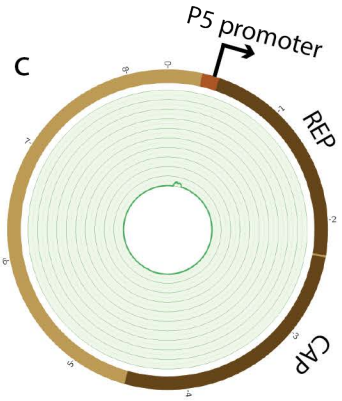
a



b

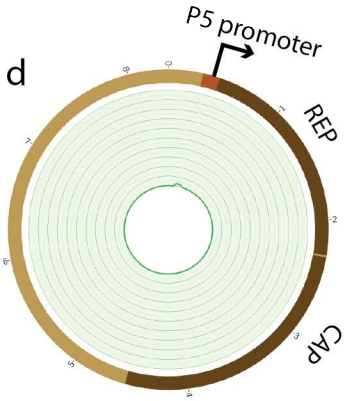
	Overall Purity %		
Prep	**P5**	**P5-HS**	***P5 (GMP)***
AAV8ssHLPhFVIII	89.468	91.528	*76.113*
AAV8scLP1hFIXco	95.728	97.720	*97.047*

c



P5-HS5_AAV8ssHLPhFVIII

d



P5-HS5_AAV8scLP1hFIXco

Figure 4.20 - P5-HS results in higher purity AAV

(a) Schematic; AAV8ssHLPhFVIII and AAV8scLP1hFIXco were produced at cell factory scale using AAV2_8 REPCAP containing either the P5 promoter or the P5-HS5 promoter. **(b)** Table of deep sequencing results showing overal purity of the side by side produced AAV preps (P5 and P5-HS produced side by side; *P5 GMP produced at earlier timepoint in GMP facility*) **(c)** Circos plot of P5-HS5 AAV2_8 REPCAP from AAV8ssHLPh-FVIII **(d)** Circos plot of P5-HS5 AAV2_8 REPCAP from AAV8scLP1hFIXco

4.3 Discussion:

AAV clinical production purity and efficiency is an important issue for the gene therapy field. In recent years the FDA and European regulators have listed AAV product purity and safety requirements that relate to the production process, including contaminating nucleic acids.[502] In order to meet these requirements, strategies to actively reduce DNA contamination of purified AAV will be required.

The rationale for using papilloma early promoters was built off extensive basic research into the relationship between HPV and AAV conducted in the 1990s. There are numerous interactions between the HPV-16 genome and AAV replication proteins, and the HPV-16 proteins E1, E2 and E6 can assist in helper function for recombinant and WT AAV replication and production.[310] The aim of the initial studies was in part to establish a mechanism behind AAV mediated inhibition of HPV-16 driven oncogenic transformation.[494] These early studies of the inhibitory effects of the AAV replication proteins on the papilloma life cycle provided insight into how an alternative production system could be built and implemented.

Harvest timepoint had no observable effect on the relative levels of contamination in a prep. This is somewhat surprising and suggests that this incorporation is occurring throughout the production incubation at a constant rate relative to the expression cassette. This finding shows that merely optimising the virus collection timepoint is not a viable strategy to improve vector purity.

The slight reduction seen in raw values of contamination outside of the ITRs could be in part due to a lower promoter activity and thus production of REP78/68 from these

promoters. The promoter activity of the promoters other than for P5-HS was not directly compared against P5. It is quite likely that the constructs that performed lower than the 50% production threshold that seems to be reached with the papilloma-based promoters could be due to lower transcriptional activity in 293T cells. This is suggested by the fact that converting the start codon from the low efficiency ACG to a standard ATG in the HPV16-P97 promoter configuration improved production from 21.5% of P5 levels to 44.5%.

The HPV-18 driven REPCAP construct included just the first 105bp of the HPV18 genome. In all likelihood this promoter, whilst mapping from the start point of the HPV-18 genome, missed elements from the upstream regulatory region, which helps regulate expression of the E6/E7 gene products of this promoter.[503] It is possible that including those elements would facilitate AAV production, as the BPV based construct, which did include the upstream regulatory region found at the end of its genome was able to do so, albeit at a lower level than the HPV16 based constructs.

It is possible that the level of P19 activation achieved is not sufficient with the papilloma-based promoters. As the REP binding site of P5 aids in the transactivation of P19, something that may not occur with these alternatives which are lacking GCTC containing REP binding sites.

The successful design for improving the purity of AAV production (P5-HS5) retained all the elements of the P5 promoter with a spacer sequence (GCAAC) inserted in between the P5 TATA box and the YY1+1 box. The spacer sequence was generated at random and the only parameters that were required were a lack of a GCTC REP binding motif or an AT dinucleotide. The use of a DNA secondary structure predictor (Probknot, University

of Rochester) indicates that this spacer sequence has the potential to contribute to an exposed loop. This GCAAC sequence was the only spacer employed for this length of sequence. Whilst contamination upstream of P5-HS5 promoter was distinctly reduced, a small number of reads upstream of P5 were still detected when P5-HS produced AAV was analysed by next generation sequencing. It would be therefore be interesting in future to test the effect of varying the sequence identity of the spacer as well as the length that is used.

Despite the 5bp and 100bp spacers appearing functionally equivalent, the 5bp spacer was used for the scale up studies. The rationale behind this was that because both insertion lengths appeared functionally equivalent, introducing a smaller change to the plasmid, in this case 5bp instead of 100bp, reduces the probability of introducing additional sequence elements to the production system that could alter its properties in an unintended manner. As this insertion is occurring within a promoter, perhaps certain sequence insertions may alter the promoter activity and regulation, which if inadvertently introduced could impact ratio of REP proteins and thus the final product. This was not directly tested with respect to P5 and P5-HS constructs but would be an interesting follow up experiment.

With regards to the in vivo infection data: Whilst the P5-HS-produced AAV produced the highest levels of circulating protein in the experiments conducted, the results did not appear to show a clear correlation between contamination levels and therapeutic protein expression. Previous research with minicircles to produce AAV has suggested that lower contamination of product yields superior cellular transduction in the self-complementary context, although not in the single stranded context.[467] No mechanism for how this would

occur has been proven, and there may be other explanations for this such as differences in full capsid ratios, or immune reactivity that is only seen at certain doses. It should be stated that we have only tested a relationship between contamination and expression in the context of the REP78/68 promoter. Indeed, improving the efficiency of other elements of the recombinant AAV genome may also have a beneficial effect on vector purity and may exhibit an improved therapeutic expression phenotype, but that is yet to be shown and will require further investigation.

Chapter 5

Serotype independent conversion of AAV producer plasmids in to P5-HS system:

5. Serotype independent conversion of AAV producer plasmids in to P5-HS system:

5.1 Introduction:

Given the proven utility of the P5-HS system in different AAV serotypes for retaining AAV production capacity and reducing P5 derived contamination, it would be ideal to ensure widespread implementation of such a system. The following chapter details a universal method that takes advantage of this to generate plasmids from any serotype and production configuration that contain the P5-HS5 spacer promoter, and a means to directly test the production titer and DNA contamination levels as compared to the plasmid it was generated from.

5.2 Universal protocol to generate and test efficacy P5-HS5 REP78/68 promoter plasmid:

5.2.1 Materials for generating and testing P5-HS5 REPCAP plasmid:

P5 containing REPCAP plasmid

SDM Forward primer: gcaacCTCCATTTTGAAGCGGGAGG - *Inserted spacer sequence is shown in red*

SDM Reverse primer: ACCCTGCGTGCTCACTCG

NEB SDM kit (New England biolabs Cat# E0554S)

PCR tubes

Thermocycler

HEK293T cells (ATCC Cat# CRL-3216)

D10 media

Applied biosystems PowerSYBR reagent (ThermoFisher Cat#4367659)

Contamination Forward primer for qPCR: TAGAGGTCCTGTATTAGAGGTCACGT

Contamination reverse primer for qPCR: GCGTGCTCACTCGGG

Sanger sequence primer: TAGAGGTCCTGTATTAGAGGTCACGT

Annealing temperature: 67 degrees Celsius.

0.5M EDTA pH8.0.

NEBuffer 3 (NEB Cat#B7003S)

DNASE1 (NEB Cat# M0303)

Proteinase K. (Cat. # AM2546, Ambion)

5.2.2 Protocol for generating and producing AAVwith P5-HS5 REPCAP plasmid:

Day 1:

Dilute purified P5 containing plasmid to 10ng/µL in $_{dd}$H20

Make up PCR mix as follows: (NEB SDM protocol E0554)

Reagent	Amount to add (µL)
2X Q5 Hot start master mix	12.5
10µM Forward primer: gcaacCTCCATTTTGAAGCGGGAGG	1
10µM Reverse primer: ACCCTGCGTGCTCACTCG	1
10ng/µL P5 containing plasmid	1
$_{dd}$H20	9.25
Total	25

Run reaction in thermocycler using the following protocol:

Phase	Temperature	Time (seconds)
Denaturation 1	98°C	120
Denaturation cycle	98°C	10
Annealing	67°C	30
Extension	72°C	255*
Final extension	72°C	120
Hold	12°C	

* Based on a maximum P5 containing plasmid size of 8.5kb. For larger plasmids use 30 seconds extra per kb.

Once PCR reaction is complete thaw competent cells on ice.

Important: Do not transform into the NEB SDM kit provided competent cell type. For accurate comparison between the starting plasmid and the mutagenized product for AAV production efficiency and relative contamination levels the SDM product should be transformed into the same competent cell type the starting plasmid was stored within.

Make the DPN1 digestion mix using the NEB KLD reagents:

Reagent	Amount to add (µL)
2X KLD reaction mix	5
SDM PCR product	1
10X KLD enzyme mix	1
$_{dd}$H20	3
Total	*10*

Mix sample and incubate at room temperature for 5 minutes.

Transform 4µL of DPN1 treated PCR product mix into competent cell type as per manufacturer's instructions. Resuspend competent cells post heat shock in 250µL of S.O.C. media.

Incubate S.O.C./ transformed bacteria mix in a shaking incubator at 37°C for 1 hour.

Streak 100µL of S.O.C./ transformed bacteria mix onto Agar plate harbouring the appropriate antibiotic.

Incubate bacteria containing Agar plates at 37°C for 16 hours.

Day 2:

Check bacterial plate for colonies:

Pick colonies X6 per construct and seed into 5ml NZY or LB media + antibiotic.

Incubate for 16 hours at 37°C

Day 3:

Remove picked colony cultures from incubator.

Take 0.5ml of media and add it to a corning tube containing 0.5ml of 50% glycerol. Transfer stock to -80°C freezer.

Purify DNA from remaining 4.5ml culture by QIAGEN miniprep kit.

Elute plasmid DNA in 35µL $_{dd}$H20

Sanger sequence verify transformed plasmid product reaction using the universal primer sequence:

Reagent	Amount to add (µL)
3µM universal P5-HS primer:	1
100ng/µL P5-HS plasmid	1
$_{dd}$H20	10

Day 4:

Plate 293 based production cells into a 6-well plate at 800k cells per well. (3 wells are required per plasmid condition)

Day 5:

Make up transfection mix as below:

Add DNA solution to PEI solution and mix by inversion 5 times.

Incubate solution at room temperature for 15 minutes.

Add transfection solution dropwise to 293 cells.

Incubate 293 cells at 37°C 10% C02 for 7 days.* - Change media on day 2 to remove leftover plasmid DNA

*After 7 days over 90% of produced AAV will be in the supernatant.[349] However, the duration of this step can be altered to match preferred production conditions.

Day 12:

Collect 100µL of supernatant from each of the AAV production wells.

Either freeze the supernatant directly at -20°C or proceed to the next step.

Remove non AAV packaged DNA from cell supernatant:

Reagent	Amount to add (µL)
AAV production supernatant	2
NEBuffer 3	10
DNASE1 (200U/µL)	3
ddH20	85

Incubate DNASE1 treatment reaction at 37°C for 1 hour.

Inactivate DNASE1 with addition of 2µL of 0.5M EDTA (pH8.0)

Incubate at 98°C for 10 minutes

Cool solution on ice.

Add 2µL of 10% SDS, 2µL of 20mg/ml Proteinase K and incubate at 55°C for 1 hour.

Incubate sample at 98°C for 10 minutes to inactivate Proteinase K.

Add 98µL of 0.01% Pfu68 to bring total volume to 200µL

Store sample at -20°C or proceed to quantification.

5.23 qPCR of AAV titer and contamination from 293 supernatants :

Materials:

10% Fluronic F-68 (cat# 24040-032, Gibco by Life Technologies)

Sample or standards diluent: 0.01%(v/v) F68 in nuclease free water

Dilute 10% Fluronic F-68 (stock) with nuclease free water to 0.01% (v/v) F68, use this as all virus diluent diluent

(2X) PCR mix

MicroAmp Optical 96-Well Reaction Plate with Barcode (Cat#4306737, Applied Biosystems)

MicroAmp Optical 8-Cap Strip (Cat# 4323032, Applied Biosystems)

Applied Biosystems 7500 Real Time PCR System

REPCAP plasmid standard:

Transgene primers: Varies upon construct

P5 contamination forward primer: TAGAGGTCCTGTATTAGAGGTCACGT

P5 contamination reverse primer: GCGTGCTCACTCGGG

Prepare 1:10 serial dilutions of standards: from stock standard linear 5E8 ssDNA to 5E2 ssDNA in diluent 0.01% F68.

Prepare a master mix for qPCR reactions:

Reagent	Per PCR reaction (µL)
SYBR Master MIX (2X)	12.5
primer F (100µM)	0.125
primer R (100µM)	0.125
Nuclease free water	7.25
Total	20

Total reaction volume is 25µL i.e. 5µL of template DNA (viral or standard)

20µL of master mix to each required well in a 96-well plate.

5µL of template sample is added to each well. (Each sample is measured in triplicate.)

qPCR plate is sealed with 8 Cap strips.

Spin down qPCR plate at 200g for 5 minutes in a benchtop centrifuge.

Place qPCR plate in Applied Biosystems 7500 machine (or equivalent)

Run qPCR program as follows: (40 cycles)

Phase	Temperature	Time (seconds)
Initial hold	60°C	120
Denaturation 1	95°C	600
Denaturation cycle	95°C	15
Annealing cycle	60°C	60
X40 cycles		

Export and analyse AAV titer and relative contamination originating at P5 promoter.

The above methodology was used to produce the AAV1 and AAV2 P5-HS variants that were tested and has been used to produce other serotype plasmids currently under investigation.

5.3 Discussion:

The majority of produced AAV utilises the P5 promoter in some form. In chapter 4 the P5-HS promoter was tested in serotypes other than AAV8; AAV1 and AAV2. The methodology presented in this chapter will allow a 5 base pair spacer to be introduced into any AAV producer plasmid that utilises the P5 promoter sequence and provides a means to test this in direct comparison to the original P5 containing plasmid without the introduction of confounding factors. The sequence upstream of P5 will vary depending on the plasmid design, any differences in amplicon amplification kinetics, and finding amplicon sets equidistant upstream of P5 in different plasmids could be a challenge when attempting to compare contamination levels between heterologous designs. Here a universal primer set contained within the P5 promoter that can successfully amplify contaminant sequences is presented. (Fwd: TAGAGGTCCTGTATTAGAGGTCACGT; Rev: GCGTGCTCACTCGGG). These primers should be used in addition to further upstream amplicons to characterise AAV preps for REPCAP contamination. Further testing of the P5-HS spacer system in multiple production contexts; suspension vs adherent, 293 vs 293T, split REPCAP vs standard design. This will be an important to ascertain if the benefit to using the P5-HS system is truly universal to AAV production or has certain constraints, be it capsid, REP gene or configuration based.

Chapter 6

Discussion

6. Discussion:

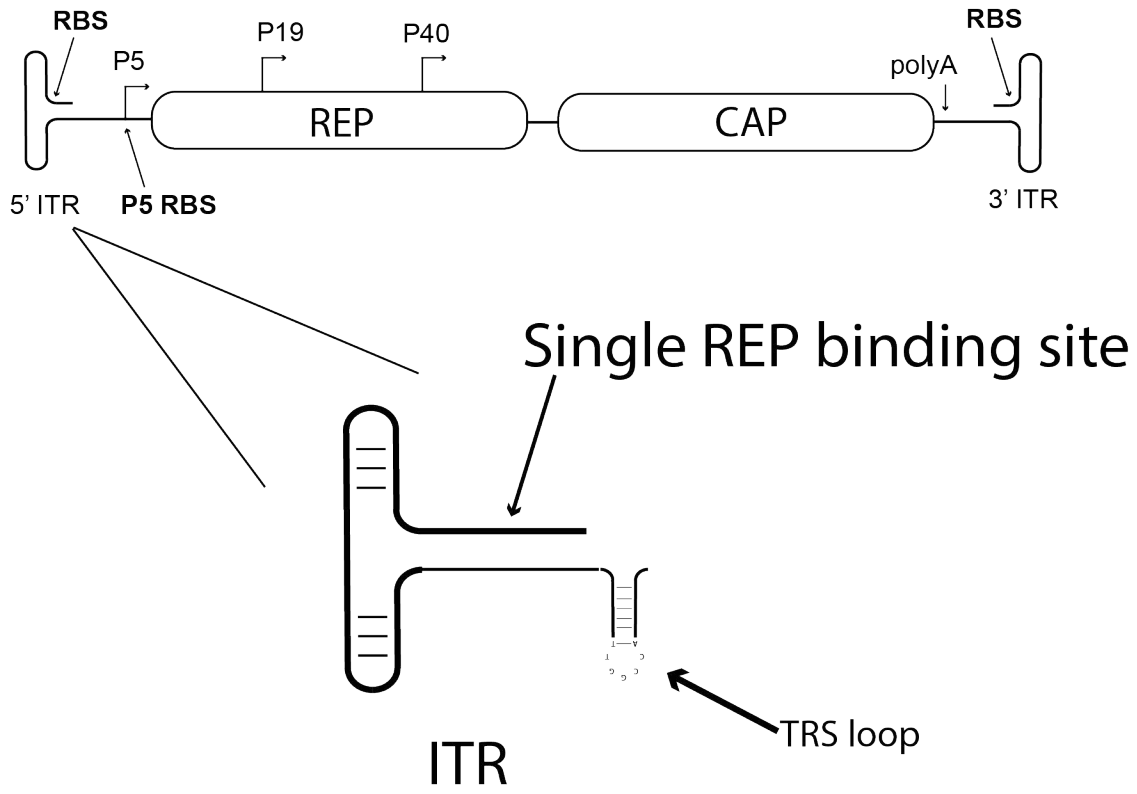
6.1 General Discussion:

AAV gene therapy stands at the forefront of novel therapies for genetic diseases. The number of clinical trials using AAV as a modality continues to increase. As such an expansion occurs, the onus lies with researchers in the field to ensure that AAV based treatments are highly reliable, safe and efficacious. A question lies within this however, on where to set benchmarks on what constitutes reliable, safe and efficacious. For instance there are monogenic diseases that are of such severity and with poor alternative treatment options, such as spinal muscular atrophy or Batten's disease, that to apply a rigid standard on certain product related impurities may deny patients the chance to survive that an AAV gene therapy approach may offer. On the other hand, there are other less mortal diseases that may require a different risk-benefit analysis for the use of AAV. Nevertheless, it is essential to understand the nature of the product that is administered in any given trial. In the research conducted for this PhD the aim was to characterise in detail the profile of non-expression cassette DNA present within AAV preparations, the potential activity of that contaminant DNA post infection, and to subsequently develop a means to minimise this. This has been achieved, with several questions answered and a new component of the AAV vector toolkit developed.

In chapter 3 the profile of DNA contamination from producer plasmids was described. The contamination profile was identified first by qPCR, and verified by deep sequencing of purified AAV preps. By mapping deep sequencing reads to producer plasmids, three

origins of contamination could be identified. DNA sequences adjacent to these sites were overrepresented in the preps; the 5’ ITR, the 3’ ITR and upstream of the P5 promoter. This profile was retained across different vector genomes, although their contribution to the total DNA within AAV varied considerably. Interestingly, self-complementary ITRs did not appear to initiate contamination efficiently, as minimal reads mapped to the regions just outside of scITRs, suggesting that this could be an additional advantage to using the self-complementary vector genome configuration where the expression cassette size is below 2.35kb. The contamination profile from the P5 promoter raises an interesting question. REP is known to be a directionally active protein, so it stands to reason that any replication activity would occur in a single direction. This is indeed seen in the P5 containing REPCAP plasmid, with contaminant incorporation being seen upstream but not downstream. In the ITR containing plasmids, unidirectional replication should be occurring in the direction that excises the vector genome from the plasmid at either ITR. Whilst this is the predominant mechanism in the expression cassette plasmid, the presence of contaminants outside of the ITRs suggest some level of initiation in the opposite direction. It is possible that this is due to the manner in which ITR containing plasmids are set up. In the P5 promoter, there is 1 REP binding site (Fig. 6.1.a). In the recombinant AAV genome there is 1 REP binding site per ITR. Yet, for the ssAAV genome to form the ITRs from a dsDNA plasmid construct, 2 REP binding sites per ITR must be present in inverted form. (Fig. 6.1.b) It is possible that some degree of directional REP binding and excision into the plasmid backbone is occurring as a result of this. A good way to test this would be via the creation of single stranded ITR flanked DNA cassettes

a



b

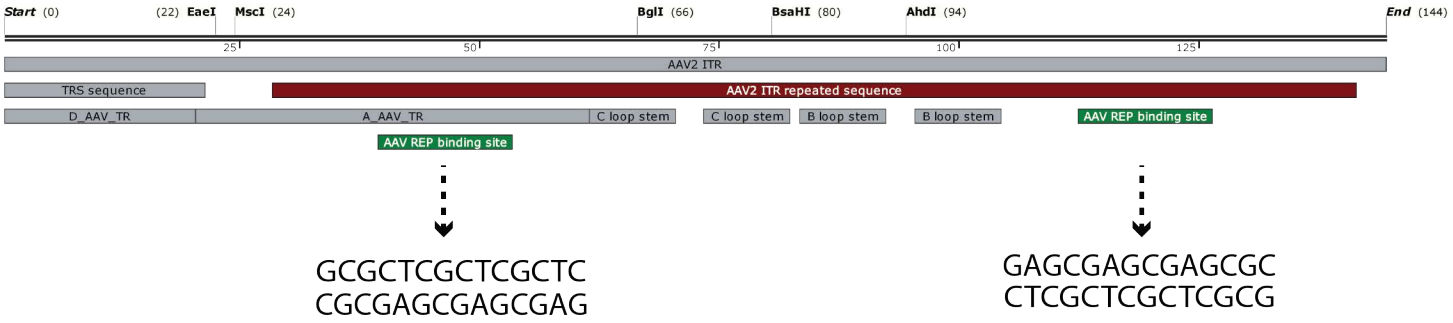


Figure 6.1 - REP binding sites in plasmid inverted terminal repeat sequences

(a) Schematic; AAV ITR showing the position of a single REP binding site in the ITR of the viral genome. **(b)** Map of ITR sequence from plasmid DNA (snapgene). The inverted nature of the terminal repeat means that in the context of double stranded plasmid DNA that there are two sequences that REP can bind to, positioned in mirrored orientation.

to transfect in AAV production, perhaps generating these by asymmetric PCR from the expression cassette plasmid.[504]

Chapter 3 also described the presence and activity of AAV producer plasmid derived contaminants post infection. The DNA profile of contaminant reads in post infection mouse hepatocytes matched that of the preps. Furthermore, the RNA profile of contaminants showed reads that mapped back to regions both upstream of the P5 promoter and outside of the ITRs. Focusing specifically on the anti-P5 sequence, it was identified that both transcription and translation could be driven from contaminants originating upstream of the P5 promoter both in vitro and in vivo. Anti-P5 promoter activity had previously been identified to drive small RNAs in the WT AAV context in HeLa cells.[272] In fact, in the recombinant AAV setting, and the cell lines tested in this work the implication of anti-P5 promoter activity appears to be amplified. Transfection experiments showed in vitro that the anti-P5 sequence had an activity equal to that of the P5 promoter, and that akin to the P5 forward sequence, its activity appeared to be repressed in the presence of REP. In recombinant AAV there is no REP gene present in the final product, and therefore no REP present to repress the activity of the contaminant sequences attached to the incorporated P5 sequence that are transferred to the cell post infection. A caveat to this would be that the 293T cells used for the transfection of the P5 fluorescent constructs contain the E1a sequence from adenovirus type 5, which in the forward direction of P5 activity removes the transcriptional repression mediated by YY1 on the P5 promoter.[297] If YY1 were to provide this same repression in the negative direction it is possible that the observed activity in 293 cells may be artificially higher than what would be seen in an infected cell. However, the regulation of anti-P5 activity by YY1 was not investigated

during this study, and the RNA sequencing data and protein studies suggest that transcription is readily occurring upstream of P5 in cells post infection in mouse hepatocytes.

In addition to the activity of sequences upstream of P5, the RNA seq profiles showed reads mapping to sequences outside of the ITRs. The AAV ITRs have previously been shown to mediate transcription and hence transgene expression.[257,478] However, transcriptional activity has not previously been investigated with the reverse of these sequences. This makes sense as in a typical setup you would not expect to observe reverse ITR sequences, and only in the presence of these ITR derived contaminants does this become an issue. It also stands to reason that the ITR would show transcriptional activity in its reverse direction due to most of the sequence being an 'inverted repeat'. We can hypothesise that the vector genome plasmid derived contaminants that exhibited transcriptional activity in the mouse RNA studies were being driven by reverse ITR sequences, however we only tested the reverse ITR promoter activity in the plasmid context. Whilst these experiments did lead to detectable transcriptional activity, this would not necessarily mimic the secondary structure of the contaminant fragment in its infected form, and therefore further work should be done to assay whether it is indeed the reverse ITR sequence that is driving contaminant transcription and whether these contaminants can be reduced or removed by modifications to the plasmid design or other strategies.

Chapter 3 also used a reporter expression cassette placed upstream of the P5 sequence to demonstrate the potential of sequences upstream of P5 in purified AAV to be translated. This was observed both in vitro and in vivo. The significance of this is not to say that every plasmid construct with P5 derived contaminants will result in protein

product, but that the potential for this exists. Different research and clinical groups will use different plasmid setups in their AAV production strategies and may not take consideration of the location of the P5 promoter when placing other required genes. For instance, if antibiotic resistance genes or adenoviral helper genes are present on the P5 containing plasmid, it is possible that they could become incorporated and expressed if proximally upstream and correctly oriented in relation to P5. This could have a negative downstream impact on the target tissue, be it via an increased immunogenic load or unintended interactions of the aberrantly produced protein product. So far, this phenomenon has only been tested in vivo with regards to liver directed gene therapy, and it is possible the risk will be lower with other target tissues for gene therapy, although this will have to be investigated further.

An interesting observation from the in vitro translation studies was the potential for persistence in dividing cells. After infection with either an ITR flanked GFP cassette or a P5-GFP contaminant cassette with no provided ITR flanked expression cassette, the amount of DNA detected in 293T cells 3 weeks post infection as compared to 72 hours post infection was not significantly different between the standard cassette and the contaminant cassette. AAV is known to integrate in the genome at a low rate.[505] It is likely that the persistent GFP expression seen in the 293T cells at 3 weeks post infection is the result of integration events. Indeed, when GFP positive contaminant cells were sorted out from the negative population at 72 hours post infection and left to grow, some wells produced GFP positive colonies. The rate by which this integration by contaminants occurs however is unclear, and no attempt was made in this study to look at potential integration sites. The results do however support the argument that these contaminants

need to be minimised, as any increase to the potential integrative load of DNA increases the probability of a disruptive insertion into the host genome. The integration profile of these identified contaminants is an important question that should be addressed in future studies. The hypothesis would be that the contaminants would integrate in a random process, but that in the presence of REP these sequences would be directed towards the characterised AAV integration sites AAVS1, S2, and S3.

The consistent observation made in this study of the transcriptional activity of AAV contaminants conflicts with the conclusions of a well cited previous report.[466] However, it does not appear that the results are incongruous. In the Hauck et al study the primary focus of assessing potential for contaminant transcription was the capsid gene and was centred around the observation of an immune response against capsid epitopes in the unsuccessful AAV2 based haemophilia B gene therapy trial.[425] It was hypothesised in that study that based upon the timing of this immune response, the capsid epitopes may be sourced not from residual AAV capsid protein after infection, but from transferred and transcribed capsid contaminant DNA. Whilst residual DNA impurities were detected post infection, transcription of capsid gene sequence was undetectable. The conclusion was therefore made that the CD8+ T cell immune response against the capsid epitopes was the result of the delivered virus and not from de novo synthesis mediated by contaminating sequences. Following what has been elucidated in the presented thesis, it is unsurprising that this 2009 paper did not identify transcription of contaminants. Firstly, presuming that a standard configuration with regards to P5 promoter position was used in these production plasmids, the amplicons in that study would have been many kb upstream of P5 and thus unlikely to be ideal for detecting contaminant transcription.

Indeed, the level of cap gene contamination was quantified at 0.00018% of the transgene, suggesting the purity of the prep used for the analysis of contaminant transcription was such that the sequences assayed by qRT-PCR were likely under the limit of detection. Similarly, the adenoviral and ampicillin resistance amplicons tested in CBL57/Bl6 mice did not yield detection of transcript in the Hauck study. Exact quantification of these amplicons is not listed in this paper but referred to as “trace amounts”. In contrast, the analysis of AAV preps in this thesis has elucidated that the sequence identities themselves are not relevant, as the P5 promoter serves as the initiation point for both contaminant incorporation and contaminant transcription. In the standard configuration this leads to the transcription of antibiotic resistance genes contained within the plasmid backbone, but not necessarily in a manner that would lead to their translation. If the P5 promoter is positioned downstream of CAP, as is common to many publicly available REPCAP plasmids of varying serotypes (Table 6.1), it is those upstream capsid sequences that will be incorporated and then potentially subject to transcription (although this would still not yield AAV capsid protein fragments being produced due to the directionality of this transcriptional activity).

Another issue with respect to AAV contamination raised in the Hauck study is the question of backbone size to reduce contamination. The study identified sequences from the ITR containing plasmid backbone to be present with abundance in the preps.[466] When the backbone size was increased from 2.5kb to over 6kb the contaminant sequences detected were reduced over 7 fold. The amplicon used to measure this mapped to the ampicillin gene, present in all backbone constructs, but did not factor in the distance of the ampicillin from an ITR in each vector setup. The conclusion was made that by

Plasmid	Serotype	Addgene ID	Notes
pAAV2_1	AAV1	112862	
pAAV2_2	AAV2	104963	2 copies of P5
pAAV2_5	AAV5	104964	2 copies of P5
pAAV2_7	AAV7	112863	
pAAV2_8	AAV8	112864	
pAAV2_9n	AAV9	112865	
pAAV2_rh10	rh10	112866	
rAAV2-retro helper	AAV2-retro	81070	2 copies of P5
7M8	7M8	64839	2 copies of P5
shh10	shh10	64867	2 copies of P5
pAnc80L65	Anc80L65	68837	2 copies of P5

Table 6.1 - Addgene AAV plasmids with P5 promoter downstream of CAP gene

AAV serotype REPCAP plasmids available from plasmid repository addgene containing the P5 promoter downstream of the capsid gene. (Analysed for P5 presence in snapgene) Some of the plasmids identified have 2 intact P5 promoter sequences; one directly upstream of REP and one positioned downstream of CAP, theoretically allowing for 2 initiation points of REPCAP contamination initiation

increasing the backbone size to over 5kb i.e. the packaging capacity of AAV, the incorporation of contaminants by “reverse packaging” could be considerably reduced. The work in this thesis builds upon this by showing that even with an oversized backbone (19kb in this thesis work) initiation of packaging in the reverse direction occurs at ITRs, meaning that though an oversized backbone may help in reducing the level of DNA contaminants, it is not sufficient to eliminate these sequences from AAV preparations.

Chapter 4 focused on developing a system that would remove the contamination phenotype upstream of P5, to yield a purer overall AAV product. It was known that a simple promoter substitution would not yield high quantities of AAV from previous work,[506] and that an effort to replace or modify P5 would ideally retain the autoregulation of AAV REP protein expression in the AAV replication process. High titer production that removed the upstream contaminant phenotype in the REP78/68 driving promoter was achieved in several constructs, that used the REP binding site present in the early promoter of HPV16 viral genome ‘P97’. However, in these constructs it was not possible to generate a viral yield that equalled P5 and on further investigation this caused an increase in DNA contamination from the other plasmid source, the vector genome plasmid. Fascinatingly, the raw numbers of contaminating sequences as determined by qPCR did not vary considerably between REPCAP constructs, and it appeared that the increase of contamination was as a function of the titer in each production context. This may have implications for transgene design and selection. Certain gene construct configurations may be packaged less efficiently than others. For instance, constructs encoding short hairpins close to the WT ITR in self-complementary AAV designs have lower yields than

those that do not.[463] Other constructs that incorporate an expression cassette which interferes with cellular processes yield lower viral quantities.[507] Notably, cassettes that are oversized result in lower yields than those that fit within the packaging capacity of AAV.[508–510] Oversized cassettes are often used in AAV production, as groups try to push the envelope on what genetic information can be delivered through AAV. Indeed, much of the work in this thesis was done by packaging a minimally oversized FVIII cassette. If contamination yield at a given scale were to remain constant as seen in our studies, then contamination as a percentage of the transgene would be greatly increased in cassettes that produce with low efficiency. Further study of this precise question will be useful in aiding future therapeutic cassette designs.

Implications also exist for gene constructs where the GOI exceeds the packaging capacity of a single AAV virion and a dual vector approach is used. Dual AAV approaches have been designed and implemented for several preclinical gene delivery strategies for disorders like congenital hearing loss, retinal disorders, and dysferlin deficiency.[511–513] Dual AAV approaches have also been applied to gene editing approaches for disorders such as Tay Sachs and Sandhoff disease.[514] One recent gene editing strategy for haemophilia B in mice used dual AAV delivery to provide saCas9 and a full FIX copy for insertion at the albumin locus.[515] The efficacy of this approach relies on the successful transduction of both sides of the dual cassette in a single cell. By its very nature the production efficiency is halved, as 2 genome containing particles are required for 1 transducing unit. As such the ratio of contaminant DNA vs full expression cassette copies would be doubled. The same principle would apply to situations in which 3 AAV particles are required, as has been developed preclinically for Alström syndrome and certain

disorders of the inner ear.[516,517] A triple AAV delivery has also been applied in the clinical setting for Zinc finger mediated gene editing to treat Mucopolysaccharidosis I (Clinical trial# NCT02702115) and Mucopolysaccharidosis II (Clinical trial: NCT03041324). In these cases, contamination from production would be tripled for each functional copy of the cassette administered.

The imperfect findings in chapter 4 relating to the substitution of the P5 promoter with sequences from the papillomavirus early promoter sequences led to the investigation of more subtle modifications to the P5 promoter. The direct substitution of the P5 RBS into the HPV16-P97 promoter did not lead to the direct upstream incorporation of contaminant sequences, suggesting that an RBS alone was not sufficient to drive contaminant incorporation. This makes sense as previous studies have shown that an RBS and a TRS nicking site are required for AAV replication and indeed integration at preferred REP dependent sites.[518] The deletion of 2 nucleotides at the TRS mimic site in the YY1+1 box of the P5 promoter saw yields drop to 50% of the intact P5 promoter. It is possible the disruption of YY1+1 binding could impair P19 activation, in which the other YY1 binding site YY1-60 aids in the P5 dependent REP mediated transactivation of P19.[175] This hypothesis was not directly tested. Alternatively, the disruption of the YY1+1 box could have directly decreased transcription from the P5 promoter. The subtle modification that retained AAV titer at normal levels retained all elements of the P5 promoter, i.e. the MLTF binding site, the YY1 boxes, the P5 TATA (includes the RBS). This design merely added a spacer between the end of the TATA box and the start of the YY1 box. Retaining the vector titer led to a product that was of higher overall purity. The implications of this are far reaching within the AAV gene therapy field. Because the modification to improve

vector purity was made to the REPCAP plasmid, any ITR flanked expression cassette construct may be used. Furthermore, as demonstrated within chapter 4, this system is applicable across AAV serotypes, as all AAV capsid genes can utilise the P5 driven AAV2 REP gene in 293 based production. The relationship observed between the higher viral titers and lower contamination brings up an interesting question with regards to serotype: A previous study demonstrated vast differences between AAV capsid identity and production titers.[519] The results in our serotype test of the P5-HS system support this and indicate that low producing serotypes such as AAV2 result in vastly increased contamination phenotype. A whole subfield exists in both the academic and commercial realm to alter the AAV capsid for desired properties, most commonly transduction efficiency of the target cell. However, if the best transducing variant produces considerably lower viral yields, then the purity of the resulting product may be far lower than an alternative serotype with a slightly lower transduction potential. Finding this balance should be a consideration for clinical decision makers and indeed regulators to ensure the best and safest product is developed. The system developed in this work should aid in achieving that goal.

The results from this study did not aid in the active removal of contaminant sequences outside of the ITRs. One crucial observation was that high efficacy AAV production reduces the proportion of contaminants in the final prep. For an individual expression cassette this means that the best way to reduce contamination is to have efficient production. The packaging of oversized genomes is known to be less efficient than genomes ≤ 4.7kb.[509,510,520] For expression cassettes that are minimally oversized, DNA contamination may be a greater concern than cassettes that package efficiently. Based

on the evidence amassed, I would hypothesise that a dual AAV approach would result in a purer overall treatment strategy than a low production efficiency single virion method that used an oversized expression cassette (Figure 6.2), although this will likely depend on how oversized the construct is. The precise effect on cassette size on DNA contamination and production efficiency should be investigated in detail in future studies.

In chapter 5 a universal method for the conversion of AAV production plasmids into the P5-HS system is described. The purpose of this is in part detailed in the previous paragraph. When making comparisons between two variables the best experiment only changes that specific variable. Plasmids used in the research setting are often transferred lab to lab depending on user need and not necessarily constructed from a previous design. In the context of AAV production this could mean REPCAP plasmids of different serotypes differing in not just the capsid sequence, but other features like P5 promoter positioning, antibiotic resistance genes, and other backbone sequence identities. By providing a stepwise means of direct conversion (5bp SDM insertion), and a robust assay for small scale (6 well) assessment of yield and purity, this P5 promoter manipulation can be easily implemented into the production plasmid system of an individual research group and compared in different settings to origin plasmids previously in use. Such truly direct comparisons may help tease out serotypes or configurations in which the benefits of the system are either heightened or minimised. Furthermore, as the field begins to look to the REP genes of AAV in the manner that it has viewed the CAP genes for many years, as sequences to modify and alter for beneficial properties, it will be worth studying whether this P5-HS configuration provides any benefit to production in scenarios in which the REP protein is not derived from WT AAV2.

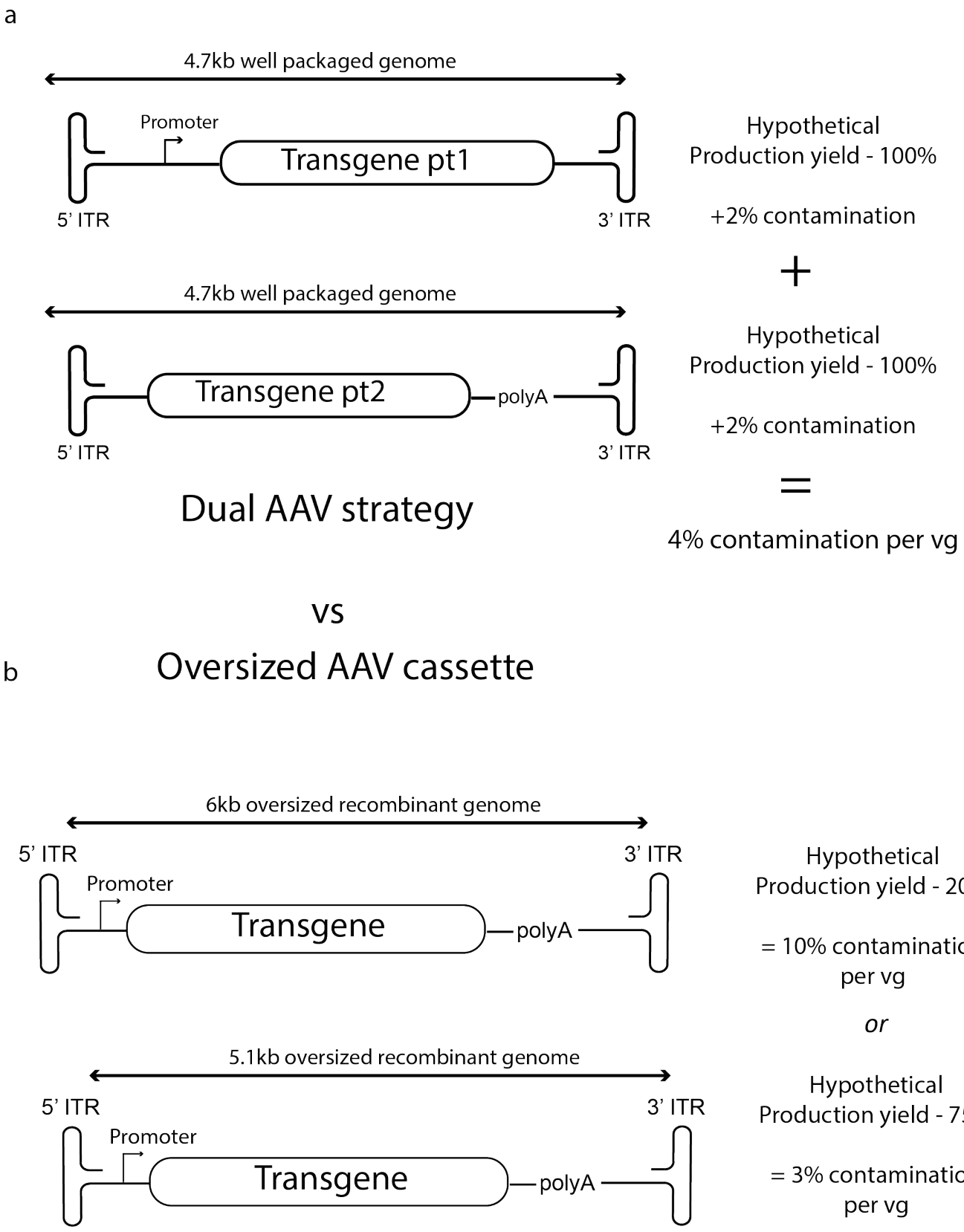


Figure 6.2 - Hypothesis of dual AAV strategy vs oversized AAV for delivery of a large transgene

(a) Schematic; If a dual AAV strategy is used with a well packaged cassette to deliver a large transgene in two portions and it is assumed both cassettes package with maximum efficiency, then the contmaination contribution will be doubled per transducing unit, as both copies are needed. (b) Schematic; If a single AAV strategy is oversized to the point that production yield drops below 50% of a well packaged genome, then the overall contamination per vector genome will be greater even though only a single cassette is being used.

The production system used in this work used a consistent methodology: All production was done in 293T cells, and all with adenovirus type 5 helper genes. Helper functions have been identified in viruses other than adenovirus, in viruses such as HSV-1, HPV16, and others,[306,310,521] and research groups often seek to take advantage of the beneficial functions of these other viruses in the production process. For instance genes from the Human Bocavirus 1 genome have recently been used to boost AAV production levels.[313,522] It is likely that further genes helpful to the production of AAV will be elucidated and implemented in production schemes. The broad strata of viruses that have already been identified to provide helper function to AAV replication indicate that more efficient helper proteins could be discovered or artificially synthesised. It will be important as the field evolves to produce virus for clinical applications in new and improved ways that vector purity is considered alongside of titer as a high priority consideration. Side by side comparative studies of the contamination profile of different AAV production systems are lacking and given the already known interaction between the AAV REP proteins and these other viruses.[202,495,523] It will be important to assess that where these elements are used in an AAV production setting, they are not inadvertently driving an active DNA contamination process that could compromise the purity of the prep. It should be noted that even plasmid DNA sequences at great distance from the identified incorporation sites are still represented at a detectable rate. It may be beneficial to build out stable cell lines that incorporate the adenoviral helper gene functions to create a system that will reduce to incorporation of any sequences from viral origin be it adenoviral or AAV based.

With regards to how this work can be implemented into improving gene therapy for haemophilia: A recent estimate put the total prevalence of haemophilia patients worldwide

at over 1.1 million.[524] Currently approved gene therapy products are debuting on the market for costs upwards of 1,000,000 USD.[525] Using this proposed economic model for haemophilia, a 1-million-dollar treatment would cost over 400 trillion to reach the 418,000 estimated haemophilia patients worldwide with severe disease.[524] Given the vast cost of the current standard of care; recombinant factor prophylaxis, this high price tag should still result in cost savings for health services over an extended period. However, the setup of the economic model is such that patients in developed nations will likely not be the majority recipients of gene therapy treatments for haemophilia in the coming years. Arguably the greatest clinical benefit of a onetime treatment would be to translate gene therapy approaches to developing nations, where far less access to recombinant factor exists, especially for prophylactic regimens. To achieve such a goal, further improvements to streamline cost of production are going to be essential if this form of therapeutic is to be implemented on a truly global scale. One aspect of the process that will help in this role is improving viral titers. Strategies exist to do such a thing, such as alteration of the P40 splice site in the REP gene.[526] or knocking down certain proteins in the production cell line such as YB1.[527] Whilst this thesis work did not yield an improvement in viral titers with the final production design, it achieved a higher level of purity whilst retaining the yield, and thus something that could be implemented without increasing overall production cost.

The recent strides made in the field of genet have opened opportunities to treat genetic disorders with a precision, durability and efficacy previously unattainable to biomedicine. This is objectively beneficial to patients. However, the rapid pace at which treatments are being developed with these tools can lend itself to a parochial deference regarding the

tools themselves. It is essential to further our understanding of the best treatment for a given disease to asses with equal intensity the merits and drawbacks of the tools. To take a parallel example; shRNA technology has revolutionised our ability to silence gene expression and has been applied to the clinic in a variety of contexts.[528] However, it was only recently that a mechanism behind potential liver and cardiotoxicity after delivery of shRNA constructs was characterized and understood. [529,530] This knowledge allows adaptation of the design; in this specific case the use of 19 nucleotide shRNAs instead of 21 for delivery to the heart to avoid toxicity.[530] Similarly, the characterisation and understanding of all the component parts of AAV production and purification is key to the future of gene therapy. Unaccounted for differences between AAV preparations could confound the results of comparisons made between different vector designs and lead to incorrect conclusions being drawn. Final purity of vectors should be a part of this consideration when performing these types of experiments. To give an example: conflicting data was referenced in the introduction about the ideal serotype for liver gene therapy delivery.[418–422] Greater characterisation of AAV construct purity in the research setting may help to settle any differences in outcome and reproducibility related to product efficiencies and should be strived for in studies that utilise AAV as a research modality regardless of experimental aims.

6.2 Future Directions:

The research conducted in this thesis has provided avenues of future research that should be followed up on:

With regards to AAV vector purity it should be highlighted that there have been other tested hypotheses to reduce DNA contamination from AAV vectors. DNA Minicircles have been previously been used for AAV production plasmids to reduce AAV contamination.[467,531] However, this modality would not be successful in preventing contamination from the REPCAP plasmid, as a constrained circle to eliminate the vector backbone would bring the P5 promoter into closer upstream proximity to the capsid genes and merely result in an increase in the incorporation of AAV capsid and potential replication gene DNA. There is also a strategy that utilises telomerase sequences to replicate DNA in bacterial cell free manner known as “doggybone” DNA. This is an interesting idea if it can be implemented into more widespread use. There would in this context be less sequence length of DNA to be incorporated upstream of P5 in a doggybone configuration, although still would not fully address the issue of the contamination initiation upstream of P5, as P5 would still be present and there would still be upstream sequences that led to the terminal “doggybone” end of DNA. It will be useful, in the future, to test this improved P5-HS setup in conjunction with other strategies that can potentially improve vector purity, such as minicircles, doggybone DNA or other future developments.

Another question that could be asked given the results of this thesis relates to plasmid purity as a factor in AAV contamination. Plasmid DNA can exist in either a covalently

closed circular (ccc) or open circle (oc) form depending on the level of DNA nicking present. Ideally for AAV production, plasmid with a high level of ccc DNA should be used. It is possible that nicked plasmid DNA could result in a higher contamination prep due to REP only having consensus binding and nicking sites at the P5 and ITR loci, thus pre-present nicks elsewhere along the plasmid would serve to complete any excision event. This DNA is likely either being excised without replicating or undergoing an incomplete replication event. Regardless of the mechanism, high consideration should be given towards eliminating variance between the production, storage and starting quality of plasmids used for AAV production in any setting

The discovery of equal activity of the reverse direction of the P5 promoter is an interesting facet of AAV biology. Firstly, it suggests that anti-P5 could have a time-sensitive role in the WT AAV replication cycle that is yet undescribed which could be further explored. Assessing the activity of the anti-P5 promoter in different cell types will also help to elucidate which gene therapy strategies have the most to be concerned about regarding off target transcriptional activity from P5 derived contaminants. Of interest would be assessing this activity in the retina.

The robust activity of anti-P5 detected in this study opens the possibility for the use of the P5 promoter in research. The P5 promoter is exceedingly short (145bp). Short promoters are exceedingly useful for incorporation into recombinant viral constructs where space is limited. There are many instances by which the expression of two genes is required. The typical way this is done from a single promoter is via the incorporation of IRES or 2A sequences.[532,533] These systems are incredibly useful but have downsides. With the IRES sequences the expression of the second gene is often far weaker than that of the first.[534]

The 2A system is much shorter (18-22 amino acids)[535] and therefore more suited to use in sequence length constrained settings. A problem with this system is the furin cleavage site from 2A sequences leaves a hanging sequence of 20-23 amino acids at the c-terminus of the first gene used in the construct, and an N-terminal proline on the second gene.[536] This could cause a problem to experiments involving proteins with highly conserved c or n-terminal regions. These issues could be avoided if a short bidirectional promoter with equivalent + and – activity could be characterised. The P5 sequence would be a candidate for this. However, the functionality of this promoter across cell types is poorly described and it may be that strong bidirectional activity of P5 is only observed in a limited manner. The strong bidirectionality activity was quantified with P5 in 293T cells, which contain Adenoviral helper genes E1a and E1b. E1a is known to play a role in the transcriptional activity of P5 and so it is possible that strong bidirectional promoter activity is diminished outside of this context. It is also possible that this activity from P5 is limited to tissue types permissive to AAV replication. Further work should focus on the comparative utility of P5/antiP5 to other bidirectional promoters of different strengths and sizes in different target cell contexts.[537,538]

Interestingly, the extent of P5 activity and characterisation in different human cell types has not been extensive, the same goes for the other AAV promoters. AAV has been shown to be widely distributed across tissues in humans.[227] It could be largely assumed that P19 and P40 activity would be retained in any cell types capable of replicating AAV. A systematic study of the cell type specificity of AAV promoters could reveal interesting insights into both AAV replication biology and provide useful information about which

tissue types are of greater concern with regards to the risk posed by any residual replication competent AAV particles transferred.

Whilst the P5-AAV2 REP system is used throughout AAV production, it has also been co-opted to aid in the production of another useful parvoviral gene therapy vector; Bocavirus. Bocavirus production also uses a transient transfection system similar to that of AAV and the P5 driven AAV2 REP is used in production of recombinant bocavirus.[539] The AAV2 ITRs are also often used in this system and therefore in future work it would be useful to assay whether bocaviral production is hampered by the same contamination profiles and indeed whether the P5-HS system could be implemented into the recombinant bocaviral production system to improve vector purity. The characterisation of contaminant DNA from other DNA based gene therapy vectors is also currently lacking. An important follow up to this study would be to look at adenoviral gene therapy preparations through a similar lens as AAV has been, especially considering recent advances that have expanded the utility of the adenoviral vector toolkit.[540,541]

Whilst in the deep sequencing analysis, contaminant reads in the producer plasmids against the reads that map to the expression cassette are quantified, the percentage of expression cassette reads may not represent a count of functional genomes. As is shown more markedly in the FVIII deep sequencing, there is a disparity between genome position and read counts. The read counts suggest there is more abundant representation of DNA towards the 3’ end of the recombinant genome. This bias has also been detected in other deep seq analysed preps and appears to be consistent. In FVIII, which is an oversized expression cassette this would make sense, the full genome cannot be packaged successfully and perhaps a shearing event occurs during the packaging

process, leading to a smearing of genomes when run on a gel.[480] Productive infection likely relies on recombination of these sequences post infection. Future studies focusing on expression cassette truncations should ask the question, what constitutes a functional AAV genome. Certainly, it does not appear that question is answered merely by the presence of the full cassette sequence. Future work with oversized cassettes could look at whether partial strands of single polarity (just + or just -) in an oversized cassette can provide any functionality in a combined setting, the hypothesis would be that they would not, and to look to strategies for either expanding the packaging capacity of AAV or to reduce the presence of smaller identity fragmented genomes and VP1 deficient particles from final AAV preps. Looking into differences in charge densities, or molecular weight differences between VP1 containing and deficient full particles could be one avenue to increasing not just apparent vector purity, but 'functional purity' of AAV preparations, and similar methods could be employed to reduce the so called 'empty' particle fractions that may contain small transgene fragments that increase the antigenic load without functionally contributing to the preparation.

The FDA states that the residual DNA from cell line fragments should have a median length less than 200bp.[1] The alkaline gel results in Chapter 4 from a contaminant only prep clearly show that product related impurities from the P5 promoter contain a substantial proportion of sequences that range in size up to the packaging capacity of around 5kb. In this case the identified DNA is more likely to originate from the P5 containing REPCAP plasmid than the producer cell line, however in any future study of the contribution of host cell DNA to AAV contamination, sequence length should also be

assessed. Whilst production cell line DNA is a lesser factor in terms of contamination contribution to AAV, identifying whether specific sequences within the production cell line genome were overrepresented in the final AAV prep is a pertinent point of future study. There are over 15,000 positions within the human genome that contain at least two adjacent GCTC tetranucleotide sequences.[447] Indeed some of these can be found near or within proto-oncogenes such as c-sis and the carcinoma associated marker TROP-2.[184] If the mere presence of the REP recognition sequences could facilitate incorporation into AAV then these sequences from the human genome could be incorporated into AAV from the human derived producer cell line. Our results suggest that more than just a REP recognition sequence is required and that a loop structure with an eligible REP cleavage site would be required for incorporation. It is likely that of these identified sequences within the human genome, some would meet these requirements.

Likewise, the Baculoviral infection of SF9 insect cells is another primary means of generating AAV for the clinic. A full genome search of spodoptera frugiperda SF9 isolate (WGS accession: NJHR01000000) reveals that there are 99 sites that contain a GCTC triplet and 8 sites that contain a GCTC quadruplet. It is likely that AAV REP will bind to these identified sequences. Interestingly, within these 107 sites there are 11 sequences that identically match the 14bp Rep binding site present in the AAV2 ITRs. The quadruplet sequences would be of particular interest to follow up on for REP binding strength and potential replication initiation because a study that identified an AAV integration site within the African Green Monkey genome showed that a GCTC quintuplet motif has a six fold higher REP68 binding affinity than the human AAVS1 site, although this was identified with the AAV4 REP and not AAV2.[542] In addition to this, early studies of the virus showed

that of a GCTC trimer binding site, the first and third repeats are the most important in retaining replicative potential.[543] GCTC/NNNN/GCTC sequences of the cell line genome should also be scanned in future studies of AAV production for potential incorporation into the virion. A tangible future direction towards improving the purity of the AAV vectors produced with both human derived cell lines and the baculoviral infection system would be to use a higher fidelity form of deep sequencing to quantify host cell genome packaging at these regions, and to develop producer cell lines that have those regions knocked out, examining whether equivalent production efficiency and higher AAV product purity could be achieved. An additional interesting facet of the contaminant deep sequencing data from AAV preps is that we did not observe even read coverage across the expression cassette. A consistent bias towards the terminal ends of the genome was observed, most markedly in the oversized FVIII cassette but also within some well packaged cassettes. A recent paper from the University of Massachusetts that was using deep sequencing to characterise truncated genomes resultant from sgRNA containing constructs also observed genome truncation at self-complementary ITRs,[464] and this has also been described by the same group in the context of single stranded shRNA containing constructs. Further study of the factors influencing the presence of truncated vector genome reads will be incredibly valuable to the field.

The information gained from this PhD could also help to further the investigation into the improvement and diversification of AAV helper sequences for production. As sequences from across DNA viral families have been shown to exhibit helper function for AAV replication, it is likely that this repertoire could be expanded as research in the area

develops, be that through discovery of novel helper sequences from other viruses, or the improvement of known helper sequences through artificial engineering strategies. By subjecting AAV preps produced with novel helper sequences to a deep sequencing protocol it can be ensured that any newly developed method does not lead to active incorporation of sequences other than the expression cassette.

6.3 Concluding statement:

This PhD study has achieved a far greater characterisation of the causes and effects of AAV DNA contamination than prior studies, overturning a well cited prior observation that contaminant AAV sequences were transcriptionally silent,[466] and the development of a universally implementable system that can produce AAV to a higher level of purity than would otherwise be achieved. The P5 spacer promoter system developed in the course of this study can be used to improve the purity of AAV genomes produced with any AAV serotype regardless of expression cassette and will be useful for both the improvement of research grade AAV for nonclinical applications and most importantly for the improvement of purity in AAV produced in a clinical grade setting.

Bibliography:

1. (FDA), U. S. F. and D. A. *FDA briefing document: vaccines and related biological products advisory committee meeting: cell lines derived from human tumors for vaccine manufacture*. (2012).
2. Sambrook, J., Westphal, H., Srinivasan, P. R. & Dulbecco, R. The integrated state of viral DNA in SV40-transformed cells. *Proc. Natl. Acad. Sci. U. S. A.* **60**, 1288–95 (1968).
3. Friedmann, T. & Roblin, R. Gene Therapy for Human Genetic Disease? *Science (80-.).* **175**, 949–956 (1972).
4. Rosenberg, S. A. *et al.* Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *N. Engl. J. Med.* **30**, 570–578 (1990).
5. Robinson, A. Gene therapy: The future touches down. *Can. Med. Assoc. J.* **150**, 377–380 (1994).
6. Dubé, D. & Cournoyer, D. Gene Therapy: Here to Stay. *Can. Med. Assoc. J.* **152**, 1605–1613 (1995).
7. Hacein-Bey-Abina, S. *et al.* Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N. Engl. J. Med.* **346**, 1185–1193 (2002).
8. Hacein-Bey-Abina, S. *et al.* A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* **348**, 255–256 (2003).
9. Howe, S. J. *et al.* Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients Find the latest version : Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following ge. *J. Clin. Invest.* **118**, 3143–3150 (2008).
10. Hacein-bey-abina, S. *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1 Find the latest version : Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J. Clin. Invest.* **118**, 3132–3142 (2008).
11. Maguire, A. M. *et al.* Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N. Engl. J. Med.* **358**, 2240–2248 (2008).
12. Nathwani, A. C. *et al.* Adenovirus-Associated Virus Vector–Mediated Gene Transfer in Hemophilia B. *N. Engl. J. Med.* **365**, 2357–2365 (2011).
13. Nathwani, A. C. *et al.* Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. *N. Engl. J. Med.* **371**, 1994–2004 (2014).

14. Zhou, S. *et al.* A self-inactivating lentiviral vector for SCID-X1 gene therapy that does not activate LMO2 expression in human T cells. *Blood* **116**, 900–909 (2019).

15. De Ravin, S. S. *et al.* Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.* **8**, 335ra57 (2016).

16. Mamcarz, E. *et al.* Lentiviral Gene Therapy Combined with Low-Dose Busulfan in Infants with SCID-X1. *N. Engl. J. Med.* **380**, 1525–1534 (2019).

17. Shahryari, A. *et al.* Development and Clinical Translation of Approved Gene Therapy Products for Genetic Disorders. *Front. Genet.* **10**, 1–25 (2019).

18. Goswami, R. *et al.* Gene Therapy Leaves a Vicious Cycle. *Front. Oncol.* **9**, 1–25 (2019).

19. Joyner, A., Keller, G., Phillips, R. A. & Bernstein, A. Retrovirus Transfer of a Bacterial Gene Into Mouse Haematopoetic Progenitor Cells. *Nature* **305**, 556–558 (1983).

20. Williams, D. A., Orkin, S. H. & Mulligan, R. C. Retrovirus-mediated transfer of human adenosine deaminase gene sequences into cells in culture and into murine hematopoietic cells in vivo ZDHFRip3HssF. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 2566–2570 (1986).

21. Chowdhury, J. R. *et al.* Long-Term Improvement of Hypercholesterolenia After ex Vivo Gene Therapy in LDLR-Deficient Rabbits. *Science (80-.).* **254**, 1802–5 (1991).

22. Grossman, M. *et al.* Successful Ex Vivo Gene Therapy Directed to Liver in a patient with Familial Hypercholesterolaemia. *Nat. Genet.* **6**, 335–341 (1994).

23. Raymon, H. K., Thode, S. & Gage, F. H. Application of Ex Vivo Gene Therapy in the Treatment of Parkinson's Disease. *Exp. Neurol.* **144**, 82–91 (1997).

24. Ridet, J. *et al.* Toward Autologous ex Vivo Gene Therapy for the Central Nervous System with Human Adult Astrocytes. *Hum. Gene Ther.* **10**, 271–280 (1999).

25. Imren, S. *et al.* Permanent and panerythroid correction of murine Beta Thalassemia by multiple lentiviral integration in hematopoietic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 14380–5 (2002).

26. Piacibello, W. *et al.* Lentiviral gene transfer and ex vivo expansion of human primitive stem cells capable of primary, secondary, and tertiary multilineage repopulation in NOD/SCID mice. *Blood* **100**, 4391–4400 (2002).

27. Milone, M. C. & O'Doherty, U. Clinical use of lentiviral vectors. *Leukemia* **32**, 1529–1541 (2018).

28. Nathwani, a C. *et al.* Efficient gene transfer into human cord blood CD34+ cells and the CD34+CD38- subset using highly purified recombinant adeno-associated viral vector preparations that are free of helper virus and wild-type AAV. *Gene Ther.* **7**, 183–95 (2000).

29. Hacein-Bey Abina, S. *et al.* Outcomes Following Gene Therapy in Patients With Severe Wiskott-Aldrich Syndrome. *Jama* **313**, 1550–63 (2015).

30. Payen, E. *et al.* Gene Therapy in a Patient with Sickle Cell Disease. *N. Engl. J. Med.* **376**, 848–855 (2017).

31. Picanço-Castro, V. *et al.* Establishment of a simple and efficient platform for car-t cell generation and expansion: from lentiviral production to in vivo studies. *Hematol. Transfus. cell Ther.* **2**, 150–158 (2019).

32. Jin, C. *et al.* Safe engineering of CAR T cells for adoptive cell therapy of cancer using long-term episomal gene transfer. *EMBO Mol. Med.* **8**, 702–711 (2016).

33. Wiesinger, M. *et al.* Clinical-Scale Production of CAR-T Cells for the Treatment of Melanoma Patients by mRNA Transfection of a CSPG4-Specific CAR under Full GMP Compliance. *Cancers (Basel)* **11**, 1–18 (2019).

34. Wang, K., Wei, G. & Liu, D. CD19: a biomarker for B cell development , lymphoma diagnosis and therapy. *Exp. Hematol. Oncol.* **1**, 36 (2012).

35. Grupp, S. A. *et al.* Chimeric Antigen Receptor–Modified T Cells for Acute Lymphoid Leukemia. *N. Engl. J. Med.* **368**, 1509–1518 (2013).

36. Garfell, A. L. *et al.* Chimeric Antigen Receptor T Cells against CD19 for Multiple Myeloma. *N. Engl. J. Med.* **373**, 1040–1047 (2015).

37. Porter, D. L., Levine, B., Kalos, M., Bagg, A. & June, C. H. Chimeric Antigen Receptor–Modified T Cells in Chronic Lymphoid Leukemia. *N. Engl. J. Med.* **365**, 725–733 (2011).

38. Fry, T. J. *et al.* CD22-targeted CAR T cells induce remission in B-ALL that is naive or resistant to CD19-targeted. *Nat. Med.* **24**, 20–28 (2018).

39. Pule, M. A. *et al.* Virus-specific T cells engineered to coexpress tumor- specific receptors: persistence and antitumor activity in individuals with neuroblastoma. *Nat. Med.* **14**, 1264–1270 (2008).

40. Raje, N. *et al.* Anti-BCMA CAR T-Cell Therapy bb2121 in Relapsed or Refractory Multiple Myeloma. *N. Engl. J. Med.* **380**, 1726–1737 (2019).

41. Ahmed, N. *et al.* HER2-Specific Chimeric Antigen Receptor–Modified Virus-Specific T Cells for Progressive Glioblastoma A Phase 1 Dose-Escalation Trial. *JAMA Oncol.* **3**, 1094–1101 (2020).

42. Yu, J. X., Hubbard, V. M. & Tang, J. The global pipeline of cell therapies for cancer. *Nat. Rev. Drug Discov.* **18**, 821–822 (2019).

43. Martinez, M. & Moon, E. K. CAR T Cells for Solid Tumors: New Strategies for Finding, Infiltrating, and Surviving in the Tumor Microenvironment. *Front. Immunol.* **10**, 1–21 (2019).

44. Pfendner, E. G. & Lucky, A. W. Junctional Epidermolysis Bullosa. in *Gene Reviews* (eds. Adam, M. P., Ardinger, H. H., Pagon, R. A. & Wallace, S. E.) 1–30

(University of Washington, Seattle, 2008).

45. Hirsch, T. *et al.* Regeneration of the entire human epidermis using transgenic stem cells. *Nature* **551**, 327–332 (2017).

46. Turner, J. C. & Mulliken, B. Parasitization of Mouse Sarcoma 180 by Vaccine Virus and Its Effect on Tumor Growth. *Cancer Res.* **7**, 774–779 (1947).

47. Turner, J. C. & Mulliken, B. Effects of Intravenous Vaccinia in mice with Sarcoma 180 or Leukemia 9417. *Cancer* **3**, 354–360 (1950).

48. Martuza, R. L., Malick, A. M. Y., Markertt, J. M., Ruffner, K. L. & Coen, D. M. Experimental Therapy of Human Glioma by Means of a Genetically Engineered Virus Mutant. *Science (80-.).* **252**, 10–13 (1991).

49. Trager, M. H., Geskin, L. J. & Saenger, Y. M. Oncolytic Viruses for the Treatment of Metastatic Melanoma. *Curr. Treat. Options Oncol.* **21**, 26 (2020).

50. Lawler, S. E., Speranza, M.-C., Cho, C.-F. & Chiocca, E. A. Oncolytic Viruses in Cancer Treatment. *JAMA Oncol.* **3**, 841–849 (2020).

51. Conry, R. M., Westbrook, B., Mckee, S. & Norwood, T. G. Talimogene laherparepvec: First in class oncolytic virotherapy. *Hum. Vaccin. Immunother.* **14**, 839–846 (2018).

52. Westbrook, B. C. *et al.* Talimogene laherparepvec induces durable response of regionally advanced Merkel cell carcinoma in 4 consecutive patients. *JAAD Case Reports* **5**, 782–786 (2019).

53. Kelly, C. M. *et al.* Objective Response Rate Among Patients With Locally Advanced or Metastatic Sarcoma Treated With Talimogene Laherparepvec in Combination With Pembrolizumab A Phase 2 Clinical Trial. *JAMA Oncol.* **6**, 402–408 (2020).

54. Kajon, A. E., Weinberg, J. B. & Spindler, K. R. Adenoviruses. in *Reference Module in Biomedical Sciences* 1–15 (2019). doi:10.1016/B978-0-12-801238-3.00086-6

55. Zabner, J., Richards, S. M. & Standaert, T. A. Repeat administration of an adenovirus vector encoding cystic fibrosis transmembrane conductance regulator to the nasal epithelium of patients with cystic fibrosis. *J. Clin. Invest.* **97**, 1504–1511 (1996).

56. Crystal, R. G. *et al.* Administration of an adenovirus containing the human CFTR cDNA to the respiratory tract of individuals with cystic fibrosis. *Nat. Genet.* **8**, 42–51 (1994).

57. Rosengart, T. K., Lee, L. Y., Patel, S. R., Kligfield, P. D. & Okin, P. M. Six-Month Assessment of a Phase I Trial of Angiogenic Gene Therapy for the Treatment of Coronary Artery Disease Using Direct Intramyocardial Administration of an Adenovirus Vector Expressing the VEGF121 cDNA. *Ann. Surg.* **230**, 466–472 (1999).

58. Wilson, J. M. Adenoviruses as gene-delivery vehicles. *N. Engl. J. Med.* **334**, 1185–1187 (1996).

59. Lehrman, S. Virus treatment questioned after gene therapy death. *Nature* **401**, 517–518 (1999).

60. Raper, S. E. *et al.* Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Mol. Genet. Metab.* **80**, 148–158 (2003).

61. Capasso, C., Garofalo, M., Hirvinen, M. & Cerullo, V. The Evolution of Adenoviral Vectors through Genetic and Chemical Surface Modifications. *Viruses* **6**, 832–855 (2014).

62. Alba, R., Bosch, A. & Chillon, M. Gutless adenovirus : last-generation adenovirus for gene therapy. *Gene Ther.* **12**, 18–27 (2005).

63. Wold, W. S. M. & Toth, K. Adenovirus Vectors for Gene Therapy, Vaccination and Cancer Gene Therapy. *Curr. Gene Ther.* **13**, 421–433 (2015).

64. Grieger, J. C. & Samulski, R. J. Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. *J. Virol.* **79**, 9933–44 (2005).

65. Atchison, R., Casto, B. & Hammon, W. Adenovirus-Associated defective virus particles. *Science (80-.).* **149**, 754–756 (1965).

66. Penaud-Budloo, M. *et al.* Adeno-associated virus vector genomes persist as episomal chromatin in primate muscle. *Journal of virology* **82**, 7875–7885 (2008).

67. Aitken, M. L. *et al.* A Phase I Study of Aerosolized Administration of tgAAVCF to Cystic Fibrosis Subjects with Mild Lung Disease. *Hum. Gene Ther.* **12**, 1907–1916 (2001).

68. Nathwani, A. C. *et al.* Adeno-Associated Virus Vector-Mediated Gene Transfer in Hemophilia B. *N. Engl. J. Med.* **365**, 2357–2365 (2011).

69. Mendell, J. R. *et al.* Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N. Engl. J. Med.* **377**, 1713–1722 (2017).

70. Yan, Z., Jr, P. B. M. & Engelhardt, J. F. Advances in gene therapy for cystic fibrosis lung disease. *Hum. Mol. Genet.* **28**, 88–94 (2019).

71. Chen, J., Guo, Z., Tian, H. & Chen, X. Production and clinical development of nanoparticles for gene delivery. *Mol. Ther. - Methods Clin. Dev.* **3**, 16023 (2016).

72. Carvalho, A. M., Cordeiro, R. A. & Faneca, H. Silica-Based Gene Delivery Systems : From Design to Therapeutic Applications. *Pharmaceutics* **12**, E649 (2020).

73. Rai, R., Alwani, S. & Badea, I. Polymeric Nanoparticles in Gene Therapy: New Avenues of Design and Optimization for Delivery Applications. *Polym.* **11**, 1–35 (2019).

74. Hardee, C. L., Arévalo-Soliz, L. M., Hornstein, B. D. & Zechiedrich, L. Advances in Non-Viral DNA Vectors for Gene Therapy. *Genes (Basel).* **8**, 1–22 (2017).

75. Liang, X., Liu, L., Wei, Y., Gao, G. & Wei, X. Clinical Evaluations of Toxicity and Efficacy of Nanoparticle-Mediated Gene Therapy. *Hum. Gene Ther.* **29**, 1227–1234 (2018).

76. Berntorp, E. & Shapiro, A. D. Modern haemophilia care. *Lancet* **379**, 1447–1456 (2012).

77. Bolton-maggs, P. H. B. & Pasi, K. J. Haemophilias A and B. *Lancet* **361**, 1801–1809 (2003).

78. Wheeler, A. P. & Gailani, D. Why Factor XI Deficiency is a Clinical Concern. *Expert Rev. Hematol.* **9**, 629–637 (2017).

79. Bolton-maggs, P. H. B. The Rare Inherited Coagulation Disorders. *Paediatr. Blood Cancer* **60**, S37–S40 (2013).

80. Dijk, K. Van, Fischer, K., Van Der Bom, J. G., Grobbee, D. E. & Van Den Berg, H. M. Variability in clinical phenotype of severe haemophilia: the role of the first joint bleed. *Haemophilia* **11**, 438–443 (2005).

81. Iorio, A., Stonebraker, J. S., Makris, M. & Coffin, D. Establishing the Prevalence and Prevalence at Birth of Hemophilia in Males A Meta-analytic Approach Using National Registries. *Ann. Intern. Med.* **171**, 540–546 (2019).

82. Aledort, L. M. History of haemophilia. *Haemophilia* **13**, 1–2 (2007).

83. Ikkala, E. *et al.* Changes in the life expectancy of patients with severe haemophilia A in Finland in 1930-79. *Br. J. Haematol.* **52**, 7–12 (1982).

84. Farr, A. D. Treatment of haemophilia by transfusion: the first recorded case. *J. R. Soc. Med.* **74**, 301–305 (1981).

85. Cohn, E. J. *et al.* Preparation and Properties of Serum and Plasma Proteins . IV . A System for the Separation into Fractions of the Protein and Lipoprotein Components of Biological Tissues and Fluids. *J. Am. Chem. Soc.* **68**, 459–75 (1946).

86. Pool, J. G., Gershgold, E. J. & Pappenhagen, A. R. High-potency Antihaemophilic Factor Concen- trate prepared from Cryoglobulin Precipitate. *Nature* **203**, 312 (1964).

87. Jason, J. *et al.* Human T-lymphotropic retrovirus type III/lymphadenopathy-associated virus antibody. Association with hemophiliacs' immune status and blood component usage. *JAMA* **253**, 3409–3015 (1985).

88. Evatt, B. L., Gomperts, E. D., McDougal, J. S. & Ramsey, R. B. Coincidental Appearance of LAV/HTLV-III Antibodies in Hemophiliacs and the Onset of the AIDS Epidemic. *N. Engl. J. Med.* **312**, 483–486 (1985).

89. Martin, D. *et al.* Expression of active human factor VIII from recombinant DNA

clones. *Nature* **312**, 330–7 (1984).

90. Kurachi, K. & Davie, E. W. Isolation and characterization of a cDNA coding for human factor IX. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 6461–6464 (1982).

91. Anson, D. S., Austent, D. E. G. & Brownlee, G. G. Expression of active human clotting factor IX from recombinant DNA clones in mammalian cells. *Nature* **315**, 683–5 (1985).

92. World Federation of Hemophilia. World Federation of Hemophilia Statement on Cryoprecipitate for Treatment of Congenital Bleeding Disorders. (2019).

93. Poon, M. & Lee, A. Individualized prophylaxis for optimizing hemophilia care: can we apply this to both developed and developing nations? *Thromb. J.* **14(Suppl 1**, 65–71 (2016).

94. Manco-Johnson, M. *et al.* Prophylaxis versus episodic treatment to prevent joint disease in boys with severe hemophilia. *N. Engl. J. Med.* **357**, 535–544 (2007).

95. Nilsson, I. M., Berntorp, E., Lofqvist, T. & Petterssont, H. Twenty-five years' experience of prophylactic treatment in severe haemophilia A and B. *J. Intern. Med.* **232**, 25–32 (1992).

96. Hara, J. O. *et al.* The cost of severe haemophilia in Europe: the CHESS study. *Orphanet J. Rare Dis.* **12**, 1–8 (2017).

97. Guh, S., Grosse, S. D., McAlister, S., Kessler, C. M. & Soucie, J. M. Healthcare expenditures for males with haemophilia and employer-sponsored insurance in the United States, 2008. *Haemophilia* **18**, 268–275 (2012).

98. Schrijvers, L. H., Uitslager, N., Schuurmans, M. J. & Fischer, K. Barriers and motivators of adherence to prophylactic treatment in haemophilia: A systematic review. *Haemophilia* **19**, 355–361 (2013).

99. Thornburg, C. D. & Duncan, N. A. Treatment adherence in hemophilia. *Patient Prefer. Adherence* **11**, 1677–1686 (2017).

100. Pierce, G. F. *et al.* First-year results of an expanded humanitarian aid programme for haemophilia in resource-constrained countries.pdf. *Haemophilia* **24**, 229–235 (2018).

101. Meeks, S. L. & Batsuli, G. Hemophilia and inhibitors : current treatment options and potential new therapeutic approaches. *Am. Soc. Hematol. Educ. Progr.* 657–662 (2016).

102. Kempton, C. L. & Meeks, S. L. Toward optimal therapy for inhibitors in hemophilia. *Blood* **124**, 3365–3372 (2014).

103. Walsh, C. E., Soucie, J. M., Miller, C. H., Center, T. U. S. H. T. & Network. Impact of inhibitors on hemophilia A mortality in the United States. *Am. J. Hematol.* **90**, 400–405 (2015).

104. Coppola, A. *et al.* Treatment of hemophilia: a review of current advances and

ongoing issues. *J. Blood Med.* **1**, 183–195 (2010).

105. Chowdary, P. Extended half-life recombinant products in haemophilia clinical practice – Expectations, opportunities and challenges. *Thromb. Res.* 1–9 (2019). doi:10.1016/j.thromres.2019.12.012

106. Metzner, H. J., Pipe, S. W., Weimer, T. & Schulte, S. Extending the pharmacokinetic half-life of coagulation factors by fusion to recombinant albumin. *Thromb. Haemost.* **110**, 931–939 (2013).

107. Kitazawa, T. *et al.* Factor VIIIa-mimetic cofactor activity of a bispecific antibody to factors IX / IXa and X / Xa , emicizumab , depends on its ability to bridge the antigens. 1348–1357 (2017).

108. Pasi, K. J. *et al.* Targeting of Antithrombin in Hemophilia A or B with RNAi Therapy. *N. Engl. J. Med.* **377**, 819–828 (2017).

109. Kay, M. A. *et al.* Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nat. Genet.* **24**, 257–261 (2000).

110. Jiang, H. *et al.* Evidence of Multiyear Factor IX Expression by AAV-Mediated Gene Transfer to Skeletal Muscle in an Individual with Severe Hemophilia B. *Mol. Ther.* **14**, 452–455 (2006).

111. 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.* **12**, 342–7 (2006).

112. Nathwani, A. C. *et al.* Safe and efficient transduction of the liver after peripheral vein infusion of self-complementary AAV vector results in stable therapeutic expression of human FIX in nonhuman primates Safe and efficient transduction of the liver after peripheral vein infu. 1414–1421 (2014). doi:10.1182/blood-2006-03-010181

113. Nathwani, A. C. *et al.* Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. *N Engl J Med* **21371**, 1994–2004 (2014).

114. Gao, G. *et al.* Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11854–11859 (2002).

115. Pipe, S., Leebeek, F. W. G., Ferreira, V., Sawyer, E. K. & Pasi, J. Clinical Considerations for Capsid Choice in the Development of Liver-Targeted AAV-Based Gene Transfer. *Mol. Ther. Methods Clin. Dev.* **15**, 170–178 (2019).

116. Nakai, H. *et al.* Unrestricted hepatocyte transduction with adeno-associated virus serotype 8 vectors in mice. *J. Virol.* **79**, 214–224 (2005).

117. Nathwani, A. C. *et al.* Safe and efficient transduction of the liver after peripheral vein infusion of self-complementary AAV vector results in stable therapeutic expression of human FIX in nonhuman primates. *Blood* **109**, 1414–1421 (2007).

118. Thomas, C. E., Storm, T. a, Huang, Z. & Kay, M. a. Rapid uncoating of vector genomes is the key to efficient liver transduction with pseudotyped adeno-associated virus vectors. *J. Virol.* **78**, 3110–3122 (2004).

119. Blacklow, N. R., Hoggan, M. D., Kapikan, A. Z., Austin, J. B. & Rowe, W. P. Epidemiology of Adenovirus-Associated Virus Infection in a nursery population. *Am. J. Epidemiol.* **88**, 368–378 (1968).

120. Boutin, S. *et al.* Prevalence of Serum IgG and Neutralizing Factors against Adeno-Associated Virus (AAV) Types 1,2,5,6,8, and 9 in the Healthy Population : Implications for Gene Therapy Using AAV Vectors. *Hum. Gene Ther.* **712**, 704–712 (2010).

121. Calcedo, R., Vandenberghe, L. H., Gao, G., Lin, J. & Wilson, J. M. Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses. *J. Infect. Dis.* **199**, 381–390 (2009).

122. Calcedo, R. *et al.* Adeno-associated virus antibody profiles in newborns, children, and adolescents. *Clin. Vaccine Immunol.* **18**, 1586–1588 (2011).

123. Nathwani, A. *et al.* Self-complementary adeno-associated virus vectors containing a novel liver- specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver. *Blood* **107**, 2653–2661 (2006).

124. Graham, J. B. *et al.* The Malmo Polymorphism of Coagulation Factor IX , An Immunologic Polymorphism Due to Dimorphism of Residue 148 That Is in Linkage Disequilibrium with Two Other F.IX Polymorphisms. *Am. J. Hum. Genet.* **42**, 573–580 (1988).

125. McCarty, D. M., Monahan, P. E. & Samulski, R. J. Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther.* **8**, 1248–1254 (2001).

126. Matsushita, T. *et al.* Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Ther.* **5**, 938–945 (1998).

127. Louis, N., Evelegh, C. & Graham, F. L. Cloning and Sequencing of the Cellular – Viral Junctions from the Human Adenovirus Type 5 Transformed 293 Cell Line. *Virology* **233**, 423–429 (1997).

128. Allay, J. A. *et al.* Good manufacturing practice production of self-complementary serotype 8 adeno-associated viral vector for a hemophilia B clinical trial. *Hum. Gene Ther.* **22**, 595–604 (2011).

129. Pasi, J. K. *et al.* Multiyear Follow-up of AAV5-hFVIII-SQ Gene Therapy for Hemophilia A. *N. Engl. J. Med.* **382**, 29–40 (2020).

130. Doshi, B. S. & Arruda, V. R. Gene therapy for hemophilia: what does the future hold? *Ther. Adv. Hematol.* **9**, 273–293 (2018).

131. Samulski, R. J. Modified factor viii and factor ix genes and vectors for gene

therapy. *US2010O284971A1* (2010).

132. Herzog, R. W. Hemophilia Gene Therapy: Caught Between a Cure and an Immune Response. *Mol. Ther.* **23**, 1411–1412 (2015).

133. Xiang, Z. *et al.* The Effect of CpG Sequences on Capsid-Specific CD8+ T Cell Responses to AAV Vector Gene Transfer. *Mol. Ther.* **28**, 1–13 (2020).

134. Wright, J. F. Human gene therapy - immunogenicity perspective. *Mednous* 26–27 (2019).

135. Simioni, P. *et al.* X-linked thrombophilia with a mutant factor IX (factor IX Padua). *N. Engl. J. Med.* **361**, 1671–1675 (2009).

136. Samelson-Jones, B. J., Finn, J. D., George, L. A., Camire, R. M. & Arruda, V. R. Hyperactivity of factor IX Padua (R338L) depends on factor VIIIa cofactor activity. *JCI Insight* **5**, e128683 (2019).

137. Nair, N. *et al.* Computationally designed liver-specific transcriptional modules and hyperactive factor IX improve hepatic gene therapy. *Blood* **123**, 3195–3199 (2014).

138. Cantore, A. *et al.* Hyperfunctional coagulation factor IX improves the efficacy of gene therapy in hemophilic mice. *Blood* **120**, 4517–4521 (2015).

139. Spronck, E. A. *et al.* Enhanced Factor IX Activity following Administration of AAV5-R338L “ Padua ” Factor IX versus AAV5 WT Human Factor IX in NHPs. *Mol. Ther. Methods Clin. Dev.* **15**, 221–231 (2019).

140. George, L. A. *et al.* Hemophilia B Gene Therapy with a High-Specifc-Activity Factor IX Variant. *N. Engl. J. Med.* **317**, 2215–2227 (2017).

141. Quellec, S. Le *et al.* Recombinant Adeno-Associated Viral Vectors Expressing Human Coagulation FIX-E456H Variant in Hemophilia B Mice. *Thromb. Haemost.* **119**, 1956–1967 (2019).

142. Kao, C. *et al.* FIX-Triple, a gain-of-function factor IX variant, improves haemostasis in mouse models without increased risk of thrombosis. *Thromb. Haemost.* **104**, 355–365 (2010).

143. Sabatino, D. E. *et al.* Efficacy and Safety of Long-term Prophylaxis in Severe Hemophilia A Dogs Following Liver Gene Therapy Using AAV Vectors. *Mol. Ther.* **19**, 442–449 (2011).

144. Crudele, J. M. *et al.* AAV liver expression of FIX-Padua prevents and eradicates FIX inhibitor without increasing thrombogenicity in hemophilia B dogs and mice. *Blood* **125**, 1553–1562 (2015).

145. Arruda, V. R. & Samelson-Jones, B. J. Gene therapy for immune tolerance induction in hemophilia with inhibitors. *J. Thromb. Haemost.* **14**, 1121–1134 (2016).

146. Cantore, A. *et al.* Liver-directed lentiviral gene therapy in a dog model of

hemophilia B. **7**, 1–12 (2015).

147. Merlin, S. *et al.* FVIII expression by its native promoter sustains long-term correction avoiding immune response in hemophilic mice. *Blood Adv.* **3**, 825–838 (2019).

148. Suwanmanee, T. *et al.* Integration-deficient Lentiviral Vectors Expressing Codon-optimized R338L Human FIX Restore Normal Hemostasis in Hemophilia B Mice. *Mol. Ther.* **22**, 567–574 (2014).

149. Shi, Q. *et al.* Platelet gene therapy corrects the hemophilic phenotype in immunocompromised hemophilia A mice transplanted with genetically manipulated human cord blood stem cells. *Blood* **123**, 395–403 (2014).

150. Zakas, P. M. *et al.* Enhancing the pharmaceutical properties of protein drugs by ancestral sequence reconstruction. *Nat. Biotechnol.* **35**, 35–37 (2017).

151. Labow, M. A., Hermonat, P. L. & Berns, K. I. Positive and Negative Autoregulation of the Adeno-Associated Virus Type 2 Genome. *J. Virol.* **60**, 251–258 (1986).

152. Mendelson, E., Trempe, J. P. & Carter, B. J. Identification of the trans-Acting Rep Proteins of Adeno-Associated Virus by Antibodies to a Synthetic Oligopeptide. *J. Virol.* **60**, 823–832 (1986).

153. Sonntag, F. *et al.* The Assembly-Activating Protein Promotes Capsid Assembly of Different Adeno-Associated Virus Serotypes ▿ .pdf. *J. Virol.* **85**, 12686–97 (2011).

154. Cao, M., You, H. & Hermonat, P. L. The X Gene of Adeno-Associated Virus 2 (AAV2) Is Involved in Viral DNA Replication. *Plos* **9**, e104596 (2014).

155. Ogden, P. J., Kelsic, E. D., Sinai, S. & Church, G. M. Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science (80-.).* **366**, 1139–1143 (2019).

156. Im, D.-S. & Muzyczka, N. Partial Purification of Adeno-Associated Virus Rep78, Rep52, and Rep40 and Their Biochemical Characterization. *J. Virol.* **66**, 1119–1128 (1992).

157. Urabe, M. *et al.* Charged-to-Alanine Scanning Mutagenesis of the N-Terminal Half of Adeno-Associated Virus Type 2 Rep78 Protein. *J. Virol* **73**, 2682–2693 (1999).

158. Walker, S. L., Wonderling, R. S. & Owens, R. A. Mutational Analysis of the Adeno-Associated Virus Type 2 Rep68 Protein Helicase Motifs. *J. Virol.* **71**, 6996–7004 (1997).

159. Trempe, J. P. & Carter, B. J. Alternate mRNA Splicing Is Required for Synthesis of Adeno-Associated Virus VP1 Capsid Protein. *J. Virol.* **62**, 3356–3363 (1988).

160. Ni, T., Zhou, X., Mccarty, D. M., Zolotukhin, I. & Muzyczka, N. In Vitro Replication of Adeno-Associated Virus DNA. *J. Virol.* **68**, 1128–1138 (1994).

161. Hölscher, C., Kleinschmidt, J. A. & Bürkle, A. High-Level Expression of Adeno-Associated Virus (AAV) Rep78 or Rep68 Protein Is Sufficient for Infectious-

Particle Formation by a rep-Negative AAV Mutant. *J. Virol.* **69**, 6880–6885 (1995).

162. Owens, R. A., Weitzman, M. D., Kyostio, S. R. M. & Carter, B. J. Identification of a DNA-Binding Domain in the Amino Terminus of Adeno-Associated Virus Rep Proteins. *J. Virol.* **67**, 997–1005 (1993).

163. Ward, P., Elias, P. & Linden, R. M. Rescue of the Adeno-Associated Virus Genome from a Plasmid Vector: Evidence for Rescue by Replication. *J. Virol.* **77**, 11480–11490 (2003).

164. Wang, X. S. & Srivastava, A. A novel terminal resolution-like site in the adeno-associated virus type 2 genome. *J Virol* **71**, 1140–1146 (1997).

165. Im, D. & Muzyczka, N. The AAV Origin Binding Protein Rep88 Is an ATP-Dependent Site-SpeMic Endcm with DNA Helicase Activity. *Cell* **61**, 447–457 (1990).

166. Smith, R. H., Spano, A. J., Kotin, R. M., Al, S. E. T. & Irol, J. V. The Rep78 Gene Product of Adeno-Associated Virus (AAV) Self-Associates To Form a Hexameric Complex in the Presence of AAV ori Sequences. *J. Virol.* **71**, 4461–4471 (1997).

167. Dignam, S. S., Correia, J. J., Nada, S. E., Trempe, J. P. & Dignam, J. D. Activation of the ATPase Activity of Adeno-Associated Virus Rep68 and Rep78. *Biochemistry* **46**, 6364–6374 (2007).

168. Mansilla-soto, J. *et al.* DNA Structure Modulates the Oligomerization Properties of the AAV Initiator Protein Rep68. *PLoS Pathog.* **5**, e1000513 (2009).

169. Weitzman, M. D., Kyöstiö, S. R. M., Kotin, R. M. & Owens, R. A. Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 5808–5812 (1994).

170. Feng, D. *et al.* A 16 bp Rep Binding Element is Sufficient for Mediating Rep-dependent Integration into AAVS1. 38–45 (2006). doi:10.1016/j.jmb.2006.01.029

171. Davis, M. D., Wu, J. & Owens, R. A. Mutational Analysis of Adeno-Associated Virus Type 2 Rep68 Protein Endonuclease Activity on Partially Single-Stranded Substrates. *J. Virol.* **74**, 2936–2942 (2000).

172. Chiorini, J. A. *et al.* Biologically Active Rep Proteins of Adeno-Associated Virus Type 2 Produced as Fusion Proteins in Escherichia coli. *J. Virol.* **68**, 797–804 (1994).

173. Cassell, G. D. & Weitzman, M. D. Characterization of a nuclear localization signal in the C-terminus of the adeno-associated virus Rep68/78 proteins. *Virology* **327**, 206–214 (2004).

174. Kyostio, S. R. M. *et al.* Analysis of Adeno-Associated Virus (AAV) Wild-Type and Mutant Rep Proteins for Their Abilities To Negatively Regulate AAV P5 and P19 mRNA Levels. *J. Virol.* **68**, 2947–2957 (1994).

175. Lackner, D. F. & Muzyczka, N. Studies of the Mechanism of Transactivation of the Adeno-Associated Virus p19 Promoter by Rep Protein. *J. Virol.* **76**, 8225–8235 (2002).

176. Hermonat, P. L., Santin, A. D., Batchu, R. B. & Zhan, D. The Adeno-Associated Virus Rep78 Major Regulatory Protein Binds the Cellular TATA-Binding Protein in Vitro and in Vivo. *Virology* **245**, 120–127 (1998).

177. François, A. *et al.* The Cellular TATA Binding Protein Is Required for Rep-Dependent Replication of a Minimal Adeno-Associated Virus Type 2 p5 Element. *J. Virol.* **79**, 11082–11094 (2005).

178. Hermonat, P. L., Santin, A. D. & Batchu, R. B. The Adeno-associated Virus Rep78 Major Regulatory/Transformation Suppressor Protein Binds Cellular Sp1 in Vitro and Evidence of a Biological Effect. *Cancer Res.* **56**, 5299–304 (1996).

179. Weger, S., Wendland, M., Kleinschmidt, J. A. & Heilbronn, R. The Adeno-Associated Virus Type 2 Regulatory Proteins Rep78 and Rep68 Interact with the Transcriptional Coactivator PC4. *J. Virol.* **73**, 260–269 (1999).

180. Bevington, J. M. *et al.* Adeno-associated virus interactions with B23/Nucleophosmin: Identification of sub-nucleolar virion regions. *Virology* **357**, 102–113 (2007).

181. Weger, S., Hammer, E., Götz, A. & Heilbronn, R. Identification of a cytoplasmic interaction partner of the large regulatory proteins Rep78 / Rep68 of adeno-associated virus type 2 (AAV-2). *Virology* **362**, 192–206 (2007).

182. Wonderling, R. S. & Owens, R. A. Binding sites for adeno-associated virus Rep proteins within the human genome. *J. Virol.* **71**, 2528–2534 (1997).

183. Hermonat, P. L. Down-regulation of the human c-fos and c-myc proto-oncogene promoters by adeno-associated virus Rep78. *Cancer Lett.* **81**, 129–136 (1994).

184. Wonderling, R. S. & Owens, R. A. The Rep68 Protein of Adeno-Associated Virus Type 2 Stimulates Expression of the Platelet-Derived Growth Factor B c-sis Proto-Oncogene. *J. Virol.* **70**, 4783–4786 (1996).

185. Batchu, R. B., Shammas, M. A., Wang, J. Y. & Munshi, N. C. Dual Level Inhibition of E2F-1 Activity by Adeno-associated Virus Rep78. *J. Biol. Chem.* **276**, 24315–24322 (2001).

186. Batchu, R. B., Shammas, M. A., Wang, J. Y. & Munshi, N. C. Interaction of Adeno-associated Virus Rep78 with p53: Implications in Growth Inhibition. *Cancer Res.* **59**, 3592–3595 (1999).

187. Saudan, P., Vlach, J. & Beard, P. Inhibition of S-phase progression by adeno-associated virus Rep78 protein is mediated by hypophosphorylated pRb. *EMBO J.* **19**, 4351–61 (2000).

188. Schmidt, M., Afione, S. & Kotin, R. M. Adeno-Associated Virus Type 2 Rep78 Induces Apoptosis through Caspase Activation Independently of p53. *J. Virol.* **74**,

9441–9450 (2000).

189. Zhou, C., Yang, Q. & Trempe, J. P. Enhancement of UV-induced cytotoxicity by the adeno-associated virus replication proteins. *Biochim. Biophys. Acta* **1444**, 371–383 (1999).

190. Hermonat, P. L. Adeno-associated Virus Inhibits Human Papillomavirus Type 16: A Viral Interaction Implicated in Cervical Cancer. *Cancer Res.* **54**, 2278–2282 (1994).

191. Hermonat, P. L. Inhibition of Bovine Papillomavirus Plasmid DNA Replication by Adeno-Associated Virus. *Virology* **189**, 329–333 (1992).

192. Liu, T. *et al.* Adeno-associated virus Rep78 protein inhibits Hepatitis B virus replication through regulation of the HBV core promoter. *Biochem. Biophys. Res. Commun.* **385**, 106–111 (2009).

193. Antoni, B. A. *et al.* Adeno-Associated Virus Rep Protein Inhibits Human Immunodeficiency Virus Type 1 Production in Human Cells. *J. Virol.* **65**, 396–404 (1991).

194. Timpe, J. M., Verrill, K. C. & Trempe, J. P. Effects of Adeno-Associated Virus on Adenovirus Replication and Gene Expression during Coinfection. *J. Virol.* **80**, 7807–7815 (2006).

195. Glauser, D. L. *et al.* Inhibition of Herpes Simplex Virus Type 1 Replication by Adeno-Associated Virus Rep Proteins Depends on Their Combined DNA-Binding and ATPase/Helicase Activities. *J. Virol.* **84**, 3808–3824 (2010).

196. Surosky, R. T. *et al.* Adeno-Associated Virus Rep Proteins Target DNA Sequences to a Unique Locus in the Human Genome. *J. Virol.* **71**, 7951–7959 (1997).

197. Smith, R. H. & Kotin, R. M. The Rep52 Gene Product of Adeno-Associated Virus Is a DNA Helicase with 3′-to-5′ Polarity. *J. Virol.* **72**, 4874–4881 (1998).

198. Maggin, J. E., James, J. A., Chappie, J. S., Dyda, F. & Hickman, A. B. The Amino Acid Linker between the Endonuclease and Helicase Domains of Adeno-Associated Virus Type 5 Rep Plays a Critical Role in DNA-Dependent Oligomerization. *J. Virol.* **86**, 3337–3346 (2012).

199. King, J. A., Dubielzig, R., Grimm, D. & Kleinschmidt, J. A. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J.* **20**, 3282–3291 (2001).

200. Collaco, R. F., Kalman-maltese, V., Smith, A. D., Dignam, J. D. & Trempe, J. P. A Biochemical Characterization of the Adeno-associated Virus Rep40 Helicase. *J. Biol. Chem.* **278**, 34011–34017 (2003).

201. Chejanovsky, N. & Carter, B. J. Mutagenesis of an AUG Codon in the Adeno-Associated Effects on Viral DNA Replication. *Virology* **173**, 120–128 (1989).

202. Needham, P. G. *et al.* Adeno-Associated Virus Rep Protein-Mediated Inhibition of Transcription of the Adenovirus Major Late Promoter In Vitro. *J. Virol.* **80**, 6207–6217 (2006).

203. Yun, A. *et al.* Characterization of the gene expression pro fi le of human bocavirus. *Virology* **403**, 145–154 (2010).

204. Ninth Report of the International Committee on Taxonomy of Viruses. Family - Parvoviridae. in *Virus Taxonomy* 405–425 (2012). doi:10.1016/B978-0-12-384684-6.00039-2

205. Rose, J. A., Maizel, J. V., Inman, J. K. & Shatkin, A. . J. Structural Proteins of Adenovirus-Associated Viruses. *J. Virol.* **8**, 766–770 (1971).

206. Becerra, S. P., Koczot, F., Fabisch, P. & Rose, J. A. Synthesis of Adeno-Associated Virus Structural Proteins Requires Both Alternative mRNA Splicing and Alternative Initiations from a Single Transcript. *J. Virol.* **62**, 2745–2754 (1988).

207. Becerra, S. P., Rose, J. A., Hardy, M., Baroudy, B. M. & Anderson, C. W. Direct mapping of adeno-associated virus capsid proteins B and C: A possible ACG initiation codon. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 7919–7923 (1985).

208. Sommer, M. *et al.* Quantification of Adeno-Associated Virus Particles and Empty Capsids by Optical Density Measurement. *Mol. Ther.* **7**, 122–128 (2003).

209. Gruntman, A. M. *et al.* Stability and Compatibility of Recombinant Adeno-Associated Virus Under Conditions Commonly Encountered in Human Gene Therapy Trials. *Hum. Gene Ther. Methods* **26**, 71–76 (2015).

210. Sonntag, F., Schmidt, K. & Kleinschmidt, J. A. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 10220–10225 (2010).

211. Qiu, J. & Brown, K. E. A 110-kDa Nuclear Shuttle Protein , Nucleolin , Specifically Binds to Adeno-Associated Virus Type 2 (AAV-2) Capsid. *Virology* **257**, 373–382 (1999).

212. Earley, L. F. *et al.* Adeno-associated Virus (AAV) Assembly-Activating Protein Is Not an Essential Requirement for Capsid Assembly of AAV Serotypes 4, 5, and 11. *J. Virol.* **91**, 1–21 (2017).

213. Mary, B., Maurya, S., Arumugam, S., Kumar, V. & Jayandharan, G. R. Post-translational modifications in capsid proteins of recombinant adeno-associated virus (AAV) 1-rh10 serotypes. *FEBS J.* **286**, 4964–4981 (2019).

214. Giles, A. R. *et al.* Deamidation of Amino Acids on the Surface of Adeno-Associated Virus Capsids Leads to Charge Heterogeneity and Altered Vector Function. *Mol. Ther.* **26**, 2848–2862 (2018).

215. Zhong, L. *et al.* Next generation of adeno-associated virus 2 vectors: Point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc.*

Natl. Acad. Sci. U. S. A. **105**, 2–7 (2008).

216. Powell, S. K., Samulski, R. J. & Mccown, T. J. AAV Capsid-Promoter Interactions Determine CNS Cell-Selective Gene Expression In Vivo. *Mol. Ther.* **28**, 1373–1380 (2020).

217. Takeuchi, Y., Myers, R. & Danos, O. Recombination and Population Mosaic of a Multifunctional Viral Gene, Adeno-Associated Virus cap. *PLoS One* **3**, 1–7 (2008).

218. Johnson, J. S. *et al.* Mutagenesis of Adeno-Associated Virus Type 2 Capsid Protein VP1 Uncovers New Roles for Basic Amino Acids in Trafficking and Cell-Specific Transduction. *J. Virol.* **84**, 8888–8902 (2010).

219. Popa-wagner, R. *et al.* Impact of VP1-Specific Protein Sequence Motifs on Adeno-Associated Virus Type 2 Intracellular Trafficking and Nuclear Entry. *J. Virol.* **86**, 9163–9174 (2012).

220. Girod, A. *et al.* The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. *J. Gen. Virol.* **83**, 973–978 (2002).

221. Lagache, T., Danos, O. & Holcman, D. Modeling the Step of Endosomal Escape during Cell Infection by a Nonenveloped Virus. *Biophys. J.* **102**, 980–989 (2012).

222. Sonntag, F., Bleker, S., Leuchs, B., Fischer, R. & Kleinschmidt, J. A. Adeno-Associated Virus Type 2 Capsids with Externalized VP1/VP2 Trafficking Domains Are Generated prior to Passage through the Cytoplasm and Are Maintained until Uncoating Occurs in the Nucleus. *J. Virol.* **80**, 11040–11054 (2006).

223. Warrington, K. H. *et al.* Adeno-Associated Virus Type 2 VP2 Capsid Protein Is Nonessential and Can Tolerate Large Peptide Insertions at Its N Terminus †. **78**, 6595–6609 (2004).

224. Münch, R. C. *et al.* Displaying high-affinity ligands on adeno-associated viral vectors enables tumor cell-specific and safe gene transfer. *Mol. Ther.* **21**, 109–18 (2013).

225. Wistuba, A., Kern, A., Weger, S., Grimm, D. & Kleinschmidt, J. A. Subcellular Compartmentalization of Adeno-Associated Virus Type 2 Assembly. *J. Virol.* **71**, 1341–1352 (1997).

226. Hoque, M. *et al.* Nuclear Transport of the Major Capsid Protein Is Essential for Adeno-Associated Virus Capsid Formation. *J. Virol.* **73**, 7912–7915 (1999).

227. Gao, G. *et al.* Clades of Adeno-Associated Viruses Are Widely Disseminated in Human Tissues. *J. Virol.* **78**, 6381–6388 (2004).

228. Schmidt, M., Katano, H., Bossis, I. & Chiorini, J. A. Cloning and Characterization of a Bovine Adeno-Associated Virus. *J. Virol.* **78**, 6509–6516 (2004).

229. Li, Y., Li, J., Liu, Y. & Shi, Z. Bat adeno-associated viruses as gene therapy vectors with the potential to evade human neutralizing antibodies. *Gene Ther.* **26**,

264–276 (2019).

230. Farkas, S. L. *et al.* A parvovirus isolated from royal python (Python regius) is a member of the genus Dependovirus. *J. Gen. Virol.* **85**, 555–561 (2004).

231. Rangarajan, S. *et al.* AAV5–Factor VIII Gene Transfer in Severe Hemophilia A. *N. Engl. J. Med.* **377**, 2519–2530 (2017).

232. Landegger, L. D. *et al.* A synthetic AAV vector enables safe and efficient gene transfer to the mammalian inner ear. *Nat. Biotechnol.* **35**, 280–284 (2017).

233. Choudhury, S. R. *et al.* In Vivo Selection Yields AAV-B1 Capsid for Central Nervous System and Muscle Gene Therapy. *Mol. Ther.* **24**, 1247–1257 (2016).

234. Antunes, S. L. M. *et al.* A novel adeno-associated virus capsid with enhanced neurotropism corrects a lysosomal transmembrane enzyme deficiency. *Brain a J. Neurol.* **141**, 2014–2031 (2018).

235. Michelfelder, S. *et al.* Peptide Ligands Incorporated into the Threefold Spike Capsid Domain to Re-Direct Gene Transduction of AAV8 and AAV9 In Vivo. *PLoS One* **6**, e23101 (2011).

236. Davidsson, M. *et al.* A systematic capsid evolution approach performed in vivo for the design of AAV vectors with tailored properties and tropism. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 27053–27062 (2019).

237. Gigout, L. *et al.* Altering AAV Tropism with Mosaic Viral Capsids. *Mol. Ther.* **11**, 856–865 (2005).

238. Judd, J. *et al.* Tunable Protease-Activatable Virus Nanonodes. *ACS Nano* **8**, 4740–4746 (2014).

239. Pearson, C. E., Zorbas, H., Price, G. B. & Zannis-Hadjopoulos, M. Inverted Repeats, Stem-Loops, and Cruciforms: Significance for Initiation of DNA Replication. *J. Cell. Biochem.* **63**, 1–22 (1996).

240. Garon, C. F., Berry, K. W. & Rose, J. A. A Unique Form of Terminal Redundancy in Adenovirus DNA Molecules. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 2391–2395 (1972).

241. Ninth Report of the International Committee on Taxonomy of Viruses. Family Poxviridae. in *Virus Taxonomy* 291–309 (2012). doi:10.1016/B978-0-12-384684-6.00028-8

242. Ninth Report of the International Committee on Taxonomy of Viruses. Phycodnaviridae. in *Virus Taxonomy* 249–262 (2012). doi:10.1016/B978-0-12-384684-6.00024-0

243. Lusby, E., Fife, K. H. & Berns, K. I. Nucleotide Sequence of the Inverted Terminal Repetition in Adeno-Associated Virus DNA. *J. Virol.* **34**, 402–409 (1980).

244. Wang, X., Qing, K., Ponnazhagan, S. & Srivastava, A. Adeno-Associated Virus Type 2 DNA Replication In Vivo : Mutation Analyses of the D Sequence in Viral

Inverted Terminal Repeats. *J. Virol.* **71**, 3077–3082 (1997).

245. Julien, L., Chassagne, J., Peccate, C., Lorain, S. & Piétri-rouxel, F. RFX1 and RFX3 Transcription Factors Interact with the D Sequence of Adeno-Associated Virus Inverted Terminal Repeat and Regulate AAV Transduction. *Sci. Rep.* **8**, 210 (2018).

246. Chiorini, J. A., Afione, S. & Kotin, R. M. Adeno-Associated Virus (AAV) Type 5 Rep Protein Cleaves a Unique Terminal Resolution Site Compared with Other AAV Serotypes. *J. Virol.* **73**, 4293–4298 (1999).

247. Ling, C. *et al.* Enhanced Transgene Expression from Recombinant Single-Stranded D-Sequence-Substituted Adeno-Associated Virus Vectors in Human Cell Lines In Vitro and in Murine Hepatocytes In Vivo. *J. Virol.* **89**, 952–961 (2015).

248. Samulski, R. J., Berns, K. I., Tan, M. & Muzyczka, N. Cloning of adeno-associated virus into pBR322 : Rescue of intact virus from the recombinant plasmid in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 1854–1858 (1982).

249. Wilmott, P., Lisowski, L., Alexander, I. E. & Logan, G. J. A User's Guide to the Inverted Terminal Repeats of Adeno-Associated Virus. *Hum. Gene Ther. Methods* **30**, 1–3 (2019).

250. Xiao, X., Xiao, W., Li, J. & Samulski, R. J. A Novel 165-Base-Pair Terminal Repeat Sequence Is the Sole cis Requirement for the Adeno-Associated Virus Life Cycle. *J. Virol.* **71**, 941–948 (1997).

251. Brister, J. R. & Muzyczka, N. Mechanism of Rep-mediated adeno-associated virus origin nicking. *J. Virol.* **74**, 7762–71 (2000).

252. Ryan, J. H., Zolotukhin, S. & Muzyczka, N. Sequence Requirements for Binding of Rep68 to the Adeno-Associated Virus Terminal Repeats. *J. Virol.* **70**, 1542–1553 (1996).

253. Yan, Z., Zak, R., Zhang, Y. & Engelhardt, J. F. Inverted Terminal Repeat Sequences Are Important for Intermolecular Recombination and Circularization of Adeno-Associated Virus Genomes. *J. Virol.* **79**, 364–379 (2005).

254. Linden, R. M., Ward, P., Giraud, C., Winocourt, E. & Berns, K. I. Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 11288–11294 (1996).

255. Rauschhuber, C., Wolf, A. & Ehrhardt, A. Transcriptional activity of inverted terminal repeats of various human adenovirus serotypes. *J. Gen. Virol.* **92**, 669–674 (2011).

256. Venkatesan, S. & Moss, B. In Vitro Transcription of the Inverted Terminal Repetition of the Vaccinia Virus Genome: Correspondence of Initiation and Cap Sites. *J. Virol.* **37**, 738–747 (1981).

257. Flotte, T. R. *et al.* Expression of the cystic fibrosis transmembrane conductance

regulator from a novel adeno-associated virus promoter. *J. Biol. Chem.* **268**, 3781–90 (1993).

258. Rubenstein, R. C., Mcveigh, U., Flotte, T. R., Guggino, W. B. & Zeitlin, P. L. CFTR gene transduction in neonatal rabbits using an adeno-associated virus (AAV) vector. *Gene Ther.* **4**, 384–392 (1997).

259. Haberman, R. P., McCown, T. J. & Samulski, R. J. Novel Transcriptional Regulatory Signals in the Adeno-Associated Virus Terminal Repeat A/D Junction Element. *J. Virol.* **74**, 8732–8739 (2000).

260. Earley, L. F. *et al.* Adeno - associated virus serotype specific inverted terminal repeat sequence role in vector transgene expression. *Hum. Gene Ther.* **31**, 151–162 (2020).

261. Hirsch, M. L. *et al.* Viral Single-Strand DNA Induces p53-Dependent Apoptosis in Human Embryonic Stem Cells. *PLoS One* **6**, 1–13 (2011).

262. Chiorini, J. A., Kim, F., Yang, L. & Kotin, R. M. Cloning and Characterization of Adeno-Associated Virus Type 5. *J. Virol.* **73**, 1309–1319 (1999).

263. Ling, C. *et al.* Strategies to generate high-titer, high-potency recombinant AAV3 serotype vectors. *Mol. Ther. - Methods Clin. Dev.* **3**, 16029 (2016).

264. Qiu, J., Nayak, R., Tullis, G. E., Pintel, D. J. & Irol, J. V. Characterization of the Transcription Profile of Adeno-Associated Virus Type 5 Reveals a Number of Unique Features Compared to Previously Characterized Adeno-Associated Viruses. *J. Virol.* **76**, 12435–12447 (2002).

265. Kyöstiö, S. R. M., Wonderling, R. S. & Owens, R. A. Negative Regulation of the Adeno-Associated Virus (AAV) P5 Promoter Involves both the P5 Rep Binding Site and the Consensus ATP-Binding Motif of the AAV Rep68 Protein. *J. Virol.* **69**, 6787–6796 (1995).

266. West, M. H. P., Trempe, J. P., Tratschin, J.-D. & Carter, B. J. Gene Expression in Adeno-Associated Virus Vectors: The Effects of Chimeric Structure, Helper Virus, and Adenovirus VA1 RNA. *Virology* **160**, 38–47 (1987).

267. Hörer, M. *et al.* Mutational Analysis of Adeno-Associated Virus Rep Protein-Mediated Inhibition of Heterologous and Homologous Promoters. *J. Virol.* **69**, 5485–5496 (1995).

268. Gordon, S., Akopyan, G., Garban, H. & Bonavida, B. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* **25**, 1125–42 (2006).

269. Seto, E., Shi, Y. & Shenk, T. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. *Nature* **354**, 241–245 (1991).

270. Carthew, R. W., Chodosh, L. A. & Sharp, P. A. An RNA Polymerase II Transcription Factor Binds to an Upstream Element in the Adenovirus Major Late Promoter. *Cell* **43**, 439–446 (1985).

271. Chang, L., Shi, Y. & Shenk, T. Adeno-Associated Virus P5 Promoter Contains an Adenovirus ElA-Inducible Element and a Binding Site for the Major Late Transcription Factor. *J. Virol.* **63**, 3479–3488 (1989).

272. Stutika, C. *et al.* A Comprehensive RNA Sequencing Analysis of the Adeno-Associated Virus (AAV) Type 2 Transcriptome Reveals Novel AAV Transcripts. *J. Virol.* **90**, 1278–1289 (2016).

273. Feng, D. *et al.* A 16-bp RBE element mediated is sufficient for mediating Rep-dependent integration into AAVS1. *J. Mol. Biol.* **358**, 38–45 (2006).

274. Philpott, N. J. *et al.* Efficient Integration of Recombinant Adeno-Associated Virus DNA Vectors Requires a p5- rep Sequence in cis. *J. Virol.* **76**, 5411–5421 (2002).

275. Y. Ogasawara, M. Urabe, K. O. The Use of Heterologous Promoters for Adeno-Associated Virus (AAV) Protein Expression in AAV Vector Production.pdf. *Microbiol. Immunol.* **42**, 177–185 (1998).

276. Beaton, A., Palumbo, P. & Berns, K. I. Expression from the Adeno-Associated Virus p5 and p19 Promoters Is Negatively Regulated in trans by the rep Protein. *J. Virol.* **63**, 4450–4454 (1989).

277. Pereira, D. J. & Muzyczka, N. The Cellular Transcription Factor SP1 and an Unknown Cellular Protein Are Required To Mediate Rep Protein Activation of the Adeno-Associated Virus p19 Promoter. *J. Virol.* **71**, 1747–1756 (1997).

278. Pereira, D. J. & Muzyczka, N. The Adeno-Associated Virus Type 2 p40 Promoter Requires a Proximal Sp1 Interaction and a p19 CArG-Like Element To Facilitate Rep Transactivation. *J. Virol.* **71**, 4300–4309 (1997).

279. McCarty, D. M. *et al.* Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein. *J. Virol.* **68**, 4988–4997 (1994).

280. Weger, S., Wistuba, A., Grimm, D. & Kleinschmidt, J. A. Control of Adeno-Associated Virus Type 2 Cap Gene Expression: Relative Influence of Helper Virus, Terminal Repeats, and Rep Proteins. *J. Virol.* **71**, 8437–8447 (1997).

281. Weger, S., Hammer, E., Gonsior, M., Stutika, C. & Heilbronn, R. A Regulatory Element Near the 3' End of the Adeno-Associated Virus rep Gene Inhibits Adenovirus Replication in cis by Means of p40 Promoter-Associated Short Transcripts. *J. Virol.* **90**, 3981–3993 (2016).

282. Pereira, D. J., McCarty, D. M. & Muzyczka, N. The Adeno-Associated Virus (AAV) Rep Protein Acts as both a Repressor and an Activator To Regulate AAV Transcription during a Productive Infection. *J. Virol.* **71**, 1079–1088 (1997).

283. Ye, C., Qiu, J. & Pintel, D. J. Efficient Expression of the Adeno-Associated Virus Type 5 P41 Capsid Gene Promoter in 293 Cells Does Not Require Rep. *J. Virol.* **80**, 6559–6567 (2006).

284. Stutika, C. *et al.* Comprehensive Small RNA-Seq of Adeno-Associated Virus (AAV)-Infected Human Cells Detects Patterns of Novel, Non-Coding AAV RNAs in

the Absence of Cellular miRNA Regulation. *PLoS One* **11**, e0161454 (2016).

285. Sonntag, F. *et al.* The Assembly-Activating Protein Promotes Capsid Assembly of Different Adeno-Associated Virus Serotypes. *J. Virol.* **85**, 12686–12697 (2011).

286. Naumer, M. *et al.* Properties of the Adeno-Associated Virus Assembly-Activating Protein. **86**, 13038–13048 (2012).

287. Maurer, A. C., Karla, A., Diaz, C. & Vandenberghe, L. H. Residues on Adeno-associated Virus Capsid Lumen Dictate Interactions and Compatibility with the Assembly-Activating Protein. *J. Virol.* **93**, 1–16 (2019).

288. Maurer, A. C. *et al.* The Assembly-Activating Protein Promotes Stability and Interactions between AAV's Viral Proteins to Article The Assembly-Activating Protein Promotes Stability and Interactions between AAV ' s Viral Proteins to Nucleate Capsid Assembly. *Cell Rep.* **23**, 1817–1830 (2018).

289. Grosse, S. *et al.* Relevance of Assembly-Activating Protein for Adeno-associated Virus Vector Production and Capsid Protein Stability in Mammalian and Insect Cells. *J. Virol.* **91**, 1–30 (2017).

290. Herrmann, A. *et al.* Impact of the Assembly-Activating Protein on Molecular Evolution of Synthetic Adeno-Associated Virus Capsids. *Hum. Gene Ther.* **30**, 21–36 (2019).

291. Le, D. T., Radukic, M. T. & Müller, K. M. Adeno-associated virus capsid protein expression in Escherichia coli and chemically defined capsid assembly. *Sci. Rep.* **9**, 1–10 (2019).

292. Hermonat, P. L. *et al.* Chromosomal latency and expression at map unit 96 of a wild-type plus adeno-associated virus (AAV)/Neo vector and identification of p81, a new AAV transcriptional promoter. *J. Hum. Virol.* **2**, 359–368 (1999).

293. Cao, M., Chiriva-internati, M. & Hermonat, P. L. AAV2 X increases AAV6 rep/cap-driven rAAV production. *Virology* **482**, 84–88 (2015).

294. Graham, F. L., Smiley, J., Russel, W. C. & Nairn, R. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *J. Gen. Virol.* **36**, 59–74 (1977).

295. Ding, L., Lu, S. & Munshi, N. C. In vitro packaging of an infectious recombinant adeno-associated virus 2. *Gene Ther.* **4**, 1167–1172 (1997).

296. Horwitz, G. A. *et al.* Adenovirus Small e1a Alters Global Patterns of Histone Modification. *Science (80-.).* **321**, 1084–1086 (2008).

297. Shi, Y., Seto, E., Chang, L. & Shenk, T. Transcriptional Repression by YY1, a Human GLI-Krüppel-Related Protein, and Relief of Repression by Adenovirus E IA Protein. *Cell* **67**, 377–388 (1991).

298. Chang, L. & Shenk, T. The Adenovirus DNA-Binding Protein Stimulates the Rate of Transcription Directed by Adenovirus and Adeno-Associated Virus Promoters.

J. Virol. **64**, 2103–2109 (1990).

299. Janik, J. E., Huston, M. M. & Rose, J. A. Efficient Synthesis of Adeno-Associated Virus Structural Proteins Requires Both Adenovirus DNA Binding Protein and VA I RNA. *Virology* **162**, 320–329 (1989).

300. Samulski, R. J. & Shenk, T. Adenovirus EBB 55-Mr Polypeptide Facilitates Timely Cytoplasmic Accumulation of Adeno-Associated Virus mRNAs. *J. Virol.* **62**, 206–210 (1988).

301. Nayak, R. & Pintel, D. J. Positive and Negative Effects of Adenovirus Type 5 Helper Functions on Adeno-Associated Virus Type 5 (AAV5) Protein Accumulation Govern AAV5 Virus Production. *J. Virol.* **81**, 2205–2212 (2007).

302. Nada, S. & Trempe, J. P. Characterization of Adeno-Associated Virus Rep Protein Inhibition of Adenovirus E2a Gene Expression. *Virology* **293**, 345–355 (2002).

303. Casper, J. M., Timpe, J. M., Dignam, J. D. & Trempe, J. P. Identification of an Adeno-Associated Virus Rep Protein Binding Site in the Adenovirus E2a Promoter. *J. Virol.* **79**, 28–38 (2005).

304. Colgrove, R. C. *et al.* History and genomic sequence analysis of the herpes simplex virus 1 KOS and KOS.1 sub-strains. *Virology* **487**, 215–221 (2016).

305. Conway, J. E., Zolotukhin, S., Muzyczka, N., Hayward, G. S. & Byrne, B. J. Recombinant Adeno-Associated Virus Type 2 Replication and Packaging Is Entirely Supported by a Herpes Simplex Virus Type 1 Amplicon Expressing Rep and Cap. *J. Virol.* **71**, 8780–8789 (1997).

306. Weindler, F. W. & Heilbronn, R. A Subset of Herpes Simplex Virus Replication Genes Provides Helper Functions for Productive Adeno-Associated Virus Replication. *J. Virol.* **65**, 2476–2483 (1991).

307. Glauser, D. L. *et al.* Live Covisualization of Competing Adeno-Associated Virus and Herpes Simplex Virus Type 1 DNA Replication : Molecular Mechanisms of Interaction □. *J. Virol.* **81**, 4732–4743 (2007).

308. You, H. *et al.* Multiple human papillomavirus genes affect the adeno-associated virus life cycle. *Virology* **344**, 532–540 (2006).

309. Cao, M., Bandyopadhyay, S., Zhu, H., You, H. & Hermonat, P. L. The HPV16 E1 Carboxyl Domain Provides a Helper Function for Adeno-Associated Virus Replication. *Intervirology* **61**, 185–192 (2019).

310. Cao, M., Zhu, H., Bandyopadhyay, S., You, H. & Hermonat, P. L. HPV-16 E1 , E2 and E6 each complement the Ad5 helper gene set , increasing rAAV2 and wt AAV2 production. *Gene Ther.* **19**, 418–424 (2012).

311. Schlehofer, J. R., Ehrbar, M. & Hausen, H. Zur. Vaccinia Virus, Herpes Simplex Virus, and Carcinogens Induce DNA Amplification in a Human Cell Line and Support Replication of a Helpervirus Dependent Parvovirus. *Virology* **152**, 110–117 (1986).

312. Moore, A. R., Dong, B., Chen, L. & Xiao, W. Vaccinia virus as a subhelper for AAV replication and packaging. *Mol. Ther. - Methods Clin. Dev.* **2**, 15044 (2015).

313. Wang, Z. *et al.* Human Bocavirus 1 Is a Novel Helper for Adeno-Associated Virus Replication. *J. Virol.* **91**, 1–18 (2017).

314. Yalkinoglu, A. Ö., Heilbronn, R., Bürkle, A., Schlehofer, J. R. & Hausen, H. Zur. DNA Amplification of Adeno-associated Virus as a Response to Cellular Genotoxic Stress. *Cancer Res.* **48**, 3123–3129 (1988).

315. Hermonat, P. L. & Muzyczka, N. Use of adeno-associated virus as a mammalian DNA cloning vector : Transduction of neomycin resistance into mammalian tissue culture cells. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 6466–6470 (1984).

316. Xiao, X., Li, J. & Samulski, R. J. Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J. Virol.* **72**, 2224–32 (1998).

317. Evers, M. M. *et al.* AAV5-miHTT Gene Therapy Demonstrates Broad Distribution and Strong Human Mutant Huntingtin Lowering in a Huntington' s Disease Minipig Model. *Mol. Ther.* **26**, 2163–2177 (2018).

318. Li, C., Xiao, P., Gray, S. J., Weinberg, M. S. & Samulski, R. J. Combination therapy utilizing shRNA knockdown and an optimized resistant transgene for rescue of diseases caused by misfolded proteins. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14258–63 (2011).

319. Foust, K. D. *et al.* Therapeutic AAV9-mediated Suppression of Mutant SOD1 Slows Disease Progression and Extends Survival in Models of Inherited ALS. *Mol. Ther.* **21**, 2148–2159 (2013).

320. Excoffon, K. J. D. a *et al.* Directed evolution of adeno-associated virus to an infectious respiratory virus. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3865–70 (2009).

321. Lisowski, L. *et al.* Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature* **506**, 382–6 (2014).

322. Sallach, J. *et al.* Tropism-modified AAV vectors overcome barriers to successful cutaneous therapy. *Mol. Ther.* **22**, 929–39 (2014).

323. Tse, L. V, Moller-tank, S. & Asokan, A. Strategies to circumvent humoral immunity to adeno-associated viral vectors. 845–855 (2015).

324. Maersch, S., Huber, A., Büning, H., Hallek, M. & Perabo, L. Optimization of stealth adeno-associated virus vectors by randomization of immunogenic epitopes. *Virology* **397**, 167–175 (2010).

325. Grimm, D. *et al.* In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *J. Virol.* **82**, 5887–911 (2008).

326. Asuri, P. *et al.* Directed evolution of adeno-associated virus for enhanced gene

delivery and gene targeting in human pluripotent stem cells. *Mol. Ther.* **20**, 329–38 (2012).

327. Ellis, B. L. *et al.* A survey of ex vivo / in vitro transduction efficiency of mammalian primary cells and cell lines with Nine natural adeno-associated virus (AAV1-9) and one engineered adeno-associated virus serotype. *Virol. J.* **10**, 1–10 (2013).

328. Raupp, C., Naumer, M., Gurda, B. L., Agbandje-mckenna, M. & Kleinschmidt, J. A. The Threefold Protrusions of Adeno-Associated Virus Type 8 Are Involved in Cell Surface Targeting as Well as Postattachment. *J. Virol.* **86**, 9396–9408 (2012).

329. Zhang, R. *et al.* Divergent engagements between adeno-associated viruses with their cellular receptor AAVR. *Nat. Commun.* **10**, 3760 (2019).

330. Lerch, T. F., Xie, Q. & Chapman, M. S. The structure of adeno-associated virus serotype 3B (AAV3V): Insights into receptor binding and immune evasion. *Virology* **403**, 26–36 (2011).

331. Bartlett, J. S., Wilcher, R. & Samulski, R. J. Infectious Entry Pathway of Adeno-Associated Virus and Adeno-Associated Virus Vectors. *J. Virol.* **74**, 2777–2785 (2000).

332. Uhrig, S. *et al.* Successful target cell transduction of capsid-engineered rAAV vectors requires clathrin-dependent endocytosis. *Gene Ther.* **19**, 210–218 (2012).

333. Xiao, P.-J. & Samulski, R. J. Cytoplasmic trafficking, endosomal escape, and perinuclear accumulation of adeno-associated virus type 2 particles are facilitated by microtubule network. *J. Virol.* **86**, 10462–73 (2012).

334. Nonnenmacher, M. E., Cintrat, J., Gillet, D. & Weber, T. Syntaxin 5-Dependent Retrograde Transport to the trans -Golgi Network Is Required for Adeno-Associated Virus Transduction. *J. Virol.* **89**, 1673–1687 (2015).

335. Kronenberg, S., Böttcher, B., Lieth, C. W. Von Der, Bleker, S. & Kleinschmidt, J. A. A Conformational Change in the Adeno-Associated Virus Type 2 Capsid Leads to the Exposure of Hidden VP1 N Termini. *J. Virol.* **79**, 5296–5303 (2005).

336. Grieger, J. C., Snowdy, S. & Samulski, R. J. Separate Basic Region Motifs within the Adeno-Associated Virus Capsid Proteins Are Essential for Infectivity and Assembly. *J. Virol.* **80**, 5199–5210 (2006).

337. Ferrari, F. K., Samulski, T., Shenk, T. & Samulski, R. J. Second-Strand Synthesis Is a Rate-Limiting Step for Efficient Transduction by Recombinant Adeno-Associated Virus Vectors. *J. Virol.* **70**, 3227–3234 (1996).

338. Brimble, M. A. *et al.* New and improved AAVenues : current status of hemophilia B gene therapy New and improved AAVenues : current status of hemophilia B gene therapy. **2598**, (2015).

339. Pillay, S. *et al.* An essential receptor for adeno-associated virus infection. *Nature* **530**, 108–112 (2016).

340. Dudek, A. M. *et al.* An Alternate Route for Adeno-associated Virus (AAV) Entry Independent of AAV Receptor. *J. Virol.* **92**, 1–15 (2018).

341. Dudek, A. M. *et al.* GPR108 Is a Highly Conserved AAV Entry Factor. *Mol. Ther.* **28**, 1–15 (2020).

342. Meisen, W. H. *et al.* Pooled Screens Identify GPR108 and TM9SF2 as Host Cell Factors Critical for AAV Transduction. *Mol. Ther. Methods Clin. Dev.* **17**, 601–611 (2020).

343. Bordet, T. & Behar-cohen, F. Ocular gene therapies in clinical practice : viral vectors and nonviral alternatives. *Drug Discov. Today* **24**, 1685–1693 (2019).

344. Aslanidi, G., Lamb, K. & Zolotukhin, S. An inducible system for highly efficient production of recombinant adeno-associated virus (rAAV) vectors in insect Sf9 cells. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 9–14 (2009).

345. Kamen, A. & Henry, O. Development and optimization of an adenovirus production process. *J. Gene Med.* **6**, 184–192 (2004).

346. Ory, D. S., Neugeboren, B. A. & Mulligan, R. C. A stable human-derived packaging cell line for production of high titer retrovirus / vesicular stomatitis virus G pseudotypes. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 11400–11406 (1996).

347. Baldi, L., Hacker, D. L., Adam, M. & Wurm, F. M. Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. *Biotechnol. Lett.* **29**, 677–684 (2007).

348. Dumont, J., Euwart, D., Mei, B., Estes, S. & Kshirsagar, R. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit. Rev. Biotechnol.* **36**, 1110–1122 (2016).

349. Piras, B. A. *et al.* Distribution of AAV8 particles in cell lysates and culture media changes with time and is dependent on the recombinant vector. *Mol. Ther. - Methods Clin. Dev.* **3**, 1–5 (2016).

350. Blessing, D. *et al.* Scalable Production of AAV Vectors in Orbitally Shaken HEK293 Cells. *Mol. Ther. Methods Clin. Dev.* **13**, 14–26 (2019).

351. Gray, S. J. *et al.* Production of Recombinant Adeno-Associated Viral Vectors and Use in In Vitro and In Vivo Administration. *Curr. Protoc. Neurosci.* **57**, 4.17.1-4.17.30 (2011).

352. Wobbe, C. R. *et al.* Replication of simian virus 40 origin-containing DNA in vitro with purified proteins. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 1834–1838 (1987).

353. Bae, D. H. *et al.* Design and Testing of Vector-Producing HEK293T Cells Bearing a Genomic Deletion of the SV40 T Antigen Coding Region. *Mol. Ther. - Methods Clin. Dev.* **18**, 631–638 (2020).

354. Clément, N. & Grieger, J. C. Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Mol. Ther. - Methods Clin. Dev.* **3**, 16002 (2016).

355. Lock, M. *et al.* Rapid, Simple, and Versatile Manufacturing of Recombinant Adeno-Associated Viral Vectors at Scale. *Hum. Gene Ther.* **21**, 1259–1271 (2010).

356. Chahal, P. S., Schulze, E., Tran, R., Montes, J. & Kamen, A. A. Production of adeno-associated virus (AAV) serotypes by transient transfection of HEK293 cell suspension cultures for gene delivery. *J. Virol. Methods* 163–173 (2013).

357. Yuan, Z., Qiao, C., Hu, P., Li, J. & Xiao, X. A Versatile Adeno-Associated Virus Vector Producer Cell Line Method for Scalable Vector Production of Different Serotypes. *Hum. Gene Ther.* **22**, 613–624 (2011).

358. Wright, J. F. Transient Transfection Methods for Clinical Adeno-Associated Viral Vector Production. *Hum. Gene Ther.* **706**, 698–706 (2009).

359. Fang, X. T., Sehlin, D., Lannfelt, L., Syvänen, S. & Hultqvist, G. Efficient and inexpensive transient expression of multispecific multivalent antibodies in Expi293 cells. *Biol. Proced. Online* **19**, 1–9 (2017).

360. Oers, M. M. Van & Vlak, J. M. Baculovirus Genomics. *Curr. Drug Targets* **8**, 1051–68 (2007).

361. Clem, R. J. & Passarelli, A. L. Baculoviruses : Sophisticated Pathogens of Insects. *PLoS Pathog.* **9**, e1003729. (2013).

362. Arif, B., Escasa, S. & Pavlik, L. Biology and Genomics of Viruses Within the Genus Gammabaculovirus. *Viruses* **3**, 2214–2222 (2011).

363. Lightner, D. V. & Redman, R. M. A baculovirus-caused disease of the penaeid shrimp, Penaeus monodon. *J. Invertebr. Pathol.* **38**, 299–302 (2011).

364. Nandakumar, S., Ma, H. & Khan, A. S. Whole-Genome Sequence of the Spodoptera fruigperda Sf9 Insect Cell Line. *Genome Announc.* **5**, 9–10 (2017).

365. Jeang, K. T., Holmgren-Konig, M. & Khoury, G. A baculovirus vector can express intron-containing genes. *J. Virol.* **61**, 1761–1764 (1987).

366. Urabe, M., Ding, C. & Kotin, R. M. Insect Cells as a Factory to Produce Adeno-Associated Virus Type 2 Vectors. *Hum. Gene Ther.* **13**, 1935–1943 (2002).

367. Savy, A., Kaikkonen, M. U., Le, A., Galibert, L. & Merten, O. Genetics instability of wtAAV2 genome and AAV promoter activities in the Baculovirus / Sf9 cells system. 1–18 (2018).

368. Negrete, A., Yang, L. C., Mendez, A. F., Levy, J. R. & Kotin, R. M. Economized large-scale production of high yield of rAAV for gene therapy applications exploiting baculovirus expression system. *J. Gene Med.* **9**, 938–948 (2007).

369. Wu, Y. *et al.* A Recombinant Baculovirus Efficiently Generates Recombinant Adeno-Associated Virus Vectors in Cultured Insect Cells and Larvae. *Mol. Ther. Methods Clin. Dev.* **10**, 38–47 (2018).

370. Mietzsch, M. *et al.* OneBac: Platform for Scalable and High-Titer Production of

Adeno-Associated Virus Serotype 1–12 Vectors for Gene Therapy. *Hum. Gene Ther.* **25**, 212–222 (2014).

371. Mietzsch, M. *et al.* OneBac 2.0: Sf9 Cell Lines for Production of AAV1, AAV2, and AAV8 Vectors with Minimal Encapsidation of Foreign DNA. *Hum. Gene Ther. Methods* **28**, 15–22 (2017).

372. Galibert, L. *et al.* Origins of truncated supplementary capsid proteins in rAAV8 vectors produced with the baculovirus system. *PLoS One* **13**, e0207414 (2018).

373. Galvin, T. A., Glasner, D. R., Shaheduzzaman, S. & Khan, A. S. Identification of a Novel Rhabdovirus in Spodoptera frugiperda Cell. *J. Virol.* **88**, 6576–6585 (2014).

374. Geisler, C. A new approach for detecting adventitious viruses shows Sf-rhabdovirus-negative Sf-RVN cells are suitable for safe biologicals production. *BMC Biotechnol.* **18**, 1–19 (2018).

375. Bryant, L. M. *et al.* Lessons Learned from the Clinical Development and Market Authorization of Glybera. *Hum. gene Ther. Clin. Dev.* **24**, 55–64 (2013).

376. Kondratov, O. *et al.* Direct Head-to-Head Evaluation of Recombinant Adeno-associated Viral Vectors Manufactured in Human versus Insect Cells. *Mol. Ther.* **25**, 2661–2675 (2017).

377. Clément, N., Knop, D. R. & Byrne, B. J. Large-Scale Adeno-Associated Viral Vector Production Using a Herpesvirus-Based System Enables Manufacturing for Clinical Studies. *Hum. Gene Ther.* **20**, 796–806 (2009).

378. Martin, J. *et al.* Generation and Characterization of Adeno-Associated Virus Producer Cell Lines for Research and Preclinical Vector Production. *Hum. Gene Ther.* **24**, 253–269 (2013).

379. Flotte, T. R. *et al.* Phase 2 clinical trial of a recombinant adeno-associated viral vector expressing α1-antitrypsin: interim results. *Hum. Gene Ther.* **22**, 1239–47 (2011).

380. Picken, R. N. & Yang, H.-L. The integration of HPV-18 into HeLa cells has involved duplication of part of the viral genome as well as human DNA flanking sequences. *Nucleic Acids Res.* **15**, 10068 (1987).

381. Gao, G. *et al.* Rep/Cap Gene Amplification and High-Yield Production of AAV in an A549 Cell Line Expressing Rep/Cap. *Mol. Ther.* **5**, 644–649 (2002).

382. Farson, D. *et al.* Development and characterization of a cell line for large-scale , serum-free production of recombinant adeno-associated viral vectors. *J. Gene Med.* **6**, 1369–1381 (2004).

383. Barajas, D. *et al.* Generation of infectious recombinant Adeno- associated virus in Saccharomyces cerevisiae. *PLoS One* **12**, e0173010 (2017).

384. Shin, J.-H., Yue, Y. & Duan, D. Recombinant Adeno-Associated Viral Vector Production and Purification. *Methods Mol. Biol.* **798**, 267–284 (2012).

385. Arden, E. & Metzger, J. M. Inexpensive, serotype-independent protocol for native and bioengineered recombinant adeno-associated virus purification. *J. Biol. Methods* **3**, 1–7 (2016).

386. Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nat. Protoc.* **1**, 1412–28 (2006).

387. Kimura, T. *et al.* Production of adeno-associated virus vectors for in vitro and in vivo applications. *Sci. Rep.* **9**, 13601 (2019).

388. Guo, P. *et al.* Rapid and simplified purification of recombinant adeno-associated virus. *J. Virol. Methods* **183**, 139–146 (2012).

389. Florencio, G. D. *et al.* Simple downstream process based on detergent treatment improves yield and in vivo transduction efficacy of adeno-associated virus vectors. *Mol. Ther. - Methods Clin. Dev.* **2**, 15024 (2015).

390. Buclez, P. *et al.* Rapid, scalable, and low-cost purification of recombinant adeno-associated virus produced by baculovirus expression vector system. *Mol. Ther. - Methods Clin. Dev.* **3**, 1–10 (2016).

391. Smith, R. H., Yang, L. & Kotin, R. M. Chromatography-based purification of adeno-associated virus. *Methods Mol. Biol.* **434**, 37–54 (2008).

392. Strobel, B. *et al.* Standardized, Scalable, and Timely Flexible Adeno-Associated Virus Vector Production Using Frozen High-Density HEK-293 Cell Stocks and CELLdiscs. *Hum. Gene Ther. Methods* **30**, 23–33 (2019).

393. De La Maza, L. M. & Carter, B. J. Molecular Structure of Adeno-associated Virus Variant DNA. *J. Biol. Chem.* **255**, 3194–203 (1980).

394. Su, Q., Sena-esteves, M. & Gao, G. Purification of Recombinant Adeno-Associated Viruses (rAAVs) by Cesium Chloride Gradient Sedimentation. *Cold Spring Harb. Protoc.* **8**, 341–347 (2020).

395. Ayuso, E. *et al.* High AAV vector purity results in serotype- and tissue-independent enhancement of transduction efficiency. *Gene Ther.* **17**, 503–510 (2010).

396. Zolotukhin, S. *et al.* Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther.* **6**, 973–985 (1999).

397. Klein, R. L., Dayton, R. D., Tatom, J. B., Henderson, K. M. & Henning, P. P. AAV8, 9, Rh10, Rh43 Vector Gene Transfer in the Rat Brain: Effects of Serotype, Promoter and Purification Method. *Mol. Ther.* **16**, 89–96 (2008).

398. Strobel, B., Miller, F. D., Rist, W. & Lamla, T. Comparative Analysis of Cesium Chloride- and Iodixanol-Based Purification of Recombinant Adeno-Associated Viral Vectors for Preclinical Applications. *Hum. Gene Ther. Methods* **26**, 147–157 (2015).

399. Zolotukhin, S. *et al.* Production and purification of serotype 1, 2, and 5

recombinant adeno-associated viral vectors. *Methods* **28**, 158–167 (2002).

400. Nass, S. A. *et al.* Universal Method for the Purification of Recombinant AAV Vectors of Differing Serotypes. *Mol. Ther. Methods Clin. Dev.* **9**, 33–46 (2018).

401. Pulicherla, N. & Asokan, A. Peptide affinity reagents for AAV capsid recognition and purification. *Gene Ther.* **18**, 1020–1024 (2011).

402. Wang, Q. *et al.* Identification of an adeno-associated virus binding epitope for AVB sepharose affinity resin. *Mol. Ther. - Methods Clin. Dev.* **2**, 15040 (2015).

403. Terova, O., Soltys, S., Hermans, P., Rooij, J. De & Detmers, F. Overcoming downstream purification challenges for viral vector manufacturing: enabling advancement of gene therapies in the clinic. *Cell Gene Ther. Insights* 101–111 (208AD). doi:10.18609/CGTI.2018.017

404. Clark, K. R., Liu, X., Grath, J. P. M. C. & Johnson, P. R. Highly Purified Recombinant Adeno-Associated Virus Vectors Are Biologically Active and Free of Detectable Helper and Wild-Type Viruses. *Hum. Gene Ther.* **10**, 1031–9 (1999).

405. Kaludov, N., Handelman, B. & Chiorini, J. A. Scalable Purification of Adeno-Associated Virus Type 2, 4, or 5 Using Ion-Exchange Chromatography. *Hum. Gene Ther.* **13**, 1235–1243 (2002).

406. Urabe, M. *et al.* Removal of Empty Capsids from Type 1 Adeno-Associated Virus Vector Stocks by Anion-Exchange Chromatography Potentiates Transgene Expression. *Mol. Ther.* **13**, 823–828 (2006).

407. Lock, M., Alvira, M. R. & Wilson, J. M. Analysis of Particle Content of Recombinant by Ion-Exchange Chromatography. *Hum. Gene Ther.* **23**, 56–64 (2012).

408. Tomono, T. *et al.* Highly Efficient Ultracentrifugation-free Chromatographic Purification of Recombinant AAV Serotype 9. *Mol. Ther. Methods Clin. Dev.* **11**, 180–190 (2018).

409. Burova, E. & Ioffe, E. Chromatographic purification of recombinant adenoviral and adeno-associated viral vectors: methods and implications. *Gene Ther.* **12**, S5–S17 (2005).

410. Brument, N. *et al.* A Versatile and Scalable Two-Step Ion-Exchange Chromatography Process for the Purification of Recombinant Adeno-associated Virus Serotypes-2 and -5. *Mol. Ther.* **6**, 678–686 (2002).

411. Gao, G. *et al.* Purification of Recombinant Adeno-Associated Virus Vectors by Column Chromatography and Its Performance in Vivo. *Hum. Gene Ther.* **11**, 2079–2091 (2000).

412. Grantham, R. Viral, Prokaryote and Eukaryote Genes Contrasted by mRNA Sequence Indexes. *FEBS Lett.* **95**, 1–11 (1978).

413. Grantham, R., Gautier, C., Gouy, M., Mercier, R. & Pavé, A. Codon Catalog

Usage and the Genome Hypothesis. *Nucleic Acids Res.* **8**, 49–62 (1980).

414. Zhou, Z. *et al.* Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 6117–6125 (2016).

415. Plotkin, J. B., Robins, H. & Levine, A. J. Tissue-specific codon usage and the expression of human genes. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12588–91 (2004).

416. Alexaki, A. *et al.* Effects of codon optimization on coagulation factor IX translation and structure: Implications for protein and gene therapies. *Sci. Rep.* **9**, 15449 (2019).

417. Pierce, G. F. & Iorio, A. Past, present and future of haemophilia gene therapy : From vectors and transgenes to known and unknown outcomes. *Haemophilia* **24**, 60–67 (2018).

418. Paulk, N. K. *et al.* Bioengineered AAV Capsids with Combined High Human Liver Transduction In Vivo and Unique Humoral Seroreactivity. *Mol. Ther.* **26**, 289–303 (2018).

419. Pekrun, K. *et al.* Using a barcoded AAV capsid library to select for clinically relevant gene therapy vectors Graphical abstract Find the latest version : Using a barcoded AAV capsid library to select for clinically relevant gene therapy vectors. **4**, (2019).

420. Wang, L. *et al.* Comparative Study of Liver Gene Transfer With AAV Vectors Based on Natural and Engineered AAV Capsids. *Mol. Ther.* **23**, 1877–1887 (2015).

421. Bissig-choisat, B. *et al.* Development and rescue of human familial hypercholesterolaemia in a xenograft mouse model. *Nat. Commun.* **6**, 1–9 (2015).

422. Shao, W. *et al.* Superior human hepatocyte transduction with adeno-associated virus vector serotype 7. *Gene Ther.* **26**, 504–514 (2019).

423. Miesbach, W. *et al.* Gene therapy with adeno-associated virus vector 5 – human factor IX in adults with hemophilia B. *Blood* **131**, 1022–1032 (2019).

424. Wu, T.-L. *et al.* CD8(+) T Cell Recognition of Epitopes Within the Capsid of Adeno-associated Virus 8-based Gene Transfer Vectors Depends on Vectors' Genome. *Mol. Ther.* **22**, 42–51 (2014).

425. Mingozzi, F. *et al.* CD8+ T-cell responses to associated virus capsid in humans. *Nat. Med.* **13**, 419–422 (2007).

426. Yuasa, K. *et al.* Injection of a recombinant AAV serotype 2 into canine skeletal muscles evokes strong immune responses against transgene products. *Gene Ther.* **14**, 1249–1260 (2007).

427. Zhu, J., Huang, X. & Yang, Y. The TLR9-MyD88 pathway is critical for adaptive

immune responses to adeno- associated virus gene therapy vectors in mice. *J. Clin. Invest.* **119**, (2009).

428. Faust, S. M. *et al.* CpG-depleted adeno-associated virus vectors evade immune detection Find the latest version : CpG-depleted adeno-associated virus vectors evade immune detection. *J. Clin. Invest.* **123**, 2994–3001 (2013).

429. Martino, A. T. *et al.* The genome of self-complementary adeno-associated viral vectors increases Toll-like receptor 9 – dependent innate immune responses in the liver. *Blood* **117**, 6459–6468 (2011).

430. Mueller, C. *et al.* 5 Year Expression and Neutrophil Defect Repair after Gene Therapy in Alpha-1 Antitrypsin Deficiency. *Mol. Ther.* **25**, 1387–1394 (2017).

431. Salas, D. *et al.* Immunoadsorption enables successful rAAV5-mediated repeated hepatic gene delivery in nonhuman primates. *Blood Adv.* **3**, 2–11 (2019).

432. Mingozzi, F. *et al.* Pharmacological Modulation of Humoral Immunity in a Nonhuman Primate Model of AAV Gene Transfer for Hemophilia B. *Mol. Ther.* **20**, 1410–1416 (2012).

433. 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.* **25**, 1831–1842 (2017).

434. Meliani, A. *et al.* Antigen-selective modulation of AAV immunogenicity with tolerogenic rapamycin nanoparticles enables successful vector re- administration. *Nat. Commun.* **9**, (2018).

435. Perocheau, D. P. *et al.* Age-Related Seroprevalence of Antibodies Against AAV-LK03 in a UK Population Cohort. *Hum. Gene Ther.* **30**, 79–88 (2019).

436. Calcedo, R. & Wilson, J. M. AAV Natural Infection Induces Broad Cross-Neutralizing Antibody Responses to Multiple AAV Serotypes in Chimpanzees. *Hum. Gene Ther.* **27**, 79–82 (2016).

437. Barnes, C., Scheideler, O. & Schaffer, D. Engineering the AAV capsid to evade immune responses. *Curr. Opin. Biotechnol.* **60**, 99–103 (2019).

438. Bertin, B. *et al.* Capsid-specific removal of circulating antibodies to adeno-associated virus vectors. *Sci. Rep.* **10**, 1–11 (2020).

439. Hinderer, C. *et al.* Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose Intravenous Administration of an Adeno-Associated Virus Vector Expressing Human SMN. **29**, 285–298 (2018).

440. Wilson, J. M. & Flotte, T. R. Moving Forward After Two Deaths in a Gene Therapy Trial of Myotubular Myopathy. *Hum. Gene Ther.* **31**, 695–697 (2020).

441. Gao, G. *et al.* Adeno-associated viruses undergo substantial evolution in primates during natural infections. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6081–6086 (2003).

442. Kotin, R. M., Menninger, J. C., Ward, D. C. & Berns, K. I. Mapping and Direct

Visualization of a Region-Specific Integration Site on Chromosome 19ql3-qter Viral DNA. *Genomics* **10**, 831–834 (1991).

443. Kotin, R. M. *et al.* Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2211–2215 (1990).

444. Samulski, R. J. *et al.* Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *EMBO J.* **10**, 3941–3950 (1991).

445. Kotin, R. M., Linden, R. M. & Berns, K. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *EMBO J.* **11**, 5071–5078 (1992).

446. Mccarty, D. M., Jr, S. M. Y. & Samulski, R. J. Integration of Adeno-Associated Virus (AAV) and Recombinant AAV Vectors. *Annu. Rev. Genet.* **38**, 819–845 (2004).

447. Huser, D. *et al.* Integration Preferences of Wildtype AAV-2 for Consensus Rep-Binding Sites at Numerous Loci in the Human Genome. *PLoS Pathog.* **6**, e1000985 (2010).

448. Dutheil, N. *et al.* Characterization of the Mouse Adeno-Associated Virus AAVS1 Ortholog. *J. Virol.* **78**, 8917–8921 (2004).

449. Janovitz, T., Oliveira, T., Sadelain, M. & Falck-pedersen, E. Highly Divergent Integration Profile of Adeno-Associated Virus Serotype 5 Revealed by High-Throughput Sequencing. *J. Virol.* **88**, 2481–2488 (2014).

450. Kaeppel, C. *et al.* A largely random AAV integration profile after LPLD gene therapy. *Nat. Med.* **19**, 889–891 (2013).

451. Nakai, H. *et al.* AAV serotype 2 vectors preferentially integrate into active genes in mice. *Nat. Genet.* **34**, 297–302 (2003).

452. Tatsuno, K. *et al.* Impact of AAV2 and Hepatitis B Virus Integration Into Genome on Development of Hepatocellular Carcinoma in Patients with Prior Hepatitis B Virus Infection. *Clin. Cancer Res.* **25**, 6217–6228 (2019).

453. Bella, T. La *et al.* Adeno-associated virus in the liver: natural history and consequences in tumour development. *Gut* **0**, 1–11 (2019).

454. Chandler, R. J. *et al.* Vector design influences hepatic genotoxicity after adeno-associated virus gene therapy. 1–11 doi:10.1172/JCI79213DS1

455. Kattenhorn, L. M. *et al.* Adeno-Associated Virus Gene Therapy for Liver Disease. *Hum. Gene Ther.* **27**, 947–961 (2016).

456. Luk, J. M. *et al.* DLK1-DIO3 Genomic Imprinted MicroRNA Cluster at 14q32 . 2 Defines a Stemlike Subtype of Hepatocellular Carcinoma Associated with Poor Survival □. *J. Biol. Chem.* **286**, 30706–30713 (2011).

457. Laughlin, A., Myers, W. & Risin, L. Defective-Interfering Particles of the Human Parvovirus Adeno-Associated Virus. *Virology* **174**, 162–174 (1979).

458. Flotte, T. R. Empty Adeno-Associated Virus Capsids : Contaminant or Natural Decoy ? *Hum. Gene Ther.* **28**, 147–148 (2017).

459. Okada, T. *et al.* Scalable Purification of Adeno-Associated Virus Serotype 1 (AAV1) and AAV8 Vectors, Using Dual Ion-Exchange Adsorptive Membranes. *Hum. Gene Ther.* **20**, 1013–1021 (2009).

460. Wright, J. F. AAV Empty Capsids: For Better or for Worse? *Mol. Ther.* **22**, 2013–2014 (2013).

461. Gao, K. *et al.* Empty virions in AAV8 vector preparations reduce transduction efficiency and may cause total viral particle dose-limiting side effects. *Mol. Ther. — Methods Clin. Dev.* **1**, 1–8 (2014).

462. Mingozzi, F. *et al.* Overcoming Preexisting Humoral Immunity to AAV Using Capsid Decoys. *Sci. Transl. Med.* **5**, 194ra92 (2013).

463. Xie, J. *et al.* Short DNA Hairpins Compromise Recombinant Adeno-Associated Virus Genome Homogeneity. *Mol. Ther.* **25**, 1–12 (2017).

464. Tran, N. T. *et al.* AAV-Genome Population Sequencing of Vectors Packaging CRISPR Components Reveals Design-Influenced Heterogeneity. *Mol. Ther. Methods Clin. Dev.* **18**, 639–651 (2020).

465. Chadeuf, G., Ciron, C., Moullier, P. & Salvetti, A. Evidence for encapsidation of prokaryotic sequences during recombinant adeno-associated virus production and their in vivo persistence after vector delivery. *Mol. Ther.* **12**, 744–753 (2005).

466. Hauck, B. *et al.* Undetectable transcription of cap in a clinical AAV vector: implications for preformed capsid in immune responses. *Mol. Ther. J. Am. Soc. Gene Ther.* **17**, 144–52 (2009).

467. Schnödt, M. *et al.* DNA Minicircle Technology Improves Purity of Adeno-associated Viral Vector Preparations. 1–11 (2016). doi:10.1038/mtna.2016.60

468. Penaud-budloo, M. *et al.* Accurate Identification and Quantification of DNA Species by Next-Generation Sequencing in Adeno-Associated Viral Vectors Produced in Insect Cells. **28**, 148–162 (2017).

469. Nony, P., Chadeuf, G., Tessier, J., Moullier, P. & Salvetti, A. Evidence for packaging of rep-cap sequences into adeno-associated virus (AAV) type 2 capsids in the absence of inverted terminal repeats: a model for generation of rep-positive AAV particles. *J. Virol.* **77**, 776–81 (2003).

470. Wang, Z. *et al.* Elimination of contaminating cap genes in AAV vector virions reduces immune responses and improves transgene expression in a canine gene therapy model. *Gene Ther.* **21**, 363–370 (2014).

471. Emmerling, V. V *et al.* Rational plasmid design and bioprocess optimization to enhance recombinant adeno-associated virus (AAV) productivity in mammalian cells. *Biotechnol. J.* **11**, 290–297 (2016).

472. Namekawa, S. H., Payer, B., Huynh, K. D., Jaenisch, R. & Lee, J. T. Two-Step Imprinted X Inactivation : Repeat versus Genic Silencing in the Mouse □ †. **30**, 3187–3205 (2010).

473. Childs, D. Patient Death Latest Setback for Gene Therapy. *ABC News* (2007). Available at: https://abcnews.go.com/Health/story?id=3421869&page=1. (Accessed: 8th September 2020)

474. Keim, B. Gene Therapy: Is Death an Acceptable Risk? *Wired Magazine* (2007). Available at: https://www.wired.com/2007/08/gene-therapy-is-death-an-acceptable-risk/. (Accessed: 8th September 2020)

475. Urcelay, E., Ward, P., Wiener, S. M., Safer, B. & Kotin, R. M. Asymmetric Replication In Vitro from a Human Sequence Element Is Dependent on Adeno-Associated Virus Rep Protein. *J. Virol.* **69**, 2038–2046 (1995).

476. Brister, J. R. & Muzyczka, N. Rep-Mediated Nicking of the Adeno-Associated Virus Origin Requires Two Biochemical Activities, DNA Helicase Activity and Transesterification. *J. Virol.* **73**, 9325–9336 (1999).

477. McCarty, D. M. *et al.* Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther.* **10**, 2112–2118 (2003).

478. Haberman, R. P., Cown, T. J. M. C. & Samulski, R. J. Novel Transcriptional Regulatory Signals in the Adeno-Associated Virus Terminal Repeat A/D Junction Element. *J. Virol.* **74**, 8732–8739 (2000).

479. Alexander, I. E., Russell, D. W. & Miller, A. D. Transfer of Contaminants in Adeno-Associated Virus Vector Stocks Can Mimic Transduction and Lead to Artifactual Results. *Hum. Gene Ther.* **8**, 1911–1920 (1997).

480. Mcintosh, J. *et al.* Therapeutic levels of FVIII following a single peripheral vein administration of rAAV vector encoding a novel human factor VIII variant Therapeutic levels of FVIII following a single peripheral vein administration of rAAV vector encoding a novel human fac. *Blood* **121**, 3335–3344 (2013).

481. Wang, J. *et al.* Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.* **33**, 1256–1263 (2015).

482. Bak, R. O. & Porteus, M. H. CRISPR-Mediated Integration of Large Gene Cassettes Using AAV Donor Vectors. *Cell Rep.* **20**, 750–756 (2017).

483. Gaj, T. *et al.* Targeted gene knock-in by homology-directed genome editing using Cas9 ribonucleoprotein and AAV donor delivery. *Nucleic Acids Res.* **45**, 1–11 (2017).

484. Ikeda, Y. *et al.* Efficient Gene Transfer to Kidney Mesenchymal Cells Using a Synthetic Adeno-Associated Viral Vector. *J. Am. Soc. Nephrol.* **29**, 2287–2297 (2018).

485. Moyzis, R. K. *et al.* A highly conserved repetitive DNA sequence (TTAGGG)n present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6622–6626 (1988).

486. Chan, Y. K. & Church, G. M. Engineered viral vector reduces induction of inflammatory and immune responses. WO2017214378A1 (2017).

487. Grieger, J. C., Soltys, S. M. & Samulski, R. J. Production of Recombinant Adeno-associated Virus Vectors Using Suspension HEK293 Cells and Continuous Harvest of Vector From the Culture Media for GMP FIX and FLT1 Clinical Vector. *Mol. Ther.* **24**, 287–297 (2016).

488. Batchu, R. B. & Hermonat, P. L. The trans-inhibitory Rep78 protein of adeno-associated virus binds to TAR region DNA of the human immunodeficiency virus type 1 long terminal repeat. *FEBS Lett.* **367**, 267–271 (1995).

489. Hermonat, P. L. Inhibition of H-ras Expression by the Adeno-associated Virus Rep78 Transformation Suppressor Gene Product. *Cancer Res.* **51**, 3373–3377 (1991).

490. Beveren, C. Van, Enami, S., Curran, T. & Verma, I. M. FBR Murine Osteosarcoma Virus. *Virology* **135**, 218–28 (1984).

491. Bentley, D. L. & Groudine, M. Novel Promoter Upstream of the Human c-myc Gene and Regulation of c-myc Expression in B-Cell Lymphomas. *Mol. Cell. Biol.* **6**, 3481–3489 (1986).

492. Müller, R. Cellular and viral fos genes: structure, regulation of expression and biological properties of their encoded products. *Biochim. Biophys. Acta* **823**, 207–225 (1986).

493. Lee, W. & Keller, E. B. Regulatory Elements Mediating Transcription of the Human Ha-ras Gene. *J. Mol. Biol.* **220**, 599–611 (1991).

494. Hermonat, P. L. Adeno-associated Virus Inhibits Human Papillomavirus Type 16: A Viral Interaction Implicated in Cervical Cancer. *Cancer Res.* **54**, 2278–2281 (1994).

495. Zhan, D. *et al.* Binding of the Human Papillomavirus Type 16 p97 Promoter by the Adeno-associated Virus Rep78 Major Regulatory Protein Correlates with Inhibition *. **274**, 31619–31624 (1999).

496. Walz, C. M., Correa-ochoa, M. M., Müller, M. & Schlehofer, J. R. Adenoassociated Virus Type 2-Induced Inhibition of the Human Papillomavirus Type 18 Promoter in Transgenic Mice. *Virology* **293**, 172–181 (2002).

497. Chon, S. K., Rim, B. M. & Im, D. S. Adeno-Associated Virus Rep78 Binds to E2-Responsive Element 1 of Bovine Papillomavirus Type 1. *IUBMB Life* **48**, 397–404 (1999).

498. Pittayakhajonwut, D. & Angeletti, P. C. Viral trans-factor independent replication of human papillomavirus genomes. *Virol. J.* **7**, 1–9 (2010).

499. Merten, O.-W., Gény-Fiamma, C. & Douar, A.-M. Current issues in adeno-associated viral vector production. *Gene Ther.* **12**, S51–S61 (2005).

500. Zhou, X., Zolotukhin, I., Im, D. & Muzyczka, N. Biochemical Characterization of Adeno-Associated Virus Rep68 DNA Helicase and ATPase Activities Downloaded from http://jvi.asm.org/ on January 24 , 2020. *J. Virol.* **73**, 1580–1590 (1999).

501. Levitt, M. How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 640–644 (1978).

502. Penaud-budloo, M., François, A., Clément, N. & Ayuso, E. Pharmacology of Recombinant Adeno-associated Virus Production. *Mol. Ther. Methods Clin. Dev.* **8**, 166–180 (2018).

503. Butz, K. & Hoppe-seyler, F. Transcriptional Control of Human Papillomavirus (HPV) Upstream Regulatory Region. *J. Virol.* **67**, 6476–6486 (1993).

504. Sanchez, J. A., Pierce, K. E., Rice, J. E. & Wangh, L. J. Linear-After-The-Exponential (LATE)–PCR: An advanced method of asymmetric PCR and its uses in quantitative real-time analysis. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 1933–8 (2004).

505. Deyle, D. R. & Russell, D. W. Adeno-associated virus vector integration. *Curr. Opin. Mol. Ther.* **11**, 442–447 (2010).

506. Li, J., Samulski, R. J. & Xiao, X. Role for Highly Regulated rep Gene Expression in Adeno- Associated Virus Vector Production. *J. Virol.* **71**, 5236–5243 (1997).

507. Strobel, B. *et al.* Riboswitch-mediated Attenuation of Transgene Cytotoxicity Increases Adeno-associated Virus Vector Yields in HEK-293 Cells. *Mol. Ther.* **23**, 1582–1591 (2015).

508. Kyostio-Moore, S. *et al.* The impact of minimally oversized adeno-associated viral vectors encoding human factor VIII on vector potency in vivo. *Mol. Ther. — Methods Clin. Dev.* **3**, 16006 (2016).

509. Dong, B., Nakai, H. & Xiao, W. Characterization of genome integrity for oversized recombinant AAV vector. *Mol. Ther.* **18**, 87–92 (2010).

510. Hirsch, M. L. *et al.* Oversized AAV Transduction Is Mediated via a DNA-PKcs-independent, Rad51C-dependent Repair Pathway. *Mol. Ther.* **21**, 2205–2216 (2013).

511. Akil, O. *et al.* Dual AAV-mediated gene therapy restores hearing in a DFNB9 mouse model. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 4496–4501 (2019).

512. Carvalho, L. S. *et al.* Evaluating Efficiencies of Dual AAV Approaches for Retinal Targeting. *Front. Neurosci.* **11**, 1–8 (2017).

513. Lostal, W. *et al.* Efficient recovery of dysferlin deficiency by dual adeno-associated vector-mediated gene transfer. *Hum. Mol. Genet.* **19**, 1897–1907

(2010).

514. Ou, L. *et al.* A novel gene editing system to treat both Tay–Sachs and Sandhoff diseases. *Gene Ther.* **27**, 226–236 (2020).

515. Wang, Q. *et al.* CRISPR-Cas9-Mediated In Vivo Gene Integration at the Albumin Locus Recovers Hemostasis in Neonatal and Adult Hemophilia B Mice. *Mol. Ther. - Methods Clin. Dev.* **18**, 520–531 (2020).

516. Maddalena, A. *et al.* Triple Vectors Expand AAV Transfer Capacity in the Retina. *Mol. Ther.* **26**, 524–541 (2018).

517. Akil, O. Dual and triple AAV delivery of large therapeutic gene sequences into the inner ear. *Hear. Res.* **26**, 524–541 (2020).

518. Murphy, M., Gomos-klein, J., Stankic, M. & Falck-pedersen, E. Adeno-Associated Virus Type 2 p5 Promoter: a Rep-Regulated DNA Switch Element Functioning in Transcription, Replication, and Site-Specific Integration. *J. Virol.* **81**, 3721–3730 (2007).

519. Westhaus, A. *et al.* High-Throughput In Vitro, Ex Vivo, and In Vivo Screen of Adeno-Associated Virus Vectors Based on Physical and Functional Transduction. *Hum. Gene Ther.* **31**, 575–589 (2020).

520. Kyostio-moore, S. *et al.* The impact of minimally oversized adeno-associated viral vectors encoding human factor VIII on vector potency in vivo. *Mol. Ther. - Methods Clin. Dev.* **3**, 16006 (2016).

521. Alazard-Dany, N. *et al.* Definition of Herpes Simplex Virus Type 1 Helper Activities for Adeno-Associated Virus Early Replication Events. *PLoS Pathog.* **5**, 1–12 (2009).

522. Wang, Z., Cheng, F., Engelhardt, J. F., Yan, Z. & Qiu, J. Development of a Novel Recombinant Adeno-Associated Virus Production System Using Human Bocavirus 1 Helper Genes. *Mol. Ther. Methods Clin. Dev.* **11**, 40–51 (2018).

523. Seyffert, M. *et al.* Adeno-Associated Virus Type 2 Rep68 Can Bind to Consensus Rep-Binding Sites on the Herpes Simplex Virus 1 Genome. *J. Virol.* **89**, 11150–11158 (2015).

524. Iorio, A. *et al.* Establishing the Prevalence and Prevalence at Birth of Hemophilia in Males A Meta-analytic Approach Using National Registries. *Ann. Intern. Med.* **171**, 540–546 (2019).

525. Keeler, A. M. & Flotte, T. R. Recombinant Adeno-Associated Virus Gene Therapy in Light of Luxturna (and Zolgensma and Glybera): Where Are We, and How Did We Get Here? *Annu. Rev. Virol.* **6**, 601–621 (2019).

526. Farris, K. D. & Pintel, D. J. Improved Splicing of Adeno-Associated Viral (AAV) Capsid Protein-Supplying Pre-mRNAs Leads to Increased Recombinant AAV Vector Production. *Hum. Gene Ther.* **19**, 1421–1427 (2008).

527. Satkunanathan, S., Wheeler, J., Thorpe, R. & Zhao, Y. Establishment of a Novel Cell Line for the Enhanced Production of Recombinant Adeno-Associated Virus Vectors for Gene Therapy. *Hum. Gene Ther.* **25**, 929–941 (2014).

528. Burnett, J. C., Rossi, J. J. & Tiemann, K. Current progress of siRNA/shRNA therapeutics in clinical trials. *Biotechnol. J.* **6**, 1130–1146 (2011).

529. Valdmanis, P. N. *et al.* RNA interference – induced hepatotoxicity results from loss of the first synthesized isoform of microRNA-122 in mice. *Nat. Med.* **22**, 557–562 (2016).

530. Course, M. M., Gudsnuk, K., Desai, N., Chamberlain, J. R. & Valdmanis, P. N. Endogenous MicroRNA Competition as a Mechanism of shRNA-Induced Cardiotoxicity. *Mol. Ther. Nucleic Acid* **19**, 572–580 (2020).

531. Chen, Z. Y., He, C. Y., Ehrhardt, A. & Kay, M. A. Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. *Mol. Ther.* **8**, 495–500 (2003).

532. Ghattas, I. R., Sanes, J. R. & Majors, J. E. The Encephalomyocarditis Virus Internal Ribosome Entry Site Allows Efficient Coexpression of Two Genes from a Recombinant Provirus in Cultured Cells and in Embryos. *Mol. Cell. Biol.* **11**, 5848–5859 (1991).

533. Szymczak, A. L. & Vignali, D. A. A. Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin. Biol. Ther.* **5**, 627–638 (2005).

534. Houdebine, L. M. & Attal, J. Internal ribosome entry sites (IRESs): reality and use. *Transgenic Res.* **8**, 157–177 (1999).

535. Liu, Z., Chen, O., Wall, J. B. J., Zheng, M. & Zhou, Y. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci. Rep.* **7**, 2193 (2017).

536. Kim, J. H. *et al.* High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines , Zebrafish and Mice. **6**, 1–8 (2011).

537. Golding, M. C. & Mann, M. R. W. A bidirectional promoter architecture enhances lentiviral transgenesis in embryonic and extraembryonic stem cells. *Gene Ther.* **18**, 817–826 (2011).

538. Vogl, T. *et al.* Engineered bidirectional promoters enable rapid multi-gene co-expression optimization. *Nat. Commun.* **9**, 3589 (2018).

539. Yan, Z. *et al.* A Novel Chimeric Adenoassociated Virus 2/ Human Bocavirus 1 Parvovirus Vector Efficiently Transduces Human Airway Epithelia. *Mol. Ther.* **21**, 2181–2194 (2013).

540. Zhang, W. *et al.* An Engineered Virus Library as a Resource for the Spectrum-wide Exploration of Virus and Vector Resource An Engineered Virus Library as a Resource for the Spectrum-wide Exploration of Virus and Vector Diversity. *Cell Rep.* **19**, 1698–1709 (2017).

541. Gao, J., Mese, K., Bunz, O. & Ehrhardt, A. State-of-the-art human adenovirus vectorology for therapeutic approaches. *FEBS Lett.* **593**, 3609–3622 (2019).

542. Amiss, T. J., Mccarty, D. M., Skulimowski, A. & Samulski, R. J. Identification and Characterization of an Adeno-Associated Virus Integration Site in CV-1 Cells from the African Green Monkey. *J. Virol.* **77**, 1904–1915 (2003).

543. Bishop, B. M., Santin, A. D., Quirk, J. G. & Hermonat, P. L. Role of the terminal repeat GAGC trimer, the major Rep78 binding site, in adeno-associated virus DNA replication. *FEBS Lett.* **397**, 97–100 (1996).

Appendices:

Appendix 1 - Detailed protocol: Caesium Chloride plasmid purification:

Caesium Chloride plasmid purification:

Materials:

1. Kanamycin Disulfate salt: cat# K1876-25G, Sigma
2. Ampicillin Sodium salt: cat# A0166-25G, Sigma
3. Isopropanol: 1L. Cat# A416-4, Fisher Scientific
4. Isopropanol: 500ml. Cat# I9516-500ml, Sigma
5. CsCl: cat# BP210-500, Fisher Scientific
6. Ethidium Bromide (10mg/ml): cat# E1510-10ml, Sigma
7. OptiSeal centrifuge tubes: 36.2ml, cat# 362183; 11.2ml, cat# 362181, Beckman Coulter
8. Needles: 18G, cat# 305195, BD. 20G, cat# 305176, BD
9. Slide-A-Lyzer Dialysis cassette (10K MWCO): 0.1-0.5ml capacity: cat# 66383; 0.5-3ml capacity: cat# 66380; 3-12ml capacity, cat# 66810, Thermo Scientific

Method:

1. Seed 1 litre of NZY media + antibiotic in flask with plasmid from glycerol stock.
2. Incubate at 37°C overnight with shaking
3. Spin down at 6000g at 4°C for 20 min.
4. Discard supernatant.
5. Freeze the cell pellet at -20°C or continue to next step.
6. Add 30ml **Cell Resuspension Solution**, resuspend the cells.
    - 50mM TRIS-HCL pH 7.5
    - 10mM EDTA
    - In ddH_20
7. Add 30ml **Cell Lysis Solution**, invert 6 times, then stand @RT for 5 min.
    - 0.2M NaOH
    - 1% SDS
    - In ddH_20
8. Add 30ml **Neutralization Solution**, invert 6 times.
    - 1.32M Potassium Acetate
    - Made to pH4.8 with glacial acetic acid
    - In ddH_20
9. Spin at 12,000 rpm for 15 min at 4°C.

10. Filter supernatant through 70um cell strainer.
11. Add 0.7 volume of Isopropanol + mix
12. After 10 min, spin at 20,000g for 30 min at 4°C.
13. Add 26ml of TE to resuspend DNA pellet by gentle shaking.
14. Add 27g of Caesium Chloride and mix.
15. Add 1ml of Ethidium Bromide (EtBr stock 10mg/ml) into a 36.2ml OptiSeal centrifuge tube
    - add the CsCl-DNA-TE mixed solution into the same OptiSeal tube
16. Balance tubes with CsCl/TE 1:1 solution, cap and mix.
17. Spin in ultracentrifuge at 47K rpm for >16 hrs @20°C. (Beckman VTI 50 rotor)
    - After spin, avoid mixing of gradient during handling of tubes.
18. Harvest the lower DNA band (plasmid) with an 18G needle and syringe, transfer DNA solution into an 11.2ml OptiSeal tube, fill with CsCl/TE 1:1 solution, and balance tubes.
19. Spin at 65K rpm for 6 hrs @20°C (Beckman NVT 65 rotor).
20. Harvest DNA band with 18G needle and syringe and transfer solution into a 50ml conical tube.
21. Remove EtBr by adding same amount of Isopropanol from an Isopropanol-CsCl-TE solution.
22. Repeat 3-5 times until liquid appears clear.
23. Transfer solution to dialysis cassette with a 20G needle and syringe
24. Dialyse with 4L of TE Buffer O/N.
25. Change out TE buffer, dialyse for 4 hours.
26. Change out TE buffer, dialyse for 4 hours.
27. Collect DNA solution from dialysis cassette in conical tube.
28. Quantitate DNA and label

Appendix 2 – List of active AAV clinical trials:

Sourced from clinicaltrials.gov search term AAV gene therapy

ClinicalTrials.gov Search Results 07/27/2020

	Title	Status	Study Results	Conditions	Interventions	Locations
1	Efficacy and Safety of AAV2-REP1 for the Treatment of Choroideremia	Active, not recruiting	No Results Available	•Choroideremia	•Genetic: AAV2-REP1	•Study Site, Los Angeles, California, United States •Study Site, Miami, Florida, United States •Study Site, Baltimore, Maryland, United States •Study Site, New York, New York, United States •Study Site, Portland, Oregon, United States •Study Site, Dallas, Texas, United States •Study Site, Madison, Wisconsin, United States •Study Site, Montréal, Canada •Study Site, Vancouver, Canada •Study Site, Glostrup, Denmark •and 7 more
2	Trial of AAV5-hFIX in Severe or Moderately Severe Hemophilia B	Active, not recruiting	No Results Available	•Hemophilia B	•Genetic: AAV5-hFIX	•uniQure Investigative Site, Copenhagen, Denmark •uniQure Investigative Site, Berlin, Germany •uniQure Investigative Site, Frankfurt, Germany •uniQure Investigative Site, Amsterdam, Netherlands •uniQure Investigative Site, Groningen, Netherlands •uniQure Investigative Site, Rotterdam, Netherlands •uniQure Investigative Site, Utrecht, Netherlands
3	Safety and Dose Escalation Study of AAV2-hCHM in Subjects With CHM (Choroideremia) Gene Mutations	Active, not recruiting	No Results Available	•Choroideremia •CHM (Choroideremia) Gene Mutations	•Biological: AAV2-hCHM	•Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, United States •University of Pennsylvania, Philadelphia, Pennsylvania, United States •Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States
4	A Phase 1/2 Open-Label, Dose Escalation Study to Determine the Optimal Dose, Safety, and Activity of AAV2hAQP1 in Subjects With Radiation-Induced Parotid Gland Hypofunction and Xerostomia	Recruiting	No Results Available	•Radiation-Induced Parotid Gland Hypofunction •Xerostomia Due to Radiotherapy •Head and Neck Cancer	•Drug: intra-parotid administration of AAV2hAQP1	•University of Louisville, Louisville, Kentucky, United States •Brigham and Women's Hospital, Boston, Massachusetts, United States •Memorial Sloan Kettering Cancer Center, New York, New York, United States •Atrium Health, Charlotte, North Carolina, United States •Health Sciences North - Northeast Cancer Center, Sudbury, Ontario, Canada
5	VRC 603: A Phase I Dose-Escalation Study of the Safety of AAV8-VRC07 (VRC-HIVAAV070-00-GT) Recombinant AAV Vector Expressing VRC07 HIV-1 Neutralizing Antibody in Antiretroviral -Treated, HIV-1 Infected Adults With Controlled Viremia.	Recruiting	No Results Available	•HIV-1 Infected Adults With Controlled Viremia	•Drug: VRC-HIVAAV070-00-GT (AAV8-VRC07)	•National Institutes of Health Clinical Center, Bethesda, Maryland, United States
6	Dose Confirmation Trial of AAV5-hFIXco-Padua	Active, not recruiting	No Results Available	•Hemophilia B	•Genetic: AAV5-hFIXco-Padua (AMT-061)	•Los Angeles Orthopedic Hospital, Los Angeles, California, United States •University of California, Davis, Sacramento, California, United States •University of California, San Diego, San Diego, California, United States •University of Michigan, Ann Arbor, Michigan, United States
7	AAV2/8-LSPhGAA in Late-Onset Pompe Disease	Recruiting	No Results Available	•Pompe Disease	•Biological: AAV2/8LSPhGAA	•Duke University, Durham, North Carolina, United States

271

	Title	Status	Study Results	Conditions	Interventions	Locations
8	Long-term Safety and Efficacy Follow-up of AAV2-REP1 for the Treatment of Choroideremia (SOLSTICE)	Enrolling by invitation	No Results Available	•Choroideremia	•Genetic: AAV2-REP1	•Study Site, Miami, Florida, United States •Study Site, Tübingen, Germany
9	A Safety Study of Retinal Gene Therapy for Choroideremia	Active, not recruiting	No Results Available	•Choroideremia	•Drug: AAV2-REP1	•Study Site, Miami, Florida, United States •Study Site, Boston, Massachusetts, United States •Study Site, Cincinnati, Ohio, United States •Study Site, Portland, Oregon, United States •Study Site, Tübingen, Germany
10	AAV2-GDNF for Advanced Parkinson s Disease	Active, not recruiting	No Results Available	•Parkinson's Disease	•Genetic: Convection enhanced delivery/ AAV2-GDNF	•National Institutes of Health Clinical Center, 9000 Rockville Pike, Bethesda, Maryland, United States
11	Long-term Follow-up Study in Subjects Who Received Voretigene Neparvovec-rzyl (AAV2-hRPE65v2)	Active, not recruiting	No Results Available	•Inherited Retinal Dystrophy Due to RPE65 Mutations	•Biological: AAV2-hRPE65v2	
12	Gene Therapy for X-linked Retinitis Pigmentosa (XLRP) Retinitis Pigmentosa GTPase Regulator (RPGR)	Active, not recruiting	No Results Available	•X-Linked Retinitis Pigmentosa	•Genetic: AAV2/5-RPGR	•Massachusetts Eye and Ear Institute, Boston, Massachusetts, United States •Kellogg Eye Center, Ann Arbor, Michigan, United States •UPMC Eye Center, Pittsburgh, Pennsylvania, United States •Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom •Moorfields Eye Hospital NHS Foundation Trust, London, United Kingdom
13	Phase 1 Follow-on Study of AAV2-hRPE65v2 Vector in Subjects With Leber Congenital Amaurosis (LCA) 2	Active, not recruiting	No Results Available	•Leber Congenital Amaurosis	•Biological: voretigene neparvovec-rzyl	•The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States
14	A Phase I/II Clinical Trial for Treatment of Aromatic L-amino Acid Decarboxylase (AADC) Deficiency Using AAV2-hAADC	Active, not recruiting	No Results Available	•Aromatic L-amino Acid Decarboxylase (AADC) Deficiency	•Drug: gene therapy	•National Taiwan University Hospital, Taipei, Taiwan
15	A Clinical Trial for Treatment of Aromatic L-amino Acid Decarboxylase (AADC) Deficiency Using AAV2-hAADC - An Expansion	Recruiting	No Results Available	•Aromatic Amino Acid Decarboxylase Deficiency	•Drug: AAV2-hAADC	•National Taiwan University Hospital, Taipei, Taiwan
16	Global Epidemiologic Study of Preexisting Immunity to AAV in Adults With Severe Hemophilia	Active, not recruiting	No Results Available	•Hemophilia A •Hemophilia B	•Other: Non-treatment, seroprevalence	•Orthopaedic Hemophilia Treatment Center, Los Angeles, California, United States •University of Colorado Hemophilia & Thrombosis Center, Aurora, Colorado, United States •Michigan State University, East Lansing, Michigan, United States •Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States •Medical University of South Carolina (MUSC), Charleston, South Carolina, United States •Gulf States Hemophilia and Thrombophilia Center, Houston, Texas, United States •University of Washington, Seattle, Washington, United States •AKH - Medizinische Universität Wien, Vienna, Austria •Hôpital Morvan, Brest Cedex, France •Groupement Hospitalier Est- Hôpital Louis Pradel, Bron cedex, France •and 13 more
17	Evaluation of Safety and Tolerability of Libella Gene Therapy for the Treatment of Aging: AAV- hTERT	Recruiting	No Results Available	•Aging	•Drug: AAV-hTERT	•IPS Arcasalud SAS, Zipaquirá, Cundinamarca, Colombia

	Title	Status	Study Results	Conditions	Interventions	Locations
18	Long-Term Follow-Up Gene Therapy Study for Leber Congenital Amaurosis OPTIRPE65 (Retinal Dystrophy Associated With Defects in RPE65)	Recruiting	No Results Available	•Leber Congenital Amaurosis (LCA) •Eye Diseases •Eye Diseases, Hereditary •Retinal Diseases	•Biological: AAV OPTIRPE65	•Moorfields Eye Hospital NHS Foundation Trust, London, United Kingdom
19	Evaluation of Safety and Tolerability of Libella Gene Therapy for Critical Limb Ischemia: AAV- hTERT	Recruiting	No Results Available	•Critical Limb Ischemia	•Drug: AAV-hTERT	•IPS Arcasalud SAS, Zipaquirá, Cundinamarca, Colombia
20	Evaluation of Safety and Tolerability of Libella Gene Therapy for Alzheimer's Disease: AAV- hTERT	Recruiting	No Results Available	•Alzheimer Disease	•Drug: AAV-hTERT	•IPS Arcasalud SAS, Zipaquirá, Cundinamarca, Colombia
21	Safety and Efficacy Study in Subjects With Leber Congenital Amaurosis	Active, not recruiting	Has Results	•Inherited Retinal Dystrophy Due to RPE65 Mutations •Leber Congenital Amaurosis	•Biological: AAV2-hRPE65v2,voretigene neparvovec-rzyl	•University of Iowa, Iowa City, Iowa, United States •Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States
22	Long-Term Follow-Up Gene Therapy Study for Achromatopsia CNGB3 and CNGA3	Recruiting	No Results Available	•Achromatopsia	•Biological: either AAV - CNGB3 or AAV - CNGA3	•Moorfields Eye Hospital NHS Foundation Trust, London, United Kingdom
23	Gene Therapy for Achromatopsia (CNGA3)	Recruiting	No Results Available	•Achromatopsia	•Biological: AAV- CNGA3	•Kellogg Eye Center, Ann Arbor, Michigan, United States •Moorfields Eye Hospital NHS Foundation Trust, London, United Kingdom
24	A Clinical Trial of Retinal Gene Therapy for X-linked Retinitis Pigmentosa	Recruiting	No Results Available	•X-Linked Retinitis Pigmentosa	•Biological: AAV8-RPGR	•Study Site, Gainesville, Florida, United States •Study Site, Miami, Florida, United States •Study Site, Portland, Oregon, United States •Study Site, Philadelphia, Pennsylvania, United States •Study Site, Dallas, Texas, United States •Study Site, Manchester, United Kingdom •Study Site, Oxford, United Kingdom •Study Site, Southampton, United Kingdom
25	HOPE-B: Trial of AMT-061 in Severe or Moderately Severe Hemophilia B Patients	Active, not recruiting	No Results Available	•Hemophilia B	•Genetic: AAV5-hFIXco-Padua	•Phoenix Children's Hospital, Phoenix, Arizona, United States •Arkansas Children's Hospital, Little Rock, Arkansas, United States •Los Angeles Orthopedic Hospital, Los Angeles, California, United States •Children's Hospital of Los Angeles, Los Angeles, California, United States •University of California, Davis, Sacramento, California, United States •University of California, San Diego, San Diego, California, United States •University of California, San Francisco, San Francisco, California, United States •University of Colorado Denver, Aurora, Colorado, United States •Children's National Medical Center Hematology and Oncology, Washington, District of Columbia, United States •University of South Florida, Tampa, Florida, United States •and 29 more

	Title	Status	Study Results	Conditions	Interventions	Locations
26	Re-administration of Intramuscular AAV9 in Patients With Late-Onset Pompe Disease	Recruiting	No Results Available	•Pompe Disease	•Genetic: Recombinant Adeno-Associated Virus Acid Alpha-Glucosidase •Drug: Rapamycin •Other: saline •Drug: Rituxan •Drug: Diphenhydramine •Drug: Acetaminophen •Drug: Lidocaine •Drug: LMX 4 Topical Cream	•Clinical and Translational Research Building (CTRB), University of Florida, Gainesville, Florida, United States
27	A Gene Therapy Study for Homozygous Familial Hypercholesterolemia (HoFH)	Active, not recruiting	No Results Available	•Homozygous Familial Hypercholesterolemia (HoFH)	•Biological: AAV directed hLDLR gene therapy	•Boca Raton location, Boca Raton, Florida, United States •Kansas City Location, Kansas City, Kansas, United States •Portland location, Portland, Oregon, United States •Philadelphia Location, Philadelphia, Pennsylvania, United States •Nashville location, Nashville, Tennessee, United States •Chicoutimi location, Chicoutimi, Quebec, Canada •Montreal location, Montreal, Quebec, Canada •Montreal location, Montreal, Quebec, Canada •Palermo location, Palermo, PA, Italy •Rome location, Roma, RM, Italy •Rotterdam location, Rotterdam, Netherlands
28	GDNF Gene Therapy for Parkinson's Disease	Recruiting	No Results Available	•Parkinson's Disease	•Biological: AAV2-GDNF	•University of California San Francisco, San Francisco, California, United States •The Ohio State University Medical Center, Columbus, Ohio, United States
29	Safety and Dose-Finding Study of DTX401 (AAV8G6PC) in Adults With Glycogen Storage Disease Type Ia (GSDIa)	Active, not recruiting	No Results Available	•GSD1	•Genetic: DTX401	•UCONN Health, Farmington, Connecticut, United States •Michigan Medicine University of Michigan, Ann Arbor, Michigan, United States •UT Health - McGovern Medical School, Houston, Texas, United States •Montreal Children Hospital, McGill University Health Centre, Montréal, Quebec, Canada •University Medical Center Groningen, Groningen, Netherlands •Complejo Hospitalario Universitario de Santiago, Santiago De Compostela, A Coruna, Spain •Hospital Universitario Cruces Servicio Pediatria, Barakaldo, Vizcaya, Spain •Hospital Universitario 12 de Octubre, Madrid, Spain
30	Long Term Follow-Up Gene Therapy Study for XLRP RPGR	Recruiting	No Results Available	•X-Linked Retinitis Pigmentosa	•Biological: AAV-RPGR	•Moorfields Eye Hospital, London, United Kingdom

	Title	Status	Study Results	Conditions	Interventions	Locations
31	Gene Therapy Study in Severe Hemophilia A Patients With Antibodies Against AAV5	Enrolling by invitation	No Results Available	•Haemophilia A •Gene Therapy •Clotting Disorders •Blood Disorder	•Biological: Valoctocogene Roxaparvovec	•Nantes University Hospital Center - Hotel Dieu Hospital, Nantes, France •Charlotte Maxeke Johannesburg Academic Hospital, Hemophilia Comprehensive Care Center, Johannesburg, South Africa •Royal Free Hospital, London, United Kingdom •University Hospital Southampton NHS Foundation Trust, Southampton, United Kingdom
32	ADVM-022 Intravitreal Gene Therapy for Wet AMD	Recruiting	No Results Available	•Wet Age-related Macular Degeneration •Neovascular Age-related Macular Degeneration	•Biological: ADVM-022	•Adverum Clinical Site, Bakersfield, California, United States •Adverum Clinical Site, Beverly Hills, California, United States •Adverum Clinical Site, Golden, Colorado, United States •Adverum Clinical Site, Deerfield Beach, Florida, United States •Adverum Clinical Site, Reno, Nevada, United States •Adverum Clinical Site, Philadelphia, Pennsylvania, United States •Adverum Clinical Site, West Columbia, South Carolina, United States •Adverum Clinical Site, Nashville, Tennessee, United States •Adverum Clinical Site, Abilene, Texas, United States •Adverum Clinical Site, Houston, Texas, United States •Adverum Clinical Site, The Woodlands, Texas, United States
33	REP1 Gene Replacement Therapy for Choroideremia	Active, not recruiting	No Results Available	•Choroideremia	•Genetic: AAV-mediated REP1 gene replacement	•Moorfields Eye Hospital NHS Foundation Trust, London, United Kingdom •Oxford University Hospitals NHS Foundation Trust, Oxford, United Kingdom
34	Safety Study in Subjects With Leber Congenital Amaurosis	Active, not recruiting	No Results Available	•Leber Congenital Amaurosis	•Biological: voretigene neparvovec-rzyl	•The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States
35	ADVM-022 Intravitreal Gene Therapy for DME	Recruiting	No Results Available	•Diabetic Macular Edema •Diabetic Retinopathy	•Biological: 6E11 vg/eye of ADVM-022 •Biological: 2E11 vg/eye of ADVM-022 •Biological: Aflibercept	•Adverum Clinical Site, Beverly Hills, California, United States •Adverum Clinical Site, Reno, Nevada, United States •Adverum Clinical Site, Philadelphia, Pennsylvania, United States •Adverum Clinical Site, West Columbia, South Carolina, United States •Adverum Clinical Site, Nashville, Tennessee, United States •Adverum Clinical Site, Abilene, Texas, United States •Adverum Clinical Site, The Woodlands, Texas, United States

	Title	Status	Study Results	Conditions	Interventions	Locations
36	A Phase 1/2 Study of Intravenous Gene Transfer With an AAV9 Vector Expressing Human <=-Galactosidase in Type I and Type II GM1 Gangliosidosis	Recruiting	No Results Available	•Lysosomal Diseases •Gangliosidosis •GM1	•Biological: AAV9-GLB1 •Drug: Rituximab •Drug: Sirolimus •Drug: Methylprednisolone •Drug: Prednisone •Diagnostic Test: Audiology assessmentwith ABR •Diagnostic Test: Bone density scan (DEXA •Diagnostic Test: Electrocardiogram (EKG) •Diagnostic Test: Echocardiogram •Other: Electroencephalogram (EEG) awake andextended overnight •and 10 more	•National Institutes of Health Clinical Center, Bethesda, Maryland, United States
37	Gene Therapy for Male Patients With Danon Disease Using RP-A501; AAV9.LAMP2B	Recruiting	No Results Available	•Danon Disease	•Biological: RP-A501	•University of California, San Diego, La Jolla, California, United States
38	Safety of a Single Administration of AAV2hAQP1, an Adeno-Associated Viral Vector Encoding Human Aquaporin-1 to One Parotid Salivary Gland in People With Irradiation-Induced Parotid Salivary Hypofunction	Recruiting	No Results Available	•Squamous Cell Head and Neck Cancer •Radiation Induced Xerostomia •Salivary Hypofunction	•Biological: AAV2hAQP1	•National Institutes of Health Clinical Center, 9000 Rockville Pike, Bethesda, Maryland, United States
39	Gene Therapy for Chinese Hemophilia B	Recruiting	No Results Available	•Hemophilia B	•Genetic: Single dose intravenous injection of BBM-H901	•Institute of Hematology & Blood Diseases Hospital, Tianjin, Tianjin, China
40	Study of RS1 Ocular Gene Transfer for X-linked Retinoschisis	Recruiting	No Results Available	•Retinoschisis •X-Linked	•Biological: RS1 AAV Vector	•National Institutes of Health Clinical Center, 9000 Rockville Pike, Bethesda, Maryland, United States
41	A Single-Stage, Adaptive, Open-label, Dose Escalation Safety and Efficacy Study of AADC Deficiency in Pediatric Patients	Recruiting	No Results Available	•AADC Deficiency	•Drug: AAV2-hAADC	•University of California San Francisco, Benioff Children's Hospital, San Francisco, California, United States
42	Gene Therapy for Haemophilia A.	Recruiting	No Results Available	•Hemophilia A	•Biological: AAV2/8-HLP-FVIII-V3	•University of Kentucky, Lexington, Kentucky, United States •St Jude's Children's Research Hospital, Memphis, Tennessee, United States •Royal Free Hospital, London, United Kingdom
43	Long-Term Follow-up to Evaluate the Safety and Efficacy of Adeno Associated Virus (AAV) Serotype 8 (AAV8)-Mediated Gene Transfer of Glucose-6-Phosphatase (G6Pase) in Adults With Glycogen Storage Disease Type Ia (GSDIa)	Enrolling by invitation	No Results Available	•Glycogen Storage Disease Type IA •Von Gierke's Disease (GSD Type Ia)	•Other: No intervention	•UCONN Health, Farmington, Connecticut, United States •Michigan Medicine University of Michigan, Ann Arbor, Michigan, United States •University of Texas Health Science Center at Houston, Houston, Texas, United States

	Title	Status	Study Results	Conditions	Interventions	Locations
44	Dose-Ranging Study of Recombinant AAV2/6 Human Factor 8 Gene Therapy SB-525 in Subjects With Severe Hemophilia A	Recruiting	No Results Available	•Hemophilia A	•Biological: SB-525	•Arkansas Children's Hospital, Little Rock, Arkansas, United States •Midtown Ambulatory Care Center, Sacramento, California, United States •UC Davis Ambulatory Care Clinic, Sacramento, California, United States •UC Davis Comprehensive Cancer Center, Sacramento, California, United States •UC Davis CTSC Clinical Research Center, Sacramento, California, United States •UC Davis Hemophilia Clinical Trials Unit, Sacramento, California, United States •UC Davis Hemophilia Treatment Center, Sacramento, California, United States •UC Davis Investigational Drug Services Pharmacy, Sacramento, California, United States •UC Davis Medical Center, Sacramento, California, United States •UCSF Medical Center, San Francisco, California, United States •and 11 more
45	Safety and Efficacy Study in Patients With Retinitis Pigmentosa Due to Mutations in PDE6B Gene	Recruiting	No Results Available	•Retinitis Pigmentosa	•Biological: AAV2/5-hPDE6B	•Clinique Ophtalmologique, CHU de Nantes, Nantes, France
46	A Patient Registry Study for Patients Treated With Voretigene Neparvovec	Active, not recruiting	No Results Available	•Confirmed Biallelic RPE65 Mutation-associated Retinal Dystrophy	•Biological: AAV2-hRPE65v2,voretigene neparvovec-rzyl	•Children's Hospital of Los Angeles, Los Angeles, California, United States •Bascom Palmer Eye Institute, Miami, Florida, United States •University of Iowa Hospitals & Clinics, Iowa City, Iowa, United States •Massachusetts Eye and Ear Institute, Boston, Massachusetts, United States •Kellogg Eye Center, Ann Arbor, Michigan, United States •Cincinnati Eye Institute, Cincinnati, Ohio, United States •Casey Eye Institute, Portland, Oregon, United States •Scheie Eye Institute, Philadelphia, Pennsylvania, United States •Cullen Eye Institute, Houston, Texas, United States

	Title	Status	Study Results	Conditions	Interventions	Locations
47	Study to Test the Safety and How Well Patients With Severe Hemophilia A Respond to Treatment With BAY 2599023 (DTX 201), a Drug Therapy That Delivers a Healthy Version of the Defective Factor VIII Gene Into the Nucleus of Liver Cells Using an Altered, Non-infectious Virus (AAV) as a "Shuttle".	Recruiting	No Results Available	•Hemophilia A	•Drug: BAY2599023 (DTX201)	•Arkansas Children's Hospital, Little Rock, Arkansas, United States •Loma Linda Children's Hospital, San Bernardino, California, United States •C.S. Mott Children's Hospital, Ann Arbor, Michigan, United States •University Hospitals Cleveland Medical Center, Cleveland, Ohio, United States •University of Wisconsin - Madison, Madison, Wisconsin, United States •SHATHD Spec. Hospi. for Active Treatm. of Haematol. Dis. EAD, Sofia, Bulgaria •Hôpital Calmette - Lille Cedex, Lille Cedex, France •Hôpital de la Timone - Marseille, Marseille, France •Hopital Necker les enfants malades - Paris, Paris, France •Hôpital Pontchaillou, Rennes Cedex, France •and 15 more
48	Study of Viral Transduction of Human Auditory Sensory Cells for the Development of Gene Therapy	Recruiting	No Results Available	•Hearing Loss, Sensorineural	•Other: Peroperative collect of inner ear cells	•Hôpital de Bicêtre, Le Kremlin-Bicêtre, Ile De France, France •Hôpital Pitié-Salpétrière, Paris, Ile De France, France
49	A Phase 1/2 Study of SHP648, an Adeno-Associated Viral Vector for Gene Transfer in Hemophilia B Subjects	Recruiting	No Results Available	•Hemophilia B	•Genetic: SHP648	•AKH - Medizinische Universität Wien, Vienna, Austria •Hôpital Morvan, Brest, France •Groupement Hospitalier Sud - Hôpital Bicêtre, Le Kremlin Bicêtre, France •Hopital Jeanne de Flande - CHI Lille, Lille, France •Hôpital de la Timone, Marseille, France •CHU de Nantes Site Hotel Dieu, Nantes, France •Hôpital Necker - Enfants Malades, Paris, France •Klinikum der Johann Wolfgang Goethe-Universität, Frankfurt, Germany •Semmelweis Egyetem, Budapest, Hungary •Chaim Sheba Medical Center, Ramat Gan, Israel •and 7 more
50	AAV Gene Therapy Study for Subjects With PKU	Recruiting	No Results Available	•Phenylketonuria (PKU)	•Drug: BMN 307	•University Hospital Birmingham NHS Foundation Trust, Birmingham, United Kingdom

	Title	Status	Study Results	Conditions	Interventions	Locations
51	Safety and Dose Escalation Study of an Adeno-Associated Viral Vector for Gene Transfer in Hemophilia A Participants	Active, not recruiting	No Results Available	•Hemophilia A	•Drug: BAX 888	•Phoenix Childrens Hospital, Phoenix, Arizona, United States •Orthopaedic Hemophilia Treatment Center, Los Angeles, California, United States •University of Colorado Hemophilia & Thrombosis Center, Aurora, Colorado, United States •Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States •Medical University of South Carolina (MUSC), Charleston, South Carolina, United States •Gulf States Hemophilia and Thrombophilia Center City, Houston, Texas, United States •AKH - Medizinische Universität Wien, Vienna, Austria •Hôpital de la Timone, Marseille Cedex 05, Bouches-du-Rhône, France •CHU de Nantes Site Hotel Dieu, Nantes Cedex 1, Loire Atlantique, France •Hopital Jeanne de Flandre - CHU Lille, Lille Cedex, Nord, France •and 6 more
52	MICRO-RNAs OF NEUTROPHILS IN RENAL ANTINEUTROPHIL CYTOPLASMIS ANTIBODY (ANCA) - ASSOCIATED VASCULITIS	Recruiting	No Results Available	•ANCA-associated Vasculitis	•Other: Non interventional study	•Nantes University Hospital, Nantes, France
53	Gene Transfer Clinical Study in X-Linked Myotubular Myopathy	Active, not recruiting	No Results Available	•X-Linked Myotubular Myopathy	•Genetic: AT132	•UCLA Medical Center, Los Angeles, California, United States •Powell Center for Rare Disease Research, Univ. of Florida, Gainesville, Florida, United States •Ann & Robert H Lurie Children's Hospital of Chicago, Chicago, Illinois, United States •National Institute of Neurological Disorders and Stroke/NIH Porter, Bethesda, Maryland, United States •Hospital for Sick Children, Toronto, Ontario, Canada •Hopital Armad Trousseau, Paris, France •Kinderklinik und Kinderpoliklinik im Dr. Von Haunerschen Kinderspital Klinikum der Universitat Munchen, München, Germany
54	A Single Dose Clinical Trial to Study the Safety of ART-I02 in Patients With Arthritis	Active, not recruiting	No Results Available	•Arthritis, Rheumatoid •Osteo Arthritis	•Genetic: ART-I02	•Centre for Human Drug Research (CHDR), Leiden, Zuid Holland, Netherlands
55	AAV9 U7snRNA Gene Therapy to Treat Boys With DMD Exon 2 Duplications.	Enrolling by invitation	No Results Available	•Duchenne Muscular Dystrophy	•Biological: scAAV9.U7.ACCA	•Nationwide Children's Hospital, Columbus, Ohio, United States
56	Exploring Durable Remission With Rituximab in Antineutrophil Cytoplasmic Antibody(ANCA)-Associated Vasculitis	Recruiting	No Results Available	•ANCA Associated Vasculitis	•Drug: Rituximab •Drug: endoxan •Drug: Methylprednisolone •Drug: Prednisolone	•Leiden University Medical Center, Leiden, Zuid-Holland, Netherlands •Hammersmith Hospital, London, United Kingdom

	Title	Status	Study Results	Conditions	Interventions	Locations
57	Study of AAVrh10-h.SGSH Gene Therapy in Patients With Mucopolysaccharidosis Type IIIA (MPS IIIA)	Active, not recruiting	No Results Available	•Mucopolysaccharidosis Type IIIA	•Drug: LYS-SAF302	•CHOC Children's, Orange, California, United States •University of Minnesota, Minneapolis, Minnesota, United States •Weill Cornell Medical College, New York, New York, United States •Baylor college of medicine / Texas children's hospital, Houston, Texas, United States •Armand Trousseau Public Hospital, Paris, France •University Medical Center Hamburg-Eppendorf, Hamburg, Germany •Amsterdam UMC, Amsterdam, Netherlands •Great Ormond Street Hospital, London, United Kingdom
58	Six Month lead-in Study to Evaluate Prospective Efficacy and Safety Data of Current FIX Prophylaxis Replacement Therapy in Adult Hemophilia B Subjects (FIX:C#2%) or Current FVIII Prophylaxis Replacement Therapy in Adult Hemophilia A Subjects (FVIII:C#1%)	Recruiting	No Results Available	•Hemophilia B •Hemophilia A	•Drug: Standard of Care FIX Replacement therapy •Drug: Standard of Care FVIII Replacement therapy	•UC Davis Ambulatory Care Clinical, Sacramento, California, United States •UC Davis Comprehensive Cancer Center, Sacramento, California, United States •UC Davis CTSC Clinical Research Center, Sacramento, California, United States •UC Davis Health, Sacramento, California, United States •UC Davis Hemophilia Clinical Trials Unit, Sacramento, California, United States •UC Davis Investigational Drug Services Pharmacy, Sacramento, California, United States •UC Davis Midtown Cancer Center, Sacramento, California, United States •University of California San Francisco - UCSF HTC, San Francisco, California, United States •University of California, San Francisco - Clinical Research Center, San Francisco, California, United States •University of California, San Francisco - Outpatient Hematology Clinic, San Francisco, California, United States •and 54 more
59	Gene Transfer Study in Patients With Late Onset Pompe Disease	Recruiting	No Results Available	•Pompe Disease (Late-onset)	•Genetic: AT845	•University of California Irvine, Department of Neurology, Orange, California, United States •Stanford University, Palo Alto, California, United States •University of Utah, Division of Medical Genetics, Salt Lake City, Utah, United States •Ludwig-Maximillians University of Munich, Munich, Germany •Newcastle Upon Tyne Hospitals Foundation Trust Clinical Research Facility, Newcastle upon Tyne, United Kingdom
60	Gene Therapy for Children With Variant Late Infantile Neuronal Ceroid Lipofuscinosis 6 (vLINCL6) Disease	Active, not recruiting	No Results Available	•Variant Late-Infantile Neuronal Ceroid Lipofuscinosis	•Genetic: AT-GTX-501	•Nationwide Children's Hosptial, Columbus, Ohio, United States
61	Ascending Dose Study of Genome Editing by Zinc Finger Nuclease Therapeutic SB-FIX in Subjects With Severe Hemophilia B	Active, not recruiting	No Results Available	•Hemophilia B	•Biological: SB-FIX	•Georgetown University Medical Center, Washington, District of Columbia, United States
62	Dose Escalation Study of Intravitreal 4D-110 in Patients With Choroideremia	Recruiting	No Results Available	•Choroideremia	•Biological: 4D-110	•Retina Foundation of the Southwest, Dallas, Texas, United States •Moran Eye Center, University of Utah, Salt Lake City, Utah, United States

	Title	Status	Study Results	Conditions	Interventions	Locations
63	Gene Therapy Study in Severe Haemophilia A Patients	Active, not recruiting	No Results Available	•Severe Haemophilia A	•Genetic: BMN 270	•Hampshire Hospitals NHS Foundation Trust, Basingstoke, United Kingdom •Queen Elizabeth Hospital Birmingham, Birmingham, United Kingdom •University Hospitals Bristol NHS Foundation Trust, Bristol, United Kingdom •Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom •Greater Glasgow Health Board, Glasgow, United Kingdom •Barts Health NHS Trust, London, United Kingdom •Guy's & St. Thomas' NHS Foundation Trust, London, United Kingdom •Imperial College Healthcare NHS Trust, London, United Kingdom •University Hospital Southampton NHS Foundation Trust, Southampton, United Kingdom
64	Safety and Efficacy Trial of AAV Gene Therapy in Patients With CNGA3 Achromatopsia	Recruiting	No Results Available	•Achromatopsia	•Biological: AGTC-402	•University of California, San Francisco, San Francisco, California, United States •VitreoRetinal Associates, Gainesville, Florida, United States •Bascom Palmer Eye Institute, Miami, Florida, United States •Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, United States •Boston Children's Hospital, Boston, Massachusetts, United States •Columbia College of Physicians and Surgeons, New York, New York, United States •Cincinnati Eye Institute, Cincinnati, Ohio, United States •Casey Eye Institute, Oregon Health and Sciences University, Portland, Oregon, United States •University of Pennsylvania, Philadelphia, Pennsylvania, United States •Baylor College of Medicine, Houston, Texas, United States •University of Wisconsin, McPherson Eye Research Institute, Madison, Wisconsin, United States •Hadassah-Hebrew University Medical Center, Jerusalem, Israel

	Title	Status	Study Results	Conditions	Interventions	Locations
65	Safety and Efficacy Trial of AAV Gene Therapy in Patients With CNGB3 Achromatopsia	Recruiting	No Results Available	•Achromatopsia	•Biological: rAAV2tYF-PR1.7-hCNGB3	•University of California, San Francisco, San Francisco, California, United States •VitreoRetinal Associates, Gainesville, Florida, United States •Bascom Palmer Eye Institute, Miami, Florida, United States •Pangere Center for Inherited Retinal Diseases, The Chicago Lighthouse for People Who Are Blind or Visually Imp, Chicago, Illinois, United States •Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, United States •Boston Children's Hospital, Boston, Massachusetts, United States •Columbia College of Physicians and Surgeons, New York, New York, United States •Duke Eye Center, Duke University Medical Center, Durham, North Carolina, United States •Cincinnati Eye Institute, Cincinnati, Ohio, United States •Cleveland Clinic, Cole Eye Institute, Cleveland, Ohio, United States •and 5 more
66	Gene Therapy in Patients With Mucopolysaccharidosis Disease	Recruiting	No Results Available	•Mucopolysaccharidosis Type VI	•Biological: AAV2/8.TBG.hARSB	•Federico II University, Napoli, Naples, Italy •Erasmus Medical Center, Rotterdam Center for Lysosomal and Metabolic disease, Rotterdam, Netherlands •Children's Hospital Hacettepe University, Ankara, Turkey
67	A Study to Evaluate the Safety and Tolerability of PF-06939926 Gene Therapy in Duchenne Muscular Dystrophy	Enrolling by invitation	No Results Available	•Duchenne Muscular Dystrophy	•Genetic: PF-06939926	•MRI Research Center, Los Angeles, California, United States •Reed Neurological Research Center, Los Angeles, California, United States •Ronald Reagan UCLA Medical Center Drug Information Center, Los Angeles, California, United States •UCLA (David Geffen School of Medicine), Los Angeles, California, United States •UCLA Mattel Children's Hospital, Los Angeles, California, United States •UCLA Medical Center, Los Angeles, California, United States •UCLA Outpatient Surgery Center, Los Angeles, California, United States •Duke Neurology, Durham, North Carolina, United States •Duke University Medical Center, Lenox Baker Children's Hospital, Durham, North Carolina, United States •Duke Biospecimen Repository & Processing Core - BPRC, Durham, North Carolina, United States •and 10 more
68	Gene Therapy for Children With CLN3 Batten Disease	Active, not recruiting	No Results Available	•CLN3 •Batten Disease	•Genetic: AT-GTX-502	•Nationwide Children's Hospital, Columbus, Ohio, United States

	Title	Status	Study Results	Conditions	Interventions	Locations
69	Maintaining or Stopping Immunosuppressive Therapy in Patients With ANCA Vasculitis and End-stage Renal Disease	Recruiting	No Results Available	•Anti-Neutrophil Cytoplasmic Antibody-Associated Vasculitis •End Stage Renal Disease	•Other: Discontinuation (or not initiation) of Immunosuppressive Therapy •Drug: Maintenance (or initiation) of immunosuppressive treatment: Imurel®, Mabthera®,Cellcept®, Cortancyl®	•Centre Hospitalier Universitaire Amiens, Amiens, France •Centre Hospitalier Angoulême, Angoulême, France •Centre Hospitalier ARRAS, Arras, France •Centre Hospitalier Avignon, Avignon, France •CHRU Besançon, Besançon, France •Centre Hospitalier Universitaire Bordeaux, Bordeaux, France •Centre Hospitalier Boulogne sur Mer, Boulogne-sur-Mer, France •Centre Hospitalier Jacques Coeur, Bourges, France •Centre Hospitalier Universitaire de Brest, Brest, France •Centre Hospitalier René Dubois - Pontoise, Cergy-Pontoise, France •and 40 more
70	THOR - Tübingen Choroideremia Gene Therapy Trial	Active, not recruiting	No Results Available	•Choroideremia	•Genetic: rAAV2.REP1	•University Hospital Tuebingen, Center for Ophthalmology, Tuebingen, Germany
71	Long Term Follow Up to Evaluate DTX301 in Adults With Late-Onset OTC Deficiency	Enrolling by invitation	No Results Available	•Ornithine Transcarbamylase (OTC) Deficiency	•Other: No Intervention	•Boston Children's Hospital, Boston, Massachusetts, United States •Icahn School of Medicine, New York, New York, United States •Alberta Children's Hospital, Calgary, Alberta, Canada •Hospital Clinico Universitario de Santiago, Santiago De Compostela, Coruna, Spain •Hospital Universitario de Cruces. Servicio de Pediatria, Barakaldo, Vizcaya, Spain •Queen Elizabeth Hospital, Department of Endocrinology, Birmingham, United Kingdom
72	Safety and Dose-Finding Study of DTX301 (scAAV8OTC) in Adults With Late-Onset OTC Deficiency	Recruiting	No Results Available	•Ornithine Transcarbamylase (OTC) Deficiency	•Genetic: scAAV8OTC •Drug: Oral prednisone	•Ronald Reagan University of Califonria Los Angeles Medical Center, Los Angeles, California, United States •University of California San Francisco, San Francisco, California, United States •The Children's Hospital Colorado, Aurora, Colorado, United States •Boston Children's Hospital, Boston, Massachusetts, United States •Icahn School of Medicine at Mount Sinai, New York, New York, United States •University Hospital Cleveland Medical Center/Case Western Reserve University, Cleveland, Ohio, United States •Oregon Health and Science University, Portland, Oregon, United States •Alberta's Children's Hospital, Calgary, Alberta, Canada •Hospital Clinico Universitario de Santiago, Santiago de Compostela, A Coruna, Spain •Hospital Universitario de Cruzes, Barakaldo, Vizcaya, Spain •National Hospital for Neurology & Neurosurgery, London, London City, United Kingdom •Queen Elizabeth Hospital, Birmingham, United Kingdom

	Title	Status	Study Results	Conditions	Interventions	Locations
73	Long-Term Safety, Tolerability, and Efficacy of DTX101 (AAVrh10FIX) in Adults With Moderate/Severe to Severe Hemophilia B	Active, not recruiting	No Results Available	•Hemophilia B		•Arkansas Children's Hospital, Little Rock, Arkansas, United States •UF CRC - Clinical Research Center, Gainesville, Florida, United States •University of Michigan Hospital and Health Systems, Ann Arbor, Michigan, United States •Haemophilia, Haemostasis & Thrombosis Centre, Basingstoke, Hampshire, United Kingdom •Manchester Haemophilia Comprehensive Care Center, Manchester, United Kingdom
74	Gene Therapy for Severe Crigler Najjar Syndrome	Recruiting	No Results Available	•Crigler-Najjar Syndrome	•Genetic: GNT0003	•Hopital Antoine BECLERE, Clamart, France •ASST Papa Giovanni XXIII, Bergame, Italy •Azienda Ospedaliera Universitaria Federico II, Napoli, Italy •AMC, Amsterdam, Netherlands
75	Single-Arm Study To Evaluate The Efficacy and Safety of Valoctocogene Roxaparvovec in Hemophilia A Patients at a Dose of 4E13 vg/kg	Active, not recruiting	No Results Available	•Hemophilia A	•Biological: Valoctocogene Roxaparvovec	•UC Davis Hemophilia Treatment Center, Sacramento, California, United States •UCSF Medical Center, San Francisco, California, United States •University of Colorado AMC, Hemophilia and Thrombosis Center, Aurora, Colorado, United States •University of Michigan, Pediatric Hematology and Oncology, Ann Arbor, Michigan, United States •University of Minnesota, Minneapolis, Minnesota, United States •Washington University School of Medicine, Department of Pediatrics, Division of Hematology/Oncology, Saint Louis, Missouri, United States •UNC Hemophilia and Thrombosis Center, Chapel Hill, North Carolina, United States •Nationwide Children's Hospital, Columbus, Ohio, United States •Hemophilia Center of Western Pennsylvania, Pittsburgh, Pennsylvania, United States •Morvan Hospital, Brest, France •and 9 more
76	Gene Therapy for APOE4 Homozygote of Alzheimer's Disease	Recruiting	No Results Available	•Alzheimer Disease •Early Onset Alzheimer Disease	•Biological: AAVrh.10hPOE2 vector	•Weill Cornell Medical College and Weill Cornell Medical Center, Department of Genetic Medicine, New York, New York, United States
77	RGX-314 Gene Therapy for Neovascular AMD Trial	Active, not recruiting	No Results Available	•Neovascular Age-related Macular Degeneration •Wet Age-related Macular Degeneration	•Genetic: RGX-314	•Santa Barbara location, Santa Barbara, California, United States •Baltimore location, Baltimore, Maryland, United States •Boston location, Boston, Massachusetts, United States •Reno location, Reno, Nevada, United States •Philadelphia location 1, Philadelphia, Pennsylvania, United States •Philadelphia location 2, Philadelphia, Pennsylvania, United States •Memphis location, Germantown, Tennessee, United States •Houston location, Houston, Texas, United States

	Title	Status	Study Results	Conditions	Interventions	Locations
78	Single-Dose Gene Replacement Therapy Clinical Trial for Patients With Spinal Muscular Atrophy Type 1	Active, not recruiting	No Results Available	•SMA	•Biological: Onasemnogene Abeparvovec-xioi	•University Hospital Ghent Neuromuscular reference center, Ghent, Belgium •Neuropédiatrie - Centre de Référence des Maladies Neuromusculaires, Liège, Belgium •Hôpital Armand Trousseau, Paris, France •Istituto Gianninia Gaslini, Genova, Italy •Policlinico "G. Martino", Messina, Italy •Carlo Besta Neurological Research Institute, Milan, Italy •University of Milan, Milan, Italy •Policlinico Gemelli, Rome, Italy •Great Ormond Street Hospital for Children, London, United Kingdom •The John Walton Muscular Dystrophy Research Centre MRC Centre for Neuromuscular Diseases at Newcastle, Newcastle Upon Tyne, United Kingdom
79	Phase I Trial of CEA Specific AAV-DC-CTL Treatment in Stage IV Gastric Cancer	Active, not recruiting	No Results Available	•Gastric Cancer	•Biological: CTL	
80	Dose-Escalation Study Of A Self Complementary Adeno-Associated Viral Vector For Gene Transfer in Hemophilia B	Active, not recruiting	No Results Available	•Hemophilia B	•Genetic: Gene Transfer •Drug: scAAV2/8-LP1-hFIXco	•Stanford Medical School, Stanford, California, United States •University of Kentucky, Lexington, Kentucky, United States •Oregon Health and Science University, Portland, Oregon, United States •Hemophilia Center of Western Pennsylvania, Pittsburgh, Pennsylvania, United States •St. Jude Children's Research Hospital, Memphis, Tennessee, United States •University of Texas Southwestern, Dallas, Texas, United States •Scott and White Memorial Hospital, Temple, Texas, United States •Katharine Dormandy Haemophilia Centre and Haemostasis Unit, University College of London, London, United Kingdom
81	Dose-finding Study of SPK-8016 Gene Therapy in Patients With Hemophilia A to Support Evaluation in Individuals With FVIII Inhibitors	Active, not recruiting	No Results Available	•Adeno-Associated Virus (AAV) •Blood Coagulation Disorder •Blood Coagulation Disorders, Inherited •Coagulation Protein Disorders •Factor VIII (FVIII) •Factor VIII (FVIII) Deficiency •Factor VIII (FVIII) Gene •Factor VIII (FVIII) Protein •Genetic Diseases, Inborn •Genetic Diseases, X-Linked •and 7 more	•Genetic: SPK-8016	•Illinois Bleeding and Clotting Disorders Institute, Peoria, Illinois, United States •Mississippi Center for Advanced Medicine, Madison, Mississippi, United States •Weill Cornell Medicine, New York, New York, United States •Oregon Health & Science University, Portland, Oregon, United States •Penn State Health, Hershey, Pennsylvania, United States •Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States •Jefferson University Hospitals, Philadelphia, Pennsylvania, United States •Virginia Commonwealth University School of Medicine, Richmond, Virginia, United States

	Title	Status	Study Results	Conditions	Interventions	Locations
82	Dose-Ranging Study of ST-920, a rAAV2/6 Human Alpha Galactosidase A Gene Therapy in Subjects With Fabry Disease	Recruiting	No Results Available	•Fabry Disease	•Biological: ST-920	•Emory University School of Medicine, Atlanta, Georgia, United States •University of Minnesota Medical Center, Minneapolis, Minnesota, United States •NYU Langone Health Neurogenetics, New York, New York, United States •Mt. Sinai School of Medicine, New York, New York, United States •Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States •Lysosomal and Rare Disorders Research and Treatment Center (LDRTC), Fairfax, Virginia, United States
83	VY-AADC02 for Parkinson's Disease With Motor Fluctuations (RESTORE-1)	Recruiting	No Results Available	•Parkinson's Disease	•Biological: VY-AADC02 •Other: Sham (Placebo) Surgery	•UC Irvine, Irvine, California, United States •UC Davis Health System, Sacramento, California, United States •San Francisco VA Medical Center, San Francisco, California, United States •University of California, San Francisco, San Francisco, California, United States •University of Colorado, Aurora, Colorado, United States •Emory University Hospital, Atlanta, Georgia, United States •Northwestern Medical Faculty Foundation, Chicago, Illinois, United States •University of Kansas Medical Center, Kansas City, Kansas, United States •Tufts Medical Center, Boston, Massachusetts, United States •Beth Israel Deaconess Medical Center (BIDMC), Boston, Massachusetts, United States •and 10 more
84	The Effect of Some Drugs Used in Treatment of Vasculitis on the Complement System in Children	Recruiting	No Results Available	•Vasculitis	•Drug: Ibuprofen •Drug: Prednisone •Drug: Methotrexate	•Assiut University Pediatric Hospital, Assiut, Upper Egypt, Egypt
85	Ascending Dose Study of Genome Editing by the Zinc Finger Nuclease (ZFN) Therapeutic SB-318 in Subjects With MPS I	Active, not recruiting	No Results Available	•MPS I	•Biological: SB-318	•UCSF Benioff Children's Hospital Oakland, Oakland, California, United States
86	Comparison of ANCA and Anti-GBM Auto-antibodies Removal Kinetics Between Plasma Exchanges and Immunoadsorption in Patients With ANCA-associated Vasculitis or Anti-GBM Disease	Recruiting	No Results Available	•Kidney Failure, Acute	•Procedure: Apheresis technics	•Assisatance Publique Hôpitaux de Marseille, Marseille, France

	Title	Status	Study Results	Conditions	Interventions	Locations
87	Pre-Symptomatic Study of Intravenous Onasemnogene Abeparvovec-xioi in Spinal Muscular Atrophy (SMA) for Patients With Multiple Copies of SMN2	Active, not recruiting	No Results Available	•Spinal Muscular Atrophy	•Biological: onasemnogene abeparvovec-xioi	•David Geffen School of Medicine at UCLA, Los Angeles, California, United States •Stanford University, Stanford, California, United States •Children's Hospital Colorado, Aurora, Colorado, United States •Nemours Children's Hospital, Orlando, Florida, United States •Center for Rare Neurological Diseases, Norcross, Georgia, United States •Massachusetts General Hospital, Boston, Massachusetts, United States •Helen DeVos Children's Hospital, Grand Rapids, Michigan, United States •Washington Unviersity School of Medicine, Saint Louis, Missouri, United States •Columbia University, New York, New York, United States •Nationwide Children's Hospital, Columbus, Ohio, United States •and 19 more
88	First in Human Study to Evaluate the Safety and Efficacy of GT005 Administered in Subjects With Dry AMD	Recruiting	No Results Available	•Dry Age-related Macular Degeneration •Macular Degeneration •Retinal Disease •Eye Diseases •Retinal Degeneration •Geographic Atrophy •Macular Atrophy	•Biological: GT005	•Bristol Eye Hospital, Bristol, United Kingdom •London Vision Clinic, London, United Kingdom •Moorfields Eye Hospital, London, United Kingdom •Manchester Eye Hospital, Manchester, United Kingdom •Oxford University Hospital, Oxford, United Kingdom •Sunderland Eye Infirmary, Sunderland, United Kingdom
89	Safety of Intra-Articular Sc-rAAV2.5IL-1Ra in Subjects With Moderate Knee OA	Recruiting	No Results Available	•Osteoarthritis, Knee	•Drug: sc-rAAV2.5IL-1Ra	•Mayo Clinic, Rochester, Minnesota, United States
90	A Long Term Follow-Up Study of Fabry Disease Subjects Treated With FLT190	Recruiting	No Results Available	•Fabry Disease •Lysosomal Storage Diseases	•Genetic: FLT190	•Royal Free London, London, United Kingdom
91	Open-Label Single Ascending Dose of Adeno-associated Virus Serotype 8 Factor IX Gene Therapy in Adults With Hemophilia B	Active, not recruiting	No Results Available	•Hemophilia B	•Biological: AskBio009	•Orthopaedic Hemophilia Treatment Center, Los Angeles, California, United States •Children's Hospital Los Angeles, Los Angeles, California, United States •University of California Davis Medical Center, Sacramento, California, United States •University of California at San Diego Medical Center, San Diego, California, United States •U of Colorado School of Medicine, Hemophilia & Thrombosis Treatment Center, Aurora, Colorado, United States •Emory University, Atlanta, Georgia, United States •Rush University Medical Center, Chicago, Illinois, United States •Children's Hospital of Boston, Boston, Massachusetts, United States •University of Minnesota, Masonic Clinical Research Unit, Clinical and Translational Science Institute, Minneapolis, Minnesota, United States •Mount Sinai Medical Center, New York, New York, United States •and 5 more

	Title	Status	Study Results	Conditions	Interventions	Locations
92	A Fabry Disease Gene Therapy Study	Recruiting	No Results Available	•Fabry Disease •Lysosomal Storage Diseases	•Genetic: FLT190	•Universita Federico II di Napoli, Napoli, Italy •Haukeland University Hospital, Bergen, Norway •Royal Free Hospital, London, United Kingdom
93	Systemic Gene Delivery Clinical Trial for Duchenne Muscular Dystrophy	Active, not recruiting	No Results Available	•Duchenne Muscular Dystrophy	•Genetic: rAAVrh74.MHCK7.micro-dystrophin	•Nationwide Children's Hospital, Columbus, Ohio, United States
94	Single-Dose Gene Replacement Therapy Using for Patients With Spinal Muscular Atrophy Type 1 With One or Two SMN2 Copies	Recruiting	No Results Available	•Spinal Muscular Atrophy Type I	•Biological: Onasemnogene Abeparvovec-xioi	•Tokyo Women's Medical University, Tokyo, Japan •Pusan National University Yangsan Hospital, Yangsan, Gyeongsangnam-do, Korea, Republic of •Seoul National University Hospital, Seoul, Korea, Republic of •National Taiwan University Hospital, Taipei, Taiwan
95	Identifying and Genotyping Homozygous Familial Hypercholesterolemia (HoFH) Patients	Active, not recruiting	No Results Available	•Homozygous Familial Hypercholesterolemia (HoFH)		•Excel Medical Clinical Trials, LLC, Boca Raton, Florida, United States
96	Safety Study of a Gene Transfer Vector (Rh.10) for Children With Late Infantile Neuronal Ceroid Lipofuscinosis (LINCL)	Active, not recruiting	No Results Available	•Batten Disease •Late-Infantile Neuronal Ceroid Lipofuscinosis	•Biological: AAVrh.10CUhCLN2 vector 9.0x10^11 genome copies •Biological: AAVrh.10CUhCLN2 vector 2.85x10^11 genome copies	•Weill Cornell Medical College, New York, New York, United States
97	Ascending Dose Study of Genome Editing by the Zinc Finger Nuclease (ZFN) Therapeutic SB-913 in Subjects With MPS II	Active, not recruiting	No Results Available	•Mucopolysaccharidosis II •MPS II	•Biological: SB-913	•UCSF Benioff Children's Hospital Oakland, Oakland, California, United States •Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, Illinois, United States •NYU School of Medicine, Neurogenetics Division, New York, New York, United States •University of North Carolina, Chapel Hill, North Carolina, United States •Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States
98	Lead-in Study to Collect Prospective Efficacy and Safety Data of Current FVIII Prophylaxis Replacement Therapy in Adult Hemophilia A Participants	Active, not recruiting	No Results Available	•Blood Coagulation Disorder •Blood Coagulation Disorders, Inherited •Coagulation Protein Disorders •Hemophilia A •Genetic Diseases, Inborn •Genetic Diseases, X-Linked •Hematologic Diseases •Hemorrhagic Disorders •Factor VIII Deficiency	•Drug: Standard of Care FVIII Replacement therapy	•University of California San Francisco, San Francisco, California, United States •University of Florida, Gainesville, Florida, United States •Emory University, Atlanta, Georgia, United States •Children's Hospital of Michigan, Detroit, Michigan, United States •Mississippi Center for Advanced Medicine, Madison, Mississippi, United States •Bloodworks Northwest, Seattle, Washington, United States •The Alfred Hospital, Melbourne, Victoria, Australia •Fiona Stanley Hospital, Murdoch, Western Austrailia, Australia •Providence Hematology/St. Paul's Hosptial, Vancouver, British Columbia, Canada •McMaster University / Royal Prince Alfred Hospital, Hamilton, Ontario, Canada •Ramathibodi Hospital, Mahidol University, Bangkok, Thailand
99	Glybera Registry, Lipoprotein Lipase Deficient (LPLD) Patients	Active, not recruiting	No Results Available	•Lipoprotein Lipase Deficiency •Familial Hyperlipoproteinemia Type 1 •Familial Hyperchylomicronemia	•Other: Observational study	•Interdisciplinary Metabolism Center, Lipid Out-Patient-Clinic, Lipid Apheresis, Charité, University of Berlin, Berlin, Germany

	Title	Status	Study Results	Conditions	Interventions	Locations
100	Trial of Ocular Subretinal Injection of a Recombinant Adeno-Associated Virus (rAAV2-VMD2-hMERTK) Gene Vector to Patients With Retinal Disease Due to MERTK Mutations	Recruiting	No Results Available	•Retinal Disease	•Genetic: Recombinant Adeno-Associated Virus	•King Khaled Eye Specialist Hospital, Riyadh, Saudi Arabia

19 additional studies not shown